

„Fascynująca, świetnie napisana i ważna książka”. YUVAL NOAH HARARI

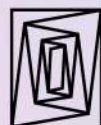
**Sztuczna inteligencja, władza
i najważniejszy dylemat ludzkości
w XXI wieku**

NADCHODZĄCA FALA

MUSTAFA SULEYMAN

współzałożyciel DEEPMIND i INFLECTION AI

oraz MICHAEL BHASKAR



Nadchodząca fala to fascynująca, znakomicie napisana i ważna książka. Rozpatruje egzystencjalne zagrożenia, jakie sztuczna inteligencja i biotechnologia stwarzają dla ludzkości, a zarazem proponuje praktyczne rozwiązania na rzecz zażegnania niebezpieczeństwa. Nadchodząca fala technologiczna niesie obietnicę obdarzenia ludzkości boską mocą kreacji, ale jeśli nie będziemy nią mądrze dysponować, może nas zniszczyć.

Yuval Noah Harari, autor bestsellera „New York Timesa”
Sapiens. Od zwierząt do bogów

To dobiegające z przyszłości ostrzeżenie przed tym, co nadchodzi i jakie mogą być tego globalne implikacje gospodarcze i polityczne. Doprawdy niezwykła, ambitna, przekonująco uargumentowana książka, której nie sposób zignorować. Jej autor, uznany ekspert w dziedzinie technologii, stworzył prawdziwy majstersztyk, który uformuje wasze spojrzenie na przyszłość i odmieni rozumienie teraźniejszości.

Nouriel Roubini, emerytowany profesor
na Uniwersytecie Nowojorskim

Wnikliwa wiedza Mustafy Suleymana jako twórcy technologii, przedsiębiorcy i wizjonera jest trudna do przecenienia. Jego książka, poparta rzetelnymi badaniami i niezwykle aktualna, oferuje fascynujący wgląd w niektóre najważniejsze wyzwania naszych czasów.

Al Gore, były wiceprezydent Stanów Zjednoczonych

W tej odważnej książce Mustafa Suleyman, jeden z prawdziwych koryfeuszy zaawansowanych technologii, skupia się na najważniejszym paradoksie naszych czasów: musimy powstrzymać technologie, których powstrzymać się nie da. Jak wyjaśnia, generatywna sztuczna inteligencja, biologia syntetyczna, robotyka i inne innowacje są coraz szybciej ulepszone i upowszechniane. Przynoszą ogromne korzyści, ale też stwarzają realne i rosnące niebezpieczeństwo. Suleyman jest na tyle mądry, by wiedzieć, że nie istnieje żaden prosty trzy punktowy plan zarządzania tym ryzykiem, i na tyle odważny, by o tym głośno mówić. Ta książka jest szczerą i pełną pasji,

nie boi mierzyć się z czymś, co bez wątpienia stanowi jedno z największych wyzwania, przed jakimi naszemu gatunkowi przyjdzie stanąć w tym stuleciu. Dzięki Suleymanowi wiemy, w jakim znajdujemy się położeniu i jakie mamy opcje. Teraz pozostało nam tylko działać.

Andrew McAfee, kierownik działu badań w Sloan School of Management w Massachusetts Institute of Technology, autor książki *The Geek Way*

Na naszych oczach dokonuje się rewolucja sztucznej inteligencji. Jak dobrze ją rozumiemy? *Nadchodząca fala* to błyskotliwy, a zarazem wnikliwy przewodnik zarówno po historii radykalnych zmian technologicznych, jak i po kolosalnych wyzwaniach politycznych, przed którymi stoimy.

Anne Applebaum, historyczka, laureatka Nagrody Pulitzera

Kiedy to wylądowało w mojej skrzynce mailowej, zamknąłem się w swoim pokoju i zabrałem do czytania. To niezwykła i szalenie potrzebna książka: szokujące jest to, że za dwadzieścia lat będzie wydawać się niemal konserwatywną wizją przyszłości, a dziś nie sposób czytać ją, nie zatrzymując się co kilka stron, by zadać sobie pytanie: czy to może stać się faktem? Geniusz tej książki polega na tym, że wyjaśnia – trzeźwo i z wycuciem – że owszem, to wszystko stanie się faktem, oraz uzasadnia, dlaczego i w jaki sposób się to stanie. Całość jest utrzymana w subtelnym, przyjaznym tonie łagodzącym przejmujący czytelnika szok. Pojawiają się momenty grozy, co wydaje się nieuniknione, gdy zdajemy sobie sprawę, że niemal wszystko, co znamy, wkrótce ulegnie głębokiej transformacji. Ostatecznie jednak na końcu lektury czujemy przypływ energii i podekscytowanie faktem, że żyjemy właśnie w tym momencie. Fala wkrótce uderzy i tak przedstawia się prognoza.

Alain de Botton, filozof, autor bestsellerów

Nadchodząca fala oferuje bardzo potrzebną dawkę konkretności, realizmu i klarowności w kwestii nieprzewidzianych, a zarazem potencjalnie katastrofalnych konsekwencji sztucznej inteligencji, biologii syntetycznej i innych zaawansowanych technologii. Ta ważna książka to sugestywny i przekonujący plan

działania, pokazujący, w jaki sposób ludzie mogą wyznaczać kierunki innowacji technologicznych, zamiast dawać się im zdominować.

Martha Minow, profesor Uniwersytetu Harvarda,
była dziekan Harvard Law School

Nikt nie był bliżej początków rewolucji sztucznej inteligencji niż Mustafa Suleyman i nikt nie ma lepszych papierów na to, by naświetlać zagrożenia i korzyści płynące z potężnych zmian technologicznych, jakich jesteśmy świadkami. To niezwykle i obowiązkowe studium wyjątkowego momentu w historii ludzkości.

Eric Schmidt, były dyrektor generalny Google,
współautor książki *Era sztucznej inteligencji*

Na kartach *Nadchodzącej fali* Mustafa Suleyman ze wszelkich miar przekonująco dowodzi, że dokonująca się właśnie potężna rewolucja technologiczna może okazać się wyjątkowo destrukcyjna. To lektura obowiązkowa, która uzmysławia nam tempo i skalę tych technologii i pokazuje, jak te będą się rozprzestrzeniać w naszym społeczeństwie, potencjalnie wstrząsając posadami instytucji organizujących nasz świat.

Ian Bremmer, założyciel Eurasia Group,
autor bestsellera *The Power of Crisis*

Ta ważna książka jest tyleż inspirująca, co przerażająca. To niezastąpiona lekcja dla tych, którzy nie rozumieją dokonujących się na naszych oczach rewolucji technologicznych, a zarazem ogromne wyzwanie dla tych, którzy je rozumieją. To rzecz o przyszłości nas wszystkich: musimy to przeczytać i wyciągnąć wnioski.

David Miliband, były minister spraw
zagranicznych Wielkiej Brytanii

Kreśląc nieupiększony obraz niebezpieczeństw, a także cudów sztucznej inteligencji, Mustafa Suleyman proponuje program niecierpiących zwłoki działań, które rządy powinny podjąć już

dziś, aby ograniczyć potencjalnie najbardziej katastrofalne zastosowania tych rewolucyjnych technologii.

Graham Allison, profesor Uniwersytetu Harvarda, autor bestsellera *Destined for War*

Wykładniczy rozwój technologii poraził nas swoją mocą i niebezpieczeństwami. Śledząc historię rozwoju przemysłu aż do zawrotnego przyspieszenia charakteryzującego najnowsze postępy w technologii, Mustafa Suleyman wyważonym, pragmatycznym i głęboko etycznym tonem poszerza naszą perspektywę. Jego droga życiowa i bagaż doświadczeń przydają książce smaczku i czynią ją fascynującą lekturą dla każdego, kto chce oderwać się od codziennego natłoku wiadomości technologicznych.

Angela Kane, była zastępczyni sekretarza generalnego ONZ i wysoki przedstawiciel Narodów Zjednoczonych ds. rozbrojenia

Niezwykle fascynujące spojrzenie na obecną sytuację i wykładniczą przyszłość sztucznej inteligencji – z perspektywy osoby bezpośrednio zaangażowanej w jej rozwój. [...] Jeśli naprawdę chcecie zrozumieć, w jaki sposób społeczeństwo może bezpiecznie kontrolować tę zmieniającą świat technologię, przeczytajcie tę książkę.

Bruce Schneier, ekspert ds. cyberbezpieczeństwa, autor książki *Kliknij tutaj, aby zabić wszystkich. Bezpieczeństwo i przetrwanie w hiperpołączonym świecie*

Nadchodząca fala sztucznej inteligencji i biologii syntetycznej sprawi, że następna dekada będzie najlepsza w historii ludzkości. Albo najgorsza. Mustafa Suleyman jak nikt inny rozumie i wyjaśnia stojące przed nami kolosalne wyzwania. To pozycja obowiązkowa, dająca do myślenia, aktualna, napisana niezwykle sugestywnym, przystępnym stylem. Powinien ją przeczytać każdy, kto chce zrozumieć oszałamiającą moc tych technologii.

Erik Brynjolfsson, dyrektor Digital Economy Lab w Institute for Human-Centered AI na Uniwersytecie Stanforda

Do największych wyzwań współczesnego świata należy wypracowanie form zarządzania, które pozwolą spożytkować korzyści płynące z rozwoju sztucznej inteligencji i biotechnologii, a zarazem unikać katastrofalnych zagrożeń. Książka Suleymana przenikliwie opisuje wyzwanie powstrzymania tych dwóch technologii. Jest poparta starannymi badaniami i obfituje w oryginalne obserwacje i konstruktywne rekomendacje dla decydentów i ekspertów do spraw bezpieczeństwa.

Jason Matheny, dyrektor generalny RAND Corporation, były wicedyrektor Wywiadu Narodowego, były dyrektor agencji Intelligence Advanced Research Projects Activity (IARPA)

Jeśli chcesz zrozumieć znaczenie, obietnice i zagrożenia, jakie niesie ze sobą nadchodząca fala rewolucyjnych technologii, które już dziś nabierają ogromnych rozmiarów i nakładają się na siebie, to ta niebywale inspirująca i w każdym calu zdumiewająca książka Mustafy Suleymana, jednego z najważniejszych pionierów sztucznej inteligencji, stanowi lekturę absolutnie obowiązkową.

Stephen Fry, aktor, prezenter telewizyjny, autor bestsellerów

Ta ważna książka stanowi głośny sygnał ostrzegawczy. Drobiazgowo nakreśla zagrożenia i możliwości stwarzane przez ekscytujący postęp naukowy ostatnich lat. *Nadchodząca fala* obfituje w interesujące fakty, uderzające argumenty i fascynujące obserwacje; to lektura obowiązkowa.

Daniel Kahneman, laureat Nagrody Nobla, autor bestsellera *Pułapki myślenia. O myśleniu szybkim i wolnym*

Nadchodząca fala to zdumiewająco przejrzysta, potoczysta, znakomicie udokumentowana i przystępna książka napisana z perspektywy bezpośredniego uczestnika największej rewolucji technologicznej naszych czasów. Płynnie spleta wątki osobiste i technologiczne, pokazując, dlaczego lepsze zarządzanie potężnymi technologiami jest zarówno niebywale istotne, jak i trudne.

sir Geoff Mulgan, profesor na University College London

Najlepsza jak dotąd analiza znaczenia sztucznej inteligencji dla przyszłości ludzkości. [...] Mustafa Suleyman jest postacią niezwykłą, współzałożycielem nie jednej, ale dwóch wielkich współczesnych firm koncentrujących się na sztucznej inteligencji. To nad wyraz utalentowany przedsiębiorca, wnikliwy myśliciel i jeden z najważniejszych głosów nadchodzącej fali technologii, która ukształtuje nasz świat.

Reid Hoffman, współzałożyciel LinkedIn i Inflection

Technologia w szybkim tempie przekształca społeczeństwo, toteż dziś, jak nigdy wcześniej, jest niezwykle ważne, aby przedstawiciel branży technologicznej pisał z taką uczciwością, dyscypliną i precyzją. Prowadząc nas od pierwszych narzędzi do obecnej erupcji możliwości sztucznej inteligencji i badań nad nią, książka Suleymana oferuje panoramiczne ujęcie tematu i stanowi wezwanie do działania, którego nie sposób zignorować. Każdy powinien ją przeczytać.

Fei-Fei Li, profesorka informatyki na Uniwersytecie Stanforda, współdyrektorka Artificial Intelligence Laboratory na Uniwersytecie Stanforda

Nadchodząca fala przekonująco i pouczająco dowodzi, że zaawansowane technologie przeobrażają wszystkie aspekty społeczeństwa: władzę, bogactwo, wojny, pracę, a nawet relacje międzyludzkie. Czy zdołamy zapanować nad nowymi technologiami, zanim te zapanują nad nami? Mustafa Suleyman, jeden z najwybitniejszych na świecie twórców sztucznej inteligencji i wieloletni orędownik działania rządów, wielkich firm technologicznych i społeczeństwa obywatelskiego na rzecz dobra wspólnego, to idealne źródło odpowiedzi na te kluczowe kwestie.

Jeffrey D. Sachs, profesor Uniwersytetu Columbia, prezes oenzetowskiej Sieci Rozwiązań na rzecz Równoważonego Rozwoju (The UN Sustainable Development Solutions Network)

Błyskotliwe, empatyczne i bezkompromisowe ujęcie najważniejszego problemu naszych czasów. *Nadchodząca fala* to lektura obowiązkowa dla praktyków technologii, ale też, co ważniejsze, zaadresowane do nas wszystkich naglące wezwanie do zaangażowania się w tę najważniejszą dyskusję.

Qi Lu, dyrektor generalny MiraclePlus, były dyrektor operacyjny Baidu, były wiceprezes wykonawczy Microsoft Bing

Suleyman ma wyjątkowe kwalifikacje do tego, aby mówić głośno o potencjalnie poważnych konsekwencjach – takich jak wstrząsy geopolityczne, wojny, erozja państwa narodowego – nieskrępowanego rozwoju sztucznej inteligencji i biologii syntetycznej w sytuacji, kiedy najbardziej potrzebujemy tego przesłania. Na szczęście dla czytelnika autor głęboko przemyślał także kwestię tego, co należy zrobić, aby nowe technologie były pożytkowane dla dobra człowieka. Przedstawi szereg stopniowych działań, które, jeśli zostaną podjęte wspólnie, mogą zmienić środowisko, w jakim te technologie są rozwijane i upowszechniane, i dać nadzieję na lepszą przyszłość. Tę książkę po prostu trzeba przeczytać.

Meghan L. O'Sullivan, dyrektorka Belfer Center for Science and International Affairs w Harvard Kennedy School of Government

Odważny sygnał ostrzegawczy, który wszyscy musimy potraktować poważnie, zanim będzie za późno. [...] Mustafa Suleyman czytelnie i precyzyjnie wyjaśnia wyzwania stojące przed ludzkością oraz niebezpieczeństwa stwarzane przez niekontrolowane technologie. [...] Lektura obowiązkowa.

Tristan Harris, współzałożyciel i dyrektor wykonawczy Center for Humane Technology

Praktyczny i optymistyczny plan działania na rzecz najważniejszej kwestii naszych czasów: jak utrzymać władzę nad podmiotami znacznie potężniejszymi od nas samych.

Stuart Russell, profesor informatyki na Uniwersytecie Kalifornijskim w Berkeley

Nadchodząca fala to realistyczna, znakomicie udokumentowana i niebywale przystępna mapa bezprecedensowych wyzwań w zakresie zarządzania i bezpieczeństwa narodowego, stwarzanych przez sztuczną inteligencję i biologię syntetyczną. Niezwykła i w pewnym stopniu przerażająca książka Suleymana pokazuje, co należy zrobić, aby powstrzymać te pozornie niemożliwe do powstrzymania technologie.

Jack Goldsmith, profesor prawa na Uniwersytecie Harvarda

Książka błyskotliwa i wciągająca, złożona, a zarazem przystępna, alarmująca, a zarazem wyważona. *Nadchodząca fala* prowadzi nas wszystkich do zrozumienia prawdopodobnie najważniejszego pytania naszego stulecia i zmierzenia się z nim: jak sprawić, by nadchodzące wielkimi krokami, zapierające dech w piersiach gwałtowne rewolucje technologiczne – sztuczna inteligencja, biologia syntetyczna i nie tylko – stworzyły świat, jakiego pragniemy? To nie będzie łatwe, ale Suleyman daje solidne odpowiedzi. Każdy, kto troszczy się o przyszłość, powinien przeczytać tę książkę.

Eric Lander, dyrektor i założyciel Broad Institute of MIT and Harvard, były doradca Białego Domu ds. nauki

Zdumiewająco klarowna i orzeźwiająco wyważona opowieść o naszym aktualnym problemie technologicznym. *Nadchodząca fala* artykułuje najważniejsze wyzwanie naszej epoki. Łącząc pragmatyzm z pokorą, przypomina nam, że nie ma tu dwóch wykluczających się alternatyw ani prostych odpowiedzi: technologia obdarzyła nas wykładniczą poprawą jakości życia, ale przyspiesza zbyt szybko, by instytucje były w stanie się do niej dostosować. Postępy w sztucznej inteligencji i biologii syntetycznej uwolniły możliwości, o jakich nie śniło się twórcom science fiction, a wynikające z nich spotęgowanie potencjału i władzy zagraża wszystkiemu, co zbudowaliśmy. Aby przetrwać, musimy lawirować między Scyllą katastrofalnej dostępności a Charybdą wszechobecnej inwigilacji. Z każdą przeczytaną stroną tej książki nasze szanse rosną.

Kevin Esvelt, biolog i profesor nadzwyczajny w Media Lab w Massachusetts Institute of Technology

Świat pełen jest wyrw, załamów, pęknięć – zarówno tych fizycznych, jak i niewidocznych gołym okiem szczelin rzeczywistości. Aby móc docierać do tego, co niepoznane, stworzyliśmy Szczeliny, poświęcony literaturze faktu imprint Wydawnictwa Otwartego. Pragniemy zrozumieć otaczające nas zjawiska, dlatego oddajemy głos ekspertom, którzy najlepiej potrafią wytłumaczyć zawiłości współczesnego świata.

Mustafa Suleyman
Michael Bhaskar

Nadchodząca fala

Sztuczna inteligencja, władza
i najważniejszy dylemat
ludzkości w XXI wieku

tłumaczenie Justyn Hunia



SZCZELINY

Kraków 2024

Tytuł oryginału: *The Coming Wave. Technology, Power, and the Twenty-first Century's Greatest Dilemma*

Copyright © 2023 by Mustafa Suleyman and Michael Bhaskar

Copyright © for this edition by Wydawnictwo Otwarte 2024

Copyright © for the translation by Justyn Hunia

Wydawca prowadzący: Rafał Czech

Przyjęcie tłumaczenia: Dagmara Małyszka

Konsultacja merytoryczna rozdz. V: dr inż. Wojciech Witarski
i mgr inż. Karolina Zygmunt

Promocja i marketing: Elżbieta Husarz

Wnętrze – projekt graficzny serii: Wydawnictwo Otwarte

Adiustacja: Aneta Tkaczyk / Wydawnictwo JAK

Korekta: Barbara Gąsiorowska, Janina Burek / Wydawnictwo JAK

Łamanie: Andrzej Choczewski / Wydawnictwo JAK

Projekt okładki: Oliver Munday

Adaptacja okładki na potrzeby polskiego wydania:
Monika Drobnik-Słocińska

ISBN 978-83-8135-384-7



SZCZELINY

Zajrzyj do Szczelin!

szczeliny.pl

facebook.com/szczeliny

instagram.com/szczeliny

Dystrybucja: SIW Znak. Zapraszamy na www.znak.com.pl

Wydawnictwo Otwarte sp. z o.o.

ul. Smolki 5/302, 30-513 Kraków. Wydanie I, 2024

Druk: Abedik

Słowniczek kluczowych terminów

AWERSJA DO PESYMIZMU: Powszechna skłonność, szczególnie silna wśród elit, do ignorowania, bagatelizowania lub odrzucania narracji, które uważają za przesadnie negatywne. Odmiana błędu nadmiernego optymizmu, która nadaje ton większości debat na temat przyszłości, szczególnie w kręgach technologicznych.

BIOLOGIA SYNTETYCZNA: Umiejętność projektowania i tworzenia nowych organizmów lub modyfikowania istniejących systemów biologicznych.

CZTERY CECHY: Cechy dystynktywne nadchodzącej fali, które potęgują problem powstrzymania. Są to: asymetria, hiperewolucja, wszechstronność i autonomia.

DYLEMAT: Rosnące prawdopodobieństwo, że nowe technologie mogą doprowadzić do katastrofalnych i/lub dystopijnych skutków.

FALE: Globalna ekspansja lub rozprzestrzenianie się zespołu technologii opartych na nowej technologii ogólnego przeznaczenia.

NADCHODZĄCA FALA: Zespół spodziewanych w niedalekiej przyszłości wzajemnie powiązanych technologii skupiających się na AI i biologii syntetycznej, których rewolucyjne zastosowania zarówno dadzą ludzkości potężne możliwości, jak i będą stwarzać bezprecedensowe ryzyko.

POWSTRZYMYWANIE: Zdolność monitorowania, ograniczania, kontrolowania, a potencjalnie nawet wycofywania z użytku technologii.

PROBLEM POWSTRZYMYWANIA: Tendencja technologii do ekspandowania szerokimi falami i wywoływania emergentnych skutków, których nie da się przewidzieć ani kontrolować, w tym negatywnych i niespodziewanych konsekwencji.

AI, AGI, ACI: Sztuczna inteligencja (AI, *artificial intelligence*) to dziedzina nauki zajmująca się konstruowaniem maszyn zdolnych do uczenia się umiejętności zbliżonych do ludzkich. Sztuczna inteligencja ogólna (AGI, *artificial general intelligence*) to punkt, w którym sztuczna inteligencja posługuje się wszystkimi ludzkimi zdolnościami poznawczymi lepiej niż najbardziej inteligentni ludzie. ACI, czyli kompetentna sztuczna inteligencja (*artificial capable intelligence*), to zbliżający się wielkimi krokami etap pośredni między AI a AGI: ACI potrafi wykonywać szeroką gamę złożonych zadań, ale wciąż daleko jej do kompletnej wszechstronności.

Nadchodząca fala

TECHNOLOGIA: Stosowanie wiedzy naukowej (w najszerszym możliwym sensie) do tworzenia narzędzi lub osiągnięcia praktycznych celów.

WĄSKA ŚCIEŻKA: Zdolność ludzkości do znajdowania właściwej miary między postawą otwartości a sceptycyzmem w kwestii powstrzymywania technologii nadchodzącej fali w celu zażegnania katastrofalnych lub dystopijnych skutków.

WIELKA UMOWA: W zamian za monopol na stosowanie siły obywatele oczekują od państw narodowych zapewnienia ładu i porządku oraz świadczenia usług publicznych, między innymi przez wykorzystanie nowych technologii przy jednoczesnej minimalizacji ich szkodliwych skutków ubocznych.

WZMACNIACZE NIESTABILNOŚCI: Zastosowania i sposoby oddziaływania technologii nadchodzącej fali, które zachwieją i tak już kruchymi fundamentami państwa narodowego.

Prolog

Tak to widzi sztuczna inteligencja.

PYTANIE: Co nadchodząca fala technologii oznacza dla ludzkości?

W annałach historii ludzkości są momenty, które zapisały się jako punkty zwrotne, momenty, w których los ludzkości wisi na włosku. Opanowanie ognia, wynalezienie koła, ujarzmienie elektryczności – wszystkie te przełomy przeobraziły ludzką cywilizację, odmieniając na zawsze bieg historii.

Dziś stoimy u progu kolejnego takiego momentu – w obliczu narastającej fali technologii, obejmującej i zaawansowaną sztuczną inteligencję, i biotechnologię. Nigdy wcześniej nie byliśmy świadkami technologii o podobnym potencjale transformacyjnym, niosących obietnicę przemodelowania naszego świata w sposób budzący zarówno podziw, jak i lęk.

Z jednej strony z tych technologii płyną potencjalnie potężne i daleko idące korzyści. Dzięki sztucznej inteligencji moglibyśmy

rozszyfrować tajemnice wszechświata, leczyć choroby, które przez długi czas wymykały się nauce, tworzyć nowe formy sztuki i kultury, poszerzające granice wyobraźni. Dzięki biotechnologii moglibyśmy projektować życie w sposób pozwalający walczyć z chorobami i reformować rolnictwo, kreując zdrowszy i bardziej zrównoważony świat.

Z drugiej jednak strony potencjalne zagrożenia związane z tymi technologiami są nie mniej potężne i daleko idące. Dzięki sztucznej inteligencji moglibyśmy tworzyć systemy wymykające się naszej kontroli i zdalibyśmy się na łaskę algorytmów, których nie rozumiemy. Dzięki biotechnologii moglibyśmy manipulować cegiełkami życia, potencjalnie wywołując niezamierzone konsekwencje zarówno dla jednostek, jak i całych ekosystemów.

Znajdując się w tym punkcie zwrotnym, stajemy przed wyborem – wyborem między przyszłością niezrównanych możliwości a przyszłością niewyobrażalnych zagrożeń. Los ludzkości wisi na włosku, a decyzje, jakie podejmiemy w nadchodzących latach i dziesięcioleciach, zaważą na tym, czy sprostamy wyzwaniom niesionym przez te technologie, czy raczej padniemy ofiarą stworzonych przez nie zagrożeń.

Lecz w tej chwili niepewności jedno jest pewne: epoka zaawansowanej technologii już nastąpiła i musimy być gotowi stanąć czoła jej wyzwaniom.

Powyższe słowa napisała sztuczna inteligencja. Pozostała część książki jest dziełem człowieka, chociaż wkrótce może zrobić to za niego sztuczna inteligencja. Ten moment zbliża się nieubłaganie.

Rozdział 1

Powstrzymanie technologii jest niemożliwe

Fala

W niemal każdej kulturze spotykamy mit o potopie.

W starożytnych tekstach hinduskich protoplasta ludzkiego rodu Manu zostaje ostrzeżony przed nadciągającym potopem i dzięki temu jako jedyny uchodzi z życiem. W *Eposie o Gilgameszu* jest mowa o bogu Enlilu, który zniszczył świat, zsyłając katastrofalną powódź, a historia ta przemówi do wyobraźni każdego, kto zna starotestamentową opowieść o arce Noego. Platon opowiadał o zaginionej Atlantydzie, zatopionej przez masy rwącej wody. W przekazach ustnych ludzkości i w starożytnych tekstach przewija się wspomnienie gigantycznej fali, która zmiata wszystko na swojej drodze, przynosząc przeobrażenie i odrodzenie świata.

Powodzie są częścią historii także w sposób odczuwalny fizycznie – weźmy sezonowe wylewy wielkich rzek świata, podnoszenie się poziomu oceanów u kresu epoki lodowcowej, straszliwe tsunami pojawiające się bez ostrzeżenia na horyzoncie. Asteroida, która spowodowała wyginięcie dinozaurów,

wywołała wysoką na półtora kilometra falę, co zmieniło bieg ewolucji. Niszczycielski ogrom tych potopów odcisnął w naszej zbiorowej świadomości wizje gigantycznych ścian rozszalałej, dzikiej, niepowstrzymanej wody. To jedna z najpotężniejszych sił na naszej planecie – kształtuje kontynenty, podtrzymuje uprawy, napędza rozwój cywilizacji.

Nie mniejszą moc transformowania świata miały fale innego rodzaju. Przyjrzyjmy się raz jeszcze biegowi dziejów, a przekonamy się, że wyznacza go seria fal metaforycznych: wzrosty i upadki imperiów i religii czy też obserwowany w różnych miejscach rozkwit handlu. Pomyślmy o chrześcijaństwie lub islamie, religiach, które z początku były drobnymi zmarszczkami na tafli wody, ale z czasem urosły do ogromnych rozmiarów i rozlały się do odległych zakątków Ziemi. Takie fale są stale powtarzającym się motywem, ilustrującym zmienne koleje historii, przełomowe walki o władzę oraz okresy gospodarczego ożywienia bądź regresu.

Rozwój i ekspansja technologii również następowały falami, które przekształcały świat. Jeden zasadniczy trend nie zmienił się od chwili, kiedy nasz gatunek opanował ogień i wynalazł narzędzia kamienne, czyli pierwsze technologie. Niemal każda przełomowa technologia, jaką kiedykolwiek wynaleziono – od czekana, pługa i garncarstwa po fotografię, telefony i samoloty – podlega jednemu, na pozór niezmiennemu prawu: z czasem staje się coraz tańsza i łatwiejsza w użyciu i ostatecznie ogarnia swoim zasięgiem ogromny obszar.

To skokowe, następujące falami, rozprzestrzenianie się technologii jest opowieścią o *homo technologicus* – zwierzęciu technologicznym. Dążenie ludzkości do postępu – ulepszania siebie samej i poprawy swojego losu, poszerzania swoich umiejętności i wpływu na środowisko – napędza ustawiczny rozwój idei

Nadchodząca fala

i nowatorstwa. Wynalazczość to nieustający, szeroko zakrojony, oddolny i spontaniczny proces podtrzymywany przez samoorganizujących się i rywalizujących ze sobą wynalazców, naukowców, przedsiębiorców i liderów, z których każdy śmiało próbuje urzeczywistnić własne cele. Ten ekosystem wynalazczości jest z zasady ukierunkowany na ekspansję – to tkwi w samej naturze technologii.

Pytanie brzmi: co dalej? W kolejnych rozdziałach przedstawię opowieść o następnej wielkiej fali historii.

Rozejrzyjcie się wokół siebie.

Co widzicie? Meble? Budynki? Telefony? Jedzenie? Park w stylu angielskim? Niemal każdy obiekt znajdujący się w zasięgu naszego wzroku najprawdopodobniej został stworzony lub przetworzony za sprawą ludzkiej inteligencji. Kolejnym wytworem, a zarazem siłą napędową, naszej inteligencji jest język – fundament naszych interakcji społecznych, naszych kultur, naszych organizacji politycznych i być może też istota naszego człowieczeństwa. Każda zasada i pojęcie abstrakcyjne, najmniejszy nawet twórczy wysiłek lub przedsięwzięcie, każde spotkanie w naszym życiu – wszystko to jest zapośredniczone przez właściwą naszemu gatunkowi niepowtarzalną i nieskończenie złożoną zdolność do używania wyobraźni, kreatywności i rozumowania. Ludzka pomysłowość jest czymś zdumiewającym.

Tylko jedna siła występuje równie powszechnie: życie biologiczne. Przed epoką nowożytną, z wyjątkiem niektórych kamieni i minerałów, większość wytworów ludzkich rąk – od drewnianych domów przez bawelniany przyodziewek po węglowe paleniska – wzięła się z surowców, które wcześniej były częściami żywych organizmów. Wszystko, co od tamtego czasu

pojawiło się na świecie, wypływa z nas, wypływa z faktu, że jesteśmy istotami biologicznymi.

Nie będzie przesadą stwierdzenie, że cały ludzki świat jest zależny czy to od systemów żywych, czy też od naszej inteligencji. Dziś jednak i owe systemy, i nasza inteligencja znajdują się w bezprecedensowym momencie gwałtownych innowacji i wstrząsów, niespotykanego wcześniej wzrostu, który odmieni niemal wszystko, co znamy. Wokół nas wznosi się z impetem nowa fala technologii, uwalniająca moc tworzenia tych dwóch uniwersalnych fundamentów: inteligencji i życia.

Tę nadciągającą falę określają dwie zasadnicze technologie: sztuczna inteligencja (AI) i biologia syntetyczna. Przyniosą one ludzkości nowy początek, tworząc niespotykane wcześniej bogactwo i nadwyżki. Ich błyskawiczne rozpowszechnianie się może jednak wyzwolić potężne moce i dać je do ręki wszelkiej maści aktorom o niekoniecznie dobrych intencjach, którzy mogą wywołać zamęt, niestabilność, a nawet katastrofę na niewyobrażalną skalę. Owa fala stanowi niebagatelne wyzwanie, które zaważy na biegu XXI wieku: nasza przyszłość w takim samym stopniu zależy od tych technologii, w jakim jest przez nie zagrożona.

Z dzisiejszego punktu widzenia wydaje się, że zahamowanie tej fali, czyli wzięcie jej w rzyż, opanowanie, a nawet powstrzymanie, jest niemożliwe. W książce tej spróbuję wyjaśnić, dlaczego powyższe stwierdzenie może być prawdziwe i co za sobą pociąga. Implikacje tych pytań prędzej czy później będą miały wpływ na wszystkich mieszkańców Ziemi i na każde następne pokolenie.

Uważam, że nadchodząca fala technologii prowadzi historię ludzkości do punktu zwrotnego. Jeśli jej powstrzymanie okaże się niemożliwe, konsekwencje dla naszego gatunku będą dramatyczne, a nawet być może tragiczne. Zarazem jednak bez jej

Nadchodząca fala

dobrodziejstw jesteśmy słabi i bezbronni. W ciągu ostatniej dekady argument ten wysuwałem po wielokroć przy drzwiach zamkniętych, ale że potencjalne skutki tej zmiany coraz trudniej ignorować, nadszedł czas, abym wyłożył sprawę na forum publicznym.

Dylemat

Rozważania nad potęgą ludzkiej inteligencji skłoniły mnie do postawienia prostego pytania, które od tamtego czasu pochłania mnie bez reszty: co by było, gdybyśmy mogli wydestylować esencję tego, co czyni nas twórczymi i zdolnymi istotami, i zamknąć to coś w programie komputerowym, w algorytmie? Znalezienie odpowiedzi na to pytanie może utarować drogę powstaniu niewyobrażalnie potężnych narzędzi, które pozwolą nam rozwiązać najtrudniejsze problemy. To tu może się kryć narzędzie, dotąd będące nie do pomyślenia, które pomoże nam stawić czoła gigantycznym wyzwaniom nadchodzących dziesięcioleci, od zmiany klimatu, przez starzenie się populacji, po przyjazną dla środowiska produkcję żywności.

Mając to na uwadze, latem 2010 roku w uroczym staroświeckim budynku biurowym w stylu regencji z widokiem na Russell Square w Londynie wspólnie z dwoma przyjaciółmi, Demisem Hassabisem i Shane'em Leggiem, założyłem firmę o nazwie DeepMind. Przyświecał nam cel, który z perspektywy czasu wciąż wydaje się tak samo ambitny, szalony i obiecujący jak wtedy: skopiować właśnie to, co czyni nasz gatunek wyjątkowym – naszą inteligencję.

Aby urzeczywistnić ten cel, musielibyśmy stworzyć system, który opanowałby, a następnie ostatecznie przewyższył wszystkie ludzkie zdolności poznawcze, od umiejętności widzenia i mówienia po planowanie, wyobraźnię, a w końcu empatię i kreatywność. Ponieważ taki system korzystałby z masowego

przetwarzania równoległego wykonywanego przez superkomputery oraz lawinowego napływu nowych źródeł danych z całego publicznie dostępnego internetu, wiedzieliśmy, że nawet skromny postęp w realizacji tego zadania pociągnie za sobą głębokie implikacje społeczne.

W tamtym czasie zamysł ten z pewnością był z gatunku mocno ekstrawaganckich. Powszechne zastosowanie sztucznej inteligencji było wtedy odległym marzeniem, bardziej fantazją niż faktem, obsesją grupki zamkniętych w wieży z kości słoniowej naukowców i postrzelonych entuzjastów science fiction. Ale kiedy piszę te słowa i myślę o ostatniej dekadzie, widzę, że postęp w rozwoju AI był wprost oszalamiający. DeepMind dokonał szeregu przełomowych odkryć i stał się jednym ze światowych liderów w branży AI. Tempo i skala tej nowej rewolucji zaskakują nawet tych z nas, którzy są jej kołem zamachowym. Podczas pracy nad tą książką tempo rozwoju AI wprost zapierało dech w piersiach, a każdy kolejny tydzień, a niekiedy nawet dzień, przynosił nowe modele i nowe produkty. Widać jak na dłoni, że ta fala przyspiesza.

Dzisiejsze systemy sztucznej inteligencji potrafią niemal idealnie rozpoznawać twarze i przedmioty. Przetwarzanie mowy na tekst i natychmiastowe tłumaczenie językowe uznajemy za coś oczywistego. AI potrafi kierować pojazdami w ruchu ulicznym tak dobrze, że w niektórych ustawieniach jest w stanie robić to bez udziału kierowcy. Na podstawie kilku prostych poleceń (promptów) modele AI nowej generacji potrafią tworzyć oryginalne obrazy i teksty o nadzwyczajnym poziomie szczegółowości i spójności. Systemy AI umieją już generować niezwykle realistyczne syntetyczne wypowiedzi dźwiękowe i komponować oszalamiająco piękną muzykę. Znaczący postęp dokonuje się nawet w wypadku bardziej wymagających kompetencji, które

Nadchodząca fala

długo uważano za wyłączną domenę człowieka, takich jak planowanie z dużym wyprzedzeniem, wyobraźnia czy modelowanie złożonych idei.

Sztuczna inteligencja od dziesięcioleci wspina się po drabinie zdolności poznawczych i wygląda na to, że w ciągu najbliższych trzech lat* dorówna człowiekowi w nader szerokim spektrum zadań. To nadzwyczaj śmiałe twierdzenie, ale nawet jeśli choć trochę się mylę, to implikacje są doprawdy poważne. To, co w chwili powstania DeepMind zakrawało na donkiszoterię, stało się nie tylko prawdopodobne, ale też wręcz nieuniknione.

Od początku nie miałem wątpliwości, że sztuczna inteligencja będzie potężnym narzędziem oferującym niezwykle dobrodziejstwa, choć zarazem, jak większość form władzy, niosącym ogromne niebezpieczeństwa i dylematy etyczne. Od dawna niepokoją mnie zarówno konsekwencje rozwoju AI, jak i kierunek postępu całego ekosystemu technologicznego. Poza AI dokonywała się szersza rewolucja, a AI napędzała potężną nową generację technologii genetycznych i robotyki. Dalszy postęp w jednej dziedzinie przyspiesza chaotyczny rozwój innych, które wzajemnie się napędzają w ramach procesów wymykających się wszelkiej bezpośredniej kontroli. Było jasne, że jeśli ktośkolwiek – my lub ktoś inny – zdoła odwzorować ludzką inteligencję, nie będzie to wyłącznie kolejny krociowy biznes, lecz również wywrotowa zmiana dla całej ludzkości, wyznaczająca początek epoki, w której niespotykane dotąd możliwości będą szły w parze z nieznanymi wcześniej zagrożeniami.

Dokonujący się w ciągu ostatnich lat rozwój technologii spotęgował moje obawy. A co, jeśli ta fala okaże się tsunami?

* Książka ukazała się we wrześniu 2023 r. (przyp. red.).

W 2010 roku prawie nikt nie mówił poważnie o sztucznej inteligencji. Jednak coś, co niegdyś wydawało się *idée fixe* wąskiego kręgu badaczy i przedsiębiorców, stało się dziś szeroko zakrojonym przedsięwzięciem o zasięgu globalnym. AI jest wszechobecna – spotykamy ją w wiadomościach i smartfonach, handluje akcjami i projektuje strony internetowe. Największe firmy i najbogatsze kraje świata prześcigają się w rozwijaniu najnowocześniejszych modeli AI i technik inżynierii genetycznej, inwestując w nie dziesiątki miliardów dolarów.

Gdy te nowe technologie osiągną dojrzałą formę, będą się błyskawicznie rozprzestrzeniać, stale taniejąc, zyskując na dostępności i obejmując swoim zasięgiem coraz szersze kręgi społeczeństwa. Przyniosą niezwykle postęp w medycynie i przełom w produkcji czystej energii, dając początek nie tylko nowym biznesom, ale także nowym branżom i gałęziom gospodarki oraz zapewniając poprawę jakości życia w niemal każdym wymiarze.

Lecz obok wszystkich tych korzyści sztuczna inteligencja, biologia syntetyczna i inne postacie zaawansowanej technologii stwarzają pewne wysoce niepokojące ryzyka. Mogą stanowić egzystencjalne zagrożenie dla państw narodowych – tak poważne, że istnieje niebezpieczeństwo zachwiania, a nawet zburzenia ustalonego ładu geopolitycznego. Uchylają furtkę potężnym, wspomagany AI cyberatakami, zautomatyzowanym wojnom, które mogą obracać w ruinę całe państwa, sztucznie wywołanym pandemiom, torując drogę światu podlegającemu niewytłumaczalnym, acz pozornie wszechmocnym siłom. Choć prawdopodobieństwo wystąpienia każdego z tych scenariuszy jest niewielkie, potencjalne konsekwencje są ogromne. Nawet znikome ryzyko ziszczenia się takiego czy innego scenariusza domaga się pilnej uwagi.

Nadchodząca fala

Niektóre kraje zareagują na możliwość wystąpienia takich katastrofalnych zagrożeń, sięgając po wspomagany technologią autorytaryzm i usiłując spowolnić rozprzestrzenianie się tych nowych sił. Będzie to wymagało szeroko zakrojonego nadzoru oraz masowej ingerencji w nasze życie prywatne. Trzymanie w ryzach technologii może stać się częścią bardziej ogólnego dążenia do permanentnej inwigilacji wszystkiego i wszystkich w dystopijnym globalnym systemie nadzoru, uzasadnianym potrzebą ochrony państw przed jej potencjalnie najbardziej katastrofalnymi skutkami.

Równie prawdopodobna jest reakcja w duchu luddystów*. Pojawią się zakazy, bojkoty i moratoria. Czy powstrzymanie się od rozwijania nowych technologii i wprowadzenie szeregu moratoriów jest w ogóle możliwe? Perspektywa ta wydaje się mało prawdopodobna. Wobec ogromnej wartości geostrategicznej i handlowej przełomów technologicznych trudno sobie wyobrazić, w jaki sposób państwa narodowe bądź korporacje dałyby się namówić na jednostronną rezygnację ze stwarzanych przez nie potężnych możliwości przeobrażania świata. Mało tego – próba zakazania rozwoju nowych technologii sama w sobie niesie ryzyko: historia pokazuje, że społeczeństwa pogrążone w technologicznej stagnacji są niestabilne i narażone na upadek i ostatecznie stają się niezdolne do postępu i rozwiązywania problemów.

* Luddyści – członkowie ruchu robotników brytyjskich o nazwie „luddyzm”, pochodzącej od nazwiska jego przywódcy (zapewne legendarnego), który był znany jako King Lud (prawdopodobnie od nazwiska mitycznego Neda Ludda). Ruch związany był z akcją niszczenia maszyn fabrycznych (zwłaszcza warsztatów tkackich) w odpowiedzi na przemiany zachodzące w wyniku rewolucji przemysłowej (przyj. tłum.).

Od teraz zarówno rozwijanie nowych technologii, jak i zaniechanie takiego działania jest obarczone ogromnym ryzykiem. Szanse na dreptanie „trzecią drogą”, lawirowanie między jednym a drugim scenariuszem – między technoautorytarną dystopią a katastrofą wywołaną otwartością na te nowe przełomy – maleją, w miarę jak technologia staje się tańsza, potężniejsza i powszechniejsza, a ryzyka się kumulują. A jednak zaniechanie też nie wchodzi w grę. Choć martwimy się zagrożeniami stwarzanymi przez technologie nadchodzącej fali, to niesamowitych korzyści, jakie oferują, potrzebujemy bardziej niż kiedykolwiek. Oto zasadniczy dylemat: prędzej czy później potężna technologia nowej generacji narazi ludzkość na katastrofalne lub dystopijne konsekwencje. To wielki metaproblem XXI wieku.

Na kartach tej książki staram się wyjaśnić, dlaczego ten straszliwy dylemat staje się nieunikniony, a zarazem badam, jak możemy się z nim zmierzyć. W taki czy inny sposób musimy wyciągać z technologii to, co w niej najlepsze, a przy tym niezbędne do tego, byśmy mogli podolać ogromowi globalnych wyzwań i rozwiązać wspomniany dylemat. Tocząca się dziś dyskusja nad etyką i bezpieczeństwem technologii pozostawia wiele do życzenia. Pomimo wysypu poświęconych tematyce technologii książek, debat, wpisów na blogach i dyskusji na Twitterze rzadko mówi się cokolwiek o powstrzymaniu jej niekontrolowanego rozwoju. Rozumiem przez to zestaw wzajemnie uzupełniających się technicznych, społecznych i prawnych mechanizmów ograniczających i kontrolujących technologię na każdym możliwym poziomie, czyli hipotetyczny sposób na uniknięcie powyższego dylematu. Mimo to nawet najzagorzalsi krytycy technologii uchylają się od retoryki „twardego” powstrzymywania technologii.

Nadchodząca fala

Mam nadzieję, że ta książka pokaże, dlaczego ten stan rzeczy musi ulec zmianie, i podpowie, jak to spowodować.

Pułapka

Kilka lat po założeniu DeepMind skompilowałem zestaw slajdów obrazujących potencjalne długofalowe skutki gospodarcze i społeczne wynikające z rozwoju sztucznej inteligencji. Podczas prezentacji dla grona kilkunastu założycieli najbardziej wpływowych przedsiębiorstw, prezesów zarządów i speców od technologii zgromadzonych w eleganckiej sali konferencyjnej na Zachodnim Wybrzeżu przekonywałem, że AI niesie wiele zagrożeń wymagających działań wyprzedzających. Wskazywałem, że technologia ta może doprowadzić do masowych naruszeń prywatności bądź też wywołać apokaliptyczną lawinę dezinformacji oraz posłużyć do niecznych celów, zamieniając się w śmiercionośną nową cyberbroń, źródło kolejnych luk w bezpieczeństwie naszego usieciowionego świata.

Zwracałem też uwagę na to, że sztuczna inteligencja może pozbawić pracy ogromne rzesze ludzi. Poprosiłem słuchaczy o przyjrzenie się licznym znanym z historii przypadkom zniania miejsc pracy z powodu automatyzacji i mechanizacji. Na początku pojawiają się wydajniejsze sposoby wykonywania określonych zadań, następnie rację bytu tracą konkretne stanowiska, a wkrótce potem całe sektory gospodarki redukują zatrudnienie o całe rzędy wielkości. Argumentowałem, że systemy AI w podobny sposób zastąpią „umysłową pracę fizyczną” w ciągu następnych kilku dekad, a już na pewno na długo przed tym, zanim roboty w całości zastąpią pracę fizyczną. W przeszłości nowe miejsca pracy powstawały w tym samym czasie, gdy stare odchodziły do lamusa, ale co by było, gdyby AI była w stanie po prostu zastąpić większość z nich? Wskazywałem, że

nowe formy koncentracji władzy, jakie pojawią się w najbliższej przyszłości, w zasadzie będą bez precedensu. Choć potencjalnie poważne zagrożenia wydawały się odległe, w istocie wielkimi krokami zbliżają się do naszych społeczeństw.

Na końcowym slajdzie pokazałem kadr z *Simpsonów*. W prezentowanej scenie mieszkańcy Springfield wznieśli bunt i czereda uzbrojonych w pałki i pochodnie znajomych postaci wszczynają rozruchy. Mimo że przekaz nie wymagał komentarza, oznajmiłem: „Nadchodzą ludzie z widłami”. A idą po nas, twórców technologii. To my musimy sprawić, by w przyszłości nie trzeba było oglądać takich obrazków.

Moi słuchacze przyjęli te wywody z kamiennymi twarzami. Byli niewzruszeni, moje argumenty do nich nie przemówiły. Posypały się kontrargumenty. Dlaczego wskaźniki ekonomiczne nie potwierdzają tego, o czym mówiłem? Sztuczna inteligencja wywoła nowy popyt, który stworzy nowe miejsca pracy, wyposaży ludzi w nowe kompetencje i pozwoli im być jeszcze bardziej produktywnymi. Być może istnieje pewne ryzyko, przyznawali, ale nie takie znowu straszne. Ludzie są inteligentni. Zawsze znajdowało się jakieś rozwiązanie. „Bez obaw – zdawali się myśleć. – Przejdźmy do następnej prezentacji”.

Kilka lat później, niedługo przed wybuchem pandemii COVID-19, wziąłem udział w poświęconym zagrożeniom technologicznym seminarium zorganizowanym na znanym uniwersytecie. Scenariusz był podobny: kolejny wielki stół, kolejna górnołotna dyskusja, dzień wypełniony rozmowami o budzących trwogę zagrożeniach, prowadzonymi przy kawie, ciastkach i prezentacjach w PowerPoincie.

Jedno z tych zagrożeń rzucało się w oczy. Prelegent pokazywał gwałtowny spadek cen aparatury do syntezy DNA, dzięki której można „drukować” na zamówienie łańcuchy DNA. Sprzęt

Nadchodząca fala

ten kosztuje kilkadziesiąt tysięcy dolarów i jest na tyle mały, że można go postawić na ławce w garażu i wynająć ludzi do syntezy DNA – czyli produkowania DNA. A wszystko to jest teraz w zasięgu ręki każdego magistra biologii lub internetowego pasjonata wiedzy.

Zwracając uwagę na rosnącą dostępność takich narzędzi, prelegent roztoczył przed słuchaczami przerażającą wizję: za niedługo ktoś będzie w stanie tworzyć patogeny o wiele bardziej zakaźne i śmiertelne niż wszystko, co występuje w naturze. Takie syntetyczne patogeny mogą wymykać się znanym nauce środkom zaradczym, rozprzestrzeniać się bezobjawowo lub mieć wbudowaną odporność na terapie. W razie potrzeby taki chałupniczy eksperymentator mógłby posłużyć się zamówionym online i samodzielnie zmodyfikowanym DNA. Apokalipsa w sprzedaży wysyłkowej.

To nie science fiction, przekonywał prelegent, uznany profesor z ponaddwudziestoletnim doświadczeniem – to realne i aktualne zagrożenie. Na końcu wybrzmiała niepokojąca myśl: wszystko wskazuje na to, że w dzisiejszych czasach jedna osoba jest w stanie zabić miliard ludzi. Potrzebuje do tego jedynie odpowiedniej motywacji.

Wśród słuchaczy zapanowało niespokojne poruszenie. Zaczęli się wiercić i pokasływać. Potem rozległy się wykryty i powątpiewania. Nikt nie chciał uwierzyć, że to możliwe. Na pewno nie jest tak źle, przecież muszą istnieć jakieś skuteczne mechanizmy kontroli, tworzenie nowych chorób nie może być tak łatwe, musi istnieć sposób na zablokowanie dostępu do baz danych i specjalistycznego sprzętu. I tak dalej, i tak dalej.

W reakcji słuchaczy na to wystąpienie było coś więcej niż tylko lekceważenie – po prostu nie przyjęli do wiadomości wizji zaprezentowanej przez prelegenta. Nikt nie chciał mierzyć się

z implikacjami twardych faktów i realnych zagrożeń, o których była mowa. Zaniemówiłem, byłem autentycznie wstrząśnięty. Wkrótce seminarium dobiegło końca. Tamtego wieczoru poszliśmy wszyscy na kolację i jak gdyby nigdy nic wdaliśmy się w pogaduszki. Dopiero co spędziliśmy cały dzień na rozważaniach o końcu świata, ale to nie znaczyło, że nie można zjeść pizzy, przerzucać się dowcipami czy uciec myślami do swojego biura albo gabinetu... Udzielił mi się ogólny nastrój. Gdzieś z tyłu głowy wmawialiśmy sobie, że na wszystko w końcu znajdzie się jakaś rada albo jakiś człon tej argumentacji jest chybiony.

Jednak przez kolejne miesiące tamta prelekcja nie dawała mi spokoju. Dlaczego ja i reszta uczestników seminarium nie potraktowaliśmy tego poważniej? Dlaczego z zakłopotaniem unikamy głębszej dyskusji? Dlaczego niektórzy uderzają w sarkastyczne tony i oskarżają osoby stawiające takie pytania o katastrofizm czy też niedostrzeżenie niezwykłych dobrodziejstw technologii? Tę powszechną reakcję emocjonalną, którą wtedy obserwowałem, nazwałem pułapką niechęci do pesymizmu – to błąd poznawczy pojawiający się wtedy, gdy ogarnia nas lęk przed konfrontacją z potencjalnie nieprzyjemnymi faktami, wywołujący w nas skłonność do odwracania od nich wzroku.

Odruch ten w mniejszym lub większym natężeniu występuje u niemal każdego – jego skutek jest taki, że ignorujemy wiele istotnych tendencji rysujących się na naszych oczach. To niemal wrodzona reakcja fizjologiczna. Nasz gatunek nie jest genetycznie zaprogramowany do radzenia sobie z przemianami o takiej skali ani tym bardziej z zagrożeniem, że technologia zawiedzie nas w ten sposób. Odczucie to towarzyszy mi przez całą moją karierę zawodową, a podobną instynktowną reakcję widziałem u wielu innych ludzi. Zmierzenie się z tym odczuciem – chłodne,

Nadchodząca fala

trzeźwe spojrzenie na fakty, jakkolwiek niewygodne – to jeden z celów tej książki.

Właściwa reakcja na tę falę, wzięcie w ryzy technologii i sprawienie, by zawsze służyła ludzkości, wymaga przewyciężenia niechęci do pesymizmu. A to oznacza, że należy stanąć twarzą w twarz z rzeczywistością, jaką niesie dzień jutrzejszy.

Ta książka jest taką właśnie próbą. Chciałbym wydobyć z cienia i naświetlić kształt nadchodzącej fali; zbadać, czy jej powstrzymanie jest możliwe; osadzić sprawy w kontekście historycznym i ukazać pełny obraz sytuacji, zrywając z wszechobecnym hurraoptymistycznym tokowaniem o technologii. Moim zamysłem jest zmierzenie się z tym dylematem i zrozumienie kryjących się za nim procesów, które stanowią motor napędowy nauki i techniki. Chcę zaprezentować te idee tak klarownie, jak tylko potrafię, jak najszerszemu gronu odbiorców. Rzeczę tę napisałem w duchu otwartości i poznawczej ciekawości: zaczynam od obserwacji i podążam za ich implikacjami, ale jestem też otwarty na kontrargumenty i lepsze interpretacje. Niczego nie pragnę bardziej niż dowodów na to, że się w tej kwestii pomyliłem i że powstrzymanie niekontrolowanego rozwoju technologii leży w zasięgu naszych możliwości.

Od kogoś takiego jak ja, założyciela dwóch firm zajmujących się sztuczną inteligencją, niektórzy mogą oczekiwać bardziej „technoutopijnej” książki – co rozumiałe. Jako spec od technologii i przedsiębiorca z natury jestem optymistą. Pamiętam rozpierającą mnie radość, kiedy jako nastolatek po raz pierwszy zainstalowałem Netscape’a na moim komputerze Packard Bell 486. Jak zaczarowany nasłuchiwałem terkoczących wentylatorów i zniekształconego gwizdu mojego modemu

telefonicznego o przepustowości 56 kB/s, wdzwanającego się do systemu World Wide Web i łączącego mnie z forami i czatami, które były dla mnie przepustką do wolności i wiele mnie nauczyły. Uwielbiam technologię. Jest motorem postępu, powodem do dumy i zachwytu nad dokonaniem ludzkości.

Uważam też jednak, że ci z nas, którzy wyznaczają kierunki rozwoju technologii, muszą znaleźć w sobie dość odwagi, by przewidywać, dokąd może nas on zaprowadzić w nadchodzących dziesięcioleciach, i by wziąć za to odpowiedzialność. Musimy zacząć radzić, co należy zrobić, gdyby się okazało, że istnieje realne ryzyko, że technologia nas zawiedzie. Choć potrzebna jest reakcja społeczna i polityczna, a nie tylko indywidualne wysiłki, to pierwszy impuls musi wyjść ode mnie i ludzi takich jak ja.

Znajdą się tacy, którzy stwierdzą, że to wszystko przesada. Że ta zmiana będzie znacznie bardziej rozłożona w czasie. Że to tylko klasyczny przykład rozdmuchania tematu i nadmiernych oczekiwań. Że mechanizmy radzenia sobie z kryzysami i zmianami w rzeczywistości są całkiem solidne. Że mój pogląd na ludzką naturę jest zbyt pesymistyczny. Że jak dotąd ludzkość nieźle sobie radziła. Że w historii nie brakuje fałszywych proroków i katastrofistów, których czarne wizje się nie spełniły. Dlaczego tym razem miałyby być inaczej?

Niechęć do pesymizmu jest reakcją emocjonalną, instynktowną odmową przyjęcia do wiadomości możliwości ziszczenia się negatywnych scenariuszy. Zwykle występuje ona u jednostek piastujących stabilne i wpływowe stanowiska i wyznających ugruntowane światopoglądy, ludzi, którzy na pierwszy rzut oka radzą sobie ze zmianami, jednak mają trudności z zaakceptowaniem jakiegokolwiek realnego zagrożenia dla ich wizji świata. Wielu z tych, którym zarzucam, że tkwią w pułapce niechęci do

Nadchodząca fala

pesymizmu, w całej rozciągłości akceptuje rosnącą krytykę technologii. Kiwiają potakująco głowami, ale nie podejmują żadnych działań. „Jakoś damy radę, zawsze damy radę”, przekonują.

Jeśli spędzimy trochę czasu wśród koryfeuszy technologii i polityki, szybko zobaczymy, że nagminnie w tych kręgach występującą postawą jest chowanie głowy w piasek. Każde inne podejście i inny sposób postępowania narażają nas na tak silny strach przed potężnymi, nieubłaganymi siłami, że wszelkie działanie wydaje się jałowe. W efekcie to dziwaczne intelektualne podziemie kisi się we własnym sosie niechęci do pesymizmu. Znam je aż za dobrze, sam tkwiłem w nim zbyt długo.

Od czasu, kiedy założyliśmy DeepMind i od tamtych prezentacji, dyskusja nieco się zmieniła. Temat automatyzacji pracy przewalowano już na dziesiątą stronę. Globalna pandemia ukazała zarówno zagrożenia, jak i potężne możliwości biologii syntetycznej. Pojawił się swego rodzaju „techlash”, w którego ramach krytycy w ośrodkach decyzyjnych w Waszyngtonie, Brukseli i Pekinie w artykułach i książkach odsądzali technologie i firmy technologiczne od czci i wiary. Marginalne dotąd lęki związane z technologią z impetem wdarły się do życia publicznego, wzrósł sceptycyzm opinii publicznej wobec niej, nasiliła się krytyka ze strony środowiska akademickiego, społeczeństwa obywatelskiego i świata polityki.

Wobec obserwowanej u technoelity silnej niechęci do pesymizmu oraz nadciągającej fali i stwarzanego przez nią wielkiego dylematu to wszystko może jednak okazać się niewystarczające.

Argumentacja

Historia ludzkości toczy się falami, a fala, o której tu mówimy, jest po prostu najnowszą z całego ich ciągu. Wielu ludziom się wydaje, że to wciąż melodia przyszłości, tak odległej

i abstrakcyjnej, że zajmującej jedynie nerdów i zwariowanych radykalnych myślicieli – że to raczej hiperbola, technologiczna dętologia i autopromocja. A to błąd. Zagrożenie jest realne, nieczym tsunami, które wylania się na horyzoncie niewzruszonego błękitnego oceanu.

To nie jest niczyja fantazja ani nie są to oderwane od rzeczywistości wydumane dywagacje. Choćbyście nie zgadzali się z moją argumentacją i uważali, że żaden z tych scenariuszy nie jest prawdopodobny, to apeluję o doczytanie tej książki do końca. Owszem, wywodzę się ze środowiska badaczy sztucznej inteligencji i siłą rzeczy patrzę na świat przez pryzmat technologii. Owszem, jestem stronnicy, gdy idzie o pytanie o znaczenie tych wszystkich zjawisk. Obserwując z bliska od półtorej dekady tę nadciągającą rewolucję, dochodzę do wniosku, że stoimy u progu najważniejszej transformacji, jaka wydarzy się za naszego życia.

Jako twórca tych technologii wierzę, że mogą one przysporzyć wiele dobra, przynieść poprawę życia niezliczonej rzeszy ludzi i wspomóc nas w stawianiu czoła fundamentalnym wyzwaniom, od nowych metod produkcji czystej energii po tworzenie tanich i skutecznych leków na nasze najtrudniejsze do wyleczenia choroby. Technologie mogą i powinny wzbogacać nasze życie; w przeszłości, co wypada tu powtórzyć, stojący za nimi wynalazcy i przedsiębiorcy byli potężnym kołem napędowym postępu, który podniósł stopę życiową miliardów ludzi.

Lecz bez ujęcia w ryzyko technologii każdy inny jej aspekt, każda dyskusja na temat jej niedostatków etycznych lub korzyści, jakie może przynieść, są pozbawione znaczenia. Pilnie potrzebujemy niepodważalnych odpowiedzi na pytanie, jak opanować nadchodzącą falę i stawiać jej tamę oraz jak utrzymać zabezpieczenia i dobrodziejstwa demokratycznego państwa

Nadchodząca fala

narodowego, ale w tej chwili nikt nie ma w odniesieniu do tych zagadnień żadnego planu. To przyszłość, której nikt z nas sobie nie życzy. Obawiam się jednak, że jest ona coraz bardziej prawdopodobna, co uzasadnię w kolejnych rozdziałach.

W części pierwszej przyjrzymy się długiej historii technologii oraz sposobowi jej rozprzestrzeniania się – wzbierającym przez tysiąclecia falom. Co napędza te fale? Co czyni je tak powszechnymi? Zapytamy też, czy istnieją przykłady społeczeństw, które świadomie mówią „nie” nowym technologiom. Kiedy spoglądamy w głąb dziejów, nie dostrzegamy przykładów odwracania się od postępu technologicznego, lecz wyraźny wzorzec jego rozprzestrzeniania się, pociągającego za sobą długie rozgałęziające się łańcuchy zarówno zamierzonych, jak i niezamierzonych konsekwencji.

Powyższą kwestię nazywam problemem powstrzymywania: jak utrzymać kontrolę nad najcenniejszymi technologiami, jakie zna historia, kiedy te tanieją i ekspandują szybciej niż wszystkie wcześniejsze?

W części drugiej przyjrzę się bliżej tej nadchodzącej fali. U jej podstaw leżą dwie technologie ogólnego zastosowania niosące ogromny potencjał, możliwości i zagrożenia: sztuczna inteligencja i biologia syntetyczna. Obie były od dawna zapowiedane, a mimo to twierdzę, że skala ich oddziaływania jest wciąż powszechnie niedoceniana. W ich otoczeniu dojrzewa szereg technologii pokrewnych, takich jak robotyka i komputery kwantowe, których trajektorie rozwojowe będą się wzajemnie krzyżować w złożony i żywiołowy sposób.

W części tej przyjrzymy się temu, jak powstawały i co potrafią, a także zastanowimy się, dlaczego tak trudno trzymać je w ryzach. Technologie, o których tu mowa, mają cztery wspólne zasadnicze cechy wskazujące na ich bezprecedensowość: są

z natury uniwersalne, a zatem mają wszechstronne zastosowania, błyskawicznie ewoluują, mają asymetryczne oddziaływanie i pod pewnymi względami stają się coraz bardziej autonomiczne.

Tworzenie ich jest pobudzane przez potężne bodźce: rywalizację geopolityczną, perspektywę krociowych profitów oraz otwartą, rozproszoną kulturę badań. Dziesiątki podmiotów państwowych i niepaństwowych będą się prześcigać w ich rozwijaniu, niezależnie od działań zmierzających do ich uregulowania i okiełznania. Ryzyko podejmowane przez uczestników tego wyścigu odbije się na nas wszystkich, czy nam się to podoba, czy nie.

W części trzeciej skupię na politycznych implikacjach gigantycznej redystrybucji władzy wywołanej przez niekontrolowaną falę technologii. Fundamentem naszego obecnego ładu politycznego – a przy tym najważniejszym aktorem zdolnym do powstrzymywania technologii – jest państwo narodowe. Choć już dziś targają nim rozmaite kryzysy, to jeszcze osłabi je seria wstrząsów potęgowanych przez nadchodzącą falę, takich jak nowe formy przemocy, zalew dezinformacji, ubytek miejsc pracy i perspektywa wypadków o katastrofalnych skutkach.

Następnie fala ta wywoła serię tektonicznych przesunięć władzy, skutkujących jednoczesną centralizacją i decentralizacją. W efekcie powstaną potężne nowe przedsiębiorstwa, umocni się autorytaryzm, a całe grupy i ruchy staną się zdolne do funkcjonowania poza tradycyjnymi strukturami społecznymi. Delikatna tkanka państwa narodowego zostanie poddana ogromnej presji dokładnie wtedy, gdy będziemy najbardziej potrzebować takich właśnie instytucji. W konsekwencji stanimy przed wspólnym dylematem.

W części czwartej dyskusja skoncentruje się na tym, co możemy na to poradzić. Czy istnieje choć cień szansy na

Nadchodząca fala

powstrzymanie technologii, na wyrwanie się z kleszczy wspomnianego dylematu? A jeśli tak, to jak tego dokonać? Przedstawię tu dziesięć kroków, przechodzących od poziomu kodu komputerowego i DNA do poziomu międzynarodowych traktatów, tworzących twardy zestaw zagnieźdzonych jedne w drugich ograniczeń, zarys planu powstrzymywania technologii.

To książka o mierzeniu się z porażkami. Technologie mogą okazać się zawodne w przyziemnym sensie niedziałania, tak jak silnik, który nie chce odpalić, albo zerwany most. Ale mogą też zawieść w szerszym znaczeniu. Jeśli technologia szkodzi ludzkiemu życiu czy też hoduje społeczeństwa przesycone krzywdą bądź nierządem z powodu rosnących wpływów wszelkiej maści uzbrojonych w nią aktorów o złych intencjach (lub nieumyślnie niebezpiecznych) – jeśli *per saldo* technologia jest szkodliwa – to można powiedzieć, że zawiodła w innym, głębszym sensie, nie spełniając obietnicy, jaką z nią wiązano. Tak rozumiana porażka nie jest podyktowana żadną obiektywną cechą technologii – wiąże się z kontekstem, w którym działa, formami rządzenia, jakim podlega, sieciami władzy oraz użytkiem, jaki z niej uczyniono.

Ta imponująca pomysłowość, która daje dziś początek wielkim przemianom, sprawia, że lepiej radzimy sobie z unikaniem sytuacji awaryjnych pierwszego rodzaju. Mniej samolotów się rozbija, samochody są mniej szkodliwe dla środowiska i mniej narażone na wypadki, komputery stają się sprawniejsze i bezpieczniejsze. Naszym wielkim wyzwaniem jest to, że wciąż nie przygotowaliśmy się na sytuacje awaryjne drugiego rodzaju.

Na przestrzeni stuleci technologia radykalnie poprawiła warunki bytowe miliardów ludzi. Dzięki nowoczesnej medycynie cieszymy się niespotykanym wcześniej zdrowiem, większość

mieszkańców planety ma zapewnione bezpieczeństwo żywnościowe, ludzie nigdy nie byli tak dobrze wykształceni, nie żyli w tak spokojnych czasach ani nie dysponowali takimi wygodami materialnymi. Te przełomowe osiągnięcia po części zawdzięczamy wielkiemu kołu zamachowemu ludzkości: nauce i tworzeniu nowych technologii. To dlatego postanowiłem poświęcić swoje życie rozważnemu doskonaleniu tych narzędzi.

Ale wszelki optymizm, jaki czerpiemy z tej niezwyklej historii, musi mieć pokrycie w nagiej rzeczywistości. Zabezpieczenie się przed utratą kontroli i sytuacjami awaryjnymi wymaga zrozumienia i ostatecznego zmierzenia się z tym, co może pójść nie tak. Musimy doprowadzić naszą linię rozumowania do jej logicznego końca, bez obawy o to, dokąd może ona zmierzać, a kiedy już do tego końca dotrzemy, musimy podjąć odpowiednie działania. Istnieje poważne ryzyko, że nadchodząca fala technologii wymknie się spod kontroli wcześniej i na szerszą skalę niż wszystko, co zna historia ludzkości. Ten stan rzeczy wymaga powszechnej uwagi całego świata. Wymaga odpowiedzi, których jeszcze nie mamy.

Na pierwszy rzut oka powstrzymanie technologii nie wydaje się możliwe. A jednak dla dobra nas wszystkich musi być możliwe.