

Sztuczna inteligencja i mrówki: rewolucja postprzemysłowa

John McClellan Marshall*

„Czasy się zmieniają”

Bob Dylan (1965), laureat Nagrody Nobla

Ludzka natura ma tendencję do refleksji i wspomina tego, co wydarzyło się w ich życiu osobistym. Z czasem przekłada się to na nostalgię za „tym, co było kiedyś”. Jest to prawdopodobnie największe złudzenie, na jakie podatne są istoty ludzkie, ponieważ ignoruje ono jedyną prawdziwą zasadę Wszechświata: zmianę. W rzeczywistości prawdopodobnie widzimy te rzeczy przez pryzmat tego kim jesteśmy¹. A zatem możemy nawet nie widzieć ich takimi, jakimi były „wtedy”, bo zależy to od naszych własnych doświadczeń życiowych. W rzeczywistości, jak zauważono: „Są tacy, którzy patrzą na rzeczy takimi, jakimi są i pytają dlaczego? Ja marzę o rzeczach, których nigdy nie było i pytam, dlaczego nie?”². Systemowe zmiany otaczającego nas świata jakie zaszły w ciągu ostatnich kilku lat, od ekonomii po COVID, stworzyły poczucie oderwania się od naszych zbiorowych „korzeni”. Na poziomie społecznym analogią byłoby wyobrażenie sobie destrukcyjnego wpływu na populację farmy mrówek, która została wywrócona do góry nogami, tak że mrówki muszą zacząć od nowa w tworzeniu swojego środowiska³. Przychodzi na myśl słowo „restrukturyzacja”, a społeczeństwo może być w trakcie wymyślania na nowo tego, co uważamy za byłe.

Dodatkowym czynnikiem jest przekonanie, że czas płynie bardzo szybko. W rzeczywistości, z perspektywy widza, czas płynie w tym samym tempie, co zawsze. Ważne jest to, że tempo zmian jest niezależne od pozornego tempa upływu czasu. Co więcej, paradygmat poznania 4E (Embodied, Embedded, Enacted lub Extended – wcielony w życie, osadzony w realiach, wprowadzony w życie lub rozszerzony) wzmacnia ten proces poprzez uwzględnienie aspektów środowiskowych w indywidualnej percepcji tego, co się wokół nas dzieje⁴. Może to prowadzić do poczucia pilności, które może zastąpić racjonalne podejmowanie decyzji w procesie „restrukturyzacji”.

W drugiej połowie XX w. świat zgodnie dążył do innowacji technologicznych w dziedzinach, które koncentrowały się na rozwiązaniach inżynierskich różnych problemów takich jak loty kosmiczne i nieprzywierające naczynia kuchenne. Dążenie to nie było ograniczone przez granic narodowych lub polityk, ale był napędzany częściowo przez ekonomię, a częściowo przez zimnowojenną politykę, ani nie towarzy-

szyły mu żadne znaczące innowacje społeczne. Jednocześnie rozwój technologii wpłynął na ludzkość w takich kwestiach, jak leczenie chorób, z których niektóre istniały od wieków, a niektóre, takie jak AIDS, były zupełnie nowe.

Jako całość, okres ten reprezentuje coś, co można nazwać rewolucją technologiczną/cyfrową. Był to do pewnego stopnia okres, w którym dążenie do celu było celem samym w sobie i nie dotyczyło długoterminowych wyników ani możliwych konsekwencji⁵. Pod pewnymi względami można to porównać do psa, który goni samochód, ponieważ pojawia się pytanie: „Co pies zrobi z samochodem, jeśli go złapie?” Do pewnego stopnia „pies” faktycznie złapał „samochód”, a rewolucja technologiczna/cyfrowa przekształciła się w rewolucję postindustrialną. Rzeczywistość XXI w. pokazuje ten dylemat dość wyraźnie w synergicznym społecznym zderzeniu innowacji technologicznych ze sztuką i naukami humanistycznymi.

Filozoficzny wpływ tej kolizji podsumowuje słowo „cyber-etyka”. Definiuje się ją jako relację między systemami etycznymi i prawnymi, które zostały opracowane, aby służyć ludzkości od czasów starożytnych do współczesności, wyrażonymi w naszych filozoficznych i etycznych systemach myślenia, w przeciwieństwie do zdolności technologii komputerowej do działania poza tymi konwencjami niemal bez ograniczeń⁶. Zasadniczo problem cyber-etyczny został przyspieszony przez tempo innowacji technologicznych, którym nie towarzyszą innowacyjne zmiany w strukturze społecznej. W rezultacie wpływ ten nie ogranicza się jedynie do upływu czasu, ale wydaje się również rozszerzać swój zakres społeczny.

Na poziomie społecznym, ekspansja rejestrowania informacji, czasami określana jako „Big Data”, podobnie doprowadziła do subtelnej zmiany w codziennej działalności

* Autor jest Sędzią w stanie spoczynku, Czternastego Sądu Okręgowego stanu Teksas. Honorowym Profesorem UMCS w Lublinie, Członkiem Międzynarodowej Akademii Astronautyki; ORCID ID: 0000-0003-4504-144X.

¹ *Anais Nin*, *Seduction of the Minotaur*, t. 5 *Cities of the Interior* (1959). Pewien profesor powiedział kiedyś: „Uczelnia nie jest tym, czym była kiedyś... i nigdy nie była”.

² *George Bernard Shaw*, *Back to Methuselah*, wypowiedziane przez Węza, później sparafrazowane przez nieżyjącego już Roberta F. Kennedy’ego.

³ Analogia do „mrówek” została zasugerowana przez profesora Nishiki Sugawara-Beda w rozmowie z autorem.

⁴ Zob. *James Carney*, „Thinking avant la lettre: A review of 4E Cognition”, *Evolutionary Studies in Imaginative Culture* (2020), doi:10.2661.3/esci/4.1.172.

⁵ Zob. *Edward Teller*, „Można powiedzieć, że przewidywania są niebezpieczne, szczególnie w odniesieniu do przyszłości. Jeśli niebezpieczeństwo związane z przewidywaniem nie jest ponoszone, nie wynikają z tego żadne konsekwencje, a zasada niepewności nie jest naruszana”, w: *Conversations on the Dark Secrets of Physics* (2013), s. 235.

⁶ Koncepcja ta została pierwotnie wyartykułowana przez autora w artykule „The Terminator Missed a Chip!: Cyberethics”, zaprezentowanym na Międzynarodowym Kongresie Astronautycznym w Oslo w 1995 r. i pierwotnie opublikowanym, za zgodą American Institute of Aeronautics and Astronautics, Inc. Udostępnione IAF/AIAA do publikacji we wszystkich formach. Następstwem tego jest zdolność technologii do wprowadzania zmian w tych konwencjach bez względu na wkład człowieka w społeczne „domyślne” działanie maszyn.

wielu zawodów⁷. Rzeczywiście, zauważono, że „przeszliśmy od epoki bogatej w znaczenia, ale ubogiej w dane, do epoki bogatej w dane, ale ubogiej w znaczenia... [Jest to rewolucja epistemologiczna równie fundamentalna jak rewolucja kopernikańska]”⁸.

W praktyce definicja „miejsca pracy”, na którą pozwala dostępność danych na szeroką skalę, doprowadziła do wzrostu zatrudnienia zdalnego „z domu” w wielu dziedzinach. Co więcej, sprzyja to tworzeniu „cyfrowych nomadów”, którzy nie chodzą do biura jako rutynowa część swojej pracy, ale są w stanie wykonywać zadania „normalnie”. Jednym z rezultatów takiego stanu rzeczy jest rosnąca separacja pracowników od siebie nawzajem, a tym samym zmniejszenie wartości interakcji międzyludzkich w procesie „wykonywania pracy”⁹.

Wady tej hybrydyzacji są najbardziej widoczne w środowisku akademickim. Oprócz potencjalnego naruszenia integralności dokumentacji akademickiej studentów, a w konsekwencji fałszowania życiorysów, gdy student wchodzi w świat pracy, pojawia się kwestia jakości samej edukacji. Podczas gdy naukowcy często pracują w wysoce zindywidualizowanej dziedzinie, proces edukacji historycznie obejmował interakcję między profesorami i studentami oraz między samymi studentami, czasami określaną jako „nauka kohortowa”. Wprowadzenie oprogramowania konferencyjnego, które pozwala na „zdalne” nauczanie, z definicji odsuwa studenta od profesora, choćby poprzez możliwość „wyciszenia” rozmowy lub wyłączenia wideo i oddalenia się od klasy. Pozwala to również na fałszowanie odpowiedzi ucznia na pytania profesora poprzez dostęp do zasobów, takich jak smartfony znajdujące się poza zasięgiem kamery. Chociaż szczegółowe badania długoterminowego wpływu tego typu środowiska nauczania jeszcze się nie pojawiły, istnieją już pewne dowody na to, że jakość produktu edukacyjnego ulega pogorszeniu, a jakość edukacji może zejść na dalszy plan.

Niestety, niektóre z przeprowadzonych badań mają tendencję do patrzenia na „efektywność kosztową” jako główny czynnik przy podejmowaniu decyzji o „zdalnym nauczaniu”¹⁰. Problem edukacji zdalnej może być potęgowany przez brak innowacji w tym formacie. W rzeczywistości technologia wydaje się być po prostu wykorzystywana do odtworzenia sali lekcyjnej bez kontaktu z ludźmi¹¹. Problem ten jest dodatkowo potęgowany przez brak kontaktu między wykładowcami a studentami, co sprawia, że studenci mają niewielki „wpływ” na studiowany materiał. Taka dysocjacja prawdopodobnie ilustruje, dla niektórych, różnicę między „uczeniem się z książek” a „edukacją”, zwłaszcza gdy uczeń przenosi się do „prawdziwego świata”. Innymi słowy, „wirtualna” rzeczywistość, niezależnie od tego, jak jest przedstawiana, po prostu nie jest „prawdziwa”, a zatem jest nieistotna.

Wśród najbardziej znanych przykładów są społeczności zaangażowane w proces sądowy i praktykę medyczną.

Do niedawna tempo innowacji było najbardziej widoczne w tych dziedzinach, które są predysponowane do tworzenia inżynierskich rozwiązań ludzkich problemów. Technologia cyfrowa zmieniła jednak kierunek takich innowacji w sposób, który rozprzestrzenił wpływ takiej kreatywności nieco szerzej¹². Na przykład, kronika rozwoju technologii na świecie została zarejestrowana najpierw na taśmie perforowanej i magnetycznej, potem na dyskach laserowych, a następnie na cyfrowych dyskach twardych. W rezultacie granica między zwykłą technologią a cyfrową rzeczywistością, w której się narodziła, stała się zamazana, jeśli nie nieistniejąca, jak w przypadku zniknięcia całych katalogów muzycznych na taśmach ośmiościeżkowych.

U swoich podstaw rewolucja postindustrialna ma potencjalnie śmiertelną, ale systemową wadę: jest nią sama baza cyfrowa. Przykładem tego w dziedzinie medycyny było opracowanie tomografii komputerowej (CT) lub techniki rezonansu magnetycznego (MRI). System ten wykonuje bardzo szczegółowe skany ludzkiego ciała, które następnie przechowuje w formie cyfrowej, umożliwiając ich komputerową analizę. W ten sposób maszyna zwiększa umiejętności diagnostyczne lekarza. Samo ulepszenie zmniejsza jednak element ludzki w procesie, co prawdopodobnie redukuje lekarza do przedłużenia maszyny. Sama obecność bazy cyfrowej w procesie CT pozwala jednak na możliwość „włamania”, które może zmienić wynik diagnostyczny ze szkodą dla pacjenta.

Kolejnym przykładem zakresu zmniejszenia czynnika ludzkiego w odpowiedzi na postęp technologiczny w kontekście instytucjonalnym jest funkcjonowanie niektórych funkcji rządowych. Przykładowo, w procesach związanych z funkcją sądowniczą, centralnym elementem jest dostępność wiarygodnych dowodów. Na wstępie należy pamiętać, że sąd jest instytucją z natury ludzką, zaprojektowaną w celu służenia społeczeństwu poprzez rozwiązywanie sporów za osoby, które nie są w stanie rozwiązać ich samodzielnie. Najnowsze badania ujawniły, że sąd ma swoje korzenie w najwcześniejszych, być może prehistorycznych, fazach obecności człowieka na naszej planecie. Rzeczywiście, koncepcja „wiarygodności” jest unikalnym ludzkim konstruktem, który leży u podstaw większości współczesnego społeczeństwa. Jednak w cyberświecie

⁷ Bryan Lufkin, „Is the great digital-nomad workforce actually coming?”, 15.6.2021 r. (BBC), <https://www.bbc.com/worklife/article/20210615-is-the-great-digital-nomad-workforce-actually-coming?ocid=ww.social.link.email>.

⁸ Boorstin, Daniel J., *Cleopatra's Nose: Essays on the Unexpected*, Nowy Jork: Random House, 1994.

⁹ Zob. Ray Bradbury, „The Pedestrian”, *The Reporter*, sierpień 1951.

¹⁰ U.S. Department of Education, Office of Educational Technology, „Understanding the Implications of Online Learning for Educational Productivity”, styczeń 2012.

¹¹ Justin Reich, *Failure to Disrupt: Why Technology Alone Can't Transform Education*, Harvard University Press, 2020.

¹² Zob. Ben Williams, „The Fourth Industrial Revolution”, 69 *Knights Templar* 44, wrzesień 2023, aby zapoznać się z problemami związanymi z używaniem oprogramowania takiego jak ChatGPT do wyciągania faktycznych wniosków w porównaniu do gromadzenia informacji.

kryteria, według których oceniana jest sama wiarygodność, mogą zostać przeddefiniowane.

Historycznie rzecz biorąc, standardem dla wiarygodności dowodów w sądzie były zeznania świadków obecnych w otwartym dla publiczności sądzie, wspomagane zdjęciami, dokumentami i przedmiotami, które przyczyniają się do przedstawienia faktów sędziemu lub ławie przysięgłych. Takie dowody odzwierciedlają również wrażenia zmysłowe świadka. Czasami świadek zmarł przed rozprawą, więc zazwyczaj dowód był prowadzony do procesu sądowego w formie pisemnego zeznania. Dzięki technologii cyfrowej osoba ta może zostać wskrzeszona i zeznawać jako projekcja hologramowa. „Fotografie” jako takie już nie istnieją. Są one reprodukowane jako „obrazy” tworzone cyfrowo przez to, co widziała „kamera”. Chociaż taki obraz może być dokładniejszy niż wzrok przefiltrowany przez ludzki mózg, jest znacznie bardziej podatny na manipulacje, wprowadzając w ten sposób już na wstępie element nieufności co do autentyczności obrazu. Podobna manipulacja w przypadku nagrań głosowych lub dokumentów potencjalnie osłabia wiarygodność dowodów, a tym samym samego procesu. Dla prawnika, zwykle wykształconego w dziedzinie nauk humanistycznych, potrzeba pewnej edukacji w zakresie możliwości cyfrowych jest niezbędna do przygotowania odpowiedniego badania dowodów lub „świadków” w kontekście procesu.

Przykładem może być niedawne pojawienie się „chatbotów”, takich jak ChatGPT, które stały się modne w codziennej praktyce prawniczej. Ta forma „asystenta badawczego” jest zarówno szybka, jak i wygodna pod względem możliwości przeszukiwania baz danych lub tworzenia dokumentów, a niektórzy prawnicy postrzegają ją jako transformację¹³. Problemem związanym z wykorzystaniem chatbotów jako „asystentów badawczych” jest ich zdolność do „halucynowania” faktów, a nawet „tworzenia” innych „książek” wewnątrz programowania, bez wiedzy użytkownika. Rezultatem takiej „halucynacji” mogą być błędne wnioski, jeśli to właśnie jest celem badania¹⁴. Ta debata, bardzo aktualna w środowisku prawniczym, już teraz wprowadza ostrożność do dyskusji na temat autentyczności tego, co odkrywa chatbot¹⁵. Ujmując rzecz abstrakcyjnie, „w rzeczywistości chodzi o przejście od wiedzy jako próby zrozumienia rzeczywistości do danych jako opisu rzeczywistości”¹⁶. W tym zakresie chatbot najwyraźniej nie przekroczył jeszcze granicy między gromadzeniem danych a „wiedzą” o tym, czym jest rzeczywistość.

Być może jeszcze bardziej niepokojący przykład możliwości technologicznej manipulacji faktami miał miejsce kilka lat temu podczas procesu sądowego. Było to związane z transferem środków z banku w Stanach Zjednoczonych na konto zagraniczne wbrew tymczasowemu nakazowi sądowemu. Adwokaci powoda, który domagał się rozstrzygnięcia, nie byli „obeznani z technologiami cyfrowymi”, więc szukali porady, jak odzyskać środki. Doradzono im, że jeśli uzyskają kod

numeryczny, za pomocą którego środki zostały wysłane i po prostu je odwrócić, pozostanie jeden zestaw „jedynek i zer”. Przy odrobinie szczęścia wszystko, co byłoby konieczne, to odwrócenie tych ostatnich cyfr, aby odzyskać pieniądze z zagranicy na amerykańskie konto bankowe. Pomimo początkowego sceptycyzmu ze strony prawników, instrukcje zostały wykonane, a pieniądze zostały odzyskane w ciągu 24 godzin od początkowego odkrycia transakcji. Fakt, że spowodowało to mały incydent dyplomatyczny, nie ma znaczenia. To, co jest istotne w tym przykładzie, to fakt, że bieg sprawiedliwości był w rzeczywistości kontrolowany przez technologię cyfrową tamtych czasów. Nie ma wątpliwości, że stwarza to obawy co do wiarygodności faktów, które są następnie przedstawiane ławie przysięgłych w pełnym procesie.

W konkretnym obszarze prawa przecięcie między technologią a naukami humanistycznymi jest dość wyraźne: chodzi o sztukę. Termin „sztuka” obejmowałby nie tylko sztuki plastyczne, ale także samą własność intelektualną, która musi być rozpatrywana w kontekście cyber-etycznym, niezależnie od tego, czy ma być stosowana na poziomie indywidualnym, czy w odniesieniu do społeczeństwa jako całości. Rzeczywiście, po kilku tysiącach systemu wartości, który sprzyjał rozwojowi materialnemu, intelektualnemu i duchowemu, pojawia się pytanie, w jakim stopniu społeczeństwo powinno odrzucić ten system na rzecz „cyberefektywności”, wiedząc, że jest to tak głęboko sprzeczne z intuicją?

W XXI w. nie ma wątpliwości, że napędzana technologią rewolucja postindustrialna i towarzyszące jej zakłócenia społeczne będą kontynuowane, prawdopodobnie z coraz większą prędkością. Powstanie więcej wynalazków, które do działania będą wymagały nowego kodowania. To, że kodowanie będzie wymagało nowych form technologii do jego przechowywania i realizacji jego celów, jest oczywiste. Połączy się to z kwestiami instytucjonalnymi, społecznymi, własności intelektualnej i praw autorskich, które już są w ruchu. W efekcie ten trend, który niesie ze sobą stopniowe zmniejszanie się udziału człowieka w codziennych czynnościach i zawodach, sprawia, że problem cybernetyczny staje się wyraźny i niepokojący. Rzeczywiście, wyścig może toczyć się między nami, ludźmi, a mrówkami, o to, które z nich jako pierwsze przywróci jakąkolwiek część tego, co było, lub po prostu pójdzie naprzód w kierunku tego, co dopiero może być.

¹³ Brenda Derouen, „Przyszłość prawa? Jak chat GPT zmienia moją praktykę prawniczą”, 86 Tex.BarJ. 542 (wrzesień 2023).

¹⁴ Zob. *Mata przeciwko Avaianca*, sprawa nr 1:22-cv-01461 (Sąd Okręgowy Stanów Zjednoczonych dla Południowego Okręgu Nowego Jorku), czerwiec 2023 r., cytowany w Law 360, 23.6.2023 r., jako przykład konsekwencji polegania na chatbotach, które mogą „halucynować”.

¹⁵ John G. Browning, „Nowe i ulepszone? Ryzyko korzystania z ChatGPT”, 86 Tex.BarJ. 544 (wrzesień 2023).

¹⁶ Emerytowany Profesor Moihai Nadin, komentarz do sędziego Marshalla, lipiec 2023 r.