

## Wstęp

Chociaż rozwiązania wykorzystujące algorytmy w ramach zautomatyzowanego podejmowania decyzji obecne są w sferze zarówno usług publicznych, jak i tych świadczonych przez sektor prywatny od dekad<sup>2</sup>, ich regulacja nadal powoduje wiele wątpliwości. Rozwiązania legislacyjne dotyczące zautomatyzowanego podejmowania decyzji są relatywnie nowe w prawie Unii Europejskiej (UE) i Rady Europy, a ich rzeczywisty wpływ na stosowanie takich metod podejmowania decyzji pozostaje dyskusyjny. Tymczasem używanie zautomatyzowanego podejmowania decyzji staje się coraz bardziej powszechne zarówno w sektorze prywatnym, jak i publicznym. W sektorze prywatnym przykładem zastosowania zautomatyzowanego podejmowania decyzji jest ocena zdolności kredytowej<sup>3</sup> dokonywana z wykorzystaniem zautomatyzowanej analizy danych czy profilowany przekaz marketingowy, którego treść jest personalizowana w zależności od tego, do jakiej grupy została przyporządkowana jednostka, do której dana reklama ma być skierowana<sup>4</sup>. Z kolei obszary wykorzystywania systemów zautomatyzowanego

---

<sup>2</sup> Na przykład już w latach 70. XX w. w Norwegii system informatyczny był używany do decydowania, czy danej jednostce przysługuje prawo do wsparcia w postaci dodatku mieszkaniowego. System wykorzystywał wprowadzony przez daną osobę numer identyfikacyjny do połączenia z innymi bazami danych i na podstawie dostępnych w nich informacji wydawał decyzje, czy należy przyznać danej osobie zasiłek, zob. J. Bing, *Code, Access and Control*, [w:] *Human Rights in the Digital Age*, red. A. Murray, M. Klang, Routledge-Cavendish, Londyn–Sydney–Portland 2005, s. 204.

<sup>3</sup> Omówienie problemu, zob. K. Johnson, F. Pasquale, J. Chapman, *Artificial Intelligence, Machine Learning, and Bias in Finance: Toward Responsible Innovation*, „Fordham Law Review” 2019, t. 88, nr 2, s. 499–530.

<sup>4</sup> Badanie empiryczne na ten temat, zob. A. Datta, M.C. Tschantz, A. Datta, *Automated Experiments on Ad Privacy Settings: A Tale of Opacity, Choice, and Discrimination*, „Proceedings on Privacy Enhancing Technologies” 2015, nr 1, s. 92–112.

podejmowania decyzji w sektorze publicznym obejmują np. działania mające na celu prewencję przestępczości, czyli sytuacje, w których analiza danych stanowi podstawę prowadzenia działań przez policję w określonych dzielnicach lub wobec danych jednostek<sup>5</sup>, wydawanie decyzji w zakresie przysługującej jednostkom pomocy społecznej<sup>6</sup>, zarządzanie mobilnością nauczycieli, czyli przyporządkowywanie przez algorytm przyszłego miejsca pracy do danej osoby<sup>7</sup>, a także przydzielanie spraw sędziom<sup>8</sup> czy zapewnianie dzieciom miejsc w żłobkach<sup>9</sup> przy użyciu decyzji wydawanych za pośrednictwem analizy dokonanej przez system informatyczny.

Rozdźwięk między rosnącym znaczeniem społecznym zautomatyzowanego podejmowania decyzji i jego coraz powszechniejszym stosowaniem a brakiem jasności w zakresie dotyczących go rozwiązań regulacyjnych stanowił główne źródło motywacji do przeprowadzania badań w tym zakresie. W książce stawiam hipotezę, że możliwe jest postrzeganie algorytmów używanych w ramach systemów zautomatyzowanego podejmowania decyzji jako informacji podlegającej prawu dostępu do informacji lub do dokumentów urzędowych w świetle prawa europejskiego<sup>10</sup>. Dlatego przedmiotem zainteresowania mojej analizy są algorytmy stosowane w systemach zautomatyzowanego podejmowania decyzji wykorzystywanego w sektorze publicznym.

Przez pojęcie systemów zautomatyzowanego podejmowania decyzji rozumie się programy komputerowe używane do wspierania lub zastępowania procesów podejmowania decyzji wykonywanych przez człowieka rozwiązaniami zautomatyzowanymi, niezależnie od tego, jak duży jest udział rozwiązań cyfrowych w tym procesie<sup>11</sup>. Samo zautomatyzowane podejmowanie

<sup>5</sup> AlgorithmWatch, Bertelsmann Stiftung, Open Society Foundations, *Automating Society. Taking Stock of Automated Decision-Making in the EU*, 2019, s. 44, 52, 90–91, 100. [https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/001-148\\_AW\\_EU-ADMreport\\_2801\\_2.pdf](https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/001-148_AW_EU-ADMreport_2801_2.pdf), dostęp: 1.04.2021.

<sup>6</sup> *Ibidem*, s. 101, 130–131.

<sup>7</sup> *Ibidem*, s. 88.

<sup>8</sup> Zob. Fundacja ePaństwo, *alGOVrithms. STATE of PLAY. Report on Algorithms Usage in Government-Citizens Relations in Czechia, Georgia, Hungary, Poland, Serbia and Slovakia*, 2019, s. 15–32. <https://epf.org.pl/en/wp-content/uploads/sites/3/2019/05/alGOVrithms-State-of-Play-Report.pdf>, dostęp: 1.04.2021.

<sup>9</sup> Zob. Centrum Cyfrowe, Klub Jagielloński, *AlgoPolska. Zautomatyzowane podejmowanie decyzji w służbie społeczeństwu*, 2019, s. 25. <https://centrumcyfrowe.pl/wp-content/uploads/sites/16/2019/12/AlgoPolska-raport.pdf>, dostęp: 1.04.2021.

<sup>10</sup> Na temat różnicy między pojęciami informacji i dokumentów urzędowych – zob. rozdział trzeci i czwarty.

<sup>11</sup> Definicja jest zbliżona do tych proponowanych w: AlgorithmWatch, Bertelsmann Stiftung and Open Society Foundations, *op. cit.*, s. 9 i Algo:aware, *State-of-the-Art Report. Algorithmic decision-making*, 2018, s. II. <https://platformobservatory.eu/app/uploads/2019/06/AlgoAware-State-of-the-Art-Report.pdf>, dostęp: 1.04.2021.

decyzji to proces, w którym jest stosowany system zautomatyzowanego podejmowania decyzji rozumiany jako rozwiązanie technologiczne, u którego podstaw stoi algorytm. Bez algorytmu nie jest możliwe wprowadzenie zautomatyzowanego podejmowania decyzji. Dlatego za każdym razem, kiedy w książce posługuję się pojęciem zautomatyzowanego podejmowania decyzji, w jego ramach mieści się stosowany w nim algorytm. Przyjmuję przy tym perspektywę neutralną technologicznie, czyli niezależną od tego, jaka konkretna metoda czy technologia jest stosowana w ramach danego systemu zautomatyzowanego podejmowania decyzji.

Aby odnieść tę definicję do konkretnych przykładów, można się posłużyć algorytmem, który był używany do profilowania osób bezrobotnych w Polsce. Na podstawie sumy punktów, przyznanych za odpowiedzi na pytania sformułowane w kwestionariuszu, osoba bezrobotna była kwalifikowana do jednego z trzech profili. Teoretycznie urzędnik miał możliwość dokonania zmiany w stosunku do profilu zaproponowanego przez system, jednak sytuacja taka występowała w mniej niż 1% przypadków profilowania<sup>12</sup>. Chociaż rozwiązanie technologiczne nie zastępowało człowieka całkowicie, w praktyce jego wpływ na podejmowane decyzje był ogromny. Ponadto przykład ten pokazuje, że stosowane rozwiązania mogą nie być zbyt wysublimowane technologicznie – algorytm nie „uczył się” na analizowanych danych, nie doskonalił metod dopasowywania pomocy do konkretnych jednostek. Odmianą sytuację ilustruje sprawa programu SyRI, który był używany w Holandii w celu identyfikacji oszustw związanych z korzystaniem z pomocy socjalnej. System mógł wykorzystywać technologie uczenia maszynowego i na podstawie analizy danych pochodzących z wielu różnych baz przewidywać ryzyko zachowań niepożądanych z punktu widzenia państwa<sup>13</sup>. W świetle przyjętej przeze mnie definicji oba te przykłady są objęte zakresem znaczenia terminu „system zautomatyzowanego podejmowania decyzji”.

Taki sposób zdefiniowania zautomatyzowanego podejmowania decyzji pozwala ponadto zauważyć, że na implementację rozwiązań służących

---

<sup>12</sup> Zob. Fundacja Panoptykon (J. Niklas, K. Sztandar-Sztanderska, K. Szymielewicz), *Profiling the Unemployed in Poland: Social and Political Implications of Algorithmic Decision Making*, Warszawa 2015, s. 17. [https://panoptykon.org/sites/default/files/leadimage-biblioteka/panoptykon\\_profiling\\_report\\_final.pdf](https://panoptykon.org/sites/default/files/leadimage-biblioteka/panoptykon_profiling_report_final.pdf), dostęp: 1.04.2021; M. Kuziemski, G. Misuraca, *AI Governance in the Public Sector: Three Tales from the Frontiers of Automated Decision-Making in Democratic Settings*, „Telecommunications Policy” 2020, t. 44, nr 6, artykuł nr 101976.

<sup>13</sup> System omówiony np. w: Agencja Praw Podstawowych Unii Europejskiej, *Artificial Intelligence, Big Data and Fundamental Rights Country Research Netherlands 2020 Report provided to FRA under contract D-SE-19-T021*, 2020, s. 28–31. [https://fra.europa.eu/sites/default/files/fra\\_uploads/fra-ai-project-netherlands-country-research\\_en.pdf](https://fra.europa.eu/sites/default/files/fra_uploads/fra-ai-project-netherlands-country-research_en.pdf), dostęp: 1.04.2021.

takiemu podejmowaniu decyzji można spojrzeć z jednej strony z perspektywy skoncentrowanej na technologii, z drugiej zaś strony – z perspektywy kładącej nacisk na zmiany zachodzące w samych procesach podejmowania decyzji i relacji między człowiekiem a maszyną<sup>14</sup>. Szczególnie rozwój uczących się algorytmów stanowi pokusę do zwiększenia efektywności podejmowanych na masową skalę decyzji, np. przez administrację publiczną. Możliwość zastąpienia żmudnych i pracochłonnych analiz dokumentów dokonywanych przez urzędników zunifikowanym kwestionariuszem, którego analizę przeprowadza maszyna i sugeruje decyzję, może istotnie przyspieszyć wiele procesów. W sytuacji, w której maszyna ta jest na dodatek zdolna – na podstawie analizy zbiorów danych dotyczących kolejnych przeprowadzanych przez nią spraw – coraz lepiej rozpoznawać, jakiego rodzaju rozwiązanie wydaje się najkorzystniejsze, zautomatyzowane podejmowanie decyzji może być postrzegane jako ergonomiczna utopia.

Wystarczy jedynie nieco przesunąć akcenty w takiej narracji, skupionej na efektywności i szybkości zautomatyzowanego podejmowania decyzji, aby okazało się, że równie blisko jej do dystopii. Podstawowym problemem związanym z zautomatyzowanym podejmowaniem decyzji są wątpliwości co do tego, na ile jest ono bezstronne. Badanie nad stosowanym w Stanach Zjednoczonych programem COMPAS, który jest używany do oceny prawdopodobieństwa, z jakim dana jednostka popełni ponownie przestępstwo, wykazało, że stosowanie programu prowadzi do wyników powielających niekorzystną sytuację osób o innym niż biały kolorze skóry, a także sugeruje surowsze traktowanie kobiet niż mężczyzn<sup>15</sup>. Podobnie konkluzje badania nad wynikami stosowania systemu PredPol, służącego do wspierania procesów decyzyjnych dotyczących patrolowania danych dzielnic, wskazują na nieproporcjonalną koncentrację na dzielnicach zamieszkałych głównie przez osoby czarnoskóre<sup>16</sup>. Chociaż istnieją badania wykazujące, że stosowanie algorytmów w miejsce ludzi może mieć pozytywne skutki w odniesieniu do traktowania osób narażonych na dyskryminację<sup>17</sup>, problem możliwej dyskry-

---

<sup>14</sup> Zob. konkluzje analizy J. Zerilli, A. Knott, J. Maclaurin, C. Gavaghan, *Algorithmic Decision-Making and the Control Problem*, „Minds and Machines” 2019, nr 29, s. 575.

<sup>15</sup> J. Angwin, J. Larsona, S. Mattu, L. Kirchner, *Machine Bias. There's Software Used across the Country to Predict Future Criminals. And It's Biased against Blacks*, „ProPublica” 2016. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>, dostęp: 1.04.2021.

<sup>16</sup> K. Lum, W. Isaac, *To Predict and Serve?*, „Significance” 2016, t. 13, nr 5, s. 14–19.

<sup>17</sup> Prowadzone w tym kontekście badania dotyczące dostępu do pożyczek udzielanych przy wykorzystaniu zautomatyzowanego podejmowania decyzji, zob. R. Bartlett, A. Morse, R. Stanton, N. Wallace, *Consumer-Lending Discrimination in the FinTech Era*, „NBER Working Paper” 2019, nr 25943, a także analiza B. Lepri, N. Oliver, E. Letouzé, A. Pentland, P. Vinck, *Fair, Transparent, and Accountable Algorithmic Decision-making*

minacji wynikającej ze stosowania systemów zautomatyzowanego podejmowania decyzji pozostaje aktualny.

Ponadto zautomatyzowane podejmowanie decyzji może oznaczać zastępowanie przez rozwiązania technologiczne procesów rządzących się do tej pory normami zawartymi w aktach prawnych lub mogących podlegać kontroli pod względem ich zgodności z normami prawnymi (np. wydawanie decyzji co do pomocy społecznej przysługującej jednostce). Można postawić pytanie o słuszność cedowania na algorytmy i ich twórców możliwości ustalenia, kto na jakie traktowanie zasługuje. Technologie tworzą konkretne osoby, np. programiści i programistki, inżynierowie i inżynierki. W przeciwieństwie do ustawodawców, osoby tworzące algorytmy nie są wybierane w demokratycznych wyborach. Mimo to, wraz z rozpowszechnianiem się zautomatyzowanego podejmowania decyzji, rośnie ich wpływ na tworzenie zasad rządzących życiem społecznym. Nawet w sytuacji, w której stworzone zostają wytyczne dotyczące kryteriów, jakimi mają się kierować twórcy algorytmów, efekt przekładania wytycznych na ostatecznie podejmowane decyzje zostaje ukryty w czarnej skrzynce<sup>18</sup> technologii. Brak demokratycznej kontroli nad oprogramowaniem, które zaczyna pełnić funkcje prawa, może zatem niepokoić.

Brak przejrzystości wprowadzanych rozwiązań sprawia, że na uwagę zasługuje pytanie o to, czy i w jaki sposób prawo może chronić jednostki przed potencjalnie dyskryminacyjnym traktowaniem wynikającym z zautomatyzowanego przetwarzania danych. To właśnie brak dostępu do informacji na temat tego, jak dany system zautomatyzowanego podejmowania decyzji działa, często uniemożliwia jednoznaczne stwierdzenie, czy prowadzi ono, czy też nie, do dyskryminacyjnego traktowania. W analizie nie kwestionuję tego, że możliwe jest tworzenie zautomatyzowanego podejmowania decyzji, którego wyniki będą mniej dyskryminujące niż te osiągnięte w wyniku działań człowieka<sup>19</sup>. Twierdzą natomiast, że ich rozwój może być inspirowany rozwiązaniami instytucjonalno-prawnymi, które będą gwarantować przejrzystość w zakresie wdrażanych zautomatyzowanych rozwiązań i tym samym tworzyć możliwość weryfikacji, czy dany algorytm prowadzi do dyskryminacyjnego traktowania. Jak okazało się w czasie przeprowadzonych przeze mnie badań, sama informacja na temat działania danego systemu zautomatyzowanego

---

*Processes. The Premise, the Proposed Solutions, and the Open Challenges*, „Philosophy and Technology” 2018, t. 31, s. 611–627.

<sup>18</sup> Metafora, która weszła do dyskursu na temat nowych technologii w związku z publikacją F. Pasquale, *The Black Box Society. The Secret Algorithms That Control Money and Information*, Harvard University Press, Cambridge–Londyn 2015.

<sup>19</sup> Przedstawienie takiej perspektywy, zob. J. Kleinberg, J. Ludwig, S. Mullainathan, C.R. Sunstein, *Discrimination in the Age of Algorithms*, „Journal of Legal Analysis” 2018, t. 10, s. 113–174.

podejmowania decyzji jest kluczowa dla możliwości oceny jego funkcjonowania w świetle zakazu dyskryminacji w prawie europejskim, czyli prawie UE oraz Rady Europy<sup>20</sup>.

Dlatego też założeniem stojącym u podstaw przeprowadzonej analizy jest przekonanie o konieczności „odczarowania” istniejących i potencjalnych „czarnych skrzynek” technologii przy wykorzystaniu narzędzi oferowanych przez prawo europejskie. Możliwe jest, że w przyszłości prawo ulegnie daleko idącym przemianom, np. pod wpływem rozwoju tzw. uczenia maszynowego, rejestrów rozproszonych czy inteligentnych kontraktów. Wtedy rozbieżności między technologią i regulującymi ją ramami prawnymi mogą utrudniać podporządkowywanie nowych rozwiązań technologicznych istniejącym ramom prawnym. Punktem wyjścia analizy jest założenie, że dopóki innowacje technologiczne funkcjonują w ramach istniejącego obecnie porządku prawnego, konieczne są próby osadzenia ich w obowiązujących normach prawnych.

Moją intencją jest przeprowadzenie analizy zautomatyzowanego podejmowania decyzji w szerszym świetle pytań o europejską aksjologię. Przede wszystkim mierzę się z zagadnieniem relacji zachodzącej między zakazem dyskryminacji a zautomatyzowanym podejmowaniem decyzji zarówno w odniesieniu do prawa UE, jak i Rady Europy. Ponadto stawiam pytanie o związek między nieprzejrzystym charakterem zautomatyzowanych rozwiązań a wartościami chronionymi przez art. 2 TUE<sup>21</sup>, w tym szczególnie demokracją i państwem prawnym<sup>22</sup>. Wartości UE mogą stanowić istotną wskazówkę, w jaki sposób myśleć o regulacji nowych technologii, a przede wszystkim o wymogach dotyczących ich przejrzystości. Po skandalu związanym z Cambridge Analytyką<sup>23</sup> i z rolą, jaką nowe technologie odegrały w kampaniach

---

<sup>20</sup> Posługuję się pojęciem „prawo europejskie”, kiedy odnoszę się do prawa UE i Rady Europy łącznie. Kiedy pisze tylko o prawie UE lub tylko o prawie Rady Europy, używam odpowiednio tych terminów precyzyjnie odnoszących się jedynie do danej organizacji. Ze względu na tak zarysowany zakres pracy nie zajmuję się analizą prawa krajowego. W zakresie polskich publikacji dotyczących jawności należy wskazać wielotomowe wydanie *Jawność i jej ograniczenia*, red. G. Szpor, C.H. Beck, Warszawa 2013–2016. Tom XI serii, zatytułowany *Standardy europejskie*, red. C. Mik, C.H. Beck, Warszawa 2016, najbliższy jest tematowi analizy i odsyłam do niego w odpowiednich miejscach.

<sup>21</sup> *Traktat o Unii Europejskiej* (TUE), Dz.Urz. EU 2012 C 326, s. 13–46.

<sup>22</sup> W odniesieniu do UE analizuję również zagadnienie dostępu do zasad stojących za systemami zautomatyzowanego podejmowania decyzji w świetle „otwartości” przywoływanej w art. 15, *Traktat o funkcjonowaniu Unii Europejskiej* (TFUE), Dz.Urz. EU 2012 C 326, s. 47–390.

<sup>23</sup> Skandal dotyczył problemu targetowanych reklam i ich wpływu na wyniki wyborcze. Dane pozyskane w celu analizy profili psychologicznych jednostek zostały w dużym stopniu pobrane bez wiedzy i zgody osób, których dotyczyły, ponieważ osoby wypełniające test stanowiący jedno z ich źródeł udostępniły jednocześnie informa-

wyborczych i referendalnych ostatnich lat, trudno kwestionować, że tworzą one poważne wyzwanie dla demokracji i praworządności.

Przy pisaniu towarzyszyło mi przekonanie, że wydarzenia te wymuszają próbę odejścia od indywidualistycznej perspektywy skupionej na pojedynczym podmiocie i zwrócenia się ku instytucjom, normom i środkom prawnym stawiającym w centrum zbiorowość i jej interes. Jak wskazuje motto książki, zautomatyzowane podejmowanie decyzji nie jest problemem dotyczącym wyłącznie jednostki – jest ono zjawiskiem dotyczącym poszczególne osoby, ale jako członków konkretnych grup społecznych lub przedstawicieli algorytmicznie wygenerowanej zbiorowości. Dlatego też odpowiedzi na wyzwania tworzone przez zautomatyzowane podejmowanie decyzji powinny uwzględniać ten kolektywny wymiar stosowanych w tym zakresie mechanizmów.

## 1. Zakres badań, hipoteza i pytania badawcze

Celem analizy jest weryfikacja hipotezy mówiącej, że możliwe jest postrzeganie algorytmów używanych w zautomatyzowanym podejmowaniu decyzji jako informacji publicznej lub dokumentów urzędowych<sup>24</sup>. Rozwiązanie takie pojawia się wśród propozycji, które są formułowane jako narzędzia mające na celu zwiększenie poziomu przejrzystości zautomatyzowanego podejmowania decyzji<sup>25</sup>. Analiza takiej możliwości została już przeprowadzona dla ram prawnych dotyczących dostępu do informacji w Stanach

---

cję na temat swoich znajomych na Facebooku. Artykuły prasowe na temat skandalu, zob. Z. Kleinman, *Cambridge Analytica: The story so far*, BBC, 21 marca 2018 r. <https://www.bbc.co.uk/news/technology-43465968>, dostęp: 1.04.2021 czy J.C. Wong, *The Cambridge Analytica Scandal Changed the World – But It Didn't Change Facebook*, „The Guardian”, 18 marca 2019 r. <https://www.theguardian.com/technology/2019/mar/17/the-cambridge-analytica-scandal-changed-the-world-but-it-didnt-change-facebook>, dostęp: 1.04.2021.

<sup>24</sup> Używam zamiennie terminów „prawo dostępu do informacji” i „prawo dostępu do informacji publicznej” ze względu na fakt, że zarówno ETPCz, jak i TSUE często ogranicza się do mówienia właśnie o „dostęp do informacji”, pomijając przymiotnik „publicznej”. Podobnie zamiennie używam terminów „prawo dostępu do dokumentów” i „prawo dostępu do dokumentów urzędowych” dla większej przejrzystości. Natomiast w zakresie różnic między znaczeniem pojęcia „dokument” i „informacja” przeprowadzam analizę dotyczącą ich znaczenia w prawie Rady Europy i prawie UE odpowiednio w trzecim i czwartym rozdziale.

<sup>25</sup> Dostęp do kodu źródłowego jako jeden ze sposobów przeciwdziałania niejawności zautomatyzowanego podejmowania decyzji pojawia się np. w raporcie Omidyar Network, Upturn, *Public Scrutiny of Automated Decisions: Early Lessons and Emerging Methods An Upturn and Omidyar Network Report*, 2018, s. 19. <https://omidyar.com/wp-content/uploads/2020/09/Public-Scrutiny-of-Automated-Decisions.pdf>, dostęp: 1.04.2021.

Zjednoczonych<sup>26</sup> oraz Australii<sup>27</sup>. Poniższa analiza ma natomiast na celu sprawdzenie, czy takie podejście do zautomatyzowanego podejmowania decyzji jest uprawnione w świetle prawa UE oraz prawa Rady Europy – a jeżeli tak, to w jakim zakresie.

Analiza jest motywowana wątpliwościami, które zautomatyzowane podejmowanie decyzji może budzić w zakresie zgodności stosowanych w nim rozwiązań z zakazem dyskryminacji w prawie europejskim. W książce skupiam się na zautomatyzowanym podejmowaniu decyzji w stosunku do jednostek w obszarach zadań publicznych wykonywanych przez państwo lub podmioty upoważnione przez państwo<sup>28</sup>. Przykładami takich sytuacji są podejmowanie decyzji dotyczącej wymiaru pomocy udzielonej osobie bezrobotnej, wdrażanie systemu zautomatyzowanego przydzielania spraw sędziom czy wykorzystywanie systemów sugerujących odpowiedni wymiar czy typ kary sędziom wydającym wyroki. Zakres badań obejmuje zatem jedynie część sytuacji, w których zautomatyzowane podejmowanie decyzji znajduje zastosowanie<sup>29</sup>.

Próba analizy algorytmów, jako informacji podlegającej prawu dostępu do informacji publicznej lub dokumentów urzędowych, wynika ze spostrzeżenia, że w demokratycznym społeczeństwie jest nie do przyjęcia, aby nieprzejrzysta technologia pełniła funkcję organu administracyjnego lub prawodawcy. Szczególnie jeśli istnieją wątpliwości co do zgodności jej stosowania z zakazem dyskryminacji. W związku z tym w celu weryfikacji głównej hipotezy niezbędne jest poszukanie odpowiedzi na wymienione poniżej pytania badawcze.

---

<sup>26</sup> Zob. H. Bloch-Wehba, *Access to Algorithms*, „Fordham Law Review” 2020, t. 88, s. 1265–1314. Analiza Bloch-Wehby oparta jest m.in. na orzecznictwie amerykańskich sądów i ukazuje brak jednolitego podejścia do kwestii umożliwiania dostępu do algorytmów jako informacji publicznej. Tekst zamyka podkreślenie konieczności wprowadzania nowych rozwiązań prawnych, które stanowiłyby bardziej adekwatną odpowiedź na wyzwania tworzone przez zautomatyzowane podejmowanie decyzji.

<sup>27</sup> Zob. A. Ray, B. Adams, D. Thampapillai, *Access to Algorithms Post-Robodebt: Do Freedom of Information Laws Extend to Automated Systems?*, „Alternative Law Journal” 2021 (online first). <https://doi.org/10.1177/1037969X211029434>, dostęp: 18.07.2021. Konkluzją autorów jest niejasny status systemów zautomatyzowanego podejmowania decyzji z punktu widzenia prawa dostępu do informacji. Wskazują oni jednak, że wydaje się możliwe traktowanie jako informacji publicznej przynajmniej części takich systemów.

<sup>28</sup> Wybór taki w żaden sposób nie powinien być odczytywany jako próba umniejszania konieczności gwarantowania ochrony praw jednostki w odniesieniu do stosowania zautomatyzowanego podejmowania decyzji w ramach sektora prywatnego. Jest on związany z charakterem prawa dostępu do informacji i do dokumentów urzędowych, których potencjał w odniesieniu do uzyskiwania dostępu do kodu komputerowego używanego w ramach zautomatyzowanego podejmowania decyzji jest przedmiotem przedstawionego badania.

<sup>29</sup> W związku z tak określonym zakresem tematycznym nie uwzględniam w książce szczegółowych rozważań doktrynalnych dotyczących części używanych przeze mnie pojęć i poruszanych zagadnień.



Po pierwsze, konieczne jest wykazanie, czy rozwój systemów zautomatyzowanego podejmowania decyzji i rozpowszechnianie ich stosowania w ramach sektora publicznego mogą być postrzegane jako procesy zmieniające prawo i zastępujące normy prawne kodem komputerowym. Część analizy poświęcona tym zagadnieniom jest osadzona w nurcie teorii przedstawiającej kod komputerowy jako prawa, a rozwijanej przez m.in. Lawrence'a Lessiga i Mireille Hildebrandt.

Po drugie, niezbędne jest przedstawienie przyczyn, w związku z którymi przejrzystość algorytmów jest istotnym zagadnieniem badawczym w odniesieniu do prawa europejskiego. Wśród nich kluczowe jest wskazanie na zagrożenia związane z brakiem przejrzystości oraz z dyskryminacyjnym traktowaniem jednostek w obszarach objętych prawem UE i w odniesieniu do praw i wolności chronionych przez EKPCz<sup>30</sup> w związku z wdrażaniem systemów zautomatyzowanego podejmowania decyzji. Na uwagę zasługuje analiza mechanizmów prowadzących do częściowego lub całkowitego zastępowania regulacji prawnych rozwiązaniami technologicznymi, czyli zautomatyzowanym podejmowaniem decyzji, które oparte są na klasyfikowaniu jednostek do grup, w stosunku do których podejmowane są dane typy decyzji. Scharakteryzowanie wyzwań tworzonych przez takie mechanizmy dla zakazu dyskryminacji oraz dla przejrzystości – która determinuje możliwość rozpoznania dyskryminacyjnego charakteru stosowanych rozwiązań – pozwala uzasadnić konieczność poszukiwania rozwiązań prawnych dotyczących nie tylko jednostkowych decyzji wydawanych z użyciem zautomatyzowanego podejmowania decyzji (obecnych w prawie ochrony danych osobowych), lecz także algorytmów używanych do ich wydawania.

Po trzecie, istotne jest sprawdzenie, czy ramy regulacyjne stworzone przez Ogólne rozporządzenie o ochronie danych (RODO)<sup>31</sup> oraz działania podejmowane przez Radę Europy w obszarze zautomatyzowanego podejmowania decyzji odnoszą się w adekwatny sposób do zagrożeń związanych z brakiem przejrzystości zautomatyzowanego podejmowania decyzji. Kryterium, które jest istotne do dokonania takiej oceny, jest m.in. precyzja przyjętych przepisów. Interesuje mnie również, na ile odnoszą się one do decyzji podejmowanych wobec danej jednostki, a na ile pozwalają ocenić system zautomatyzowanego podejmowania decyzji sam w sobie. Odpowiedź na te

---

<sup>30</sup> *Konwencja o ochronie praw człowieka i podstawowych wolności sporządzona w Rzymie dnia 4 listopada 1950 r., zmieniona następnie Protokołami nr 3, 5 i 8 oraz uzupełniona Protokołem nr 2 (EKPCz)*, Dz.U. 1993 nr 61 poz. 284.

<sup>31</sup> *Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2016/679 z dnia 27 kwietnia 2016 r. w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych i w sprawie swobodnego przepływu takich danych oraz uchylenia dyrektywy 95/46/WE (ogólne rozporządzenie o ochronie danych) (Tekst mający znaczenie dla EOG) (RODO)*, Dz.Urz. UE 2016 L 119, s. 1–88.

pytania pozwala skonfrontować rozwiązania problemów związanych z zautomatyzowanym podejmowaniem decyzji przyjmowane w obszarze ochrony danych osobowych z kolektywnym wymiarem używania zautomatyzowanych metod podejmowania decyzji.

Po czwarte, należy wskazać, czy RODO pozostawia część zagadnień związanych z możliwością ochrony jednostki przed nieprzejrzystym i potencjalnie dyskryminacyjnym charakterem zautomatyzowanego podejmowania decyzji w zakresie ustawodawstwa państw. Dyskrecjonalność przyjmowanych środków ochrony praw jednostek może prowadzić do efektu sprzecznego z celem przyjęcia rozporządzenia, czyli tworzeniem różnych standardów ochrony w ramach UE w odniesieniu do zautomatyzowanego podejmowania decyzji. Luki w istniejących ramach prawnych, dotyczących ochrony praw jednostki w odniesieniu do używania przez państwa zautomatyzowanych systemów podejmowania decyzji, stanowiły dla mnie motywację do poszukiwania innych perspektyw pozwalających zmierzyć się z problemami wynikającymi z ich stosowania.

Po piąte, konieczne jest dokonanie analizy warunków, pod którymi możliwe jest korzystanie z prawa dostępu do informacji zgodnie z wykładnią EKPCz. Przedstawienie testu umożliwiającego sprawdzenie, czy w danej sytuacji dostęp do informacji objęty jest prawem dostępu do informacji, pozwala stwierdzić, czy istnieją przypadki, w których zasady działania systemów służących do zautomatyzowanego podejmowania decyzji w postaci kodu komputerowego mogłyby być postrzegane jako informacja publiczna w świetle orzecznictwa ETPCz.

Po szóste, niezbędne jest zestawienie podejścia prezentowanego przez ETPCz z wykładnią odnoszącą się do prawa dostępu do informacji dokonywaną przez TSUE. W tym celu konieczne jest przedstawienie kształtowania się prawa dostępu do dokumentów urzędowych w prawie UE oraz orzecznictwa TSUE dotyczącego tej kwestii. Wskazanie rozbieżności między perspektywą prezentowaną w orzecznictwie ETPCz a podejściem przyjętym przez TSUE pozwala wykazać ograniczenia unijnych regulacji dotyczących dostępu do dokumentów urzędowych oraz prawa dostępu do informacji, a także ich konsekwencje dla przejrzystości zautomatyzowanego podejmowania decyzji.

Ponadto prezentowane przez TSUE podejście do zagadnienia prawa dostępu do informacji oraz prawa dostępu do dokumentów urzędowych rodzi pytania o relację między tymi prawami a wartościami stojącymi u podstaw UE, dlatego staram się zbadać związek między „demokratyzacją”, której służyć miało m.in. wprowadzenie prawa dostępu do dokumentów urzędowych UE, a „demokracją” jako wartością wymienianą w art. 2 TUE. W efekcie stawiam pytanie, czy narzędzie mające w zamyśle służyć demokratyzacji UE może być postrzegane jako element niezbędny dla zgodności z zasadą demokracji i praworządności porządków prawnych państw członkowskich?

Czy możliwe jest wymaganie od państw członkowskich, aby działały one w zgodzie z zasadą przejrzystości postrzeganej jako niezbędny element demokracji? Próba odpowiedzi osadza pracę w szerszym kontekście pytań o specyfikę europejskiego podejścia do regulacji nowych technologii i jego związku z wartościami chronionymi przez prawo europejskie.

## 2. Metody i materiał badawczy

Analiza została przeprowadzona z wykorzystaniem metody formalno-dogmatycznej oraz komparatystycznej<sup>32</sup>. Pomocniczo, przede wszystkim w odniesieniu do historii rozwoju wykładni prawa dostępu do informacji, stosowana jest metodologia historyczno-prawna. Materiał, na którym opierają się badania, obejmuje akty prawne UE i Rady Europy dotyczące ochrony danych osobowych oraz dostępu do dokumentów urzędowych i dostępu do informacji, a także relewantne dla tematu orzecznictwo TSUE i ETPCz. Wykorzystuję również dokumenty i raporty przygotowane przez UE i Radę Europy oraz na ich zlecenie. Kluczowym punktem odniesienia dla analizy są ustalenia na temat zautomatyzowanego podejmowania decyzji prezentowane przez badaczy, a więc sfera doktrynalna. Książka opiera się głównie na analizie przeprowadzonej w latach 2017–2020. W lipcu 2021 r. została ona uzupełniona o informacje o najnowszych dokumentach, orzecznictwie i literaturze, istotnych z punktu widzenia jej tematu.

Wśród źródeł, z których korzystam pomocniczo, znajdują się raporty i artykuły prasowe przygotowywane przez rozmaite fundacje, stowarzyszenia, oddolne inicjatywy pozarządowe czy dziennikarzy na temat algorytmów i zautomatyzowanego podejmowania decyzji. Powodem wykorzystywania tych publikacji jest z jednej strony ich interdyscyplinarność, która pozwala spojrzeć na badane przeze mnie zagadnienia z szerszej perspektywy – niezbędnej w świetle tematu badania, z drugiej zaś strony – powolne tempo publikacji w czasopiśmie naukowych. Rozwój technologiczny sprawia, że badane przeze mnie zagadnienia podlegają niezwykle dynamicznym zmianom i dyskusje na ich temat odbywają się niekiedy poza obiegiem naukowym zależnym od powolnego procesu wydawniczego. Źródła te są użyteczne ze względu na to, że zawierają np. przystępnie opisane przykłady problemów związanych z zautomatyzowanym podejmowaniem decyzji, czyli sferę faktów dotyczących analizowanej technologii.

---

<sup>32</sup> W zakresie, w jakim rozwiązania przyjęte w ramach UE są porównywane z tymi, które pojawiają się w ramach działalności Rady Europy i w wyniku wykładni EKPCz.

### 3. Streszczenie rozdziałów

Na książkę składają się cztery rozdziały. W rozdziale pierwszym, po pierwsze, dokonuję przeglądu teoretycznych conceptualizacji znaczenia rozwoju zautomatyzowanego przetwarzania danych. Przedstawione są w nim wyzwania związane z intensyfikacją interakcji między prawem i technologią oraz wynikające z nich zagrożenia dla przejrzystości i – szerzej – demokratycznych procedur stanowienia i stosowania prawa. Po drugie, wykazuję źródła braku przejrzystości zautomatyzowanego podejmowania decyzji oraz dyskryminacji algorytmicznej. Przedstawiony wywód ma za zadanie udzielenie odpowiedzi, czy przejawy dyskryminującego potencjału zautomatyzowanego podejmowania decyzji mogą dotyczyć obszarów objętych prawem europejskim i w rezultacie prowadzić do tworzenia zagrożeń dla zgodności z zakazem dyskryminacji. Po trzecie, przedstawiam przyjęte przeze mnie ramy teoretyczne. Kładę nacisk na konieczność postrzegania zagrożeń związanych z rozwojem zautomatyzowanego podejmowania decyzji w odniesieniu do zakazu dyskryminacji, jako problemu dotyczącego całych grup i wynikającego z samych systemów zautomatyzowanego podejmowania decyzji. Dążenie do zwiększenia ich przejrzystości jest krokiem w stronę zagwarantowania możliwości sprawdzenia, czy nie naruszają one zakazu dyskryminacji.

Aby ocenić, czy konieczne jest poszukiwanie nowych rozwiązań problemów pojawiających się w odniesieniu do relacji między zautomatyzowanym podejmowaniem decyzji a zakazem dyskryminacji, niezbędne jest przedstawienie relewantnych rozwiązań regulacyjnych wypracowanych do tej pory, czemu służy rozdział drugi. Po pierwsze, odpowiada on na pytanie badawcze dotyczące tego, w jakim stopniu ramy regulacyjne stworzone przez prawo ochrony danych osobowych odpowiadają na zagrożenia związane z potencjałem dyskryminacyjnym zautomatyzowanego podejmowania decyzji. Odpowiedź na to pytanie jest udzielona przede wszystkim na podstawie analizy przepisów RODO, w tym zwłaszcza art. 13–15 oraz art. 22. Ponadto przedstawiam analizę innych praw podmiotów danych oraz rozwiązań odnoszących się do systemowego wymiaru zautomatyzowanego podejmowania decyzji (ocena skutków dla ochrony danych – DPIA i wiążące reguły korporacyjne – BCR) w RODO. Ponieważ wciąż nie ma orzecznictwa relewantnego dla zagadnienia, ogromną rolę dla prowadzonej analizy odgrywają debata akademicka tocząca się na temat zautomatyzowanego podejmowania decyzji i jego regulacji w RODO, jak również wskazówki wypracowane w tym zakresie przez Grupę Roboczą Art. 29 i EROD<sup>33</sup>. Po drugie, przedstawiam działania podjęte przez Radę Europy w obszarze zautomatyzowanego podejmowania

---

<sup>33</sup> Grupa Robocza Art. 29 została zastąpiona w związku z reformą unijnego prawa ochrony danych osobowych przez EROD.

decyzji, czyli propozycje zmian w treści *Konwencji 108*<sup>34</sup> (tzw. *Konwencja 108+*<sup>35</sup>). Jednym z nowych praw sformułowanych w protokole zmieniającym *Konwencję 108* jest prawo do niepodlegania decyzji opartej na zautomatyzowanym przetwarzaniu danych. Analiza wzajemnego oddziaływania rozwiązań z prawa unijnego i tych przyjmowanych w ramach działalności Rady Europy pozwala zarysować pełny obraz stanu prawnego dotyczącego zautomatyzowanego podejmowania decyzji w prawie europejskim.

Celem rozdziału trzeciego jest częściowa weryfikacja głównej hipotezy. Poświęcony jest on próbie zbadania i wykazania, czy prawo dostępu do informacji może stanowić narzędzie prowadzące do zwiększania poziomu przejrzystości algorytmów, a jeśli tak, to pod jakimi warunkami. Rozdział poświęcony jest analizie charakteru prawa dostępu do informacji w ramach działalności Rady Europy. Rozpoczyna go krótkie wprowadzenie dotyczące roli odgrywanej przez prawo dostępu do informacji w historii rozwoju międzynarodowych umów dotyczących praw człowieka oraz w działalności Rady Europy. Analiza art. 10 EKPCz wymaga sięgnięcia do historii międzynarodowych inicjatyw wiążących się z wolnością informacji, ponieważ to dyskusje na jej temat i negocjacje międzynarodowych umów dotyczących praw człowieka stanowiły inspirację dla treści EKPCz. Dlatego w rozdziale została zarysowana historia udziału ONZ w kształtowaniu standardów dotyczących dostępu do informacji, a na jej tle przedstawiono kontrowersje związane z art. 10 EKPCz i prawem dostępu do informacji. Ze względu na wagę perspektywy porównawczej w odniesieniu do zrozumienia wątpliwości wiążących się z obecnością prawa dostępu do informacji w art. 10 EKPCz w tym fragmencie analizy została zastosowana metoda komparatystyczna oraz do pewnego stopnia historyczno-prawna. Obecność tego podrozdziału znajduje uzasadnienie m.in. w argumentacji przytaczanej przez ETPCz w orzecznictwie odnoszącym się do prawa do informacji. W związku z wagą wyroków ETPCz dla tej kwestii, kluczową częścią rozdziału trzeciego jest analiza orzecznictwa dotyczącego prawa dostępu do informacji. Zostały w niej przedstawione zarówno warunki, których spełnienie determinuje możliwość powołania się na prawo do informacji jako część dorobku europejskiego systemu ochrony praw człowieka, jak i możliwe ograniczenia tego prawa. Uwzględniona została również możliwość traktowania prawa dostępu do informacji jako narzędzia służącego realizacji innych praw gwarantowanych w EKPCz. Rozdział przedstawia również analizę przeprowadzoną ściśle w odniesieniu do

<sup>34</sup> Polska wersja językowa: *Konwencja Nr 108 Rady Europy o ochronie osób w związku z automatycznym przetwarzaniem danych osobowych, sporządzona w Strasburgu dnia 28 stycznia 1981 r.* (Konwencja 108), Dz.U. 2003 nr 3 poz. 25 (*Convention for the Protection of Individuals with Regard to Automatic Processing of Personal Data*, ETS No. 108).

<sup>35</sup> *Protocol amending the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data* (Konwencja 108+), CETS 223.

zautomatyzowanego podejmowania decyzji: ma ona pomóc w dokładnym przedstawieniu możliwych zastosowań prawa dostępu do informacji do algorytmów i tym samym doprowadzić do weryfikacji głównej hipotezy w zakresie dotyczącym działalności Rady Europy.

Pierwsza część rozdziału czwartego przedstawia kolejno zarys historyczny rozwoju prawa dostępu do dokumentów w UE oraz szerszy kontekst reform instytucjonalnych UE i ewolucji horyzontów aksjologicznych integracji europejskiej w zakresie istotnym dla wykazania ich znaczenia dla prawa dostępu do dokumentów. Taka struktura podrozdziału jest podyktowana chęcią skupienia się na tych elementach prawa dostępu do dokumentów w prawie UE, które są istotne z punktu widzenia analizy, w odniesieniu zaś do innych zagadnień ogranicza się do odsyłania do literatury przedmiotu. Drugą część rozdziału otwiera analiza *dyrektywy 2019/1024*<sup>36</sup>. Jest ona poświęcona dostępowi do dokumentów i informacji w prawie UE w świetle zautomatyzowanego podejmowania decyzji wprowadzanego przez państwa członkowskie. Znajduje się w niej również analiza osadzona na szerszym tle zagadnienia zakresu ochrony praw podstawowych w UE oraz związku między wartościami z art. 2 TUE a prawem dostępu do dokumentów i prawem dostępu do informacji.

Każdy z rozdziałów zamyka krótkie podsumowanie opisanych w nim zagadnień, natomiast książkę wieńczy zakończenie stanowiące zestawienie wniosków wynikających z weryfikacji głównej hipotezy i odpowiedzi na stawiane pytania badawcze.

#### 4. Terminologia

Przedmiotem zainteresowania analizy są algorytmy służące do zautomatyzowanego podejmowania decyzji (*automated decision-making*). Z perspektywy istniejących regulacji należy odnotować, że pojęcie zautomatyzowanego podejmowania decyzji jest terminem używanym w prawie ochrony danych osobowych, które reguluje wybrane sytuacje posługiwania się takimi rozwiązaniami. W świetle hipotezy weryfikowanej w tej książce niezbędne jest jednak nie tylko zarysowanie znaczenia tego terminu zgodnie z jego charakterystyką zawartą w RODO, lecz również uzupełnienie go o definicję pojęć takich jak algorytm i sztuczna inteligencja. Przyjrzenie się temu, czym są algorytmy, czyli rozwiązania stojące za systemami zautomatyzowanego podejmowania decyzji, dostarcza argumentów za możliwością postrzegania zautomatyzowanego podejmowania decyzji jako narzędzia mogącego kształtować

---

<sup>36</sup> *Dyrektywa Parlamentu Europejskiego i Rady (UE) 2019/1024 z dnia 20 czerwca 2019 r. w sprawie otwartych danych i ponownego wykorzystywania informacji sektora publicznego* (dyrektywa 2019/1024), Dz.Urz. EU 2019 L 172, s. 56–83.

sytuację prawną jednostki i w związku z tym zmieniającego rozumienie, czym jest prawo.

#### 4.1. Zautomatyzowane podejmowanie decyzji i profilowanie

W związku z brakiem definicji legalnej terminu „zautomatyzowane podejmowanie decyzji” punktem wyjścia do analizy jest definicja, którą można wywodzić z treści art. 22 RODO zatytułowanego „Zautomatyzowane podejmowanie decyzji w indywidualnych przypadkach, w tym profilowanie”<sup>37</sup>. Stwierdza on, że osoba, której dane dotyczą, ma prawo nie podlegać decyzji opierającej się wyłącznie na zautomatyzowanym przetwarzaniu, w tym profilowaniu, i wywołującej wobec tej osoby skutki prawne lub w podobny sposób istotnie na nią wpływającej. Taki sposób patrzenia na zautomatyzowane podejmowanie decyzji jest neutralny technologicznie: obejmuje on zarówno prosty algorytm, jak i zaawansowany technologicznie system wykorzystujący uczenie maszynowe, jak długo decyzja wydawana z jego użyciem spełnia warunki wymienione w art. 22 RODO.

W związku z niejasnym znaczeniem określeń użytych w art. 22, takich jak „całkowicie zautomatyzowany” oraz „decyzja wywołująca skutki prawne”, należy odwołać się do wytycznych na temat zautomatyzowanego podejmowania decyzji przygotowanych przez Grupę Roboczą Art. 29<sup>38</sup>. Wytyczne Grupy Roboczej Art. 29 nie mają co prawda mocy prawnie

<sup>37</sup> Art. 22, RODO. Przepis ten analizuję szczegółowo w rozdziale drugim, tutaj jedynie zaznaczam kluczowe elementy definicyjne niezbędne do określenia zakresu analizy. Rozwiązania dotyczące zautomatyzowanego podejmowania decyzji wprowadzone w RODO zostały powtórzone w art. 24, *Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2018/1725 z dnia 23 października 2018 r. w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych przez instytucje, organy i jednostki organizacyjne Unii i swobodnego przepływu takich danych oraz uchylenia rozporządzenia (WE) nr 45/2001 i decyzji nr 1247/2002/WE (Tekst mający znaczenie dla EOG)* (rozporządzenie 2018/1725), Dz.Urz. UE 2018 L 295, s. 39–98. W nieco odmienny sposób są one regulowane w art. 11, *Dyrektywa Parlamentu Europejskiego i Rady (UE) 2016/680 z dnia 27 kwietnia 2016 r. w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych przez właściwe organy do celów zapobiegania przestępczości, prowadzenia postępowań przygotowawczych, wykrywania i ścigania czynów zabronionych i wykonywania kar, w sprawie swobodnego przepływu takich danych oraz uchylenia decyzji ramowej Rady 2008/977/WSiSW* (dyrektywa 2016/680), Dz.Urz. UE 2016 L 119, s. 89–131. Na problemy związane z taką strategią regulacji ochrony danych osobowych zwracała uwagę już na etapie prac nad projektami tych aktów prawnych A. Grzelak, *Projekt reformy ochrony danych osobowych – czy rzeczywiście powstanie jednolity i spójny system?*, „Kwartalnik Kolegium Ekonomiczno-Społecznego. Studia i Prace” 2014, nr 4, s. 93–125.

<sup>38</sup> Grupa Robocza Art. 29, *Wytyczne w sprawie zautomatyzowanego podejmowania decyzji w indywidualnych przypadkach i profilowania do celów rozporządzenia 2016/679/UE*, 3 października 2017 r., ostatnio zmienione i przyjęte 6 lutego 2018 r., 17/PL WP 251 rev.01. Rozważania doktryny na temat art. 22 przedstawiam w rozdziale drugim.

wiążące, stanowią natomiast istotną wskazówkę dla wykładni pojęć pojawiających się w RODO. Zgodnie z wytycznymi Grupy Roboczej Art. 29, aby przetwarzanie danych nie zostało uznane za całkowicie zautomatyzowane, to interwencja ludzka musi pozwalać na istotny nadzór nad podejmowaniem decyzji. Ponadto działanie musi być podejmowane przez osobę, która ma uprawnienia i kompetencje do zmiany decyzji, i musi uwzględniać wszystkie dane istotne dla danego podejmowania decyzji<sup>39</sup>. Pozorny udział człowieka w podejmowaniu decyzji, polegający np. jedynie na zatwierdzeniu decyzji wskazywanej przez algorytm, nie będzie zatem stanowić przesłanki do wyłączenia danego podejmowania decyzji z zakresu zastosowania zakazu z art. 22 RODO.

Wytyczne pozwalają ponadto uniknąć interpretacji problematycznego pojęcia decyzji, przez uwzględnienie całych fraz użytych w RODO, czyli „decyzja wywołująca skutki prawne” lub „podobny istotny wpływ”. Za decyzję taką należy uznać dokonanie wyboru, który „ma [...] wpływ na prawa danej osoby przewidziane w przepisach, takie jak wolność do zrzeszania się, głosowania w wyborach lub wystąpienia na drogę sądową” lub „na status prawny danej osoby lub na prawa przysługujące jej na mocy umowy”<sup>40</sup>. Proponowany przez Grupę Roboczą Art. 29 zakres znaczenia zautomatyzowanego podejmowania decyzji w indywidualnych przypadkach jest zatem bardzo szeroki. Zdaniem Grupy Roboczej Art. 29 zakaz z art. 22 powinien odnosić się do wszelkich praw jednostki, które są zagwarantowane w przepisach<sup>41</sup>.

Zakres sytuacji objętych zakazem z art. 22 poszerza dodatkowo uwzględnienie kategorii decyzji, które na sytuację jednostki mają „podobnie istotny wpływ”. Przykłady takich sytuacji pojawiają się w motywie 71 RODO (automatyczne odrzucenie elektronicznego wniosku kredytowego, elektroniczne

---

<sup>39</sup> Grupa Robocza Art. 29, *Wytyczne w sprawie zautomatyzowanego podejmowania decyzji...*, *op. cit.*, s. 23.

<sup>40</sup> *Ibidem*, s. 24.

<sup>41</sup> Przykładami są np. kwestie udzielania świadczeń społecznych czy mechanizmy rekrutacyjne zdefiniowane przepisami prawa. W podobnym duchu, zob. G. Malgieri, G. Comandé, *Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation*, „International Data Privacy Law” 2017, t. 7, nr 4, s. 252. To, co pozostaje niejasne, to kwestia, czy te elementy praw jednostki zagwarantowanych w przepisach, które nie są w przepisach zdefiniowane, mogą być również objęte zakresem stosowania art. 22 RODO. Jeśli np. system zautomatyzowanego podejmowania decyzji jest używany do określania dzielnic będących priorytetem dla kontroli policji opartej na kryteriach dyskryminacyjnych, to czy można uznać, że stanowi on naruszenie jakiegoś prawa jednostki gwarantowanego przepisami prawa, mimo że to konkretne, dane zagadnienie nie jest regulowane w przepisach? W świetle RODO sytuacje takie powinny być prawdopodobnie uznane za mające „podobnie istotny wpływ” na jednostkę.



metody rekrutacji bez interwencji ludzkiej<sup>42</sup>). Wytyczne Grupy Roboczej Art. 29 sugerują konieczność szerokiej interpretacji tej kategorii jako uwzględniającej sytuację, w których decyzja stwarza ryzyko „wywarcia istotnego wpływu na sytuację danej osoby, jej zachowanie lub podejmowane przez nią wybory, wywarcia długotrwałego lub trwałego wpływu na osobę, której dane dotyczą lub w najbardziej ekstremalnym przypadku, doprowadzenia do wykluczenia lub dyskryminacji osób fizycznych”<sup>43</sup>. W ich obrębie mogą znaleźć się nawet decyzje dotyczące marketingu *online*, w sytuacji, w której będą one prowadzić do istotnego wpływu na jednostkę, np. w odniesieniu do jej sytuacji finansowej<sup>44</sup>.

Należy podkreślić, że w art. 22 RODO została uwzględniona również szczególna kategoria zautomatyzowanego podejmowania decyzji, czyli zautomatyzowane podejmowanie decyzji oparte na profilowaniu<sup>45</sup>. Termin „profilowanie”, zdefiniowany w art. 4 RODO, oznacza dowolną formę zautomatyzowanego przetwarzania danych osobowych, które polega na wykorzystaniu danych osobowych do oceny niektórych czynników osobowych osoby fizycznej<sup>46</sup>. Definicja profilowania z art. 4 nie pokrywa się całkowicie z zakresem sytuacji, do których odnosi się art. 22. Nie jest to dostrzegane przez część polskich autorów komentujących RODO, np. Marcin Zadrożny nie do końca trafnie określa prawa jednostki zdefiniowane w art. 22 RODO jako „prawa związane z profilowaniem”<sup>47</sup>. Podobnie Witold Chomiczewski odwołuje się również do pojęcia profilowania, referując przepisy dotyczące zautomatyzowanego podejmowania decyzji<sup>48</sup>. Natomiast różnicę między zakresem znaczeniowym tych pojęć odnotowuje np. Arwid Mednis, który zauważa, że „[...] profilowanie i podejmowanie zautomatyzowanych decyzji to dwie

<sup>42</sup> Motyw 71, RODO.

<sup>43</sup> Grupa Robocza Art. 29, *Wytyczne w sprawie zautomatyzowanego podejmowania decyzji...*, *op. cit.*, s. 24.

<sup>44</sup> Taki sposób rozumienia sytuacji używania zautomatyzowanego podejmowania decyzji podlegających regulacji RODO stanowi również argument za przyjęciem perspektywy szerokiej, w której są analizowane możliwości otrzymania dostępu do algorytmów oparte na prawie dostępu do informacji lub prawie dostępu do dokumentów urzędowych i którą przyjmuję – jak zauważyłam powyżej – w książce.

<sup>45</sup> Przykładem analizy wychodzącej od pojęcia profilowania i charakteryzującej ramy prawne dotyczącego tego zagadnienia w obszarze prawa ochrony danych osobowych jest: M. Ciechomska, *Prawne aspekty profilowania oraz podejmowania zautomatyzowanych decyzji w ogólnym rozporządzeniu o ochronie danych osobowych*, „Europejski Przegląd Sądowy” 2017, nr 5, s. 37–42.

<sup>46</sup> Art. 4 ust. 4, RODO.

<sup>47</sup> A. Dmochowska, M. Zadrożny, *Unijna reforma ochrony danych osobowych. Analiza zmian*, C.H. Beck, Warszawa 2016, s. 49–52.

<sup>48</sup> W. Chomiczewski w *Profilowanie w ogólnym rozporządzeniu o ochronie danych*, [w:] *Polska i Europejska Reforma Ochrony Danych Osobowych*, red. E. Bielak-Jomaa, D. Lubasz, Wolters Kluwer Polska, Warszawa 2016, s. 127–138.

różne czynności, które mogą być ze sobą powiązane, ale mogą też występować niezależnie od siebie”<sup>49</sup>.

Zakres sytuacji objętych treścią art. 22 RODO obejmuje np. profilowanie oparte całkowicie na zautomatyzowanym podejmowaniu decyzji. Natomiast zgodnie z definicją profilowania z art. 4 „musi [ono] obejmować pewną formę zautomatyzowanego przetwarzania – interwencja ludzka nie musi jednak powodować, że dane działanie wykracza poza definicję profilowania”<sup>50</sup>. Znacząca interwencja człowieka w podejmowanie decyzji będzie natomiast sprawiała, że dana sytuacja nie będzie objęta zakazem zautomatyzowanego podejmowania decyzji sformułowanym w art. 22. Grupa Robocza Art. 29 podsumowała zagadnienie relacji między zautomatyzowanym podejmowaniem decyzji i profilowaniem:

Zautomatyzowane podejmowanie decyzji ma inny zakres niż profilowanie, może jednak częściowo pokrywać się z profilowaniem lub z niego wynikać. [...] Zautomatyzowane podejmowanie decyzji może, ale nie musi, obejmować profilowanie, z kolei profilowanie może zachodzić bez podejmowania zautomatyzowanych decyzji. Profilowanie i zautomatyzowane podejmowanie decyzji nie muszą jednak stanowić odrębnych działań. Bywa, że działanie, które początkowo stanowi zwykły proces zautomatyzowanego podejmowania decyzji, może przekształcić się w działanie oparte na profilowaniu, co będzie zależeć od sposobu wykorzystania danych<sup>51</sup>.

Należy zauważyć, że zautomatyzowane podejmowanie decyzji i profilowanie zostało uregulowane – poza RODO – w *dyrektywie 2016/680* oraz w *rozporządzeniu 2018/1725*. *Dyrektywa 2016/680* dotyczy ochrony osób fizycznych w związku z przetwarzaniem danych osobowych przez właściwe organy do celów zapobiegania przestępczości, prowadzenia postępowań przygotowawczych, wykrywania i ścigania czynów zabronionych i wykonywania kar. Przepisy *dyrektywy 2016/680* zawierają rozwiązanie dotyczące zautomatyzowanego podejmowania decyzji nieopierającego się na szczególnych kategoriach danych. Z zasady są one zakazane, z wyjątkiem sytuacji, w której dopuszcza je prawo Unii lub państwa członkowskiego, któremu podlega administrator oraz które przewiduje odpowiednie zabezpieczenia praw i wolności osoby, której dane dotyczą, a przynajmniej prawo do uzyskania interwencji ludzkiej ze strony administratora (art. 11 ust. 1). Definicja profilowania została natomiast ujęta art. 3 ust. 4 *dyrektywy 2016/680* i stanowi

<sup>49</sup> A. Mednis, *Prawo ochrony danych osobowych wobec profilowania osób fizycznych*, PRESSCOM, Warszawa 2019, s. 179.

<sup>50</sup> Grupa Robocza Art. 29, *Wytyczne w sprawie zautomatyzowanego podejmowania decyzji...*, *op. cit.*, s. 7.

<sup>51</sup> *Ibidem*, s. 8–9.

powtórzenie definicji profilowania sformułowanej w RODO. W art. 11 ust. 3 *dyrektywy 2016/680* zabronione zostało profilowanie skutkujące dyskryminacją osób fizycznych na podstawie danych osobowych szczególnych kategorii<sup>52</sup>. *Rozporządzenie 2018/1725* natomiast powtarza przepisy dotyczące zautomatyzowanego podejmowania decyzji pojawiające się w RODO<sup>53</sup>, dotyczy ono jednak przetwarzania danych osobowych przez instytucje, organy i jednostki organizacyjne UE. Z perspektywy ściśle definicyjnej te akty prawne nie tworzą zatem dodatkowych wyzwań interpretacyjnych<sup>54</sup>.

Problemy prawne związane ze sposobem sformułowania zakazu zautomatyzowanego podejmowania decyzji w RODO obejmują bardzo różnorodne kwestie, począwszy od wspomnianej niejasności pojęć użytych w RODO, dla których rozumienia wytyczne Grupy Roboczej Art. 29 stanowią jedynie wskazówkę, nie zaś wiążącą ich wykładnię, przez to jak dopuszczone w treści rozporządzenia wyjątki ograniczają rzeczywisty zakres zastosowania zakazu, po trudności związane z uwzględnieniem kolektywnego wymiaru wpływu zautomatyzowanego podejmowania decyzji. Perspektywa opierająca się na prawie ochrony danych osobowych sprawia, że z dwóch elementów niezbędnych do zautomatyzowanego podejmowania decyzji, czyli danych, na podstawie których decyzja jest podejmowana, oraz algorytmu wykorzystywanego do jej podjęcia, uwaga unijnego ustawodawcy skupia się na danych. Z wyjątkiem „prawa do wyjaśnienia”<sup>55</sup> prawa podmiotu danych dotyczą danych osobowych i dostępu do nich. Natomiast w analizie koncentruję się na algorytmach używanych do podejmowania zautomatyzowanych decyzji oraz prawie dostępu do informacji. Stąd też definicja, którą przedstawiłam na początku wstępu, nie jest identyczna z tym, jak może być rozumiane zautomatyzowane podejmowanie decyzji na tle przepisów RODO.

---

<sup>52</sup> Zgodnie z art. 10 *dyrektywy 2016/680* są to: „dane osobowe ujawniające pochodzenie rasowe lub etniczne, poglądy polityczne, przekonania religijne lub światopoglądowe lub przynależność do związków zawodowych oraz przetwarzanie danych genetycznych, danych biometrycznych w celu jednoznacznego zidentyfikowania osoby fizycznej, danych dotyczących zdrowia lub danych dotyczących seksualności i orientacji seksualnej osoby fizycznej”.

<sup>53</sup> Art. 24 *rozporządzenia 2018/1725* jest odpowiednikiem art. 22 RODO, profilowanie zdefiniowane jest natomiast w art. 3 ust. 5 *rozporządzenia 2018/1725*.

<sup>54</sup> Nieco inaczej kwestia przedstawia się w odniesieniu do materialnej treści przepisów *dyrektywy 2016/680*. Jest ona jednak przedmiotem zainteresowania analizy jedynie w zakresie, w jakim może stanowić wskazówkę interpretacyjną dotyczącą niejasności pojawiających się w związku z wykładnią przepisów RODO (w rozdziale drugim).

<sup>55</sup> Również analizowanym w rozdziale drugim.

## 4.2. Algorytm, kod i program komputerowy

Przez pojęcie algorytmu określa się łańcuch poleceń, których wykonanie ma prowadzić do „osiągnięcia danego celu pod danymi warunkami”<sup>56</sup>. Definicja taka jest neutralna technologicznie – „algorytm nie musi być oprogramowaniem: w szerokim sensie są to zakodowane procedury służące zmianie danych wprowadzanych w pożądaną wynik, w oparciu o sprecyzowane wyliczenia”<sup>57</sup>. Z tej perspektywy przez pojęcie algorytmu można rozumieć zarówno schemat blokowy zapisany na kartce, jak i algorytm determinujący treść facebookowego *newsfeeda*. Taki sposób myślenia o algorytmach jest coraz rzadszy w związku z upowszechnieniem rozumienia algorytmów jako immanentnie związanych z cyfrowością. W dyskursie publicznym algorytmy wiążą się z oprogramowaniem, systemami informacyjnymi, aplikacjami<sup>58</sup>. To semantyczne przesunięcie jest widoczne również w definicjach stworzonych przez organizacje międzynarodowe:

Każdy zestaw zasad dotyczący zmiany cyfrowego wkładu w cyfrowy wynik jest „algorytmem”. Algorytm to oprogramowanie służące do podejmowania decyzji. Jak w przypadku innych typów podejmowania decyzji, jakość decyzji – wynik – zależy od tego, jakiego rodzaju dane zostały wprowadzone do algorytmu oraz od jakości algorytmu<sup>59</sup>.

Definicja ta stanowi przykład dualnej struktury zautomatyzowanego podejmowania decyzji i ilustruje, czym jest w nim algorytm: działanie algorytmu jest zależne od danych przez niego używanych, ale i sam algorytm determinuje sposób, w jaki dane te są używane.

Fakt, że zgodnie z definicją zautomatyzowanego podejmowania decyzji z RODO musi się ono odbywać wyłącznie automatycznie, sprawia, że algorytm używany do takiego podejmowania decyzji musi mieć formę cyfrową. Nie zmienia to jednak faktu, że pozostaje on „zestawem zasad”, jak np. zgodnie z definicją Solona Barocasa i Andrew D. Selbsta, którzy określają algorytm jako „formalnie sprecyzowaną sekwencję operacji logicznych, która

<sup>56</sup> R.K. Hill, *What an Algorithm Is*, „Philosophy & Technology” 2016, nr 29, s. 47. Przegląd literatury dotyczącej definicji algorytmu, zob. M. Kullmann, *Platform Work, Algorithmic Decision-Making, and EU Gender Equality Law*, „International Journal of Comparative Labour Law and Industrial Relations” 2018, t. 34, nr 1, s. 6.

<sup>57</sup> T. Gillespie, *The Relevance of Algorithms*, [w:] T. Gillespie, P. Boczkowski, K. Foot, *Media Technologies, Essays on Communication, Materiality, and Society*, MIT Press, Cambridge–Londyn 2014, s. 167.

<sup>58</sup> Zob. B.D. Mittelstadt, P. Allo, M. Taddeo, S. Wachter, L. Floridi, *The ethics of algorithms: Mapping the debate*, „Big Data & Society” 2016, t. 3, nr 2, s. 2.

<sup>59</sup> OECD (Directorate for Financial and Enterprise Affairs Competition Committee), *Algorithms and Collusion – Note from the European Union*, DAF/COMP/WD (2017) 12, s. 2.

dostarcza komputerowi instrukcji w kolejnych krokach, w celu dokonywania operacji na danych i w związku z tym automatyzacji decyzji”<sup>60</sup>. Cyfrowa forma algorytmów używanych do zautomatyzowanego podejmowania decyzji sprawia z kolei, że zaciera się różnica między tym, jak rozumiany jest algorytm, a tym, czym jest program komputerowy, czego świadectwem jest przykładowa definicja programu komputerowego:

Powszechnie za program komputerowy uważa się zbiór instrukcji (rozkazów, poleceń), wyrażonych w postaci zrozumiałej dla komputera (słowami, symbolami matematycznymi, znakami graficznymi) i mający na celu wywołanie zamierzonych przez twórcę programu działań komputera, prowadzących do osiągnięcia określonego rezultatu<sup>61</sup>.

Bliski związek pojęć algorytmu i programu lub kodu komputerowego w przypadku systemów zautomatyzowanego podejmowania decyzji pozwala traktować je – dla celów analizy<sup>62</sup> – jako do pewnego stopnia tożsame. Zautomatyzowane podejmowanie decyzji zakłada używanie programu komputerowego, czyli wymusza zapis procedur rozwiązywania danego problemu (algorytmu) w formie cyfrowej (kodu komputerowego). Ze względu na to, że algorytmy w ramach systemów zautomatyzowanego podejmowania decyzji muszą być sformułowane w postaci kodu komputerowego, żądanie dostępu do nich jako do informacji publicznej lub dokumentów urzędowych może oznaczać żądanie dostępu do kodu, w postaci którego zostały one ujęte. Nie wyklucza to oczywiście wnioskowania o inne, bardziej abstrakcyjne formy zapisu stosowanych algorytmów (np. dokumenty opisujące je w formie tekstu pisanego językiem naturalnym lub formułami matematycznymi). Jak wyjaśniam w rozdziałach trzecim i czwartym pracy, zaletą podejścia koncentrującego się na zapisie algorytmów w formie kodu komputerowego jest konieczność istnienia zapisu stosowanych w zautomatyzowanym podejmowaniu decyzji algorytmów w formie kodu komputerowego – bez niego nie byłoby bowiem możliwe stosowanie takich rozwiązań. Z tych powodów używam zamiennie pojęć algorytm, kod komputerowy i program komputerowy,

<sup>60</sup> S. Barocas, A.D. Selbst, *Big Data's Disparate Impact*, „California Law Review” 2016, t. 104, s. 674.

<sup>61</sup> F. Rodaniewicz, *Ochrona programów komputerowych w prawie UE (cz. I)*, „Europejski Przegląd Sądowy” 2009, nr 3, s. 24.

<sup>62</sup> Chociaż należy odnotować inne interpretacje relacji między programem komputerowym a algorytmem obecne w literaturze, np.: „uwzględniając poglądy zgłaszane w literaturze przedmiotu, poza zakresem ochrony prawnoautorskiej należałoby w typowej sytuacji pozostawić algorytm, rozumiany jako pewien abstrakcyjny sposób rozwiązania problemu”. *Prawo Internetu*, red. P. Podrecki, Z. Okoń, P. Litwiński, M. Świerczyński, T. Targosz, M. Smycz, D. Kasprzycki, wyd. 2, Lexis Nexis, Warszawa 2007, s. 440. Więcej na ten temat piszę w sekcji poświęconej problemom z przejrzystością algorytmów.

najchętniej posługując się jednak terminem „algorytm”, ponieważ ma on najszersze znaczenie i najbardziej pozwala podkreślić fakt, że mamy do czynienia z zapisem (w wypadku zautomatyzowanego podejmowania decyzji: w formie cyfrowej) pewnych procedur<sup>63</sup>.

Perspektywa, która stoi u podstaw głównej hipotezy, opiera się bowiem na spojrzeniu na algorytm z punktu widzenia neutralności technologicznej. Niezależnie od tego, czy jest on wyrażony w zapisie cyfrowym, czy na papierze, powinien być postrzegany jako zestaw zasad. Klasycznym przykładem przypominającym o tej neutralności technologicznej może być zdjęcie przedstawiające Margaret Hamilton, która stoi obok równie wysokiej jak ona góry papieru. Na papierze znajduje się kod, dzięki któremu możliwe stało się wysłanie człowieka na księżyc<sup>64</sup>. Perspektywa taka – odrywająca algorytmy od postrzegania technologii jako tajemniczej czarnej skrzynki – umożliwia analizę prawną algorytmu jako informacji lub dokumentu, który w określonych przypadkach powinien podlegać prawu dostępu do informacji lub dokumentów. Zważywszy na odnotowany powyżej wpływ zarówno algorytmu, jak i danych na jakość podejmowanej w sposób zautomatyzowany decyzji, stawiana hipoteza, skupiona na samych algorytmach, nie stanowi uniwersalnej i wyczerpującej odpowiedzi na problemy związane z rozwijanymi w tym obszarze technologiami. Jest ona raczej poszukiwaniem uzupełnienia rozwiązań wprowadzonych w RODO. Stawiam w jej centrum zainteresowania drugi, obok danych, ale również niezbędny, element zautomatyzowanego podejmowania decyzji.

### 4.3. Sztuczna inteligencja

Temat analizy jest częściowo związany z zagadnieniem sztucznej inteligencji i jej wpływu na prawo. W książce unikam jednak posługiwania się terminem SI (AI), stąd też istnieje konieczność podania przyczyn stojących za takim wyborem terminologicznym, który mógłby się wydawać konieczny dla przedstawionej analizy. Definicje słownikowe<sup>65</sup> wskazują, że SI jest pojęciem opisującym dwa typy fenomenów. Po pierwsze, odnosi się ono do dziedziny wiedzy i nauki<sup>66</sup>. Po drugie, opisuje ono programy lub systemy komputerowe, które

<sup>63</sup> Za wskazanie konieczności bardziej szczegółowego opisu tego zagadnienia dziękuję dr hab. prof. UŁ Monice Namysłowskiej.

<sup>64</sup> Zdjęcie dostępne w Wikipedii: [https://en.wikipedia.org/wiki/Margaret\\_Hamilton\\_\(scientist\)#/media/File:Margaret\\_Hamilton\\_-\\_restoration.jpg](https://en.wikipedia.org/wiki/Margaret_Hamilton_(scientist)#/media/File:Margaret_Hamilton_-_restoration.jpg), dostęp: 1.04.2021.

<sup>65</sup> *Oxford Dictionary*: [https://en.oxforddictionaries.com/definition/artificial\\_intelligence](https://en.oxforddictionaries.com/definition/artificial_intelligence), dostęp: 1.04.2021; *Słownik języka polskiego*: <https://sjp.pwn.pl/sjp/sztuczna-inteligencja;2466532.html>, dostęp: 1.04.2021.

<sup>66</sup> W definicji podanej w *Słowniku języka polskiego* jest postrzegana ściśle jako „dział informatyki”, co w świetle interdyscyplinarności badań prowadzonych nad sztuczną inteligencją jest dość kontrowersyjnym stwierdzeniem i przykładem traktowania technologii

mają na celu wykonywanie różnego typu zadań „normalnie wymagających inteligencji ludzkiej”<sup>67</sup>, takich jak np. rozpoznawanie mowy, tłumaczenie czy podejmowanie decyzji. Możliwość wykonywania przez SI tych zadań wiąże się z pewnym stopniem autonomii i zdolnością do uczenia się, która charakteryzuje algorytmy wykorzystujące np. uczenie maszynowe. Jednocześnie drugi obszar znaczeniowy tego pojęcia wskazuje na wiele problemów związanych z wyborem precyzyjnej definicji SI: czym jest sama inteligencja?<sup>68</sup> Jaki jest poziom zadań, które wymagają inteligencji ludzkiej? Jak traktować to kryterium w świetle rozwoju technologicznego?

Z perspektywy prawnej pojawia się szersze pytanie: czy istnieje możliwość lub konieczność objęcia SI regulacjami? Dyskusje prowadzone w zakresie tej kwestii już w drugiej połowie XX stulecia<sup>69</sup> koncentrują się z jednej strony na określeniu relacji między sztuczną inteligencją a ciałami fizycznymi, które dzięki niej nabyć mogą pewien stopień autonomii (np. roboty), z drugiej zaś na rozwiązaniach technologicznych wykorzystujących sztuczną inteligencję, które nie wiążą się z przybieraniem przez dane wynalazki formy fizycznej wykazującej się pewną autonomią. Dwoistość ta uwidoczniła została w definicji SI, która pojawia się w komunikacie Komisji Europejskiej „Sztuczna inteligencja dla Europy”:

Termin sztuczna inteligencja odnosi się do systemów, które wykazują inteligentne zachowanie dzięki analizie otoczenia i podejmowaniu działań – do pewnego stopnia autonomicznie – w celu osiągnięcia konkretnych celów. Systemy SI mogą być oparte na oprogramowaniu, działając w świecie wirtualnym (np. asystenci głosowi, oprogramowanie do analizy obrazu, wyszukiwarki, systemy rozpoznawania mowy i twarzy), lub mogą być wbudowane w urządzenia (np. zaawansowane roboty, samochody autonomiczne, drony lub aplikacje internetu rzeczy)<sup>70</sup>.

W związku z brakiem przyjętej definicji legalnej SI to właśnie komunikaty i wnioski legislacyjne Komisji Europejskiej stanowią główne źródło

---

informatycznych jako osobnej rzeczywistości, nie zaś – czym w rzeczywistości one są – elementu współkształtującego również świat fizyczny.

<sup>67</sup> Oxford Dictionary: [https://en.oxforddictionaries.com/definition/artificial\\_intelligence](https://en.oxforddictionaries.com/definition/artificial_intelligence), dostęp: 1.04.2021.

<sup>68</sup> Kompleksowy przegląd definicji „inteligencji”: S. Legg, M. Hutter, *A Collection of Definitions of Intelligence*, [w:] B. Goertzel, P. Wang, *Frontiers in Artificial Intelligence and Applications: Volume 157: Advances in Artificial General Intelligence: Concepts, Architectures and Algorithms*, IOS Press, Amsterdam 2007, s. 17–24.

<sup>69</sup> Zob. S.N. Lehman-Wilzig, *Frankenstein Unbound: Towards a Legal Definition of Artificial Intelligence*, „Futures” 1981, t. 13, nr 6, s. 442–457.

<sup>70</sup> Komisja, *Komunikat Komisji do Parlamentu Europejskiego, Rady Europejskiej, Europejskiego Komitetu Ekonomiczno-Społecznego i Komitetu Regionów: Sztuczna inteligencja dla Europy*, COM (2018) 237 final, s. 1.

informacji dotyczących tego, co przez to pojęcie jest określone przez instytucje unijne<sup>71</sup>. Należy również odnotować prowadzenie prac przez Grupę ekspertów wysokiego szczebla ds. sztucznej inteligencji powołaną przez Komisję Europejską, która w kwietniu 2019 r. opublikowała raport na temat etycznego wymiaru rozwoju SI<sup>72</sup>. Definicja proponowana przez ekspertów z jednej strony przypomina tę sformułowaną przez Komisję przez określenie SI jako systemów charakteryzujących się pewnymi cechami, z drugiej jednak uwzględnia tradycyjne rozumienie SI jako obszaru wiedzy i zbioru rozwiązań programistycznych. Pojawia się w niej również techniczny wymiar rozwoju SI<sup>73</sup>.

Jego obecność w pracach Grupy ekspertów wysokiego szczebla ds. sztucznej inteligencji mogła mieć wpływ na uwzględnienie technicznych aspektów definicji SI w opublikowanej w kwietniu 2021 r. propozycji wniosku legislacyjnego *Aktu w sprawie SI*<sup>74</sup>. Wniosek legislacyjny zawiera propozycję definicji legalnej SI, która opiera się na wymienieniu w Załączniku I technik i podejść, które miałyby być traktowane jako SI, np. mechanizmy uczenia maszynowego, w tym uczenie nadzorowane, uczenie się maszyn bez nadzoru i uczenie przez wzmacnianie, z wykorzystaniem szerokiej gamy metod, w tym uczenia głębokiego<sup>75</sup>. Zgodnie z wnioskiem *Aktu w sprawie SI* przez pojęcie systemu sztucznej inteligencji należy rozumieć oprogramowanie opracowane przy użyciu co najmniej jednej spośród technik i podejść wymienionych w Załączniku I do rozporządzenia, które może:

---

<sup>71</sup> Pierwsze zdanie definicji przedstawionej w komunikacie „Sztuczna inteligencja dla Europy” zostało powtórzone w komunikacie „Skoordynowany plan wobec sztucznej inteligencji” („Coordinated Plan on Artificial Intelligence”): Komisja, *Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions: Coordinated Plan on Artificial Intelligence*, COM (2018) 795 final, s. 1.

<sup>72</sup> Grupa ekspertów wysokiego szczebla ds. sztucznej inteligencji, *Wytyczne w zakresie etyki dotyczące godnej zaufania sztucznej inteligencji*, Bruksela 2019 – oparte na szkicu: The European Commission’s High-Level Expert Group on Artificial Intelligence, *Draft Ethics Guidelines For Trustworthy AI. Working Document for stakeholders’ consultation*, Bruksela 2018.

<sup>73</sup> Zob. Grupa ekspertów wysokiego szczebla ds. sztucznej inteligencji, *Wytyczne w zakresie etyki dotyczące godnej zaufania sztucznej inteligencji*, *op. cit.*, s. 36. Dodatkowo Grupa przygotowała dokument: Grupa ekspertów wysokiego szczebla ds. sztucznej inteligencji, *Definicja SI: główne funkcje i dyscypliny naukowe*, Bruksela 2019.

<sup>74</sup> Komisja, *Wniosek: Rozporządzenie Parlamentu Europejskiego i Rady ustanawiające zharmonizowane przepisy dotyczące Sztucznej Inteligencji (Akt w sprawie Sztucznej Inteligencji) i zmieniające niektóre akty ustawodawcze Unii*, COM/2021/206 final. *Pre-print kompleksowego artykułu omawiającego Akt w sprawie SI*: M. Veale, F. Zuiderveen Borgesius, *Demystifying the Draft EU Artificial Intelligence Act*, SocArXiv 2021. <https://doi.org/10.31235/osf.io/38p5f>, dostęp: 18.07.2021.

<sup>75</sup> Zob. Załącznik I do *Aktu w sprawie SI*.



[...] dla danego zestawu celów określonych przez człowieka – generować wyniki, takie jak treści, przewidywania, zalecenia lub decyzje wpływające na środowiska, z którymi wchodzi w interakcję<sup>76</sup>.

Taka strategia definicyjna stanowi nowe podejście do problemu określenia, czym jest SI w świetle prawa UE. Nie jest jednak oczywiste, że proponowana przez Komisję definicja znajdzie się w ostatecznej formie rozporządzenia. W momencie przygotowywania książki do druku dyskusja nad *Aktem w sprawie SI* dopiero się rozpoczyna, doświadczenia zaś z prac nad RODO podpowiadają, że będzie ona długa i obfitująca w niespodzianki.

Brak konsensu w odniesieniu do znaczenia terminu SI, kontrowersje wokół znaczenia pojedynczych elementów proponowanych definicji, brak wiążących prawnie aktów w prawie europejskim, które odnosiłyby się bezpośrednio do tego zagadnienia, a przede wszystkim brak obowiązującej w czasie przygotowywania tej książki definicji legalnej SI w prawie europejskim sprawiają, że posługiwanie się tym pojęciem może generować niejasności terminologiczne. Próby określenia, czym z perspektywy prawnej jest SI, stanowią osobny przedmiot badań, pozostający poza zakresem zainteresowania tej pracy<sup>77</sup>. Tymczasem zjawisko zautomatyzowanego podejmowania decyzji w rozumieniu zgodnym z RODO jest czymś niezależnym od SI, jednocześnie obejmującym część sytuacji, w których SI znajduje lub może znaleźć zastosowanie. Systemy zautomatyzowanego podejmowania decyzji nie muszą – chociaż mogą – charakteryzować się zdolnością do uczenia się i autonomii. Biorąc pod uwagę zasadę neutralności technologicznej<sup>78</sup> wyrażoną w RODO<sup>79</sup>, fakt używania lub nieużywania technologii określanych terminem SI w zakresie zautomatyzowanego podejmowania decyzji ma

<sup>76</sup> Art. 3 ust. 1, *Akt w sprawie SI*.

<sup>77</sup> Przegląd definicji oraz zagadnień dotyczących podmiotowości prawnej SI, zob. M. Jankowska, *Podmiotowość prawna sztucznej inteligencji?*, [w:] *O czym mówią prawnicy, mówiąc o podmiotowości*, red. A. Bielska-Brodziak, Uniwersytet Śląski, Katowice 2015, s. 171–196; A. Chłopecki, *Sztuczna inteligencja – szkice prawnicze i futurologiczne*, C.H. Beck, Warszawa 2018, s. 1–6.

<sup>78</sup> Zasada neutralności technologicznej zgodnie z definicją M. Hildebrandt: „Prawo neutralne technologicznie to zasadniczo apel normatywny, aby tworzyć prawo, które nie faworyzuje żadnej technologii, bo to szkodzi innowacyjności, lub aby powstrzymać się od przyjmowania prawa, które odnosi się tylko do konkretnych technologii, jako że to czyni prawo mniej możliwym do egzekwowania na dłuższą metę, lub też jako apel, aby brać pod uwagę to, jak dane technologie wpływają na zmniejszenie efektywności danych praw podstawowych, poprzez kompensację tych efektów środkami prawnymi”. M. Hildebrandt, *Smart Technologies and the End(s) of Law. Novel Entanglements of Law and Technology*, Edward Elgar Publishing, Cheltenham 2015, s. 174.

<sup>79</sup> Art. 2 ust. 1 oraz wprost: motywy 15, RODO.

ograniczony wpływ na stosowanie rozporządzenia wobec zautomatyzowanego podejmowania decyzji<sup>80</sup>. Jak już wcześniej podkreśliłam, kryteria definicyjne zautomatyzowanego podejmowania decyzji nie odnoszą się do rodzajów technologii zastosowanych do ich stworzenia. Z tych powodów to przede wszystkim algorytmy i zautomatyzowane podejmowanie decyzji są terminami, którymi posługuję się w analizie.

---

<sup>80</sup> Problem ten omawiam w rozdziale drugim.