DOI: 10.31648/sp.12572

Martyna Kaczmarczyk
Uniwersytet Warmińsko-Mazurski w Olsztynie
ORCID: 0000-0001-6169-9466
martyna.kaczmarczyk@uwm.edu.pl

Prawo do wyjaśnienia decyzji algorytmicznej jako element ochrony godności człowieka

Wstęp

Dynamiczny rozwój technologii cyfrowych doprowadził do szerokiego wykorzystywania systemów sztucznej inteligencji w procesach podejmowania decyzji dotyczących podstawowych praw jednostek. Algorytmy analizujące duże zbiory danych są obecnie stosowane w wielu obszarach życia społecznego i gospodarczego. Systemy te umożliwiają szybkie przetwarzanie ogromnych ilości informacji oraz identyfikowanie zależności, które byłyby trudne do wykrycia przy użyciu tradycyjnych metod analitycznych. Zastosowanie systemów algorytmicznych w procesach decyzyjnych przynosi liczne korzyści, w tym zwiększenie efektywności działania instytucji oraz ograniczenie kosztów operacyjnych. Jednak rozwój technologii sztucznej inteligencji rodzi szereg nowych wyzwań natury prawnej i etycznej. Jednym z problemów jest brak przejrzystości działania algorytmów, zwłaszcza tych opartych na metodach uczenia maszynowego.

W wielu przypadkach systemy algorytmiczne funkcjonują jako tzw. czarne skrzynki, których sposób działania jest trudny do zrozumienia zarówno dla użytkowników, jak i dla samych twórców technologii. Oznacza to, że osoba, której dotyczy decyzja podjęta przez system, często nie ma możliwości poznania przesłanek doprowadzających do określonego rozstrzygnięcia. Sytuacja ta rodzi istotne pytania dotyczące ochrony praw jednostki w warunkach społeczeństwa cyfrowego. Brak przejrzystości procesów decyzyjnych utrudnia bowiem kontrolę nad działaniem instytucji oraz ogranicza możliwość zakwestionowania decyzji podejmowanych przez systemy informatyczne. W konsekwencji pojawia się ryzyko naruszenia podstawowych wartości systemu prawnego, w tym poszanowania godności człowieka.

W odpowiedzi na te wyzwania unijny ustawodawca wprowadził koncepcję prawa do wyjaśnienia decyzji algorytmicznej. Prawo to polega na umożliwieniu

osobie, której dotyczy decyzja podejmowana w sposób zautomatyzowany, uzyskania informacji o logice działania systemu oraz czynnikach, które doprowadziły do określonego rozstrzygnięcia. Celem niniejszego artykułu jest analiza znaczenia prawa do wyjaśnienia decyzji algorytmicznej w kontekście ochrony godności człowieka. W obecnym stanie prawnym europejski standard zapewniania przejrzystości decyzji podejmowanych przez systemy sztucznej inteligencji nie stanowi skutecznego elementu realizacji zasady poszanowania godności jednostki w warunkach rosnącej automatyzacji procesów decyzyjnych.

Godność człowieka jako fundament ochrony praw jednostki

Godność ludzka jest wartością nadrzędną w systemach prawnych wielu państw oraz fundamentem międzynarodowego systemu praw człowieka. Jej znaczenie polega na uznaniu jednostki za podmiot autonomiczny, posiadający niezbywalne prawa i wolności, które należy chronić niezależnie od warunków społecznych, ekonomicznych czy politycznych. Współczesna teoria prawa konstytucyjnego i prawa międzynarodowego podkreśla, że godność człowieka stanowi podstawę wszystkich praw jednostki i jest źródłem zarówno wolności osobistych, jak i ochrony prawnej przed dyskryminacją, arbitralnością czy nadużyciami władzy[1].

Podstawowym dokumentem konstytuującym godność człowieka w prawie międzynarodowym jest Powszechna deklaracja praw człowieka, która w art. 1 określa, że „wszyscy ludzie rodzą się wolni i równi pod względem godności i praw”[2]. Regulacja ta stanowi fundament doktryny praw człowieka i normę interpretacyjną dla innych instrumentów prawa międzynarodowego. Kolejnym kluczowym dokumentem jest Międzynarodowy pakt praw obywatelskich i politycznych, który w art. 7 chroni przed torturami i nieludzkim traktowaniem, odwołując się pośrednio do godności człowieka, a w art. 26 gwarantuje równość wobec prawa, zapewniając, że każdy ma prawo do ochrony prawnej w sposób respektujący jego fundamentalną wartość jako osoby[3]. W ramach prawa europejskiego godność człowieka zajmuje centralne miejsce również w Karcie praw podstawowych Unii Europejskiej, gdzie w art. 1 stwierdzono, że „godność ludzka jest nienaruszalna; należy ją szanować i chronić”[4].

---

[1] L. Garlicki, *Polskie prawo konstytucyjne. Zarys wykładu*, Warszawa 2021, s. 41–45.

[2] Art. 1 Powszechnej deklaracji praw człowieka, uchwalonej przez Zgromadzenie Ogólne Organizacji Narodów Zjednoczonych dnia 10 grudnia 1948 r.

[3] Art. 7 i art. 26 Międzynarodowego paktu praw obywatelskich i politycznych otwarty do podpisu w Nowym Jorku dnia 19 grudnia 1966 r. (Dz.U. z 1977 r., Nr 38, poz. 167).

[4] Art. 1 Karty praw podstawowych Unii Europejskiej z dnia 7 grudnia 2000 r. (Dz. Urz. UE C 202/389, 7.6.2016).

Wszystkie prawa jednostki w Unii Europejskiej są interpretowane w kontekście poszanowania godności, co wpływa na ocenę polityk publicznych, decyzji administracyjnych czy stosowania technologii w życiu obywateli.

Europejski system ochrony praw człowieka stanowi jeden z podstawowych fundamentów ochrony godności ludzkiej. Chociaż Europejska konwencja praw człowieka nie wymienia godności jako samodzielnego prawa, to standard ten został wyprowadzony przede wszystkim z prawa do poszanowania życia prywatnego oraz rozwinięty w orzecznictwie Europejskiego Trybunału Praw Człowieka (dalej jako ETPC)[5]. Art. 8 EKPC gwarantuje każdemu prawo do poszanowania życia prywatnego i rodzinnego, mieszkania oraz korespondencji[6]. W kontekście ochrony danych osobowych w orzecznictwie ETPC przyjmuje się, że gromadzenie, przechowywanie, wykorzystywanie i udostępnianie danych osobowych stanowi ingerencję w sferę prywatności, która musi być przewidziana ustawą, realizować uzasadniony cel oraz spełniać wymóg konieczności i proporcjonalności w demokratycznym społeczeństwie[7]. Przepis ten stanowi zatem podstawowy filar europejskiej ochrony danych osobowych. Art. 6 EKPC gwarantuje prawo do rzetelnego procesu i ma znaczenie pośrednie. W kontekście przetwarzania danych osobowych wymaga zapewnienia jednostce skutecznych gwarancji proceduralnych, w tym możliwości poznania podstaw rozstrzygnięcia oraz zakwestionowania decyzji wpływających na jej prawa i obowiązki, co nabiera szczególnego znaczenia w przypadku decyzji opartych na zautomatyzowanym przetwarzaniu danych[8]. Art. 13 EKPC ustanawia prawo do skutecznego środka odwoławczego przed organem krajowym w razie naruszenia praw konwencyjnych, w tym prawa do prywatności. W obszarze ochrony danych osobowych oznacza to konieczność zapewnienia efektywnych mechanizmów dochodzenia roszczeń oraz kontroli legalności przetwarzania danych[9]. Natomiast art. 14 EKPC zakazuje dyskryminacji przy korzystaniu z praw i wolności gwarantowanych konwencją. W odniesieniu do przetwarzania danych osobowych przepis ten ma znaczenie zwłaszcza wtedy, gdy wykorzystanie danych lub systemów algorytmicznych prowadzi do nierównego traktowania osób ze względu na cechy chronione, takie jak płeć, pochodzenie, religia czy niepełnosprawność[10].

Uzupełnieniem standardu EKPC jest Konwencja Rady Europy nr 108+, będąca pierwszym wiążącym traktatem międzynarodowym poświęconym

---

[5] Konwencja o ochronie praw człowieka i podstawowych wolności, sporządzona w Rzymie dnia 4 listopada 1950 r. (Dz.U. z 1993 r. Nr 61, poz. 284 ze zm.), dalej jako EKPC.

[6] Ibidem, art. 8.

[7] Wyroki ETPC: z 16 lutego 2000 r. w sprawie *Amann v. Szwajcaria*, skarga nr 27798/95, §§ 65–69; z 4 grudnia 2008 r. w sprawie *S. i Marper v. Zjednoczone Królestwu*, skargi nr 30562/04 i 30566/04, §§ 67–68, 95–103.

[8] Art. 6 EKPC.

[9] Ibidem, art. 13.

[10] Ibidem, art. 14.

wyłącznie ochronie danych osobowych[11]. Konwencja 108+ modernizuje dotychczasowe standardy, wprowadzając m.in. zasadę legalności i przejrzystości przetwarzania, minimalizacji danych, rozliczalności administratorów, ochrony danych już w fazie projektowania oraz rozszerzone prawa osób, których dane dotyczą[12]. Jej postanowienia wywarły istotny wpływ na rozwój europejskich regulacji dotyczących ochrony danych, wzmacniając ochronę godności jednostki w warunkach cyfryzacji i rosnącego wykorzystania systemów opartych na sztucznej inteligencji.

W krajowym porządku prawnym Konstytucja RP jednoznacznie określa godność człowieka jako wartość nienaruszalną i fundament praw jednostki. Art. 30 ustawy zasadniczej stanowi, że „godność człowieka jest nienaruszalna; jest ona źródłem wolności i praw człowieka i obywatela"[13], a art. 31 podkreśla zakaz naruszania godności oraz zobowiązuje państwo do ochrony praw jednostki w sposób zapewniający poszanowanie jej podmiotowości[14]. W art. 32 ust. 1 Konstytucji RP została zagwarantowana równość wobec prawa i zakaz dyskryminacji z jakiejkolwiek przyczyny, co stanowi praktyczną realizację zasady poszanowania godności[15].

Godność człowieka pełni funkcję fundamentalnej zasady interpretacyjnej i normatywnej, będąc źródłem wszystkich praw i wolności jednostki[16]. Jej praktyczne znaczenie można rozważać w trzech wymiarach. Po pierwsze, stanowi podstawę ochrony elementarnych praw i wolności, takich jak prawo do życia, wolność osobista, bezpieczeństwo, wolność sumienia, prawo do prywatności czy ochrona danych osobowych. Po drugie, wprowadza zakaz dyskryminacji i nieludzkiego traktowania, nakładając na państwo i podmioty prywatne obowiązek traktowania jednostki w sposób sprawiedliwy i równy wobec prawa. Po trzecie, zapewnia podmiotowość i autonomię jednostki, gwarantując prawo do decydowania o sobie oraz aktywnego uczestnictwa w procesach prawnych, społecznych i gospodarczych[17].

Znaczenie godności człowieka w prawie międzynarodowym i krajowym przekłada się na obowiązek państw do tworzenia i egzekwowania prawa chroniącego życie, wolność i prawo do prywatności, zapewnienia równości wobec prawa oraz przeciwdziałania dyskryminacji. Godność jednostki powinna

---

11 Konwencja Rady Europy z dnia 28 stycznia 1981 r. o ochronie osób w związku z automatycznym przetwarzaniem danych osobowych (ETS No. 108), zmieniona Protokołem zmieniającym z dnia 18 maja 2018 r. (CETS No. 223), dalej jako Konwencja 108+.

12 Ibidem, art. 5–10.

13 Art. 30 Konstytucji Rzeczypospolitej Polskiej (Dz.U. z 1997 r., Nr 78, poz. 483).

14 Ibidem, art. 31.

15 Ibidem, art. 32 ust. 1.

16 R. Kozłowski, D. Bieńkowska, *Prawa człowieka i ludzka godność. Ochrona i aksjologia*, Poznań 2020, s. 8–12.

17 P. Wilczyński, *W kwestii godności człowieka jako kategorii prawa,* „Acta Universitatis Lodziensis" 2023, nr specjalny, s. 342–343.

stanowić punkt odniesienia przy tworzeniu i interpretacji nowych regulacji prawnych z zakresu technologii i procesów zautomatyzowanych. Naruszenie autonomii lub instrumentalizacja jednostki w decyzjach opartych na systemach sztucznej inteligencji może być traktowane jako naruszenie godności i podstawowych praw człowieka[18]. Godność ludzka jest fundamentem ochrony praw jednostki zarówno w prawie międzynarodowym, jak i krajowym. Uznanie jej nadrzędnej wartości pozwala chronić jednostki przed dyskryminacją, arbitralnością, nieludzkim traktowaniem oraz ograniczeniem autonomii. W polskim prawie konstytucyjnym godność stanowi źródło wolności i praw człowieka, a jej ochrona jest obowiązkiem państwa. W prawie międzynarodowym godność pełni rolę kluczowej zasady interpretacyjnej, kształtując prawa jednostki i wyznaczając standardy ich ochrony[19]. Godność człowieka w erze globalizacji i cyfryzacji, pozostaje niezbywalnym fundamentem, który pozwala zapewnić respektowanie praw jednostki w każdej sferze życia.

Decyzja algorytmiczna i jej znaczenie w procesach społecznych

Systemy sztucznej inteligencji coraz częściej wykorzystywane są w procesach decyzyjnych dotyczących jednostek. Decyzje algorytmiczne mogą mieć charakter całkowicie zautomatyzowany lub stanowić wsparcie dla rozstrzygnięć podejmowanych przez człowieka. Przykłady zastosowań takich systemów obejmują m.in. ocenę zdolności kredytowej klientów w sektorze finansowym, analizę dokumentów aplikacyjnych w procesach rekrutacyjnych, ocenę ryzyka ubezpieczeniowego czy identyfikowanie potencjalnych nadużyć w systemach świadczeń publicznych. Wykorzystanie algorytmów w procesach decyzyjnych wiąże się z szeregiem korzyści. Systemy te mogą analizować ogromne ilości danych w krótkim czasie, identyfikować złożone zależności statystyczne oraz ograniczać wpływ subiektywnych ocen podejmowanych przez ludzi. Jednocześnie jednak pojawiają się zagrożenia związane z możliwością występowania błędów systemowych oraz uprzedzeń zawartych w danych wykorzystywanych do trenowania algorytmów. Jeżeli dane historyczne zawierają określone nierówności społeczne, systemy uczące się mogą je powielać lub nawet wzmacniać[20].

---

[18] M. Kuciński, *Sztuczna inteligencja a godność człowieka*, „Fides, Ratio Et Patria. Studia Toruńskie" 2021, nr 14, s. 221–223.

[19] K. Orzeszyna, M. Skwarzyński, R. Tabaszewski, *Prawo międzynarodowe praw człowieka*, Warszawa 2022, s. 15; J. Mrozek, *Godność osoby ludzkiej jako źródło praw człowieka i obywatela*, „Civitas et Lex" 2014, nr 1, s. 41.

[20] M. Kubiak, M. Kudła, *Stronniczość i uprzedzenia sztucznej inteligencji (AI Bias) – ryzyko i sposoby jego mitygacji*, „Prawo Nowych Technologii" 2023, nr 2, s. 40–42.

W literaturze przedmiotu zwraca się uwagę na zjawisko tzw. dyskryminacji algorytmicznej, polegające na tym, że decyzje podejmowane przez systemy sztucznej inteligencji mogą prowadzić do nieuzasadnionego różnicowania sytuacji jednostek ze względu na cechy takie jak płeć, pochodzenie etniczne czy status społeczny[21]. Szczególnym problemem jest w tym kontekście brak przejrzystości działania algorytmów. W wielu przypadkach ustalenie dokładnych przyczyn podjęcia konkretnej decyzji jest zawiłe, co utrudnia identyfikację potencjalnych błędów lub uprzedzeń w działaniu systemu.

Istota prawa do wyjaśnienia decyzji algorytmicznej

Prawo do wyjaśnienia decyzji algorytmicznej stanowi wyzwanie w dobie gwałtownego rozwoju technologicznego i ma na celu zapewnienie przejrzystości działania systemów sztucznej inteligencji[22]. Jego istotą jest umożliwienie osobie, której dotyczy decyzja podjęta w sposób zautomatyzowany, uzyskania informacji o mechanizmach prowadzących do określonego rozstrzygnięcia. Funkcja ta wykracza poza samą informację – z założenia umożliwia jednostce kontrolę nad procesem decyzyjnym oraz zakwestionowanie decyzji, które mogą naruszać jej prawa lub interesy. Jednak w praktyce prawo to nie realizuje standardów przejrzystości, odpowiedzialności i sprawiedliwości proceduralnej, które stanowią instrument ochrony zarówno godności, jak i równości jednostki wobec prawa[23].

Od momentu wejścia w życie przepisów rozporządzenie Parlamentu Europejskiego i Rady (UE) 2024/1689 z dnia 13 czerwca 2024 r. w sprawie ustanowienia zharmonizowanych przepisów dotyczących sztucznej inteligencji (tzw. *AI Act*), czyli od 1 sierpnia 2024 r., prawo do wyjaśnienia decyzji algorytmicznej zyskało własną, dedykowaną systemom sztucznej inteligencji podstawę prawną[24]. W art. 86 unijny ustawodawca ustanowił prawną możliwość uzyskania wyjaśnień w kontekście stosowania w procesie decyzyjnym

---

[21] R. Rejmaniak, *Bias in Artificial Intelligence Systems*, „Białostockie Studia Prawnicze" 2021, nr 3, s. 26.

[22] E. Papadimitriou, *The right to explanation in the processing of personal data with the use of AI systems*, „International Journal of Law in Changing World" 2023, nr 2, s. 44 i nast.; K. Nizioł, *The right to obtain an explanation of the decision-making procedure in an individual case under Article 86(1) of the Artificial Intelligence Act – selected problems*, „Studia Prawnicze KUL" 2025, nr 4, s. 78 i nast.

[23] A. Powałowski, *Aksjologia nowych technologii w administracji publicznej*, „Dyskurs Prawniczy i Administracyjny" 2024, nr 3, s. 16–17.

[24] Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2024/1689 z dnia 13 czerwca 2024 r. w sprawie ustanowienia zharmonizowanych przepisów dotyczących sztucznej inteligencji oraz zmiany rozporządzeń (WE) nr 300/2008, (UE) nr 167/2013, (UE) nr 168/2013, (UE) 2018/858, (UE) 2018/1139 i (UE) 2019/2144 oraz dyrektyw 2014/90/UE, (UE) 2016/797 i (UE) 2020/1828, dalej jako *AI Act*.

systemów AI, szczególnie w odniesieniu do rozstrzygnięć mających wpływ na osoby fizyczne[25]. Zakres podmiotowy regulacji obejmuje osoby, których życie, zdrowie, bezpieczeństwo lub prawa podstawowe mogą zostać istotnie dotknięte przez decyzję opartą na systemie AI, z wyłączeniem systemów infrastruktury krytycznej[26]. Przepis ten podkreśla potrzebę ochrony praw jednostki w sytuacjach, gdy decyzje mają charakter władczy lub wywołują skutki prawne. W katalogu przesłanek unijny ustawodawca powołał się na doniosłe wartości, takie jak życie, zdrowie, bezpieczeństwo i prawa podstawowe. Niewątpliwie podane przesłanki mają szeroki zakres i w szczególności poprzez użycie sformułowania „prawa podstawowe" katalog ten jest niezwykle pojemny, co może prowadzić do nadinterpretacji. Ponadto art. 86 ust. 1 zawiera zwroty niedookreślone, które dodatkowo mogą przysparzać trudności w ocenie ewentualnych naruszeń praw i wolności. Sformułowanie „oddziałuje na tę osobę na tyle znacząco, że uważa ona, iż ma to niepożądany wpływ (…)"[27] ma charakter *stricte* subiektywny. Odczucia jednostki mogą być oparte na osobistej perspektywie, uczuciach, a nie na weryfikowalnych faktach.

Art. 86 *AI Act* pełni zatem funkcję przepisu gwarancyjnego, zapewniając osobom, których dotyczą decyzje oparte na systemach AI wysokiego ryzyka, prawo do uzyskania jasnego i zrozumiałego wyjaśnienia roli systemu opartego na sztucznej inteligencji w procesie decyzyjnym. Możliwość realizacji tego prawa opiera się na obowiązkach wynikających z innych przepisów rozporządzenia. Art. 13 stanowi podstawę informacyjną dla art. 86, ponieważ zobowiązuje dostawców do zapewnienia odpowiedniej dokumentacji i informacji o działaniu systemu[28]. Art. 14 uzupełnia ten obowiązek, wymagając skutecznego nadzoru człowieka, który umożliwia interpretację wyników systemu i ich wyjaśnienie osobie, której dotyczy decyzja[29]. Z kolei art. 26 nakłada na podmioty stosujące system AI obowiązek korzystania z niego zgodnie z instrukcjami, prowadzenia monitoringu i wykorzystywania rejestrowanych informacji, co pozwala odtworzyć przebieg procesu decyzyjnego[30]. Art. 27 wspiera realizację art. 86 poprzez obowiązek przeprowadzenia oceny wpływu na prawa podstawowe, która ułatwia uzasadnienie zastosowania systemu AI i identyfikację ryzyk dla jednostki[31]. Natomiast art. 50 wzmacnia ogólną zasadę transparentności, zobowiązując do informowania o korzystaniu z określonych

[25] Ibidem, art. 86.

[26] Ibidem, art. 86 ust. 1; z wyjątkiem pkt 2 załącznika III: „Infrastruktura krytyczna: systemy AI przeznaczone do wykorzystywania jako związane z bezpieczeństwem elementy procesów zarządzania krytyczną infrastrukturą cyfrową, ruchem drogowym i procesów ich działania lub zaopatrzenia w wodę, gaz, ciepło lub energię elektryczną".

[27] Ibidem, art. 86 ust. 1.

[28] Ibidem, art. 13.

[29] Ibidem, art. 14.

[30] Ibidem, art. 26.

[31] Ibidem, art. 27.

systemów AI, choć nie zastępuje indywidualnego prawa do wyjaśnienia przewidzianego w art. 86[32]. Wreszcie art. 85 zapewnia mechanizmy nadzoru i egzekwowania obowiązków wynikających z *AI Act*, wzmacniając praktyczną skuteczność prawa do wyjaśnienia[33]. W konsekwencji art. 86 stanowi zwieńczenie systemu obowiązków dotyczących transparentności, dokumentowania i nadzoru nad systemami AI wysokiego ryzyka, przekształcając je w konkretne uprawnienie jednostki do uzyskania wyjaśnienia decyzji wspomaganej przez sztuczną inteligencję.

Kluczowym celem regulacji jest zwiększenie przejrzystości działania systemów sztucznej inteligencji oraz ochrona praw jednostki. W praktyce oznacza to, że osoby, których dotyczą decyzje podejmowane lub wspierane przez AI, powinny mieć możliwość uzyskania zrozumiałych informacji o tym, jak i dlaczego dana decyzja została podjęta. Przepis ten wpisuje się w szerszy nurt regulacyjny Unii Europejskiej, rozszerzając istniejące mechanizmy, takie jak prawo do wyjaśnienia wynikające z rozporządzenie Parlamentu Europejskiego i Rady (UE) 2016/679 z dnia 27 kwietnia 2016 r. w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych i w sprawie swobodnego przepływu takich danych (RODO)[34]. W odróżnieniu jednak od RODO, które koncentruje się na ochronie danych osobowych, *AI Act* rozszerza perspektywę na funkcjonowanie samych systemów AI, niezależnie od tego, czy przetwarzają dane osobowe.

Analizując europejski porządek prawny dotyczący decyzji algorytmicznych, nie sposób pominąć Konwencji ramowej Rady Europy o sztucznej inteligencji, prawach człowieka, demokracji i rządach prawa[35]. Choć dokument ten nie posługuje się pojęciem decyzji algorytmicznej ani zautomatyzowanego podejmowania decyzji, postanowienia w nim zawarte pozostają istotne dla oceny dopuszczalności stosowania systemów AI w procesach decyzyjnych. Konwencja przyjmuje podejście oparte na ochronie praw człowieka, nakładając na państwa obowiązek zapewnienia odpowiednich gwarancji proceduralnych w przypadku wykorzystywania systemów sztucznej inteligencji. Szczególne znaczenie ma art. 8, który przewiduje obowiązek zapewnienia osobom dotkniętym działaniem systemów AI odpowiednich informacji, możliwości uzyskania wyjaśnień oraz skutecznych środków zaskarżenia i ochrony

[32] Ibidem, art. 50.

[33] Ibidem, art. 85.

[34] Art. 22 rozporządzenia Parlamentu Europejskiego i Rady (UE) 2016/679 z dnia 27 kwietnia 2016 r. w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych i w sprawie swobodnego przepływu takich danych (Dz. Urz. UE L 119, 4.5.2016), dalej jako RODO.

[35] Konwencja ramowa Rady Europy o sztucznej inteligencji, prawach człowieka, demokracji i rządach prawa (Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, CETS No. 225), przyjęta w Wilnie dnia 17 maja 2024 r.

prawnej[36]. W konsekwencji konwencja nie ustanawia odrębnego prawa do wyjaśnienia decyzji algorytmicznej, lecz tworzy ogólne standardy ochrony proceduralnej, stanowiące punkt odniesienia dla bardziej szczegółowych regulacji zawartych w art. 22 RODO oraz art. 86 *AI Act*. Jednocześnie jej znaczenie polega na wyznaczeniu minimalnego poziomu ochrony jednostki w sytuacjach, w których systemy AI wpływają na prawa i wolności człowieka. Konwencja akcentuje konieczność zapewnienia przejrzystości i rozliczalności stosowania systemów AI, co pośrednio wzmacnia wymóg wyjaśnialności decyzji podejmowanych z ich udziałem. Wprowadza ona również perspektywę systemową, w której ochrona jednostki nie ogranicza się do pojedynczej decyzji, lecz obejmuje cały cykl życia systemu AI. W tym sensie stanowi ona uzupełnienie standardów unijnych, koncentrując się na aksjologicznych fundamentach ochrony praw człowieka. Jej postanowienia sprzyjają interpretacji przepisów RODO i *AI Act* w sposób prowolnościowy i ukierunkowany na ochronę autonomii jednostki. W rezultacie konwencja pełni funkcję ramową, wzmacniającą spójność europejskiego systemu ochrony praw człowieka w kontekście rozwoju technologii algorytmicznych.

Prawo do wyjaśnienia decyzji algorytmicznej powinno spełniać kilka istotnych funkcji. Po pierwsze, umożliwiać jednostce zrozumienie procesów decyzyjnych wpływających na jej sytuację; po drugie, pozwalać na skuteczne zakwestionowanie decyzji, która została podjęta na jej niekorzyść; po trzecie, sprzyjać budowaniu zaufania społecznego do nowych technologii[37]. Zakres prawa do wyjaśnienia nie jest jednoznacznie określony. Pojawia się pytanie, czy wyjaśnienie powinno obejmować szczegółowe informacje dotyczące struktury algorytmu, czy też jedynie ogólne podanie czynników, które miały wpływ na decyzję. Z jednej strony zbyt ogólne wyjaśnienia mogą okazać się niewystarczające dla realnej ochrony praw jednostki. Brak przejrzystości algorytmów i brak możliwości poznania przyczyn decyzji prowadzą do traktowania jednostki instrumentalnie – jako zbioru danych – zamiast jako podmiotu prawa. Z drugiej strony ujawnienie szczegółowych informacji o działaniu algorytmu może prowadzić do naruszenia tajemnicy przedsiębiorstwa oraz własności intelektualnej. W związku z tym w doktrynie coraz częściej podkreśla się potrzebę poszukiwania kompromisu między ochroną praw jednostki a interesami podmiotów rozwijających technologie sztucznej inteligencji[38].

---

[36] Ibidem, art. 8.

[37] A. Szot, *Algorithm in the structure of administrative decision-making: a model of explainability and burden of proof for an effective legal remedy within discretionary power in public administration*, „Studia Iuridica Lublinensia" 2025, nr 5, s. 277 i nast.

[38] U. Mylly, *Transparent AI? Navigating between rules on trade secrets and access to information*, „International Review of Intellectual Property and Competition Law" 2023, nr 54, s. 19–22.

Algorytmiczna dyskryminacja a zasada równości

Algorytmiczna dyskryminacja jest konsekwencją nieuwzględnienia prawa do wyjaśnienia decyzji algorytmicznej. Polega na sytuacji, w której decyzje podejmowane przez systemy informatyczne prowadzą do nieuzasadnionego różnicowania osób ze względu na cechy chronione prawem, takie jak płeć, pochodzenie etniczne, wiek czy status społeczny. Zjawisko to rodzi pytania o zgodność decyzji algorytmicznych z zasadą równości i innymi fundamentalnymi wartościami prawa[39]. Brak przejrzystości algorytmów i ich uprzedzenia mogą prowadzić do naruszenia prawa do równego traktowania, a skuteczne mechanizmy kontroli i odpowiedzialności prawnej są kluczowe dla ochrony tej zasady.

Zasada równości, będąca fundamentem współczesnych systemów prawa, znajduje wyraz zarówno w prawie krajowym, jak i międzynarodowym. W Konstytucji RP w art. 32 ust. 1 ustanowiono, że „wszyscy są wobec prawa równi; wszyscy mają prawo do równego traktowania przez władze publiczne i prawo do ochrony przed dyskryminacją"[40]. Równość jednostek jest kluczowym elementem ładu międzynarodowego, zapewniającym ochronę praw człowieka i sprawiedliwość w stosunkach międzynarodowych. Jej fundament stanowi Powszechna deklaracja praw człowieka, która w art. 1 i art. 7[41] gwarantuje równość wobec prawa i zakaz dyskryminacji. Również Międzynarodowy pakt praw obywatelskich i politycznych odwołuje się do zasady równości, bowiem w art. 26 zapewnia równe korzystanie z praw człowieka[42]. Zasada równości pełni funkcję ochronną, gwarantując zakaz dyskryminacji oraz programową, wyznaczając kierunki działań państwa mające na celu niwelowanie nierówności społecznych[43]. W kontekście algorytmów szczególnie istotna jest funkcja ochronna, gdyż decyzje podejmowane przez systemy automatyczne mogą bezpośrednio wpływać na prawa jednostki i jej sytuację życiową.

Algorytmiczna dyskryminacja może mieć charakter jawny lub ukryty. Jawna występuje, gdy algorytm *explicite* wykorzystuje niedozwolone kryteria, np. płeć lub pochodzenie etniczne, w celu podjęcia decyzji. Ukryta ma miejsce wtedy, gdy algorytm pozornie neutralny prowadzi do nierównego traktowania, np. poprzez analizę danych historycznych zawierających strukturalne nierówności społeczne[44]. Mechanizmy dyskryminacji często wynikają

---

[39] J. Adams-Prassl, R. Binns, A. Kelly-Lyth, *Directly discriminatory algorithms*, „The Modern Law Rewiev" 2023, nr 1, s. 144–146.

[40] Art. 32 ust. 1 Konstytucji RP.

[41] Art. 1 i 7 Powszechnej deklaracji praw człowieka.

[42] Art. 26 Międzynarodowego paktu praw obywatelskich i politycznych.

[43] A. Chorążewska, *The principle of equality as a challenge to the democratic rule of law*, „Scientific Journal of Bielsko-Biala School of Finance and Law" 2024, nr 4, s. 8.

[44] M. Lingyu, *From algorithm bias to algorithm discrimination: research on the responsibility of algorithmic discrimination*, „Journal of Northeastern University" 2022, nr 1, s. 1.

z błędów w danych treningowych. Jeśli system rekrutacyjny jest trenowany na historycznych danych, w których dominowała określona płeć lub grupa społeczna, algorytm może preferować kandydatów podobnych do wcześniejszych pracowników, powielając uprzedzenia społeczne. Innym źródłem dyskryminacji może być niewłaściwe projektowanie funkcji celu algorytmu, który optymalizuje decyzje według kryteriów wydajnościowych, nie uwzględniając wartości etycznych ani norm prawnych. W konsekwencji decyzje systemu informatycznego mogą być formalnie neutralne, lecz prowadzić do wykluczenia lub marginalizacji określonych grup[45].

Praktyczne przykłady algorytmicznej dyskryminacji są liczne. W sektorze finansowym algorytmy oceny zdolności kredytowej mogą nieświadomie uprzywilejowywać osoby z określonych dzielnic miast lub o określonym pochodzeniu, bazując na historycznych danych finansowych. W rekrutacji systemy selekcji kandydatów mogą preferować osoby podobne do wcześniejszych pracowników, co skutkuje dyskryminacją ze względu na płeć lub wiek. W systemach opieki zdrowotnej algorytmy przewidujące ryzyko chorób mogą nie doszacować ryzyka w grupach mniejszościowych, jeśli dane treningowe nie są reprezentatywne. Wymienione przykłady pokazują, że algorytmy mogą nieświadomie powielać istniejące nierówności społeczne, nawet jeśli twórcy systemów nie mają intencji dyskryminacyjnych.

Prawo krajowe i międzynarodowe oferuje mechanizmy przeciwdziałania algorytmicznej dyskryminacji. W prawie Unii rozporządzenie RODO przewiduje m.in. prawo do wyjaśnienia decyzji zautomatyzowanych oraz prawo do sprzeciwu wobec takich decyzji[46]. *AI Act* definiuje systemy wysokiego ryzyka (ang. *high-risk AI systems*) i wymaga oceny ryzyka dyskryminacji przed ich wdrożeniem[47]. Dodatkowo prawo antydyskryminacyjne, w tym konwencje ONZ i normy ILO, nakładają obowiązek zapewnienia równego traktowania osób w różnych sferach życia[48]. Mechanizmy kontroli obejmują audyty algorytmiczne, testy na obecność uprzedzeń, obowiązek dokumentowania procesów decyzyjnych oraz procedury odwoławcze.

Mimo istniejących regulacji problem algorytmicznej dyskryminacji pozostaje aktualny. Identyfikacja uprzedzeń wymaga zaawansowanych narzędzi statystycznych i technik *explainable AI*, czyli tzw. wyjaśnialnej sztucznej inteligencji[49]. Ponadto wciąż niejasna pozostaje kwestia odpowiedzialności

---

45 R. Rejmaniak, op. cit., s. 27 i nast.

46 Art. 22 RODO.

47 Art. 27 I 86 *AI Act*.

48 Konwencja ONZ w sprawie likwidacji wszelkich form dyskryminacji rasowej (CERD) (Dz.U. z 1969 r., Nr 11, poz. 76); Konwencja w sprawie likwidacji wszelkich form dyskryminacji kobiet (CEDAW) (Dz.U. z 1981 r., Nr 16, poz. 87); Konwencja Międzynarodowej Organizacji Pracy nr 111 w sprawie dyskryminacji w zatrudnieniu i zawodzie (Dz.U. z 1972 r., Nr 25, poz. 160).

49 Ł. Sobczyk, *Zgodność algorytmów i samouczących się systemów z wybranymi aspektami prawa Unii Europejskiej*, „Prawo Mediów Elektronicznych. Kwartalnik Naukowy" 2024, nr 4, s. 66–69.

prawnej za decyzje algorytmiczne: kto ponosi odpowiedzialność – producent, programista czy użytkownik? W związku z tym niezbędna jest obowiązkowa ocena ryzyka dyskryminacji przed wdrożeniem systemów AI, projektowanie algorytmów uwzględniających kryteria równości, zapewnienie możliwości odwołania od decyzji algorytmicznej oraz wprowadzenie niezależnych mechanizmów audytu i kontroli.

Algorytmiczna dyskryminacja stanowi realne zagrożenie dla zasady równości w społeczeństwie cyfrowym. Systemy AI, choć pozornie neutralne, mogą powielać i wzmacniać istniejące nierówności społeczne, jeśli dane treningowe i funkcje celu nie uwzględniają wartości praw człowieka. Skuteczna ochrona zasady równości wymaga połączenia regulacji prawnych, mechanizmów kontroli oraz odpowiedzialnego projektowania systemów AI. Szczególnie ważne jest prawo do wyjaśnienia decyzji algorytmicznej, audyty algorytmiczne i odpowiedzialność twórców systemów, które pozwalają zapewnić, że automatyzacja decyzji nie prowadzi do nierównego traktowania obywateli.

Wyzwania regulacyjne w kontekście ochrony godności

Prawo do wyjaśnienia decyzji algorytmicznej, przewidziane w unijnych rozporządzeniach, powinno zapewniać jednostce możliwość poznania kryteriów decyzyjnych, logiki algorytmu oraz skutków decyzji automatycznej. Art. 22 RODO ustanawia prawo osoby do niepodlegania decyzji opierającej się wyłącznie na zautomatyzowanym przetwarzaniu, w tym profilowaniu, jeżeli wywołuje ona wobec niej skutki prawne lub w podobny sposób istotnie na nią wpływa[50]. Skuteczna realizacja tego prawa jest uzależniona od obowiązków informacyjnych przewidzianych w art. 13–15 RODO.

Art. 13 i 14 zobowiązują administratora do przekazania osobie, której dane dotyczą, informacji o stosowaniu zautomatyzowanego podejmowania decyzji, w tym profilowania, oraz do przedstawienia istotnych informacji o zasadach jego działania, a także o znaczeniu i przewidywanych konsekwencjach takiego przetwarzania[51]. Natomiast art. 15 przyznaje prawo dostępu do tych informacji już na etapie wykonywania prawa dostępu do danych[52]. W konsekwencji art. 13–15 RODO zapewniają przejrzystość procesu zautomatyzowanego podejmowania decyzji i umożliwiają osobie ustalenie, czy może skorzystać z ochrony przewidzianej w art. 22 RODO. Przepisy te pełnią zatem funkcję informacyjną i gwarancyjną wobec prawa do sprzeciwu wobec decyzji opartych wyłącznie na zautomatyzowanym przetwarzaniu oraz

[50] Art. 22 RODO.
[51] Art. 13 i 14 RODO.
[52] Art. 15 RODO.

do żądania interwencji człowieka, wyrażenia własnego stanowiska i zakwestionowania takiej decyzji. Jednak skuteczna implementacja zanalizowanych gwarancji napotyka szereg problemów regulacyjnych, zarówno natury prawnej, technologicznej, jak i praktycznej.

Pierwszym istotnym wyzwaniem jest niejednoznaczność prawna pojęcia „prawo do wyjaśnienia". W związku z nadal stosunkowo nową koncepcją prawa do wyjaśnienia decyzji algorytmicznej, ustanowioną w art. 86 *AI Act*, polemika naukowa w zakresie tego tematu dotyczy głównie prawa do wyjaśnienia decyzji zautomatyzowanych, przewidzianych w rozporządzeniu RODO. W literaturze przedmiotu pojawiają się spory, czy art. 22 rozporządzenia gwarantuje rzeczywisty „prawny obowiązek udzielenia wyjaśnień", czy jedynie możliwość uzyskania ogólnych informacji o logice algorytmu i rodzaju wykorzystywanych danych[53]. Autorzy, tacy jak Sandra Wachter, Brent Mittelstadt i Luciano Floridi, podkreślają, że brak precyzyjnych definicji prowadzi do sytuacji, w której podmioty stosujące AI mogą ograniczyć się do minimalnego udostępnienia informacji, co nie zapewnia jednostce realnej kontroli nad decyzją[54]. Rozważania te stanowią cenne spojrzenie również na wykładnię art. 86 *AI Act*.

Drugim wyzwaniem jest złożoność techniczna systemów AI. Wiele algorytmów, zwłaszcza opartych na uczeniu głębokim, działa jako tzw. „czarne skrzynki", których mechanizm decyzyjny jest trudny do odtworzenia nawet dla ekspertów[55]. Regulacje prawne powinny uwzględniać ten aspekt i określić standard, w jakim stopniu wyjaśnienia powinny być dostępne. Kolejnym problemem jest konflikt między ochroną praw jednostki a interesem przedsiębiorcy lub państwa. Ujawnienie pełnej logiki algorytmu może prowadzić do naruszenia tajemnicy handlowej lub bezpieczeństwa systemów informatycznych. Regulacje powinny zatem znaleźć równowagę między zapewnieniem przejrzystości a ochroną własności intelektualnej i bezpieczeństwa, co w praktyce jest trudne do osiągnięcia. Konieczne staje się wprowadzenie mechanizmów, które pozwolą udostępniać wyjaśnienia w formie zrozumiałej dla jednostki, bez ujawniania wrażliwych szczegółów technologicznych[56]. Innym wyzwaniem jest egzekwowanie prawa do wyjaśnienia. W praktyce jednostka może uzyskać informacje o decyzji, ale brak jasnych procedur odwoławczych i mechanizmów odpowiedzialności ogranicza realną ochronę jej praw. Wciąż jednak nie wypracowano jednolitych standardów, które określałyby, jakie

---

[53] L. Edwards, M. Veale, *Slave to the algorithm? Why a 'right to an explanation' is probably not the remedy you are looking for*, „Duke Law & Technology Review" 2017, vol. 16, s. 20 i nast.

[54] S. Wachter, B. Mittelstadt, L. Floridi, *Why a right to explanation of automated decision-making does not exist in the general data protection regulation*, „International Data Privacy Law" 2017, s. 4–5.

[55] J. Burrell, *How the machine 'thinks': understanding opacity in machine learning algorithms*, „Big Data & Society" 2016, vol. 3, s. 10.

[56] U. Mylly, op. cit., s. 20 i nast.

sankcje ponosi podmiot stosujący algorytm w przypadku naruszenia prawa do wyjaśnienia. Z tego powodu coraz częściej postulowane są obowiązkowe audyty algorytmów, rejestry decyzji automatycznych i procedury nadzorcze, które w sposób praktyczny umożliwią kontrolę zgodności decyzji z prawem[57].

Kolejnym wyzwaniem regulacyjnym jest różnorodność dziedzin stosowania systemów AI. Prawo do wyjaśnienia musi funkcjonować zarówno w sektorze finansowym, ubezpieczeniowym, administracyjnym, medycznym oraz innych. Każdy z tych sektorów generuje inne ryzyka i wymaga innych standardów przejrzystości. Regulacje uniwersalne mogą być niewystarczające, co prowadzi do potrzeby tworzenia ram sektorowych lub uzupełniających wytycznych. Konieczne jest zatem rozwijanie współpracy międzynarodowej i wypracowywanie standardów dotyczących procesów decyzyjnych opartych na systemach sztucznej inteligencji[58].

Podsumowując, wyzwania regulacyjne dla prawa do wyjaśnienia decyzji algorytmicznej obejmują zarówno problemy prawne, jak i techniczne oraz organizacyjne. Należą do nich m.in. niejednoznaczność definicji prawa do wyjaśnienia, złożoność algorytmów, konflikt między przejrzystością a tajemnicą handlową, ograniczona egzekwowalność prawa, różnorodność sektorów stosowania AI oraz brak harmonizacji regulacji na poziomie międzynarodowym. Skuteczne przeciwdziałanie tym wyzwaniom wymaga synergii prawa, technologii i praktyki, wprowadzenia procedur audytu i odpowiedzialności, a także stworzenia jasnych standardów interpretacyjnych, które umożliwią jednostce realny dostęp do wyjaśnień decyzji automatycznych, przy jednoczesnym zabezpieczeniu interesów podmiotów stosujących algorytmy. Prawo do wyjaśnienia decyzji algorytmicznej przewiduje możliwość uzyskania informacji o kryteriach stosowanych w procesie zautomatyzowanego podejmowania decyzji oraz o przewidywanych skutkach takich decyzji. Ma zatem charakter ograniczony, ponieważ nie obejmuje pełnej możliwości kwestionowania decyzji, a jedynie umożliwia uzyskanie ogólnych informacji o logice algorytmu i danych wejściowych[59].

Podsumowanie

Choć nowe technologie mogą przynosić liczne korzyści społeczne i gospodarcze, ich stosowanie wiąże się z ryzykiem naruszenia praw podstawowych, w tym godności ludzkiej. Prawo do wyjaśnienia decyzji algorytmicznej powinno stanowić istotny instrument ochrony praw podstawowych jednostki w erze

[57] S. Wachter, B. Mittelstadt, L. Floridi, op. cit., s. 21–23.

[58] J. Kabza, A. Sokołowska, *O regulowaniu AI z perspektywy law and economics w kontekście Rozporządzenia AI*, „Przedsiębiorczość – Edukacja" 2025, nr 1, s. 297, 308.

[59] B. Goodman, S. Flaxman, *European Union regulations on algorithmic decision making and a „right to explanation"*, „AI Magazine" 2017, nr 3, s. 53–56.

cyfryzacji i automatyzacji procesów decyzyjnych. Zapewnienie transparentności rozstrzygnięć, umożliwienie kontroli legalności decyzji, przeciwdziałanie dyskryminacji i wzmacnianie poczucia podmiotowości oraz autonomii człowieka stanowią kluczowe elementy pełnej realizacji prawa do wyjaśnienia. W świetle regulacji unijnych prawo to powinno realizować konstytucyjne i europejskie standardy ochrony godności, równości i sprawiedliwości proceduralnej. Jego skuteczna implementacja wymaga połączenia ram prawnych, technologicznych narzędzi oraz mechanizmów odpowiedzialności za decyzje automatyczne, co pozwala chronić godność ludzką i prawa podstawowe jednostki w społeczeństwie opartym na systemach AI.

Prawo do wyjaśnienia decyzji algorytmicznej powinno być traktowane bardziej jako postulat symboliczny niż realnie skuteczny instrument ochrony praw i godności jednostki, ponieważ w praktyce napotyka na szereg poważnych ograniczeń. Problemem jest sama złożoność współczesnych modeli algorytmicznych, zwłaszcza systemów opartych na uczeniu maszynowym. W konsekwencji wyjaśnienie sprowadza się często do ogólnikowego opisu działania modelu, który nie pozwala jednostce realnie zrozumieć, dlaczego została potraktowana w określony sposób. Co więcej, prawo to rodzi istotne napięcia z ochroną tajemnicy przedsiębiorstw i własności intelektualnej. Podmioty stosujące algorytmy mogą ograniczać zakres udzielanych informacji, powołując się na konieczność ochrony swoich rozwiązań technologicznych. W efekcie wyjaśnienia są niepełne, selektywne i pozbawione kluczowych danych, co podważa sens całej instytucji.

Obecna treść prawa do wyjaśnienia stwarza ryzyko pozorności ochrony prawnej. Samo otrzymanie wyjaśnienia nie oznacza jeszcze możliwości skutecznego zakwestionowania decyzji. Jednostki często nie dysponują wiedzą techniczną ani środkami, aby ocenić poprawność działania algorytmu, a tym bardziej wykazać jego błędy lub dyskryminujący charakter. Prawo do wyjaśnienia może więc pełnić funkcję fasadową, dając jedynie iluzję kontroli. Dodatkowo, w praktyce pojawia się problem standaryzacji wyjaśnień. Brak jednolitych wymogów co do ich formy i szczegółowości prowadzi do sytuacji, w której różne podmioty przedstawiają wyjaśnienia o skrajnie różnej jakości – od bardzo technicznych i niezrozumiałych, po nadmiernie uproszczone i mało użyteczne. W tym ujęciu prawo do wyjaśnienia decyzji algorytmicznej jawi się jako instytucja obarczona istotnymi wadami – trudna do realizacji, łatwa do obejścia i często nieskuteczna w zapewnianiu rzeczywistej ochrony praw jednostki.

Wykaz literatury

Adams-Prassl J., Binns R., Kelly-Lyth A., *Directly discriminatory algorithms*, „The Modern Law Review" 2023, nr 1.

Burrell J., *How the machine 'thinks': understanding opacity in machine learning algorithms*, „Big Data & Society" 2016, vol. 3.
Chorążewska A., *The principle of equality as a challenge to the democratic rule of law*, „Scientific Journal of Bielsko-Biala School of Finance and Law" 2024, nr 4.
Edwards L., Veale M., *Slave to the algorithm? Why a 'right to an explanation' is probably not the remedy you are looking for*, „Duke Law & Technology Review" 2017, vol. 16.
Garlicki L., *Polskie prawo konstytucyjne. Zarys wykładu*, Wolters Kluwer, Warszawa 2021.
Goodman B., Flaxman S., *European Union regulations on algorithmic decision making and a „right to explanation"*, „AI Magazine" 2017, nr 3.
Kabza J., Sokołowska A., *O regulowaniu AI z perspektywy law and economics w kontekście Rozporządzenia AI*, „Przedsiębiorczość – Edukacja" 2025, nr 1.
Kozłowski R., Bieńkowska D., *Prawa człowieka i ludzka godność. Ochrona i aksjologia*, Wyd. Nauk. Silva Rerum, Poznań 2020.
Kubiak M., Kudła M., *Stronniczość i uprzedzenia sztucznej inteligencji (AI Bias) – ryzyko i sposoby jego mitygacji*, „Prawo Nowych Technologii" 2023, nr 2.
Kuciński M., *Sztuczna inteligencja a godność człowieka*, „Fides, Ratio Et Patria. Studia Toruńskie" 2021, nr 14.
Lingyu M., *From algorithm bias to algorithm discrimination: research on the responsibility of algorithmic discrimination*, „Journal of Northeastern University" 2022, nr 1.
Mrozek J., *Godność osoby ludzkiej jako źródło praw człowieka i obywatela*, „Civitas et Lex" 2014, nr 1.
Mylly U., *Transparent AI? Navigating between rules on trade secrets and access to information*, „International Review of Intellectual Property and Competition Law" 2023, nr 54.
Nizioł K., *The right to obtain an explanation of the decision-making procedure in an individual case under Article 86(1) of the Artificial Intelligence Act – selected problems*, „Studia Prawnicze KUL" 2025, nr 4.
Orzeszyna K., Skwarzyński M., Tabaszewski R., *Prawo międzynarodowe praw człowieka*, C.H. Beck, Warszawa 2022.
Papadimitriou E., *The right to explanation in the processing of personal data with the use of AI systems*, „International Journal of Law in Changing World" 2023, nr 2.
Powałowski A., *Aksjologia nowych technologii w administracji publicznej*, „Dyskurs Prawniczy i Administracyjny" 2024, nr 3.
Rejmaniak R., *Bias in Artificial Intelligence Systems*, „Białostockie Studia Prawnicze" 2021, nr 3.
Sobczyk Ł., *Zgodność algorytmów i samouczących się systemów z wybranymi aspektami prawa Unii Europejskiej*, „Prawo Mediów Elektronicznych. Kwartalnik Naukowy" 2024, nr 4.
Szot A., *Algorithm in the structure of administrative decision-making: a model of explainability and burden of proof for an effective legal remedy within discretionary power in public administration*, „Studia Iuridica Lublinensia" 2025, nr 5.
Wachter S., Mittelstadt B., Floridi L., *Why a right to explanation of automated decision-making does not exist in the general data protection regulation*, „International Data Privacy Law" 2017.
Wilczyński P., *W kwestii godności człowieka jako kategorii prawa,* „Acta Universitatis Lodziensis" 2023, nr specjalny.

Summary

The right to an explanation of algorithmic decisions as an element of the protection of human dignity

Keywords: human rights, artificial intelligence; algorithmic decision-making; right to explanation; human dignity; algorithmic transparency; automated decision-making; human rights protection.

The article addresses the issue of protecting human dignity in the context of the growing use of artificial intelligence systems in decision-making processes. The aim of the study is to analyse the importance of the right to an explanation of algorithmic decisions as a tool for protecting individual rights in the face of increasing decision automation. The main research problem is to determine whether, and to what extent, the right to an explanation of decisions made by algorithmic systems can be an effective mechanism for protecting human dignity. The article discusses the concept of dignity in the human rights protection system, characteristics of algorithmic decisions and threats arising from the lack of transparency in algorithmic operations. Functions of the right to an explanation and its importance in ensuring the transparency of decision-making processes are also analysed. The main thesis of the article is that the right to an explanation of algorithmic decisions, in its current legal status, does not constitute an effective instrument for protecting human dignity. The analysis demonstrated that, in practice, this right does not fully guarantee the possibility of understanding the mechanisms affecting an individual's legal situation and does not provide adequate opportunities for oversight and challenge of decisions made by artificial intelligence systems.