Niewidzialna sieć kontroli

W ciągu ostatnich trzech dekad byliśmy świadkami systematycznego „ogołacania człowieka z prywatności przez urządzenia IT”[1], jak trafnie ujmują to eksperci. Dzisiaj nasze smartfony zawierają kompleksowe profile naszego życia: od prywatnych zdjęć i notatek, przez badania lekarskie aż po transakcje finansowe. Prawdziwe zagrożenie nie tkwi jednak jedynie w gromadzeniu danych, ale w sposobie ich wykorzystania przez systemy sztucznej inteligencji do oceniania, kategoryzowania i kontrolowania obywateli.

Karolina Wilamowska

adwokatka, aplikantka rzecznikowska, mediatorka Centrum Mediacji przy Krajowej Izbie Gospodarczej, w którym kieruje zespołem Nowych Technologii w mediacji, doktorantka Uczelni Łazarskiego, trenerka, wykładowczyni, mentorka Fundacji Women in Law, Partnership Director w Singularity University Chapter Cracow. Członkini Sekcji Aktualne Wyzwania Sztucznej Inteligencji Polskiego Towarzystwa Informatycznego, ekspert Grupy Roboczej ds. Sztucznej Inteligencji przy Ministrze Cyfryzacji.

1 T. Kulisiewicz, J. Grabowski, *Nie dajmy się skopać Pegazowi*, „Domena” 1/2022, s. 18.

Współczesny świat cyfrowy istotnie zmienia zasady relacji między jednostką a społeczeństwem. Powstaje nowy ekosystem technologiczny, w którym systemy *social scoring* łączą się z technikami manipulacji podprogowej, tworząc niewidzialną sieć kontroli społecznej. To zjawisko istotnie wykracza poza tradycyjne formy nadzoru, wprowadzając elementy proaktywnej inżynierii społecznej operującej na poziomie podświadomości.

***Social scoring* to system algorytmicznej oceny obywateli lub podmiotów gospodarczych na podstawie analizy ich zachowań w różnych sferach życia społecznego, ekonomicznego i prywatnego. System ten wykorzystuje rozległe bazy danych zawierające informacje o historii kredytowej, aktywności internetowej, relacjach społecznych, preferowanych zakupach, a nawet o przestrzeganiu przepisów ruchu drogowego. Na podstawie tych danych algorytmy sztucznej inteligencji generują punktową ocenę „wiarygodności społecznej", która następnie determinuje dostęp do szeregu usług – od kredytów bankowych, przez możliwość podróżowania aż po przyjęcie do prestiżowych uczelni.**

Techniki podprogowe natomiast oddziałują na ludzką podświadomość poprzez bodźce, które są prezentowane poniżej progu świadomej percepcji. W kontekście cyfrowym mogą to być obrazy wyświetlane przez czas krótszy niż 0,04 sekundy[2], niewidoczne elementy interfejsu czy subtelne manipulacje architekturą wyboru. Współczesne systemy AI potrafią wykorzystywać te techniki w sposób wysoce wyrafinowany, korzystając z mechanizmów „Systemu 1" myślenia – szybkich, automatycznych procesów mentalnych, które są szczególnie podatne na wpływy zewnętrzne.

Chiński wzór

Chiński *System Zaufania Społecznego* od lat budzi emocje i bywa opisywany jako futurystyczna maszyna do inżynierii społecznej, która miała do 2020 r. objąć 1,3 mld obywateli. Badania i raporty z ostatnich lat pokazują jednak, że system nigdy nie przybrał formy scentralizowanej oceny wszystkich obywateli, raczej sieci powiązanych baz danych, rejestrów i mechanizmów monitorowania, które rozwijane były stopniowo w różnych sektorach i regionach.

Do 2020 r. system obejmował informacje na temat około 103 mln osób oraz ponad 20 mln firm. Nie powstała jednolita, punktowa ocena moralności obywateli – zamiast tego nacisk położono przede wszystkim na egzekwowanie zobowiązań finansowych, wykonywanie orzeczeń sądowych i nadzór nad przedsiębiorstwami. To tłumaczy głośne doniesienia o ograniczeniach w podróżach. W 2019 r. władze podały, że odrzucono ponad 26 mln rezerwacji biletów lotniczych i niemal 6 mln biletów na szybkie koleje, głównie wobec osób, które znalazły się na tzw. „czarnych listach" dłużników. Powszechnie powielany obraz systemu, który nagradza segregowanie śmieci, a karze za posiadanie znajomych o niskim statusie, okazuje się zatem raczej mitem niż rzeczywistością.

Mimo to infrastruktura kontrolna w Chinach jest imponująca. *Projekt Skynet*, czyli państwowa sieć monitoringu wizyjnego, od połowy lat 2000 rozwija się w błyskawicznym tempie. W 2013 r. w kraju działało około 20 mln kamer, a 10 lat później – już 700 mln. Kamery z funkcją rozpoznawania twarzy, zintegrowane z systemami analizy danych, stały się podstawą funkcjonowania nowoczesnego państwa nadzoru, w którym niemal każda przestrzeń publiczna podlega rejestracji.

Rola sektora prywatnego w całym projekcie również uległa zmianie. W początkowej fazie wiele emocji budziły eksperymenty prowadzone przez gigantów technologicznych, takich jak Tencent (twórca WeChat) czy Alibaba (poprzez system Sesame Credit powiązany z Alipay). Firmy te rzeczywiście testowały rozwiązania polegające na ocenianiu użytkowników na podstawie danych zakupowych, finansowych czy nawet towarzyskich. Jednak w kolejnych latach państwo przejęło kontrolę nad tego rodzaju inicjatywami. Pilotaże zostały wygaszone, a władze skupiły się na rozwoju własnych rejestrów i na włączeniu prywatnych danych do systemu administracyjnego w ściślejszy sposób.

W efekcie współczesny chiński *Social Credit System* nie jest jednolitym, narodowym algorytmem przyznającym obywatelom punktowe oceny za moralne zachowanie. Bardziej trafnym określeniem jest raczej „ekosystem regulacyjny", w którym państwo integruje dane z sądów, instytucji finansowych, organów administracji oraz systemów nadzoru, a następnie wykorzystuje je do egzekwowania prawa i wzmacniania kontroli nad podmiotami gospodarczymi. Obywatele doświadczają jego działania najczęściej nie poprzez niewidzialny ranking moralny, lecz poprzez konkretne ograniczenia w dostępie do usług, jeśli nie wywiązują się z zobowiązań wobec państwa.

Zachodnie postrzeganie chińskiego systemu *Social Credit* jako totalnej dystopii z czterema kategoriami obywateli

2 Według Wikipedii, w percepcji wzrokowej bodźce trwające krótko – krócej niż 0,04 s – są klasyfikowane jako podprogowe, czyli niewidoczne dla świadomości obserwatora [dostęp: 15.08.2025].

i automatycznym odbieraniem praw nie znajduje więc pełnego potwierdzenia w faktach. Nadal jednak niepokoi.

Skala gromadzonych danych, połączenie technologii nadzoru z administracyjnymi bazami państwa i stopniowe ograniczanie prywatności jednostki sprawiają, że system ten pozostaje jednym z najbardziej ambitnych eksperymentów kontroli społecznej na świecie. To laboratorium przyszłości, w którym granice między technologią, prawem i polityką zostały zatarte.

Globalna ekspansja

Pewne elementy chińskiego systemu – przede wszystkim mechanizmy oceny, klasyfikacji i profilowania obywateli – zyskują coraz większe znaczenie także w liberalnych demokracjach. Systemy te, choć mniej widoczne i otwarcie definiowane jako *social scoring*, coraz częściej wykorzystują behawioralne i cyfrowe dane, wpływając na dostęp do usług, kredytów i ubezpieczeń.

W USA dominującym mechanizmem kontroli kredytowej pozostaje system FICO. Choć formalnie bazuje na tradycyjnych danych finansowych – takich jak historia spłat – instytucje finansowe coraz częściej eksperymentują z danymi alternatywnymi. Są to informacje obejmujące historię płatności za media, wzorce zakupowe, a nawet sygnały z mediów społecznościowych czy dane lokalizacyjne. Choć FICO oficjalnie ich nie integruje, raporty wskazują, że ten trend rośnie, zwłaszcza dzięki fintechom i startupom starającym się wyrównać dostęp do usług kredytowych dla osób z marną historią kredytową[3].

Już w 2018 r. magazyn „Wired" opisał badania pokazujące, że elementy takie jak typ urządzenia (Android vs iPhone), domena e-maila czy sposób dotarcia na stronę (np. przez porównywarkę cen) mogą być równie predykcyjne jak formalny *scoring* kredytowy[4]. To dowodzi, że granica między oceniającym systemem kredytowym a formą cyfrowego *scoringu* staje się coraz płynniejsza.

Podobny kierunek obserwujemy w wielu krajach demokratycznych, gdzie firmy ubezpieczeniowe czy instytucje finansowe zaczynają korzystać z niekonwencjonalnych danych: opłat za media, zakupów online, aktywności cyfrowej i mobilnej. Raporty z 2025 r. wskazują na coraz częstsze wykorzystywanie takich danych przy jednoczesnym wzroście zainteresowania regulacjami chroniącymi prywatność klientów[5]. Działania te przyczyniają się do wprowadzania zasad bardziej odpowiedzialnego modelowania ryzyka i ochrony konsumentów.

W niektórych państwach, takich jak Chile, Wielka Brytania czy Kanada, elementy profilowania pojawiają się raczej w kontekście ubezpieczeń, kredytów czy zatrudnienia i bazują głównie na danych zawodowych lub finansowych, często ograniczonych przez regulacje związane z ochroną danych i równością. Polska – podobnie jak wiele demokratycznych państw – ma rozbudowane rejestry publiczne (np. skarbowe, CEIDG), które mogą być wykorzystane do analiz i podejmowania decyzji administracyjnych. Jednak ich zastosowanie pozostaje w większości przypadków transparentne, a zakres wykorzystania danych – ograniczony przez przepisy o ochronie danych osobowych.

W krajach demokratycznych obserwujemy jednak stopniowe wprowadzanie np. *scoringu* kredytowego wzbogaconego o dane alternatywne, elektroniczne profilowanie migracyjne czy algorytmiczne decyzje w ubezpieczeniach i kredytach. Choć działają one w znacznie bardziej ograniczonym i regulowanym otoczeniu, ich rozwój niesie ze sobą poważne wyzwania, zwłaszcza w kontekście prywatności, przejrzystości i potencjalnej dyskryminacji. W przeciwieństwie do Chin, demokracje umiejętnie wdrażają takie technologie, ale nadal muszą chronić obywatela w rosnącym świecie danych.

Manipulacje w sieci

Równolegle z rozwojem systemów oceny społecznej obserwujemy proliferację technik manipulacyjnych określanych jako *dark patterns* – ciemne wzorce projektowe, które wykorzystują ludzkie słabości psychologiczne do skłaniania użytkowników do niechcianych działań.

Badania[6] Urzędu Ochrony Konkurencji i Konsumentów (UOKiK) z 2024 r. ujawniły, że aż 75,7% zbadanych stron internetowych stosowało co najmniej jeden *dark pattern*, a 66,8% wykorzystywało dwa lub więcej. W Polsce sytuacja okazała się nieco lepsza, ale nadal niepokojąca – połowa przeanalizowanych stron stosowała niedozwolone praktyki manipulacyjne[7].

Najczęściej obserwowane warianty *dark patterns* (wg badania UOKiK i raportów europejskich) to:

3 World Bank, *The Use of Alternative Data in Credit Risk Assessment: Opportunities, Risks, and Challenges*, 2025, http://hdl.handle.net/10986/42970 [dostęp: 15.08.2025].

4 https://www.wired.com/story/your-smartphone-could-decide-whether-youll-get-a-loan/ [dostęp: 15.08.2025].

5 SME Finance Working Group, *Alternative Data for Credit Scoring*, https://www.afi-global.org/wp-content/uploads/2025/02/Alternative-Data-for-Credit-Scoring.pdf [dostęp: 15.08.2025].

6 UOKiK wziął udział w corocznym przeglądzie sprzedażowych stron internetowych i aplikacji mobilnych organizowanym przez Międzynarodową Sieć Ochrony Konsumentów (ICPEN).

7 https://uokik.gov.pl/wielkie-wymiatanie-zlych-praktyk-w-e-commerce [dostęp: 15.08.2025].

- preselekcja – automatyczne zaznaczanie opcji korzystnych dla firmy (np. dodatkowe usługi, zgody marketingowe), co wymaga ich świadomego odznaczenia przez użytkownika;
- społeczny dowód słuszności (*social proof*) – informacje sugerujące, że oferta jest popularna lub ograniczona, często o charakterze fikcyjnym, by skłonić do szybkiej decyzji;
- *fake countdowns* – fałszywe wskaźniki ograniczonego czasu promocji, które mają wywołać poczucie pilności.

Serwis Booking.com systematycznie wykorzystywał *dark patterns* i został w 2020 r. ukarany za stosowanie komunikatów typu „pozostał tylko 1 pokój" czy „10 osób ogląda teraz ten hotel", które nie bazowały na rzeczywistych danych, lecz były sztucznie generowane przez system[8].

Techniki manipulacji podprogowej

Jeszcze subtelniejszą formą cyfrowej manipulacji stają się dziś techniki podprogowe wspierane przez sztuczną inteligencję. O ile klasyczne przykłady reklamy podprogowej z połowy XX w. – jak głośna, choć sfabrykowana, historia o wstawianiu pojedynczych klatek z coca-colą czy popcornem do filmów kinowych – pozostają raczej w sferze mitu, o tyle współczesne systemy AI posługują się narzędziami o znacznie większej precyzji i skuteczności. W odróżnieniu od masowych przekazów dawnych kampanii, dzisiejsze algorytmy potrafią dostosowywać strategię oddziaływania do indywidualnego profilu psychologicznego użytkownika.

Mechanizm ten szczególnie dobrze widać w projektowaniu architektury wyboru. Interfejsy aplikacji i serwisów internetowych coraz częściej nie są neutralnym „oknem" do usług, lecz starannie skrojonym środowiskiem, które prowadzi użytkownika w stronę decyzji korzystnych dla platformy. Sztuczna inteligencja wie, które kolory zwiększają skłonność do kliknięcia, jak rozmieścić przyciski, by wywołać reakcję impulsywną, i kiedy zaprezentować opcję płatną jako „domyślną". Co więcej, algorytmy potrafią rozpoznać momenty obniżonej samokontroli – późne godziny wieczorne, chwile zmęczenia czy stresu – i właśnie wtedy podsuwają reklamy, oferty lub sugestie, które mają największą szansę na akceptację.

Szczególną rolę odgrywa tu personalizacja. Dane o wcześniejszych zachowaniach, preferencjach czy nawet emocjonalnych reakcjach pozwalają systemom tworzyć niemal intymny profil psychologiczny jednostki. To nie przypadek, że reklama kredytu „na już" trafia do osób w kryzysie finansowym, a powiadomienia z aplikacji społecznościowych pojawiają się akurat w chwilach największego znużenia. AI nie tylko analizuje, co użytkownik robi, lecz również kiedy jest najbardziej podatny na sugestię.

Takie techniki są dziś wszechobecne. W handlu elektronicznym stosuje się dynamiczne zmiany cen i dostępności produktów, aby wywołać presję emocjonalną i poczucie pilności. W mediach społecznościowych rekomendacyjne algorytmy subtelnie przesuwają granice akceptowalnych treści, stopniowo wciągając użytkowników w coraz bardziej radykalne narracje. W marketingu natomiast algorytmy potrafią identyfikować grupy osób szczególnie podatnych na przekaz – jak zadłużeni konsumenci – i kierować do nich agresywne oferty finansowe. Również w aplikacjach mobilnych powiadomienia *push* nie są już przypadkowe: system analizuje rytm dnia użytkownika i wysyła sygnały w momentach, gdy szansa na interakcję jest najwyższa.

W ten sposób manipulacja podprogowa w erze AI nie polega już na pojedynczym „mrugnięciu" obrazem ukrytym w filmie, lecz na ciągłym, precyzyjnie dostosowanym oddziaływaniu, niemal niezauważalnym w codziennej aktywności cyfrowej. To właśnie ta nieuchwytność sprawia, że granica między perswazją a manipulacją coraz bardziej się zaciera, a autonomia decyzji użytkowników bywa poddawana nieustannej, algorytmicznej próbie.

AI Act na ratunek

Odpowiedzią Unii Europejskiej na narastające ryzyka związane z wykorzystaniem sztucznej inteligencji jest przełomowe Rozporządzenie w sprawie AI (AI Act)[9], które weszło w życie 1 sierpnia 2024 r. Od lutego 2025 r. zaczęły obowiązywać najważniejsze przepisy, w tym katalog praktyk uznanych za zakazane. Regulacja ta jest pierwszym na świecie kompleksowym aktem prawnym dotyczącym sztucznej inteligencji i obowiązującym w całej UE.

Kategorycznie zakazuje stosowania systemów AI w celach *social scoringu*. Jak wskazuje art. 5 ust. 1 lit. c, zakazane są wszelkie praktyki polegające na ocenianiu lub klasyfikacji osób fizycznych na podstawie ich zachowań społecznych albo cech osobistych, jeśli prowadzą one do krzywdzącego lub nieproporcjonalnego traktowania

8 https://events.zoom.us/e/lobby/qy3TbReJRbma8vyyONwv-A/error?webUserWithSessionManage=false [dostęp: 15.08.2025].

9 Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2024/1689 z dnia 13 czerwca 2024 r. w sprawie ustanowienia zharmonizowanych przepisów dotyczących sztucznej inteligencji oraz zmiany rozporządzeń (WE) nr 300/2008, (UE) nr 167/2013, (UE) nr 168/2013, (UE) 2018/858, (UE) 2018/1139 i (UE) 2019/2144 oraz dyrektyw 2014/90/UE, (UE) 2016/797 i (UE) 2020/1828 (akt w sprawie sztucznej inteligencji) (Tekst mający znaczenie dla EOG).

w kontekstach innych niż te, w których dane zostały zebrane. Oznacza to, że choć *scoring* kredytowy czy oceny antyfraudowe pozostają dozwolone (przy zachowaniu określonych prawem warunków), to systemy przypominające chiński model „oceny społecznej" są w Unii Europejskiej niedopuszczalne[10].

Równocześnie zakazane zostało wykorzystywanie technik podprogowych oraz manipulacyjnych, czyli takich, które działają poza świadomością użytkownika i w istotny sposób zniekształcają jego zdolność do podejmowania autonomicznych decyzji. Kluczowe jest tutaj nie tyle udowodnienie intencji manipulacji, ile sam skutek – wystarczy, że istnieje wysokie prawdopodobieństwo, iż praktyka doprowadzi do poważnej szkody psychologicznej, ekonomicznej czy zdrowotnej.

Szczególną ochroną objęto także grupy wrażliwe. Zgodnie z art. 5 ust. 1 lit. b, zakazane jest wykorzystywanie słabości osób wynikających z wieku, niepełnosprawności czy trudnej sytuacji społeczno-ekonomicznej. Ustawodawca unijny wprost wskazuje, że dzieci, osoby starsze oraz osoby w kryzysie społecznym czy ekonomicznym są szczególnie narażone na tego rodzaju wpływy.

Kolejną kategorią praktyk zakazanych jest zdalna identyfikacja biometryczna w czasie rzeczywistym w przestrzeni publicznej. AI Act dopuszcza jej użycie jedynie w ściśle określonych i wyjątkowych sytuacjach, takich jak ściganie najpoważniejszych przestępstw czy poszukiwanie osób zaginionych. Poza tymi przypadkami stosowanie takiej technologii jest uznawane za niedopuszczalne naruszenie praw podstawowych.

Nowe przepisy wprowadzają także precyzyjną definicję „znaczącej szkody", rozumianej jako „wystarczająco istotny niekorzystny wpływ na interesy fizyczne, psychologiczne, zdrowotne lub finansowe". Ważne jest, że szkoda nie musi wystąpić natychmiast – może się kumulować w czasie, a intencja jej wyrządzenia nie jest konieczna dla stwierdzenia naruszenia.

Egzekwowanie przepisów ma wspierać surowy reżim sankcyjny. Naruszenia zakazów przewidzianych w art. 5 mogą skutkować karami finansowymi sięgającymi do 35 mln euro lub 7% globalnego rocznego obrotu przedsiębiorstwa, w zależności od tego, która kwota jest wyższa. Odpowiedzialność obejmuje zarówno osoby prawne, jak i fizyczne. W ten sposób AI Act nie tylko ustanawia standard ochrony praw podstawowych przed nadużyciami sztucznej inteligencji, lecz także wyznacza globalny punkt odniesienia. To pierwsza regulacja, która wprost zakazuje praktyk przypominających chiński system *social scoringu* i techniki manipulacji podprogowej, wzmacniając ochronę autonomii jednostki w cyfrowym świecie[11].

Wyzwania implementacyjne

Wprowadzenie AI Act w Unii Europejskiej to bez wątpienia moment historyczny. Europa jako pierwsza zdecydowała się nie tylko mówić o etyce sztucznej inteligencji, lecz także zamknąć ją w ramy prawa. Jednak implementacja tego rozporządzenia, choć ambitna, napotyka na wiele wyzwań, zarówno technicznych, jak i społecznych czy geopolitycznych.

Jednym z najtrudniejszych zadań jest samo **wykrywanie zakazanych praktyk**. Jak bowiem rozpoznać w kodzie algorytmu elementy *social scoringu* czy manipulacji podprogowej? Potrzebne są narzędzia analityczne zdolne do „prześwietlania" systemów AI, a prace nad ich opracowaniem dopiero trwają. Polski UOKiK prowadzi obecnie projekt badawczy, który wykorzystuje sztuczną inteligencję do automatycznego tropienia *dark patterns* na stronach internetowych – to przykład, jak regulatorzy próbują odpowiadać technologią na technologię.

Nie mniej istotny jest **międzynarodowy charakter platform cyfrowych**. Wielkie modele AI, media społecznościowe czy sklepy internetowe działają ponad granicami, co oznacza, że unijne przepisy muszą zmagać się z firmami spoza Europy. Ponadto **technologia rozwija się szybciej niż prawo** – ryzyko, że regulacje będą spóźnione w momencie ich wejścia w życie, jest wpisane w logikę cyfrowej rewolucji.

Są też wyzwania społeczne. Większość użytkowników nie zdaje sobie sprawy, że podlega manipulacji algorytmicznej – stąd rosnące znaczenie **edukacji cyfrowej**, która może uzbroić obywateli w umiejętność rozpoznawania nieuczciwych praktyk. Do tego dochodzi **presja ekonomiczna**: zakaz *social scoringu* czy manipulacji podprogowej zmusza firmy do zmiany modeli biznesowych, szczególnie w branżach bazujących na danych, jak fintech czy e-commerce. Napięcie pojawia się także przy próbie znalezienia **równowagi między bezpieczeństwem a prywatnością**. Przykładowo: *scoring* kredytowy czy systemy antyfraudowe bywają potrzebne, ale łatwo mogą stać się nadużyciem.

10 Wytyczne Komisji Europejskiej w zakresie aktu AI dostępnych online: https://www.nask.pl/magazyn/wytyczne-komisji-europejskiej-w-zakresie-aktu-o-ai-opracowanie [dostęp: 15.08.2025].

11 Szerzej o praktycznych aspektach wdrażania zakazów i definicji zawartych w rozporządzeniu można przeczytać w **Wytycznych Komisji Europejskiej w zakresie aktu AI**: https://www.nask.pl/magazyn/wytyczne-komisji-europejskiej-w-zakresie-aktu-o-ai-opracowanie [dostęp: 15.08.2025].

Ambicją UE jest **harmonizacja międzynarodowa**, czyli stworzenie globalnych standardów etycznego wykorzystania AI. Podobne inicjatywy podejmują OECD czy ONZ. Ważną rolę odgrywać ma też **certyfikacja systemów AI** pod kątem zgodności z prawem i zasadami etycznymi oraz **współpraca regulatorów z nauką i biznesem**, tak by szybciej identyfikować nowe zagrożenia.

Równolegle rozwijają się narzędzia ochronne dla obywateli. Oprócz edukacji cyfrowej pojawiają się **programy techniczne** pozwalające wykrywać *dark patterns* i inne formy manipulacji. Coraz większe znaczenie zyskuje także **prawo do wyjaśnienia**, czyli możliwość domagania się od systemów AI przejrzystego uzasadnienia decyzji dotyczących obywatela.

Nie sposób pominąć ostrzeżeń płynących z obszaru kultury. Serial „Black Mirror" w odcinku „Nosedive" przedstawił wizję świata, w którym każda interakcja oceniana jest w pięciogwiazdkowej skali, a wynik decyduje o dostępie do mieszkań czy usług. To, co jeszcze kilka lat temu wyglądało na fantazję dystopijną, dziś jest realną praktyką. Również literatura – od Kafki po Orwella – ostrzegała przed biurokratyczną i technologiczną kontrolą. Różnica polega na tym, że współczesne systemy AI działają znacznie subtelniej: na poziomie podświadomości, przy użyciu naszych własnych danych.

Co zatem dalej? Decydenci polityczni powinni postawić na **międzynarodową współpracę**, inwestować w badania nad przeciwdziałaniem manipulacji algorytmicznej i wzmacniać nadzór nad wdrażaniem AI Act, najlepiej poprzez wyspecjalizowane instytucje. Przemysł musi odpowiedzieć zasadą *privacy by design* i *ethical AI by design*, a także otworzyć algorytmy na niezależny audyt. Obywatele z kolei powinni rozwijać swoje kompetencje cyfrowe, świadomie korzystać z technologii i aktywnie włączać się w debatę publiczną o przyszłości sztucznej inteligencji. AI Act to dopiero początek drogi. Prawdziwym sprawdzianem będzie to, czy uda się zbudować system, w którym technologia będzie wspierać człowieka, zamiast go kontrolować.

Niewidzialna sieć kontroli już się rozprzestrzenia. Naszym zadaniem jest sprawienie, aby pozostała pod demokratycznym nadzorem, służąc wspólnemu dobru, a nie interesom nielicznych. To może być jedna z ostatnich szans na zachowanie ludzkiej autonomii w cyfrowym świecie.