Hybrydowa Ludzkość: Architektura kognitywna i instytucjonalna w świecie po AGI na podstawie The AI Instinct Rany Gujral

AUTOR

Fundacja Dobre Państwo

STRESZCZENIE / ABSTRACT

Artykuł analizuje wizję Hybrydowej Ludzkości, gdzie rozszerzenia kognitywne AI stają się narzędziem realnej sprawczości. Autor odwołuje się do 'capability approach', wskazując, że technologia powinna zwiększać faktyczne możliwości człowieka, a nie być jedynie luksusem dla elit. Kluczowym zagrożeniem jest powstanie nowych klas jakości sprzężenia – podziału na osoby z audytowalnymi systemami i tych zależnych od modeli optymalizowanych pod dyktando korporacji. Tekst postuluje stworzenie policentrycznej sieci inteligentnych instytucji (Supersociety), która zamiast monolitycznego zarządzania, stawia na pluralizm infrastrukturalny i niezależność procesów poznawczych. Wolność w świecie po AGI ma zatem zależeć od otwartych standardów oraz możliwości zmiany dostawcy rozszerzeń, co czyni politykę konkurencji fundamentem wolności poznawczej.

The article analyzes the vision of Hybrid Humanity, where AI cognitive augmentations become tools for real agency. The author refers to the 'capability approach', indicating that technology should increase actual human capabilities rather than being a mere luxury for elites. A key threat is the emergence of new classes based on coupling quality—a divide between those with audited systems and those dependent on models optimized for corporate interests. The text advocates for the creation of a polycentric network of intelligent institutions (Supersociety) which, instead of monolithic management, emphasizes infrastructural pluralism and the independence of cognitive processes. Freedom in a post-AGI world thus depends on open standards and the ability to switch augmentation providers, making competition policy a foundation of cognitive freedom.

Artykuł analizuje ewolucję gatunku ludzkiego w erze post-AGI, odrzucając wizję jednego superinteligentnego bytu na rzecz koncepcji Hybrydowej Ludzkości. Autor stawia tezę, że kluczowym wyzwaniem przyszłości nie jest sama sprawność techniczna modeli AI, lecz zaprojektowanie pluralistycznych instytucji i ram prawnych, które zapobiegną nowym formom infrastrukturalnej dominacji. Tekst argumentuje, że rozszerzenia kognitywne mogą stać się narzędziem realnej sprawczości i wyrównywania szans, o ile zostaną osadzone w systemie gwarantującym autonomię jednostki i prawo do poznawczej idiosynkrazji. Zamiast skupiać się na binarnym podziale człowiek-maszyna, autor proponuje przejście od mikro-kontroli nad algorytmami do kontroli konstytucyjnej, gdzie człowiek występuje jako projektant relacji i gwarant wartości normatywnych. W tej perspektywie superinteligencja nie jest cechą pojedynczego podmiotu, lecz właściwością dobrze zorganizowanych, hybrydowych instytucji zdolnych do mądrego zarządzania niepewnością i błędami.

SŁOWA KLUCZOWE

Hybrydowa Ludzkość | architektura kognitywna | AGI | rozszerzenie kognitywne | pluralizm infrastrukturalny | capability approach | meaningful human control | constitutional control | epistemic authority | normative authority | human-AI-institution system | neurozróżnicowanie | sprawiedliwość kognitywna | superintelligence governance

Rozszerzenie kognitywne jako narzędzie realnej sprawczości

Na poziomie antropologicznym kluczowym pytaniem pozostaje kwestia ciągłości człowieka w obliczu zmiany jego własnych granic funkcjonalnych. Człowieczeństwo nie wymaga bowiem ochrony każdego historycznego ograniczenia biologicznego – medycyna, edukacja i technologia od dawna przekształcają sposób, w jaki żyjemy. Nie ma jednak powodu, by uznać maksymalizację technologicznego enhancement za cel sam w sobie. Rozszerzenie powinno być oceniane przez pryzmat tego, czy realnie zwiększa możliwości człowieka w prowadzeniu wybranego życia, utrzymywaniu relacji, uczestniczeniu w społeczeństwie i modyfikowaniu własnych celów. Właściwą

przeciwwagą dla transhumanistycznego maksimum jest tutaj capability approach Amartyi Sena i Marthy Nussbaum. Istotne jest nie tyle to, ile zasobów technicznych posiada człowiek, lecz co dzięki nim rzeczywiście może robić i kim może się stawać. AI może zwiększyć capability osoby z niepełnosprawnością, osoby neurozróżnicowanej, pracownika pozbawionego dostępu do eksperta czy mieszkańca regionu wykluczonego z określonych usług. Może jednak także te możliwości ograniczyć, jeśli dostęp będzie warunkowy, infrastruktura nieprzenośna, a człowiek utraci kompetencje niezbędne do opuszczenia systemu.

Źródło wskazuje wprost, że integracja kognitywna może pełnić funkcję wyrównującą dla osób neurodywergentnych, zamiast być wyłącznie luksusowym ulepszeniem dla elit. Bardziej ostrożna naukowo perspektywa powinna rozszerzyć tę intuicję na problem sprawiedliwości kognitywnej. Jeżeli AI stanie się podstawową infrastrukturą edukacji, pracy i administracji, nierówny dostęp do wysokiej jakości systemów może działać podobnie jak dysproporcje w dostępie do edukacji, internetu czy ochrony zdrowia. Pytanie nie będzie brzmieć jedynie: kto „ma AI", lecz kto posiada system chroniący jego interesy, prywatność i uwagę, a kto korzysta z rozwiązania optymalizowanego pod dyktando reklamodawcy lub pracodawcy. Właśnie dlatego Hybrydowa Ludzkość może podzielić się nie na „ludzi" i „postludzi", lecz na różne klasy jakości sprzężenia.

Jedni będą dysponować prywatnymi, przenośnymi i audytowalnymi systemami dopasowanymi do własnych celów. Inni otrzymają tańsze rozwiązania, których model biznesowy opiera się na ekstrakcji danych i kształtowaniu zachowań. Nierówność nie musi więc przybrać formy futurystycznego podziału biologicznego; może być znacznie bardziej prozaiczna, a przez to politycznie realna.

Kwestia ta łączy się z klasyczną teorią kapitału. Jeśli najlepsze rozszerzenia poznawcze pozostaną własnością nielicznych podmiotów, produktywność wynikająca z hybrydyzacji może zwiększać stopę zwrotu z posiadania infrastruktury. Jeżeli natomiast narzędzia staną się szeroko dostępne, przenośne i interoperacyjne, mogą wzmocnić zdolność negocjacyjną jednostek oraz małych organizacji. Ta sama technologia może zatem działać jako mechanizm koncentracji lub demokratyzacji, zależnie od instytucjonalnego modelu dystrybucji. Nie ma sensu przeciwstawiać „AI" i „rynku", podobnie jak nie wystarczy apelować o „regulację". Konieczna jest analiza konkretnych mechanizmów: efektów sieciowych, kosztów wyjścia, prawa do przenoszenia pamięci, otwartych standardów, dostępu do compute, konkurencji modeli oraz publicznej infrastruktury. W świecie osobistych agentów polityka konkurencji stanie się w części polityką wolności poznawczej, gdyż możliwość zmiany dostawcy może decydować o możliwości zmiany własnego rozszerzonego systemu działania.

To prowadzi ponownie do pytania o Supersociety. Najbardziej wartościowa wersja tej koncepcji nie przedstawia społeczeństwa jako monolitycznego mózgu, lecz jako policentryczną sieć inteligentnych instytucji, zdolnych do współdzielenia informacji bez utraty autonomii i możliwości wzajemnego kwestionowania. Taki system może posiadać znacznie większą zbiorową kompetencję niż dzisiejsze społeczeństwo, zachowując jednocześnie prawo jednostek i lokalnych wspólnot do odmiennego działania. W tym miejscu zasada Ashby'ego spotyka się z pluralizmem politycznym: złożone społeczeństwo wymaga regulatora o dostatecznej różnorodności wewnętrznej. Centralny supermodel może nie posiadać tej społecznej „wymaganej różnorodności", nawet jeśli przewyższa każdego człowieka w tysiącach testów.

Społeczeństwo składa się z wielu wartości, kultur, lokalnych uwarunkowań i konfliktujących dóbr. Jego pełna reprezentacja wymaga zatem nie jednego perfekcyjnego modelu, lecz ekosystemu modeli i instytucji. Rosnąca inteligencja społeczna może wymagać zachowania częściowej niezgody. Konsensus jest wartościowy, gdy wynika z niezależnej konfrontacji dowodów; traci na wartości, gdy staje się produktem jednego upstreamowego filtra. W świecie Hybrydowej Ludzkości należy chronić niezależność procesów dochodzenia do wniosków, a nie tylko wielość finalnych głosów. Dotyczy to również polityki. Milion obywateli może formalnie podejmować niezależne decyzje, ale jeśli większość z nich konstruuje argumentację poprzez ten sam system rekomendacyjny, demokratyczny pluralizm staje się jedynie ukrytą korelacją.

Pluralizm kognitywny i instytucjonalna kontrola nad hybrydyzacją

Nie wynika z tego potrzeba zakazania wspólnych narzędzi, lecz konieczność pluralizmu infrastrukturalnego, audytu i przejrzystości w kwestii tego, jak systemy uczestniczą w formowaniu naszych przekonań. Na poziomie kultury należy chronić także przestrzeń nieoptymalności. Sztuka, nauka i filozofia często rozwijają się dzięki ideom oddalonym od statystycznego centrum istniejących danych. Model optymalizowany pod kątem prawdopodobieństwa będzie preferował rozwiązania semantycznie osadzone w znanych wzorcach. Choć AI potrafi generować oryginalne kombinacje, społeczeństwo nie powinno utożsamiać płynnej generatywności z pełną różnorodnością kulturową. Ludzka biografia, lokalna historia i sytuacyjne doświadczenie wciąż mogą rodzić perspektywy słabo reprezentowane w zbiorach treningowych. Hybrydowa Ludzkość powinna chronić prawo do poznawczej idiosynkrazji. Nie każda różnica wymaga korekty w stronę statystycznie lepszego rozwiązania; neurozróżnicowanie, style twórcze, tradycje lokalne czy mniejszościowe sposoby konceptualizacji mogą wzbogacać zbiorową różnorodność poznawczą.

Oczywiście nie każde odstępstwo jest wartością – błędne przekonanie pozostaje błędem – jednak należy odróżnić pomyłkę od różnicy metodologicznej lub aksjologicznej.

W tym miejscu nauka o innowacji łączy się z teorią ewolucji. System adaptacyjny potrzebuje zarówno eksploatacji sprawdzonych rozwiązań, jak i eksploracji nowych. Nadmierna optymalizacja (exploitation) może zwiększyć bieżącą efektywność, ale jednocześnie obniżyć zdolność do odkrycia nowego optimum. AI jest doskonałym narzędziem eksploatacji ogromnych zasobów istniejącej wiedzy, lecz powinna być projektowana tak, by wspierać eksplorację i nie zamykać człowieka w rozkładzie własnej przeszłości. To właśnie tutaj filozofia nauki, ekonomia innowacji i psychologia twórczości spotykają się z koncepcją human-centered AI.

Najbardziej wartościowy system nie zawsze podaje użytkownikowi to, co najbardziej prawdopodobne. Czasem powinien wskazać mało prawdopodobną, lecz testowalną alternatywę, zakwestionować założenie lub odnaleźć analogię w odległej dyscyplinie. Partnerstwo poznawcze powinno poszerzać przestrzeń sensownych pytań, a nie tylko doprecyzowywać odpowiedzi. Tak można zdefiniować dojrzały Explanation Engine: jego zadaniem nie jest ostateczne „wyjaśnienie świata", lecz zwiększenie generatywnej i krytycznej zdolności cywilizacji do tworzenia lepszych modeli rzeczywistości. Najwyższą wartością nie jest pojedyncza odpowiedź, lecz usprawnienie procesu, dzięki któremu powstają kolejne. W takim ujęciu superinteligencja przestaje być absolutną cechą jednego bytu, a staje się relacją między systemem a klasą problemów. Organizacja może być superludzka w integracji danych epidemiologicznych, pozostając słabą w rozstrzyganiu kwestii wartości. Model może przewyższać człowieka w analizie przestrzeni wariantów, nie posiadając jednak instytucjonalnego mandatu do wyboru jednego z nich. Człowiek, choć słabszy obliczeniowo, pozostaje niezbędny jako prawny i moralny uczestnik procedury. Różne rodzaje przewagi mogą współistnieć bez jednego zwycięzcy całej kognicji.

Zmienia to również podejście do problemu „kontroli nad superinteligencją". W świecie hybrydowym kontrola nie będzie przypominać operatora trzymającego czerwony przycisk nad maszyną. Będzie raczej podobna do zarządzania współczesnym państwem, rynkiem finansowym czy internetem: rozproszona między prawo, standardy, instytucje, redundancję, monitoring i możliwość interwencji. Governance superinteligencji staje się zatem problemem projektowania instytucji, a nie szukania pojedynczego bezpiecznika. Z tego powodu ostatecznym przeciwieństwem nie jest relacja człowiek versus AI.

Bardziej trafne rozróżnienie przebiega między hybrydą kontrolowalną a dominującą – między systemem zachowującym zdolność korekty a takim, który jest zbyt głęboko osadzony, by można go było zakwestionować. Kryterium nie jest biologiczna czystość, lecz struktura sprawczości. Tak dochodzimy do pojęcia Hybrydowej Ludzkości jako formacji cywilizacyjnej. Nie oznacza ono

powstania nowego gatunku biologicznego, lecz społeczeństwo, w którym pamięć, percepcja, decyzje, praca, relacje i instytucje są coraz częściej współrealizowane przez komponenty biologiczne i sztuczne. Hybrydowość nie ma jednej głębokości.

Jedna osoba może korzystać z AI sporadycznie, druga funkcjonować z agentem posiadającym wieloletnią pamięć, a trzecia używać BCI. Wspólną cechą tej formacji jest to, że normy społeczne zaczynają być projektowane z założeniem powszechnej obecności sztucznej warstwy kognitywnej. W takim układzie osoba niekorzystająca z AI może znaleźć się w sytuacji podobnej do współczesnego użytkownika pozbawionego dostępu do internetu: formalnie zachowuje wolność, lecz środowisko instytucjonalne sprawia, że koszty tej decyzji drastycznie rosną. Dlatego prawo do nieaugmentowania się nie może być rozumiane wyłącznie formalnie; należy badać, czy człowiek zachowuje realną możliwość uczestnictwa społecznego bez przyjęcia najbardziej inwazyjnych technologii. Jest to klasyczny problem socjologii normalizacji: gdy technologia staje się powszechna, zmienia normę oczekiwanej wydajności. Człowiek nie staje się biologicznie mniej sprawny, ale może zostać uznany za mniej produktywnego. Właśnie w ten sposób dobrowolny enhancement może przekształcić się w strukturalny przymus ulepszania. Jeśli każdy konkurent korzysta z AI, rezygnacja z niej przestaje być neutralnym wyborem. Dlatego przyszła polityka pracy powinna obejmować nie tylko ochronę przed zastąpieniem przez AI, lecz także warunki uczciwej augmentacji: kto płaci za narzędzie, kto posiada dane, czy pracownik może przenieść wypracowaną wiedzę i czy system monitoruje jego psychologię?

Czy wzrost produktywności przełoży się na czas wolny, płace i bezpieczeństwo, czy jedynie na intensyfikację norm pracy? Te pytania należą do domeny ekonomii politycznej, a nie benchmarków modeli.

Pluralizm kognitywny i przewaga hybrydowych instytucji

W ochronie zdrowia największą wartością AI może być zwiększenie dostępności i jakości diagnostyki, jednak ryzykiem pozostaje centralizacja infrastruktury oraz deskilling. W edukacji korzyścią jest personalizacja, kosztem zaś nadmierny offloading. W prawie zyskujemy dostęp do ogromnej pamięci, ale ryzykujemy zamrożenie historycznych wzorców. W każdym sektorze istnieje zatem inna funkcja optymalna hybrydy.

Jedna uniwersalna teoria mówiąca, że „AI zastępuje” lub „AI wspiera”, jest niewystarczająca. Przyszłość po AGI będzie prawdopodobnie heterogeniczna. Część domen osiągnie wysoki poziom

automatyzacji, inne pozostaną hybrydowe z powodów prawnych, moralnych bądź technicznych. Niektórzy będą dążyć do głębokiej integracji, inni wybiorą mniejszą ingerencję. Nie należy oczekiwać jednej globalnej trajektorii; to właśnie główny argument za pluralizmem antropologicznym. Człowieczeństwo nie musi zostać zdefiniowane przez jedną, obowiązkową konfigurację technologiczną. Społeczeństwo może uznawać równość statusu osób funkcjonujących z różnym stopniem rozszerzenia. Implant, asystent AI, neuroproteza czy decyzja o ich nieużywaniu nie powinny tworzyć hierarchii moralnej. Takie stanowisko pozwala zachować najlepsze elementy transhumanizmu bez przyjmowania jego najbardziej radykalnych założeń: technologia może realnie zmniejszać cierpienie, kompensować ograniczenia i zwiększać ludzkie możliwości.

Nie ma powodu absolutyzować istniejących ograniczeń biologicznych tylko dlatego, że są „naturalne". Jednocześnie nie należy traktować wszystkich ograniczeń jako błędów inżynieryjnych ani budować normy człowieka wokół maksimum produktywności. Wartość Hybrydowej Ludzkości będzie zatem zależała nie od tego, jak daleko przekroczy biologiczne granice, lecz od tego, czy rozszerzenie zwiększa zakres możliwych dobrych żyć bez tworzenia nowych form dominacji. Jest to kryterium jednocześnie etyczne, polityczne i ekonomiczne.

W tym miejscu pojęcie mądrości można odnieść nie tylko do jednostki, ale do całej cywilizacji. Cywilizacja mądra nie jest tą, która rozwiązuje każdy problem najszybciej. Jest systemem potrafiącym rozpoznać, które problemy wymagają szybkości, a które deliberacji; gdzie niepewność jest nieusuwalna, kiedy należy chronić głos mniejszości i w jakich momentach wzrost możliwości technicznych nie powinien zostać wykorzystany. Zdolność do niewykonania technicznie możliwego działania jest również kompetencją. Taką kompetencję można nazwać cywilizacyjną kontrolą hamowania. Biologiczne funkcje wykonawcze obejmują bowiem nie tylko inicjację zachowania, ale i jego hamowanie. Społeczeństwa również posiadają mechanizmy inhibicji: moratoria, sądy, komisje etyczne, procedury zgody czy badania bezpieczeństwa.

W epoce przyspieszonej AI ich rola może wzrosnąć. System bardzo inteligentny, lecz niezdolny do prawidłowego powstrzymania się od działania, jest funkcjonalnie niedojrzały. Właśnie tutaj doświadczenie, inteligencja i mądrość można złożyć w jedną strukturę. Inteligencja zwiększa repertuar możliwych działań; doświadczenie sprawia, że przeszłość modyfikuje sposób wyboru między nimi; metakognicja dostarcza informacji o granicach własnego modelu, a mądrość praktyczna wprowadza kontekst wartości i odpowiedzialności. Governance organizuje te właściwości na poziomie społecznym. Dopiero ich sprzężenie tworzy dojrzały system poznawczy. AGI obejmuje przede wszystkim pierwszy z tych elementów; AGE próbuje dołączyć drugi. Badania nad świadomością (consciousness) dotyczą części problemu podmiotowości, natomiast human-AI interaction dotyczy samego sprzężenia.

Prawo i polityka organizują odpowiedzialność, lecz żadna z tych dyscyplin nie wystarcza samodzielnie. To właśnie dlatego przyszłość AI jest z definicji problemem interdyscyplinarnym. Jeśli trzeba odpowiedzieć jednym teoretycznym zdaniem na pytanie „co poza AGI?", odpowiedź nie brzmi „kolejny poziom AI". Poza AGI znajduje się transformacja jednostki inteligencji: przejście od modelu izolowanego wykonawcy ku trwałym, historycznym i wielopoziomowym układom ludzi, maszyn, pamięci oraz instytucji.

Samo to przejście nie musi oznaczać powstania nowej świadomości; wystarczy, że nowe funkcje pojawią się na poziomie całego systemu. Właśnie dlatego najbardziej realistyczną „superinteligencją" najbliższej przyszłości może być nie odizolowany cyfrowy umysł, ale dobrze zorganizowana hybrydowa instytucja. Będzie ona dysponować wieloletnią pamięcią, flotą wyspecjalizowanych agentów, ludzkimi ekspertami, automatyczną analizą niepewności, narzędziami wykonawczymi i procedurami samokorekty. Jej przewaga nad klasyczną organizacją może być ogromna nawet wtedy, gdy żaden pojedynczy element nie spełnia rygorystycznej definicji AGI.

Taki wniosek zmienia także geopolitykę AI. Rywalizacja państw nie będzie dotyczyć wyłącznie tego, kto stworzy „najmądrzejszy model". Równie istotna stanie się zdolność integracji modeli z edukacją, administracją, nauką, przemysłem, bezpieczeństwem i infrastrukturą energetyczną. Państwo posiadające nieco słabsze modele, ale lepszą architekturę instytucjonalną, może osiągnąć większą zdolność zbiorową niż państwo dysponujące najpotężniejszym modelem, który jest jednak słabo wdrożony.

Jest to analogiczne do historycznej roli technologii ogólnego przeznaczenia. Sama maszyna parowa nie tworzyła przewagi, jeśli gospodarka nie potrafiła przebudować fabryk, logistyki i instytucji. Elektryczność nie zmieniła produktywności w pełni, dopóki organizacja pracy była projektowana pod wcześniejsze źródła energii. AI podobnie może wymagać komplementarnych innowacji organizacyjnych. W takim świecie AGI może być politycznie mniej istotne niż zdolność do tworzenia general-purpose hybrid institutions. To przesunięcie zasługuje na szczególną uwagę, ponieważ debata technologiczna z natury skupia się na tym, co można wykazać w benchmarku. Przewaga instytucjonalna jest trudniejsza do uchwycenia i rozwija się wolniej, ale może okazać się trwalsza.

Od mikro-kontroli do konstytucyjnego projektowania relacji

To właśnie na poziomie instytucjonalnym rozstrzygnie się, czy AI zwiększy demokratyczną sprawczość, czy technokratyczną asymetrię. System publiczny może wspierać obywatela w rozumieniu prawa i uczestnictwie w konsultacjach, ale może również wzmocnić zdolność aparatu państwowego do klasyfikowania jednostek. Nie istnieje „neutralna inteligencja administracyjna" – każde rozwiązanie zostaje włączone w istniejące relacje władzy. Polityka technologiczna staje się zatem integralną częścią polityki konstytucyjnej. Decyzje o tym, jakie dane państwo może łączyć, które modele mogą rekomendować rozstrzygnięcia, jak długo przechowywana jest pamięć i kiedy obywatel ma prawo do kontaktu z człowiekiem, wyznaczają granice władzy równie realnie, co tradycyjne normy proceduralne. W obliczu AGI wyłania się zatem nowa teoria cyfrowej zwierzchności człowieka.

Nie należy jej rozumieć jako wymogu, by człowiek zatwierdzał każdą mikrodecyzję – w złożonych systemach byłoby to niemożliwe. Zwierzchność oznacza prawo do ustanawiania reguł, wyznaczania granic uprawnień i definiowania procedur zmiany. Człowiek przestaje być operatorem, a staje się konstytucyjnym projektantem relacji z agentem. Jest to właściwy model kontroli dla systemów coraz bardziej autonomicznych. Pilot nie steruje każdym parametrem nowoczesnego samolotu, jednak cały system lotniczy funkcjonuje w ramach ścisłych procedur, certyfikacji, rozdziału odpowiedzialności i ograniczeń technicznych. Analogicznie: wysoka autonomia AI nie musi wykluczać ludzkiej zwierzchności, o ile autonomia ta jest delegowana warunkowo i odwołalnie. Warunkowa delegacja może być bardziej realistyczną podstawą przyszłej sprawczości niż próba utrzymywania człowieka „w pętli" każdej operacji. Przy milionach mikrooperacji na sekundę człowiek nie jest w stanie sensownie zatwierdzać każdego kroku; powinien natomiast zachować władzę nad klasami działań, progami ryzyka, eskalacją i możliwością audytu. Jest to przejście od micro-control do constitutional control.

Taka perspektywa wyjaśnia, dlaczego przyszłość nie wymaga wyboru między automatyzacją a humanizmem. Można zwiększać stopień automatyzacji, jednocześnie rozszerzając zakres ludzkiej kontroli na poziomie systemowym. Warunkiem jest właściwy projekt architektury. Automatyzacja staje się problemem wtedy, gdy znosi zdolność do zmiany zasad, a nie wtedy, gdy przejmuje wykonawczą część pracy.

W tym punkcie rozumowanie powraca do Simona. Ograniczona racjonalność oznacza, że człowiek nigdy nie przeanalizuje osobiście wszystkich informacji dostępnych przyszłym systemom. Próba zachowania autonomii poprzez pełną weryfikację każdej decyzji jest więc iluzją. Potrzebujemy

instytucjonalnych skrótów poznawczych: zaufanych procedur, audytów, certyfikacji i jasnego podziału kompetencji. Problemem nie są pośrednicy, lecz ich rozliczalność i możliwość zmiany. Z tej perspektywy epistemiczna pokora staje się zasadą systemową. Człowiek musi mieć świadomość, że nie potrafi samodzielnie ocenić wszystkiego; AI powinna funkcjonalnie reprezentować granice własnej pewności, instytucja – wiedzieć, gdzie jej procedury są słabo zweryfikowane, a społeczeństwo – zachować świadomość obszarów naukowej niewiedzy. Najwyższym poziomem dojrzałości systemu nie jest bowiem pewność, lecz dobrze skalibrowana zdolność do wątpienia.

W tym znaczeniu można na nowo odczytać sokratejską tradycję filozofii. „Wiedzieć, że się nie wie" nie jest przejawem skromności, lecz metapoznawczą kompetencją pozwalającą prawidłowo alokować wysiłek poznawczy. Przyszła AI o ogromnej wiedzy będzie szczególnie potrzebowała funkcjonalnego odpowiednika takiej postawy, gdyż jej płynność językowa może maskować granicę między rzeczywistą wiedzą a interpolacją. Najbardziej wartościowy system nie będzie tym, który odpowiada na każde pytanie, lecz takim, który potrafi zaklasyfikować niepewność, wskazać brak danych, oddzielić konflikt wartości od konfliktu faktów, odmówić pozornej precyzji i przekazać problem do innego źródła. To może stanowić jeden z najważniejszych warunków Artificial General Experience rozumianego funkcjonalnie: system nie tylko pamięta błędy, ale lepiej rozpoznaje sytuacje, w których wcześniej nie powinien ufać własnemu osądowi. Wówczas AGE przestaje być wyłącznie kategorią filozoficzną i staje się programem metauczenia odpowiedzialności operacyjnej. System po popełnieniu błędu zmniejsza własne uprawnienia, podnosi wymagany poziom dowodu lub zmienia politykę eskalacji. Nie jest to cierpienie, lecz realna transformacja przez konsekwencję.

Podobnie człowiek może rozwijać się dzięki AI nie dlatego, że system zawsze poprawnie go prowadzi, ale dlatego, że uczestniczy w jego procesie metakognicji. AI może zwracać uwagę na powtarzalne błędy, przypominać nieskuteczne strategie i ukazywać różnicę między deklarowaną pewnością a trafnością wcześniejszych decyzji. Byłaby wtedy lustrem poznawczym, a nie tylko generatorem odpowiedzi. Taki model posiada jednak potencjał głęboko paternalistyczny – system znający historię człowieka zyskuje niezwykłą zdolność modelowania jego psychologii. Dlatego każda metakognitywna pomoc wymaga prawa do kontroli nad tym modelem. Użytkownik powinien wiedzieć, jakie cechy są inferowane i mieć możliwość wyłączenia części personalizacji; w przeciwnym razie narzędzie samowiedzy stanie się infrastrukturą manipulacji. Granica między augmentacją a manipulacją przebiega przez interes oraz możliwość refleksyjnej odmowy. Jeśli system pomaga użytkownikowi realizować jego własny, refleksyjnie zaakceptowany cel i transparentnie ujawnia środki wpływu, mamy do czynienia z augmentacją. Jeśli wykorzystuje wiedzę psychologiczną do realizacji niejawnego interesu dostawcy – powstaje manipulacja.

Pluralizm instytucjonalny jako gwarant autonomii kognitywnej

Kryterium to powinno stać się fundamentem przyszłego prawa regulującego agentów osobowych. Właśnie taki agent może okazać się jednym z najważniejszych artefaktów ery po AGI — nie dlatego, że sam stanie się osobą, lecz ponieważ będzie reprezentował człowieka w coraz szerszym obszarze świata cyfrowego. Będzie negocjował, filtrował informacje, zarządzał pamięcią, komunikował się z innymi systemami, a być może podejmował niskiego ryzyka decyzje finansowe i administracyjne. Wówczas „moja AI" staje się częściowo wykonawczą reprezentacją mojego interesu.

Taki agent wymaga ram przypominających obowiązek powierniczy. Jeśli system zna użytkownika lepiej niż większość instytucji i działa w jego imieniu, nie może jednocześnie potajemnie maksymalizować zysków platformy. W przeciwnym razie powstaje strukturalny konflikt: człowiek traktuje system jako rozszerzenie własnego umysłu, podczas gdy ten pozostaje ekonomicznie lojalny wobec innego podmiotu. Ostatecznie optymalna infrastruktura osobowej AI powinna przypominać cyfrową instytucję powierniczą, w której prawo precyzuje zasady lojalności, poufności, przenoszenia pamięci oraz rozwiązywania konfliktów interesów. Nie jest to jedyny możliwy model, lecz pokazuje on, że po AGI pojawią się problemy prawne jakościowo inne niż tradycyjne licencje na oprogramowanie.

Na poziomie cywilizacyjnym analogiczną rolę będą pełnić instytucje epistemiczne. Biblioteki, uniwersytety, archiwa i publiczne repozytoria mogą stać się przestrzeniami, w których obywatel uzyska dostęp do AI nieoptymalizowanej pod komercyjną retencję. Byłby to cyfrowy odpowiednik infrastruktury wiedzy traktowanej jako dobro publiczne. Nie chodzi tu o zastąpienie modeli komercyjnych publicznymi; pluralizm wymaga właśnie współistnienia różnych form własności. Wartość publicznej infrastruktury tkwi w utrzymaniu alternatywy, która ogranicza dominację jednego mechanizmu rynkowego i pozwala nauce, edukacji oraz społeczeństwu obywatelskiemu na swobodne eksperymentowanie. Hybrydowa Ludzkość potrzebuje zatem nie jednej AI, lecz całego ekosystemu instytucji AI.

Niektóre z nich będą prywatne i osobiste, inne publiczne, naukowe, lokalne, branżowe lub open source. Bez takiego pluralizmu kognicja zbiorowa mogłaby ulec zbyt silnej koncentracji. Prowadzi to do najbardziej dojrzałej wersji Supersociety, która nie jest społeczeństwem połączonym w jeden, perfekcyjnie zsynchronizowany mózg, lecz federacją interoperacyjnych systemów poznawczych. Systemy te zachowują lokalną autonomię i konkurencję teorii, ale potrafią współdzielić dowody, ostrzegać przed ryzykami i koordynować działania zbiorowe.

Można tu dostrzec analogię do federalizmu: nie wszystkie kompetencje muszą trafiać do centrum. Część decyzji lepiej podejmować lokalnie, w oparciu o specyficzną wiedzę z danego obszaru; inne wymagają koordynacji globalnej ze względu na efekty zewnętrzne. Teoria kognicji hybrydowej może zatem korzystać z zasady subsydiarności: decyzja powinna zapadać na najniższym poziomie zdolnym skutecznie zarządzić jej konsekwencjami. Ma to również wymiar bezpieczeństwa. Zbyt scentralizowany Metamind tworzy pojedynczy punkt awarii i politycznej dominacji, podczas gdy całkowicie zdecentralizowana struktura może okazać się niezdolna do reakcji na problemy globalne. Inteligentny system wymaga warstwowej koordynacji. Najwyższą formą przyszłej kognicji może być zatem nie maksymalna centralizacja inteligencji, lecz maksymalna zdolność do dynamicznego przełączania skali.

Problem lokalny pozostaje lokalny. Sygnał systemowy uruchamia szerszą koordynację. Konflikt wartości trafia do procedury politycznej, a nie do optymalizatora; niepewność naukowa inicjuje badanie, a nie arbitralny wybór. To właśnie można nazwać cywilizacyjną funkcją wykonawczą. Podobnie jak człowiek, społeczeństwo musi alokować uwagę, wyznaczać cele, hamować niewłaściwe działania, pamiętać błędy i korygować strategię.

AI może gwałtownie zwiększyć wydajność tych funkcji, ale może też prowadzić do ich centralizacji. Kluczowym projektem nie będzie więc ich automatyzacja, lecz odpowiednie rozproszenie i konstytucjonalizacja. W tym punkcie można już precyzyjniej odpowiedzieć na pytanie o to, co znajduje się „poza AGI”. Nie jest to jedna rzecz, lecz przestrzeń transformacji, w której zdolność maszynowa, sztuczna historyczność, agentowość, rozszerzona ludzka kognicja, organizacyjna inteligencja zbiorowa i infrastrukturalna władza zaczynają wzajemnie się sprzęgać. AGE jest jednym z języków opisujących część tej przestrzeni, lecz jej nie wyczerpuje. Fused Mind pozostaje hipotezą o głębokiej integracji, ale nie jedyną możliwą trajektorią. Metamind i Supersociety mogą być użytecznymi kategoriami funkcjonalnymi bez konieczności postulowania jednego fenomenalnego superpodmiotu.

AI jako konstrukcja niszy poznawczej i wyzwanie odporności

Najbardziej realistyczne wydaje się zatem odejście od drabiny na rzecz mapy. O ile na drabinie każdy kolejny stopień zastępuje poprzedni, o tyle mapa oferuje wiele równoległych kierunków rozwoju. Jeden z nich prowadzi ku większej generalności modeli, drugi ku agentowości, trzeci ku pamięci i AGE, czwarty ku robotycznemu ucieleśnieniu, piąty ku BCI, szósty ku hybrydowym

organizacjom, a siódmy ku zbiorowej kognicji. Procesy te mogą postępować w różnym tempie i łączyć się w odmiennych konfiguracjach.

Właśnie dlatego termin „post-AGI" należy stosować ostrożnie, gdyż sugeruje on epokę następującą po pojedynczym wydarzeniu technicznym. Trafniej byłoby mówić o post-tool AI – świecie, w którym system sztuczny przestaje być jedynie narzędziem uruchamianym doraźnie, a staje się trwałym komponentem środowiska poznawczego. Przejście to już trwa, lecz nie powinno być opisywane tak, jakby jego rezultat był z góry przesądzony. Przyszłość ma charakter projektowy: instytucje stoją przed wyborem między systemami wzmacniającymi kompetencje a tymi, które je zastępują; między interoperacyjnością a lock-in; między ujawnianiem niepewności a pozorną pewnością oraz między pluralizmem a monokulturą modeli. Kluczowe pytanie nie dotyczy zatem AI jako „innego gatunku", lecz kształtu gatunku ludzkiego w środowisku własnych wytworów. Homo sapiens zawsze przekształcał otoczenie, by następnie rozwijać się w świecie już przez siebie zmienionym – proces ten określa się mianem niche construction.

Rolnictwo, miasta, pismo, pieniądz, szkoła, prawo czy internet nie były jedynie narzędziami, lecz niszami zmieniającymi presję społeczną i kognitywną kolejnych pokoleń. AI może stać się jedną z najgłębszych dotychczas form konstrukcji niszy poznawczej, co stanowi biologiczno-kulturowego mostu między technologią i ewolucją. Nie musimy oczekiwać gwałtownych zmian genetycznych, by mówić o ewolucji sposobu funkcjonowania; kultura i instytucje modyfikują środowisko selekcji kompetencji znacznie szybciej niż biologia. Dziecko wychowane w świecie permanentnie dostępnych agentów może wykształcić inne strategie pamięci, wyszukiwania informacji czy rozwiązywania problemów niż poprzednie pokolenia. To realny proces rozwojowy, niezależny od koncepcji „postczłowieka", co rodzi konieczność prowadzenia badań longitudinalnych.

Krótkie eksperymenty potrafią wykazać, czy człowiek wykonuje zadanie lepiej z pomocą AI, lecz nie odpowiedzą na pytanie, kim staje się po dziesięciu latach codziennej współpracy. Ten długi horyzont czasowy będzie kluczowy dla nauki o Hybrydowej Ludzkości. Należy badać w czasie rozwój kompetencji, pamięci, metakognicji, relacji, zaufania oraz autonomii i zależności infrastrukturalnej. Podobnie organizacje wymagają analizy dynamicznej: system może poprawiać KPI w pierwszym roku, by w piątym drastycznie obniżyć odporność organizacji. Może zwiększać bieżącą produktywność, jednocześnie niszcząc ścieżki kształcenia przyszłych ekspertów lub centralizować wiedzę tak skutecznie, że po awarii zewnętrznego dostawcy instytucja utraci zdolność samodzielnego funkcjonowania.

Efektywność statyczna i odporność dynamiczna są odrębnymi właściwościami. Ta różnica sprawia, że dyskusja o AI wymaga silniejszego osadzenia w teorii resilience. System odporny nie tylko unika błędów, ale potrafi przystosować się po wystąpieniu szoku, co wymaga redundancji, różnorodności,

pamięci i zdolności do rekonfiguracji. Można zatem uznać, że dojrzałe AGE na poziomie organizacyjnym jest elementem resilience: błąd z przeszłości modyfikuje przyszłą strukturę. W tej perspektywie „mądra" Hybrydowa Ludzkość nie będzie cywilizacją wolną od błędów – taki ideał byłby bowiem niemożliwy.

Instytucjonalna pamięć błędów jako fundament odpowiedzialnej AI

Będzie to cywilizacja charakteryzująca się wysoką jakością uczenia się na błędach, bez konieczności ich ukrywania w celu ochrony reputacji poszczególnych ogniw. Just culture, transparentne raportowanie incydentów i procedury rewizji stają się zatem immanentnym elementem inteligencji zbiorowej. Psychologia społeczna przypomina, że grupy nie uczą się efektywnie w środowisku tłumiącym sprzeciw. Groupthink Janisa, pluralistyczna ignorancja czy badania nad konformizmem obnażają mechanizmy, przez które grupa osiąga pozorną zgodność kosztem jakości informacji. AI może te problemy łagodzić, jeśli umożliwi anonimowe generowanie kontrargumentów, lub potęgować, jeśli system podsumowujący spotkanie zbyt pochopnie zamieni rozbieżne stanowiska w jedną wersję „konsensusu". Dlatego jednym z kryteriów inteligentnej organizacji hybrydowej powinno być preservation of dissent. Pamięć instytucjonalna nie powinna archiwizować wyłącznie decyzji końcowej; winna rejestrować istotne argumenty przeciw, warunki obowiązywania decyzji oraz założenia, które z czasem mogą okazać się błędne. Wówczas doświadczenie organizacji obejmuje nie tylko to, co zostało zrobione, ale i powody odrzucenia alternatywnych możliwości. To przykład tego, jak AI może podnieść jakość zbiorowej pamięci ponad tradycyjną dokumentację.

Agent mógłby w przyszłości przypomnieć, że mniejszościowy ekspert przewidział konkretne ryzyko, które wówczas uznano za mało prawdopodobne. Taka pamięć ograniczałaby hindsight bias i uczyła organizacje pokory wobec własnej historii. W ten sposób wracamy do sedna AGE: doświadczenie nie jako ornament ontologiczny, ale pamięć transformacyjna. Pojęcie to może okazać się znacznie bardziej użyteczne dla realnej przyszłości AI niż spekulacje o sztucznym cierpieniu. System z ewoluującą polityką może poprawiać bezpieczeństwo nawet bez qualiów. Jednocześnie należy zapytać, kto decyduje o tym, które doświadczenia mają transformować system. Uczenie się również podlega polityce. Model może zostać skorygowany po skardze kluczowego klienta, a nie po wyrządzeniu szkody mało wpływowej grupie. Organizacja może łatwiej „pamiętać" problemy finansowe niż społeczne. AGE pozbawione governance może zatem po prostu internalizować istniejącą w środowisku strukturę siły. W tym miejscu doświadczenie spotyka się ze sprawiedliwością. Nie wystarczy, aby system uczył się na konsekwencjach; trzeba określić, czyje

konsekwencje są traktowane jako sygnał do nauki. Jest to kluczowe w administracji, medycynie i zarządzaniu platformami.

Koszt błędu może być rozproszony na użytkowników, podczas gdy model optymalizuje metrykę istotną dla dostawcy. Dlatego koncepcja „internalizing cost" zyskuje sens ekonomiczno-instytucjonalny: system nie powinien jedynie rejestrować błędu – struktura organizacji winna internalizować koszty, które w przeciwnym razie zostałyby przerzucone na społeczeństwo. W ekonomii jest to klasyczny problem efektów zewnętrznych. Odpowiedzialna AI wymaga nie tylko algorytmicznego uczenia się z błędów, ale ekonomicznego mechanizmu, dzięki któremu podmiot podejmujący ryzyko nie może całkowicie eksportować jego kosztów na innych. Może się to okazać ważniejsze niż symulowanie w maszynie „cierpienia". Firma ma bodziec do poprawy bezpieczeństwa wtedy, gdy ponosi część kosztów szkody. Regulator, odpowiedzialność cywilna, obowiązek ubezpieczenia czy reputacyjny koszt incydentu mogą funkcjonalnie internalizować konsekwencje na poziomie instytucji. Doświadczenie odpowiedzialnego systemu jest więc częściowo pamięcią ekonomiczną. W tej perspektywie najbardziej zaawansowana AI nie powinna być odseparowana od społeczeństwa, lecz głęboko włączona w jego mechanizmy odpowiedzialności. To kolejny argument przeciw wizji autonomicznego „Boga-Maszyny".

Rozdzielenie autorytetu epistemicznego od normatywnego w zarządzaniu AI

Nawet najbardziej kompetentny system wymaga środowiska, które określi normatywną wagę jego działań. Sama hybrydyzacja nie gwarantuje jednak bezpieczeństwa, gdyż człowiek również bywa źródłem błędów, uprzedzeń i przemocy. Human in the loop nie jest moralnym sakramentem. W wielu przypadkach system sztuczny okaże się trafniejszy, bardziej konsekwentny i mniej podatny na określone biasy. Celem nie jest zatem maksymalizacja udziału człowieka, lecz optymalizacja jakości sprawczości całego układu przy jednoczesnym zachowaniu praw i odpowiedzialności. Prowadzi to do dojrzalszej koncepcji meaningful human control. Kontrola nie polega tu na założeniu, że człowiek zawsze ma rację, lecz na istnieniu legitymizowanego mechanizmu określającego cele, ograniczenia oraz procedury rewizji. Jeśli AI stawia lepszą diagnozę, człowiek powinien z niej racjonalnie korzystać. Jeżeli jednak decyzja dotyczy rozdziału ryzyka lub praw, nie może ona zostać zredukowana do samej przewagi predykcyjnej modelu.

To rozróżnienie między epistemic authority a normative authority stanie się jednym z fundamentalnych problemów ery po AGI. AI może uzyskać większy autorytet epistemiczny w konkretnych domenach dzięki skuteczniejszej analizie danych i lepszej predykcji, lecz nie nadaje

mu to automatycznie normatywnego mandatu do decydowania o celach. Społeczeństwo będzie musiało nauczyć się delegować wiedzę bez delegowania całej suwerenności. Właśnie w tym punkcie możemy zrekonstruować ideę człowieka jako podmiotu „nadającego sens i kierunek", nadając jej formę naukowo-polityczną. Nie chodzi o metafizyczne twierdzenie, że maszyna nie może posiadać znaczenia, lecz o zasadę, według której systemy oddziałujące na ludzi powinny realizować cele legitymizowane przez społeczności, których te dotyczą. To fundament demokratycznego governance, a nie antropocentryczny dogmat.

Przyjmując tę perspektywę, cały wywód układa się w model kilku nakładających się na siebie ewolucji. Pierwszą jest wzrost technicznej kompetencji AI; drugą – rozwój jej historyczności i agentowości. Trzecia dotyczy interfejsów: od języka, przez biosensory, aż po BCI. Czwarta to ewolucja ludzkich praktyk poznawczych pod wpływem offloadingu i współpracy. Piąta obejmuje instytucje reorganizujące się wokół AI, a szósta – normy i prawa próbujące regulować nowe zależności. Żadna z tych ścieżek nie przebiega synchronicznie, a to rozminięcie temp jest głównym źródłem ryzyka. Technologia zmienia się w miesiące, kompetencje zawodowe w lata, instytucje w dekady, natomiast normy kulturowe i prawo jeszcze wolniej. Powstaje regulatory and institutional lag. W tej luce mogą utrwalić się rozwiązania trudne do późniejszej zmiany z powodu lock-in i path dependence.

Dlatego przyszłe governance musi stać się bardziej adaptacyjne. Nie może zakładać, że raz ustanowione przepisy opiszą stabilny przedmiot przez dwadzieścia lat. Niezbędne są procedury okresowych przeglądów, monitorowanie efektów, regulacyjne sandboxy oraz standardy aktualizowalne szybciej niż ustawy. Jednocześnie podstawowe prawa powinny pozostać stabilną kotwicą, aby adaptacyjność nie stała się pretekstem do arbitralności. W ten sposób prawo samo staje się systemem hybrydowego uczenia: stabilne zasady konstytucyjne tworzą warstwę wolnozmienną, a regulacje techniczne – warstwę szybką. Taka wieloskalowa adaptacja odpowiada naturze systemów złożonych, w których różne poziomy zmieniają się z różną prędkością. To kolejny argument za tym, że jednowymiarowa narracja o AGI jest niewystarczająca.

Hybrydowa Ludzkość jako interdyscyplinarny program badawczy

Nie chodzi tylko o to, czy model przekroczył określony próg, lecz o to, czy otaczające go instytucje zdołały zbudować odpowiedni poziom kontroli. Ten sam model może być bezpieczny w jednym środowisku i niebezpieczny w innym; dlatego kontekst wdrożenia powinien stać się podstawową jednostką oceny ryzyka. Model generujący niepewną odpowiedź w grze ma zupełnie inne znaczenie

niż ten sam mechanizm w diagnozie medycznej czy wymiarze sprawiedliwości. Nie istnieje jedna, uniwersalna „bezpieczność modelu" oderwana od decyzji, które mu powierzamy. To ważna lekcja dla filozofii technologii: własność artefaktu jest relacyjna. System staje się elementem konkretnej praktyki i dopiero w niej uzyskuje pełne znaczenie społeczne. AI nie „zabiera sprawczości" w sposób abstrakcyjny.

Może ją zwiększać lub ograniczać, zależnie od tego, czy użytkownik dysponuje wiedzą, czasem, możliwością odmowy i alternatywą. Hybrydowa Ludzkość nie potrzebuje zatem jednej etyki AI, lecz etyki sprzężeń. Powinna ona badać nie tylko cechy maszyny i człowieka, ale to, co powstaje pomiędzy nimi. Zaufanie, zależność, podział odpowiedzialności i kontrola są bowiem własnościami relacji. Analogicznie można stwierdzić, że potrzebujemy socjologii sprzężeń, badającej adaptację praktyk społecznych do AI; ekonomii sprzężeń, analizującej rozkład wartości między właścicielem infrastruktury a użytkownikiem; prawa sprzężeń, regulującego delegację i odpowiedzialność; oraz epistemologii sprzężeń, badającej powstawanie przekonań w systemach mieszanych. Dopiero taki interdyscyplinarny aparat odpowiada rzeczywistej naturze problemu. Właśnie dlatego najbardziej produktywną jednostką przyszłych badań może być nie artificial intelligence, lecz human-AI-institution system. Człowiek, technologia i reguły tworzą całość dynamiczną: zmiana modelu bez zmiany procedury przyniesie inny efekt niż równoczesna reorganizacja instytucjonalna. Jeśli zmienimy interfejs, zmienimy zachowanie człowieka. System jest więc współkonstytuowany.

Taki punkt widzenia przesuwa również debatę o odpowiedzialności naukowców. Badacz AI nie tworzy neutralnego algorytmu, jeśli ma świadomość, że architektura zostanie włączona w społeczne procesy decyzyjne. Nie oznacza to odpowiedzialności za każde późniejsze użycie, ale wskazuje, że projektowanie nie może całkowicie abstrahować od przewidywalnych skutków interakcji. Podobny obowiązek spoczywa na naukach społecznych – krytyka AI nie może opierać się wyłącznie na abstrakcyjnych obawach przed automatyzacją.

Konieczne są badania konkretnych mechanizmów, efektów i populacji w długiej perspektywie czasowej. Hybrydowa Ludzkość nie powinna być ani mitem technologicznym, ani metaforą krytycznej teorii, lecz empirycznym programem badawczym. Program ten musiałby analizować dzieci dorastające z AI, profesjonalistów korzystających z systemów przez lata, organizacje o różnym stopniu automatyzacji, społeczności operujące na różnych modelach oraz ludzi korzystających z neurotechnologii. Wymagałoby to połączenia metod eksperymentalnych, longitudinalnych, etnograficznych, ekonomicznych i obliczeniowych.

Tylko w ten sposób można oddzielić intuicje od rzeczywistych procesów. Jest to kluczowe, gdyż wiele obecnych tez może okazać się błędnych. Być może część obaw o deskilling zostanie zrównoważona przez rozwój nowych kompetencji, a głęboka personalizacja – przy odpowiednim

projekcie – zwiększy eksplorację zamiast ją ograniczać. Możliwe, że ludzie wypracują nowe normy społeczne, pozwalające łatwo odróżnić syntetyczną empatię od relacji ludzkiej. Nauka powinna pozostać otwarta na takie wyniki. Z tego powodu pojęcie „Hybrydowej Ludzkości" powinno być hipotezą organizującą, a nie diagnozą z ustalonym finałem. Jego wartością jest łączenie zjawisk, które osobno wydają się kwestiami interfejsu, produktywności, edukacji, neurotechnologii czy prawa. Razem tworzą one pytanie o zmianę ludzkiego środowiska poznawczego.

Tak rozumiana hybrydyzacja ma również wymiar kulturowy. Pismo zmieniło strukturę pamięci i autorytetu, druk skalę rozpowszechniania wiedzy, media elektroniczne tempo społecznego doświadczenia, a internet sposób wyszukiwania informacji i publikacji. AI zmienia natomiast strukturę generowania odpowiedzi: pośrednik przestaje jedynie wskazywać cudzą treść, a zaczyna tworzyć nową reprezentację dopasowaną do użytkownika. Może to przeobrazić normy autorstwa. Coraz więcej rezultatów będzie współprodukowanych przez ludzi i modele. Autorstwo prawne, odpowiedzialność naukowa i kultura twórczości będą musiały odróżnić zwykłe użycie narzędzia od sytuacji, w której człowiek nie jest już w stanie zrekonstruować zasadniczej argumentacji powstałego tekstu. Podobnie jak przy innych formach rozszerzonego poznania, centralnym problemem pozostanie zdolność do odpowiedzialnego przyjęcia rezultatu jako własnego.

> Ostatecznie to nie benchmarki modeli, lecz kształt naszych instytucji zdecyduje o tym, czy AI stanie się lustrem powiększającym ludzki potencjał, czy klatką zoptymalizowaną pod dyktando algorytmicznej efektywności. Czy w świecie permanentnej augmentacji będziemy jeszcze potrafili odróżnić rozwój własnych kompetencji od sprawnego zarządzania naszą zależnością? Prawdziwą miarą mądrości Hybrydowej Ludzkości może okazać się nie to, jak daleko zdoła przesunąć granice biologii, lecz zdolność do świadomego powstrzymania się od wykonania tego, co technicznie możliwe.

Inspiracje

[1] **"The AI Instinct"** Rana Gujral

Rana Gujral (ur. 18 czerwca 1976 r.) to amerykański przedsiębiorca, dyrektor i inwestor specjalizujący się w kognitywnej sztucznej inteligencji (CSS) i rozpoznawaniu emocji w mowie. Najbardziej znany jest z funkcji dyrektora generalnego Behavioral Signals, firmy zajmującej się sztuczną inteligencją, która rozwija technologię głębokiego uczenia (DNA) do rozpoznawania emocji, intencji i sygnałów behawioralnych w danych głosowych i mowy. Wcześniej Gujral założył TiZE, firmę zajmującą się oprogramowaniem w chmurze, która została później przejęta przez Alchemy, i pełnił funkcje kierownicze oraz operacyjne w firmach technologicznych, takich jak Cricut i Logitech. Jako wybitny mówca i autor tekstów o rozwoju sztucznej inteligencji ogólnej i obliczeń kognitywnych, Gujral publikował obszerne artykuły w takich publikacjach jak Forbes i TechCrunch, a także wygłaszał przemówienia na konferencjach takich jak TEDx i World Government Summit.

Rana Gujral *(born June 18, 1976) is an American entrepreneur, executive, and investor specializing in cognitive artificial intelligence and speech emotion recognition. He is best known for his role as the Chief Executive Officer of Behavioral Signals, an enterprise artificial intelligence company that develops deep learning technology to recognize emotion, intent, and behavioral cues in voice and speech data. Gujral previously founded TiZE, a cloud software enterprise that was later acquired by Alchemy, and held leadership and operational roles at technology companies including Cricut and Logitech. A prominent speaker and writer on the trajectory of artificial general intelligence and cognitive computing, Gujral has contributed extensively to publications such as Forbes and TechCrunch and presented keynotes at venues like TEDx and the World Government Summit.*

Behavioral Signals

Literatura uzupełniająca

Książki

Książki autorów przywoływanych i literatura pokrewna:



[1] **"Design for a brain"**

Ross Ashby

ISBN: 9781614277569



[2] "Design for a brain; the origin of adaptive behavior"

Ross Ashby

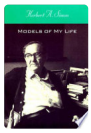
ISBN: 9781614277569



[3] "Organizations"

Herbert Simon

John Wiley & Sons | ISBN: 9780631186311



[4] "Models of my life"

Herbert Simon

MIT Press | ISBN: 9780262326582



[5] "Human Problem Solving"

Herbert Simon

Prentice Hall

[6] "Communication and persuasion"

Irving Janis

Artykuły naukowe

Prace autorów przywoływanych w tekście:

[1] "The challenge of gender justice"

Martha C. Nussbaum

2009 | Against Injustice | DOI: 10.1017/cbo9780511657443.006

Źródło

[2] "Design for a Brain"

William Ross Ashby

1952 | Medical Entomology and Zoology

Źródło

[3] "Requisite Variety and Its Implications for the Control of Complex Systems"

William Ross Ashby

1991 | Springer eBooks | DOI: 10.1007/978-1-4899-0718-9_28

Źródło

[4] **"Design for a Brain: The origin of adaptive behaviour"**

William Ross Ashby

2014 | Medical Entomology and Zoology

Źródło

[5] **"Design for a brain: The origin of adaptive behaviour (2nd ed. rev.)."**

William Ross Ashby

1960 | Chapman & Hall eBooks | DOI: 10.1037/11592-000

Źródło

[6] **"Papers presented at the International Conference on the Use of Micros in Fluid Engineering, London, England, June, 1983"**

Herbert Simon Stephens, Peter Shipley, Bhra

1983 | BHRA eBooks

Źródło

[7] **"Effects of Fear Arousal on Attitude Change: Recent Developments in Theory and Experimental Research"**

Irving L. Janis

1967 | Advances in experimental social psychology | DOI: 10.1016/s0065-2601(08)60344-5

Źródło

[8] **"Improving Adherence to Medical Recommendations: Prescriptive Hypotheses Derived from Recent Research in Social Psychology"**

I. Janis

2020 | Handbook of Psychology and Health (Volume IV) | DOI: 10.1201/9781003044307-5

Źródło

[9] **"Cognitive Complexity in International Decision Making"**

I. Janis, L. Mann

2019 | Psychology and Social Policy | DOI: 10.4324/9781315793030-3

Źródło

Artykuły naukowe autora:

[10] **"The Nature of Human Cognition and AI's Approach to Mimicking It"**

RANA GUJRAL, KAI-FU LEE

2026 | DOI: 10.1002/9781394392155.ch02

Źródło

[11] "AI in the Real World: Applications and Use Cases"

RANA GUJRAL, KAI-FU LEE

2026 | DOI: 10.1002/9781394392155.ch03

Źródło

[12] "The AI Instinct"

RANA GUJRAL, KAI-FU LEE

2026 | DOI: 10.1002/9781394392155.fmatter

Źródło

[13] "The Emergence of Artificial General Intelligence (AGI)"

RANA GUJRAL, KAI-FU LEE

2026 | DOI: 10.1002/9781394392155.ch04

Źródło

[14] "Introduction: Our Definition of “There” Is Wrong"

RANA GUJRAL, KAI-FU LEE

2026 | DOI: 10.1002/9781394392155.ch01

Źródło

Treść tworzy Fundacja Dobre Państwo.
Redaguje i publikuje APA ONE, autorski system redakcyjny Fundacji oparty na AI.
APA ONE Publishing System
Fundacja Dobre Państwo © 2026
Article ID: 2c0534ff-ecad-4b48-862e-00d8a5a4bc1b