Anna Mazur*

Wyzwania związane z zastosowaniem sztucznej inteligencji w rekrutacji a ochrona prywatności

DOI: 10.26399/meip.1(84).2025.04/a.mazur

Wprowadzenie do sztucznej inteligencji w rekrutacji: kontekst i kluczowe wyzwania

Sztuczna inteligencja (AI) odgrywa coraz większą rolę w procesach rekrutacyjnych, oferując automatyzację, lepszą analizę danych i potencjalnie bardziej efektywne dopasowanie kandydatów do stanowisk. Jednak zastosowanie AI w rekrutacji wiąże się z poważnymi wyzwaniami, zwłaszcza w kontekście ochrony prywatności, przejrzystości decyzji oraz ryzyka algorytmicznej dyskryminacji[1]. Brak odpowiednich mechanizmów kontroli i regulacji może prowadzić do naruszeń praw kandydatów oraz utrwalać systemowe nierówności w dostępie do zatrudnienia.

Jednym z głównych problemów związanych z AI w rekrutacji jest problem „nieprzeniknionej czarnej skrzynki" – sposób działania algorytmów pozostaje nieprzejrzysty dla kandydatów i nie daje im dostępu do kryteriów selekcji[2]. Modele AI, szkolone na danych historycznych, mogą wzmacniać istniejące uprzedzenia, faworyzując określone grupy społeczne lub eliminując kandydatów na podstawie nieistotnych, lecz łatwo mierzalnych cech, takich jak słowa kluczowe w CV, miejsce zamieszkania czy luki w zatrudnieniu. Brak precyzyjnych regulacji i nadzoru nad tymi systemami zwiększa ryzyko naruszeń prywatności i niesprawiedliwych decyzji rekrutacyjnych.

Istnieje widoczna luka w literaturze dotycząca kompleksowego podejścia do AI w rekrutacji, które uwzględniałoby zarówno aspekty technologiczne, jak i prawne oraz

* Uczelnia Łazarskiego, e-mail: anna.mazur@lazarski.edu.pl, ORCID: 0000-0002-5562-2734.

1 L.M. Wilkins, *Artificial intelligence in the recruitinf process: Identifying perceptions of bias*, SSRN; Z.U. Oman, A. Siddiqua, R. Noorain, *Artificial Intelligence and its ability to reduce recruitment bias*, „World Journal of Advanced Research and Reviews" 2024, nr 24(1), s. 551–564.

2 2. C. O'Neil, *Broń matematycznej zagłady. Jak big data zwiększa nierówności społeczne i zagraża demokracji*, tłum. M.Z. Zieliński, Wydawnictwo Naukowe PWN, Warszawa, 2017, s. 41–45.

etyczne. Większość badań koncentruje się na technicznych aspektach działania algorytmów – ich wydajności, selekcji danych czy wynikach predykcyjnych[3]. Mniej uwagi poświęca się natomiast wpływowi tych technologii na prawa kandydatów, transparentność procesów rekrutacyjnych oraz mechanizmy ochrony prywatności.

Cel i zakres artykułu

Celem artykułu jest analiza wyzwań związanych z wykorzystaniem AI w rekrutacji w kontekście ochrony prywatności oraz zaproponowanie rozwiązań regulacyjnych i etycznych, które mogą ograniczyć ryzyko naruszeń. W szczególności artykuł:

1. omawia problem nieprecyzyjnych definicji AI, które wpływają na skuteczność regulacji i standardy etyczne, prowadząc do nadregulacji lub niedoregulowania niektórych technologii;
2. analizuje algorytmy rekrutacyjne pod kątem ryzyka dyskryminacji, w tym mechanizmy powielania uprzedzeń historycznych i schematów preferowania określonych grup kandydatów;
3. wprowadza koncepcję hybrydowej prywatności, łączącej różne wymiary ochrony: informacyjny, dostępności fizycznej i wirtualnej oraz decyzyjny;
4. ocenia obecne regulacje prawne (RODO, AI Act) pod kątem ich skuteczności w zapewnieniu transparentności procesów rekrutacyjnych oraz ochrony kandydatów;
5. proponuje rozwiązania etyczne i prawne dla organizacji, działów HR i ustawodawców, które mogą pomóc w wypracowaniu standardów sprawiedliwej rekrutacji opartej na AI.

Oryginalność ujęcia tematu artykułu opiera się na kilku kluczowych aspektach, które wyróżniają go na tle dotychczasowych badań:

1. Nowa perspektywa na ochronę prywatności w rekrutacji AI – w artykule wprowadzono koncepcję hybrydowej prywatności, która integruje różne aspekty ochrony danych, w tym prywatność informacyjną, decyzyjną i dostępności fizycznej oraz wirtualnej. To podejście wydaje się interesujące, ponieważ większość literatury skupia się wyłącznie na ochronie danych osobowych, pomijając szeroki kontekst wpływu AI na prywatność kandydatów;
2. Uwzględnienie luki regulacyjnej i jej konsekwencji – w literaturze istnieją analizy dotyczące wpływu RODO i innych regulacji na ochronę danych, ale artykuł zwraca uwagę na niedostateczność tych regulacji w kontekście

[3] H. Ghazanfar, A.U. Hag, *Ethical and legal implications of AI in Human Rerource Management*, „Journal of Social and Organizational Matters" 2025, nr 4(2), s. 417–428; R. Binns, *Algorithmic accountability and public reason*, „Philosophy & Technology" 2018, nr 31(4), s. 1–14.

transparentności decyzji algorytmicznych oraz etycznych implikacji AI w rekrutacji. Takie podejście wypełnia lukę między analizami technicznymi a praktycznymi konsekwencjami dla kandydatów i organizacji HR;

3. Krytyczna analiza algorytmicznych uprzedzeń i ich wpływu na różnorodność zatrudnienia – artykuł nie tylko podkreśla problem algorytmicznej dyskryminacji, ale idzie o krok dalej, wskazując, w jaki sposób historyczne dane i modele AI mogą systematycznie powielać nierówności w dostępie do pracy. W literaturze pojawiają się wzmianki o tych zagrożeniach, jednak w artykule przedstawiono bardziej kompleksową analizę tego, jak AI może „dziedziczyć" systemowe uprzedzenia i jakie działania można podjąć, aby temu zapobiec;
4. Zaproponowanie praktycznych rozwiązań – artykuł nie ogranicza się do diagnozy problemów, ale również formułuje konkretne rekomendacje dla HR, ustawodawców i organizacji, mające na celu poprawę przejrzystości i etyki stosowania AI w rekrutacji. To podejście wyróżnia go na tle wielu badań, które pozostają na poziomie teoretycznym;
5. Analiza definicji AI i jej konsekwencji regulacyjnych – wprowadzenie krytycznej refleksji nad ewolucją definicji AI i jej wpływem na regulacje prawne to kolejny interesujący aspekt artykułu. Zwrócenie uwagi na to, że brak precyzyjnych definicji prowadzi do nadregulacji lub niedoregulowania AI w rekrutacji, pokazuje, że problem nie dotyczy wyłącznie technologii, ale również polityki prawnej i standardów etycznych;
6. Interdyscyplinarne podejście – artykuł łączy perspektywę technologiczną, prawną i etyczną, co rzadko pojawia się w analizach AI w rekrutacji, które zwykle skupiają się na jednym z tych aspektów;
7. Dynamiczna analiza definicji AI – wprowadzenie krytycznej refleksji nad ewolucją pojęcia AI oraz jej konsekwencjami dla regulacji i praktyki HR. Artykuł pokazuje, jak zmieniające się definicje wpływają na praktyczne wdrożenia technologii w firmach;
8. Akcent na transparentność i kontrolę użytkowników – wskazanie konkretnych mechanizmów poprawy przejrzystości AI, takich jak możliwość odwołania od decyzji algorytmu, co zwiększa kontrolę kandydatów nad procesem rekrutacji.

Struktura artykułu

Artykuł rozpoczyna się od omówienia ewolucji definicji AI, wskazując na ich ograniczenia oraz wpływ na regulacje i etykę. Następnie przedstawia problemy związane z algorytmiczną rekrutacją, koncentrując się na ryzyku algorytmicznych uprzedzeń i nieracjonalnych schematów selekcji kandydatów. W kolejnej części wprowadzona

zostaje koncepcja hybrydowej prywatności, uwzględniająca różne wymiary ochrony danych osobowych w kontekście AI. Dalej artykuł analizuje obowiązujące regulacje prawne, wskazując ich ograniczenia i konieczność reformy. Na końcu zaprezentowane zostają praktyczne rekomendacje dla organizacji i ustawodawców, mające na celu poprawę przejrzystości algorytmicznych procesów rekrutacyjnych i ochronę praw kandydatów.

Definicje, zastosowania AI w rekrutacji i regulacje prawne

Problemy i wyzwania związane z definiowaniem AI

Definiowanie sztucznej inteligencji stanowi trudne wyzwanie, ponieważ istnieje wiele różnych ujęć tego pojęcia, obejmujących zarówno filozoficzne pytania o naturę inteligencji, jak i techniczne zagadnienia związane z granicami tego, co możemy uznać za AI, a także zmieniające się oczekiwania społeczne i technologiczne. Sztuczna inteligencja to nie tylko kwestia techniczna, ale również filozoficzna, ponieważ jej definicja zależy od tego, jak rozumiemy inteligencję oraz jakie mamy oczekiwania wobec maszyn[4]. Zmieniające się oczekiwania dotyczące tej technologii sprawiają, że definicje AI muszą być dostosowywane do postępu technologicznego. Klasyfikacja systemów jako AI zmieniła się w czasie. W latach 80. XX w. systemy ekspertowe były uważane za AI. Dziś są traktowane jako klasyczne algorytmy, bo nie uczą się na danych. Podobnie kiedyś rozpoznawanie tekstu przez komputer uchodziło za AI, podczas gdy dziś jest to standardowa technologia. O tym, czy dany system może być uznany za AI, decyduje kilka czynników, takich jak regulacje prawne, standardy branżowe, perspektywa naukowa oraz sposób, w jaki dany system działa i jest wykorzystywany.

Ewolucja definicji sztucznej inteligencji (AI) stanowi kluczowy temat, który ma istotne znaczenie w kontekście jej zastosowań. Termin „sztuczna inteligencja" nie ma jednej, uniwersalnej definicji, a jego znaczenie zmieniało się na przestrzeni lat. W literaturze przedmiotu wyróżnia się kilka głównych typów definicji AI, które pełnią odmienne funkcje w kontekście zrozumienia i implementacji.

1. Definicje operacyjne – koncentrują się na tym, co systemy AI potrafią robić, niezależnie od ich wewnętrznej struktury. Klasycznym przykładem jest test Turinga, definiujący AI poprzez zdolność maszyny do przejawiania zachowań

[4] N.J. Nilsson, *The Quest for Artificial Intelligence*, Cambridge University Press, Cambridge 2009, s. 13–15, 20–23; M.A. Boden, *AI: Its Nature and Future*, Oxford University Press, 2016, s. 1–4.

inteligentnych w sposób nieodróżnialny od człowieka[5], a także pierwotna propozycja projektu badawczego z Dartmouth[6].

2. Definicje funkcjonalne – definiują AI poprzez konkretne zadania i funkcje, które realizują systemy, np. podejmowanie decyzji, rozwiązywanie problemów lub wykonywanie określonych działań. Przykładem takich definicji są prace Simona i Newella[7], które koncentrują się na zdolności systemów do logicznego wnioskowania i osiąganie celów.
3. Definicje realne – starają się opisać rzeczywiste mechanizmy działania AI, jej strukturę i architekturę. Klasycznym przykładem jest koncepcja „The Society of Mind" autorstwa Minsky'ego (1986)[8], przedstawiająca AI jako system złożony z interakcji prostych agentów.
4. Definicje sprawozdawcze – klasyfikują i porządkują różne podejścia do AI, opisując paradygmaty, algorytmy i zastosowania w sposób systematyczny. Takim przykładem jest współczesny podręcznik Russella i Norviga[9], który prezentuje przegląd metod i koncepcji w nauce o AI.
5. Definicje regulujące – mają na celu określenie ram prawnych i etycznych dla rozwoju i zastosowania AI, uwzględniając zarówno ochronę danych osobowych, jak i skutki społeczne czy etyczne implementacji technologii. Przykładem współczesnych prac w tym zakresie są analizy Wachter i Mittelstadt (2019)[10], które podkreślają konieczność tworzenia nowych mechanizmów ochrony jednostki, takich jak „prawo do rozsądnych wnioskowań".
6. Definicje prototypowe – opisują AI poprzez zestaw cech charakterystycznych, które częściowo definiują inteligentne zachowanie systemów, przy czym żadna cecha nie jest wymagana jako absolutna. Takie podejście pozwala uchwycić elastyczność w rozumieniu AI i jest szczególnie przydatne w kontekście

---

5 A. Turing, *Computing machinery and intelligence*, „Mind" 1950, nr 59(236), s. 433–460, https://doi.org/10.1093/mind/LIX.236.433.

6 J. McCarthy et all, *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*, Darmouth College, Hanover (NH) 1955; M. Minsky, *The society of mind*, MIT Press, Cambridge (MA) 1986, s. 17–42.

7 H.A. Simon, A. Newell, *Human Problem Solving*, Prentice-Hall, Englewood Cliffs, NJ 1972.

8 M. Minsky, *The Society of Mind*, MIT Press, Simon & Schuster, New York 1986, s. 17–42.

9 S.J. Russell, P. Norvig, *Artificial Intelligence: A Modern Approach* (4th ed.), Pearson, Hoboken (NJ) 2020.

10 S. Wachter, B. Mittelstadt, *A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI*, „Columbia Business Law Review" 2019, nr 2, s. 494–620.

zastosowań praktycznych, np. w systemach rekrutacyjnych, gdzie różne funkcje AI mogą występować w różnym zakresie[11].

Każdy z tych typów definicji pełni odmienną rolę i jest przydatny w innych kontekstach – zarówno teoretycznych, jak i praktycznych. Jednak nie wszystkie definicje są adekwatne do współczesnych wyzwań związanych z AI, co może prowadzić do nieporozumień i błędów w jej stosowaniu, zwłaszcza w obszarach takich jak rekrutacja.

Definiowanie sztucznej inteligencji stanowi kluczowy element dyskusji nad tą dziedziną, ale wiąże się z wieloma trudnościami, zarówno terminologicznymi, jak i koncepcyjnymi. Jak wcześniej wspomniano, w literaturze przedmiotu wyróżnia się wiele rodzajów definicji AI, które pełnią odmienne funkcje i mają różne zastosowania. Już samo pojęcie „inteligencji" jest trudne do jednoznacznego zdefiniowania, co sprawia, że jest ono rozmyte i wieloznaczne. W tym kontekście często odwołuje się do poglądów Johna Searle'a, który wskazywał, że złożone aspekty umysłu, takie jak świadomość i intencjonalność nie poddają się prostym definicjom[12]. Searle, znany ze swojego eksperymentu myślowego „Chiński Pokój" podkreślał, że choć maszyny mogą wykonywać zadania, które wyglądają na inteligentne, to nie oznacza, że naprawdę rozumieją, co robią. To stanowisko, choć wpływowe, nie jest jedynym w debacie nad naturą inteligencji. W opozycji do niego Daniel Dennett w swoich pracach, takich jak *Consciousness Explained* argumentuje, że ludzka świadomość może być rozumiana jako rodzaj „wirtualnej maszyny", w której subiektywne doświadczenie wynika z równoległej architektury mózgu. Z tego punktu widzenia brak subiektywnego doświadczenia w maszynach nie wyklucza ich zdolności do inteligentnego działania[13]. Dennett zauważa, że to, co z pozoru może wyglądać na „brak rozumienia" w zachowaniu maszyny, w rzeczywistości może być wyrazem innego rodzaju inteligencji.

Inteligencja, zarówno ludzka, jak i sztuczna, jest pojęciem, które nie daje się łatwo zdefiniować. W klasycznym ujęciu często mówi się, że inteligencja to to, co mierzy test na inteligencję, jak w przypadku testów IQ. Może być to użyteczne w określonych sytuacjach, ale nie wyczerpuje pełnego zakresu tego pojęcia. Jednak pojęcie inteligencji jest dużo bardziej złożone i obejmuje różne aspekty, takie jak zdolność rozwiązywania problemów, rozumowanie, uczenie się czy kreatywność, które nie są adekwatne do tradycyjnych testów IQ. Sztuczna inteligencja wciąż jest w fazie rozwoju, więc jej zdolności w tych obszarach są różne w zależności od konkretnego systemu. Właśnie dlatego

11 N.J. Nilsson, *Artificial Intelligence: A New Synthesis*, Morgan Kaufmann, San Francisco 1998, s. 2–10.

12 J. Searle, *Mind, language and society: Philosophy in the real world*, Basic Books, New York 1998, s. 40.

13 D. Dennett, *Consciousness explained*, Little, Brown and Company New York, Boston, London 1991, s. 210.

nie istnieje powszechnie akceptowana definicja, ponieważ AI może przejawiać się w różnorodny sposób – od prostych algorytmów po zaawansowane systemy uczące się.

Na początku badań nad AI, jak zauważa Duch, celem było nauczenie maszyn tego, co potrafi człowiek, poprzez tworzenie algorytmów, które miały imitować ludzkie procesy poznawcze. Jednak ze względu na złożoność niektórych problemów nie wszystko da się opisać za pomocą algorytmów. W XXI w. naukowcy zmienili podejście i zamiast uczyć maszyny naśladowania ludzkich zdolności, koncentrują się na budowaniu systemów, które potrafią uczyć się samodzielnie. Kluczową rolę odgrywają dziś neurokognitywne technologie, które umożliwiają rozwój takich samodzielnych systemów, adaptujących się do nowych sytuacji i danych[14].

Te różne perspektywy na inteligencję i sztuczną inteligencję prowadzą do konieczności uwzględnienia ewolucji definicji AI, która wciąż dostosowuje się do nowych wyzwań i osiągnięć w tej dziedzinie.

Ewolucja definicji AI

Pierwsza definicja „sztucznej inteligencji", która miała istotny wpływ na rozwój tej dziedziny, pochodzi od Alana Turinga. W 1950 r., w artykule *Computing Machinery and Intelligence*, zaproponował on test Turinga – metodę oceny czy maszyna może myśleć w sposób nieodróżnialny od człowieka. Turing uznał, że sztuczna inteligencja to zdolność maszyny do naśladowania ludzkich reakcji, takich jak rozmowa, rozwiązywanie problemów czy rozumowanie. Stwierdził: „Jeśli maszyna potrafi przejść test, w którym nie można odróżnić jej odpowiedzi od odpowiedzi człowieka, to można ją uznać za inteligentną"[15]. Definicja ta jest przykładem definicji operacyjnej, ponieważ koncentruje się na tym, co system potrafi zrobić, a nie na tym, czym inteligencja jest. Z tego powodu test Turinga, mimo historycznej wartości, podlega współczesnej krytyce, np. w 2023 r. eksperci wskazali, że modele językowe, takie jak ChatGPT, mogą „zdawać" test Turinga, nie wykazując rzeczywistego rozumienia ani zdolności do działania w świecie realnym. M. Suleyman, proponuje w tym kontekście „nową wersję testu Turinga", w którym AI otrzymuje abstrakcyjny cel do zrealizowania (np. pomnożenie określonej sumy pieniędzy), co lepiej odzwierciedla zdolność systemu do planowania i inteligentnego działania[16]. Sama imitacja ludzkiej rozmowy nie wystarcza,

[14] W. Duch, *Informatyka neurokognitywna. Stan obecny, zastosowania, perspektywy*, Seminarium, Politechnika Wrocławska, Wrocław 10 lutego 2021, https://staff-ksi.pwr.edu.pl/seminariumITT/pdf/05-02-21.pdf [dostęp: 10.08.2025].

[15] A. Turing, *Computing machinery and intelligence*, „Mind" 1950, nr 59(236), s. 433–460.

[16] M. Suleyman, komentarz w podkaście Have a Nice Future, WIRED, 16 sierpnia 2023, https://www.wired.com/story/have-a-nice-future-podcast-18/ [dostęp: 16.11.2025]; M. Suleyman, *My new Turing Test*, Mustafa Suleyman Official Website, 2023, https://mustafa-suleyman.ai/my-new-turing-test [dostęp16.11.2025].

aby stwierdzić, że maszyna posiada świadomość czy zdolność refleksji – w rzeczywistości porusza się jedynie w ramach predykcji i symulacji[17].

Kolejna definicja, zaproponowana przez Johna McCarthy'ego w 1955 r., określa sztuczną inteligencję jako „sztuczną konstrukcję, której celem jest wykonywanie zadań uznawanych za inteligentne przez ludzi"[18]. W tym ujęciu AI jest technologią, która ma naśladować zdolności człowieka w rozwiązywaniu problemów, podejmowaniu decyzji i rozumowaniu. Ta definicja najlepiej odpowiada definicjom funkcjonalnym, ponieważ koncentruje się na zadaniach i funkcjach realizowanych przez systemy, a nie na naturze inteligencji jako takiej. Choć jest użyteczna w kontekście projektowania i klasyfikowania systemów AI, nie wyjaśnia, czym jest sama inteligencja, co pokazuje niedoskonałość tego podejścia w rozumieniu AI w pełni. Definicja McCarthy'ego prowadzi jednak do kilku błędów: antropomorfizacji, utożsamiania inteligencji z zachowaniem oraz błędu *ignotum per ignotum*, ponieważ zakłada znajomość pojęcia „inteligencji", nie precyzując go. W efekcie, definicja nie rozróżnia między AI a tradycyjnym oprogramowaniem realizującym złożone operacje, co utrudnia budowanie spójnych ram teoretycznych.

Marvin Minsky, pionier sztucznej inteligencji i współtwórca MIT AI Lab, wniósł istotny wkład zarówno w rozwój definicji, jak i filozofii sztucznej inteligencji. W artykule *Steps Toward Artificial Intelligence* (1960) wskazywał, że jednym z kluczowych zadań AI jest konstruowanie maszyn zdolnych do heurystycznego rozwiązywania problemów, uczenia się i planowania[19]. Jego ujęcie ma głównie charakter funkcjonalny – opisuje zdolności, jakie powinny posiadać systemy, nie definiując samej inteligencji. W książce *The Society of Mind* (1986) Minsky zaproponował natomiast definicję realną, opartą na architekturze systemu. Inteligencja jest efektem interakcji wielu prostych „agentów", których kooperacja prowadzi do złożonych procesów poznawczych[20]. Jakkolwiek innowacyjne, podejście to bywa interpretowane jako oparte na założeniu, że inteligencja to suma „czynności podobnych do ludzkich", co ponownie prowadzi do błędu *ignotum per ignotum*.

W latach 70. i 80. XX w. Nils J. Nilsson rozwinął pojęcie funkcjonalne i praktyczne. W *Principles of Artificial Intelligence* (1980) argumentował, że AI to dziedzina zajmująca

---

17 M. Suleyman, *DeepMind co-founder suggest new Turing test for AI chatbots*, „Business Insider", 6 czerwca 2023, https://www.businessinsider.com/deepmind-co-founder-suggests-new-turing-test-ai-chatbots-report-2023-6 [dostęp: 16.11.2025].

18 J. McCarthy, M. Minsky, N. Rochester, C.E. Shannon, *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*, „AI Magazine" 2006, nr 27(4), s. 12–14, https://ojs.aaai.org/aimagazine/index.php/aimagazine/article/view/1904 [dostęp:16.11.2025].

19 M. Minsky, *Steps Toward Artificial Intelligence*, Dept. of Mathematics & Research Lab. of Electronics, MIT, 1960, https://www.cs.bham.ac.uk/research/projects/cogaff/misc/minsky/minsky-steps.pdf [dostęp: 16.11.2025].

20 M. Minsky, *The society..., op. cit.*,, s. 339.

się projektowaniem systemów zdolnych do wykonywania zadań wymagających analizy, planowania i rozumowania – choć nie definiuje inteligencji jako odrębnej cechy, a raczej wskazuje, jakie czynności są uznawane za inteligentne[21]. Definicja ta ma charakter sprawozdawczy, ponieważ opisuje paradygmaty działania systemów AI, nie oferując jednak ogólnej teorii inteligencji.

Współczesny podręcznik Stuarta Russella i Petera Norviga *Artificial Intelligence: A Modern Approach* (2021) jest jednym z ważnych przykładów definicji sprawozdawczej. Autorzy podkreślają, że AI to nauka i technologia tworzenia maszyn zdolnych do rozwiązywania problemów, uczenia się, rozumowania i podejmowania racjonalnych decyzji[22]. Wprowadzają przy tym klasyfikację systemów AI według czterech typów: (1) myślących jak ludzie, (2) myślących racjonalnie, (3) działających jak ludzie oraz (4) działających racjonalnie. Ich podejście pomaga porządkować różne podejścia do AI, ale – podobnie jak w przypadku Nilssona – nie dostarcza definicji inteligencji samej w sobie, co może prowadzić do nieporozumień przy próbach interpretacji, czym jest AI i czym powinna być w praktyce.

W polskim kontekście Wodzisław Duch proponuje ujęcie realne, koncentrując się na modelowaniu wiedzy i tworzeniu systemów zdolnych do inteligentnego rozwiązywania problemów. Wskazuje, że AI powinna być rozumiana jako nauka o systemach przetwarzających wiedzę, przy czym kluczowe staje się wykorzystanie uczenia maszynowego[23]. Duch zwraca uwagę na ważną rolę uczenia maszynowego w procesach decyzyjnych, co pozwala na bardziej zaawansowaną i adaptacyjną rolę AI, np. w rekrutacji, gdzie najnowsze algorytmy mogą uczyć się na podstawie złożonych zbiorów danych o kandydatach i podejmować decyzje bardziej dostosowane do rzeczywistych potrzeb organizacji.

Z kolei OpenAI koncentruje się na rozwoju *Artificial General Intelligence* (AGI) – ogólnej sztucznej inteligencji, która mogłaby dorównywać lub przewyższać ludzi w większości zadań poznawczych. OpenAI, definiuje AI (AGI) jako systemy zdolne do wykonywania szerokiego zakresu zadań poznawczych porównywalnych z ludzkimi, obejmujących m.in. rozumowanie, uczenie się, generowanie treści i analizę danych[24].

Opis OpenAI ma charakter definicji funkcjonalnej, ponieważ określa AGI poprzez zakres realizowanych funkcji, a nie poprzez wskazanie jej istoty czy mechanizmów działania. Ujęcie OpenAI jest przydatne operacyjnie, lecz nie definiuje inteligencji samej w sobie, co rodzi ryzyko błędnych interpretacji.

---

21 N.J. Nilsson, *Principles of artificial intelligence*, Springler-Verlag, Berlin, Heidelberg 1982, s. 476.

22 S.J. Russell, P. Norvig, *Artificial Intelligence: A Modern Approac*h, Pearson, Hoboken, NJ: Pearson 2020, s. 1136.

23 W. Duch, *Informatyka neurokognitywna…*, op. cit.

24 OpenAI, Artificial general intelligence (AGI), 2023, https://openai.com/research/agi [dostęp: 10.08.2025].

Może też sprzyjać błędnym interpretacjom, w których funkcjonalna sprawność systemu mylona jest z rozumieniem, intencjonalnością czy innymi cechami właściwymi ludzkiej inteligencji. W praktyce wpływa to m.in. na rozwój narzędzi opartych na AI, wykorzystywanych także w obszarze rekrutacji, takich jak systemy generowania tekstów, kodu, obrazów czy analizy danych. Brak wyraźnie określonych granic pojęcia AGI utrudnia jednak jednoznaczne rozstrzyganie, które technologie rzeczywiście spełniają kryteria sztucznej inteligencji w tym szerokim ujęciu.

W ujęciu teoretycznym warto także wskazać podejście prototypowe, które zakłada, że pojęcia funkcjonują na zasadzie przykładów, a nie ostrych granic. Przykładem są klasyczne prace Nilssona (1998), wskazujące na prototypowe właściwości systemów AI. Zamiast próbować znaleźć uniwersalną definicję, badacze koncentrują się na analizie systemów uznawanych za „inteligentne" w określonym kontekście[25]. Podejście to pozwala traktować AI jako spektrum systemów o różnej bliskości do „prototypowej inteligencji", co jest bardziej elastyczne niż dążenie do sztywnej definicji, jak w przypadku pojęcia „ptaka", gdzie wróbel jest bardziej typowy niż kura czy pingwin[26].

Podejście prototypowe znajduje podstawy we współczesnej kognitywistyce i filozofii nauki, które wskazują na to, że wiele pojęć nie ma ostrych granic[27]. Koncepcja „podobieństwa rodzinnego" Wittgensteina pokazuje, że pojęcia są definiowane przez powiązane cechy, a nie sztywne warunki konieczne i wystarczające. Z kolei filozofowie tacy jak Karl Popper[28] i Thomas Kuhn[29] wskazują, że brak jednoznacznej definicji nie jest błędem, lecz wynikiem dynamicznego rozwoju tej dziedziny.

Podejście prototypowe umożliwia elastyczniejsze traktowanie systemów AI, choć nie likwiduje trudności związanych z klasyfikacją i implementacją systemów. AI może być rozumiane na różne sposoby, np. jako system uczący się na podstawie danych (uczenie maszynowe), podejmujący decyzje w złożonych środowiskach (systemy ekspertowe) czy przetwarzający język naturalny (rozpoznawanie mowy). Każdy z tych systemów pełni inną funkcję, ale niekoniecznie odpowiada jednej uniwersalnej definicji. Zamiast szukać jednego wyjaśnienia, bardziej produktywne jest badanie różnorodności form inteligencji maszynowej i ich funkcji w określonych kontekstach.

---

25 K. Frankish, W. Ramsey, *The Cambridge Handbook of Artificial Intelligence*, Cambridge University Press, Cambridge 2014, s. 1–14.

26 G. Lakoff, *Kobiety, ogień i rzeczy niebezpieczne: Co kategorie mówią nam o umyśle*, tłum. T. Brzezińska, Wydawnictwo Universitas, Kraków 2011, s. 62–63.

27 L. Wittgenstein, *Dociekania filozoficzne*, tłum. B. Wolniewicz, Wydawnictwo Naukowe PWN, Warszawa 2000, s. 61–68.

28 K. Popper, *Logika odkrycia naukowego,* tłum. J.M. Kowalski, Wydawnictwo Naukowe PWN, Warszawa 2009, s. 40–55.

29 T.S. Kuhn, *Struktura rewolucji naukowych*, tłum. M.M. Ławniczak, Wydawnictwo Znak, Kraków 2019, s. 45–68.

Pojęcie „sztucznej inteligencji" nie poddaje się klasycznemu definiowaniu. Brak jednoznacznej definicji prowadzi do różnic w podejściu do regulacji prawnych i zarządzania AI oraz do odmiennego rozumienia tego pojęcia przez różnych autorów i w różnych dziedzinach nauki[30]. Tradycyjne podejście, oparte na precyzyjnych i jednoznacznych granicach pojęć, staje się coraz mniej adekwatne w kontekście złożoności i dynamicznego charakteru AI[31]. Definicje McCarthy'ego[32] i Minsky'ego[33] traktują AI jako zdolność maszyn do wykonywania zadań wymagających inteligencji, jednak krytyka tych definicji wskazuje na brak precyzyjnych granic tego pojęcia.

Podobne trudności w definiowaniu AI przenoszą się także na sferę prawa i regulacji, gdzie brak precyzyjnych kryteriów utrudnia jednoznaczne klasyfikowanie systemów oraz określanie odpowiedzialności za ich działanie.

Definicyjna niejednoznaczność AI a uregulowania prawne

Trudności w jednoznacznym zdefiniowaniu sztucznej inteligencji mają istotne konsekwencje dla regulacji prawnych oraz zarządzania nią. Różne podejścia do definicji (w zależności od kontekstu technologicznego, naukowego czy prawnego) wpływają na zakres technologii objętych regulacjami, mechanizmy kontrolne oraz potencjalne luki prawne. Nieprecyzyjna definicja może prowadzić do sytuacji, w której regulacje obejmują technologie niemające związku z AI, podczas gdy zaawansowane systemy mogą pozostawać poza kontrolą. Przykładowo, w prawodawstwie mogą występować różnice w tym, jak traktuje się systemy AI w kontekście odpowiedzialności za błędy czy bezpieczeństwo, co może wpływać na sposób ich implementacji i regulacji.

Przykładem kontrowersyjnego ujęcia definicji AI jest propozycja IBM. Na stronie IBM *What is Artificial Intelligence?* autorzy wskazują, że sztuczna inteligencja to technologia umożliwiająca komputerom i maszynom symulowanie ludzkich procesów poznawczych, w tym uczenia się, rozumienia, rozwiązywania problemów, podejmowania decyzji, kreatywności oraz autonomii[34]. Choć definicja obejmuje wiele kluczowych aspektów inteligencji, jej szeroki zakres może prowadzić do niejednoznaczności w praktyce, zwłaszcza w kontekście rekrutacji i HR, gdzie firmy mogą określać swoje produkty jako AI, mimo że w rzeczywistości wykorzystują jedynie ograniczone algorytmy, np. do filtrowania CV. Brak wyraźnych kryteriów pozwala

30 L. Floridi, *The Ethics of Artificial Intelligence*, Oxford University Press, Oxford 2023, s. 3–13.

31 S.J. Russell, P. Norvig, *Artificial Intelligence…*, op. cit.

32 J. McCarthy, *Dartmouth Summer Research…*, op. cit.

33 M. Minsky, *The Society..., op. cit.*.

34 IBM, *Artificial intelligence in practice*, 2024, https://www.ibm.com/artificial-intelligence-in-practice [dostęp: 10.08.2025]; C. Stryker, E. Kavlakoglu, *What is Artificial Intelligence?*, „Think", https://www.ibm.com/think/topics/artificial-intelligence [dostęp: 10.08.2025].

również na różne interpretacje pojęć, takich jak kreatywność czy autonomia, które nie są obecne we wszystkich systemach AI.

Podejście IBM dobrze ilustruje problem szerokich definicji AI: obejmują one wiele potencjalnych funkcji i zastosowań, ale nie wyznaczają jasnej granicy między AI a zaawansowanym oprogramowaniem. W praktyce może to powodować trudności w klasyfikacji systemów, określaniu odpowiedzialności i tworzeniu regulacji prawnych dla różnych form sztucznej inteligencji. Może to skutkować trudnościami w określeniu, które systemy podlegają regulacjom, dotyczącym przejrzystości i etyki oraz ryzykiem mylących praktyk marketingowych. Definicja IBM nie odnosi się do kluczowych technologii AI, takich jak uczenie maszynowe czy sieci neuronowe, co utrudnia legislatorom określenie, które systemy wymagają kontroli. Może to też umożliwić firmom uchylanie się od regulacji, twierdząc, że ich rozwiązania nie są „prawdziwą AI". Sugerowana przez IBM autonomia AI może prowadzić do błędnych interpretacji odpowiedzialności prawnej. Jeśli odrzuci kandydata na podstawie nieprzejrzystości kryteriów, pojawia się pytanie, kto ponosi odpowiedzialność. W ramach AI Act (2024) i RODO wymaga się przejrzystości decyzji systemów rekrutacyjnych.

Brak odniesienia do ryzyka błędów AI i uprzedzeń w danych może prowadzić do systemowej dyskryminacji w rekrutacji, co wiąże się z koniecznością audytowania modeli AI. Problemy te pokazują, jak szeroka i nieprecyzyjna definicja AI może utrudniać skuteczne regulowanie tej technologii. Komisja Europejska[35] oraz AI Act[36] zaproponowały precyzyjniejsze definicje prawne.

Definicje legalne AI: porównanie podejść KE i AI Act

Komisja Europejska oraz AI Act zaproponowały definicje regulujące AI, które mają kluczowe znaczenie dla skuteczności prawa. Różnią się one pod względem skuteczności egzekwowania, adaptacyjności oraz określenia granic regulacyjnych, co wpływa na ich efektywność. Obie definicje stanowią jednak wstępne próby regulacji AI, które z pewnością będą podlegały dalszym zmianom w miarę rozwoju technologii.

---

35 Komisja Europejska, *Proposal for a regulation laying down harmonised rules on artificial intelligence (AI Act)*, *COM (2021) 206 final*, Bruksela, 21 kwietnia 2021, https://op.europa.eu/en/publication-detail/-/publication/e0649735-a372-11eb-9585-01aa75ed71a1 [dostęp: 10.08.2025].

36 Parlament Europejski i Rada, Rozporządzenie (UE) 2024/1689 z dnia 13 czerwca 2024 r. w sprawie ustanowienia zharmonizowanych przepisów dotyczących sztucznej inteligencji (AI Act), Dziennik Urzędowy Unii Europejskiej I, nr 1689, 12 lipca 2024, https://www.si-dla-sprawiedliwosci.gov.pl/publikacja-ai-act [dostęp: 16.11.2025].

Tabela 1.
Definicje AI w regulacjach KE vs. AI Act

Źródło	Definicja	Ocena regulacyjna	Skuteczność egzekwowania	Adaptacyjność	Granice regulacyjna
Komisja Europejska (2021)	„Sztuczna inteligencja oznacza oprogramowanie, które może dla określonego zestawu celów, określonych przez człowieka, generować wyniki, takie jak przewidywania, rekomendacje lub decyzje, które wpływają na rzeczywistość lub wirtualne środowisko"	Zbyt szeroka, obejmuje systemy, które nie są AI. Brak wyraźnej granicy między AI a klasycznym oprogramowaniem	Ograniczona – niejasność definicji utrudnia skuteczną regulację	Średnia – obejmuje przyszłe systemy, ale ryzyko nadregulacji	Niejasne – może obejmować klasyczne systemy eksperckie i algorytmy decyzyjne
AI Act (2024)	„System sztucznej inteligencji oznacza oprogramowanie opracowane przy użyciu technik i podejść sztucznej inteligencji, które może generować treści, przewidywania, rekomendacje lub podejmować decyzje wpływające na środowisko fizyczne lub wirtualne"	Uwzględnia uczenie maszynowe i logikę. Obejmuje wszystkie istotne systemy. Mimo to nadal nie rozwiązuje wszystkich problemów związanych z regulacją AI, jak ostre granice między AI a zaawansowanymi algorytmami	Wysoka – klasyfikacja ryzyka ułatwia egzekwowanie	Wysoka – obejmuje różne podejścia do AI, ale wymaga aktualizacji	Lepsze – ale granice nadal nieprecyzyjne

Źródło: opracowanie własne.

Analiza definicji regulacyjnych zaproponowanych przez Komisję Europejską i AI Act pokazuje, że każda z tych definicji ma swoje ograniczenia, które mogą wpłynąć na skuteczność regulacji w rzeczywistości. Chociaż definicje regulacyjne są bardziej praktyczne niż realne lub nominalne (ze względu na konieczność dostosowania do prawodawstwa i praktyki stosowania), to wciąż nie są wolne od błędów i luk, które mogą prowadzić do trudności w egzekwowaniu, adaptacji i ustalaniu precyzyjnych granic.

1. Problemy definicji AI zaproponowanej przez Komisję Europejską

Definicja jest szeroka, ale problematyczna, ponieważ obejmuje systemy, które nie są AI. Przykładowo, może obejmować klasyczne algorytmy i systemy eksperckie, które nie spełniają kryteriów inteligencji maszynowej. Taka szerokość definicji może prowadzić do niezamierzonej nadregulacji, ponieważ klasyczne systemy algorytmiczne mogą zostać objęte regulacjami opracowanymi dla systemów AI. Brak wyraźnej granicy między systemami AI a tradycyjnym oprogramowaniem. Zamiast opierać się na szerokiej definicji, lepiej byłoby zbudować prototypy, które testują konkretne techniki lub funkcje AI w różnych scenariuszach, aby zobaczyć, które aspekty regulacji wymagają uwagi. Dobrze zaprojektowane prototypy mogłyby wyjaśnić granice technologii i jej zastosowań.

Głównym problemem tej definicji jest brak odniesienia do mechanizmów uczenia maszynowego i samo-uczenia się systemów. Jak podkreśla Duch, definicja KE koncentruje się na efektach działania systemu, zamiast precyzować jego wewnętrzne mechanizmy, co prowadzi do sytuacji, w której za AI mogą zostać uznane tradycyjne systemy rekomendacyjne lub modele statystyczne[37]. Taki brak precyzji utrudnia skuteczne egzekwowanie prawa i może powodować niezamierzone konsekwencje, takie jak objęcie regulacjami technologii, które nie niosą ze sobą ryzyka właściwego dla systemów AI.

Dodatkowo definicja KE nie uwzględnia wprost systemów uczących się na podstawie danych użytkowników, co oznacza, że bardziej zaawansowane modele AI mogą funkcjonować poza zakresem regulacji. Duch wskazuje, że brak odniesienia do mechanizmów AI może prowadzić do sytuacji, w której systemy analizujące dane osobowe, ale niewytwarzające przewidywań czy rekomendacji, nie będą podlegać odpowiednim ograniczeniom prawnym, co stwarza zagrożenie dla ochrony prywatności użytkowników.

2. Problemy definicji AI zaproponowanej przez AI Act (2024)

Definicja zawężona jest do systemów, które opierają się na technikach AI, takich jak uczenie maszynowe, co jest bardziej precyzyjne niż definicja Komisji Europejskiej. Jednak wciąż występują problemy związane z brakiem ostrej granicy między systemami AI a zaawansowanymi algorytmami, które nie wykorzystują jej technik, ale mają podobny wpływ na środowisko (np. algorytmy decyzyjne oparte na regułach).

Definicja jest wciąż nieprecyzyjna, jeśli chodzi o granice między AI a bardziej tradycyjnymi algorytmami. Może to prowadzić do niejednoznaczności w regulowaniu systemów, które nie są „pełnoprawnym" AI. Budowanie prototypów mogłoby pomóc w identyfikowaniu, które systemy rzeczywiście wymagają regulacji, a które mogą być traktowane jako zaawansowane narzędzia decyzyjne, ale nie AI. Przykładowo, można

37 W. Duch, *Informatyka neurokognitywna…*, op. cit.

opracować prototypy systemów AI i tradycyjnych algorytmów decyzyjnych, aby sprawdzić, jak są traktowane w praktyce przez istniejące regulacje.

Brak podziału na słabą i silną AI może skutkować objęciem jednolitymi regulacjami wszystkich systemów – od prostych chatbotów po zaawansowane modele autonomiczne – co nie oddaje rzeczywistego ryzyka tych technologii. Klasyfikacja ryzyka w AI Act (podział na niskie, wysokie i nieakceptowalne ryzyko) to krok w dobrą stronę, ale wymaga doprecyzowania. Brakuje wytycznych dotyczących konkretnych technologii AI wysokiego ryzyka oraz standardów audytu, co może prowadzić do niejednoznaczności w egzekwowaniu przepisów i różnych interpretacji prawa w krajach UE.

Zarówno definicje Komisji Europejskiej, jak i AI Act są użyteczne w kontekście regulacji, ale każda z nich ma swoje słabe punkty, które mogą prowadzić do niezamierzonych konsekwencji w praktyce. Budowa prototypów zamiast stosowania tradycyjnych definicji nominalnych lub realnych może być bardziej efektywna, ponieważ pozwala na testowanie i dostosowanie regulacji do rzeczywistych przypadków użycia AI. Prototypy pozwolą na określenie, które systemy powinny podlegać regulacjom, a które mogą zostać pominięte, dzięki czemu regulacje staną się bardziej precyzyjne i skuteczne.

Regulacje AI łączą podejścia techniczne, etyczne i prawne, co sprawia, że nie można ich jednoznacznie zakwalifikować jako zamknięty system regulacyjny. Z jednej strony AI Act wprowadza precyzyjne kategorie ryzyka, ale z drugiej – pozostawia otwarte kwestie dotyczące odpowiedzialności, audytów i przyszłych modyfikacji jej definicji. Ponadto system regulacyjny AI jest w istocie eksperymentalny – legislatorzy testują jego skuteczność w praktyce, co oznacza, że mechanizmy kontroli i egzekwowania mogą wymagać dalszych iteracji. Można więc mówić o prawie w fazie prototypu, które będzie podlegało ciągłym aktualizacjom wraz z rozwojem technologii AI.

Chociaż AI Act dostarcza pierwszej kompleksowej regulacji AI w UE, jego skuteczność będzie wymagała praktycznej weryfikacji. System regulacyjny AI można traktować jako dynamiczny proces, który prawdopodobnie będzie podlegał dalszym modyfikacjom, w miarę jak technologie AI będą się rozwijać. Wyzwania związane z definicją AI, hybrydowością regulacji oraz ich prototypowym charakterem wskazują, że ostateczny kształt przepisów wciąż wymaga dalszych prac i dostosowań.

Zastosowania AI w procesach rekrutacyjnych

Po zaprezentowaniu przeglądu definicji sztucznej inteligencji oraz ich ewolucji, warto przyjrzeć się praktycznym zastosowaniom AI w rekrutacji. Różne ujęcia tego pojęcia wpływają na wybór i implementację algorytmów wykorzystywanych w procesach selekcji kandydatów. Systemy rekrutacyjne oparte na prostych regułach również mogą być uznawane za AI, jeśli zawierają elementy predykcyjne i adaptacyjne. Przykładem są ATS (*Applicant Tracking Systems*), takie jak WorkDay Recruit, które stosują zarówno

filtrowanie CV na podstawie słów kluczowych, jak i algorytmy uczenia maszynowego do porządkowania kandydatów.

AI znajduje zastosowanie na różnych etapach rekrutacji – od analizy CV po ocenę kompetencji miękkich kandydatów. Technologie te zwiększają efektywność i redukują koszty procesów rekrutacyjnych, ale jednocześnie mogą prowadzić do poważnych wyzwań, w tym uprzedzeń algorytmicznych oraz naruszeń prywatności kandydatów[38].

Ajunwa[39] podkreślała, że automatyzacja decyzji rekrutacyjnych, mimo że bywa postrzegana jako interwencja redukująca uprzedzenia, może w praktyce replikować i wzmacniać dyskryminację ze względu na swoje strukturalne cechy.

Zastosowania AI w procesie rekrutacyjnym:

1. Automatyzacja przetwarzania CV. AI przyspiesza selekcję kandydatów, analizując tysiące CV i identyfikując tych, którzy spełniają określone wymagania. Przykładem jest system HireVue, który analizował również nagrania wideo kandydatów, oceniając ton głosu i ekspresję twarzy. Jednak w 2020 r. firma wycofała analizę twarzy po krytyce dotyczącej prywatności i uprzedzeń[40].
2. Analiza danych kandydatów. Wykorzystanie NLP (*Natural Language Processing*) pozwala algorytmom AI na ocenę treści CV, listów motywacyjnych oraz profili na LinkedIn. JPMorgan wdrożyło system rekrutacyjny oparty na AI, który preferował absolwentów prestiżowych uczelni, co podnosiło kwestie sprawiedliwości i ograniczenia różnorodności kandydatów[41].
3. Chatboty rekrutacyjne. Firmy, takie jak Unilever czy McDonald's korzystają z chatbotów AI do prowadzenia wstępnych rozmów rekrutacyjnych, odpowiadając na pytania kandydatów i eliminując tych, którzy nie spełniają podstawowych kryteriów. Choć AI usprawnia proces, może prowadzić do odrzucenia wartościowych kandydatów na podstawie błędnych algorytmów oceny, tj. algorytmów, które nie uwzględniają pełnej gamy kompetencji lub które dyskryminują na podstawie nieistotnych cech[42].
4. Ocena umiejętności i dopasowania kulturowego. Niektóre systemy rekrutacyjne analizują mowę ciała i wyraz twarzy, oceniając ich kompetencje miękkie. Amazon stosował AI do oceny aplikacji na stanowiska techniczne, jednak

38 M. Bogen, A. Rieke, *Help Wanted: An Examination of Hiring Algorithms, Equity, and Bias*, Upturn, 2018, s. 3–47.

39 I. Ajunwa, *The Paradox of Automation as Anti-Bias Intervention*, „Cardozo Law Review" 2020, nr 41(3), s. 1671–1741, https://larc.cardozo.yu.edu/clr/vol41/iss5/2 [dostęp: 16.11.2025].

40 I. Ajunwa, *The Paradox of…*, op. cit., s. 1671–1741.

41 C. O'Neil, *Weapons of math destruction…*, op. cit.

42 A. Berdowska, *Wspomaganie procesu rekrutacji pracowników za pomocą chatbotów – analiza wybranych rozwiązań*, „Projektowanie i analiza komunikacji w organizacji" 2018, nr 5(124), s. 93–112.

system wykazywał uprzedzenia płciowe, faworyzując mężczyzn. Został wycofany z uwagi na te kontrowersje[43].

5. Predykcja sukcesu zawodowego. AI prognozuje, którzy kandydaci mają największe szanse na sukces w danej roli, wykorzystując informacje historyczne i identyfikując wzorce sukcesu. Jednak analiza twarzy i głosu, jak w przypadku HireVue, może prowadzić do nieuzasadnionej dyskryminacji kandydatów[44].
6. Optymalizacja procesów rekrutacyjnych. LinkedIn Recruiter wykorzystuje algorytmy AI do sugerowania najbardziej odpowiednich kandydatów, co pomaga rekruterom efektywnie zarządzać talentami. Jednak rodzi to pytania o przejrzystość kryteriów rekomendacji kandydatów[45].

Te przypadki pokazują, że mimo wielu korzyści AI w rekrutacji, systemy te niosą ze sobą również ryzyko. Zastosowanie AI usprawnia proces selekcji, ale wiąże się z poziomem ingerencji w prywatność kandydatów[46]. Różne technologie rekrutacyjne różnią się pod względem zakresu zbieranych danych, stopnia automatyzacji decyzji oraz możliwości kandydatów do ich korekty.

---

43 J. Dastin, *Amazon scraps secret AI recruiting tool that showed bias against women*, „Reuters", 10.10. 2018, https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G [dostęp: 10.08.2025].

44 M. Raghavan, S. Barocas, J. Kleinberg, K. Levy, *Mitigaiting Bias in Algorithmic Hiring: Evaluating Claims and Practices*, [w:] *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT*'20)*, New York: ACM 2020, s. 469–481, https://doi.org/10.1145/3351095.3372828 [dostęp: 16.11.2025].

45 M. Bogen, A. Rieke, *Help Wanted…*, op. cit.

46 I. Ajunwa, *The Paradox of…*, op. cit.

Tabela 2.
Poziom ingerencji w prywatność w systemach AI stosowanych w rekrutacji

Kryterium	Analiza CV	Monitoring wideo	Social media screening	Profilowanie AI	Analiza mikroekspresji	Testy kompetencyjne (śledzenie aktywności)
Ocena	5/10	8/10	7/10	9/10	10/10	6/10
Świadoma zgoda	Kandydat świadomie dostarcza CV	Kandydat wie o nagrywaniu, ale nie zawsze o analizie danych	Kandydat często nie wie, że analizowany jest jego profil	Kandydaci rzadko wiedzą, jak działa system	Kandydat nie zawsze wie o analizie ekspresji	Kandydat wie o ocenie, ale nie zawsze o analizie aktywności
Zakres danych	Ograniczony do informacji podanych przez kandydata	Analizowana treść rozmowy, ekspresja, gesty, emocje	Dane z mediów społecznościowych, w tym opinie i interakcje	Szeroki – CV, testy, zachowania online	Mimika, gesty, emocje	Wyniki testu, czas reakcji, sposób nawigacji w systemie
Rodzaj danych	Głównie dane obiektywne (doświadczenie, wykształcenie)	Dane behawioralne, mogą być wrażliwe	Mogą obejmować dane wrażliwe (poglądy, religia)	Kombinacja danych obiektywnych i behawioralnych	Bardzo wrażliwe dane o stanie emocjonalnym	Dane behawioralne i poznawcze, mniej inwazyjne niż analiza emocji
Automatyczność decyzji	Wspomaga decyzję, ale zwykle nie jest jedynym czynnikiem	AI wspiera ocenę, ale decyzję podejmuje człowiek	AI klasyfikuje kandydatów na podstawie aktywności	AI może eliminować kandydatów bez udziału człowieka	AI sugeruje wynik na podstawie mikroekspresji	AI często automatycznie ocenia wyniki testów
Możliwość korekty	Kandydat może edytować swoje CV i ponownie aplikować	Brak wpływu na analizę po nagraniu	Kandydat nie ma wpływu na interpretację testu	Brak przejrzystości decyzji, utrudnione odwołanie	Kandydat nie może kontrolować reakcji	Kandydat może przystąpić do testu ponownie, ale nie ma wpływu na algorytm
Czas przechowywania	Zależy od polityki firmy, często przechowywane na przyszłość	Wideo może być przechowywane przez firmę rekrutacyjną	Dane mogą być wykorzystywane długo po aplikacji	Dane mogą być używane do przyszłych analiz	Analiza może być archiwizowana i używana przez AI	Dane testowe mogą być przechowywane przez firmę lub platformę

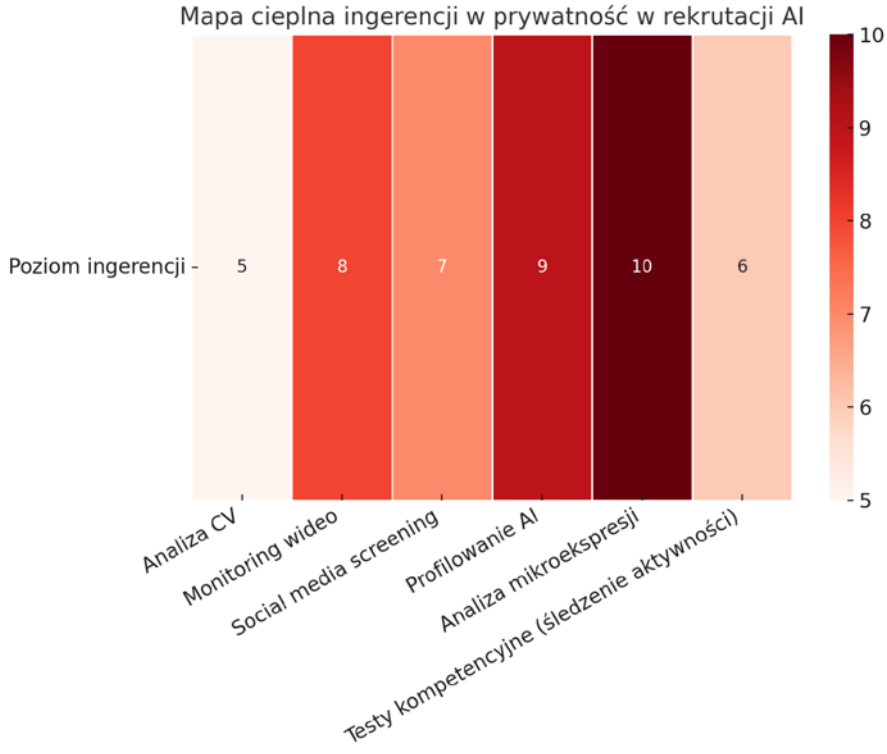
Źródło: opracowanie własne.

Tabela 2 ukazuje kluczowe aspekty wpływu AI na prywatność, w tym:

1. zakres zbieranych danych – od podstawowych informacji zawartych w CV po zaawansowaną analizę behawioralną, w tym mikroekspresję twarzy i aktywność w mediach społecznościowych;
2. automatyczność decyzji – niektóre systemy AI wspomagają rekruterów, podczas gdy inne całkowicie eliminują kandydatów bez udziału człowieka;
3. możliwość korekty – różne technologie umożliwiają kandydatom większy lub mniejszy wpływ na interpretację ich danych.

Analiza ta pokazuje, że niektóre metody selekcji są bardziej inwazyjne niż inne, co rodzi istotne pytania o etykę i przejrzystość algorytmów rekrutacyjnych. Aby lepiej zobrazować poziom ingerencji poszczególnych technologii w prywatność kandydatów, poniżej przedstawiono mapę cieplną, która ilustruje, które metody rekrutacyjne niosą ze sobą największe ryzyko w zakresie automatyzacji decyzji, zakresu zbieranych danych oraz możliwości korekty. Im intensywniejsze zaznaczenie, tym większa ingerencja w prywatność kandydata.

Wykres 1.
Mapa cieplna ingerencji w prywatność w rekrutacji AI



Źródło: opracowanie własne.

Przykłady te pokazują, że mimo wielu korzyści, AI w rekrutacji niesie również ryzyko. Brak regulacji i przejrzystości algorytmów może prowadzić do naruszeń praw kandydatów, uprzedzeń w ocenie oraz inwigilacji[47]. W odpowiedzi na te wyzwania Unia Europejska wprowadziła regulacje, takie jak AI Act, które mają zapewnić większą przejrzystość algorytmów rekrutacyjnych. RODO również nakłada obowiązek informowania kandydatów o sposobie przetwarzania ich danych. Kluczowym wyzwaniem pozostaje jednak skuteczne wdrażanie etycznych standardów oraz monitorowanie wpływu algorytmów na decyzje rekrutacyjne[48].

Paradoks racjonalności w algorytmicznych decyzjach rekrutacyjnych

Współczesne systemy rekrutacyjne mogą wzmacniać uprzedzenia, np. preferować kandydatów z określonych uczelni lub firm. Mogą też prowadzić do nieoptymalnych wyborów, jeśli opierają się na sztywnych regułach decyzyjnych.

Rozważmy algorytm stosujący preferencje leksykograficzne, gdzie decyzje opierają się na hierarchii kryteriów[49]. Najpierw porównuje się kandydatów do pracy pod kątem IQ, a dopiero przy minimalnych różnicach, brane jest pod uwagę doświadczenie zawodowe. Jeśli algorytm opiera się na ściśle określonej hierarchii kryteriów, może dojść do nielogicznych lub indyferentnych wyborów.

Załóżmy, że preferujemy kandydatów wg IQ, a przy nierozróżnialnych wynikach decyduje doświadczenie: IQ (inteligencja) jest uważane za ważniejsze niż doświadczenie. Doświadczenie zawodowe rozstrzyga wybór tylko wtedy, gdy różnice w IQ są minimalne, tj. poniżej błędu pomiaru.

Załóżmy, że:

1. kandydaci x, y, z różnią się poziomem IQ i doświadczenia;
2. IQ jest ważniejsze niż doświadczenie, więc zawsze preferujemy kandydata z wyższym IQ;
3. gdy IQ dwóch kandydatów jest nierozróżnialne (różnica mniejsza niż błąd pomiaru η), decyduje doświadczenie;
4. pomiar zarówno IQ, jak i doświadczenia obarczony jest błędem – η dla IQ i $\eta*$ dla doświadczenia;
5. preferencje są określone według leksykograficznej reguły decyzyjnej (najpierw IQ, potem doświadczenie).

---

47 I. Ajunwa, *The Paradox of…*, op. cit.

48 C. O'Neil, *Weapons of math destruction…*, op. cit.

49 P. Świskak, *Filozofia nauki i nauki społeczne po okresie pozytywizmu logicznego*, „Edukacja Filozoficzna" 1998, nr 26, s. 21–35.

Jeśli algorytm stosuje zasadę „IQ decyduje, a doświadczenie decyduje tylko w razie remisu", to nie może dokonać jednoznacznego wyboru.

Teraz przeanalizujmy wybór kandydatów:

1. kandydat x i kandydat y: ich IQ jest nierozróżnialne ($|IQx - IQy| < \eta$), więc wybieramy x, bo ma większe doświadczenie;
2. kandydat y i kandydat z: ich IQ również jest nierozróżnialne, więc wybieramy y, bo ma większe doświadczenie;
3. kandydat z i kandydat x: tutaj IQ z jest wyraźnie wyższe niż IQ x (poza granicą błędu η), więc preferujemy z.

Powstaje sprzeczność: Preferujemy x nad y i y nad z, ale jednocześnie preferujemy z nad x. Otrzymujemy cykliczną strukturę preferencji: x preferujemy nad y i y preferujemy nad z, ale z preferujemy nad x, co narusza warunek przechodniości, który jest podstawowym wymogiem racjonalności decyzji[50].

W efekcie algorytm podejmie nieracjonalny lub indyferentny wybór. Chociaż każda pojedyncza decyzja jest podejmowana w zgodzie z regułami racjonalności, całościowo prowadzi do niemożności dokonania jednoznacznego wyboru. Oznacza to, że postępując racjonalnie, nie możemy podjąć racjonalnej decyzji, co stanowi paradoks i podważa możliwość wyboru najlepszego kandydata.

Możemy zaprojektować system rekrutacyjny AI, który:

1. oceni kandydatów na podstawie dwóch cech – np. IQ i doświadczenia;
2. przypisze większą wagę IQ – jeśli IQ kandydata A jest wyższe niż B, zawsze preferuje A, niezależnie od doświadczenia;
3. gdy IQ jest w granicach błędu eta, AI porównuje doświadczenie – jeśli różnice w IQ są nierozróżnialne, AI wybiera kandydata z większym doświadczeniem;
4. wprowadzi błędy pomiaru IQ i doświadczenia – wartości mogą się różnić w zależności od losowego błędu eta i eta*;
5. użyje danych, które prowadzą do cyklicznych preferencji – np. AI oceni:
    a. $x \approx y$ pod względem IQ $\rightarrow$ wybiera x, bo ma większe doświadczenie;
    b. $y \approx z$ pod względem IQ $\rightarrow$ wybiera y, bo ma większe doświadczenie;
    c. $z > x$ pod względem IQ $\rightarrow$ wybiera z.

W efekcie system preferuje x nad y, y nad z, ale jednocześnie z nad x, co prowadzi do niespójności decyzyjnej.

Konsekwencje tego paradoksu obejmują:

1. niespójne wybory – algorytm może podejmować różne decyzje dla tych samych kandydatów;

[50] Ibidem.

2. arbitralność decyzji (brak transparentności) – AI może losowo preferować jednego kandydata przy minimalnych różnicach (mniejszych niż wielkość błędu). Może to prowadzić do sytuacji, w której wynik zależy nie od kwalifikacji, ale od przypadkowych fluktuacji w danych;
3. zagrożenie dla sprawiedliwości rekrutacji – trudno wyjaśnić, dlaczego AI odrzuca lub akceptuje kandydatów.

AI, opierając się na regułach logicznych (np. IQ > doświadczenie), może powielać ten sam problem. Może nie być w stanie wybrać najlepszego kandydata, bo błędy pomiaru (eta i eta*) są mniejsze niż margines błędu. W tej sytuacji, AI może podjąć sprzeczne decyzje, np. preferując A nad B, B nad C, ale potem C nad A, co zapętla proces decyzyjny[51].

Stosowanie probabilistycznych modeli oceny kandydatów (tj. traktowanie IQ i doświadczenia jako rozkładów prawdopodobieństwa) nie rozwiąże problemu, ponieważ wciąż może występować sytuacja, w której różnice między kandydatami są na tyle niewielkie, że decyzje podejmowane przez model są arbitralne lub losowe, co wciąż prowadzi do cyklu preferencji. Podobnie wprowadzenie elastycznych reguł decyzyjnych (elastyczne ważenie kryteriów w zależności od kontekstu danej rekrutacji) nie eliminuje całkowicie problemu. Elastyczność w regułach może prowadzić do dodawania nowych, arbitralnych kryteriów, co może wprowadzać niezamierzoną stronniczość i faworyzowanie określonych grup kandydatów. Dodatkowo nie rozwiązuje problemu przechodniości, ponieważ wciąż może dojść do sytuacji, w której kandydaci są nierozróżnialni na podstawie kryteriów, a decyzje są subiektywne lub losowe. AI powinno wyjaśniać podjęte decyzje, natomiast firmy powinny monitorować, czy jej decyzje są spójne i sprawiedliwe. Ta metoda wprowadza mechanizm nadzoru, który pozwala na identyfikowanie i eliminowanie przypadków niespójności i dyskryminacji. Tylko regularny audyt oraz zapewnienie przejrzystości i odpowiedzialności w procesach decyzyjnych pozwolą zapobiec nierównościom i uprzedzeniom, które mogą wystąpić.

Nowoczesne algorytmy rekrutacyjne rzadziej wykorzystują sztywne reguły hierarchiczne, stosując:

- modele probabilistyczne – zamiast sztywnego „IQ > doświadczenie", analizują, jak każda cecha wpływa na skuteczność kandydata w pracy;
- uczenie nadzorowane – algorytm uczy się na podstawie danych historycznych, które wskazują, jakie kombinacje cech prowadziły do sukcesu w danej firmie;
- modele hybrydowe – łączą elementy reguł ekspertowych z podejściem statystycznym, dostosowując się do realiów rynku pracy.

[51] Ibidem.

Choć nowoczesne algorytmy rekrutacyjne rzadziej bazują na sztywnych regułach, problem preferencji leksykograficznych nie zniknął całkowicie. Systemy hybrydowe łączą tradycyjne metody selekcji z AI, integrując ręcznie ustawione reguły z dynamicznymi algorytmami. Przykładem są systemy ATS, jak eRecruiter, który automatycznie analizuje CV pod kątem zgodności z wymaganiami. Chatboty rekrutacyjne oceniają odpowiedzi kandydatów, łącząc pytania kwalifikacyjne z analizą AI. Z kolei Pymetrics analizuje nagrania rozmów, uwzględniając nie tylko treść, ale także mowę ciała i ton głosu.

Paradoks preferencji leksykograficznych

Nowe podejścia do rekrutacji, zwłaszcza w kontekście zaawansowanych systemów hybrydowych, zmieniają sposób, w jaki rozumiemy problem preferencji leksykograficznych w praktyce rekrutacyjnej. Odwołując się do stwierdzenia Fishburna, który sugerował, że problem nieprzechodniości preferencji pojawia się głównie w modelach teoretycznych i nie jest powszechnie stosowany w praktycznych zastosowaniach, warto zauważyć, że współczesne podejście do rekrutacji, w tym wykorzystywanie systemów hybrydowych, zmienia ten obraz. Fishburn zakładał, że w praktyce preferencje leksykograficzne zazwyczaj są rozwiązywane lub modyfikowane, aby uniknąć cykli. Z perspektywy ówczesnych technologii i podejść do decyzji było to uzasadnione. Jednak obecnie, w kontekście rosnącego wykorzystywania sztucznej inteligencji i systemów hybrydowych w rekrutacji, problem nieprzechodniości preferencji staje się coraz bardziej aktualny i może występować częściej niż zakładał Fishburn[52].

Paradoks preferencji leksykograficznych, który w teorii rozwiązywał problem nieprzechodniości, w rzeczywistości staje się wyzwaniem dla systemów rekrutacyjnych opartych na algorytmach. Jest to realny problem w procesach rekrutacyjnych dużych przedsiębiorstw, który jest często niewidoczny, maskowany lub nieodpowiednio adresowany przez osoby odpowiedzialne za projektowanie i monitorowanie systemów AI. Szacuje się, że całkowite prawdopodobieństwo wystąpienia cykli preferencji w rekrutacji może wynosić od 5 do 40%, w zależności od rodzaju systemu rekrutacyjnego, specyfiki branży i organizacji, jakości algorytmów oraz nadzoru nad procesami decyzyjnymi[53]. W przypadku dużych organizacji, które prowadzą wiele procesów rekrutacyjnych jednocześnie, ryzyko pojawienia się cyklu preferencji staje się realne. Paradoks preferencji

---

52 P.C. Fishburn, *Utility Theory for Decision Making*, Wiley, New York 1970.

53 Opracowanie własne autorki na podstawie analizy literatury i charakterystyki systemów rekrutacyjnych opartych na algorytmach. Wartości prawdopodobieństwa cykli preferencji (5–40%) mają charakter szacunkowy i hipotetyczny.

leksykograficznych może wystąpić szczególnie wtedy, gdy różnice między kandydatami są minimalne, co jest częste w branżach takich jak marketing, technologie, niektóre role w HR, a mniej prawdopodobne w IT, zdrowiu czy administracji publicznej.

W literaturze często pojawia się przekonanie, że paradoks preferencji leksykograficznych jest wyłącznie problemem teoretycznym, który nie wpływa na praktykę[54]. Jednak, jak pokazują współczesne badania, takie podejście jest niepełne, szczególnie w kontekście systemów rekrutacyjnych[55]. Arrow i Sen omawiali ten problem w ramach ogólnych teorii decyzji, lecz nie wskazywali jednoznacznie, że ten paradoks nie może wpływać na rzeczywiste procesy decyzyjne. Przykład Piotra Świstaka pokazuje, że w kontekście rekrutacji na prestiżowe uczelnie – takich jak University of Maryland, gdy kandydaci są nieodróżnialni na podstawie formalnych kryteriów (np. IQ, doświadczenie), paradoks preferencji leksykograficznych staje się istotnym wyzwaniem w procesie podejmowania decyzji. Taki przypadek stanowi konkretne potwierdzenie, że paradoks ten może mieć realny wpływ na praktykę rekrutacyjną, zwłaszcza w złożonych procesach selekcyjnych, gdzie różnice między kandydatami są minimalne[56].

Paradoks preferencji leksykograficznych w praktyce rekrutacyjnej prowadzi do szeregu negatywnych konsekwencji. Po pierwsze, generuje niesprawiedliwość i arbitralność wyborów. W sytuacjach, gdzie kandydaci są nierozróżnialni na podstawie przyjętych kryteriów (np. IQ i doświadczenie), system selekcji – niezależnie od tego, czy oparty na sztywnych regułach, czy algorytmach AI – może wybierać losowo lub na podstawie sztywnych przesłanek, co podważa obiektywność i sprawiedliwość procesu decyzyjnego. Po drugie, w wyniku trudności z dokonaniem jednoznacznego wyboru, systemy rekrutacyjne często dodają nowe kryteria, takie jak „dopasowanie kulturowe" czy „kompetencje miękkie", co może prowadzić do ukrytej dyskryminacji. Jeśli takie kryteria nie są istotne dla stanowiska, ich wprowadzenie nie rozwiązuje problemu i może go wręcz maskować, prowadząc do faworyzowania kandydatów o podobnych cechach do rekrutera lub innych grup. Może to także prowadzić do niezamierzonych błędów, co dodatkowo komplikuje proces selekcji.

W systemach opartych na sztucznej inteligencji paradoks preferencji leksykograficznych nadal stanowi istotne wyzwanie, także w algorytmach hybrydowych. Nawet jeśli AI analizuje szereg zmiennych i stara się uwzględnić elastyczność w ocenie kandydatów, to brak jednoznaczności w różnicach między nimi może skutkować losowym wyborem lub podjęciem decyzji na podstawie arbitralnych przesłanek, które nie mają merytorycznego uzasadnienia. Taki proces nie tylko podważa transparentność całego

54 K.J. Arrow, *Social choice and individual values* (2nd ed.), Wiley & Sons, New York, London, 1951; A. Sen, *Collective choice and social welfare*, Holden-Day, San Francisco 1970.

55 Techsetter, *Dyskryminacja w rekrutacji: Dlaczego AI powiela stereotypy?*, https://techsetter.pl/dyskryminacja-w-rekrutacji-dlaczego-ai-powiela-stereotypy/. [dostęp: 10.08.2025].

56 P. Świskak, *Filozofia nauki i nauki społeczne…*, op. cit.

procesu rekrutacyjnego, ale także może prowadzić do naruszeń etycznych i prawnych związanych z równością szans w rekrutacji.

Warto zwrócić uwagę na to, że klasyczne systemy ekspertowe, które były szeroko stosowane w rekrutacji, bazowały na regułach, jednak nowoczesne modele uczenia maszynowego uczą się na podstawie danych, co zmniejsza sztywność zasad. Jednak wciąż stosuje się systemy regułowe w procesach rekrutacji, szczególnie w takich sektorach jak finanse, zdrowie, sektor edukacyjny, administracja publiczna oraz sektor IT i technologiczny, które wymagają ścisłej zgodności z politykami oraz przepisami wewnętrznymi. Preferuje się w nich takie systemy, ponieważ umożliwiają pełną kontrolę nad procesem decyzyjnym, co jest niezbędne w branżach, w których decyzje muszą być zgodne z określonymi normami prawnymi, etycznymi lub regulacjami dotyczącymi równości szans i przeciwdziałania dyskryminacji.

Szacunkowo w Polsce preferencje leksykograficzne mogą stanowić od 5 do 20% wszystkich procesów rekrutacyjnych, zwłaszcza w przypadkach, gdy stosuje się systemy regułowe, oparte na hierarchii kryteriów, takich jak IQ, doświadczenie zawodowe czy inne sztywne parametry[57]. Z kolei, w kontekście systemów hybrydowych, które łączą preferencje leksykograficzne z bardziej zaawansowanymi metodami uczenia maszynowego można zauważyć, że takie podejście znajduje zastosowanie w rekrutacji, gdzie organizacje muszą balansować między spełnieniem określonych wymagań formalnych a elastycznością w ocenie kandydatów. Dzięki połączeniu ustalonych reguł (np. wymagania dotyczące wykształcenia, doświadczenia czy umiejętności) z dynamicznymi analizami opartymi na danych (np. preferencje kulturowe, potencjał rozwoju), proces rekrutacyjny staje się bardziej kompleksowy i dopasowany do realiów współczesnego rynku pracy.

W praktyce w sektorach takich jak IT, technologie, edukacja czy administracja publiczna, które wymagają zarówno zgodności z przepisami, jak i dużej elastyczności w dopasowaniu kandydatów, zastosowanie systemów hybrydowych w rekrutacji staje się coraz bardziej powszechne. Hybrydowe podejścia pozwalają na zastosowanie „twardych" kryteriów, które są niezbędne do spełnienia norm prawnych czy regulacyjnych, a jednocześnie pozwala na

[57] Szacowanie odsetka preferencji leksykograficznych w rekrutacji opiera się na kilku czynnikach. Przede wszystkim dotyczy to branż i sektorów, w których procesy rekrutacyjne są silnie regulowane przez przepisy prawne lub wewnętrzne polityki. W takich branżach, jak finanse, zdrowie, edukacja, administracja publiczna czy IT, preferencje leksykograficzne mogą stanowić od 5 do 20% procesów rekrutacyjnych, głównie w przypadkach, gdy kluczowe są ściśle określone wymagania (np. minimalny poziom wykształcenia, doświadczenie zawodowe). W takich przypadkach preferencje leksykograficzne pomagają zapewnić porządek i przejrzystość w decyzjach rekrutacyjnych, spełniając normy zgodności z przepisami dotyczącymi równości szans, przeciwdziałania dyskryminacji i innych regulacji prawnych.
Uwaga. Szacunki te nie obejmują branż, w których systemy rekrutacyjne są bardziej zróżnicowane, a decyzje są podejmowane na podstawie bardziej dynamicznych i elastycznych algorytmów AI.

bardziej dynamiczne i elastyczne podejście do oceny kandydata, uwzględniając zmieniające się potrzeby organizacji i rynku pracy. Szacunkowo w sektorze IT, technologii i edukacji systemy hybrydowe mogą stanowić od 15 do 30% procesów rekrutacyjnych. W administracji – mogą obejmować ok. 10–15% procesów rekrutacyjnych. Biorąc pod uwagę różne sektory, w których wprowadzenie hybrydowych systemów rekrutacyjnych z wykorzystaniem preferencji leksykograficznych jest możliwe, szacunkowy odsetek procesów rekrutacyjnych, opartych na takich systemach w Polsce może wynosić od 20 do 40%[58].

Pietruszkiewicz i inni proponują hybrydowe podejście do podejmowania decyzji w systemach, które łączą inteligencję ludzką i AI, poprawiając dokładność procesów decyzyjnych, jednocześnie zachowując przejrzystość i odpowiedzialność. W szczególności ich model może być przydatny w kontekście preferencji leksykograficznych, ponieważ uwzględnia ludzki nadzór nad rekomendacjami AI[59].

Sposoby rozwiązania problemu paradoksu preferencji leksykograficznych w systemach hybrydowych

1. Optymalizacja algorytmów decyzji – systemy hybrydowe mogą wykorzystywać zaawansowane algorytmy, które analizują różnorodne dane i uwzględniają niejednoznaczności w preferencjach, zamiast polegać na prostych regułach leksykograficznych. Algorytmy te mogą automatycznie eliminować cykle preferencji poprzez stosowanie rozwiązań takich jak optymalizacja wielokryterialna lub podejście probabilistyczne, które uwzględniają niepełne informacje.
2. Włączenie elementu losowości – gdy system napotyka na nierozróżnialnych kandydatów, wprowadzenie losowego elementu może pomóc w złamaniu cykli preferencji. Choć nie jest to idealne rozwiązanie, może stanowić sposób na utrzymanie sprawiedliwości procesu, gdy inne kryteria są niewystarczające do dokonania wyboru.

[58] Szacunek opiera się na analizie sektorów, w których stosowanie systemów rekrutacyjnych wykorzystujących preferencje leksykograficzne jest najbardziej prawdopodobne. Dane dotyczące udziału systemów regułowych w sektorach takich, jak finanse, administracja publiczna, zdrowie, edukacja czy IT, sugerują, że mogą one stanowić od 5 do 20% wszystkich procesów rekrutacyjnych. Dodatkowo, w sektorach intensywnie wykorzystujących AI i zaawansowane analizy danych, systemy hybrydowe mogą obejmować od 20 do 40% rekrutacji. Estymacja ta opiera się na analizach trendów w automatyzacji procesów HR oraz badaniach dotyczących wykorzystania algorytmów w rekrutacji w Europie i na świecie.
Uwaga. Szacunkowe dane nie obejmują wszystkich branż, w których wykorzystywane są zaawansowane techniki AI, a także mogą się różnić w zależności od regionu geograficznego i specyfiki konkretnej organizacji.

[59] W. Pietruszkiewicz, M. Twardochleb, M. Roszkowski, *Hybrid approach to supporting decision making processes in companies*, „Control and Cybernetics" 2011, nr 40(1), s. 126–133.

3. Zaangażowanie ekspertów ludzkich – w przypadkach, w których algorytmy napotykają trudności w podjęciu decyzji (np. w sytuacjach granicznych), eksperci mogą wprowadzić kontekst lub dodatkowe informacje, które umożliwią wyjaśnienie lub rozstrzygnięcie trudnych decyzji, eliminując nieprzechodniość w preferencjach.
4. Testowanie i monitorowanie systemu – warto przeprowadzać regularne testy systemów AI pod kątem ich tendencji do generowania cyklicznych preferencji. Analiza wyników oraz korekta algorytmów w czasie rzeczywistym może pomóc w eliminowaniu niepożądanych skutków tych cykli.
5. Hybrydowe podejście łączące inteligencję ludzką i AI – Pietruszkiewicz i inni proponują hybrydowe podejście do podejmowania decyzji w systemach, które łączą inteligencję ludzką i AI, poprawiając dokładność procesów decyzyjnych, jednocześnie zachowując przejrzystość i odpowiedzialność. W szczególności ich model może być przydatny w kontekście preferencji leksykograficznych, ponieważ uwzględnia ludzki nadzór nad rekomendacjami AI. Taki system pozwala na lepsze zarządzanie decyzjami, zwłaszcza w przypadkach, gdy tradycyjne algorytmy mogą napotkać trudności w eliminowaniu cyklicznych preferencji[60].

Monitorowanie decyzji AI powinno obejmować:

1. testowanie spójności algorytmów. Regularne analizowanie wyników podejmowanych przez systemy rekrutacyjne w kontekście tego, czy decyzje są przejrzyste i spójne;
2. ewaluację błędów pomiarowych. Zrozumienie, jak różnice w kryteriach (np. IQ czy doświadczeniu) są traktowane przez systemy, i czy te różnice nie prowadzą do niespójnych preferencji w wyniku błędów pomiarowych;
3. przegląd procesów decyzyjnych. Zapewnienie, że każda decyzja rekrutacyjna jest uzasadniona na podstawie rzeczywistych wymagań stanowiska, a nie nieistotnych lub dodatkowych kryteriów.

[60] Ibidem.

Tabela 3.
Porównanie modeli rekrutacyjnych

Kryterium	Systemy ekspertowe (klasyczne, oparte na wiedzy)	Preferencje leksykograficzne	Uczenie maszynowe (nowoczesne AI)	Modele hybrydowe
Sposób działania	Reguły logiczne opracowane przez ekspertów	Hierarchiczne podejmowanie decyzji: najważniejsze kryterium decyduje, pozostałe są drugorzędne	Analiza dużych zbiorów danych i optymalizacja na podstawie wzorców	Połączenie reguł ekspertowych z uczeniem maszynowym
Elastyczność	Niska – działa wg ustalonych zasad, trudne do dostosowania	Bardzo niska – raz ustalona hierarchia nie zmienia się w zależności od kontekstu	Wysoka – dostosowuje się do danych, może uwzględniać interakcje między cechami	Średnia do wysokiej – można dostosować reguły, ale elastyczność zależy od implementacji
Przejrzystość decyzji	Wysoka – jasne reguły i wyjaśnienia	Wysoka – łatwo zrozumieć, dlaczego kandydat został wybrany	Niska do średniej – modele „czarnej skrzynki" mogą być trudne do interpretacji	Średnia – można wyjaśnić część decyzji, ale elementy AI mogą być nieprzejrzyste
Ryzyko błędnych decyzji	Średnie – zależy od jakości reguł	Wysokie – brak elastyczności może prowadzić do nielogicznych wyników	Średnie do wysokie – może uczyć się błędnych wzorców	Średnie – system może ulegać zarówno błędom ekspertów, jak i modelu AI
Zastosowanie w rekrutacji	Reguły oceny kandydatów oparte na doświadczeniu HR	Hierarchiczne filtrowanie kandydatów, np. IQ > doświadczenie	Dynamiczna analiza cech kandydata oparta na danych historycznych	Łączy reguły ekspertowe z analizą AI, np. ustalanie progów dla modeli ML
Zalety	Transparentność decyzji. Możliwość manualnego dostosowania	Łatwa do interpretacji. Szybkość działania	Dostosowuje się do realiów rynku pracy. Może wykrywać nieszablonowych kandydatów	Balans między przejrzystością a adaptacyjnością. Możliwość ograniczenia ryzyka stronniczości modeli AI

Kryterium	Systemy ekspertowe (klasyczne, oparte na wiedzy)	Preferencje leksykograficzne	Uczenie maszynowe (nowoczesne AI)	Modele hybrydowe
Wady	Trudne do skalowania i aktualizacji. Wymaga ręcznego wprowadzania reguł	Może prowadzić do nielogicznych i sprzecznych wyborów. Nie uwzględnia błędów pomiaru	Brak pełnej przejrzystości. Ryzyko uczenia się niesprawiedliwych wzorców	Ryzyko nieprzechodniości preferencji (np. AI może nadpisywać reguły w sposób sprzeczny). Złożoność integracji i potrzeba stałej kontroli
Popularność w systemach rekrutacyjnych	Tak, ale coraz rzadziej	Tak, nadal stosowane w niektórych systemach	Tak, dominujące podejście	Tak, coraz częściej stosowane w celu ograniczenia ryzyka stronniczości modeli AI

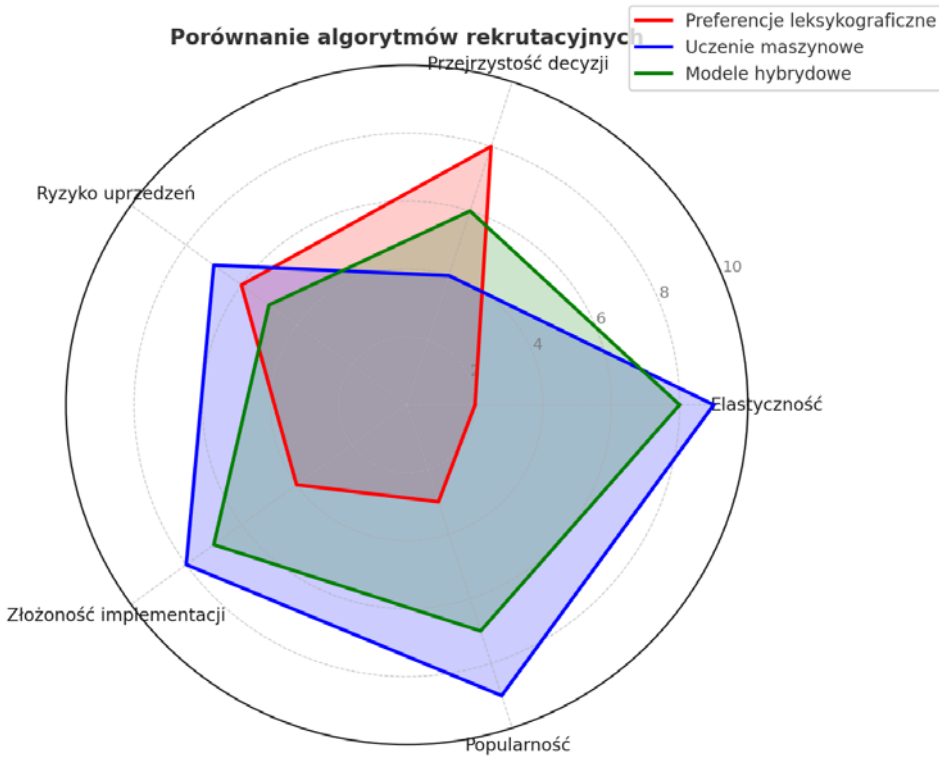
Źródło: opracowanie własne.

Algorytmy oparte na regułach preferencji leksykograficznych nie spełniają definicji uczenia się sztucznej inteligencji, ponieważ nie uczą się na podstawie danych, ani nie dostosowują swoich wyników. Jednak współczesne systemy łączą te podejścia – klasyczne reguły mogą stanowić element bardziej złożonych algorytmów decyzyjnych.

Modele hybrydowe, które mogą wykorzystywać klasyczne reguły oparte na preferencjach leksykograficznych, ale z dodatkową zdolnością do adaptacji i zmiany tych reguł są najbardziej obiecujące. Hybrydowe modele, które integrują deterministyczne reguły z adaptacyjną analizą AI, są szczególnie obiecujące.

Na wykresie 2 porównano algorytmy rekrutacyjne pod względem elastyczności, przejrzystości decyzji, ryzyka uprzedzeń, złożoności implementacji i popularności.

Wykres 2.
Porównanie algorytmów rekrutacyjnych



Źródło: opracowanie własne.

Porównanie algorytmów rekrutacyjnych:

1. preferencje leksykograficzne – bardzo przejrzyste, ale mało elastyczne i rzadziej stosowane we współczesnych systemach rekrutacyjnych;
2. uczenie maszynowe – elastyczne i popularne, ale trudniejsze do wyjaśnienia oraz bardziej wymagające pod względem wdrożenia;
3. modele hybrydowe – kompromis między przejrzystością a elastycznością, łączący zalety obu podejść.

Paradoks wyboru pokazuje, że nawet matematycznie racjonalne decyzje mogą prowadzić do niespójnych wyników. W kontekście rekrutacji opartej na AI oznacza to, że:

1. AI może powielać ludzkie błędy w ocenianiu kandydatów, jeśli zostanie zaprogramowane według sztywnych reguł;
2. błędy pomiarowe mogą powodować arbitralność decyzji, jeśli algorytm nie uwzględnia ich wpływu. Systemy AI są trenowane na danych historycznych,

które mogą odzwierciedlać wcześniejsze praktyki rekrutacyjne oraz uprzedzenia społeczne[61];

3. AI musi być projektowane w sposób umożliwiający adaptację i uwzględnienie kontekstu, zamiast ślepo stosować jedną regułę.

Aby uniknąć tych problemów, konieczne jest: wprowadzenie transparentności w decyzjach podejmowanych przez AI, zapewnienie audytu algorytmów oraz błędnych wyborów rekrutacyjnych.

Wyzwania dotyczące ochrony prywatności w rekrutacji AI

Hybrydowa prywatność – model, ograniczenia i konsekwencje prawne

Współczesna debata nad prywatnością pokazuje, że tradycyjne rozumienie tego pojęcia nie wystarcza. Dotychczas traktowano ją głównie jako jednorodne zjawisko, lecz dynamiczne zmiany technologiczne i społeczne ujawniają jej wielowymiarowy i ewoluujący charakter. Daniel Solove proponuje, aby zamiast jednej, uniwersalnej definicji, traktować prywatność jako sieć powiązanych problemów, zmieniających się w zależności od kontekstu społecznego i technologicznego[62]. Carissa Véliz podkreśla, że w cyfrowej erze prywatność to nie tylko indywidualne prawo, ale także wartość społeczna, ponieważ dane jednostek wpływają na całe społeczności[63].

Nowe technologie, takie jak sztuczna inteligencja, analityka predykcyjna czy personalizacja treści, wprowadzają nowe wyzwania w ochronie prywatności. W kontekście rekrutacji wykorzystanie AI może ograniczać autonomię kandydatów i prowadzić do nadmiernego profilowania, co rodzi pytania o etyczne granice przetwarzania danych[64]. Algorytmy rekrutacyjne analizują nie tylko CV, ale także zachowania kandydatów podczas rozmów wideo czy ich aktywność w mediach społecznościowych. Brak przejrzystości działania tych algorytmów oraz trudności w wyjaśnianiu decyzji (tzw. *black box problem*) rodzą obawy o prywatność i sprawiedliwość selekcji[65].

Podejście hybrydowe w analizie prywatności pozwala uwzględnić zarówno aspekty jednostkowe, jak i strukturalne. Z jednej strony kładzie nacisk na indywidualną kontrolę

---

61 M. Bogen, A. Rieke, *Help Wanted…*, op. cit.

62 D.J. Solove, *Understanding privacy*, Harvard University Press, Cambridge, MA, London 2008, s. 20–30.

63 C. Véliz, *Privacy is power: Why and how you should take back control of your data*, Bantam Press, London 2020, s. 51.

64 I. Ajunwa, *The Paradox of…*, op. cit.

65 M. Bogen, A. Rieke, *Help Wanted…*, op. cit.

nad danymi, co jest charakterystyczne dla klasycznych koncepcji prywatności opartych na zgodzie i transparentności. Z drugiej strony podkreśla wpływ systemów AI na społeczeństwo, pokazując, że dane jednostek są wykorzystywane do profilowania grup i przewidywania zachowań na masową skalę[66]. W kontekście rekrutacji hybrydowe podejście wymaga zastosowania zarówno technicznych zabezpieczeń (np. anonimizacji danych), jak i regulacji prawnych, które zapobiegną nadużyciom i zapewnią równość szans w procesie zatrudnienia.

Hybrydowa koncepcja prywatności obejmuje trzy główne wymiary: informacyjny, dostępności oraz decyzyjny. Na przecięciu tych wymiarów znajduje się hybrydowość prywatności, czyli złożona struktura ochrony danych, dostępu i ochrony decyzyjnej kandydatów w rekrutacji AI. Wymaga to nowych mechanizmów i uregulowań, które zapewniają równowagę między efektywnością algorytmów a poszanowaniem praw kandydatów do pracy.

Wymiary prywatności są kwestią teoretycznie i empirycznie otwartą. Alan Westin[67] wskazywał na kontrolę jednostki nad informacjami o sobie, podczas gdy Ruth Gavison[68] podkreślała trzy kluczowe elementy prywatności: anonimowość, niewiedzę i niedostępność. Z kolei Ferdinand Schoeman zwracał uwagę na etyczny wymiar prywatności i jej relacje z innymi wartościami społecznymi[69]. Choć stanowiska te różnią się, łączy je przekonanie, że prywatność jest kategorią dynamiczną, zależną od kontekstu społeczno-technologicznego.

Analiza literatury pozwala na wyróżnienie trzech głównych wymiarów prywatności: informacyjnego, dostępności oraz ekspresyjnego, które przedstawiono poniżej.

Prywatność informacyjna

Prywatność informacyjna oznacza kontrolę nad zakresem informacji, które jednostka chce zachować dla siebie[70]. W procesie rekrutacji dotyczy to kontroli nad danymi, które kandydaci udostępniają rekruterom oraz algorytmom AI. W tradycyjnych koncepcjach prywatności chodziło o ochronę danych osobowych i prawo do ich ujawniania, modyfikowania i usuwania[71], jednak w dobie sztucznej inteligencji pojawiają się nowe wyzwania, związane z automatycznym przetwarzaniem danych i ich dalszym wykorzystaniem.

---

66 S. Zuboff, *The age of surveillance capitalism: The fight for a human future at the new frontier of power*, PublicAffairs, New York 2019, s. 210–230.

67 A.F. Westin, *Privacy and freedom*, Atheneum, New York 1967, s. 7–10.

68 R. Gavison, *Privacy and the limits of law*, „Yale Law Journal" 1980, nr 89(3), s. 428–432.

69 F.D. Schoeman, *Privacy and social values: Theories and concepts*, „The Stanford Journal of Philosophy" 1984, nr 1(2), s. 30–35.

70 A.F. Westin, *Privacy and freedom…*, op. cit., s. 7–10.

71 Ibidem, s. 7–10.

Kandydaci często nie mają pełnej świadomości, jakie informacje są gromadzone, jakie są przetwarzane i do jakich celów wykorzystywane. Algorytmy rekrutacyjne analizują nie tylko CV i listy motywacyjne, ale również historię zatrudnienia, aktywność online oraz sposób interakcji w rozmowach wideo. Ponadto systemy oparte na uczeniu maszynowym mogą wykorzystywać te dane do profilowania kandydatów, co rodzi ryzyko niejawnej dyskryminacji i ograniczenia ich szans na rynku pracy[72]. Przykładem jest analiza semantyczna aplikacji, która może faworyzować określone grupy kandydatów, jeśli AI było trenowane na danych zawierających nieuświadomione uprzedzenia[73].

Prywatność dostępności

Prywatność dostępności odnosi się do możliwości kontrolowania dostępu do własnej osoby, zarówno w wymiarze fizycznym, jak i wirtualnym, ale także do prawa do zachowania odrębności czasowej. Klasyczne rozumienie tego wymiaru wiązało się z fizyczną izolacją jednostki oraz ochroną przed niechcianym nadzorem[74], natomiast w erze cyfrowej obejmuje również sferę wirtualną.

Algorytmy rekrutacyjne analizują również aktywność kandydatów w mediach społecznościowych, ich ślad cyfrowy oraz dane biometryczne pozyskane z nagrań wideo. Stosowane są technologie śledzenia mikroekspresji twarzy, tonu głosu czy ruchów gałek ocznych, które mogą oceniać „autentyczność" kandydata lub jego poziom zaangażowania[75]. Coraz częściej firmy rekrutacyjne używają oprogramowania monitorującego aktywność komputerową kandydatów podczas testów kompetencyjnych, co rodzi pytania o granice dopuszczalnej kontroli w procesie selekcji. Podczas testów kompetencyjnych AI może śledzić sposób pisania, czas reakcji, a nawet ruchy myszki, co może być postrzegane jako naruszenie prywatności. AI może zbierać informacje o kandydatach na podstawie ich historii przeglądania, aktywności na forach czy komentarzy na mediach społecznościowych, co może prowadzić do nieetycznych praktyk oceny kandydatów.

Prywatność decyzyjna (ekspresyjna)

Prywatność decyzyjna dotyczy autonomii jednostki w podejmowaniu decyzji dotyczących różnych sfer życia. W kontekście rekrutacji chodzi o autonomię w kształtowaniu kariery oraz swobodę wyrażania opinii i tożsamości bez obawy przed konsekwencjami. Ferdinand Schoeman podkreślał, że prywatność ekspresyjna ogranicza wpływ

---

72 M. Bogen, A. Rieke, *Help Wanted…*, op. cit., s. 7–15.

73 S. Barocas, A.D. Selbst, *Big data's disparate impact*, „California Law Review" 2016, nr 104(3), s. 671–732.

74 R. Gavison, *Privacy and the limits of law*, „Yale Law Journal" 1980, nr 89(3), s. 421–471.

75 I. Ajunwa, *The Paradox of…*, op. cit., s. 4–8.

zewnętrznej kontroli społecznej na wybory jednostki, chroniąc jej swobodę decydowania o własnym rozwoju zawodowym[76].

W kontekście algorytmicznej rekrutacji pojawiło się ryzyko „bańki algorytmicznej", czyli sytuacji, w której systemy AI ograniczają kandydatom dostęp do różnorodnych ścieżek kariery, sugerując określone oferty pracy na podstawie wcześniejszych decyzji, profilu zawodowego i analizy behawioralnej[77]. Może to prowadzić do efektu predestynacji zawodowej, gdzie jednostki są kierowane w stronę określonych sektorów czy stanowisk, a ich możliwość eksploracji alternatywnych ścieżek jest ograniczona. Ponadto analiza treści generowanych przez kandydatów w mediach społecznościowych może wpływać na ocenę ich „dopasowania kulturowego" do organizacji, co skutkuje autocenzurą i ograniczeniem pluralizmu w miejscu pracy[78].

Granice między poszczególnymi wymiarami prywatności mogą być nieostre, co sugeruje, że hybrydowe podejście do prywatności jest elastyczne i ewoluuje wraz z postępem technologicznym. W praktyce nie zawsze da się oddzielić różne wymiary prywatności, ponieważ są one ze sobą powiązane i mogą się przenikać. To rozmycie granic jest charakterystyczne dla hybrydowego podejścia do prywatności, które jest dynamiczne i zmienia się wraz z rozwojem technologii. Oznacza to, że definicje i granice poszczególnych wymiarów prywatności mogą ewoluować w odpowiedzi na zmieniające się warunki technologiczne, co sprawia, że nie są one stałe, ani łatwe do określenia w każdym przypadku.

Na diagramie Venna przedstawiono koncepcję hybrydowej prywatności, która ilustruje, jak różne wymiary prywatności mogą się przenikać i prowadzić do wielowymiarowych naruszeń w rekrutacji opartej na AI.

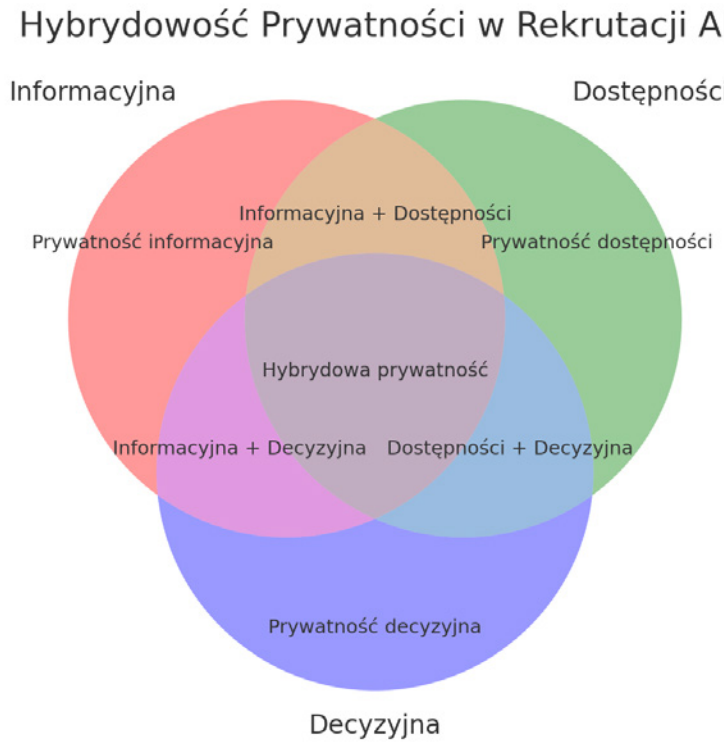
---

76 F.D. Schoeman, *Privacy and social values: Theories and concepts*, „The Stanford Journal of Philosophy" 1984, nr 1(2), s. 28–52.

77 R. Calo, *The boundaries of privacy harm*, „Indiana Law Journal" 2011, nr 86(3), s. 1131–1162.

78 S. Zuboff, *The age of surveillance capitalism*…, op. cit., s. 63–64.

Diagram 1.
Wizualizacja koncepcji hybrydowej prywatności
w procesach rekrutacyjnych opartych na AI



Źródło: opracowanie własne.

Poniżej znajdują się przykłady takich naruszeń, które obejmują różne aspekty prywatności.

Naruszenie wyłącznie prywatności informacyjnej

Przykładem naruszenia prywatności informacyjnej, bez ingerencji w prywatność dostępności i decyzyjną, jest sprzedaż lub udostępnienie danych osobowych kandydata, bez jego wiedzy i zgody. W tym przypadku dane są wykorzystywane w sposób niezgodny z intencjami osoby, której dotyczą, ale nie ogranicza to jej dostępu do tych informacji ani nie wpływa na podejmowanie decyzji rekrutacyjnych.

Naruszenie wyłącznie prywatności decyzyjnej

Z kolei naruszenie prywatności decyzyjnej, ale bez naruszenia prywatności informacyjnej i dostępu, może obejmować sytuację, w której algorytm używany w rekrutacji

analizuje dane o historii zawodowej kandydata, odrzucając aplikację osoby, która wcześniej pracowała w konkurencyjnej firmie. Kandydat nie ma możliwości zmiany tej decyzji, nie ma wpływu na dane, które zostały zebrane, ani na proces ich przetwarzania.

Nie została naruszona prywatność informacyjna, ponieważ kandydat udostępnił dane dotyczące swojej historii zawodowej, które są wykorzystywane zgodnie z zasadami informowania o przetwarzaniu danych. Nie została też naruszona prywatność dostępności. W tym przypadku problematyczne jest to, że decyzja zapadła automatycznie, na podstawie kryteriów, na które kandydat nie miał wpływu i nie miał pełnej świadomości, że takie parametry są stosowane i mogą go dyskwalifikować.

Naruszenie wyłącznie prywatności dostępności

Przykład naruszenia prywatności dostępności, bez naruszenia prywatności informacyjnej i decyzyjnej, może dotyczyć sytuacji, w której pracownik jest monitorowany przez firmę pod kątem swojej aktywności online, np. śledzenie czasu spędzanego na stronach internetowych, bez wyraźnej zgody. Jednak firma nie wykorzystuje tych danych do podejmowania decyzji o pracy, co oznacza, że nie ma naruszenia prywatności informacyjnej, ani też decyzyjnej. W tym przypadku chociaż nie zbiera się danych osobowych ani nie podejmuje decyzji na ich podstawie, fakt, że pracownik nie jest świadomy monitorowania swojej aktywności, stanowi naruszenie jego prawa do prywatności dostępności.

Naruszenie prywatności informacyjnej i dostępności

Jeśli system monitoruje aktywność kandydata w mediach społecznościowych bez jego zgody, ingerując w wirtualną przestrzeń prywatną, to mamy do czynienia z naruszeniem zarówno prywatności informacyjnej, jak i dostępu. Przykładem może być sytuacja, w której dane są zbierane bez zgody kandydata i wykorzystywane w procesie rekrutacji. Jednak w tym przypadku kandydat wciąż ma kontrolę nad decyzjami dotyczącymi swojej kariery (prywatność decyzyjna).

Naruszenie prywatności informacyjnej i decyzyjnej

Przykładem naruszenia jednocześnie prywatności informacyjnej i dostępności może być sytuacja, w której pracodawca monitoruje aktywność kandydatów w mediach społecznościowych i wykorzystuje te informacje w rekrutacji, nie informując ich o tym fakcie. Kandydat nie został poinformowany, że gromadzone i analizowane są dane z mediów społecznościowych w kontekście procesu rekrutacyjnego. Nie ma również możliwości kontrolowania, jakie informacje są wykorzystywane i w jaki sposób wpływają na ocenę jego kandydatury (naruszenie prywatności informacyjnej).

Pracodawca ingeruje w przestrzeń wirtualną kandydata, zbierając dane o jego aktywności bez jego wiedzy i zgody. Jest to naruszenie prywatności dostępności, ponieważ kandydat nie ma świadomości, że jego profil jest monitorowany, ani nie może ograniczyć dostępu do swoich treści w kontekście rekrutacji. Nie mamy tu do czynienia

z naruszeniem prywatności decyzyjnej, ponieważ kandydat nadal ma możliwość podejmowania decyzji dotyczących swojej kariery, a system nie automatyzuje całkowicie procesu selekcji kandydatów.

Naruszenie prywatności decyzyjnej i dostępności

Przykładem naruszenia jednocześnie prywatności decyzyjnej i dostępności może być sytuacja, w której firma rekrutacyjna stosuje algorytm automatycznie odrzucający kandydatów na podstawie określonych kryteriów, jednocześnie uniemożliwiając im dostęp do procesu oceny i decyzji.

Kandydat nie ma wpływu na decyzję rekrutacyjną, ponieważ algorytm automatycznie odrzuca jego aplikację na podstawie wcześniej zdefiniowanych reguł, np. miejsca zamieszkania lub historii zatrudnienia. Nie ma również możliwości zakwestionowania tej decyzji ani uzyskania wyjaśnienia, dlaczego jego aplikacja została odrzucona (naruszenie prywatności decyzyjnej).

Kandydat nie ma dostępu do pełnych informacji na temat procesu rekrutacji – nie wie, jakie kryteria są stosowane przez system oraz jakie czynniki miały wpływ na decyzję. Może się również zdarzyć, że system nie pozwala mu na ponowne aplikowanie lub ogranicza jego możliwość kontaktu z osobą rekrutującą, co dodatkowo pogłębia brak przejrzystości i kontroli nad własnym procesem rekrutacyjnym (naruszenie prywatności dostępu).

Nie ma tu naruszenia prywatności informacyjnej, ponieważ system może działać na podstawie danych, które kandydat sam udostępnił w swoim CV lub formularzu aplikacyjnym, bez dodatkowego gromadzenia informacji bez jego wiedzy.

Naruszenie wszystkich trzech rodzajów prywatności

Centralna część diagramu reprezentuje najbardziej problematyczne przypadki, gdy prywatność jest naruszana jednocześnie w wymiarze informacyjnym, dostępu i decyzyjnym. Przykładem może być firma rekrutacyjna wykorzystująca AI do analizy danych kandydatów na stanowisko specjalisty ds. marketingu cyfrowego. System ten zbiera dane z różnych źródeł, takich jak CV, aktywność w mediach społecznościowych, analiza treści blogów osobistych i obecność w sieci. Kandydaci nie są w pełni świadomi zakresu zbieranych informacji, a ich zgoda na przetwarzanie tych danych jest niejasna i trudna do wycofania. System AI ocenia ich w oparciu o preferencje leksykograficzne, najpierw analizując doświadczenie zawodowe w marketingu internetowym (poziom znajomości SEO, SEM, Google Analitics), a następnie w sytuacji, gdy kandydaci są nierozróżnialni, bierze pod uwagę inne kryteria (obecność w mediach społecznościowych, aktywności blogowe czy opinie na temat ich działalności zawodowej). Kandydaci nie mają pełnej kontroli nad swoimi danymi, nie wiedzą, jakie dokładnie informacje są wykorzystywane w procesie decyzyjnym, a system nie pozwala im na zakwestionowanie wyników.

W tym przypadku naruszenie prywatności dotyczy trzech wymiarów:

- prywatności informacyjnej – kandydat nie jest świadomy, jakie dane są zbierane ani jak są wykorzystywane;
- prywatności dostępności (czyli kontrola nad tym, kto i w jakim zakresie ma dostęp do danych osobowych) – kandydat nie ma kontroli nad swoimi danymi ani nie ma możliwości wycofania zgody na ich dalsze przetwarzanie;
- prywatności decyzyjnej – decyzja o zatrudnieniu podejmowana jest przez system AI, a kandydat nie ma możliwości jej zakwestionowania.

Hybrydowa prywatność to koncepcja dynamiczna, która ewoluuje w odpowiedzi na rozwój technologii oraz zmiany w regulacjach prawnych. W kontekście rekrutacji opartych na AI, wymaga elastycznego podejścia, które uwzględnia specyfikę zastosowania technologii oraz różne regulacje prawne. Koncepcja ta obejmuje podejście holistyczne, które chroni prywatność kandydatów w różnych wymiarach, jednocześnie uwzględniając wpływ technologii na rynek pracy i społeczeństwo.

Jednym z rozwiązań może być wielopoziomowy model ochrony prywatności, w którym poszczególne wymiary prywatności są chronione w różnym stopniu, zależnie od specyfiki procesu rekrutacyjnego. Przykładowo może to obejmować ściślejszą kontrolę nad sposobem zbierania i przetwarzania danych (wymiar informacyjny), transparentność algorytmów (wymiar decyzyjny) oraz mechanizmy umożliwiające kandydatom dostęp do informacji o kryteriach rekrutacyjnych oraz ograniczenie nadzoru nad aktywnością kandydatów (wymiar dostępności).

Dostosowanie ochrony prywatności do kontekstu zastosowania AI w rekrutacji powinno opierać się na elastycznych i adaptacyjnych regulacjach, które odpowiadają za zmieniające się zagrożenia związane z wykorzystaniem tych technologii.

Dynamiczna równowaga w regulacjach ochrony prywatności

Dynamiczna równowaga w regulacji ochrony prywatności jest kluczowa dla opracowania systemu, który równocześnie chroni prawa jednostki i umożliwia innowacje technologiczne[79]. Dążenie do kompromisu między sztywnymi a elastycznymi regulacjami pozwala na zachowanie zarówno ochrony prywatności, jak i elastyczności w odpowiedzi na zmiany technologiczne. Tradycyjne podejście, mimo że gwarantuje ścisłe normy ochrony danych, może okazać się niewystarczające w obliczu szybko rozwijających się technologii, w tym AI. Z drugiej strony elastyczne regulacje umożliwiają szybszą

[79] J. Wieczorkowski, *Big Data a prywatność. Naruszenie prywatności w świecie wirtualnym – wyniki badań*, „Roczniki Kolegium Analiz Ekonomicznych" 2017, nr 45, s. 33–43, https://rocznikikae.sgh.waw.pl/p/roczniki_kae_z45_02.pdf [dostęp:16.11.2025].

adaptację do nowych wyzwań, ale mogą prowadzić do niepewności w interpretacji, co z kolei stwarza ryzyko nadużyć.

Pojęcie „hybrydowej prywatności" podkreśla konieczność uwzględnienia wzajemnego przenikania różnych wymiarów ochrony prywatności, takich jak bezpieczeństwo danych, kontrola nad informacjami osobistymi oraz ich wykorzystywanie w kontekście nowych technologii. Regulacje dotyczące prywatności muszą być nie tylko elastyczne, ale również adaptacyjne, by mogły odpowiadać na zmieniające się wymagania technologiczne i społeczne. Ważne jest, aby przepisy oparte były na ogólnych zasadach, które umożliwiają dostosowywanie ich do nowych wyzwań, takich jak wykorzystanie AI w rekrutacji[80].

Jednym z obszarów, który nie jest jeszcze dostatecznie uregulowany, są neurodane – informacje pozyskiwane z aktywności mózgu, które mogą zawierać wrażliwe dane dotyczące procesów myślowych i emocji jednostki. Obecne regulacje, takie jak RODO czy AI Act, nie odnoszą się wprost do neurodanych, co może prowadzić do luk w ochronie prywatności w kontekście rozwoju neurotechnologii. Tu pojawia się nowe zagrożenie – możliwość odczytywania i interpretowania myśli oraz stanów psychicznych. Nieautoryzowane wykorzystanie neurodanych mogłoby prowadzić do wykluczania kandydatów na podstawie niejawnych cech psychicznych, co budzi poważne wątpliwości etyczne. Wprowadzenie odpowiednich mechanizmów prawnych jest niezbędne, aby zapobiegać potencjalnym nadużyciom, związanym z analizą i wykorzystywaniem tego typu danych[81]. Podobnie nieuregulowanym zagrożeniem jest kradzież tożsamości. Nowoczesne systemy AI mogą zbierać i analizować dane kandydatów na podstawie ich cyfrowych śladów, co stwarza ryzyko wykorzystania fałszywych lub skradzionych tożsamości do manipulowania procesem selekcji. Brak jednoznacznych regulacji dotyczących weryfikacji tożsamości w zautomatyzowanych procesach rekrutacyjnych, może prowadzić do sytuacji, w których firmy podejmują decyzje rekrutacyjne na podstawie sfałszowanych danych, a kandydaci tracą kontrolę nad sposobem, w jaki ich dane są wykorzystywane[82]. To kolejny przykład na to, dlaczego dynamiczna i adaptacyjna regulacja ochrony prywatności jest konieczna.

Aby zobrazować różnice między podejściem sztywnym a elastycznym w regulacjach ochrony prywatności, tabela 4 prezentuje kluczowe kryteria, które odzwierciedlają, jak oba podejścia różnią się w praktyce, zwłaszcza w kontekście ochrony prywatności i technologii zmieniających się w szybkim tempie.

---

80 Ibidem.

81 L. Słocka, *Aktualność unijnego system ochrony danych osobowych w świetle przetwarzania neurodanych*, „Przegąd prawa medycznego" 2021, nr 3(4), s. 79–97; A.S. Jwa, R.A. Poldrack, *Adressing Privacy Risk in neuroscience data: from data protection to harm prevention*, „Journal of Law & the Bioscience" 2022, nr 9(2); European Data Protection Supervisor (EDPS), TechDispatch – Neurodata, 3 czerwca 2024, s. 1–8.

82 R. Calo, *The boundaries of privacy harm…*, op. cit.

Tabela 4.
Sztywne vs. elastyczne podejście do regulacji prawnych

Kryterium	Sztywne regulacje	Elastyczne regulacje
Definicja prywatności	Jednoznaczna, zamknięta	Dynamiczna, dostosowana do kontekstu
Dostosowanie do technologii	Powolne, wymaga częstych nowelizacji prawa	Oparte na zasadach ogólnych, bardziej elastyczne
Interpretacja	Literalna, zależna od istniejących przepisów	Kontekstowa, uwzględniająca aktualne wyzwania
Skutki dla prywatności	Może być niewystarczające w szybko zmieniających się warunkach	Lepsza ochrona przed nowymi zagrożeniami
Zastosowanie w rekrutacji AI	Ograniczone do ściśle określonych sytuacji	Uwzględnia nowe technologie i sposoby wykorzystywania AI

Źródło: opracowanie własne.

Optymalne podejście łączy oba modele, uwzględniając:

- podstawowe zasady ochrony prywatności, tj. elastyczne fundamenty, np. prawo do wyjaśnienia decyzji AI, minimalizacja zbieranych danych, ograniczenia stosowania AI do niezbędnych celów, aby zapewnić ochronę prywatności już na etapie projektowania systemów (*privacy by design*);
- szczegółowe wytyczne dla konkretnych zastosowań, np. precyzyjne regulacje dotyczące AI w rekrutacji, w tym normy dla automatycznego przetwarzania danych osobowych, zasady transparentności procesów selekcyjnych oraz mechanizmy monitorowania i audytu algorytmów w celu wczesnego wykrywania i eliminowania ewentualnych uprzedzeń;
- dostosowanie regulacji do dynamicznego rozwoju technologii, tj. opracowanie elastycznych mechanizmów, które pozwalają na bieżąco reagować na zmiany w sposobach wykorzystywania AI w rekrutacji, uwzględniając potencjalne zagrożenia związane z niezamierzonymi skutkami ubocznymi;
- współpracę między ustawodawcami, firmami i organizacjami badawczymi w celu zapewnienia, że zarówno przepisy, jak i implementacja technologii będą rozwijały się równolegle, przy uwzględnieniu opinii wszystkich interesariuszy, a także przeprowadzanie regularnych audytów zewnętrznych w celu oceny wpływu technologii na prywatność i równość szans w procesie rekrutacyjnym.

Obecnie większość regulacji, takich jak Rozporządzenie o Ochronie Danych Osobowych (RODO w UE), koncentruje się na ochronie prywatności informacyjnej, czyli zabezpieczeniu danych osobowych kandydatów. RODO określa zasady zgody na

przetwarzanie danych, prawa dostępu do informacji oraz mechanizmy ochrony przed ich niewłaściwym wykorzystaniem (Rozporządzenie Parlamentu Europejskiego i Rady [UE] 2016/679). Jednak prywatność to pojęcie szersze. AI Act idzie dalej, regulując także przejrzystość i odpowiedzialność systemów AI. Nakłada na firmy obowiązek projektowania systemów w sposób minimalizujący ryzyko błędów i uprzedzeń oraz zapewniający przejrzystość dla kandydatów[83]. AI Act jest propozycją regulacji, a jego implementacja w różnych krajach członkowskich UE może wymagać dalszych dostosowań. Choć RODO zapewnia solidne ramy ochrony danych osobowych, AI Act stawia dodatkowe wymagania dotyczące przejrzystości procesów AI, co ma kluczowe znaczenie w kontekście rekrutacji.

AI Act koncentruje się głównie na ochronie prywatności informacyjnej, ale częściowo odnosi się także do prywatności decyzyjnej. W zakresie prywatności informacyjnej reguluje przetwarzanie danych osobowych w systemach AI, szczególnie w kontekście systemów wysokiego ryzyka. Odwołuje się do zasad RODO, wymagając minimalizacji danych, transparentności oraz zapewnienia mechanizmów kontrolnych nad danymi użytkowników. W zakresie prywatności dostępu, choć to nie jest głównym celem AI Act, to pośrednio wpływa na ten wymiar, wymagając np. systemów nadzoru nad działaniami AI oraz określając, kto może mieć dostęp do danych i w jaki sposób powinny być zabezpieczone[84]. W zakresie prywatności decyzyjnej AI Act wprowadza pewne mechanizmy ochrony autonomii, szczególnie w kontekście systemów wysokiego ryzyka, jak rekrutacja. Wymaga ludzkiego nadzoru nad decyzjami AI, zakazu stosowania niektórych form manipulacji i dyskryminacji systemów AI, które mogłyby wpływać na decyzje użytkowników, np. zakaz manipulacyjnych interfejsów „*dark patterns*"[85].

Granice ingerencji w prywatność kandydatów

Problemy regulacyjne wpływają na kształtowanie standardów etycznych w organizacjach. Gdy prawo nie nadąża za postępem technologicznym, normy etyczne stają się kluczowym mechanizmem ochrony jednostek. Firmy, szczególnie te wykorzystujące AI w rekrutacji, powinny wdrażać wewnętrzne regulacje obejmujące:

- przejrzystość algorytmów – zapewnienie kandydatom dostępu do informacji na temat sposobu działania systemów AI;
- sprawiedliwość w podejmowaniu decyzji – eliminowanie uprzedzeń i dyskryminacji w procesie selekcji;
- zapewnienie użytkownikom kontroli nad danymi – umożliwienie kandydatom decydowania o zakresie i sposobie wykorzystywania ich danych[86].

[83] Komisja Europejska, *Proposal for a regulation…*, op. cit.

[84] Ibidem.

[85] R. Calo, *The boundaries of privacy harm…*, op. cit., s. 1131–1162.

[86] L. Floridi, B. Mittelstadt, P. Allo, *The ethics of artificial intelligence: A primer*, *The Cambridge Handbook of Artificial Intelligence*, Cambridge University Press, Cambridge 2019, s. 33–50.

Brak zgodności standardów z regulacjami międzynarodowymi rodzi jednak ryzyko ich arbitralnego stosowania i nierównego traktowania kandydatów[87].

Poniżej zaprezentowano ocenę ingerencji w prywatność w procesach rekrutacyjnych opartych na AI. Sztuczna inteligencja w rekrutacji obejmuje różne metody selekcji kandydatów, takie jak analiza CV, testy kompetencyjne, wideorozmowy czy analiza aktywności online. Każda z nich różni się pod względem przejrzystości, sprawiedliwości i ryzyka dyskryminacji, co jest przedmiotem analiz badaczy, organizacji i ustawodawców.

Kontrola odnosi się do możliwości zarządzania własnymi danymi – im mniejsza kontrola, tym większe ryzyko naruszenia prywatności, zwłaszcza gdy kandydaci nie mają wpływu na sposób przetwarzania informacji. Ryzyko dyskryminacji wynika z błędów algorytmów, które mogą nieświadomie faworyzować określone grupy, np. w metodach takich jak analiza aktywności online, social media screening czy wideorozmowy. Przejrzystość oznacza poziom wiedzy kandydata o sposobie wykorzystywania jego danych – im bardziej niejawny proces, tym wyższe ryzyko naruszenia prywatności.

Stopień ingerencji w prywatność oceniono w skali od 0 do 10, gdzie 0 oznacza brak naruszenia prywatności, a 10 – największy możliwy stopień naruszenia prywatności. Skala ta obejmuje trzy wymiary: kontroli, dyskryminacji i przejrzystości.

Tabela 5.
Ocena stopnia naruszenia prywatności w procesach rekrutacyjnych w zależności od metody selekcji

Metoda	Kontrola nad danymi (0–10)	Dyskryminacja (0–10)	Przejrzystość (0–10)	Średnia ważona ryzyka naruszenia prywatności
Analiza CV	7	6	8	4,8
Testy kompetencyjne	8	5	7	4,2
Wideorozmowy analizowane przez AI	3	9	4	7,4
Analiza aktywności online	2	8	2	7,2
Profilowanie AI	4	7	3	6,2
Social Media Screening	4	8	3	6,4

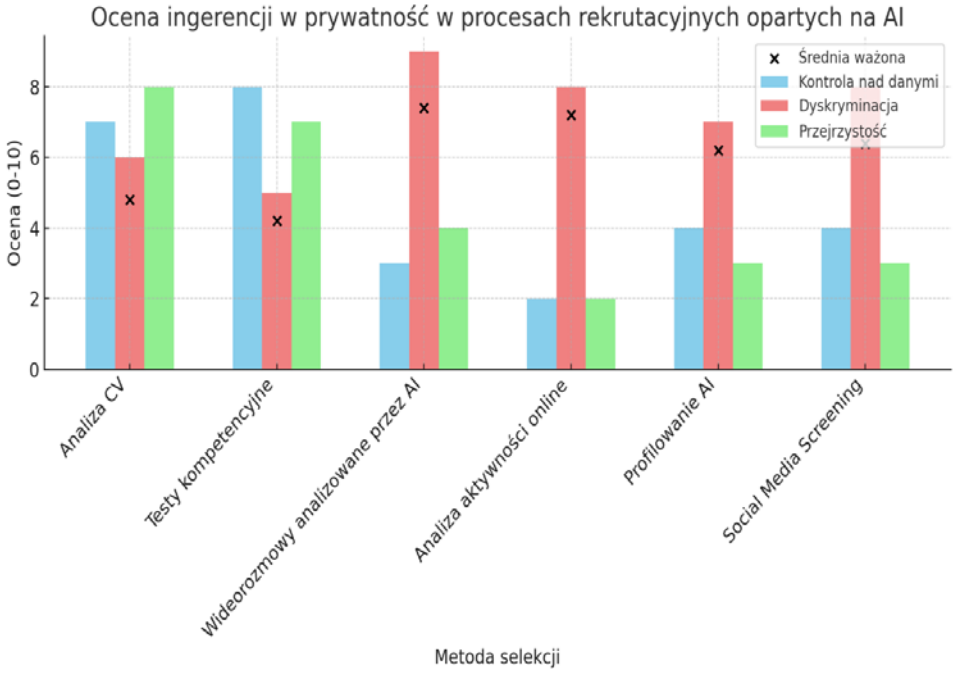
Źródło: opracowanie własne na podstawie wymogów prawnych i oczekiwań społecznych.

87 B. Mittelstadt, L. Floridi, *The ethics of big data: Current and foreseeable issues in biomedical context*, „Science and Engineering Ethics" 2016, nr 22(2), s. 303–341.

Do wymiarów kontroli i przejrzystości zastosowano odwrotną skalę (tzn. 10 minus wartość w danym wymiarze), podczas gdy dla dyskryminacji skala pozostała niezmieniona. Na tej podstawie obliczono łączną ocenę ryzyka naruszenia prywatności dla każdej metody, korzystając ze wzoru na średnią ważoną:
Średnia ważona = (Kontrola × 0,4) + (Dyskryminacja × 0,4) + (Przejrzystość × 0,2)

Waga kontroli (0,4) i dyskryminacji (0,4) została przyznana na podstawie ich kluczowego wpływu na prywatność, szczególnie w kontekście potencjalnych skutków społecznych. Waga przejrzystości (0,2) jest mniejsza, ponieważ choć pomaga w zrozumieniu procesu, ma mniejszy bezpośredni wpływ na prywatność. Wykres 3 przedstawia ocenę tych trzech aspektów w różnych metodach rekrutacyjnych opartych na AI.

Wykres 3.
Metody selekcji a prywatność kandydatów do pracy



Źródło: opracowanie własne na podstawie wymogów prawnych i oczekiwań społecznych.

Wykres 3 prezentuje poziom ingerencji metod selekcyjnych w prywatność kandydatów uwzględniając trzy wymiary: kontrolę danych, ryzyko dyskryminacji i przejrzystość. Zaznaczone punkty reprezentujące średnią ważoną dla każdej z metod (czarne kropki), pokazują ogólne ryzyko naruszenia prywatności, z uwzględnieniem wag. Metody takie jak analiza aktywności online i social media screening otrzymują

najniższe oceny w kategorii kontroli i przejrzystości, ale wysokie oceny w dyskryminacji, co skutkuje wyższym ryzykiem naruszenia prywatności. Z kolei analiza CV i testy kompetencyjne osiągają lepsze wyniki w przejrzystości i kontroli, przez co są mniej inwazyjne.

Poniżej przedstawiono interpretację metod selekcyjnych opartych na AI, pokazując, jak różne technologie wpływają na prywatność, przejrzystość procesów oraz sprawiedliwość decyzji:

1. Analiza CV – przejrzysta metoda, zwłaszcza gdy algorytmy klasyfikacji są odpowiednio udokumentowane. Kandydat ma kontrolę nad tym, co zawiera w CV, ale pewne dane, jak płeć, wiek czy miejsce zamieszkania mogą nieświadomie wpływać na decyzje (7). Istnieje ryzyko dyskryminacji, zwłaszcza, gdy AI bazuje na danych obciążonych uprzedzeniami (7). CV to jedna z najbardziej przejrzystych metod (8).
2. Testy kompetencyjne – charakteryzują się wysoką sprawiedliwością i przejrzystością (8), zwłaszcza jeśli są standaryzowane. Kandydat ma pełną kontrolę nad wynikami testów, ale mogą wystąpić różnice w szansach w zależności od poziomu edukacji (5). Kandydat zwykle wie, że jest testowany, ale może nie znać wpływu wyników na proces rekrutacji (7).
3. Wideorozmowy analizowane przez AI – niski poziom przejrzystości i wysokie ryzyko dyskryminacji, ponieważ systemy mogą oceniać kandydatów na podstawie cech niezwiązanych z kompetencjami, takich jak sposób mówienia czy mimika. Kandydat ma minimalną kontrolę nad tymi danymi (3), a analiza emocji, tonu głosu czy mimiki może prowadzić do błędnych interpretacji, zależnych od kultury, wieku czy płci (9). Kandydaci nie zawsze wiedzą, w jaki sposób AI analizuje ich rozmowy. Proces ten jest mało przejrzysty (4).
4. Analiza aktywności online – najtrudniejsza metoda, realizowana często bez zgody kandydata. Niski poziom przejrzystości, sprawiedliwości oraz wysoka dyskryminacja i naruszenie prywatności. Kandydat nie ma praktycznie kontroli nad tym, co jest zbierane z jego aktywności online (2), a oceny mogą faworyzować określone grupy (8). Proces jest mało przejrzysty (2).
5. Profilowanie AI – analiza wzorców zachowań i preferencji kandydata z różnych źródeł danych. Może być użyteczne w ocenie dopasowania do kultury organizacyjnej, ale wiąże się z ryzykiem błędnej interpretacji danych, niską przejrzystością i nieświadomą dyskryminacją. Kandydat ma ograniczoną kontrolę nad tym, jakie dane są wykorzystywane (4), a proces jest trudny do zrozumienia (3). Duże ryzyko dyskryminacji (7), ponieważ AI może oceniać na podstawie cech, które nie są bezpośrednio związane z rekrutacją.

6. Social media screening – analiza aktywności w mediach społecznościowych, która ma ujawniać zachowania, opinie i wartości kandydata, ale wiąże się z ryzykiem naruszenia prywatności oraz subiektywną oceną treści, co może prowadzić do niesprawiedliwego traktowania. Kandydat ma częściową kontrolę nad swoimi danymi (4), ale może nie zdawać sobie sprawy z konsekwencji udostępniania ich. Oceny mogą być obarczone ryzykiem dyskryminacji (8), a proces jest mało przejrzysty (3).

Wszystkie te metody różnią się pod względem poziomu przejrzystości, sprawiedliwości i ryzyka dyskryminacji oraz potencjalnego naruszenia prywatności, co podkreśla konieczność wprowadzenia odpowiednich regulacji i standardów etycznych w rekrutacji opartej na AI. Niezbędne są audyty algorytmów, aby zminimalizować ryzyko dyskryminacji i zapewnić etyczne wykorzystanie AI w procesach rekrutacji.

Wieczorkowski w swojej pracy *Big data a prywatność. Naruszenie prywatności w świecie wirtualnym – wyniki badań* zwraca uwagę na ogromną rolę technologii Big Data w naruszeniu prywatności użytkowników. Gromadzenie i analiza gigantycznych zbiorów danych użytkowników, szczególnie w kontekście rekrutacji i selekcji, stwarzają zagrożenia związane z brakiem pełnej kontroli nad danymi osobowymi. Wskazuje na ryzyko manipulacji wynikami analiz oraz potencjalne błędy w interpretacji danych osobowych, które mogą skutkować naruszeniem prywatności kandydatów i prowadzić do dyskryminacji. Technologie oparte na Big Data mogą także prowadzić do nieautoryzowanego gromadzenia informacji, które w tradycyjnych procesach rekrutacyjnych byłyby uznane za nieistotne[88].

Aby lepiej zrozumieć stopień ingerencji poszczególnych technologii w prywatność kandydatów, poniżej przedstawiono macierz cieplną. Ilustruje ona, które metody rekrutacyjne niosą ze sobą największe ryzyko w zakresie automatyzacji decyzji, zakresu zbieranych danych oraz możliwości korekty. Im intensywniejsze zaznaczenie na macierzy, tym większa ingerencja w prywatność kandydata.

[88] P. Wieczorkowski, *Big Data a prywatność…*, op. cit.

Tabela 5.
Macierz ryzyka naruszenia prywatności w zależności od technologii rekrutacyjnych

Technologie/Standardy etyczne	Zgodność z regulacjami unijnymi	Ryzyko naruszenia prywatności	Rekomendacje
Analiza zachowań online	Wysoka (jeśli przestrzegane są zasady przejrzystości i świadomej zgody użytkowników)	Wysokie (ryzyko inwigilacji)	Zwiększenie kontroli użytkownika nad danymi, przejrzystość algorytmów
Analiza tonu głosu	Średnia (potrzebne są szczegółowe regulacje w zakresie przetwarzania danych dźwiękowych)	Wysokie (niezdefiniowane granice prywatności)	Wprowadzenie wysokich standardów w zakresie przetwarzania danych audio
Ocena cech osobowości	Niska (zbyt ogólne przepisy, brak odpowiednich regulacji)	Wysokie (ryzyko manipulacji)	Konieczność szczegółowych regulacji dotyczących oceny cech osobowości
Przejrzystość algorytmów	Wysoka (zwykle zgodność z RODO, wymaga audytu algorytmów)	Niskie (przejrzystość pomaga w zarządzaniu prywatnością)	Zwiększenie wymagań dotyczących audytów algorytmów w firmach
Sprawiedliwość w podejmowaniu decyzji	Wysoka (powinna być zgodna z zasadami RODO, jeśli dotyczy danych osobowych)	Niskie (sprawiedliwość zmniejsza ryzyko nadużyć)	Implementacja procedur weryfikujących sprawiedliwość procesów decyzyjnych
Kontrola użytkowników nad danymi	Wysoka (zgodność z RODO w zakresie prawa do usunięcia danych)	Niskie (zapewnienie kontroli chroni prywatność)	Wzmocnienie mechanizmów zapewniających kontrolę nad danymi przez użytkowników

Źródło: opracowanie własne.

Opis poszczególnych kategorii:

1. Technologie/standardy etyczne przedstawiają konkretne technologie, które mogą ingerować w prywatność użytkowników (np. analiza tonu głosu, zachowań online, ocena osobowości) oraz standardy etyczne, które firmy mogą implementować, aby zapewnić zgodność z regulacjami. Zgodność z regulacjami unijnymi to ocena, jak dane technologie i standardy etyczne wpisują się w obowiązujące regulacje, takie jak RODO, które nakładają obowiązki na firmy związane z ochroną danych osobowych. Wysoka zgodność

oznacza, że dane rozwiązanie jest dobrze dostosowane do przepisów prawnych UE.

2. Ryzyko naruszenia prywatności określa, jakie ryzyko niesie dana technologia w kontekście prywatności użytkowników. Na przykład analiza cech osobowości może prowadzić do naruszenia prywatności, jeżeli firma zbiera dane bez pełnej zgody osoby lub nie zachowuje pełnej przejrzystości co do sposobu wykorzystywania tych danych.
3. Rekomendacje to propozycje działań, które firmy mogą podjąć w celu minimalizacji ryzyk związanych z danym rozwiązaniem, jak np. zwiększenie kontroli nad danymi przez użytkowników czy wprowadzenie szczegółowych audytów algorytmów.

Pracodawcy dążą do maksymalizacji efektywności rekrutacji i redukcji ryzyka błędnych decyzji, co często prowadzi do stosowania zaawansowanych narzędzi analitycznych, takich jak profilowanie AI czy analiza aktywności online. Z kolei kandydaci mają prawo do ochrony swoich danych i prywatności, szczególnie gdy technologie wykorzystywane do oceny wykraczają poza tradycyjne kryteria kompetencji zawodowych.

To napięcie prowadzi do fundamentalnego konfliktu między prawem do prywatności jednostki a prawem pracodawcy do informacji o kandydacie. AI w rekrutacji analizuje nie tylko kwalifikacje i doświadczenie, ale także aspekty osobowościowe czy zachowania w sieci, co rodzi pytania o granice dopuszczalnej ingerencji.

AI w rekrutacji analizuje nie tylko formalne kompetencje kandydatów, ale także ich aktywność online, sposób komunikacji oraz cechy osobowości[89]. Oznacza to, że pracodawcy mogą ingerować w życie prywatne kandydatów w sposób znacznie szerszy niż dotychczas. Choć ochrona interesów przedsiębiorstwa jest uzasadniona, nadmierna inwigilacja może prowadzić do naruszeń prawa do prywatności oraz negatywnych konsekwencji społecznych. Wachter i Mittelstadt podkreślają, że etyka AI nie może ograniczać się jedynie do zgodności z prawem – powinna także uwzględniać kontekst społeczny, kulturowy oraz potencjalne skutki technologii dla jednostek i społeczeństwa[90]. Autorzy wskazują, że współczesne systemy sztucznej inteligencji i analizy Big Data generują nieintuicyjne i nieweryfikowalne wnioski dotyczące zachowań, preferencji czy cech osobistych jednostek, które często pozostają poza zasięgiem świadomej kontroli osób, których dotyczą. Takie wnioski mogą być wykorzystywane do podejmowania

---

89 M. Raghavan, S. Barocas, J. Kleinberg, *Mitigating bias in algorithmic hiring. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 1–12*, 2020, https://doi.org/10.1145/3313831.3376313.

90 S. Wachter, B. Mittelstadt, *A right to reasonable inferences: Re-thinking data protection law in the Age of Big Data and AI*, „Columbia Business Law Review" 2019, nr 2, s. 494–620.

decyzji mających realne konsekwencje, np. w rekrutacji czy w marketingu, a jednocześnie mogą ingerować w prywatność, reputację i autonomię jednostek.

W związku z tym samo przestrzeganie przepisów prawa ochrony danych nie wystarcza, ponieważ obecne regulacje, w tym RODO, nie zapewniają pełnej ochrony przed skutkami wyciągania wniosków inferencyjnych. Wachter i Mittelstadt wskazują, że jednostki mają ograniczoną kontrolę nad tym, jak ich dane są wykorzystywane do formułowania wniosków, a prawa takie, jak dostęp do informacji, sprostowanie, sprzeciw czy przenoszenie danych w praktyce często nie obejmuje tego rodzaju inferencji. Oznacza to, że etyczne podejście do AI musi wykraczać poza wymogi prawne i obejmować także przejrzystość procesów decyzyjnych, możliwość weryfikacji wniosków oraz mechanizmy umożliwiające kwestionowanie decyzji opartych na algorytmach.

W konsekwencji autorzy postulują wprowadzenie nowego prawa ochrony danych – „prawa do rozsądnych wnioskowań", które wymagałoby od administratorów danych uzasadnienia *ex-ante* dla każdego wniosku wysokiego ryzyka, wskazując, dlaczego wnioski są normatywnie dopuszczalne, relewantne dla danego celu przetwarzania i oparte na wiarygodnych danych i metodach. Takie podejście pozwalałoby nie tylko na zwiększenie przejrzystości i odpowiedzialności algorytmicznej, ale także uwzględniałoby społeczne i etyczne konsekwencje technologii, czyniąc AI narzędziem bardziej sprawiedliwym i bezpiecznym dla jednostek i całych społeczności.

W związku z tym Wachter i Mittelstadt[91] wskazują, że etyczne podejście do AI powinno wykraczać poza wymogi prawne i uwzględniać zarówno wpływ technologii na jednostki, jak i mechanizmy kontroli nad sposobem wykorzystywania danych osobowych. To podejście staje się szczególnie istotne w kontekście rozwoju technologii rekrutacyjnych, gdzie sztuczna inteligencja i cyfrowe boty umożliwiają automatyzację procesów, personalizację doświadczeń kandydatów czy przewidywanie wyników zawodowych. Jednocześnie takie rozwiązania wprowadzają nowe wyzwania w zakresie ochrony prywatności, ponieważ mogą generować wnioski o charakterze inferencyjnym i nieweryfikowalnym, które wykraczają poza dotychczasowe ramy regulacyjne[92]. Wykorzystywanie zaawansowanych narzędzi, takich jak analiza mikroekspresji, tonacji głosu czy śladów cyfrowych kandydatów podkreśla, że technologie te niosą ze sobą zarówno ogromne możliwości, jak i istotne ryzyka związane z ingerencją w prywatność oraz potencjalną manipulacją danymi[93]. Jednocześnie takie technologie pozwalają na zbieranie danych w czasie rzeczywistym, co stwarza możliwość nie tylko analizy umiejętności i doświadczenia, ale także przewidywania stanu zdrowia psychicznego

91 Ibidem, s. 497–502.

92 R. Calo, *Robotics and the lessons of cyberlaw*, „California Law Review" 2015, nr 103(3), s. 513–563, https://doi.org/10.15779/Z38P69X.

93 Ibidem; C. O'Neil, *Weapons of math destruction*…, op. cit.

czy emocjonalnego kandydatów[94]. Tego rodzaju analizy mogą prowadzić do bardziej precyzyjnego zrozumienia kandydatów, ale także stwarzać zagrożenie, że rekruterzy będą podejmować decyzje oparte na danych, które wykraczają poza to, co jest dostępne w tradycyjnych formularzach aplikacyjnych.

Rozwój neurodanych i technologii rozpoznawania twarzy, jak również zaawansowane techniki analizy mikroekspresji i tonu głosu, wchodzą na coraz bardziej powszechny poziom w procesach rekrutacyjnych. Dzięki tym technologiom rekruterzy mogą uzyskać głębszy wgląd w emocje i intencje kandydatów, co daje im możliwość lepszego dopasowania osób do określonych ról w organizacji. Niemniej jednak, takie podejście rodzi poważne pytania o granice prywatności. Jakie dane mogą być wykorzystywane w procesach selekcji? Jakie informacje są wciąż uznawane za prywatne, a jakie granice powinny być postawione w zakresie analizy emocji, mikroekspresji i innych aspektów psychicznych kandydatów[95]?

W kontekście tych wyzwań dynamiczna równowaga staje się kluczowa. Prywatność nie jest wartością stałą, lecz zmieniającą się w czasie i przestrzeni, w zależności od kontekstu społecznego, technologicznego oraz regulacyjnego[96]. Technologie zmieniają sposób, w jaki postrzegamy granice prywatności, a procesy rekrutacyjne oparte na AI pokazują, jak łatwo te granice mogą zostać przesunięte w stronę bardziej inwazyjnych metod analizy i oceny. W związku z tym konieczne jest, aby regulacje prawne, takie jak RODO czy AI Act, były elastyczne i dynamiczne, umożliwiając dostosowanie do nowych wyzwań, które stają się nieuniknione w erze cyfrowej.

Najnowsze wyzwania związane z rozwojem sztucznej inteligencji, analityki danych, neurotechnologii oraz nowych technik rozpoznawania i przewidywania emocji w rekrutacji wymagają dynamicznego dostosowywania regulacji prawnych, które uwzględniają zmieniające się zagrożenia dla prywatności i indywidualnych wolności[97]. Zwiększona obecność technologii rozpoznawania twarzy, analiza mikroekspresji, a także przewidywanie zdrowia psychicznego w kontekście selekcji kandydatów, otwiera drogę do nowych form inwigilacji, które mogą wpłynąć na wolność jednostki i sprawiedliwość procesów rekrutacyjnych. Wykorzystywanie technologii AI w rekrutacji niesie ze sobą ogromny potencjał, ale wiąże się także z ryzykiem stworzenia systemów, które mogą nie tylko

---

94 V. Eubanks, *Automating inequality: How high-tech tools profile, police, and punish the poor*, St. Martin's Press, New York 2018.

95 S. Albrecht, M. Lauristin, V. Jourová, *The ethics of using AI in hiring processes*, 2023, [Unpublished manuscript/Working paper], https://btu.edu.ge/wp-content/uploads/2023/03/The-Ethics-of--Using-Artificial-Intelligence-in-Hiring-Processes.pdf [dostęp: 16.11.2025].

96 W.A. Parent, *Privacy, Morality, and the Law*, „Philosophy and Public Affairs" 1983, nr 12(4), s. 273.

97 S. Zuboff, *The age of surveillance capitalism…*, op. cit.

naruszać prywatność kandydatów, ale także wprowadzać niezamierzoną stronniczość czy dyskryminację[98].

Z tego względu przyszłość prywatności w kontekście rekrutacji i technologii AI, chmurowych oraz neurodanych wymaga elastycznego podejścia, które umożliwi ochronę interesów jednostki i sprawiedliwe traktowanie wszystkich kandydatów, niezależnie od tego, jak zaawansowane narzędzia będą wykorzystywane w procesach selekcji[99]. Zamiast tworzyć sztywne, ostre granice w regulacjach prawnych, powinny one opierać się na ogólnych zasadach, które będą mogły się elastycznie dostosowywać do dynamicznie zmieniającego się krajobrazu technologicznego w obszarze rekrutacji. Tufekci wskazuje, że w kontekście analizy danych z mediów społecznościowych, użytkownicy często angażują się w praktyki, takie jak *subtweeting* czy *screen capturing*, które mają na celu ochronę ich prywatności lub unikanie negatywnej oceny[100]. Praktyki te mogą prowadzić do błędnych analiz i decyzji rekrutacyjnych podejmowanych przez algorytmy, ponieważ ukrywają rzeczywiste intencje użytkowników lub zmieniają sposób ich postrzegania przez systemy, co utrudnia dokładną interpretację danych. Takie działania mają wpływ na wyniki analiz, ponieważ algorytmy mogą wyciągać mylne wnioski, które nie odzwierciedlają rzeczywistej sytuacji, co może skutkować niesprawiedliwym traktowaniem kandydatów w procesach selekcji.

Rekomendacje prawne i etyczne dotyczące wykorzystania AI w rekrutacji

Rekomendacje dla ustawodawców w zakresie ochrony prywatności kandydatów w procesach rekrutacyjnych wspieranych przez AI

Ogólna rekomendacja dotycząca wprowadzenia regulacji prawnych dotyczących AI w rekrutacji

Ustawodawcy powinni opracować jasne i precyzyjne przepisy regulujące stosowanie sztucznej inteligencji w rekrutacji, obejmujące kwestie przejrzystości, odpowiedzialności, ochrony prywatności oraz przeciwdziałania dyskryminacji. Regulacje te powinny uwzględniać trzy wymiary ochrony prywatności: informacyjny, dostępu oraz decyzyjny, aby zapewnić kompleksową ochronę osób ubiegających się o pracę.

98 S.U. Noble, *Algorithms of oppression: How search engines reinforce racism*, NYU Press, New York 2018.

99 Z. Tufekci, *Big data: Pitfalls and perils*. *Social Science Research Network*, 2014, https://doi.org/10.2139/ssrn.2410974.

100 Ibidem.

Uzasadnienie – zrównoważone regulacje prawne będą zapewniały, że technologie AI będą wykorzystywane w sposób odpowiedzialny, zgodny z zasadami etyki oraz prawami osób ubiegających się o pracę. Obecna ochrona danych osobowych (prywatność informacyjna) nie wystarcza, ponieważ AI w rekrutacji może wpływać na inne aspekty prywatności, takie jak kontrola nad dostępem do danych, transparentność decyzji algorytmicznych czy możliwość wyrażania zgody. Ochrona prywatności w szerokim ujęciu pozwala uniknąć niekontrolowanego wykorzystania danych, a także zmniejsza ryzyko dyskryminacji i nierówności w procesach rekrutacyjnych.

Szczegółowe rekomendacje dotyczące ochrony prywatności kandydatów w rekrutacji AI

Aby zapewnić pełną ochronę prywatności kandydatów w procesach rekrutacyjnych wspieranych przez sztuczną inteligencję, konieczne jest wprowadzenie regulacji dotyczących trzech głównych wymiarów prywatności: informacyjnego, dostępu oraz decyzyjnego. Te wymiary stanowią fundament dla ochrony danych osobowych kandydatów, jednocześnie umożliwiając odpowiednią kontrolę nad procesami rekrutacyjnymi.

1. W zakresie prywatności informacyjnej

Rekomendacja – wprowadzenie regulacji zapewniających kandydatom pełną kontrolę nad swoimi danymi osobowymi, m.in. nad tym, w jaki sposób są zbierane, przechowywane, analizowane i wykorzystywane w procesach rekrutacyjnych. Kandydaci powinni mieć możliwość wycofania zgody na przetwarzanie ich danych w dowolnym momencie.

Uzasadnienie – zgodnie z zasadami ochrony prywatności, kandydaci powinni mieć pełną kontrolę nad swoimi danymi osobowymi. Przepisy te są zgodne z regulacjami RODO, ale wymagają dostosowania do specyfiki procesów rekrutacyjnych, które z wykorzystaniem AI mogą zbierać dane w sposób bardziej inwazyjny, w tym analizować aktywność w Internecie, mediach społecznościowych czy wykorzystywać cechy biometryczne. Zapewnienie prawa do wycofania zgody i dostępu do swoich danych jest kluczowe dla utrzymania autonomii kandydatów.

2. W zakresie prywatności dostępności

Rekomendacja 1 – decyzje podejmowane przez algorytmy rekrutacyjne powinny być przejrzyste i zrozumiałe dla kandydatów. Powinna być zapewniona możliwość uzyskania informacji na temat danych wykorzystanych przez algorytmy oraz metodologii, na podstawie której zapadają decyzje.

Uzasadnienie – transparentność w procesach decyzyjnych opartych na AI jest niezbędna dla budowania zaufania do systemu rekrutacyjnego. Kandydaci powinni rozumieć, w jaki sposób ich dane wpływają na wynik rekrutacji, co pozwoli im podejmować świadome decyzje oraz zgłaszać ewentualne zastrzeżenia do procesu. Wprowadzenie

takiej regulacji zwiększy odpowiedzialność firm rekrutacyjnych i pomoże uniknąć ukrytych form dyskryminacji lub nieuczciwych praktyk.

Rekomendacja 2 – kandydaci powinni być informowani o wszelkich formach monitoringu, takich jak analiza wideo, biometria czy inne metody śledzenia. Powinna być im także zapewniona możliwość wyrażenia zgody na stosowanie takich technologii w procesach rekrutacyjnych.

Uzasadnienie – rekrutacja oparta na AI nie może naruszać przestrzeni prywatnej kandydatów. Technologie takie jak biometria, analiza mikroekspresji czy śledzenie aktywności fizycznej mogą stanowić poważne naruszenie prywatności, jeśli nie są stosowane w sposób zgodny z prawem i etyką. Wprowadzenie regulacji w tym zakresie pozwoli na ochronę kandydatów przed nieuzasadnionym nadzorem, który może wpływać na ich decyzje zawodowe oraz ograniczać ich wolność wyboru.

3. W zakresie prywatności decyzyjnej

Rekomendacja 1 – kandydaci muszą wyrazić świadomą zgodę na przetwarzanie ich danych, a także mieć pełną wiedzę, w jaki sposób te dane będą wykorzystywane. Zgoda powinna być dobrowolna, jednoznaczna i możliwa do wycofania w dowolnym momencie.

Uzasadnienie – bez świadomej zgody na przetwarzanie danych osobowych, procesy rekrutacyjne mogą być uznane za nieetyczne lub niezgodne z obowiązującymi przepisami. Kandydaci powinni rozumieć skutki udzielenia zgody, a także mieć pełną kontrolę nad tym, jak ich dane są używane, w tym w kontekście analizy predykcyjnej czy systemów rekomendacyjnych. Gwarancja prawa do wycofania zgody na każdym etapie procesu jest kluczowa, aby chronić autonomię kandydatów i ich prawa do prywatności.

Rekomendacja 2 – algorytmy wykorzystywane w procesie rekrutacyjnym powinny zapewniać kandydatom możliwość kwestionowania decyzji algorytmicznych, co pozwoli im na kontrolowanie wyników, które mają bezpośredni wpływ na ich życie zawodowe. Proces rekrutacyjny powinien oferować mechanizmy zapewniające równowagę i eliminujące nieuzasadnione decyzje.

Uzasadnienie – zautomatyzowane procesy rekrutacyjne mogą prowadzić do sytuacji, w których kandydaci tracą kontrolę nad swoimi ścieżkami kariery. Algorytmy oparte na wcześniejszych danych mogą zamknąć kandydatów w „bańce algorytmicznej", sugerując jedynie te ścieżki zawodowe, które były wcześniej wybierane lub przewidywane przez system. Wprowadzenie regulacji, które umożliwią kwestionowanie decyzji oraz zapewnią równowagę w traktowaniu kandydatów, pozwala na ochronę ich autonomii i zapobiega negatywnym konsekwencjom związanym z automatyzowaniem rekrutacji.

Zgodnie z omawianymi wymiarami prywatności, każdemu kandydatowi powinno przysługiwać prawo do pełnej kontroli nad swoimi danymi osobowymi, przejrzystości w procesie decyzyjnym oraz pełnej odpowiedzialności za wszelkie decyzje podejmowane

przez algorytmy. Wprowadzenie odpowiednich regulacji w zakresie prywatności informacyjnej, dostępu oraz decyzyjnej (ekspresyjnej) jest niezbędne, aby procesy rekrutacyjne były etyczne, sprawiedliwe i zgodne z prawami kandydatów.

Chociaż trzy wymiary prywatności (informacyjny, dostępu oraz ekspresyjny) obejmują kluczowe kwestie ochrony danych osobowych w kontekście AI w rekrutacji, nie wyczerpują one wszystkich istotnych aspektów związanych z odpowiedzialnością i nadzorem nad wykorzystywaniem tych technologii. Aspekty takie jak przejrzystość algorytmów czy odpowiedzialność za decyzje podejmowane przez AI są zagadnieniami, które należy traktować osobno, ponieważ wykraczają poza zakres ochrony prywatności, a są kluczowe dla zapewnienia sprawiedliwości, odpowiedzialności oraz równowagi w procesach rekrutacyjnych.

Dlatego konieczne jest wprowadzenie dodatkowych regulacji w tych obszarach:

1. Zwiększenie wymogów dotyczących audytów etycznych i przejrzystości algorytmów

Rekomendacja – wprowadzenie wymogu przeprowadzania regularnych audytów etycznych oraz publikacji wyników tych audytów przez firmy używające AI w rekrutacji. Algorytmy wykorzystywane w procesach rekrutacyjnych powinny być przejrzyste i dostępne do oceny przez odpowiednie organy, a także udostępniane w sposób zrozumiały dla kandydatów.

Uzasadnienie – regularne audyty etyczne pozwolą na identyfikację i eliminowanie potencjalnych nieprawidłowości, uprzedzeń oraz błędów w działaniu systemów AI. Takie audyty przyczynią się do zwiększenia równości i sprawiedliwości procesów rekrutacyjnych, zapobiegając dyskryminacji kandydatów. Wymóg przejrzystości algorytmów i publikacja wyników audytów pozwoli na zbudowanie zaufania do technologii AI w rekrutacji, a także umożliwi skuteczny nadzór nad ich działaniem.

2. Zachowanie odpowiedzialności za decyzje podejmowane przez AI

Rekomendacja – wprowadzenie przepisu, który określi, że firmy muszą ponosić pełną odpowiedzialność za decyzje podejmowane przez AI w procesach rekrutacyjnych. Ostateczna odpowiedzialność za wynik rekrutacji powinna leżeć po stronie pracodawcy, a nie algorytmu. Firmy powinny zapewnić, że decyzje podejmowane przez AI są zgodne z obowiązującymi przepisami prawa oraz zasadami etycznymi.

Uzasadnienie – tego rodzaju przepisy zapewnią, że odpowiedzialność za decyzje podejmowane w ramach procesów rekrutacyjnych pozostaje w rękach ludzi, a nie maszyn. Dzięki temu, jeśli dojdzie do kontrowersyjnych lub nieuczciwych decyzji, będzie można pociągnąć do odpowiedzialności konkretne osoby, a nie system AI. Uregulowanie tej kwestii pomoże w budowaniu odpowiedzialności za wykorzystywanie technologii AI i zapewni, że będą one stosowane w sposób sprawiedliwy i zgodny z prawami kandydatów.

Dwie ostatnie rekomendacje dotyczą kwestii, które nie są w pełni objęte wcześniej omawianymi wymiarami ochrony prywatności, ale mają kluczowe znaczenie dla odpowiedzialnego wykorzystania AI w rekrutacji. Zwiększenie wymogów dotyczących audytów etycznych oraz wprowadzenie odpowiedzialności za decyzje podejmowane przez algorytmy przyczyniają się do sprawiedliwości i przejrzystości procesów rekrutacyjnych, a także zwiększają zaufanie kandydatów do technologii AI.

Rekomendacje dla przedsiębiorstw w zakresie ochrony prywatności i rozwijania kompetencji etycznych

Wdrażanie sztucznej inteligencji w procesach rekrutacji wymaga odpowiedzialnego podejścia i rozwijania kompetencji etycznych. Poniższe rekomendacje pozwolą firmom zarządzać AI w sposób sprawiedliwy, transparentny oraz zgodny z normami etycznymi i prawnymi.

1. Budowanie zespołu ds. etyki AI

Rekomendacja – powołanie interdyscyplinarnego zespołu specjalistów, obejmującego ekspertów z zakresu technologii, prawa, HR i etyki. Opracowanie wewnętrznych polityk, regulujących wykorzystanie algorytmów.

Uzasadnienie – zapewnienie różnorodnych perspektyw pozwoli na identyfikację potencjalnych zagrożeń i lepsze zarządzanie ryzykiem związanym z AI.

2. Szkolenie z etyki AI dla pracowników

Rekomendacja – organizowanie cyklicznych szkoleń dla zespołów HR, IT i menedżerów, dotyczących algorytmicznej stronniczości, przejrzystości decyzji AI oraz ochrony prywatności.

Uzasadnienie – pracownicy działów HR powinni być przeszkoleni w zakresie wykorzystywania AI w rekrutacji oraz jego etycznych implikacji. Kandydaci powinni być informowani o tym, jak AI wpływa na proces rekrutacji oraz jakie dane są zbierane i wykorzystywane. Świadomość pracowników w zakresie etyki AI zmniejszy ryzyko nieświadomej dyskryminacji i poprawi jakość podejmowanych decyzji.

3. Monitorowanie i audytowanie algorytmów AI

Rekomendacja – regularne przeprowadzanie audytów pod kątem wskaźnika *Disparate Impact Ratio* (DIR) oraz testowanie algorytmów na różnych grupach demograficznych. Dodatkowo analiza wyników w kontekście fałszywie pozytywnych (*false positive*) i fałszywie negatywnych (*false negative*) w celu oceny skuteczności oraz potencjalnych błędów systemu.

Uzasadnienie – pozwoli na wykrycie potencjalnej dyskryminacji oraz niesprawiedliwych wzorców oceny kandydatów. Analiza fałszywie pozytywnych i fałszywie negatywnych wyników umożliwia lepsze zrozumienie czy system niesłusznie eliminuje lub promuje niektórych kandydatów, co wpływa na rzetelność i uczciwość procesu rekrutacyjnego.

4. Mechanizmy kontroli, „wyjaśnialności" i dokumentacji decyzji AI

Rekomendacja – wprowadzenie mechanizmów umożliwiających ręczną korektę decyzji AI przez rekruterów oraz mechanizmów odwoławczych dla kandydatów.

Wdrażanie procesów *Explainaible AI* (EAI), które pozwolą na wyjaśnianie decyzji podejmowanych przez system. Dodatkowo, opracowanie standardów dokumentowania decyzji AI oraz ich uzasadnień, tak, aby w razie potrzeby możliwa była ich ocena i korekta.

Uzasadnienie – AI może popełniać błędy lub nie uwzględniać kontekstu, dlatego potrzebna jest ludzka interwencja w uzasadnionych przypadkach. Mechanizmy XAI zwiększają przejrzystość decyzji algorytmicznych, co umożliwia lepszą kontrolę i przeciwdziała potencjalnym nadużyciom. Dokumentowanie decyzji pozwoli na przeprowadzenie audytu i poprawę procesów rekrutacyjnych.

5. Transparentność procesów rekrutacyjnych

Rekomendacja – informowanie kandydatów o wykorzystaniu AI w rekrutacji, udostępnianie kryteriów oceny oraz informacji zwrotnej na temat decyzji.

Uzasadnienie – pozwoli budować zaufanie kandydatów i zmniejszyć ryzyko nieporozumień oraz oskarżeń o dyskryminację.

6. Ochrona danych osobowych i prywatności kandydatów w ramach hybrydowej koncepcji prywatności

Rekomendacja – firmy powinny przyjąć bardziej rygorystyczne standardy ochrony prywatności, zamiast polegać jedynie na regulacjach prawnych. Proponuje się wdrożenie trójwymiarowej koncepcji ochrony prywatności, która obejmuje następujące wymiary:

- prywatność informacyjną – zapewnienie kontroli nad gromadzeniem, przetwarzaniem i udostępnianiem danych osobowych;
- prywatność dostępności – regulacja tego, kto i w jakim zakresie może uzyskiwać dostęp do danych kandydatów. Dotyczy to zarówno danych świadomie udostępnianych przez kandydata, jak i tych zbieranych bez jego pełnej wiedzy, np. poprzez śledzenie aktywności w Internecie czy analizę cech biometrycznych;
- prywatność decyzyjna – zapewnienie autonomii jednostki w podejmowaniu decyzji dotyczących jej danych osobowych oraz procesu rekrutacyjnego.

Hybrydowy charakter tej koncepcji oznacza, że naruszenie jednego z wymiarów prywatności (np. decyzyjnej) może prowadzić do konsekwencji w innym wymiarze (np. informacyjnym). Dlatego ochrona prywatności powinna być traktowana całościowo, uwzględniając wzajemne oddziaływanie tych 3 wymiarów. Takie podejście wykracza poza ochronę danych osobowych, jak przewiduje RODO i odnosi się również do szerszych aspektów interakcji człowieka z systemami AI, co częściowo reguluje AI Act, który odnosi się do kwestii transparentności, wyjaśnialności oraz nadzoru ludzkiego, które są istotne dla ochrony prywatności decyzyjnej i dostępu. Jednak kompleksowa ochrona wymaga podejścia wykraczającego poza te regulacje.

Uzasadnienie – szerokie podejście do prywatności jest kluczowe dla ochrony kandydatów do pracy. Wzmacnia ono ich poczucie kontroli nad procesem rekrutacyjnym, minimalizując ryzyko nadużyć i zapewniając większą przejrzystość decyzji. Choć AI Act reguluje niektóre aspekty ochrony prywatności, takie jak transparentność algorytmów, czy kontrola nad danymi, nie obejmuje on wielu kwestii związanych z prywatnością dostępności i decyzyjną. Zwiększenie przejrzystości stosowanych algorytmów.

Rekomendacja 2 – firmy powinny publikować szczegóły dotyczące wykorzystywanych algorytmów, zasad podejmowania decyzji i źródeł danych treningowych.

Uzasadnienie – przejrzystość pozwoli budować zaufanie wśród kandydatów i pracowników oraz minimalizować ryzyko nieetycznego wykorzystania AI.

7. Szkolenie pracowników HR z zakresu etyki AI

Rekomendacja – pracownicy HR powinni być przeszkoleni w zakresie wykorzystywania AI, a kandydaci powinni otrzymywać informacje na temat działania AI w rekrutacji.

Uzasadnienie – szkolenie zwiększy świadomość użytkowników AI i pozwoli na bardziej odpowiedzialne podejście do procesu rekrutacji.

8. Ocena procesów rekrutacyjnych

Rekomendacja – regularne analizy wpływu AI na różne grupy kandydatów w celu wykrycia potencjalnych przypadkowych fluktuacji, błędów systemowych i uprzedzeń.

Uzasadnienie – pozwoli na szybkie korygowanie ewentualnych błędów systemowych i zwiększy efektywność rekrutacji.

9. Zastosowanie hybrydowego modelu decyzji rekrutacyjnej

Rekomendacja – AI powinno wspierać, a nie zastępować decyzje ludzi. Ostateczna decyzja powinna należeć do człowieka.

Uzasadnienie – łączenie zalet AI (szybkość, obiektywność w analizie danych) z doświadczeniem i intuicją pracownika HR prowadzi do bardziej trafnych i sprawiedliwych decyzji rekrutacyjnych.

Rekomendacje stanowią kompleksowy przewodnik dla przedsiębiorstw, które chcą wdrażać AI w rekrutacji w sposób etyczny i odpowiedzialny.

Podsumowanie i wnioski

W artykule przeanalizowano wyzwania związane z wykorzystaniem sztucznej inteligencji w procesach rekrutacyjnych. Omówiono różne definicje AI, wskazując na ich niejednoznaczność i konsekwencje dla praktyki rekrutacyjnej. Szczególną uwagę poświęcono zagrożeniom związanym z automatyzacją decyzji oraz ryzykiem dyskryminacji wynikającej z algorytmicznych uprzedzeń. Podkreślono również istnienie paradoksów decyzyjnych, takich jak problem preferencji leksykograficznych, które mogą prowadzić do nieoczekiwanych i potencjalnie nieetycznych wyników rekrutacji.

Wyzwania związane z zastosowaniem sztucznej inteligencji w rekrutacji są wielowymiarowe i wymagają uwzględnienia nie tylko technologicznych, ale i społecznych, prawnych oraz etycznych aspektów, zwłaszcza w kontekście ochrony prywatności. Prywatność kandydatów do pracy staje się jednym z kluczowych zagadnień w tym procesie, a AI, przetwarzając dane osobowe kandydatów, może stwarzać ryzyko naruszenia tej prywatności. Ochrona prywatności kandydatów do pracy w erze nowych technologii wymaga nie tylko przestrzegania prawa, ale także poszanowania indywidualnych i społecznych wartości. W artykule omówiono znaczenie prywatności w erze rozwoju technologii oraz wpływ, jaki AI wywiera na procesy rekrutacyjne, w tym na sposób, w jaki zbierane i przetwarzane są dane osobowe kandydatów.

Prywatność jest pojęciem złożonym, które wymyka się jednoznacznemu definiowaniu. Może być rozumiana zarówno jako prawo, jak i jako wartość – zarówno indywidualna, jak i społeczna. W kontekście AI, prywatność nie tylko dotyczy ochrony danych, ale także zapewnienia autonomii jednostki, ochrony tożsamości oraz zaufania społecznego. Przesunięcie granic prywatności, zwłaszcza w obliczu rosnącej analityki danych, prowadzi do sytuacji, w których naruszenia prywatności pociągają za sobą konsekwencje w wielu wymiarach, takich jak utrata autonomii, bezpieczeństwa danych.

W przypadku hybrydowej ochrony prywatności wyzwania związane z jej rozumieniem są podobne do problemów z definicją sztucznej inteligencji – granice między różnymi wymiarami prywatności są rozmyte. Analiza wykazała, że prywatność obejmuje różne aspekty, takie jak dane osobowe, kontrola nad danymi, czy autonomia decyzji AI. W kontekście AI w rekrutacji, pytania dotyczące etyczności procesów, przejrzystości algorytmów i kontroli nad procesem decyzyjnym stają się kluczowe, szczególnie gdy naruszenia prywatności mogą prowadzić do niezamierzonych konsekwencji, takich jak dyskryminacja lub naruszenie autonomii kandydatów.

Podobnie jak w przypadku prywatności, również definicja sztucznej inteligencji nie jest jednoznaczna i stała. Przemiany w technologii, które AI wprowadza, zmieniają sposób, w jaki nauka postrzega to pojęcie. Prototypowa definicja AI, oparta na dynamicznych, zmieniających się właściwościach technologii, pokazuje, jak trudne staje się tworzenie jednoznacznych definicji w nauce. Współczesna definicja AI musi uwzględniać nie tylko jej techniczne aspekty, ale także implikacje społeczne, etyczne i kulturowe, które mają znaczący wpływ na nasze życie codzienne. W tym sensie AI to nie tylko narzędzie, ale także obszar, w którym tradycyjne definicje stają się nieadekwatne do wyzwań, przed którymi stoimy.

Hybrydowa prywatność to koncepcja, która ewoluuje w odpowiedzi na rozwój technologii oraz zmiany w regulacjach prawnych. W kontekście rekrutacji opartych na AI, wymaga elastycznego podejścia, które uwzględnia specyfikę zastosowania technologii oraz różne regulacje prawne. Prototyp – jako narzędzie testujące różne modele ochrony prywatności – może odegrać kluczową rolę w praktycznej ocenie efektywności strategii ochrony danych, pozwalając na adaptację systemów rekrutacyjnych do zmieniającego się kontekstu prawnego i technologicznego.

Koncepcja ta obejmuje podejście kompleksowe, które chroni prywatność kandydatów w różnych wymiarach, jednocześnie uwzględniając wpływ technologii na rynek pracy i społeczeństwo. Jednym z rozwiązań może być wielopoziomowy model ochrony prywatności, w którym poszczególne wymiary prywatności są chronione w różnym stopniu, zależnie od specyfiki procesu rekrutacyjnego. Np. może to obejmować ściślejszą kontrolę nad sposobem zbierania i przetwarzania danych (wymiar informacyjny), transparentność algorytmów (wymiar decyzyjny) oraz mechanizmy umożliwiające kandydatom dostęp do informacji o kryteriach rekrutacyjnych oraz ograniczenie nadzoru nad aktywnością kandydatów (wymiar dostępności). Testowanie takich rozwiązań przy użyciu prototypów pozwala na lepsze dostosowanie systemów AI do rzeczywistych wymagań ochrony prywatności.

Podkreślenie potrzeby ochrony prywatności hybrydowej ma kluczowe znaczenie w kontekście nowoczesnych technologii, ponieważ zapewnia elastyczność w definiowaniu tego, co stanowi naruszenie prywatności w różnych kontekstach. Stosowanie AI w rekrutacji wiąże się z koniecznością wyważenia między potrzebą wykorzystania danych a ochroną praw jednostki do prywatności. Przyszłe rozwiązania muszą uwzględniać zarówno aspekty techniczne, jak i społeczne, z zachowaniem pełnej transparentności i odpowiedzialności za decyzje podejmowane przez algorytmy.

Dostosowanie ochrony prywatności do kontekstu zastosowania AI w rekrutacji powinno opierać się na elastycznych i adaptacyjnych regulacjach, które odpowiadają za zmieniające się zagrożenia związane z wykorzystaniem tych technologii, a prototypy mogą stanowić istotny element tych regulacji, umożliwiając eksperymentowanie z różnymi scenariuszami przed ich pełną implementacją. W tym kontekście, zarówno

organizacje, jak i twórcy technologii, powinni zrozumieć, że kwestia prywatności nie jest już tylko zagadnieniem ochrony danych, ale fundamentalnym elementem tworzenia systemów, które mogą być akceptowane przez użytkowników. Wyzwania związane z zastosowaniem sztucznej inteligencji w rekrutacji muszą być rozwiązane w sposób, który zapewnia równowagę między efektywnością procesów rekrutacyjnych a ochroną prywatności kandydatów do pracy.

Wnioski

1. Definicyjna niejednoznaczność AI stanowi kluczowe wyzwanie. Różne podejścia do definiowania sztucznej inteligencji (m.in. perspektywa Floridiego, Russella i Norviga czy model kategorii naturalnych Roscha) wpływają na sposób jej rozumienia i implementacji w rekrutacji. Brak jednej, spójnej definicji utrudnia zarówno regulację prawną, jak i ocenę etyczną tych systemów. Zamiast precyzyjnego definiowania sztucznej inteligencji, które może okazać się niemożliwe lub niepraktyczne, bardziej sensowne jest podejście funkcjonalne. Zamiast pytać, czym jest AI, możemy zapytać, jakie funkcje realizuje. Takie podejście może pomóc w tworzeniu bardziej praktycznych regulacji.
2. Zastosowanie AI w rekrutacji wiąże się z ryzykiem dyskryminacji. Systemy oparte na uczeniu maszynowym mogą nieświadomie wzmacniać istniejące uprzedzenia, co stanowi istotne zagrożenie dla równości szans kandydatów. Wskazano na konieczność stosowania transparentnych metod oceny algorytmów oraz regularnych audytów etycznych.
3. Paradoksy decyzyjne w AI mają realne konsekwencje w procesach rekrutacyjnych, szczególnie w przypadku problemu preferencji leksykograficznych, gdzie AI wybiera kandydatów na podstawie jednej, priorytetowej cechy, pomijając inne istotne aspekty.
4. Sztuczna inteligencja oferuje znaczące możliwości usprawnienia procesów rekrutacyjnych, jednak jej stosowanie wymaga ostrożności. Zidentyfikowane wyzwania wskazują na konieczność interdyscyplinarnego podejścia, łączącego perspektywę technologiczną, etyczną i prawną, aby zapewnić sprawiedliwość i równość szans w rekrutacji.
5. Oprócz wyzwań związanych z hybrydową prywatnością, rozwiązaniem może być opracowanie i testowanie prototypów, które mogą pomóc w ocenie różnych strategii ochrony prywatności w kontekście rekrutacji opartych na AI. Taki prototyp pozwoli na dokładne zrozumienie konsekwencji prawnych i praktycznych rozwiązań.

Ograniczenia analiz

Prezentowana analiza koncentrowała się na wybranych aspektach zastosowania AI w rekrutacji, nie obejmując pełnego spektrum wpływu tej technologii. W związku z ciągłym rozwojem AI, niektóre wnioski mogą wymagać późniejszej aktualizacji, a także uwzględnienia perspektywy kandydatów do pracy, co mogłoby dostarczyć bardziej złożonego obrazu problemu. Ponadto, analiza nie uwzględniała pełnych wyzwań związanych z hybrydową ochroną prywatności, która jako koncepcja dynamiczna, ewoluuje w odpowiedzi na rozwój technologii oraz zmieniające się regulacje prawne. Granice tej ochrony, które mogą różnić się w zależności od kontekstu zastosowania AI, pozostają wciąż niejednoznaczne i wymagają dalszego badania. Prototypy testujące różne modele ochrony prywatności stanowią dodatkowy obszar, w którym mogą wystąpić trudności związane z oceną ich efektywności w realnych warunkach rekrutacyjnych. Warto zauważyć, że wprowadzenie takich rozwiązań może prowadzić do nowych wyzwań, takich jak konieczność dostosowania prototypów do zmieniającego się kontekstu prawnego i technologicznego, co może wpływać na ich zdolność do odpowiedniego reagowania na specyficzne zagrożenia związane z prywatnością kandydatów.

Kierunki przyszłych badań

Aby lepiej zrozumieć i zarządzać wpływem AI na rekrutację, konieczne są dalsze badania w następujących obszarach:

1. Opracowanie elastycznej i funkcjonalnej definicji AI – dostarczenie sensownej klasyfikacji AI w kontekście zastosowań praktycznych. Zamiast stawiać na precyzyjne definicje, bardziej sensowne wydaje się przyjęcie podejścia funkcjonalnego, które koncentruje się na tym, do jakich celów służy AI i jakie funkcje pełni w danym kontekście. Zamiast pytać: „Czym jest AI?", lepiej zadać pytanie: „Do czego AI może być używane?", a także: „W jaki sposób systemy AI realizują swoje zadania?". Pozwala to na bardziej elastyczne i praktyczne podejście do regulacji, które może różnić się w zależności od zastosowań i kontekstu technologicznego.
2. Rozwój transparentnych algorytmów – wdrożenie metod zapewniających przejrzystość decyzji podejmowanych przez AI w procesach rekrutacyjnych. Badania nad tworzeniem algorytmów, które umożliwiają wgląd w proces podejmowania decyzji, w tym mechanizmy objaśniania i audytowania działań AI, staną się kluczowe dla zapewnienia sprawiedliwości i odpowiedzialności.
3. Analiza wpływu AI na rynek pracy – badanie długoterminowych konsekwencji stosowania AI w rekrutacji, w tym wpływu na strukturę zatrudnienia

i dynamikę zawodową. Ważne będzie także zrozumienie, jak AI wpływa na różne grupy zawodowe i branże oraz jakie zmiany mogą wystąpić w wymaganiach dotyczących umiejętności.

4. Hybrydowa ochrona prywatności w kontekście AI – konieczne będą badania nad dynamicznymi modelami ochrony prywatności, które uwzględniają zmieniające się wymagania technologiczne i prawne. Ważne jest, aby opracować elastyczne i adaptacyjne regulacje, które zapewnią odpowiednią ochronę danych osobowych w kontekście AI, jednocześnie umożliwiając zastosowanie tej technologii w rekrutacji. Będą to również badania nad prototypami, które pozwolą testować i udoskonalać modele ochrony prywatności w warunkach rzeczywistych.
5. Testowanie i optymalizacja prototypów ochrony prywatności – prototypy mogą odgrywać kluczową rolę w testowaniu różnych strategii ochrony prywatności w systemach AI wykorzystywanych w rekrutacji. Badania nad projektowaniem, wdrażaniem i testowaniem takich rozwiązań pozwolą na lepsze zrozumienie ich skuteczności w kontekście rzeczywistych procesów rekrutacyjnych oraz w kontekście zmieniającego się prawa i technologii.
6. Etyczne i społeczne implikacje AI w rekrutacji – konieczne będzie zbadanie etycznych aspektów stosowania AI w rekrutacji, zwłaszcza w kontekście decyzji algorytmicznych, które mogą prowadzić do dyskryminacji, nierówności czy niezamierzonych konsekwencji. Będzie to obejmować także badania nad budowaniem zaufania społecznego do technologii oraz nad metodami, które mogą zapewnić większą równość i sprawiedliwość w procesach rekrutacyjnych.

AI w rekrutacji otwiera nowe możliwości, ale i rodzi poważne wyzwania etyczne, prawne i technologiczne. Kluczowym zadaniem na przyszłość będzie opracowanie rozwiązań zapewniających sprawiedliwość i transparentność procesów rekrutacyjnych, przy jednoczesnym wspieraniu innowacyjności i efektywności organizacyjnej. Badania w powyższych obszarach pozwolą nie tylko na lepsze zrozumienie i zarządzanie wpływem AI na rynek pracy, ale także na opracowanie bardziej elastycznych i dostosowanych do rzeczywistych potrzeb systemów ochrony prywatności oraz regulacji prawnych.

Bibliografia

Ajunwa I., *The paradox of automation as anti-bias intervention*, „Cardozo Law Review" 2020, nr 41(3), s. 1671–1741 https://larc.cardozo.yu.edu/clr/vol41/iss5/2 [dostęp: 16.11.2025].

Albrecht S., Lauristin M., Jourová V., *The ethics of using AI in hiring processes*, „Journal of Ethical AI Practices" 2023, [unpublished manuscript/working paper], https://btu.edu.ge/wp-content/uploads/2023/03/The-Ethics-of-Using-Artificial-Intelligence--in-Hiring-Processes.pdf. [dostęp: 16.11.2025]

Arrow K.J., *Social choice and individual values* (2nd ed.), Wiley & Sons, New York, London 1951.

Barocas S., Selbst A.D., *Big data's disparate impact*, „California Law Review" 2016, nr 104(3), s. 671–732.

Berdowska A., *Wspomaganie procesu rekrutacji pracowników za pomocą chatbotów – analiza wybranych rozwiązań*, „Projektowanie i analiza komunikacji w organizacji" 2018, nr 5(124), s. 93–112.

Binns R., *Algorithmic accountability and public reason*, „Philosophy & Technology" 2018, nr 31(4), s. 1–14.

Boden M.A., *AI: Its Nature and Future*, Oxford University Press, Oxford 2016, s. 1–4.

Bogen M., Rieke A., *Help wanted: An examination of hiring algorithms, equity, and bias*, Upturn 2018, https://www.upturn.org/ [dostęp: 16.11.2025].

Calo R., *Robotics and the lessons of cyberlaw*, „California Law Review" 2015, nr 103(3), s. 513–563, https://doi.org/10.15779/Z38P69X.

Calo R., *The boundaries of privacy harm*, „Indiana Law Journal" 2011, nr 86(3), s. 1131–1162.

Dastin J., *Amazon scraps AI recruiting tool that showed bias against women*, Reuters 2018, https://www.reuters.com/article/us-amazon-com-jobs-automation-insight--idUSKCN1MK08G [dostęp: 10.08.2025].

Dennett D., *Consciousness explained*, Little, Brown and Company New York, Boston, London 1991.

Duch W., *Informatyka neurokognitywna. Stan obecny, zastosowania, perspektywy*, Seminarium, Politechnika Wrocławska, Wrocław 10 lutego 2021, https://staff--ksi.pwr.edu.pl/seminariumITT/pdf/05-02-21.pdf [dostęp: 10.08.2025].

Eubanks V., *Automating inequality: How high-tech tools profile, police, and punish the poor*, [w:] *Law, Technology and Humans*, F. Gordon (ed.), St. Martin's Press, New York 2018.

European Commission, *Proposal for a regulation of the European Parliament and of the Council on artificial intelligence. COM (2021) 206 final*, 2021 https://ec.europa.eu/info/publications/210421-ai-act_en.

European Commission, *Proposal for a regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (AI Act)*, 2021, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206 [dostęp: 16.11.2025].

European Commission, *Artificial intelligence act. Regulation (EU) 2024/123*, 2024 https://europa.eu/legislation [dostęp: 16.11.2025].

European Data Protection Supervisor (EDPS), TechDispatch – Neurodata, 3 czerwca 2024, s. 1–8.

Fishburn P.C., *Utility theory for decision making*, Wiley, New York, 1970.

Floridi L., Mittelstadt B., Allo P., *The ethics of artificial intelligence: A primer*, *The Cambridge Handbook of Artificial Intelligence*, Cambridge University Press, Cambridge 2019.

Floridi L., *The Ethics of Artificial Intelligence*, Oxford University Press, Oxford 2023.

Frankish K., Ramsey W., *The Cambridge Handbook of Artificial Intelligence*, Cambridge University Press, Cambridge 2014.

Gavison R., *Privacy and the limits of law*, „Yale Law Journal" 1980, nr 89(3), s. 421–471.

Ghazanfar H., Hag A.U., *Ethical and legal implications of AI in Human Rerource Management*, "Journal of Social and Organizational Matters", 2025, 4(2), s. 417–428.

IBM, *Artificial intelligence in practice*, 2024, https://www.ibm.com/artificial-intelligence--in-practice [dostęp: 10.08.2025].

Jwa A.S., Poldrack R.A., Adressing Privacy Risk in neuroscience data: from data protection to harm prevention, „Journal of Law & the Bioscience" 2022, nr 9(2).

Komisja Europejska, *Proposal for a regulation laying down harmonised rules on artificial intelligence (AI Act)*, *COM (2021) 206 final*, Bruksela, 21 kwietnia 2021, https:// op.europa.eu/en/publication-detail/-/publication/e0649735-a372-11eb-9585--01aa75ed71a1 [dostęp: 10.08.2025].

Kuhn T.S., *Struktura rewolucji naukowych*, tłum. M.M. Ławniczak, Wydawnictwo Znak, Kraków 2019, s. 45–68.

Lakoff G., *Kobiety, ogień i rzeczy niebezpieczne: Co kategorie mówią nam o umyśle*, tłum. T. Brzezińska, Wydawnictwo Universitas, Kraków 2011.

McCarthy J., Minsky M., Rochester N., Shannon C.E., *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*, „AI Magazine" 2006, nr 27(4), s. 12–14, Dartmouth College, Hanover (NH), https://ojs.aaai.org/aimagazine/ index.php/aimagazine/article/view/1904 [dostęp: 16.11.2025].

Minsky M., *Steps Toward Artificial Intelligence*, Dept. of Mathematics & Research Lab. of Electronics, MIT, 1960, https://www.cs.bham.ac.uk/research/projects/cogaff/ misc/minsky/minsky-steps.pdf. [dostęp: 16.11.2025]

Minsky M., *The Society of Mind*, Simon & Schuster, New York 1986.

Mittelstadt B., Floridi L., *The ethics of big data: Current and foreseeable issues in biomedical context*, „Science and Engineering Ethics" 2016, nr 22(2), s. 303–341.

Nilsson N.J., *Principles of artificial intelligence*, „ZAMM – Journal of Applied Mathematics and Mechanics" 1983, nr 63(11), s. 476.

Nilsson N.J., *The Quest for Artificial Intelligence*, Cambridge University Press, Cambridge 2009.

Noble S.U., *Algorithms of oppression: How search engines reinforce racism*, NYU Press, New York 2018.

O'Neil C., *Broń matematycznej zagłady. Jaak big data zwiększa nierówności społeczne i zagraża demokracji*, tłum. M.Z. Zieliński, Wydawnictwo Naukowe PWN, Warszawa 2017.

Oman Z.U., Siddiqua A., Noorain R., *Artificial Intelligence and its ability to reduce recruitment bias*, „World Journal of Advanced Research and Reviews" 2024, nr 24(1), s. 551–564.

Oman A., Siddiqua S., *Ethical and legal implications of AI in hiring processes*, „Human Resource Management Review" 2024, nr 39(2), s. 1–18.

OpenAI, Artificial general intelligence (AGI), 2023, https://openai.com/research/agi [dostęp: 10.08.2025].

Parent W.A., *Privacy, Morality, and the Law*, „Philosophy and Public Affairs" 1983, nr 12(4), s. 273.

Parlament Europejski i Rada, Rozporządzenie (UE) 2024/1689 z dnia 13 czerwca 2024 r. w sprawie ustanowienia zharmonizowanych przepisów dotyczących sztucznej inteligencji (AI Act), Dziennik Urzędowy Unii Europejskiej I, nr 1689, 12 lipca 2024, https://www.si-dla-sprawiedliwosci.gov.pl/publikacja-ai-act [dostęp: 16.11.2025].

Pietruszkiewicz W., Twardochleb M., Roszkowski M., *Hybrid approach to supporting decision making processes in companies*, „Control and Cybernetics" 2011, nr 40(1), s. 126–133.

Popper K., *Logika odkrycia naukowego,* tłum. J.M. Kowalski, Wydawnictwo Naukowe PWN, Warszawa 2009.

Raghavan M., Barocas S., Kleinberg J., *Mitigating bias in algorithmic hiring. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 1–12*, 2020, https://doi.org/10.1145/3313831.3376313.

Raghavan M., et al., *Mitigating AI bias in hiring algorithms: A critical perspective*, „Journal of AI Ethics" 2020, nr 3(2), s. 34–50.

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation, GDPR). Official Journal of the European Union, L119/1.

Rosch E., *Principles of Categorization*, [w:] E. Rosch, B.B. Lloyd (red.), *Cognition and Categorization*, Lawrence Erlbaum Associates, miejsce wydania 1978, s. 27–48.

Russell S.J., Norvig P., *Artificial Intelligence: A Modern Approach* (4th ed.), Hoboken NJ: Pearson 2021.

Schoeman F.D., *Privacy and social values: Theories and concepts*, „The Stanford Journal of Philosophy" 1984, nr 1(2), s. 28–52.

Searle J., *Mind, language and society: Philosophy in the real world*, Basic Books, New York 2005.

Sen A., *Collective choice and social welfare*, Holden-Day, San Francisco 1970.

Simon H.A., Newell A., *Human Problem Solving*, Prentice-Hall, Englewood Cliffs, NJ 1972.

Słocka L., *Aktualność unijnego system ochrony danych osobowych w świetle przetwarzania neurodanych*, „Przegąd prawa medycznego" 2021, nr 3(4), s. 79–97.

Solove D.J., *Understanding privacy*, Harvard University Press, Cambridge 2008.

Stryker C., Kavlakoglu E., *What is Artificial Intelligence?*, „Think", https://www.ibm.com/think/topics/artificial-intelligence [dostęp: 10.08.2025].

Suleyman M., *DeepMind co-founder suggest new Turing test for AI chatbots*, „*Business Insider"*, 6 sierpnia 2023, https://www.businessinsider.com/deepmind-co-founder-suggests-new-turing-test-ai-chatbots-report-2023-6 [dostęp: 16.11.2025].

Suleyman M., komentarz w podkaście Have a Nice Future, WIRED, 16 sierpnia 2023, https://www.wired.com/story/have-a-nice-future-podcast-18/ [dostęp: 16.11.2025].

Suleyman M., *My new Turing Test*, Mustafa Suleyman Official Website, 2023, https://mustafa-suleyman.ai/my-new-turing-test [dostęp16.11.2025].

Świskak P. *Filozofia nauki i nauki społeczne po okresie pozytywizmu logicznego*, „Edukacja Filozoficzna" 1998, nr 26, s. 21–35.

Techsetter (n.d.), *Dyskryminacja w rekrutacji: Dlaczego AI powiela stereotypy?* [dostęp: 10.08.2025].

Tufekci Z., *Big data: Pitfalls and perils*. *Social Science Research Network*, 2014, https://doi.org/10.2139/ssrn.2410974.

Turing A., *Computing machinery and intelligence*, „Mind" 1950, nr 59(236), s. 433–460, https://doi.org/10.1093/mind/LIX.236.433.

Véliz C., *Privacy is power: Why and how you should take back control of your data*, Bantam Press, London 2020.

Wachter S., Mittelstadt B.D., *A right to explanation? An exploration of the impact of the General Data Protection Regulation on AI*, „International Data Privacy Law" 2019, nr 9(3), s. 178–190, https://doi.org/10.1093/idpl/ipz014.

Westin A.F., *Privacy and freedom*, Atheneum, New York 1967.

Wieczorkowski P., *Big Data a prywatność. Naruszenie prywatności w świecie wirtualnym – wyniki badań*, „Roczniki Kolegium Analiz Ekonomicznych" 2017, nr 45, s. 33–43.

Wilkins L., *Artificial intelligence in recruitment: Challenges and risks*, „Journal of HR Technology" 2021, nr 45(3), s. 12–29.

Wittgenstein L., *Dociekania filozoficzne*, tłum. J. Kijak, Wydawnictwo Naukowe PWN, Warszawa 1972.

Zuboff S., *The age of surveillance capitalism: The fight for a human future at the new frontier of power*, PublicAffairs, New York 2019.

Wyzwania związane z zastosowaniem sztucznej inteligencji w rekrutacji a ochrna prywatności

Streszczenie

Artykuł omawia ewolucję definicji sztucznej inteligencji (AI) oraz wyzwania związane z jej zastosowaniem w rekrutacji w kontekście ochrony prywatności kandydatów do pracy. Kluczowym problemem jest nieprecyzyjność definicji AI, która wpływa na skuteczność regulacji i zapewnienie standardów etycznych. Nieprecyzyjne definicje mogą prowadzić do nadmiernej lub niewystarczającej regulacji, obejmując technologie niezwiązane z AI lub pomijając zaawansowane systemy o wysokim ryzyku. Dodatkowo AI w rekrutacji stwarza nowe wyzwania związane z prywatnością, które wykraczają poza tradycyjną ochronę danych osobowych. Aby sprostać tym wyzwaniom, zaproponowano koncepcję hybrydowej prywatności, łączącą różne wymiary jej ochrony: informacyjny, dostępności fizycznej i wirtualnej oraz decyzyjny.

Ewolucja definicji AI, od testu Turinga (1950) po współczesne podejścia pokazuje, jak zmienia się rozumienie tej technologii i jakie ma to konsekwencje dla etyki oraz regulacji prawnych. Obecne definicje AI i regulacje wymagają dopracowania, aby nadążyć za dynamicznym rozwojem technologii. Nieprecyzyjne definicje prowadzą do problemów interpretacyjnych, które utrudniają skuteczne wdrażanie prawa.

Wnioski sugerują potrzebę elastycznych regulacji oraz wypracowania standardów etycznych, które chronią prywatność kandydatów, a jednocześnie uwzględniają interesy pracodawców. Kluczowe jest znalezienie dynamicznej równowagi między sztywnymi a elastycznymi regulacjami, które mogą dostosować się do zmieniającego się charakteru technologii AI i związanych z nią wyzwań etycznych, takich jak balansowanie między prywatnością kandydatów a interesem pracodawców.

Słowa kluczowe: sztuczna inteligencja, algorytmiczna rekrutacja, prototypowość definicji i regulacji AI, hybrydowa ochrona prywatności, prywatność informacyjna, dostępności, decyzyjna, etyka i prawo AI, decyzje AI, uprzedzenia algorytmiczne, nieracjonalne schematy wyboru

Challenges of Artificial Intelligence in Recruitment and Privacy Protection

Abstract

This article discusses the evolution of artificial intelligence (AI) definitions, and the challenges associated with its application in recruitment, particularly in the context of protecting candidates' privacy. A key issue is the imprecision of AI definitions, which affects the effectiveness of regulations and the establishment of ethical standards. Inaccurate definitions can lead to excessive or insufficient regulation, encompassing technologies unrelated to AI or overlooking advanced systems with high risks. Moreover, AI in recruitment raises new privacy concerns that go beyond traditional data protection. To address these challenges, the concept of hybrid privacy has been proposed, integrating various dimensions of privacy protection: informational, physical and virtual accessibility, and decision-making.

The evolution of AI definitions, from Turing's test (1950) to contemporary approaches, illustrates how the understanding of this technology has changed and what implications this has for ethics and legal regulations. Current AI definitions and regulations require further refinement to keep pace with the dynamic development of technology. Inaccurate definitions lead to interpretive issues that hinder effective implementation of the law.

The conclusions suggest the need for flexible regulations and the development of ethical standards that protect candidates' privacy while also taking into account the interests of employers. A dynamic balance is crucial between rigid and flexible regulations that can adapt to the evolving nature of AI technology and its associated ethical challenges, such as balancing candidates' privacy with employers' interests.

Keywords: artificial intelligence, algorithmic recruitment, prototypicality of AI definitions and regulations, hybrid privacy protection, informational privacy, accessibility, decision-making privacy, AI ethics and law, AI decisions, algorithmic bias, irrational selection schemes.

Cytuj jako:

Mazur A., *Wyzwania związane z zastosowaniem sztucznej inteligencji w rekrutacji a ochrona prywatności*, „Myśl Ekonomiczna i Polityczna" 2025, nr 1(84), s. 93–159 DOI: 10.26399/meip.1(84).2025.04/a.mazur

Cite as:

Mazur A. (2025). 'Challenges of Artificial Intelligence in Recruitment and Privacy Protection'. *Myśl Ekonomiczna i Polityczna* 1(84), 93–159 DOI: 10.26399/meip.1(84).2025.04/a.mazur