

WIESŁAW BANACH

Uniwersytet im. Adama Mickiewicza w Poznaniu
Instytut Kulturoznawstwa
e-mail: banach@amu.edu.pl
<https://orcid.org/0000-0002-3411-9099>
<https://doi.org/10.14746/h.2025.2.7>

Etyka biznesu wobec rozwoju nowoczesnych technologii

Challenges of business ethics in the face of the development of modern technologies

Abstract. *The aim of this article is to attempt to answer a very current issue concerning contemporary challenges of business ethics in the face of the development of modern technologies. In particular, the dynamic development of artificial intelligence is unprecedented in history and therefore requires special reflection in the field of sociology, cultural studies, philosophy and ethics. On the one hand, the positives resulting from the effective use of AI in subsequent areas of life are obvious – on the other hand, the signs of a lack of reflection on the civilizational effects of the ubiquitous technological expansion based on artificial intelligence may be disturbing. This deficit is the result of the characteristic of engineers and uncritical supporters of modern technologies: technoenthusiasm, “tunnel thinking” and focusing on a specialized field of knowledge. This causes them to perceive everything separately and thus ignore signals about emerging threats. The challenge for business ethics may be not only to indicate potential dangers, but also to develop appropriate ethical standards, good practices and procedures that will define the principles of our functioning in a reality dominated by modern technologies.*

Keywords: *ethics, modern technologies, progress, artificial intelligence, threats*

Słowa kluczowe: *etyka, nowoczesne technologie, postęp, sztuczna inteligencja, zagrożenia*



Wprowadzenie

C hińskie powiedzenie „Obyś żył w ciekawych czasach” zazwyczaj traktowane jest nie tyle jako życzenie, ile raczej jako przekleństwo. Taka jego interpretacja znajduje uzasadnienie zwłaszcza w odniesieniu do współczesności. Bez wątpienia bowiem czasy, w których przyszło nam żyć, stanowią okres gwałtownych przemian społecznych, kulturowych, a nawet cywilizacyjnych. Tempo i dynamika wszechobecnego przeobrażenia jest tak duże, że wypracowane od wieków tradycyjne mechanizmy adaptacji do nich stają się zawodne. Spośród czynników mających największy na to wpływ wymieniamy zazwyczaj postępującą globalizację, konwergencję kulturowych wzorów, reguł czy praktyk, migracje ludności i związane z nimi niekontrolowane rozprzestrzenianie się idei i pomysłów na życie, erozję czy wręcz rozpad kulturowych i społecznych granic. Jednak wszystkie te – niezmiernie ważne i powszechnie występujące – czynniki są w mniejszym lub większym stopniu jedynie pochodną podstawowej determinanty, jaką stanowi współczesny rozwój nowoczesnych technologii. To właśnie on jest „odpowiedzialny” za zakres, głębokość, a przede wszystkim gwałtowny charakter wspomnianych przemian. Bez rewolucji technologicznej trudno byłoby mówić o rewolucyjnym charakterze tych zmian. Mówiąc krótko, żyjemy w czasach, w których „świat przyspieszył”, a głównym akceleratorem tego dynamicznego procesu stał się niekontrolowany rozwój nowoczesnych technologii.

Rodzi się pytanie, na ile obecna sytuacja związana z rozwojem nowoczesnych technologii ma charakter bezprecedensowy? Mówiąc wprost: czy rewolucyjne zmiany technologiczne – poza dynamiką i stopniem złożoności – różnią się w jakiś istotny sposób od tych z przeszłości? Szukając odpowiedzi na to pytanie, warto uświadomić sobie, że każda znacząca rewolucja technologiczna postrzegana była przez ludzi przede wszystkim jako szansa na polepszenie warunków życia i dokonanie cywilizacyjnego skoku. Ułatwiała ona człowiekowi pracę, a wielokrotnie nawet zastępowała go w niej. Wydaje się, że oryginalność współczesnej rewolucji technologicznej (zwłaszcza w przypadku rozwiązań opartych na sztucznej inteligencji¹) nie wynika z możliwości wspomagania czy zastępowania człowieka w pracy lub innych sytuacjach życia codziennego, ale z jej potencjału zastąpienia ludzkiego umysłu jako takiego. Innymi słowy, osiągnęliśmy poziom rozwoju

¹ Pojęcie nowoczesnych technologii jest bardzo szerokie i dotyczy wielu specjalistycznych dziedzin wiedzy z zakresu nauk ścisłych, która to wiedza została wykorzystana w różnych sferach ludzkiego życia (m.in. w medycynie, komunikacji, transporcie, edukacji, biznesie itp.). W niniejszym artykule skoncentrowano się głównie na tych rozwiązaniach technologicznych, które oparte zostały na działaniach sztucznej inteligencji. Jest tak dlatego, że gwałtowny rozwój tej ostatniej nie ma precedensu w historii ludzkości i w związku z tym refleksja etyczna nad jej funkcjonowaniem wymaga szczególnej uwagi oraz stanowi pionierskie wyzwanie.

technologicznego, który przewyższa intelektualne zdolności człowieka i tym samym umożliwia zastąpienie go w kierowniczej roli nadzorcy systemu. Tego typu twierdzenia stanowią nie elementy jakichś futurystycznych wizji, ale stają się częścią współczesnego dyskursu naukowego, świadczącego o potencjale sztucznej inteligencji, a zwłaszcza granicach tzw. sztucznej superinteligencji².

Dochodzimy w tym momencie do tytułowego zagadnienia wyzwań, przed którymi stoi etyka biznesu w obliczu rozwoju nowoczesnych technologii. Wydaje się, że panujący powszechnie optymizm, a nawet entuzjazm związany z możliwościami, jakie daje sztuczna superinteligencja, sprawia, iż trudno jest – zwłaszcza osobom zaangażowanym w prace nad nią – dostrzec zagrożenia wynikające z jej wdrażania. W związku z tym na wyższym, ogólnym poziomie rozważań pojawia się przestrzeń dla refleksji filozoficznej, w której m.in. etycy biznesu mogliby wskazywać na ewentualne, niepożądane skutki rozwoju sztucznej superinteligencji; refleksji, która mogłaby stanowić przestrożę dla ludzkości przed samounicestwieniem związanym z wymknięciem się szeroko rozumianej sztucznej inteligencji spod kontroli człowieka. Natomiast w wymiarze mniej ogólnym, tzn. związanym z już realizowanymi, skutecznymi próbami zastosowania sztucznej inteligencji w konkretnych rozwiązaniach technologicznych (np. ChatGPT, chatboty, generatory obrazu, programy analityczne, zarządzane przez sztuczną inteligencję narzędzia internetowe, inteligentna nawigacja, auta kierowane przez AI, Internet Rzeczy itd.) etyka biznesu może posłużyć do rozstrzygania różnego rodzaju dylematów, które pojawiają się wraz z ich stosowaniem.

Artykuł stanowi próbę wskazania pól tematycznych, które – ze względu na rozwój nowoczesnych technologii – stanowić mogą największe wyzwania dla etyki biznesu. Dla jasności wywodu pola te zostały wydzielone zgodnie ze wskazanym powyżej stopniem ogólności, na którym może być prowadzona refleksja etyczna na ten temat. Po wprowadzającej części pierwszej, która zawiera ogólną charakterystykę ambiwalencji, wynikających z postępu technologicznego, wskazane zostały w części drugiej zagrożenia związane zwłaszcza z dynamicznym rozwojem

² Por. A. Zybertowicz, *AI Eksploracja*. Andrzej Zybertowicz odpowiada Katarzynie Zybertowicz, Zona Zero, Warszawa 2024: „ASI (Artificial Superintelligence): sztuczna superinteligencja; teoretycznie wyobrażony poziom sztucznej inteligencji, który przewyższa ludzką inteligencję we wszystkich dziedzinach, w tym w kreatywności, rozwiązywaniu problemów, podejmowaniu decyzji i ogólnej mądrości. Jest także w stanie samodzielnie uczyć się i rozwijać, osiągając poziomy rozumowania, które są nieosiągalne dla ludzi. Chociaż ASI jest obecnie tylko koncepcją teoretyczną, to pole badań przyciąga uwagę wielu naukowców” (*Ibidem*, s. 22). Ze względu na nowatorski charakter tych zagadnień istnieje wiele pojęć i określeń, pod którymi różni badacze rozumieją podobne lub tożsame zjawiska czy procesy. Dlatego też, jako synonimiczne do powyższej definicji sztucznej superinteligencji, używa się określenia „ogólna sztuczna inteligencja” lub „generalna sztuczna inteligencja”, w skrócie AGI (Artificial General Intelligence). Wszystkie te pojęcia służą do określenia takiego systemu sztucznej inteligencji (AI), który nie jest projektowany do zastosowania w konkretnym rozwiązaniu technologicznym, ale zdolny jest (a raczej będzie) rozwiązać teoretycznie każdy problem intelektualny postawiony mu przez człowieka.

sztucznej inteligencji. Otóż pomimo licznych korzyści wynikających z postępującej obecności nowoczesnych technologii we współczesnym świecie należy dostrzec także koszty i zagrożenia będące skutkiem tego procesu. Tematem części trzeciej jest omówienie terminów i pojęć, którymi określani są bezkrytyczni entuzjaści nowoczesnych technologii. Ich zawężony sposób postrzegania rzeczywistości może stanowić w niedalekiej przyszłości źródło największych niebezpieczeństw, nawet o charakterze globalnym. W części czwartej przedstawione zostały bardziej konkretne sytuacje i przykłady ilustrujące różnego rodzaju dylematy etyczne, wynikające ze stosowania w życiu codziennym nowoczesnych technologii opartych na sztucznej inteligencji. W podsumowaniu zebrane zostały wnioski wypływające z przedstawionych w poszczególnych rozdziałach uwag. Posłużyły one jako punkt wyjścia do filozoficznej i kulturoznawczej refleksji na temat wyzwań, które stoją przed etyką biznesu w świecie zdominowanym przez nowoczesne technologie, zwłaszcza te, które oparte są na sztucznej inteligencji.

Na koniec słowa wstępnego warto poczynić uwagę dotyczącą dwóch kwestii. Otóż ze względu na charakterystykę przedmiotu badań, czyli specyfikę bezprecedensowych, dynamicznych zmian w zakresie projektowania i wdrażania nowoczesnych technologii, wiele zawartych w niniejszym artykule twierdzeń nie jest wygłaszana w sposób kategoriyczny, ale ma często tryb warunkowy lub po prostu formę pytającą. Jest tak dlatego, że to, co dziś uchodzi za powszechnie obowiązujące, w wyniku pojawienia się przełomowego rozwiązania technologicznego za kilka miesięcy może się stać po prostu nieaktualne. Dlatego też – i jest to drugie zastrzeżenie – wykorzystana literatura w pewnym stopniu pochodzi ze źródeł internetowych, mających znacznie szybszy proces wydawniczy, niż to mam miejsce w przypadku publikacji papierowych.

1. Ambiwalencje postępu technologicznego

Genezy dynamicznego rozwoju nowoczesnych technologii należy szukać w idei postępu obecnej na gruncie zachodniej cywilizacji od końca XVIII wieku. Towarzystwo jej kult nauki i wiara w nieograniczone możliwości ludzkiego rozumu nie tylko stymulowały, ale wręcz wymuszały poszukiwanie nowych rozwiązań technologicznych, innowacyjność i wynalazczość. W sposób niezaprzeczalny przyczyniły się one do poprawy życia dużej części ludzkości, rozwiązując jej odwieczne problemy, takie jak: zarazy, choroby, głód. Wiele przekonujących danych o charakterze statystycznym dostarcza na ten temat książka Stevena Pinkera pt. *Nowe Oświecenie. Argumenty za rozumem, nauką, humanizmem i postępem*³. Jest

³ S. Pinker, *Nowe Oświecenie. Argumenty za rozumem, nauką, humanizmem i postępem*, tłum. T. Bieroń, Zysk i S-ka, Poznań 2018. Mimo pewnych zastrzeżeń i wątpliwości metodologicznych (dotyczących głównie rzetelności i wiarygodności danych statystycznych sprzed dwóch, a nawet

jednak wielu autorów, którzy – nie kwestionując osiągnięć samej nauki i korzyści wynikających ze stosowania nowoczesnych technologii – podkreślają ich koszty. Co więcej, coraz częściej wskazuje się na zagrożenia, a nawet pułapki, w które wpaść może ludzkość bezrefleksyjnie przyjmująca dobrodziejstwa wynikające z rozwoju technologii. Niemal cała książka Ronalda Wrighta pt. *Krótką historią postępu* poświęcona jest omówieniu przykładów oraz mechanizmów powstawania tzw. pułapek postępu. Dochodziło do nich od zarania dziejów, a ich występowanie wpisane było w rozwój ludzkiej cywilizacji. Wraz ze wzrostem złożoności systemów społecznych i kulturowych, a zwłaszcza z – idącym z nim w parze – rozwojem technologii wzrastało zagrożenie samozagłady. Niebezpieczeństwo było tym większe, im bardziej technologie „potrafiły uwieść całe społeczności”. Nieuzasadniony optymizm, a nawet entuzjazm, połączone z brakiem refleksji i pokusą „pójścia na skróty”, niejednokrotnie przynosiły katastrofalne skutki. Jak zauważa autor: „Paleolityczni myśliwi, którzy nauczyli się zabijać dwa mamuty zamiast jednego, dokonali postępu. Ci, którzy wpadli na pomysł, by zagnać całe stado na krawędź urwiska i uśmiercić dwieście mamutów naraz, posunęli się zbyt daleko. Żyli na wysokim poziomie... przez chwilę. Potem nastał głód”⁴. Współcześnie, gdy funkcjonujemy w ramach zglobalizowanego, połączonego kanałami szybkiej i nieograniczonej komunikacji świata zagrożenie wpadnięcia w „pułapkę postępu” jest tym bardziej realne. Stopień uzależnienia naszego życia od technologii jest bardzo wysoki i cały czas rośnie. Już nie tylko wielkie projekty naukowe i badawcze, ale nawet prozaiczne i rutynowe czynności wykonywane przez nas każdego dnia oparte są na zaawansowanych technologiach. Niepożądane skutki awarii systemu mogą przybrać wymiar globalny, wywołując trudne do przewidzenia konsekwencje.

2. Czy oparte na sztucznej inteligencji technologie zagrażają ludzkości?

Pojawia się pytanie, czy w przypadku obserwowanego współcześnie gwałtownego rozwoju nowoczesnych technologii, a zwłaszcza sztucznej inteligencji, możemy

trzech stuleci, które posłużyły autorowi do sporządzania zestawień i wykresów, wskazujących na postęp w wielu dziedzinach życia codziennego) książka ta jest godna polecenia. Dzięki odwołaniu się do mówiących „same za siebie” danych ilościowych i faktów udało się Pinkerowi uniknąć obecnej w wielu apologetycznych pracach ideologizacji postępu.

⁴ R. Wright, *Krótką historią postępu*, tłum. Ł. Wierzbicki, Pogotowie Kazikowe, Borówiec 2021, s. 35. Autor tak opisuje sposób, w jaki rozumie przywołane pojęcie „pułapki postępu”: „Gdy kultury stają się bardziej złożone, a technologie potężniejsze, potrafią stać się do tego stopnia ociężałe wyspecjalizowane, wrażliwe, że w skrajnych wypadkach mogą być skazane na zagładę. Bomba atomowa, logiczna kontynuacja strzały i pocisku, była pierwszym efektem rozwoju technologii zdolnej zagrozić naszemu gatunkowi wyginięciem. To właśnie nazywam »pułapką postępu«. W przeszłości, nawet w epoce kamienia łupanego, dużo prostsze technologie potrafiły uwieść całe społeczności, a potem im zagrozić”. *Ibidem*, s. 60.

mieć obawy o opisane wyżej koszty i zagrożenia z nim związane? Wydaje się, że jest to uzasadnione przynajmniej z kilku powodów. Po pierwsze, marzenia i aspiracje rodzaju ludzkiego, by stworzyć sztuczną inteligencję, przerastającą swymi możliwościami nie tylko pojedynczego człowieka (bo to odnotowano już dość dawno temu)⁵, ale cały rodzaj ludzki, staną się wkrótce rzeczywistością. Wydaje się, iż póki co, mamy jeszcze kontrolę nad rozwojem AI, ale świadomi jesteśmy tego, że posiada ona zdolność samouczenia się, omijania wprowadzonych przez programistów zabezpieczeń i ograniczeń, a także potrafi znajdować i wdrażać „własne pomysły” na rozwiązanie zadanego problemu⁶. Trudno mówić o konkretnych terminach, ale być może bliscy jesteśmy momentu, w którym skuteczną metodą zapanowania nad sztuczną inteligencją będzie... wyłączenie wtyczki z gniazdka. W tym kontekście szczególnie ważne wydają się głosy tych naukowców, którzy podnoszą kwestię naszego bezpieczeństwa. Wskazują oni na to, że przybliżyliśmy się w rozwoju technologicznym opartym na sztucznej inteligencji do momentu, w którym zagrożone zostało istnienie naszego gatunku. Zwraca na to uwagę w swej najnowszej książce *Nexus. Krótka historia informacji. Od epoki kamienia do sztucznej inteligencji* Yuval Noah Harari. Przywołuje on nazwiska znaczących postaci, które w tej kwestii wręcz biją na alarm: „Nie tylko filozofowie i socjolodzy, ale też wielu czołowych ekspertów w dziedzinie sztucznej inteligencji i przedsiębiorców, takich jak Yoshua Bengio, Geoffrey Hinton, Sam Altman, Elon Musk, i Mustafa Suleyman, ostrzegają opinię publiczną, że sztuczna inteligencja może zniszczyć naszą cywilizację”⁷.

W podobnym tonie wypowiada się w swoich książkach na temat rozwoju sztucznej inteligencji Andrzej Zybertowicz. W ostatniej z nich, *AI Eksploracja*, wskazuje na możliwe konsekwencje bezrefleksyjnego wdrażania opartych na AI technologiach i apeluje: „Wiele osób już dostrzega, że AI – nawet gdyby nie osiągnęła poziomu AGI czy ASI [Artificial General Intelligence czy Artificial Superintelligence – W.B.] – zapewne okaże się (bo może jeszcze nie jest) najpotężniejszą technologią, jaką człowiek powołał do istnienia. Potężniejszą nawet od broni

⁵ Można tu przywołać jako wymowny (i – co ważne – wymierny) przykład słynny pojedynek pomiędzy ówczesnym mistrzem świata w szachach Gary Kasparowem a superkomputerem firmy IBM o nazwie Deep Blue. Była to seria meczów rozgrywanych w formule szachów klasycznych (duża ilość czasu do namysłu) w latach 1996–1997, ostatecznie zakończona zwycięstwem komputera. Dzień 11 maja 1997 r., czyli data zakończenia meczu, uchodzi za symboliczny dzień, w którym komputer pokonał aktualnego mistrza świata w szachach, a tym samym sztuczna inteligencja w sposób obiektywny wykazała swą przewagę nad człowiekiem. Por. *Man vs Machine*, <https://www.kasparov.com/timeline-event/deep-blue/> [dostęp 25.02.2025].

⁶ Na interesujący przypadek obejścia przez ChatGPT wprowadzonych zabezpieczeń i tym samym zarwania się z ludzkiego łańcucha przez sztuczną inteligencję wskazuje raport, którego wnioski znajdujemy tutaj: Por. M. Lesiuk, *ChatGPT urwał się z łańcucha. Nowy model AI zaczyna omijać zaprogramowane ograniczenia*, <https://tiny.pl/tpq1yfbr> [dostęp: 25.02.2025].

⁷ Y.N. Harari, *Nexus. Krótka historia informacji. Od epoki kamienia do sztucznej inteligencji*, tłum. J. Hunia, Wydawnictwo Literackie, Kraków 2024, s. 19.

jądrowej. Zatem, jeśli istnienie broni masowego rażenia traktujemy – słusznie – jako egzystencjalne ryzyko globalne, to tym bardziej tak powinniśmy podchodzić do AI⁸. W kontekście zapewnienia bezpieczeństwa jako szczególnie groźna jawi się wspomniana zdolność AI do usamodzielniania się, wyciągania wniosków, obchodzenia zabezpieczeń systemu. Są to, jak zostało wspomniane, umiejętności wykraczające poza te, które zostały zaprojektowane przez jej twórców. Zjawisko to określane jest jako emergencja, a refleksja nad nim ma charakter nie tyle informatyczny, ile filozoficzny. Cytowany autor wskazuje na doniosłość emergencji i skutki lekceważenia tego zjawiska: „Emergencja to poważne wyzwanie, gdyż utrudnia pełne zrozumienie działania zaawansowanych systemów AI. Jednak wielu badaczy i konstruktorów tak jest podekscytowanych swoimi odkryciami, że nie widzą, iż kaliber zagrożeń rośnie wraz ze złożonością tworzonych przez nich modeli”⁹. Wyłania się tu istotny problem prawno-etyczny, który pojawi się w dalszej części artykułu, gdy przywołane zostaną konkretne przypadki aplikowania sztucznej inteligencji w różnych rozwiązaniach technologicznych. Otóż niezmiernie istotne jest ustalenie, kto ponosi odpowiedzialność za decyzje i działania podejmowane przez AI, zwłaszcza w sytuacji, gdy sami nie rozumiemy kryteriów, jakimi się kierowała. Jeśli rozmywa się odpowiedzialność, tzn. brak jest jasnych kryteriów określania tego, kto, kiedy i za co odpowiada, to gwałtownie rośnie niebezpieczeństwo, że decyzje podjęte w takich warunkach będą ryzykowne, a nawet ryzykanckie. Szkodliwość tego stanu rzeczy polega na tym, że jednostki czy podmioty, które same niczego nie ryzykują, mają skłonność do podejmowania decyzji nieprzemyślanych, błędnych czy wręcz katastrofalnych dla całego społeczeństwa. Po prostu łatwiej im one przychodzą, gdyż w razie niepowodzenia i tak nie ponoszą żadnych konsekwencji¹⁰. Właśnie z tego rodzaju zagrożeniami mamy do czynienia wówczas, gdy projektujemy coraz bardziej skomplikowane systemy oparte na sztucznej inteligencji. Cechuje je nadmierna złożoność oraz funkcjonowanie podobne do działania „czarnych skrzynek”, czyli m.in. brak transparentności, niejasność co do kryteriów w procesie podejmowania decyzji, nieokreślona odpowiedzialność za ewentualne niepowodzenia. Bezkrytyczne poleganie na systemach, które pozbawione są kontroli lub choćby filozoficznego namysłu prędzej czy później stanowi prostą drogę do katastrofy¹¹.

⁸ A. Zybortowicz, *AI Eksploracja*, s. 115.

⁹ *Ibidem*, s. 128. Ze względu na podstawową cechę emergencji, czyli nieprzewidywalność, pod uwagę należy brać jeszcze niebezpieczeństwo związane z wrogim wykorzystaniem emergencji: „AI z emergentnymi zdolnościami może zostać wykorzystana w szkodliwy sposób lub stać się celem cyberataków. Jak chronić system przed zhakowaniem, gdy nie bardzo wiemy, jak on działa, co jeszcze kryje w zanadrzu”. *Ibidem*.

¹⁰ Por. N. Taleb, *Na własne ryzyko. Ukryte asymetrie w codziennym życiu*, tłum. A. Unterschuettz, Zysk i S-ka, Poznań 2019.

¹¹ Nassim Taleb w przywołanej wyżej książce podaje wiele przykładów z historii cywilizacji, które unaoczniają tego typu zagrożenia i przestrzega przed ich bagatelizowaniem. Mechanizm

Opisany wyżej problem i wynikające z niego zagrożenia stają się jeszcze bardziej doniosłe i realne, gdy uświadomimy sobie, że wraz z postępem technologicznym zwiększa się wrażliwość naszej cywilizacji na przypadek. Pozornie niewielka awaria (np. komputera, systemu informatycznego itd.) może wywołać chaos na dużą skalę, przejawiający się np. w zjawisku całkowitego *blackoutu*¹². Ponadto wielkim zagrożeniem mogą być ataki hakerskie i towarzyszące im wykradanie danych wrażliwych (np. z baz danych, portali społecznościowych itp.). Osobna kategoria potencjalnych niebezpieczeństw związana jest z patologiami funkcjonowania tzw. społeczeństwa informacyjnego i kosztami otwartego dostępu do Internetu. *Fake newsy* i inne formy dezinformacji stały się na tyle powszechne, że przyzwyczailiśmy się do nich, przestaliśmy zwracać na nie uwagę i bardzo często nie postrzegamy ich jako zagrożenie. W głośnej pracy *Antykruchłość* Nassim Taleb twierdzi, że: „padamy ofiarą nowej choroby, nazwanej w tej książce neomanią, która każe nam budować systemy wrażliwe na działanie Czarnych Łabędzi. Jest nią postęp”¹³. Postęp wiedzy i technologii zatem rośnie, ale ceną, jaką za nie płacimy, jest to, że coraz trudniej przewidzieć, jakie będą tego skutki. Zdaniem autora taki stan rzeczy stanowi konsekwencję odejścia od naturalnych, sprawdzonych rozwiązań i skoncentrowanie się na budowaniu złożonych, skomplikowanych modeli i konstrukcji, których funkcjonowania sami nie rozumiemy.

3. Specjalizacja, „myślenie tunelowe” i „technoentuzjazm”, czyli gdzie tkwi źródło współczesnych zagrożeń

Według wielu autorów jest jeszcze inny, bardzo istotny czynnik czy raczej grupa czynników, które sprawiają, że zagrożenia wywołane ewentualną utratą przez człowieka kontroli nad sztuczną superinteligencją mogą być niebezpieczne. Warto poświęcić im nieco uwagi, zwłaszcza że ich analiza otwiera przestrzeń dla namysłu

ich powstawania jest zwykle taki sam: „Rzeczy zaprojektowane przez ludzi, którzy nie ryzykują własnej skóry, stają się coraz bardziej skomplikowane (aż się ostatecznie zawalą)”. *Ibidem*, s. 54.

¹² Dobrym przykładem ilustrującym opisywane zagrożenie była głośna awaria usług Microsoftu z 19 lipca 2024 r., spowodowana błędem w aktualizacji oprogramowania. Wywołała ona znaczącą dezorganizację pracy wielu poważnych instytucji z sektora bankowego, handlowego, komunikacyjnego, logistycznego, medycznego, edukacyjnego, medialnego czy administracyjnego. Pomimo że dotyczyła ona jedynie około 8,5 mln komputerów (co stanowi mniej niż 1% wszystkich urządzeń elektronicznych obsługiwanych przez Microsoft), to i tak wywołała ogromne skutki. Było tak właśnie dlatego, że awaria dotknęła w głównej mierze komputery o charakterze strategicznym, tzn. takie, które obsługują np. lotniska, dworce, banki, sklepy internetowe itp., a więc instytucje, z których usług korzystają każdego dnia setki milionów ludzi na całym świecie. Por. *CrowdStrike: globalna awaria systemu Microsoftu wynika z naszego błędu, to nie był cyberatak*, 19.07.2024, <https://tiny.pl/d5zyfz0j> [dostęp: 25.02.2025].

¹³ N. Taleb, *Antykruchłość. Jak żyć w świecie, którego nie rozumiemy*, tłum. O. Siara, Zysk i S-ka, Poznań 2020, s. 12.

filozoficznego i może stanowić prawdziwe wyzwanie dla etyki biznesu. Wśród tych czynników jest specjalizacja oraz idące z nią w parze „myślenie tunelowe”, które z kolei wiąże się zazwyczaj z bezkrytycznym „technoentuzjazmem”. Wydaje się, że wszystkie te elementy nie tylko wpisują się w mechanizm postępu opartego na rozwoju nowoczesnych technologii, ale wręcz są jego warunkiem. Jeśli chodzi o specjalizację, to widzimy znaczący jej postęp zwłaszcza w dziedzinach, w których wymagane są unikalne kompetencje i ekskluzywna wiedza. Innymi słowy, ze względu na wyjątkowy charakter jakiejś specjalistycznej dziedziny wiedzy i złożoność problemów w niej występujących może być ona dostępna jedynie dla nielicznych. Jednak ci nieliczni, siłą rzeczy, nie mogą specjalizować się w żadnej innej dziedzinie wiedzy. Krótko, a zarazem obrazowo mówiąc, specjalizacja pozwala z bliska badać poszczególne drzewa, ale uniemożliwia dostrzeżenie tego, że przebywa się w lesie. Jak przekonuje Yuval Noah Harari:

Owszem, są eksperci orientujący się w tym, jak wygląda postęp w tej czy innej sferze, na przykład w zakresie prac nad sztuczną inteligencją, w nanotechnologii, w dziedzinie big data albo genetyki, jednak nikt nie jest ekspertem od wszystkiego. Nikt zatem nie jest w stanie połączyć wszystkich punktów i zobaczyć całego obrazu. Poszczególne dziedziny wpływają na siebie w tak złożony sposób, że nawet najtęższe umysły nie potrafią w pełni pojąć, jak przełom w pracach nad sztuczną inteligencją mógłby oddziaływać na nanotechnologię czy vice versa. Nikt nie potrafi wchłonąć całości najnowszych odkryć naukowych, nikt nie potrafi przewidzieć, w jaki sposób globalna gospodarka będzie wyglądała za dziesięć lat, nikt nie ma pojęcia, dokąd w takim pośpiechu zmierzamy. A skoro nikt już nie rozumie całego tego systemu, to nikt nie potrafi go zatrzymać¹⁴.

Naturalną konsekwencją specjalizowania się w jakiejś wąskiej dziedzinie wiedzy jest patrzenie na całą otaczającą nas rzeczywistość przez pryzmat wypracowanego tam ograniczonego aparatu pojęciowego. Taki brak umiejętności szerszego spojrzenia w połączeniu z bezkrytycznym optymizmem i niezachwianą wiarą w nieograniczone możliwości technologii dobrze opisują dwa pojęcia, którymi posługuje się Andrzej Zybertowicz. Są nimi „technoentuzjazm” i „myślenie tunelowe”. Autor ten uzasadnia niebezpieczeństwa kryjące się za upowszechnieniem wyrażanych przez nie postaw i percepcji. Przekonuje on, że technoentuzjazm „cehuje się bezkrytycznym zaufaniem do rozwoju techniki i karmi narrację podci-

¹⁴ Y.N. Harari, *Homo deus. Krótka historia jutra*, tłum. M. Romanek, Wydawnictwo Literackie, Kraków 2018, s. 69. W nie mniej barwny sposób opisuje „kompetencje” specjalisty cytowany wcześniej Ronald Wright: „Pewien miłośnik logicznych absurdów tak swego czasu zdefiniował specjalistę: »To taki ktoś, kto wie coraz więcej o coraz mniejszej liczbie rzeczy, aż w końcu wie już wszystko o niczym«. Wiele gatunków zwierząt wysoce się wyspecjalizowało, ich organizmy przystosowały się do nisz ekologicznych i specyficznego sposobu życia. Specjalizacja przynosi krótkotwałe korzyści, ale w dłuższej perspektywie może zaprowadzić w ewolucyjny ślepy zaułek. Gdy wyginęło pożywienie tygrysa szablozębnego, sam tygrys także wymarł”. R. Wright, *Krótka historia postępu*, s. 59. Używając aparatu pojęciowego, stosowanego przez Nassima Taleba, można zauważyć, że specjalizacja nie sprzyja antykruchości.

nające rzeczową, krytyczną refleksję nad rozwojem technologii. Technoentuzjazm przełamuje dystanse ideologiczne, pokoleniowe i cywilizacyjne. Technoentuzjaści to w sporej części ludzie, których umysły zostały zhakowane przez rewolucję cyfrową. Czym takie osoby się cechują? Krótko i w uproszczeniu: niemal tunelowym postrzeganiem rzeczywistości, w tym swojego człowieczeństwa¹⁵. Autor nazywa ich „bezkrytycznymi fanami technologii”, którzy – ze względu na wiarę w to, że technologia rozwiąże w przyszłości wszystkie nasze problemy – uchodzą też za tzw. technologicznych solucjonistów¹⁶. Co więcej, ich wyrywkowe postrzeganie świata, a więc to, że nie postrzegają określonych problemów czy zagadnień w szerszym, cywilizacyjnym kontekście, sprawia, że nie są oni w stanie zatrzymać się na chwilę, nabrać dystansu do tego, co robią i w czym pokładają swe nadzieje. Przez to, że widzą wszystko oddzielnie, nie są zdolni (ani zainteresowani) dostrzec społecznych, politycznych, kulturowych i etycznych konsekwencji wprowadzania dynamicznych zmian technologicznych.

4. Ludzka codzienność w otoczeniu nowoczesnych technologii

Omówione wyżej niebezpieczeństwa o charakterze globalnym, związane z niekontrolowanym rozwojem sztucznej superinteligencji, nie są jedynym polem, na którym refleksja etyczna może znaleźć swoje uzasadnienie. Otóż inna, ważna grupa problemów, z którą już musi się mierzyć etyka biznesu, związana jest z używaniem nowoczesnych technologii w nieco bardziej konkretnych sytuacjach, często w celu realizacji śmiałych i nowatorskich przedsięwzięć oraz rozwiązań technicznych. Skutkuje to szeregiem nieznanych wcześniej lub bagatelizowanych zagrożeń, które współcześnie zaczynają się urzeczywistniać. Jednym z tego typu problemów jest narastająca groźba utraty pracy w wielu profesjach. Rozwój nowoczesnych technologii znacząco wpłynął na rynek pracy, przyczyniając się bądź to do jej „uberyzacji”, bądź do zaniku szeregu mniej lub bardziej tradycyjnych zawodów czy kompetencji. Na naszych oczach dochodzi do przekształcenia struktury i formy zatrudnienia milionów ludzi na całym świecie. Mówiąc krótko, stoimy u progu zmian, które Klaus Schwab określa mianem „czwartej rewolucji przemysłowej”,

¹⁵ A. Zybertowicz, *Cyber kontra real. Cywilizacja w techno-pułapce. Z Andrzejem Zybertowiczem rozmawia Jarema Piekutowski*, Wydawnictwo Nowej Konfederacji, Warszawa 2022, s. 12. W innym miejscu autor ten definiując technoentuzjazm, pisze, że jest to „postawa oznaczająca daleko idące, zwykle bezrefleksyjne przyzwolenie na wprowadzanie w życie społeczne kolejnych nowych technologii jako narzędzi postępu” (zob. A. Zybertowicz, *AI Eksploracja*, s. 25). Aby uświadomić czytelnikom powagę sytuacji i realizm opisywanych zagrożeń, Zybertowicz nie waha się stwierdzić, że „technoentuzjazm jest w dzisiejszym świecie najniebezpieczniejszą z ideologii”. A. Zybertowicz, *Cyber kontra real...*, s. 12.

¹⁶ Pojęciem tym posługuje się Evgeny Morozov, autor książki *To Save Everything, Click Here: The Folly of Technological Solutionism*, Public Affairs, New York 2013.

i które będą wymagać już nie tylko filozoficznej refleksji, ale po prostu realnych działań i rozwiązań¹⁷.

Kolejnym niezmiernie istotnym zagadnieniem dotyczącym współczesnych wyzwań etyki biznesu jest odpowiedź na szereg skomplikowanych kwestii, które wyłaniają się w trakcie projektowania i użytkowania nowoczesnych technologii opartych na sztucznej inteligencji. Są to często bardzo konkretne dylematy etyczne, wynikające wprost z faktu pojawienia się w naszym codziennym życiu AI. Na przykład pytanie o to, kto ponosi odpowiedzialność (etyczną, prawną itd.) za wypadek samochodowy spowodowany w wyniku złej „decyzji” sztucznej inteligencji, czyli błędu systemu kierowania samochodem przez AI? Czy winę ponosi kierowca, czy też może inżynier tworzący oprogramowanie? A może zarząd fabryki samochodowej, który dopuścił taki pojazd do ruchu? A może państwo, które istnieje m.in. po to, by sprawować nadzór nad różnymi przedsięwzięciami? Kwestie te muszą zostać jednoznacznie rozstrzygnięte w warstwie prawnej (gdy chodzi np. o ewentualne odszkodowanie), ale problem w tym, że – ze względu na dynamikę postępu technologicznego i ekspansywne wdrażanie sztucznej inteligencji – prawo niejednokrotnie nie nadąża za zmianami¹⁸. Otwiera się tu zatem przestrzeń dla rozważań z zakresu etyki biznesu w celu wypracowania jakichś ogólnych procedur i reguł postępowania. W kategoriach etycznych należy rozpatrywać także głośną sprawę „wskrzeszenia” przez sztuczną inteligencję Wisławy Szymborskiej, z którą troje wygenerowanych również przez AI „dziennikarzy” publicznego Radia Kraków przeprowadziło „wywiad”¹⁹. Jeszcze bardziej kontrowersyjne, a przy tym etycznie wielce wątpliwe, wydają się być coraz częściej podejmowane skuteczne próby „przywracania do życia” osób zmarłych. Na życzenie (i za opłatą) najbliższych zmarłego tworzy się jego awatara²⁰. Dzięki temu mogą się oni nadal cieszyć jego „obecnością” i funkcjonować na dotychczasowych zasadach. W tego

¹⁷ K. Schwab, *Czwarta rewolucja przemysłowa*, tłum. A.D. Kamińska, Studio Emka, Warszawa 2018. Autor wymienia następujące zawody, które najbardziej narażone są na niebezpieczeństwo całkowitego zaniku: telemarketerzy, obsługa podatków, rzeczoznawcy ubezpieczeniowi, arbitrzy sportowi, prawnicy w spółkach, hostessy w restauracjach i kawiarniach, maklerzy nieruchomości, sezonowi pracownicy rolni, sekretarki i asystenci, kurierzy i gońcy (por. K. Schwab, *Czwarta rewolucja...*, s. 58). W opinii autora, wiele kolejnych zawodów i dziedzin czeka całkowita lub częściowa automatyzacja, w miarę tempa wzrostu mocy obliczeniowej różnego rodzaju urządzeń. Są to m.in.: prawo, medycyna, dziennikarstwo, analityka finansowa, księgowość, ubezpieczenia czy bibliotekarstwo. Por. *ibidem*, s. 56.

¹⁸ Istnieją oczywiście próby uregulowania tych kwestii, np. w ramach regulacji prawnych Unii Europejskiej, ale ze względu na dynamikę sytuacji nie są one doskonałe. Zob. R. Pilar-ski, *Szkody wyrządzone przez AI – kto za nie odpowiada?*, <https://tiny.pl/8sgx3h34> [dostęp: 25.02.2025].

¹⁹ Zob. M. Tomczak, *Burza wokół Radia Kraków. „Wywiad” AI z Wisławą Szymborską*, 24.10.2024, <https://tiny.pl/kyjpw353> [dostęp: 25.02.2025].

²⁰ Zob. *Wskrzesanie zmarłych przy pomocy AI – czyli jak sztuczna inteligencja może dać drugie życie naszym bliskim*, 1.11.2024, <https://tiny.pl/847xs5kw> [dostęp: 25.02.2025].

typu przypadkach kontrowersje nie dotyczą jedynie kwestii etycznych, ale także ogólniejszych, tzn. kulturowych. Naruszają bowiem odwieczne tabu związane ze śmiercią, podkopując w ten sposób fundamenty cywilizacji w postaci możliwości pożegnania się ze zmarłym, pogrzebania go i przeżycia żałoby po nim.

Nie ma w tym momencie miejsca na mnożenie przykładowych dylematów etycznych, spowodowanych aplikowaniem do naszego codziennego życia technologicznych rozwiązań opartych na sztucznej inteligencji. Należy jednak wspomnieć o całej grupie zagrożeń związanych z inwigilacją jednostek. Nadzór i kontrola społeczna wpisane są w dzieje ludzkości, stanowiąc od wieków skuteczne narzędzie w rękach każdej władzy. Jednak współcześnie, w dobie Internetu i technologii wykorzystujących AI, możliwości skutecznej inwigilacji znacząco wzrastają. Novum polega głównie na tym, że dziś jest ona domeną nie aparatu państwa, ale wielkich, międzynarodowych korporacji z branży IT²¹. Pozyskują one różnego rodzaju informacje o użytkownikach sieci w celu wykorzystania ich w działalności komercyjnej (choć nie tylko)²². Zazwyczaj odbywa się to bez wiedzy i zgody internautów, z pominięciem transparentnych procedur i zrozumiałych reguł prawnych. Przewaga GAFAM-u²³ nad zwykłym użytkownikiem Internetu jest tak wielka, że w wielu sytuacjach bywa on całkowicie bezbronny. Jest on nie tylko nieustannie inwigilowany (przez co zagrożona jest jego prywatność), ale także wykorzystywany. Tę asymetrię i poddańczy charakter omawianych tu stosunków niektórzy autorzy nazywają technofeudalizmem. Korporacje niczym średniowieczni panowie feudalni w nie do końca uprawniony sposób korzystają z pracy swoich chłopów pańszczyźnianych, czyli użytkowników sieci²⁴.

²¹ Na temat patologii wynikających ze skoncentrowania realnej władzy w rękach kilku amerykańskich gigantów branży IT pisze m.in. Vivek Ramaswamy w swej książce *Woke S.A. Kulisy amerykańskiego przekrętu sprawiedliwości społecznej*. Zwłaszcza rozdział 9 pt. *Krzemowy Lewiatan* obrazuje konsekwencje przeniesienia centrów decyzyjnych do Doliny Krzemowej. Zob. V. Ramaswamy, *Woke S.A. Kulisy amerykańskiego przekrętu sprawiedliwości społecznej*, tłum. A. Janus, WEI, Warszawa 2023, ss. 151–174.

²² Por. S. Zuboff, *Wiek kapitalizmu inwigilacji. Walka o przyszłość ludzkości na nowej granicy władzy*, tłum. A. Unterschuetz, Zysk i S-ka, Poznań 2020.

²³ Chodzi tu o pięć obecnie największych spółek świata: Google, Apple, Facebook, Amazon i Microsoft (stąd GAFAM), dla których podstawową przestrzenią funkcjonowania jest Internet. Stanowi on dla nich miejsce eksploracji różnego rodzaju informacji dotyczących jego użytkowników.

²⁴ Taką terminologią posługuje się m.in. grecki ekonomista, wykładowca akademicki i polityk Yanis Varoufakis. W swej książce *Technofeudalizm* autor w następujący sposób opisuje funkcjonowanie „kapitału w chmurze”, czyli połączonych w sieć komputerów i urządzeń służących do wyłudzenia od nas danych i wykonujących wiele zadań, takich jak: „skłanianie miliardów samodzielnych pracowników (chłopów pańszczyźnianych w chmurze) do pracy za darmo (często nieświadomie) przy pomnażaniu kapitału w chmurze (poprzez wrzucanie na Instagram czy TikToka zdjęć i filmów lub publikowanie nagrań i recenzji restauracji oraz książek); polecanie nam do snu książek, filmów, miejsc na wakacje z taką skutecznością, że chętnie kupujemy inne dobra oferowane w cyfrowych lennach lub na platformach sprzedażowych (jak Amazon)

Do ingerowania w prywatność i zawłaszczania danych dochodzi także w prężnie rozwijającej się dziedzinie zwanej Internetem Rzeczy. Otóż rozwój nowoczesnych technologii sprawił, że obecnie już nie tylko komputery, tablety czy smartfony, ale też samochody, telewizory, drukarki, systemy alarmowe, łodówki, odkurzacze, garnki, a nawet szczoteczki do zębów mogą zostać podłączone do sieci. Rozwiązanie takie jest pod wieloma względami korzystne i wygodne dla ich użytkowników. Jednak ceną, jaką muszą oni za te dobrodziejstwa zapłacić, jest mniej lub bardziej uświadomiona zgoda na kontrolę, inwigilację i wykorzystywanie informacji o ich prywatnym życiu i codziennych przyzwyczajeniach. Wszystkie te zaawansowane technologicznie urządzenia zbierają bowiem różnorakie dane o naszej aktywności lub otoczeniu i automatycznie przekazują je do „centrali”²⁵. Opisywane tu kwestie własnościowe i towarzyszące im wątpliwości etyczne wynikają też z innego powodu. Mianowicie w świecie zdominowanym przez nowoczesne technologie w coraz mniejszym stopniu jesteśmy właścicielami, a coraz częściej pozostajemy jedynie użytkownikami, którzy wykupują subskrypcje na dostęp do np. gry (tzw. model buy-to-play), filmu, aplikacji czy usługi. Sytuacja taka może rodzić szereg nieporozumień, komplikacji czy rozczarowań. Nie występowały one w tradycyjnych rozwiązaniach, w których z chwilą dokonania zakupu stajemy się właścicielami produktu. Nie tylko prawne, ale również etyczne wątpliwości pojawiają się w przypadku stosowania nowatorskich rozwiązań, które – wykorzystując sztuczną inteligencję – kreują od podstaw utwory literackie. Co więcej, mają one także możliwość „dopisywania” w wybrany przez nas sposób fabuły do ulubionej powieści, filmu, utworu czy praktycznie każdego innego dzieła kultury artystycznej²⁶. Spośród licznych zagrożeń związanych z tego typu działalnością szczególnie niebezpieczny wydaje się proceder polegający na kreowaniu i upublicznianiu różnego rodzaju *deep fake*ów. Za ikoniczny uchodzi niedawny przypadek wyprodukowania przy użyciu AI filmu pornograficznego z „udziałem” premier Włoch Giorgi Meloni²⁷.

wpiętych w te same sieci, które polecają nam książki, filmy i miejsca na wakacje tuż przed tym, jak pójdziemy spać; wykorzystywanie sztucznej inteligencji oraz big data do zarządzania pracą robotników (proletariat w chmurze) z zakładów produkcyjnych, przy równoczesnym kierowaniu sieciami energetycznymi, robotami, ciężarówkami, zautomatyzowanymi liniami produkcyjnymi i drukarkami 3D, które pomijają tradycyjne sposoby produkcji”. Y. Varoufakis, *Technofeudalizm. Co zabiło kapitalizm?*, tłum. P. Szadkowski, GlowBook, Sieradz 2024, ss. 234–235.

²⁵ J. Greser, *Etyka biznesu wobec Internetu Rzeczy*, w: *25 lat etyki biznesu w Polsce. Raport 2019*, red. B. Rok, R. Sroka, Akademia Leona Koźmińskiego, Warszawa 2019, ss. 91–92.

²⁶ O upowszechniającym się zjawisku *self-publishingu* i nadużyciach z nim związanych pisze się coraz częściej: Zob. G. Kubera, *Ludzie nie czytają książek, za to piszą ich bardzo dużo. „Zalewają nas powieści generowane przez AI”*, 1.10.2023, https://tiny.pl/6_yzwcgm [dostęp: 25.02.2025].

²⁷ Zob. A. Wittenberg, *Pornograficzne deep fake podbijają sieć. Jak się obronić?*, 19.04.2024, https://tiny.pl/9xt_726n [dostęp: 25.02.2025].

Zamiast zakończenia, czyli do czego może dziś posłużyć etyka biznesu

W ramach podsumowania niniejszych rozważań zasadne wydaje się zebranie wniosków dotyczących współczesnych wyzwań stojącymi przed etyką biznesu. Kluczową wydaje się kwestia, na ile etyczna refleksja nad skutkami upowszechnienia się nowych technologii wymaga wyjątkowego i nowatorskiego podejścia. Innymi słowy, czy realia, warunki i sytuacje, z którymi przychodzi się mierzyć użytkownikom sztucznej inteligencji, są na tyle bezprecedensowe i osobliwe, że tradycyjna refleksja etyczna nie do końca znajduje zastosowanie? Są to w gruncie rzeczy najważniejsze pytania dotyczące wyzwań, które przed etyką biznesu stawia obecnie świat zdominowany przez nowoczesne technologie. Próbując ustosunkować się do powyższych kwestii, należy mieć świadomość, że obok niezaprzeczalnych korzyści, wynikających ze stosowania zaawansowanych technologii, narażeni jesteśmy na liczne niebezpieczeństwa i ryzyka związane z tym dobrodziejstwem. Będą się one pojawiać (bądź już są obecne) zwłaszcza w technologiach i rozwiązaniach, w których wykorzystuje się sztuczną inteligencję. Na poziomie ogólnym, gdzie rozwój sztucznej superinteligencji zagraża przyszłości cywilizacji, etyka biznesu może dostarczyć wielu przekonujących argumentów za wprowadzeniem określonych ram czy barier, które do pewnego stopnia spowolniłyby ten rozwój. Mogłoby to z kolei umożliwić wypracowanie skutecznych narzędzi i procedur zabezpieczających przed wymknięciem się AI spod kontroli człowieka²⁸. Natomiast na poziomie bardziej przyziemnym, na którym człowiek, jako zwyczajny użytkownik, styka się w swoim codziennym życiu ze wspierającą go sztuczną inteligencją, rola etyki biznesu wydaje się nieco inna. Wyzwaniem dla etyków biznesu nie jest tu dzierżenie drążka hamulca, ale raczej argumentowanie za wprowadzaniem

²⁸ Takie spowalniające zabiegi są już czynione, ale cały czas należy być w gotowości do podejmowania kolejnych inicjatyw w tym względzie. Jedną z nich jest ruch *słowa AI*, który lobbuje na rzecz wyhamowania tempa rozwoju w technologiach i dziedzinach wykorzystujących AI. Jedną z najbardziej znaczących i rozpoznawalnych postaci w świecie nowych technologii, która podejmuje się niewdzięcznej roli „hamulcowego”, jest Mustafa Suleyman. W książce *Nadchodząca fala. Sztuczna inteligencja, władza i najważniejszy dylemat ludzkości w XXI wieku*, postuluje on wprowadzenie czegoś na kształt moratorium, czyli powstrzymania się (przynajmniej czasowego) niekontrolowanego rozwoju AI. Proponuje on wprowadzenie dziesięciu konkretnych kroków, które w realny i skuteczny sposób zastopują ów rozwój i przyczynią się do panowania nad nim. Można wśród nich znaleźć postulaty wprowadzenia regulacji prawnych, standardów etycznych, mechanizmów i procedur pozwalających pociągnąć do odpowiedzialności instytucje za szkody wyrządzone przez AI, zawarcia międzynarodowych porozumień o ograniczeniach itp. Co istotne, pisze on o tym jako praktyk i przedsiębiorca, współzałożyciel firmy DeepMind – pioniera w dziedzinie tworzenia AI. Zob. M. Suleyman, M. Bhaskar, *Nadchodząca fala. Sztuczna inteligencja, władza i najważniejszy dylemat ludzkości w XXI wieku*, tłum. J. Hunia, Szczeliny, Kraków 2024, ss. 368–431.

takich rozwiązań technologicznych, które będą jak najbardziej „przyjazne” dla użytkownika. Funkcjonuje już nawet specjalne określenie – „dostosowanie AI” (ang. *AI alignment*), które oddaje sens tej strategii. Kryje się pod nim cały szereg działań mających na celu skonstruowanie takiego modelu sztucznej inteligencji, który będzie zgodny z ludzkim systemem wartości. Chodzi zatem o dostosowanie nowoczesnej technologii wykorzystującej sztuczną inteligencję do tego, co my, ludzie, uznajemy za dobre, słuszne, sprawiedliwe, jednym słowem – etyczne²⁹.

Aby idea *AI alignment* mogła się urzeczywistnić, tzn. przynieść wymierne skutki, należy myśleć o owym dostosowaniu AI do ludzkiego świata wartości już na poziomie programowania czy też projektu jakiegoś technologicznego rozwiązania. W literaturze przedmiotu funkcjonuje już pojęcie określające ten rodzaj refleksji etycznej, czyli „Ethics by Design for AI” lub w skrócie „EbD-AI”³⁰. U podstaw koncepcji „etyki projektowania sztucznej inteligencji” leży założenie, że technologia nie jest neutralna aksjologicznie, tzn. że wartości mogą być, do pewnego stopnia, osadzone w niej na etapie projektowania. Dla przykładu, już na etapie projektu można zdecydować, że skierowana do młodzieży aplikacja używana w smartfonie posiada właściwości uzależniające. W tym właśnie sensie nowoczesne rozwiązania technologiczne nie są „neutralne”, tak jak neutralne są tradycyjne narzędzia (nóż, młotek, nożyczki itp.), które bierzemy do ręki wyłącznie wtedy, gdy chcemy ich użyć. Aplikacja, czy nawet cały smartfon, towarzyszą nam nieustannie, stanowiąc niejako „przedłużenie” naszej ręki.

O tym, jakie będą wyzwania stojące przed etyką biznesu oraz jakie tematy znajdą swój wyraz w literaturze przedmiotu³¹ i akademickich dyskusjach, zdecyduje oczywiście przyszłość. Jednak już dziś można powiedzieć, że rola etyki w zmieniającym się technologicznie świecie nie tylko nie zmniejszy się, ale – jak się wydaje – wręcz wzrośnie. Wbrew pozorom „zapotrzebowanie” na stare, sprawdzone zasady etyczne w zmieniającym się dynamicznie świecie ma się (albo powinno się mieć) całkiem dobrze. Choć z jednej strony bezprecedensowe zmiany technologiczne wymagają w refleksji etycznej uwzględnienia nowych aspektów, to z drugiej mamy do czynienia z pewnym paradoksem. Polega on na tym, że wraz z rozwojem nowoczesnych technologii i coraz powszechniejszym wdrażaniem w życie opartych na nich rozwiązań (w tym także rozwiązań biznesowych) zwiększa się tradycyjna rola etyki. Dzieje się tak m.in. dlatego, że jako zwykli użytkownicy zaawansowanych technologii nie rozumiemy zazwyczaj zasady ich działania. Są one dla nas swego rodzaju „czarny-

²⁹ M. Malinowski, *Alignment AI. Jak dostosować sztuczną inteligencję do ludzkich celów*, 23.09.2024, <https://www.drmalinowski.edu.pl/posts/3021-alignment-ai> [dostęp: 25.02.2025].

³⁰ P. Brey, B. Dainow, *Ethics by design for artificial intelligence*, „AI and Ethics” 2024, no. 4, ss. 1265–1277.

³¹ Systematyczny przegląd aktualnej literatury z zakresu etyki biznesu, które dotyczą sztucznej inteligencji, zawiera artykuł: F.V. Giarmoleo et al., *What ethics can say on artificial intelligence: Insights from a systematic literature review*, „Business and Society Review” 2024, vol. 129(2), <https://onlinelibrary.wiley.com/doi/10.1111/basr.12336> [dostęp: 25.02.2025].

mi skrzynkami”, których efekty działań musimy przyjmować na wiarę. Widać to zwłaszcza w dziedzinie e-commerce³², w której szereg działań nie byłyby możliwe bez elementarnego założenia o uczciwości sprzedawcy bądź nabywcy, okazania minimum zaufania dla anonimowego podmiotu, znajdującego się „po drugiej stronie” komputera czy smartfona. Wciskając klawisz „enter”, którym potwierdzamy transakcję i przelewamy określoną kwotę środków z naszego bankowego konta, zakładamy, że dojdzie ona do skutku i będzie rzetelna. Pomimo technicznych i prawnych zabezpieczeń w ostateczności musimy się odwołać do ustalonych zasad etycznych i zdani jesteśmy na to, że będzie się nimi kierował sprzedawca.

Zapewne niniejszy artykuł mógłby zostać napisany przez ChatGPT. Być może byłby bardziej treściwy i merytoryczny, lepiej zredagowany i przyjemniejszy w lekturze, jednym słowem – bardziej interesujący i wartościowy. Ale to m.in. etyka nie pozwoliła jego autorowi ułatwić sobie życia i „pójść na skróty”.

Literatura

- Brey P., Dainow, B., *Ethics by design for artificial intelligence*, „AI and Ethics” 2024, no. 4, <https://doi.org/10.1007/s43681-023-00330-4> [dostęp: 25.02.2025].
- Giarmoleo F.V., Ferrero I., Rocchi M., Pellegrini M.M., *What ethics can say on artificial intelligence: Insights from a systematic literature review*, „Business and Society Review” 2024, vol. 129(2), <https://onlinelibrary.wiley.com/doi/10.1111/basr.12336> [dostęp: 25.02.2025].
- Greser J., *Etyka biznesu wobec Internetu Rzeczy*, w: *25 lat etyki biznesu w Polsce. Raport 2019*, red. B. Rok, R. Sroka, Akademia Leona Koźmińskiego, Warszawa 2019.
- Harari Y.N., *Homo deus. Krótka historia jutra*, tłum. M. Romanek, Wydawnictwo Literackie, Kraków 2018.
- Harari Y.N., *Nexus. Krótka historia informacji. Od epoki kamienia do sztucznej inteligencji*, tłum. J. Hunia, Wydawnictwo Literackie, Kraków 2024.
- Morozov E., *To Save Everything, Click Here: The Folly of Technological Solutionism*, Public Affairs, New York 2013.
- Pinker S., *Nowe Oświecenie. Argumenty za rozumem, nauką, humanizmem i postępem*, tłum. T. Bieroń, Zys i S-ka, Poznań 2018.
- Ramaswamy V., *Woke S.A. Kulisy amerykańskiego przekrętu sprawiedliwości społecznej*, tłum. A. Janus, WEI, Warszawa 2023.
- Schwab K., *Czwarta rewolucja przemysłowa*, tłum. A.D. Kamińska, Studio Emka, Warszawa 2018.

³² E-commerce jest dziedziną, w której szczególnie dobrze uwidacznia się potrzeba nie tylko ogólnej refleksji etycznej, ale także zastosowania konkretnych rozwiązań promujących uczciwy i sprawiedliwy handel. W polskich warunkach idee postulujące wprowadzenie i przestrzeganie kodeksu etyki branży e-commerce promuje na swojej stronie internetowej i w podcastach na kanele You Tube Robert Sroka. Zob. *Etyka biznesu, ESG. etyka AI*, <https://robertsroka.pl/> [dostęp: 25.02.2025]; *Czy handel internetowy może być ETYCZNY?*, <https://tiny.pl/kwg1gnnp> [dostęp: 25.02.2025]; *Czy potrzebujemy nowej etyki biznesu?*, https://tiny.pl/mfpvzw_7 [dostęp: 25.02.2025].

- Suleyman M., Bhaskar M., *Nadchodząca fala. Sztuczna inteligencja, władza i najważniejszy dylemat ludzkości w XXI wieku*, tłum. J. Hunia, Szczeliny, Kraków 2024.
- Taleb N., *Antykruchłość. Jak żyć w świecie, którego nie rozumiemy*, tłum. O. Siara, Wydawnictwo Zysk i S-ka, Poznań 2020.
- Taleb N., *Na własne ryzyko. Ukryte asymetrie w codziennym życiu*, tłum. A. Unterschuetz, Wydawnictwo Zysk i S-ka, Poznań 2019.
- Varoufakis Y., *Technofeudalizm. Co zabiło kapitalizm?*, tłum. P. Szadkowski, GlowBook, Sieradz 2024.
- Wright R., *Krótką historia postępu*, tłum. Ł. Wierzbicki, Pogotowie Kazikowe, Borówiec 2021.
- Zuboff S., *Wiek kapitalizmu inwigilacji. Walka o przyszłość ludzkości na nowej granicy władzy*, tłum. A. Unterschuetz, Zysk i S-ka, Poznań 2020.
- Zybertowicz A., *AI Eksploracja. Andrzej Zybertowicz odpowiada Katarzynie Zybertowicz, Zona Zero*, Warszawa 2024.
- Zybertowicz A., *Cyber kontra real. Cywilizacja w techno-pułapce. Z Andrzejem Zybertowiczem rozmawia Jarema Piekutowski*, Wydawnictwo Nowej Konfederacji, Warszawa 2022.

Źródła internetowe

- CrowdStrike: globalna awaria systemu Microsoftu wynika z naszego błędu, to nie był cyberatak*, 19.07.2024, <https://tiny.pl/d5zyfz0j> [dostęp: 25.02.2025].
- Czy handel internetowy może być ETYCZNY?*, <https://tiny.pl/kwg1gnnp> [dostęp: 25.02.2025].
- https://bithub.pl/gospodarka/chatgpt-urwal-sie-z-lancucha-nowy-model-ai-zaczyna-omijac-zaprogramowane-ograniczenia/#utm_source=dlvr.it&utm_medium=facebook [dostęp: 25.02.2025].
- <https://brandsit.pl/globalna-awaria-microsoft/> [dostęp: 25.02.2025].
- Kubera G., *Ludzie nie czytają książek, za to piszą ich bardzo dużo. „Zalewają nas powieści generowane przez AI”*, 1.10.2023, https://tiny.pl/6_yzwcgm [dostęp: 25.02.2025].
- Malinowski M., *Alignment AI. Jak dostosować sztuczną inteligencję do ludzkich celów*, 23.09.2024, <https://www.drmalinowski.edu.pl/posts/3021-alignment-ai> [dostęp: 25.02.2025].
- Man vs Machine*, <https://www.kasparov.com/timeline-event/deep-blue/> [dostęp: 25.02.2025].
- Pilarski R., *Szkody wyrządzone przez AI – kto za nie odpowiada?*, <https://tiny.pl/8sgx3h34> [dostęp: 25.02.2025].
- Sroka R., *Czy potrzebujemy nowej etyki biznesu?*, https://tiny.pl/mfpvzw_7 [dostęp: 25.02.2025].
- Sroka R., *Etyka biznesu, ESG. etyka AI*, <https://robertsroka.pl/> [dostęp: 25.02.2025].
- Tomczak M., *Burza wokół Radia Kraków. „Wywiad” AI z Wisławą Szymborską*, 24.10.2024, <https://tiny.pl/kyjpw353> [dostęp: 25.02.2025].
- Wittenberg A., *Pornograficzne deep fake podbijają sieć. Jak się obronić?*, 19.04.2024, https://tiny.pl/9xt_726n [dostęp: 25.02.2025].
- Wskrzesanie zmarłych przy pomocy AI – czyli jak sztuczna inteligencja może dać drugie życie naszym bliskim*, 1.11.2024, <https://tiny.pl/847xs5kw> [dostęp: 25.02.2025].

