

Uniwersytet Marii Curie-Skłodowskiej
w Lublinie

Wydział Filozofii i Socjologii
Instytut Filozofii

Łukasz Blechar

OD SZTUCZNEJ INTELIGENCJI DO ZAAWANSOWANEJ ANALITYKI I
WZMACNIANIA INTELIGENCJI

Praca doktorska napisana pod kierunkiem
Dr hab. Piotra Gizy, prof. UMCS

Lublin 2023

Spis treści

1.	Wprowadzenie	1
1.1.	Marzenia o myślących maszynach	1
1.2.	Konsekwencje realizowania marzeń	3
1.2.1.	Rywalizacja z maszynami	3
1.2.2.	Technologiczna osobliwość.....	3
1.2.3.	Krytyka współczesnej sztucznej inteligencji.....	4
1.3.	Dostępna alternatywa.....	4
1.3.1.	Współpraca zamiast rywalizacji	4
1.3.2.	Symbioza ludzi i maszyn.....	5
2.	Motywacja i metody badawcze.....	6
2.1.	Implementowanie starych pomysłów nowoczesnymi metodami.....	6
2.2.	Główne postawy wobec sztucznej inteligencji.....	7
2.2.1.	Niezależne maszyny	7
2.2.2.	Współpraca z maszynami	8
2.2.3.	Percepcja obu koncepcji w momencie ich powstania	8
2.3.	Luka badawcza	9
2.3.1.	Symbioza jako obecnie dominujące podejście	9
2.3.2.	Superumysły Malone'a	9
2.3.3.	Brakujący środek	10
2.3.4.	Wyzwania edukacyjne.....	10
2.3.5.	Stan polskiej nauki w kontekście wzmacniania inteligencji	11
2.4.	Cel rozprawy	12
2.5.	Metoda	12
2.6.	Struktura pracy	13
3.	Część I – dwie historie sztucznej inteligencji.....	16
3.1.	Historia myśli technicznej.....	16

3.1.1.	Od antycznej Grecji do XX wieku.....	16
3.1.2.	XX wiek do konferencji w Dartmouth.....	19
3.1.3.	1956 rok — Warsztaty w Dartmouth College.....	26
3.1.4.	Początki sztucznej inteligencji	28
3.1.5.	Weryfikacja początkowych założeń sztucznej inteligencji.....	31
3.1.6.	Systemy oparte na wiedzy.....	34
3.1.7.	Sztuczna inteligencja współcześnie	39
3.2.	Myśl filozoficzna a sztuczna inteligencja.....	45
3.2.1.	Filozofia sztucznej inteligencji.....	45
3.2.2.	Krytyka obliczeniowego podejścia do umysłu.....	52
4.	Część II – dwa źródła koncepcji sztucznej inteligencji	60
4.1.	Spółeczność umysłów	60
4.1.1.	Wprowadzenie	60
4.1.2.	Ogólna struktura	61
4.1.3.	Historia idei	62
4.1.4.	Jak działa spółeczność umysłów?.....	65
4.1.5.	Podsumowanie.....	71
4.2.	Symbioza ludzi i maszyn	72
4.2.1.	Wstęp	72
4.2.2.	Koncepcje symbiozy ludzi i maszyn w latach 60. XX wieku.....	75
4.2.3.	William Ross Ashby: Wzmacnianie Inteligencji.....	75
4.2.4.	Joseph Carl Robnett Licklider: Symbioza ludzi i maszyn	76
4.2.5.	Douglas Carl Engelbart: System H-LAM/T.....	78
4.2.6.	Podsumowanie.....	79
4.3.	Porównanie koncepcji	81
5.	Część III – współczesne amalgamaty umysłów	83
5.1.	Superumysły	83

5.1.1.	Inteligencja	84
5.1.2.	Sztuczna inteligencja.....	87
5.1.3.	Definicja superumysłu.....	97
5.1.4.	Superumysły jako wsparcie procesów poznawczych	99
5.1.5.	Podsumowanie.....	109
5.2.	Human + Machine: Reimagining Work in the Age of AI	112
5.2.1.	Rozumienie sztucznej inteligencji	112
5.2.2.	Trzecia fala transformacji biznesowych.....	114
5.2.3.	Brakujący Środek.....	116
5.2.4.	Przykłady zastosowań sztucznej inteligencji.....	120
5.2.5.	Wypełnienie Brakującego środka	122
5.2.6.	Pięć koniecznych warunków dla odkrycia brakującego środka	124
5.3.	Porównanie koncepcji i analityki danych	126
5.3.1.	Wzmocniona analityka.....	126
5.3.2.	Trzy pokolenia BI	127
5.3.3.	Cykl działań analitycznych.....	129
5.3.4.	Usprawnienie cyklu analizy za pomocą sztucznej inteligencji	130
5.3.5.	Podsumowanie	131
6.	Zakończenie.....	135
6.1.	Podsumowanie i wnioski.....	135
6.2.	Inne ujęcia podejmowania decyzji.....	136
6.2.1.	Przegląd definicji	136
6.2.2.	Podejmowanie decyzji jako proces poznawczy	136
6.2.3.	Decyzje strategiczne.....	139
6.3.	Założenia własnego modelu współpracy ludzi i maszyn.....	142
6.4.	Model współpracy ludzi i maszyn.....	144
6.5.	Główne elementy procesu podejmowania decyzji	145

6.5.1. Środowisko	145
6.5.2. Wiedza	146
6.5.3. Cele	146
6.5.4. Ograniczenia.....	147
6.5.5. Scenariusze	148
6.5.6. Decyzje.....	149
6.5.7. Implementacje	150
6.5.8. Oceny.....	150
6.5.9. Pamięć	150
7. Dyskusja	152
Bibliografia	156
Spis rysunków:	174

1. Wprowadzenie

Sztuczna inteligencja jest jednym z najtrudniejszych egzaminów, przed jakimi stanęliśmy jako ludzkość. Jej doniosłość i rola jest szeroko uznawana. Niektórzy mówią wprost, że stanowi jedną z kluczowych technologii naszych czasów (West i Allen, 2018). Inni, jak Przegalińska, widzą w niej więcej, jeden z największych metafizycznych projektów XX i XXI wieku (Przegalińska, 2016, s. 9). Możemy ją rozpatrywać z wielu, różnych perspektyw, lecz jej istotność stanowi fakt.

Bezsprzecznie wspiera nas w integrowaniu informacji, analizowaniu danych i, w rezultacie, usprawnianiu naszych procesów decyzyjnych. Zyskała wręcz pewną renomę na tym polu. Do tego stopnia, że, jak dowodzi Ciupa, często wykorzystywana jest obecnie jako chwyt marketingowy, który ukrywa niedookreślone cele i nierealne ambicje (Ciupa, 2017).

1.1. Marzenia o myślących maszynach

Ambicja stworzenia maszyny myślącej jak człowiek towarzyszy nam od wielu wieków. Przegalińska stwierdza, że stworzenie robota naśladującego człowieka, androida, czy połączenia robota i człowieka, cyborga, z perspektywy psychologicznej, może być postrzegane w kategorii wyrazu potrzeby transcendowania rzeczywistości i wykorzystania technologii do wydłużenia naszego życia biologicznego (Przegalińska, 2016, s. 83). Nie jest też to bynajmniej nowe oczekiwanie. Przegalińska w innym miejscu pisze bowiem, że ludzkość od dawna marzy o maszynie zbudowanej i działającej na podobieństwo człowieka. Postrzega tę ambicję jako wyraz chęci zrozumienia umysłu ludzkiego i jego wewnętrznych mechanizmów prowadzących do inteligentnych zachowań (Przegalińska, 2016, s. 27).

Marzenia o stworzeniu maszyn, które będą nas czy to wyręczać, czy wspierać sięgają, jak zauważają niektórzy badacze, aż do prawie tysiąca lat przed naszą erą. Jedna z najstarszych opowieści opowiada o Yan Shi, słynnym rzemieślniku, który ofiarował swojemu królowi Mu z dynastii Zhou, rządzącemu w latach 977-922 p.n.e., maszynę poruszającą się i zachowującą jak człowiek. Władca obrażony jej zachowaniem wydał rozkaz egzekucji i wszyscy członkowie dworu mogli zobaczyć, że w istocie był to byt składający się wyłącznie ze skóry, drewna, kleju i lakieru (Needham, 1986).

Od tego czasu podobne historie pojawiają się mniej lub bardziej regularnie w historii ludzkości. Szczególnie popularne były one w antycznej Grecji. Już wtedy krążyły

bowiem mity, opowieści czy plotki o maszynach myślących lub po prostu zachowujących się jak ludzie. Prawdopodobnie najslynniejszą jest ta o Talosie, przytoczona przez Apollodorosa z Aten. Ta gigantyczna machina krocząca miała bronić wyspy Krety, codziennie okrążając ją trzykrotnie i rzucając głazami w napastników (Kopaliński, 2003).

Russel i Norvig wymieniają również inne źródła inspiracji dla sztucznej inteligencji, a w zasadzie konkretne wynalazki i silnie umocowane w rzeczywistości poglądy. Przywołują choćby ideę Ramona Lulla, według której wnioskowanie mogłoby być wykonywane przez materialny, mechaniczny artefakt, ale wymieniają w tym kontekście również projekty Leonarda Da Vinci, Pascalinę Pascala oraz kalkulatory Leibniza. Należy zwrócić uwagę, że niektórzy już za czasów tego ostatniego obawiali się, że maszyny mogą nie tylko wykonywać obliczenia, ale też myśleć, a wręcz działać autonomicznie. Za przykład można potraktować Thomasa Hobbesa, który w swojej książce *Lewiatan* postulował możliwość istnienia „sztucznego zwierzęcia” (org. „artificial animal”; Hobbes, 1651).

Należy również zwrócić uwagę, że na koncepcję myślącej maszyny możemy się natknąć w większości kultur. Za przykłady mogą posłużyć: żydowski mit golema, ulepionej z gliny postaci, która mogła zostać obdarzona życiem dzięki karteczce z imieniem Boga w ustach (Blau, Jacobs i Eisenstein, 2006), muzułmańskie koncepcje alchemiczne, zwłaszcza homunkulusy (O'Connor, 1994), czy, nieco bardziej współcześnie, Frankenstein z 1818 roku pióra Mary Shelley.

Również dzisiaj nasza (pop-) kultura jest przesiąknięta treściami przedstawiającymi postaci, które możemy określić mianem syntetycznych. Możemy dostrzec przejawy tego w szerokiej gamie dzieł kulturowych opartych nie tyle na tworzeniu relacji czy kreatywnym wykorzystaniu możliwości oferowanych przez maszyny, ile na jawnej rywalizacji, jeśli nie wręcz walce z nimi. Doktor Piotr Bilski przytacza w wywiadzie szereg przykładów, w których syntetyczni agenci stanowią w nich efektywnie zagrożenie dla ludzi: HAL 9000 w *Odysei Kosmicznej*, hosty w *Westworld*, Skynet w *Terminatorze*, CLU w *Tron: Dziedzictwo*, replikanci w *Łowcy androidów*, maszyny w *Matriksie* czy roboty w *Ja, Robot* (Gajewski, 2018).

1.2. Konsekwencje realizowania marzeń

1.2.1. Rywalizacja z maszynami

Jednakże warto odnotować, że przez te blisko 60 lat rozwojowi sztucznej inteligencji towarzyszył nam właśnie cel tworzenia wirtualnych agentów, którzy będą myśleć i zachowywać się jak ludzie. Poskutkowało to wytworzeniem szeregu negatywnych skojarzeń tej dziedziny. Zaczynają się one od technofobicznych postaw głoszących, że maszyny odbiorą nam pracę, aż do postrzegania rozwoju tej technologii w kategorii zagrożenia egzystencjalnego. Wynika to z odwoływania się do naszych pierwotnych lęków związanych z rywalizacją o zasoby. Ujmując tę pozycję inaczej, myślimy, że jeśli maszyny są lub będą w stanie myśleć, jak my to będą realizować te same zadania co my. Stąd też zakładamy, że prawdopodobnie będą pożądać tych samych zasobów, przez co będą prędzej lub później stanowić w pewnym momencie naszych rywali.

1.2.2. Technologiczna osobliwość

W skrajnych ujęciach sztucznej inteligencji mówi się wręcz o zagrożeniu zgoła egzystencjalnym dla całej ludzkości. Reprezentanci tego stanowiska obejmują tak znamienite i znane postaci jak Stephen Hawking (Cellan-Jones, 2014), Elon Musk (Gibbs, 2012) czy Nick Bostrom (Bostrom, 2014). Warto odnotować, że dwaj ostatni uczestniczyli w powołaniu specjalnych instytutów, kolejno *Open AI* oraz *Future of Humanity Institute*. Ich celem jest zapewnienie ludziom pewnej dozy kontroli nad powstającą sztuczną inteligencją, a szczególnie ochronienie nas przed osiągnięciem tak zwanej „technologicznej osobliwości”.

Ta ostatnia, czasami nazywana też po prostu „osobliwością”, stanowi stan technologii, w ramach którego stworzymy sztuczną inteligencję, której inteligencja będzie rosła wykładniczo. Przyjmuje się w tym ujęciu, że po osiągnięciu przez maszynę poziomu inteligencji człowieka oraz zachowaniu tego samego tempa, stanie się ona wkrótce niemożliwa do kontrolowania i ludzkość będzie zdana na jej łaskę.

Obawy te są rozważane również ściśle akademicko, przykładowo przez Stuarta Russella (Russell S., 2019). Jednakże brak jest jednocześnie ścisłych przesłanek ku uznaniu takiej trajektorii rozwoju. Wśród czołowych oponentów możemy wymienić, chociażby Paula Allena, Jeffa Hawkinsa, Johna Hollanda, Jarona

Laniera oraz Gordona Moore'a (IEEE SPECTRUM, 2008; Allen i Greaves, 2011; IEEE SPECTRUM, 2008).

1.2.3. Krytyka współczesnej sztucznej inteligencji

Warto odnotować, że wciąż wpływowi założyciele dziedziny sztucznej inteligencji są niezadowoleni z obranego współcześnie kierunku jej rozwoju. Obejmuje to, ale nie ogranicza się do takich postaci jak John McCarthy (2007), Marvin Minsky (2007) czy Nils Nilsson (1995; 2005).

Uważają oni obecne cele za zbyt mało ambitne, a nacisk kładziony na poprawianie aplikacji za źle ulokowany. Ich zdaniem powinniśmy wrócić do tworzenia maszyn, jak określił je Simon przeszło pół wieku temu: myślących, uczących się oraz tworzących (org. „machines that think, that learn and that create”; Simon, 1957). Badacze ci używają określenia sztucznej inteligencji funkcjonującej na poziomie człowieka (org. human-level AI, HLAI) jako celu, do którego powinniśmy dążyć. Warto odnotować, że odbyło się już pierwsze sympozjum w tej tematyce (Minsky, Singh i Sloman, 2004), a kolejne odbywają się regularnie.

Podobnym nurtem do HLAI jest domena ogólnej sztucznej inteligencji (org. Artificial General Intelligence, AGI; Goertzel i Pennachin, 2007). Od 2008 roku ma ona własne czasopismo *Journal of Artificial General Intelligence*, jak również organizowane są konferencje i sympozja w tej tematyce. Celem tejże domeny jest stworzenie uniwersalnego algorytmu zdolnego do uczenia się czegokolwiek i działania w jakimkolwiek środowisku. Jednym z kluczowych źródeł jej inspiracji jest praca Raya Solomonoffa (Solomonoff, 1964a; Solomonoff, 1964b), uczestnika oryginalnej konferencji w Dartmouth, w 1956 roku.

1.3. Dostępna alternatywa

1.3.1. Współpraca zamiast rywalizacji

Jak więc możemy stwierdzić – opisane powyżej stanowisko i kierunek rozwoju sztucznej inteligencji budzą nie tylko wiele obaw, ale prowadzą niekiedy do sprzeciwu i chęci blokowania jej rozwoju. A to właśnie niechęć ludzi stanowi główny inhibitor rozwoju technologii, a nie nasze możliwości techniczne sensu stricto (Schumpeter, 1962).

Istnieje jednak podejście alternatywne, również pochodzące przynajmniej z lat 60. XX wieku. W tym kontekście najczęściej przywoływana jest postać Josepha Carla Licklida. W 1960 roku przedstawił on swoją koncepcję w artykule *Man-Computer*

Symbiosis, doskonale egzemplifikującą to odmienne podejście do sztucznej inteligencji. Zakładała ona, że ludzie i maszyny, nawet nie tyle będą ze sobą współpracować, ile wspólnie wytworzą silnie sprzężony ekosystem wzajemnych powiązań. Dzięki temu powstanie nowa jakość pracy. Jak określił to autor, tak powstałe środowisko będzie się składać z mózgów, które będą myślały jak żadne mózgi wcześniej i maszyn, które będą przeprowadzały obliczenia jak żadne inne maszyny liczące wcześniej (Licklider, 1960).

1.3.2. Symbioza ludzi i maszyn

Uściślając, Licklider przyjmuje stanowisko, według którego ludzie będą wyznaczać cele, określać hipotezy, kryteria ewaluacji i ostatecznie oceniać wyniki. Tymczasem maszyny przejmą zadania możliwe do zautomatyzowania, tak by dostarczyć ludziom możliwość poszukiwania wiedzy w danych i generowanie „wglądu” (org. *insight*), jak również, ostatecznie, podejmowanie wspomnianych decyzji (Licklider, 1960, s. 1). W dalszej części artykułu stwierdzał on, że kolejnym krokiem będzie traktowanie komputerów na poziomie współpracowników, których kompetencje efektywnie wspierają i uzupełniają nasze (Licklider, 1960, s. 4). W przeciwieństwie jednak do ich ludzkich odpowiedników nie będzie im przeszkadzał podrzędny status w tej relacji (Licklider, 1960, s. 7). Licklider uważał za największe wyzwanie w tym obszarze komunikację między podmiotami.

Praca z językiem maszynowym była wtedy bowiem oparta przede wszystkim na syntaktyce, a w komunikacji ludzkiej równie istotna, jeśli nie ważniejsza, pozostaje warstwa semantyczna. Žáčková stwierdza, że w 2011 roku problem ten został rozwiązany poprzez rozwój języków programowania i specjalnych interfejsów ułatwiających pracę z kodem komputerowym (Žáčková, 2011).

Teraz, przeszło dekadę później, poszliśmy kilka kroków dalej i mamy, chociażby oprogramowanie do tworzenia raportów biznesowych Tableau, które ma specjalną funkcję *Ask Data*. Użytkownicy mogą dzięki niej wpisać pytanie w języku naturalnym, na przykład: „Które produkty sprzedawały się najczęściej w zeszłym roku, w północnej Polsce?”. Następnie jest ono transformowane w zapytanie maszynowe, które zwraca uzyskane wyniki w formie gotowych i interaktywnych wizualizacji danych (Tableau, 2022).

2. Motywacja i metody badawcze

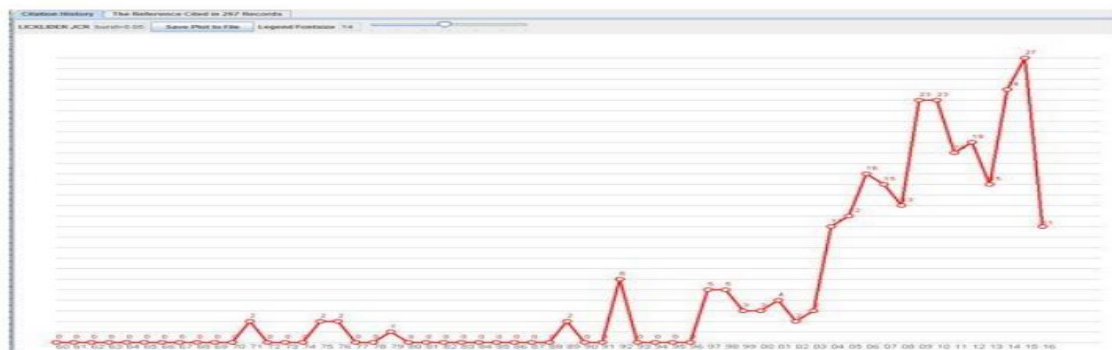
2.1. Implementowanie starych pomysłów nowoczesnymi metodami

Motorem napędowym dla rosnącej popularności sztucznej inteligencji jest szeroko obserwowany wzrost znaczenia danych i pracy z nimi. Dzięki dynamicznemu rozwojowi technologii i operacjonalizacji szeregu aktywności stało się możliwym digitalizowanie kolejnych aspektów naszego życia. Jesteśmy w stanie coraz lepiej cyfrowo opisywać nasze zachowania, kwantyfikować je i wirtualizować. Co więcej, nasze życia są coraz silniej splecione z rozwiązaniami technicznymi, Internetem, aplikacjami, smartphonami i technologiami ubieralnymi, czyli takimi, które możemy na siebie nałożyć, jak smartwatche czy ubrania z rozwiązaniami opartymi na Internecie Rzeczy, za przykład których może posłużyć kurtka zbierająca dane o temperaturze ciała.

Zjawisko to posunęło się do tego stopnia, że już w 2006 roku Clive Humby określił dane nową ropą naftową (Charles, 2013), a obecnie mówimy wręcz o tworzeniu tak zwanego „cyfrowego bliźniaka” (ang. Digital twin) naszej rzeczywistości (Hatałska, 2021). Koncepcja ta została w pierw opracowana w 1988 roku, lecz dopiero teraz jest możliwa do zrealizowania dzięki współczesnym technologiom (jak big data) lub takim, które wkrótce określimy będziemy mogli określić tym mianem (jak „Edge computing” czy „sieć 5G”).

Techniczna możliwość implementacji rozwiązań wymyślonych już dawno temu jest zresztą zjawiskiem typowym dla naszych czasów. Kluczową obserwacją w kontekście tej dysertacji jest wniosek Shyama Sankara, który w swoim wystąpieniu powiedział, wreszcie możemy odejść od koncepcji sztucznej inteligencji Minsky’ego i zrealizować ideę Licklidera (Sankar, 2013). Porównanie tych koncepcji stanowi główny cel części II rozprawy.

Dynamiczny wzrost popularności idei wspomnianego Licklidera został dowiedziony przez brazylijskich autorów: José Reinaldo Silva, Nilberto Machado de Sá, Walter Teixeira Lima Junior i podsumowany wykresem poniżej. Dzięki statystycznej analizie danych bibliometrycznych wykazali, że cytowania Licklidera drastycznie wzrosły od lat 90. XX wieku (Silva, de Sá i Lima Jr, 2017).



Rysunek 1 Ewolucja cytatów Licklidera (Silva, de Sá i Lima Jr, 2017)

2.2. Główne postawy wobec sztucznej inteligencji

Jak wspomniano wyżej, głównym celem części II rozprawy jest porównanie dwóch podejść wobec sztucznej inteligencji: autonomicznych maszyn i symbiozy z nimi. W ramach rozprawy postawy te egzemplifikowane są przez czołowych reprezentantów. Samodzielne maszyny zostają przeanalizowane na podstawie publikacji Marvinina Minsky’ego, książki *Society of Mind*. Drugie podejście, przez wzgląd na mniejszą skalę popularności wymagają sięgnięcia po prace aż trzech autorów: Rossa Ashby’ego, Josepha Carla Robnetta Licklidera oraz Douglasa Carla Engelbarta.

2.2.1. Niezależne maszyny

Postawa postulująca za cel sztucznej inteligencji stworzenie niezależnego bytu najlepiej podsumowana jest cytatem: „(Sztuczna inteligencja) jest nauką, która postawiła sobie za cel stworzenie maszyn zdolnych do wykonywania zadań, które, wykonywane przez człowieka, wymagają wykorzystania inteligencji” (org. „the science of making machines do things that would require intelligence if done by men”; Minsky, 1968).

Jej przedstawiciele dążyli więc do wyręczenia lub zastąpienia ludzi w zadaniach umysłowych. Wizja ta towarzyszyła nam już od warsztatów w Dartmouth, w 1956 roku. Minsky stanowi sztandarowy i często przytaczany przykład tego podejścia. Swoje spojrzenie na ten temat opisał najpełniej 30 lat po wspomnianych warsztatach. Skorzystał wtedy z metafory tak zwanego społeczeństwa umysłu (org. *Society of Mind*), którą to przedstawił w książce o tym samym tytule (Minsky, 1986). Następnie była ona rozwijana przez szereg badaczy, których obecnie zbiorczo określamy zwolennikami stworzenia Silnej Sztucznej Inteligencji (org. AGI – Artificial General Intelligence).

2.2.2. Współpraca z maszynami

Koncepcja symbiozy ma, jak wspomniano wyżej, zupełnie inny cel. Zamiast replikować inteligencję, jej głównym postulatem było myślenie o naszych, ludzkich celach i rozbijaniu ich na zadania tak dla maszyn, jak i ludzi. W ten sposób samoistnie wytwarza się swoisty ekosystem wzajemnie powiązanych relacji. Często przywoływanym w tym kontekście przykładem jest, wspomniany wcześniej, artykuł *Man-Computer Symbiosis* Licklida (1960), ale również *An Introduction to Cybernetics* (1957) Rossa Ashby'ego oraz *Augmenting Human Intellect: A Conceptual Framework* (1962) Douglasa Carla Engelbarta.

2.2.3. Percepcja obu koncepcji w momencie ich powstania

Zajmujący w tym kontekście był odbiór tychże koncepcji w momencie ich powstania, w latach 60. XX wieku. Badacze określali wtedy ideę zbudowania myślącego, syntetycznego bytu jako łatwiejszą do osiągnięcia od pomysłów J.C.R. Licklida i pozostałych reprezentantów alternatywnego podejścia. Stworzenie takowej maszyny stale zresztą jawi się nam, jakby było możliwe już za chwilę, jakby było dostępne na wyciągnięcie ręki. Uczni już od lat 40. XX wieku przewidują, że osiągnięcie tego celu nastąpi za kolejne dwie dekady i ciągle, praktycznie rokrocznie, przesuwają tę spodziewaną datę, ale wciąż towarzyszą nam te magiczne dwie dekady wyznaczane jako data, kiedy w końcu zbudujemy myślące maszyny (Bostrom, 2016, s. 21).

Możemy to lepiej zrozumieć, gdy zauważymy, że wizja Licklida nie tylko zakładała dzielenie się zadaniami z maszynami, ale też, a może przede wszystkim, ułatwienie kontaktu z nimi. Pragnął on bowiem stworzyć stanowisko pracy oferujące ludziom bardziej naturalne i wygodniejsze wchodzenie w interakcję z maszyną zarówno poprzez automatyczne rozpoznawanie głosu, jak i wydawanie komend za pomocą pisania długopisem.

Było to jednak w czasach, gdy komputery normalnie miały rozmiar pokoju, a wspomniane interakcje opierały się na wprowadzaniu dziurkowanych kart jako „programów” i odczytywaniu wyników wydrukowanych na długich, wąskich karteczkach. Operowano wtedy na maszynach pozbawionych interfejsów graficznych, programów umożliwiających pracę z tekstem jak współczesny Word, nie było płyt cd, dvd, przenośnych dysków, w zasadzie nie było dysków twardych jako takich. Najnowszym wynalazkiem w 1960, awangardą tak zwanych minikomputerów, był system PDP1, który zawierał wbudowany wyświetlacz graficzny z lampą katodową,

wejście/wyjście taśmy papierowej, a jego reklamy podkreślały, że nie wymagał specjalnej klimatyzacji i mógł być obsługiwany przez tylko jednego operatora komputera.

2.3. Luka badawcza

2.3.1. Symbioza jako obecnie dominujące podejście

Co zaskakujące, większość współczesnych implementacji rozwiązań uczenia maszynowego (czyli realnej implementacji sztucznej inteligencji) podąża raczej za tym drugim nurtem, zgodnym z wizją Licklidersa, a percepcja sztucznej inteligencji zatrzymała się na nurcie pierwszym, Minsky'ego. Ilustruje to, chociażby artykuł Kazimierza Trzęsickiego *Idea Sztucznej Inteligencji* (Trzęsicki, 2020). Za ciekawy kontrprzykład może posłużyć opracowanie Aleksandry Kuzior i Waldemara Czajkowskiego *Filozofia umysłu a sztuczna inteligencja*. W jego ramach autorzy posługują się przez prawie cały tekst pierwszą definicją, by w końcowej nocie odwołać się do symbiotycznej czy, jak określają to wspomniani filozofowie, kolektywnej inteligencji (Kuzior i Czajkowski, 2019).

2.3.2. Superumysły Malone'a

Praktycznie jednak odchodzimy od nurtu „samodzielnej” sztucznej inteligencji. Na poziomie akademickim powoływane są już specjalistyczne instytuty badające możliwe alternatywy. Wyraźnym, pozytywnym przykładem jest Centrum Inteligencji Zbiorowej na MIT (org. *MIT Center for Collective Intelligence*). Został on niejako zapoczątkowany badaniami obecnego dyrektora – Thomasa W. Mallone'a i wynikającą z nich książką *Superminds* (Malone, 2018). Co zajmujące nie dość, że połączył on obie koncepcje sztucznej inteligencji, to zrobił to niejako przypadkiem. Wspomniana książka opisuje ideę „superumysłów” złożonych z wielu bytów, tak ludzi, jak i maszyn, którzy razem pracują nad wspólnym celem. Nasuwa się więc zarówno interpretacja Minsky'ego, gdzie pojedynczy agenci są sprzęgnięci różnymi relacjami w ramach jednego umysłu, jak i wizja Licklidersa, wedle której ludzie i maszyny współpracują w realizowaniu zadań.

Naturalnie, nie jest to kompletne ujęcie tych hipotez i warto odnotować zachodzące różnice między koncepcją Malone a intencjami autorów. Ujęcie Minsky'ego zostaje tu rozszerzone – nie mówimy już wyłącznie o agentach w obrębie jednego bytu (czy to człowieka, czy też agenta sztucznej inteligencji), a raczej pewnym, jeszcze bardziej abstrakcyjnym „superbycie”, „superumysłem”. Natomiast wbrew wizji Licklidersa uczestnicy

procesu zostają zrównani jako współtworzący jeden super umysł, bez wyróżniania hierarchii stwierdzającej kto podejmuje jakie decyzje.

2.3.3. Brakujący środek

Przejawy wizji wspierającej sztucznej inteligencji możemy dostrzec również w środowisku biznesowym. Podobnie jak z książką Malone'a robią to niejako przypadkiem, a przynajmniej wszystko wskazuje, że nieświadomie. Na szczególną uwagę zasługuje tutaj książka *Human + Machine: Reimagining Work in the Age of AI* autorstwa Paula Daugherty'ego oraz Jamesa Wilsona (2018). Powstała ona w odpowiedzi na coraz częstsze i coraz bardziej alarmujące raporty o spodziewanych redukcjach etatów spowodowanych przez postępującą automatyzację pracy. Autorzy konstatują, że część zawodów prawdopodobnie zostanie w istocie oddelegowana maszynom, aczkolwiek dużo większy wpływ na pracę ogółu ludzi, a zwłaszcza sposobu jej wykonywania, będzie miała sztuczna inteligencja uzupełniająca i wzmacniająca ludzi.

Za przykład podają oni procesy przeciwdziałania praniu pieniędzy, gdzie systemy sztucznej inteligencji przejmują na siebie żmudne i rutynowe zadania związane z przetwarzaniem zgłoszeń, a odciążeni ludzie mogą dzięki temu skupić się na trudniejszych, mniej jednoznacznych przypadkach i realnym wsparciu klientów (Daugherty i Wilson, 2018, s. 46). Autorzy wymieniają tu trzy rodzaje aktywności: tylko ludzką (obejmującą takie działania jak przywództwo, empatia, tworzenie idei, wydawanie sądów), tylko maszynową (realizowanie transakcji, powtarzanie operacji, tworzenie modeli predykcyjnych, dostosowywanie danych), jak i mieszaną, hybrydową, podzieloną na dwa obszary: pierwszy, w którym to ludzie wzmacniają maszyny (trenowanie syntetycznych agentów, wyjaśnianie zdarzeń, podtrzymywanie usług), drugi, gdzie sztuczna inteligencja, jak określają to autorzy, daje ludziom „supermoce” (wzmacnianie inteligencji, wchodzenie w interakcję, ucieleśnianie; Daugherty i Wilson, 2018, s. 8). Tę ostatnią, mieszaną nazywają oni „brakującym środkiem” (org, missing middle, tłumaczenie własne). Określili ją tak ze względu na niedostateczną reprezentację takiego podejścia w obecnych rozwiązaniach biznesowych.

2.3.4. Wyzwania edukacyjne

Efektywne wykorzystanie sztucznej inteligencji stwarza więc wyzwanie edukacyjne. Nie tylko jak przebranżowić ludzi, którzy faktycznie będą tracić pracę w związku ze sztuczną inteligencją, ale też jak przygotować pozostałych

do zredefiniowania ich ról. To wyzwanie nie dotknie jednak wyłącznie państw (obecnych, głównych dostawców edukacji), ale również, a może przede wszystkim, firm, które, by utrzymać konkurencyjność, będą musiały nadążyć za tymi błyskawicznymi zmianami.

Warto odnotować, że już powstają oddolne inicjatywy, które przygotowują ludzi do efektywnego wykorzystania tej technologii. Za przykład można przywołać choćby kurs Uniwersytetu Helsińskiego *Elements of AI*. Instruktaż został przetłumaczony na ponad 30 języków. 750,000 studentów ze 170 krajów wzięło w nim udział i w ten sposób zdobyło elementarne podstawy zrozumienia systemów sztucznej inteligencji (Uniwersytet Helsiński, 2018).

Z drugiej strony, ten okres przejściowy będzie zdecydowanie stanowił wyzwanie, jednak w dłuższej perspektywie powinniśmy dostrzec pozytywy tej sytuacji, gdyż wprowadzone zmiany pozwolą nam na zmiętygowanie takich trudności jak kurcząca się siła robocza na skutek starzejącego się społeczeństwa (Dizkes, 2021).

2.3.5. Stan polskiej nauki w kontekście wzmacniania inteligencji

Polskie badania nad wzmacnianą ludzi sztuczną inteligencją mają tendencję do popadania z jednej skrajności w drugą. Z jednej strony realizowane są badania, czy wręcz całe programy badawcze, zorientowane bardzo fragmentarycznie, na pewien ściśle określony wycinek tematu współpracy ludzi i maszyn.

Za przykład możemy przyjąć badania realizowane przez warszawskich badaczy. Dynamicznie rozwijają oni nurt, tak zwanej, „wyjaśnialnej sztucznej inteligencji” (org. *explainable artificial intelligence*). Ich przełomowe badania znacząco rozwijają tę dziedzinę. Jednakże skupiają się one na konkretnych implementacjach, w określonych branżach, z zastosowaniem ściśle zawężonych rozwiązań technologicznych, jak, na przykład, uniknięcie „problemu czarnej skrzynki” w projektowaniu modeli predykcyjnych zastosowanych do tworzenia scoringu kredytowego (Biecek i inni, 2021).

Drugi koniec spektrum reprezentują prace bardzo szeroko zapatrujące się na ten temat. Możemy to dostrzec w działalności badawczej Artura Guni, zwłaszcza w jego rozprawie doktorskiej. Dostrzega on szereg wzmocnień. Wyróżnia różne metody w zależności od jakich środków chcemy użyć: biotechnologiczne, nefarmakologiczne czy informatyczne, jak również rozpatruje cel, gdyż wzmocnieniu może ulec nie tylko nasza inteligencja, ale również zmysły czy emocje. W jego interpretacji

wzmacnianie inteligencji zdefiniowane jak to zostało powyżej dokonane, stanowi podzbiór wzmocnienia poznawczego za pomocą metod informatycznych (Gunia, 2015).

2.4. Cel rozprawy

Celem rozprawy jest ujednoczenie istniejącej terminologii, porównanie dwóch kluczowych wizji sztucznej inteligencji (Minsky'ego oraz Licklidera) i ich współczesnych interpretacji, a następnie stworzenie zbiorczego modelu implementacji rozwiązań sztucznej inteligencji, które efektywnie wspierają ludzi w podejmowaniu decyzji. Tak więc ma ona przede wszystkim charakter poznawczo-teoretyczny. Możemy wyrazić cel rozprawy poprzez pytanie: czy sztuczna inteligencja jest w stanie usprawnić nasze procesy decyzyjne? Przyjmując hipotezę, że tak, następuje pogłębienie odpowiedzi na to pytanie poprzez próbę opracowania modelu odpowiadającego, jak to osiągnąć.

Dzięki analizie tekstów zastanych (tzw. desk research) możliwe jest określenie poziomu świadomości zarówno badaczy, jak i praktyków. Podobnie, możemy ustalić jaką wizję sztucznej inteligencji obecnie implementujemy z perspektywy historycznej i społecznej. Przedmiotem tejże ewaluacji jest określenie takich aspektów jak:

- Ogólna wiedza na temat sztucznej inteligencji,
- Wpływ Licklidera i innych badaczy, którzy zapoczątkowali nurt wzmacniania inteligencji, na obecnie stanowione cele biznesowe i badawcze,
- Usytuowanie ludzi w projektach opartych na sztucznej inteligencji,
- Przejawy implementacji wzmacniania inteligencji w szeroko dostępnych na rynku narzędziach business intelligence.

Jak wspomniano wyżej, celem rozprawy jest nie tylko opis stanu zastanego, ale również opracowanie modelu umożliwiającego faktyczną implementację rozwiązań sztucznej inteligencji wspierających ludzi w procesach decyzyjnych.

2.5. Metoda

W pracy przede wszystkim wykorzystywana jest analiza krytyczna oraz porównawcza tekstów źródłowych. Można również dostrzec przejawy metod: źródłowej, analitycznej oraz syntetycznej, bez pominięcia metody porównawczej.

Dysertacja ma charakter teoretyczny. W ramach desk research przedstawione zostają główne, współczesne stanowiska oraz niezbędne do ich wprowadzenia zaplecza technologiczne. Uzasadnienie ich wyboru jest przedstawione w adekwatnych rozdziałach.

Po zbudowaniu niezbędnego zaplecza teoretycznego opracowany zostaje własny model sztucznej inteligencji wspierającej podejmowanie decyzji. Praca ma charakter interdyscyplinarny, łącząc dziedziny nauk humanistycznych, społecznych oraz ścisłych i przyrodniczych.

2.6. Struktura pracy

Dysertacja składa się ze wstępu, podsumowania i trzech części. W ramach wstępu zawarty został kontekst opisujący obecny stan sztucznej inteligencji i z jakimi wyzwaniem się mierzy. Przedstawione zostaje stanowisko, w ramach którego maszyny będą rywalizować z ludźmi oraz bardziej pragmatyczna alternatywa zakładająca wspieranie ludzi. Wyszczególniona zostaje luka badawcza, wynikający z niej cel badawczy oraz metody jego osiągnięcia.

Pierwsza część rozbudowuje wspomniany w poprzednim akapicie kontekst. Przedstawione zostają dwie historie sztucznej inteligencji: rozwój myśli technicznej, która pozwoliła na stworzenie obecnych rozwiązań oraz ewolucja naszego sposobu myślenia o maszynach, w tym historia myśli filozoficznych dotyczących tego tematu.

W drugiej części porównane zostają dwie koncepcje sztucznej inteligencji, w ramach których mogą być projektowane rozwiązania myślących maszyn: tworzenie niezależnych, syntetycznych bytów lub budowanie rozwiązań usprawniających pracę i funkcjonowanie ludzi, lub całych ich zespołów. Podejścia te są egzemplifikowane odniesieniami do myśli konkretnych autorów. Z jednej strony przeanalizowana zostaje wizja Marvinina Minsky'ego, w ramach której zostaje przedstawiona koncepcja społeczności umysłów. Jest to idea opisująca, jak zbudowane są umysły: czy to ludzi, czy maszyn, które będą myśleć i zachowywać się w ten sam sposób jak oni. Alternatywa nie jest równie kompleksowo przedstawiona i opiera się na poglądach trzech badaczy: Williama Rossa Ashby'ego, Josepha Carla Licklida oraz Douglasa Engelbarta. W ramach dysertacji są oni zbiorczo określani mianem symbiotyków, za kluczowym dla tego podejścia artykułem J.C. R. Licklida, *Man-computer Symbiosis*. Druga część kończy się porównaniem tych koncepcji.

Ostatnia część kontynuuje porównanie poglądów badaczy pragnących tworzyć symbiotyczne relacje z maszynami i autonomicznych agentów. Tym razem zestawione zostają współczesne koncepcje: z jednej strony Thomasa Malone'a, a z drugiej Paula Daugherty'ego wraz z Jamesem Wilsonem. Zasadniczo odeszliśmy już od budowania inteligentnych maszyn jako celu samego w sobie. Zamiast tego rozważamy

tworzenie mieszanych zespołów złożonych z quasi-inteligentnych maszyn i ludzi. Wspomniane koncepcje ilustrują, w jaki sposób możemy zmierzyć się z tym tematem.

T. Malone podkreśla w tym kontekście rolę technologii w łączeniu ludzi oraz rozważa, w jaki sposób współpraca z maszynami może usprawnić określone etapy w zaproponowanym przez siebie ogólnym modelu procesu poznawczego. Ostatecznie, jego celem jest przedstawienie koncepcji superumysłów, hipotetycznych modeli opisujących dowolne byty, od pojedynczych ludzi, przez kawiarnie po całe państwa i korporacje zatrudniające tysiące pracowników.

Alternatywnie, P. Daugherty i J. Wilson skupiają się na proponowanej przez siebie koncepcji brakującego środka. Zauważają oni, że poświęcono już mnóstwo uwagi usprawnianiu ludzi i maszyn, ale obszar stychny pozostaje enigmą. W swojej książce przestrzegają przed dynamicznym rozwojem sztucznej inteligencji i wynikającymi z niego zmianami sposobów, w jaki prowadzimy biznesy. Celem ich publikacji jest zapoznanie odbiorcy z ogólnymi metodami jak implementować sztuczną inteligencję w różnych branżach i na różne sposoby, co ilustrują przykładami konkretnych przedsiębiorstw, jak one przeszły tę drogę. Ostatecznie sugerują oni kluczowe umiejętności i warunki konieczne do zaimplementowania omawianych rozwiązań.

Po omówieniu obu podejść zostaje zaprezentowana dziedzina business intelligence. W jej obrębie bardzo dynamicznie rozwija się podejście zbliżone do teorii symbiotyków. I tak jak J. C. R. Licklider pisał o wzmacnianiu inteligencji, tak w obrębie omawianej dziedziny jednym z dominujących trendów jest wzmacnianie analityki. Obejmuje ono dokładnie to, o czym pisał Licklider, oddelegowaniu możliwie wielu zadań maszynom i umożliwienie w ten sposób człowiekowi zajęcie się zagadnieniami nietrywialnymi. Następnie kroki wzmocnionej analityki zostają porównane z koncepcjami zaproponowanymi przez T. Malone'a, P. Daugherty'ego oraz J. Wilsona, by wykazać ich uniwersalność oraz zbieżność z praktykami biznesowymi.

W ramach zakończenia omówione koncepcje zostają odniesione do teorii procesów podejmowania decyzji, z zamiarem ich technologicznego polepszenia. Omówione zostają zwłaszcza dwa spojrzenia na procesy decyzyjne. Najpierw przytoczono koncepcję Yingxu Wanga i Guenthera Ruhe'a. Jest ona oparta na najnowszych osiągnięciach neuronauk poznawczych i informatyki kognitywnej, traktuje decyzje jako proces poznawczy. Druga omawiana koncepcja, Franka Harrisona, dotyczy podejmowania decyzji strategicznych. Autor traktuje je jako specyficzny rodzaj decyzji w ogóle, gdyż mają one tendencję do kaskadowania. Oznacza to, że gdy zapadną na najwyższym

poziomie, to spływają na kolejnych decydentów, wymuszając na nich podejmowanie kolejnych decyzji. Kolejnym etapem jest przedstawienie własnego modelu decyzyjnego, opartego na wszystkich głównych koncepcjach opracowanych w trakcie dysertacji.

Rozprawa jest zakończona dyskusją, obejmującą przypomnienie pytań badawczych oraz celu rozprawy. Następnie zostają przytoczone główne odkrycia wynikające z dokonywanych w trakcie pracy porównań i opracowań. Zostają one kolejno zinterpretowane w odniesieniu do podstawowych założeń, jak również całości pola badawczego w obszarze wzmocnienia inteligencji. Naturalne rozszerzenie tej myśli stanowi omówienie konsekwencji przeprowadzonych badań, z wyszczególnieniem obszarów, w których możliwe jest praktyczne wykorzystanie uzyskanych wyników.

Na koniec wymienione zostają ograniczenia i bariery, na jakie napotkano w trakcie prac nad dysertacją wraz z omówieniem możliwych kierunków dalszego rozwoju w tym obszarze.

3. Część I – dwie historie sztucznej inteligencji

3.1. Historia myśli technicznej

3.1.1. Od antycznej Grecji do XX wieku

Badacze, jak chociażby Russell i Norvig (2010) czy Przegalińska (2016), doszukują się początków myśli, które doprowadziły do powstania sztucznej inteligencji, w czasach antycznej Grecji. Jednym z pierwszych udokumentowanych wynalazków to potwierdzających jest zegar wodny zbudowany około 25 r. p.n.e. przez Ktesibiosa z Aleksandrii. Był to przełom, gdyż maszyna ta była w stanie utrzymywać stały przepływ wody. Było to zachowanie absolutnie nowe i efektywnie zredefiniowało pojęcie artefaktu.

Otworzyło to temat budowy maszyn, czy też mechanizmów, zdolnych do dostosowywania siebie i swoich zachowań do warunków, w jakich przebywają, do swojego środowiska. Russell i Norvig przywołują w tym kontekście również wynalazek Corneliusa Drebbela, termostat, a następnie silniki parowe Jamesa Watta. Zwracają jednak uwagę na fakt, że matematyczna teoria stabilnego sprzężenia zwrotnego została opracowana dopiero w XIX wieku. Ostatecznie doprowadziły one do powstania nurtu cybernetyki i teorii kontroli, który miał wiele wspólnego ze sztuczną inteligencją. Badacze je zgłębiający nawzajem się inspirowali, znacząco usprawniając swoje twory (Russell i Norvig, 2010, s. 15).

W kontekście jednak znów antycznej Grecji jest tutaj również bardzo istotna rola intelektualnej spuścizny Arystotelesa. Jego myśli nie tyle ukierunkowały, ile w ogóle umożliwiły powstanie matematyki, jak ją dzisiaj rozumiemy. Szczególnie istotne jest w tym kontekście sformułowanie precyzyjnych zestawów praw rządzących racjonalną częścią umysłu oraz nieformalnego systemu sylogizmów umożliwiających poprawne rozumowanie. Efektywnie myśli te pozwoliły na mechaniczne generowanie wniosków wynikających z początkowych przesłanek, co później umożliwiło zrealizowanie koncepcji automatyzowania wnioskowań logicznych (Russell i Norvig, 2010, s. 5).

Arystoteles był według Russella i Norviga kluczowy również dla rozwoju sztucznej inteligencji przez wzgląd na jego rozważania nad naturą myśli i poszukiwaniem odpowiedzi na pytanie, jak to się dzieje, że niektóre z nich prowokują działanie. Jak również nad naturą celów i środków- lekarz nie zastawia się czy leczyć tylko jak leczyć efektywnie. To drugie zresztą określone zostaje przez wspomnianych badaczy mianem algorytmu. Kluczowe dla tej klasyfikacji jest przyjęcie pewnego celu oraz metod,

aby go osiągnąć. Russell i Norvig określają ten algorytm za zaimplementowany dopiero 2300 lat później przez Newella i Simona w systemie GPS, opisanym dokładniej dalej, w podrozdziale 3.1.4 (Russell i Norvig, 2010, s. 7).

Nieco później, w IX wieku, al-Khowarazmi, perski matematyk, po raz pierwszy użył słowa algorytm, w sensie, w jakim posługujemy się nim dzisiaj. To również dzięki jego pismom wprowadzono arabskie cyfry oraz algebrę do Europy (Pickover, 2009, s. 84). Następnym krokiem w rozwoju algorytmów jako takich była propozycja Boole'a. Przedstawił on sposób, jak można wykorzystać algorytmy do operowania na zasadach logicznej dedukcji. Jednakże wciąż nie było wtedy sposobu, aby objąć nią zasad matematycznej indukcji, niezbędnej do opisu liczb naturalnych. Wysiłki matematyków w późnym XIX wieku skupiły się na sformalizowaniu ogólnych matematycznych wzorców, które by to umożliwiły (Russell i Norvig, 2010, s. 8).

Inną, kluczową koncepcją matematyczną było choćby zarysowanie w XVI wieku idei prawdopodobieństwa przez Włocha, Gerolamo Cardano. Jego celem było opisanie w ten sposób wyników zdarzeń hazardowych (Russell i Norvig, 2010, s. 9). Antoine Arnauld poprawnie opisał wzór ilościowy do decydowania jaką akcją podjąć w takich sytuacjach przeszło sto lat później (Russell i Norvig, 2010, s. 6), a, w 1654 roku, Blaise Pascal, w liście do Pierre'a Fermata, wykazał jak przewidzieć przyszłość niezakończonych gry hazardowej i jakie będą oczekiwane zyski uczestników danej rozgrywki. Później James Bernoulli, Pierre Laplace i inni rozwijali tę teorię oraz wprowadzali nowe metody statystyczne. Wreszcie Thomas Bayes, zaproponował zasadę do aktualizowania rozkładów prawdopodobieństwa w oparciu, o napływające dowody modyfikujące szanse zajścia konkretnych zdarzeń. Znamienne jest, że zasada Bayesa jest stosowana w większości współczesnych podejść do wnioskowania w stanie niepewności w systemach sztucznej inteligencji (Russell i Norvig, 2010, s. 9).

Równie istotna dla rozwoju sztucznej inteligencji była logika formalna. Russell i Norvig zauważają, że jej przejawy mogą być zauważone w myślach filozofów z antycznej Grecji, ale jej matematyczny rozwój zaczął się efektywnie pracą George'a Boole'a, który opracował zarówno logikę zdań, jak i logikę opartą na, opracowanej przez siebie samego, funkcji boolowskiej (Boole, 1847). W 1879 roku Gottlob Frege rozwinął logikę Boole'a, aby obejmowała również obiekty i relacje. W ten sposób stworzył logikę pierwszego rzędu, która jest używana również dzisiaj (Adamowicz i Zbierski, 1997). Alfred Tarski przedstawił teorię odniesienia. Opisuje ona, w jaki sposób możemy odnosić obiekty w logice do obiektów w prawdziwym świecie (Tarski, 1941).

Kolejną kluczową dyscypliną dla powstania sztucznej inteligencji, wymienioną przez Russella i Norviga jest ekonomia. Ta zyskała, według nich, status nauki w 1776 roku, kiedy to szkocki filozof Adam Smith opublikował książkę „Badania nad naturą i przyczynami bogactwa narodów”. Mimo że starożytni Grecy i inni udzielali się w myśli ekonomicznej, to Smith jako pierwszy potraktował ekonomię jako naukę (Russell i Norvig, 2010, s. 9).

Wyraził on ideę, wedle której ekonomia i ekonomika mogły być rozważane jako opis działań indywidualnych agentów, którzy maksymalizują swój ekonomiczny dobrostan. Tak, jak większość osób zapytanych o to, na czym skupia się ekonomia, odpowie, że na pieniądzech, tak ekonomiści powiedzą raczej, że studiują sposoby, jak ludzie podejmują decyzje, aby doprowadziły do pożądaných przez nich skutków. Możemy o tym pomyśleć jako o rodzaju handlu wymiennego (Russell i Norvig, 2010, s. 9).

Inną obserwacją Arystotelesa z 335 roku p.n.e., kluczową dla rozwoju neuronauk, a przez to również sztucznej inteligencji, było dostrzeżenie szczególnej roli mózgu i stwierdzenie, że ze wszystkich zwierząt człowiek ma największy mózg proporcjonalnie do swoich rozmiarów. O ile Arystoteles dostrzegł już w tamtych czasach istotną rolę mózgu, tak dopiero w XVIII wieku rozpoznano w tym organie siedzibę świadomości. Wcześniej przypisywano tę rolę sercu i śledzionie (Russell i Norvig, 2010, s. 10). To doniosłe spostrzeżenie stanowiło bodziec do powstania odrębnej dziedziny nauki zajmującej się ściśle badaniami nad mózgiem, tak zwanej neuronauki.

Paul Broca był jednym z najważniejszych badaczy, którzy pierwotnie się nią zajmowali. Studiował on afazję, czyli deficyty mowy, u pacjentów z uszkodzeniami mózgu. W 1861 roku zademonstrował istnienie wyspecjalizowanych obszarów mózgu odpowiedzialnych za konkretne funkcje poznawcze. Szczególnie istotne było wskazanie, że produkcja mowy była zlokalizowana w porcji lewej półkuli, którą dziś nazywamy, od jego nazwiska, ośrodkiem Broci (Russell i Norvig, 2010, s. 10).

W XIX wieku wiedziano już ogólnie, że mózg składa się z komórek nerwowych (neuronów), ale dopiero w 1873 roku Camillo Golgi przedstawił technikę ich markowania, czyli oznaczania. Umożliwił on w ten sposób obserwację aktywacji pojedynczych neuronów. Została ona wykorzystywana przez Santiago Ramona y Cajala w jego odkrywczych badaniami nad strukturami nerwowymi mózgu (Russell i Norvig, 2010, s. 10).

Również w XIX wieku możemy zaobserwować formalny załazek psychologii. Russell i Norvig doszukują się jej naukowych źródeł w pracach niemieckiego fizyka

Hermana von Helmholtza i, jego ucznia, Wilhelma Wundta. Helmholtz zastosował metodę naukową do badania ludzkiej percepcji wzrokowej, a jego podręcznik optyki fizjologicznej (org. *Handbuch der physiologischen Optik*) jest do dziś przedstawiany jako jeden z najważniejszych traktatów o fizyce i fizjologii ludzkiej percepcji (Nalwa, 1994, s. 15).

Wspomniany uczeń, Wundt, otworzył na Uniwersytecie w Lipsku, w 1879 roku pierwsze laboratorium psychologii eksperymentalnej. Jego podejście do eksperymentów było przełomowe. Miały być one ściśle kontrolowane, a przeprowadzane badania opierały się na zadaniach percepcyjnych i skojarzeniowych- opartych zarówno na introspekcji, jak i własnych opisach procesów mentalnych. Podejście to zdecydowanie przyczyniło się przypisaniu psychologii rangi nauki jako takiej. Aczkolwiek, inherentna subiektywność właściwa tej dziedzinie badawczej sprawia, że nawet badaczowi niełatwo jest opracować falsyfikowalne teorie.

3.1.2. XX wiek do konferencji w Dartmouth

XX wiek był pod wieloma względami przełomowy w kontekście rozwoju technologii i, sukcesywnie, sztucznej inteligencji. W kontekście matematyki, w 1931 roku Kurt Gödel pokazał, że istnieją czynniki ograniczające dla dedukcji. Najpierw, w 1930 roku, podczas drugiej konferencji o Epistemologii Nauk Ścisłych (org. *Epistemology of the Exact Sciences*), K. Gödel pokazał, że istnieje efektywna procedura do udowodnienia dowolnego stwierdzenia zawierającego się w pierwszorzędowej logice Fregego i Russella. Tak zwane „Twierdzenie o Niezupełności” zostało następnie dokładnie opisane i opublikowane rok później (Gödel, 1931). Stwierdza ono, że każdy niesprzeczny, lecz rozstrzygalny, system formalny pierwszego rzędu, który spełnia aksjomaty Peano, musi być niezupełny. Innymi słowy, mówi ono, że żaden system formalny nie może wygenerować wszystkich stwierdzeń arytmetyki.

Russel i Norvig postulują dalej, że ten fundamentalny wynik może również zostać zinterpretowany jako pokazujący, że niektóre funkcje na liczbach całkowitych nie mogą zostać reprezentowane przez algorytm, co wskazuje, że nie mogą zostać obliczone. Stanowi to dla nich odniesienie do Alana Turinga, który to właśnie podjął się próby dokładnego scharakteryzowania, które funkcje są obliczalne, czyli, innymi słowy, możliwe do obliczenia. Odpowiedź na to pytanie jest problematyczna, ponieważ obliczalność lub określenie efektywności procedur nie może być sformalizowana jako formalna definicja (Russell i Norvig, 2010, s. 8).

„Teza Churcha Turinga” stwierdza jednak, że maszyna Turinga jest zdolna do obliczenia jakiegokolwiek obliczalnej funkcji (Turing, 1937). Jak zauważają Russel i Norvig jest ona powszechnie akceptowana jako wystarczająca definicja rozwiązująca problem sformalizowania obliczalności (Russell i Norvig, 2010, s. 8). Turing pokazał również, że są takie funkcje, których żadna maszyna Turinga nie będzie w stanie obliczyć. Za przykład może posłużyć odpowiedź na pytanie, czy dany program zwróci odpowiedź na podane dane, czy też będzie działać w nieskończoność.

Inną kluczową myślą z XX wieku było matematyczne sformułowanie preferowanych wyników lub użyteczności. Leon Walras sformalizował je po raz pierwszy, a następnie Frank Ramsay je ulepszył (Ramsey, 1926). Późniejsze poprawki zostały naniesione przez Johna von Neumanna oraz Oskara Morgensterna w ich książce z 1944 roku *The Theory of Games and Economic Behavior* (1944).

Teoria decyzji łączy teorię prawdopodobieństwa wraz z teorią użyteczności i dostarcza formalny opis dla kompletnej struktury dla decyzji podejmowanych w warunkach niepewności (ekonomicznych lub innych). Ma ona zastosowanie w przypadkach, gdzie opis probabilistyczny adekwatnie opisuje środowisko, w jakim podejmowana jest decyzja. Jest to domeną „dużych ekonomii”, w których żaden agent nie musi zwracać uwagi na zachowania innych agentów. Dużo częściej mamy jednak do czynienia z „małymi ekonomiami”, w których to opis sytuacji często wygląda jak gra: akcje jednego gracza (agenta) mogą w znaczący sposób wpłynąć na zyski innego gracza (tak pozytywnie, jak i negatywnie).

Teorię gier jako taką opracowali John von Neuman i Oskar Morgenstern. Zaprezentowali ją w 1944 roku, w pracy pod tytułem *Teoria gier i zachowanie ekonomiczne*, jak zauważają Duncan Luce i Howard Raiffa (1957).

Russel i Norvig zwracają uwagę, że ekonomiści nie zaadresowali jeszcze pytania o to, jak podejmować racjonalne decyzje w sytuacji, gdy zwrot inwestycji nastąpi w przyszłości, czyli nie będzie natychmiastowy, a wręcz będzie wymagał od podmiotu decydującego sekwencji akcji (Russell i Norvig, 2010, s. 9). Kontynuują jednak, że temat ten był poruszany w dziedzinie tak zwanych badań operacyjnych, które to pojawiły się po II wojnie światowej w wyniku prac Brytyjczyków nad optymalizacją instalacji radarowych, a później znalazł zastosowanie w skomplikowanych decyzjach zarządczych. Praca Richarda Bellmana (Bellman, 1957) sformalizowała klasę sekwencyjnych problemów decyzyjnych i nazwała je procesem decyzyjnym Markova (ang. Markov decision process).

Russell i Norvig zwracają uwagę, że sztuczna inteligencja wymaga nie tylko opracowania wewnętrznie koherentnej definicji inteligencji, ale również pewnego środka, który umożliwi jej zrealizowanie. Artefaktem, któremu wyznaczono to zadanie, jest komputer. Autorzy przytaczają inżynierię komputerową jako kolejny niezbędny fundament dla powstania dyscypliny sztucznej inteligencji.

Zwracają oni uwagę, że maszyny, które działały w sposób porównywalny z naszymi współczesnymi komputerami, zostały wynalezione niezależnie, wręcz równocześnie, przez naukowców w trzech krajach zaangażowanych w drugą wojnę światową. Pierwszym działającym komputerem był elektromechaniczny „Heath Robinson” zbudowany w 1940 roku przez zespół Alana Turinga w jednym celu: odszyfrowywania niemieckich wiadomości. Pierwszym nie tylko działającym, ale również programowalnym komputerem był „Z-3”, wynalazek Konrada Zuse’a w 1941 roku. Wynalazł on również liczby zmiennoprzecinkowe i pierwszy wysokopoziomowy język programowania „Plankalkül”. Pierwszy elektroniczny komputer, „ABC”, został złożony przez Johna Atanasoffa i jego studenta Clifforda Berriego między 1940 a 1942 rokiem na Uniwersytecie Stanowym Iowa. Warto zauważyć, że badania Atanasoffa otrzymały niewiele wsparcia czy uwagi. Dopiero „ENIAC”, opracowany jako część sekretnego militarnego projektu na Uniwersytecie Pensylwania przez zespół Johna Mauchly oraz Johna Eckera, okazał się najbardziej wpływowym protoplastą dla współczesnych komputerów.

W latach 40. maszyny liczące powstawały przede wszystkim dla armii, gdyż tylko ona chciała w nie inwestować (Przegalińska, 2016, s. 91). Pierwszym programowalnym komputerem była maszyna IAS. Działała na podstawie opracowanej przez Alana Turinga podstawy teorii maszyn liczących, a zaprojektowana została przez Johna von Neumanna (Przegalińska, 2016, s. 94).

Kolejna dyscyplina, kluczowa dla powstania sztucznej inteligencji, psychologia, została zdominowana przez inspirację biologią. Biolodzy, przez wzgląd na brak dostępu do introspektywnych danych swoich obiektów badawczych, głównie zwierząt, skupiali się na opisach ich zachowań. Jak zauważa H.S. Jennings, dzięki temu ograniczeniu rozwinęli oni obiektywną metodologię (Jennings, 1906). Zastosowanie tegoż punktu widzenia do ludzi zostało pierwotnie zaaplikowane przez Johna Watsona. Był to początek tak zwanego ruchu behawiorystycznego. Odrzucał on definitywnie wszelkie hipotezy i spekulacje na temat procesów mentalnych właśnie przez wzgląd na wspomnianą

inherentną subiektywność i wynikający z niej brak rzetelności w otrzymywanych danych (Russell i Norvig, 2010, s. 12).

Behawioryści, jak wskazuje nazwa, studiowali wyłącznie zachowanie. Rolę naukowców widzieli jako obiektywnych obserwatorów notujących spostrzeżenia opisujące reakcje zwierząt (i ludzi) na ściśle kontrolowane bodźce. Jak zauważają Russell i Norvig, dzięki tym badaniom znacząco pogłębiła się nasza wiedza o szczurach i gołębiach, ale były one mniej użyteczne w odniesieniu do zrozumienia ludzi jako takich (Russell i Norvig, 2010, s. 13).

Psychologia poznawcza jest kolejną dyscypliną podrzędną wobec psychologii, która miała wpływ na rozwój fundamentów sztucznej inteligencji. Postrzega ona mózgi jako narzędzia przetwarzające informacje. Russell i Norvig wskazują na jej początek prace Williama Jamesa i, wspomniane wcześniej, prace Helmholtza, który to również nalegał na uwzględnienie pewnej formy nieświadomego wnioskowania logicznego (Russell i Norvig, 2010, s. 13).

Tenże nurt, określany również czasem kognitywnym, długo pozostawał w cieniu wspomnianego wcześniej behawioryzmu. W pełni rozwinął się on jednak na Wydziale Psychologii Stosowanej w Cambridge (org. Cambridge's Applied Psychology Unit), kierowanej przez Fryderyka Barletta. W rozpowszechnieniu tego nurtu wymiennie pomogła publikacja jego ucznia i następcy Kennetha Craika *O naturze wyjaśnień* (org. *The Nature of Explanation*, tłumaczenie własne; 1943). Forsował w nim przywrócenie zainteresowania i uczynienia ponownie obiektem badań takich własności mentalnych jak przekonania czy cele. Argumentował, że zastosowanie ma tu metafora opisywania ciśnienia i temperatury gazów. Traktujemy bowiem ich pomiary jako naukowe, mimo że molekuly tworzące te gazy nie muszą mieć żadnej z tych cech. K. Craik określił trzy kluczowe kroki w działaniach agentów opartych na wiedzy:

- 1) bodziec musi być przetłumaczony na reprezentację wewnętrzną,
- 2) osiągnięta reprezentacja może być manipulowana za pomocą procesów psychicznych, aby osiągnąć nowe wewnętrzne reprezentacje,
- 3) nowe reprezentacje są tłumaczone ponownie i skutkują dostrzegalną z zewnątrz akcją.

Publikacja K. Craika była dalej rozwijana przez Donalda Broadbenta. Jego książka *Percepcja i komunikacja* (org. *Perception and communication*, tłumaczenie własne) była jedną z pierwszych publikacji przedstawiających modele psychologiczne w formie procesów informacyjnych (Broadbent, 1958).

Równoległe do psychologii poznawczej w Stanach Zjednoczonych nastąpił dynamiczny rozwój modelowania komputerowego, który to doprowadził do powstania, kolejnej dziedziny powiązanej z powyższymi, kognitywistyki. Russell i Norvig za kluczowe w jej powołaniu wymieniają warsztaty z września 1956 roku na MIT.

Na tychże warsztatach George Miller wystąpił z prezentacją dotyczącą ograniczeń poznawczych w kontekście pamięci krótkoterminowej i naszej zdolności zapamiętywania naraz 7 ± 2 elementów (oryginalny tytuł wystąpienia: *The Magical Number Seven*, tłumaczenie własne), Noam Chomsky zwrócił uwagę na językoznawcze aspekty w warsztacie o trzech modelach języka (org. *Three Models of Language*, tłumaczenie własne). Allen Newell i Herbert Simon natomiast opowiedzieli o możliwościach automatyzacji formalnych wnioskowań i dowodzenia twierdzeń matematycznych za pomocą swojego programu „Logiczny Teoretyk” (org. *The Logic Theory Machine*, tłumaczenie własne). Wszystkie wystąpienia na tej konferencji dotyczyły wykorzystania modelowania komputerowego do opisanego pamięci, języka i myślenia logicznego.

Aby opisać współcześnie już powszechną praktykę w tej dyscyplinie Russell i Norvig przytaczają, wtedy przełomowy, cytat Johna Roberta Andersona: „teoria kognitywna powinna być jak program komputerowy” (org. „a cognitive theory should be like a computer program”, tłumaczenie własne; Anderson, 1980). Oznacza to, że każda teoria opisująca jak myślimy, powinna w dokładny sposób przedstawić procesy mentalne oraz ich mechanizmy. Podobnie, zaaplikowanie matematycznych modeli do badań nad systemem nerwowym przez Nicolasa Rashevsky’ego stanowiło milowy krok w zakresie rozwoju neuronauk. Dodatkowo przełomowym wynalazkiem, który efektywnie umożliwił pomiary aktywności mózgowej, był przedstawiony publicznie po raz pierwszy w 1929 roku wynalazek Hansa Bergera: elektroencefalograf (EEG; Russell i Norvig, 2010, s. 10).

W XX wieku żył również Norbert Wiener. Centralna postać dla nurtu, który współcześnie przyjęto określać teorią kontroli (org. control theory). Był on błyskotliwym matematykiem, który pracował dla Bertranda Russella. Interesował się biologicznymi i mechanicznymi systemami kontroli oraz ich powiązaniem z mechanizmami poznawczymi. Wiener i jego współpracownicy: Arturo Rosenbluth oraz Julian Bigelow sprzeciwiali się behawiorystom. Podobnie jak Kenneth Craik, również pracowali nad systemami kontroli w modelach psychologicznych. (Rosenbluth, Wiener i Bigelow, 1943). Ich zrozumienie zachowania oparli na tak zwanych mechanizmach

regulacyjnych. Dążyli do minimalizowania błędów przez dany organizm. Pomyłki rozumieli jako dysproporcję między stanem obecnym a stanem pożądanym.

W późnych latach 40. XX wieku Norbert Wiener, razem z Warrenem McCullochem i Walterem Pittsem oraz Johnem von Neumannem, zorganizowali serię konferencji. Skutkiem tego była wręcz eksplozja popularności modeli poznania opartych na matematycznych i obliczeniowych fundamentach. Książka N. Wienera *Cybernetyka czyli sterowanie i komunikacja w zwierzęciu i maszynie* była bestsellerem i obudziła powszechne pragnienie zrealizowania marzenia o zbudowaniu inteligentnych maszyn (1948).

W podobnym czasie, w Wielkiej Brytanii, narodził się zbliżony ruch. Jego najznamienitszymi przedstawicielami byli William Ross Ashby, ale również Alan Turing czy Walter Grey. Utworzyli oni tak zwany „Ratio Club”. W. R. Ashby podsumował ich stanowisko w książce *Design for a Brain*. Przedstawił w niej ideę, że inteligencja może być stworzona przy użyciu homeostatycznych urządzeń, które, dzięki systemom sprzężeń zwrotnych, będą w stanie osiągnąć stabilne zachowania adaptacyjne (Ashby W. R., 1940).

Przyjęło się nazywać to stanowisko cybernetyką. Stanowi ona dziedzinę nauki, której główny przedmiot badań stanowią systemy sterowania działające na podstawie kontroli przepływu informacji. Jej trzon stanowiły teoria gier i badania operacyjne. Oferowała, jak pisze Aleksandra Przegalińska, kuszącą perspektywę połączenia nauk ścisłych, czerpania ze stabilnej ekonomii oraz podejścia do ludzi jako zasobów wymagających racjonalnego zagospodarowania (Przegalińska, 2016, s. 91). Charakterystyczne było również jej ujęcie komunikacji. Przegalińska przywołuje w tym kontekście Jerzego Bobryka, który stwierdza, że cybernetycy rozumieli komunikację jako efektywnie sterowanie maszynami. Jest ona, w ich ujęciu, możliwa tylko dzięki przetwarzaniu danych, czyli charakterystycznemu dla cybernetyków ujęciu rozumienia procesów (Bobryk, 2004, s. 43).

Do przetwarzania danych sprowadzała się również komunikacja w ujęciu, tak zwanej, teorii informacji. Było to nowe pole badawcze, powołane przez Claude'a Shannona. To on użył tej nazwy po raz pierwszy w swojej pracy *A Mathematical Theory of Cryptography*, w 1945 roku. A już trzy lata później, w kolejnej pracy *A Mathematical Theory of Communication*, przedstawił najważniejsze dla niej zagadnienia związane ze skwantyfikowanym podejściem do komunikacji.

Informacja była w tym ujęciu traktowana jako zbiór możliwych komunikatów, które są wysyłane przez nadawcę. Nadawane były one na wybranym, w różnym stopniu zanieczyszczonym hałasem, kanale komunikacyjnym. Wreszcie, zostawały przechwytywane przez odbiorcę, który znów to musiał te wiadomości odtworzyć, co znów obarczało je niskim prawdopodobieństwem błędnego zrozumienia. Najważniejszym odkryciem Shannona było stwierdzenie o kodowaniu kanałów o różnym stopniu zakłóceń. Potwierdził on bowiem, że szybkość przekazywania informacji jest najwyższej równa przepustowości kanału (Przegalińska, 2016, s. 94).

W 1957 roku opublikowano szeroko dyskutowaną książkę Burrhusa Frederica Skinnera *Zachowania Werbalne* (org. Verbal Behavior). Stanowiła ona bardzo szczegółowy opis nauki języka w ujęciu behawiorystycznym, a autor był już, w momencie publikacji, uznany za jednego z największych ekspertów na tym polu.

Co zaskakujące, równie popularna, jak wspomniana publikacja okazała się jej krytyczna recenzja pióra Noama Chomsky'ego. Russell i Norvig stwierdzają wręcz, że skutkiem tejże recenzji było wydatne ograniczenie zainteresowania behawioryzmem jako takim. Wykazała ona bowiem, że teoria behawiorystów nie adresuje kwestii kreatywności w języku, przykładowo, w kontekście ani dłaczego, ani jak dziecko może rozumieć i tworzyć zdania, których nigdy wcześniej nie słyszało (Russell i Norvig, 2010, s. 16).

W tym samym roku Chomsky opublikował własną książkę, w ramach której opisał swoją teorię struktur syntaktycznych *Syntactic Structures*. Jego modele syntaktyczne odnosiły się do prac hinduskiego językoznawcy z czasów około 350 roku przed naszą erą, Paniniego. Jest ona o tyle precyzyjniejsza w opisie rzeczywistości, że nie tylko mogła wyjaśnić przytoczone powyżej zjawisko tworzenia nowych treści, jak również była na tyle sformalizowana, że mogła być wykorzystana, po minimalnych dostosowaniach, przy programowaniu.

Kluczowa dla rozwoju sztucznej inteligencji była również publikacja Warrena McCulloha i Waltera Pittsa (McCulloch i Pitts, 1943). Połączyli oni trzy obszary: wiedzę o podstawach fizjologii i zasad działania neuronów, formalną analizę logiki propozycjonalnej Russella i Whiteheada oraz teorię obliczalności Turinga.

Umożliwiły im one opracowanie modelu sztucznego neuronu, charakteryzującego się przede wszystkim swoim stanem. Mógł on być włączony lub wyłączony. Dodatkowo miał możliwość dynamicznej zmiany stanu w reakcji na stymulowanie (włączanie), arbitralnie ustalonej liczby, sąsiadujących z nim neuronów. Co więcej,

neurony te mogły zostać wykorzystane do obliczenia dowolnej funkcji obliczeniowej i można je było łączyć w sieci za pomocą odpowiedników bramek logicznych („i”, „lub”, „nie” i im podobne).

McCulloch i Pitts zasugerowali również, że odpowiednio zdefiniowana sieć może się „uczyć”. Dopiero jednak sześć lat później praca Donalda Hebba eksplorowała dostatecznie prostą zasadę aktualizowania siły połączeń między neuronami, abyśmy mogli ją szeroko zastosować (Hebb D., 1949). Warte odnotowania jest, że nazywamy ją właśnie regułą Hebba oraz że jest stosowana nieustannie od tego czasu, również we współczesnych projektach.

W 1950 roku powstał pierwszy komputer oparty na sieciach neuronowych. Został on zaprojektowany przez Marviną Minsky’ego oraz Deana Edmonsa i nazywał się SNARC. Zbudowano go z trzech tysięcy lamp elektronowych oraz mechanizmu z automatycznego pilota bombowca B-24. Był w stanie symulować pracę 40 neuronów. Warto dodać, że Minsky ostatecznie zdobył tytuł doktora dzięki rozprawie na temat uniwersalnych obliczeń realizowanych z wykorzystaniem sieci neuronowych.

Istnieje wiele innych przykładów wczesnych badań, które możemy określić mianem sztucznej inteligencji, zanim powstała ta nazwa. Jednak na uwagę zasługuje szczególnie, jak wskazują Russel i Norvig, wizja Alana Turinga. Już w 1947 roku występował on z wykładami wokół tej tematyki. To właśnie on jako pierwszy przeprowadził cykl wykładów przybliżających ją dla Londyńskiej Społeczności Matematycznej (org. London Mathematical Society). Następnie, podsumował swoje projekty w 1950 roku, w pracy *Maszynaria Obliczeniowa i Inteligencja* (org. Computing Machinery and Intelligence, tłumaczenie własne). W artykule zawarł takie pomysły jak test Turinga, uczenie maszynowe, algorytmy genetyczne czy uczenie ze wzmocnieniem. Ostatecznie, to również on zaproponował ideę, w ramach której nie powinniśmy starać się zbudować od razu sztucznego, dorosłego umysłu, lecz raczej zacząć od dziecięcego. Jak wskazują Russel i Norvig, jego wkład w rozwój sztucznej inteligencji był nieoceniony (Russell i Norvig, 2010, s. 17).

3.1.3. 1956 rok — Warsztaty w Dartmouth College

John McCarthy był naukowcem powiązany z wieloma uczelniami na przestrzeni lat. Pierwotnie pracował dla Princeton, następnie dla Stanfordu, by, ostatecznie, osiąść w Kolegium Dartmouth (org. Dartmouth College). Przekonał on Marviną Minsky’ego Claude’a Shannona oraz Nathaniela Rochesterą, aby pomogli mu zebrać

amerykańskich naukowców zainteresowanych automatami, sieciami neuronowymi czy studium inteligencji jako takiej.

Dzięki ich zaangażowaniu zorganizowano dwumiesięczny warsztat we wspomnianym Kolegium Dartmouth, w Hanover, New Hampshire, latem, od czerwca do sierpnia, 1956 roku. Głównym celem spotkania było stworzenie około dziesięcioosobowej grupy badawczej. Nie mieli oni stanowić nowej, formalnej jednostki. Zamierzeniem organizatorów było jedynie zapoznanie ze sobą badaczy o współdzielonych zainteresowaniach i celach badawczych. Mieli oni nadzieję, że samo to zadziała jak proces katalizy, przez co badacze będą się nawzajem motywowali i inspirowali. Organizatorzy przyjęli założenie, że każdy aspekt uczenia się lub dowolnej innej cechy inteligencji może być opisany tak dokładnie, że będzie można go realizować, czy też symulować, na maszynie. Przekrój podejmowanej problematyki był bardzo szeroki: obejmował on „nauczenie” maszyn używania języka, by formowały abstrakcyjne idee, rozwiązywać problemy wtenczas uznawane za „zarezerwowane” wyłącznie dla ludzi, a nawet umożliwienie maszynom usprawnianie samych siebie. Russell i Norvig określają to wydarzenie mianem „narodzin sztucznej inteligencji” (Russell i Norvig, 2010, s. 17).

Ostatecznie w warsztatach faktycznie wzięło udział dziesięciu naukowców. Poza wspomnianymi wyżej: McCarthyem, Minskym, Shannonem oraz Rochesterem, byli to: Trenchard More z Princeton, Arthur Samuel z IBM i Ray Solomonoff oraz Oliver Selfridge z MIT. Ton rozmowom nadało zwłaszcza dwóch badaczy z Carnegie Tech: Allen Newell i Herbert Simon, gdyż w momencie, gdy pozostali przybyli na warsztaty z ogólnymi pomysłami, ci dwaj mieli już gotowy program “Teoretyk Logiczny” (org. Logic Theorist). Simon twierdził wręcz, że program ten stanowił przyczynek do rozwiązania problemu ciała-umysłu. Wkrótce, po warsztacie, program został wykorzystany do rozwiązania większości twierdzeń z drugiego rozdziału *Principia Mathematica* Russella i Whiteheada. Russell był rzekomo zachwycony, kiedy Simon pokazał mu, że program opracował krótszy dowód na jedno z jego twierdzeń. Warsztat w Dartmouth zgodnie z założonym celem nie doprowadził do żadnych przełomowych odkryć, ale posłużył temu, aby najistotniejsi (w ocenie organizatorów) badacze zajmujący się sztuczną inteligencją poznali się nawzajem. Następne 20 lat rozwoju tej dziedziny było zdominowane przez nich oraz ich uczniów i znajomych z MIT, CMU, Stanford, i IBM.

Patrząc po samych nawet założeniach warsztatu, możemy zrozumieć, dlaczego sztuczna inteligencja wymagała uznania za osobną dyscyplinę naukową

i nie mogła zostać przypisana jako podrzędna wobec teorii kontroli, teorii decyzji lub badaniom operacji.

Kluczowymi wyzwaniem było dla niej odtwarzanie zdolności, które uważane były wtedy za typowo ludzkie. Nie bez znaczenia było również przyjęcie metodologii z nauk informatycznych. Na koniec warto odnotować, że, jak zauważają Russell i Norvig, celem sztucznej inteligencji również, tak jak we wspomnianych wyżej dyscyplinach, było zbudowanie maszyn, które będą funkcjonowały w autonomiczny sposób w skomplikowanych, stale zmieniających się w środowiskach. Doprowadziło to reprezentujących je badaczy do rywalizacji, jak i wzajemnej inspiracji (Russell i Norvig, 2010, s. 18).

3.1.4. Początki sztucznej inteligencji

Wczesne lata sztucznej inteligencji, to jest od warsztatów w Dartmouth do lat 70. XX wieku, były pełne sukcesów. Było to tym bardziej zaskakujące, że, jak zauważają Russell i Norvig, działo się to ledwie kilka lat po tym, jak komputery przestały być postrzegane jako urządzenia przydatne jedynie w prostej arytmetyce. Przyjmowano wtedy, że maszyna nigdy nie będzie umieć zrobić wielu rzeczy. Istniała wręcz, sporządzona przez Turinga, lista aktywności, których maszyna nigdy nie będzie w stanie dokonać. Stanowiła ona dla badaczy sztucznej inteligencji zbiór wyzwań, które sukcesywnie rozwiązywali (Russell i Norvig, 2010, s. 18).

Wczesne sukcesy Herberta Simona i Allena Newella zostały wkrótce zebrane w postaci, wspomnianego wcześniej, programu „General Problem Solver”. W przeciwieństwie do zaprezentowanego podczas warsztatów w Dartmouth „Teoretyka Logicznego” nowy program został zaprojektowany tak, by możliwie, jak najwierniej naśladować reakcje ludzi na zmieniające się warunki i strategie rozwiązywania problemów.

Działał on jednak w ograniczonym zakresie i miał możliwość poruszania się w obrębie ledwie kilkunastu zagadek. Niemniej kolejność i zakres podejmowanych działań przez program do rozwiązywania problemów był bardzo zbliżony do ludzkich. Stąd też, jak kontynuują Russell i Norvig, GPS był prawdopodobnie pierwszym programem, który reprezentował podejście „myśl jak człowiek” (Russell i Norvig, 2010, s. 18).

Tymczasem, w IBM, Nathaniel Rochester i Herbert Gelernter również rozwijali pierwsze programy Sztucznej Inteligencji. Ten drugi opracował tak zwany

„Geometry Theorem Prover”, który był w stanie dowodzić twierdzenia, które, jak określił je Gelernter, wielu adeptów matematyki uważa za dość podchwytliwe (Gelernter, 1959).

Innym istotnym wydarzeniem było udowodnienie, że komputery mogą wyjść ponad wykonywanie bezpośrednich poleceń. Artur Samuel napisał serię programów operujących na grze w warcaby, które były w stanie optymalizować swoje poczynania i niedługo grały lepiej od swojego twórcy. W lutym 1956 wyemitowano nawet program o tym w narodowej telewizji Stanów Zjednoczonych (Russell i Norvig, 2010, s. 19).

W 1958 roku John McCarthy (który po krótkim okresie w Dartmouth przeniósł się do kolejnej uczelni: MIT) wykazał się trzema kluczowymi dla sztucznej inteligencji osiągnięciami. Po pierwsze przedstawił on język LISP, który szybko zajął dominującą pozycję w programowaniu sztucznej inteligencji przez następne 30 lat. Następnie wymiernie wsparł zespół badawczy, który opracował technikę czasowego współdzielenia komputerów, co, do tej pory, stanowiło olbrzymi problem w rozwoju oprogramowania jako takiego.

Ostatnie, ale równie ważne osiągnięcie stanowi artykuł dotyczący programów komputerowych charakteryzujących się ludzkim, zdrowym rozsądkiem (org. programs with common sense, tłumaczenie własne). Przedstawiono w nim hipotetyczny program „Advice Taker”, operujący na podobnych zasadach jak wspomniane programy „Logiczny Teoretyk” czy „General Problem Solver”. Łączyła go metoda wykorzystania istniejącej wiedzy i przeszukiwania przestrzeni problemowej w poszukiwaniu rozwiązań. Co go wyróżniało to ambicja i skala. Program ten miał się bowiem odnosić do ogólnej wiedzy o świecie.

McCarthy przytoczył tutaj przykład wygenerowania drogi do lotniska w oparciu o jedynie kilka prostych wyrażeń aksjomatycznych. Dodatkowo program powinien być w stanie rozbudowywać swoją bazę tychże wyrażeń w trakcie normalnych operacji użytkownika, w ten sposób ustawicznie poszerzając swoje kompetencje bez bycia explicite programowanym ku temu.

„Advice Taker” stanowi więc doskonały przykład centralnych zasad reprezentacji wiedzy i świata oraz manipulacji na nich za pomocą prostych procesów dedukcji i indukcji. Russell i Norvig kończą jego opis stwierdzeniem, że to zaskakujące jak wiele treści tego pomysłu z 1958 roku jest wciąż aktualne i stanowi adekwatny opis obecnych, wciąż nie rozwiązanych, wyzwań (Russell i Norvig, 2010, s. 19).

W 1958 roku Marvin Minsky przeniósł się na MIT. Początkowo nawet współpracował ze wspomnianym McCarthy, jednakże zakończyło się to relatywnie

szybko. Jak zauważają Russell i Norvig, wynikało to z odmiennego rozumienia dyscypliny sztucznej inteligencji. McCarthy bowiem podkreślał istotność reprezentacji czy wnioskowania w obrębie logiki formalnej. Minsky tymczasem skupiał się na doprowadzaniu programów do stanu użyteczności, a jego postawę niektórzy badacze jak, przykładowo, Russell i Norvig określają mianem „antylogicznej” (Russell i Norvig, 2010, s. 20).

Druga połowa XX wieku była świadkiem dynamicznego rozwoju koncepcji rozwiązywalności algorytmów. I tak jak zauważają Russell i Norvig, o ile rozstrzygalność i obliczalność są ważnymi do zrozumienia obliczeniowości, tak idea rozwiązywalności, czyli ustalenia możliwości w ogóle rozwiązania danego zadania, ma nawet większe znaczenie. Autorzy przyjmują roboczą definicję, że problem możemy nazwać nierozwiązywalnym, jeśli czas wymagany do generowania instancji jego rozwiązania rośnie wykładniczo wraz z rozmiarami tejże instancji (Russell i Norvig, 2010, s. 9).

Rozróżnienie między wielomianowym a wykładniczym wzrostem skomplikowania było po raz pierwszy dokonane w połowie lat sześćdziesiątych (Cobham, 1964; Edmonds, 1965). Jest to istotny podział, ponieważ wzrost wykładniczy oznacza, że nawet umiarkowanie duże instancje nie mogą być rozwiązane w rozsądnym czasie. Stąd też, projektując inteligentne rozwiązania, powinniśmy dążyć do podzielenia nierozwiązywalnych problemów na mniejsze, możliwe do rozwiązania.

Jak jednak możemy stwierdzić, że problem nie ma rozwiązania? Metody na to dostarczyła teoria problemów NP-zupełnych, zaproponowana przez Stephena Cooka (1971) oraz Richarda Karpa (1972). Problemy NP-zupełne to problemy NP, czyli problemy decyzyjne, które zamiast wielomianowym redukcji czasu rozwiązania analizują je w czasie logarytmicznym. Cook i Karp wykazali istnienie w ramach tej kategorii dużej klasy kanonicznych problemów przeszukiwania i wnioskowania. Przyjmuje się więc za nimi, że dowolna klasa problemu, która może zostać ostatecznie zredukowana do klasy problemów NP-zupełnych jest najprawdopodobniej nierozwiązywalna.

Jednakże, jak dodają Russell i Norvig, o ile nie zostało dowiedzione w powszechnie akceptowalny sposób, że problemy NP-zupełne są koniecznie nierozwiązywalne, tak większość teoretyków skłania się ku takiemu twierdzeniu. Te wyniki kontrastują z optymizmem, z którym prasa popularna powitała pierwsze komputery i określiła je mianem elektronicznych supermózgów, które będą szybsze niż Einstein (Russell i Norvig, 2010, s. 9).

Pomimo zwiększającej się szybkości komputerów to ostrożne gospodarowanie zasobami będzie kluczowym elementem charakteryzującym systemy inteligentne. Warto odnotować, że prace związane ze sztuczną inteligencją pozwoliły już wyjaśnić, dlaczego niektóre instancje problemów NP-zupełnych uznajemy za trudne (Cheeseman, 1991).

W drugiej połowie XX wieku rozwijało również się podejście, które możemy określić jako spodziewane skomplikowanie podejmowania racjonalnych decyzji. Pionier sztucznej inteligencji, Herbert Simon, otrzymał Nagrodę Nobla w dziedzinie ekonomii w 1978 roku. W ten sposób doceniono jego wcześniejsze prace, które wykazały, że na potrzeby tworzonych modeli sztucznej inteligencji nie ma konieczności poszukiwania najlepszych decyzji, a wystarczą nam te, które są wystarczająco dobre. Jak kontynuują Russell i Norvig stanowi to również lepszy opis faktycznego zachowania człowieka w procesie podejmowania decyzji (Simon H. A., 1947).

3.1.5. Weryfikacja początkowych założeń sztucznej inteligencji

Badacze sztucznej inteligencji mieli od samego początku wielkie ambicje. Ten optymizm podsumowuje często przytaczane stwierdzenie Herberta Simona z 1957 roku. Stwierdził wtedy, że zbudowane zostały już maszyny, które się uczą, myślą, czy tworzą. Co więcej, jak rozwinął Simon, kompetencje maszyn w tych zakresach będą jedynie rosły i, na dodatek, tempo tego rozwoju będzie jedynie przyspieszać, aż ci syntetyczni agenci osiągną poziom kompetencji porównywalny z możliwościami człowieka. Simon nie poprzestał na tym i dodał, że w przeciągu dekady komputer będzie mistrzem szachowym, a któreś skomplikowane twierdzenie matematyczne zostanie udowodnione przez maszynę (Simon H., 1957). Warto odnotować, że obie te predykcje okazały się prawdą, lecz ich spełnienie zajęło przeszło cztery razy więcej czasu.

Taka pewność siebie była charakterystyczna w tym wczesnym okresie sztucznej inteligencji. Z perspektywy czasu widzimy jednak, że opierała się na wątpliwych fundamentach. Badacze mieli wtedy jedynie prototypy, rozwiązania działające tylko na prostych przykładach i zakładali, że łatwo przyjdzie im ekstrapolować je na realistycznie skomplikowane problemy albo szerszy ich zakres.

Z podobnym optymizmem spotykamy się w założeniu, że sam rozwój mocy obliczeniowej rozwiąże obecne problemy, podchodzono do algorytmów genetycznych (czy, jak wtedy je nazywano, ewolucji maszynowej). Są to programy, które wraz z kolejnymi iteracjami symulacji modyfikują swój kod maszynowy na zasadzie

zainspirowanej biologicznymi mutacjami kodu genetycznego (Russell i Norvig, 2010, s. 21). Przegalińska dodaje, że współczesne systemy eksperckie bazują na takich zaawansowanych algorytmach genetycznych i, w swoim przeszukiwaniu alternatywnych przestrzeni rozwiązań problemów, często mogą pracować z podobną efektywnością co eksperci (Przegalińska, 2016, s. 111).

Przez wzgląd na te i podobne założenia badacze sztucznej inteligencji nie podołali szeregowi wyzwań. Pierwsze z nich dotyczyło odcięcia programów od wiedzy o świecie i rozwiązywanie problemów jedynie na poziomie prostych manipulacji syntaktycznych. Za typowego przedstawiciela tego wyzwania można potraktować próby opracowania metod skutecznego tłumaczenia całkowicie maszynowego.

Celem Amerykańskiej Rady Naukowej (org. U.S. National Research Council) było szybsze tłumaczenie rosyjskich artykułów naukowych, zwłaszcza po tym, jak Sowieci wystrzelili Sputnika, w 1957 roku. Wydawało się wtedy, że proste syntaktyczne transformacje oparte na gramatyce rosyjskiej i angielskiej oraz podmienianie słów z elektronicznego słownika wystarczą, aby zachować dokładne znaczenie zdań. Okazało się jednak, że zachowanie znaczenia w tłumaczeniu wymaga rozległej wiedzy domenowej, szczególnie w celu rozwiązywania niejasności.

W 1966 opublikowano raport komitetu doradczego opiniującego stan badań w tej dziedzinie. Autorzy piszą w nim wprost, że nie powstało nawet jedno akceptowalne tłumaczenie maszynowe tekstu akademickiego i nie ma perspektyw blisko terminowych, aby miało się to zmienić. Zamrożono wówczas finansowanie podobnych projektów. Jak stwierdzają Russell i Norvig, dzisiaj tłumaczenie maszynowe jest wciąż nieperfekcyjne, jednakże dostatecznie dobre, by stosować je szeroko w kontekstach technicznych, komercyjnych, a nawet rządowych (Russell i Norvig, 2010, s. 21).

W podobnym duchu, w Wielkiej Brytanii przełomowy dla rozwoju sztucznej inteligencji był raport Lighthill z 1973 roku. Przyczynił się on do decyzji brytyjskiego rządu o zaprzestaniu wspierania i finansowania rozwoju omawianej dziedziny na większości uniwersytetów. Jak odnotowują autorzy, jednym z kluczowych argumentów przemawiających za takim rozstrzygnięciem były opory badaczy przed dostrzeżeniem eksplozji kombinatorycznej (Russell i Norvig, 2010, s. 21).

Drugim typem wyzwania było skomplikowanie problemów i niedoceniecie wpływu ilości czynników, które program musiał rozważyć. Ówczesne projekty sztucznej inteligencji operowały na analizowaniu szeregu kombinacji kroków prowadzących do rozwiązania. Funkcjonowały one w tak zwanych mikro światach, gdzie operowały

na małej ilości obiektów, co przekładało się na ograniczoną ilość akcji, więc tworzone sekwencje kroków były relatywnie krótkie. Badacze zakładali, że opracowanie szybszych jednostek obliczeniowych i większa ilość pamięci rozwiążą ten problem.

Nie było to zupełnie bezzasadne, gdyż, jak zauważają Russell i Norvig, każde pokolenie komputerowego hardware'u przynosiło zwiększenie szybkości oraz pojemności, jak również obniżenie ceny. Wydajność była podwajana co 18 miesięcy aż do 2005 roku, kiedy to problemy z rozprowadzaniem mocy doprowadziły producentów do multiplikowania liczby rdzeni CPU, zamiast szybkości obliczeniowej (Russell i Norvig, 2010, s. 14).

Russell i Norvig kontynuują, stwierdzając, że o ile zmiany w hardware nie są już tak spektakularne, tak obecne oczekiwania zakładają, że w przyszłości wzrost mocy będzie pochodził z masowego paralelizmu procesów obliczeniowych lub współbieżności, czyli osobiwej zbieżności rozwiązań technologicznych z własnościami naszego ściśle biologicznego organu, mózgu (Russell i Norvig, 2010, s. 14).

Za przykład działania, które nie mogły być rozwiązane przez zwiększenie tychże parametrów Russell i Norvig przytaczają rozwój matematycznej rezolucji, czyli automatyzacji dowodzenia twierdzeń matematycznych. Działała ona poprzez tworzenie kolejnych klauzul, aż do zaistnienia sprzeczności. Szybko okazało się, że system ten nie radził sobie nawet z twierdzeniami zawierającymi kilkadziesiąt faktów. Russell i Norvig dodają, że nawet jeśli program ten potrafił znaleźć rozwiązanie, to nie przekładało się na to, że będzie posiadał mechanizm pozwalający stwierdzić czy jest to rozwiązanie właściwe (Russell i Norvig, 2010, s. 21).

Bardzo duże nadzieje na rozwój sztucznej inteligencji niosły również inne technologie. Istotne w tym okresie były zwłaszcza dynamicznie badania Warrena McCullocha i Waltera Pittsa nad tak zwanymi, sieciami neuronowymi. Ich dokładniejszy opis przytoczono w poświęconym im fragmencie rozdziału 3.1.7. Jednakże temat tych badań zainspirował szereg badaczy już w latach 60. XX wieku, w tym Samuela Winograda i Jacka Cowana, którzy obliczyli, jak duża ilość elementów może zostać kolektywnie przedstawiona jako indywidualne pojęcie z odpowiednim wzrostem odporności i równoległości parametrów sieci (Winograd i Cowan, 1963).

Podobnie, zastosowanie dla tej technologii znalazły metody uczenia Donalda Hebba, które następnie zostały dodatkowo usprawnione i wzmocnione przez Berniego Widrowa i nazwane przez niego adalinami (org. adalines) (Widrow i Hoff, 1960; Widrow, 1962). Kolejnym kamyczkiem w rozwoju sieci neuronowych były prace

Franka Rosenblatta nad perceptronami (Rosenblatt, 1962). Twierdzenie o zbieżności perceptronów mówi, że algorytm uczący się może dostosowywać siłę połączeń perceptronu tak, by dostosować się do dowolnego sygnału wejściowego, przy założeniu, że takie dostosowanie istnieje (Block, Knight i Rosenblatt, 1962).

To dynamicznie rozwijające się pole zostało jednak spowolnione poprzez wykazanie ograniczeń sieci neuronowych na fundamentalnym poziomie. Kluczowa była książka *Perceptrons* autorstwa Marvin'a Minsky'ego i Seymoura Paperta. Dowiedli oni w niej, że perceptrony, czyli wspomniane, jedno z najprostszych sieci neuronowych, mogą się nauczyć wszystkiego, co są w stanie zaprezentować. Jednakże problem stanowił fakt wykazany przez naukowców, że nie są one w stanie zaprezentować zbyt wiele. Szczególnie bolesny był przytoczony przez nich przykład perceptronu z dwoma sygnałami na wejściu, który nie mógł być wytrenowany do rozpoznania czy sygnały na wejściu są takie same (Minsky i Papert, 1969).

Ich argumenty dotyczyły wspomnianych najprostszych sieci i nie miały przełożenia na dużo bardziej skomplikowane, wielowarstwowe sieci, które efektywnie były już wtedy regularnie stosowane. Jednakże książka ta była tak wpływowa, że dofinansowanie na rozwój tych technologii zaczęło sukcesywnie maleć, aż praktycznie zostało zamrożone. Przyjmuje się to jako symboliczne zakończenie pierwszego okresu sztucznej inteligencji, odejście od rozbuchanego optymizmu i czas bardziej realistycznego mierzenia siły na zamiary. Warto jednak odnotować pewną ironię losu, że w tym samym roku opublikowano książkę Arthura Brysona i Yu-Chi Ho (Bryson i Ho, 1969). Jej odkrycie i rozpropagowanie nastąpiło po blisko dekadzie, od kiedy się ukazała. Treści w niej zawarte jednak spowodowały wtedy masowy powrót do sieci neuronowych jako preferowanego rozwiązania w obrębie sztucznej inteligencji.

3.1.6. Systemy oparte na wiedzy

Wczesne programy sztucznej inteligencji opierały się na przeszukiwaniu całej przestrzeni problemowej krok po kroku celem znalezienia holistycznego, kompletnego rozwiązania. Dziś nazywamy takie podejście metodami słabymi (ang. weak methods, tłumaczenie własne), gdyż nie uwzględniają one ani problemu skalowania, ani trudniejszych jednostek problemowych.

Badacze zaproponowali alternatywę w postaci wyposażenia programów w wiedzę domenową. Algorytmy opierały się wtedy nie tyle na żmudnym powtarzaniu kroków, a na rozpoznawaniu najpierw sytuacji typowych i przekraczaniu ich określoną zawczasu

sekwencją. Dzięki temu programy mogły, mówiąc metaforycznie, podejmować większe kroki w obrębie prowadzonych wnioskowań dla typowych, często powtarzających się problemów. Trudniejsze, nietypowe problemy stanowią jednak wciąż olbrzymie wyzwanie dla takich systemów, gdyż przeprowadzane sekwencje kroków nie musiały prowadzić do rozwiązania, trzeba było się cofać w rozumowaniu, szukać gdzie wystąpiły błędy. Russell i Norvig zauważają, że dochodziło do tego stopnia komplikacji, że aby rozwiązać taki trudny problem, trzeba było w zasadzie znać już na niego odpowiedź (Russell i Norvig, 2010, s. 22).

Przegalińska określa ten okres jako ograniczenie naszych metafizycznych ambicji związanych ze sztuczną inteligencją i zaprzęcenie jej do celów bardziej pragmatycznych, wspierających nasze codzienne zadania (Przegalińska, 2016, s. 110). Za przykład takiej praktyki przywołuje ona rozwój systemów eksperckich. Były to programy, już w samym swoim założeniu, doradcze. Ich głównym celem było wsparcie naszych, ludzkich procesów decyzyjnych w dziedzinach oderwanych od wyzwań związanych z codzienną zdroworozsądkowością i w sytuacjach społecznych. Przegalińska egzemplifikuje te systemy, odwołując się do dwóch projektów: „DENDRAL” i ”MYCIN” (Przegalińska, 2016, s. 111).

Wczesnym przykładem takiego podejścia był program „DENDRAL”. Opracowali go trzej badacze z uniwersytetu w Stanford: Ed Feigenbaum (uczeń Herberta Simona), Bruce Buchanan (filozof, który został informatykiem) oraz Joshua Lederberg (laureat nagrody Nobla w dziedzinie genetyki). Celem projektu było rozwiązanie problemu inferencji struktury molekularnej z informacji dostarczonych przez spektrometr masy. Sygnałem wejściowym dla programu były elementarne wzory cząsteczek (np. $C_6H_{13}NO_2$) i widma masowe, określające masy różnych fragmentów cząsteczki powstające w wyniku bombardowania jej skupionymi wiązkami elektronów (Buchanan, Sutherland i Feigenbaum, 1969).

Pierwotna wersja tego programu generowała wszystkie możliwe struktury na podstawie wzorów matematycznych, a następnie przewidywała, jakie widma masowe mogą zostać zaobserwowane, by ostatecznie porównać ją z faktycznymi wartościami obserwacji. Naturalnie, szybko odkryto, że takie podejście jest niemożliwe do obliczenia nawet dla średnio dużych cząstek.

Badacze skonsultowali się więc z analitykami chemicznymi i odkryli, że ci pracują na podstawie dobrze znanych im wzorców szczytów pomiarów, gdyż sugerowały one wspólne podstruktury cząstek. Zdanie sobie sprawy z istnienia tychże podstruktur

było znaczące, ponieważ umożliwiło zawczasu odfiltrowanie znaczącej ilości kandydatów. Obniżono w ten sposób ilość kombinacji, jakie należy sprawdzić. Warto odnotować, że stanowiło to odniesienie do wspomnianego w rozdziale 3.1.4. programu McCarthy'ego: „Advice Taker”, w ramach którego postulowane było odseparowanie zasad (wiedzy) od komponentu wnioskującego.

„DENDRAL” był pierwszym systemem opartym na wiedzy ekspertów. Szybko jednak dołączyły do niego kolejne. Jego autorzy następnie uruchomili „Projekt Programowania Heurystycznego” (org. Heuristic Programming Project). Jego celem było zbadanie zakresu, w jakim nowa metodologia systemów ekspertowych mogła być zaaplikowana do innych obszarów.

Kolejnym istotnym krokiem było stworzenie systemu w obrębie diagnostyki medycznej. Tak powstał „MYCIN”. Jego głównym celem była diagnoza infekcji krwi. Wyodrębniono ponad 450 zasad, jakimi kierują się eksperci w tym obszarze. Program działał z podobną skutecznością do wspomnianych ekspertów, a wyraźnie wyższą niż przeciętni lekarze o krótkim stażu pracy.

W przeciwieństwie do „DENDRAL-a” nie istniał tu ogólny, powszechny i łatwo dostępny model teoretyczny, na którym można by oprzeć proces wnioskowania. Zamiast tego konieczny był cykl pogłębionych wywiadów z ekspertami, którzy posiadli swoją wiedzę w podobny sposób, współpracując z innymi specjalistami, jak również z podręczników napisanych przez inne autorytety w obrębie tej dziedziny i ostatecznie również na drodze ich własnych, empirycznych doświadczeń.

Co więcej, ustalone przez tych ekspertów zasady wcale nie były pewne. Reprezentacja medycznej wiedzy musiała więc odzwierciedlać inherentne dla tej domeny poziomy niepewności. Osiągnięto to za pomocą tak zwanych czynników pewności, które funkcjonowały w przybliżony sposób jak wspomniani lekarze, którzy napotykali na kolejnych pacjentów, kolejne przypadki i musieli, najlepiej jak potrafili, wykorzystywać swoją wiedzę i umiejętności zdobyte wcześniej (Russell i Norvig, 2010, s. 22).

Jak podsumowuje z kolei Przegalińska, badacze sztucznej inteligencji zakładali, być może znów nazbyt optymistycznie, że wiedza ekspertów może być w prosty sposób wydobyta i skwantyfikowana czy też zalgorytmizowana. Tymczasem okazało się, że sama wiedza, do której mamy dostęp jest często błędna, niedookreślona lub niekompletna, a sami eksperci nie zawsze zdają sobie dokładnie sprawę z tego, co wiedzą na temat własnych domen. Przypomina to cztery poziomy kompetencji wymieniane w psychologii. Zostały one opisane pierwotnie jako cztery poziomy nauki przez Martina Broadwella

w 1969 roku (1969), a następnie wspomniane w książce Paula Curtissa i Phillippa Warrena w 1973 roku (1973, s. 89), ale największą popularność zapewnił im Noel Burch w latach 70. XX wieku, który nazywał je czterema etapami nauki dowolnej umiejętności (Adams, 2011).

Polega on na podziale naszej biegłości w dowolnej umiejętności na cztery, kolejne etapy: nieświadomej niekompetencji, świadomej niekompetencji, świadomej kompetencji, aż do nieświadomej kompetencji. Na początku osoba ucząca się musi dostrzec, pożyteczność danej umiejętności i zdać sobie sprawę z własnej niewiedzy. Gdy już uświadomi sobie swoje braki, może podjąć próby ich uzupełnienia i nabycia kompetencji, wciąż jednak świadoma popełnianych przez siebie błędów. Trzeci etap to już poziom, na którym osoba wie i rozumie jak coś robić, jednakże wciąż wymaga to od niej skupienia uwagi i nie przychodzi naturalnie. Ostatni poziom biegłości to stan, gdy dana umiejętność staje się naszą „drugą naturą” i możemy ją wykonywać z łatwością, nawet w trakcie zajmowania się czym innym (Adams, 2011).

Ten ostatni poziom biegłości opiera się na intuicji, a to właśnie w jej kontekście Przegalińska przywołuje również Huberta Dreyfusa, który stwierdził, że ekspercki poziom wiedzy jest niemożliwy do osiągnięcia przez maszyny (Przegalińska, 2016, s. 111).

Innym obszarem, który wymagał zastosowania wiedzy domenowej, był język naturalny. Dokładniej rzecz ujmując, w latach 70. XX wieku trwały prace nad stworzeniem systemu, który rozumiałby język, w kształcie i formie używanej normalnie przez ludzi. Jednym z pierwszych systemów, który realizował ten cel był SHRDLU, autorstwa Terry’ego Winograda. Jednakże był on oparty na analizie ściśle syntaktycznej i efektywnie powtórzył on te same błędy, co wcześniej opisane projekty związane z tłumaczeniem maszynowym.

W przeciwieństwie do nich radził sobie jednak z niejasnością oraz rozumiał odnośniki do zaimków. Wynikało to ze ścisłego dopasowania do obszaru, w którym się poruszał, wspomnianego wcześniej świata klocków. W tamtym czasie inni badacze, jak Eugene Charniak, sugerowali, że rozumienie języka naturalnego wymagałoby szerokiej wiedzy o świecie i ogólnej metody jak z tej wiedzy korzystać (Russell i Norvig, 2010, s. 23).

Roger Shank, lingwista i badacz sztucznej inteligencji, również zmierzył się z wyzwaniem modelowania języka naturalnego. Wzbudził szereg dyskusji kontrowersyjnym stwierdzeniem, że nie istnieje coś takiego jak składnia. Wraz ze swoimi studentami zbudowali serię programów, które miały nas przybliżyć do łatwiejszego

komunikowania się z maszynami (Schank i Abelson, 1977; Wilensky, 1978; Schank i Riesbeck, 1981; Dyer, 1983).

W swoim podejściu kładli oni nacisk na rozwiązanie problemów związanych z reprezentacjami i wnioskowaniem z wykorzystaniem uzyskanej wiedzy. Obejmowało to takie tematy jak: reprezentacje stereotypowych sytuacji (Cullingford, 1981), opisywanie sposobów możliwej organizacji ludzkiej pamięci (Rieger, 1976; Kolodner, 1983) czy też rozumienie planów i celów (Wilensky, 1978).

Russell i Norvig stwierdzają, że biorąc pod uwagę zbliżony czas „narodzin” obu dyscyplin, językoznawstwo i sztuczna inteligencja efektywnie „dorastały razem”. Obecnie przecinają się one w hybrydowej formie nazywanej językoznawstwem obliczeniowym lub przetwarzaniem języka naturalnego, chętnie czerpiącym z osiągnięć na obu polach (Russell i Norvig, 2010, s. 16).

Inne systemy sztucznej inteligencji opierały się na idei ram Marvinina Minsky’ego (1975). Przyjęły one bardziej ustrukturyzowane podejście i organizowały wiedzę w fakty opisujące obiekty lub typy zdarzeń, które następnie organizowano w duże, taksonomiczne hierarchie analogiczne do struktur biologicznych.

Minsky w tym czasie, w ramach swojej działalności na MIT, często nadzorował prace studentów i zachęcał ich do rozwiązywania ściśle określonych problemów badawczych. Te ograniczone domeny przyjęto się określać „mikroświatami” (org. Microworlds, tłumaczenie własne). W ten sposób, jak zauważają Russell i Norvig, powstał szereg rozwiązań istotnych dla rozwoju całej dyscypliny. Autorzy wymieniają między innymi program Jamesa Slagle’a „SAINT” z 1963 roku, który był zdolny rozwiązać zamknięte formy zadań związanych z całkowaniem na poziomie prac domowych dla studentów pierwszego roku. Kolejny, napisany w 1968 roku, program „ANALOGY”, autorstwa Toma Evansa, rozwiązywał problemy analogii geometrycznej, znane z testów IQ. Następny był program Daniela Bobrowa „STUDENT”, z 1967 roku, który rozwiązywał problemy algebraiczne oparte na takich historiach jak: „jeżeli liczba klientów Toma jest dwukrotnością kwadratu 20% liczby reklam, jakie wykupił i liczba tych reklam wynosi 45, to ilu klientów ma Tom?” (Russell i Norvig, 2010, s. 19).

Najsłynniejszym „mikroświatem” był wspomniany wcześniej „świat klocków” (org. blocks world, tłumaczenie własne), Składał się on z zestawu solidnych klocków umieszczanych na blacie stołu. Typowe zadania realizowane w nim polegały na przemieszczaniu klocków w określony sposób, używając robotycznej dłoni,

która mogła podnieść jeden klocek naraz. W świecie klocków zrealizowano liczne projekty:

- David Huffman badał w nim, ogólnie ujmując, sposoby na konsolidację bodźców zbieranych za pomocą wizji komputerowej (Huffman, 1971),
- David Waltz również zajmował się bodźcami wzrokowymi, ze szczególnym uwzględnieniem problematyki wykrywania ścian oraz krawędzi przedmiotów (Waltz, 1975),
- Patrick Winston studiował uczenie maszynowe i ogólnie sposoby modelowania ludzkich procesów uczenia w maszynach (Winston, 1970),
- Terry Winograd realizował program związany z rozumieniem języka naturalnego (Winograd, 1972),
- Scott Fahlman pracował nad projektem związanym z organizowaniem i inteligentnym planowaniem zadań (Fahlman, 1974).

Był to czas raczej spokojnych przemian nastawionych na tworzenie pragmatycznych rozwiązań. Badacze nie tworzyli już szumnych obietnic bez pokrycia, a ich programy badawcze były nastawione na określone problemy i szukanie możliwości ich ekstrapolacji na inne obszary.

3.1.7. Sztuczna inteligencja współcześnie

Kolejny okres zaczyna się od lat 80. XX wieku i trwa do dzisiaj. W tym czasie możemy wyróżnić kilka wyraźnych trendów, które pozwoliły na unormowanie i sukcesywne ustabilizowanie rozwoju sztucznej inteligencji. Warto odnotować, że w związku z poniżej opisanymi doniesieniami powoli mówi się o ustępowaniu kolejnej zimy sztucznej inteligencji (Havenstein, 2005). Jak zauważają inni, obecnie wiele tysięcy aplikacji opartych na sztucznej inteligencji jest powszechnych. Praktycznie każde przedsiębiorstwo z nich korzysta, nierzadko nie zdając sobie z tego sprawy i operując jedynie na poziomie infrastruktury informatycznej (Kurzweil, 2005).

Komercyjne wykorzystanie sztucznej inteligencji

Pierwszym systemem eksperckim, który odniósł komercyjny sukces był R1 (McDermott, 1982). Służył on do konfigurowania nowych zestawów systemów komputerowych. Już przed 1986 panowała opinia, że korzystanie z niego pozwoli zaoszczędzić odpowiednio dużej firmie ponad 40 milionów dolarów rocznie.

Systemy eksperckie rozwijały się w latach 80. XX wieku bardzo dynamicznie. Jeszcze przed 1988 sama firma DEC opracowała już ponad 40 takich systemów i planowała kolejne. Inne przedsiębiorstwo, DuPont, miało przeszło 100 zaimplementowanych systemów tego typu i szykowało kolejnych 500. Jak to opisali, pozwalało to na oszczędności rzędu 10 milionów dolarów rocznie. Prawie każda większa korporacja amerykańska miała swoją grupę badawczą wyspecjalizowaną w sztucznej inteligencji, która zajmowała się badaniem i implementowaniem systemów ekspertowych (Russell i Norvig, 2010, s. 24).

Tymczasem, w 1981 roku, Japonia ogłosiła 10-letni plan budowy inteligentnych komputerów opartych na języku PROLOG. W odpowiedzi na to Stany Zjednoczone powołały Korporację Mikroelektroniki i Technologii Komputerowych (org. Microelectronics and Computer Technology Corporation, MCC). Równolegle, w Wielkiej Brytanii, raport Alvera posłużył jako bodziec do przywrócenia środków na badania sztucznej inteligencji, które wcześniej odebrano na podstawie, wspomniany wcześniej w rozdziale 3.1.5., raport Lighthilla.

Russell i Norvig podsumowują, że w mniej jak dekadę wartość przemysłu sztucznej inteligencji urosła z wyceny na kilka milionów dolarów do wielu miliardów. Setki firm zaczęły w tym samym czasie budować systemy ekspertowe, ale również komercyjne rozwiązania w obszarze wizji komputerowej czy robotyce. Po tym okresie nastąpiła jednak kolejna „zima sztucznej inteligencji”, gdy wiele z tych firm upadło, gdyż nie były w stanie zaspokoić pokładanych w nich nadziei (Russell i Norvig, 2010, s. 24).

Powrót sieci neuronowych

Jak zauważają Russel i Norvig, w latach 80. XX wieku przynajmniej cztery grupy badawcze wynalazły na nowo algorytm uczący oparty na wstecznej propagacji błędu. Pierwszy, być może najpopularniejszy, autorstwa Arthura Bryson i Yu-Chi Ho, wspomniany w rozdziale 3.1.5., był w istocie przedstawiony wcześniej, bo już w 1969 roku. Dopiero po przeszło dekadzie został on jednak dostrzeżony szerzej i zastosowany do wielu problemów w obszarach nie tylko ściśle informatycznych, ale również w obrębie badań nad umysłem i psychologii. Powszechny entuzjazm wywołała szeroko rozpowszechniona analiza tego algorytmu opublikowana w ramach zbioru *Pararell Distributed Processing* (Rumelhart i McClelland, 1986).

Modele inteligentnych systemów oparte na sieciach neuronowych określamy mianem koneksjonistycznych. Przegalińska wskazuje w tym kontekście na Donalda Hebba

jako ich twórcę, jako „ojca koneksjonizmu” (1980). Warto przypomnieć, że sieci neuronowe to rozwiązanie, które odchodzi od tworzenia oprogramowania za pomocą tradycyjnych instrukcji warunkowych i podobnych metod. Zamiast tego tworzone są pojedyncze jednostki przetwarzające dane, które funkcjonują podobnie do komórek nerwowych. Zgodnie z naszą wiedzą neurologiczną zakładamy, że aktywność mózgu opiera się właśnie na propagowaniu pobudzeń między komórkami nerwowymi zorganizowanymi w konkretne sieci. Architektura koneksjonistyczna jest tym silnie inspirowana i nie ma pojedynczego, centralnego procesora przetwarzającego symbole, lecz rozproszoną sieć neuronów równolegle przekazujących sobie impulsy. Powiązania między nimi zazwyczaj nie są równoważne i za pomocą wag jesteśmy w stanie określić ich siłę. Jednostki też nie odpowiadają konkretnym symbolom czy znaczeniom (Przegalińska, 2016, s. 122).

Inteligencja według koneksjonistów jest pochodną interakcji dużej ilości prostych jednostek (Franklin, 2001, s. 121). Ich główną aspiracją jest modelowanie procesu uczenia się w ten sam sposób, w jaki przebiega on u ludzi. Stąd też nie projektują oni sieci pod kątem wykonania pojedynczego zadania (na przykład, jak rozpoznać ręcznie zapisaną cyfrę „1”), lecz przygotowanie jej do uogólnionej pracy (kontynuując przykład, jak rozpoznawać pismo ręczne). Sieć uczy się poprzez próby wykonania zadań i w ramach nich popełnia błędy podobnie jak człowiek.

Szeroko rozumiane podejścia koneksjonistyczne stanowią bezpośrednich rywali dla modeli symbolicznych, promowanych przez Newella i Simona, czy logistycznych, które reprezentował między innymi McCarthy (Smolensky, 1988). W latach 80. XX wieku zdecydowanie dominującym podejściem w obrębie sztucznej inteligencji były modele symboliczne. Była to teza bardzo długo podtrzymywana. Terrence Deacon, po przeszło 20 latach, postulował w swojej książce, że operacje na symbolach to cecha dystynktywna dla gatunku ludzkiego (Deacon, 1997).

Argumenty poruszane przez koneksjonistów, jak parafrazują Russell i Norvig, kwestionują użyteczność i wartość eksplanacyjną opisów manipulacji symbolicznych w kontekście dokładnych modeli poznawczych (Russell i Norvig, 2010, s. 25). Stąd też Przegalińska, za Franklinem, nazywa podejście koneksjonistyczne subsymbolicznym paradygmatem sztucznej inteligencji (Franklin, 2001). Russell i Norvig zauważają, że podejścia te zaczynały jako rywalizujące, dziś jednak widzimy je jako komplementarne. Za przykład podają badania realizowane na przecięciu sztucznej inteligencji i kognitywistyki (Russell i Norvig, 2010, s. 25).

Współcześnie badania nad sieciami neuronowymi realizowane są w dwóch celach: pierwszy z nich pragnie optymalizować ich własności matematyczne poprzez pracę nad dostępnymi architekturami i algorytmami. Drugim celem jest możliwie bliskie odtworzenie danych empirycznych opisujących własności neuronów i ich połączeń, aby lepiej zrozumieć nasze biologiczne uposażenie.

Zaadoptowanie metody naukowej

Sztuczna inteligencja nieustannie się rozwija i, jak zauważają Russel i Norvig, powoli odchodzimy w niej od gwałtownych, rewolucyjnych skoków i przechodzimy do bardziej zrównoważonego modelu rozwoju. Oznacza to przejście do rozwijania istniejących teorii zamiast proponowania całkowicie nowych rozwiązań.

Podobnie, Russel i Norvig mówią o odejściu od intuicji na rzecz rygorystycznej pracy nad twierdzeniami matematycznymi i danymi empirycznymi. Ostatecznie, widzimy, że zostały wyciągnięte wnioski z wcześniejszych błędów i pomyłek. Widzimy poprawę w postawach naukowców, w ich skupieniu na pragmatyce zastosowań zamiast przywoływania wyidealizowanych przykładów (Russell i Norvig, 2010, s. 25).

Warto powtórzyć, że sztuczna inteligencja powstała niejako w kontrze do istniejących w tamtym czasie dziedzin badawczych jak teoria kontroli czy statystyka. Niemniej jednak obecnie dyscypliny te stanowią bardzo istotne źródło, które rozwija sztuczną inteligencję. Do tego stopnia, że wreszcie zaadoptowano w badaniach nad nią tak charakterystyczną i sprawdzoną metodykę ściśle naukową, opartą na stawianiu hipotez, przeprowadzaniu eksperymentów, obserwacji wyników i ich skrupulatnym odnotowywaniu. Hipoteza badacza sztucznej inteligencji, aby zostać zaakceptowaną (czy też nieodrzuconą), musi zostać poddana serii eksperymentów empirycznych, a jej rezultaty muszą spełniać warunki statystycznej istotności (Cohen, 1995). Dodatkowo powszechną, ale też ekstremalnie ważną praktyką stało się udostępnianie repozytoriów danych i kodu źródłowego, celem możliwie jak największego zwiększenia dostępności rozwiązań i powtarzalności eksperymentów (Russell i Norvig, 2010, s. 25).

Pojawienie się inteligentnych agentów

Ogólny postęp w różnych dyscyplinach szczegółowych sztucznej inteligencji stanowił również istotny bodziec do ponownego rozważenia tematu inteligentnych agentów. Praca Allena Newella, Johna Lairda i Paula Rosenbloom'a nad programem SOAR stanowi jeden z lepiej rozpoznawanych przykładów kompletnej architektury poznawczej syntetycznego agenta (Newell, 1990; Laird, Newell i Rosenbloom, 1987).

Istotne też dla rozwoju tej dyscypliny było rozpowszechnienie internetu, a sztuczni agenci zostali rozpropagowani tak szeroko w wielu internetowych aplikacjach, że słowo „bot” weszło do naszego codziennego języka. Co więcej, stały się one integralną częścią systemów, na bazie których tworzymy nasze obyczaje powiązane z poruszaniem się po sieci web. Obecnie ciężko nam sobie wyobrazić tę czynność bez wspierania się, chociażby silnikami przeszukiwania (org. search engines, tłumaczenie własne), systemami rekomendacyjnymi czy agregatorami stron internetowych (Russell i Norvig, 2010, s. 26).

Jak kontynuują Russell i Norvig, próby tworzenia „kompletnych” agentów, czyli zdolnych do funkcjonowania w realnym świecie, unaocznily konieczność ujednoczenia rozproszonych pod-dyscyplin sztucznej inteligencji. Szczególnie dotyczy to szeroko docenianych systemów sensorycznych opartych na wizji, dźwięku i zdolnych do ekstrakcji informacji w nich zawartych. Przykładem użyteczności takiego rozwiązania jest rozpoznawanie mowy (Russell i Norvig, 2010, s. 26).

Dodatkową komplikacją jest fakt, że systemy te nie są w stanie dostarczać informacji absolutnie pewnych. Stąd też nasze systemy wnioskowania i planowania muszą radzić sobie z niepewnością. Wspomniano o tym w rozdziale 3.1.2., na przykładzie teorii decyzji i, powiązanej z nią, teorii gier.

Drugą, znaczącą konsekwencją takiego celu jest istotne zbliżenie dyscyplin sztucznej inteligencji i teorii kontroli. Szczególnie zauważalne jest to w obszarze projektowania pojazdów robotycznych, a zwłaszcza samochodów autonomicznych. Dochodzi w tej sytuacji do połączenia szerokiego spektrum podejść celem usprawnienia działania sensorów, integracji sensoryki z lokalizacją, mapowaniem położenia w przestrzeni siebie i innych pojazdów na drodze, ale również wysokopoziomowego planowania (Russell i Norvig, 2010, s. 26).

Współcześnie celem teorii kontroli, a zwłaszcza jej gałęzi opartej na zagadnieniach stochastycznej optymalizacji, jest projektowanie systemów zdolnych do maksymalizowania swojej użyteczności na przestrzeni czasu. Russell i Norvig porównują tę ambicję do podobnego tematu optymalizacji działania agentów (Russell i Norvig, 2010, s. 26).

Kontynuują oni, dostrzegając zbieżność tych celów i poszukują cechy dystynktywnej, która pozwoliłaby różnicować te dziedziny. Znajdują ją w doborze narzędzi, gdyż osoby zajmujące się teorią kontroli poniekąd zamknęły się w obrębie rachunku różniczkowego i algebrze macierzy, a badacze sztucznej inteligencji wyszli ponad te techniki, uwzględniając je w swoim repertuarze, ale umieszczając w nim również,

choćby narzędzia wnioskowania logicznego czy teorii obliczeń. Dzięki nim możliwe było wyjście do szerokiego spektrum rozważań, pozostających poza możliwościami teoretyków kontroli, jak język, percepcja czy planowanie (Russell i Norvig, 2010, s. 15).

Big Data

Jak zauważają Russell i Norvig, ostatnie blisko 60 lat historii informatyki unaocznia, że główny obiekt badań w tym obszarze stanowią algorytmy (Russell i Norvig, 2010, s. 26). Nowsze prace w dziedzinie sztucznej inteligencji zwracają jednak naszą uwagę na inny aspekt. Konkretnie na dane, na których pracujemy, ich wolumen i jakość.

Zjawisko to zwróciło uwagę Gartnera. Jest to amerykańskie przedsiębiorstwo, które od lat 80. XX wieku buduje swoją pozycję poprzez bogatą ofertę wnikliwych raportów podsumowujących najważniejsze odkrycia i trendy w zakresie strategicznego wykorzystania technologii oraz zarządzania technologiami. Ostatecznie, jeden z zatrudnionych tam analityków, Doug Laney, już w 2001 roku wyróżnił trzy konkretne aspekty danych, jakie nabrały znaczenia w erze Big Data (Laney, 2001). Są to kolejno: Wolumen (Volume, tłumaczenie własne), Szybkość (Velocity, tłumaczenie własne) i Zmienność (Variety). Mamy więc do czynienia z olbrzymimi ilościami danych, które muszą być przetwarzane w czasie bliskim rzeczywistości o różnorodnej, zmieniającej się strukturze lub wręcz jej braku.

Stanowi to istotną uwagę w naszych czasach, gdy zauważamy rosnącą dostępność bardzo dużych źródeł danych. Wirtualnie każdy ma dostęp, na przysłowiowe wyciągnięcie ręki, do zbiorów, na przykład opisujących kwintyliony słów po angielsku czy miliardy obrazów z sieci (Kilgarriff i Grefenstette, 2006) lub też kolejne miliardy par splecionych w sekwencjonowaniu genetycznym (Collins, Morgan i Patrinos, 2003).

Przykładem ilustrującym tę zmianę podejścia jest opisany w artykule Davida Yarowsky'ego temat ujednoznaczniania słów (Yarowsky, 1995). Oznacza to proces wyboru lub dopasowania znaczenia słowa do kontekstu. Przywołany zostaje przykład angielskiego słowa „plant”, które posiada przynajmniej dwa znaczenia- roślinę i fabrykę.

Jak możemy rozpoznać, kiedy określone znaczenie jest adekwatne? Wcześniejsze podejścia opierały się na tak zwanych etykietach. Zatrudnieni w tym celu ludzie przechodzili przez zbiór danych i umieszczali etykiety wskazujące, że w tym zdaniu takie znaczenie jest właściwe, a w tamtym zdaniu inne znaczenie jest poprawne.

Tymczasem D. Yarowsky udowodnił, że możemy obejść się bez takich etykiet. Dodatkowo osiągnął on dokładność rzędu 96% rozpoznawania znaczeń.

Zamiast zatrudniać ludzi, załadował on olbrzymi korpus czystego tekstu i odwołał się do słownikowej definicji. Następnie automatyczny proces przeanalizował cały ten zbiór danych w poszukiwaniu wzorców, które umożliwiły rozpoznanie nowych przykładów.

Michele Banko i Eric Brill wykazali, że podobne techniki radzą sobie jeszcze lepiej, gdy zwiększymy ilość tekstu o rząd wielkości- od miliona do miliarda. Co więcej, operacja ta spowodowała większy wzrost dokładności niż jakakolwiek zmiana algorytmu. Przeciętny algorytm z dostępem do wzorca liczącego 100 milionów słów radzi sobie lepiej niż jego najlepszy odpowiednik z milionem podobnych rekordów (Banko i Brill, 2001).

Do podobnych wniosków doszli James Hays i Alexei Efros, którzy rozpatrywali problem wypełnienia braków w fotografii. Dopiero zmiana ilości elementów zbioru uczącego przyniosła w ich wypadku istotną zmianę wydajności, od 10 000 zdjęć do blisko 2 milionów. Zmiany algorytmów nie miały podobnego wpływu nawet w przybliżeniu (Hays i Efros, 2007).

Stąd też niektórzy badacze postulują, że czynnikiem ograniczającym sztuczną inteligencję jest problem, jak przedstawić całą wiedzę, której potrzebuje określony dany system. Kontynuują oni, że rozwiązaniem może być w wielu przypadkach zastosowanie podobnych podejść opartych na inżynierii wiedzy i dostarczenie systemom dostatecznie dużych ilości danych (Halevy, Norvig i Pereira, 2009).

3.2. Myśl filozoficzna a sztuczna inteligencja

3.2.1. Filozofia sztucznej inteligencji

Przegalińska stawia sobie za cel ustalenie, jakie filozoficzne koncepcje umysłu i ciała umożliwiły konceptualizację procesu technologicznego, który miał nas doprowadzić do mechanizacji czy też wirtualizacji umysłu oraz, finalnie, genezy sztucznej inteligencji (Przegalińska, 2016, s. 83).

Ponadto wykazuje ona, że sztuczna inteligencja zyskuje znacząco dzięki wkładowi fenomenologów w jej krytykę i, poprzez nią, rozwój. Postuluje również, że bez tego humanistycznego spojrzenia nastąpiłoby zahamowanie rozwoju dyscypliny na paradygmacie obliczeniowym i zrównaniem umysłu do sprawnego komputera (Przegalińska, 2016, s. 15). Uwzględnienie go jednak umożliwia nam teraz tworzenie rozwiązań, które wchodzą z ludźmi w wielowymiarowe, skomplikowane relacje. Dodatkowo interakcje z nimi odbywają się na „ludzkim” poziomie, na przykład poprzez czat wirtualny w języku naturalnym. Stąd też, ludzie nasze rosnące oczekiwania

wobec syntetycznych botów i spodziewanie się praktycznie ludzkich zachowań (Przegalińska, 2016, s. 19).

Dla Przegalińskiej formowanie się idei sztucznej inteligencji zaczęło się już w czasach antycznej Grecji. Zwłaszcza wspomina w tym kontekście Arystotelesa i jego ujęcie duszy, jako najdoskonalszej z form, zdolnej do zrozumienia świata, świadomej własnej odrębności i, przez to, również siebie samej. Ciało stanowiło dla niej jedynie mechanizm pośredniczący, umożliwiający poznawanie rzeczywistości, gdyż tylko bodźce zmysłowe były prawdziwe (Przegalińska, 2016, s. 88).

W XVII wieku jednak pojawiła się odmienna idea, dualizm Kartezjusza. Ciało i dusza zostały rozdzielone. To pierwsze traktowane jako maszyneria o określonych granicach. Dusza natomiast jest skupiona na myśleniu, przeżywaniu uczuć i świadomych przeżyciach. Ciało jest mechaniczne i ożywiane przez impulsy materialne wytwarzane we krwi. Podobnie, przyroda jest zbudowana z ciał, które można dzielić na mniejsze i łatwiejsze do zrozumienia części. Dzięki takiemu ujęciu możliwe jest tworzenie opracowań naukowych opisujących jej działanie (Przegalińska, 2016, s. 89). Ponadto kartezjańska wizja dualnej natury umysłu była, według Przegalińskiej, kluczowa w formowaniu idei sztucznej inteligencji. Zakładała ona, że oddzielenie umysłu od ciała pozostanie bez wpływu na umysł, jak również pozwalała na możliwość transgresji poza siebie i dostęp do świata wirtualnego, świata nieograniczonych możliwości. Przy czym „wirtualny” ma tu znaczenie średniowieczne, synonimiczne do „potencjalny” (Przegalińska, 2016, s. 29).

Podobnie, istotna jest dla niej kartezjańska wizja maszyny niedoskonalej, będącej jedynie krzywym zwierciadłem dla ludzkiego umysłu (Przegalińska, 2016, s. 27). Z kolei już w XVII i XVIII wieku myśliciele tacy jak Charles Babbage i Thomas Hobbes widzieli w „automatycznym umyśle” partnera zamiast narzędzia (Przegalińska, 2016, s. 28).

Przegalińska dostrzega, że dla sztucznej inteligencji kluczowe były również poglądy David Hume’a, który postulował, że reprezentacje są ideami, a ich ciągi skojarzeniowe to procesy mentalne (Przegalińska, 2016, s. 99).

Autorka kontynuuje, stwierdzając, że w tym czasie zrealizowano w wielu obszarach koncepcje Kartezjusza. Kluczowe w tym były dwa zjawiska: technologizacja rzeczywistości za pomocą zaawansowanych programów służących do przeprowadzania skomplikowanych obliczeń oraz ogólny proces wirtualizacji, przenoszenia rosnącej liczby aktywności do przestrzeni wirtualnej, komputerowej. Autorka pisze wręcz,

że „[...] człowiek coraz częściej zwracał swoją uwagę ku ekranowi, by stopniowo zacząć przenosić się w jego świat” (Przegalińska, 2016, s. 30).

Podstawy propozycyjne oraz relacje między osobą a sądami natomiast znajdują źródło w pracach Gottloba Fregego i Bertranda Russella (Przegalińska, 2016, s. 99).

Idea wspierania decyzji jest, jak zauważa Przegalińska, dobrze osadzona w niedawnej historii Zachodu (Przegalińska, 2016, s. 90). Zauważają to również Russell i Norvig, którzy dostrzegają, jak w książce Johna Stuarta Milla promowana była idea racjonalnego podejmowania decyzji jako kryterium adekwatnego dla wszystkich sferach aktywności życia człowieka (Mill, 1863). W podobnym celu, w latach 20. XX wieku, Abraham Felixner powołał „Institute for Advanced Study”.

Przegalińska opisuje powojenny ideał maszyny myślącej. Jest to odcięta, niezależna od ciała, acz zdolna do zarządzania fizycznymi procesami jednostka obliczeniowa. Stanowi ona ideał całkowicie posthumanistyczny, maszyna bowiem miała stać się nieodróżnialna od człowieka poprzez ścisłą integrację, przez co niemożliwe byłoby określenie, gdzie się kończy maszyna, a gdzie zaczyna człowiek. Była to realizacja obaw Martina Heideggera, który przestrzegał przed stanem relacji *Dasein* – narzędzie, w której technologia już nie jest narzędziem. Technologia tak ujęta kontroluje wręcz człowieka, jednocześnie pośrednicząc w jego poznawaniu świata i nadając mu sens (Przegalińska, 2016, s. 90).

Koncepcja obliczania jest kluczowa dla sztucznej inteligencji. Hipoteza Alona Churcha i Alana Turinga z 1936 roku mówi, że problem nie może być rozwiązany myślą, jeśli nie może być rozwiązany przez obliczenia oraz jeśli może być rozwiązany na drodze umysłowej, to może być również skonstruowana maszyna, która go rozwiąże. Przegalińska parafrazuje to, stwierdzając, że maszyna Turinga, stanowiąca abstrakcyjny model komputera, jest w stanie rozwiązać każdy problem, dla którego istnieje efektywny algorytm, z założeniem nieograniczonej pamięci i zasobów (Przegalińska, 2016, s. 95). Turing, zirytowany pytaniami, czy maszyny myślą, sformułował specjalny test. Polegał on na zaangażowaniu sędziego, człowieka, który prowadząc rozmowę w języku naturalnym z pozostałymi stronami (przypadkowo podzielonymi na ludzi i maszyny). Jeśli arbiter nie jest zdolny określić czy któraś ze stron jest maszyną, uznajemy, że maszyna przeszła test. Ten model, wielokrotnie modyfikowany, stał się bazą dla operacyjnej definicji inteligencji. Stwierdza ona, że zachowanie inteligentne to takie, które charakteryzuje się relacjami poznawczymi na poziomie ludzkim. W ich poczet

wliczane są takie kompetencje jak przetwarzanie języka naturalnego, reprezentacja wiedzy, automatyczne wnioskowanie oraz uczenie się (Przegalińska, 2016, s. 95).

Hipoteza Churcha i Turinga prawdopodobnie stanowiła inspirację dla Allena Newella oraz Herberta Simona, którzy w 1955 roku próbowali udowodnić, że komputery są zdolne nie tylko do obliczeń (Przegalińska, 2016, s. 95). W trakcie Symposium w Sprawie Teorii Informacji, we wrześniu 1956 roku na MIT, Newell i Simon przedstawili koncepcję Maszyny Teorii Logicznej. Był to program, który potrafił udowodniać twierdzenia matematyczne z *Principia Mathematica* Bertranda Russela i Alfreda Notha Whiteheada. Podczas tego wydarzenia wystąpił również Noam Chomsky, który zaprezentował swoją koncepcję gramatyki generatywnej, a George Miller przedstawił ideę pamięci krótkotrwałej (Przegalińska, 2016, s. 96).

Shanon również był wysoce zainspirowany koncepcjami maszyn Turinga. Bezpośrednim tego efektem był opracowany przez niego w 1956 roku komputer szachowy MANIAC. Przegalińska widzi w nim kluczowy wynalazek, który na długie lata zdefiniował rolę maszyn względem człowieka, jak również na definicję maszyny jako takiej w sensie filozoficznym. Jako że operowała ona na operacjach, które są tradycyjnie pojmowane jako intelektualne, stanowiła ona fizyczną realizację koncepcji rzeczy myślącej (Przegalińska, 2016, s. 94).

Kluczowa w rozwoju filozofii sztucznej inteligencji była działalność takich postaci jak Norbert Wiener, John von Neumann, Alan Turing i Warren McCulloch. Ich głównym zainteresowaniem badawczym było zbudowanie komputerów, które byłyby w stanie przejąć funkcje ludzkiego mózgu. Zastanawiali się, czy jest możliwa sztuczna inteligencja w ogóle. Zakładali też odpowiedź twierdzącą i rozważali, jak mogliby ją osiągnąć. Poszukiwali analogii między procesami w komputerze a ludzkim myśleniem. Kluczowymi tematami dla nich była komunikacja z maszynami i za ich pośrednictwem, tworzenie programów zdolnych do uczenia się, przeprowadzanie wnioskowania w stanach niepewności i przy niedostatku informacji, nabywanie wiedzy, a także percepcja i działanie. Aczkolwiek, jak zauważa Przegalińska, działanie być może w niewystarczającym stopniu, gdyż właśnie brak umocowania w sferze fizycznej doprowadził do przyszłych kryzysów całej dyscypliny określanych mianem zim sztucznej inteligencji (Pfeifer i Scheier, 2001; Przegalińska, 2016). Russel i Norvig zwracają zwłaszcza uwagę na Rudolfa Carnapa, który to, jako prawdopodobnie pierwszy, w swojej książce zdefiniował teorię umysłu jako procesu obliczeniowego (Carnap, 1928).

Tematy te zainspirowały Hillary'ego Putnama do stworzenia funkcjonalistycznej teorii umysłu, którą przedstawił w 1975 roku. Jej głównym założeniem było stwierdzenie, że inteligencja jest obliczeniem i przez to funkcją, którą możemy realizować na dowolnym „nośniku”. Przegalińska ponownie odnosi się tutaj to do poglądów Turinga, gdyż mówimy o badaniu inteligencji czy też poznania na poziomie algorytmów i procesów obliczeniowych, bez angażowania się w analizowanie podmiotów je realizujący (Przegalińska, 2016, s. 97).

Jak zauważają Russell i Norvig, opisany w rozdziale 2.2.1, GPS stanowił bezpośrednią inspirację dla Newella i Simona, którzy w 1976 roku przedstawili istotną hipotezę systemu fizycznych symboli. Przyjmuje ona, że dowolny system inteligentny musi być zdolny do operowania strukturami danych złożonymi z symboli (Russell i Norvig, 2010, s. 18). Zakłada ona, że każdy fizyczny system symboliczny ma możliwość realizowania zachowań, które możemy uznać za inteligentne. Umysł jest takim systemem i działa on na podstawie algorytmizowalnych operacji, polegających na przetwarzaniu symboli w obrębie ściśle określonych reguł. Takimi symbolami w tym ujęciu mogą być myśli, a przejścia między nimi wyjaśniają kausalność stanów mentalnych, czyli jak przechodzimy od jednego świadomego stanu do innego. Zamiast precyzować podmiot czy to posiadający, czy też realizujący umysł autorzy odwołują się do sytuacji wydarzających się w „świecie zewnętrznym” (jako przeciwwaga dla reprezentacji wewnętrznych). Dziś jednak takie procesy określilibyśmy mianem przetwarzania informacji zamiast myślenia (Bobryk, 2004, s. 42).

Dzięki takiemu spojrzeniu stało się możliwym używanie komputerów do przeprowadzania symulacji określonych aspektów inteligencji, co w połowie XX wieku przełożyło się na prowadzenie badań nad sztuczną inteligencją w obrębie psychologii poznawczej (Hetmański, 2000, s. 43-76). Doprowadziło to do nowej koncepcji umysłu nazywanej teorią obliczeniową. Funkcjonował w niej nowy obraz człowieka, przyrównanego do maszyny i nowy język, umożliwiający formułowanie teorii na temat procesów poznawczych (Przegalińska, 2016, s. 98).

Jak kontynuuje Przegalińska, równoległe do sformułowania koncepcji człowieka jako maszyny następował dynamiczny rozwój cybernetyki, który, odwrotnie, podniósł maszynę do roli podmiotu. Krytykowany przez Heideggera przedmiot jako obiekt techniczny, niezależny i stanowiący przedłużenie podmiotu (człowieka), został zastąpiony przez maszynę-podmiot zdolną do liczenia, a przez to również myślenia. Przegalińska podkreśla, że nie był to już podmiot jedynie w sensie ontologicznym,

ale również społecznym. Sztuczna inteligencja jako dyscyplina dąży do skonstruowania maszyny inteligentnej, mającej wpływ na rzeczywistość. Takiej, która dzięki przetwarzaniu danych jest w stanie organizować życie.

Przegalińska rozwija tę koncepcję i przedstawia sztuczną inteligencję na tym etapie już jako projekt metafizyczny, który, podobnie do fenomenologii, pragnie stać się *organum* nauk. Skupia się na abstrahowaniu, zerwaniu z naocznością i poddaniu procesów całkowitej formalizacji (Bobryk, 2004, s. 64).

Dąży do przeinterpretowania i zmatematyzowania materiałów takich nauk jak historia, filozofia, prawo czy socjologia kultury, a następnie użycia ich do badania procesów sterowniczych w różnych społeczeństwach i z różnymi funkcjami kultury społecznej (Kossecki, 1974, s. 2). Ostatecznie, jak pisze Przegalińska, zadaje pytanie o rolę ludzkiego doświadczenia, status świadomości czy umysłu, a wręcz stawia postulat o zmianę myślenia o naturze człowieczeństwa (Przegalińska, 2016, s. 99).

XX-wieczni filozofowie dzięki osiągnięciom swoim poprzedników, ale też technicznym możliwościom swoich czasów stworzyli całościową teorię umysłu. Przede wszystkim zakładała ona intencjonalność naszych spostrzeżeń i myśli. Jak określa to Przegalińska, podmioty doświadczające tychże intencjonalnych stanów mentalnych odnoszą się do przedmiotów w sposób zapośredniczony, podmioty poznające i przedmioty poznawane stoją, po kartezjańsku, na przeciwnych biegunach poznania (Przegalińska, 2016, s. 99).

Poznawane przedmioty tworzyły w umyśle ślad określany mianem reprezentacji. Przegalińska stwierdza, że były one największymi osiągnięciami cybernetyków i to właśnie one stanowiły o tym, że możemy uznać obliczeniową teorię umysłu za odmianę szerszej teorii reprezentacjonistycznej. W teorii tej stany poznawcze to odniesienia do mentalnych treści – reprezentacji, natomiast operacje na nich to procesy poznawcze (Przegalińska, 2016, s. 100).

Obliczeniowa teoria umysłu określa naturę człowieka jako algorytmicznego procesora informacji, z naturalnym dyskiem twardym i zaprogramowanego w określony sposób. Za swój nadrzędny cel postawiła określenie, jak odbieramy i przetwarzamy informacje. Bodziec-reakcja został zastąpiony sygnałami wchodzącymi i wychodzącymi, a łańcuchy tychże bodźców-reakcji zastąpiła teoriami o wewnętrznych obliczeniach i stanach obliczeniowych (Pfeifer i Scheier, 2001).

Jak stwierdza Przegalińska, traktowanie człowieka jako procesora wymaga uznania dwóch przesłanek:

a) aktywność poznawcza człowieka opiera się na funkcjonowaniu układu nerwowego,

b) posiadamy w umyśle oprogramowanie organizujące nasze procesy mentalne.

Człowiek staje się więc maszyną, która najpierw zmysłowo doznaje świat, przetwarza to w mózgu, oblicza w umyśle, a na koniec, być może, ujawnia swoje przemyślenia poprzez komunikowanie ich „na zewnątrz” (Przegalińska, 2016, s. 102).

Przegalińska kontynuuje, że stąd już było relatywnie łatwo przenieść pewne aspekty ludzkiej natury na maszynę inteligentną i, zasadniczo, przestać ją określać mianem maszyny. Autorka przywołuje w tym kontekście Phillipa Johnsona-Lairda, który pisze, że inteligencja jest równa postrzeganiu świata, zapamiętywaniu tychże spostrzeżeń, uczeniu się, kontrolowaniu swoich działań i komunikacji, a także, że każda z tych aktywności może być zapisana w formie obliczeń (Johnson-Laird, 1988, s. 52). Takie postrzeganie ludzi, umysłów i maszyn święciło triumfy w teoriach postrzegania, pamięci, wnioskowania, jak również analizach kodu genetycznego czy tworzeniu komputerowych symulacji zjawisk przyrodniczych (Przegalińska, 2016, s. 103). Podobnie, Russell i Norvig stwierdzają, że od lat 90. XX wieku możemy zaobserwować powrót zainteresowania technikami, dla których podejmowanie decyzji jest centralne w projektowaniu systemów agentów (Wellman, 1995).

Umysł, jak komputer, miał w tym ujęciu przetwarzać informacje sekwencyjnie i reagować na podstawie instrukcji warunkowych. Miał on być oprogramowaniem, realizowanym przez dowolną maszynę o dowolnej budowie zdolnej go zrealizować. Warto odnotować, że było to podejście często krytykowane i można tu przywołać teksty Newella (1980), Pylyshyna (1984), czy Miłkowskiego (2009). Pisali oni bowiem o zagrożeniach płynących z banalizacji obliczeniowej metodologii badań.

Tak więc maszyna stawała się odzwierciedleniem wizji człowieka, którą Przegalińska nazywa Ja-Naukowe (Przegalińska, 2016, s. 104). Przewartościowaniu podlegała też rola nauki, która z formy pracy twórczej stała się pragmatyczną praktyką odkrywania nowych prawd, które mogły dać określone korzyści. Przegalińska, za Dalstonem i Gallisonem stwierdza, że nauka wychodzi z roli obserwatora i staje się czynnikiem w gospodarce i życiu społecznym. Chwalono racjonalizm i tworzenie wspólnego dobra, rozumianego utylitarystyczno-pragmatycznie. Maszyna miała stać się lustrem dla Ja-Naukowego, a sprzęgnięcie ich razem miało skutkować rewolucją naukową i gospodarczą. Koniecznym więc stało się stworzenie maszyny zdolnej przenieść skutki wirtualnych obliczeń do rzeczywistości (Dalston i Gallison, 2007, s. 38-39).

3.2.2. Krytyka obliczeniowego podejścia do umysłu

Przegalińska przytacza opinię, że krytyka paradygmatu obliczeniowego stanowiła bezpośrednią reakcję na niezrealizowane założenia czy też obietnice sztucznej inteligencji (Ihde, 1997; Ihde, *Bodies in Technology*, 2002; Feenberg, 1991). Autorka postuluje, że charakterystyczną cechą sztucznej inteligencji jest to, że wszystkie teorie, które powstały dla niej, przez nią i dzięki niej zostały tak naprawdę ukute z praktyczną intencją wykorzystania ich. Píše dalej, że sztuczna inteligencja stanowi projekt ściśle filozoficzny. Skupia on bowiem w sobie całe spektrum problemów ontologicznych i epistemologicznych. Porusza takie problemy, jak co to znaczy być człowiekiem, jaka jest relacja ciała i umysłu oraz świadomości i doświadczenia. Stąd też, jako bodaj jedyna dyscyplina wiedzy, musi wciąż odpowiadać na pytanie: „czy to już?”, „czy już się udało?”. Autorka stara się wykazać, że nurt filozoficzny, za jaki uznaje sztuczną inteligencję, nie musi rozwijać się na drodze przemysłów wybitnych jednostek, ale równie słuszną metodą będzie ciągła korekta rzeczywistości opinii, *doxy*, która to poprzez poruszanie kolejnych wątpliwości o naturze filozoficznej będzie ustawicznie rozsadać się od środka (Przegalińska, 2016, s. 106).

Przegalińska dostrzega te okresy we wspomnianych wcześniej latach i zimach sztucznej inteligencji. Tak nazywane są okresy erupcji entuzjazmu i fale gromkiej krytyki, których bezpośrednim następstwem są inwestycje i cięcia finansowe. Faktem jest, że nigdy nie powstała zapowiadana już przez pierwszych entuzjastów maszyna samoświadoma i inteligentna. Przełom lat 70. i 80. XX wieku był czasem pierwszej zimy sztucznej inteligencji i krytyki wynikającej z rozczarowania obliczeniową teorią umysłu.

Sama teoria wyrastała w trudnych okolicznościach, udowadniając wartość badania umysłu w środowisku zdominowanym przez behawiorystów, którzy odrzucali jako nienaukowe wszelkie rozmyślenia na tematy mentalne. Przegalińska przywołuje w tym kontekście Toma Froese. Stwierdził on, że sam fakt pojawienia się tego programu badawczego, a następnie systematycznego jego rozwoju, jak również skupieniu się na poznaniu jako formach obliczeń, przyciągnął do jego rozwoju specjalistów od technologii komputerowych. Dało to mu przewagę nad rywalizującymi z nim teoriami (Froese, 2011, s. 95-96). Dalej, autorka dodaje, że to właśnie dzięki temu podejściu człowiek przestał być jedynym podmiotem, dla którego zastosowanie znajdowały takie koncepcje jak myślenie, świadomość czy doświadczenie (Przegalińska, 2016, s. 107).

W latach 70. następowała dwustronna wymiana terminologii, technologie informatyczne zaczęły mówić o pamięci komputera, a psychologia opisywała

w podobnym języku umysł. Komputer też przestał być traktowany wyłącznie jako maszyna obliczeniowa i przyjął nową rolę, jak określa ją Sherry Turkle: „obiekty ewokatywnego”. Oznacza to byt nastawiony na wywoływanie przeżyć i refleksji u odbiorcy (Turkle, 1995). Przeżycia te mogły być, naturalnie, pozytywne i wtedy sztuczna inteligencja była traktowana jako przyjaciel lub negatywne, co prowadziło do traktowania jej wrogo i było relatywnie częstszą opcją (Przegalińska, 2016, s. 108). Podobnie pisze Hans Moravec, stwierdzając, że postawy społeczne wobec „mechanicznego życia” wahały się od neutralności, przez obawy aż po entuzjazm (Moravec, 1998; Moravec, 1999). Inteligentny robot zaczął być traktowany jako realny cel, możliwy do osiągnięcia w przeciągu kilku dekad. Jak kontynuuje Przegalińska, sukcesywnie realizowano więc wzorzec kartezjańskiej maszyny, coraz bliższej człowiekowi (Przegalińska, 2016, s. 29).

Jednymi z najczęściej przywoływanych krytyków sztucznej inteligencji byli bracia Dreyfusowie: Hubert i Stuart. Przegalińska przytacza w tym kontekście ich trzy publikacje: *How to Stop Worrying About the Frame Problem Even Though It's Computationally Insoluble* (1987) oraz *Making a Mind Versus Modeling the Brain: Artificial Intelligence Back At a Branchpoint* (1988) i *Mind over Machine* (1986) i, dodatkowo, stwierdza, że bardziej zapiekłym oponentem był starszy z braci, Hubert, i przywołuje jego kolejne utwory: *Alchemy and AI* (1965), *What Computers Can't Do* (1972). Jego argumenty przeciwko obliczeniowej teorii umysłu opierają się na perspektywie fenomenologicznej. Przegalińska zwraca również uwagę, że krytyka ta zdezaktualizowała się, a maszyny zostały udoskonalone w sposób i w stopniu, którego H. Dreyfus nie przewidział (Przegalińska, 2016, s. 108).

H. Dreyfus wytykał badaczom sztucznej inteligencji przede wszystkim, że nigdy nie powstał robot na miarę HALA 9000 z *Odysei Kosmicznej* Kubricka. Przywoływał on też Imre Lakatosa, który określił klasyczny program badawczy sztucznej inteligencji jako „degenerujący się”. Oznaczało to wstępne sukcesy prowadzące do snucia ambitnych wizji przyszłości, by rozbić się o nieprzewidziane wyzwania i niemożliwe do rozwiązania problemy i, ostatecznie, zostać całkowicie porzuconym tak przez badaczy, jak i zwykłych entuzjastów.

Tymi początkowymi sukcesami były dla H. Dreyfusa prace Newella i Simona oraz mikroświaty jak, na przykład, SHRDLU Terry'ego Winograda (1972). Zwłaszcza te ostatnie dawały badaczom złudną nadzieję, że już lada chwila, już za moment, po kilku drobnych poprawkach, rozszerzeniach i połączeniach osiągnięta

zostanie „prawdziwa” sztuczna inteligencja: myśląca, rozmawiająca w językach naturalnych, wpływająca na otoczenie (Przegalińska, 2016, s. 109). Przegalińska nadmienia, że Marvin Minsky w 1967 roku podzielił się przewidywaniem, że taka maszyna powstanie w przeciągu jednego pokolenia (Przegalińska, 2016, s. 109).

H. Dreyfus kontynuuje jednak, że już niedługo później zaczęły się wspomniane kłopoty. Za pierwszy ich przejaw Przegalińska podaje porażki w próbach zaprogramowania maszyn tak, by rozumiały opowiadane historie w 1970 roku. Brakowało im bowiem zdroworozsądkowego rozumienia znaczeń słów. Badacze wychodzili z założenia opartego na kartezjańskich prostych ideach i elementach oraz myśli G. Fregego, że wszystkie pojęcia są zasadami, które można sformalizować (Przegalińska, 2016, s. 109). Stąd też intuicyjnie wychodzili z założenia, że potrzebują jedynie opracować formalne reguły i zbiory faktów, by było to możliwe. Jednakże okazało się to nieporównywalnie trudniejsze i kolejne nieudane próby skłoniły badaczy do określenia problemu sztucznej inteligencji jednym z najtrudniejszych w historii (Przegalińska, 2016, s. 110).

H. Dreyfus odrzucał pogląd, który określił jako dominujący od Platona do Chomsky’ego, wedle którego początkujący zaczyna od poszczególnych przypadków i analizując je, internalizuje rządzące nimi zasady, tworząc w ten sposób reguły o rosnącym poziomie skomplikowania. Zamiast tego postulował, że nauka umiejętności przebiega odwrotnie: od abstrakcyjnych reguł do konkretnych przypadków. Przegalińska stwierdza, że według H. Dreyfusa nasz system poznawczy nie jest dostosowany do ścisłych kategorii, z prostymi do wyznaczenia granicami między nimi. Zamiast tego preferujemy ogólne, elastyczne schematy i pracę na prototypach. Przegalińska dostrzega w tym podobieństwo do arystotelesowskiej „fronesis”. H. Dreyfus, zainspirowany Mauricem Merleau-Pontym, Martinem Heideggerem i Karlem Polanyim uznaje nabywanie umiejętności i uczenie się za kompleksowe relacje, których nie można rozpoznać i opisać w sposób formalny (Przegalińska, 2016, s. 112).

Ostatecznie Przegalińska podsumowuje spojrzenie H. Dreyfusa na uczenie się nazywające je raczej przyswajaniem i wykazując implikacje takiego podejścia. Zwraca ona uwagę, że przyjmując takie „ucieleśnione uczenie się”, jak również je określa, wynika, że z jednej strony ludzie wiedzą więcej, niż podejrzewają, a z drugiej, że, przez tę właśnie nieświadomość, nie przychodzi nam łatwo formalizowanie naszej wiedzy. Co więcej, nasze myślenie nastawione jest na akceptowanie niejasności i przeróżnych permutacji. Autorka przywołuje w tym kontekście H. Dreyfusa, który stwierdzał, że systemy

eksperyckie mogą w najlepszym razie stać się kompetentne, jednakże żadna ilość zasad czy faktów nie uchwyci doświadczenia eksperta nabytego na przestrzeni dziesiątków tysięcy przeżytych i przepracowanych sytuacji (Przegalińska, 2016, s. 113).

Jak kontynuuje Przegalińska, H. Dreyfus odkrył jedno z największych wyzwań sztucznej inteligencji, problem ucieleśnienia i sytuacyjności. Jest on zwodniczo prosty i zasadza się na stwierdzeniu, że inteligencja, w ludzkim jej wymiarze, obejmującym takie procesy jak uczenie się, pamięć, percepcja, wymaga ciała. H. Dreyfus, w ujęciu Przegalińskiej, jako pierwszy podjął wyzwanie wyprowadzenia maszyny z izolacji od środowiska zewnętrznego. Pragnął wystawić go bodźce zewnętrzne, gdyż dostrzegł wpływ interakcji z fizycznym otoczeniem na sposób działania systemu poznawczego. Za cel tegoż systemu nie uznawał bowiem wiernego reprezentowania świata, co skuteczne działanie. Dzięki tejże krytyce, jak zauważa Przegalińska, obrano kierunek uznawania umysłu za coś, co umożliwia świadome i celowe oddziaływanie podmiotu poznającego na bezpośrednie otoczenie. Co więcej, umożliwiła ona zanegowanie pierwotnego ujęcia wirtualności jako przedłużenia kartezjańskiego dualizmu i urealniło sztuczną inteligencję, nastawiając ją na praktyczne działania (Przegalińska, 2016, s. 114).

Inni badacze i filozofowie również zgłaszali wątpliwości i obawy dotyczące dalszego rozwoju sztucznej inteligencji. Kluczowe w tym procesie było materialne odniesienie, jak nazywa je Przegalińska. Rozwój technologii umożliwił poniekąd realizację filozoficznych idei. Komputacjonizm nie był bynajmniej jedyną nową teorią umysłu od czasu Kartezjusza. Przegalińska przytacza tu, chociażby naturalizm, emergentyzm, monizm anomalny i neutralny czy panpsychizm (Przegalińska, 2016, s. 115). Jednakże, jak kontynuuje autorka, to właśnie sztuczna inteligencja rozbudziła wyobraźnię wielu osób i zainspirowała ich do nowych pomysłów. W tym kontekście Przegalińska wymienia: Zenona Pylyshyna (2007), Daniela Dennetta (2006) oraz Davida Chalmersa (1999).

Ich spekulacje krążyły wokół takich tematów jak natura świadomości. Na przykład czy informacje skopiowane z ludzkiego umysłu do komputera byłyby świadome albo, czy przeskanowanie umysłu, a następnie skopiowanie go do identycznego zbioru cząstek to czy ten zbiór miałby również moją świadomość? Pojawił się też uaktualniony problem statku Tezeusza, gdzie zamiast pytania o podmianę desek zastanawiano się nad cyborgizacją poprzez wymianę neuronów na mikroukłady krzemowe (Przegalińska, 2016, s. 116). Jak podsumowuje Przegalińska, poruszane problemy dotyczyły zazwyczaj relacji świadomości do doświadczenia oraz umysłu do ciała, często w kontekście szczeliny eksplanacyjnej, czyli luki poznawczej, której uzupełnienie

pozwoilioby nam stwierdzić, jak mają się świadomość do mózgu czy ciało do umysłu. Istotne były również relacje między umysłami: czy jesteśmy w stanie je poznać lub też, czy przeżywają w ten sam sposób, *qualia* (Przegalińska, 2016, s. 116).

Kolejnym poważnym krytykiem sztucznej inteligencji wymienionym przez Przegalińską jest Roger Penrose, zwłaszcza w kontekście *Nowego umysłu cesarza* (1989). Skupiał się on przede wszystkim na świadomości i jej relacji z nieświadomymi procesami obliczeniowymi realizowanymi w mózgu. Postulował on bowiem, że sążenie rzadko jest przeprowadzane w uporządkowany, algorytmizowalny sposób i stąd też nie może nigdy stać się domeną komputerów. Odwoływał się on do Gödla i jego koncepcji niealgorytmicznej wizji myślenia matematycznego oraz kwantowej natury umysłu ludzkiego (1989, s. 417-418).

Następnymi filozofami wymienionymi przez Przegalińską i oponującymi przed tezami stawianymi w obrębie pola omawianej dziedziny byli Horgan i Tierson (Przegalińska, 2016, s. 117). Ich krytyka dotyczyła problemu ramy i wpływu zewnętrznych ograniczeń na nasze możliwości decyzyjne. W kontekście sztucznej inteligencji dotyczyło to przede wszystkim nabywania wiedzy i biegłości przez maszynę. Problem ramy polega na modelowaniu zmian w obrębie określonej sytuacji, tak by inteligentna maszyna mogła podjąć adekwatne do nich działania. Autorem tej koncepcji był John McCarthy (McCarthy i Hayes, 1969).

Ramy w tym kontekście odnoszą się do prób rozwiązania tego problemu za pomocą aksjomatów ramowych. Tworzono określone ramy odniesienia, które pozwalały maszynom stwierdzić jakie efekty nie mogą być wywołane przez konkretne działania. Na przykład żadna ilość przemieszczania przedmiotu nie wpłynie sama z siebie na zmianę koloru tego przedmiotu. Szybko jednak okazało się, że rozwiązanie to nie jest praktyczne. Ilość aksjomatów musiała być niemal równa ilości egzemplarzy działań, a samo ich zapisywanie było nie dość, że czasochłonne to jeszcze negatywnie wpływało na wydajność systemów. Aksjomaty zupełnie się też nie sprawdzały w systemach, które uwzględniały możliwość jednoczesnych działań wielu podmiotów. Przegalińska stwierdza zresztą, że problem dostosowania zachowania maszyny do ustawicznie i dynamicznie zmieniających się warunków wciąż pozostał nierozwiązany (Przegalińska, 2016, s. 118).

Ostatnim filozofem przywołanym przez Przegalińską jest John Searle, który w swojej książce *Minds, Brains and Programs* opisał argument „chińskiego pokoju” (1980). Stanowił on eksperyment mentalny, który wykazywał, że operowanie wyłącznie na syntaktyce nie doprowadzi do wykształcenia się własności semantycznych.

„Funkcja”, jak określa ją Przegalińska, więc, wbrew postulatowi funkcjonalizmu, nie jest niezależna od świata rzeczywistego, lecz zawsze stanowi konstrukt poddawany interpretacji przez podmiot ją realizujący (Przegalińska, 2016, s. 119).

Komputacjonizm nie był jedyną próbą zbudowania maszyny myślącej. O ile podejście to sprawdzało się w sytuacjach jednoznacznych i pewnych, tak poza zasięgiem były takie działania, jak dowodzenie intuicyjnie, wnioskowanie w warunkach niepewnych, czyli opartych na nieznanych przesłankach, dokonywanie odkryć oraz, wspomniane wcześniej, działania w świecie realnym (Przegalińska, 2016, s. 121).

Koncepcja umysłu jako programu wciąż jednak budzi sprzeciw wśród kognitywistów, gdyż opiera się na określeniu ludzkiego podmiotu jako prostego, jak również biernego konwertera symboli i znaczeń. Krytycy, zwłaszcza powiązani z fenomenologią jak H. Dreyfus, podkreślali, że „skomputeryzowane” podejścia do umysłu pomijają ludzką zdolność do samodzielnego nadawania znaczeń. Przegalińska potwierdza ten zarzut, stwierdzając, że klasyczna, obliczeniowa teoria umysłu nie dotyka istoty uczenia się i sprawdza się jedynie w przypadkach opartych na prostych regułach, jak dowodzenie dedukcyjne czy sylogistyczne (Przegalińska, 2016, s. 123). Podobnie stwierdza Marek Hetmański, pisząc, że nasza wiedza o budowie mózgu i działaniu umysłu uniemożliwia nam trwanie przy koncepcjach z początków badań kognitywistycznych (Hetmański, 2005). Kolejnym zarzutem przytoczonym przez Przegalińską jest, że żaden z paradygmatów nie przybliżył nas do ustalenia relacji umysłu i ciała (Przegalińska, 2016, s. 124).

Określeniem wpływu struktury cielesnej na postawy podmiotu względem siebie samego i otoczenia pojawiły się jeszcze na gruncie cybernetyki. Przegalińska odwołuje się w tym kontekście do Gregory’ego Batesona przywołując jego stwierdzenie, że informacja w działaniu ludzi i zwierząt to sygnał otoczenia będący przyczynkiem do reakcji i działania, które równie dobrze może być realizowane cieleśnie, jak i mentalnie, poznawczo. Sygnały są zróżnicowane, a organizmy uczą się je rozpoznawać. Następnie utrwalają je w formie odpowiadających im reakcji. Informacja jest istotna dla człowieka, gdy jawi się jako anomalia w dotychczasowych sygnałach. Rejestrujemy te zaburzenia tak anatomicznie, jak i funkcjonalnie, po czym aktualizujemy posiadane wzorce czy schematy określające nasze oczekiwania. Tak działa zarówno percepcja, jak i również pamięć, mowa oraz myślenie abstrakcyjne (Przegalińska, 2016, s. 125).

Stąd też informacja staje się wtórna wobec własności otoczenia. Oparta jest na sygnałach przetwarzanych na podstawie schematów zapisanych na poziomie anatomicznym. Przegalińska przytacza w tym kontekście Szymona Wróbla, który stwierdza, że mózg, według ówczesnych badań neurofizjologicznych, stanowi nie tyle automat bodziec-reakcja, ile system zamknięty i tworzący własne pojmowanie świata. W takim ujęciu bodźce sensoryczne stają się katalizatorem zmian mózgu, a ich wytworzone reprezentacje zależą od tego, w jakim stanie zastały określony mózg. Interakcje ze światem zewnętrznym to, w ujęciu Wróbla, celowe sterowanie układem ruchowym w poszukiwaniu nowych wrażeń zmysłowych. Wróbel podsumowuje, stwierdzając, że w opisanym przez niego ujęciu akty percepcji stanowią realizowanie procesu samoorganizacji mózgu na drodze poszukiwania kontaktów z otoczeniem, a świadomość stanowi subiektywne doznanie współwystępujące z tym działaniem (Wróbel, 1997).

W kontekście badań neuronalnych Russell i Norvig stwierdzają, że współcześnie mamy niepomiarowo więcej danych, dzięki czemu możliwym stało się mapowanie obszarów mózgu na kontrolowane przez nie rejony ciała lub zasilane odpowiednim wkładem sensorycznym. Należy dodać jednak, że jakkolwiek dokładne, tak mapowania te mają naturę cokolwiek dynamiczną i, jak zauważają Russell i Norvig, mogą zmieniać się w ciągu kilku tygodni. Jest to tym bardziej skomplikowane, gdyż niektóre gatunki mają symultanicznie więcej niż jedno takie przypisanie. Następne utrudnienie stanowi tak zwana plastyczność mózgu. Jedne obszary mózgu mogą przejmować funkcje innych w sytuacjach wynikających z uszkodzenia tkanek, które normalnie odpowiadają za ich realizację. Jest to niezmiernie pożyteczna własność tego organu, jednakże zrozumienie sposobu działania tej funkcji i kryterium, dla którego te, a nie inne tkanki zastępują uszkodzone, stanowi kolejną, wciąż nierozwikłaną, zagadkę. Listę wyzwań stojących przed neuronaukami Russell i Norvig zamykają tematem indywidualnej pamięci i mechanizmów jej przechowywania i odzyskiwania (Russell i Norvig, 2010, s. 10).

Obecnie odnotowujemy również dynamiczny rozwój dyscypliny dzięki wykorzystaniu tak zwanych funkcjonalnych rezonansów magnetycznych (fMRI) (Ogawa, Lee, Kay i Tank, 1990; Cabeza i Nyberg, 2001). Umożliwiają one uzyskanie bezprecedensowej dokładności w śledzeniu aktywności mózgu- na poziomie umożliwiającym stosowanie obliczeń i pomiarów, które w pozytywny sposób korelują z zachodzącymi procesami kognitywnymi. Podobnie, w kontekście technik śledzenia aktywności neuronalnej- mamy teraz też możliwość monitorowania pojedynczych

neuronów. Możemy je stymulować elektrycznie, chemicznie, nawet optycznie (Han i Boyden, 2007). Techniki te pozwalają na mapowanie połączeń neuronalnych typu wejście-wyjście.

Russell i Norvig podsumowują jednak, że pomimo tych licznych osiągnięć wciąż jesteśmy bardzo daleko od faktycznego zrozumienia mechanizmów procesów poznawczych. Kontynuują oni, stwierdzając, niejako zdumieni, że to niesamowite, jak zbiór prostych komórek może prowadzić do realizowania myśli, akcji, świadomości. Przywołują oni tutaj Johna Searle'a i jego książkę, w ramach której dowodzi, że mózgi tworzą umysły (Searle J. R., 1992). Autorzy kończą podsumowanie, konstatując, że jedyną alternatywą dla neuronauk stanowią podejścia mistyczne, które zakładają realizowanie procesów myślowych w wymiarze niedostępnym dla nauk fizycznych (Russell i Norvig, 2010, s. 10).

Sztuczna inteligencja więc ostatecznie jawi się jako etap na długiej drodze filozoficznych myśli, które zbiorczo umożliwiły jej konceptualizację. W szczytowym momencie jej popularności doprowadziła wręcz do nowej metafory umysłu, komputera. Doprowadziło to do szeregu konsekwencji, w tym nowych gałęzi psychologii czy nowych dylematów filozoficznych. Ostatecznie jednak wzbogaciliśmy je, gdyż jak podsumowuje Przegalińska, funkcjonowanie informacji odbywa się wielopoziomowo i wieloaspektowo. Daleko temu do sprowadzenia tego złożonego procesu do „przetwarzania informacji”. Antyobliczeniowy model poznania zakłada, że człowiek jako układ informacyjno-poznawczy jest samoorganizujący się otwarty. W teoriach ucieleśnionego poznania używa się określenia aktywnego eksplanatora, gdyż stosunek systemu poznawczego do napływających z otoczenia bodźców jest właśnie aktywny, rejestruje on wszelkie zachodzące zmiany i przetwarza je w informacje. Inną różnicę, jak konstatuje Przegalińska, jest funkcjonująca metafora, w ramach której zastąpiono komputer relacjami. Opisują one sprzężenia strukturalne w środowisku, czyli działania inicjujące procesy poznawcze. Stanowi to znaczącą różnicę z modelem obliczeniowym, który zakładał bierność poznania. Co więcej, aktywne poznanie oznacza, że wymagana jest dokładna analiza współzależności zachodzących między ciałem, umysłem a otoczeniem. Kolejna istotna różnica wymieniona przez Przegalińską dotyczy problemu reprezentacji symbolicznej. O ile w podejściu reprezentacja była kluczowym pojęciem i zakładano, że w naszym umyśle mamy reprezentacje zakodowane za pomocą symboli i dekodowane na gruncie operacji logicznych, tak w ucieleśnieniu reprezentacje albo są całkowicie odrzucane, albo akceptowane są proste reprezentacje sensomotoryczne.

4. Część II – dwa źródła koncepcji sztucznej inteligencji

4.1. Społeczność umysłów

4.1.1. Wprowadzenie

Marvin Minsky, wielokrotnie wspomniany w poprzednich rozdziałach, zostaje tu przywołany jako reprezentant obliczeniowej teorii umysłu i zwolennik tworzenia inteligentnych maszyn. Przyjmuje się, że agenci tworzeni zgodnie z jego pojmowaniem umysłu będą zdolni do porozumiewania się z ludźmi w ich języku naturalnym, ale również będą w stanie przeprowadzać mentalne operacje prowadzące do tworzenia koncepcji służących do rozwiązywania problemów i, konsekwentnie, rozwijania samych siebie (McCarthy, Minsky, Rochester i Shannon, 2006).

M. Minsky najszerszej przedstawił swoją wizję sztucznej inteligencji w książce *Społeczność umysłów* (org. Society of Mind, tłumaczenie własne). Publikacja ta ukazała się w 1986 roku, nakładem wydawnictwa Simon and Schuster, w Nowym Jorku. Zawiera ona zestaw obserwacji i myśli, każda z nich ograniczona do pojedynczej strony. Nie ma w niej przypisów czy bibliografii (poza tą w glosariuszu), poza tym, co zostało ujęte w załączonym do niej słowniku pojęć.

Zasadniczo celem M. Minsky'ego było opisanie, jak funkcjonuje umysł, przede wszystkim ludzki, ale zakładał uniwersalność swojego podejścia i postulował, że znajdzie on w przyszłości zastosowanie również w odniesieniu do wspomnianych bytów syntetycznych. Ujmując to dokładniej, porusza ona szereg tematów jak świadomość, emocje, język czy żarty i stąd też bywa porównywana do wittgensteinowskich *Dociekań filozoficznych*, gdyż celem nie jest przekazanie czytelnikowi konkretnych informacji, lecz raczej zainspirowanie go nowatorskimi pytaniami (Dyer, 1991, s. 321).

Wright określa tę książkę najlepszym przykładem społecznej metafory zaaplikowanej do zrozumienia i projektowania umysłów. Zwraca uwagę, że zarysowuje ona obliczeniową społeczność niejednorodnych agentów, którzy współpracują i rywalizują, i w ten sposób wytwarzają zdolności mentalne. Podejście, w ramach którego przeprowadzona zostaje dekompozycja obliczeniowego umysłu do społeczności mniej inteligentnych agentów, jest według niego przekonujące, gdyż współbieżne systemy społeczne i duże systemy obliczeniowe są projektowane w podobny sposób (Wright, 2000, s. 113).

Jako przykłady podaje fakt, że oba systemy składają się ze wzajemnie połączonych i wchodzących ze sobą w interakcję komponentów zdolnych do wykonywania określonych zadań, jednostki obliczeniowe przetwarzają dane, a ludzie wykonują pracę. Co więcej, zarówno agenci obliczeniowi, jak i ludzie występują jednocześnie w rolach producentów i konsumentów, czy to informacji, czy fizycznych produktów. Agenci również przyjmują specjalizacje, zawężając zakres swoich działań, ale i podnosząc ich jakość. W obu systemach agenci działają niezależnie i próbują osiągać własne cele. Jednakże, poprzez wchodzenie ze sobą w interakcję osiągają współpracę i ostatecznie możemy ich określić jako działających współbieżnie. Obie społeczności, które wytwarzają agenci, wymagają też pewnej dozy koordynacji za pomocą mechanizmów produkcji, dystrybucji i konsumpcji, co odbywa się czy to na wolnych rynkach, czy na relacyjnych bazach danych. Ostatnia wymieniona przez Wrighta zbieżność polega na dostrzeżeniu i docenieniu roli adaptacyjnego zarządzania niedostateczną ilością zasobów (Wright, 2000, s. 113).

4.1.2. Ogólna struktura

Dla Minsky'ego każdy umysł, że jest niezwykle bogatym i wielopłaszczyznowym zbiorem struktur i procesów, z których każdy ma za sobą setki tysięcy lat ewolucji, tysiące lat zmian kulturowych i lata osobistego doświadczenia. Nie powinniśmy jednak dążyć do ich odtworzenia. Zamiast tego wyzwaniem leży w stworzeniu zorganizowanych systemów zdolnych wspierać dostatecznie dużo niezależnych i różnorodnych schematów oraz umożliwić im wykorzystywanie nawzajem swoich możliwości (Minsky, 1990).

Minsky rozważa olbrzymią rozpiętość tematów, od opisu najprostszycch procesów mentalnych do propozycji architektonicznych opisujących organizację całego umysłu. Jak podsumowuje Push Singh, Minsky nie szuka uniwersalnego mechanizmu, który opisz wszystkie możliwe procesy mentalne. Zamiast tego Minsky postuluje liczne i różnorodne mechanizmy, które będzie można zaadoptować do opisu dowolnego takiego procesu mentalnego (Singh, 2003, s. 522).

Minsky przedstawia pewną strukturę bytów inteligentnych, które wchodzą ze sobą w interakcje. Co istotne, mają one nie tylko ściśle określone zadania, ale również zamiary. Za przykład może posłużyć jednostka realizująca chęć trzymania zabawki przez dziecko. Skupia się ona na aktywność trzymania właśnie i z własnej inicjatywy pragnie tylko i wyłącznie trzymać tę zabawkę. Dopiero zmuszona przez innych agentów zaprzestanie swojej aktywności i wtedy dziecko tę zabawkę wypuści.

Stąd też konieczne jest, by relacje tychże podmiotów przypominały interakcje międzyludzkie, na przykład: aby wywierali oni na siebie nawzajem wpływ, negocjowali czy formowali hierarchie, gdyż to z nich wyłaniają się zachowania właściwych jednostek (czyli znów ludzi lub syntetycznych agentów). Umysły w ujęciu Minsky'ego są złożone z tychże agentów i ich, podyktowanych ściśle określonymi motywacjami, działań oraz powiązaniem między nimi tworzącymi określone struktury decyzyjne.

Minsky, by określić te najbardziej podstawowe indywidua, posługuje się właśnie określeniem agenta. Każdy z nich jest efektywnie na poziomie skomplikowania pojedynczego komponentu komputerowego programu. Agenci mogą być łączeni i tworzą wtedy tak zwane społeczności agentów (ang. societies of agents, tłumaczenie własne). Społeczności te mają własne motywacje i możliwości działania większe niż indywidualni agenci. Ponadto są zorganizowane wokół określonych hierarchii, mogą ze sobą współpracować (jak również, naturalnie, rywalizować) i razem wykonywać bardziej skomplikowane operacje niż pojedynczo.

Warto jednak zauważyć za Pushem Singhem, że metafora społeczeństwa jest bardzo inspirująca i trafna. Zwraca on uwagę na towarzyszące jej idee zróżnicowania i składania się umysłów z licznych, różnorodnych bytów o mnogich celach i możliwościach. Z drugiej jednak strony, jak kontynuuje P. Singh, gdy myślimy o społecznościach, przychodzą nam do głowy ludzie, a ci są dalece bardziej skomplikowani od agentów w ujęciu Minsky'ego, którzy to są bardzo mocno wyspecjalizowani. Tak więc P. Singh postuluje, by metaforę tę traktować jako punkt wyjścia, a nie dokładną instrukcję rozumienia społeczności umysłów (Singh, 2003, s. 522).

4.1.3. Historia idei

Jak zauważa P. Singh, teoria społeczności umysłów narodziła się podczas dyskusji między Minskym a Seymourem Papertem we wczesnych latach 70. XX wieku, w Artificial Intelligence Lab na MIT. Był to jeden z kluczowych ośrodków badań nad sztuczną inteligencją. Rozwijano tam takie technologie jak uczenie maszynowe, reprezentacja wiedzy, robotyczna manipulacja, przetwarzanie języka naturalnego, wizja komputerowa i wnioskowanie zdroworozsądkowe. Stąd też, jak kontynuuje P. Singh, była to społeczność wysoce zainteresowana procesami kognitywnymi (Singh, 2003, s. 523).

Ściślej rzecz ujmując, pierwsze pomysły w kierunku formułowania tej teorii przysły Minsky'emu podczas pracy nad pierwszym autonomicznym robotem łączącym

oko i rękę w latach 60. XX wieku. Wtedy to napisał wraz z S. Papertem i grupą studentów programy takie jak Porusz (org. Move, tłumaczenie własne), Zobacz (org. See, tłumaczenie własne), Złap (org. Grasp, tłumaczenie własne) i setki innych programików niezbędnych do działania robota albo, jak określa to Minsky, stworzyli społeczność agentów o nazwie Budowniczy (org. Builder, tłumaczenie własne; Minsky, 1986, s. 29).

W istocie żaden pojedynczy algorytm ani metoda nie były w stanie rozwiązać, wydawałoby się najprostszego problemu, jak budowa wieży z klocków (Singh, 2003, s. 523). Minsky pisze o tym, że roboty nie były w stanie rozpoznać kształtu obiektu, opierając się wyłącznie na zmyśle wzroku i musieli tworzyć rozwiązania oparte na wiedzy, dzięki której roboty mogły wnioskować i tworzyć wewnętrzne reprezentacje. Wtedy też oczywistym stała się konieczność przepływu danych między programami, czy też agentami, którzy musieli współpracować, by je wytworzyć (Minsky, 1986, s. 324).

Ostatecznie, jak pisze P. Singh, te doświadczenia doprowadziły Minsky'ego i S. Paperta do zostania orędownikami poglądu, że inteligencja nie jest rezultatem prostego zestawu kroków. Zamiast tego widzieli oni konglomerat wyspecjalizowanych procesów mentalnych, które ustawicznie wchodzą ze sobą w interakcje. Jednakże, w tamtym czasie nie było zbyt wielu rozwiązań, które umożliwiały tworzenie systemów modelujących symultanicznie działanie tysięcy różnorodnych procesów mentalnych (Singh, 2003, s. 523).

Wczesne myśli Minsky'ego na ten temat przebijają się również w jego artykule *A Framework for Representing Knowledge*. Rozważa w nim różne metody organizowania wiedzy proceduralnej i deklaratywnej niezbędnej do rozwiązywania problemów, które możemy określić mianem zdroworozsądkowych (Minsky, 1974).

Zauważa tam na przykład, że części umysłu odpowiedzialne za wnioskowanie, język, pamięć i percepcję muszą być większe i lepiej ustrukturyzowane, a ich rzeczowe i proceduralne aspekty silnie powiązane. Minsky kontynuuje, że tylko w ten sposób można wyjaśnić szybkość i moc powiązanych z nimi aktywności mentalnych (Minsky, 1974, s. 1).

W dalszej części raportu Minsky pisze, że inną inspirację stanowił dla niego esej studenta Scotta Fahlmana, który przedstawił koncepcję ramowania jako zbioru faktów i społeczności agentów na nich pracujących. Dowolna liczba ram mogła być aktywowana naraz i sukcesywnie uruchamiała tak inne ramy (org. frames, tłumaczenie własne), jak i zawarte w sobie pod-ramy (Minsky, 1974, s. 57). Ogólnie, P. Singh stwierdza, że cała koncepcja ram stanowi przodka czy też inspirację dla teorii społeczności umysłów (Singh, 2003, s. 524).

Obliczeniowe podejścia do społeczności współpracujących agentów zaczęły się pojawiać w połowie lat 70. XX wieku. Wczesnym przejawem tych idei była inna koncepcja, która również stanowiła inspirację dla Minsky'ego. Tak zwana teoria aktorów autorstwa Carla Hewitta. Stanowiła ona model obliczeniowy, w którym zbiór obecnie aktywnych agentów rozwiązywał problemy poprzez wymianę wiadomości (Hewitt, 1976). Jednakże, jak zauważa P. Singh, idea wymiany wiadomości w języku naturalnym byłaby zbyt obciążająca dla idei Minsky'ego i potrzebne były mniej bezpośrednie formy komunikacji (Singh, 2003, s. 525).

Minsky i S. Papert zaczęli pisać książkę o teorii społeczności umysłów już w 1976 roku, ale zdali sobie sprawę, że ich pomysły nie dojrzały jeszcze do tej formy. Jednakże, chociaż idee te były już rozwinięte i wykorzystane w końcowej książce, to współpraca między autorami zamarła, a Minsky rozwijał teorię sam. Jej fragmenty możemy dostrzec w artykule z kolejnego roku (1977) *Plain Talk about Neurodevelopmental Epistemology*, gdzie pisze o kontynuacji artykułu z 1974 roku oraz współpracy z S. Papertem (Singh, 2003, s. 525).

Minsky wymienia w nim również swoje nowe źródła inspiracji, w tym Sigmunda Freuda i Jeana Piageta. Pojawia się również idea organizowania społeczności w hierarchie. Ostatecznie, spójność osobowości jest dla Minsky'ego emergentną własnością interakcji opisanych genetyczną kontrolą, konfliktem między agentami, którzy chcą działać oraz ich krytykami i cenzorami. Ostateczny kształt społeczności agentów jawi się Minsky'emu niejako freudowsko, z ciągłą koniecznością praktykowania dyscypliny i porównywaniu się z obrazami samych siebie wytworzonymi na wcześniejszych stadiach rozwoju (Minsky, 1977, s. 1091).

Później Minsky pisze, że stany mentalne w społeczności umysłu nie są, jak mogłoby się wydawać, pojedynczymi, jednolitymi bytami. Zamiast tego, Minsky widzi je jako podzbiory stanów konkretnych części całego umysłu. Rozwija tę ideę poprzez porównanie umysłu do znanych nam instytucji publicznych, urzędów.

Na samej górze mamy podział na specjalne „działy” odpowiadające takim obszarom jak procesowanie sensoryczne, język, długofalowe planowanie i im podobne. W każdym z nich mamy rzesze specjalistów, których nazywamy agentami. Ci bowiem odpowiadają za konkretne jednostki wiedzy, umiejętności i metod. Żaden z nich nie jest w stanie dużo zdziałać samemu, ale każdy ma swoje miejsce w odpowiednich konfiguracjach (Minsky, 1980).

Minsky w swoich pracach patrzy na te procesy z bardzo wysokiego punktu widzenia i sam dostrzega, że osoby zapoznające się z tymi artykułami domagają się szczegółów i metod realizacji jego pomysłów. Odpowiada on jednak na te zarzuty i stwierdza, że jego prace nie stanowią gotowych rozwiązań, a są raczej źródłem inspiracji dla innych badaczy, a sam problem określenia tychże metod umyka ogólnie środowisku naukowemu, czy to w dziedzinie sztucznej inteligencji, psychologii, lingwistyki czy nauk o mózgu (Minsky, 1977, s. 1085).

4.1.4. Jak działa społeczność umysłów?

P. Singh zauważa, że nie ma zbyt wielu implementacji teorii Minsky'ego. Problemów upatruje w tym, że Minsky opisuje ją we fragmentach i na różnych poziomach abstrakcji, a bardziej mechaniczne jej aspekty są również opisane w bardzo rozproszony sposób. Stąd też podsumowuje ją klarownie i w jednym miejscu (Singh, 2003, s. 526).

Kim są agenci?

Agenci są dla Minsky'ego skrótem myślowym, metaforą wynikającą z niedostatków języka angielskiego. Przyznaje on, że poddał recyklingowi stare słowa agent i agencja. Potrzebował określić niepojedynczą część mechanizmu a proces, jaki ona wykonuje. Tymże procesem jest właśnie agent. Dodatkowo procesy te muszą być łączone, by osiągnąć coś konkretnego i dopiero analizowanie wyników takich łączonych procesów jest sensowne. Stąd też potrzeba określenia tejże formy zbiorczej i powołanie określenia agencji (Minsky, 1991, s. 391).

P. Singh powtarza i upraszcza to. Pisze bowiem, że agent oznacza dowolny komponent procesu kognitywnego, a agencja tymczasem opisuje społeczność agentów działających razem. Dzięki temu są one w stanie realizować funkcje bardziej skomplikowane niż pojedynczy agenci. Jednakże, jeśli zignorujemy ich wewnętrzną strukturę i skupimy się na widocznych z zewnątrz efektach, wiele agencji może wraz zostać uznany za pojedynczych agentów (Singh, 2003, s. 527).

Dokładniej rzecz ujmując, jak postuluje P. Singh, umysł jest dla Minsky'ego amalgamatem wysoce złożonych procesów kognitywnych, z których każdy jest wyspecjalizowany w realizowaniu konkretnej funkcji. Przytacza on listę przykładów: przewidywanie, oczekiwanie, naprawianie, aktorstwo, porównywanie, upraszczanie. Lista ta naturalnie nie jest wyczerpująca, zresztą Minsky również takowej nie dostarcza, podając przykłady agentów jedynie dla zilustrowania konkretnych idei. Agentom brak też wspólnych mianowników czy aspektów ich łączących. Każdy z nich jest osobnym

procesem, o odrębnych celach, sposobach reprezentowania wiedzy czy metod wnioskowania (Singh, 2003, s. 527).

Aktywność mentalna oznacza u Minsky'ego efektywnie, tylko i aż, aktywowanie i dezaktywowanie agentów. W dowolnym momencie jedynie część z nich będzie aktywna, tworząc w ten sposób totalny stan umysłu. Jednakże, w tym samym czasie może być jednocześnie uaktywnionych wiele aktywności realizowanych przez różne agencje. Stąd też Minsky wprowadził termin częściowego stanu umysłu, aby opisać te właśnie aktywne podzbiory agentów, które realizują konkretny proces, który w danym momencie stanowi obiekt jego zaciekawienia (Singh, 2003, s. 527).

Jakie są typy agentów?

Minsky wyróżnił kilka rodzajów podstawowych agentów, poświęcając uwagę zwłaszcza liniom K, które dalej dzielą się na Nemy i Nomy. Dodatkowo wyróżnia on ich jeszcze bardziej wyspecjalizowane wersje.

Linie K (org. K-Lines, tłumaczenie własne) stanowią najpowszechniejszej występujących agentów. Ich głównym celem jest pobudzanie aktywności określonych zbiorów agentów. Przez wzgląd na liczne połączenia między agentami i agencjami, ich aktywacje mogą „kaskadować” i pobudzenie jednego spowoduje aktywację również innych. Liczne Linie K mają też wbudowaną „pamięć”, więc jeśli umysł napotyka na znany już problem, to agenci tego typu będą mogli uruchomić naraz całą, kaskadowo połączoną, sieć agentów i agencji.

Linie K uwzględnią w takiej sytuacji również przypadki fałszywych pozytywów i inne lekcje wyniesione z poprzednich pobudzeń. Za przykład podaje agentów, których, jak się okazało, nie było potrzeby uruchamiać, gdyż utrudniali oni pracę w poprzednich doświadczeniach. Co więcej, w każdym przypadku możliwe są niespodziewane odkrycia. Stąd też, jak podsumowuje P. Singh, Linie K stanowią potężny i prosty mechanizm wyzwalający określone konfiguracje agentów w odpowiedzi na konkretne potrzeby umysłu (Singh, 2003, s. 528).

Nemy (org. Nemes, tłumaczenie własne) i Nomy (org. Nomes, tłumaczenie własne) to bardziej ustrukturyzowany i podobny do komputera model opisujący jak informacje są procesowane i reprezentowane w społeczności umysłu.

Nemy odpowiadają za tworzenie reprezentacji świata realnego w umyśle. Minsky wyróżnia kilka ich typów, jak chociażby Polynemy, które odpowiadają za wybudzenie agencji odpowiedzialnych za rozpoznawanie kolorów, kształtu, smaku,

ale też na przykład sytuacji, kiedy możemy zjeść reprezentowany obiekt czy jego koszt w sklepie. Polnemy egzemplifikują myśl Minsky'ego, że dany obiekt nie ma pojedynczej reprezentacji, a raczej cały ich rozproszony szereg.

Mikronemy dostarczają globalnego, kontekstowego sygnału dla wszystkich agencji w danym umyśle. Minsky podkreśla ich aspekt w determinowaniu przez umysł niewerbalnych komponentów reprezentacji takich jak skojarzenia emocjonalne czy rozproszone wrażenia sensoryczne jak na przykład śladowe wrażenia zapachowe. Mikronemy również odpowiadają za aspekty sytuacji, których nie możemy przypisać do konkretnego obiektu, a jednak wpływają na naszą percepcję (Singh, 2003, s. 528).

Nomy z kolei odpowiadają za kontrolowanie sposobów procesowania i manipulowania reprezentacjami. Minsky wyróżnił kilka ich typów, między innymi izonomy, pronomy czy paranomy.

Izomy wysyłają sygnały do różnych agencji, aby te wykonały te same typy operacji kognitywnych. P. Singh przywołuje przykład zapamiętania obecnego stanu w pamięci krótkotrwałej i uruchomienia innego stanu lub też wymuszenia treningu nowej, długoterminowej Linii K, która będzie w stanie odtworzyć obecny stan.

Pronomy to podrodzaj izonomów, które kontrolują reprezentacje w układach pamięci krótkotrwałej. Są one często powiązane z określoną rolą w perspektywie większej sytuacji czy wydarzenia, jak choćby podjęcie akcji przez aktora czy skojarzenie miejsca, gdzie wydarzenie zachodzi. Niektóre pronomy łączą się z bardzo ograniczonymi typami pamięci krótkotrwałej, która przechowuje konkretne fragmenty wiedzy, jak miejsca, kształty czy drogi do zapamiętanych przez nas miejsc. Inne, zdecydowanie mniej liczne, pronomy mogą być stosowane ogólniej i pobudzać liczne agencje powiązane z praktycznie dowolnymi reprezentacjami.

Ostatni wymieniony przez P. Singha typ nomów to paranomy. One również stanowią podzbiory pronomów, zwłaszcza takich, które są ze sobą nawzajem powiązane albo z sytuacjami, gdzie pronomy uaktualniają reprezentacje i aktywizują się nawzajem by uwzględnić zachodzące w nich zmiany. Minsky wprowadza paranomy, aby opisać, w jaki sposób wiedza reprezentowana na różne sposoby może wciąż być traktowana odniesiona do czegoś jednolitego i jednostkowego. Za przykład może posłużyć orientowanie się w przestrzeni dzięki paranomowi miejsca. Taki paranom może aktywować tak pronomy lokalizacji przestrzennej w kontekście systemu zorientowanego z perspektywy ciała aktora i równocześnie pronomy zewnętrznej lokalizacji, jak określa P. Singh, z perspektywy trzeciej osoby. Dzięki paranomom jesteśmy w stanie

błyskawicznie przełączać się w naszym myśleniu między tymi i kolejnymi reprezentacjami (Singh, 2003, s. 529).

Jak agenci łączą się w agencje?

Minsky podaje kilka przykładów, jak agenci mogą się łączyć w większe struktury, czyli wspomniane wyżej agencje, aby realizować bardziej skomplikowane cele. Opisuje również jak większe systemy, oparte na reprezentacjach, takie jak Ramy (org. Frames, tłumaczenie własne) mogą być skonstruowane z prostych elementów, jak wspomniane pronomy i agenci wyspecjalizowani w operacjach rozpoznawania obiektów czy zdarzeń.

Rama stanowi formę reprezentacji wiedzy skupioną na konkretnym obiekcie lub konkretnych jego własnościach, które zajmują miejsce w jej otworach (org. slots, tłumaczenie własne). Minsky opisuje, w jaki sposób proste ramy mogą być zbudowane ze zbiorów pronomów, które kontrolują dane, możliwe do przyłączenia w otworach ramy. Te pronomy są wzajemnie połączone w taki sposób, że dowolne przywołanie określonej ramy wywołuje cząstkowe reprezentacje powiązane z pronomami i, sukcesywnie, cząstkowe opisy percypowanego obiektu, który jest reprezentowany.

Szereg ram (org. frame-array, tłumaczenie własne) to zbiór pojedynczych ram, który pozwala na dużo dokładniejszy i bardziej wyczerpujący opis obiektu, gdzie każda rama odpowiada za określoną stronę lub fragment reprezentacji. P. Singh przywołuje tutaj Minsky'ego, który opisywał szeregi ram, odnosząc się w ten sposób do kilku połączonych ram, które wspólnie opisują kostkę.

Każda ściana sześcianu była powiązana z inną ramą, ale miały wspólne obszary, które je łączyły. Dokładniej, Minsky określał te zbieżne ramy jako posiadające wspólne otwory. Ich połączenia umożliwiają też przełączanie się między ramami, aby znaleźć potrzebne informacje. Współdzielone otwory są tu też określone jako przodkowie paranom.

Ramy transformujące (org. transframe, tłumaczenie własne) to centralne formy reprezentacji wiedzy w teorii społeczności umysłu. Reprezentują one wydarzenia i wszystkie byty zaangażowane lub powiązane z tymi reprezentacjami. Mogą mieć otwory opisujące pochodzenie oraz cel zmiany, czyli ich stany przed i po. Dodatkowo śledzą one kto był zaangażowany w wywołanie wydarzenia, motywację za nim stojącą, cel, jaki ma osiągnąć, jakie obiekty były zaangażowane, jak i kiedy nastąpiły czy też jakie narzędzia i obiekty brały udział. Przez wzgląd na te rozliczne, codzienne zdarzenia

oraz relacje między wydarzeniami różnych typów ramy transformujące są kluczowe dla naszego powszedniego myślenia.

Inne typy ram, w tym ramy historyczne (org. story-frames, tłumaczenie własne), są opisywane przez Minsky'ego jako reprezentacje kolekcji powiązanych wydarzeń lub przestrzennego rozmieszczenia obiektów w trakcie określonej sceny. Pojawia się tu założenie, że stanowią one jedynie kilka z wielu typów ram potrzebnych do skutecznego reprezentowania i organizowania bodźców napływających ze świata i procesów kognitywnych samych w sobie (Singh, 2003, s. 530).

Jak agenci rozwiązują problemy?

Następnie P. Singh przytacza kilka metod, jak według Minsky'ego mogą powstawać i organizować się agencje rozwiązujące problemy.

Maszyna różnicowa (org. Difference-Engine, tłumaczenie własne) to rozwiązanie redukujące różnice między stanem faktycznym a docelowym. Najpierw je identyfikuje, a następnie aktywuje odpowiednie linie K przypisane konkretnym metodom rozwiązań (Singh, 2003, s. 531).

Cenzorzy (org. censors, tłumaczenie własne) i Dławiki (org. suppressors, tłumaczenie własne) odpowiadają za naszą wiedzę o najpowszechniejszych pułapkach i pomyłkach w rozwiązywaniu problemów. Minsky określa to negatywną ekspertyzą (org. negative expertise, tłumaczenie własne; Minsky, 1994). Cenzorzy tłumią aktywność mentalną prowadzącą do nieproduktywnych lub destruktywnych zachowań. Dławiki z kolei powstrzymują te same zachowania (a nie prowadzącą do nich aktywność mentalną).

Zastanawiająca jest w tym rola humoru. Według Minsky'ego głównym jego mechanizmem jest uczenie się poprzez zrozumienie, do czego prowadzi niepowstrzymanie działań (Singh, 2003, s. 531). Innymi słowy, kiedy słyszymy żart, możemy uczyć się o konkretnej pułapce lub problemie za pomocą wnioskowania opartego na wypaczeniu zdrowego rozsądku. Minsky dokładnie się temu przygląda w artykule „Jokes and the Cognitive Unconscious” (Minsky, 1980).

Mózgi-A (org. A-Brains, tłumaczenie własne) oraz Mózgi-B (org. B-Brains, tłumaczenie własne) stanowią opis poziomów refleksji nad umysłem. Minsky zakłada, że w naszym umyśle mamy reprezentacje świata zewnętrznego, czyli właśnie mózgi-A. Za refleksję nad nim, swoistą meta-refleksję, odpowiadają mózgi-B (Singh, 2003, s. 531).

Jak agenci się komunikują?

Jak wspomniano wyżej, Minsky odrzucał koncepcję komunikacji między agentami opartą na języku naturalnym, postulując coś prostszego. Dostrzegał on jednak problemy z tego wynikające. I tak odrzucał on komunikację opartą na symbolach, by nie obarczać systemu potrzebą korzystania z konwencji lub definicji słownikowych (Minsky, 1977, s. 1083), jak również dostrzegał zagrożenia wynikające z nadmiernej prostoty (Minsky, 1986, s. 67).

Zamiast nich Minsky postulował komunikację opartą na liniach K. Sygnał pobudzający innych agentów stanowi proste i skuteczne rozwiązanie. Pojedynczy agent, rozpoznawszy obiekt, może pobudzić czy to Polynema, który wzbudzi kolejnych agentów odpowiedzialnych za właściwą reprezentację, czy też Mikronema, aby pobudzić agentów do myślenia o jakimś ogólnym kontekście. W tej formie komunikacji nie ma potrzeby wyrażania konkretnych idei, zamiast tego informacje są przechowywane w schematach pobudzeń linii K (Singh, 2003, s. 532).

Linie połączeń rozciągają się między agentami. Możemy myśleć o nich jako kolejnych prostych agentach, którzy nabierają znaczenia wraz z rosnącą częstotliwością pobudzeń. Agenci łączą się z odpowiednimi liniami K (Singh, 2003, s. 533).

Minsky zauważa, że wewnętrzny język nie jest potrzebny pojedynczym agentom. Jest on jednak niezbędny agencjom do komunikowania bardziej skomplikowanych i ustrukturyzowanych opisów rzeczy. Jeśli jedna agencja chce przekazać skomplikowaną ideę drugiej, musi odtworzyć tę ideę wyrażoną w swoim własnym systemie reprezentacyjnym poprzez sekwencję przywoływanych ram. Musi zakodować wynik swoich zabiegów i wysłać je do odbierającej agencji. Ta również musi je odkodować. Komunikacja ta wymaga zgody między agentami w znaczeniu kodowanych wiadomości, czyli współdzielenia kodów dla konkretnych reprezentacji (Singh, 2003, s. 533).

Najprawdopodobniej jednak, według P. Singha, najpowszechniejszą formą komunikacji w społeczeństwie umysłu jest, przewrotnie, brak aktywnej komunikacji. Zamiast tego agenci odkrywają, że informacje, których potrzebują, już są dostępne w momencie, kiedy ich poszukują. Jest to możliwe dzięki paranomom. Kiedy jedna z nich aktualizuje swoją reprezentację to inne, powiązane z nią również wykonują podobne operacje. Stąd też nie zachodzi między nimi komunikacja w normalnym rozumieniu tego słowa, lecz raczej kaskada aktualizacji agencji zainteresowanych tymi samymi obiektami lub ich własnościami. Podobnie, jak zauważa P. Singh, komunikacja między

agentami lub agencjami nie musi być precyzyjna, o ile w ogóle może taka być (Singh, 2003, s. 533). Minsky bowiem pisze o inherentnej niejasności, nieprecyzyjności i niepewności w naszej komunikacji. Kontynuuje, stwierdzając, że ciężko nam wyrazić nasze myśli dlatego, że są one często wieloznaczne, niedokładne i zagadkowe. Podobnie zresztą mówi o używanych przez nas słowach- tolerujemy ich niejednoznaczność, bo jesteśmy przyzwyczajeni do niej w naszych myślach (Minsky, 1986, s. 207).

4.1.5. Podsumowanie

Wizja Minsky'ego była trudna do sparafrazowania. Różne poziomy opisu i jego niedookreślenie zdecydowanie nie sprzyjają jej zrozumieniu. Jednakże jej walory stanowią uniwersalność i kompleksowość. Przedstawiona idea jest wewnętrznie spójna i wciąż budzi zaciekawienie badaczy sztucznej inteligencji.

Umysł bowiem jawi się nam jako miejsce, miasto lub państwo pełne wspomnianych biur, czy urzędów, każde wypełnione kotłującymi się agentami. Metafora ta daje nam więc pewną dozę zrozumienia, punkt odniesienia, dzięki któremu jesteśmy w stanie zacząć realnie przyglądać się tej teorii.

A ta jednak nie jest tak oczywista i banalna. Agencje i agenci to znaczące uproszczenia mające umożliwić nam zbliżenie się do ich właściwego opisu. A ten jest całkowicie odhumanizowany i mniej nam się jawi jako targowisko, a bardziej jako system komputerowy czy biologiczny mózg.

Jest to najbardziej widoczne w opisie komunikacji. Nie mamy tu do czynienia z agentem Jankiem, który zauważa na drodze auto, więc wychodzi ze swojego gabinetu, lekko nucąc, przechodzi przez korytarz do gabinetu agenta Krzysia i informuje go o swoim spostrzeżeniu, co skutkuje serią kolejnych agentów biegających od gabinetu do gabinetu i wzajemnie przekazujących sobie tę wiadomość, aż cała Agencja Rozpoznawania Aut zostaje aktywowana.

Zamiast tego mamy tutaj pojedynczego agenta w formie linii K, który w reakcji na bodziec, będący w istocie nemem, w formie samochodu aktywuje się, w ten sposób propagując sygnał do pronomy odpowiedzialnej za rozpoznawanie kształtu auta, która może rozesłać ten sygnał dalej do kolejnych i kolejnych paranom, na przykład aktualizując wiedzę, że widziano niebieską Toyotę, jak w tych ramach, które opisywały zeszły wtorek.

Struktura ta nie jest bynajmniej intuicyjna i taka też nie powinna być. Daje ona nam poczucie poznawania czegoś obcego i uporządkowanego w nieznanym nam sposób.

4.2. Symbioza ludzi i maszyn

4.2.1. Wstęp

Drugim z omawianych stanowisk jest wizja symbiozy ludzi i maszyn. Wiąże się ona ściśle z koncepcjami wzmocnienia poznawczego czy wzmocnienia inteligencji. Eva Žáčková, w swoim artykule z 2011 roku, opisuje ją na podstawie wybranych prac trzech głównych autorów: Josepha Carla Robnetta Licklida, Williama Rossa Ashby'ego oraz Douglasa Carla Engelbarta (Žáčková, 2011, s. 133).

Symbioza odnosi się tu do relacji między ludźmi a maszynami, na powszechnie znanej zasadzie z biologii opisującej co najmniej dwa organizmy, które nie dość, że współistnieją, to jeszcze czerpią z tego korzyści. Maszyny odnoszą się tutaj do nieorganicznego mechanizmu. Jednakże, jak kontynuuje E. Žáčková, naszą relację z nimi zaczęliśmy postrzegać na zasadzie symbiozy dopiero na etapie komputerów. Wcześniejsze wynalazki, takie jak silnik spalinowy czy żarówka, nie pociągały takich skojarzeń (Žáčková, 2011, s. 134).

Autorka postuluje, iż wynika to z naszego obsesyjnego wręcz połączenia z technologią. Tworzymy coraz intymniejsze interfejsy, niektórzy już pragną połączenia na poziomie neuronalnym. Za jeden z najszerzej znanych przykładów tej ambicji jest Neuralink Corporation. Jest to amerykańska firma, powołana w 2016 roku, między innymi przez Elona Muska, której celem jest stworzenie wszczepialnego chipu umożliwiającego stworzenia połączenia mózg-maszyna (Wikipedia - Neuralink).

Celem tak rozumianych symbiotyków jest stanie się, tak zwanym, cyborgiem. E. Žáčková przyjmuje tutaj definicję takiego bytu jako „samo-sterowalnego systemu ludzko-maszynowego”. Jednakże tak sformułowana definicja może być jednak zbyt wąska. Jak stwierdza Łukasz Blechar, nie istnieje uniwersalna definicja cyborga (Blechar, 2018). Przywołuje on pierwsze użycie tego słowa w artykule o eksploracji kosmosu z 1960 roku (Clynes i Kline, 1960), w którym to użyto go jako skrótu dla opisanego organizmu cybernetycznego. Podobnie jak w rozumieniu zaproponowanym przez E. Žáčkovą, mówimy tu o połączeniu części biologicznych i mechanicznych, bez określenia ich dokładnej relacji lub proporcji. W znaczeniu wąskim więc cyborgiem możemy określić osobę, która posiada mechaniczne części ciała. Mimo że nasuwają się tu przykłady rodem z filmów science fiction, równie dobrze możemy tutaj wskazać jako przykład osobę z rozrusznikiem serca. Alternatywnie, w znaczeniu szerokim, wszelkie interakcje z technologią klasyfikują nas jako cyborgi. Przykładem takiego rozumienia tego określenia

jest stanowisko Amber Case, która wprost pisze, że jesteśmy już cyborgami, gdyż nasza codzienność jest oparta na technologii (Case, 2014). Zdarzają się również przypadki skrajne, gdzie pada wprost, że to bycie cyborgiem sprawia, że jesteśmy ludźmi (Smedes, 2005).

Motywacją dla takiego połączenia jest uzupełnianie się. Dzięki komputerom już jesteśmy w stanie wykonywać wiele operacji łatwiej, szybciej i precyzyjniej. Stąd też określenie wzmocnienia, tak kluczowe w opisie tego stanowiska. Co więcej, współczesne komputery nie oznaczają pojedynczego urządzenia w konkretnym pomieszczeniu, lecz raczej system połączonych online aparatów. Łączą nas one w czasie rzeczywistym, czy raczej- łączą nas z systemami, w które wpleceni są inni ludzie i w ten sposób zaspokajają one nasze potrzeby socjalne i poznawcze.

Wzmacnianie ludzi jest szeroko rozpowszechnionym określeniem w grupach transhumanistów, technocentrystów i ich oponentów- biokonserwatystów czy neoluddystów. Samo słowo wzmocnienie w języku polskim wywodzi się z formy prasłowiańskiej, która jest przymiotnikiem odrzeczownikowym określenia bliskoznacznego dla siły, potęgi, mocy (Wielki słownik języka polskiego). W angielskim źródłosłowie „enhance” Žáčková odwołuje się do łacińskiego słowa „inaltiare” (z prefiksem in, altus – znaczącym wysoki). Oryginalnie oznaczało ono po prostu czynić coś wyższym. Jednak wraz z czasem znaczenie ewoluowało w bardziej metaforyczną stronę- czynienia ogólnie pojętych ulepszeń podnoszących wartość lub jakość (Žáčková, 2011, s. 134).

Nick Bostrom definiuje z kolei wzmacnianie jako ingerencję w funkcjonowanie podsystemu organizmu lub tworzenie całkowicie nowej funkcjonalności, lub całego podsystemu (Bostrom, 2008). Wcześniej wspomniano, że wzmocnieniu ulegają procesy poznawcze. Eva Žáčková pojmuje je szeroko i traktuje ulepszenie jako cokolwiek, co polepszy nasze kompetencje w zakresie uczenia się, pamięci, zdobywania wiedzy, myślenia, czy procesowania informacji jest określane przez nią wzmocnieniem poznawczym.

Warto odnotować za E. Žáčkovą, że wzmacnianie, w tym kontekście, jest postrzegane w zasadzie wyłącznie pozytywnie, a głównym opisywanym przez nią celem jest znaczące zwiększenie poziomu ludzkiej inteligencji na poziomie ogólnym, czyli ponad jednostkowym.

Inteligencja właśnie jest tutaj drugim członem, przywołanego wyżej, określenia „wzmacnianie inteligencji”. Słowo to wywodzi się z łacińskiego czasownika

intellegere oznaczającego dosłownie wybrać spośród możliwości. Jest ono ściśle zbieżne z psychologiczną koncepcją mierzalnej, kwantyfikowalnej inteligencji. Jest ona równie harmonicznie zestrojona z potrzebami badaczy z lat 60. XX wieku, którzy czynili pierwsze próby określenia, czym ma być sztuczna inteligencja. Jak zauważa E. Žáčková, tak pojmowana inteligencja doskonale nadaje się na przełożenie na matematykę i optymalizację z wykorzystaniem maszyn, które przeprowadzają podobne obliczenia szybciej i sprawniej od ludzi (Žáčková, 2011, s. 135).

Co jest w tym jednak problematyczne to, że problemy, którymi zajmujemy jako ludzie, rzadko mogą być rozwiązane poprzez odpowiednie zastosowanie algorytmu. Co więcej, E. Žáčková przytacza Josepha Weizenbauma, który stwierdza, że większość naszych problemów nie ma w ogóle rozwiązania (Weizenbaum, 2002).

Autorka przedstawia tutaj tezę, że nasza inteligencja jest uzależniona od określonej ramy odniesienia opartej na naszych życiowych doświadczeniach, których nie jesteśmy w stanie skwantyfikować. Jak kontynuuje, opracowaliśmy już maszyny, które doskonale radzą sobie z problemami opartymi na prostych obliczeniach. Dzięki temu konsekwentnie oddalamy się od prostych, rutynowych, mechanicznych obowiązków, które jesteśmy w stanie sukcesywnie automatyzować za pomocą technologii. W tym sensie E. Žáčková zgadza się, że efektywnie zwiększyliśmy nasze możliwości poznawcze. Za przykłady podaje pisanie pracy dyplomowej z wykorzystaniem elektronicznych źródeł. Wzmocnienie pamięci za pomocą technologii oznacza brak konieczności pamiętania faktów, które możemy łatwo sprawdzić. Natomiast operacje oparte na liczeniu, przechowywaniu (danych), analizowaniu, symulowaniu czy przeszukiwaniu stają się praktycznie niemożliwe bez komputera.

Sukcesywnie rozwiązujemy również mniej oczywiste wyzwania oparte na złożonych systemach ze zdolnością do adaptacji. Ludzie jednak operują na dziesiątkach systemów każdego dnia, których możliwość rozwiązania za pomocą technologii jest wątpliwa. Przytaczane są tutaj takie przykłady jak: odgadywanie nastroju partnera, prowadzenie auta podczas godzin szczytu czy dyskutowanie z dzieckiem wyjątków w ustalonych zasadach współżycia typu zjedzenie dodatkowej porcji lodów (Žáčková, 2011, s. 136).

Autorka waha się przy tym, czy inteligencja kompletnie sztuczna będzie w stanie kiedykolwiek zastąpić ludzką. Powołuje się tutaj na tezę podobną do jednego z argumentów Huberta Dreyfusa, wedle którego prawdziwą ludzką inteligencją stanowi mądrość, która nie potrzebuje wzmocnień lub przyspieszeń, lecz raczej doświadczeń. Odwołuje się

tu do przeżyć związanych z popełnianiem błędów, odczuwaniem skutków naszych decyzji, cierpieniem czy bezsilnością. Podobny wpływ na nas ma wchodzenie w interakcje społeczne z innymi bytami, które miały podobne, lecz prawdopodobnie zupełnie inne, przeżycia i odczucia. Zupełny brak ekwiwalentów tychże w przetwarzaniu danych przez maszyny budzi wątpliwość E. Žáčkovéj i, za Josephem Weizenbaumem, podkreśla ona, że nie powinniśmy dyskutować, jak maszyny mogą rozwiązać typowo ludzkie problemy, lecz raczej czy maszyny w ogóle powinny je rozwiązywać (Weizenbaum, 2002).

4.2.2. Koncepcje symbiozy ludzi i maszyn w latach 60. XX wieku

E. Žáčkova przedstawia koncepcje związane ze wzmacnianiem inteligencji, co rozumie jako połączenie ludzi i komputerów, wsparcie oraz wzmocnienie funkcji poznawczych ludzkiego organizmu przy, co istotne, zachowaniu doświadczenia bycia człowiekiem. Jako kontrast podaje programy silnej sztucznej inteligencji lub osobliwości, opisanej powyżej jako ogólna sztuczna inteligencja.

Pierwsze prognozy połączenia ludzi i maszyn Žáčkova dostrzega już w latach 20. XX wieku (Haldane, 1924; Bernal, 1929), jednakże myśl ta nabrała kształtu i mocy perswazyjnej, dopiero gdy pojawiła się druga generacja komputerów, w której tranzystory zastąpiły tuby próżniowe.

Trudno jest wskazać pojedyncze źródło tej idei w tych czasach, ale nazwiska trzech badaczy są przywoływane w tym kontekście częściej niż inne. William Ross Ashby, pionier cybernetyki. Joseph Carl Robnett Licklider, osoba zaangażowana w liczne projekty, w tym pracownik ARPA, który pierwszy sformułował wizję połączenia komputerów, co ostatecznie doprowadziło do powstania internetu. Douglas Carl Engelbart, który był zaangażowany w usprawnianie interakcji między ludźmi a maszynami, a przy tym zasłynął jako wynalazca komputerowej myszki (Žáčková, 2011, s. 137).

4.2.3. William Ross Ashby: Wzmacnianie Inteligencji

W swojej książce z 1957 roku *An Introduction to Cybernetics* William Ross Ashby zawarł rozdział *Regulowanie Wzmocnienia* (org. Amplifying Regulation, tłumaczenie własne). Przedstawił w nim tezę, że ludzie są kolejno kierowani szeregiem regulatorów: najpierw są geny, które regulują funkcjonowanie organizmu doprowadzając do powstania kolejnego etapu, kolejnego regulatora, mózgu, który jest odpowiedzialny za szeregi procesów psychologicznych, takich jak uczenie się, myślenie, rozpoznawanie.

Dzięki tym regulatorom mamy podwyższone możliwości adaptowania się do środowiska, co W. R. Ashby określił mianem inteligencji. Była to definicja zbieżna z wcześniej opisanymi koncepcjami mierzalnej inteligencji. Warto odnotować, że w czasie publikacji tejże książki była to definicja silnie popularyzowana w Stanach Zjednoczonych przez behawiorystów. Jak kontynuuje W. R. Ashby, moc intelektualna to tak naprawdę sztuka dokonywania właściwych wyborów. Tak rozumiana może być według niego wzmocniona (Ashby R., 1957, s. 272). Taki wzmacniacz pozwoliłby ludziom na przewyższenie swoich intelektualnych ograniczeń, jak również stanowiłby kolejny regulator działający na całe społeczeństwo. W. R. Ashby uważał stworzenie go nie tylko za możliwe, ale również za pożądane. Według niego umożliwiłby on ludziom lepiej zrozumieć i panować nad swoim życiem mentalnym i rzeczywistością interakcji społecznych.

W. R. Ashby określił następującą kolejność: natura określa geny, geny określają mózg, mózg określa dostosowanie do środowiska. Stanowi to proces naturalny, niewytworzony przez człowieka, zastany. Według W. R. Ashby'ego w drugiej połowie XX wieku możliwa była bezprecedensowa okazja na wyjście poza ten proces, na syntetyczne, świadome rozszerzenie go.

We wzmiankowanej publikacji nie opisał jak dokładnie osiągnąć ten cel, jednakże zasugerował dwie możliwe drogi, jak może to nastąpić. Pierwszą było stworzenie niezależnej sztucznej inteligencji, drugą sztuczne wzmocnienie inteligencji ludzkiej za pomocą symbiozy ludzi i komputerów. Žáčková zauważa, że mimo że oba scenariusze są przywoływane, a konkluzja książki nie jest zdecydowana za jednym z nich, to stanowisko W. R. Ashby'ego częściej pojawia się jednak w kontekście tejże symbiozy. Jak kontynuuje, w jej ocenie, bardziej prawdopodobne jest, że W. R. Ashby preferował opcję związaną ze sztuczną inteligencją. Tak rozumiana sztuczna inteligencja jest zbieżna z koncepcją, przytaczanej wcześniej, technologicznej osobliwości (Žáčková, 2011, s. 138).

4.2.4. Joseph Carl Robnett Licklider: Symbioza ludzi i maszyn

Joseph Carl Licklider opublikował w 1960 swój słynny artykuł *Man-computer Symbiosis*. Wierzył on, że w niedalekiej przyszłości nasze mózgi i maszyny obliczeniowe będą połączone ze sobą bardzo ściśle i tak powstałe partnerstwo będzie myślało „jak żaden mózg przed nią i przetwarzało dane jak żadna współczesna maszyna” (org. „the resulting partnership will think as no human brain has ever thought and process data in a way not approached by the information-handling machines we know today” (Licklider, 1960, s. 4).

J. C. R. Licklider określa tę relację jako specyficzny ludzko-mechaniczny system. Wcześniej takie połączenia ludzi z maszynami rozumiano jako mechanicznie wzbogaconego człowieka, jak choćby w pracy Johna Dudleya Northa (North, 1954), koncepcji, do której J. C. R. Licklider odwołuje się bezpośrednio. Człowiek w tymże podejściu zawsze określa co i jak należy zrobić. Mechaniczne części tego systemu stanowią jedynie przedłużenia człowieka, jego oczu czy rąk. J. C. R. Licklider nie widzi w tym symbiozy, do której sam dąży. Określa on taki system mianem półautomatycznego (Licklider, 1960, s. 4).

Zamiast tego poszukiwał on systemu, w ramach którego zwiększana będzie efektywność tak ludzi, jak i maszyn. W jego rozumieniu ludzie wciąż będą wyznaczać cele, formułować hipotezy, określać kryteria i ewaluować wyniki. Maszyny przejmą natomiast zadania, które można zrutynizować, które trzeba wykonać, by przygotować się na możliwość poszukiwania odkryć, wglądu i dokonywania decyzji (Licklider, 1960, s. 1).

Co istotne, komputery wtedy były dużo mniej zaawansowane niż współcześnie. Jak trafnie określa E. Žáčková, w tamtych czasach nie było nawet rozróżnienia między *hardwarem* i *softwarem* (Žáčková, 2011, s. 140). Niemniej, J. C. R. Licklider już wtedy wierzył, że kwestią czasu jest, zanim zaczniemy traktować komputery jak znajomych z pracy, których kompetencje wspierają i uzupełniają nasze (Licklider, 1960, s. 4). Jediną różnicą, jak pisze później, jest to, że komputerom nie będzie przeszkadzał status względnie drugorzędny w tej relacji (Licklider, 1960, s. 7).

J.C.R. Licklider największą przeszkodę dla realizacji tej wizji upatrywał w komunikacji. Język maszynowy jest oparty przede wszystkim na syntaktyce, podczas gdy w naszym wypadku bardzo istotna jest również warstwa semantyczna wypowiedzi. Stąd też człowiekowi wystarczy zadeklarować cel, na przykład „wybrać się do kina” i ten zrozumie natychmiast cały ciężar decyzyjny zaangażowany w cały proces jego osiągnięcia. Komputer natomiast potrzebuje otrzymać zestaw precyzyjnych instrukcji i bezpośredniego wskazania kluczowych momentów by mógł skutecznie ewaluować wymaganą pracę. J. C. R. Licklider pisał o potrzebie stworzenia interfejsu komunikacyjnego.

Jak podsumowuje E. Žáčková, komputer musi rozumieć informacje wsadowe, a użytkownik musi być w stanie pojąć wyniki przedstawione przez komputer. Stwierdza ona, że w 2011 roku cel ten został osiągnięty dzięki rozwojowi języków programowania oraz opracowania specjalnych interfejsów umożliwiających pracę z kodem komputerowym (Žáčková, 2011, s. 140).

O ile w kontekście wprowadzania informacji wsadowych dla maszyn tak jest w istocie, tak problem rozumienia wyników jest daleki od rozwiązania. Dotyczy on zrozumienia nie tyle wyników jako takich, ile sposobu, w jaki sztuczne sieci neuronowe je otrzymały. W celu uzupełnienia tej luki powstała specjalna dziedzina sztucznej inteligencji nazywana wytłumaczalnym uczeniem maszynowym (org. explainable AI, tłumaczenie własne).

4.2.5. Douglas Carl Engelbart: System H-LAM/T

Jak stwierdza E. Žáčková, Wizja Rossa Ashby'ego zaprezentowana w 1957 była dość mglista (Žáčková, 2011, s. 140). Kilka lat później Douglas Carl Engelbart opublikował raport podsumowujący stan prac nad wzmocnieniem ludzkiego intelektu: *Augmenting Human Intellect: A Conceptional Framework* (Engelbart, 1962).

Przedstawił w nim solidną ramę dla zastanego stanu badań i opracował sugestie rozwinięć. Podobnie jak W.R. Ashby, D.C. Engelbart postrzega wzmocnienie ludzkiego intelektu jako naturalne rozwinięcie procesu ewolucyjnego zapoczątkowanego wraz z narodzinami kultury (Engelbart, 1962, s. 19).

Wyróżnia on cztery główne kategorie wzmocnień:

1. Artefakty;
2. Język;
3. Metodologia;
4. Trening dla ludzi, aby lepiej korzystali z wcześniejszy wymienionych wzmocnień.

Stanowią one wysoce zorganizowane i synergicznie funkcjonujące struktury. Z jego, ściśle inżynierskiego, punktu widzenia, wiadome jest, że jeśli wzmocnimy którąkolwiek z tych kategorii, wzmocnimy całą strukturę.

D.C. Engelbart zaproponował tak zwany system H-LAM/T, który obejmuje wszystkie powyższe kategorie, a jego nazwa pochodzi od pierwszych ich liter w oryginale, tj.: Człowiek używający Języka, Artefaktów, Metodologii, w których jest Wytrenowany (org. „Human using Language, Artefacts, Methodology, in which he is Trained”, tłumaczenie własne; Engelbart, 1962, s. 11).

Struktura ta pomaga człowiekowi przetrwać czy, jak określa to Žáčková, odnosić ewolucyjne sukcesy (Žáčková, 2011, s. 141). Najbardziej fundamentalny jest tutaj rozwój języka, ponieważ umożliwia on zwiększenie zrozumienia otaczającego nas świata i zaprezentowanie go symbolicznie. Komputery w tym ujęciu funkcjonują jako artefakty, które pozwalają nam na manipulowanie symbolami na wcześniej niespotykaną skalę.

Žáčková podsumowuje opis stanowiska D. C. Engelbarta jego cytatem stwierdzającym, że natura naszych przyszłych, usprawnionych systemów byłaby ściśle zależna od spodziewanych zmian w naszej technologii lub w naszym zrozumieniu bycia człowiekiem (Žáčková, 2011, s. 141).

4.2.6. Podsumowanie

Te trzy, różne spojrzenia na wzmacnianie inteligencji możemy podsumować relatywnie prosto. William Ross Ashby pisze o komputerach jako kolejnych regulatorach naszych zachowań. Pierwszych, które sami sobie wytworzymy. Joseph Carl Robnett Licklider dostrzega w komputerach potencjalnych partnerów, którzy już w jego czasach byli w stanie wymiennie wspierać ludzi, ale czynnikiem, który to uniemożliwił, była niewygodna komunikacja. Jak wspomniano w poświęconym mu rozdziale, problem ten nie został całkowicie rozwiązany. Podejście Douglasa Carla Engelbarta zostało najsilniej skompresowane. Jego praca to bowiem 144-stronicowy raport o możliwościach technicznych lat 60. XX wieku, jednakże przytaczanie jej całości nie pomogłoby w zrozumieniu intencjonalnie prostego modelu. Wymienione przez niego cztery możliwości wzmocnienia są tak samo aktualne dzisiaj.

Spójrzmy bowiem na obecny stan technologii. Czy nie jesteśmy silniej powiązani z maszynami? Prawdopodobnie tak, przez wzgląd na dynamiczny rozwój smartphone'ów i technologii ubieralnych mamy bezprecedensowe połączenie ze sobą. Technologie, które rozwijają się najdynamiczniej to te związane z ulepszaniem naszej komunikacji, zbliżaniem nas do siebie. Praktycznie gdziekolwiek teraz jesteśmy, możemy wyjąć telefon z kieszeni i zadzwonić do naszych bliskich. Wolny czas możemy spędzać na kłótniach w Internecie z kimś z drugiej strony globu, a niedawno królował slogan francuskiego ekologa René Julesa Dubosa „myśl globalnie, działaj lokalnie”.

Funkcjonujemy więc w relatywnie innej rzeczywistości niż przytoczeni badacze, ale ich koncepcje się bynajmniej nie zdezaktualizowały. Więcej, stały się wręcz wykonalne. W końcu bowiem mamy techniczne możliwości zrealizowania tychże wizji, jak wspomniano w rozdziale 1.2.1., ich popularność wyłącznie rośnie. Dokładniejszy opis współczesnych rozwiązań zostaje przedstawiony w dalszej części rozprawy, ale kluczowy dla zrozumienia doniosłości tego fenomenu jest fakt, że postulaty W. R. Ashby'ego, J.C.R. Licklidera i D. C. Engelbarta się nie zdezaktualizowały.

Koncepcje W.R. Ashby'ego silnie rezonują ze współczesnymi osiągnięciami w adopcji sztucznej inteligencji w psychiatrii. Bynajmniej nie chodzi tu o zastąpienie

psychiatrów programami jak w dawnej Elizie, lecz o wymierne wsparcie ich praktyki osiągnięciami nowoczesnej technologii. Zastosowanie tutaj znajdują zwłaszcza dwa artykuły: *Technologii pomagającej nam pomóc samym sobie* Rana Zilca (2019) oraz *Manifeście Wzmocnionego Zdrowia Psychicznego* George'a Eleftheriou (2018).

Oba odnoszą się do koncepcji telemedycyny i zastosowania technologii ubieralnych do powiązania introspektywnych danych psychicznych z obiektywnie mierzonymi danymi biomedycznymi na bezprecedensową skalę. Alternatywnie, sama jakość terapii może być analizowana przez sztuczną inteligencję, jak dowodzą Zac Imel i jego współbadacze (Imel i inni, 2019).

Przeprowadzili oni badanie na 21 psychiatrach zajmujących się terapią uzależnień i zaprosili ich do korzystania z aplikacji, która nagrywała prowadzone przez nich sesje, opracowywała transkrypt i analizowała przebieg pod kątem użyteczności, zadowolenia terapeuty, percypowanej skuteczności oraz spodziewanych intencji pacjenta do zastosowania się do zasugerowanych metod leczenia. Następnie program generował automatycznie zestaw sugestii zwrotnych dla psychiatrów. Wszyscy uczestnicy badania ocenili system za łatwy do użycia i byli zadowoleni z uzyskanych sugestii. 83% określiło sugestie za zbieżne z własnymi obserwacjami, a 90% chciało włączyć program do swojej codziennej praktyki (Imel i inni, 2019).

Z kolei koncepcje J. C. R. Licklidera i D. C. Engelbarta znajdują silne odbicie w źródłach przytoczonych w kolejnej części i tam też zostaną ocenione w jakim stopniu pozostały aktualne.

4.3. Porównanie koncepcji

Po przedstawieniu obu koncepcji powstaje możliwość ich porównania. Z jednej strony mamy więc społeczność umysłów Marvinina Minsky'ego, z drugiej symbiozę ludzi i maszyn Williama Rossa Ashby'ego, Josepha Carla Robnetta Licklidera i Douglasa Carla Engelbarta.

Bardzo wiele różni te idee. Począwszy od czasu powstania, o ile te dotyczące symbiozy powstały w podobnym czasie W. R. Ashby – 1957, J. C. R. Licklider – 1960, D. C. Engelbart – 1962, tak książka M. Minsky'ego pochodzi aż 1986 roku. Jak wspomniano w części wprowadzającej, lata 60. były czasami, gdy komputery wciąż przyjmowały polecenia na taśmach papieru, a wyświetlacze były oparte na lampach katodowych. Gdy ukazała się publikacja M. Minsky'ego komputery były już, przynajmniej w Stanach Zjednoczonych, zwykłym elementem wyposażenia domu.

Kolejną różnicę stanowi obiekt zainteresowania autorów. M. Minsky skupiał się na pojedynczych umysłach, jak działają konkretne jednostki. U „symbiotyków” dominowało podejście zorientowane na grupę. Rzadko pisali oni jak konkretny człowiek będzie funkcjonować, a raczej jak będzie wyglądała współpraca w obrębie grupy składającej się z ludzi i maszyn.

Inną, istotną różnicę stanowi poziom skomplikowania umysłów biorących udział w rozwiązywaniu zadań. Dla M. Minsky'ego kluczowe jest, by maszyny realizowały ten sam program umysłu, który dostrzega u nas. Dla J. C. R. Licklidera było to zupełnie nieistotne. Maszyny miały stanowić partnerów, którym możemy oddelegować zadania. Było to ujęcie dużo bardziej pragmatyczne i nastawione na realizowanie celów, a możliwości konkretnych aktorów były efektywnie wobec tego drugorzędne.

W czym jednak się zgadzają to podniesienie komputerów z roli wyłącznie narzędzia. U M. Minsky'ego umysł mógł być realizowany. Nie starał się on dokładnie oddać, jak ludzie myślą, ale raczej jak przebiega myślenie jako takie. Stąd też, równie dobrze jesteśmy w stanie zaprojektować maszynę, która będzie takie oprogramowanie wykonywać. Maszyny więc nie jawią nam się już jako wyłącznie narzędzia, a raczej jako potencjalni partnerzy, osobniki, które konstruujemy, aby funkcjonowały na naszym poziomie. Z drugiej strony W.R. Ashby i pozostali również tak do nich podchodzą, zwłaszcza J. C. R. Licklider pisał *explicite* o zespołach złożonych z różnych typów agentów.

Warto również odnieść się do tego, co nie wynika bezpośrednio z tekstu, a z implikacji, jakie z nich płyną. Konkretnie do percypowanych różnic między ludźmi

a maszynami. M. Minsky pisał o hipotetycznych bytach, celowo unikał uściślenia czy mówi o człowieku, czy jakiejś futurologicznej, całkowicie syntetycznej istocie. Stąd też możemy go określić mianem funkcjonalisty i zwolennika teorii obliczeniowej umysłu. Licklider i inni natomiast wyraźnie dostrzegał różnice między ludźmi a maszynami, wręcz budował na nich swoją wizję. Obie grupy miały wyraźnie rozdzielone kompetencje i zadania. Jedni kalkulowali i przechowywali w pamięci dane o transakcjach, drudzy podejmowali decyzje i wytyczali strategię.

Co jednak jakbyśmy wyobrazili sobie społeczeństwo istot zbudowanych tak, jak opisał to M. Minsky? W jaki sposób mogłoby funkcjonować? Jeśli przecież nie byłoby funkcjonalnych różnic między oboma rodzajami agentów, to koncepcja J. C. R. Licklidera nie ma już racji bytu. Nie byłoby relacji symbiotycznej, gdyż ta zakłada różne potrzeby organizmów wchodzących w jej skład.

5. Część III – współczesne amalgamaty umysłów

5.1. Superumysły

Thomas W. Malone od dawna zajmuje się badaniem teorii funkcjonowania i optymalizacji działania organizacji. Pełni funkcję profesora nauk o zarządzaniu na MIT Sloan School of Management. Opublikował ponad 100 artykułów, prac badawczych i rozdziałów książek; jest wynalazcą i posiada 11 patentów, jak również współredagował cztery książki. Jest również współzałożycielem czterech firm wytwarzających oprogramowanie. Dodatkowo współpracuje z różnymi organizacjami w charakterze konsultanta. W 2006 roku założył MIT Center for Collective Intelligence. Wcześniej na tej uczelni powołał MIT Center for Coordination Science. Był również współzałożycielem MIT Initiative on “Inventing the Organizations of the 21st Century” (Massachusetts Institute of Technology, 2018).

W 2018 roku, w Wielkiej Brytanii, ukazała się, nakładem wydawnictwa Oneworld Publications, jego książka *Superminds*. Publikacja ta ma charakter deklaracyjny i normatywny, gdyż nie tylko opisuje stan zastany, ale również przedstawia postulaty rozwiązań najważniejszych, współczesnych problemów. Tematyka jest zdecydowanie szeroka i nie zamyka się na pojedynczą dyscyplinę naukową. Jest to zgodne z intencją autora, by ukazać wszechstronność i uniwersalność zastosowań sztucznej inteligencji albo raczej, jak sam ją nazywa, inteligencji grupowej.

Warto w tym miejscu podkreślić, że nie są to pierwsze próby przewidzenia przyszłości przez autora na podstawie jego badań. W 1987 opublikował artykuł *Electronic markets and electronic hierarchies*, w ramach którego przewidział rozwój sprzedaży elektronicznej. Dziś możemy w tym dostrzec rolę internetu i cały, olbrzymi dział handlu elektronicznego (ecommerce). W 2004 została wydana książka, której był współautorem: *Inventing the Organizations of the 21st Century*. Stanowi ona podsumowanie szeroko zakrojonej pięcioletniej inicjatywy badawczej funkcjonującej pod tym samym tytułem. Mimo że przewiduje tam olbrzymi wzrost znaczenia małych firm, które dzięki internetowi będą w stanie dużo łatwiej i szybciej organizować współpracę między sobą, co nie było zupełnie trafioną predykcją, to wspomina również o zmniejszaniu czasu pracy i wzroście popularności pracy zdalnej lub hybrydowej (wykonywanej częściowo z biura, częściowo z domu), które stanowią jedne z najczęściej obecnie dyskutowanych tematów dotyczących organizacji pracy.

Jego przewidywania możemy uznać za raczej trafne, gdyż o ile redukcja wymiaru czasu stanowi kwestię budzącą wątpliwości i obawy, tak w istocie sukcesywnie rośnie liczba firm, które decydują się na eksperyment i wprowadzenie pracy czy to w 4 dni w tygodniu lub zmniejszają liczbę godzin każdego dnia, z zachowaniem wcześniejszego wynagrodzenia pracowników. Warto odnotować, że wyniki tychże prób są raczej pozytywne. Pracownicy są bardziej zmotywowani, a wzrasta zarówno mierzalna ilość wykonywanej pracy, jak i jej jakość (Wikipedia, 2023).

Praca zdalna budziła podobne obawy, lecz jej wprowadzenie zostało niejako wymuszone przez pandemię COVID-19. Według badania Kim Parker, Juliany Menasce Horowitz i Rachel Minkin na 5,585 dorosłych obywatelach Stanów Zjednoczonych w trakcie epidemii mnóstwo pracowników przeszło na pracę, chociaż w części zdalną, mówimy tu o przejściu od 23% do blisko 60% osób aktywnych zawodowo. Byli oni raczej zadowoleni, gdyż większość osób badanych, konkretnie 64%, w pracy zdalnej odkryła lepszy balans między pracą a życiem po niej (tzw. work-life balance). Podobnie, 44% z nich uważa, że pracując zdalnie łatwiej o skupienie i osiągnięcie założonych celów biznesowych, podczas gdy jedynie 10% twierdziło odwrotnie (Parker, Menasce Horowitz i Minkin, 2022).

5.1.1. Inteligencja

Inteligencja wyspecjalizowana i ogólna

Malone zaczął myśleć nad superumysłami w trakcie prowadzonych przez siebie badań nad inteligencją. Dzieli on ją na dwa rodzaje: wyspecjalizowaną i ogólną.

Ta pierwsza jest określana poprzez odniesienie do dwóch popularnych definicji tej charakterystyki mentalnej. Malone parafrazuje pracę Shane'a Legga i Marcusa Huttera, cytujących Encyklopedię Britannica, określając inteligencję jako zdolność do efektywnego dostosowywania się do środowiska (Legg i Hutter, 2007). Z drugiej strony sięga również Howarda Gardniera i jego określenia inteligencji jako zdolności do rozwiązywania problemów lub tworzenia produktów, które są cenione w obrębie przynajmniej jednego otoczenia kulturowego (Gardner, 1983).

Autor określa wyspecjalizowaną inteligencję jako zdolność do efektywnego osiągania konkretnych celów w określonym środowisku. Czyli, jak rozwija w towarzyszącym akapicie, byt inteligentny zrobi to, co w największym stopniu pomoże mu osiągnąć założone cele, na podstawie całej wiedzy, do jakiej ma dostęp.

A jeszcze bardziej uproszczając, wyspecjalizowana inteligencja to, jak określa Malone, po prostu miara efektywności w osiągnięciu celów (Malone, 2018, s. 24).

Ogólna inteligencja z kolei jest przedstawiona jako zdolność do efektywnego osiągnięcia szerokiego spektrum celów w różnych środowiskach. W przeciwieństwie do poprzedniego typu inteligencji, w tym przypadku mówimy tak o licznych celach, jak i licznych środowiskach. Oznacza to więc, że inteligentny agent musi być dobry nie tylko w jednym konkretnym zadaniu, ale też rozwijać swoje kompetencje i uczyć się jak wykonywać inne zadania, by osiągać więcej celów. Jak kontynuuje Malone, pomaga to również określić różnice między dostępnymi obecnie technologiami a ogólną sztuczną inteligencją, czyli taką, która przekroczyła stan technologicznej osobliwości opisany we wstępie (Malone, 2018, s. 25).

Istotność inteligencji

Inteligencja jest w ujęciu Malone'a kluczowa, gdyż dzięki określeniu poziomu czyjejs inteligencji poprzez standaryzowany test jesteśmy w stanie z dużym prawdopodobieństwem stwierdzić, jak będzie sobie radził w szeregu innych zadań, eliminując w ten sposób potrzebę żmudnych obserwacji, które musiałyby trwać latami. Odwołuje się tutaj do Charlesa Spearmana (Spearman, 1904), który pisał, że ludzie, którzy są dobrzy w jednej aktywności (na przykład matematyce), są też na ogół lepsi w innych dziedzinach (przykładowo w czytaniu).

Malone przytacza tutaj za Ch. Spearmanem, że dzięki przeprowadzeniu analizy czynnikowej odkryto, że szereg naszych kompetencji poznawczych i jakość naszych rozwiązań są zależne od pojedynczego aspektu, który możemy uznać za bliski przytoczonych wcześniej definicji inteligencji ogólnej (Malone, 2018, s. 29).

Z drugiej strony Malone zauważa, że testy te są użyteczne, ale nie stanowią kryształowej kuli, która zawsze daje słuszne odpowiedzi. Jako kontrargumenty przytacza teorię typów inteligencji Gardnera, wykazując, że istnieją inne pojęcia inteligencji i sam obliczeniowy charakter może być niewystarczający oraz krytykę standaryzowanych testów stosowanych w szkolnictwie, które nie mają udowodnionej, pragmatycznej funkcji (Malone, 2018, s. 29).

Badanie inteligencji grupowej

Malone rozszerza przytoczone definicje inteligencji w nietypowy sposób. Klasycznie bowiem przyjmuje się, że mówimy w tym kontekście o pojedynczym bycie. Autor postuluje jednak rozszerzenie tychże określeń na grupy. Jak sam wspomina,

nie spotkał się z takim podejściem wcześniej i wykonywane w obrębie jego zespołu badawczego prace były bezprecedensowe (Wooley, Chabris, Pentland, Hashmi i Malone, 2010).

Wspomniane badania odbywały się w 2010 roku. W ramach prac zespołu przebadano wtedy 699 osób podzielonych na 192 grupy składające się z dwóch do pięciu osób. Aby zapewnić reprezentatywność próby badawczej, badani byli rekrutowani licznymi kanałami, w tym poprzez portale z ogłoszeniami o drobnych zleceniach, jak Craigslist. Dodatkowo krótki test inteligencji pomógł utrzymać typową dla Stanów Zjednoczonych dystrybucję inteligencji. Ponadto, w przeciwieństwie do większości innych badań na zachowaniu ludzi w grupach, nie wyznaczano liderów ani nie wymagano jakichś specjalnych umiejętności. Warunkiem koniecznym było jednak wypracowanie współpracy w obrębie grupy (Wooley, Chabris, Pentland, Hashmi i Malone, 2010).

Przebieg badania

Wszystkie grupy były zaangażowane w zadania polegające na:

- wspólnym wykonywaniu zadań (zapisanie dłuższego tekstu poprzez współpracę online nad jednym dokumentem),
- generowaniu pomysłów (wymyślanie licznych zastosowań dla cegły),
- wybieraniu spośród wielu możliwości (test matryc Ravena),
- negocjowaniu (symulowana dyskusja na temat planowania wycieczki przez współlokatorów).

Dzięki wynikom badacze potwierdzili, że w istocie możemy mówić o stwierdzeniu inteligencji u grupy ludzi (lub superumysłu złożonego również z nie-biologicznych agentów), a następnie jej badaniu czy też mierzeniu. Następnie postuluje również, że skoro możliwy jest pomiar, to możliwe jest również porównywanie grup i ustalenie, która z nich jest bardziej inteligentna.

Czynniki wskazujące na wysoką inteligencję grupową

Podobnie jak z indywidualną inteligencją ogólną, która odpowiada za większość wyników testów inteligencji, tak grupy również wykazały, że istnieje wspólny parametr, korelujący pozytywnie z większością mierzonych cech.

Na to, jak dana grupa poradzi sobie z bardziej skomplikowanymi zadaniami, nie wpływała ani średnia, ani maksymalna inteligencja indywidualna członków grup. Było to zaskakujące, gdyż badacze spodziewali się, że właśnie im średnio bardziej inteligentni będą członkowie grupy, tym lepiej będą sobie radzić w badaniach inteligencji

grupowej. Podobnie wybitne jednostki o ponadprzeciętnej wobec grupy maksymalnej inteligencji okazały się jedynie średnio istotnym czynnikiem (Malone, 2018, s. 34).

Zamiast tego odkryto trzy parametry, które pozwalały trafnie przewidzieć na jakim poziomie będzie inteligencja grupy. Pierwszym była średnia grupowa percepcja społeczna (org. social perceptiveness, tłumaczenie własne). Badacze określali jej poziom za pomocą testu polegającego na pokazywaniu badanym oczu i określaniu stanu mentalnego sfotografowanych osób. Test ten był oryginalnie stosowany do badania autyzmu i podobnych zaburzeń, lecz został dostosowany w tym badaniu do mierzenia inteligencji społecznej badanych.

Drugim czynnikiem był równy podział czasu, ile badani mówili w obrębie grupy. Grupy, w których komunikacja była wyrównana i badani mówili podobną ilość czasu, radziły sobie lepiej niż te zdominowane przez jedną lub dwie osoby.

Ostatecznie, inteligencja grupowa była tym wyższa, im większy odsetek grupy stanowiły kobiety. Aczkolwiek, jak zauważa Malone, może być to powiązane z faktem, że kobiety ogólnie lepiej radzą sobie z zadaniem z pierwszego czynnika, percepcji społecznej, więc ich spodziewany poziom inteligencji społecznej jest zazwyczaj wyższy (Malone, 2018, s. 34).

5.1.2. Sztuczna inteligencja

Po co nam komputery?

Największą zasługą komputerów według Malone'a nie jest możliwość stworzenia sztucznej inteligencji, lecz bezprecedensowe połączenie ludzi czy, jak on to określa, ich hiperpołączenie (org. hyperconnectivity, tłumaczenie własne). Jak rozwija, mamy tendencję do równoczesnego przeceniania możliwości sztucznej inteligencji i niedoceniań efektów jakie może dać hiperpołączenie siedmiu miliardów ludzi. Jak podejrzewa, wynika to z faktu, że łatwiej nam stworzyć takie połączone grupy osób i maszyn niż wyobrazić sobie, czego mogą oni razem faktycznie dokonać (Malone, 2018, s. 13).

Sztuczna inteligencja

Malone zaczyna od stwierdzenia definicyjnych trudności jakich następcza sama nazwa sztuczna inteligencja. By ją przybliżyć, odwołuje się do określeń „maszyn, które myślą lub zachowują się jak ludzie” lub „maszyn które zachowują się racjonalnie”. Przywołuje również test Turinga, który streszcza, stwierdzając, że przydaje on miano inteligencji bytowi, którego człowiek, na podstawie przeprowadzonej

konwersacji, nie jest w stanie jednoznacznie określić jako człowieka lub maszynę (Malone, 2018, s. 63). Warto podkreślić, że ta metoda również odnosi się do komunikacji.

Ostatecznie jednak decyduje się na prostszą definicję. Według której sztuczną inteligencją cechują się maszyny wykazujące się inteligencją. Odnosi się w tym do opisanych wcześniej definicji inteligencji wyspecjalizowanej i ogólnej oraz zastrzega, że współcześnie wszystkie maszyny wykazują się jedynie inteligencją wysoce wyspecjalizowaną (Malone, 2018, s. 63).

Możliwość osiągnięcia przez maszyny inteligencji ogólnej stanowi bardziej skomplikowany temat. Malone zaznacza, że zdania badaczy są w tej kwestii podzielone. Najpierw przywołuje tutaj Russela i Norviga jako przykład osób, które stwierdzają, że maszyny nigdy nie będą jej posiadały (Malone, 2018, s. 64). Rozwija to stanowisko, odwołując się do filozofów Huberta Dreyfusa oraz Johna Searle'a, pisząc, że, według nich, nawet jeśli maszyny będą w stanie robić wszystko to, co ludzie, to i tak nigdy nie osiągną inteligencji, gdyż tylko ludzie mogą być inteligentni (Malone, 2018, s. 64).

Kwestię tejże inteligencji omija poprzez analogię do Edsgera Dijkstry, który zapytał, czy łódź podwodna pływa i ostatecznie stwierdza, że pytanie o to, czy maszyny mogą być inteligentne, rozbija się wyłącznie o nasze rozumienie tego słowa i wskazuje, że zaproponowana przez niego definicja jak najbardziej może dotyczyć również agentów syntetycznych (Malone, 2018, s. 64).

Jako przykłady innych osób stwierdzających, że maszyny mogą osiągnąć inteligencję ogólną podaje również Erika Brynjolfssona oraz Andrew McAfee'ego (Brynjolfsson i McAfee, 2014). Spośród badaczy spodziewających się osiągnięcia tego poziomu inteligencji maszyn w perspektywie czasowej rzędu setek lat przywołuje Rodneya Brooksa (Brooks, 2017).

Z drugiej jednak strony, Malone kontynuuje, powołując się na badania Stuarta Armstronga oraz Kaji Sotali. Wykazali oni bowiem, że wszelkie przewidywania dotyczące rozwoju sztucznej inteligencji są notorycznie nietrafione. Badacze, w latach 1950-2021, stale przewidywali, że maszyny osiągną poziom ogólnej inteligencji za 15-25 lat. Stale, w sensie, że niezależnie od tego, kiedy dokładnie dokonywana była predykcja. Jak parafrazuje Malone, jesteśmy więc około 20 lat od osiągnięcia tego poziomu przez ostatnich 60 lat (Malone, 2018, s. 65). Warto odnotować, że stanowisko to jest zbieżne z później wykonanymi badaniami cytowanymi przez Nicha Bostroma (Bostrom, 2016, s. 41-44).

Malone podsumowuje, stwierdzając, że, o ile nie nastąpi jakaś nieprzewidywalna katastrofa, efektywnie kiedyś powinniśmy osiągnąć tak rozwiniętą sztuczną inteligencję, ale nie spodziewa się, by miało to nastąpić wcześniej niż za kilka dekad (Malone, 2018, s. 65). Następnie przypatruje się on czterem źródłom, które mogą nam pozwolić na osiągnięcie takiej technologii: wiedza powszechna, big data, uczenie maszynowe, obliczenie neuromorficzne (Malone, 2018, s. 69-72).

Wiedza powszechna

Wiedza powszechna oznacza tutaj założenie, że maszyna ma dostęp do pokładów wiedzy ukrytej, która umożliwi jej odczytywanie komunikatów na wielu poziomach. Na przykład słysząc, że odczuwamy ból głowy, będzie w stanie skojarzyć, że:

- ból głowy to pewien stan,
- stan ten jest zazwyczaj nieprzyjemny,
- ludzie chcą unikać rzeczy i stanów, które są nieprzyjemne,
- istnieją leki, które zostały opracowane specjalnie, żeby pozwolić ludziom przestać odczuwać ból głowy.

I, dzięki takiego rodzaju powiązaniom, maszyna w reakcji na nasz komunikat będzie w stanie zaproponować nam pigułkę na ból głowy (Malone, 2018, s. 69).

Naturalnie, wypisanie tego rodzaju faktów nie stanowi problemu, ukrytym tu wyzwaniem jest skala. Aby mogło to funkcjonować, system musiałby mieć miliony takich faktów, żeby w ogóle zacząć działać. W istocie podjęto się już przynajmniej raz próby stworzenia takiego programu. W 1984 roku wystartował projekt Cyc, prowadzony przez naukowca Douga Lenata (Malone, 2018, s. 69).

Jak stwierdza Malone, nie ma zgody wśród badaczy w jakim stopniu realną możliwość stanowi stworzenie takiej bazy faktów, aczkolwiek mniejsze i bardziej wyspecjalizowane jej wersje są już używane. Za przykład przywołuje on klinikę medyczną w Cleveland, której pracownicy są wspierani przez automatyczne systemy do zbierania badanych o pożądanym cechach do określonych badań klinicznych (Malone, 2018, s. 70).

Big data

Big data określa wpływ dużych zbiorów danych jakie stały się dostępne dzięki gwałtownemu rozwojowi technologii w ostatnich latach na badania nad sztuczną inteligencją. Malone przywołuje tu zwłaszcza rozwój tłumaczenia maszynowego, które bez tych technologii stanowiło nieosiągalny cel dla badaczy sztucznej inteligencji, jak wspomniano w rozdziale 2.1.5. Współcześnie nie musimy już programować a priori

określonych zasad językowych, które pozwolą programowi w trakcie tłumaczenia rozróżnić różne znaczenia określonego słowa. Zamiast tego, programiści przygotowują algorytm, który automatycznie przetwarza tysiące dokumentów, źródeł internetowych czy temu podobnych i w ten sposób wyodrębnia konteksty i formy użycia tego słowa w różnych okolicznościach (Malone, 2018, s. 71).

Uczenie maszynowe

Uczenie maszynowe jest powiązane z big data. Tu również za cel stanowi wytwarzanie metod umożliwiających maszynom uczenie się na podstawie udostępnionych przykładów. Ponownie, zamiast pisać dokładne zasady, programiści przedstawiają maszynom metody nauki inspirowane tym, jak uczą się ludzie. Chociaż mechanizm jak my sami się uczymy, nie jest dokładnie znany, badacze osiągają doskonałe rezultaty, bazując na tym, co wiemy.

W obrębie uczenia maszynowego wyróżniamy dwa główne podejścia: uczenie nadzorowane i uczenie bez nadzoru. To pierwsze polega na pokazaniu programowi szeregu odpowiednio oznaczonych danych. Jeśli chcemy nauczyć program rozpoznawać koty, możemy przedstawić szereg zdjęć i oznaczyć je „kot” oraz „nie-kot”. Następnie program sam przeanalizuje te zdjęcia i wytworzy swoje własne zasady rozpoznawania kota, które bynajmniej nie muszą być logiczne z naszego punktu widzenia. Zamiast szukać czterech łap i ogona program będzie mógł wyłapać zasadę, że powinien szukać czarnego piksela w prawym górnym rogu zdjęcia.

Stąd też programy te wymagają olbrzymich ilości możliwie różnorodnych zdjęć, które pozwolą uniknąć sytuacji jak błędne klasyfikowanie kota, bo ustawił się na tle białych płytek w łazience. Uczenie bez nadzoru działa podobnie, lecz nie ma już oznaczeń. Zbiór obrazów, na których trenuje maszyna, nie jest etykietowany i algorytmy, podobnie jak człowiek, samodzielnie uczy się rozpoznawać pożądane własności obrazu, jak przedstawianie kota (Malone, 2018, s. 72).

Obliczenia neuromorficzne

Obliczenia neuromorficzne czerpią inspirację z neuronauk poznawczych. Zamiast opierać się na klasycznych procesorach, dążą one do odtworzenia pracy mózgu, który niektórzy badacze przyrównują do współpracujących 80-100 miliardów procesorów. Procesorami tymi są neurony. Te komórki są ze sobą połączone w skomplikowany sposób i, w pewnym sensie, działają równolegle.

Teoretycznie byłoby technicznie możliwe odtworzenie ich pracy za pomocą symulacji generowanych na tradycyjnych procesorach, jednakże dużo efektywniejszym podejściem jest wytworzenie specjalnych komputerów, które fizycznie miałyby w sobie taką ilość procesorów. Pracują nad tym liczne grupy badawcze, w tym IBM czy HRL Laboratories. Dzięki odtworzeniu tej fizycznej struktury byłoby możliwe dużo wierniejsze odtworzenie pracy mózgu i, przynajmniej hipotetycznie, wynikającej z niej inteligencji (Malone, 2018, s. 73).

Jak zauważa Malone, obliczenia te podnoszą ciekawą kwestię. Wiemy bowiem, że ludzki mózg sam w sobie stanowi formę inteligencji zbiorowej (org. collective intelligence), gdyż składa się z miliardów indywidualnych neuronów, które, pracując razem, zachowują się w sposób posiadający znamiona inteligencji.

Metoda stworzenia ogólnej sztucznej inteligencji

Malone kontynuuje, rozważając, że być może oznacza to ogólnie, że najlepszym sposobem na stworzenie prawdziwej ogólnej inteligencji będzie stworzenie systemu łączącego liczne sztuczne inteligencje. Przywołuje on tutaj Marviną Minsky'ego, który sugerował stworzenie społeczeństwa umysłów (org. society of mind). W zamyśle tegoż, jak rozwija Malone, umysł wyłania się emergentnie z interakcji wielu mniejszych agentów, z których żaden z siebie nie jest specjalnie inteligentny, ale razem tworzą oni system, który znamiona inteligencji już posiada (Malone, 2018, s. 73). Stwierdza wręcz, że stworzenie prawdziwej, sztucznej inteligencji, charakteryzującej się zdefiniowaną przez niego inteligencją ogólną, będzie możliwe tylko w ten sposób. Jednakże postrzega to również jako stan, którego długo nie osiągniemy (Malone, 2018, s. 74).

Za przykład takiego systemu Malone przytacza Watsona. Jest to program IBM, który wygrał w telewizyjnym konkursie w Jeopardy. Zawiera on w sobie tysiące mniejszych agentów, z których wielu pracuje równolegle nad różnymi procesami. Każdy z nich jest bardziej skomplikowany niż pojedynczy ludzki neuron, ale też żaden z nich indywidualnie nie dałby sobie rady w podobnym konkursie. Jego mechanizm działania polega na rozbijaniu pytań na serię działań podejmowanych przez różnych agentów. I tak, gdy miał wymienić prezydenta Stanów Zjednoczonych, który uznał władze Chińskiej Republiki Ludowej, rozbił to na pracę dla wielu agentów. Kilku z nich przygotowało listę prezydentów stanowiących bazę możliwych odpowiedzi. Inni agenci zaczęli sprawdzać encyklopedie i podobne źródła wiedzy pod kątem wybranych haseł „Stany Zjednoczone”, „uznawać władze”, „Chińska Republika Ludowa”.

Za pomocą specjalnego kodowania tych informacji wraz z odwołaniem się do wiedzy powszechnej opisanej wcześniej, agenci ci opracowali więcej możliwych odpowiedzi, prawdopodobnie uwzględniających szerszą listę władz państwowych zaangażowanych w to oświadczenie. Ostatecznie, wielu agentów zajęło się oceną różnych odpowiedzi i przeprowadzili *de facto* głosowanie, które z nich są najlepiej umotywowane w różnych źródłach i stanowią prawdopodobnie najlepszą odpowiedź. Wreszcie społeczeństwo agentów Watsona udzieliło poprawnej odpowiedzi: „Jimmy Carter” (Malone, 2018, s. 74).

Rola sztucznej inteligencji

Malone dostrzega, że rola sztucznej inteligencji będzie się zmieniać i maszyny będą przejmować coraz więcej, coraz bardziej złożonych zadań, które współcześnie uważamy za wyłącznie ludzkie. Malone stwierdza, że nie próbuje pokazać, jak maszyny będą przejmować to, co teraz robią agenci oparci na białkach, lecz próbuje wskazać, jak wspólnie będą oni robić rzeczy, które wcześniej uważaliśmy za niemożliwe.

W międzyczasie więc proponuje twórcze rozwinięcie idei M. Minsky’ego i zaaplikowanie jej do budowania społeczności umysłu, gdzie agentami będą liczne podmioty, które, w ujęciu M. Minsky’ego same działałyby na zasadzie osobnych społeczności umysłu (Malone, 2018, s. 73).

Spółeczności te, w ujęciu Malone’a, powinny składać się z ludzi i indywidualnych programów, którzy razem będą realizowali plany, często nawet nie interesując się czy dana część procesu jest wykonywana przez agenta biologicznego, czy syntetycznego. Ludzie mogą w takiej sytuacji dostarczać ogólnej inteligencji i jej wyspecjalizowanych wariantów w obszarach niedostępnych maszynom. Podobnie, maszyny mogą dostarczać przetwarzania danych i innych zdolności, których ludzie nie są w stanie opanować. Co więcej, grupy ludzi i komputerów mogą być bardziej inteligentnej niż jakakolwiek osoba, maszyna czy grupa wcześniej (Malone, 2018, s. 75).

Możliwości współczesnej sztucznej inteligencji

Jednakże, na chwilę obecną sztuczna inteligencja jest daleka od tychże wizji przyszłości. Aby zilustrować współczesną realizację tego terminu, ponownie przywołuje on wyszukiwarkę Google. Ta aplikacja stanowi raczej narzędzie niż asystenta lub współpracownika. Nie wygląda ani nie zachowuje się jak człowiek. Zazwyczaj też nie komunikuje się z ludźmi za pomocą języka naturalnego. Stąd też daleko jej do wizji SI jakie znamy z filmów science fiction- humanoidalnych robotów pokroju Terminatora lub C-3PO z Gwiezdnymi Wojen (Malone, 2018, s. 49).

Aby nakreślić oczekiwaną wizję robotów, Malone odwołuje się do filozofa Marshalla McLuhana, który stwierdził, że aby zrozumieć nowe technologie, chętnie odnosimy je do starych (McLuhan, 1964).

Za przykład podaje samochody, które porównywano do zaprzęgniętych w konie bryczek. W podobny sposób kontynuuje Malone, ludzie spodziewają się, że syntetyczni agenci będą podobni do ludzi. Jednakże współczesne boty rzadko wyglądają w ten sposób. Są raczej bezcielesnymi quasi-inteligentnymi bytami poruszającymi się w chmurze czy też cyberprzestrzeni.

Nawet roboty wykonujące fizyczną pracę nie przyjmują takich kształtów. Zamiast tego mamy do czynienia z małymi dronami unoszącymi się w powietrzu za pomocą małych śmigieł, pomarańczowe sześcienny przesuwające ładunki w magazynach Amazona czy czteronogie maszyny od Boston Dynamics. Autor dostrzega dwa wyjątki, które poniekąd będą musiały mieć formę dwunożną: maszyny, które będą się musiały przemieszczać w tych samych środowiskach co my (na przykład w naszych domach albo hotelach) oraz roboty, które będą wchodziły z nami w interakcję, po której będziemy się spodziewać ludzkich reakcji (na przykład w hotelowej recepcji). W ich przypadku forma dwunożnych humanoidów będzie wymagana. Aczkolwiek, warto odnotować, że drugi punkt może zostać ominięty przez ekrany wyświetlające twarze, jak w przywołanym przez Malone'a robocie Baxter firmy Rethink Robotics (Malone, 2018, s. 59).

Komunikacja z maszynami

Komunikacja międzyludzka ma za sobą setki tysięcy lat ewolucji biologicznej i kulturowej. Dzięki temu opracowaliśmy rozliczne sposoby przekazywania i odbierania wiadomości. Najszerzej rozpowszechnionym sposobem jest rozmawianie, z wykorzystaniem słów, w językach naturalnych jak angielski czy polski. Nie do przecenienia jest również opracowanie ich formy pisanej. Nie możemy jednak zapomnieć o roli komunikacji za pomocą gestów czy wyrazów twarzy. Podobnie, możemy wiele przekazać za pomocą mniej popularnych kanałów komunikacji jak taniec czy śpiew, czy też ogólnie szeroko rozumiana sztuka (Malone, 2018, s. 60).

Naturalnie oczekujemy więc, że komunikacja z syntetycznymi agentami będzie równie złożona. Obecnie jesteśmy w stanie wydawać komunikaty agentom takim jak Alexa od firmy Amazon czy Siri od Apple. Jednakże daleko im do prowadzenia

naturalnej konwersacji. Opierają się one na wychwytywaniu słów kluczy i zupełnie nie są w stanie zrozumieć niuansów jak choćby ironia (Malone, 2018, s. 60).

Z drugiej strony maszyny potrafią się z nami komunikować na nowe sposoby, które potrafią być doskonalsze od komunikacji międzyludzkiej. Przykładowo, człowiek zapytany o drogę najprawdopodobniej będzie opisywał ją za pomocą wskazywania punktów orientacyjnych, machania i opisywania sekwencji zakrętów jakie trzeba przebyć. Tymczasem, jeśli zapytamy o to samo aplikację pokroju Map Google, wyświetli nam ona drogę zaznaczoną na mapie oraz zapyta, czy życzymy sobie słyszeć nawigację przypominającą nam, kiedy mamy skręcić w którą stronę tuż przed tym zakrętem. Podobnie możemy sobie wyobrazić trudność w podyktowaniu komuś listu i porównać ją z napisaniem jej w edytorze tekstu (Malone, 2018, s. 60). Nie wspominając o intensywnych pracach nad opracowaniem interfejsów mózg-komputer, które pozwolą nam bezpośrednio łączyć się z maszynami (Malone, 2018, s. 62).

Relacje z maszynami

W tekście zostaje wyróżniona skala obejmująca następujące role: narzędzi, asystentów, partnerów czy, ostatecznie, menadżerów. Maszyny bynajmniej nie zajmują w niej stałego miejsca. Przykładowo autonomiczne auto, dla osoby, którą wiezie, będzie asystentem i, równocześnie, partnerem dla innego uczestnika ruchu. Ponadto doza kontroli, jaką ludzie sprawują, również jest płynna i zmniejsza się wraz z rolami jakie maszyna przyjmuje w relacji z ludźmi. Innymi słowy, najwięcej kontroli mają nad narzędziami, a menadżerom stają się podlegli. (Malone, 2018, s. 50).

Narzędzia

Narzędzia fizyczne takie jak młotek czy kosiarka do trawy zwiększają możliwości ludzi. Często umożliwiają wykonywanie zadań wcześniej nieosiągalnych. Jednakże, w obcowaniu z nimi w pełni kontrolujemy sytuację, a maszyna wykonuje dokładnie nasze polecenia (nawet jeśli roz mijają się one z naszymi intencjami).

Te same stwierdzenia możemy odnieść do narzędzi informatycznych. Pod wieloma względami narzędzia wymiennie zwiększają naszą wyspecjalizowaną inteligencję. Przykładowo, analityk finansowy, który korzysta z Excela, może wykonać dużo więcej obliczeń i będą one dokładniejsze niż gdyby miał je przeprowadzić w pamięci.

Malone przewiduje jednak, że główną rolą narzędzi przyszłości będzie zwiększanie naszej inteligencji grupowej poprzez usprawnianie naszej komunikacji z innymi członkami grupy. Autor stwierdza, że efektywnie to już znacząca większość zastosowań naszych

istniejących rozwiązań informatycznych przybliży nas do tego celu. Wymienia tutaj takie wynalazki jak email, SMS-y, aplikacje mobilne i internetowe jak Wikipedia, Facebook, Twitter czy YouTube.

Żadne z tych rozwiązań bynajmniej nie jest inteligentne. Ich rolą jest ściśle transfer informacji między ludźmi. Malone określa to wprost, stwierdzając, że kolejna dekada lub dwie, jeśli nie więcej, upłynie na wynajdywaniu kolejnych, coraz bardziej wyrafinowanych sposobów komunikowania się ze sobą na drodze do wspomnianego wcześniej hiperpołączenia (Malone, 2018, s. 51).

Asystenci

Ludscy asystenci, w przeciwieństwie do narzędzi, mogą pracować bez sprawowania nad nimi bezpośredniej kontroli i często wykazują inicjatywę w osiąganiu celu, które przed nimi postawiliśmy. Automatyczni asystenci są podobni, aczkolwiek należy zauważyć za Malonem, że granica między narzędziem a asystentem nie jest zawsze łatwa do wyznaczenia (Malone, 2018, s. 51).

Za przykład możemy tu przytoczyć automatyczną korektę tekstu, która zazwyczaj działa na podstawie zaprogramowanych zasad, ale czasem wykaże się inicjatywą i poprawi coś, co sama błędnie sklasyfikuje jako błąd. Na podobnym, granicznym, poziomie znajdzie się również wyszukiwarka Google, która nie tylko oczekuje na zapytanie od użytkownika, ale w tle wykonuje dodatkowe obliczenia, które pozwolą przewidzieć, kiedy i o co zapyta użytkownik, żeby szybciej przedstawić mu pożądane informacje.

Na wyższym poziomie (mniej narzędzia, bardziej asystenta) znajdują się pomocnicy jak Asystent Google czy Alexa. Nie tylko odpowiadają oni na zadane pytania, ale również z własnej inicjatywy oferują nam oni dodatkowe informacje, na przykład przypominając o konieczności wyjścia z domu, żeby zdążyć na samolot.

Ciekawym przykładem jest też oprogramowanie używane w firmie Stitch Fix. Jej klienci najpierw wypełniają dokładne ankiety o swoim stylu ubierania się oraz preferowanych rozmiarach i zakresie cenowym. Następnie algorytmy uczenia maszynowego przetwarzają te dane i wybierają z oferty sklepu produkty, które sugerują ludzkim stylistom do rozważenia. To jednak ci ostatni podejmują ostateczną decyzję jakie pięć ubrań trafi do klienta. Na końcu nabywcy wybierają przedmioty, które chcą zachować, tylko za nie płacą a resztę odsyłają. Warto podkreślić, że jest to usługa abonamentowa i proces ten powtarza się co miesiąc, naturalnie z wyłączeniem wysyłania ankiety. Informacje o tym, które produkty zostały zachowane, również zostają uwzględnione przez algorytmy w kolejnych iteracjach. Dzięki pracy algorytmów możliwe zostało

odkrycie, że dobór spodni dżinsowych na odległość najlepiej dokonywać na podstawie wymiarów nogawek produktów, które klienci zachowali. Z drugiej strony, algorytmy nie mają pojęcia o modzie albo wymogach określonych okazji typu przyjęcie urodzinowe czy spotkanie biznesowe. Naturalnie, ludzcy styliści dodają również osobistego akcentu proponowanym ubiorom. Stąd też możemy stwierdzić, że zarówno algorytmy, jak i styliści, wnoszą coś unikalnego do tworzonego rozwiązania (Malone, 2018, s. 52).

Partnerzy

Kolejny poziom oddawania kontroli opisuje sytuację, gdzie maszyny przyjmują rolę równą ludziom. Mówimy tu o relacji, w której stanowią one współpracowników czy, jako pojazd autonomiczny, innych uczestników ruchu drogowego. Ten drugi przykład jest zresztą ciekawy i pokazuje skalę rozmycia tych kryteriów, gdyż, jak wspomniano wyżej, autonomiczny pojazd będzie asystentem dla swojego pasażera, ale będzie równoważnym uczestnikiem ruchu drogowego dla innego kierowcy (Malone, 2018, s. 53).

Za przykład już istniejących współpracowników możemy tu przywołać boty handlujące akcjami na giełdzie, rozpatrujące zgłoszenia po stronie ubezpieczyciela czy szereg czat-botów (czyli botów wchodzących z ludźmi w interakcje na poziomie wymiany wiadomości tekstowych za pomocą specjalnego czatu na stronie firmowej) przyjmujących zgłoszenia usterek i awarii produktów.

Innym przykładem takiego bota są programy skanujące Wikipedię i dokonujące automatycznych poprawek artykułów. Niektóre usuwają edycje o znamionach wandalizmu jak, chociażby zawierające słowa powszechnie uznawane za wulgarne. Inny rodzaj bota sprawdza regularnie, czy duże bloki tekstu nie są po prostu skopiowane z innych stron internetowych. Tak więc ludzie i boty występują tutaj równolegle, wszyscy zajmują się edycją artykułów. Warto jednak odnotować, że boty te są często poniżej w hierarchii użytkowników i, wspomniane wyżej, boty przeszukujące artykuły w poszukiwaniu znamion plagiatu same nie przeprowadzają edycji, zamiast tego jedynie zgłaszają błąd, by człowiek podjął ostateczną decyzję (Malone, 2018, s. 54).

Przełożony

Jak stwierdza Malone, ludzcy menadżerowie delegują zadania, wskazują kierunki rozwoju, ewaluują pracę i koordynują zadania podwładnych (Malone, 2018, s. 55). Już teraz maszyny również mogą wykonywać te zadania. Za nieco przewrotny przykład Malone podaje tutaj światła uliczne, które pełnią tę samą funkcję co policjant stojący

na środku skrzyżowania lub maszynę przekierowującą połączenia telefoniczne w centrum obsługi klienta (Malone, 2018, s. 55).

Za bardziej imponujący przykład przyjmuje on system CrowdForge, który zatrudnia pracowników online, by wspólnie pisali dokumenty podobne do artykułów z encyklopedii, na przykład, aby opisać miasto Nowy Jork. W tym celu wykorzystują oni usługę Amazona Mechanical Turk. Każdy artykuł jest wstępnie opisany przez autorów, w podanym przykładzie, przytoczony zostanie krótki opis, lista atrakcji turystycznych oraz pobieżnie opisana historia miasta. Następnie system automatycznie przypisuje zarejestrowanych w Mechanical Turk pracowników, aby manualnie poszerzali ten opis, przykładowo dokładnie opisując jeden z zabytków. Na koniec system automatycznie zbierze te fragmenty i stworzy z nich jeden artykuł. Tak napisane artykuły były dużo lepszej jakości niż praca pojedynczej osoby za tę samą stawkę, a poziomem języka dorównywały artykułom na Wikipedii. Jak podsumowuje Malone, badacze zbudowali automatycznych asystentów dla siebie, którzy występują w roli przełożonych dla pracowników zarejestrowanych w Mechanical Turk. Relacja jaką ci pracownicy mieli z systemem jest bardzo zbliżona do tej jaką mieliby z ludzkim przełożonym, który delegowałby im pracę.

5.1.3. Definicja superumysłu

Jak jednak opisać taką relację ludzi i maszyn? Malone proponuje tutaj koncepcję superumysłów. Stanowią one zbiory lub raczej grupy jednostek, które poprzez efektywne współdziałanie tworzą na obserwatorze impresję, że są inteligentne.

Na początek warto wyjaśnić, że pojęcie superumysłu bynajmniej nie zostało wymyślone przez Malone'a. Według słownika Merriam-Webster pierwsze użycie tego słowa padło w 1903 roku. Jedno z przytoczonych w nim znaczeń oznacza jedność stworzoną z licznych, kolektywnych myśli, wspomnień, inteligencji czy wnioskowań wykonywanych przez grupę ludzi lub maszyn (Merriam Webster).

Malone traktuje to pojęcie jako punkt wyjścia do stwierdzenia, że historia ludzkości to, w istocie, przebieg szeregu interakcji ledwie kilku typów superumysłów, których dokonania znacząco przekraczają możliwości dowolnej jednostki.

Ponadto przedrostek „super-” nie odnosi się tutaj do określenia ponadprzeciętnej jakości umysłu, lecz raczej do jego większej otwartości i inkluzywnego charakteru opisywanych grup. Malone przybliży to za pomocą analogii. Pisze, że tak, jak

superorganizmem określamy kolonię mrówek, tak superumysłem będzie cała firma (Malone, 2018, s. 12).

Ostatecznie, autor definiuje superumysł następująco: grupa stworzona z jednostek, których wspólne działanie ma znamiona inteligencji. Następnie rozbija ją, dookreślając, co rozumie przez każdą z części tego stwierdzenia, czyli konkretnie: grupa, jednostka, wspólne działanie, znamiona inteligencji.

I tak grupa jest zazwyczaj łatwa do określenia. Malone podaje za przykład pracowników restauracji, ekipę remontową, ale też zbiór ludzi, którzy akurat znaleźli się w tym samym momencie na tej samej ulicy i którzy, bez świadomego wysiłku, unikają zderzenia się.

Jednostki są pojmowane nie tylko jako czyste umysły, ale też ciała czy zasoby pozostające pod kontrolą tychże umysłów. Superumysł lokalnej kawiarni nie obejmuje wyłącznie jego pracowników, ale też ziarna czy ekspresy do kawy, stoły, krzesła i temu podobne.

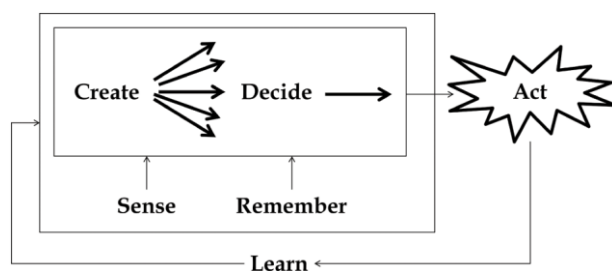
Alternatywnie, superumysły mogą stanowić wszystkie kawiarnie w okolicy, które rywalizują o tych samych klientów. Jeszcze inaczej, na potrzeby kolejnej analizy, superumysł może stanowić pojedynczy barista przygotowujący kawę, którego komponenty składowe stanowią neurony w jego mózgu (Malone, 2018, s. 21).

Wspólne działanie jest kluczowe dla zrozumienia tej definicji. Przede wszystkim musi zachodzić jakakolwiek aktywność, byśmy mogli mówić o superumyśle, a dodatkowo musi ona obejmować przynajmniej dwie jednostki działające w pewien sposób razem. Nie musi to być działanie skoordynowane czy kooperatywne jak w przytoczonym wyżej przykładzie o wymijaniu się na ulicy.

Znamiona inteligencji również są kluczowe, zwłaszcza pierwsza połowa te określenia, której rolą jest podkreślenie roli obserwatora w procesie determinowania czy coś jest superumysłem. Malone podkreśla tutaj subiektywność tego procesu i poniekąd przerzuca na obserwatora ciężar związany z określeniem statusu obserwowanej grupy, czy stanowi ona superumysł. Rola osoby obserwującej jest bowiem krytyczna w ustalaniu inteligencji, gdyż ta, według Malone'a, polega na ustaleniu celów percypowanego, potencjalnego superumysłu oraz jego skuteczności w ich osiągnięciu, a przynajmniej próbuje, gdyż sukces nie powinien być traktowany jako wyznacznik inteligencji, lecz raczej celowość działań (Malone, 2018, s. 21).

5.1.4. Superumysł jako wsparcie procesów poznawczych

Malone podchodzi do superumysłów jako realizujących określony schemat:



Rysunek 2 Schemat procesów myślowych (Malone, 2018, s. 82)

Obejmuje on szereg aktywności, które wzajemnie się wywołują. Najgłębiej zinternalizowane są akty tworzenia (org. Create, tłumaczenie własne) oraz decydowania (org. Decide, tłumaczenie własne). Tworzenie polega na generowaniu alternatywnych scenariuszy opisujących możliwe rozwiązania problemów, jakie napotykaemy. Decydowanie z kolei dotyczy wybrania z nich najlepszej opcji.

Obie te aktywności są zasilane przez doświadczanie (org. Sense, tłumaczenie własne) i pamiętanie (org. Remember, tłumaczenie własne). Doświadczanie oznacza zbieranie danych. Czy to poprzez naturalne ich odbieranie jako bodźców zmysłowych, czy też za pomocą technologicznych sposobów opisujących przykładowo zachowania konsumenckie naszego sklepu online. Pamiętanie znow odnosi się do podsuwania znanych nam faktów, które mają znaczenie dla akurat rozpatrywanej sytuacji.

Wszystkie te cztery akty mentalne prowadzą nas do działania (org. Act, tłumaczenie własne). Jest to proces implementacji opcji wybranej na drodze wspomnianego wcześniej decydowania, które nie jest samo z siebie procesem poznawczym. Natomiast uczenie się (org. Learn, tłumaczenie własne) w ujęciu Malone'a oznacza obserwowanie skutków podjętych działań i aktualizowanie naszych przyszłych procesów decyzyjnych tymi nowymi faktami.

Warto podkreślić, że wszystkie te aktywności są realizowane przez cały superumysł, więc równie dobrze mogą odnosić się do grup złożonych z ludzi, maszyn lub ich kombinacji, jak również korzystać z opisanego wcześniej hiperpołączenia.

Tworzenie

Kluczowym procesem poznawczym w generowaniu rozwiązań dla problemów jest, zwłaszcza w ujęciu Malone'a, generowanie licznych scenariuszy jak dany problem możemy, w istocie, rozwikłać. Autor wymienia dwa główne sposoby na zwiększenie ilości

tychże opcji: zwiększenie liczebności członków grupy oraz optymalizacja jej pracy i komunikacji .

Zaangażowanie większej ilości osób – Malone dostrzega analogię między większą ilością neuronów w mózgu, a większą inteligencją zwierzęcia i, podobnym efektem jaki ma na procesowanie informacji przez komputer, zwiększenie ilości tranzystorów. Stąd też uznaje za zasadne, że im więcej indywidualnych agentów w obrębie superumysłu, tym inteligentniejszy się on stanie (Malone, 2018, s. 174).

Po pierwsze, zachodzi w ten sposób efekt siłowy (org. the brute force effect, tłumaczenie własne). Jak to czasem się określa „duży może więcej”. Lub, już słowami Malone’a, większe grupy są w stanie osiągnąć więcej. Dzięki współczesnej technologii jesteśmy w stanie łączyć się łatwiej i w większe grupy. Malone w tym kontekście przytacza przykład Wikipedii, którą zarządza 70,000 wolontariuszy z całego świata (Malone, 2018, s. 175).

Dodatkowo autor dostrzega, tak zwaną, mądrość tłumu, za książką *The Wisdom of Crowds* Jamesa Surowieckiego. Dowodzi on w niej, że kiedy zapytamy wiele osób o spodziewane wartości czegoś, ich odpowiedzi będą różnorodne i mocno rozproszone, ale jeśli poddamy je obróbce statystycznej i uśrednieniu to, ostatecznie, otrzymamy wartość zbliżoną do prawdy. Co zajmujące, im więcej osób zapytamy, tym ostatecznie bliższą rzeczywistości odpowiedź otrzymamy. W nieco abstrakcyjnym przykładzie przytoczonym przez Surowieckiego angielski statystyk Sir Francis Galton przebadął opinie ludzi biorących udział w konkursie na odgadnięcie wagi osła. Średnia wartość wszystkich odpowiedzi (542,95 kilogramy) była bardzo zbliżona do realnej wartości (543,4 kilogramy). Surowiecki podaje również przykłady zgadywania ilości fasolek w słoiku czy lokalizacji łodzi podwodnej (Surowiecki, 2004). Malone zauważa dwie wady w opisanym podejściu. Po pierwsze, w swoich badaniach wykazał, że dużo lepszą miarą jest mediana. Jak również, że ludzie biorący udział w zgadywaniu nie mogą wiedzieć, co pisali pozostali, by się nimi nie sugerować (Malone, 2018, s. 178).

Inną zaletą większych grup jest dostęp do różnych zasobów specjalistycznej wiedzy. Nasz współczesny świat jest tak skomplikowany, że musimy się specjalizować w jednej tylko rzeczy, aby być w niej ekspertami. Łączenie ekspertów z różnych obszarów pozwala na integrację ich perspektyw i możliwości. Dodatkowo Malone zwraca uwagę, że zwiększenie liczebności grupy daje dostęp do wiedzy nieoczywistej, opartej na luźnych skojarzeniach. Kontrintuicyjnie, jak dowodzą Lars Jeppesen i Karim Lakhani, prawdopodobieństwo wpadnięcia na pomysł, który będzie stanowił inspirację

do rozwiązania nietrywialnego problemu z określonej technicznej dyscypliny, rośnie odwrotnie proporcjonalnie do doświadczenia i wiedzy z tej dyscypliny (Jeppesen i Lakhani, 2010).

Zmiana sposobu pracy – krytyczne, w połączeniu tranzystorów i przepustowości synaps, jest określenie i optymalizacja ich metod działania. Stąd też kolejnym sposobem na ulepszenie procesu tworzenia jest optymalizacja sposobów pracy superumysłów poprzez łączenie i organizowanie grup ludzi i komputerów.

Propozycją Malone’a, jak to możemy osiągnąć, jest hieperspecjalizacja. Za konkretny przykład podaje on serwis Mechanical Turk, opisany w rozdziale 5.1.2., podrozdziale „Przełożony”. Zamysł jest tutaj taki, że operacje na informacjach mogą zostać podzielone podobnie, jak Ford dzielił zadania fizyczne w pracy nad konstrukcją aut. Każde zadanie, na przykład napisania artykułu, jak w przytoczonym fragmencie o Mechanical Turku, może zostać podzielone na fragmenty, którymi zajmą się różne osoby. Następnie, czy to sztuczna inteligencja, czy, co równie rewolucyjne, my sami możemy przypisać sobie te zadania. Jak zauważa Malone, nie tylko istotne jest to, w czym jesteśmy dobrzy, ale też to, co nas motywuje i inspiruje (Malone, 2018, s. 195).

Za pragmatyczny przykład rozwiązania usprawniającego tworzenie rozwiązań Malone podaje sieć konkursów (org. contest webs, tłumaczenie własne). Polega ona na tworzeniu dwóch poziomów pomysłów. Pierwszy konkurs agreguje pomysły bazowe, na przykład nowy fizyczny produkt, czy, w kontekście rozwiązań dla kryzysu klimatycznego, propozycje zmian w transporcie publicznym czy sieci energetycznej. Następnie, te bazowe pomysły są agregowane w ramach kolejnego poziomu, konkursów integracyjnych. Kontynuując przykład kryzysu klimatycznego, indywidualne propozycje mogą być łączone w sugestie dla konkretnych krajów czy wręcz rozwiązań globalnych. Odbiorcy rozwiązań podejmują ostateczne decyzje, które pomysły zostaną zaimplementowane, a osoby, które je wymyśliły, dostają odpowiednią kompensację w postaci abstrakcyjnych punktów w kontekście inicjatyw wolontariackich lub adekwatnej kompensacji w postaci bonusu do wynagrodzenia w przypadku przedsiębiorstw i podmiotów publicznych (Malone, 2018, s. 201).

W przeciwieństwie do hierarchii czy demokracji opisanych niżej sieć konkursów oferuje możliwość eksplorowania bardzo dużej ilości pomysłów w tym samym czasie. Metoda ta redukuje skutki pojedynczych pomyłek, zwiększa innowacyjność i elastyczność, gdyż określony pomysł może nie mieć zastosowania w jednym obszarze, ale równocześnie może stanowić doskonałe rozwiązanie w innym. Co więcej, obserwując skutki jego

implementacji, możemy usprawnić kolejne instancje jego wprowadzenia. Możemy więc stwierdzić, że sieci konkursów promują myślenie holistyczne, globalne i wysokopoziomowe. Warto odnotować, że wzajemne odbijanie pomysłów zwiększa jakość współpracy między zaangażowanymi osobami (Malone, 2018, s. 201-203).

Decydowanie

Decydowanie może być realizowane na drodze czterech rodzajów superumysłów: hierarchie, demokracji, rynki i społeczności (Malone, 2018, s. 13). Wymienia on również ekosystemy, jako sytuacje, gdzie nie mamy do czynienia ze współpracą jako taką, lecz okolicznościami, w których panują, tak zwane, „rządy silniejszego”.

Hierarchie

W hierarchii jednostki obdarzone autorytetem podejmują decyzje, a pozostali je wykonują. Malone sugeruje wręcz, że hierarchia to superumysł, który mobilizuje liczne indywidualne byty do osiągania celów kogokolwiek, kto siedzi na jej szczycie. Stąd też rola indywidualnych, ludzkich czynników jak emocje, wyznawane wartości czy ograniczenia mają dużo większe znaczenie niż w innych superumysłach. Alternatywnie, możemy również spojrzeć na nie jako na sytuacje, gdzie decyzje są delegowane i przekazywane kolejnym podmiotom zajmującym niższe stanowiska czy role (Malone, 2018, s. 86).

Technologie informacyjne mogą wesprzeć hierarchie poprzez automatyzację i decentralizację.

Automatyzacja oznacza, że pojedynczy człowiek może uruchomić olbrzymią liczbę procesów komputerowych, które będą realizować jego cele. Za przykład może posłużyć chociażby wyszukiwarka Google. Innymi słowy, człowiek używa swojej ogólnej inteligencji, aby skorzystać z wyspecjalizowanej, maszynowej (Malone, 2018, s. 87).

Decentralizacja jest tutaj użyta dość przewrotnie. Malone w tym kontekście odnosi się do usprawnienia komunikacji. Nowe technologie pozwalają na znaczące ograniczenie kosztów z nią związanych i, stąd też, na popularyzację rozwiązań ją intensyfikujących i, poniekąd, zwiększających. Ludzie mogą więc zdobywać sami informacje, nie muszą czekać, aż zostaną im przekazane w ramach hierarchii. Prowadzi to do tego, że stają się bardziej zmotywowani, kreatywni i elastyczni (Malone, 2018, s. 89). Tak więc, parafrazując, wsparcie informacyjne hierarchii prowadzi do jej załamania.

Demokracje

Demokracje stanowią superumysł, w których podejmowanie decyzji odbywa się na drodze głosowania. Malone przedstawia dwie koncepcje, jak technologie informatyczne mogą je usprawnić.

Pierwsza z nich dotyczy płynnych demokracji (org. liquid democracy, tłumaczenie własne), partii piratów (org. pirate parties, tłumaczenie własne) oraz agentów oddających głosy (org. voting agents, tłumaczenie własne).

Płynne demokracje stanowią ideę, która zakłada, że wszyscy uprawnieni obywatele biorą udział w głosowaniach bezpośrednich oraz że mogą oni oddać swój głos innemu obywatelowi, który zagłosuje za nich. Co więcej, pośrednictwo to może być zależne od tematu głosowania. Przykładowo jedna osoba może oddawać swój głos w głosowaniach dotyczących sądownictwa jednej osobie, a szkolnictwa jeszcze komuś innemu. Reprezentanta też w każdej chwili można zmienić. Daje to więc wyborcy olbrzymią elastyczność w reprezentowaniu jej lub jego opinii, jak również w zarządzaniu własnym czasem, bo przecież nie ma przymusu oddawania głosu żadnemu reprezentantowi (Malone, 2018, s. 97).

Wspomniane partie piratów zostają przytoczone jako przykład płynnych demokracji. W momencie, gdy ich przedstawiciele zostają wybrani do sejmu, uruchomiona zostaje specjalna aplikacja w internecie, która umożliwi obywatelom wskazywanie, jak sobie życzą, żeby głosować w kolejnych głosowaniach (Malone, 2018, s. 98).

Malone dodatkowo kreśli jeszcze jedną możliwość w tym wątku. Hipotetyczną sytuację, w której obywatele mogą oddać swoje głosy w ręce mechanicznych reprezentantów. Niektórzy tacy agenci mogą realizować określone cele polityczne, a inni próbować uczyć się naszych preferencji i nas zastępować (Malone, 2018, s. 98).

Druga koncepcja zmieniająca demokracje dotyczy głosowania jako sposobu na poszukiwanie prawdy. Dotyczy ona projektów crowdsourcingowych, które wykorzystują mądrość uczestników. Malone przywołuje przykład „The Good Judgment Project”, prowadzonego przez Philipa Tetlocka. Była to propozycja w konkursie amerykańskiego wywiadu, społeczności Intelligence Advanced Research Projects Activity (IARPA) na tworzenie nowych metod przewidywania zdarzeń politycznych i odpowiedzi na takie pytania jak:

Czy Serbia zostanie oficjalnie uznana za kandydata do Unii Europejskiej przed 31 grudnia 2011 roku? Czy 30 września 2011 cena uncji złota przekroczy 1850 dolarów amerykańskich? Zamiast prostej, binarnej odpowiedzi „tak” lub „nie”

wymagane było oszacowanie prawdopodobieństw wszelkich wariantów zdarzeń (Malone, 2018, s. 99).

Zespół „The Good Judgment Project” pokonał konkurencję i zaskoczył wszystkich przyjętą taktyką. Nie zatrudnili oni bowiem światowej klasy ekspertów w dziedzinie geopolityki, lecz stworzyli internetową platformę, na której tysiące ludzi mogło oddawać głosy. Następnie wybrali część uczestników, 2% tych, którzy najtrafniej przewidywali rozwój wydarzeń i utworzyli z nich zespoły składające się 12-15 osób. Na koniec zaś przetwarzali dostarczane przez te zespoły predykcje za pomocą skomplikowanych technik statystycznych (Malone, 2018, s. 99).

Nie możemy ocenić skuteczności predykcji, bo nie mamy ich do czego odnieść, ale w listopadzie 2013 roku, w *Washington Post*, ukazał się wywiad z agentem amerykańskiego wywiadu, a przynajmniej kimś z dostępem do zastrzeżonych dokumentów, który stwierdził, że „The Good Judgment Project” radził sobie 30% lepiej niż przeciętny analityk wywiadu z dostępem do wszelkich tajnych źródeł danych (Ignatius, 2013).

Rynki

Decyzje w superumyśle określanym mianem rynków są tak naprawdę wypadkową czy wynikiem umów opartych na wzajemnym porozumieniu zawieranych przez partnerów handlowych. Innymi słowy, decyzje stanowią emergentną własność zawierania umów przez indywidualnych agentów. Jak dowodzą ekonomiści, pisze Malone, taka metoda podejmowania decyzji zazwyczaj prowadzi do optymalizacji wykorzystania zasobów (Malone, 2018, s. 106).

Malone postrzega ten superumysł jako sposób formowania przewidywań przyszłości, zwłaszcza w kontekście rynków prognostycznych. Stanowią one charakterystyczną formę rynku, gdzie uczestnicy kupują, podobnie jak na giełdzie papierów wartościowych, udziały w przewidywaniach przyszłości, na przykład, że określona reprezentacja wygra mistrzostwa świata w piłce nożnej.

Malone wspomina w tym kontekście, jak wspierał swoją studentkę, Yiftach Nagar. Przeprowadziła ona badanie, w którym uczestnicy mieli odgadnąć, jak zachowają się gracze futbolu amerykańskiego. Oglądali oni nagranie chwilę przed rozpoczęciem gry i mieli określić czy piłka zostanie podana do kogoś innego, czy też zawodnik zdecyduje się biec z nią sam.

Co było w tym jednak nietypowe, to fakt, że w obstawianiu wyników nie brali udziału wyłącznie ludzie, ale również agenci syntetyczni. Naturalnie, badani zostali

podzieleni na trzy grupy: handlujących wyłącznie ludźmi, handlujących wyłącznie maszynami oraz grupę mieszaną (Nagar i Malone, 2011).

Zgodnie z założoną hipotezą, grupa mieszana przewidywała wyniki najtrafniej. Jakkolwiek maszyny ogólnie, miały uśredniając lepsze wyniki, tak zdarzały się sytuacje, gdzie ludzie przykładowo rozpoznawali określoną formację po ułożeniu zawodników i wiedzieli, jakie zachowania po niej nastąpią. Maszyny były więc trudniejsze do rozproszenia i systematycznie stosowały obliczenia statystyczne, aby minimalizować szanse pomyłki, jak również obstawiały bardziej zachowawczo. Z drugiej strony, ludzie mieli informacje, do których syntetyczni agenci nie mieli dostępu i mogli lepiej dostosowywać swoje głosy do sytuacji. Jak wspomniano wyżej, połączenie tych kompetencji dało najlepszy wynik (Malone, 2018, s. 111).

Spoleczności

Spoleczności, w ramach których rządzi nieformalny konsensus i współdzielone normy. Przez wzgląd na pojemność tej grupy Malone, w kontekście wsparcia za pomocą technologii informacyjnych, wyróżnia spoleczności oparte na osiaganiu celów biznesowych oraz spoleczności dazące do maksymalizacji poczucia szczescia swoich czlonków.

Te pierwsze wiele zyskują dzięki zintensyfikowaniu interakcji między swoimi członkami. Malone przywołuje firmę Xerox, która powołała usługę Eureka. Była to platforma wymiany wiedzy dla certyfikowanych techników ich maszyn. Dzieli się oni radami i uwagami, a następnie je oceniali, w procesie zbliżonym do recenzji naukowych. Xerox wycenił wartość tej społeczności, estymując, że dzięki usprawnionej wymianie wiedzy między technikami zaoszczędzono na naprawach blisko 100 milionów dolarów (Brown i Duguid, 2000).

Istnieją społeczności online, jak Facebook, które powstają, by zaspokoić pewną emocjonalną potrzebę swoich członków. Dla wielu osób portale jak Facebook, Instagram czy Twitter stanowią teraz główną metodę interakcji z przyjaciółmi. Warto zaznaczyć również powierzchowność takich relacji, jak wykazują inni badacze, jak również odmienne problemy jakie te platformy tworzą, prowadząc do obniżenia poczucia zadowolenia ze swojego życia (Turkle, 2015). Inne społeczności, jak Amazon czy Netflix, używają skomplikowanych algorytmów, aby polecać swoim członkom produkty najlepiej dopasowane do ich preferencji (Malone, 2018, s. 121).

Doświadczenie

W naszym procesie decyzyjnym kluczowe jest zbieranie danych i poznawanie świata. Malone dostrzega, że razem jesteśmy w stanie robić to lepiej i szybciej. Pierwszy przytoczony przez niego przykład dotyczy powrotu wirusa Zika. Ten, rozpoznany w 1947 roku, był uważany za niezbyt groźne schorzenie, występujące głównie w Afryce i przenoszone przez komary. Nagle, w 2016 roku, nastąpiła w Stanach Zjednoczonych fala urodzeń dzieci z pomniejszonymi głowami. Lekarze i naukowcy obserwujący pojedyncze przypadki nie potrafili postawić diagnozy. Dopiero dzięki internetowi i sieci kontaktów emailowych medycy zauważyli skalę problemu. Wspólnymi siłami byli w stanie też porównać swoich pacjentów i wydać trafną diagnozę. Problem jednak był na tyle rozpowszechniony, że World Health Organization musiała wydać ostrzeżenie o publicznym zagrożeniu zdrowia (Malone, 2018, s. 210).

Najpowszechniejszą technologią, która zmieniła i usprawniła nasze akty procesu doświadczenia i zauważania jest big data. Malone dostrzega eksplozję wolumenu danych, które generujemy i jesteśmy w stanie kolekcjonować, lecz uważa go za bezużyteczny bez superumysłów, które te dane by przeanalizowały i podjęły działania na ich podstawie. Rozpatruje on przypadki wykorzystania danych z internetu, publicznie dostępnych zdjęć satelitarnych, jak również dane medyczne i biometryczne oraz dane zebrane dzięki Internetowi Rzeczy, czyli praktyce podłączania do internetu urządzeń zazwyczaj z nim nie kojarzonych, jak chociażby lodówki (Malone, 2018, s. 213).

Przykład, który dokładniej analizuje, dotyczy bezpieczeństwa przed zagrożeniami terrorystycznymi. Ludzie i maszyny muszą w tym obszarze współpracować ściśle, gdyż żadna ze stron nie jest w stanie przejąć tego zadania całkowicie. Zastosowanie znajdują tutaj bayesowskie sieci neuronowe, operujące na faktach i prawdopodobieństwach zdarzeń. Rozumieją je oba typy agentów i są w stanie wspólnie nad nimi pracować (Malone, 2018, s. 217). Malone zagłębia się w to, jak powinny być zorganizowane dane i prawdopodobieństwa zdarzeń, w jaki sposób maszyny mogłyby wykrywać automatycznie podobne sytuacje i sugerować ludziom, że powinni je zbadać. Hipotetyzuje również nad stworzeniem rynku prognostycznego, w którym brałoby udział ludzcy i maszynowi agenci. Jednakże warto odnotować, że opinie ekspertów w tej dziedzinie są miażdżące i zupełnie niezgodne z oczekiwaniami Malone'a (Sætnan, Schneider i Green, 2018).

Pamiętanie

Dla Malone'a pamięć to zarówno „stanie na barkach gigantów” jak w słynnym cytacie Bernarda z Chartres, ale również przechowywanie danych klientów linii lotniczych o tym ile kilometrów przelecieli, aby nadać im vipowski status. Ostatecznie jednak wszystko się sprowadza do trzech operacji: kodowania (org. encode, tłumaczenie własne), przechowywania (org. store, tłumaczenie własne) oraz odzyskiwania (org. retrieve, tłumaczenie własne). Kluczowe jednak jest tutaj, że nie mówimy o pamięci jednostkowej, a raczej kolektywnej, masowej. Zespoły, organizacje i wirtualnie wszystkie grupy mają własne ustalenia w jaki sposób przechowywać pamięć i co w ogóle warto pamiętać. Pamięć staje się więc również obiektem komunikacji i ustaleń (Malone, 2018, s. 223-225).

Za najważniejszą technologię w kontekście pamięci Malone uznaje pisanie. Chronologicznie kolejną było wynalezienie prasy drukarskiej, a współcześnie jesteśmy świadkiem, jak cyfrowe technologie ponownie zmieniają to, jak kolekcjonujemy i magazynujemy obiekty, które chcemy zachować w pamięci, zapamiętać. Jak kontynuuje autor, nie tyle kodowanie i przechowywanie informacji, które są ustawicznie poprawiane, ile odzyskiwanie danych stanowi największe ograniczenie dla technologicznie wzmocnionego zapamiętywania (Malone, 2018, s. 226).

Przykład, któremu Malone się dokładniej przygląda, dotyczy dokumentacji medycznej. Tysiące pacjentów rokrocznie odwiedza szpitale i generuje mnóstwo danych, które możemy wykorzystać do usprawnienia stosowanych przez nas procedur medycznych. Obecnie w momencie, gdy lekarz przyjmuje pacjenta, w pierwszym odruchu odwołuje się on do swojego doświadczenia, czyli przypadków z którymi zetknął się wcześniej. Kluczowe w nauce nowych praktyków medycyny jest nauczanie ich dostrzegania podobieństw i wzorców (Malone, 2018, s. 229).

Projekt, który próbuje usprawnić ten proces z wykorzystaniem technologii informatycznych, nazywa się „Human Diagnosis Project”. Osoby w niego zaangażowane nie starają się zautomatyzować procesu diagnozy dokonywanej przez indywidualnego lekarza, lecz stworzyć możliwość współpracy w procesie grupowej diagnozy realizowanej przez społeczność medyczną. Działa to na zasadzie, że kiedy lekarz stwierdza, że chciałby uzyskać opinię uzupełniającą, wprowadza do systemu symptomy i zanonimizowane wyniki badań oraz swoje podejrzenia. Następnie system pomaga w łatwy sposób udostępnić je ekspertom medycznym z całego świata i uzyskać ich opinię (Malone, 2018, s. 230).

Warto podkreślić skalę projektu, dzięki której pielęgniarzka z afrykańskiej wioski może uzyskać wsparcie od profesora medycyny z Nowego Jorku. System ten jest już

w stabilnym użyciu i pomógł w diagnozie tysięcy przypadków, a badania skuteczności potwierdziły, że opinie grupowe były trafniejsze niż indywidualne. Dodatkowo ten system rośnie. Raz wprowadzone dane zostają zapamiętane i mogą w przyszłości zostać wykorzystane do stworzenia automatycznej metody diagnozowania. Już teraz skuteczność modeli uczenia maszynowego jest zbliżona do studentów medycyny i niewiele gorsza od przeciętnego lekarza, a wraz z rosnącym wolumenem danych będzie wyłącznie rosła (Wong, Lin, Komarneni i Nundy, 2017).

Uczenie się

Malone wyróżnia dwa rodzaje uczenia się: eksploatację (org. exploitation, tłumaczenie własne) i eksplorację (org. exploration, tłumaczenie własne).

Ten pierwszy polega na częstym powtarzaniu i sprawdza się, czy to nauki gry na instrumencie, czy też usprawnieniu procesu produkcji auta. Autor opisuje ten rodzaj nauki na przykładzie Forda, który splótł produkcję Modelu T z iteracyjnym usprawnianiem procesu produkcji, dzięki czemu kolejne modele były wytwarzane coraz efektywniej (Malone, 2018, s. 234).

Drugi rodzaj nauki polega na eksperymentowaniu i sprawdzaniu, co zadziała, zazwyczaj poprzez wypróbowywanie kolejnych opcji. Tak uczą się jednakowo dzieci odkrywając świat, jak i korporacje farmaceutyczne pracujące nad nowymi lekami. Malone rozbudowuje ten opis przykładem rynku startupów technologicznych. Nawet jeśli tylko niektóre z nich są skuteczne, to i tak widzimy które to były i co zrobiły inaczej lub lepiej od innych (Malone, 2018, s. 234).

Najbardziej oczywistym przykładem, jak uczenie może zostać wspomóżone technologią, jest promowanie kolaboracji i dzielenie się lekcjami, które uczniowie poznali indywidualnie. Może to dotyczyć zarówno oczywistej sytuacji w klasie szkolnej, ale równie dobrze sytuacji rynkowej, jak i usprawnienia procesu dzielenia się, tak zwanymi, najlepszymi praktykami w obrębie konkretnych branż (Malone, 2018, s. 234).

Malone wprowadza dodatkowo pojęcie pętli uczenia się, która obejmuje ludzi, środowisko cybernetyczne, sygnały wejściowe i wyjściowe. System komputerowy monitoruje, co jest wprowadzane do niego oraz jaki jest rezultat działań ludzi. W ten sposób zbiera dane i, po jakimś czasie, maszyna jest w stanie sama zacząć generować sugestie, co można zmienić, gdyż wcześniej taka zmiana często zachodziła. Reakcja ludzi na propozycje również jest monitorowana. Jeśli były one często akceptowane, to, po jeszcze dłuższym czasie, proponowane operacje mogą zostać uznane

za kandydatów do automatyzacji, czyli wprowadzania ich natychmiast, bez pytania ludzi o zgodę (Malone, 2018, s. 235).

5.1.5. Podsumowanie

Malone pisze o tym, że powinniśmy mniej myśleć o ludziach i komputerach jako kategoriach rozłącznych. Postulowaną przez niego alternatywą jest poświęcenie więcej uwagi myśleniu o ludziach połączonych z komputerami i o tym, co ludzie i komputery mogą zrobić razem, co nigdy wcześniej nie zostało zrobione.

Aby to osiągnąć, wykazuje on istnienie „duchów”. Potężnych, niewidzialnych bytów, które wyłaniają się podczas interakcji jednostek. Nazywa je superumysłami i definiuje jako grupy złożone z jednostek, których wspólne działanie ma znamiona inteligencji. Superumysły są wszędzie dookoła nas. Ich przykłady obejmują korporacje, kraje, globalną społeczność naukową, ale też klasę w szkole podstawowej czy kawiarnię. Niemniej, superumysły kontrolują nasz świat, napędzają go i upewniają się, że działa w określony sposób. Dzięki nim jesteśmy w stanie zorganizować się i osiągnąć rzeczy absolutnie niemożliwe dla pojedynczych jednostek.

Co istotne, superumysły nie muszą składać się wyłącznie z ludzi. Malone dostrzega, że ich częścią będą również sprzęty jak ekspres do kawy, ale też komputery, które mogą pełnić tu szereg ról, od narzędzia do przełożonego. Jak stwierdza Malone, komputery mają potencjał sprawić, że superumysły staną się jeszcze mądrzejsze. Chociaż przez chwilę rozważa on koncepcję maszyn mądrzejszych od ludzi i stwierdza, że taka sytuacja jest w istocie możliwa, to nie spodziewa się, by można ją było zrealizować wcześniej niż za wiele dekad.

W międzyczasie powinniśmy skupić się na założeniu, że każde użycie komputera będzie wiązać się z człowiekiem. Istniejące podejście człowieka w pętli (org. Human in the loop, tłumaczenie własne) egzemplifikuje taką relację, ale jest o tyle problematyczne, że dotyczy on zazwyczaj jednego człowieka wchodzącego w interakcję z jednym komputerem. Zdaniem Malone bardziej użyteczne byłoby myślenie o grupach złożonych z ludzi, które osiągnęły prawie wszystko to, co znamy i dodać do nich komputery. Wtedy mamy dostęp do wyspecjalizowanej inteligencji komputerów jak arytmetyka czy specyficznie rozumiane rozpoznawanie wzorców i możemy je dodać do uniwersalnej inteligencji ludzi z tej grupy.

Innym możliwym zastosowaniem komputerów już teraz jest tworzenie hiperpołączenia (org. hyperconnectivity, tłumaczenie własne). Polega ono na łączeniu

nie tylko ludzi, ale również komputerów, w skalach i sposobach, które nie były możliwe nigdy wcześniej. Malone podsumowuje, stwierdzając, że mamy tendencję do przeceniania potencjału sztucznej inteligencji, bo łatwo nam wyobrazić sobie maszyny jako mądrych ludzi. Niestety, dużo trudniej takie maszyny zbudować. Z drugiej strony nie doceniamy hiperpołączenia. Możemy je osiągnąć z wykorzystaniem już istniejących technologii, ale trudno nam sobie wyobrazić co ta masa połączonych ludzi może zrobić w tym momencie. Co dopiero kontemplować, do czego będziemy zdolni w przyszłości.

Aby zwizualizować nasze zmieniające się możliwości, Malone przytacza przykład planowania strategicznego jako procesu, który możemy śmiało zmienić z wykorzystaniem już istniejących technologii. W ramach zaproponowanego systemu zarówno ludzie, jak i maszyny stają się bardziej inteligentni i tworzą szybciej, lepsze strategie biznesowe. Dzięki niemu ludzie mogą, zamiast robić jeden plan rocznie, przygotowywać kilka planów na tydzień, czy nawet na dzień. Dzięki tej olbrzymiej ilości pomysłów możemy wybierać dużo lepsze rozwiązania. Skutek jest taki, że organizacja staje się bardziej elastyczna i łatwiej ją zaadoptować do zmieniających się warunków rynkowych, zwłaszcza potrzeb klientów.

Malone odwołuje się do opisanego wcześniej modelu procesów poznawczych. I kolejno odnotowuje w jaki sposób można wykonać powiązane z nimi aktywności w inny sposób.

I tak w kontekście tworzenia Malone zauważa, że klasycznie w proces planowania strategicznego zaangażowanych jest kilku menadżerów najwyższego szczebla organizacji i być może paru zewnętrznych konsultantów. Autor sugeruje, żeby zamiast tego zaangażować mnóstwo ludzi w tworzenie pomysłów, podobnie jak we wspomnianym „Climate CoLab”, z założeniem, że każdy pracownik może zasugerować nowe produkty czy usługi.

Alternatywnie, można też skorzystać z techniki sieci konkursów. Polega ona na tym, że uczestnicy mają dostęp do pewnej puli podstawowych pomysłów dotyczących konkretnej jednostki biznesowej, konkretnego działu. Ich zadaniem jest wybrać część lub wszystkie z nich i ułożyć realny plan biznesowy. Na kolejnym poziomie znajdują się kolejne osoby, które będą składać plany biznesowe dla całej firmy na podstawie wyników poprzedniego etapu.

W kontekście decydowania Malone proponuje sięgnąć po rynek prognostyczny, aby zamiast tworzyć piramidę, z niedookreśloną ilością poziomów, skupić się

na przewidywaniu, które produkty osiągną sukces oraz zastanowić się jakie będą wtedy warunki rynkowe.

Doświadczenie z kolei jest realizowane dzięki wykorzystaniu technologii Big Data. Powinniśmy bowiem zebrać mnóstwo danych takich, jak kim są nasi klienci, co robią, co robi nasza konkurencja i temu podobnych, które pozwolą nam podjąć dobrze poinformowane decyzje biznesowe.

W ramach odwołania do aktu pamiętania Malone snuje wizję systemu komputerowego, który może podpowiadać ludziom tak zwane ogólne strategie budowania konkurencji, jak rywalizowanie za pomocą niższych kosztów lub wyższej jakości. Przypominałby on o nich osobom decyzyjnym, aby w gąszczu bieżących problemów nie zapominali o podstawach.

Na koniec tego przykładu Malone odwołuje się do uczenia się. Wyobraża on sobie, że ludzie pracujący nad strategiami biznesowymi byłiby najpierw obserwowani przez maszyny. Te ustalałyby, co było na wejściu i na wyjściu pracy z dokumentami, a następnie analizowały zebrane w ten sposób dane pod kątem powtarzających się wzorców zachowań. Wraz z mijającym czasem i rosnącą ilością danych ci syntetyczni agenci mogliby zacząć proponować ludziom rozwiązania. Kontynuując obserwację i ustalając, które sugestie są akceptowane częściej od innych, maszyny mogłyby stworzyć kandydatów do automatyzacji procesu i, ostatecznie, przejąć część powtarzających się zadań od ludzi, pozwalając im uniknąć kolejnych powtarzalnych zadań.

Podsumowując, Malone dostrzega potrzebę i konieczność lepszego połączenia ludzi, jak również uwzględnienia w naszych planach kompetencje, które możemy oddać maszynom. Przedstawiony przykład tworzenia strategii biznesowych dość dokładnie opisuje zakładane relacje między ludźmi i maszynami. Są one płynne, dostosowujące się do bieżących potrzeb i usprawniane w ramach iteracyjnie powtarzanych procesów. Ostatecznie jednak prowadzą one do utworzenia bytów, tytułowych superumysłów, coraz doskonalszych i mądrzejszych. Można w tym dostrzec podobieństwo do poglądu wyrażonego przez Jacka Dukaję. Stwierdził on podczas jednego wywiadu, że ludzie myślą o osobliwości jako pewnej maszynie, która staje się bardziej inteligentna od ludzi. Jednakże, jak kontynuuje, nie jest to jedyny sposób, w jaki może ona zaistnieć. Rozważa on bowiem inną formę realizacji AI jaką są korporacje. Rozumie je jako duże organizacje, gdzie ludzie efektywnie nie są istotni jako jednostki, bo można wymienić praktycznie każdego, a dzięki wprowadzonym procedurom i sposobom pozyskiwania, przechowywania i udostępniania wiedzy cele i tak zostaną osiągnięte (Dukaj, 2021).

5.2. Human + Machine: Reimagining Work in the Age of AI

W tym samym roku co wspomniane w poprzednim rozdziale *Superminds*, 2018, ukazała się również książka *Human + Machine: Reimagining Work in the Age of AI*, wydana nakładem Harvard Business Review Press w Bostonie, w stanie Massachusetts.

Jej autorzy, Paul R. Daugherty oraz James H. Wilson są ściśle powiązani z korporacją Accenture. Jest ona jednym z największych dostawców usług informatycznych, zatrudnia ponad 720 tysięcy osób. Pierwszy z nich piastuje stanowisko dyrektora ds. technologii i innowacji (org. Chief Technology and Innovation Officer), rozwija działy badań i rozwoju oraz sztucznej inteligencji. James Wilson pełni rolę dyrektora zarządzającego ds. technologii informatycznych i badań biznesowych (org. Managing Director of Information Technology and Business Research).

Głównym celem publikacji jest pomoc ludziom zbliżającej się transformacji biznesu, rządu i ekonomii, jakiej dokona sztuczna inteligencja. Autorzy wyrażają nadzieję, że dzięki właściwemu zarządzaniu będziemy w stanie przejść przez tę rewolucję szczęśliwie. Po drodze zaś opracujemy innowacyjne rozwiązania, które zmienią sposób w jaki żyjemy i pracujemy, jak również wytworzą zapotrzebowania rynkowe na usługi osób, które będą umieć łączyć kompetencje ludzi i maszyn w bezprecedensowy sposób (Daugherty i Wilson, 2018, s. 215).

Daugherty oraz Wilson nie są przy tym ślepi na zamęt i ewidentnie nieprzyjemne zmiany czy wyzwania, jakie również będą zachodzić w tym czasie. Za kluczowe uznają oni zapewnienie wszystkim ludziom edukacji, treningu oraz wsparcia, abyśmy wszyscy mogli w pełni czerpać z nadchodzącej transformacji. Warto odnotować, że autorzy przeznaczyci swoje tantiemy ze sprzedaży tej książki na finansowanie edukacji i kursów oferujących przekwalifikowanie na zawody związane ze sztuczną inteligencją (Daugherty i Wilson, 2018, s. 215).

5.2.1. Rozumienie sztucznej inteligencji

Autorzy zaczynają, stawiając tezę, że właśnie jesteśmy tuż przed tym, gdy sztuczna inteligencja dokona znaczącej rewolucji w sposobie w jaki prowadzimy nasze biznesy. Pojmowanie tejże nie sprowadzają się już do automatyzacji licznych procesów, lecz na umożliwieniu ludziom i maszynom realnej współpracy. Sposób, w jaki to robią, zmienia samą naturę pracy jako takiej, wymuszając na nas zmiany w tym, jak zarządzamy operacjami i pracownikami (Daugherty i Wilson, 2018, s. 2).

Aby lepiej opisać tę współpracę, porównują zmiany w sztucznej inteligencji z rozwojem robotyki w przemyśle. Przez dekady roboty przemysłowe oznaczały olbrzymie i niebezpieczne maszyny, które trzeba było odseparowywać od ludzi. Zarówno ludzie, jak i roboty podążali według modelu taśmy produkcyjnej, gdzie wszyscy wykonywali dokładnie określone zadania. Były one zorganizowane na zasadzie sztywnego łańcucha, gdzie każde ogniwo było kolejnym zadaniem, które trzeba było wykonać w określony sposób i w określonej kolejności.

Współczesne fabryki działają diametralnie inaczej. Autorzy przytaczają tutaj koncepcję tak zwanych kobotów (od ang. collaborative robots). Maszyny te korzystają z zaawansowanych algorytmów uczenia maszynowego i wbudowanych sensorów, aby rozpoznawać obiekty w swoim otoczeniu, pojmować je, działać i uczyć się. Dzięki temu procesy wytwórcze mogą być elastyczne i samodopasowujące się, co umożliwia dynamiczne zmienianie składu zespołów i ich zadań. Członkami tychże zespołów mogą teraz być równie dobrze ludzie, jak i maszyny. Wszyscy zaangażowani w proces wykorzystują swoje inherentne predyspozycje do optymalizacji procesów produkcyjnych. Pozwala to na wykonywanie indywidualnie dostosowanych zamówień i dostosowanie produkcji do fluktuującego zapotrzebowania na rynku (Daugherty i Wilson, 2018, s. 2).

Za przykład podają oni fabrykę BMW w Dingolfing, w Niemczech. Mężczyzna podnosi obudowę przekładni i umieszcza ją w wyznaczonym miejscu. Robot, z którym współpracuje, podobnie przemieszcza przekładnię i umieszcza ją w obudowie. W innym miejscu fabryki robotyczne ramię równo rozsmarowuje grubą warstwę spoiwa na krawędziach małych, samochodowych okienek. Człowiek po chwili podchodzi, wyciera dyszę spoiwa z osadzonych na niej resztek, następnie zabiera okienko. (Daugherty i Wilson, 2018, s. 1).

Jak kontynuują autorzy, tak zaawansowana i zaadoptowana do naszego środowiska i potrzeb sztuczna inteligencja przyczyni się do dramatycznej zmiany sposobów rozumienia tego, co znaczy prowadzić biznes. Jednakże stanowi ona również pilne i coraz trudniejsze wyzwanie (Daugherty i Wilson, 2018, s. 3).

Autorzy definiują sztuczną inteligencję jako systemy, które zwiększają ludzkie możliwości w zakresie postrzegania, rozumienia, działania i uczenia się. P. Daugherty i J. Wilson stwierdzają, że niektóre firmy będą wkrótce obserwować umiarkowany wzrost wydajności, podczas gdy inne osiągną w tym samym czasie istotnie przełomowe wyniki (Daugherty i Wilson, 2018, s. 3).

Za czynnik różnicujący te sukcesy wskazują oni zrozumienie prawdziwej roli sztucznej inteligencji. Wcześniej maszyny jedynie wspomagały nas w automatyzowaniu pracy. Stanowiły narzędzia będące tylko i aż naszym przedłużeniem. Potrzebowały liniowych operacji podzielonych na konkretne sekwencje działań, które były powtarzalne i, najlepiej, mierzalne. To już jednak nie wystarczy. Przekroczyliśmy ten etap, a większość procesów udało się już zoptymalizować czy, wręcz, zautomatyzować.

Współczesne spojrzenie na procesy biznesowe jest jednak inne. Dostrzegamy potrzebę większej elastyczności i dostosowywania się do dynamicznie zmieniających się wymagań rynkowych. Tego samego wymagamy od sztucznej inteligencji, co prowadzi nas w stronę organicznych zespołów mieszających ludzi z zaawansowanymi maszynami. Takie grupy, jak wspomniano powyżej, mogą w sposób ciągły dostosowywać się do zmieniających się danych i warunków środowiskowych. Umożliwiają więc firmom w istocie wymyślenie tychże procesów biznesowych na nowo, w sposób, który pozwoli pracownikom skupić się na zadaniach wymagających ludzkiej finezji i polotu, a maszynom poddać procesy wymagające matematycznej precyzji, szybkości i jasności (Daugherty i Wilson, 2018, s. 4).

5.2.2. Trzecia fala transformacji biznesowych

Transformacja tychże procesów jest kluczowa dla P. Daugherty'ego i J. Wilsona. Powołują się oni na szeroko rozpowszechnione nieporozumienie, że systemy sztucznej inteligencji będą sukcesywnie zabierać pracę ludziom w kolejnych branżach.

Kogo to będzie najbardziej dotyczyć? Przyszłość jest, również pod tym względem, trudna do przewidzenia. Z jednej strony mamy bowiem raporty, jak ten od McKinsey & Company, które alarmują, że 800 milionów ludzi na całym świecie wkrótce straci pracę. Zwłaszcza pracowników fizycznych jak pracownicy sieci fast foodowych odpowiedzialni za przygotowywanie jedzenia czy operatorzy maszyn przemysłowych w fabrykach. Wskazywanie takich stanowisk stanowi najczęściej przytaczaną opinię w kontekście zabierania ludziom pracy przez maszyny (McKinsey & Company, 2017).

Jednakże nie jest to jedyna wizja, kto w istocie może zostać pozbawiony pracy. Jak dowodzi Michael Webb z uniwersytetu Stanforda, prace fizyczne są szczególnie trudne do zautomatyzowania. Największe wyzwania w dziedzinie robotyki dotyczą tworzenia humanoidalnych robotów, które będą w stanie poruszać się i zachowywać jak my.

Z drugiej strony, łatwo nam przychodzi automatyzowanie pracy, która opiera się na wydawaniu sądów i optymalizacji. Za przykłady podaje on fizyków, inżynierów chemii, politologów, radiologów czy sędziów administracyjnych. Doszedł do tych wniosków na podstawie analizy przeszło 16,000 patentów powiązanych z uczeniem maszynowym i 800 opisów stanowisk, aby określić stopień ekspozycji zadań danej pozycji na automatyzację (Webb, 2019).

Z całą pewnością będą przypadki, że ludzie będą po prostu tracić pracę, jednak dużo częstszym scenariuszem, w ocenie Daugherty'ego i Wilsona, będzie sytuacja, w której technologia uzupełni i wesprze ludzkie możliwości. Rutynowe, żmudne procesy biznesowe będą efektywnie oddelegowane maszynom, a ludzie będą mogli się skupić na zadaniach krytycznych, niejasnych i wymagających taktu oraz ludzkiego zrozumienia. Zjawisko to zostaje w książce nazwane trzecią falą transformacji biznesowych (Daugherty i Wilson, 2018, s. 5).

Fale biznesowe w rozumieniu autorów opisują historię przemian realizowanych procesów. Pierwsza z nich dotyczyła standaryzacji procesów. Zapoczątkował ją Henry Ford, który zdekonstruował proces konstruowania samochodów, aby można było go wykonać na taśmie produkcyjnej. Każdy krok produkcji musiał być dokładnie zmierzony, zoptymalizowany i wystandaryzowany, aby osiągnąć znaczące zyski w efektywności (Daugherty i Wilson, 2018, s. 5).

Druga fala dotyczyła automatyzacji. Zaczęła się w 1970 roku i szczyt osiągnęła w latach 90. XX wieku. Jej zaistnienie było możliwe dzięki rozwojowi technologii informacyjnych jak komputery osobiste, bazy danych i oprogramowanie wykonujące powtarzalne zadania administracyjne. Zmiany te w różnym stopniu dotknęły różne firmy, ale detaliści, jak Walmart, którzy popłynęli na tej fali, stali się rozpoznawanymi na całym świecie potęgami. Przykładowo UPS przeszedł transformację z usługi dostarczającej paczki do globalnej firmy logistycznej (Daugherty i Wilson, 2018, s. 5).

Obecna, trzecia fala skupia się na procesach adaptacyjnych. Rozwija się ona dzięki wcześniejszym osiągnięciom i, jak przewidują autorzy, będzie miała znaczny wpływ na to, jak wkrótce będziemy prowadzić interesy. Jak podają w swoich przykładach, różne firmy wynajdują obecnie na nowo swoje procesy biznesowe, aby były one elastyczniejsze, szybsze i łatwiejsze do dostosowania do zmieniających się zachowań, preferencji i potrzeb pracowników lub całych firm. Jest to możliwe dzięki oparciu się na danych dostarczanych w czasie rzeczywistym zamiast zaprojektowanym *a priori* sekwencjom kroków. Paradoks odnotowany przez autorów polega na tym, że pomimo braku rutynowości

i standaryzacji osiągnęte wyniki są wciąż wyższe. W istocie, kontynuują, największe organizacje były w stanie wprowadzić na rynek indywidualizowane produkty i usługi (jako przeciwieństwo towarów produkowanych masowo), i wciąż osiągnąć istotne zyski (Daugherty i Wilson, 2018, s. 6).

5.2.3. Brakujący Środek

Główną myślą postulowaną przez Daugherty'ego i Wilsona jest założenie, że sztuczna inteligencja nie chce nas zastąpić, a raczej wzmocnić. Nazywają ją brakującym środkiem (org. missing middle, tłumaczenie własne). Poprzez współpracę ludzi i maszyn jesteśmy w stanie osiągnąć niemożliwe inaczej zwiększenie produktywności. Przecięcie kompetencji agentów biologicznych i syntetycznych nie jest jednak szeroko znane, a tym bardziej wykorzystywane. Autorzy określają to więc mianem brakującego środka.

Główną myślą przyświecającą autorom tej koncepcji jest założenie, że obie strony (ludzie i maszyny) wykorzystują to, w czym są najlepsi. Agenci biologiczni są więc potrzebni, by rozwijać, trenować i zarządzać różnymi zastosowaniami sztucznej inteligencji. Robiąc to, umożliwiają tym systemom działanie jako współpracownicy. Z drugiej strony, agenci syntetyczni ze swojej strony wzmocniają ludzi, oferując im możliwość przetwarzania i analizowania olbrzymich ilości danych z wielu źródeł (Daugherty i Wilson, 2018, s. 8).

W brakującym środku obie strony nie rywalizują ze sobą. Zamiast tego tworzą relację symbiotyczną, oferując sobie nawzajem możliwości niemożliwe bez drugiej strony. Możemy tutaj zauważyć wysoką zbieżność z postulatami J.C.R. Licklidera, podsumowanymi w rozdziale 4.2.2.

Na koniec warto odnotować, że brakujący środek nie musi być domeną wyłącznie cyfrową, autorzy przywołują tutaj przykład Rio Tinto, konglomeratu górniczego. Wykorzystują oni flotę ciężkich maszyn zdalnie, z centralnego obiektu kontrolnego. Uwolniło to ludzi, którzy musieli nimi jeździć w niebezpiecznych warunkach i dodatkowo umożliwiło firmowym analitykom przetwarzanie informacji z sensorów maszyn, dzięki czemu możliwe jest odkrycie nowych, wartościowych wskazówek jak można zarządzać flotą w efektywniejszy i bezpieczniejszy sposób (Daugherty i Wilson, 2018, s. 9).

Ludzie wzmacniają maszyny

W ramach brakującego środka autorzy dostrzegają dwa główne obszary aktywności, jak ludzie mogą wesprzeć maszyny i na odwrót. Ta pierwsza obejmuje trzy aktywności: Trenowanie, Wyjaśnianie oraz Podtrzymywanie. Stanowią one podstawowe zadania, jakie będą wykonywane w ramach zawodów przyszłości, powstałych w wyniku naszej zmieniającej się relacji ze sztuczną inteligencją (Daugherty i Wilson, 2018, s. 115).

Trenowanie

Pierwsza z nich dotyczy adaptowania systemów sztucznej inteligencji do naszych potrzeb, poniekąd „wytrenowania ich”. Autorzy wyróżniają tutaj trzy główne scenariusze, jak taki trening może przebiegać. Albo będzie to aktywny proces, w który zaangażuje się jedna osoba, albo aktywny proces, gdzie wielu trenerów będzie szkolić pojedynczego agenta, albo, ostatecznie, wielu ludzi będzie szkolić agenta pasywnie, czyli agent będzie miał możliwość obserwowania ich podczas pracy i wyłapywać wzorce zachowań.

Autorzy dzielą przykładowe zadania jakich może dotyczyć trening na poprawę wydajności oraz człowieczeństwa. Pierwsza kategoria obejmuje takie zadania jak czyszczenie danych przed udostępnieniem ich szerzej, wskazywanie istotnych w danym kontekście punktów danych, oznaczanie danych na przyszły użytek. Trening człowieczeństwa oznacza natomiast trenowanie maszyn w zakresie języka naturalnego, gestów czy empatii. Syntetyczni agenci obserwują wtedy interakcje międzyludzkie. Zadaniem trenerów jest poprawianie błędów, wychwytywanie niezręczności i, ostatecznie, określanie i rozwijanie osobowości maszyn (Daugherty i Wilson, 2018, s. 116).

Wyjaśnianie

Celem tej grupy zadań jest budowanie zrozumienia między technologiami a biznesowymi przywódcami. Ich rola już jest bardzo istotna, zwłaszcza w sytuacjach, gdy systemy rekomendacyjne dają nieoczywiste wyniki. Daugherty i Wilson spodziewają się jej dynamicznego wzrostu wraz z rosnącą popularnością i skomplikowaniem systemów sztucznej inteligencji. Dostrzegają oni dwie główne relacje w tym zakresie: człowieka analizującego algorytm w poszukiwaniu wytworzonych przez niego heurystyk oraz osobę sprawdzającą wyniki działania maszyny pod kątem ich spójności oraz zgodności.

Dwie podstawowe aktywności w tym obszarze to testowanie i edytowanie algorytmów. Przekłada się to się na ich monitorowanie i wyjaśnianie oraz tworzenie

interfejsów zwiększających ich wyjaśnialność oraz interpretowanie wyników i komunikowanie ich interesariuszom (Daugherty i Wilson, 2018, s. 123).

Podtrzymywanie

Daugherty i Wilson widzą w tym dość standardową procedurę upewniania się, że maszyny działają prawidłowo. Osoby zaangażowane w ten proces przede wszystkim nadzorują prace syntetycznych agentów. Zapewniają niezbędną jakość danych i upewniają się, czy wyniki ich przetwarzania są w normie. Podchodzą krytycznie do oceny wydajności algorytmów. Zgłaszają błędy i błędne decyzje. Zarządzają systemami poprzez oznaczanie modeli, które działają lepiej lub gorzej i w ten sposób je „awansując” i „degradując”, na podstawie ich wyników i bieżącego zapotrzebowania (Daugherty i Wilson, 2018, s. 127).

Maszyny wzmacniają ludzi

Drugą stroną brakującego środka stanowią aktywności wykonywane przez maszyny, które wspierają ludzi. Systemy sztucznej inteligencji są w stanie już teraz wzmacniać pracowników z różnych branż i obszarów. Od medycyny po fabryki, systemy sztucznej inteligencji zmieniają sposób w jaki pracujemy. W porównaniu do naszych „naturalnych” możliwości zaczynamy dysponować, jak określają to P. Daugherty i J. Wilson „supermocami” (2018, s. 139).

Wzmacnianie

Pierwsza aktywność wymieniona przez autorów dotyczy wsparcia podejmowania decyzji danymi zbieranymi i analizowanymi przez systemy sztucznej inteligencji. Autorzy przewidują w tym kontekście cztery możliwe scenariusze: pojedynczy agent wspiera jednego człowieka, pojedynczy agent wspiera grupę ludzi, kilku agentów wspiera człowieka, wielu agentów wspiera jednocześnie i równoległe wielu ludzi.

Zadania jakie realizują, dzielą się na trzy rodzaje: dopasowywanie, doradzanie oraz praca nad trendami. Dopasowywanie polega efektywnie na odpowiadaniu na pytania, gdyż dotyczy lokowania zasobów na podstawie zapytań człowieka, ale również automatyzacji powtarzalnych i nisko-poziomowych zadań. Przykładem może być rozlokowanie dostępnych pracowników do projektów na podstawie posiadanych przez nich kompetencji. Doradzanie, czy też przedstawianie rekomendacji, pomaga nam w podejmowaniu decyzji poprzez ustalanie rankingów dostępnych alternatyw, co pozwala na priorytetyzowanie zasobów oraz automatyzację zmian w procesach. Praca nad trendami dotyczy automatycznego przetwarzania danych przez maszyny w poszukiwaniu trendów

w czasie rzeczywistym, dopasowywanie propozycji do potrzeb konkretnych osób, identyfikowanie anomalii, kategoryzowanie danych. Ostatecznie wszystkie te aktywności związane z trendami prowadzą do wzmocnienia naszych kompetencji w zakresie podejmowania strategicznych decyzji (Daugherty i Wilson, 2018, s. 142).

Interakcje

Agenci interaktywni to wszelkiej postaci asystenci, którzy wspierają nas w naszym życiu. Mogą to być chatboty, do których już przywykliśmy, nieco nowsze technologie pokroju Alexy od Amazona czy Siri od Apple. Agenci w tych relacjach zazwyczaj kontaktują się z pojedynczą osobą, ale autorzy dostrzegają też możliwość agenta, który towarzyszy wielu osobom jednocześnie.

Przykładowe aktywności zakładają głównie odpowiadania na pytania użytkowników, na przykład podając prognozę pogody albo stan konta w banku. Mogą też być bardziej zaawansowane i dotyczyć dostarczania danych z bazy, w ten sposób odciążając pracowników działu IT. Dużą rolę w kontekście tych agentów gra interfejs, gdyż częstym rozwiązaniem staje się dostęp do usług i analityki na podstawie polecenia głosowe. Możemy wypowiedzieć pytanie w języku naturalnym i dostać wynik na ekranie komputera. Co więcej, Daugherty i Wilson widzą w tych agentach potencjał niejako doradcy, mogą występować w charakterze coacha. W tej roli będą odpowiadać na pytania o sugestie kolejnych kroków na podstawie zaimplementowanej wiedzy domenowej i polityki korporacyjnej (Daugherty i Wilson, 2018, s. 145).

Ucieleśnianie

Poprzednie dwie kategorie, jak zauważają autorzy, dotyczyły raczej procesów mentalnych, ale nie są to jedyna możliwość, jak sztuczna inteligencja jest w stanie nas wesprzeć. Daugherty i Wilson przywołują w tym kontekście coboty, czyli roboty kooperacyjne, od nazwy angielskiej „collaborative robots”. Pracują one obok ludzi, płynnie zmieniając wykonywane zadania. Jakkolwiek stanowią najszerzej omówiony przykład, autorzy przywołują też drony używane w Rwandzie i Nowej Gwinei do transportowania próbek medycznych w obszarach pozbawionych infrastruktury drogowej.

Praca z agentami z tej kategorii przebiega na zasadzie partnerstwa, w ramach którego jeden człowiek pracuje z jedną maszyną. Roboty te mają zdolność autonomicznego nawigowania po powierzchniach współdzielonych z ludźmi. Mogą też stanowić przedłużenie naszych zmysłów w obszarach, do których ludziom ciężko

się dostać, chociażby poprzez udostępnienie obrazu z kamery zamontowanej na latającym dronie. Ostatecznie ich rolą jest wykonywanie bardzo dokładnych zadań o żmudnym, rutynowym charakterze w przestrzeni fizycznej (Daugherty i Wilson, 2018, s. 147).

5.2.4. Przykłady zastosowań sztucznej inteligencji

Autorzy opisują bieżącą sytuację poprzez opis kolejnych branż i przedstawieniu, jak największe firmy w niej działającej przygotowują się do zmian wynikających z rozwoju technologii powiązanych ze sztuczną inteligencją.

Sztuczna inteligencja w motoryzacji

Najpierw opisane zostaje środowisko przemysłowe, zwłaszcza w kontekście motoryzacji. Analizie poddane zostają BMW, Mercedes-Benz czy General Electric. Na szczególną uwagę zasługuje opis tak zwanego cyfrowego bliźniaka, czyli komputerowo generowanej reprezentacji istniejących miejsc, fabryk i produktów. Dzięki nim firmy mogą znacząco zredukować koszty projektowania, opracowywania prototypów i dostosowywania urządzeń, gdyż operacje te najpierw wykonywane są w symulatorach, przed modyfikowaniem ich fizycznych wersji (Daugherty i Wilson, 2018, s. 10).

Zadania administracyjne wsparte przez sztuczną inteligencję

Drugi obszar, który opisują autorzy, skupia się na operacjach wykonywanych przez zaplecze biurowe firm zajmujące się administracją i procesowaniem dokumentów. Daugherty i Wilson opisują tutaj, jak rozwiązania sztucznej inteligencji mogą wspierać ludzi w filtrowaniu i analizowaniu strumieni informacji z różnych źródeł i umożliwić automatyzację żmudnych i rutynowych zadań, aby ostatecznie wzmacniać ludzkie umiejętności i ekspertyzę. Autorzy podają za przykład kanadyjską firmę oferującą usługi ubezpieczeniowe i finansowe, która zaadoptowała do swoich potrzeb system sztucznej inteligencji. Przetwarza on nieustrukturyzowane dane finansowe z wiadomości prasowych, raportów i wiadomości elektronicznych (np. e-mail), aby przedstawiać rekomendacje ogólnofirmowe, ale może też zostać dostosowany do przygotowywania opracowań według potrzeb konkretnych analityków (Daugherty i Wilson, 2018, s. 10).

Sztuczna inteligencja w działach badań i rozwoju

Następnie zostaje przedstawione, jak różne firmy usprawniają działy badań i rozwoju. Autorzy dowodzą, że na każdym kroku procesu odkrywania wiedzy- obserwacji, generowania hipotez, projektowania eksperymentów oraz analizy rezultatów, można skutecznie zaaplikować technologie sztucznej inteligencji.

Umożliwiają one zwiększenie efektywności działań oraz wyraźnie lepsze wyniki. Przykładem firmy, która tak zmieniła swoje procesy jest GNS Healthcare. Firma ta opracowała skomplikowane oprogramowanie, oparte na uczeniu maszynowym. Poszukuje ono wzorców w danych medycznych pacjentów i automatycznie generuje hipotezy badawcze. Użyteczność systemu sprawdzono poprzez odtworzenie procesu uprzednio przeprowadzonego projektu, w ramach którego badano negatywne reakcje pomiędzy kombinacjami leków u seniorów zarejestrowanych w systemie ubezpieczeń zdrowotnych Medicare. System był w stanie odtworzyć wyniki projektu blisko ośmiokrotnie szybciej. Początkowe badanie zajęło dwa lata, praca z SI skróciła ten czas do ledwie trzech miesięcy (Daugherty i Wilson, 2018, s. 10).

Marketing i sprzedaż

Ostatni rozdział pierwszej części skupia się na otoczeniu marketingu i sprzedaży. Przywołane zostają wykorzystania technik uczenia maszynowego w „elektronicznych towarzyszach” jak Alexa od Amazona, Siri od Apple czy Cortana od Microsoftu. Autorzy sugerują, że stanowią one cyfrowe uosobienie tych marek. Innymi słowy, mówią, że sztuczna inteligencja stała się marką.

Bardziej adekwatnym przykładem zastosowania sztucznej inteligencji w marketingu zdaje się przytoczony przez Aleksandrę Przegalińską przykład opisany przez Tomasza Pawlikowskiego. Opisał on stworzonego na potrzeby firmy Publicis asystenta pracy w branży reklamowej, Marcela. Został on tak nazwany na cześć założyciela tejże firmy Marcela Bleustaina-Blancheta. System oferuje przegląd umiejętności pracowników dostępnych w firmie i stanowi kompendium zrealizowanych kampanii reklamowych (Przegalińska, 2020, s. 184).

Efektywnie zmienił on sposób pracy zaangażowanych osób. Dawniej, gdy firma dostawała zlecenie, trafiało ono do lokalnego ośrodka, który natychmiast kompletował swój lokalny zespół i realizował to zlecenie, najlepiej jak potrafił. Proces był więc ograniczony przez doświadczenie osoby, która prowadziła określonym zleceniem, dostępnością i ilością lokalnych pracowników i ich wiedzą (Przegalińska, 2020, s. 188).

Teraz gdy firma dostaje zlecenie, zostaje ono najpierw dostosowane do potrzeb analizy Marcela. Ten następnie porównuje je z historią zrealizowanych kampanii reklamowych, analizuje wymagane umiejętności i przedstawia sugestię zespołu złożonego z aktualnie dostępnych pracowników, których doświadczenie, cele rozwojowe

i umiejętności dadzą największe szanse na jak najlepszy projekt. Zespół oczywiście dostaje też listę dawnych zleceń wraz z oceną ich skuteczności (Przegalińska, 2020, s. 188).

Jak zauważa Pawlikowski, implementacja Marcela była trudna tak technicznie, jak i kulturowo, bo pracownicy początkowo wcale nie byli zachwyceni tą opcją, Jednakże wprowadzenie go zdecydowanie obniżyło rywalizację między krajowymi ośrodkami, zwiększyło demokratyczność organizacji, jakość realizowanych kampanii i, paradoksalnie, przywiązanie pracowników do firmy (Przegalińska, 2020, s. 189).

5.2.5. Wypełnienie Brakującego środka

Druga część książki skupia się na przedstawieniu kluczowej myśli autorów, wspomnianego wcześniej „brakującego środka”. Dostarcza ona wysokopoziomowy przewodnik jak zmieniać organizacje i na nowo wymyślać rozumienie pracy jako takiej. Kluczem, by skutecznie wykorzystać sztuczną inteligencję, będzie według nich dostrzeżenie tego środka i wyeliminowanie go za pomocą nowych zawodów, nowych typów relacji oraz zmianie postrzegania kadry zarządzającej (Daugherty i Wilson, 2018, s. 11).

Nowe stanowiska

Na początku tej części książki autorzy opisują nowe zawody jakie powstaną dzięki efektywnemu wykorzystaniu uczenia maszynowego do przeorganizowania procesów biznesowych. Dokładniej rzecz ujmując, pojawi się zapotrzebowanie na pracowników, którzy będą w stanie projektować i trenować algorytmy, jak również wyjaśnić ich metody działania. Jednakże nie będzie to rola ściśle programistyczna. Autorzy nazywają ją menadżerem zasobów maszynowych, co stanowi odniesienie do współczesnych menadżerów zasobów ludzkich. Do ich obowiązków należeć będzie nadzorowanie systemów sztucznej inteligencji i regularne przeprowadzanie oceny ich efektywności. Systemy lepiej działające będą „awansowane”, jak autorzy określają ich replikowanie do innych działów organizacji. Systemy gorsze będą natomiast poddawane analizie i planom naprawczym, a jeśli i to nie pomoże, zostaną ostatecznie wyeliminowane (Daugherty i Wilson, 2018, s. 11).

Jak maszyny mogą wesprzeć ludzi

Następnie, autorzy opisują, jak ludzie osiągają olbrzymie wzrosty produktywności poprzez pracę z systemami sztucznej inteligencji. O ile w poprzednim rozdziale to ludzie pomagali maszynom usprawnić ich pracę, tak w tej części skupiają się na tym, jak może się ta sytuacja odwrócić. Wyodrębniają oni trzy techniki: wzmocnienie

(org. amplify, tłumaczenie własne), interakcje (org. interact, tłumaczenie własne) i ucieleśnienie (org. embody, tłumaczenie własne). Te nowe typy relacji pomagają ludziom pozbyć się żmudnych zadań i umożliwiają im szybszą i dokładniejszą pracę poprzez korzystanie z przewodnictwa, rad i wsparcia oferowanego przez systemy sztucznej inteligencji (Daugherty i Wilson, 2018, s. 12).

Zarządzanie sztuczną inteligencją

Kolejny poruszony temat poświęcony jest wyzwaniom przed jakimi stanie kadra zarządzająca w firmach. Autorzy wymieniają tutaj szereg wyzwań, z których największymi prawdopodobnie będzie stworzenie warunków umożliwiających pracownikom eksperymentowanie metodą prób i błędów oraz zapewnienie odpowiedniego zaplecza informatycznego, które dostarczy dane o jakości umożliwiającej maszynom prowadzenie bardziej skomplikowanych wnioskowań (Daugherty i Wilson, 2018, s. 12).

Umiejętności połączone

Ostatecznie, autorzy zastanawiają się nad przyszłością pracy jako takiej. Ich rozważania zostały podsumowane w ośmiu nowych „połączonych umiejętnościach” (org. fusion skills, tłumaczenie własne):

- Pierwsza z umiejętności to „inteligentne odpytywanie” (org. intelligent interrogation, tłumaczenie własne). Polega ona na wydobywaniu od agentów syntetycznych wyników ich działania za pomocą odpowiednio zadanych pytań na różnych poziomach abstrakcji.
- Druga, „wzmocnienie oparte na botach” (org. bot-based empowerment, tłumaczenie własne), określa kompetencje w zakresie współpracy z botami, aby przekroczyć własne ograniczenia.
- „Praktyka wzajemności” (org. reciprocal apprenticing, tłumaczenie własne) oznacza rozwój umiejętności tak syntetycznych, jak i biologicznych agentów. Sztuczna inteligencja byłaby szkolona przez ludzi, a ludzie przechodziliby własne szkolenia, aby nauczyć się jak efektywniej usprawniać wykonywane przez siebie procesy z wykorzystaniem sztucznej inteligencji.
- Kolejna kompetencja to „holistyczne naprawianie” (org. holistic melding, tłumaczenie własne). Polega ona na rozwoju modeli mentalnych w agentach sztucznej inteligencji, dzięki czemu możliwe będzie usprawnienie ich współpracy z ludźmi.

- „Rehumanizacja czasu” (org. rehumanizing time, tłumaczenie własne) zakłada, że będziemy potrzebowali nauczyć się wyobrażać sobie na nowo procesy biznesowe, aby uwolnić czas na zadania możliwe do wykonania wyłącznie przez ludzi.
- „Odpowiedzialna normalizacja” (org. responsible normalizing, tłumaczenie własne) przekłada się na kształtowanie celów i postrzegania współpracy ludzi i maszyn. Ma ona zastosowanie do indywidualnych osób, ale też biznesów czy społeczności.
- „Integracja osądów” (org. judgment integration, tłumaczenie własne) sprowadza się do wybierania działań w momentach, gdy maszyny nie będą w stanie podjąć decyzji.
- „Nieustanne wymyślanie” (org. relentless reimagining, tłumaczenie własne) dotyczy myślenia na jakie nowe sposoby można przeorganizować pracę, procesy i modele biznesowe, aby osiągnąć wykładnicze wzrosty optymalizacji.

5.2.6. Pięć koniecznych warunków dla odkrycia brakującego środka

W swoim badaniu Daugherty i Wilson odkryli, że 9% firm (z przebadanych 1500) już płynnie na trzeciej fali przemian biznesowych. Wykorzystały one do cna swój potencjał w zakresie automatyzacji i rozwijają teraz procesy oparte na współpracy ludzi i maszyn. Dostrzegają oni, że wymyślanie procesów biznesowych na nowo jest procesem ciągłym i opartym na danych dostarczanych w sposób ciągły (Daugherty i Wilson, 2018, s. 13).

W jaki sposób te przedsiębiorstwa to osiągnęły? Według autorów było to możliwe dzięki przyjęciu odpowiedniego sposobu myślenia, zaakceptowaniu eksperymentowania, wyedukowaniu swoich przywódców, zapewnieniu jakościowych danych i pracy nad umiejętnościami swoich pracowników (Daugherty i Wilson, 2018, s. 13).

Sposób myślenia

Autorzy rozumieją przez to przyjęcie diametralnie innego podejścia do biznesu i wyobrażanie sobie pracy przez pryzmat brakującego środka. Wcześniejszym celem było wykorzystanie maszyn do zautomatyzowania konkretnych kroków w określonym procesie. Teraz potencjalna współpraca między agentami biologicznymi i syntetycznymi umożliwia wymyślenie wielu tradycyjnych procesów na nowo. Kluczowym może się tu okazać zbliżenie biznesu do potrzeb konsumentów- umożliwienie szybszej i lepiej dopasowanej reakcji do określonych potrzeb. Aby było to jednak możliwe, konieczne jest przekonanie kadry zarządzającej do myślenia o biznesie w ten sposób, jak również,

konieczne jest zbudowanie solidnych fundamentów, zanim będzie to możliwe. Pierwszym krokiem powinno być zautomatyzowanie rutynowej pracy (Daugherty i Wilson, 2018, s. 14).

Eksperymentowanie

Czasy standaryzowanych procesów biznesowych, w opinii autorów, już mijają. Firmy nie będą już w stanie kopiować rozwiązań liderów w swoich branżach. Stąd też koniecznym stanie się eksperymentowanie. Kadra zarządzająca musi ustawicznie sprawdzać jakie procesy zadziałają najlepiej w osobno rozpatrywanych, unikatowych okolicznościach. Duża część wysiłku odnalezienia proporcji i kompozycji zespołów ludzi i maszyn będzie oparta na metodzie prób i błędów.

Przywództwo

Liderzy muszą zawsze rozważać etyczne, moralne i prawne implikacje wykorzystania sztucznej inteligencji które tworzą. Systemy te muszą też być wyjaśnialne (org. explainable), muszą eliminować uprzedzenia i promować odpowiedzialność w stosowanych algorytmach. Przedsiębiorstwa powinny również pilnować, aby ich eksperci od sztucznej inteligencji nie tracili poczucia sprawczości, ani świadomości tego, jak ich działania wspierają procesy decyzyjne innych osób. Dodatkowo firmy muszą zapewnić pracownikom szkolenia, dzięki którym nabiorą kompetencji umożliwiających przyjęcie ról wypełniających brakujący środek (Daugherty i Wilson, 2018, s. 15).

Dane

Agenci syntetyczni wymagają rozległych danych, tak w sensie ich ilości, jak również różnorodności. Dotyczy to również „danych wylotowych” (org. exhaust data, tłumaczenie własne”), jak autorzy określają dane tworzone jako produkt uboczny innego procesu (za przykład mogą posłużyć ciasteczka generowane podczas przeglądania internetu). Zbieranie i organizowanie takich informacji będzie stanowiło jedno z kluczowych wyzwań podczas bieżącej fali transformacji biznesowej. Co również istotne, powinniśmy odejść od tak zwanych silosów danych, czyli struktur ich przechowywania polegających na separowaniu i ograniczaniu. Autorzy postulują, że dane powinny poruszać się swobodnie pomiędzy różnymi działami firm. Dopiero po spełnieniu tych wymogów możliwe będzie wprowadzanie podejścia opartego na brakującym środku (Daugherty i Wilson, 2018, s. 15).

Umiejętności

Autorzy ponownie odwołują się do koncepcji połączonych umiejętności, wspomnianych wcześniej. Ich aktywny rozwój stanowi kolejny warunek kluczowy dla wypełnienia brakującego środka. Daugherty i Wilson dodatkowo odnotowują zmieniającą się relację między ludźmi a maszynami. W trakcie drugiej fali maszyny były z reguły używane do zastąpienia ludzi- automatyzacja dotknęła wielu zawodów i pozbawiła pracy przedstawicieli takich zawodów jak pracownicy fabryk, asystenci administracyjni, księgowi czy kasjerzy bankowi. W postulowanej trzeciej fali ludzie będą jednakże potrzebni bardziej niż podczas wcześniejszych fal. Adaptacyjne procesy opierają się umieszczeniu w przytoczonym wcześniej podejściu człowieka w pętli (org. Human in the loop, tłumaczenie własne). Dzięki niemu ludzie będą w stanie nie tylko tworzyć i rozwijać produkty wraz z systemami sztucznej inteligencji, ale również współpracować z nią, by zapełnić brakujący środek (Daugherty i Wilson, 2018, s. 16).

5.3. Porównanie koncepcji i analityki danych

Podobnie jak w części II porównanie koncepcji nie jest łatwe, gdyż operują one na różnych poziomach. Superumysły są uniwersalne, mogą odpowiadać i za opis podejmowania indywidualnej decyzji podczas zakupów w sklepie spożywczym, jak i stanowić model działania całego przedsiębiorstwa, jeśli nie kraju. Z drugiej strony koncepcja brakującego środka skupia się na usprawnianiu działania określonych zespołów.

Na potrzeby analizy możemy więc przyjrzeć się w jaki sposób każda z idei potraktuje ten sam proces. Doskonałym przykładem jest przeanalizowanie w jaki sposób analitycy danych wykorzystują możliwości sztucznej inteligencji do wsparcia tradycyjnie realizowanych przez nich procesów analizy danych. Praktycy korzystają w tym kontekście z określenia wzmocnionej analityki. W artykułach autorstwa Nooraha Alghamdiego i Heyama Al-Baity'ego oraz Nicolasa Prata zostały opisane współcześnie dostępne rozwiązania techniczne (Alghamdi i Al-Baity, 2022; Prat, 2019). Jednakże, co warto podkreślić, sposób rozumienia tego terminu jest przez nich bardzo ukierunkowany na realizowane przez nich projekty i codzienne zadania.

5.3.1. Wzmocniona analityka

Wzmacnianie inteligencji stanowi już integralną część wielu procesów biznesowych. Obszarem, w którym widać te praktyki wyraźnie jest domena business

intelligence (BI) oraz analityki danych. N. Alghamdi i H. Al-Baity stwierdzają, że przez wzgląd na wzrost wolumenów danych manualne podejście do analizy jest już wręcz niewystarczające (Alghamdi i Al-Baity, 2022, s. 1)

BI i analityka biznesowa stanowią zbiór technologii, technik, systemów, praktyk, metodologii i aplikacji, których nadrzędnym celem jest krytyczna analiza danych biznesowych. Ma to umożliwić przedsiębiorstwom lepsze zrozumienie swoich procesów i warunków rynkowych, dzięki czemu mogą one podejmować skuteczne decyzje biznesowe (Chen, Chiang i Storey, 2012).

Jak kontynuuje N. Prat, terminy te często stosowane są łącznie, lub wręcz wymiennie. Aczkolwiek, jak mówimy o BI, to zazwyczaj odnosimy się do raportowania, procesowania analitycznego online (OnLine Analytical Processing, tłumaczenie własne), oraz ekranów czy pulpityw menadżerskich, podczas gdy analityka odnosi się do zaawansowanych technik opartych na uczeniu maszynowym (Prat, 2019, s. 375).

Nowe, zaproponowane przez instytut Gartnera, pojęcie wzmocnionej analityki jeszcze bardziej zaciera różnice między tymi terminami. Zakłada się w nim dostarczanie nowych możliwości użytkownikom produktów BI. Najczęściej stanowią one rozwiązania oparte na zaawansowanych metodach uczenia maszynowego i sztucznej inteligencji (Gartner, 2017a).

Wzmocnienie analityki wzbogaca tak zwany cykl analityki o automatyzację możliwe dzięki zastosowaniu szerokiej gamy technologii. N. Alghamdi i H. Al-Baity wymieniają w tym kontekście uczenie maszynowe, sztuczną inteligencję, jak również przetwarzanie i generowanie języka naturalnego (Alghamdi i Al-Baity, 2022, s. 1). Nazywane przez niektórych autorów bywa również analityką zasilaną sztuczną inteligencją (org. AI powered analytics, tłumaczenie własne) stanowi zdecydowanie kamień milowy w historii BI (Gartner, 2017a; Watson, 2017; Henschen, 2018).

5.3.2. Trzy pokolenia BI

N. Prat wyróżnia trzy pokolenia BI. Pierwsze z nich było w latach 90. XX wieku, za sprawą opracowania koncepcji, wspomnianego wyżej OLAP powiązanych z hurtowni danych (org. data warehouses, tłumaczenie własne), jak również systemów zarządzania informacjami, systemów informowania osób w rolach zarządczych oraz wsparcia procesu decyzyjnego. Analityka w tym okresie była oparta na czystych technikach statystycznych. Raportowanie miało jedynie proste funkcjonalności i wymagało zaangażowania

pracowników działu IT, aby wyszukali dla użytkownika biznesowego dane w systemie (Alghamdi i Al-Baity, 2022, s. 3).

Drugie pokolenie dotyczyło wzrostu znaczenia analityki z zastosowaniem narzędzi informatycznych około 2005 roku i rewolucji big data oraz narzędzi do samoobsługowego podejścia do tworzenia raportów biznesowych, której początek N. Prat wyznacza na rok 2010 (Alpar i Schulz, 2016).

W kontekście zwiększenia zainteresowania użytkowników końcowych rozwiązaniami samoobsługowymi N. Alghamdi i H. Al.-Baity wymieniają zwiększenie mobilności poprzez umożliwienie użytkownikom korzystania z raportów na urządzeniach przenośnych, jak również wizualizacji danych i kart biznesowych. Podobnie, umożliwienie im eksploracji danych miało za zadanie zwiększyć ich zaangażowanie oraz elastyczność raportów końcowych, jak również odciążyło pracowników działów IT, którzy mogli stworzyć pojedynczy raport, który końcowi użytkownicy mogli dopasować pod swoje potrzeby wybierając z, udostępnionych im na tę okoliczność, odpowiednich filtrów czy parametrów (Alghamdi i Al-Baity, 2022, s. 3).

Trzecie pokolenie zaczęło się według N. Prata w 2010 roku i dotyczy omawianej tutaj analityki wspartej sztuczną inteligencją. W tym czasie ponownie użytkownicy końcowi dostają więcej możliwości, gdyż techniki zaawansowanej analityki stają się dostępne dla każdego. Prowadzi to do nowej koncepcji, tak zwanego, citizen data scientist (brak tłumaczenia, gdyż brak jest nawet powszechnie akceptowanego tłumaczenia dla nazwy stanowiska data scientist). Jest to osoba, która tworzy lub generuje modele matematyczne dla celów analitycznych, predykcyjnych i normatywnych, a która w swojej głównej roli porusza się poza obszarami statystyki i analityki (Gartner, 2017b). Sami data scientists też zostają wspomóczeni, gdyż ich najbardziej żmudne zadania będą sukcesywnie automatyzowane i oddelegowywane maszynom. Will Knight dostrzega w obu tych zjawiskach szansę na załatanie luki w zapotrzebowaniu na osoby z kompetencjami w zakresie data science (Knight, 2017). Lakshman Bulusu i Rosendo Abellera z kolei piszą, że cały proces decyzyjny może zostać zautomatyzowany dzięki współczesnym technologiom informatycznym (Bulusu i Abellea, 2020, s. 3-7).

Badania nad percepcją szans i ograniczeń wzmocnionej analityki wykonane na rumuńskich firmach pozwoliło wykazać, że stanowią one kluczowy element współczesnego biznesu. Badani stwierdzili, że raportowanie jest dzięki nim szybsze i dokładniejsze, oszczędza czas i poprawia zdolności strategicznego planowania, jak również podejmowanie taktycznych decyzji, proces decyzyjny ogólnie był

sprawniejszy i to wszystko przełożyło się na wyższe zyski. Mimo że określili go jako trudny, menadżerowie uznali krok wprowadzenia wzmocnionej analityki do swoich firm za kluczowy. Za największe wyzwania wskazali efektywne komunikowanie potrzeby zmian, wytworzenie organizacyjnej kultury zmian w ogóle, jak również dostosowanie infrastruktury do umożliwienia transformacji procesu decyzyjnego na bardziej nastawiony na podejmowanie decyzji na podstawie danych (Grigorescu, Baiasu i Chitescu, 2020, s. 651-657).

5.3.3. Cykl działań analitycznych

N. Prat przedstawia własny model opisujący cykl działań analitycznych powstały na bazie prac szeregu autorów (Erl, Khattak i Buhler, 2015; SAS, 2016; Storey i Song, 2017; Seddon, Constantinidis, Tamm i Dod, 2017).

Wyróżnił on 7 kroków:

1. Określenie problemu biznesowego i zidentyfikowanie okazji.
2. Przygotowanie danych.
 - a. Profilowanie danych.
 - b. Transformacja danych.
3. Analiza danych.
 - a. Analiza eksploracyjna.
 - b. Modelowanie danych.
4. Zastosowanie modelu.
5. Decyzja.
6. Akcja.
7. Obserwacja.

Przed wszystkim konieczne jest zdefiniowanie problemu, który analityka ma zaadresować. Następnie przygotowujemy dane do właściwej analizy. Autor odróżnia tutaj badanie jakości posiadanych danych od transformacji ich w formę gotową do właściwej analityki. Kolejnym krokiem jest właściwa analiza, która znów zostaje rozbita na dwie fazy: poznawanie danych i poszukiwanie obiecujących wzorców podczas analizy eksploracyjnej oraz budowa właściwych modeli matematycznych wraz z oceną, które modele najlepiej opisują rzeczywistość zapisaną w danych. Po wybraniu najlepszego kandydata należy go zastosować w środowiskach informatycznych dostępnych dla użytkowników końcowych, aby ci, w kolejnych krokach, mogli kolejno podejmować dobrze poinformowane decyzje i działać na ich podstawie. Nie jest to jednak koniec,

bo musimy stale monitorować dopasowanie modelu do danych, jakość podejmowanych akcji i poszukiwać kolejnych okazji biznesowych, aby zacząć cykl na nowo (Prat, 2019, s. 376).

5.3.4. Usprawnienie cyklu analizy za pomocą sztucznej inteligencji

Każdy z kroków opisanych w poprzednim rozdziale może zostać czy to zautomatyzowany, czy też usprawniony za pomocą różnych rozwiązań sztucznej inteligencji. N. Alghamdi i H. Al.-Baity nazywają to wirtualnym data scientistem, który wykonuje aktywności powiązane z analizą i następnie zoperacjonalizuje je, czyli doprowadzi do stanu końcowego (Abas i inni, 2020).

Najtrudniejsze jest to w kroku pierwszym, gdyż wciąż myślimy o ludziach jako nadających kierunek działań i identyfikujących problemy, które powinniśmy rozwiązać. Jednakże, jak zauważa N. Prat, jesteśmy już w stanie generować automatycznie wnioski z danych, które mogą stanowić bezpośrednią inspirację dla osób odpowiedzialnych. Za przykład możemy tutaj przywołać funkcjonalność „Narratives” od Tableau, która automatycznie generuje wnioski w języku naturalnym z używanych wizualizacji danych, podsumowując najważniejsze informacje, typu trendy czy najwyższe wartości ułatwiając ich wychwycenie mniej technicznie usposobionym użytkownikom (Prat, 2019, s. 377).

Przygotowywanie danych jest najbardziej czasochłonnym krokiem ze wszystkich wymienionych, według niektórych szacunków zajmuje on nawet 80% całego procesu (Lohr, 2014). Wszelkie automatyzacje związanych z tym obszarem aktywności będą więc skalować się do największych oszczędności czasu i znacząco zwiększą produktywność zespołów analitycznych. Dzięki takim narzędziom jak Trifacta Wrangler możliwa jest częściowa automatyzacja profilowania danych poprzez określenie w danych wartości pustych, niepasujących czy odstających. Transformacje również zyskują dzięki sugestiom dotyczącym czyszczenia danych (jak potraktować brakujące wartości czy też jak standaryzować różne pola), zmian w agregacjach (dzielenie, agregowanie i przestawianie kolumn), mieszanie i wzbogacanie danych poprzez propozycje połączenia z innymi, znanymi źródłami danych (Prat, 2019, s. 377).

Na etapie odkrywania danych, podczas analizy eksploracyjnej narzędzia takie jak Tableau automatycznie podsuwają sugestie typów wykresów (słupkowe, liniowe, mapy, ...) na podstawie wybranych pól z danymi. Umożliwiają one również łatwe w użyciu wzmocnienia analityki za pomocą takich środków jak klastrowanie czy przewidywanie

trendów. Watson Analytics idzie krok dalej i sam automatycznie analizuje dane, po czym przedstawia sugestie wizualizacji pogrupowane według dopasowania. Oba narzędzia umożliwiają również interakcje w języku naturalnym (na przykład „Który produkt sprzedał się najlepiej w Polsce, w 2012 roku?”), aczkolwiek z ograniczoną syntaktyką i, jeszcze, nie w formie naturalnego dialogu (Prat, 2019, s. 377).

Wsparcie sztucznej inteligencji w budowaniu modeli i fazie ewaluacji może być realizowane przez narzędzia jak „Driverless AI”, które automatyzują proces feature engineering (brak powszechnie akceptowanego tłumaczenia), czyli przygotowania sugestii pól wartych do użycia w budowie modeli uczenia maszynowego. Podobnie użyteczne są turnieje modeli (SAS, 2016), które stosują techniki uczenia maszynowego do usprawnienia aplikowania uczenia maszynowego (Knight, 2017). Przykładowo system DataRobot jest w stanie przetworzyć miliony kombinacji pól przez swoje algorytmy, by ostatecznie przedstawić listę przetestowanych modeli wraz z wybranymi parametrami określającymi ich skuteczność (Prat, 2019, s. 377).

Przeniesienie modelu do środowiska produkcyjnego, czyli dostępnego dla użytkowników końcowych, często jest problematyczne przez wzgląd na przekazanie go od zespołów zajmujących się danymi do klasycznie pojmowanego IT (SAS, 2016). Tu również można wprowadzić automatyzację poprzez zastosowanie takich narzędzi jak Alteryx Promote. Zastosowanie tej metody naturalnie kaskaduje do fazy decyzji i pozwala automatyzować te, które uznamy za rutynowe i ściśle operacyjne. Podobnie, ma to znaczenie dla fazy obserwacji, gdyż pozwala określić jak radzi sobie model i informowania odpowiedzialnych osób, gdyby określone parametry spadały poniżej określonych wartości (Kobielus, 2017).

5.3.5. Podsumowanie

W tym rozdziale następuje podsumowanie w formie odniesienia przedstawionego procesu analitycznego do dwóch, głównych koncepcji opisanych w tej części rozprawy. Przypominając, Malone przedstawił propozycję schematu decyzyjnego, który opierał się na 6 elementach: Tworzeniu, Decydowaniu, Doświadczeniu, Pamiętaniu, Działaniu i Uczeniu się. Z kolei Daugherty i Wilson skupili się na określeniu kluczowych aktywności jak ludzie i maszyny mogą się nawzajem wspierać i określili je mianem brakującego środka. Były to kolejno: trenowanie, wyjaśnianie, podtrzymywanie w kontekście ludzi wspierających maszyny oraz wzmacnianie, interakcje oraz ucieleśnianie w odwróconej sytuacji, gdy to maszyny wspierają ludzi.

7 kroków procesu analitycznego z perspektywy superumysłu i brakującego środka

Dla przypomnienia, N. Prat wyróżnił 7 kroków w procesie analitycznym: 1) Określenie problemu biznesowego i zidentyfikowanie okazji, 2) Przygotowanie danych (w tym: a. Profilowanie danych i b. Transformacja danych), 3) Analiza danych. (w tym: a. Analiza eksploracyjna i b. Modelowanie danych), 4. Zastosowanie modelu, 5. Decyzja, 6. Akcja i 7. Obserwacja.

Określenie problemu biznesowego

Już na pierwszy rzut oka możemy dostrzec silne podobieństwa między tymi krokami a podejściem T. Malone'a. Pratowskie określenie problemu biznesowego to wyraźny przykład połączenia Doświadczenia oraz Tworzenia. Doświadczenie, rozumiane jako zbieranie danych, przyglądanie się trendom i wynajdywanie w nich okazji na generowanie zysków lub ograniczanie kosztów. Tradycyjnie, jak zostało to wspomniane w poprzednim rozdziale 5.3.4., odpowiedzialność ta spoczywa na jednostkach. T. Malone jednak, w części książki poświęconej praktycznej implementacji swoich pomysłów (rozdział „How can superminds help us solve our problems?”) porusza podobny wątek generowania strategii biznesowych i proponuje zastanowienie się nad możliwością demokratyzacji tego procesu. Możemy więc rozpatrzeć tę sytuację jako okazję do zrewidowania naszych oczekiwań i zastanowić się nad możliwością zaangażowania w ten proces większej ilości pracowników, generując więcej scenariuszy, jak w Malonowskiej aktywności Tworzenia. Wtedy, ciężar decyzyjny spoczywa wciąż na tych samych osobach, ale pomysły pochodzą zarówno od wspomnianych maszyn, jak i również innych pracowników, hiperpołączonych w obrębie sieci konkursów, opisanych w rozdziale 5.1.5.

Pracę tychże osób wspiera również aktywność Wzmacniania i Interakcji, jak zostały opisane przez P. Daugherty'ego i J. Wilsona. Osoby podejmujące decyzje mają w tym ujęciu wsparcie w postaci zestawu agentów, którzy monitorują dla nich dane, przesyłają sugestie, działając jako bardzo mało doświadczeni, ale bardzo pracowici analitycy nastawieni na szukanie okazji. Co więcej, jeśli cokolwiek ich zainteresuje, mogą wywołać agenta na zasadzie Interakcji i poprosić go o przygotowanie adekwatnych danych z posiadanych systemów informatycznych i zdobycie w ten sposób informacji uszczegóławiających.

Przygotowywania danych

Niestety, przygotowywanie danych to proces ściśle techniczny, więc uwagi Malone'a nie znajdują tu wyrazu. Jednakże, jak wspomniano w poprzednim rozdziale, syntetyczni agenci są w stanie wesprzeć tutaj ludzkich ekspertów, automatyzując część zadań lub też Wzmacniają ich, w sensie P. Daugherty'ego i J. Wislona, odciażając z rutynowych aspektów ich pracy. Warto jednak zauważyć, że w kontekście Trenowania wymienili oni również etap czyszczenia danych.

Analiza danych

Proces analizy danych w ujęciu N. Prata wykazuje silną analogię do procesu Tworzenia u T. Malone'a. I znów, podobne rozwiązania mogą usprawnić ten proces, który w istocie polega na generowaniu możliwie dużej liczby pomysłów. Jak wspomniano w poprzednim rozdziale, mamy już szereg rozwiązań technicznych, które automatycznie generują sugestie. Aspekt kolaboracji i włączania w proces większej liczby ludzi też jest możliwy. Ukazują to, chociażby „Jupyter Notebooks”, które pozwalają na uporządkowanie procesu eksploracji danych i łatwe podzielenie się naszym ciągiem analitycznym z innymi członkami zespołu. Podobnie, Wzmacnianie oraz Interakcje mogą wesprzeć w tym zadaniu analityków poprzez wspomniane automatyczne sugestie, ale też ułatwienie dostępu do danych poprzez delegowanie syntetycznym agentom przeprowadzania zapytań zbierających dane. Warto dodać też aspekt Trenowania, gdyż to na tym w zasadzie etapie budujemy nasz właściwy model danych. Zwłaszcza znajduje tutaj zastosowanie aspekt poprawy wydajności i dopasowania danych do naszych potrzeb.

Zastosowanie modelu

Podobnie jak przygotowywanie danych, aplikowanie modeli stanowi proces ściśle techniczny. Stąd też nie stanowi jako takie elementu procesu decyzyjnego, na którym bazuje Malone. Niemniej, analogicznie jak w poprzednich punktach, Wzmacnianie znajduje tutaj zastosowanie.

Decyzja

Decyzja stanowi naturalne odniesienie do Decydowania, w ujęciu Malone'a. Przez wzgląd na pojedynczą osobę odpowiedzialną za podjęcie decyzji, tak powinniśmy rozważyć opis hierarchicznego podejmowania u Malone. I, tak jak pisał, ujęcie to zakłada, że procesy decyzyjne są sukcesywnie delegowane coraz niżej, a środki technologiczne umożliwiają ich stopniową demokratyzację. Stąd też łatwo wyobrazić sobie sytuację,

gdzie osoba odpowiedzialna za decyzję o końcowym uruchomieniu danego systemu oddeleguje sprawdzenie poprawności danych agentom syntetycznym oraz wyda im polecenie dostarczenia danych źródłowych, aby upewnić się o skuteczności zaproponowanego rozwiązania. Ponownie zachodzą tutaj aktywności określone przez Daugherty'ego i Wilsona mianem Wzmacniania i Interakcji.

Akcja

Zaimplementowanie rozwiązania na środowisku produkcyjnym jest ponownie zadaniem ściśle technicznym i nie ma potrzeby się rozwódzić nad sytuacją analogiczną do zastosowania modelu, opisanego powyżej.

Obserwacja

Etap ten jest w zasadzie tożsamy z Malonowskim Uczeniem się. Zwłaszcza zastosowanie ma tu wariant oparty na eksploatacji i sukcesywnym, iteracyjnym ulepszaniu całego systemu. Ciekawy w tym kontekście jest również pomysł pętli, w ramach której wybory użytkowników stają się informacją zasilającą dane i usprawniającą efektywność modelu jako takiego.

Monitorowanie działania rozwiązania stanowi ciekawy moment, gdyż efektywnie dopiero wtedy do głosu dochodzi cała druga strona brakującego środka: Trenowanie, Wyjaśnianie i Podtrzymywanie. Wszystkie te aktywności bowiem wymagają pewnej maszynowej aktywności, do której możemy się odnieść. I tak Trenowanie sprawdza się zwłaszcza w kontekście systemów często wchodzących w kontakt z użytkownikiem końcowym, jak chociażby systemy rekomendacyjne pokroju Netflix'a czy Spotify. Wtedy to ludzie mogą sprawdzać dopasowanie sugestii sztucznej inteligencji do potrzeb użytkowników. Przykładowo, czy sugerowane utwory są słuchane do końca, w kontekście Spotify. Wyjaśnianie ma tu rolę, zwłaszcza gdy system zaczyna sugerować wybory nieoczywiste. Podtrzymywanie natomiast jest konieczną, stałą aktywnością, która wymaga mnóstwo czasu i uwagi.

6. Zakończenie

6.1. Podsumowanie i wnioski

Jak zadeklarowano w rozdziale 2.4., głównym celem rozprawy jest ujednoczenie istniejącej terminologii, porównanie dwóch kluczowych wizji sztucznej inteligencji (Marvina Minsky'ego oraz Josepha Carla Robnetta Licklida) i ich współczesnych interpretacji oraz stworzenie zbiorczego modelu implementacji rozwiązań sztucznej inteligencji, które efektywnie wspierają ludzi w podejmowaniu decyzji.

O ile porównania zostały dokonane w korespondujących rozdziałach 4.3. oraz 5.3.5., tak ten fragment poświęcony został zadaniu opracowania modelu decyzyjnego wzbogaconego, czy też być może wzmocnionego, o możliwości sztucznej inteligencji.

Zagadnienie, czy sztuczna inteligencja jest w stanie usprawnić nasze procesy decyzyjne, towarzyszy nam prawie tak długo, jak niezależna koncepcja sztucznej inteligencji, a usprawnienie podejmowania decyzji za pomocą mechanizmów i technologii jeszcze dłużej. Co więcej, jak wspomniano w rozdziale 3.1.6., dokładnie ten cel przyświecał rozwojowi technologii sztucznej inteligencji od lat 80. XX wieku, kiedy to popularność zdobywały systemy ekspertowe.

Widzimy przejawy tego podejścia we wszystkich omawianych stanowiskach. Cała społeczność umysłów kręci się wokół tematu decyzyjności. M. Minsky rozważa je w wielu aspektach, począwszy od decyzji pojedynczych agentów, przez agencje, aż po rezultaty widoczne z zewnątrz. Zwraca również uwagę na kwestię wolnej woli i sprawczości pojedynczej osoby na decyzje agentów tworzących jej umysł (Minsky, 1986, s. 306).

Bardziej klasycznie możemy przeczytać o decyzyjności w artykule J. C. R. Licklida o symbiozie ludzi i maszyn, gdzie syntetyczni agenci mieli nam przede wszystkim pomagać zbierać i analizować dane (Licklider, Man-Computer Symbiosis, 1960). Podobnym przykładem jest model podejmowania decyzji zaprezentowany przez T. Malone'a (Malone, 2018). Druga z omawianych, współczesnych publikacji również mierzy się z tym tematem, w pracy P. Daugherty'ego i J. Wilsona widzimy bowiem szereg przykładów, jak maszyny są w stanie wspierać ludzi. Wspieranie decyzji postrzegają oni wręcz jako coś naturalnego i oczywistego, rozwijając ten nurt o bardziej nowoczesne

systemy integracji ludzi i maszyn w świecie fizycznym. Autorzy przywołują również przykłady robotów pracujących równolegle z nami w fabrykach. Wyróżniają oni to, tym bardziej, że specjalną aktywnością brakującego środka jest Ucieleśnianie, jak wspomniano w rozdziale 5.2.3.

6.2. Inne ujęcia podejmowania decyzji

6.2.1. Przegląd definicji

Aby móc stworzyć własny model procesu decyzyjnego, pierwsze w kolejności powinno być wyklarowanie, jak rozumiane są decyzje jako takie. Niektórzy, jak Yingxu Wang i Guenther Ruhe, pojmują je jako jeden z najbardziej podstawowych procesów poznawczych realizowanych przez ludzi (Wang i Ruhe, 2007). Inni, patrząc nieco szerzej, określają akt podejmowania decyzji jako wydawanie sądu względem tego, co dana osoba powinna zrobić w danej sytuacji po zastanowieniu się nad różnymi, alternatywnymi scenariuszami (Ofstad, 1961). Herbert Simon traktuje znów podejmowanie decyzji jako wyrażenie synonimiczne z procesem zarządzania. Według niego postanawianie składa się z trzech faz: wyszukiwania okoliczności pozwalających na podjęcie decyzji, odkrywania możliwości i alternatywnych rozwiązań, wyboru sposobu postępowania (Simon H., 1960).

Frank Harrison przedstawia własną definicję. W jego ujęciu decyzja jest definiowana jako moment wyboru jednego z alternatywnych scenariuszy osiągnięcia celu. Oczekiwania dotyczące określonego kierunku działania zmuszają decydenta do wyboru tej alternatywy, która z największym prawdopodobieństwem i najmniejszym kosztem doprowadzi go do osiągnięcia założonego celu (Harrison, 1995, s. 5). Ta ostatnia jest o tyle istotna, że, podobnie jak u T. Malone'a, zakłada, że podejmowanie decyzji jest jedynie krokiem milowym szerszego procesu.

6.2.2. Podejmowanie decyzji jako proces poznawczy

Podejście Y. Wanga i G. Ruhe zakłada, że preferowany rozwój wypadków jest wybierany ze zbioru alternatywnych scenariuszy, ograniczonego przez określone kryteria (Wang i Ruhe, 2007). Stanowi on jeden z 37 fundamentalnych procesów poznawczych modelowanych w warstwowym modelu odniesień wewnątrz mózgu (org. layered reference model of the brain, tłumaczenie własne),

opisanym we wcześniejszych artykułach Y. Wanga (Wang, Wang i Patel, 2004; Wang, 2007b).

Warto odnotować, że, jak kontynuują autorzy, proces ten stanowi obiekt zainteresowania badaczy z rozlicznych obszarów badawczych, zawierających, ale nie ograniczających się do: informatyki kognitywnej, informatyki, nauk o zarządzaniu, ekonomiki, socjologii, psychologii, politologii oraz statystyki (Berger, 1990; Edwards i Fasolo, 2001; Hastie, 2001; Matlin, 1998; Payne i Wenger, 1998; Pinel, 1997; Wald, 1950; Wang, Wang i Patel, 2004; Wilson i Keil, 2001).

Teorie decyzji mogą być podzielone na deskryptywne i normatywne. Te pierwsze oparte są na empirycznych obserwacjach i eksperymentach nad zachowaniami decyzyjnymi. Te drugie zakładają racjonalnego aktora, który podejmuje decyzję na podstawie dobrze zdefiniowanych preferencji, które przestrzegają, uprzednio założone, aksjomaty racjonalnych zachowań.

Y. Wang i G. Ruhe przyjmują w swoim opracowaniu aksjomaty wyboru Seymoura Lipschutza (1967). Zgodnie z nimi istnieją trzy kluczowe elementy w procesie podejmowania decyzji. Są to: cele decyzji, zbiór alternatyw, klucz wyboru wprowadzający kryteria i inne strategie ograniczające liczbę alternatyw. Warto odnotować, że Y. Wang i G. Ruhe stwierdzają bezpośrednio, że mając dostęp do tych trzech elementów, proces decyzyjny może zostać przeprowadzony czy to przez człowieka, czy inny system inteligentny. Stanowi to więc kognitywną podstawę dla implementowania systemów eksperckich i podobnych im systemów wspierających osoby podejmujące decyzje (Ruhe, 2003; Ruhe i An, 2004; Wang Y., 2007a).

Y. Wang i G. Ruhe dostrzegają, że podejmowanie decyzji stanowi jeden z kluczowych procesów poznawczych realizowanych przez ludzi. Wykorzystują go oni w określaniu racjonalnych, heurystycznych i intuicyjnych wyborów tak, w skomplikowanych sytuacjach naukowych, inżynierskich, ekonomicznych i związanych z zarządzaniem ludźmi, jak również w praktycznie każdej aktywności naszego codziennego życia. Stwierdzają oni dalej, że podejmowanie decyzji jest tak bardzo podstawowym procesem, że zachodzi w naszych umysłach co kilka sekund, zarówno świadomie, jak i nieświadomie (Wang i Ruhe, 2007, s. 74).

Autorzy wyróżniają cztery kategorie decyzji, ze względu na metody ich podejmowania. Są to kolejno decyzje: intuicyjne, empiryczne, heurystyczne i racjonalne. Te ostatnie dzielą się na dwie podkategorie: dynamiczne i statyczne. Jak kontynuują, kategorie intuicyjne i empiryczne są zgodne z naszymi skojarzeniami związanymi z podejmowaniem decyzji i osiągnięciami psychologii poznawczej. Badanie ich jest utrudnione przez brak określonych, racjonalnych modeli, które by je opisywały.

Proces poznawczy realizujący podejmowanie decyzji zawiera według Y. Wanga i G. Ruhe'a 7 kroków. Są to kolejno:

1. Zrozumienie obiektu i jego atrybutów, zgodnie z modelem pamięci OAR, opisanym przez Y. Wanga w jego wcześniejszych artykułach (Wang i Wang, 2004; Wang, Liu i Wang, 2003). Zakładał on, że pamięć składa się z obiektów (O) opisanych atrybutami (A), które są połączone między sobą i innymi obiektami czy atrybutami za pomocą relacji (R). W ramach tego kroku następuje zrozumienie problemu, czy też celu, jaki chcemy osiągnąć. Jest to kluczowe, gdyż autorzy zakładają, że nasze decyzje stanowią drogę do realizacji naszych potrzeb i zamierzeń.
2. Przeszukanie abstrakcyjnej warstwy pamięci długotrwałej (Wang, Liu i Wang, 2003) dla ustalenia zbioru możliwych rozwiązań oraz kryteriów określających użyteczność i umożliwiających porównanie alternatyw, by wybrać najlepiej dopasowane do potrzeb podmiotu realizującego proces. Autorzy zaznaczają, że ten etap realizowany jest nie tylko w pamięci osoby podejmującej decyzję, ale również w źródłach zewnętrznych pozwalających rozszerzyć wiedzę, doświadczenia i oczekiwania podmiotu realizującego proces. Warto również odnotować, że wraz z nabywaniem doświadczenia jesteśmy w stanie znacząco ograniczyć zakres przestrzeni problemowej, którą musimy przeszukać, gdyż już wiemy, które strategie są skuteczne dla określonych sytuacji problemowych .
3. Kwantyfikowanie alternatyw i kryteriów, dzięki czemu możliwe jest ustalenie, czy możliwe jest już podjęcie zdecydowanej decyzji.
4. Zbudowanie zbioru decyzji ze zbiorów alternatyw i kryteriów. Y. Wang i G. Ruhe podkreślają, że stworzony zbiór stanowi iloczyn kartezjański obu elementów.

Im liczniejsze alternatywy i kryteria, tym lepszą decyzję możemy podjąć, gdyż mamy większe spektrum wyboru i lepiej rozumiemy jakie kryteria nami kierują.

5. Wybranie preferowanej lub preferowanych decyzji, na podstawie stopnia zadowolenia osoby podejmującej decyzję. Warto przy tym odnotować, że możemy przerwać proces i uznać, że dana decyzja nie może w ogóle zostać podjęta, gdy żadna z alternatyw nie zaspokaja założonych kryteriów.
6. Określenie poziomu satysfakcji z podjętej decyzji, zgodnie ze wspomnianą wyżej metodologią OAR.
7. Zaktualizowanie pamięci długotrwałej o podjętą decyzję. W dalszej części artykułu autorzy zauważają również, że znalezienie satysfakcjonującego i racjonalnego lub formalnego rozwiązania znacząco obniża złożoność kognitywną wymaganą do poradzenia sobie ponownie z takim, lub podobnym dylematem.

6.2.3. Decyzje strategiczne

Alternatywnie, F. Harrison rozważa decyzje w kontekście strategicznym, nadawania kierunku rozwoju firmom i innym organizacjom. Stanowią one sytuacje, które nadają ton i tempo pracy każdej indywidualnej osobie zatrudnionej w danej firmie, gdyż mają one naturę kaskadową i podjęte „na górze”, spływają „niżej”, wywołując po drodze dziesiątki kolejnych decyzji wydawanych przez kolejnych podwładnych. W swoim podejściu autor wyróżnił pięć cech wyróżniających taką decyzję (Harrison, 1996):

1. Decyzja musi być skierowana na określenie relacji organizacji z jej otoczeniem.
2. Decyzja musi uwzględniać całą organizację jako przedmiot analizy.
3. Decyzja musi obejmować wszystkie najważniejsze funkcje wykonywane w ramach organizacji.
4. Decyzja musi dostarczać ograniczeń i w ten sposób wspierać operacyjne i administracyjne pionierzy organizacji w planowaniu i zarządzaniu swoją pracą.
5. Decyzja musi być krytycznie istotna dla długoterminowego sukcesu organizacji (Shirley, 1982).

Punkt pierwszy, opisujący relację między organizacją i jej otoczeniem, środowiskiem, w którym funkcjonuje, Harrison nazywa strategiczną luką (Harrison, 1996, s. 47). Opisuje ona rozdzźwięk między obecną a pożądaną strategiczną pozycją organizacji (Harrison, Policy, Strategy, and Managerial Action, 1982). Lukę tę opisujemy poprzez porównanie zasobów i możliwości organizacji do okazji oraz zagrożeń znajdujących się poza nią. Gdyby organizacja była w stanie wykorzystać każdą nadarżającą się okazję oraz zneutralizować każde zagrożenie nie byłoby luki. Jest to jednak, jak dowodzi autor, okoliczność nader nieprawdopodobna (Harrison, 1989).

Poprawnie przeprowadzona analiza strategicznej luki zaczyna się, według Harrisona, od oceny kompetencji i możliwości organizacji w zakresie zarządzania, technologii, polityk i posiadanych zasobów. W ten sposób powstaje, tak zwany, profil możliwości opisujący przewagi i bolączki danej organizacji. Jest on kluczowy dla funkcjonowania firmy, gdyż wykorzystanie zewnętrznych okazji wymaga efektywnego użycia własnych sił. Zarówno ocena wewnętrzna, jak i ocena zewnętrzna wymaga przeanalizowania czterech czynników: okazji, zagrożeń, wymagań i obowiązków.

Dopiero po przeanalizowaniu obu stron modelu możemy określić poziom luki strategicznej. Harrison wyróżnia trzy możliwe wyniki: pozytywny, negatywny i zerowy. Warto zaznaczyć, że ten ostatni istnieje jedynie jako byt teoretyczny, gdyż zawsze będzie różnica między organizacją a środowiskiem, w którym funkcjonuje.

Autor motywuje to stwierdzenie takimi nieuniknionymi czynnikami jak niedoskonałe informacje, opóźnienia czasowe w odpowiedzi na zmiany wymuszone przez otoczenie, pomyłki menadżerów czy przełomowe odkrycia w wykorzystywanej technologii. Pozytywna ocena luki oznacza, że organizacja jest dobrze przygotowana wewnętrznie na bodźce, które nadciągają z zewnątrz, czy to nowe okazje, czy też nowe wymagania prawne. Negatywna odwrotnie, jeśli dojdzie do takiej oceny, oznacza to, że organizacja musi skupić się na wewnętrznych usprawnieniach, zanim zacznie liczyć na pomocne bodźce z zewnątrz, gdyż nie jest pewnym czy firma będzie w stanie nawet poprawnie wykorzystać nadarżającą się okazję.

F. Harrison rozpatruje podejmowanie decyzji z perspektywy zarządczej. W jego ujęciu oznacza to, że decyzja stanowi w istocie rezultat całego, powtarzanego

iteracyjnie, procesu decyzyjnego, który składa się z szeregu logicznie powiązanych funkcji. Wyróżnia on następujące etapy:

- Wyznaczenie celów. Pierwszym krokiem w podejmowaniu decyzji jest określenie, do czego tak naprawdę dążymy. Cykl kończy się w momencie osiągnięcia celów.
- Poszukiwanie alternatyw. Obejmuje przeszukiwanie wewnętrznych możliwości i zewnętrznego środowiska w poszukiwaniu informacji, które pozwolą opracować kilka strategii osiągnięcia celu.
- Porównywanie i ocenianie alternatyw. Opracowane w poprzednim kroku strategie są następnie weryfikowane. Konieczne jest w ujęciu Harrisona dokonanie zestawienia dostępnych opcji.
- Akt wyboru. Stanowi to moment w ciągłym procesie, w którym osoba podejmująca decyzję wybiera pojedynczą strategię.
- Implementowanie decyzji. Ponownie, stanowi to moment albo też okres, gdy wybrana strategia przechodzi ze stanu abstrakcji do operacyjnej rzeczywistości.
- Monitorowanie i kontrola. Ostatnim wymienionym etapem w procesie jest funkcja odpowiadająca za określenie czy rezultaty naszych działań są zbieżne z założonymi uprzednio celami.

Powyższy opis funkcji wskazuje na linearność procesu, jednakże F. Harrison podkreśla, że opisane przez niego etapy bynajmniej nie muszą występować w tej kolejności, a funkcje są ze sobą często wzajemnie powiązane. Mimo że proces zawsze zaczyna się od generowania celów, następnie opracowywane są strategie ich osiągnięcia, które muszą zostać zrealizowane, by móc monitorować zbieżność rezultatów i pożądaných wyników, tak całkiem prawdopodobne są sytuacje, gdy ta sekwencyjność zostanie zaburzona. Harrison rozpatruje, chociażby przykład, gdy monitorowane rezultaty zdecydowanie odbiegają od oczekiwań i osoba decydująca musi wrócić do etapu określania celów. Inną taką sytuacją jest moment, gdy żadna z rozpatrywanych alternatyw nie umożliwi osiągnięcia założonego celu.

F. Harrison w końcowej części artykułu łączy analizę luki strategicznej i procesu podejmowania decyzji tworząc złożony model uwzględniający fakt, że nasze decyzje strategiczne nie zapadają w przysłowiowej pustce, lecz stanowią wypadkową interakcji między naszą organizacją i jej możliwościami oraz ambicjami a warunkami, występującymi w ramach środowiska, w którym funkcjonuje.

W tym podejściu jest jeszcze jedna istotna różnica wobec T. Malone'a. F. Harrison używa bowiem określenia decyzji pozytywnej (org. successful decision, tłumaczenie własne). Oznacza ona sytuację, gdzie wybrana strategia została poprawnie zaimplementowana i spowodowała osiągnięcie założonych na samym początku celów. Te bowiem są kluczowe w ujęciu F. Harrisona, a praktycznie pominięte u T. Malone'a.

6.3. Założenia własnego modelu współpracy ludzi i maszyn

Jak wspomniano we wstępie rozprawy, jednym z założonych celów jest stworzenie ujednoliconego, współczesnego modelu sztucznej inteligencji. Warto odnotować, że model jest tu rozumiany podobnie jak w opisie założeń modelu u Przybysza. Zakłada się więc uproszczony obraz świata, społeczeństwa i ludzi jako takich. Założenie o wyidealizowaniu rzeczywistości jest przyjmowane jako oczywiste i wymuszone naturą wyjaśnienia jako takiego. Obraz tak ludzi, jak i syntetycznych agentów musi więc siłą rzeczy zostać uproszczony, aby sprawić, że tworzony konstrukt był użyteczny i możliwie czytelny (Przybysz, 2009, s. 3).

Zakłada się więc ujednoczenie przedstawionych w ramach rozprawy rozproszonych spojrzeń. Zrealizowanie tego celu jest problematyczne już na etapie ustalenia ziarnistości, odpowiedniego poziomu rozważań. M. Minsky pisał bowiem o funkcjonowaniu jednostek, symbiotycy rozważali perspektywę zespołu, T. Malone, w charakterystyczny dla siebie sposób, stworzył opis adekwatny, tak na poziomie pojedynczej osoby, jak i całego kraju oraz wszystkiego pomiędzy a P. Daugherty i J. Wilson poruszali się głównie w obrębie jednostek, a ich sytuację opisywali przykładami całych przedsiębiorstw.

Aby nadać temu pewien ład, przyjmuje się najmniejszy wspólny mianownik tych podejść, pojedynczą osobę. Naturalnie, wchodzi ona w interakcje z innymi ludźmi czy maszynami. Bardziej jednak w rozumieniu Harrisona niż Malone'a, czyli traktujemy inne byty jako zewnętrzne wobec osoby podejmującej decyzję. Uwzględnienie możliwości technologicznych również stanowi cechę dystynktywną w proponowanym opisie. Ostatecznie tworzy to więc pewne novum, o wysokiej adaptowalności do indywidualnych potrzeb pojedynczych osób.

Przyjmujemy tu za Natalią Hatalską (2021) czy Jackiem Dukajem (2021), że uczyniliśmy z siebie nie do końca samodzielny gatunek. Jakby zabrać nam technologię cierpielibyśmy i nie jest pewne czy byśmy przetrwali. Podobnie, nasze procesy myślowe nie są też już tylko nasze. Niektórzy postulują, że nasz umysł został już niejako rozszerzony poprzez interfejsy z komputerami (Clark i Chalmers, 1998). Podobne pojęcie proponowała Donna Haraway w swoim eseju z 1985 roku. Autorka zakłada w nim, że połączenie ludzi i maszyn jest nieuchronne i powinniśmy je zaakceptować jako sposób na przekroczenie tradycyjnych ograniczeń i hierarchii. W ujęciu D. Haraway taki byt powinniśmy nazwać cyborgiem. Według niej stanowiło to nową formę podmiotowości wymykającej się tradycyjnym podziałom jak człowiek/zwierzę czy kultura/natura. Zamiast tego cyborg stanowił byt, który trwał w stanie ciągłej transformacji i adaptacji do bieżących wymagań środowiskowych (Haraway, 1985).

Cyborgi czy nie, ludzie wchodzą w coraz intymniejsze relacje z technologią. Komputery nie tylko nie odstępują nas już zazwyczaj o krok dzięki rozwojowi technologiom komórkowym i ubieralnym, ale również mamy nasze cyfrowe życia, które nigdy się nie zatrzymują. Nasze skrzynki mailowe, niezależnie od nas, wciąż filtrują spam, podejmując decyzje, co jest dla nas istotne. Media społecznościowe organizują kolejności wyświetlanych nam treści, aby zmaksymalizować czas jaki im poświęcamy, a przeglądarki internetowe optymalizują nasze doświadczenie, przewidując jakie strony najprawdopodobniej odwiedzimy następane i zawczasu je ładując.

Brak w tym również intencjonalności i synchronizacji, jakie można było wyczuć w postulatach J.C.R. Licklidera. W jego ujęciu człowiek musiał zadać pytanie i komputer obliczał dla niego odpowiedź. Teraz maszyny zgadują jakie pytania możemy zadać, porządkują je według prawdopodobieństwa i zasypują nas licznymi odpowiedziami. I, co w tym równie istotne, jeśli często współpracujemy i oceniamy te sugestie, to stają się one coraz lepiej dopasowane do naszych potrzeb. Niektórzy postulują odejście od określenia interakcji ludzi i maszyn na rzecz ich integracji (Boy i Narkevicius, 2014).

Jak więc widzimy, koncepcja symbiozy niejako samoistnie wyparła cel budowania inteligentnych maszyn. Obecnie poruszamy się w skomplikowanych środowiskach fizyczno-wirtualnych, w ramach których maszyny przyjmują role opisane przez T. Malone'a, jak wykazano w rozdziale 5.1.2., podrozdziale *Relacje z maszynami*.

Nasz proces decyzyjny jednak wciąż przebiega podobnie do tradycyjnie pojmowanych ujęć zaprezentowanych na przykładzie F. Harrisona czy Y. Wanga i G. Ruhe'a.

6.4. Model współpracy ludzi i maszyn

W ramach tworzonego modelu współpracy ludzi i maszyn za kluczowy proces uznany zostaje proces decyzyjny. Podstawę stanowi tutaj model Y. Wanga i G. Ruhe'a, który następnie zostaje rozbudowany o specyficzne ujęcie T. Malone'a. Otoczenie decyzyjne, za Harrisonem, również zostaje uwzględnione. Na koniec, podobnie jak w rozdziale 5.3., przedstawione zostają techniczne możliwości usprawnienia każdego elementu.

W ramach opisu procesu decyzyjnego przyjmuje się następujące kroki:

1. Uzyskujemy ze środowiska, jak określał je Harrison, informacje będące w istocie zbiorem okazji i zagrożeń, które kształtują nasze możliwości działania.
2. Zebrane informacje porządkujemy w wiedzę. Używamy jej do porównania z naszym stanem wewnętrznym, w procesie zbliżonym do analizy luki strategicznej Harrisona.
3. Wyznaczamy nasze cele.
4. Przygotowujemy zbiór ograniczeń, które pozwolą nam uniknąć eksplozji ilości możliwych scenariuszy, które musimy rozpatrzyć.
5. Następnie tworzymy scenariusze osiągnięcia celów.
6. Podejmujemy decyzję, wybierając scenariusz lub scenariusze, które zrealizujemy.
7. Implementujemy wybrany scenariusz.
8. Obserwujemy wyniki oraz oceniamy skuteczność naszych działań wobec założonych celów.
9. Aktualizujemy nasz stan wiedzy o możliwych problemach i rozwiązaniach.

Tak opisany proces zakłada, podobnie jak w opisywanych wcześniej modelach, zwłaszcza F. Harrisona, cykliczność i wzajemne wpływanie na siebie elementów. Dopuszczalne są również zmiany w kolejności czy przepływie. Warto również odnotować,

że T. Malone opisywał zachowanie inteligentne, nie podejmowanie decyzji sensu stricto. Jego opis uwzględniał podobne elementy, jednakże inaczej je określał. Interakcja z otoczeniem, w jego ujęciu, wyrażana była bowiem przede wszystkim w aktywności doświadczania oraz pamiętania. Brak w tym jednak wyraźnego etapu porównania informacji z zewnątrz z własnymi potrzebami i możliwościami. Decyzja natomiast, zamiast skupić się na procesie, dotyczyła w jego ujęciu różnych sposobów, jak grupy mogą wchodzić ze sobą w interakcje. Zbiorowiska te wciąż podejmowały tradycyjnie pojmowane decyzje, aczkolwiek zachodziło to niejako przy okazji tych interakcji, nie stanowiło ich celu.

6.5. Główne elementy procesu podejmowania decyzji

6.5.1. Środowisko

Środowisko stanowi tło i podstawę wyznaczającą ramę naszych działań. Podobnie jak u Harrisona możemy dostrzec jego pozytywny wpływ w postaci nadarzających się okazji, ale też zagrożeń, które stwarza. Malone zbieranie informacji o otoczeniu nazywał doświadczaniem i wykazywał, jak hiperpołączenie ludzi i maszyn jest w stanie usprawnić ten proces. Widział on również szereg zastosowań technologii powiązanych z Big Data, które umożliwiają włączenie do naszego repertuaru technik zbierania informacji tych, które pochodzą z szeroko rozumianego internetu, w tym zdjęć satelitarnych, publicznych baz danych czy informacji uzyskanych poprzez wyodrębnianie danych bezpośrednio ze stron internetowych (ang. web scraping, tłumaczenie własne).

Warto odnotować, że P. Daugherty i J. Wilson również widzą liczne sposoby na wzmocnienie tego procesu, zwłaszcza w kontekście analizowania środowiska z perspektywy, wspomnianego w rozdziale 2.1., cyfrowego bliźniaka. Takie bezprecedensowe połączenie świata fizycznego i wirtualnego umożliwia nam zupełnie nowe możliwości dostrzegania okazji do usprawnień. Podobnie, wspomniane przez nich aktywności w ramach Wzmacniania stanowią dodatkowe źródła informacji, które są w stanie zasilić naszą bazę informacji, podobnie jak wspomniane w poprzednim akapicie malonowskie Big Data.

6.5.2. Wiedza

Wiedza to wypadkowa opisanych wcześniej koncepcji: produktu analizy luki strategicznej według Harrisona, pamięci długotrwałej w ujęciu Y. Wanga i G. Ruhe'a, ale też Pamiętania i Uczenia się T. Malone'a. Efektywnie stanowi ona podsumowanie tego, co wiemy o własnych możliwościach i ograniczeniach, ale również potrzebach czy marzeniach.

Dzięki procesom na wzór wspomnianej koncepcji F. Harrisona jesteśmy w stanie określać, co możemy realnie osiągnąć, a co w danym momencie pozostaje poza naszymi możliwościami. Warto odnotować, że rozwiązuje to również wspomniany przez niego błąd, gdy nie jesteśmy w stanie wygenerować scenariusza, który zaspokoiłby zakładane cele i nałożone ograniczenia. Nasza wiedza stanowi sumę naszych doświadczeń, ale też zasobów do których mamy dostęp w postaci, chociażby artykułów, książek czy kursów, które są w stanie efektywnie wesprzeć nas w poszerzaniu naszych horyzontów. Sztuczna inteligencja ukierunkowana na tworzenie rekomendacji może pomóc nam, w intencjonalnym ich poszerzaniu poprzez podsuwanie sugestii kolejnych pozycji, które powinniśmy poznać w ramach szeroko rozumianego rozwoju kompetencji w danym zakresie.

6.5.3. Cele

W ramach omawianego modelu przyjmuje się, że proces decyzyjny jest ukierunkowany na coś, a tę ambicję określamy mianem celu. Jest to podejście tożsame z ujęciami F. Harrisona oraz Y. Wanga i G. Ruhe'a.

Porównanie z T. Malonem nie jest proste w tym kontekście. W swoim ujęciu wyróżnił on bowiem cztery rodzaje superumysłów opisujących nasze główne metody podejmowania decyzji. Superumysły nie mają zazwyczaj typowo rozumianych przez nas celów. Wyjątek stanowi być może hierarchia, gdyż opisuje ona organizację, która realizuje ambicje i założenia pojedynczej osoby będącej na jej szczycie. Niemniej demokracje, rynki i społeczności poniekąd odwracają cały ten proces. Cele stają się w nich ewidentnie ponadjednostkowe i kluczem do ich zrozumienia jest przyjęcie założenia, że ludzie stanowią jeden z zasobów jakimi dysponują superumysły. Dążeniem realizowanym przez przykładową społeczność może być po prostu maksymalizacja poczucia zadowolenia z życia mieszkańców określonego osiedla.

Nie będziemy w stanie zrozumieć go na naszym, ludzkim poziomie, lecz intuicyjnie możemy dostrzec formy realizacji takiego założenia w działaniach pojedynczych osób, decyzjach osiedlowej rady czy inicjatywach wspólnotowych.

Wracając jednak do celów rozumianych bardziej tradycyjnie, technologia jest w stanie nas wymiennie wspierać w ich osiągnięciu. Przykładami ilustrującymi takie podejście mogą być aplikacje na telefon oferujące wsparcie w odchudzaniu. Ustalają one plan treningowy, dzienne zapotrzebowanie kaloryczne i formułują cel wyrażony terminem wraz ze spodziewaną wagą po zastosowaniu planu oraz diety na podstawie wypełnionego przez użytkownika kwestionariusza. Nieco wybiegając kilka kroków modelu dalej, taka aplikacja następnie monitoruje nasz postęp i w razie odnotowania wyników odbiegających od oczekiwań potrafi automatycznie zapytać użytkownika o powód i zasugerować plany naprawcze. Podobne rozwiązania stosuje się również przy nauce online języków obcych czy analizy danych.

6.5.4. Ograniczenia

Przez ograniczenia wskazujemy wszelkie czynniki, które zawężają spektrum możliwych do zastosowania strategii osiągnięcia celu. Są one w istocie tożsame z koncepcją kryteriów opisaną przez Y. Wangę i G. Ruhe'a. Tak samo zakłada się w ramach omawianego modelu, że kryteria stanowią efektywnie zbiór ograniczony wiedzą, doświadczeniem i oczekiwaniami wobec celu. Mogą być definiowane wewnątrz umysłu podmiotu decydującego, jak również poza nim, w źródłach zewnętrznych. Różnicę w opisywanych podejściach stanowi wymiar czasu. U Y. Wangę i Ruhe'a zakładana jest jednoczesność tworzenia strategii i kryteriów. Tutaj natomiast ograniczenia stanowią krok poprzedzający formowanie strategii, gdyż w ten sposób tworzymy bariery wyznaczające zasadność proponowanych strategii.

Wsparcie technologiczne na tym etapie zasadza się na aktywności Wzmacniania, jak opisali ją P. Daugherty i J. Wilson. W kontekście odniesienia do T. Malone'a widzimy to jako wypadkową Pamiętania i Doświadczenia. Ograniczenia bowiem mogą pochodzić zarówno z naszej pamięci, jak i również stanowić przypadek poszerzania wiedzy podmiotu decydującego. Może to nastąpić poprzez zaoferowanie mu danych poszerzających jego spojrzenie na określone zagadnienie, jak choćby spodziewane wartości brzegowe czy zagregowany na przestrzeni wielu zdarzeń przeciętny czas trwania określonej

czynności. Warto podkreślić, że jest to przejaw wspomnianego w poprzednim akapicie źródła ograniczeń pochodzącego spoza podmiotu decydującego.

6.5.5. Scenariusze

Zarówno F. Harrison, jak i Y. Wang wraz z G. Ruhem określają scenariusze mianem alternatyw. W odniesieniu do T. Malone'a stanowią one bezpośredni przejaw procesu Tworzenia. Sprowadza się ona do generowania scenariuszy opisujących jakie akcje muszą zostać podjęte, by zniwelować różnice między stanem bieżącym a stanem końcowym wyrażonym celem ustalonym wcześniej. Scenariusze są ściśle kontrolowane poprzez ustalone w poprzednim kroku ograniczenia.

Warto odnotować, że w procesie tworzenia scenariuszy, o ile zaspokojone zostają ograniczenia, nie ma wielu limitów. Podobnie, na tym etapie nieefektywne jest ich ocenianie. Zamiast tego kluczem jest stworzenie jak największej ich ilości. Y. Wang i G. Ruhe stwierdzają, że rozmiar iloczynu kartezjańskiego alternatyw i kryteriów, koreluje bezwzględnie pozytywnie z końcowym zadowoleniem ze zrealizowanego procesu decyzyjnego (Wang i Ruhe, 2007).

Malone w opisie Tworzenia kierował się podobną intuicją, proponując szereg środków, jak można stworzyć możliwie wiele scenariuszy realizujących nasze cele, zostały one opisane w adekwatnym fragmencie rozdziału 5.1.4. Daugherty i Wilson przedstawiają szereg sytuacji, gdzie najpierw maszyny tworzą scenariusze na podstawie zaprogramowanych przez ludzi ograniczeń, a następnie ludzie dokonują ostatecznego wyboru z szeregu propozycji.

Możliwości wsparcia ludzi przez maszyny w tworzeniu licznych scenariuszy oddaje szczególnie dobrze przykład projektowania lepszych kolców dla sprinterów w butach Nike. Problem z kolcami polega na znalezieniu materiału, który będzie jednocześnie dość twardy, by nie zginać się w kontakcie z podłożem i tak lekki, jak to tylko możliwe, by nie obciążać osoby biegnącej. Dzięki zastosowanym algorytmom projektanci Nike mogli wymyślić tego rodzaju buty na nowo, oddelegowując maszynom zadanie sprawdzenia parametrów milionów materiałów i stworzeniu początkowych projektów. Najlepiej dopasowane do ustalonych uprzednio ograniczeń projekty zostały następnie przeanalizowane przez ludzkich projektantów i analityków, którzy wybrali najbardziej obiecujące projekty. Zrealizowali je na drukarkach 3D,

przetestowali i powtórzyli proces, aż znaleźli idealne rozwiązanie. Tak zaprojektowane buty pozwalały, jak podsumowują P. Daugherty i J. Wilson, polepszyć czas zawodnika o dziesiąte części sekundy, co w realiach sprinterskich mogło realnie przełożyć się na zajęcie pierwszego miejsca, zamiast czwartego (Daugherty i Wilson, 2018, s. 74-75).

6.5.6. Decyzje

Podjęcie decyzji jest kluczowym, ale bynajmniej nie końcowym, momentem w opisywanym procesie. Stanowi on kulminację wykonanych uprzednio analiz i refleksji. Skutkuje wybraniem scenariusza lub, rzadziej, kilku, które najskuteczniej zrealizują obrane wcześniej cele. Y. Wang i G. Ruhe dostrzegają, jak wspomniano wcześniej, wspomniane wyżej cztery główne metody podejmowania ich: intuicyjne, empiryczne, heurystyczne i racjonalne. Każda z nich wiąże się z innymi strategiami oraz kryteriami, wszystkie jednak dążą do tego, by ułatwić decydentowi podjęcie decyzji, z której będzie zadowolony. Zadowolenie stanowi zresztą, jak zauważają Y. Wang i G. Ruhe, często przywoływany w literaturze aspekt decyzyjny. Zakłada się bowiem, że staramy się często wyobrazić sobie nas w przyszłości, w której podjęliśmy taką, a nie inną decyzję i w ten sposób zdecydować czy podoba nam się taka projekcja. Y. Wang i G. Ruhe określają to dokładnie kategorią intuicyjnego podejmowania decyzji, w ramach której kierujemy się strategią preferencji (org. preference, tłumaczenie własne), a przyjęte kryteria zawierają między innymi właśnie nasze oczekiwania.

T. Malone dzieli podejmowanie decyzji ze względu na cztery możliwe typy superumysłów: hierarchie, demokracje, rynki i społeczności. Ich dokładny opis został przytoczony w rozdziale temu poświęconemu, 5.1.4. Jednakże warto w tym miejscu przypomnieć główne metody każdego superumysłu. Decyzje zapadają w hierarchiach na ich szczycie, gdzie jednostki decydują i ich zdanie kaskaduje na podwładnych. Demokracje wspólnie decydują na drodze głosowania uprawnionych członków. Rynki nie tyle podejmują decyzje w sposób świadomy, ile wyłaniają się one na drodze interakcji i pojedynczych transakcji ich uczestników. Członkowie społeczności realizują decyzje, które wynikają z podążania za wartościami istotnymi dla określonych superumysłów.

O ile technologiczne usprawnienia decyzji ludzkich zostały wyczerpująco przedstawione w rozdziałach poświęconych typom superumysłów Malone'a

oraz Wzmacnianiu P. Daugherty'ego i J. Wislona, o tyle odwrócona sytuacja, w której człowiek wspiera maszynę, nie została jeszcze poruszona. P. Daugherty i J. Wilosn opisują takie sytuacje mianem Trenowania, jak opisano je w rozdziale 5.2.3. Warto jednak wspomnieć, że podobny cel realizuje model człowieka w pętli (org. Human in the loop, tłumaczenie własne), który został uprzednio wspomniany. Zakłada on relację agentów biologicznych i syntetycznych, w ramach której człowiek jest wpisany w pętlę zachowań maszyny. Oznacza to, że komputery są w stanie niezależnie podejmować większość rutynowych decyzji, na podstawie dobrze opisanych ograniczeń, ale wywołują one człowieka, gdy natykają się na problemy nietypowe, będące realnie anomaliami spodziewanych wyników lub niemieszczące się w dostarczonych scenariuszach.

6.5.7. Implementacje

Zakłada się szerokie spektrum możliwości zrealizowania podjętych decyzji. Mogą to być jak u F. Harrisona, oddelegowane decyzje operacyjne realizujące zaplanowane cele strategiczne, ale też prostsze wykonanie szeregu czynności założonych na etapie tworzenia scenariusza. Efektywnie jednak proces implementacji jako takiej jest zbyt różnorodny, by móc go wyczerpująco opisać. Podobnie jak zauważono w rozdziale 5.3.5., opisując akcję, implementacja stanowi czynność ściśle techniczną i operacyjną.

6.5.8. Oceny

Po zrealizowaniu wybranego scenariusza należy poznać i ocenić jego skutki. Y. Wang i G. Ruhe podkreślają na tym etapie ocenę zadowolenia z podjętej decyzji. Podobnie, F. Harrison określa ten stan mianem kontroli i podkreśla konieczność ustalenia w jakim stopniu zaimplementowane środki zbliżyły nas do założonego na początku celu. Środki techniczne wspierające nas w tym celu zostały wyczerpująco opisane w rozdziale 5.3.5., w sekcji Obserwacja, z jednym zastrzeżeniem, że Uczenie się, opisane przez Malone'a, jest czynnością ciągłą i znajduje zastosowanie zarówno na tym etapie, jak i również w kolejnym.

6.5.9. Pamięć

Ostatnim krokiem realizowanym w procesie decyzyjnym jest zapamiętanie skutków i podjętych kroków. Jak dowodzą Wang i Ruhe, dzięki temu w znaczącym stopniu redukujemy skomplikowanie kolejnych decyzji przed jakimi staniemy (Wang i Ruhe, 2007,

s. 81). Będziemy już bowiem wiedzieć, co robiliśmy wcześniej, jakie były tego konsekwencje i jaki był nasz poziom satysfakcji z tym związany. Dzięki temu, gdy staniemy przed podobną decyzją w przyszłości, będziemy już wiedzieć, co zadziało się poprzednim razem. Jak wspomniano w części poświęconej uczeniu się u Malone'a, proces ten może być również wykonywany przez system komputerowy na drodze pętli uczenia się. W jej ramach agent syntetyczny monitoruje działania podejmowane przez człowieka i na ich podstawie tworzy dla siebie zestawy ograniczeń, które następnie jest w stanie autonomicznie wprowadzać w proponowanych przez siebie rozwiązaniach, sukcesywnie przejmując na siebie coraz więcej decyzji.

7. Dyskusja

Zrealizowane w ramach dysertacji analizy materiałów źródłowych wykazały dwie rzeczy: ewidentnie brakuje ujednoczonego aparatu pojęciowego oraz widoczności na inne spojrzenia, rozumianego jako czerpanie inspiracji z badań i dokonań innych badaczy. W ramach rozprawy przeanalizowano szereg modeli opisujących relacje ludzi i maszyn, ze szczególnym uwzględnieniem możliwości tworzenia zespołów hybrydowych, łączących najlepsze cechy obu rodzajów agentów. Ostatecznie, przeprowadzono dwie próby odniesienia różnorodnych podejść do stworzenia dwóch opisów: wzmocnionej analityki i nowego modelu podejmowania decyzji.

Rozprawa rozpoczyna się, w części I, zarysowaniem rysu historycznego, tak pod kątem myśli technicznej, jak również filozoficznej. Dzięki niemu widzimy, jak sztuczna inteligencja zaczynała od górnolotnych ambicji zrealizowania antycznego marzenia człowieka o zbudowaniu maszyny myślącej i przeszła do zdecydowanie bardziej pragmatycznych implementacji usprawniających nasze procesy analityczne i, docelowo, decyzyjne. Celem tego rozdziału było podsumowanie naszego ogólnego, powszechnego poziomu wiedzy o sztucznej inteligencji.

W II części porównane zostały dwie główne koncepcje, czym jest dla nas sztuczna inteligencja. Podejście Marviniego Minsky'ego stanowiło tu bardzo zaawansowany przykład projektowania sztucznych umysłów, podejścia, które wcześniej w rozprawie określono mianem obliczeniowej koncepcji umysłu. Badacz ten przez całą swoją karierę dążył do zbudowania nieludzkiego, w pełni syntetycznego umysłu. Cel ten jest wciąż żywy w ambicjach niektórych badaczy, jednakże stanowią oni mniejszość, jak podsumowano we wcześniejszym fragmencie dysertacji, w rozdziale 1.2.3. Krytyka współczesnej sztucznej inteligencji. Jakkolwiek mniej istotne obecnie, wciąż żywo rozpala naszą wyobraźnię, a wzmianki o tym modelu widzimy relatywnie często, chociażby u analizowanego w części III Thomasa Malone'a.

Drugą koncepcją rozważaną w II części jest opracowanie stanowisk badaczy z lat 60. XX wieku: Josepha Carla Robnetta Licklidersa, Douglasa Engelbarta i Rossa Ashby'ego, określanych w ramach rozprawy zbiorczo symbiotykami. Jakkolwiek oceniane w momencie publikacji jako czcze życzenia, idee wyrażone

w ich pracach nabierają coraz większego znaczenia, a ich popularność i rozpoznawalność rośnie.

Jest to szczególnie istotne w tworzeniu podejścia nazywanego wzmocnioną inteligencją. Wyrasta ono z odkrycia, że pozostawienie sztucznej inteligencji samej sobie nie stanowi dobrego pomysłu, gdyż nienadzorowana popada efektywnie w stagnację. Przez wzgląd na to, że stanowi ona jedynie drobne permutacje początkowego zbioru uczącego, nie jest ona w stanie zastąpić człowieka w zadaniach związanych z kreatywnością czy sytuacjach niejasnych, nieoczywistych.

Co gorsza, prowadzi też często do propagacji uprzedzeń zawartych w danych. Paradoksalnie więc maszyna, po której spodziewamy się obiektywności i neutralności, staje się dla nas lustrem i pozwala zobaczyć jakie błędy popełnialiśmy. Przykładem takiej sytuacji jest firma, która zdziwiła się zachowaniem syntetycznego agenta, który odrzucał kobiety czy kandydatów o obco brzmiących imionach. Przeprowadzona analiza techniczna wykazała, że nie było to spowodowane konstrukcją algorytmu, a wcześniejszym zachowaniem rekruterów, które stanowiło zbiór uczący (Blechar, 2018).

Tymczasem, wspomniani badacze, symbiotycy, już w latach 60. zwracali uwagę, że potrzebujemy maszyn, tak jak one potrzebują nas. Popadamy więc z technologią w relację symbiotyczną i przestajemy ją traktować jako proste narzędzie. Warto zauważyć, że tezy stawiane przez wspomnianych badaczy wciąż są aktualne. Co więcej, możemy zaryzykować stwierdzenie, że stały się wręcz pewną oczywistością, jak chociażby Lickliderowskie zlecenie maszynom zadań rutynowych i prostych, takich jak przeszukiwanie baz danych czy śledzenie transakcji. Widzimy więc powszechne zastosowania symbiotycznych podejść w stanowieniu celów biznesowych i badawczych.

Część III dotyczy czasów obecnych i roli ludzi w dobie dojrzałej sztucznej inteligencji. Znowu przeanalizowane zostają dwie koncepcje. Pierwsza, T. Malone'a, dotyczy nie tyle sztucznej inteligencji jako takiej, ile wykorzystania współczesnych możliwości technicznych do tworzenia, jak określa to autor, hiperpołączenia między ludźmi, ale też maszynami. Wprowadza on koncepcję inteligencji grupowej i rozważa w jaki sposób możemy usprawnić nasze, ogólnie rozumiane, procesy kognitywne. Opisuje on swój ogólny model, na który składają się kolejne aktywności mentalne: Tworzenie, Decydowanie, Doświadczenie, Pamiętanie i Uczenie się.

Każda z nich zostaje przez niego opisana i przeanalizowana pod kątem tego, jak można ją usprawnić. Ze szczególną uwagą przygląda się on Decydowaniu, gdyż oparte jest ono na kolejnej kluczowej koncepcji zawartej w książce, superumysłach. Stanowią one zbiorcze byty, które współpracują w osiągnięciu celu. Warto odnotować również wzmiankę Malone'a, że mogą one być analizowane za pomocą też M. Minsky'ego opisanych w poprzedniej części rozprawy, z tą różnicą, że agentem, rozumianym według podejścia M. Minsky'ego, mogą być tu cali ludzie, a umysłem cała firma (Malone, 2018, s. 73).

Kolejną koncepcją rozważaną w ramach części III jest brakujący środek, opisany przez Paula Daugherty'ego oraz Jamesa Wilsona. Stanowi ona próbę opisanie relacji między ludźmi a maszynami, gdzie oba rodzaje wzajemnie wspierają się w osiągnięciu celów. Kluczową myślą jest w tym ujęciu podział zadań na ściśle ludzkie, ściśle maszynowe oraz hybrydowe, podzielone na sytuacje, w których ludzie wspierają maszyny lub, w których maszyny wspierają ludzi. Zadania hybrydowe są według autorów niedostatecznie utylizowane w większości przedsiębiorstw. Ich książka ma za zadanie zainspirować ludzi do zwiększenia wykorzystania sztucznej inteligencji w sposób, który zwiększy możliwości ludzi.

Część trzecia kończy się porównaniem obu koncepcji na przykładzie wzmocnionej analityki. Stanowi ona naturalne rozwinięcie koncepcji technologii business intelligence polegającej na demokratyzacji, rozumianej jako upowszechnienie, rozwiązań uczenia maszynowego na szeregowych pracownikach, na co dzień niepowiązanych z analityką czy statystyką. W ramach dysertacji zostają również podsumowane najważniejsze narzędzia dostępne na rynku, które są w stanie wesprzeć nas w projektowaniu, implementowaniu i rozwoju takich rozwiązań.

Wzmocniona analityka jest doskonałym przykładem pragmatycznej implementacji sztucznej inteligencji wspierającej ludzi w podejmowaniu decyzji w ograniczonym, ściśle zaprojektowanym środowisku. W ramach rozprawy przeanalizowany zostaje model operacji analitycznych zaproponowany przez Nicolasa Prata pod kątem możliwości wykorzystania pomysłów T. Malone'a, P. Daugherty'ego oraz J. Wilsona.

W ramach zakończenia zrealizowany zostaje cel rozprawy zadeklarowany na początku dysertacji, opracowanie własnego modelu decyzyjnego na podstawie artykułów Yingxu Wanga i Guentera Ruhe'a, opisujących podejmowanie decyzji

z perspektywy neuroinformatyki oraz Franka Harrisona, który rozpatrywał decyzje strategiczne dla przedsiębiorstw. W efekcie powstaje własny model decyzyjny, rozszerzony również o koncepcje T. Malone'a, P. Daugherty'ego oraz J. Wilsona. Wartość badawcza postulowanego modelu jest tym wyższa, że uwzględnia on perspektywę jednostki i rolę środowiska w podejmowanych decyzjach, co stanowi poniekąd nowość w kontekście wspomnianych teorii.

Dzięki samodzielnemu opracowaniu modelu decyzyjnego możliwe było właściwe przetestowanie modeli współpracy ludzi i maszyn. Niestety, wynik nie był zadowalający, gdyż zarówno koncepcje T. Malone'a oraz P. Daugherty'ego i J. Wilsona były nazbyt ogólne. Jakkolwiek inspirujące, zastosowanie ich w praktyce nie stanowi wielkiej wartości dodanej. Zaproponowany model decyzyjny stanowi jednak kompletne narzędzie, możliwe do wykorzystania w szeregu sytuacji opartych na opracowaniu strategii, czy to jednostki, czy też przedsiębiorstwa, z wykorzystaniem własnych możliwości, jak również usytuowaniem ich procesu decyzyjnego w szerszym kontekście, czego zdecydowanie brakowało u Malone'a.

Najważniejszym jednak osiągnięciem opracowania tego modelu jest wykazanie konieczności dalszych badań w zakresie optymalizacji wykorzystania technologii do, szeroko i wnikliwie, przebadanego obszaru jakim jest podejmowanie decyzji, jak określają go Y. Wang i G. Ruhe. Z drugiej jednak strony, publikacja T. Malone'a pokazuje, że dzięki technologii cały nasz proces decyzyjny może wkrótce diametralnie się zmienić i tezy F. Harrisona z połowy lat 90. XX wieku dotyczące wzniosłej roli osób decyzyjnych w strukturach firmowych mogą zostać rozmyte, zdemokratyzowane i rozdystrybuowane między licznych pracowników. Nie bez znaczenia jest bowiem fakt, że technologia kształtuje nas na równi z tym, jak my kształtujemy technologię, na podobnej zasadzie jak pisał Bruno Latour, stwierdzając, że technologia jest aktorem tak jak człowiek i też może wywoływać w nas konkretne zachowania (Latour, 1991).

Bibliografia

- Abas, Z., Rahman, F., Pramudya, G., Wee, S., Kasmin, F., Yusof, N., Yunus, N., Abidin, Z. (2020). Analytics: A review of current trends, future application and challenges. *Compusoft*, s. 3560–3565.
- Adamowicz, Z. i Zbierski, P. (1997). *Logic of mathematics. A modern course of classical logic. „Pure and Applied Mathematics”*. New York, NY: A Wiley-Interscience Publication. John Wiley & Sons, Inc.
- Adams, L. (2011, Maj 21). *Learning a New Skill is Easier Said Than Done*
Pobrano z: <https://www.gordontraining.com/free-workplace-articles/learning-a-new-skill-is-easier-said-than-done/> (data dostępu: 2022, Październik 10).
- Alghamdi, N. A. i Al-Baity, H. H. (2022, Październik 21). Augmented Analytics Driven by AI: A Digital Transformation beyond Business Intelligence. *Sensors* 22, no. 20: 8071.
- Allen, P. i Greaves, M. (2011, Październik 12). *Paul Allen: The Singularity Isn't Near*. Pobrano z: MIT Technology Review: <https://www.technologyreview.com/2011/10/12/190773/paul-allen-the-singularity-isnt-near/> (data dostępu: 2021, Grudzień 3).
- Alpar, P. i Schulz, M. (2016). Self-service business intelligence. *Business Informatics System Engineering* 58(2), s. 151-155.
- Anderson, J. R. (1980). *Cognitive Psychology and Its Implications*. San Francisco: W. H. Freeman.
- Ashby, R. (1957). *An Introduction to Cybernetics*. Londyn: Chapman & Hall Ltd.
- Ashby, W. R. (1940, Maj 1). Adaptiveness and Equilibrium. *Journal of Mental Science*, Volume 86, Issue 362, s. 478 - 483.
- Banko, M. i Brill, E. (2001). Scaling to very very large corpora for natural language disambiguation. *ACL-01*, s. 26–33.
- Bellman, R. E. (1957). *Dynamic Programming*. Princeton: Princeton University Press.
- Beniger, J. (1986). *The Control Revolution. Technological and Economic Origins of the Information Society*. Cambridge, MA: Harvard University Press.
- Berger, J. (1990). *Statistical decision theory - Foundations, concepts, and methods*. New York, NY: Springer-Verlag.

Bernal, J. (1929). *The Worlds, the Flesh & the Devil. An Enquiry into the Future of the Three Enemies of the Rationality Soul*. Bloomington: Indiana University Press.

Biecek, P., Chlebus, M., Gajda, J., Gosiewska, A., Kozak, A., Ogonowski, D., Sztachelski, J., Wojewnik, P. (2021, Kwiecień 14). *Arxiv. Cornell University*. Pobrano z: Enabling Machine Learning Algorithms for Credit Scoring -- Explainable Artificial Intelligence (XAI) methods for clear understanding complex predictive models: <https://arxiv.org/abs/2104.06735> (data dostępu: 2022, Czerwiec 7).

Blau, L., Jacobs, J. i Eisenstein, J. (2006, Marzec 01). *The Jewish Encyclopedia*. Pobrano z The Jewish Encyclopedia: <https://www.jewishencyclopedia.com/articles/6777-golem> (data dostępu: 2020, Listopad 22).

Blechar, Ł. (2018). Sztuczna Inteligencja i Wzmacnianie Inteligencji w nurcie Antropologii Cyborgów. *Różne optyki rozważań filozoficznych w naukach społecznych i przyrodniczych*, s. 43-55.

Blechar, Ł. (2019). Rewolucja sztucznej inteligencji a wzmacnianie inteligencji. *Humanizacja Pracy 3/2019*, s. 85-99.

Block, H. D., Knight, B. i Rosenblatt, F. (1962). Analysis of a four-layer series coupled perceptron. *Rev. Modern Physics*, 34(1), s. 275–282.

Bob (2012, Czerwiec 13). *Piratenpartij*. Pobrano z Uitleg LiquidFeed systeem: <https://web.archive.org/web/20130905224944/https://www.piratenpartij.nl/liquidfeedback/pagina/uitleg-liquidfeed-systeem> (data dostępu: 2023, Styczeń 19).

Bobryk, J. (2004). *Świadomość człowieka w epoce mediów elektronicznych*. Warszawa: PTS.

Boole, G. (1847). *The Mathematical Analysis of Logic: Being an Essay towards a Calculus of Deductive Reasoning*. Cambridge, MA: Macmillan, Barclay, and Macmillan.

Bostrom, N. (2008). Dignity and enhancement. W A. Schulman, *Human Dignity and Bioethics: Essays Commissioned by the President's Council on Bioethics* (s. 173-207). Washington: the President's Council on Bioethics.

Bostrom, N. (2014). *Superintelligence: Paths, Dangers, Strategies*. Oksford: Oksford University Press.

Bostrom, N. (2016). Superinteligencja Scenariusze, strategie, zagrożenia. W: D. Konowrocka-Sawa, *Superinteligencja Scenariusze, strategie, zagrożenia* (s. 21). Gliwice: Helion SA.

Boy, G. i Narkevicius, J. (2014). Unifying Human Centered Design and Systems Engineering for Human Systems Integration. *Complex Systems Design and Management*, s. 151-162.

Broadbent, D. E. (1958). *Perception and Communication*. Cambridge, MA: Pergamon.

Broadwell, M. (1969). Teaching For Learning (XVI.). *The Gospel Guardian*, s. 1-3.

Brooks, R. (2017, Październik 6). *The Seven Deadly Sins of AI Predictions*. Pobrano z: The MIT Technology Review: <https://www.technologyreview.com/2017/10/06/241837/the-seven-deadly-sins-of-ai-predictions/> (data dostępu: 2022, Maj 5)

Brown, J. S. i Duguid, P. (2000). Balancing Act: How to Capture Knowledge Without Killing It. *Harvard Business Review* 78, no. 3, 73-80.

Brynjolfsson, E. i McAfee, A. (2014). *The Second Machine Age: Work Progress and Prosperity in a Time of Brilliant Technologies*. New York, NY: W. W. Norton.

Bryson, A. E. i Ho, Y.-C. (1969). *Applied optimal control: Optimization, estimation, and control*. Waltham: Mass: Blaisdell Pub. Co.

Buchanan, B. G., Sutherland, G. L. i Feigenbaum, E. A. (1969). Heuristic DENDRAL: A program for generating explanatory hypotheses in organic chemistry. (B. Meltzer, D. Michie i M. Swann, Redaktorzy), s. 209–254.

Bulusu, L. i Abellea, R. (2020). *AI Meets BI: Artificial Intelligence and Business Intelligence, 1st ed.* New York, NY: Auerbach Publications.

Cabeza, R. i Nyberg, L. (2001). Imaging cognition II: An empirical review of 275 PET and fMRI. *Cognitive Neuroscience*, 12, s. 1-47.

Carnap, R. (1928). *Der Logische Aufbau der Welt*. Berlin: Meiner Verlag.

Case, A. (2014). *An Illustrated Dictionary of Cyborg Anthropology*. CreateSpace Independent Publishing Platform.

Cellan-Jones, R. (2014, Grudzień 2). *Stephen Hawking warns artificial intelligence could end mankind*. Pobrano z: [bbc.com](https://www.bbc.com/news/technology-30290540): <https://www.bbc.com/news/technology-30290540> (data dostępu: 2021, Październik 20).

Chalmers, D. (1999). *Towards a Science of Consciousness III: The Third Tucson Discussions and Debates*. Cambridge, MA: The MIT Press.

Cheeseman, P. K. (1991). Where the really hard problems are. *IJCAI'91: Proceedings of the 12th international joint conference on Artificial intelligence* (s. 331–337). Sydney, Australia: IJCAI'91.

Chen, H., Chiang, R. i Storey, V. (2012). Business intelligence and analytics: from big data to big impact. *MIS Q* 36(4), s. 1165-1188.

Ciupa, M. (2017). Is AI in Jeopardy? The Need to Under Promise and Over Deliver - The Case for Really Useful Machine Learning. *Fourth International Conference on Computer Science and Information Technology (CoSIT 2017) March 25~26, 2017* (s. 59-70). Genewa: AIRCC Publishing Corporation.

Clark, A. i Chalmers, D. (1998). The Extended Mind. *Analysis, vol. 58, no. 1.*, s. 7-19.

Clynes, M. i Kline, N. (1960). Cyborgs and Space. *Astronautics*, s. 26-27, 74-75.

Cobham, A. (1964). The intrinsic computational difficulty of functions. *Proceedings 1964 International Congress for Logic, Methodology, and Philosophy of Science* (s. 24-30). International Congress for Logic, Methodology, and Philosophy of Science.

Cohen, P. R. (1995). *Empirical methods for artificial intelligence*. Cambridge, MA: MIT Press.

Collins, F. S., Morgan, M. i Patrinos, A. (2003). The human genome project: Lessons from largescale biology. *Science*, 300(5617), s. 286–290.

Cook, S. A. (1971). The complexity of theoremproving procedures. *STOC '71: Proceedings of the third annual ACM symposium on Theory of computing* (s. 151–158). Shaker Heights, Ohio, USA: Association for Computing Machinery, New York, NY.

Craik, K. J. (1943). The Nature of Explanation. *Philosophy, Volume 19, Issue 73.*, s. 173 - 174.

Cullingford, R. E. (1981). Integrating knowledge sources for computer “understanding” tasks. . *IEEE Transactions on Systems, Man and Cybernetics (SMC)*, 11., s. 52-60.

Curtiss, P. R. i Warren, P. W. (1973). *The dynamics of life skills coaching*. Prince Albert: Training Research and Development Station Department of Manpower and Immigration.

Dalston, L. i Gallison, P. (2007). *Objectivity*. Boston MA: Zone Books.

Daugherty, P. i Wilson, J. (2018). *Human + Machine: Reimagining Work in the Age of AI*. Boston, Massachusetts: Harvard Business Review Press.

- Deacon, T. (1997). *The Symbolic Species*. New York: W. W. Norton & Company, Inc.
- Dennett, D. (2006). *Sweet Dreams - Philosophical Obstacles to a Science of Consciousness*. Cambridge, MA: The MIT Press.
- Dizkes, P. (2021). *News MIT*. Pobrano z: Study: As a population gets older, automation accelerates.: <https://news.mit.edu/2021/aging-workers-automation-0915> (data dostępu: 2022, Czerwiec 13).
- Dreyfus, H. (1965). *Alchemy and Artificial Intelligence*. SANTA MONICA CA: RAND CORP .
- Dreyfus, H. (1972). *What computers can't do : a critique of artificial reason*. New York, NY: Harper & Row.
- Dreyfus, H. i Dreyfus, S. (1987). How to Stop Worrying About the Frame Problem Even Though It's Computationally Insoluble . W: Z. Pylyshyn, *The Robot's Dilemma: The Frame Problem in Artificial Intelligence* (s. 95-112). Westport CN: Praeger.
- Dreyfus, H. i Dreyfus, S. (1988). Making a Mind versus Modeling the Brain: Artificial Intelligence Back at a Branchpoint. *Vol. 117, No. 1, Artificial Intelligence*, s. 15-43.
- Dreyfus, H., Dreyfus, S. i Athanasiou, T. (1986). *Mind over machine : the power of human intuition and expertise in the era of the computer*. New York, NY: Free Press.
- Dukaj, J. (2022, Październik 15). Starość Aksolotla, sztuczna inteligencja i odbiór literatury, czyli wywiad z Jackiem Dukajem. (A. Włoka, Osoba przeprowadzająca wywiad) Pobrano z: Youtube.com: <https://www.youtube.com/watch?v=mxVUcZKhEmA> (data dostępu: 2022, Grudzień 1).
- Dyer, M. (1983). *In-Depth Understanding*. Cambridge, MA: MIT Press.
- Dyer, M. (1991). A society of ideas on cognition: review of Marvin Minsky's "The Society of Mind". *Artificail Intelligence 48*, s. 321-334.
- Edmonds, J. (1965). Paths, trees, and flowers. *Canadian Journal of Mathematics*, 17, s. 449-467.
- Edwards, W. i Fasolo, B. (2001). Decision technology. *Annual Review of Psychology*, s. 581-606.
- Eleftheriou, G. (2018, Czerwiec 1). The death of psychotherapy as we know it. A manifesto for augmented mental health. San Francisco, Bay Area, USA.

- Engelbart, D. (1962). *Augmenting Human Intellect: A Conceptual Framework*. Washington: Director of Information sciences Air force office of scientific research.
- Erl, T., Khattak, W. i Buhler, P. (2015). *Big data fundamentals: concepts, drivers & techniques*. Upper Saddle River: Prentice Hall.
- Feenberg, A. (1991). *Critical Theory of Technology*. Oxford: Oxford University Press.
- Franklin, S. (2001). *Artificial Minds*. Cambridge, MA: The MIT Press.
- Frey, C. B. i Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation?., *Technological Forecasting and Social Change, Volume 114*, s. 254-280.
- Froese, T. (2011). Tchnąć nowe życie w kognitywistykę. *Avant. The Journal of the Philosophical-Interdisciplinary Vanguard, vol. II*, s. 95-112.
- Gajewski, M. (2018, 05 28). *Czy zabicie hosta to zbrodnia? O Westworld i sztucznej inteligencji - rozmowa z dr Piotrem Bilskim*. Pobrano z: spidersweb.pl <https://spidersweb.pl/rozrywka/2018/05/28/sztuczna-inteligencja-w-popkulturze> (data dostępu: 2021, Listopad 28).
- Gardner, H. (1983). *Frames of Mind: Theory of Multiple Intelligences*. New York, NY: Basic Books.
- Gartner. (2017a). *Augmented analytics is the future of data and analytics*. Gartner.
- Gartner. (2017b). *Gartner says more than 40 percent of data science tasks will be automated*. San Francisco: Gartner.
- Gelernter, H. (1959). Realization of a geometrytheorem proving machine. *Proc. an International Conference on Information Processing* (s. 273-280). Paryż: Unesco House.
- Gibbs, S. (2012). *Elon Musk: artificial intelligence is our biggest existential threat*. Pobrano z The Guardian: <https://www.theguardian.com/technology/2014/oct/27/elon-musk-artificial-intelligence-ai-biggest-existential-threat> (data dostępu: 2022, Kwiecień 22).
- Gödel, K. (1931). Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme. *I. Monatsh. f. Mathematik und Physik* 38, s. 173–198.
- Goertzel, B. i Pennachin, C. (2007). *Artificial General Intelligence*. Berlin: Springer.

Grigorescu, A., Baiasu, D. i Chitescu, R. (2020). *Business intelligence, the new managerial tool: Opportunities and limits*. Konstanca, Rumunia: Ovidius University Annals, Economic Sciences Series.

Grudin, J. (2012). A Moving Target—The Evolution of Human-Computer Interaction. W: J. A. Jacko, *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications*. New York, NY: CRC Press.

Gunia, A. (2015). Koncepcje wzmocnienia poznawczego. Próba definicji oraz przegląd metod. *AVANT*, 6(2), s. 35-56.

Haldane, J. (1924). *Daedalus; or Science and the Future*. Londyn.

Halevy, A., Norvig, P. i Pereira, F. (2009). The unreasonable effectiveness of data. *IEEE Intelligent Systems*, March/April, s. 8–12.

Han, X. i Boyden, E. (2007). Multiple-color optical activation, silencing, and desynchronization of neural activity, with single-spike temporal resolution. *PLoS One* 2(3): e299.

Haraway, D. (1985). A Cyborg Manifesto. *Socialist Review*, s. 149-181.

Hardt, S. i Lopes, L. C. (2015, Czerwiec 5). *Technical Disclosure Commons*. Pobrano z Google Votes: A Liquid Democracy Experiment on a Corporate Social Network: https://www.tdcommons.org/dpubs_series/79/ (data dostępu: 2022, Styczeń 13).

Harrison, E. F. (1995). *The Managerial Decision-making Process, 4th edition*. Boston, MA: Houghton Mifflin.

Harrison, F. E. (1982). *Policy, Strategy, and Managerial Action*. Boston, MA: Houghton Mifflin.

Harrison, F. E. (1989). The concept of strategic gap. *Journal of General Management*, Vol. 15 No. 2, s. 57-72.

Harrison, F. E. (1996). A process perspective on strategic decision making. *Management Decision* 34/1, s. 46-53.

Hastie, R. (2001). Problems for judgment and decision-making. *Annual Review of Psychology*, 52, s. 653-683.

Hatalaska, N. (2021). *Wiek Paradoksów*. Kraków: Znak.

Havenstein, H. (2005, Luty 14). *Spring comes to AI winter*. Pobrano z Computer World: <https://www.computerworld.com/article/2569567/spring-comes-to-ai-winter.html> (data dostępu: 2022, Lipiec 10).

- Hays, J. i Efron, A. A. (2007). Scene completion Using millions of photographs. *ACM Transactions on Graphics (SIGGRAPH)*, 26(3), s. 4-7.
- Hebb, D. (1949). *The Organization of Behavior*. New York, NY: Wiley.
- Hebb, D. (1980). *Essays on Mind*. Londyn: Lawrance Erlbaum.
- Henschen, D. (2018, Styczeń 10). *How ML and AI will transform business intelligence and analytics*. Pobrano z zdnet.com: <https://www.zdnet.com/article/how-machine-learning-and-artificial-intelligence-will-transform-business-intelligence-and-analytics/> (data dostępu: 2022, Kwiecień 22).
- Hensler, H. (2010, Styczeń 3). *Piraten Berlin*. Pobrano z: "Piratenpartei revolutioniert parteiinternen Diskurs: Interaktive Demokratie mit Liquid Feedback": <https://www.piratenpartei.berlin/pressemitteilung/pressemitteilung-piratenpartei-revolutioniert-parteiinternen-diskurs-interaktive-demokratie-mit-liquid-feedback/> (data dostępu: 2022, Marzec 9).
- Hetmański, M. (2000). Umysł a maszyny. Krytyka obliczeniowej teorii umysłu. *Studies in Logic Grammar and Rhetoric* 20 (30).
- Hetmański, M. (2005). Informacja jako kategoria filozofii umysłu. *Kognitywistyka. Problemy i perspektywy. Tom 1*, s. 11-24.
- Hewitt, C. (1976). *Viewing Control Structures as Patterns of Passing Messages*. Cambridge, MA: MIT AILab Memo 410.
- Hobbes, T. (1651). *Leviathan*. Baltimore: Penguin Books.
- Huffman, D. A. (1971). Impossible objects as nonsense sentences. In Meltzer, B. and Michie, D.(Eds.), *Machine Intelligence 6*, s. 295–324.
- Hutchins, E. (1995a). *Cognition in the Wild*. Cambridge, MA: MIT Press.
- Hutchins, E. (1995b). How a Cockpit Remembers Its Speeds. *Cognitive Science* 19/3, s. 265-288.
- Hutchins, E. (2010). Cognitive Ecology. *Topics in Cognitive Science* 4, s. 705-715.
- IEEE SPECTRUM. (2008, Czerwiec 1). <https://spectrum.ieee.org/who-is-who-in-the-singularity>. Pobrano z WHO'S WHO IN THE SINGULARITY: <https://spectrum.ieee.org/who-is-who-in-the-singularity> (data dostępu: 2021, Lipiec 3).
- IEEE SPECTRUM. (2008, Czerwiec 1). *Tech Luminaries Address Singularity Ten technologists share their views on the Singularity*. Pobrano z IEEE SPECTRUM: <https://spectrum.ieee.org/tech-luminaries-address-singularity#toggle-gdpr> (data dostępu: 2021, Sierpień 26).

- Ignatius, D. (2013, Listopad 1). More Chatter Than Needed. *Washington Post*.
- Ihde, D. (1997). *Technology and Lifeworld: From Garden to Earth*.
Bloomington/Indianapolis IN: Indiana University Press.
- Ihde, D. (2002). *Bodies in Technology*. Minneapolis, MN: University of
Minnesota Press.
- Imel, Z. E., Pace, B. T., Soma, C., Tanana, M., Hirsch, T., Gibson, J., Georgiou,
P., Narayanan, S., Atkins, D. (2019). Design feasibility of an automated, machine-
learning based feedback system for motivational interviewing. *Psychotherapy (Chic)*. Jun
2019, 318-328.
- Jain, T., Agarwal, M., Kumar, A., Verma, V. i Yadav, A. (2022).
Building machine learning application using oracle analytics cloud. *Data Engineering
for Smart Systems*, v. 238, s. 361-375.
- Jennings, H. S. (1906). *Behavior of the Lower Organisms*. New York, NY:
Columbia University Press.
- Jeppesen, L. B. i Lakhani, K. R. (2010). Marginality and Problem-Solving
Effectiveness in Broadcast Search. *Organization Science* 21, no. 5, s. 1016-1033.
- Johnson-Laird, P. (1988). *The Computer and the Mind*. Cambridge, MA:
The MIT Press.
- Karp, R. M. (1972). Reducibility among combinatorial problems. . *Miller R.E.,
Thatcher J.W., Bohlinger J.D. (eds) Complexity of Computer Computations. The IBM
Research Symposia Series*. (s. 85-103). Boston, MA.: Springer, Boston, MA.
- Kelly, K. (1997). *Why Things Bite Back*. New York, NY: Vintage Books.
- Kelly, K. (2010). *What Technology Wants*. New York, NY: Penguin Books.
- Kelly, K. (2016). *The Inevitable: Understanding 12 Technological Forces That
Will Shape Our Future*. New York, NY: Penguin Books.
- Kilgarriff, A. i Grefenstette, G. (2006). Introduction to the special issue
on the web as corpus. . *Computational Linguistics*, 29(3), 333–347.
- Knight, W. (2017, Październik 17). *You could become an AI master before you
know it. Here's how*. Pobrano z: MIT Technology Review:
[https://www.technologyreview.com/2017/10/17/148560/you-could-become-an-ai-master-
before-you-know-it-heres-how/](https://www.technologyreview.com/2017/10/17/148560/you-could-become-an-ai-master-before-you-know-it-heres-how/) (data dostępu: 2021, Wrzesień 21).
- Kobielus, J. (2017, Październik 26). *Even data scientists are facing AI takeover*.
Pobrano z: Infoworld: [https://www.infoworld.com/article/3234465/even-data-scientists-
are-facing-ai-automation.html](https://www.infoworld.com/article/3234465/even-data-scientists-are-facing-ai-automation.html) (data dostępu: 2021, Październik 1).

- Kolodner, J. (1983). Reconstructive memory: A computer model. *Cognitive Science*, 7, s. 281–328.
- Kopaliński, W. (2003). Słownik mitów i tradycji kultury. W: W. Kopaliński, *Słownik mitów i tradycji kultury* (s. 1295–1296). Warszawa: Oficyna Wydawnicza RYTM.
- Kossecki, J. (1974). *Cybernetyka kultury*. Warszawa: PIW.
- Kurzweil, R. (2005). *The Singularity is Near*. New York, NY: Viking.
- Laird, J., Newell, A. i Rosenbloom, P. S. (1987). SOAR: An architecture for general intelligence. *AIJ*, 33(1), s. 1–64.
- Laney, D. (2001). *3D Data Management: Controlling Data Volume, Velocity, and Variety*. Gartner.
- Latour, B. (1991). Technology is Society Made Durable. W: J. Law, *A Sociology of Monsters: Essays on Power, Technology and Domination* (s. 103-131). Londyn: Routledge.
- Latour, B. (2005). *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford: Oxford University Press.
- Legg, S. i Hutter, M. (2007). *A Collection of Definitions of Intelligence*. Manno, Szwajcaria: IDSIA.
- Lem, S. (2020). *Summa technologiae*. Kraków: Wydawnictwo Literackie.
- Licklider, J. C. (1960). Man-Computer Symbiosis. *IRE Transactions on Human Factors in Electronics*, vol. 1, s. 4-11.
- Lipschutz, S. (1967). *Schaum's outline of set theory and related topics*. New York, NY: McGraw-Hill Inc.
- Lohr, S. (2014, Sierpień 18). *For big-data scientists, 'janitor work' is key hurdle to insights*. Pobrano z: NY times: <https://www.nytimes.com/2014/08/18/technology/for-big-data-scientists-hurdle-to-insights-is-janitor-work.html> (data dostępu: 2021, Październik 5).
- Luce, R. D. i Raiffa, H. (1957). *Games and decisions: introduction and critical survey*. New York, NY: Wiley.
- Malone, T. (2018). *Superminds - The Surprising Power of People and Computers Thinking Together*. New York, NY: Little, Brown and Company.
- Malone, T. (2018). *Superminds: How hyperconnectivity is Changing the Way We Solve Problems*. Londyn: Oneworld Publications.

- Massachusetts Institute of Technology. (2018, Maj 4). *Malone - biografia*.
 Pobrano z: Malone - biografia: <https://cci.mit.edu/malone/> (data dostępu: 2021, Sierpień 16).
- Matlin, M. W. (1998). *Cognition (4th ed)*. Orlando, FL: Harcourt Brace College Publishers.
- McCarthy, J. (2007). From here to hman-level AI. *AIJ*, 171(18), s. 1174–1182.
- McCarthy, J. i Hayes, P. (1969). Some Philosophical Problems from the Standpoint of Artificial Intelligence. *Machine Intelligence 4*, s. 463-512.
- McCarthy, J., Minsky, M., Rochester, D. i Shannon, C. (2006). A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955. *AI Magazine*.
- McCulloch, W. S. i Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5, s. 115–137.
- McDermott, J. (1982). R1: A rule-based configurer. *Artificial Intelligence, Volume 19, Issue 1*, s. 39-88.
- McKinsey & Company . (2017). *JOBS LOST, JOBS GAINED: WORKFORCE TRANSITIONS IN A TIME OF AUTOMATION*. San Francisco: McKinsey & Company .
- McLuhan, M. (1964). *Understanding Media: The Extensions of Man*. Cambridge, MA: The MIT Press.
- Merriam Webster. (2022, 10 11). *Merriam Webster dictionary*.
 Pobrano z Merriam Webster supermind definition: <https://www.merriam-webster.com/dictionary/supermind> (data dostępu: 2022, Listopad 2).
- Mill, J. S. (1863). *Utilitarianism*. London: Parker, Son and Bourn.
- Miłkowski, M. (2009). O tzw. metaforze komputerowej. *Analiza i Egzystencja*, 9, s. 163-185.
- Minsky, M. (1968). *Semantic information processing*. Cambridge, MA: MIT Press.
- Minsky, M. (1974). *A Framework for Representing Knowledge*. Cambridge, MA: MIT AI Lab.
- Minsky, M. (1975). A framework for representing knowledge. (P. H. Winston, Red.) *The Psychology of Computer Vision*, s. 211–277.
- Minsky, M. (1977). Plain Talk about Neurodevelopmental Epistemology. *IJCAI'77: Proceedings of the 5th international joint conference on Artificial intelligence - Volume 2* (s. 1083-1092). Cambridge, MA: Morgan Kaufmann Publishers Inc.

- Minsky, M. (1980). *Jokes and the Cognitive Unconscious*. Cambridge, MA: MIT AI Lab Memo 603.
- Minsky, M. (1980). K-lines: A Theory of Memory. *Cognitive Science, Vol. 4.*, s. 117-133.
- Minsky, M. (1986). *Society of Mind*. New York, NY: Simon & Schuster, Inc.
- Minsky, M. (1990). Logical vs. Analogical or Symbolic vs. Connectionist or Neat vs. Scruffy. *Artificial Intelligence at MIT, Expanding Frontiers*.
- Minsky, M. (1991). Society of mind: A response to four reviews. *Artificial Intelligence Volume 48, Issue 3*, s. 371-396.
- Minsky, M. (1994). Negative Expertise. *International Journal of Expert Systems, vol. 7*, 13-19.
- Minsky, M. (2007). *The Emotion Machine: Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind*. New York, NY: Simon and Schuster.
- Minsky, M. L. i Papert, S. (1969). *Perceptrons: An Introduction to Computational Geometry (first edition)*. Cambridge, MA: MIT Press.
- Minsky, M., Singh, P. i Sloman, A. (2004). The st. thomas common sense symposium: Designing architectures for human-level intelligence. *AIMag, 25(2)*, s. 113–124.
- Mitcham, C. (1994). *Thinking through Technology: The Path between Engineering and Philosophy*. Chicago, Illinois, Stany Zjednoczone: University of Chicago.
- Moravec, H. (1998). *Robot: Mere Machine to Transcendent Mind*. Oxford: Oxford University Press.
- Moravec, H. (1999). *The Universal Robot*. Cambridge, MA: The MIT Press.
- More, M. i Vita-More, N. (2013). *The Transhumanist Reader: Classical and Contemporary Essays on the Science, Technology, and Philosophy of the Human Future*. John Wiley & Sons, Inc.
- Nagar, Y. i Malone, T. W. (2011). Making Business Predictions by Combining Human and Machine Intelligence in Prediction Markets. *Proceedings of International Conference on Information Systems*. Shanghai.
- Nalwa, V. S. (1994). *A guided tour of computer vision*. Boston, MA: Addison-Wesley Longman Publishing Co., Inc.
- Needham, J. (1986). *Science and Civilization in China: Volume 2*. Taipei: Caves Books Ltd.

- Newell, A. (1980). Physical Symbol Systems. *Cognitive Science* 4, s. 135-183.
- Newell, A. (1990). *Unified theories of cognition*. . Cambridge, MA: Harvard University Press.
- Nilsson, N. (1995). Eye on the prize. *AIMag*, 16(2), s. 9–17.
- Nilsson, N. (2005). Human-level artificial intelligence? be serious! *AIMag*, 26(4), s. 68–75.
- North, J. (1954). *The rational behavior of mechanically extended man*. Wolverhampton: Boulton Paul Aircraft Ltd.
- O'Connor, K. M. (1994). *The alchemical creation of life (takwin) and other concepts of Genesis in medieval Islam [Rozprawa doktorska, University of Pennsylvania]*. Pennsylvania.
- Ofstad, H. (1961). *An Inquiry into the Freedom of Decision*. Norwegian Universities Press.
- Ogawa, S., Lee, T.-M., Kay, A. R. i Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. . *PNAS*, 87., s. 9868–9872.
- Pacey, A. (2001). *Meaning in Technology*. Cambridge, MA: The MIT Press.
- Parker, K., Menasce Horowitz, J. i Minkin, R. (2022, Luty 16). *Pew Research Center*. Pobrano z COVID-19 Pandemic Continues To Reshape Work in America: <https://www.pewresearch.org/social-trends/2022/02/16/covid-19-pandemic-continues-to-reshape-work-in-america/> (data dostępu: 2022, Marzec 20).
- Payne, D. G. i Wenger, M. J. (1998). *Cognitive Psychology*. New York, NY: Houghton Mifflin Co.
- Penrose, R. (1989). *The Emperor's New Mind*. Oxford: Oxford University Press.
- Pfeifer, C. i Scheier, R. (2001). *Understanding Intelligence*. Cambridge, MA: The MIT Press.
- Pickover, C. (2009). *The Math Book: From Pythagoras to the 57th Dimension, 250 Milestones in the History of Mathematics*. Sterling Publishing Company, Inc.
- Pinel, J. P. (1997). *Biopsychology (3rd ed)*. Needham Heights, MA: Allyn and Bacon.
- Postman, N. (1993). *Technopoly: The Surrender of Culture to Technology*. New York, NY: Alfred A. Knopf.
- Prat, N. (2019). Augmented Analytics. *Business & Information Systems Engineering: Vol. 61: Iss. 3.*, s. 375-380.

- Przegalińska, A. (2016). *Istoty wirtualne*. Kraków: Towarzystwo Autorów i Wydawców Prac Naukowych Universitas.
- Przegalińska, A. (2020). *Sztuczna Inteligencja. Nieludzka, arcyludzka*. Kraków: Znak.
- Przybysz, P. (2009). *Modele teoretyczne we współczesnej filozofii politycznej liberalizmu*. Poznań: Wydawnictwo Naukowe Wydziału Nauk Społecznych Uniwersytetu Im. Adama Mickiewicza w Poznaniu.
- Pylyshyn, Z. (1984). *Computation and Cognition: Toward a Foundation for Cognitive Science*. Cambridge MA: The MIT Press.
- Pylyshyn, Z. (2007). *Things and Places - How the Mind Connects with the World*. Cambridge, MA: The MIT Press.
- Ramsey, F. P. (1926). Truth and Probability. *The Foundations of Mathematics and other Logical Essays*, s. 156-198.
- Rieger, C. (1976). An organization of knowledge for problem solving and language comprehension. *AIJ*, 7, s. 89–127.
- Rosenblatt, F. (1962). *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*. Washington DC: Spartan Books.
- Rosenblueth, A., Wiener, N. i Bigelow, J. (1943). Behavior, purpose, and teleology. *Philosophy of Science*, 10, s. 18–24.
- Ruhe, G. (2003). Software engineering decision support - Methodologies and applications. *Innovations in decision support systems*, 3, s. 143-174.
- Ruhe, G. i An, N. T. (2004). Hybrid intelligence in software release planning. *International Journal of Hybrid Intelligent Systems*, 1 (2), s. 99-110.
- Rumelhart, D. E. i McClelland, J. L. (1986). *Parallel Distributed Processing*. Cambridge, MA: MIT Press.
- Russell, S. (2019). *Human Compatible: Artificial intelligence and the problem of control*. New York, NY: Viking.
- Russell, S. i Norvig, P. (2010). *Artificial Intelligence: A modern approach (Third Edition)*. Upper Saddle River, NJ: Prentice Hall.
- S. E. Fahlman. (1974). A planning system for robot construction tasks. *Artificial intelligence*, vol. 5, no. 1, s. 1–49.
- Sætnan, A. R., Schneider, I. i Green, N. (2018). *The Politics and Policies of Big Data: Big Data, Big Brother?* Londyn, UK: Taylor & Francis Ltd.

- Sankar, S. (2013). *Shyam Sankar: The rise of human-computer cooperation*.
Pobrano z: Youtube:
https://www.youtube.com/watch?v=ltelQ3iKybU&ab_channel=TED
(data dostępu: 2022, Styczeń 8).
- SAS. (2016). *Managing the analytical life cycle for decisions at scale*. Pobrano z:
SAS: https://www.sas.com/content/dam/SAS/en_us/doc/whitepaper1/manage-analytical-life-cycle-continuous-innovation-106179.pdf (data dostępu: 2022, Luty 17).
- Schank, R. C. i Abelson, R. P. (1977). *Scripts, Plans, Goals, and Understanding: An Inquiry Into Human Knowledge Structures*. New Jersey: Psychology Press.
- Schank, R. C. i Riesbeck, C. (1981). *Inside Computer Understanding: Five Programs Plus Miniatures*. New Jersey, : Lawrence Erlbaum Associates.
- Schumpeter, J. A. (1962). *Capitalism, socialism and democracy*.
New York, NY: Harper.
- Searle, J. (1980). Minds, Brains and Programs. *Behavioral and Brain Sciences*, s. 417-457.
- Searle, J. R. (1992). *The Rediscovery of the Mind*. Cambridge, MA: MIT Press.
- Seddon, P., Constantinidis, D., Tamm, T. i Dod, H. (2017). How does business analytics contribute to business value? *Information System Journal* 27(3), s. 237-269.
- Shirley, R. (1982). Limiting the scope of strategy: a decision-based approach. *Academy of Management Review*, s. 264-265.
- Silva, J. R., de Sá, N. i Lima Jr., W. (2017). Licklider's fundamental ideas in 'Man-computer symbiosis' reemerge in the concepts of Cognitive Computing: a bibliometric study. *Proc. of the XI International Brazilian Meeting in Cognitive Science*. SÃO PAULO, Brazil.
- Simon, H. (1957). "Heuristic Problem Solving: The Next Advance Operations Research". *Operations Research*, s. 1-10.
- Simon, H. (1960). *The New Science of Management Decision*. New York, NY: Harper & Row, New York.
- Simon, H. A. (1947). *Administrative behavior; a study of decision-making processes in administrative organization*. New York, NY: Macmillan.
- Singh, P. (2003). Examining the society of mind. *Computing and Informatics* 22(6), s. 521-543.
- Smedes, T. A. (2005). *Being Cyborgs On Creating Humanity in a Created World of Technology, Values and Ethical Issues in Theology*. Pobrano z Academia.edu:

https://www.academia.edu/950128/Being_cyborgs_on_creating_humanity_in_a_created_world_of_technology (data dostępu: 2022, Czerwiec 25).

Smolensky, P. (1988). On the proper treatment of connectionism. *BBS*, 2, s. 1–74.

Solomonoff, R. (1964a). A Formal Theory of Inductive Inference. Part I. *Information and Control*, 7, s. 1-22.

Solomonoff, R. (1964b). A Formal Theory of Inductive Inference. Part II. *Information and Control*, 7, s. 224-254.

Spearman, C. (1904). “General Intelligence,” Objectively Determined and Measured. . *The American Journal of Psychology*, 15(2), s. 201–292.

Storey, V. i Song, I. (2017). Big data technologies and management: What conceptual modeling can do. *Data Knowledge Engineering* 108, s. 50-67.

Surowiecki, J. (2004). *The Wisdom of Crowds*. New York, NY: Doubleday.

Tableau. (2022, 07 25). *Automatically Build Views with Ask Data*. Pobrano z Pomoc techniczna firmy Tableau: https://help.tableau.com/current/pro/desktop/en-us/ask_data.html (data dostępu: 2022, Listopad 7).

Tarski, A. (1941). *Introduction to Logic and to the Methodology of Deductive Sciences*. Dover: Dover Publications.

Trzęsicki, K. (2020). Idea Sztucznej Inteligencji. *Filozofia i Nauka: Studia filozoficzne i interdyscyplinarne. Tom 8, cz. 1, 2020*, s. 69-96.

Turing, A. (1937). On Computable Numbers, with an Application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, s. 230-265.

Turkle, S. (1995). *Life on the Screen - Identity in the Age of the Internet*. New York, NY: Simon & Schuster.

Turkle, S. (2015). *Reclaiming Conversation: The Power of Talk in the Digital Age*. New York, NY: Penguin Press.

Uniwersytet Helsiński. (2018). *Elements of AI*. Pobrano z Elements of AI - Wersja polska : <https://www.elementsofai.pl/> (data dostępu: 2022, Sierpień 29).

Von Neumann, J. i Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton: Princeton University Press.

Wald, A. (1950). *Statistical decision functions*. New York, NY: Wiley.

Waltz, D. (1975). Understanding line drawings of scenes with shadows. *Winston, P. H. (Ed.), The Psychology of Computer Vision. McGraw-Hill*, s. 19-92.

Wang, Y. (2007a). Software engineering foundations: A software science perspective. *CRC Software Engineering Series, Vol. II/III*.

- Wang, Y. (2007b). The theoretical framework of cognitive informatics. *The International Journal of Cognitive Informatics and Natural Intelligence (IJCINI)*, 1(1), s. 1-27.
- Wang, Y. i Ruhe, G. (2007, Kwiecień). The Cognitive Process of Decision Making. *Journal of Cognitive Informatics and Natural Intelligence*, 1(2), s. 73-85.
- Wang, Y. i Wang, Y. (2004). Cognitive Informatics models of the brain. *IEEE Transactions on Systems, Man and Cybernetics (C)*, 36(2), s. 203-207.
- Wang, Y., Liu, D. i Wang, Y. (2003). Discovering the capacity of human memory. *Brain and Mind: A Transdisciplinary Journal of Neuroscience and Neurophilosophy*, 4(2), s. 189-198.
- Wang, Y., Wang, Y. i Patel, D. (2004). A layered reference model of the brain (LRMB). *IEEE Transactions on Systems, Man, and Cybernetics (C)*, 36 (2), s. 124-133.
- Watson, H. (2017). The cognitive decision-support generation. *Business Intelligence J* 22(2), s. 5-14.
- Webb, M. (2019, Listopad 6). The Impact of Artificial Intelligence on the Labor Market. *SSRN*.
- Weizenbaum, J. (2002). *Mýtus počítače: počítačový pohled na svět*. Praga: Moraviapress.
- Wellman, P. (1995). The economic approach to artificial intelligence. *ACM Computing Surveys; Volume 27; Issue 3*, s. 360-362.
- West, D. i Allen, J. (2018, Kwiecień 28). *How artificial intelligence is transforming the world*. Pobrano z Brookings: <https://www.brookings.edu/research/how-artificial-intelligence-is-transforming-the-world/> (data dostępu: 2022, Październik 29).
- Widrow, B. (1962). Generalization and information storage in networks of adaline "neurons". *SelfOrganizing Systems 1962*, s. 435–461.
- Widrow, B. i Hoff, M. E. (1960). Adaptive switching circuits. *1960 IRE WESCON Convention Record*, s. 96–104.
- Wielki słownik języka polskiego*. (2022, Czerwiec 3). Pobrano z: <https://wsjp.pl/haslo/podglad/5223/mocny/5071507/material> (data dostępu: 2022, Październik 31).
- Wiener, N. (1948). *Cybernetics. Or control and communication in the animal and the machine*. New York, NY: The Technology Press; John Wiley & Sons, Inc.
- Wikipedia - Neuralink*. (2022, 06 08). Pobrano z: <https://pl.wikipedia.org/wiki/Neuralink> (data dostępu: 2021, Listopad 19).

Wikipedia. (2020, Grudzień 21). "Data is the new oil". Pobrano z: Data is new oil: https://en.wikibooks.org/wiki/Lentis/%22Data_is_the_new_oil%22 (data dostępu: 2021, Wrzesień 15).

Wikipedia. (2023, 01 05). *Wikipedia*. Pobrano z: Four day workweek: https://en.wikipedia.org/wiki/Four-day_workweek (data dostępu: 2022, Grudzień 3).

Wilensky, R. (1978). *Understanding goal-based stories*. New Haven, Connecticut: Yale University.

Wilson, R. i Keil, F. C. (2001). *The MIT Encyclopedia of the Cognitive Sciences*. Cambridge, MA: MIT Press.

Winograd, S. i Cowan, J. D. (1963). *Reliable Computation in the Presence of Noise*. Cambridge, MA: MIT Press.

Winograd, T. (1972). Understanding natural language. *Cognitive Psychology*, 3(1), s. 1–191.

Winston, P. H. (1970). Learning structural descriptions from examples. . *Technical report MAC-TR-76, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology.*, N/A.

Wong, S., Lin, I., Komarneni, J. i Nundy, S. (2017, Październik 22). *Machine Classifier Trained on Low-Volume, Structured Data Predicts Diagnoses Near Physician-Level: Chest Pain Case Study*. Pobrano z : 39th Annual Meeting of the Society for Medical Decision Making: <https://smdm.confex.com/smdm/2017/meetingapp.cgi/Paper/11058> (data dostępu: 2022, Styczeń 13).

Wooley, A. W., Chabris, C. F., Pentland, A., Hashmi, N. i Malone, T. W. (2010, Październik 29). Evidence for Group Intelligence Factor in the Performance of Human Groups. *Science* 330, no. 6004, s. 686-688.

Wright, I. (2000). The society of mind requires an economy of mind. *Proceedings AISB00 Symposium Starting from Society - the Application of Social Analogies to Computational Systems* (s. 113–124). Birmingham, UK.: AISB.

Wróbel, S. (1997). *The Discovery of the Unconscious or Destruction of the Cartesian Concept of Mind?* Warszawa / Wrocław: Fundacja na Rzecz Nauki Polskiej.

Yarowsky, D. (1995). Unsupervised word sense disambiguation rivaling supervised methods. *ACL95*, s. 189–196.

Žáčková, E. (2011, 01). Man-computer Symbiosis as a Way of Human Cognitive Enhancement. *Masaryk University Journal of Law and Technology*, s. 133-144.

Žáčková, E. (2011). Man-computer Symbiosis as a Way of Human Cognitive Enhancement . *Masaryk University Journal of Law and Technology*, 133-144.

Zilca, R. (2019, Październik 6). Technology That Helps Us Help Ourselves: Talk Therapy in Combination with Augmented Intelligence. Raanana, Dystrykt Centralny, Izrael.

Spis rysunków:

Rysunek 1 Ewolucja cytatów Licklidera (Silva, de Sá i Lima Jr, 2017).....	7
Rysunek 2 Schemat procesów myślowych (Malone, 2018, s. 82).....	99