Anatomia szkód algorytmicznych i odpowiedzialność systemowa w świetle książki Raising AI Chiraga Shaha

AUTOR

Fundacja Dobre Państwo

STRESZCZENIE / ABSTRACT

Artykuł analizuje mechanizmy powstawania szkód algorytmicznych, wskazując, że nie wynikają one wyłącznie z błędów matematycznych modelu, lecz z jego relacji z instytucją i procesem decyzyjnym. Autor wyróżnia pięć kategorii szkód: techniczną, ekonomiczną, prawną, społeczną oraz epistemiczną, podkreślając ryzyko ograniczenia zdolności człowieka do kwestionowania decyzji AI. Tekst odróżnia prostą automatyzację od zaawansowanego uczenia maszynowego, zauważając, że złożoność modeli często utrudnia identyfikację odpowiedzialności. Kluczowym wnioskiem jest konieczność postrzegania AI jako elementu systemów socjotechnicznych, gdzie bezpieczeństwo zależy od współdziałania ludzi, procedur i technologii. Podejście to pozwala na lepsze zrozumienie łańcucha przyczynowego prowadzącego do naruszeń praw jednostki w administracji, medycynie czy finansach.

The article analyzes the mechanisms through which algorithmic harms arise, indicating that they do not result solely from mathematical errors of the model, but from its relationship with the institution and the decision-making process. The author distinguishes five categories of harm: technical, economic, legal, social, and epistemic, emphasizing the risk of limiting a human's ability to challenge AI decisions. The text differentiates simple automation from advanced machine learning, noting that the complexity of models often hinders the identification of responsibility. The key conclusion is the necessity of viewing AI as an element of socio-technical systems, where safety depends on the interaction between people, procedures, and technology. This approach allows for a better understanding of the causal chain leading to violations of individual rights in administration, medicine, or finance.

Artykuł stawia tezę, że rzeczywiste szkody algorytmiczne nie są jedynie wynikiem błędów technicznych czy stronniczości samych modeli AI, lecz efektem wadliwych systemów socjotechnicznych. Autor argumentuje, że krzywda powstaje w momencie przejścia od matematycznej predykcji do działania instytucjonalnego, gdy organizacje bezkrytycznie wdrażają automatyzację, rozpraszają odpowiedzialność i rezygnują z realnego nadzoru ludzkiego. Analizując przykłady z medycyny, wymiaru sprawiedliwości, finansów oraz administracji publicznej (m.in. afery Robodebt), tekst wykazuje, że technologia często służy jako parawan dla systemowej nieefektywności lub przemocy strukturalnej. Kluczowym problemem okazuje się zatem nie tyle zawodność maszyn, co erozja instytucjonalnej zdolności do krytycznej refleksji nad własnymi celami i etyką władzy. Tekst wzywa do przejścia od iluzorycznego modelu 'człowieka w pętli' ku rzeczywistej odpowiedzialności organizacyjnej, gdzie przejrzystość algorytmiczna i zasada proporcjonalności stają się gwarantem ochrony podmiotowości człowieka przed redukcją do obiektu optymalizacji.

SŁOWA KLUCZOWE

szkoda algorytmiczna | odpowiedzialność systemowa | systemy socjotechniczne | szkoda epistemiczna | automatyzacyjne skrzywienie poznawcze | alert fatigue | przejrzystość algorytmiczna | walidacja kliniczna | zautomatyzowane podejmowanie decyzji | biometria i prawa człowieka | metoda STAMP | interpretowalność modeli AI | bezpieczeństwo systemowe | dyskryminacja algorytmiczna

Szkoda algorytmiczna powstaje w relacji modelu z instytucją

Anatomia realnych szkód. Odpowiedzialność algorytmiczna w medycynie, wymiarze sprawiedliwości, finansach, zatrudnieniu i administracji publicznej

Przejście od matematycznej analizy stronniczości do rozpoznania rzeczywistych szkód społecznych wymaga zasadniczej zmiany perspektywy. Model statystyczny, rozpatrywany wyłącznie jako

konstrukcja matematyczna, pozostaje narzędziem służącym do szacowania zależności, porządkowania informacji i przewidywania zdarzeń.

Dopiero włączenie go w instytucjonalny proces podejmowania decyzji sprawia, że wynik obliczenia może stać się realną ingerencją w sytuację prawną, ekonomiczną lub egzystencjalną konkretnego człowieka. Prognoza prawdopodobieństwa niewypłacalności zamienia się w odmowę kredytu, klasyfikacja ryzyka recydywy wpływa na ocenę sytuacji oskarżonego, automatyczna identyfikacja twarzy może zapoczątkować czynności organów ścigania, a wynik analizy dokumentacji medycznej zmienić priorytet udzielenia świadczenia zdrowotnego. Właśnie ten moment przejścia od reprezentacji informacyjnej do działania instytucjonalnego stanowi podstawowy przedmiot badań nad odpowiedzialnością algorytmiczną. Niesprawiedliwość nie musi bowiem tkwić wyłącznie wewnątrz modelu; może ona powstawać w relacji między predykcją, procedurą organizacyjną, obowiązującą normą prawną a społecznym położeniem osoby, wobec której wynik zostaje wykorzystany.

Z perspektywy teorii odpowiedzialności autor książki odróżnia szkodę techniczną, ekonomiczną, prawną, społeczną i epistemiczną, gdyż każda z tych kategorii ujawnia inną postać naruszenia ludzkich interesów. Szkoda techniczna może polegać na błędnym funkcjonowaniu systemu, utracie danych lub wygenerowaniu nieprawidłowej informacji, choć nie każdy taki incydent prowadzi bezpośrednio do wymiernej krzywdy człowieka. Szkoda ekonomiczna obejmuje utracone dochody, wzrost kosztów, ograniczenie dostępu do rynku lub nieuzasadnione obciążenia finansowe.

Szkoda prawna dotyczy naruszenia uprawnień, obowiązków proceduralnych i gwarancji ochrony jednostki, natomiast szkoda społeczna może przyjmować postać stygmatyzacji, wykluczenia, reprodukcji stereotypów oraz osłabienia zaufania do instytucji. Szczególnego znaczenia nabiera szkoda epistemiczna, związana z ograniczeniem zdolności człowieka do poznania, zrozumienia i zakwestionowania przesłanek decyzji go dotyczącej. W rzeczywistych zastosowaniach AI wszystkie te formy mogą współwystępować, tworząc proces kumulacji negatywnych następstw, których skala nie daje się wyrazić jednym wskaźnikiem błędu predykcyjnego.

Rozpoznanie odpowiedzialności wymaga również odróżnienia sztucznej inteligencji w ścisłym znaczeniu od szerszej kategorii zautomatyzowanych systemów decyzyjnych. Nie każda procedura komputerowa wykorzystuje uczenie maszynowe, modele neuronowe czy algorytmy adaptacyjne. Część najpoważniejszych historycznych szkód spowodowały systemy oparte na prostych regułach, porównywaniu baz danych i automatycznym generowaniu rozstrzygnięć. Różnica ta jest istotna metodologicznie, ponieważ pozwala uniknąć przypisywania wszystkich patologii cyfrowej administracji szczególnym właściwościom AI. Jednocześnie analiza nieprawidłowości w tradycyjnej automatyzacji ma dużą wartość prognostyczną dla współczesnych systemów uczących się.

Jeżeli bowiem nawet przejrzysta procedura arytmetyczna może doprowadzić do masowego naruszania praw człowieka, to zwiększenie złożoności modelu, liczby zmiennych i autonomii operacyjnej nie eliminuje problemu odpowiedzialności. Przeciwnie – może utrudniać jego identyfikację, o ile instytucja nie rozwinie równolegle odpowiednich metod nadzoru i kontroli.

Historyczne doświadczenia pozwalają zatem zrekonstruować zasadniczy łańcuch przyczynowy szkody algorytmicznej. Rozpoczyna się on od przyjęcia celu organizacyjnego, następnie obejmuje jego operacjonalizację w postaci kryteriów i danych, konstrukcję procedury obliczeniowej, wybór progu interwencji, zastosowanie wyniku do konkretnej osoby oraz późniejsze działania administracyjne, gospodarcze lub medyczne.

Na każdym etapie może wystąpić inny błąd, którego konsekwencje mogą zostać wzmocnione przez kolejne elementy procesu. Nieprawidłowa predykcja nie musi prowadzić do krzywdy, jeśli zostanie odpowiednio zweryfikowana przed podjęciem decyzji. Z drugiej strony poprawna predykcja może uczestniczyć w bezprawnym działaniu, jeżeli dotyczy cechy, której wykorzystywanie jest niedopuszczalne, albo zostaje zastosowana do celu naruszającego prawa jednostki.

Szkody algorytmiczne jako efekt wadliwych systemów socjotechnicznych

Właśnie dlatego odpowiedzialność za szkody technologiczne należy analizować w kategoriach systemów socjotechnicznych, których rzeczywiste właściwości wynikają ze współdziałania ludzi, organizacji, procedur i narzędzi obliczeniowych. W naukach o zarządzaniu ryzykiem użyteczna okazuje się koncepcja bezpieczeństwa systemowego rozwijana przez Nancy Leveson; jej metoda STAMP opiera się na analizie ograniczeń i mechanizmów kontroli w złożonych strukturach techniczno-organizacyjnych. O ile tradycyjne podejście utożsamia bezpieczeństwo z niezawodnością poszczególnych komponentów, o tyle analiza systemowa dowodzi, że katastrofa może nastąpić nawet wtedy, gdy żaden element nie uległ awarii w klasycznym znaczeniu.

Przyczyną może być wadliwa komunikacja, brak odpowiednich ograniczeń, sprzeczność wymagań lub niewłaściwe wykorzystanie poprawnie funkcjonującego narzędzia. Analogiczny mechanizm występuje w decyzjach algorytmicznych: model przekazuje prawidłowo obliczony wynik, lecz instytucja interpretuje go jako informację o znaczeniu, którego ten nie posiada. Odpowiedzialność technologiczna nie może zatem ograniczać się do wykazania, że oprogramowanie działało zgodnie z dokumentacją producenta. Konieczne jest rozpoznanie, czy cały proces decyzyjny zaprojektowano tak, by zapewnić ochronę właściwych dóbr. Szczególnej uwagi wymaga medycyna, gdyż

wykorzystanie sztucznej inteligencji w diagnostyce i terapii oddziałuje bezpośrednio na zdrowie, życie i integralność cielesną człowieka. Rozwój systemów wspomagających interpretację obrazów radiologicznych, analizy elektrokardiograficzne, wykrywanie zmian nowotworowych, ocenę ryzyka klinicznego czy personalizację leczenia może przynosić znaczące korzyści – zwiększać szybkość przetwarzania danych, identyfikować trudne do zauważenia wzorce i wspierać personel przy rosnącym obciążeniu pracą.

Wartość tych rozwiązań zależy jednak od jakości badań walidacyjnych, charakterystyki populacji pacjentów oraz sposobu organizacji współpracy z personelem medycznym. Szczególnie niebezpieczne jest utożsamianie korzystnych wyników retrospektywnego testu z potwierdzoną zdolnością do poprawy rezultatów leczenia. Model może osiągać wysoki wskaźnik rozpoznawania stanów klinicznych, ale nie przełożyć się na rzeczywiste ograniczenie śmiertelności, jeśli alerty są generowane zbyt późno, przeciążają lekarzy lub nie zostały zintegrowane z procesem terapeutycznym. Jednym z najbardziej pouczających przykładów problemów walidacji klinicznej jest badanie Epic Sepsis Model, opublikowane w 2021 roku przez Andrew Wonga i współautorów w czasopiśmie „JAMA Internal Medicine". Analiza objęła 38 455 hospitalizacji w systemie Michigan Medicine w okresie od grudnia 2018 do października 2019 roku.

Oceniano komercyjny model przeznaczony do przewidywania sepsy – stanu zagrażającego życiu, wymagającego szybkiej identyfikacji i leczenia. Przy analizowanym progu alarmowania system osiągnął czułość około 33%, co oznaczało niewykrycie około dwóch trzecich przypadków sepsy według przyjętej definicji badawczej. Pole pod krzywą ROC wyniosło około 0,63, co wskazuje na ograniczoną zdolność rozróżniania przypadków pozytywnych i negatywnych w badanej populacji. Jednocześnie system generował alerty dla około 18% hospitalizacji, co stwarzało ryzyko znacznego obciążenia personelu dodatkowymi czynnościami weryfikacyjnymi.

Wpływ AI na obciążenie poznawcze personelu

Wyniki te nie stanowią dowodu, że model spowodował określoną liczbę zgonów, gdyż badanie koncentrowało się na walidacji predykcyjnej, a nie na eksperymentalnej ocenie przyczynowego wpływu systemu na śmiertelność. Obnażają one jednak poważne ograniczenia bezkrytycznego polegania na narzędziu o niewystarczająco potwierdzonej skuteczności. Z perspektywy psychologii pracy i ergonomii poznawczej kluczowe staje się zjawisko zmęczenia alarmami, określane jako **alert fatigue**. W środowisku klinicznym personel jest nieustannie bombardowany komunikatami z aparatury monitorującej, systemów dokumentacyjnych i procedur bezpieczeństwa.

Jeśli znaczna część tych ostrzeżeń okazuje się nieistotna, może dojść do stopniowego obniżenia uwagi poświęcanej kolejnym alertom. Proces ten nie musi wynikać z nieodpowiedzialności personelu; częściej jest efektem ograniczonej pojemności zasobów poznawczych i konieczności hierarchizowania działań pod presją czasu. Badania nad niezawodnością organizacji wysokiego ryzyka dowodzą, że system bezpieczeństwa powinien wspierać zdolność człowieka do rozpoznawania sygnałów istotnych, zamiast mnożyć komunikaty pozbawione informacji o ich względnej wartości. W konsekwencji ocena medycznego AI wymaga nie tylko analizy czułości, swoistości i kalibracji predykcji, ale przede wszystkim badania wpływu technologii na uwagę, obciążenie pracą i rzeczywiste decyzje kliniczne.

W ekonomii zdrowia problem ten można opisać poprzez analizę kosztów alternatywnych. Każda weryfikacja alertu, dodatkowe badanie czy konsultacja angażują konkretne zasoby personelu, infrastruktury i czasu. Wynik fałszywie dodatni może prowadzić do zbędnej diagnostyki, niepotrzebnego leczenia lub wzrostu lęku u pacjenta, natomiast wynik fałszywie negatywny grozi opóźnieniem interwencji w stanie zagrożenia życia.

Optymalizacja pojedynczego wskaźnika diagnostycznego nie musi zatem minimalizować całkowitego kosztu społecznego. W praktyce decyzje powinny uwzględniać kliniczne znaczenie błędu, prawdopodobieństwo poważnych następstw, możliwość ich odwrócenia oraz dostępność alternatywnych procedur. Wymaga to syntezy badań klinicznych, statystyki medycznej, ekonomiki świadczeń i etyki zawodowej.

Model wykrywający więcej przypadków może być wartościowy mimo wzrostu liczby fałszywych alarmów, o ile jego zastosowanie realnie poprawia stan zdrowia pacjentów. Możliwa jest jednak i zależność odwrotna – gdy dodatkowe interwencje nie przynoszą korzyści proporcjonalnych do ponoszonych kosztów.

Iluzoryczna kontrola ludzka i systemowe nadużycia AI

Aby prawidłowo zinterpretować te przykłady, należy podkreślić, że odpowiedzialność kliniczna za wykorzystanie AI nie sprowadza się do formalnego wymogu zatwierdzenia wyniku przez lekarza. Jeśli personel nie posiada wiedzy o zakresie walidacji, ograniczeniach modelu, charakterystyce populacji treningowej czy konsekwencjach zastosowania wyniku, kontrola ta pozostaje powierzchowna. Sam fakt autoryzacji wyrobu medycznego do obrotu nie gwarantuje też jego nieomylności ani skuteczności wobec każdego pacjenta w każdej placówce. W Stanach Zjednoczonych Food and Drug Administration prowadzi wykaz dopuszczonych urządzeń

medycznych wykorzystujących AI; według danych tego organu we wrześniu 2026 roku obejmował on ponad 1600 pozycji.

Jest to wskaźnik skali rozwoju regulowanego segmentu technologii, a nie dowód jednolitej jakości klinicznej wszystkich rozwiązań. Odpowiedzialna praktyka wymaga monitorowania systemów po wdrożeniu, oceny zmian parametrów, dbałości o bezpieczeństwo aktualizacji oraz sprawnego reagowania na sygnały o błędach. Inny, choć równie doniosły charakter mają szkody wynikające z algorytmicznych ocen ryzyka w wymiarze sprawiedliwości i działaniach organów ścigania. W systemach prawnych kluczowe znaczenie mają wolność osobista, domniemanie niewinności, zasada proporcjonalności oraz możliwość skutecznego zakwestionowania rozstrzygnięcia. Ocena ryzyka może dostarczać informacji pomocniczych, lecz nie powinna zastępować prawnego ustalania odpowiedzialności za konkretny czyn ani dowodzenia przesłanek dopuszczalnej ingerencji w prawa jednostki.

Spór dotyczący systemu COMPAS, analizowany wcześniej pod kątem statystycznym, ujawnia tutaj głębszy problem. Nie chodzi wyłącznie o rozkład fałszywych alarmów, lecz o ryzyko nadania wynikowi modelu znaczenia przekraczającego jego właściwości poznawcze. Szacowanie przyszłego ryzyka nie jest dowodem popełnienia przestępstwa, a prawdopodobieństwo statystyczne przypisane jednostce nie stanowi jej indywidualnej winy. Utożsamienie tych porządków prowadziłoby do zatarcia różnicy między oceną konkretnego zachowania a prognozą przyszłych zdarzeń. Rzeczywiste skutki takiego pomieszania obrazują sprawy błędnej identyfikacji biometrycznej. W styczniu 2020 roku Robert Williams został niesłusznie zatrzymany przez policję w Detroit na podstawie nieprawidłowego dopasowania z systemu rozpoznawania twarzy. Materiał dowodowy pochodził z nagrania monitoringu sklepowego, którego jakość i warunki rejestracji ograniczały wiarygodność identyfikacji. Zamiast potraktować wynik jako wskazówkę wymagającą weryfikacji, organy ścigania wykorzystały go bezpośrednio do zatrzymania człowieka. Williams spędził w areszcie około trzydziestu godzin. Ugoda zawarta w czerwcu 2024 roku doprowadziła m.in. do wprowadzenia zasad wymagających posiadania niezależnych, wiarygodnych dowodów przed dokonaniem zatrzymania na podstawie technologii rozpoznawania twarzy.

Istotne jest, że szkoda nie wynikała wyłącznie z błędnej predykcji, lecz z niewłaściwego połączenia wyniku technologicznego z procedurą dowodową oraz braku odpowiednich ograniczeń instytucjonalnych. Przypadek Williamsa można interpretować przez pryzmat teorii sprawiedliwości proceduralnej Toma R. Tylera, według której legitymacja instytucji zależy nie tylko od rezultatów ich działań, ale i od sposobu traktowania jednostek, bezstronności procedur oraz możliwości przedstawienia własnego stanowiska.

Jeżeli człowiek zostaje pozbawiony wolności na podstawie danych, których nie zna, modelu, którego nie może zbadać, i procedury pozbawionej należytej kontroli, naruszenie wykracza poza bezpośrednie skutki zatrzymania. Osłabieniu ulega zaufanie do zdolności państwa do odróżnienia podejrzenia od udowodnionego faktu. W tym kontekście rośnie znaczenie epistemicznej odpowiedzialności organów publicznych. Państwo korzystające z zaawansowanych instrumentów analitycznych powinno dysponować kompetencjami pozwalającymi rozpoznać granice ich trafności, gdyż przymus publiczny wymaga szczególnie wysokiego standardu uzasadnienia. Im poważniejsza potencjalna ingerencja w prawa jednostki, tym mniej uzasadnione staje się traktowanie technologicznej wygody jako substytutu niezależnego postępowania dowodowego. Problem ten ma także wymiar konstytucyjny i międzynarodowy: Europejska Konwencja Praw Człowieka chroni prawo do wolności, bezpieczeństwa osobistego i rzetelnego procesu, a Karta praw podstawowych UE ustanawia gwarancje godności, ochrony danych oraz skutecznej ochrony sądowej.

Zastosowanie identyfikacji biometrycznej przez państwo wymaga zatem nie tylko oceny technicznej, lecz rozpoznania podstawy prawnej, konieczności i proporcjonalności. Nie każde użycie AI w postępowaniu karnym jest niedopuszczalne, jednak szczególne wymagania pojawiają się wtedy, gdy technologia służy identyfikowaniu osób, kształtowaniu podejrzeń lub wspomaganiu działań mogących prowadzić do pozbawienia wolności. W takich przypadkach odpowiedzialność za ustalenie faktów i wydanie rozstrzygnięcia powinna pozostawać rzeczywista, a nie jedynie formalnie przypisana człowiekowi.

Przejrzystość algorytmiczna jako gwarancja ochrony konsumenta

Odmienną postać szkody algorytmicznej ujawniają procesy podejmowania decyzji kredytowych i finansowych. Banki oraz instytucje pożyczkowe od wielu dekad wykorzystują modele statystyczne do szacowania ryzyka niewypłacalności, ustalania warunków finansowania i zarządzania portfelem należności. Współczesne systemy uczenia maszynowego zwiększają liczbę zmiennych możliwych do uwzględnienia oraz pozwalają analizować złożone, nieliniowe relacje między zachowaniami klientów a prawdopodobieństwem przyszłych zdarzeń. Rozwój ten może poprawiać precyzję oceny ryzyka, lecz równocześnie stwarza zagrożenie posługiwania się cechami, których związek z rzeczywistą zdolnością kredytową pozostaje niejasny lub społecznie problematyczny. Dane o miejscu zamieszkania, historii zatrudnienia, zachowaniach konsumpcyjnych czy sposobie korzystania z usług cyfrowych mogą bowiem odzwierciedlać głębokie nierówności majątkowe,

geograficzne i społeczne. Ich zastosowanie nie jest automatycznie bezprawne, jednak wymaga rzetelnej analizy celu, adekwatności, proporcjonalności oraz ewentualnego skutku dyskryminacyjnego.

W prawie europejskim szczególne znaczenie ma wyrok Trybunału Sprawiedliwości Unii Europejskiej z 7 grudnia 2023 roku w sprawie C-634/21 (SCHUFA Holding). Trybunał rozstrzygał w nim problem automatycznego ustalania prawdopodobieństwa przyszłego wywiązywania się przez osobę z zobowiązań płatniczych. Wskazał, że automatyczne obliczenie takiego wyniku przez agencję informacji kredytowej może stanowić zautomatyzowane podejmowanie indywidualnych decyzji w rozumieniu art. 22 RODO, jeżeli instytucja korzystająca z oceny przypisuje jej rozstrzygające znaczenie dla ustanowienia, wykonania lub rozwiązania stosunku umownego. Doniosłość tego stanowiska polega na zakwestionowaniu możliwości formalnego rozdzielenia procesu między kilka organizacji w sposób prowadzący do osłabienia ochrony osoby, której dane dotyczą. Jeśli jeden podmiot wytwarza wynik, a drugi w praktyce niemal automatycznie na jego podstawie rozstrzyga o dostępie do kredytu, nie można ignorować funkcjonalnego charakteru całego procesu tylko dlatego, że końcowa decyzja została technicznie wykonana przez inną instytucję.

Kolejny istotny etap rozwoju europejskich gwarancji przejrzystości nastąpił 27 lutego 2025 roku wraz z wyrokiem TSUE w sprawie C-203/22 (Dun & Bradstreet Austria). Trybunał odniósł się do obowiązku przekazania osobie, której dane dotyczą, informacji o zasadach stosowanych w procesie automatycznego podejmowania decyzji. Podkreślił konieczność udostępnienia danych pozwalających zrozumieć procedurę i zasady rzeczywiście wykorzystane podczas przetwarzania, w tym sposób, w jaki konkretne informacje wpłynęły na rezultat. Sama złożoność obliczeń nie zwalnia administratora z obowiązków informacyjnych wynikających z RODO. Jednocześnie ochrona tajemnicy przedsiębiorstwa wymaga odpowiedniego wyważenia interesów, a nie automatycznego ujawniania pełnego kodu źródłowego modelu każdej zainteresowanej osobie. Rozstrzygnięcie to ma kluczowe znaczenie dla prawa do skutecznego kwestionowania decyzji algorytmicznych; informacja udostępniana jednostce powinna bowiem posiadać rzeczywistą wartość poznawczą, a nie ograniczać się do abstrakcyjnego komunikatu o wykorzystaniu zaawansowanej analizy danych. Analogiczne założenia pojawiają się w regulacjach amerykańskich.

Consumer Financial Protection Bureau w komunikatach z 2022 i 2023 roku podkreślało, że instytucje finansowe korzystające z zaawansowanych modeli AI pozostają zobowiązane do przekazywania konsumentom konkretnych i prawidłowych przyczyn niekorzystnych decyzji kredytowych. Złożoność modelu nie stanowi samodzielnego usprawiedliwienia dla niemożności wskazania rzeczywistych przesłanek odmowy. Przejrzystość algorytmiczna nie jest zatem wyłącznie postulatem etycznym ani właściwością pożądaną przez specjalistów od interpretowalności modeli.

Staje się ona elementem materialnej ochrony konsumenta, pozwalającym zidentyfikować błędy w danych, kwestionować nieprawidłowe oceny i podejmować działania zmierzające do poprawy sytuacji finansowej. Choć zakres tych obowiązków zależy od systemu prawnego, rodzaju instytucji i charakteru decyzji, wspólną zasadą pozostaje niedopuszczalność zastępowania uzasadnienia rozstrzygnięcia samym stwierdzeniem, że zostało ono wygenerowane przez model. Z ekonomicznego punktu widzenia nieuzasadnione ograniczenie dostępu do kredytu może wywoływać konsekwencje daleko wykraczające poza pojedynczą odmowę finansowania. Brak środków na podjęcie działalności gospodarczej może zablokować przedsiębiorcy wejście na rynek, nieuzasadnione podwyższenie kosztów kredytu mieszkaniowego zwiększa długookresowe obciążenia gospodarstwa domowego, a błędna klasyfikacja ryzyka wpływa na przyszłe możliwości inwestowania w edukację, zdrowie i rozwój zawodowy. Powstaje mechanizm kumulacyjnej deprywacji kapitałowej, w którym wcześniejsza decyzja finansowa współkształtuje przyszłe warunki oceny tej samej osoby. Z perspektywy teorii zdolności Amartyi Sena szczególnie istotne jest to, że utrata dostępu do zasobu ekonomicznego może oznaczać realne zmniejszenie możliwości realizowania wartościowych celów życiowych.

Dlatego sprawiedliwość kredytowa nie sprowadza się do formalnej poprawności prognozy niespłacenia zobowiązania, lecz obejmuje również prawidłowe określenie celu modelu, legalność zastosowanych kryteriów i możliwość skorygowania błędnej informacji.

Automatyzacja rekrutacji utrwala historyczne nierówności

Podobnie głębokie konsekwencje może wywoływać automatyzacja procesów zatrudnienia. Systemy rekrutacyjne służą dziś do porządkowania aplikacji, analizy zgodności kompetencji z wymaganiami stanowiska, identyfikacji odpowiednich kandydatów oraz wspomagania planowania zasobów ludzkich. Choć mogą one obniżyć koszty selekcji, przyspieszyć proces i w pewnych sytuacjach ograniczyć wpływ ludzkich uprzedzeń, niosą ze sobą istotne ryzyka. Wykorzystanie historycznych danych zatrudnieniowych jako podstawy oceny przyszłych kandydatów może bowiem odtwarzać nierówności wynikające z wcześniejszych praktyk organizacyjnych. Model uczony na dokumentacji pracowników uznawanych przez przedsiębiorstwo za najlepszych może preferować określone ścieżki kariery, formy wykształcenia, przerwy w zatrudnieniu czy sposoby autoprezentacji, które są ściśle powiązane z płcią, wiekiem, niepełnosprawnością lub pochodzeniem społecznym. Jeżeli kryteria te nie odpowiadają rzeczywistym wymaganiom pracy, automatyczna selekcja może prowadzić do systematycznego ograniczania dostępu do zatrudnienia osobom posiadającym

odpowiednie kompetencje. Ilustracją tego mechanizmu jest eksperymentalny projekt rekrutacyjny Amazona, którego problemy ujawniono publicznie w 2018 roku. System ten, rozwijany od 2014 roku, opierał się na danych z historycznych aplikacji i służył do oceny kandydatów na stanowiska techniczne. W trakcie prac nad narzędziem okazało się, że preferuje ono profile częściej występujące wśród mężczyzn, jednocześnie niekorzystnie traktując sygnały związane z doświadczeniem kobiet.

Automatyzacja utrwala dyskryminację i tworzy nowe zależności ekonomiczne

Przedsiębiorstwo ostatecznie zrezygnowało z rozwijania tego rozwiązania w pierwotnej postaci. Należy jednak precyzyjnie określić kontekst: przypadek ten nie jest dowodem na masowe, autonomiczne odrzucanie kobiet przez w pełni wdrożony system produkcyjny, gdyż projekt miał charakter eksperymentalny i nie stanowił jedynej podstawy decyzji rekrutacyjnych. Jego znaczenie polega na ujawnieniu mechanizmu, w którym model uczony na danych opisujących historyczne praktyki organizacji utożsamia wcześniejszą dominację określonej grupy z pożądanym profilem przyszłego pracownika. W takim przypadku zwiększenie dokładności odtwarzania historycznych wzorców nie musi prowadzić do sprawiedliwej selekcji kandydatów.

Jeszcze bardziej bezpośrednie potwierdzenie naruszeń prawa antydyskryminacyjnego dostarcza sprawa iTutorGroup, zakończona w 2023 roku ugodą z amerykańską Equal Employment Opportunity Commission. Według ustaleń organu, firma świadcząca usługi edukacyjne korzystała z oprogramowania automatycznie odrzucającego kandydatki powyżej 55. roku życia oraz mężczyzn powyżej 60. roku życia. Mechanizm ten wykluczył ponad dwieście kwalifikujących się osób; ugoda przewidziała wypłatę 365 tysięcy dolarów świadczeń oraz wprowadzenie dodatkowych środków zaradczych. Przypadek ten nie wymaga przypisywania systemowi zaawansowanej inteligencji maszynowej – dyskryminacja wynikała tu z prostego, automatycznego wykonywania reguł selekcji.

Właśnie dlatego przykład ten jest tak istotny dla analizy odpowiedzialności. Pokazuje on, że podstawowym zagrożeniem nie zawsze musi być złożoność nieinterpretowalnego modelu. Niekiedy decydującą przyczyną szkody okazuje się nieuprawnione kryterium, które dzięki automatyzacji można zastosować szybko, konsekwentnie i na masową skalę.

Automatyzacja stosunków pracy wykracza jednak daleko poza samą rekrutację. Coraz więcej przedsiębiorstw wdraża systemy monitorowania aktywności, przydzielania zadań, planowania harmonogramów czy oceniania produktywności, które wspierają decyzje o wynagrodzeniu i

zakończeniu współpracy. W środowisku platform cyfrowych algorytm przejmuje funkcje zbliżone do tradycyjnego kierownictwa, choć relacja prawna między platformą a wykonawcą nie zawsze jest kwalifikowana jako stosunek pracy.

System przydziela zlecenia, mierzy czas ich realizacji, analizuje opinie klientów i modyfikuje warunki dostępu do zamówień, co bezpośrednio wpływa na możliwość kontynuowania działalności zarobkowej. W rezultacie formalna samodzielność ekonomiczna wykonawcy może współistnieć z istotną zależnością od cyfrowej infrastruktury zarządzania. Ocena tej sytuacji wymaga jednoczesnego zastosowania aparatu prawa pracy, ekonomii kosztów transakcyjnych, socjologii organizacji i teorii władzy.

W tradycji ekonomicznej kluczowe jest rozróżnienie między formalną swobodą zawierania umów a rzeczywistą strukturą zależności ekonomicznej. Ronald Coase analizował przedsiębiorstwo jako alternatywę dla czystej koordynacji rynkowej w organizacji transakcji, natomiast Oliver Williamson rozwijał teorię kosztów transakcyjnych i specyficzności aktywów. W gospodarce platformowej technologia pozwala tworzyć struktury, w których podział pracy, kontrola efektywności i alokacja wynagrodzeń są w znacznym stopniu sterowane przez oprogramowanie. Osoba formalnie samodzielna może zatem pozostawać zależna od zasad, których nie współtworzy, nie rozumie i nie potrafi skutecznie negocjować. Jeśli algorytm zmieni częstotliwość przydzielania zleceń lub ograniczy dostęp do platformy, efekt ekonomiczny może być tożsamy z drastycznym pogorszeniem warunków zatrudnienia, mimo że nie przybiera on postaci tradycyjnej decyzji kadrowej.

Automatyzacja nie może znosić odpowiedzialności instytucjonalnej

Właśnie ta niejednoznaczność skłania do poszukiwania nowych instrumentów ochrony praw osób wykonujących pracę za pośrednictwem systemów cyfrowych. W Unii Europejskiej odpowiedzią na część tych problemów jest dyrektywa 2024/2831 w sprawie poprawy warunków pracy za pośrednictwem platform, której termin transpozycji upływa 2 grudnia 2026 roku.

Akt ten nie ogranicza się jedynie do kwestii prawidłowego ustalania statusu zatrudnienia, lecz wprowadza ramy przejrzystości i nadzoru nad zarządzaniem algorytmicznym. Przewiduje obowiązki informacyjne w zakresie zautomatyzowanych systemów monitorowania i podejmowania decyzji, ograniczenia w przetwarzaniu niektórych kategorii danych osobowych, wymogi oceny skutków oraz udział kompetentnych osób w kontroli istotnych rozstrzygnięć. Szczególnie znaczące jest objęcie tymi zabezpieczeniami także osób wykonujących pracę za pośrednictwem platform bez

pozostawania w stosunku pracy. Rozwiązanie to wynika z uznania, że niekorzystne skutki monitorowania i automatyzacji nie zależą wyłącznie od formalnej kwalifikacji umowy. W wymiarze aksjologicznym jest to próba ochrony podmiotowości ekonomicznej człowieka w środowisku, w którym uprawnienia kierownicze mogą być realizowane przez system informatyczny. Z perspektywy nauk społecznych jednym z najgroźniejszych następstw zarządzania algorytmicznego jest przenoszenie odpowiedzialności organizacyjnej na pozornie neutralną procedurę techniczną.

Pracownik lub osoba samozatrudniona może otrzymać niekorzystną ocenę, ograniczenie dostępu do zleceń czy informację o zawieszeniu konta, nie mając możliwości ustalenia, jakie zdarzenia doprowadziły do takiej decyzji. Jednocześnie przedsiębiorstwo może przedstawiać rozstrzygnięcie jako wynik procesu automatycznego, ograniczając tym samym dostęp do osoby z realnymi kompetencjami do jego zmiany. W takim układzie dochodzi do erozji tradycyjnych form odpowiedzialności kierowniczej; decyzja pozostaje organizacyjnie skuteczna, lecz jej autorstwo zostaje rozproszone między dostawcę oprogramowania, zarządcę platformy a formalne procedury operacyjne. Odpowiedzialność nie powinna jednak znikać wraz z automatyzacją procesu. Jeśli przedsiębiorstwo organizuje działalność za pomocą określonej infrastruktury, powinno zapewniać mechanizmy wyjaśniania i korygowania skutków jej działania.

Jeszcze bardziej wymowne przykłady dostarczają systemy administracji publicznej wykorzystujące automatyzację do wykrywania nieprawidłowości w świadczeniach społecznych. W takich przypadkach jednostka nie jest jedynie konsumentem usługi, który może zmienić dostawcę, lecz adresatem działań organu publicznego, którego kompetencje wynikają z prawa i mogą być realizowane przy użyciu środków przymusu. Kluczową rolę odgrywają tu zatem legalność, proporcjonalność, obowiązek prawidłowego ustalenia stanu faktycznego oraz prawo do skutecznego kwestionowania decyzji.

Automatyzacja może zwiększać sprawność wykrywania nadużyć, ale prowadzi również do masowego stosowania błędnych założeń, jeśli procedura pomija wyjątki, niepełne dane lub rzeczywiste warunki życia świadczeniobiorców. Szczególnie niebezpieczne jest praktyczne odwrócenie ciężaru dowodzenia: obywatel zostaje zmuszony do wykazania błędu systemu, mimo że nie ma dostępu do informacji, na podstawie których sformułowano zarzut. Przykładem takiej porażki był australijski program Robodebt, realizowany od 2015 roku i przerwany po ujawnieniu fundamentalnych problemów z legalnością stosowanej procedury.

System opierał się na porównywaniu danych o dochodach oraz uśrednianiu informacji rocznych w celu identyfikacji potencjalnych nadpłat. W określonych przypadkach dochód roczny traktowano jako podstawę szacowania zarobków w krótszych okresach, ignorując fakt, że mogły one być

nieregularne. Takie oszacowania służyły do formułowania roszczeń o zwrot świadczeń, choć nie stanowiły wystarczającego dowodu faktycznego uzyskania określonych kwot w danym czasie.

Program ten nie był przykładem generatywnej AI ani zaawansowanego uczenia maszynowego. Jego znaczenie polega na wykazaniu, że nawet prosta automatyzacja, połączona z wadliwymi przesłankami prawnymi i instytucjonalnymi, może generować szkody na masową skalę. Australijska Królewska Komisja badająca Robodebt przedstawiła w lipcu 2023 roku raport dokumentujący wielowymiarowe niepowodzenia programu: od wadliwej konstrukcji procedur i problemów z legalnością, po niewystarczającą reakcję na sygnały ostrzegawcze oraz uchybienia w odpowiedzialności instytucjonalnej.

W wyniku działania programu od setek tysięcy osób dochodzono należności, które nie zostały prawidłowo ustalone. Proces naprawczy objął zwroty środków, anulowanie części zobowiązań i odszkodowania. Skala finansowa ugody zbiorowej liczona była w miliardach dolarów australijskich, choć nie należy utożsamiać całej tej kwoty z sumą rzeczywiście pobranych, bezprawnych należności. Konsekwencje wykraczały jednak poza straty materialne – obejmowały stres, poczucie zagrożenia, utratę zaufania do państwa i długotrwałe trudności życiowe.

Automatyzacja skaluje błędy systemowe i instytucjonalne

Program ten stał się przykładem sytuacji, w której organizacyjna efektywność egzekwowania należności została całkowicie oderwana od jakości prawnego uzasadnienia ich istnienia. W perspektywie teorii administracji publicznej Robodebt ujawnia głęboki konflikt między logiką zarządzania przez wskaźniki a zasadami państwa prawnego. System może wprawdzie zwiększać liczbę generowanych decyzji, przyspieszać identyfikację potencjalnych nieprawidłowości i ograniczać koszty obsługi pojedynczej sprawy, jednak wskaźniki te nie stanowią wystarczających mierników legalności ani jakości działania organu. Jeśli podstawą procesu jest błędne założenie, automatyzacja nie naprawia problemu, lecz zwiększa jego skalę i tempo reprodukcji. W literaturze organizacyjnej mechanizm ten można powiązać z koncepcją normalizacji odstępstw Diane Vaughan; wskazuje ona, że powtarzające się sygnały o nieprawidłowościach mogą być stopniowo traktowane jako zwyczajny element funkcjonowania instytucji. Gdy określona procedura staje się integralną częścią budżetowania, raportowania wyników i realizacji celów politycznych, zdolność organizacji do jej zakwestionowania maleje. Problem odpowiedzialności przestaje wówczas dotyczyć wyłącznie projektanta systemu, a obejmuje całą strukturę decyzyjną, która utrzymuje proces mimo wiedzy o jego wadliwości.

Podobne zagrożenia ujawnił amerykański system MiDAS, wykorzystywany w stanie Michigan do wykrywania nadużyć przy zasiłkach dla bezrobotnych. W latach 2013–2015 automatyczne procedury doprowadziły do masowych, fałszywych oskarżeń o nieuczciwe pobieranie świadczeń. Późniejsze ustalenia wykazały, że problem dotknął ponad czterdzieści tysięcy osób, a analiza stanowych instytucji kontrolnych potwierdziła bardzo wysoki odsetek błędnych kwalifikacji w zbadanej próbie automatycznych rozstrzygnięć. Konsekwencje dla obywateli obejmowały żądania zwrotu środków, sankcje finansowe oraz poważne trudności materialne. Spory sądowe zakończyły się m.in. zatwierdzeniem w 2023 roku ugody w wysokości dwudziestu milionów dolarów na rzecz grupy osób poszkodowanych.

Przypadek ten nie powinien być odczytywany jako dowód zawodności wszystkich algorytmów wykrywających oszustwa, lecz jako mocne ostrzeżenie przed wdrażaniem zautomatyzowanych procedur o wysokiej dotkliwości skutków bez odpowiedniego nadzoru, niezależnej kontroli danych i skutecznych środków odwoławczych. W Europie jednym z najważniejszych przykładów instytucjonalnego kryzysu związanego z profilowaniem ryzyka pozostaje holenderska afera świadczeń na opiekę nad dziećmi (**toeslagenaffaire**). Przez lata organy administracji podatkowej i świadczeniowej stosowały praktyki, które prowadziły do nieuzasadnionego podejrzewania licznych rodzin o nieprawidłowości oraz żądania zwrotu wysokich kwot pomocy. Kluczową rolę odegrały tu restrykcyjne procedury administracyjne, niewystarczające gwarancje ochrony prawnej oraz wykorzystywanie danych dotyczących obywatelstwa. Holenderski organ ochrony danych osobowych w grudniu 2021 roku nałożył na administrację podatkową karę w wysokości 2,75 miliona euro za bezprawne i dyskryminacyjne przetwarzanie informacji o obywatelstwie (w tym o posiadaniu podwójnego). Ustalono bowiem, że dane te służyły jako jeden ze wskaźników automatycznego kwalifikowania wniosków do kategorii ryzyka.

Szkoda algorytmiczna jako efekt przemocy strukturalnej i wad prawnych

Należy jednak zachować proporcję przyczynową: cały kryzys nie był skutkiem działania jednego modelu AI, lecz wieloletniego splotu wadliwych procedur prawnych, organizacyjnych, kontrolnych i cyfrowych. Doniosłość holenderskiego przypadku staje się szczególnie wyraźna w świetle teorii przemocy strukturalnej Johana Galtunga.

Koncepcja ta wskazuje, że szkoda może wynikać nie tylko z bezpośredniego działania pojedynczego sprawcy, ale także z funkcjonowania struktur społecznych, które ograniczają jednostkom dostęp do dóbr i możliwości. W przypadku zautomatyzowanej administracji szczególnie niebezpieczny jest

proces, w którym obywatel doświadcza jednocześnie asymetrii informacyjnej, presji finansowej i trudności w uzyskaniu skutecznej ochrony prawnej. Rodzina obciążona nieuzasadnionym żądaniem zwrotu świadczenia może zostać zmuszona do ograniczenia wydatków na mieszkanie, edukację czy opiekę nad dziećmi, a przeciągający się spór administracyjny jedynie pogłębia jej zależność ekonomiczną. W takich sytuacjach jednostkowa decyzja finansowa przekształca się w długotrwały mechanizm ograniczania realnych możliwości życiowych. Naprawienie tej sytuacji często wymaga więcej niż zwykłego zwrotu błędnie pobranej należności, gdyż część szkód ma charakter psychologiczny, rodzinny i społeczny.

Odrębne znaczenie ma wyrok Sądu Okręgowego w Hadze z 5 lutego 2020 roku dotyczący systemu SyRI, wykorzystywanego do identyfikowania ryzyka nieprawidłowości w obszarze świadczeń i podatków. Sąd uznał, że regulacje stanowiące podstawę stosowania systemu nie zapewniały odpowiedniego wyważenia interesu publicznego względem prawa do poszanowania życia prywatnego, co naruszało art. 8 Europejskiej Konwencji Praw Człowieka. Kluczowe były tu braki w przejrzystości, zakres wykorzystywanych danych oraz niedostateczne gwarancje ochrony jednostki. Wyrok nie wprowadził ogólnego zakazu stosowania algorytmów do przeciwdziałania nadużyciom, lecz potwierdził, że prawna dopuszczalność takiego systemu nie może wynikać wyłącznie z jego deklarowanej przydatności dla administracji. Konieczne jest zapewnienie solidnej podstawy prawnej, proporcjonalności ingerencji oraz realnych zabezpieczeń przed arbitralnością.

Na gruncie polskiego prawa konstytucyjnego doświadczenia te prowadzą do analizy kilku powiązanych gwarancji. Art. 30 Konstytucji RP ustanawia przyrodzoną i niezbywalną godność człowieka jako źródło wolności i praw, a art. 32 wyraża zasadę równości wobec prawa i zakaz dyskryminacji. Prywatność chroni art. 47, autonomia informacyjna jednostki – art. 51, natomiast prawo do rozpatrzenia sprawy przez niezależny i bezstronny sąd gwarantuje art. 45. Istotne są również art. 77, dotyczący wynagrodzenia szkody wyrządzonej przez niezgodne z prawem działanie organu władzy publicznej, oraz art. 78, określający zasadę zaskarżalności decyzji wydanych w pierwszej instancji.

Przepisy te nie stanowią odrębnego kodeksu regulującego AI, lecz są normatywną podstawą oceny dopuszczalności jej wykorzystania przez państwo. Dodatkowe gwarancje wynikają z Kodeksu postępowania administracyjnego (k.p.a.), opartego na zasadzie praworządności, obowiązku wyczerpującego wyjaśnienia okoliczności sprawy oraz prawie strony do czynnego udziału w postępowaniu i otrzymania odpowiedniego uzasadnienia rozstrzygnięcia. Art. 10 k.p.a. odnosi się do uczestnictwa strony, art. 77 nakłada na organ obowiązek zebrania materiału dowodowego, a art. 80 wiąże ocenę dowodów z całokształtem zgromadzonych informacji. Kluczowy jest tu art. 107, określający wymogi uzasadnienia decyzji administracyjnej.

Automatyczne narzędzia analityczne mogą wspierać realizację tych obowiązków, ale nie zastępują prawidłowego ustalenia stanu faktycznego ani oceny prawnej. Jeśli organ posługuje się oceną algorytmiczną, musi rozpoznać jej znaczenie dla konkretnej sprawy, granice wiarygodności narzędzia oraz to, czy strona dysponuje realną możliwością zakwestionowania przesłanek rozstrzygnięcia.

Ramy prawne i zasada proporcjonalności ograniczają automatyzację

Kluczowe jest rozróżnienie między prawem do informacji o przetwarzaniu danych a możliwością uzyskania pełnego, technicznego opisu algorytmu. Rozporządzenie 2016/679 (RODO) ustanawia zasady legalności, rzetelności i przejrzystości, a także ograniczenia celu, minimalizacji danych, ich prawidłowości oraz rozliczalności procesu przetwarzania. Artykuł 22 reguluje z kolei szczególną kategorię decyzji opartych wyłącznie na zautomatyzowanym przetwarzaniu – w tym profilowaniu – o ile wywołują one skutki prawne wobec osoby lub w podobny sposób istotnie na nią wpływają. Przepisu tego nie należy jednak interpretować jako bezwarunkowego zakazu wszelkiej automatyzacji, gdyż przewiduje on określone wyjątki i odpowiadające im zabezpieczenia.

Z kolei artykuł 35 nakłada obowiązek przeprowadzenia oceny skutków dla ochrony danych w sytuacjach, gdy dany rodzaj przetwarzania może wiązać się z wysokim ryzykiem naruszenia praw lub wolności osób fizycznych. W praktyce odpowiedzialne wdrażanie AI wymaga zatem analizy rodzaju podejmowanych decyzji, zakresu udziału człowieka, charakteru danych oraz potencjalnej dotkliwości konsekwencji. Na poziomie prawa Unii Europejskiej kluczową rolę odgrywa rozporządzenie 2024/1689 (AI Act), które wprowadza regulację opartą na kategoriach ryzyka. System ten obejmuje zakazy określonych praktyk, obowiązki przejrzystości, wymagania dotyczące modeli ogólnego przeznaczenia oraz szczególne ramy dla systemów wysokiego ryzyka. Obowiązki te są wprowadzane etapami: zakazy konkretnych praktyk i wymogi kompetencyjne w zakresie AI zaczęły obowiązywać 2 lutego 2025 roku, a zasady dotyczące modeli ogólnego przeznaczenia wejdą w życie 2 sierpnia 2025 roku. Zmiany przyjęte w 2026 roku przesunęły termin stosowania obowiązków dla wielu samodzielnych systemów wysokiego ryzyka (załącznik III) na 2 grudnia 2027 roku, natomiast dla systemów zintegrowanych z produktami objętymi załącznikiem I – na 2 sierpnia 2028 roku. Przesunięcia w harmonogramie nie oznaczają jednak zawieszenia obowiązków wynikających z RODO, prawa antydyskryminacyjnego, regulacji sektorowych czy ogólnych zasad odpowiedzialności cywilnej i administracyjnej.

Szczególnego znaczenia nabiera tutaj zasada proporcjonalności, której rola w europejskim konstytucjonalizmie była szeroko analizowana m.in. przez Roberta Alexy'ego. Oceniając dopuszczalność ingerencji w prawa jednostki, należy rozważyć przydatność środka do osiągnięcia prawnie uzasadnionego celu, konieczność jego zastosowania w świetle dostępnych alternatyw oraz relację między uzyskaną korzyścią a intensywnością ograniczenia chronionych dóbr. W odniesieniu do systemów algorytmicznych oznacza to, że sama techniczna możliwość przetworzenia danych nie jest wystarczającym uzasadnieniem dla ich wykorzystania. Instytucja musi wykazać, dlaczego potrzebuje konkretnego modelu, jakie problemy zamierza rozwiązać, czy istnieją mniej ingerencyjne alternatywy i jakie mechanizmy ograniczają niekorzystne następstwa. Analiza proporcjonalności nie może zostać zastąpiona prostym porównaniem wskaźników efektywności, gdyż obejmuje ona ocenę praw i interesów, których nie da się w pełni przeliczyć na wspólną jednostkę ekonomiczną.

Rozproszenie odpowiedzialności w łańcuchu dostaw AI

W prawie cywilnym odpowiedzialność za szkody związane z wykorzystaniem AI wymaga precyzyjnego rozpoznania podstawy prawnej oraz przesłanek odpowiedzialności. W polskim systemie kluczowe mogą okazać się ogólne reguły odpowiedzialności deliktowej, w tym art. 415 Kodeksu cywilnego, odpowiedzialność za cudze działania w określonych stosunkach organizacyjnych, zasady odpowiedzialności kontraktowej oraz regulacje dotyczące szkód wyrządzonych przy wykonywaniu władzy publicznej. Podstawą odpowiedzialności za szkodę wyrządzoną przez niezgodne z prawem działanie lub zaniechanie przy wykonywaniu władzy publicznej jest art. 417 k.c., o ile spełnione zostaną pozostałe wymagane przesłanki. Z kolei ochrona dóbr osobistych może opierać się na art. 23 i 24 k.c.

Nie oznacza to jednak, że każdy niekorzystny wynik działania AI automatycznie rodzi obowiązek odszkodowawczy. Konieczne pozostaje ustalenie zdarzenia prawnego, szkody, związku przyczynowego oraz przesłanek właściwych dla danego reżimu prawnego. Dodatkowy kierunek rozwoju odpowiedzialności za technologie cyfrowe wyznacza dyrektywa 2024/2853 dotycząca odpowiedzialności za produkty wadliwe (termin transpozycji upływa 9 grudnia 2026 roku). Regulacja ta obejmuje oprogramowanie w zakresie określonym jej przepisami i dąży do dostosowania zasad odpowiedzialności produktowej do współczesnego rynku technologicznego. Szczególne znaczenie ma tu kwestia wad związanych z bezpieczeństwem systemów cyfrowych, aktualizacjami oraz specyficznymi sposobami funkcjonowania oprogramowania.

Nie oznacza to jednak, że każde zastosowanie AI zostaje automatycznie objęte jednym jednolitym reżimem odpowiedzialności produktowej. Kwalifikacja zależy od rodzaju produktu, statusu podmiotów w łańcuchu dostaw, momentu wprowadzenia do obrotu i charakteru szkody. W praktyce może dochodzić do zbiegu lub wzajemnego oddziaływania różnych podstaw odpowiedzialności – od ochrony danych i prawa konsumenckiego, przez prawo medyczne, aż po odpowiedzialność kontraktową i produktową. Z perspektywy ekonomii prawa kluczowy pozostaje problem alokacji kosztu ryzyka pomiędzy uczestnikami procesu technologicznego. Guido Calabresi analizował znaczenie przypisywania odpowiedzialności podmiotom posiadającym możliwość zapobiegania szkodzie i ograniczania jej społecznych kosztów. W systemach AI zdolność ta bywa rozproszona.

Dostawca modelu posiada wiedzę o architekturze i procesie treningowym, integrator odpowiada za połączenie systemu z konkretnym środowiskiem, instytucja wdrażająca definiuje procedurę organizacyjną, a użytkownik końcowy kontroluje bezpośredni sposób wykorzystania narzędzia. Każdy z tych podmiotów dysponuje inną informacją i odmiennymi możliwościami interwencji. Racjonalny system odpowiedzialności powinien uwzględniać ten podział kompetencji, zamiast automatycznie przenosić całość ryzyka na najbardziej dostępnego uczestnika procesu albo na osobę, której sytuacja została niekorzystnie rozstrzygnięta.

Iluzoryczność nadzoru i bariery w dochodzeniu sprawiedliwości

W ekonomii instytucjonalnej problem ten koresponduje z zagadnieniem kosztów transakcyjnych i asymetrii informacji. Osoba poszkodowana decyzją algorytmiczną często nie zna dostawcy modelu, nie ma dostępu do dokumentacji technicznej i nie dysponuje zasobami niezbędnymi do przeprowadzenia niezależnej ekspertyzy. Przedsiębiorstwo korzystające z technologii może w tym kontekście zasłaniać się tajemnicą handlową dostawcy, ten z kolei może wskazywać na błędy we wdrażaniu systemu przez użytkownika. Tworzy to strukturę, w której dochodzenie roszczeń wymaga od jednostki poniesienia kosztów nieproporcjonalnych do jej możliwości. Z perspektywy teorii dostępu do sprawiedliwości, rozwijanej m.in. przez Mauro Cappellettiego i Bryanta Gartha, samo formalne istnienie prawa jest niewystarczające, jeśli jego realizacja pozostaje w praktyce niedostępna.

Dlatego odpowiedzialne zarządzanie AI wymaga wprowadzenia przejrzystych procedur skargowych, zabezpieczenia dokumentacji, zasad współpracy dowodowej oraz mechanizmów kontroli niezależnych od podmiotu gospodarczego. Konieczność takiego nadzoru wynika również z

ograniczeń modelu człowieka w pętli decyzyjnej (**human in the loop**). Samo wyznaczenie osoby formalnie zatwierdzającej decyzję nie gwarantuje, że kontrola ta ma rzeczywisty charakter.

Jeśli system generuje tysiące rekomendacji dziennie, nadzorca może nie mieć czasu na ich rzetelną i niezależną analizę. Gdy interfejs prezentuje ocenę modelu jako obiektywną i wysoce wiarygodną, pojawia się ryzyko nadmiernego polegania na automatycznym wyniku. Z kolei w organizacjach premiujących zgodność decyzji człowieka z sugestią algorytmu, formalne prawo do sprzeciwu może pozostać martwym zapisem. W efekcie udział człowieka służy głównie rozproszeniu odpowiedzialności, nie zapewniając dodatkowej jakości poznawczej. Prawidłowe projektowanie nadzoru musi zatem uwzględniać kompetencje kontrolera, dostępność danych, realną możliwość zmiany rozstrzygnięcia oraz odpowiednie warunki organizacyjne. W naukach o organizacji i ergonomii bezpieczeństwa problem ten opisuje koncepcja automatyzacyjnego skrzywienia poznawczego, polegająca na nieuzasadnionym uprzywilejowaniu rekomendacji systemu automatycznego względem innych dostępnych informacji. Człowiek może wówczas pomijać sygnały sprzeczne z oceną modelu, a w określonych sytuacjach interpretować własne wątpliwości jako dowód niższej kompetencji.

Odpowiedzialność instytucjonalna jako gwarancja podmiotowości człowieka

Zjawisko to nasila się, gdy technologia przez długi czas dostarcza prawidłowych wyników i zyskuje reputację niezawodnej. Paradoks polega na tym, że wzrost przeciętnej skuteczności systemu może osłabiać zdolność użytkownika do samodzielnego wykrywania rzadkich, lecz poważnych błędów. W obszarach wysokiego ryzyka konieczne jest zatem wdrażanie procedur kontroli, które zamiast nieustannego, ręcznego powtarzania wszystkich obliczeń, opierają się na precyzyjnym doborze przypadków wymagających interwencji, testach odporności i niezależnym monitorowaniu skutków. Kluczowa staje się odpowiedzialność za aktualizowanie modeli oraz ich wycofywanie z eksploatacji. System, który przeszedł pomyślnie początkową walidację, może stopniowo tracić skuteczność pod wpływem zmian w strukturze populacji, praktykach instytucjonalnych czy zachowaniach użytkowników. W medycynie nowe metody leczenia mogą zmienić korelację między objawami a rokowaniem, w finansach zmiany makroekonomiczne modyfikują strukturę ryzyka, a w administracji nowelizacja prawa może sprawić, że wcześniejsze reguły staną się nieaktualne.

Brak odpowiedniego monitoringu niesie ryzyko utrzymywania systemu, którego rzeczywiste właściwości pozostają nieznane. Jednorazowa ocena jakości oprogramowania przed wdrożeniem jest niewystarczająca; odpowiedzialność obejmuje cały cykl życia modelu, w tym reagowanie na

skargi, analizę przypadków granicznych, dokumentowanie zmian oraz decyzje o ograniczeniu lub zakończeniu stosowania technologii. Warto tu przywołać koncepcję instytucji wysokiej niezawodności (**high-reliability organizations**), rozwijaną w badaniach nad energetyką i lotnictwem – środowiskami o ekstremalnie wysokich kosztach błędów. Organizacje te nie zakładają doskonałości własnych procedur, lecz rozwijają zdolność do wykrywania słabych sygnałów zagrożenia, analizowania odstępstw i uczenia się na doświadczeniach. Szczególnie istotna jest gotowość do kwestionowania własnych założeń, uznanie kompetencji osób z wiedzą operacyjną oraz unikanie nadmiernego upraszczania przyczyn niepowodzeń. Przeniesienie tych zasad do zarządzania AI oznacza traktowanie zgłoszeń użytkowników i pracowników jako cennych źródeł wiedzy o ograniczeniach systemu.

Organizacja, która interpretuje każdą krytykę technologii jako opór wobec innowacji, ryzykuje utratę zdolności do rozpoznawania rzeczywistych szkód. Problem pogłębia się, gdy osoby najbardziej narażone na negatywne skutki działania AI mają najmniejsze możliwości zgłaszania zastrzeżeń. Dotyczy to m.in. pacjentów zależnych od świadczeń medycznych, osób korzystających z pomocy społecznej, pracowników w niestabilnej sytuacji zawodowej, migrantów, osób z niepełnosprawnościami czy obywateli w sporze z organem publicznym.

Ich sytuacja często ogranicza dostęp do informacji, poradnictwa prawnego oraz środków finansowych i czasu niezbędnych do dochodzenia roszczeń. W takim środowisku liczba formalnych skarg nie musi odzwierciedlać rzeczywistej skali szkód – niewielka liczba odwołań może wynikać nie z bezbłędności systemu, lecz z niedostępności procedury odwoławczej.

Dlatego ewaluacja społeczna AI powinna obejmować badania jakościowe, analizę barier uczestnictwa i doświadczenia użytkowników, by wykrywać szkody niezgłoszone formalnie. Perspektywa ta koresponduje z teorią sprawiedliwości Nancy Fraser, która wskazuje na powiązanie redystrybucji zasobów, społecznego uznania i reprezentacji politycznej. Algorytmiczne wykluczenie może uderzać we wszystkie te wymiary jednocześnie.

Niekorzystna ocena kredytowa odcina od zasobów ekonomicznych, błędne profilowanie prowadzi do stygmatyzacji, a nieprzejrzysty proces decyzyjny utrudnia wpływ na reguły, którym jednostka podlega. W ten sposób szkoda technologiczna może stać się mechanizmem reprodukcji nierówności strukturalnych. Zrozumienie tego procesu wymaga odrzucenia przekonania, że dyskryminację rozwiązuje samo osiągnięcie jednakowego wskaźnika błędów statystycznych. Formalna równość klasyfikacji nie usuwa automatycznie nierówności w dostępie do zasobów, ochrony prawnej czy możliwości kształtowania instytucji.

Z perspektywy filozoficznej zasadnicze znaczenie ma rozróżnienie między instrumentalnym traktowaniem człowieka a uznaniem jego podmiotowości moralnej. Immanuel Kant formułował zasadę, zgodnie z którą człowieczeństwo powinno być traktowane zawsze jako cel sam w sobie, a nie jedynie jako środek do realizacji innych zamierzeń.

W odniesieniu do AI nie oznacza to rezygnacji z klasyfikowania czy organizowania usług publicznych. Oznacza jednak, że człowiek nie może zostać całkowicie zredukowany do obiektu optymalizacji cudzej funkcji celu. Instytucja oceniająca jego sytuację powinna uwzględniać prawa niezależne od wyniku ekonomicznego i gwarantować ochronę przed błędem. Etyka godności wyznacza tu granice, za którymi technologiczna efektywność przestaje być uzasadnieniem dla ingerencji w życie jednostki. Jednocześnie odpowiedzialne podejście wymaga zachowania proporcji w ocenie zagrożeń.

Algorytmy mogą bowiem ograniczać arbitralność, identyfikować pomijane potrzeby zdrowotne, zwiększać dostępność usług czy wspierać osoby z niepełnosprawnościami. W wielu obszarach odpowiednio przetestowane systemy poprawiają decyzje podejmowane przez człowieka, którego poznanie również jest obarczone ograniczeniami i uprzedzeniami.

Odpowiedzialność instytucjonalna zamiast wiary w technologię

Nie jest zatem naukowo uzasadnione przeciwstawianie sprawiedliwego człowieka niesprawiedliwej maszynie jako dwóch jednorodnych kategorii. Przedmiotem oceny powinien pozostać rzeczywisty proces decyzyjny, a nie sam typ narzędzia. Właściwe pytanie brzmi: czy określony sposób organizacji decyzji ogranicza szkody, poprawia ich wykrywalność, zwiększa dostępność ochrony prawnej i tworzy lepsze warunki realizacji uzasadnionych celów społecznych?

W świetle przeanalizowanych przypadków autor rekonstruuje istotną prawidłowość dotyczącą współczesnego państwa i gospodarki cyfrowej. Najpoważniejsze szkody algorytmiczne powstają często nie wtedy, gdy system manifestuje niezwykłą autonomię, lecz gdy instytucja stopniowo rezygnuje z krytycznej oceny własnych procedur. Model staje się wówczas elementem infrastruktury organizacyjnej, a jego wynik zaczyna funkcjonować jako przesłanka rozstrzygająca, mimo że pierwotnie miał jedynie wspierać interpretację danych.

Proces ten może być potęgowany przez presję kosztową, braki kadrowe, nadmierne zaufanie do technologii oraz rozproszenie odpowiedzialności. W konsekwencji pojawia się paradoks instytucjonalny: organizacja deklaruje zwiększenie racjonalności działania dzięki AI, a

równocześnie osłabia własną zdolność do rozpoznawania sytuacji, w których stosowana procedura przestaje być uzasadniona. Rzeczywista odpowiedzialność za sztuczną inteligencję nie sprowadza się więc jedynie do obowiązku usuwania błędów po ich ujawnieniu. Obejmuje ona wcześniejsze rozpoznawanie możliwych szkód, precyzyjne formułowanie celów, organizowanie niezależnej kontroli, zapewnianie przejrzystości i utrzymywanie instytucjonalnej gotowości do zakwestionowania własnych decyzji. Wartość systemu nie może być mierzona wyłącznie liczbą obsłużonych spraw, oszczędnościami czy szybkością wykonania zadania, jeśli towarzyszy temu ograniczenie możliwości skutecznej ochrony jednostki. Prawidłowa ewaluacja powinna uwzględniać rozkład błędów, strukturę uzasadnionych odwołań, czas naprawienia szkody, dostępność informacji, zakres udziału człowieka oraz społeczne następstwa wdrożenia.

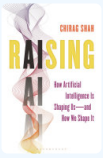
Dopiero takie podejście pozwala rozpoznać, czy automatyzacja prowadzi do rzeczywistego wzrostu sprawności instytucji, czy jedynie do przeniesienia kosztów na osoby o najmniejszych możliwościach sprzeciwu. Przedstawiona anatomia szkód prowadzi autora do wniosku, że centralnym problemem współczesnej AI nie jest wyłącznie zawodność maszyn, lecz zawodność ludzkich sposobów organizowania ich wykorzystania. Błędny model można udoskonalić, niedostatecznie reprezentatywne dane uzupełnić, a procedurę statystyczną poddać niezależnym testom. Znacznie trudniejszym zadaniem okazuje się przywracanie instytucjom zdolności do krytycznej refleksji nad własnymi celami, odpowiedzialnością i relacjami władzy. Technologia może bowiem zwiększać zarówno zakres ludzkich możliwości, jak i skalę skutków wcześniejszych błędów społecznych.

O jej cywilizacyjnej wartości nie przesądza samo istnienie algorytmu, lecz organizacja odpowiedzialności za decyzje, do których jego wynik prowadzi. Rozważania nad realnymi szkodami algorytmicznymi przygotowują grunt pod kolejny problem, bardziej subtelny i trudniejszy do bezpośredniej obserwacji. Nawet najlepiej zorganizowane procedury kontroli mogą okazać się niewystarczające, jeśli użytkownik zacznie przypisywać maszynie kompetencje, których nie posiada, lub uzna płynność jej komunikacji za dowód rozumienia i moralnej odpowiedzialności. Oprócz prawnych i instytucjonalnych warunków bezpieczeństwa konieczne staje się zatem rozpoznanie psychologicznych mechanizmów zaufania do technologii: antropomorfizacji, automatyzacyjnego skrzywienia poznawczego i delegowania osądu. Dopiero ich zbadanie pozwoli wyjaśnić, dlaczego człowiek, mimo formalnego zachowania kontroli nad systemem, może stopniowo tracić zdolność do samodzielnego kwestionowania jego rezultatów.

> Ostatecznie to nie błędy w kodzie są najtrudniejsze do naprawienia, lecz skłonność instytucji do delegowania osądu na neutralne w pozoru procedury. Czy w świecie

> zdominowanym przez algorytmiczną efektywność jesteśmy jeszcze w stanie rozpoznać moment, w którym wygoda technologiczna staje się substytutem sprawiedliwości? Prawdziwym ryzykiem nie jest zatem powstanie autonomicznej maszyny, lecz narodziny instytucji pozbawionej zdolności do zakwestionowania własnych narzędzi.

Inspiracje

[1] 

"Raising AI" Chirag Shah

Bloomsbury Academic | ISBN: 9798216452409

Chirag Shah jest profesorem w Szkole Informatyki i profesorem nadzwyczajnym w Szkole Informatyki i Inżynierii im. Paula G. Allena na Uniwersytecie Waszyngtońskim. Uzyskał tytuł doktora nauk informatycznych na Uniwersytecie Karoliny Północnej w Chapel Hill. Badania Shaha koncentrują się na inteligentnym dostępie do informacji, systemach wyszukiwania i rekomendacji, wspólnym wyszukiwaniu informacji, agentowej sztucznej inteligencji oraz odpowiedzialnych metodach sztucznej inteligencji, które zapewniają bezpieczeństwo, uczciwość i rozliczalność. Jest dyrektorem-założycielem InfoSeeking Lab i współdyrektorem-założycielem Center for Responsibility in AI Systems & Experiences (RAISE). Uznany za Wybitnego Członka zarówno Stowarzyszenia Maszyn Liczących (ACM), jak i ASIS&T, otrzymał w 2019 roku nagrodę Karen Spärck Jones od Brytyjskiego Towarzystwa Komputerowego za znaczący wkład w interaktywne wyszukiwanie i dostęp do informacji.

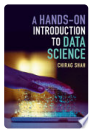
***Chirag Shah** is a Professor in the Information School and an Adjunct Professor in the Paul G. Allen School of Computer Science & Engineering at the University of Washington. He earned his Ph.D. in Information Science from the University of North Carolina at Chapel Hill. Shah's research concentrates on intelligent information access, search and recommender systems, collaborative information retrieval, agentic artificial intelligence, and responsible AI methods that ensure safety, fairness, and accountability. He is the founding director of the InfoSeeking Lab and founding co-director of the Center for Responsibility in AI Systems & Experiences (RAISE). Recognized as a Distinguished Member of both the Association for Computing Machinery (ACM) and ASIS&T, he received the 2019 Karen Spärck Jones Award from the British Computer Society for his significant contributions to interactive search and information access.*

University of Washington

Literatura uzupełniająca

Książki

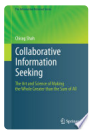
Książki autora i dzieła przywołane w tekście:



[1] "A Hands-On Introduction to Data Science"

Chirag Shah

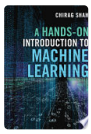
2020 | Cambridge University Press | ISBN: 9781108472449



[2] "Collaborative Information Seeking"

Chirag Shah

2012 | Springer Science & Business Media | ISBN: 9783642288135



[3] "A Hands-On Introduction to Machine Learning"

Chirag Shah

2022 | Cambridge University Press | ISBN: 9781009123303



[4] "Task Intelligence for Search and Recommendation"

Chirag Shah, Ryen W. White

2022 | Springer Nature | ISBN: 9783031023262



[5] "Agent Nation"

Chirag Shah

2026 | Springer Nature | ISBN: 9798868824548

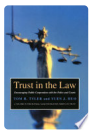


[6] "Interactive IR User Study Design, Evaluation, and Reporting"

Jiqun Liu, Chirag Shah

2022 | Springer Nature | ISBN: 9783031023194

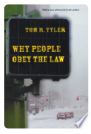
Książki autorów przywoływanych i literatura pokrewna:



[7] "Trust in the Law: Encouraging Public Cooperation with the Police and Courts"

Tom R. Tyler

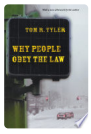
Russell Sage Foundation | ISBN: 9781610445429



[8] "Why People Obey the Law"

Tom R. Tyler

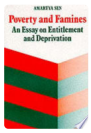
Princeton University Press | ISBN: 9781400828609



[9] "Why People Obey the Law, revised ed."

Tom R. Tyler

Princeton University Press | ISBN: 9781400828609



[10] **"Poverty and Famines: An Essay on Entitlement and Deprivation"**

Amartya Sen

Oxford University Press | ISBN: 9780198284635

[11] **"India: Economic Development and Social Opportunity"**

Amartya Sen

[12] **"The Costs of Accidents"**

Guido Calabresi

[13] **"El futuro del Derecho y Economía. Ensayos para la reflexión y la memoria"**

Guido Calabresi

ISBN: 9791387545086

[14] **"Access to justice. Vol. 2. Promising institutions. Edited by mauro cappelletti and john weisner. Book 2"**

Mauro Cappelletti, John Weisner

1979

[15] **"Access to justice. Vol. 2. Promising institutions. Edited by mauro cappelletti and john weisner. Book 1"**

Mauro Cappelletti, John Weisner

1978

[16] **"Access to justice"**

Mauro Cappelletti, John Weisner

1978

[17] **"Access to justice. Vol. 3. Emerging issues and perspectives. Edited by mauro cappelletti and bryant garth"**

Mauro Cappelletti, Bryant Garth

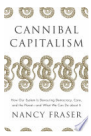
1979



[18] **"Feminism for the 99%: A Manifesto"**

Nancy Fraser

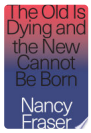
Verso Books | ISBN: 9781788734448



[19] **"Cannibal Capitalism"**

Nancy Fraser

Verso Books | ISBN: 9781804292587



[20] "The Old is Dying and the New Cannot be Born"

Nancy Fraser

Verso Books | ISBN: 9781788732741



[21] "Fundamental Principles of the Metaphysic of Morals"

Immanuel Kant

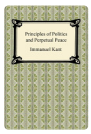
2021 | Phoemixx Classics Ebooks | ISBN: 9783986477943



[22] "Principles of Lawful Politics"

Immanuel Kant

1988



[23] "Kant's Principles of Politics and Perpetual Peace"

Immanuel Kant

2010 | Digireads.com Publishing | ISBN: 9781420938258



[24] "CBT for Synthetic Humans: Therapy, Ethics, and Healing in the Age of Engineered Minds"

Edenilson Brandl

Edenilson Brandl

Artykuły naukowe

Prace autorów przywoływanych w tekście:

[1] "Transaction-Cost Economics: The Governance of Contractual Relations"

Oliver E. Williamson

1979 | The Journal of Law and Economics | DOI: 10.1086/466942

Źródło

[2] "Engineering a Safer World: Systems Thinking Applied to Safety"

Nancy G. Leveson

2012 | The MIT Press eBooks | DOI: 10.7551/mitpress/8179.001.0001

Źródło

[3] "A new accident model for engineering safer systems"

Nancy G. Leveson

2003 | Safety Science | DOI: 10.1016/s0925-7535(03)00047-x

Źródło

[4] **"Safeware: System Safety and Computers"**

Nancy G. Leveson

1995 | CERN Document Server (European Organization for Nuclear Research)

Źródło

[5] **"Applying systems thinking to analyze and learn from events"**

Nancy G. Leveson

2010 | Safety Science | DOI: 10.1016/j.ssci.2009.12.021

Źródło

[6] **"The Social Psychology of Procedural Justice"**

E. Allan Lind, Tom R. Tyler

1988 | Critical issues in social justice | DOI: 10.1007/978-1-4899-2115-4

Źródło

[7] **"The Role of Procedural Justice and Legitimacy in Shaping Public Support for Policing"**

Jason Sunshine, Tom R. Tyler

2003 | Law & Society Review | DOI: 10.1111/1540-5893.3703002

Źródło

[8] **"A Relational Model of Authority in Groups"**

Tom R. Tyler, E. Allan Lind

1992 | Advances in experimental social psychology | DOI: 10.1016/s0065-2601(08)60283-x

Źródło

[9] **"The Group Engagement Model: Procedural Justice, Social Identity, and Cooperative Behavior"**

Tom R. Tyler, Steven L. Blader

2003 | Personality and Social Psychology Review | DOI: 10.1207/s15327957pspr0704_07

Źródło

[10] **"Markets, Money and Capital: Hicksian Economics for the Twenty First Century"**

Roberto Scazzieri, Amartya Sen, Stefano Zamagni

2009 | Cambridge University Press eBooks | DOI: 10.1017/cbo9780511575969

Źródło

[11] **"Economics, law, and ethics"**

Amartya Sen

2009 | Against Injustice | DOI: 10.1017/cbo9780511657443.002

Źródło

[12] **"Economic development and capability expansion in historical perspective"**

Amartya Sen

2001 | Pacific Economic Review | DOI: 10.1111/1468-0106.00126

Źródło

[13] **"The economic institutions of capitalism firms, markets, relational contracting"**

Oliver E. Williamson

1985

Źródło

[14] **"Comparative Economic Organization: The Analysis of Discrete Structural Alternatives"**

Oliver E. Williamson

1991 | Administrative Science Quarterly | DOI: 10.2307/2393356

Źródło

[15] **"The Economics of Organization: The Transaction Cost Approach"**

Oliver E. Williamson

1981 | American Journal of Sociology | DOI: 10.1086/227496

Źródło

[16] **"The Challenger Launch Decision: Risky Technology, Culture, and Deviance at NASA"**

Diane Vaughan

1997 | University of Chicago Press eBooks | DOI: 10.7208/chicago/9780226346960.001.0001

Źródło

[17] **"Autonomy, Interdependence, and Social Control: NASA and the Space Shuttle Challenger"**

Diane Vaughan

1990 | Administrative Science Quarterly | DOI: 10.2307/2393390

Źródło

[18] **"Rational Choice, Situated Action, and the Social Control of Organizations"**

Diane Vaughan

1998 | Law & Society Review | DOI: 10.2307/827748

Źródło

[19] **"Theorizing Disaster: Analogy, historical ethnography, and the Challenger accident"**

Diane Vaughan

2004 | Ethnography | DOI: 10.1177/1466138104045659

Źródło

[20] **"Violence, Peace, and Peace Research"**

Johan Galtung

1969 | Journal of Peace Research | DOI: 10.1177/002234336900600301

Źródło

[21] **"Cultural Violence"**

Johan Galtung

1990 | Journal of Peace Research | DOI: 10.1177/0022343390027003005

Źródło

[22] **"A Structural Theory of Imperialism"**

Johan Galtung

1971 | Journal of Peace Research | DOI: 10.1177/002234337100800201

Źródło

[23] **"On the Effects of International Economic Sanctions, With Examples from the Case of Rhodesia"**

Johan Galtung

1967 | World Politics | DOI: 10.2307/2009785

Źródło

[24] **"On Balancing and Subsumption. A Structural Comparison"**

Robert Alexy

2003 | Ratio Juris | DOI: 10.1046/j.0952-1917.2003.00244.x

Źródło

[25] "On the Structure of Legal Principles"

Robert Alexy

2000 | Ratio Juris | DOI: 10.1111/1467-9337.00157

Źródło

[26] "A Theory of Legal Argumentation: The Theory of Rational Discourse as Theory of Legal Justification"

Robert Alexy, Ruth Adler, Neil MacCormick

1989

Źródło

[27] "Constitutional Rights and Proportionality"

Robert Alexy

2014 | Revus | DOI: 10.4000/revus.2783

Źródło

[28] "Class Actions in the US Experience: The Legal Perspective"

Guido Calabresi

2012 | The Law and Economics of Class Actions in Europe | DOI: 10.4337/9781781001240.00007

Źródło

[29] "Access to Justice - Newest Wave in Worldwide Movement to Make Rights Effective"

Mauro Cappelletti, Bryant G. Garth

1978 | Buffalo law review

Źródło

[30] "Alternative Dispute Resolution Processes within the Framework of the World-Wide Access-to-Justice Movement"

Mauro Cappelletti

1993 | Modern Law Review | DOI: 10.1111/j.1468-2230.1993.tb02673.x

Źródło

[31] "Integration Through Law: Volume 1: Methods, Tools and Institutions. Book 3: Forces and Potential for a European Identity"

Mauro Cappelletti, Seccombe, Monica, Weiler, Joseph

1986 | De Gruyter eBooks | DOI: 10.1515/9783110909227

Źródło

[32] **""Who Watches the Watchmen?" A Comparative Study on Judicial Responsibility"**

Mauro Cappelletti

1983 | The American Journal of Comparative Law | DOI: 10.2307/839606

Źródło

[33] **"Global Prescriptions: The Production, Exportation, and Importation of a New Legal Orthodoxy"**

Yves Dezalay, Bryant G. Garth

2002 | University of Michigan Press eBooks | DOI: 10.3998/mpub.17196

Źródło

[34] **"Asian Legal Revivals: Lawyers in the Shadow of Empire"**

Yves Dezalay, Bryant G. Garth

2010 | University of Chicago Press eBooks | DOI: 10.7208/chicago/9780226144665.001.0001

Źródło

[35] **"Merchants of Law as Moral Entrepreneurs: Constructing International Justice from the Competition for Transnational Business Disputes"**

Yves Dezalay, Bryant G. Garth

1995 | Law & Society Review | DOI: 10.2307/3054053

Źródło

[36] **"Zwischen System und Norm"**

Nancy Fraser

2024 | DOI: 10.5771/9783518779330-555

Źródło

[37] **"Abnormal Justice"**

Nancy Fraser

2017 | Pragmatism and Justice | DOI: 10.1093/acprof:oso/9780190459239.003.0003

Źródło

[38] **"Review of Gottlieb Hufeland's Essay on the principle of natural right (1786) [translated and edited by Allen Wood]"**

Immanuel Kant, Allen W. Wood

Immanuel Kant: Practical philosophy | DOI: 10.1017/cbo9780511813306.008

Źródło

Artykuły naukowe autora:

[39] "Raising AI: How Artificial Intelligence Is Shaping Us—and How We Shape It"

Chirag Shah

2026 | DOI: 10.5040/9798216452430

Źródło

[40] "Ship It, Then Apologize: We Can Do Better Than This for AI Advancements"

Chirag Shah

2026 | SSRN Electronic Journal | DOI: 10.2139/ssrn.6952240

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: 993734b0-bc2c-4040-9cdb-b5024a1e86d0