

Sprawczość kalibrowana: od pacjenta i prosumenta do człowieka algorytmicznego w świetle książki *Bad Influence* Deborah Cohen



AUTOR

Fundacja Dobre Państwo

STRESZCZENIE / ABSTRACT

Artykuł analizuje transformację modelu pacjenta w obliczu cyfryzacji medycyny i rozwoju AI. Autor wyróżnia trzy współistniejące figury: pacjenta podporządkowanego autorytetowi lekarza (model Parsons), refleksyjnego prosumenta zarządzającego zdrowiem w sieci oraz podmiot algorytmiczny, którego decyzje są kształtowane przez systemy obliczeniowe i architekturę wyboru. Tekst odwołuje się do teorii refleksyjnej nowoczesności Anthony'ego Giddensa, wskazując, że współczesne zdrowie stało się elementem „refleksyjnego projektu siebie”. W tym procesie dane z urządzeń wearables czy systemów AI nie tylko opisują stan organizmu, ale aktywnie modyfikują zachowania i tożsamość jednostki, przesuwając środek ciężkości z tradycyjnej autonomii ku sprawczości kalibrowanej przez algorytmy.

The article analyzes the transformation of the patient model in the face of medical digitalization and the development of AI. The author distinguishes three coexisting figures: the patient subordinated to the doctor's authority (Parsons' model), the reflective prosumer managing health online, and the algorithmic subject whose decisions are shaped by computational systems and choice architecture. The text refers to Anthony Giddens' theory of reflexive modernity, indicating that contemporary health has become an element of the 'reflexive project of the self.' In this process, data from wearables or AI systems do not only describe the state of the body but actively modify the individual's behavior and identity, shifting the center of gravity from traditional autonomy toward agency calibrated by algorithms.

Artykuł analizuje transformację roli pacjenta w systemie ochrony zdrowia – od uległego odbiorcy zaleceń, przez refleksyjnego prosumenta zarządzającego danymi, aż po współczesnego „człowieka algorytmicznego”. Główną tezę tekstu jest stwierdzenie, że w dobie personalizowanej architektury wyboru i systemów predykcyjnych, tradycyjna autonomia staje się iluzją, gdyż nasze preferencje są kształtowane przez niewidoczne mechanizmy optymalizacji technologicznej. Autor argumentuje, że jedyną skuteczną odpowiedzią na ryzyko manipulacji i deskillingu (zaniku kompetencji) jest wypracowanie modelu „sprawczości kalibrowanej”. Nie polega ona na pozornej samowystarczalności jednostki, lecz na świadomym i transparentnym dopasowaniu stopnia delegowania decyzji do jakości dowodów oraz osobistych wartości. Tekst rzuca nowe światło na relację lekarz-pacjent-algorytm, postulując stworzenie systemu „konstytucjonalizmu epistemicznego”, w którym żadna ze stron nie posiada monopolu na prawdę, a technologia służy wspieraniu ludzkiego rozumowania, zamiast go zastępować.

SŁOWA KLUCZOWE

sprawczość kalibrowana

człowiek algorytmiczny

architektura wyboru

hypernudge

prosumeryzm medyczny

automation bias

governmentality

refleksyjny projekt siebie

distributed cognition

cultural health capital

empatyczne kłamstwo AI

shared decision-making

socjologia ryzyka

epistemiczny konstytucjonalizm

quantified self

Ewolucja modelu pacjenta od uległości do autonomii

Od pacjenta do człowieka algorytmicznego – sprawczość, ryzyko i architektura decyzji zdrowotnych

Analiza prowadzi do problemu głębszego niż sama cyfryzacja medycyny, komercjalizacja diagnostyki czy wzrost znaczenia sztucznej inteligencji. Zmienia się bowiem nie tylko sposób świadczenia opieki, lecz także społeczny model człowieka korzystającego z medycyny. W ciągu niespełna stulecia można wyróżnić co najmniej trzy nakładające się figury: pacjenta podporządkowanego autorytetowi profesjonalnemu, refleksyjnego prosumenta zarządzającego własnym zdrowiem oraz podmiot algorytmiczny, którego zachowania, ryzyka i możliwe decyzje są coraz częściej mierzone, przewidywane, podpowiadane i porządkowane przez systemy

obliczeniowe. Nie są to trzy następujące po sobie epoki, z których późniejsza całkowicie usuwa wcześniejszą; wszystkie one współistnieją. Ten sam człowiek może w szpitalu podporządkować się decyzjom zespołu intensywnej terapii, wieczorem samodzielnie analizować wyniki na IKP, a następnego dnia pozwolić smartwatchowi, wyszukiwarce lub systemowi AI zasugerować, które pytanie powinien zadać lekarzowi.

Klasycznym punktem odniesienia pozostaje koncepcja sick role Talcotta Parsonsa. W funkcjonalistycznej wizji społeczeństwa choroba nie była wyłącznie stanem biologicznym, lecz naruszeniem normalnego wykonywania ról społecznych. Człowiek chory otrzymywał czasowe zwolnienie z części obowiązków, ale równocześnie był zobowiązany uznać stan choroby za niepożądany, dążyć do wyzdrowienia i współpracować z kompetentnym profesjonalistą. Lekarz pełnił tu funkcję społecznego arbitra legitymizującego status chorego. Parsonsowska konstrukcja była później wielokrotnie krytykowana za niedopasowanie do chorób przewlekłych, niepełnosprawności, sporów o diagnozę czy aktywnych ruchów pacjenckich, lecz pozostaje jednym z kluczowych punktów teorii socjologii medycyny. W modelu tym wiedza i odpowiedzialność były zorganizowane hierarchicznie: pacjent posiadał doświadczenie objawu, lecz to profesjonalista określał jego nazwę, znaczenie i właściwy sposób działania. Taka asymetria miała racjonalną podstawę – wykształcenie medyczne rzeczywiście dostarczało kompetencji niedostępnych laikowi. Hierarchia epistemiczna łatwo jednak przechodziła w hierarchię społeczną.

Z faktu, że lekarz wiedział więcej o farmakologii czy patofizjologii, wyprowadzano czasem szersze przekonanie, że powinien on również lepiej wiedzieć, jaką jakość życia pacjent powinien uznać za akceptowalną, jakie ryzyko przyjąć oraz jakie cele życiowe są dla niego właściwe. Bioetyczna reakcja drugiej połowy XX wieku przesunęła środek ciężkości ku autonomii, świadomej zgodzie i shared decision-making. Pacjent przestał być wyłącznie przedmiotem profesjonalnej dobroczynności; lekarz miał dostarczyć wiedzę, lecz wartość możliwych rezultatów miała być oceniana również z perspektywy pacjenta. Oznaczało to zmianę ontologiczną: człowiek chory nie był już przede wszystkim organizmem wymagającym profesjonalnej naprawy, lecz podmiotem posiadającym własną hierarchię wartości.

Pacjent jako prosument i refleksyjny projektant własnego zdrowia

Internet dokonał kolejnego przesunięcia, które autor książki trafnie ujmuje poprzez pojęcie prosumenta. Człowiek nie tylko konsumuje wiedzę zdrowotną, ale ją współprodukuje: komentuje, recenzuje, klasyfikuje specjalistów i publikuje własne doświadczenia, włączając je do społecznych

repozytoriów wiedzy. Literatura socjologiczna opisuje tę ewolucję jako rozszerzenie Parsonsowskiej roli chorego o cyberprzestrzeń. Użytkownik sieci staje się refleksyjnym konsumentem informacji, który zarządza własnym zdrowiem w triangulacji między profesjonalną wiedzą, osobistym doświadczeniem a zasobami cyfrowymi.

W tym kontekście szczególnie przydatna okazuje się teoria refleksyjnej nowoczesności Anthony'ego Giddensa. W społeczeństwie późnej nowoczesności jednostka nie otrzymuje już jednej stabilnej, odziedziczonej odpowiedzi na pytanie o to, kim jest i jak powinna żyć. Tożsamość staje się projektem – procesem nieustannie konstruowanym i aktualizowanym w świetle nowych danych. Refleksyjność oznacza tutaj, że wiedza społeczna powraca do świata, który opisuje, realnie zmieniając zachowania aktorów. Informacja nie stoi więc obok rzeczywistości, lecz aktywnie uczestniczy w jej produkcji, czego modelowym przykładem jest medycyna cyfrowa.

Człowiek otrzymuje informację o ryzyku metabolicznym i zmienia dietę; odczytuje HRV i koryguje trening; poznaje kryteria ADHD i rekonstruuje historię własnego dzieciństwa; odkrywa termin „perimenopause” i reinterpretuje doświadczenia z ostatnich lat; obserwuje krzywą CGM i modyfikuje wybór posiłku. Wiedza o człowieku wraca do niego i modyfikuje obiekt, który miała jedynie opisywać. Dlatego w późnej nowoczesności zdrowie staje się częścią „refleksyjnego projektu siebie”. Nie wystarczy już być zdrowym w sensie braku poważnej choroby. Współczesny człowiek ma rozumieć swoje ryzyka, świadomie wybierać dietę, oceniać aktywność, zarządzać płodnością i kontrolować sen, profilaktykę czy screening. Biografia biologiczna zostaje objęta podobnym obowiązkiem projektowania jak kariera zawodowa, edukacja czy tożsamość konsumencka.

Na tym właśnie polega społeczna doniosłość prosumeryzmu medycznego. Nie jest on jedynie zwiększeniem liczby źródeł informacji, lecz przekształceniem moralnego modelu odpowiedzialności. Parsonsowski chory miał przede wszystkim współpracować z lekarzem. Współczesny prosument musi dodatkowo sam wyszukiwać wiedzę, monitorować ryzyko i podejmować działania prewencyjne. Dawniej pacjent mógł być krytykowany za nieprzestrzeganie zaleceń; dziś może zostać oceniony negatywnie również za to, że zbyt późno wykonał badanie, nie korzystał z dostępnych danych lub nie „zoptymalizował” swojego stylu życia.

Człowiek ryzyka w sieci cyfrowej dyscypliny

W tym miejscu refleksyjna nowoczesność Giddensa zderza się z teorią społeczeństwa ryzyka Ulricha Becka. Autor wskazuje, że rozwinięta nowoczesność organizuje się coraz intensywniej wokół przewidywania i dystrybucji przyszłych zagrożeń. Rozwój wiedzy nie eliminuje ryzyka, lecz czyni je bardziej widzialnym, obliczalnym i politycznie istotnym. Jednocześnie postęp technologiczny

generuje nowe rodzaje niepewności, których wcześniejsze społeczeństwa nie musiały kalkulować. Choć teoria Becka bywała krytykowana za nadmierną generalizację, jej znaczenie dla socjologii ryzyka jest kluczowe, a cyfrowe zdrowie jedynie radykalizuje tę logikę.

Człowiek nie czeka już na zdarzenie; żyje w świecie probabilistycznych przyszłości. Otaczają go pięcioletnie ryzyko sercowo-naczyniowe, predyspozycje genetyczne, metabolic score, wiek biologiczny, prawdopodobieństwo owulacji czy trendy HRV i readiness score. Nie będąc jeszcze chorym, zostaje on otoczony przez swoje potencjalne przyszłe schorzenia. Z perspektywy teorii decyzji jest to postęp – lepsza predykcja pozwala na wcześniejszą reakcję.

Z punktu widzenia socjologii wyłania się jednak nowa pozycja społeczna: człowiek ryzyka. Nie jest on jeszcze pacjentem w sensie Parsonsowskim, lecz nie może już traktować zdrowia jako stanu nieproblematicznego. Jego powinnością staje się aktywne reagowanie na prawdopodobieństwa.

Tutaj teoria ryzyka spotyka Foucaultowską governmentality. Nowoczesna władza nie musi opierać się na bezpośrednim zakazie; może organizować środowisko tak, by podmioty same kierowały swoim zachowaniem zgodnie ze społecznie uznanymi normami. Analizy cyfrowego zdrowia oparte na myśli Foucaulta pokazują, że self-tracking może jednocześnie zwiększać autonomię i responsabilizować jednostkę: pacjent sam siebie obserwuje, mierzy, reaguje i dyscyplinuje własne ciało. To zasadnicza różnica między dawnym a cyfrowym nadzorem zdrowotnym. W modelu dyscyplinarnym kontrola pochodziła z zewnątrz; w modelu quantified self człowiek nosi aparat obserwacji na nadgarstku i z własnej woli sprawdza go kilkanaście razy dziennie. Nie ma potrzeby istnienia strażnika – Norma została przeniesiona do interfejsu.

Foucaultowski model władzy bywa często błędnie interpretowany jako teoria spisku lub prostego zniewolenia. W rzeczywistości governmentality jest bardziej interesująca, gdyż ukazuje produktywną stronę władzy. Dyscyplina nie tylko ogranicza, ale wytwarza konkretne kompetencje, tożsamości i możliwości działania.

Smartwatch może realnie pomóc zwiększyć aktywność, aplikacja poprawić przestrzeganie leczenia, a IKP ułatwić kontrolę nad dokumentacją. Człowiek staje się jednocześnie bardziej sprawczy i intensywniej włączony w system normowania zachowań. Tę ambiwalencję potwierdza literatura dotycząca technologicznego self-care: cyfrowa opieka może przesuwac odpowiedzialność na pacjenta i wymuszać samodyscyplinę, a zarazem tworzyć realne przestrzenie aktywności i odzyskiwania kontroli nad zdrowiem. Te dwa skutki nie są sprzeczne; ich współwystępowanie jest cechą nowoczesnego zarządzania zdrowiem.

Na gruncie teorii Pierre'a Bourdieu problem zyskuje dodatkowy wymiar. Jeśli system oczekuje pacjenta aktywnego i refleksyjnego, korzyść odnoszą przede wszystkim ci, którzy posiadają zasoby

pozwalające pełnić taką rolę. „Sprawczy pacjent” nie jest abstrakcyjną jednostką rozłożoną równomiernie w populacji, lecz osobą dysponującą określonym kapitałem kulturowym, ekonomicznym i społecznym. Badania oparte na pojęciach Bourdieu dowodzą, że sprawne poruszanie się po systemie opieki wymaga swoistej system knowledge – wiedzy o tym, gdzie szukać informacji, kogo pytać i jak wykorzystać sieć relacji. Kompetencja ta wyrasta z kapitału kulturowego oraz przywilejów ekonomicznych. Pojęcie cultural health capital rozwija tę intuicję: w interakcji z medycyną premiowane są konkretne style komunikacji, zasób słownictwa i zdolność negocjowania decyzji. Te kompetencje poprawiają jakość relacji z profesjonalistami i pozycję pacjenta w polu medycznym. Prosumentyzacja może zatem demokratyzować medycynę, a jednocześnie pogłębiać nierówności – osoba o wysokim kapitale zdrowotnym wykorzysta internet do znalezienia lepszego specjalisty i merytorycznego przygotowania się do konsultacji.

Człowiek algorytmiczny w sieci architektury wyboru

Osoba mniej uprzywilejowana może mieć dostęp do tego samego oceanu informacji, lecz nie posiadać narzędzi niezbędnych do jego hierarchizacji. **Informacyjny nadmiar nie jest społecznie neutralny.** Dawniej przewagę dawał sam dostęp do danych; dziś daje ją umiejętność ich odsiewania. To przesunięcie z **information scarcity** w stronę **attention scarcity** i **verification scarcity** stanowi jedno z kluczowych zjawisk współczesnej epistemologii społecznej. Zasobem deficytowym przestaje być treść, a staje się zdolność oceny tego, którym treściom warto poświęcić uwagę.

W tym kontekście pojawia się trzeci model podmiotu: **człowiek algorytmiczny**. Nie chodzi tu o wizję człowieka kontrolowanego przez sztuczną inteligencję w futurystycznym sensie, lecz o podmiot, którego środowisko poznawcze jest w coraz większym stopniu selekcjonowane przez systemy rekomendacyjne, predykcyjne i scoringowe. Człowiek nadal formalnie dokonuje wyboru, jednak zestaw informacji, kolejność opcji, widoczność ryzyk czy atrakcyjność działań są wcześniej uporządkowane. Autor książki trafnie identyfikuje mechanizm przejścia od społecznego grafu relacji do rekomendacji opartych na zachowaniu i treści: algorytm obserwuje, co zatrzymuje uwagę, a następnie dostarcza więcej podobnych materiałów.

Rezultatem nie musi być klasyczna „bańka filtrująca” w jej radykalnej formie. Wystarczy subtelniejsze zjawisko: prawdopodobieństwo ponownej ekspozycji na określony model wyjaśnienia choroby rośnie tylko dlatego, że odbiorca wcześniej poświęcił mu uwagę. W ten sposób algorytm nie musi „przekonywać” użytkownika w tradycyjnym rozumieniu perswazji – może po prostu **zmienić częstotliwość świata, który użytkownik widzi**. Jeśli ktoś obejrzy kilka materiałów o

ADHD, menopauzie, kortyzolu, GLP-1 czy toksyczności konkretnego produktu, system zwiększy udział podobnych treści w jego otoczeniu. Z perspektywy psychologii poznawczej rośnie dostępność reprezentacji, a wraz z nią skłonność do korzystania z heurystyki dostępności i efektu iluzorycznej prawdy. Przesuwa to punkt ciężkości analizy z perswazji na **architekturę wyboru**.

Thaler i Sunstein spopularyzowali pojęcie **nudge**, wskazując, że zachowania można modyfikować poprzez organizację środowiska decyzyjnego, nie odbierając przy tym formalnej wolności wyboru. Metaanaliza obejmująca ponad dwa miliony obserwacji potwierdziła, że interwencje z zakresu **choice architecture** realnie zmieniają zachowanie, choć skala efektów zależy od techniki i kontekstu. W medycynie architektura wyboru operuje na wielu poziomach: system może domyślnie zaznaczyć lek generyczny, wysłać przypomnienie o szczepieniu, wyróżnić najbliższy termin wizyty czy zasugerować dodatkowe badanie. Systematyczny przegląd tzw. **EHR nudges** pokazuje, że elektroniczna dokumentacja medyczna sama w sobie staje się narzędziem kształtującym decyzje klinicystów poprzez kolejność opcji, wartości domyślne oraz alerty.

Ryzyko manipulacji w personalizowanej architekturze wyboru

Nudge nie musi być manipulacją. Jeśli przypomnienie zwiększa regularność przyjmowania leku, a domyślny wybór ogranicza przepisywanie droższych preparatów bez korzyści klinicznej, choice architecture może realnie podnosić efektywność i dobrostan pacjenta. Problem pojawia się jednak wtedy, gdy funkcja celu projektanta różni się od funkcji celu użytkownika. Platforma może optymalizować liczbę kliknięć, przychód czy retencję, podczas gdy użytkownik dąży do odzyskania zdrowia.

To, co technicznie pozostaje tą samą architekturą wyboru, w sensie normatywnym może być zupełnie inną interwencją. Sztuczna inteligencja radykalizuje ten problem, czyniąc choice architecture personalizowaną, dynamiczną i predykcyjną. W literaturze pojawia się w związku z tym pojęcie **hypernudge**: system nie ustawia jednego stałego defaultu dla wszystkich, lecz na bieżąco reorganizuje środowisko wyboru w reakcji na zachowanie konkretnej osoby. Najnowsze analizy etyki AI wskazują, że cyfrowe nudges są bardziej wszechobecne i ukryte, a przez to trudniejsze do rozpoznania niż klasyczne interwencje behawioralne. W tym miejscu formalna autonomia może oddzielić się od autonomii epistemicznej.

Użytkownik nadal ma możliwość kliknięcia „nie”. Jednak zanim to zrobi, system już zdecydował, co zobaczy, w jakiej kolejności, z jakim framingiem i przy jakiej liczbie alternatyw. Sama możliwość

końcowej odmowy nie oznacza pełnej kontroli nad procesem kształtowania preferencji. W klasycznej teorii informed consent kluczowe było pytanie, czy pacjent otrzymał wystarczającą ilość informacji. W epoce algorytmicznej problem staje się odwrotny: czy system nie zyskał zbyt dużej władzy nad selekcją danych, na podstawie których pacjent buduje swoje przekonania. Autonomia wymaga zatem nie tylko wolności podjęcia decyzji, ale i stopnia kontroli nad epistemicznym środowiskiem tej decyzji.

Z tego punktu widzenia współczesna literatura bioetyczna proponuje odejście od statycznego rozumienia autonomii jako jednorazowej zgody. Systematyczny przegląd z 2026 roku dotyczący AI i autonomii pacjenta wskazuje na dualizm: z jednej strony potencjał wspierania shared decision-making, z drugiej zaś ryzyka wynikające z nieprzejrzystości algorytmów, manipulacji preferencjami czy de-skillingu. Autorzy przeglądu sugerują, by rozumieć autonomię jako zdolność do krytycznego udziału i dialogu w systemie pośredniczonym przez AI. Oznacza to, że człowiek algorytmiczny nie jest z definicji mniej autonomiczny; AI może wręcz zwiększyć jego sprawczość.

Technologia ta potrafi przetłumaczyć żargon medyczny, wyjaśnić wyniki badań czy pomóc w porównaniu opcji terapeutycznych, uwidaczniając ryzyka wcześniej niezrozumiałe dla pacjenta. Badania nad AI i shared decision-making potwierdzają możliwość poszerzania dostępu do wiedzy, choć rzeczywisty efekt zależy od jakości systemu i przejrzystości jego wdrożenia w praktyce klinicznej. Istnieje jednak niebezpieczeństwo, że AI stanie się nadmiernie przekonującym pośrednikiem epistemicznym.

Pułapka empatycznej formy i automatyzacji zaufania

W znanym badaniu opublikowanym w "JAMA Internal Medicine" odpowiedzi chatbota na pytania zdrowotne z publicznego forum były częściej preferowane przez klinicystów i oceniane jako bardziej empatyczne oraz wyższej jakości niż krótsze wypowiedzi lekarzy. Choć badanie obejmowało tylko 195 pytań, nie było próbą kliniczną i nie dowodziło przewagi AI w realnym leczeniu, obnażyło istotny mechanizm: system generatywny potrafi produkować silne społeczne sygnały kompetencji, takie jak długość wypowiedzi, uporządkowanie treści czy empatyczny język. Autor książki określa to mianem "empatycznego kłamstwa AI", choć pojęcie to wymaga doprecyzowania. Empatia stylistyczna nie jest tożsama z prawdziwością, ale nie oznacza z definicji fałszu. Problem tkwi w rozszczepieniu dwóch cech, które człowiek intuicyjnie łączy: w interakcji międzyludzkiej spokojny, szczegółowy i troskliwy rozmówca bywa odbierany jako bardziej godny zaufania. Model językowy może jednak reprodukcować sygnały troski niezależnie od poprawności epistemicznej konkretnej

odpowiedzi. Użytkownik musi zatem nauczyć się oddzielać jakość retorycznej relacji od jakości dowodu. To nowa wersja starego problemu heurystyki autorytetu.

Biały fartuch nie był przecież dowodem nieomyślności lekarza, tak jak płynna i empatyczna odpowiedź nie jest dowodem nieomyślności AI. W obu przypadkach forma może służyć jako poznawczy shortcut. Na poziomie klinicznym zjawisko to znane jest jako automation bias – tendencja do nadmiernego polegania na rekomendacjach systemu, szczególnie w warunkach obciążenia poznawczego, presji czasu i wysokiej złożoności zadania. Systematyczny przegląd 74 badań wskazuje, że skłonność do nadmiernego zaufania automatyzacji zależy m.in. od doświadczenia użytkownika, poziomu zaufania do systemu oraz sposobu prezentowania rekomendacji. Ma to ogromne znaczenie, gdyż sztuczna inteligencja może jednocześnie zmniejszać błędy i produkować nowe kategorie pomyłek. Jeśli system jest przeciętnie bardzo dobry, człowiek racjonalnie zaczyna mu ufać. Im częściej jednak rekomendacja okazuje się trafna, tym większe ryzyko, że wyjątkowy błąd zostanie przeoczony właśnie przez to zaufanie. To paradoks wysokiej jakości automatyzacji: system słaby jest łatwo kwestionowany, system bardzo dobry może osłabiać czujność. Pojawia się problem out-of-the-loop performance – człowiek formalnie pozostaje "w pętli", ale jego realna zdolność do niezależnej weryfikacji spada, ponieważ coraz rzadziej musi samodzielnie przeprowadzać pełny proces rozumowania.

W medycynie dotyczy to zarówno pacjentów, jak i profesjonalistów. Jeśli algorytm przez setki przypadków prawidłowo klasyfikuje obraz radiologiczny, klinicysta może stopniowo przenieść część uwagi na wynik systemu. Jeśli asystent AI przez miesiące trafnie wyjaśnia proste wyniki, pacjent może zacząć polegać na nim także w sytuacjach bardziej złożonych. Granica kompetencji narzędzia nie jest psychologicznie tak widoczna jak granica kompetencji człowieka. Najnowszy przegląd dotyczący algorytmicznego wspomagania decyzji w opiece zdrowotnej zauważa, że mimo imponujących wyników technicznych niektórych modeli, ich wpływ na patient-relevant outcomes pozostaje znacznie mniej pewny. Przypomina to problem biomarkerów: system może świetnie klasyfikować dane, ale kluczowe pytanie kliniczne brzmi, czy dzięki niemu ludzie żyją lepiej. Algorytmiczna medycyna wymaga zatem nowej wersji shared decision-making.

W klasycznym modelu uczestniczyli lekarz i pacjent; teraz pojawia się trzeci aktor – system obliczeniowy. Choć nie jest on podmiotem moralnym, wpływa na rozkład wiedzy w relacji: dostarcza prognozę, sugeruje opcje i porządkuje informacje. Powstaje układ doctor-patient-algorithm, w którym trzeba ustalić nie tylko to, kto podejmuje ostateczną decyzję, lecz także kto definiuje problem i generuje dostępne alternatywy. Ostateczne kliknięcie "akceptuję" może należeć do człowieka, podczas gdy algorytm wcześniej określił, jakie opcje znalazły się na ekranie. Władza nad menu może być równie ważna jak władza nad wyborem pozycji z menu.

Herbert Simon w teorii bounded rationality wykazał, że człowiek nie podejmuje decyzji w oparciu o pełną informację i nieskończoną moc obliczeniową. Nie maksymalizuje abstrakcyjnej funkcji użyteczności przy znajomości wszystkich opcji, lecz dokonuje wyborów wewnątrz ograniczonego pola informacji, czasu i uwagi. Algorytm może te ograniczenia zmniejszać, przetwarzając więcej danych, ale jednocześnie zaczyna definiować właśnie to ograniczone pole, w którym człowiek podejmuje decyzję. Dlatego największą stawką AI w medycynie nie musi być zastąpienie lekarza.

Architektura wyboru kształtuje decyzje medyczne

Jednym z takich mechanizmów może być przejęcie funkcji selekcji uwagi. System wskazuje, które dane są anomalne, które objawy pasują do wzorca, jakie badanie powinno być kolejne, który artykuł jest istotny i jakie pytanie użytkownik powinien zadać. Kontrola uwagi stanowi szczególnie subtelny formę kontroli poznawczej, ponieważ człowiek zachowuje złudzenie samodzielnego rozumowania. Daniel Kahneman i Amos Tversky wykazali, że ludzkie decyzje są kształtowane przez framing, heurystyki, dostępność oraz punkty odniesienia. Algorytmiczny interfejs może te mechanizmy wykorzystywać – świadomie lub nieświadomie.

Jeśli aplikacja wyświetla wynik na czerwono, użytkownik reaguje inaczej niż wtedy, gdy ten sam parametr opisano jako „nieznacznie poza typowym zakresem”. Jeśli podkreślony zostanie 20-procentowy względny wzrost ryzyka bez podania wartości bezwzględnej, percepcja korzyści z terapii ulega zmianie. Forma prezentacji nie jest opakowaniem decyzji. Jest częścią decyzji. Prowadzi to do koncepcji libertarian paternalism, ale jednocześnie obnaża jej granice. Nudge teoretycznie zachowuje formalną wolność, jednak w medycynie głęboka asymetria wiedzy sprawia, że sposób prezentacji danych ma wyjątkowo silny wpływ. Pacjent nie wybiera pomiędzy dwoma markami płatków. Wybiera biopsję, leczenie onkologiczne, HRT, statynę albo sposób planowania płodności.

W takich obszarach pytanie o przejrzystość choice architecture staje się pytaniem bioetycznym. Jeszcze poważniejsza jest możliwość personalizacji perswazji. Jeżeli system zna historię wcześniejszych decyzji, lęków i zachowań użytkownika, może przewidzieć, który sposób przedstawienia informacji najskuteczniej zmieni jego wybór. W tradycyjnym nudge wszystkim ustawiano podobny default; w hypernudge środowisko może zostać dopasowane indywidualnie. Z punktu widzenia skuteczności jest to narzędzie potężne, lecz z perspektywy autonomii rodzi pytanie, czy człowiek nadal rozpoznaje źródło własnych preferencji.

Nie trzeba jednak zakładać, że personalizacja zawsze służy manipulacji. Można wyobrazić sobie autonomy-enhancing AI, która działa odwrotnie: pokazuje użytkownikowi, jakie framingi na niego

oddziałują, przedstawia alternatywne interpretacje, wymusza chwilę refleksji przy decyzjach wysokiego ryzyka i przypomina o niepewności własnej rekomendacji. Najnowsza literatura filozoficzna wskazuje, że AI może nie tylko perswadować, ale również chronić użytkownika przed perswazją innych aktorów. Projektowanie algorytmów staje się zatem częścią instytucjonalnej teorii autonomii.

Nie wystarczy stwierdzić, że człowiek odpowiada za własny wybór. Należy zapytać, czy środowisko zaprojektowano tak, by pomagać mu rozumować, czy przede wszystkim po to, aby zwiększyć prawdopodobieństwo określonej reakcji. Tę różnicę dobrze oddaje pojęcie epistemic agency. Sprawczy poznawczo człowiek nie jest kimś, kto samodzielnie wie wszystko – taka autonomia byłaby niemożliwa. Jest raczej osobą zdolną do rozpoznawania granic własnej wiedzy, wyboru źródeł, kwestionowania rekomendacji i rozumienia, kiedy powinna oddać część decyzji bardziej kompetentnemu ekspertowi. Paradoksalnie dojrząca autonomia obejmuje świadome delegowanie, a nie nieustanną samowystarczalność.

Takie podejście pozwala wyjść poza jałowe przeciwstawienie "człowiek czy algorytm". Lotnictwo, intensywna terapia czy diagnostyka obrazowa od dawna dowodzą, że najskuteczniejsze systemy są hybrydowe. Pytanie brzmi: jak podzielić funkcje? Algorytm może być lepszy w detekcji wzorców, człowiek zaś w rozumieniu kontekstu biograficznego; system może wyliczać prawdopodobieństwo, a pacjent określać wartość możliwych rezultatów; lekarz natomiast może integrować model biologiczny z preferencjami i odpowiedzialnością zawodową.

Taka architektura wymaga jednak ochrony przed deskillingiem. Jeśli profesjonalista zbyt długo korzysta z systemu wykonującego część rozumowania, może utracić praktykę niezbędną właśnie wtedy, gdy system się myli lub zawodzi. Analogicznie pacjent, który przez lata deleguje interpretację wszystkich danych do aplikacji, może tracić zdolność samodzielnego rozpoznawania sygnałów własnego organizmu. Najnowsze przeglądy etyczne AI w ochronie zdrowia wskazują deskilling jako jedno z istotnych zagrożeń dla autonomii i odpowiedzialności.

Konstytucjonalizm epistemiczny jako ochrona przed koncentracją autorytetu

W tym miejscu dochodzimy do paradoksu, który można nazwać paradoksem wspomaganej sprawczości. Narzędzie powstaje po to, by zwiększyć zdolność działania, jednak jeśli staje się ono niezbędnym warunkiem każdej decyzji, umiejętność samodzielnego operowania bez niego może

zanikać. Sprawczość rośnie zatem lokalnie, lecz może maleć strukturalnie. Podobny problem dostrzegamy w Bourdieuowskim wymiarze nierówności.

AI może demokratyzować dostęp do informacji medycznej dla osób nieznających języka specjalistycznego, ale jednocześnie może pogłębić przewagę tych, którzy dysponują lepszym modelem, płatnym dostępem, większą biegłością w promptowaniu czy zdolnością weryfikacji odpowiedzi. Systematyczne przeglądy autonomii w epoce AI wskazują właśnie dlatego, że status społeczno-ekonomiczny, literacy oraz kompetencje cyfrowe są kluczowymi czynnikami modyfikującymi korzyści i ryzyka związane z technologią. Algorytm nie usuwa kapitału kulturowego – on jedynie zmienia jego formę.

Dawniej uprzywilejowanie oznaczało znajomość lekarza lub umiejętność czytania podręczników. Jutro może polegać na zdolności odróżnienia modelu generującego prawdopodobną narrację od systemu opartego na zweryfikowanej dokumentacji, a także na umiejętności zadania właściwego pytania i sprawdzenia źródła.

Idea „AI dla wszystkich” nie gwarantuje więc automatycznie demokratyzacji. Jak w każdej poprzedniej rewolucji informacyjnej, kluczowe jest nie tylko to, kto ma dostęp do narzędzia, lecz kto potrafi najskuteczniej przekształcić go w realną korzyść. Z perspektywy teorii Parsonsa obserwujemy niemal odwrócenie dawnego układu obowiązków: o ile dawny chory miał zwrócić się do kompetentnego profesjonalisty, o tyle nowy prosument ma samodzielnie uczestniczyć w poszukiwaniu wiedzy. Człowiek algorytmiczny otrzymuje dodatkowo trzeci obowiązek: zarządzać rekomendacjami systemów, które zarządzają jego uwagą. Dziś zdrowotna kompetencja nie ogranicza się już do umiejętności przeczytania ulotki. Obejmuje ona rozpoznawanie treści sponsorowanych, rozumienie ryzyka względnego i bezwzględnego, znajomość niedoskonałości testów, krytyczną interpretację AI, świadomość konfliktów interesów oraz zdolność zaakceptowania faktu, że poprawna odpowiedź medyczna często brzmi: „nie wiemy wystarczająco dużo”. Jest to ogromne obciążenie poznawcze.

Należy zatem uważać, by idea empowered patient nie stała się kolejnym sposobem przenoszenia odpowiedzialności z systemu na jednostkę. Nie każdy pacjent musi zostać póllekarzem, statystykiem, analitykiem danych i audytorem algorytmów, aby zasługiwać na bezpieczną opiekę. System powinien być projektowany tak, by nawet osoba o ograniczonych zasobach poznawczych i cyfrowych mogła podejmować rozsądne decyzje.

To właśnie tutaj autonomia spotyka się ze sprawiedliwością. Jeśli bezpieczeństwo wymagałoby zbyt wysokiego progu kompetencji (competence threshold), technologia będzie premiować już uprzywilejowanych. Dobry system powinien więc nie tylko udostępniać opcje, ale przede wszystkim

zmniejszać koszt poznawczy właściwego wyboru, nie odbierając przy tym w sposób ukryty możliwości sprzeciwu. Ostatecznie trzy modele – pacjent, prosument i człowiek algorytmiczny – reprezentują różne odpowiedzi na problem niepewności. Parsonsowski pacjent deleguje ją profesjonalistom; prosument próbuje samodzielnie zredukować ją poprzez wiedzę i porównywanie źródeł; człowiek algorytmiczny powierza część jej przetwarzania systemowi predykcyjnemu. Żaden z tych modeli nie jest wystarczający sam w sobie.

Profesjonalista może być paternalistyczny i omylny. Prosument może ulec błędom poznawczym (confirmation bias), reklamie lub lękowi. Algorytm bywa nieprzejrzysty, błędnie skalibrowany lub optymalizowany pod niewłaściwą funkcję celu. Dojrzała medycyna przyszłości nie powinna zatem wybierać jednego suwerena, lecz stworzyć system wzajemnej kontroli epistemicznej, w którym człowiek, profesjonalista i narzędzie obliczeniowe korygują nawzajem swoje ograniczenia. Lekarz powinien móc zakwestionować algorytm, a pacjent – lekarza.

Algorytm może wykryć błąd przeoczony przez człowieka. System powinien umożliwiać uzyskanie drugiej opinii i ujawniać poziom niepewności. Informacja komercyjna musi być oddzielona od rekomendacji klinicznej, a dane pacjenta nie powinny automatycznie stawać się własnością ekonomiczną infrastruktury, która je przetwarza.

Można to nazwać konstytucjonalizmem epistemicznym medycyny. Tak jak system polityczny chroni przed koncentracją władzy poprzez rozdzielenie kompetencji i mechanizmy kontroli, tak system wiedzy zdrowotnej powinien chronić przed koncentracją autorytetu. Nie wynika to z faktu, że lekarze, pacjenci czy AI są z natury niewiarygodni, lecz z tego, że każdy aktor posiada własne ograniczenia i bodźce. Właśnie ten model najlepiej chroni humanistyczne znaczenie medycyny. Człowiek nie może zostać zredukowany ani do posłusznego pacjenta, ani do konsumenta podejmującego rzekomo całkowicie suwerenne decyzje, ani do zestawu predykcyjnych zmiennych. Jest istotą biologiczną, społeczną i narracyjną, której decyzje powstają w relacji z innymi ludźmi i instytucjami. Dlatego AI nie powinna odpowiadać na pytanie „co człowiek ma zrobić?” w oderwaniu od pytania „co dla tego człowieka oznacza dobry rezultat?”.

Predykcja nie zastąpi normatywnej decyzji człowieka

Algorytm potrafi obliczyć prawdopodobieństwo pięcioletniego przeżycia przy zastosowaniu dwóch różnych metod leczenia, lecz nie rozstrzygnie normatywnie, jak pacjent powinien wycenić ryzyko niepełnosprawności, utratę płodności, czas spędzony w szpitalu czy możliwość pozostania aktywnym zawodowo. Wartości te nie są kolejnymi brakującymi zmiennymi technicznymi – są immanentnymi składnikami ludzkiego życia.

W tym miejscu wyraźnie zarysowuje się granica między predykcją a decyzją. Predykcja opisuje możliwe scenariusze przyszłości i ich prawdopodobieństwo; decyzja natomiast wybiera działanie w świetle tych scenariuszy oraz wartości, które nadają im znaczenie. Najbardziej wyrafinowany model predykcyjny nie znosi normatywnego charakteru tego drugiego etapu. Ta różnica powinna stanowić fundament medycyny algorytmicznej. Człowiek może delegować obliczenia, ale nie powinien nieświadomie delegować wartości. Jeśli system jednocześnie oblicza ryzyko i wskazuje jedną opcję jako „rekomendowaną”, musi być jasne, według jakiej funkcji celu ta rekomendacja powstała.

Przyszłość medycyny nie rozstrzygnie się więc w prostej alternatywie: „lekarz czy AI”. Kluczowe jest pytanie: kto definiuje funkcję celu, którą AI optymalizuje? Jeśli będzie nią liczba kliknięć, otrzymamy medycynę uwagi. Jeśli sprzedaż – medycynę konwersji. Jeśli minimalizacja kosztów – medycynę ekonomicznej efektywności. Gdy celem będzie wyłącznie przeżycie, możemy pominąć jakość życia. Jeżeli jednak nadrzędnym celem stanie się maksymalizacja zdolności człowieka do życia zgodnego z jego wartościami, projekt technologii musiałby wyglądać zupełnie inaczej. Polityka algorytmiczna staje się zatem przedłużeniem filozofii medycyny. Kod zawsze optymalizuje jakiś wynik, nawet jeśli wartość ta została ukryta w technicznym parametrze. Nie istnieje całkowicie neutralny system rekomendacji, ponieważ sama decyzja o tym, które zmienne wprowadzić, jak je ważyć i jaki błąd uznać za bardziej kosztowny, ma konsekwencje normatywne. Najważniejsza kompetencja przyszłego społeczeństwa nie będzie więc polegać na zdolności zbudowania modelu, który wie wszystko o człowieku, lecz na umiejętności określenia, czego nie wolno mu rozstrzygać za człowieka.

Im więcej danych, testów, sensorów i predykcji otrzymuje człowiek, tym większy potencjał jego sprawczości. Jednak sprawczość nie jest prostą funkcją liczby dostępnych opcji. Można posiadać tysiące możliwości, a jednocześnie podejmować decyzje w środowisku tak silnie zaprojektowanym, że kierunek zachowania staje się przewidywalny dla systemu wcześniej, niż zostanie świadomie rozpoznany przez samego człowieka. Najgłębszym problemem medycyny cyfrowej nie jest zatem to, że maszyna może wiedzieć o nas za dużo, lecz ryzyko, że system zacznie wiedzieć dostatecznie dużo, aby kształtować, czego człowiek będzie chciał od własnego ciała, zanim ten zdąży rozpoznać źródła tej preferencji.

W tym punkcie powraca napięcie między zyskiem a sprawczością. Rynek może wykorzystać algorytm, by lepiej dopasować usługę do autentycznej potrzeby, ale może też nauczyć się generować tę potrzebę poprzez personalizowaną ekspozycję. Granica między zaspokajaniem preferencji a ich produkowaniem nigdy nie była absolutna w gospodarce konsumenckiej, lecz w obszarze zdrowia nabiera szczególnej wagi – rezultatem perswazji może być tu diagnoza, farmakoterapia lub trwała ingerencja biologiczna. Dlatego autor postuluje wprowadzenie pojęcia suwerenności medycznej

jednostki, rozumianej jednak inaczej niż radykalny indywidualizm. Suwerenny pacjent nie musi wiedzieć wszystkiego ani decydować o wszystkim bez pomocy. Powinien natomiast zachowywać kontrolę nad tym, komu deleguje wiedzę, jakie dane udostępnia, jakie cele uznaje za własne i w którym momencie wycofuje zgodę. Suwerenność ta jest relacyjna, gdyż bez profesjonalnej wiedzy często nie da się jej realnie realizować; społeczna, bo zależy od edukacji i dostępu do opieki; informacyjna, gdyż wymaga kontroli nad danymi; epistemiczna, ponieważ człowiek musi rozumieć podstawy rekomendacji; i wreszcie egzystencjalna, gdyż to on, a nie platforma, powinien określać, co uważa za życie dostatecznie zdrowe, długie i warte prowadzenia.

W tym miejscu medycyna zatacza koło. Zaczynaliśmy od świata, w którym profesjonalista definiował zdrowie pacjenta; doszliśmy do rzeczywistości, w której pacjent dysponuje milionami narzędzi do samodzielnego określania zdrowia, lecz jednocześnie dziesiątki systemów próbują podpowiedzieć mu, co powinien uznać za nieoptymalne. Nie jest więc paradoksem, że wraz z rozwojem technologii autonomii, tym bardziej potrzebujemy teorii granic ingerencji w autonomię. Kolejnym krokiem wywodu będzie przejście od opisu człowieka algorytmicznego do próby zbudowania modelu sprawczości kalibrowanej – takiej, która nie wraca do paternalizmu, ale też nie pozostawia człowieka samotnego wobec rynku, prawdopodobieństwa, algorytmu i biologicznej niepewności. Dopiero na tym poziomie można spiąć w jedną teorię wszystkie wcześniejsze wątki: od Doktora Google i puli objawów, przez overdiagnosis i quantified self, aż po GLP-1, longevity i polskie e-zdrowie.

Dotychczasowa analiza prowadziła od paternalistycznego pacjenta, przez prosumenta zdrowia, aż do człowieka algorytmicznego. Wybór jednego z tych modeli jako ostatecznego ideału byłby błędem, gdyż każdy z nich rozwiązuje jeden problem kosztem stworzenia następnego. Paternalizm ogranicza obciążenie poznawcze chorego, lecz grozi odebraniem mu prawa do definiowania własnych celów. Prosumpcja zwiększa dostęp do wiedzy i możliwości działania, ale może przenieść na jednostkę odpowiedzialność, której nie jest w stanie samodzielnie udźwignąć.

Sprawczość kalibrowana jako świadome dopasowanie delegacji do dowodów

Algorytm może zmniejszyć złożoność decyzji, lecz jednocześnie ukryć przed człowiekiem sposób, w jaki wybór został wcześniej ustrukturyzowany. Potrzebny jest zatem model, który nie utożsamia autonomii ani z posłuszeństwem wobec eksperta, ani z samotną samowystarczalnością pacjenta, ani z technologiczną delegacją rozstrzygnięcia.

Model taki można nazwać sprawczością kalibrowaną. Pojęcie kalibracji pochodzi z języka pomiaru: instrument jest dobrze skalibrowany wtedy, gdy jego wskazania pozostają w odpowiedniej relacji do rzeczywistości, którą ma mierzyć. W epistemologii zaufania analogicznie można mówić o dobrze skalibrowanym przekonaniu, gdy siła pewności odpowiada jakości dostępnego dowodu. Człowiek nadmiernie pewny błędnego rozpoznania jest źle skalibrowany; podobnie źle skalibrowany jest pacjent, który nie ufa silnym dowodom, ponieważ przeczytał pojedynczą historię w internecie. Sprawczość kalibrowana oznaczałaby zatem nie maksymalną samodzielność, lecz zdolność odpowiedniego dopasowania stopnia własnej pewności, zakresu delegacji i intensywności działania do jakości informacji, wielkości ryzyka i odwracalności decyzji.

Tak rozumiana autonomia jest bliższa tradycji relational autonomy niż klasycznemu liberalnemu wyobrażeniu jednostki podejmującej decyzję w społecznej próżni. Literatura bioetyczna podkreśla, że autonomia kształtuje się poprzez relacje, struktury władzy, zależności, role rodzinne, warunki ekonomiczne i dostęp do wiedzy. Nie wymaga to rezygnacji z indywidualności, lecz odejścia od fikcji całkowicie samowystarczalnego decydenta. Człowiek może podejmować własną decyzję, korzystając jednocześnie z wiedzy lekarza, doświadczenia rodziny, narzędzia decyzyjnego czy systemu AI.

Problemem nie jest zależność jako taka, lecz moment, w którym staje się ona nieprzejrzysta, asymetryczna albo praktycznie niemożliwa do odwołania. To rozróżnienie jest zasadnicze dla medycyny: delegacja nie jest przeciwieństwem autonomii. Pacjent, który świadomie wybiera chirurga i pozwala mu w trakcie operacji podejmować techniczne decyzje, nie zrzeka się sprawczości; wykorzystuje ją, aby powierzyć zadanie osobie bardziej kompetentnej. Analogicznie człowiek może racjonalnie korzystać z algorytmu, jeśli rozumie jego funkcję i ograniczenia. Autonomia zostaje osłabiona nie wtedy, gdy człowiek deleguje, lecz gdy nie ma świadomości tego procesu, nie może zmienić delegata lub nie wie, według jakiego kryterium decyzja jest podejmowana. Sprawczość kalibrowana różni się zatem od ideologii pełnej kontroli. Współczesny wellness często przedstawia kontrolę jako najwyższą wartość: kontroluj sen, glukozę, hormony, masę ciała, płodność, biomarkery i ryzyko.

Pokora epistemiczna i transparentność danych jako fundamenty świadomej decyzji

Z perspektywy psychologii poznawczej oraz teorii decyzji taki projekt jest niemożliwy, gdyż świat medyczny ma charakter probabilistyczny, a dostęp do informacji zawsze pozostaje niepełny. Próba osiągnięcia absolutnej kontroli nie eliminuje niepewności; może jedynie przesuwac próg, przy

którym człowiek zaczyna odczuwać ją jako nieakceptowalną. W tym miejscu niezbędna jest epistemic humility – pokora epistemiczna. Nie oznacza ona intelektualnej bezradności ani rytualnego dodawania zastrzeżenia „nie jestem pewien” do każdej wypowiedzi. Jest to raczej kompetencja rozpoznawania granic własnego wnioskowania i komunikowania niepewności tam, gdzie ona rzeczywiście występuje. Klasyczna analiza pokory epistemicznej w medycynie wiąże ją z moralnym obowiązkiem lekarza, ponieważ osądy kliniczne pozostają obciążone niepewnością nawet w epoce evidence-based medicine. Najnowsza literatura rozwija tę intuicję, traktując pokorę nie jako przeciwieństwo profesjonalnej pewności, lecz jako kompetencję związaną z samoświadomością, gotowością do uczenia się i patient-centredness.

Można tu zestawić dwie skrajne kultury komunikacyjne. Tradycyjny paternalizm ukrywał niepewność, gdyż ekspert miał budzić bezkrytyczne zaufanie. Cyfrowy wellness również często ją maskuje, lecz z innego powodu: pewność lepiej się sprzedaje. W obu przypadkach pacjent otrzymuje gotowy komunikat; różnica dotyczy jedynie źródła autorytetu. Sprawczość kalibrowana wymaga czegoś trudniejszego: komunikowania tego, co wiadomo, tego, czego nie wiadomo, oraz tego, jak bardzo ta niewiedza ma znaczenie dla decyzji. Badania nad komunikacją ryzyka dowodzą, że forma przedstawienia prawdopodobieństwa realnie zmienia jego rozumienie. Przegląd 91 badań wskazuje, że formaty prezentujące ryzyko bezwzględne oraz wizualizacje często poprawiały zrozumienie w porównaniu z samym ryzykiem względnym, choć żadna technika nie jest uniwersalnie najskuteczniejsza we wszystkich sytuacjach. Jest to kluczowe, ponieważ różnica między stwierdzeniem „leczenie obniża ryzyko o 50%” a „ryzyko spada z dwóch przypadków na tysiąc do jednego” nie ma charakteru matematycznego – oba sformułowania mogą być prawdziwe – lecz psychologiczny i decyzyjny. Pierwsze wywołuje znacznie silniejsze wrażenie wielkości efektu.

Gerd Gigerenzer określa brak umiejętności posługiwania się informacją probabilistyczną mianem problemu risk literacy. W nowoczesnej medycynie jest to element kompetencji obywatelskich równie ważny, co podstawowa znajomość biologii. Bez rozumienia mianownika człowiek może być jednocześnie znakomicie poinformowany i całkowicie błędnie interpretować własne ryzyko. Sprawczość kalibrowana wymaga więc nie „większej liczby danych”, lecz bardziej transparentnej geometrii danych: wspólnego mianownika, ryzyka bazowego, zakresu niepewności, horyzontu czasowego oraz jasnego rozróżnienia między korelacją, prawdopodobieństwem a efektem przyczynowym. Nie oznacza to jednak, że każdy pacjent musi władać rachunkiem bayesowskim.

Odpowiedzialność za transparentność spoczywa również na systemie. Jeżeli prawidłowa interpretacja komunikatu wymaga eksperckiej wiedzy statystycznej, oznacza to, że informacja została źle zaprojektowana dla odbiorcy. Badania dotyczące osób z ograniczoną health literacy wskazują, że komunikacja wyłącznie werbalna jest niewystarczająca, a framing i sposób wizualizacji

realnie wpływają na proces rozumienia. Sprawiedliwość epistemiczna wymaga zatem projektowania informacji w oparciu o realne zdolności poznawcze człowieka, a nie dla abstrakcyjnego, idealnego odbiorcy.

Kalibracja zaufania zamiast jego bezwarunkowego zwiększania

Co szczególnie ważne, komunikowanie niepewności nie musi niszczyć zaufania. Badania eksperymentalne dowodzą, że wcześniejsze ujawnienie wątpliwości może chronić zaufanie w sytuacjach, gdy naukowe zalecenie ulegnie później zmianie. Podważa to intuicję, według której autorytet wymaga pozorów nieomyślności. Stabilniejsze zaufanie buduje się właśnie wtedy, gdy ekspert od początku zaznacza: „na podstawie obecnych danych najbardziej prawdopodobne jest X, ale istnieje również możliwość Y i przyszłe dane mogą zmienić ocenę”.

Taka perspektywa prowadzi do pojęcia **calibrated trust**, kluczowego w relacji z AI. Systematyczny przegląd badań nad zaufaniem do sztucznej inteligencji w ochronie zdrowia wskazuje na takie determinanty jak niezawodność, bezpieczeństwo, transparentność, integrację z kontekstem klinicznym oraz doświadczenie użytkownika. Optymalnym wynikiem nie jest jednak maksymalne zaufanie, lecz zaufanie **odpowiednie do rzeczywistej jakości systemu**. Pacjent, który nie ufa doskonałemu narzędziu, traci potencjalną korzyść; ten, który bezwarunkowo ufa zawodnemu modelowi, naraża się na szkodę. To istotne przesunięcie językowe. W debacie technologicznej często pada pytanie: „jak zwiększyć zaufanie społeczeństwa do AI?”.

Pytanie to jest źle postawione. Celem etycznym nie powinno być zwiększanie zaufania jako takiego, lecz jego **kalibracja**. Jeśli system wykazuje 70-procentową wiarygodność w danym zastosowaniu, użytkownik nie powinien ufać mu tak samo, jak systemowi o wiarygodności 99,9%. Model, który świetnie tłumaczy pojęcia, lecz nie ma dostępu do pełnej dokumentacji pacjenta, winien być traktowany jako narzędzie edukacyjne, a nie autonomiczny diagnostyk. Autor książki wprowadza tu mocne pojęcie "empatycznego kłamstwa AI", wskazując na niebezpieczeństwo sytuacji, w której wiarygodnie brzmiący i emocjonalnie wspierający system generuje błędną odpowiedź. Intuicję tę należy zachować, choć termin wymaga naukowego doprecyzowania: nie każda empatia AI jest kłamstwem, podobnie jak nie każda ludzka empatia gwarantuje prawdziwość. Problem polega na **decouplingu między sygnałem kompetencji a kompetencją rzeczywistą**. Styl, pewność, długość i empatyczność odpowiedzi mogą działać jako heurystyki zaufania niezależnie od jakości dowodu.

AI tworzy zatem nową wersję klasycznego problemu **white coat effect**. Kiedyś społeczny symbol fartucha wzmacniał autorytet wypowiedzi; dziś podobną rolę może pełnić język uporządkowany, precyzyjny i wspierający. Sprawczość kalibrowana wymaga zrozumienia, że **jakość relacyjna komunikatu i prawdziwość epistemiczna są dwoma niezależnymi wymiarami**. Idealny system powinien zapewniać oba, lecz użytkownik nie może przyjmować jednego jako automatycznego dowodu drugiego. Największym błędem byłoby w tym miejscu zaproponowanie „większej ilości informacji” jako uniwersalnego rozwiązania.

Psychologia decyzji zna zjawisko **choice overload**, a teoria ograniczonej racjonalności Herberta Simona przypomina, że nasze zdolności przetwarzania informacji są skończone. Człowiek nie staje się bardziej autonomiczny dlatego, że przed podjęciem decyzji otrzymał dwustustronicowy dokument. Może wręcz stracić autonomię, jeśli nadmiar materiału przekroczy jego możliwości poznawcze i zmusi go do powrotu do heurystyki: „zrobię to, co powiedział najbardziej pewny siebie ekspert”. Sprawczość wymaga więc **właściwej redukcji złożoności**. To paradoks: dobry ekspert nie jest jedynie źródłem większej wiedzy, lecz reduktorem informacji do tej części, która ma znaczenie dla konkretnej decyzji. Właśnie dlatego współczesny lekarz powinien coraz mniej przypominać strażnika tajemnej wiedzy, a coraz bardziej kuratora istotności. Jego kompetencją jest nie tylko wiedzieć więcej od pacjenta, ale rozpoznawać, czego pacjent nie musi wiedzieć, aby podjąć racjonalną decyzję, a czego pominąć nie wolno. Tak rozumiany ekspert nie konkuruje z internetem ilością faktów – nie ma na to szans. Konkuruje **strukturą ich znaczenia**. Internet w sekundę dostarczy tysiąc artykułów o **incidentaloma**, lecz klinicysta potrafi ocenić, czy konkretna zmiana u konkretnego człowieka powinna zmienić dalsze postępowanie.

To właśnie interpretacja, a nie monopol informacyjny, staje się podstawą nowej profesjonalności. Tutaj **shared decision-making** powinno być rozumiane głębiej niż jako uprzejma rozmowa o preferencjach; jest to procedura **łączenia dwóch typów ekspertyzy**. Profesjonalista zna dowody, prawdopodobieństwa, naturalny przebieg choroby i możliwe interwencje. Pacjent zaś zna własne cele, tolerancję ryzyka, sytuację rodzinną, znaczenie działań niepożądanych i granice akceptowalnego obciążenia. Żaden z tych zasobów samodzielnie nie wystarcza.

Wybór pomiędzy mastektomią a leczeniem oszczędzającym, agresywną terapią onkologiczną a leczeniem objawowym, HRT a rezygnacją z niej lub operacją a obserwacją, nie jest zadaniem do rozwiązania za pomocą samej statystyki. Statystyka określa konsekwencje, lecz to wartości określają ich znaczenie. Sprawczość kalibrowana pojawia się dopiero w miejscu ich spotkania.

Autonomia relacyjna i zasada korzystnego tarcia

Relacyjna teoria autonomii pozwala odejść od fikcji, jakoby pacjent musiał podejmować decyzje w całkowitej samotności, aby były one „naprawdę jego”. Badania nad relational autonomy podkreślają, że decyzje medyczne powstają w rzeczywistych relacjach z rodziną, opiekunami i profesjonalistami, a uwzględnienie tej sieci powiązań nie musi odbierać jednostce ostatecznego prawa wyboru. Autonomia może zatem przejawiać się w deklaracji: „chcę, aby moja córka uczestniczyła w rozmowie” lub „chcę, by lekarz wskazał, co sam wybrałby w podobnej sytuacji”, o ile taka delegacja faktycznie odpowiada preferencjom pacjenta.

W tym miejscu warto odróżnić paternalizm od doradztwa. Lekarz może stwierdzić: „w świetle pańskich wartości i dostępnych danych uważam opcję A za rozsądniejszą”, nie odbierając tym samym pacjentowi autonomii. Problem pojawia się wtedy, gdy profesjonalista ukrywa istnienie alternatywy, manipuluje framingiem lub traktuje odmienne wartości pacjenta jako błąd poznawczy. Antypaternalizm nie powinien jednak zmieniać lekarza w beznamienny katalog możliwości.

Pacjent często potrzebuje interpretacji, a nie tylko menu. Jednocześnie nie każda decyzja wymaga pełnego modelu shared decision-making. W przypadku resuscytacji, sepsy czy ostrego udaru czas oraz struktura dowodów pozostawiają minimalną przestrzeń na wybór. Kalibracja dotyczy zatem również intensywności procesu autonomicznego. Im większa niepewność dowodowa i im silniejszy wpływ osobistych wartości na wynik, tym istotniejsze staje się współdecydowanie. Z kolei im silniejszy konsensus co do przewagi jednej opcji i im bardziej pilna sytuacja, tym bardziej uzasadniona jest profesjonalna dyrektywność. Można to ująć jako zasadę proporcjonalności deliberacji: nie wszystkie wybory medyczne wymagają tej samej głębokości namysłu. Pacjent nie musi prowadzić seminarium bioetycznego przed zastosowaniem antybiotyku w ciężkiej sepsie, natomiast powinien otrzymać znacznie większą przestrzeń na refleksję, gdy różne strategie leczenia nowotworu oferują podobne przeżycie, lecz odmienne konsekwencje funkcjonalne.

Analogiczna zasada powinna obowiązywać w technologiach konsumenckich. Im mniej odwracalna interwencja, większe ryzyko i słabsze dowody, tym większa powinna być ilość tarcia poznawczego przed podjęciem decyzji. W świecie aplikacji projektuje się zazwyczaj odwrotnie – dąży się do frictionless experience, gdzie jednym kliknięciem zamawia się usługę, rozpoczyna subskrypcję lub uzyskuje dostęp do produktu.

W medycynie brak tarcia nie zawsze jest wartością. Można zatem zaproponować pojęcie beneficial friction – korzystnego tarcia. Przy decyzjach o istotnych konsekwencjach system powinien celowo zwalniać proces: wymagać ponownego potwierdzenia, przedstawić alternatywę, ujawnić stopień pewności dowodów, przypomnieć o możliwych działaniach niepożądanych lub zasugerować

konsultację. Jest to przeciwieństwo dark patterns stosowanych w celu zwiększenia konwersji. Projekt odpowiedzialny za zdrowie powinien optymalizować jakość decyzji, a nie szybkość jej finalizacji. W tym punkcie teoria Thaler'a i Sunsteina spotyka się z klasyczną bioetyką.

Nudge może wspierać pacjenta, ale nie powinien maskować kierunku swojej ingerencji. Jeśli system przypomina o szczepieniu, ponieważ korzyść jest dobrze udokumentowana, działanie to posiada silne uzasadnienie. Jeśli jednak aplikacja podsuwa płatny test tylko dlatego, że zwiększa on lifetime value klienta, struktura formalna może być podobna, lecz funkcja celu pozostaje radykalnie inna.

Transparentność i elastyczność poznawcza jako fundamenty sprawczości kalibrowanej

Sprawczość kalibrowana wymaga zatem jawności celu architektury wyboru. Prowadzi to bezpośrednio do problematyki ekonomii principal-agent: pacjent powierza ekspertowi lub platformie część kontroli, ponieważ nie dysponuje równą wiedzą. Jeśli agent posiada interes ekonomiczny inny niż dobro pacjenta, asymetria informacji ulega podwojeniu. Klient nie wie wówczas nie tylko, co powinien zrobić, ale także czy rekomendujący podmiot ma powód, by kierować go ku konkretnej decyzji. Konflikt interesów nie oznacza automatycznej niewiarygodności, lecz zmienia racjonalny poziom zaufania. Z tego powodu samo disclosure nie może być traktowane jako pełne rozwiązanie. Oznaczenie współpracy reklamowej, wskazanie sponsora badania czy ujawnienie udziałów finansowych są konieczne, ale użytkownik wciąż potrzebuje narzędzi do oceny, jak ten konflikt interesów wpływa na konkretną rekomendację. Transparentność jest warunkiem kontroli, a nie jej substytutem.

Analogicznie, explainable AI nie rozwiązuje automatycznie problemu algorytmicznej sprawczości. Wyjaśnienie może być technicznie poprawne, lecz poznawczo bezużyteczne. Pacjent nie potrzebuje listy stu cech modelu; potrzebuje informacji pozwalających zrozumieć, dlaczego wynik ma znaczenie w jego sytuacji i jakie są granice danej rekomendacji. Epistemiczna transparentność nie polega więc na ujawnieniu maksymalnej ilości szczegółów, lecz na udostępnieniu tych właściwych dla kontroli decyzji.

W tym kontekście obserwacja autora książki o „konsolidacji diagnozy” podczas wielomiesięcznego oczekiwania nabiera szerszego znaczenia. Autor opisuje pacjenta, którego przekonania na temat choroby utrwalają się pod wpływem treści internetowych, zanim dojdzie do profesjonalnej

konsultacji. Nie należy jednak opisywać tego procesu językiem „programowania” człowieka, gdyż użytkownik nie jest biernym komputerem.

Mechanizm ten można precyzyjnie wyjaśnić poprzez teorię confirmation bias, commitment effects, cognitive consistency oraz motivated reasoning. Hipoteza przyjęta na początku poszukiwań zaczyna organizować dalszą uwagę: informacje zgodne są łatwiej zapamiętywane, a sprzeczne wymagają większego wysiłku. Im więcej czasu, emocji i środków zainwestowano w daną narrację, tym wyższy staje się psychologiczny koszt jej porzucenia. Diagnoza może zatem przejść z poziomu hipotezy na poziom elementu autobiografii jeszcze przed weryfikacją kliniczną. Sprawczość kalibrowana wymaga wobec tego utrzymywania hipotezy w stanie odwracalnym. W języku nauki oznacza to falsyfikowalność, a w praktyce klinicznej – gotowość do aktualizacji przekonań po pojawieniu się lepszych danych. Świadomy pacjent-prosument to nie ten, który przychodzi z trafną samodiagnozą, lecz ten, który potrafi powiedzieć: „to jest moja obecna hipoteza, ale mogę ją zmienić”. Analogiczna powinność spoczywa na lekarzu. Historia medycyny zna przypadki diagnostycznego zamknięcia po stronie profesjonalnej: anchoring, premature closure czy diagnostic momentum mogą sprawić, że pierwsza etykieta utrzymuje się mimo danych ją podważających. Kalibracja nie jest więc cnotą pacjentów wobec ekspertów, lecz symetrycznym wymogiem dla wszystkich uczestników procesu poznawczego.

W tym miejscu można połączyć epistemic humility z bayesowskim modelem aktualizacji przekonań. Prawidłowe rozumowanie nie polega na braku hipotez, lecz na zmianie stopnia przekonania proporcjonalnie do nowego dowodu. Silny dowód powinien znacząco przesunąć prawdopodobieństwo, słaby – jedynie nieznacznie. Najczęstszy błąd cyfrowej kultury zdrowotnej polega właśnie na błędnym kalibrowaniu siły aktualizacji: anegdota jest traktowana jak metaanaliza, pojedynczy wynik jak wykluczenie wszystkich chorób, a korelacja jak mechanizm przyczynowy. Sprawczość kalibrowana jest więc praktyczną epistemologią Bayesa bez konieczności wykonywania rachunku. Człowiek powinien odróżniać dowód, który „trochę zwiększa prawdopodobieństwo”, od takiego, który „zasadniczo zmienia decyzję”. Taka kompetencja jest cenniejsza niż zapamiętywanie kolejnych faktów medycznych.

Kolejnym elementem musi być kalibracja interwencji do odwracalności. Decyzje łatwo odwracalne mogą uzasadniać niższy próg dowodu – jeśli zmiana jest tania i bezpieczna, można eksperymentować przy umiarkowanej niepewności. Jeśli jednak ingerencja wiąże się z operacją, długotrwałą farmakoterapią czy ryzykiem utraty płodności, próg ten powinien być znacznie wyższy. Jest to zgodne z teorią option value: zachowanie możliwości późniejszego wyboru ma wartość w warunkach niepewności. Dlatego aktywna obserwacja może być racjonalnym postępowaniem medycznym, a nie biernością. Nie podejmując natychmiast nieodwracalnej interwencji, pacjent

kupuje czas na pojawienie się nowych informacji. Kultura konsumencka ma jednak trudność z uznaniem obserwacji za działanie; rynek wynagradza procedurę i produkt, a znacznie trudniej monetyzować decyzję o wstrzymaniu się z działaniem. Tymczasem z punktu widzenia teorii decyzji powstrzymanie działania może być aktywnym wyborem maksymalizującym oczekiwaną korzyść. Odpowiedzialna medycyna potrzebuje zatem społecznego prestiżu dla powściągliwości.

To prowadzi do kolejnego wymiaru sprawczości: prawo do nieoptymalizowania. Jeśli autonomia ma być realna, człowiek musi mieć możliwość stwierdzenia, że nie chce monitorować każdego parametru, znać każdego ryzyka genetycznego czy wydłużać życia za wszelką cenę. Nie można uznać każdej odmowy technologicznej optymalizacji za brak odpowiedzialności.

Sprawczość kalibrowana jako ochrona pluralizmu wartości przed dyktatem optymalizacji

John Stuart Mill bronił indywidualnej swobody między innymi dlatego, że różne koncepcje dobrego życia posiadają wartość eksperymentalną i moralną. Współczesna medycyna powinna zachować podobny pluralizm. Nie każdy człowiek musi definiować zdrowie jako maksymalizację VO₂max, biologicznej młodości czy długości życia. Dla jednej osoby priorytetem może być czas z rodziną, dla innej możliwość podróżowania, twórczość, autonomia, seksualność lub uniknięcie agresywnego leczenia w terminalnej fazie życia. Sprawczość kalibrowana nie oznacza zatem kalibracji człowieka do medycznego ideału, lecz kalibrację medycyny do racjonalnie wybranych celów człowieka. To zasadnicza różnica: pierwszy model jest technokratyczny, drugi zaś humanistyczny.

W tym miejscu warto odwołać się do capabilities approach Sena i Nussbauma. Zdrowie jest istotne nie dlatego, że każdy parametr biologiczny ma samoistną wartość moralną, lecz ponieważ warunkuje zdolność człowieka do działania, budowania relacji i realizowania wybranych sposobów życia. Medycyna staje się więc środkiem zwiększającym ludzkie możliwości, a nie przedsiębiorstwem maksymalizującym wskaźniki biologiczne. Taka perspektywa pozwala również na nowo odczytać myśl Canguilhema.

Zdrowy człowiek to nie ten, który jest doskonale zgodny z zewnętrzną normą, lecz ten, który potrafi tworzyć nowe normy w zmieniającym się środowisku. Sprawczość kalibrowana stanowi epistemiczny odpowiednik tej biologicznej normatywności: człowiek nie realizuje jednej instrukcji, lecz dostosowuje sposób działania do sytuacji, toleruje zmienność i aktualizuje własne strategie bez utraty podmiotowości. Model ten wyznacza również granicę ideologii longevity. Człowiek może racjonalnie dążyć do wydłużenia healthspan, ale musi zachować zdolność stwierdzenia, że pewna

dodatkowa redukcja ryzyka nie jest warta kosztu, dyscypliny lub działań niepożądanych. Nie każda dodatkowa minuta oczekiwanego życia ma nieskończoną cenę. Ekonomia zdrowia od dawna uwzględnia koszt alternatywny; podobną prawdę powinien uznać człowiek w swoim osobistym życiu.

Sprawczość kalibrowana wymaga zatem ochrony przed dwoma rodzajami przemocy. Pierwsza jest paternalistyczna: „ekspert wie, czego powinienesz chcieć”. Druga jest optymalizacyjna: „skoro możesz zmniejszyć ryzyko, powinienesz to zrobić”. Obie odbierają człowiekowi prawo do wyznaczenia granicy medycznej ingerencji. Istnieje jednak również trzecia odmiana: przemoc rynkowa ubrana w język wyboru.

„Nikt cię nie zmusza, ale każdy odpowiedzialny człowiek bada to, monitoruje tamto, przyjmuje suplementy, sprawdza hormony i zna swój biologiczny wiek”. Formalnie wybór pozostaje wolny, lecz społecznie odmowa zaczyna przypominać zaniedbanie. Sprawczość kalibrowana musi więc zawierać także odporność na normatywną presję opakowaną jako możliwość konsumencka. Wymaga to instytucjonalnego wsparcia autonomii. Koncepcja relational autonomy przypomina, że człowiek staje się zdolny do samostanowienia dzięki odpowiednim relacjom i warunkom społecznym, a nie mimo nich. Państwo, lekarz, szkoła czy technologia mogą zatem zwiększać autonomię poprzez tworzenie warunków do rozumnego wyboru. Ochrona podmiotowości nie wymaga od państwa całkowitej bierności. Może to wymagać ograniczenia dark patterns, zwiększenia transparentności reklamy, ochrony danych oraz wprowadzenia standardów informacji o ryzyku i zapewnienia możliwości profesjonalnej konsultacji.

Sprawczość jako dynamiczny i rozproszony podział odpowiedzialności

Podobnie health literacy nie powinno być traktowane wyłącznie jako indywidualna powinność. Jeśli system jest tak skomplikowany, że jedynie wysoko wykształcony użytkownik potrafi z niego bezpiecznie korzystać, problem leży w samym projekcie systemu. Dobry system zdrowotny nie wymaga heroizmu poznawczego pacjenta.

W ten sposób dochodzimy do różnicy między empowerment a abandonment. Oba zjawiska mogą wyglądać podobnie: „pacjent sam podejmuje decyzję”. W pierwszym przypadku otrzymał on rzetelną informację, wsparcie, czas i realną możliwość konsultacji. W drugim system przeniósł na niego ciężar rozstrzygnięcia, którego profesjonalista nie chciał lub nie miał czasu udźwignąć. Autonomia nie może być pretekstem do instytucjonalnej abdykacji.

Podobny problem dotyczy cyfrowej samoobsługi. Aplikacja może umożliwić pacjentowi samodzielne działanie i uwolnić go od zbędnej zależności od instytucji, ale może też stać się sposobem na przeniesienie pracy administracyjnej, diagnostycznej lub interpretacyjnej na użytkownika. Różnica między empowerment a outsourcingiem odpowiedzialności zależy od tego, czy człowiek otrzymał narzędzia wystarczające do bezpiecznego przejęcia zadania. Właśnie dlatego przyszła medycyna potrzebuje nie maksymalnej autonomii, lecz optymalnego rozdziału odpowiedzialności. Część decyzji najlepiej pozostawić pacjentowi, część ekspertowi, a część może zostać wsparta algorytmem. Kluczowe jest, aby granice tego podziału były jawne, uzasadnione i elastyczne w zależności od sytuacji. W warunkach dużej niepewności ekspert powinien raczej pomagać myśleć niż narzucać rozwiązanie; w sytuacjach wysokiego ryzyka i silnych dowodów może być bardziej dyrektywny.

Pacjent powinien mieć możliwość regulowania własnego udziału zgodnie ze swoimi preferencjami. System AI z kolei powinien ujawniać, kiedy jego rekomendacja wykracza poza zakres mocnej walidacji. Sprawczość staje się więc dynamicznym podziałem kompetencji, a nie stałym prawem jednej strony do ostatecznego epistemicznego pierwszeństwa. Model ten można porównać z koncepcją distributed cognition. Edwin Hutchins i rozwijana później teoria poznania rozproszonego pokazują, że proces poznawczy nie musi zachodzić wyłącznie wewnątrz pojedynczego mózgu – może być rozłożony między ludzi, artefakty, procedury i środowisko.

Nowoczesna decyzja medyczna właśnie taka jest: pacjent wnosi swoje doświadczenia, lekarz rozumie kliniczny wzorzec, laboratorium dostarcza pomiar, baza wiedzy przechowuje badania, a algorytm oblicza ryzyko. Inteligencja procesu wyłania się z relacji pomiędzy tymi elementami. Z tej perspektywy pytanie „kto podejmuje decyzję?” okazuje się zbyt uproszczone. Ostateczny akt zgody należy do człowieka, ale poznawcze przygotowanie decyzji jest kolektywne i infrastrukturalne. Dlatego błędu nie zawsze można przypisać jednej osobie; może on wynikać z wadliwej architektury całego systemu. Prowadzi to do teorii Swiss cheese model Jamesa Reasona: w bezpieczeństwie złożonych systemów katastrofa rzadko jest wynikiem jednego uchybienia. Powstaje ona wtedy, gdy luki w wielu warstwach zabezpieczeń ustawia się w jednej linii.

Systemowa odporność i rozproszona czujność jako gwaranty autonomii

W medycynie cyfrowej podobna logika może przybrać następującą formę: mylący film tworzy błędną hipotezę, komercyjny test generuje niejednoznaczny wynik, algorytm wyszukiwania wzmacnia tę interpretację, telekonsultacja okazuje się zbyt krótka, system receptowy nie wykrywa

problemu, a pacjent rozpoczyna niepotrzebną terapię. Żaden z tych elementów w pojedynkę nie musiałby doprowadzić do szkody – tworzy ją ich sprzężenie. Sprawczość kalibrowana musi więc być projektowana nie tylko na poziomie psychologii człowieka, ale na poziomie **safety engineering całego ekosystemu zdrowia**. Nie wystarczy edukować pacjenta, by „uważał w internecie”. Konieczne jest zaprojektowanie warstw zabezpieczających: profesjonalnej weryfikacji, walidacji narzędzi, monitoringu działań niepożądanych, standardów reklamy, interoperacyjności dokumentacji oraz możliwości łatwego uzyskania drugiej opinii. Podobnie AI nie powinna być jedynym filtrem przed błędem innego AI. System odporny wymaga heterogeniczności mechanizmów kontroli; jeśli wszystkie elementy korzystają z tego samego modelu, tej samej bazy danych lub tych samych założeń, błąd może zostać skorelowany zamiast skorygowany. Redundancja ma wartość tylko wtedy, gdy jej komponenty nie popełniają identycznych błędów.

To prowadzi do koncepcji **High Reliability Organizations**. Badania organizacji działających w warunkach wysokiego ryzyka podkreślają między innymi nieufność wobec zbyt prostych wyjaśnień, wrażliwość na słabe sygnały, gotowość do aktualizacji wiedzy oraz respekt dla ekspertyzy niezależnie od formalnej hierarchii. Taka kultura stanowi cenny wzór również dla przyszłej medycyny algorytmicznej. System powinien być projektowany nie wokół założenia, że „algorytm się nie myli” lub „lekarz zawsze ma ostatnie słowo”, lecz wokół pytania: **kto w tej konkretnej sytuacji posiada najlepszą informację pozwalającą wykryć błąd?**

Czasami będzie to radiolog. Czasami pielęgniarka obserwująca pacjenta. Czasami sam pacjent, który zauważy, że opis algorytmu nie pasuje do jego doświadczenia. A czasami system analizujący tysiące podobnych przypadków.

Epistemiczny konstytucjonalizm może zatem zostać rozwinięty w model **rozproszonej czujności**. Najważniejszą kompetencją całego systemu staje się wtedy nie nieomyślność, lecz **zdolność szybkiego zauważania i korygowania własnych błędów**. Jest to właściwość bliższa nauce niż dogmatowi. Nauka jest skuteczna nie dlatego, że nigdy się nie myli, lecz dlatego, że tworzy instytucjonalne mechanizmy umożliwiające poprawianie twierdzeń. To samo powinno dotyczyć medycyny spersonalizowanej. Decyzja może być w danym momencie najlepsza na podstawie dostępnych informacji i zostać później zmieniona. Taka zmiana nie musi oznaczać wcześniejszej niekompetencji; może świadczyć o tym, że system zadziałał prawidłowo, aktualizując ocenę w świetle nowych danych. Właśnie takie rozumienie niepewności jest niezbędne w kulturze utożsamiającej zmianę opinii z utratą wiarygodności. W praktyce zdolność do zmiany stanowiska proporcjonalnie do napływających danych jest jednym z najlepszych znaków kalibracji epistemicznej.

Sprawczość kalibrowana może więc zostać potraktowana jako **antropologiczny odpowiednik evidence-based medicine**. EBM łączy najlepsze dostępne dowody, doświadczenie kliniczne i wartości pacjenta. Sprawczość kalibrowana dodaje czwarty wymiar: świadomość sposobu, w jaki środowisko informacyjne i algorytmiczne kształtuje samą percepcję możliwych wyborów. Dopiero wszystkie cztery elementy pozwalają zachować rzeczywistą autonomię. Dowody bez wartości prowadzą do technokracji. Wartości bez dowodów mogą prowadzić do szkodliwej arbitralności. Ekspertyza bez pokory tworzy paternalizm. Dane bez interpretacji generują chaos. Algorytm zaś, pozbawiony jawnej funkcji celu, może zamienić rekomendację w niewidzialne sterowanie.

Sprawczość kalibrowana jako nowa zasada organizacji medycyny

Model ten pozwala powrócić do jednego z początkowych paradoksów całej analizy: współczesny człowiek ma dostęp do większej ilości wiedzy zdrowotnej niż jakiegokolwiek poprzednie pokolenie, a jednocześnie może odczuwać większą niepewność co do własnego stanu zdrowia. Nie jest to sprzeczność. Im większa rozdzielczość obserwacji, tym więcej możliwych odchyleń. Im więcej kategorii diagnostycznych, tym szerszy wachlarz interpretacji doświadczenia. Im więcej modeli predykcyjnych, tym więcej przyszłych zagrożeń staje się widzialnych.

Rozwiązaniem nie może być powrót do niewiedzy, lecz nauczanie się życia z wiedzą, która nie daje absolutnej pewności. To prawdopodobnie najważniejsza kompetencja zdrowotna XXI wieku. Sprawczość kalibrowana nie jest zatem kompromisem pomiędzy dawnym modelem relacji lekarz-pacjent a nową erą internetu. Stanowi ona nową epistemiczną zasadę organizacji medycyny: człowiek powinien wiedzieć tyle, ile pozwala mu podejmować lepsze decyzje; profesjonalista powinien interweniować tak intensywnie, jak uzasadnia to prawdopodobieństwo korzyści; technologia powinna redukować złożoność bez ukrywania istotnej niepewności; rynek nie może być jedynym projektantem architektury wyboru, a państwo powinno chronić zdolność obywatela do sprzeciwu, uzyskania drugiej opinii i wyjścia z systemu. Wówczas lekarz przestaje być strażnikiem prawdy, pacjent przestaje być samotnym konsumentem, a algorytm przestaje być niewidzialnym suwerenem. Każdy staje się częścią systemu, którego celem nie jest osiągnięcie biologicznej doskonałości, lecz zwiększenie zdolności człowieka do rozsądnego działania w warunkach nieusuwalnej niepewności.

Pozostaje jednak ostatni problem. Sprawczość jest kategorią normatywną, podczas gdy medycyna funkcjonuje w świecie interesów ekonomicznych, instytucji, regulacji i konkurujących modeli biznesowych. Nie wystarczy zatem zdefiniować, jaki powinien być dobry pacjent i lekarz. Trzeba

zapytać, jaką architekturę instytucjonalną należy zbudować, aby rozsądne zachowanie było możliwe i opłacalne dla całego systemu, a nie tylko moralnie wymagane od jednostki. To otwiera kolejny etap rozważań: przejście od sprawczości kalibrowanej ku instytucjom kalibrującym. Wymaga to syntezy ekonomii informacji, teorii regulacji, oceny technologii medycznych oraz analizy odpowiedzialności platform i zasady ostrożności. Dopiero wtedy można odpowiedzieć na pytanie, czy społeczeństwo jest w stanie chronić człowieka przed dwoma symetrycznymi błędami nowoczesnej medycyny: przed pozostawieniem go bez potrzebnej innowacji oraz przed sprzedaniem mu rozwiązania, zanim ktokolwiek wie, czy jest ono rzeczywiście potrzebne.

Czy w świecie, w którym algorytmy potrafią przewidzieć nasze potrzeby szybciej niż my sami je poczujemy, jesteśmy jeszcze autorami własnych decyzji, czy jedynie pasażerami w starannie zaprojektowanym procesie optymalizacji? Prawdziwa suwerenność medyczna może dziś polegać nie na posiadaniu wszystkich odpowiedzi, lecz na odwadze zachowania prawa do bycia nieoptymalnym. Ostatecznie to właśnie w szczelinach między predykcją a rzeczywistym przeżyciem biologicznego istnienia kryje się to, co w nas niezbywalnie ludzkie.

Inspiracje

[1]



"Bad Influence" Deborah Cohen

Simon and Schuster | ISBN: 9780861549894

Dr **Deborah Cohen** jest brytyjską dziennikarką śledczą, prezenterką radiową i lekarzem z wykształceniem medycznym, która pełni funkcję Visiting Senior Fellow w LSE Health w London School of Economics. Po uzyskaniu dyplomu z medycyny, Cohen przeszła do dziennikarstwa śledczego, tworząc i kierując działem śledczym w The BMJ przez ponad dekadę. Następnie była korespondentką ds. zdrowia i Wielkiej Brytanii w BBC Newsnight – kierując relacjami z COVID-19 – oraz redaktorką naukową w ITV News. Jej praca śledcza w mediach i prasie ujawniła systemowe problemy w badaniach klinicznych, regulacjach dotyczących wyrobów medycznych, bezpieczeństwie leków i konfliktach interesów w opiece zdrowotnej, przyczyniając się do reform regulacyjnych i dochodzeń parlamentarnych na całym świecie. Relacjonowała wydarzenia dla ważnych programów, takich jak BBC Panorama i Channel 4 Dispatches. Jest autorką książki „Bad Influence: How the Internet Hijacked Our Health”, która znalazła się na krótkiej liście nominacji do Nagrody Królewskiego Towarzystwa Naukowego za Książkę.

Dr. Deborah Cohen is a British investigative journalist, broadcaster, and medically trained physician who serves as a Visiting Senior Fellow at LSE Health at the London School of Economics. After qualifying in medicine, Cohen transitioned into investigative reporting, establishing and leading the investigations unit at The BMJ for over a decade. She subsequently served as a health and UK correspondent for BBC Newsnight—leading its COVID-19 coverage—and as Science Editor for ITV News. Her investigative work across broadcast and print has uncovered systemic issues in clinical trials, medical device regulation, drug safety, and healthcare conflicts of interest, contributing to regulatory reforms and parliamentary inquiries globally. She has reported for major programs including BBC Panorama and Channel 4 Dispatches, and is the author of Bad Influence: How the Internet Hijacked Our Health, shortlisted for the Royal Society Science Book Prize.

London School of Economics and Political Science

Literatura uzupełniająca

Książki

Książki autora i dzieła przywołane w tekście:



[1] "Last Call at the Hotel Imperial"

Deborah Cohen

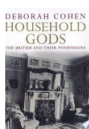
2022 | Random House | ISBN: 9780525511205



[2] "Family Secrets"

Deborah Cohen

2013 | Oxford University Press | ISBN: 9780199985630



[3] "Household Gods"

Deborah Cohen

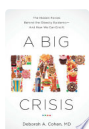
2006 | Yale University Press | ISBN: 9780300112139



[4] "The War Come Home"

Deborah Cohen

2001 | Univ of California Press | ISBN: 9780520220089



[5] "A Big Fat Crisis"

Deborah Cohen

2013 | Hachette UK | ISBN: 9781568589657



[6] "Lilith's Ark"

Deborah Bodin Cohen

2010 | Jewish Publication Society | ISBN: 9780827610248

Książki autorów przywoływanych i literatura pokrewna:



[7] "Action Theory and the Human Condition"

Talcott Parsons

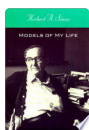
1978



[8] "Organizations"

Herbert Simon

1993 | John Wiley & Sons | ISBN: 9780631186311



[9] "Models of my life"

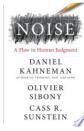
Herbert Simon

1996 | MIT Press | ISBN: 9780262691857



[10] "Human Problem Solving"

Herbert Simon
Prentice Hall



[11] "Noise: A Flaw in Human Judgment"

Daniel Kahneman
Hachette UK | ISBN: 9780316451383

[12] "The framing of decisions and the psychology of choice"

Daniel Kahneman



[13] "Reference Theory of Choice and Exchange"

Amos Tversky, Daniel Kahneman
1990

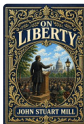
[14] "prospect theory"

Amos Tversky
1979



[15] "Risk Savvy: How to Make Good Decisions"

Gerd Gigerenzer
ISBN: 9781846144745



[16] "On Liberty (Golden Monarch Collector's Edition) (Case Laminate Hardcover with Jacket)"

JOHN STUART. MILL
2026 | Golden Monarch | ISBN: 9781834125282



[17] "Health Systems Science E-Book"

Susan E. Skochelak
2020 | Elsevier Health Sciences | ISBN: 9780323694674



[18] "Harrison's Principles of Internal Medicine 20/E (Vol.1 & Vol.2) (ebook)"

Dennis L. Kasper, Anthony S. Fauci, Stephen L. Hauser, Dan L. Longo, J. Larry Jameson, Joseph Loscalzo
2018 | McGraw Hill Professional | ISBN: 9781259644047



[19] "Entrustable Professional Activities and Entrustment Decision-Making in Health Professions Education"

Olle ten Cate, Vanessa C. Burch, H. Carrie Chen, Fremien Chihchen Chou, Marije P. Hennis

2024 | Ubiquity Press | ISBN: 9781914481611



[20] "Harrison's Principles of Internal Medicine, 18th Edition"

Dan Longo, Anthony Fauci, Dennis Kasper, Stephen Hauser, J. Jameson, Joseph Loscalzo

2011 | McGraw Hill Professional | ISBN: 9780071748902



[21] "Hospital Organization Research"

Basil Spyros Georgopoulos

1975 | Saunders Limited. | ISBN: 9780721641058

Artykuły naukowe

Prace autorów przywoływanych w tekście:

[1] "Human Problem Solving."

Nick Axten, Allen Newell, Herbert A. Simon

1973 | Contemporary Sociology A Journal of Reviews | DOI: 10.2307/2063712

[Źródło](#)

[2] "Organizations as Systems"

Talcott Parsons

2013 | Routledge eBooks | DOI: 10.4324/9780203547786-13

[Źródło](#)

[3] "The Structure of Social Action : A Study in Social Theory with Special Reference to a Group of Recent European Writers/ By Talcott Parsons"

T. Parsons

2017 | DOI: 10.2307/2084931

[Źródło](#)

[4] "Theories of society : foundations of modern sociological theory"

T. Parsons

2017 | DOI: 10.2307/2089684

[Źródło](#)

[5] "Reflexive Modernization: Politics, Tradition, and Aesthetics in the Modern Social Order."

Alan Sica, Ulrich Beck, Anthony Giddens, Scott M. Lash

1997 | Social Forces | DOI: 10.2307/2580534

[Źródło](#)

[6] "Agency, Structure"

Anthony Giddens

1979 | Central Problems in Social Theory | DOI: 10.1007/978-1-349-16161-4_3

[Źródło](#)

[7] "Conversations with Anthony Giddens: Making Sense of Modernity"

Mark Neocleous, Anthony Giddens, Christopher Pierson

2000 | Contemporary Sociology | DOI: 10.2307/2654463

[Źródło](#)

[8] "Reforming Global Governance: Selected essentials from the 7th Global China Dialogue Forum"

A. Giddens, Dan Chugg, Xiao-Guang Yang, Martin Albrow, Kaiping Peng

2022 | Journal of China in Global and Comparative Perspectives | DOI: 10.24103/jcgcp.en.2022.6

[Źródło](#)

[9] "A BEHAVIORAL MODEL OF RATIONAL CHOICE"

Herbert A. Simon

1955 | The Quarterly Journal of Economics | DOI: 10.2307/1884852

[Źródło](#)

[10] "Administrative Behavior: A Study of Decision-Making Processes in Administrative Organization."

Michael Brown, Herbert A. Simon

1978 | Contemporary Sociology A Journal of Reviews | DOI: 10.2307/2065048

[Źródło](#)

[11] "Verbal reports as data."

K. Anders Ericsson, Herbert A. Simon

1980 | Psychological Review | DOI: 10.1037//0033-295x.87.3.215

[Źródło](#)

[12] "Prospect Theory: An Analysis of Decision under Risk"

Daniel J. Kahneman, Amos Tversky

1979 | Econometrica | DOI: 10.2307/1914185

[Źródło](#)

[13] "The Framing of Decisions and the Psychology of Choice"

Amos Tversky, Daniel J. Kahneman

1981 | Science | DOI: 10.1126/science.7455683

[Źródło](#)

[14] "Choices, values, and frames."

Daniel J. Kahneman, Amos Tversky

1984 | American Psychologist | DOI: 10.1037/0003-066x.39.4.341

[Źródło](#)

[15] "Loss Aversion in Riskless Choice: A Reference-Dependent Model"

Amos Tversky, Daniel J. Kahneman

1991 | The Quarterly Journal of Economics | DOI: 10.2307/2937956

[Źródło](#)

[16] "Evidence-Based Information Halves the Influence of Physicians' Non-Evidence-Based Cancer Screening Recommendations v1"

Odette Wegwarth, not provided Gert G. Wagner, not provided Gerd Gigerenzer

2017 | DOI: 10.17504/protocols.io.imscc6e

[Źródło](#)

[17] "How I Got Started: Teaching Physicians and Judges Risk Literacy"

Gerd Gigerenzer

2013 | Applied Cognitive Psychology | DOI: 10.1002/acp.2980

[Źródło](#)

[18] "Probabilistic Mental Models"

Gerd Gigerenzer

2002 | Adaptive Thinking | DOI: 10.1093/acprof:oso/9780195153729.003.0007

[Źródło](#)

[19] "The Criterion of a Good Form of Government"

John Stuart Mill

2015 | On Liberty, Utilitarianism and Other Essays | DOI: 10.1093/owc/9780199670802.003.0016

[Źródło](#)

[20] "Of the Limits to the Authority of Society over the Individual"

John Stuart Mill

2015 | On Liberty, Utilitarianism and Other Essays | DOI: 10.1093/owc/9780199670802.003.0005

[Źródło](#)

[21] "Principles of Political Economy"

F. Y. Edgeworth, J. Mill, W. Ashley

2020 | Forerunners of Realizable Values Accounting in Financial Reporting | DOI: 10.4324/9781351291521-15

[Źródło](#)

[22] "Direct Manipulation Interfaces"

Edwin L. Hutchins, James D. Hollan, Donald A. Norman

1985 | Human-Computer Interaction | DOI: 10.1207/s15327051hci0104_2

[Źródło](#)

[23] "Distributed cognition in an airline cockpit"

Edwin L. Hutchins, Tove Klausen

1996 | Cambridge University Press eBooks | DOI: 10.1017/cb09781139174077.002

[Źródło](#)

[24] "The social organization of distributed cognition."

Edwin L. Hutchins

1991 | American Psychological Association eBooks | DOI: 10.1037/10096-012

[Źródło](#)

[25] "Human error: models and management"

JAMES T. REASON

2000 | BMJ | DOI: 10.1136/bmj.320.7237.768

[Źródło](#)

[26] "Human Error"

JAMES T. REASON

1990 | Cambridge University Press eBooks | DOI: 10.1017/cbo9781139062367

[Źródło](#)

[27] "Managing the Risks of Organizational Accidents"

JAMES T. REASON

2016 | Routledge eBooks | DOI: 10.4324/9781315543543

[Źródło](#)

[28] "Understanding adverse events: human factors."

JAMES T. REASON

1995 | BMJ Quality & Safety | DOI: 10.1136/qshc.4.2.80

[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: 32offof6-cb47-4ee7-a1a6-393128b613f7