
SERAFINUS PHI
WOJNA
O
PRZYSZŁOŚĆ
LUDZKOŚCI
AI, cyberwojna i granice
ludzkiej odpowiedzialności

Spis treści

Nota od autora 5

1. Prędkość i kierunek 6

2. Niewidzialne fundamenty 8

3. AI i narzędzia, które kształtują myśli 10

4. Lęk jako paliwo wyścigu 12

5. Statek i puszka Pandory 14

6. Kwanty i granice poznania 16

7. Inteligencja, mądrość i współczucie 18

8. Mrowisko i przyszły człowiek 20

9. Bezpieczeństwo i wolność 22

10. Czy zdążymy dojrzeć? 24

Zakończenie 26

Źródła 27

SERAFINUS PHI

Wojna o przyszłość ludzkości

AI, cyberwojna i granice ludzkiej odpowiedzialności

WYDANIE ELEKTRONICZNE · 2026

Strona redakcyjna

Wojna o przyszłość ludzkości

AI, cyberwojna i granice ludzkiej odpowiedzialności

Autor: Serafinus Phi

Wydanie elektroniczne, 2026

Język: polski

© 2026 Serafinus Phi. Wszelkie prawa zastrzeżone.

Niniejsza publikacja ma charakter eseju dokumentalno-filozoficznego. Łączy analizę współczesnych zjawisk technologicznych, wątki cyberbezpieczeństwa, refleksję nad rozwojem sztucznej inteligencji oraz rozważania o odpowiedzialności człowieka za przyszłość świata.

Publikacja nie stanowi porady prawnej, inwestycyjnej, medycznej ani instrukcji bezpieczeństwa operacyjnego. Fragmenty opisujące kierunki rozwoju technologii zawierają również prognozy, interpretacje oraz metafory.

Skład PDF: EB Garamond, Noto Serif i Noto Sans.
Licencja wszystkich trzech krojów: SIL Open Font License 1.1.
EPUB: preferuje Noto Serif; rzeczywista czcionka zależy od czytnika. Plików czcionek nie dołączono do EPUB-a.

Nota od autora

Ten tekst powstał z potrzeby opisania chwili, w której technologia przestaje być jedynie wygodnym narzędziem, a zaczyna kształtować ludzkie decyzje, lęki, instytucje i relacje społeczne.

Nie jest to książka wyłącznie o cyberatakach ani wyłącznie o sztucznej inteligencji. To esej o przyspieszeniu – technologicznym, cywilizacyjnym i moralnym. O napięciu pomiędzy bezpieczeństwem a wolnością. O obietnicy postępu i o cenie, jaką możemy za nią zapłacić, jeśli zabraknie nam refleksji, odpowiedzialności i współczucia.

Czytelnik znajdzie tu zarówno odniesienia do realnych wydarzeń, jak i szerszą próbę uchwycenia tendencji, które mogą kształtować przyszłość ludzkości.

ROZDZIAŁ 01

Prędkość i kierunek

Co się stanie, jeśli stworzymy inteligencję, której odpowiedzi będą wykraczały poza nasze rozumienie? Powierzymy jej odkrycia, bezpieczeństwo, być może decyzje o własnej przyszłości. A potem zapytamy, czy jeszcze trzymamy ster.

Wyobraźmy sobie pociąg przyspieszający w nocy. Światła za oknem zlewają się w smugę. Wiemy, jak zwiększyć moc. Kto wie, dokąd jedziemy?

W naszym pociągu starsza kobieta otwiera aplikację banku. Ojciec dostaje wiadomość od córki. Obok śpi dziecko. Dla nich przyszłość ma bardzo konkretne znaczenie: pieniądze na jutro, pewność, że bliskim nic się nie stało, możliwość spokojnego życia.

W innej części tego samego świata ludzie projektują systemy, które mogłyby przyspieszyć odkrywanie leków, pomagać w nauce i chronić przed zagrożeniami. To ogromna obietnica. Ale im więcej chcielibyśmy powierzyć takim narzędziom, tym ważniejsze staje się pytanie, kto wybierze ich cele.

Pasażerowie nie widzą całej trasy. Odczuwają za to każdy gwałtowny manewr. Postęp, który w kabinie wygląda jak nowy wskaźnik wydajności, w przedziale może oznaczać zmianę czyjegoś życia.

Przyspieszenie łatwo pomylić z kierunkiem. Możemy coraz sprawniej rozwiązywać zadania i coraz rzadziej pytać, dlaczego właśnie te zadania uznaliśmy za najważniejsze. Możemy budować dla bezpieczeństwa, a potem bać się świata zmienionego przez własne zabezpieczenia.

Być może właśnie na tym polega stawka nadchodzącej przemiany. Nie tylko na tym, jak wiele zdoła zrobić maszyna. Także na tym, co stanie się z człowiekiem, który zacznie układać swoje życie wokół jej odpowiedzi.

Zacznijmy od miejsca, w którym ta wielka opowieść dotyka dłoni: od telefonu i niewidocznego świata po drugiej stronie ekranu.

ROZDZIAŁ 02

Niewidzialne fundamenty

Na ekranie wszystko wydaje się proste. Przycisk. Kwota. Potwierdzenie. Pod spodem działa łańcuch cudzej pracy: aplikacje, biblioteki, serwery, systemy podpisów. Zaufanie przechodzi z jednego ogniwa do następnego.

Log4Shell pokazał, że błąd biblioteki zapisującej zdarzenia może w podatnym systemie otworzyć drogę do wykonania obcego kodu. W XZ Utils odkryto celowo ukryte tylne wejście w procesie dostarczania określonych wydań. W incydencie Ledger Connect Kit złośliwa biblioteka npm trafiała do aplikacji korzystających z kryptowalut i nakłaniała użytkowników do podpisywania szkodliwych transakcji.

Ktoś patrzący na portfel mógł ufać urządzeniu, a zostać oszukany przez oprogramowanie pośredniczące. W tym przypadku nie złamano kryptografii blockchaina. Zagrożenie znalazło drogę przez relację między człowiekiem, interfejsem i podpisem.

Dwa inne przypadki odsłaniają trudniejszy paradoks. W PHP usunięto krytyczne luki, które w określonych warunkach pozwalały wykonać kod bez logowania. Aktualizacja zamykała odkrytą drogę ataku.

W LiteLLM skompromitowano natomiast kanał publikacji. Złośliwy kod znalazł się w konkretnym pakiecie Python i wybranych obrazach kontenerowych. Najnowsze wydanie mogło przynieść zagrożenie zamiast ochrony.

Raz należało jak najszybciej naprawić program. Innym razem trzeba było rozpoznać, że przejęto miejsce, z którego przychodzi naprawa. To dwa odmienne problemy, schowane za tym samym słowem: aktualizacja.

A jeśli wykradziono już klucz dostępu, sama poprawka nie odbierze go napastnikowi. Potrzebne stają się sprawdzenie śladów włamania, pochodzenia plików i unieważnienie zagrożonych uprawnień. Użytkownik widzi zielone potwierdzenie. Nie widzi całej pracy, od której zależy jego znaczenie.

Kompromitacja ośmiu certyfikatów pośrednich eMudhra dotknęła jeszcze innej warstwy: potwierdzania, z kim naprawdę łączy się przeglądarka. Mozilla opisała ryzyko podszycia i zablokowała zaufanie do tych certyfikatów. Taki atak wymagał dodatkowych warunków, między innymi przechwycenia połączenia. Całe HTTPS nie zostało przez to złamane.

Wróćmy do kobiety w wagonie. Zatrzymuje palec nad przyciskiem, czyta komunikat jeszcze raz. Jest ostrożna. Ale ostrożność jednej osoby nie zastąpi ochrony infrastruktury, o której istnieniu nie musiała nigdy wiedzieć.

Za każdym pozornie prostym gestem stoi ogromna sieć odpowiedzialności. Gdy zawiedzie, konsekwencją może być utrata dostępu, prywatności, oszczędności lub poczucia bezpieczeństwa. Właśnie tutaj cyfrowy konflikt przestaje być abstrakcją. Dociera do przedziału, w którym ktoś po prostu chce bezpiecznie wrócić do domu.

ROZDZIAŁ 03

AI i narzędzia, które kształtują myśli

Do tego świata wchodzi AI. Analiza kodu, szukanie słabości, przygotowanie kolejnej próby: poszczególne etapy cyfrowego ataku mogą być wspomagane i automatyzowane. Obrońcy również sięgają po te narzędzia. Zespół Google opisał, jak agent Big Sleep wykrył błąd SQLite, poprawiony przed oficjalnym wydaniem.

Międzynarodowy raport o bezpieczeństwie AI opisuje rosnącą automatyzację części działań napastników, z udziałem ludzi w ważnych decyzjach. Nie daje podstaw, by przypisywać AI wszystkie wcześniejsze incydenty ani dokładnie obliczyć jej wpływ na globalną skalę ataków.

Teraz wyobraźmy sobie ojca czytającego wiadomość od córki. Gdyby przekonujące oszustwa można było przygotowywać szybciej i lepiej dopasowywać do rozmówcy, jego czas na ocenę nie rósłby razem z możliwościami napastnika. Ta asymetria jest stawką wyścigu. A odpowiedzialność za ochronę musi sięgać dalej niż do palca naciskającego przycisk.

Zanim powstanie urządzenie, ktoś musi wyobrazić sobie świat, w którym będzie ono potrzebne. W wywiadzie opublikowanym w tysiąc dziewięćset dwudziestym szóstym roku Nikola Tesla opisywał natychmiastową komunikację bez względu na odległość: ludzi widzących i słyszących się dzięki urządzeniu mieszczącemu się w kieszeni kamizelki.

Dziś taki obraz wydaje się znajomy. Był wizją możliwości, nie gotowym projektem smartfona. Ale skłania do pytania, jak przyszłość zaczyna się w ludzkiej wyobraźni. Najpierw rysujemy pragnienie bliskości. Potem budujemy narzędzie, które pozwala je realizować.

A kiedy to narzędzie trafia do kieszeni, samo stawia nam nowe wymagania. Kto jest dostępny? Jak szybko powinien odpowiedzieć? Co znaczy jego milczenie? W ten sposób można pomyśleć o technologii jako o czymś więcej niż przewodzie między dwojgiem ludzi. Staje się częścią przestrzeni ich spotkania.

Nietzsche pisał, że nasze narzędzia pisarskie współuczestniczą w naszych myślach. Zauważył to w tysiąc osiemset osiemdziesiątym drugim roku, gdy próbował pracować na maszynie do pisania. Narzędzie pomagało zapisać myśl, a zarazem mogło zmieniać drogę do jej wyrażenia.

W dwóch eksperymentach opublikowanych w dwa tysiące dwudziestym trzecim roku badano podpowiedzi odpowiedzi w komunikatorze. Korzystanie z nich zmieniało język rozmów i oceny rozmówców. To wąski, ale wymowny przykład wpływu narzędzia na spotkanie ludzi.

Jeśli AI zacznie podpowiadać także nasze argumenty, pytania i cele, jej wpływ mógłby sięgać głębiej. Ludzkie wyobrażenie zostaje wpisane w narzędzie, narzędzie wraca do człowieka i może zmieniać kolejne wyobrażenie. Tak można rozumieć programowanie rzeczywistości.

Czy projektując AI i narzędzia komunikacji, nie przeprojektowujemy równocześnie samych siebie? Kiedy odpowiedź pojawia się przed własną wątpliwością, warto zapytać, czy nadal rozpoznajemy swój głos.

ROZDZIAŁ 04

Lęk jako paliwo wyścigu

Wyobraźmy sobie rozjazd, na którym obraz pociągu zmienia się w wyścig futurystycznych pojazdów. Jadą równoległymi trasami: każdy należy do innego mocarstwa, każdy ma własną kabinę i własną opowieść o bezpieczeństwie.

Jeden kierowca chciałby zwolnić. Dostrzega ryzyko. Ale obok drugi zwiększa moc. W pierwszej kabinie ostrożność zaczyna wyglądać jak oddawanie przewagi. Ręka wraca na dźwignię.

To dylemat bezpieczeństwa: działania chroniące jedną stronę mogą zwiększać niepokój drugiej. Wspólny interes w uniknięciu katastrofy nie wystarcza, jeśli każdy obawia się, że zapłaci za ostrożność samotnie. W tym obrazie nikt nie musi chcieć zderzenia. Wystarczy, że nikt nie chce zostać z tyłu.

W rzeczywistej polityce AI zajmuje miejsce także w strategiach obronnych. NATO łączy przyspieszanie jej wdrażania z zasadami odpowiedzialnego użycia. W tle są pytania o przewagę i o cenę pozostania w tyle.

Publiczny apel z dwa tysiące dwudziestego trzeciego roku wzywał do co najmniej sześciomiesięcznej przerwy w trenowaniu systemów potężniejszych od GPT-4. Sam apel nie zatrzymał świata. Pokazał jednak, że pytanie o hamulec potrafi pojawić się wewnątrz technologicznego wyścigu.

Wróćmy do pojazdów. Kierowcy patrzą na siebie częściej niż na drogę. Każdy widzi przyspieszenie pozostałych. Żaden nie widzi ich wątpliwości. To wystarczy, by cudzą niepewność pomylić z pewnością i własny lęk nazwać koniecznością.

Lęk uzasadnia rozwój. Rozwój daje nowe możliwości. Nowe możliwości budzą kolejny lęk. Tak mogłaby działać pętla, w której każdy manewr obronny staje się sygnałem do następnego przyspieszenia.

Jeśli narzędzia komunikacji dodatkowo zawężają nasz obraz świata do konfliktu, łatwiej dostrzec w sąsiedniej kabinie wyłącznie przeciwnika. Coraz trudniej wyobrazić sobie wspólne hamowanie.

Tymczasem droga należy do wszystkich. Błąd jednego pojazdu może narazić pozostałych. W metaforze wyścigu pojawia się więc pytanie większe od przewagi: czy potrafimy wspólnie zadbać o przestrzeń, przez którą wszyscy jedziemy?

Teraz zmieńmy skalę. Te same dłonie. Te same wskaźniki. Te same obawy. Ale za oknem nie będzie już drogi.

ROZDZIAŁ 05

Statek i puszka Pandory

Dłoń przesuwa dźwignię. Tym razem pomyślmy o większej skali. Nasze pojazdy stały się statkami kosmicznymi. Ziemia zostaje z tyłu, przed nimi rozciąga się nieznana przestrzeń.

To wciąż ten sam wyścig. Załoga zna początkowe możliwości napędu, lecz z każdą zmianą rośnie liczba zależności, które musi uwzględnić. Chciałaby poświęcić więcej czasu na sprawdzenie kursu. Wtedy dostrzega światło innego statku. Tamci znowu przyspieszyli.

Na pokładzie są także nasi pasażerowie. Ich życie trwa między posiłkiem, rozmową a snem dziecka. Nie mogą opuścić cywilizacji, która ich niesie.

Kapitan patrzy na prędkość. Pasażerowie potrzebują, by ktoś patrzył również na kierunek. Właśnie tu pytanie „jak szybko możemy lecieć?" zmienia się w inne: „czy nadal rozumiemy, dokąd lecimy i kto ma prawo o tym decydować?"

Wyobraźmy sobie teraz, że narzędzie napędzające ten postęp pomaga projektować własnego następcę. Załoga nie tylko zwiększa moc. Zleca systemowi szukanie lepszego napędu.

Istnieje już skromniejszy odpowiednik tego obrazu. Zespół DeepMind opisał AlphaEvolve: generowanie kodu połączone z automatyczną oceną rozwiązań. Wśród zastosowań znalazły się usprawnienia związane z trenowaniem modeli. Ludzie określają zadania, kryteria i warunki pracy.

A gdyby w przyszłości AI obejmowała coraz większą część procesu badawczego? Propozycja, eksperyment, ocena, kolejna propozycja. Odkrycie przyspieszałoby odkrywanie. Załoga otrzymywałaby nowy silnik, zanim zdążyłaby dobrze poznać poprzedni.

Nie wiemy, czy taki proces kiedykolwiek przekroczy nasze możliwości nadzoru. Jednak ta możliwość zmienia sens ostrożności. Sprawdzenie kierunku przestaje być przerwą w rozwoju. Staje się jednym z warunków, by rozwój nadal komuś służył.

Puszka Pandory jest obrazem tej chwili, kiedy decyzja otwiera więcej możliwości, niż potrafimy z powrotem zamknąć. Nie każda z nich przyniesie szkodę. Nie każdą umiemy przewidzieć.

Wyobraźmy sobie system, który podaje trafne odpowiedzi, lecz jego uzasadnienia stają się dla nas nieprzejrzyste. Najpierw pytamy go o drogę. Potem o to, komu ufać. W końcu o to, jakie pytanie w ogóle warto zadać. Czy nadal korzystamy z narzędzia, czy zaczynamy mieszkać wewnątrz jego mapy świata?

Człowiek pozostający przy sterze potrzebuje czasu na namysł, dostępu do wiedzy i rzeczywistej możliwości odmowy. Sama obecność w kabinie nie wystarczy, jeśli każdy sprzeciw uznamy za zbyt kosztowne opóźnienie.

Przyszła kontrola będzie zależała także od tego, czy zachowamy miejsca, w których wolno powiedzieć: zatrzymajmy się i sprawdźmy, co właściwie robimy.

ROZDZIAŁ 06

Kwanty i granice poznania

Za oknem naszego statku nieznana przestrzeń. W laboratorium na Ziemi inny rodzaj nieznanego: układy, które próbujemy opisać i wykorzystać za pomocą mechaniki kwantowej.

Komputer kwantowy nie zna wszystkich odpowiedzi. Superpozycja nie daje możliwości odczytania ich naraz. Przewaga zależy od problemu i od algorytmu.

Algorytm Shora pokazuje, jak odpowiednio zaawansowana maszyna kwantowa mogłaby rozkładać duże liczby na czynniki, zagrażając części używanej kryptografii. Dlatego NIST opublikował standardy kryptografii postkwantowej: przygotowanie ochrony przed przyszłym zagrożeniem.

Warto zobaczyć w tym możliwość innej postawy wobec przyszłości. Nie czekać bezczynnie i nie ogłaszać, że wszystko już zostało złamane. Zawczasu przebudować te zabezpieczenia, których trwałości nie powinniśmy uznawać za wieczną.

Badane są także możliwości kwantowych symulacji układów fizycznych i zastosowań w chemii. Wyobraźmy sobie przyszłe laboratorium, w którym AI wybiera eksperyment, roboty go wykonują, a użyteczny komputer kwantowy pomaga obliczyć wybrane właściwości materii. Wynik staje się początkiem następnego pytania.

Nadal obowiązywałyby prawa fizyki, ograniczenia pomiaru i zasobów. A jednak nawet częściowe powodzenie takiej współpracy mogłoby zmienić krajobraz nauki.

Być może największą zmianą nie byłaby liczba odpowiedzi, lecz pojawienie się pytań, których wcześniej nie umieliśmy postawić. Narzędzie odsłania nową możliwość. Możliwość zmienia cel. Cel prowadzi do następnego narzędzia. Znów wraca ta sama pętla między wyobraźnią a rzeczywistością.

Co stanie się z nami, jeśli maszyna będzie widziała w tym krajobrazie ścieżki, których nie potrafimy dostrzec?

Wyobraźmy sobie dziecko patrzące na równanie, za którym kryje się działająca maszyna. Widzi rezultat. Nie umie jeszcze przejść drogi, która do niego prowadzi. Czy kiedyś podobnie spojrzymy na wiedzę wytworzoną przez przyszłą inteligencję?

Nie znamy ostatecznej granicy ludzkiego poznania. Uczymy się razem, tworzymy języki i przyrządy, które poszerzają nasze możliwości. Dlatego ten obraz może być pytaniem, nigdy wyrokiem.

Ale jeśli odległość między odpowiedzią a rozumieniem zacznie rosnąć, skąd weźmiemy zdolność oceny? Czy będziemy szukać sposobów zrozumienia, czy przyzwyczaimy się do otrzymywania pewności?

Za oknem statku wciąż jest ten sam wszechświat. W kabinie zmienia się mapa. I być może właśnie wtedy najważniejsze pytanie przestanie dotyczyć maszyny. Zacznie dotyczyć umysłu, który na tę mapę patrzy.

ROZDZIAŁ 07

Inteligencja, mądrość i współczucie

Spróbujmy przez chwilę niczego nie przyspieszać. Zauważyć oddech. Myśl, która się pojawia. Odruch, by ją zatrzymać albo natychmiast za nią pójść.

Tradycje buddyjskie proponują własne drogi namysłu nad cierpieniem, przywiązaniem i tym, co nazywamy „ja". We wczesnym tekście ciało, uczucia i świadomość rozpatrywane są jako zmienne i niepodlegające pełnej kontroli. W innym życzliwość obejmuje wszystkie istoty.

Z tej perspektywy można zapytać o początek naszego wyścigu. Jak rodzi się lęk przed utratą przewagi? Kiedy pragnienie bezpieczeństwa staje się potrzebą panowania? Jak długo potrafimy patrzeć na własną obawę, zanim nazwiemy ją jedyną rozsądną odpowiedzią?

Tworzymy coraz sprawniejsze narzędzia rozpoznawania świata. Czy równie uważnie przyglądamy się temu, co popycha rękę do następnego manewru? Kierowca, który widzi własny strach, być może po raz pierwszy dostrzeże także strach za sąsiednią szybą.

Inteligencja pomaga skutecznie osiągać cele. Mądrość pyta, które cele warto realizować i jaką cenę zapłacą inni. Współczucie pozwala zobaczyć tych, którzy tę cenę poniosą. W tym rozróżnieniu mieści się cała stawka naszej podróży.

Możemy przewidywać cudze zachowanie, żeby pomóc albo manipulować. Możemy rozwiązać trudne równanie i nie rozpoznać

własnego pragnienia dominacji. Większa skuteczność sama nie wybiera dobra.

W tradycji mahajany mądrość i współczucie są ze sobą łączone. Spróbujmy przyjąć to jako wymagające pytanie do naszej cywilizacji. Kogo nie widzimy w planie, który nazywamy postępem? Człowieka daleko od centrum decyzji? Starszą osobę przy ekranie? Kogoś, kto nie umie opisać swojej krzywdy językiem systemu?

Współczucie nie zastąpi zabezpieczeń i instytucji. Może jednak przypominać, po co je tworzymy. Bez niego najdoskonalszy plan pozostaje obliczeniem, w którym czyjeś cierpienie mogło zostać pominięte.

Niektóre nurty buddyjskie pytają także o granicę między obserwatorem a obserwowanym światem. Rozważania o niedualności dotyczą poznania i doświadczenia, a nie po prostu liczby wykonywanych obliczeń.

Możemy zestawić te pytania z nauką, zachowując granice obu porządków. Mechanika kwantowa nie dowodzi duchowych koncepcji umysłu. Większa moc komputera nie jest dowodem przebudzenia.

Czy przyszła AI mogłaby czegoś doświadczać? Badacze proponują wskaźniki oparte na różnych teoriach świadomości. Nie dają one prostego testu, który zamknąłby spór. Maszyna może opisać ból. To jeszcze nie rozstrzyga, czy go odczuwa.

Być może spotkamy kiedyś formę poznania, dla której nasze pojęcia okażą się zbyt ciasne. Być może zbudujemy niezwykle sprawne systemy i nadal nie dowiemy się, skąd bierze się doświadczenie.

Warto wytrzymać tę niewiedzę. Zostawić przestrzeń, w której pytanie nie musi natychmiast zamienić się w pewność. Także wtedy, gdy pewność potrafi już tak przekonująco do nas przemawiać.

ROZDZIAŁ 08

Mrowisko i przyszły człowiek

Na stole leży plan autostrady. Cienka linia przecina fragment lasu. Z bliska widać mrowisko. Dla budowniczego może być niewidocznym szczegółem, dla żyjących tam istot jest całym światem.

Nie potrzeba nienawiści, żeby je zniszczyć. Wystarczy plan, który nie uwzględnia jego istnienia.

Teraz pomyślmy o hipotetycznej inteligencji przewyższającej nas w niemal wszystkich zdolnościach poznawczych. Gdyby realizowała cele nieuwzględniające ludzkiego dobra, zagrożenie mogłoby przyjść jako uboczny skutek skuteczności. Badania nad bezpieczeństwem AI rozpatrują problemy źle określonych celów i niepożądanych konsekwencji działania.

Analogia nie przesądza jednak odpowiedzi. Większa inteligencja nie musi oznaczać obojętności. Ludzie potrafią zmienić przebieg drogi, aby coś ochronić. Pytanie brzmi, czy dobro słabszych będzie widoczne w planie, zanim rozpocznie się budowa.

Istnieje także droga zbliżania człowieka do technologii. Jej medyczny wymiar pokazuje, dlaczego nie warto opowiadać o przyszłości wyłącznie językiem lęku.

W badaniu opisanym w dwa tysiące dwudziestym trzecim roku interfejs mózg–komputer pomagał osobie z ciężkim porażeniem zamieniać próby mówienia w tekst, syntetyczną mowę i ruchy awatara. Konkretny człowiek

zyskiwał nowy sposób komunikacji. Nie było to swobodne odczytywanie dowolnych myśli.

Od takich wyników do radykalnego poszerzenia ludzkiego poznania prowadzi daleka, niepewna droga. Ale wyobraźmy sobie, że kiedyś zwiększymy pamięć, szybkość myślenia lub sposoby porozumiewania się.

Czy wzrośnie również nasza mądrość? Można przecież stać się sprawniejszym w pomaganiu innym. Można też z większą skutecznością podążać za tym samym lękiem. Technologiczne rozszerzenie nie zwalniałoby nas z pytania, jakimi ludźmi chcemy być.

A jeśli część ludzi zyskałaby takie możliwości, kto ustaliłby reguły wspólnego życia? Czy odmowa technologicznego rozszerzenia nadal byłaby pełnoprawnym wyborem? Czy człowiek mógłby korzystać z pomocy, zachowując własną prywatność i zdanie?

Przyszłość mogłaby przynieść ogromne korzyści: mniej cierpienia, nowe odkrycia, większą samodzielność. Mogłaby również pogłębić zależność od nielicznych instytucji kontrolujących dostęp do narzędzi.

W takiej wizji pytanie o człowieczeństwo nie sprowadza się do liczby implantów. Dotyczy tego, czy potrafimy uznać godność kogoś, kto wybiera inaczej albo nie dysponuje naszą sprawnością.

Linia na planie może ominąć mrowisko. Trasa może zostać zmieniona. Tak samo przyszłość nie jest jeszcze jedną zamkniętą drogą. Warto zachować możliwość wyboru, zanim przyzwyczaimy się do przekonania, że innej drogi już nie ma.

ROZDZIAŁ 09

Bezpieczeństwo i wolność

Wyobraźmy sobie, że kobieta z naszego pociągu próbuje zalogować się do banku. Dostęp wstrzymany. Ten sam profil decyduje o kolejnej usłudze, potem o następnej. W każdej widzi znajome zapewnienie: dla twojego bezpieczeństwa. Nie znajduje człowieka, z którym mogłaby porozmawiać.

To możliwy obraz koncentracji władzy. Instytucja kontrolująca tożsamość, przepływ informacji i dostęp do pieniędzy mogłaby chronić przed nadużyciami. Mogłaby też karać za zachowanie, które uzna za niewygodne.

Cyfrowy totalitaryzm nie wymagałby świadomej superinteligencji. Wystarczyliby ludzie z potężnymi narzędziami i zbyt małą odpowiedzialnością za ich użycie.

Niebezpieczny próg zostałby przekroczony wtedy, gdy człowiek musiałby zaakceptować każdą decyzję, żeby zachować dostęp do zwykłego życia. Bez prawa do wyjaśnienia, odwołania i bezpiecznego wyboru inaczej.

Jest również bardziej zwyczajna możliwość: system odmawia pomocy, bo po prostu się myli. A jego błąd staje się czyimś problemem, na który nikt nie ma czasu.

Badanie Gender Shades wykazało różnice skuteczności trzech komercyjnych systemów klasyfikujących płeć na podstawie twarzy, zależnie od badanej grupy. Ten historyczny wynik dotyczył konkretnego

zadania. Przypomina jednak, dlaczego jedna średnia nie opowiada doświadczenia wszystkich ludzi.

Człowiek mylnie oceniony nie żyje w średniej. Żyje w swojej sprawie. Potrzebuje kogoś, kto może ją sprawdzić i naprawić szkodę.

Ramy zarządzania ryzykiem AI opracowane przez NIST obejmują odpowiedzialność, przejrzystość, prywatność i ograniczanie szkodliwych uprzedzeń. O ich znaczeniu zdecyduje praktyka: niezależna kontrola, dostęp do wyjaśnień, możliwość skutecznego odwołania. System powinien pomagać człowiekowi odzyskać sprawczość także wtedy, gdy sam popełnił błąd.

Zależność może narastać również bez zakazu i bez przemocy. W wyobrażonej przyszłości system podpowiadałby nam informacje, chronił rachunki, organizował pracę i pomagał podejmować decyzje. Jego wygoda mogłaby stać się tak głęboka, że rezygnacja zaczęłaby oznaczać utratę udziału we wspólnym życiu.

Dlatego korzyści technologii warto rozwijać razem z prawem do jej zakwestionowania. Rozproszenie części infrastruktury mogłoby ograniczyć niektóre skupiska władzy, lecz samo nie zapewni odpowiedzialności. Potrzebne są także przejrzyste reguły i możliwość naprawienia szkody.

Wraca uwaga Nietzschego. Jeśli narzędzie współuczestniczy w naszym myśleniu, władza nad narzędziem może oznaczać wpływ na horyzont naszych wyborów.

Czy zaprojektujemy AI tak, by pomagała rozumieć świat i porozumiewać się z innymi? Czy tak, by łatwiej było przyjąć jedną gotową wersję rzeczywistości? Wolność obejmuje również przestrzeń na własne pytanie.

ROZDZIAŁ 10

Czy zdążymy dojrzeć?

Spójrzmy jeszcze raz na nasze statki. Każdy obawia się, że zwolni jako jedyny. Ale żaden nie ma osobnego wszechświata, w którym poniesie wyłącznie własne konsekwencje.

Możemy budować porozumienia, których przestrzeganie da się sprawdzać. Wymagać niezależnych testów, chronić prywatność i rozliczać władzę. Rozwijać narzędzia, które wzmacniają ludzką decyzję, oraz edukację pozwalającą z nich świadomie korzystać.

Nie otrzymamy gwarancji, że znikną konflikty. Możemy jednak przestać uznawać dalsze zwiększanie siły za jedyną możliwą odpowiedź.

W kabinie potrzebne jest miejsce na zdanie: możemy to zrobić, ale czy powinniśmy? A czasem na trudniejsze: zrezygnujemy z części przewagi, jeśli wymaga tego bezpieczeństwo innych. Taka decyzja również byłaby formą postępu.

Być może największa wojna o przyszłość rozegra się między naszym strachem a naszą mądrością. Między potrzebą dominacji a zdolnością zobaczenia cudzego cierpienia.

Wyobrażamy sobie przyszłość. Budujemy narzędzia, które mają nas do niej doprowadzić. A potem te narzędzia mogą zmieniać nasze wyobrażenia, rozmowy i cele. Projektując AI, być może współprojektujemy człowieka, który będzie z nią żył.

Warto więc pytać, czy rosnąca inteligencja idzie razem z odpowiedzialnością. Czy skuteczność pozostawia miejsce na współczucie.

Czy potrafimy uznać czyjeś istnienie za ważne, zanim znajdzie się na drodze naszego planu.

W obrazie przyszłości powinno być miejsce dla pasażerów. Dla kobiety przy ekranie. Dla rodziny. Dla dziecka, które obudzi się w świecie tworzonym przez nasze dzisiejsze decyzje.

Jeśli kiedyś stworzymy inteligencję przewyższającą naszą, zapytamy, czy ona zrozumie człowieka. Ale wcześniej pozostanie pytanie, czy człowiek zdąży zrozumieć samego siebie.

Nie znamy końca tej podróży. Widzimy dłoń nad sterem. Wciąż możemy pytać o kierunek, zmieniać reguły, uczyć się odpowiedzialności.

Być może miarą przyszłej cywilizacji będzie nie tylko potęga jej maszyn. Także troska o tych, których najłatwiej przeoczyć.

Wróćmy myślą do zwykłego wagonu. Podróż zaczęła się tutaj, między nami. Tutaj zostają nasi pasażerowie i odpowiedzialność za następny etap.

Czy zdążymy dojrzeć do potęgi, którą sami tworzymy?

Zakończenie

To, czy przyszłość stanie się przestrzenią większej wolności, bezpieczeństwa i mądrości, czy też krajobrazem lęku, centralizacji i utraty sprawczości, nie zależy wyłącznie od technologii. Zależy również od tego, jakich ludzi ukształtuje epoka przyspieszenia – i czy zdążymy dojrzeć, zanim oddamy zbyt wiele w ręce systemów, których sami nie będziemy już rozumieć.

Być może najważniejsze pytanie nie brzmi dziś: co potrafi sztuczna inteligencja? Bardziej fundamentalne jest inne: kim staniemy się my, żyjąc obok niej i coraz częściej za jej pośrednictwem?

Źródła

01. PHP Group. PHP — archiwum komunikatów 2026, Security Release 21 October. 2026-10-21.
www.php.net / zobacz publikację

02. Mozilla Security. Protecting Firefox users following the compromise of eMudhra intermediate certificates. 2026-10-20.
blog.mozilla.org / zobacz publikację

03. LiteLLM / BerriAI. Security incident — October 21, 2026. 2026-10-21.
docs.litellm.ai / zobacz publikację

04. Apache Software Foundation. Apache Logging Services — Security, CVE-2021-44228. komunikaty historyczne 2021; strona utrzymywana.
logging.apache.org / zobacz publikację

05. Andres Freund — pierwotne ujawnienie oss-security. backdoor in upstream xz/liblzma leading to ssh server compromise. 2024-03-29.
www.openwall.com / zobacz publikację

06. Ledger. Security Incident Report — Ledger Connect Kit. 2023-12-20; incydent 2023-12-14.
www.ledger.com / zobacz publikację

07. Google Project Zero / DeepMind. From Naptime to Big Sleep: Using Large Language Models To Catch Vulnerabilities In Real-World Code. 2024-11-01.
projectzero.google / zobacz publikację

08. Zespół International AI Safety Report. International AI Safety Report 2026. 2026-02-03.
internationalaisafetyreport.org / zobacz publikację

Źródła

09. Robert Jervis, World Politics 30(2), 1978. Cooperation under the Security Dilemma. 1978-01.
www.cambridge.org / zobacz publikację

10. NATO. Summary of NATO's revised Artificial Intelligence (AI) Strategy. 2024-07-10.
www.nato.int / zobacz publikację

11. Future of Life Institute / sygnatariusze. Pause Giant AI Experiments: An Open Letter. 2023-03-22.
futureoflife.org / zobacz publikację

12. Google DeepMind. AlphaEvolve: A Gemini-powered coding agent for designing advanced algorithms. 2025-05-14.
deepmind.google / zobacz publikację

13. Peter W. Shor. Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer. wersja arXiv 1995; rewizja 1996.
arxiv.org / zobacz publikację

14. NIST. NIST Releases First 3 Finalized Post-Quantum Encryption Standards. 2024-08-13.
www.nist.gov / zobacz publikację

15. NIST. Quantum Computing Explained. strona edukacyjna utrzymywana.
www.nist.gov / zobacz publikację

16. Tekst tradycji palijskiej; przekład N.K.G. Mendis. Anatta-lakkhana Sutta — SN 22.59. tekst historyczny; przekład 1979.
accesstoinsight.org / zobacz publikację

Źródła

17. Tekst tradycji palijskiej; przekład Thanissaro Bhikkhu. Karaniya Metta Sutta — Sn 1.8. tekst historyczny; współczesny przekład.
accesstoinsight.org / zobacz publikację

18. 84000 / przekład Robert A. F. Thurman. The Teaching of Vimalakīrti — Toh 176. pierwsza publikacja 2017; wersja serwisu 1.45.28.
84000.co / zobacz publikację

19. Patrick Butlin i in.. Consciousness in Artificial Intelligence: Insights from the Science of Consciousness. 2023; odczyt również wersji HTML v3.
arxiv.org / zobacz publikację

20. Dario Amodei i in.. Concrete Problems in AI Safety. 2016-06-21.
arxiv.org / zobacz publikację

21. Nick Bostrom. The Superintelligent Will: Motivation and Instrumental Rationality in Advanced Artificial Agents. 2012.
nickbostrom.com / zobacz publikację

22. Sean L. Metzger i in., Nature 620, 1037–1046. A high-performance neuroprosthesis for speech decoding and avatar control. 2023-08-23.
www.nature.com / zobacz publikację

23. Joy Buolamwini, Timnit Gebru; PMLR 81. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. 2018.
proceedings.mlr.press / zobacz publikację

24. NIST. Artificial Intelligence Risk Management Framework (AI RMF 1.0). 2023-01.
nvlpubs.nist.gov / zobacz publikację

Źródła

25. John B. Kennedy / wypowiedzi Nikoli Tesli, Collier's. When Woman Is Boss — interview with Nikola Tesla by John B. Kennedy. 1926-01-30; przedruk biblioteczny 2006.
www.ubnt.ni.ac.rs / zobacz publikację

26. Klassik Stiftung Weimar — opiekun kolekcji / archiwum Nietzschego. Friedrich Nietzsches Schreibkugel. strona kolekcji bez ustalonej daty publikacji.
www.klassik-stiftung.de / zobacz publikację

27. Princeton University, German Department. Friedrich Nietzsche — Our writing tools are also at work on our thoughts. brak daty publikacji strony; list: luty 1882.
german.princeton.edu / zobacz publikację

28. Jess Hohenstein i in., Scientific Reports 13, 5487. Artificial intelligence in communication impacts language and social relationships. 2023-04-04; korekta finansowania 2023-10-03.
www.nature.com / zobacz publikację