

Mateusz Bartoszek
Uniwersytet Warszawski
Wydział Prawa i Administracji

Zastosowanie sztucznej inteligencji w sądownictwie w świetle zasady skutecznej ochrony sądowej

Application of artificial intelligence in judiciary
in view of the principle of effective judicial protection

Streszczenie

Przedmiotem artykułu jest analiza zaproponowanych przez Komisję Europejską przepisów aktu w sprawie sztucznej inteligencji w kontekście potencjalnych zagrożeń naruszenia praw podstawowych, a dekodowanych z zasady skutecznej ochrony sądowej, które wynikać mogą z zastosowania sztucznej inteligencji w sądownictwie. Przytaczane przykłady zastosowania systemów sztucznej inteligencji zaczerpnięte są zarówno z państw członkowskich Unii Europejskiej, jak i spoza UE. Przedstawiono etapy prac nad strategią zmierzającą do stworzenia proponowanego rozporządzenia oraz konkretne zawarte w treści projektu rozwiązania, które mają zapewnić bezpieczeństwo działania systemów sztucznej inteligencji. Analizę projektu zestawiono z wnioskami na temat potencjalnych zagrożeń naruszenia zasady skutecznej ochrony sądowej oraz przedmiotowymi stanowiskami organizacji pozarządowych i międzynarodowych.

Słowa kluczowe

prawa podstawowe, ochrona sądowa, sztuczna inteligencja, akt w sprawie sztucznej inteligencji

Abstract

The subject of the article is the analysis of the draft of Artificial Intelligence Act in the context of potential threats of violation of fundamental rights included in the principle of effective judicial protection that may arise from the use of artificial intelligence in the judiciary. Examples of artificial intelligence systems are based both in EU member states and in states outside of EU. The article presents the stages of work on the strategy for the creation of the draft regulation, as well as the specific solutions included in the draft, which aim to ensure the safety of use of artificial intelligence systems. The analysis of the draft is juxtaposed with conclusions on the potential threats to the principle of effective judicial protection and the opinions of non-governmental and international organizations.

Keywords

fundamental rights, judicial protection, artificial intelligence, Artificial Intelligence Act

1. Wstęp

Sztuczna inteligencja (SI) to nowe zagadnienie, z którym mierzy się ustawodawstwo każdego systemu prawnego: międzynarodowego i państwowego, a także unijnego. Kilka lat temu zauważono zainteresowanie, jakie sztuczna inteligencja wzbudziła w środowisku naukowym, a także niezliczoną ilość możliwych zastosowań tej technologii. Należy wspomnieć, że wśród sektorów, w których sztuczna inteligencja może potencjalnie odgrywać dużą rolę, są m.in.: transport (pojazdy autonomiczne), medycyna (wykorzystanie inteligentnych maszyn do wspomagania pracy lekarzy), rozrywka (systemy sztucznej inteligencji są w stanie pokonać nawet największych mistrzów w grach strategicznych) oraz przemysł zbrojeniowy (systemy automatycznego namierzania, drony autonomiczne itp.)¹.

Również w sferze publicznej coraz więcej jest przykładów zastosowania systemów sztucznej inteligencji. We Włoszech, w regionie Emilia-Romagna, działa system pod nazwą RiskER, którego zadaniem jest przewidywać ryzyko hospitalizacji². Zgodnie z raportami z jego działania pozwolił on skutecznie zapobiegać hospitalizacji poprzez wcześniejsze nawiązanie kontaktu i udzielenie porad osobom zaliczonym przez RiskER do najwyższej kategorii ryzyka³. Postępująca cyfryzacja i wysiłki prowadzące do przyspieszenia i ułatwienia postępowań sądowych przyczyniły się do wprowadzenia systemów zautomatyzowanego podejmowania decyzji i systemów sztucznej inteligencji także w sądach. Wskazać można między innymi system COMPAS, wprowadzony w sądach stanowych m.in. Nowego Jorku, Wisconsin i Kalifornii, a służący do oceny ryzyka recydywy oskarżonych w postępowaniach karnych⁴.

Kompetencja Unii Europejskiej do podejmowania działań zmierzających do ustanowienia prawa regulującego wprowadzanie do obrotu i stosowanie systemów sztucznej inteligencji wynika wprost z przepisu art. 114 Traktatu o funkcjonowaniu Unii Europejskiej (dalej TFUE). Na jego mocy Parlament Europejski i Rada mają kompetencję do przyjęcia środków dotyczących zbliżenia przepisów państw członkowskich, które mają na celu usta-

¹ Szerzej na temat zagrożeń związanych z pojazdami autonomicznymi: G. Dede, R. Hamon, H. Junklewitz, R. Naydenov, A. Malatras, M. Sanchez, *Cybersecurity challenges in the uptake of Artificial Intelligence in Autonomous Driving*, Publications Office of the European Union, Luxembourg 2021. Szerzej na temat systemów sztucznej inteligencji wyspecjalizowanych do grania w gry strategiczne: C. Metz, *The Sadness and Beauty of Watching Google's AI Play Go*, "Wired", 3 listopada 2016 r., <https://www.wired.com/2016/03/sadness-beauty-watching-googles-ai-play-go/> [dostęp: 20.11.2021].

² Zob. AlgorithmWatch, *Automating Society Report 2019*, 1 stycznia 2019 r., s. 89, <https://algorithmwatch.org/en/automating-society/> [dostęp: 20.11.2021].

³ *Ibidem*, s. 89.

⁴ Szerzej zob.: F. Rahman, *COMPAS Case Study: Fairness of a Machine Learning Model*, "Towards Data Science", 7 września 2020 r., <https://towardsdatascience.com/compas-case-study-fairness-of-a-machine-learning-model-f0f804108751> [dostęp: 20.11.2021].

nowienie i funkcjonowanie rynku wewnętrznego⁵. Powszechny jednolity rynek cyfrowy, którego elementem są rozwiązania dotyczące cyfryzacji (w tym sztucznej inteligencji), jest właśnie częścią rynku wewnętrznego. Regulacja rynku wewnętrznego, w myśl art. 4 TFUE, należy do kompetencji dzielonych pomiędzy UE i państwa członkowskie (art. 4 pkt 2 lit. a). Sztuczna inteligencja stwarza ogromne możliwości dla usprawniania lub upraszczania wielu obszarów działalności ludzkiej, co w pełni uzasadnia przyjęcie środków prawnych zmierzających do jej regulacji lub przynajmniej harmonizacji przepisów na poziomie unijnym. W przeciwnym wypadku brak koordynacji ustawodawstw państw członkowskich skutkowałaby brakiem spójności przepisów obowiązujących w państwach członkowskich, spadkiem pewności prawa i mocno utrudnionym obrotem transgranicznym pomiędzy państwami członkowskimi UE.

Wprowadzanie systemów niezależnych w swoim działaniu od człowieka do tak kluczowych dla demokracji i praworządności instytucji jak sądy wiąże się z wieloma zagrożeniami nie tylko dla zakazu dyskryminacji, ale także dla niezawisłości sędziowskiej, prawa dostępu do sądu i prawa do rzetelnego procesu. W związku z tym przyjąłem zasadę skutecznej ochrony sądowej, obowiązującą w prawie Unii Europejskiej, a także na mocy Konwencji o ochronie praw człowieka i podstawowych wolności jako wzorzec, którego potencjalne naruszenia i ochronę należy rozważyć.

Istotą zasady skutecznej ochrony sądowej jest gwarancja prawa do skutecznego środka prawnego przed sądem dla każdego, kto uważa, że jego prawa zostały naruszone w jakikolwiek sposób. Dookoła tego rdzenia jest jednak wiele dodatkowych elementów, które mają albo wzmacniać prawo jednostki, albo też nakładać obowiązki na instytucje wymiaru sprawiedliwości⁶. Wśród tych elementów wymienić należy: prawo do sprawiedliwego rozpatrzenia sprawy, prawo do rozpatrzenia sprawy w rozsądnym terminie, zasadę niezawisłości i bezstronności sędziowskiej i wymóg ustanowienia sądu na mocy ustawy.

Niniejszy artykuł ma dwa cele. Pierwszym jest przedstawienie zagrożeń naruszenia praw podstawowych wchodzących w skład zasady skutecznej ochrony sądowej, które wynikają z zastosowania systemów sztucznej inteligencji w sądownictwie. Drugim jest odpowiedź na pytanie, czy proponowane przez Komisję Europejską przepisy rozporządzenia Parlamentu Europejskiego i Rady ustanawiającego zharmonizowane przepisy dotyczące sztucznej inteligencji (akt w sprawie sztucznej inteligencji) i zmieniające nie-

⁵ Zob. art. 114, pkt 1, *Traktat o funkcjonowaniu Unii Europejskiej*, Dz. Urz. UE C 326/47 z 26.10.2012 r.

⁶ Źródło, z którego wywodzi się zasada skutecznej ochrony sądowej, stanowią wspólne tradycje konstytucyjne państw członkowskich, a potwierdzeniem tej zasady są przepisy art. 6 i art. 13 EKPC i art. 47 KPP. Zob. wyrok Trybunału z dnia 13 marca 2007 r. w sprawie C-432/05 *Unibet (London) Ltd i Unibet (International) Ltd przeciwko Justitiekanslern*, ECLI:EU:C:2007:163.

które akty ustawodawcze Unii (Artificial Intelligence Act) mogą służyć skutecznemu zapobieganiu tymże potencjalnym zagrożeniom⁷.

2. Przykłady wykorzystania sztucznej inteligencji w sądownictwie i wynikające z tego zagrożenia

Przedstawienie dokładnej definicji sztucznej inteligencji do dziś jest niemalże niemożliwe. Sztuczna inteligencja traktowana jest jako osobna gałąź nauki, a zamknięcie samego pojęcia w ramach definicji niweczone jest kolejnymi postępami w rozwoju technologicznym⁸. W motywach projektu AIA Komisja Europejska opisała sztuczną inteligencję jako „szybko rozwijającą się grupę technologii, które muszą przyczyniać się do wielu korzyści społeczno-ekonomicznych we wszystkich gałęziach przemysłu i obszarach działalności gospodarczej”⁹. W treści proponowanych przepisów AIA zdefiniowano pojęcie systemu sztucznej inteligencji: „oprogramowanie opracowane przy użyciu co najmniej jednej z technik i podejść wymienionych w załączniku I, które może – dla danego zestawu celów określonych przez człowieka – generować wyniki, takie jak treści, przewidywania, zalecenia lub decyzje wpływające na środowiska, z którymi wchodzi w interakcję”¹⁰. W wyliczeniu technik i podejść objętym załącznikiem I zawarto m.in. technologie uczenia maszynowego, rozumowania symbolicznego i różnych metod wyszukiwania, które stanowią cechy charakterystyczne systemów sztucznej inteligencji.

Potencjalne zastosowania systemów sztucznej inteligencji w sądownictwie są niezliczone. Wszędzie tam, gdzie obecnie używane są systemy teleinformatyczne, możliwe jest wdrożenie systemów sztucznej inteligencji jako „ulepszenia”. Usprawnienie funkcjonowania sądów jest prawdopodobnie najczęściej używanym argumentem mającym przemawiać za wprowadzaniem sztucznej inteligencji do użytku. W teorii systemy sztucznej inteligencji mogą przejąć zadania na niemal każdym poziomie funkcjonowania sądu: od przydzielania spraw sędziom, poprzez ocenę kompletności pisma procesowego

⁷ Zob. Wniosek Komisji Europejskiej, *Rozporządzenie Parlamentu Europejskiego i Rady ustanawiające zharmonizowane przepisy dotyczące sztucznej inteligencji (Akt w sprawie sztucznej inteligencji) i zmieniające niektóre akty ustawodawcze Unii*, 21 kwietnia 2021 r., COM/2021/206 final, <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021PC0206> [dostęp: 20.11.2021], dalej AIA.

⁸ Aktualność zachowuje definicja stworzona przez J. McCarthy’ego, jednego z ojców sztucznej inteligencji, w 1956 r.: „nauka i inżynieria tworzenia inteligentnych maszyn” (zacytowana ze strony internetowej: <https://www.sztuczna-inteligencja.org.pl/definicja/sztuczna-inteligencja/> [dostęp: 20.11.2021]). Dwaj z najwybitniejszych naukowców zajmujących się współcześnie sztuczną inteligencją, P. Norvig i S. Russel, opracowali definicję bardziej złożoną, ale nadal prostą do zrozumienia: „sztuczna inteligencja to maszyny imitujące funkcje kognitywne charakterystyczne dla ludzkiego mózgu, takie jak np.: uczenie się, rozwiązywanie problemów, analizowanie języka naturalnego, dokonywanie obliczeń”. S. Russel, P. Norvig, *Artificial Intelligence: a modern approach*, Pearson Education Inc., Berkeley 1994.

⁹ Zob. AIA, s. 22, motyw 3.

¹⁰ Zob. Załącznik nr I do AIA.

wego pod względem formalnym, aż po analizę materiału dowodowego. Spośród przykładów zastosowania systemów sztucznej inteligencji w sądownictwie przedstawiam dwie kategorie systemów istotnie wpływających na proces orzekania, tj.: systemy analizujące dane i wspierające pracę sędziego oraz systemy mogące zastąpić sędziego.

Przykładem pierwszego ze wspomnianych modeli jest funkcjonujący w Stanach Zjednoczonych system mający służyć ocenie ryzyka ponownego złamania prawa przez oskarżonego w procesie karnym. Zaprojektowany i stworzony przez Northpoint Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) pozwala ocenić ryzyko recydywy, przez co wspomaga podejmowanie przez sędziego decyzji, czy wydać wyrok w zawieszeniu. COMPAS ocenia ryzyko w trzech kategoriach: ryzyko recydywy przed zakończeniem procesu, ogólne ryzyko recydywy, ryzyko recydywy z użyciem przemocy. Dane, których używa COMPAS pochodzą częściowo z kartoteki oskarżonego, a częściowo z formularza uzupełnianego w trakcie wywiadu z oskarżonym¹¹. Formularz zawiera 137 pytań, pośród których są np. pytania o historię kryminalną członków rodziny, o uczestnictwo w bójkach w szkole i o liczbę przestępstw popełnianych w sąsiedztwie domu oskarżonego¹². Kontrowersje dotyczyły między innymi postulatów dotyczących powielania przez COMPAS uprzedzeń, a nawet dyskryminacji¹³.

Eric Loomis, jeden ze skazanych w sprawach, w których wyrokowanie było wspierane oceną ryzyka przeprowadzoną przez program COMPAS, po nieudanej apelacji złożył skargę kasacyjną do Sądu Najwyższego Stanu Wisconsin, w której zarzucił m.in. naruszenie jego prawa do rzetelnego procesu, ponieważ sędzia w wydawaniu wyroku kierował się wskazówkami programu, którego działanie jest objęte tajemnicą handlową. Sąd Najwyższy Stanu Wisconsin oddalił skargę, uznając, że sądy mogą wspomagać się oceną ryzyka uzyskaną dzięki programom komputerowym, ale nie mogą polegać na takiej ocenie. Konieczne jest uzasadnienie wyroku argumentami innymi niż tylko ocena ryzyka ponownego popełnienia przestępstwa przez oskarżonego¹⁴. Ponadto Sąd Naj-

¹¹ Zob. wyrok Sądu Najwyższego Stanu Wisconsin z 13 lipca 2016 r. w sprawie *State przeciwko Loomis*, sygn. 881 N.W.2d 749 (2016), pkt 13 i 14.

¹² J. Angwin, J. Larson, S. Mattu, L. Kirchner, *Machine Bias There's software used across the country to predict future criminals. And it's biased against blacks*, ProPublica, 23 maja 2016 r., <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> [dostęp: 20.11.2021]. Krytyki tego artykułu podjęli się A. Flores, C. Lowenkamp, K. Bechtel w *False Positives, False Negatives, and False Analyses: A Rejoinder to "Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And it's Biased Against Blacks"*, "Federal Probation Journal" 2016, t. 80, nr 2, ale artykuł wydany w ProPublica zawiera skany formularza z pytaniami używanego do oceny ryzyka przez program COMPAS. Skany te dostępne są pod adresem internetowym <https://www.documentcloud.org/documents/2702103-Sample-Risk-Assessment-COMPAS-CORE.html> [dostęp: 20.11.2021].

¹³ Jedną z najszerzej dyskutowanych publikacji dotyczących tego problemu była: J. Angwin, J. Larson, S. Mattu, L. Kirchner, *op. cit.*

¹⁴ Zob. wyrok Sądu Najwyższego Stanu Wisconsin z 13 lipca 2016 r. w sprawie *State przeciwko Loomis*, sygn. 881 N.W.2d 749 (2016) pkt 123-125.

wyższy Stanu Wisconsin zaznaczył, że wyniki ocen ryzyka według programu COMPAS oparte są na danych grupowych, przez co nie są całkowicie miarodajne dla pojedynczych osób¹⁵. Fakt, że COMPAS jest prywatną własnością, co uniemożliwia publiczne ogłoszenie algorytmu programu i ujawnienia oskarżonym lub nawet sądowi, w jaki sposób program oblicza ryzyko i jak dokładnie dane udostępnione programowi są ważne, został uznany za istotny i wskazano, że sądy wspomagające się podczas wyrokowania oceną ryzyka dokonaną przez COMPAS muszą informować o tym oskarżonych¹⁶.

Francuskie sądy apelacyjne w Rennes i Douai w 2017 r. brały udział w pilotażowym programie testowym systemu Predictice¹⁷. Program ten miał być pierwszym krokiem do cyfryzacji sądownictwa we Francji. Jeden z twórców Predictice, Louis Larret-Chahine, zadeklarował, że „położy koniec nieprzewidywalnemu, przypadkowemu i zróżnicowanemu nawet w obrębie jednego kraju wymiarowi sprawiedliwości i ruszy w kierunku czegoś bardziej logicznego, naukowego, lub przynajmniej bardziej możliwego do kontrolowania”¹⁸. Predictice był używany wyłącznie w celu obliczania wysokości odpraw zasądzanych w sprawach pracowników zwolnionych bez „rzeczywistej i poważnej przyczyny”. Docelowo miało to doprowadzić do redukcji różnic w wysokości zasądzanych świadczeń i ujednolicić kryteria przyjmowane jako istotne w poszczególnych sprawach, przyjmując za źródło takiego celu zasadę równości obywateli wobec prawa¹⁹. Wynik tego testu był jednak negatywny. Systemowi udostępniono w celu nauki tylko fragmenty orzeczeń, czego efektem był brak umiejętności analizy wszystkich informacji dostępnych w bieżących sprawach i niezadowalające efekty działania programu. Francuskie Ministerstwo Sprawiedliwości w komunikacie na temat rezultatów programu pilotażowego mającego testować działanie Predictice oświadczyło, że program „wymaga znaczących popraw [...] i obecnie jego użycie nie jest żadną korzyścią dla sędziów”²⁰.

Do drugiej kategorii zaliczyć trzeba pierwszy w Europie w pełni zautomatyzowany system rozstrzygania sporów. Stosowany w Królestwie Niderlandów od 2010 r. system e-Court jest używany w rozstrzyganiu spraw windykacyjnych. Niderlandzki system wymiaru sprawiedliwości oprócz sądów publicznych obejmuje również sądy prywatne, będące sądami arbitrażowymi. Sąd e-Court jest właśnie takim sądem prywatnym. W całym

¹⁵ *Ibidem*, pkt 74.

¹⁶ *Ibidem*, pkt 66.

¹⁷ R. Letteron, *Digital access to the law*, “Digital Issues of Les Annales des Mines” 2018, nr 3.

¹⁸ *Predictive justice: when algorithms pervade the law*, “Paris Innovation Review” 2017, nr 9, <http://parisinnovationreview.com/articles-en/predictive-justice-when-algorithms-pervade-the-law> [dostęp: 20.11.2021].

¹⁹ Europejska Komisja na rzecz skutecznego wymiaru sprawiedliwości, *European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment*, przyjęta w Strasbourgu 4 grudnia 2018 r., s. 42.

²⁰ R. Letteron, *op. cit.*

postępowaniu interwencja człowieka jest wymagana tylko na ostatnim etapie – wydania tytułu egzekucyjnego²¹.

Całość postępowania odbywa się online, strony zamieszczają pisma procesowe na specjalnie do tego przeznaczonej platformie, gdzie są one dostępne dla obu stron przez cały czas trwania postępowania. Rozstrzygnięcie sprawy jest wydawane przez system sztucznej inteligencji na podstawie obowiązujących przepisów oraz imitacji rozumowania sędziowskiego prowadzonej na podstawie przeszłych spraw, których akta zostały udostępnione systemowi na początku jego funkcjonowania. Niderlandzkie prawo krajowe nadal nie zawiera regulacji dotyczących sędziego-sztucznej inteligencji, wobec czego zgodnie ze zmianami w prawie dotyczącym sądów prywatnych (arbitrażowych), które weszły w życie 1 stycznia 2015 r., system wydaje rozstrzygnięcie w imieniu sędziego. Zadaniem sędziego jest natomiast kontrola rozstrzygnięć wydawanych przez system. Do końca 2015 r. nie wystąpiła sytuacja, w której sędzia musiałby poprawiać lub zmieniać orzeczenie wydane przez e-Court²².

W 2019 r. powstał w Estonii projekt mający na celu stworzenie sztucznej inteligencji, która zastąpi sędziów w sprawach cywilnych o drobne roszczenia, nieprzekraczające 7000 euro. Warto zauważyć, że nawet Estonia, kraj będący w czołówce cyfryzacji administracji publicznej i wymiaru sprawiedliwości, który jako jeden z pierwszych krajów na świecie wprowadził wybory powszechne przez Internet²³, nie podjęła kroków w celu całkowitego zastąpienia sędziów systemami sztucznej inteligencji²⁴. Sztuczna inteligencja w Estonii, podobnie jak e-Court w Holandii, ma rozstrzygać jedynie sprawy o drobne roszczenia. Nasuwa się wniosek, że stan prac nad rozwojem sztucznej inteligencji nie jest jeszcze wystarczająco zaawansowany, żeby móc zastąpić sędziego. Podobnie uważają J. Morison i A. Harkens, którzy w artykule *Re-engineering justice? Robot judges, computerised courts and (semi) automated legal decision-making*²⁵ stwierdzili: „Obecnie jednak, algorytmy pozostają inteligentnymi doradcami i narzędziami. Gdyby algorytmy wykorzystujące uczenie maszynowe [...] przejęły rolę tworzenia i egzekwowania prawa bez nadzoru, nie ma gwarancji, że kontynuowałyby te same wzory przewidywań. W takiej sytuacji, rozwoju nowych sposobów rozumowania i argumentowania, opartych

²¹ H. Nakad-Weststrate, T. Jongbloed, H. van der Herik, A. Salem, *Digitally Produced Judgements in Modern Court Proceedings*, „International Journal of Digital Society” 2015, nr 4, s. 1102.

²² *Ibidem*, s. 1006-1008.

²³ <https://e-estonia.com/solutions/e-governance/i-voting/> [dostęp 15.07.2021].

²⁴ Oficjalnym argumentem wykorzystywanym przez rząd do promowania tego projektu jest zbyt duża liczba spraw sądowych i zbyt mała liczba sędziów. Statystyki dostępne na oficjalnej stronie [e-Estonia.com](https://e-estonia.com) wskazują, że liczba sędziów w Estonii jest obecnie taka sama jak 20 lat temu, natomiast liczba spraw sądowych podwoiła się. Skuteczność i prędkość działania sądów jest więc według władz Estonii najważniejszym argumentem przemawiającym za wdrożeniem programu sędziego-sztucznej inteligencji.

²⁵ J. Morison, A. Harkens, *Re-engineering justice? Robot judges, computerised courts and (semi) automated legal decision-making*, Queen’s University Belfast, „Legal Studies” 2019, t. 39, nr 4, s. 618-635.

o rozpoznawanie wzorców i język niezrozumiały dla podmiotów prawa, prawo mogłoby stawać się coraz bardziej niepewne i podważalne”²⁶.

Najbardziej zaawansowane w kwestii imitacji rozumowania sędziowskiego są projekty naukowe dotyczące przewidywania przez systemy sztucznej inteligencji rozstrzygnięć sądów. Na szczególną uwagę zasługuje tutaj projekt przewidywania orzeczeń ETPC prowadzony w 2016 r.²⁷ Model sztucznej inteligencji skonstruowany w ramach tego projektu używał uczenia maszynowego i przetwarzania języka naturalnego w celu analizy akt spraw wnoszonych przed Trybunał, procedury stosowanej przez Trybunał i właściwego prawa. Na podstawie tych informacji system przewidywał rozstrzygnięcie. Analizie poddano akta 584 spraw dotyczących potencjalnych naruszeń trzech różnych przepisów Europejskiej Konwencji Praw Człowieka. Rezultatem projektu było osiągnięcie średnio 79% skuteczności w przewidywaniu rozstrzygnięć ETPC (największą, 84% skuteczność system osiągnął w przewidywaniu rozstrzygnięć dotyczących art. 6 EKPC).

Wskazane przykłady pozwalają wyznaczyć kilka podstawowych problemów związanych z wykorzystaniem sztucznej inteligencji w sądownictwie i potencjalnych zagrożeń dla praw podstawowych tworzących zasadę skutecznej ochrony sądowej. Problemy to: brak transparentności systemów SI działających w sądownictwie („efekt czarnej skrzynki”); całkowita zależność wyników pracy systemów od danych, które są mu udostępniane; bezpośredni wpływ twórców systemów SI na ich działanie; brak możliwości zrozumienia działania systemów SI przez sędziów oraz brak możliwości uzyskania pełnej kontroli nad systemami przez sądy. Z powodu tych problemów zagrożona jest m.in. zasada niezawisłości sędziowskiej, ponieważ system SI stanowi element *de facto* zewnętrzny w stosunku do sędziego. Jeżeli system jest pod faktyczną kontrolą podmiotu prywatnego lub nawet podmiotu publicznego, zewnętrznego w stosunku do sądu, jak zagwarantować, że wpływ systemu na proces orzekania nie będzie naruszał niezawisłości sędziego²⁸. Również zasada sprawiedliwego rozpatrzenia sprawy może zostać naruszona poprzez wykorzystanie systemów SI w sądach. Przykładem takiego naruszenia może być brak umiejętności zrozumienia i wyjaśnienia przez sędziego pracy systemu, co uniemożliwia realizację prawa do uzyskania uzasadnienia wyroku²⁹. Wreszcie, jak

²⁶ *Ibidem*, s. 627.

²⁷ N. Aletras, D. Tsarapatsanis, D. Preotiuc-Pietro, V. Lampos, *Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective*, “PeerJ Computer Science” 2016, t. 93.

²⁸ Zob. wyrok ETPC z dnia 22 października 1984 r. w sprawie *Sramek przeciwko Austrii*, skarga nr 8790/79, pkt 42. Por. także wyrok ETPC z dnia 28 czerwca 1984 r. w sprawie *Campbell i Fell przeciwko Zjednoczonemu Królestwu*, skargi nr 7819/77 i 7878/77, pkt 81.

²⁹ Zob. opinia nr 11 Rady Konsultacyjnej Sędziów Europejskich (CCJE) w sprawie jakości orzecznictwa, 18 grudnia 2008 r., pkt 34-35; także wyrok ETPC z dnia 30 listopada 1987 r. w sprawie *H. przeciwko Belgii*, skarga nr 8950/80, pkt 53; także wyrok Trybunału z dnia 6 września 2012 r. w sprawie *C-619/10 Trade Agency Ltd przeciwko Seramico Investments Ltd*, ECLI:EU:C:2012:531, pkt 53.

zagwarantować rzetelne rozpatrzenie sprawy w sytuacji, gdy analiza stanu faktycznego należy w całości do systemu SI, a nie do człowieka³⁰.

3. Na drodze do stworzenia unijnej regulacji sztucznej inteligencji

Wstępem do prac nad AIA były *Wytyczne w zakresie etyki dotyczące godnej zaufania sztucznej inteligencji* przygotowane przez grupę ekspertów wysokiego szczebla do spraw sztucznej inteligencji (*Wytyczne*), opublikowane 8 kwietnia 2019 r.³¹ Dokument ten określa ramy dla godnej zaufania sztucznej inteligencji, jednocześnie wskazując kierunki dalszych prac zmierzających ostatecznie do stworzenia regulacji prawnych. Podstawą tego podejścia jest pogląd, że zaufanie jest fundamentem budowania społeczeństwa i systemów ekonomicznych, a więc pełny potencjał wykorzystania sztucznej inteligencji może być osiągnięty dopiero przy pełnym zaufaniu do systemów SI na wszystkich etapach ich cyklu życia: tworzenia, wdrażania i używania³². *Wytyczne* zawierają zbiór zasad etycznych, a także postulaty dotyczące technicznej solidności systemów sztucznej inteligencji, które muszą być spełnione, aby systemy SI mogły zostać uznane za godne zaufania³³. Wizję grupy ekspertów najlepiej streszcza wyrażona w podsumowaniu *Wytycznych* nadzieja na „stworzenie kultury «godnej zaufania sztucznej inteligencji dla Europy», dzięki której wszyscy będą mogli czerpać korzyści z SI w sposób zapewniający poszanowanie naszych podstawowych wartości: praw podstawowych, demokracji i praworządności”³⁴.

Jak podkreśla M. Safjan, prawo do sądu, a więc także zasada skutecznej ochrony sądowej, której prawo do sądu jest nieodłącznym elementem, ma kluczowe znaczenie nie tylko dla skutecznej ochrony wszystkich innych praw podstawowych, ale także dla stosowania całości prawa Unii Europejskiej jako narzędzie urzeczywistniające wartości i zasady wynikające z art. 2 Traktatu o Unii Europejskiej³⁵. Nasuwa się zatem wniosek, że w celu zrealizowania wizji grupy ekspertów zawartej w podsumowaniu *Wytycznych* konieczne jest, aby systemy sztucznej inteligencji wykorzystywane w procesie ochrony praw podstawowych, a więc m.in. w sądownictwie, zapewniały zgodność z zasadą sku-

³⁰ Zob. P. Hofmański, A. Wróbel, [w:] Garlicki L. (red.), *Konwencja o Ochronie Praw Człowieka i Podstawowych Wolności. Tom I. Komentarz do artykułów 1-18*, Warszawa 2010, komentarz do art. 6.

³¹ Zob. niezależna grupa ekspertów wysokiego szczebla do spraw sztucznej inteligencji powołana przez Komisję Europejską w czerwcu 2018 r., *Wytyczne w zakresie etyki dotyczące godnej zaufania sztucznej inteligencji*, 8 kwietnia 2019 r., <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai> [dostęp: 20.11.2021].

³² Szerzej na temat idei zaufania do sztucznej inteligencji, zob. S. Thiebes, S. Lins, A. Sunyaev, *Trustworthy artificial intelligence*, „Electronic Markets” 2021, t. 31, nr 2, s. 447-464.

³³ Zob. *Wytyczne w zakresie etyki dotyczące godnej zaufania sztucznej inteligencji...*, s. 8-9.

³⁴ Zob. *ibidem*, s. 47.

³⁵ Zob. M. Safjan, *Prawo do skutecznej ochrony sądowej – refleksje dotyczące wyroku TSUE z 19.11.2019 r. w sprawach połączonych C-585/18, C-624/18, C 625/18*, „Palestra” 2020, nr 5, s. 5-29.

tecznej ochrony sądowej. W przeciwnym wypadku ochrona praw podstawowych może być utrudniona lub nawet niemożliwa.

Po *Wytycznych* następnym krokiem w strategii Komisji Europejskiej na temat rozwoju i regulacji sztucznej inteligencji jest *Biała Księga w sprawie sztucznej inteligencji. Europejskie podejście do doskonałości i zaufania (Biała Księga)*, opublikowana 19 lutego 2020 r.³⁶ Komisja w ślad za grupą ekspertów ponownie zwraca w tym dokumencie szczególną uwagę na zaufanie społeczeństwa do systemów sztucznej inteligencji, wymieniając brak zaufania, obok braku inwestycji i umiejętności, jako jeden z głównych czynników powstrzymujących upowszechnianie sztucznej inteligencji³⁷.

W *Białej Księdze* Komisja po raz pierwszy porusza także temat praw podstawowych, dla ochrony których potencjalnie niebezpieczne jest wykorzystanie sztucznej inteligencji. Komisja wskazuje na prawo do rzetelnego procesu, czyli element zasady skutecznej ochrony sądowej, podkreślając jednocześnie, że w związku z rozwojem technologicznym i poszerzającym się wykorzystaniem systemów sztucznej inteligencji coraz częściej będą miały miejsce sytuacje, w których zarówno osoby fizyczne, jak i prawne będą podlegać działaniom takich systemów. Wynika z tego coraz poważniejsza potrzeba utworzenia ram legislacyjnych, które zapewnią m.in. skuteczną ochronę i egzekwowanie prawa³⁸.

Biała Księga wskazuje na niezwykle istotny zwrot w podejściu Komisji do problematyki regulacji prawnych dotyczących sztucznej inteligencji. Zasady etyczne i postulat godnej zaufania sztucznej inteligencji stanowią dla Komisji drogowskazy, którymi należy się kierować podczas tworzenia regulacji prawnych, ale jednocześnie pojawia się idea oparcia proponowanych przepisów na analizie ryzyka. Komisja wskazuje, że wykorzystanie analizy ryzyka pozwoli na stworzenie regulacji prawnych proporcjonalnych do zagrożeń stwarzanych przez różne sposoby wykorzystania systemów sztucznej inteligencji. Projektowana kategoria sztucznej inteligencji wysokiego ryzyka oparta jest na dwóch kryteriach.

Pierwszym kryterium jest wykorzystanie systemu sztucznej inteligencji w sektorach, w których zazwyczaj podejmowane działania charakteryzują się możliwością wystąpienia znaczącego ryzyka. Komisja postuluje stworzenie wykazu takich sektorów i okresowe poddawanie go przeglądowi i aktualizacji w celu dostosowania do szybkiego tempa rozwoju technologicznego³⁹.

³⁶ Zob. Komisja Europejska, *Biała Księga w sprawie sztucznej inteligencji, Europejskie podejście do doskonałości i zaufania*, Bruksela, 19 lutego 2020 r., COM/2020/65.

³⁷ *Ibidem*, s. 10.

³⁸ *Ibidem*, s. 13.

³⁹ W treści *Białej Księgi* Komisja podaje przykłady sektorów spełniających pierwsze kryterium: opieka zdrowotna, transport, energetyka, polityka azylowa, sądownictwo, zabezpieczenie społeczne.

Drugie kryterium odnosi się do konkretnego zastosowania systemów sztucznej inteligencji w ramach tych sektorów. Nie wszystkie sposoby wykorzystania SI wiążą się bowiem z wysokim ryzykiem właściwym dla sektora. Treść *Białej Księgi* wskazuje między innymi na zastosowania SI, wywierające skutki prawne lub podobnie istotne skutki w odniesieniu do praw jednostki lub przedsiębiorstwa⁴⁰. Systemy SI, których działanie nie ma wystarczająco doniosłego wpływu na osoby dotknięte tym działaniem, nie mogą zostać uznane za systemy wysokiego ryzyka.

W ramach konsultacji społecznych, którym poddana została *Biała Księga*, stanowiło wyrażała Rada Adwokatur i Stowarzyszeń Prawniczych Europy (CCBE). CCBE zwróciła uwagę na problem zastosowania systemów sztucznej inteligencji w sądownictwie, wskazując na szereg problemów i zagrożeń oraz zaproponowała rozwiązania mające na celu przeciwdziałanie tym zagrożeniom⁴¹.

CCBE postuluje między innymi ustanowienie obowiązku informowania stron i uczestników postępowania o każdym użyciu systemu sztucznej inteligencji mającym wpływ na proces rozstrzygania sprawy. W kontekście zgodności z zasadą skutecznej ochrony sądowej taka regulacja prowadzić może do znacznego zwiększenia przejrzystości zastosowań systemów SI w sądownictwie i przez to zapobiegać potencjalnym naruszeniom prawa do uzyskania uzasadnienia orzeczenia i prawa do sprawiedliwego rozpatrzenia sprawy. W porównaniu do postulatów wypracowanych przez grupę ekspertów i Komisję rozwiązania proponowane przez CCBE są konkretne i odpowiadają na realne zagrożenia charakterystyczne dla sądownictwa.

Również Rada Europy pracuje nad strategią utworzenia ram prawnych regulujących sztuczną inteligencję. Prace te, w przeciwieństwie do dokumentów publikowanych przez organy UE, nie dotyczą jednak sztucznej inteligencji ogólnie, ale skupiają się właśnie na zastosowaniu systemów sztucznej inteligencji w sądownictwie. Europejska Komisja na rzecz Efektywnego Wymiaru Sprawiedliwości (CEPEJ) w grudniu 2018 r. opublikowała *Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment*. Dokument ten przedstawia pięć zasad, których przestrzeganie jest konieczne do prawidłowego funkcjonowania systemów sztucznej inteligencji w sądownictwie: poszanowanie praw podstawowych, zakaz dyskryminacji, jakość i bezpieczeństwo, przejrzystość i obiektywność oraz nadzór użytkownika⁴².

⁴⁰ Zob. *Biała Księga*, s. 19-21.

⁴¹ Zob. Rada Adwokatur i Stowarzyszeń Prawniczych Europy, *CCBE Response to the consultation on the European Commission's White Paper on Artificial Intelligence*, 5 czerwca 2020 r., s. 6-7, <https://www.ccbe.eu/documents/position-papers/> [dostęp: 20.11.2021].

⁴² Zob. European Commission for the Efficiency of Justice, *European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their Environment*, 4 grudnia 2018 r., s. 7.

Różnice pomiędzy *Wytycznymi* i *Białą Księgą* a *Ethical Charter* prowadzą do wniosku, że w celu zagwarantowania najwyższego poziomu ochrony praw podstawowych, a przez to także zgodności z zasadą skutecznej ochrony sądowej, konieczne może być przyjęcie przez organy ustawodawcze UE aktów prawnych regulujących w sposób konkretny zastosowanie systemów sztucznej inteligencji w sądownictwie, a nie tylko ogólnych, horyzontalnych rozwiązań dotyczących wszelkich możliwych zastosowań sztucznej inteligencji.

4. Proponowany akt w sprawie sztucznej inteligencji

Skutkiem omówionych powyżej prac jest przedłożony 21 kwietnia 2021 r. projekt Aktu w sprawie sztucznej inteligencji. Po rozpatrzeniu wszystkich odpowiedzi uzyskanych w ramach konsultacji społecznych i przeprowadzeniu oceny pod kątem skutków gospodarczych i społecznych możliwych wariantów regulacji Komisja przedstawiła projekt horyzontalnego instrumentu legislacyjnego UE, uwzględniającego proporcjonalne podejście oparte na analizie ryzyka, uzupełniony o kodeksy postępowania dotyczące systemów sztucznej inteligencji, które nie są obciążone wysokim ryzykiem⁴³.

Komisja nie podjęła postulowanej np. przez CCBE pracy nad rozwiązaniami sektorowymi, które mogłyby być lepiej dostosowane do konkretnych zagrożeń związanych z wykorzystaniem systemów sztucznej inteligencji w różnych sferach, ale przedstawiła projekt rozporządzenia ustanawiającego wymogi dotyczące wszystkich systemów sztucznej inteligencji. Podjęto także decyzję o powiększeniu zapowiadanej początkowo liczby kategorii ryzyka z dwóch do czterech: niedopuszczalne ryzyko, wysokie ryzyko, ograniczone ryzyko i minimalne ryzyko.

Powstanie kategorii systemów sztucznej inteligencji niedopuszczalnego ryzyka, których użytek jest całkowicie zakazany, otwiera drogę do wprowadzenia zakazu zastosowania niektórych rodzajów systemów sztucznej inteligencji również w sądownictwie. Zakaz taki może okazać się w przyszłości konieczny do zapewnienia zgodności z zasadą skutecznej ochrony sądowej, zwłaszcza biorąc pod uwagę szybki rozwój technologiczny w zakresie przewidywania treści orzeczeń sądowych przez systemy sztucznej inteligencji⁴⁴.

Sprzeciw wśród komentatorów wzbudza jednak niemal nieistniejąca regulacja systemów sztucznej inteligencji ograniczonego i minimalnego ryzyka. W przypadku ryzyka ograniczonego jedynym proponowanym wymogiem, jaki musiałby spełnić system SI, jest obowiązek informowania użytkownika lub odbiorcy, że stosowane systemy (*chatboty* lub programy wykorzystujące technologię *deepfake*) są systemami sztucznej inteli-

⁴³ Zob. AIA, s. 11.

⁴⁴ Zob. N. Aletras, D. Tsarapatsanis, D. Preotiuc-Pietro, V. Lampos, *op. cit.*

gencji. Systemy ryzyka minimalnego nie są obejmowane żadnymi wymogami. Nie da się jednak wykluczyć, że systemy obecnie zaliczane wstępnie do kategorii ograniczonego lub minimalnego ryzyka mogą po wprowadzeniu na rynek lub udostępnieniu do użytku okazać się niebezpieczne dla zdrowia lub ochrony praw podstawowych⁴⁵.

Niezwykle istotne dla stosowania proponowanych przepisów jest nadanie AIA formy rozporządzenia, a nie dyrektywy. Zgodnie z przepisem art. 288 TFUE rozporządzenie wiąże w całości i jest bezpośrednio stosowane. TSUE, m.in. w wyroku w sprawie *Politi*, zaznaczył, że chodzi o pełny skutek bezpośredni, wertykalny i horyzontalny⁴⁶. Obecnie, według najlepszej wiedzy autora, polski system prawny nie zawiera żadnych przepisów regulujących tworzenie i zastosowanie sztucznej inteligencji.

5. Kategoria systemów sztucznej inteligencji wysokiego ryzyka

Dominująca większość przepisów proponowanego rozporządzenia AIA dotyczy systemów sztucznej inteligencji wysokiego ryzyka. Przepisy art. 6 i art. 7 AIA definiują systemy wysokiego ryzyka poprzez odesłanie do list umieszczonych w załącznikach. Zgodnie z przepisem art. 6 AIA systemami sztucznej inteligencji wysokiego ryzyka są:

- „systemy sztucznej inteligencji przeznaczone do wykorzystywania jako związane z bezpieczeństwem elementy produktu objętego unijnym prawodawstwem harmonizacyjnym wymienionym w załączniku II lub będące takimi produktami, jeżeli jednocześnie produkt, którego elementem jest system sztucznej inteligencji lub sam system podlegają, na mocy tego prawodawstwa harmonizacyjnego, ocenie zgodności przeprowadzanej przez osobę trzecią w celu wprowadzenia go do obrotu lub oddania do użytku;
- systemy, o których mowa w załączniku III”.

Pierwsza grupa dotyczy wielu rodzajów produktów: zabawek, wyrobów medycznych, urządzeń linii kolejowych, pojazdów silnikowych, statków powietrznych itp., które są przedmiotem prawodawstwa unijnego stwarzającego ramy prawne dla ich certyfikacji i dopuszczania do obrotu⁴⁷. Propozycja AIA zapowiada nowelizację odpowiednich aktów harmonizujących w tych sektorach, w celu dołączenia również wymogów proponowanych w AIA. Druga grupa obejmuje natomiast listę 8 obszarów i specyficznych sposobów

⁴⁵ F. Reinhold, A. Muller, *Algorithm Watch's response to the European Commission's proposed regulation on Artificial Intelligence – A major steps with major gaps*, 22 kwietnia 2021 r., <https://algorithmwatch.org/en/response-to-eu-ai-regulation-proposal-2021/> [dostęp: 20.11.2021].

⁴⁶ Zob. wyrok Trybunału z dnia 14 grudnia 1971 r. w sprawie C-43/71 *Politi przeciwko Ministero delle finanze*, ECLI:EU:C:1971:122, pkt 9.

⁴⁷ Zob. Załącznik II do AIA.

wykorzystania systemów sztucznej inteligencji w tych obszarach. Jednym z tych obszarów są właśnie „sprawowanie wymiaru sprawiedliwości i procesy demokratyczne”⁴⁸.

Propozycja AIA zawiera w obszarze sprawowania wymiaru sprawiedliwości i procesów demokratycznych tylko jedną kategorię: systemy sztucznej inteligencji, które mają służyć organowi sądowemu pomocą w badaniu i interpretacji stanu faktycznego i przepisów prawa oraz w stosowaniu prawa do konkretnego stanu faktycznego⁴⁹. Przykład systemu e-Court i plan utworzenia sędziego-sztucznej inteligencji w Estonii wskazują natomiast, że wykorzystanie SI w sądownictwie rozwija się w kierunku wykraczającym znacząco poza służenie organowi sądowemu pomocą w procesie rozpatrywania sprawy.

Proponowany przez Komisję horyzontalny instrument legislacyjny, który ma regulować ogólnie zastosowanie systemów sztucznej inteligencji, wydaje się nakierowany bardziej na wzmocnienie regulacji sektorowych dotyczących bezpieczeństwa produktów niż na skuteczne zabezpieczenie sektora publicznego, w tym sądownictwa, przed potencjalnymi naruszeniami praw podstawowych. Przemawia za tym odwołanie się do długiej listy aktów prawnych UE, które mają na celu harmonizację prawa w odniesieniu do konkretnych rodzajów produktów. Także motywy AIA, w szczególności motyw 5, podkreślają, że proponowane rozporządzenie ma umożliwić rozwój, wykorzystanie i upowszechnienie SI na rynku wewnętrznym, a ochrona praw podstawowych wymieniana jest tylko jako dodatkowy warunek mający wspierać funkcjonowanie rynku wewnętrznego.

6. Projektowane wymogi dotyczące systemów sztucznej inteligencji wysokiego ryzyka

Projekt AIA zawiera w przepisach art. 9-15 szereg wymogów, które musi spełnić system sztucznej inteligencji wysokiego ryzyka, aby zostać dopuszczony do obrotu lub użytku. Wymogi te wywiedzione zostały z *Wytycznych* oraz z *Białej Księgi*, a motywy AIA podkreślają ich znaczenie dla ograniczenia zagrożenia dla zdrowia, bezpieczeństwa i praw podstawowych. Większość przepisów tych artykułów rozpoczyna się od słów: „systemy sztucznej inteligencji wysokiego ryzyka projektuje się i opracowuje się w taki sposób, aby [...]”. Wskazuje to na obarczenie odpowiedzialnością za spełnienie tych wymogów twórców systemów SI wysokiego ryzyka. Potwierdza to treść przepisu art. 16 AIA, który jasno wskazuje, że zapewnienie zgodności z wymogami ustanawianymi przez AIA jest obowiązkiem dostawców systemów. „Dostawca” jest pojęciem

⁴⁸ Zob. Załącznik III do AIA.

⁴⁹ Zob. *ibidem*.

normatywnym, zdefiniowanym w przepisie art. 2 AIA jako: „osoba fizyczna lub prawna, organ publiczny, agencja lub inny podmiot, które opracowują system sztucznej inteligencji lub zlecają jego opracowanie w celu wprowadzenia go do obrotu lub oddania go do użytku pod własną nazwą handlową lub własnym znakiem towarowym – odpłatnie lub nieodpłatnie”. W przypadku systemów tworzonych przez podmioty prywatne, przykładem czego może być amerykański system COMPAS, odpowiedzialność spoczywa na tych właśnie podmiotach, twórcach systemów. Jeżeli natomiast stworzenie systemu zostaje zlecone przez organ publiczny, to właśnie ten organ jest odpowiedzialny za zgodność takiego systemu z przepisami AIA⁵⁰.

Spełnienie wymogów proponowanych w projekcie AIA ma zapewnić bezpieczeństwo systemów sztucznej inteligencji wysokiego ryzyka. Nie ulega wątpliwości, że Komisja stara się zrealizować wszystkie postulaty opracowane we wcześniejszych dokumentach politycznych dotyczących sztucznej inteligencji. Szczególnie istotne dla ochrony praw podstawowych, w tym zasady skutecznej ochrony sądowej, są wymogi dotyczące jakości danych, bezpieczeństwa oraz transparentności funkcjonowania systemów SI.

Dane treningowe, walidacyjne i testowe są istotne w przypadku systemów SI stosowanych w sądownictwie. To na ich analizie opiera się przyszłe działanie systemu. Błędy i usterki obecne w tych zbiorach danych, np. powielony brak obiektywizmu lub tendencyjność, mogą bezpośrednio przełożyć się na wyniki pracy systemu po oddaniu do użytku. Najdonioślejszym zagrożeniem wynikającym z takich błędów jest naruszenie zasady bezstronności sędziowskiej. System SI powinien być jedynie narzędziem, ale jeżeli narzędzie to powielać będzie tendencje zawarte w danych treningowych, sędzia może, nawet nieświadomie, w ślad za systemem, również powielać takie tendencje w swoich orzeczeniach. Zapewnienie odpowiedniej kontroli danych i przeciwdziałanie wymienionym problemom minimalizuje ryzyko naruszenia praw podstawowych, ale zastrzeżenia zgłaszane między innymi przez Swedsoft wskazują, że konieczna może być zmiana brzmienia przepisu art. 10 AIA w celu lepszego oddania rzeczywistych możliwości technologicznych⁵¹.

W kontekście pozostałych wymienionych wcześniej potencjalnych naruszeń zasady skutecznej ochrony sądowej poprzez wykorzystanie systemów sztucznej inteligencji

⁵⁰ Zob. art. 8-15 AIA.

⁵¹ Komentarz do proponowanego AIA opublikowany przez organizację pozarządową Swedsoft wskazuje wyraźnie, że wymogi adekwatności, reprezentatywności, wolności od błędów i kompletności danych są nierealistyczne. Osiągnięcie takich perfekcyjnych danych jak dotąd się nie udało, pomimo prób podejmowanych przez projektantów systemów sztucznej inteligencji. Ponadto, nawet weryfikacja spełnienia tych wymogów przez zbiory danych, w poszukiwaniu potencjalnych naruszeń, może okazać się niemożliwa przed dopuszczeniem do obrotu lub udostępnieniem do użytku. Zob. Swedsoft, *Comments Regarding the European Commission's proposal for an Artificial Intelligence Act*, 24 czerwca 2021 r., s. 4, <https://www.swedsoft.se/en/2021/06/29/comments-regarding-the-artificial-intelligence-act/> [dostęp: 20.11.2021].

w sądownictwie należy wskazać, że wymogi proponowane przez Komisję zmierzają do wprowadzenia mechanizmów, które miałyby dużą szansę zapobiegać tym zagrożeniom, w szczególności w odniesieniu do braku transparentności i „efektu czarnej skrzynki”. Wymóg wprowadzenia funkcji rejestracji zdarzeń oraz wymóg przejrzystości pozwalają oczekiwać zapewnienia przez dostawcę, że użytkownik będzie w stanie poznać i zrozumieć procedury rządzące funkcjonowaniem systemu sztucznej inteligencji oraz zinterpretować wyniki ich pracy pod kątem przesłanek prowadzących do tych wyników. Przepis art. 12 ust. 4 określa szczegółowo minimalne wymogi funkcji rejestracji zdarzeń w odniesieniu do systemów SI wysokiego ryzyka wykorzystywanych do identyfikacji biometrycznej osób fizycznych w czasie rzeczywistym, ale dla pozostałych sektorów charakteryzujących się wysokim ryzykiem, wymienionych w załączniku III do AIA, minimalne wymogi nie zostały określone. Wskazano jedynie, że poziom identyfikowalności funkcjonowania systemu musi być odpowiedni do jego przeznaczenia. W przypadku zastosowania systemów SI w sądownictwie odpowiedni poziom powinien odpowiadać gwarancjom i obowiązkom wynikającym z praw podstawowych (w tym zasady skutecznej ochrony sądowej) oraz przepisów proceduralnych. Użycie w art. 12 ust. 2 sformułowania „odpowiedni do” jest jednym z podkreślanych w komentarzu organizacji Swedsoft przykładów niejasnego sformułowania wymogów, które mogą potencjalnie prowadzić do nadmiernego obciążania odpowiedzialnością podmiotów tworzących systemy SI⁵². Wprowadzenie funkcji rejestrowania zdarzeń poprzez umożliwienie identyfikacji funkcjonowania systemu, a co za tym idzie, lepszego poznania i zrozumienia jego działania, odpowiada między innymi na potencjalne zagrożenia naruszenia prawa do uzyskania uzasadnienia orzeczenia. Kreuje to bowiem dla użytkownika systemu (sędziego) narzędzie do wyjaśnienia przebiegu i efektu pracy systemu. Także zasady niezawisłości i bezstronności sędziowskiej znajdują dodatkowe zabezpieczenie w istnieniu rejestru zdarzeń. Wprowadzenie kryterium przejrzystości działania systemów SI i jego spełnienie pozwala bowiem dokładniej monitorować ich pracę i szybko wykrywać potencjalne elementy stanowiące zagrożenie dla niezawisłości i bezstronności sędziego korzystającego z pomocy systemu SI, np. brak obiektywności lub potencjalne zewnętrzne ataki.

Proponowany przez Komisję wymóg nadzoru ze strony człowieka obejmuje możliwość podjęcia decyzji o zaniechaniu użycia systemu sztucznej inteligencji, przerwaniu już podjętej przez system pracy, a także „ręczną” zmianę lub odwrócenie wyniku tej pracy. Szczególnie możliwość podjęcia decyzji o nieużywaniu systemów SI jest istotna, ponieważ wskazuje, że działanie systemu sztucznej inteligencji powinno być jedynie fakultatywną możliwością, a nie koniecznością. W kontekście sądownictwa oznacza to,

⁵² *Ibidem*, s. 7.

że sprawy zawsze muszą być rozpatrywane przez sędziego-człowieka, który musi mieć możliwość podjęcia decyzji o rozpatrzeniu sprawy wyłącznie na podstawie własnej wiedzy i umiejętności, bez wykorzystania sztucznej inteligencji. Zapewnia to dodatkową możliwość rezygnacji z zastosowania przez sędziów wskazówek dostarczanych przez systemy sztucznej inteligencji. Jednocześnie Komisja wskazuje, że środki służące sprawowaniu nadzoru nad systemami SI ze strony człowieka mają zapewnić sposobność monitorowania działalności systemu tak, aby można było szybko wykrywać nieprawidłowe funkcjonowanie i nieoczekiwane wyniki. Ponownie należy wskazać, że tak jak w przypadku ogólnie sformułowanego wymogu wprowadzenia funkcji rejestracji zdarzeń, tak i w proponowanych przepisach dotyczących wymogu nadzoru stopień konkretyzacji jest zbyt niski. Brak jest bowiem jakichkolwiek propozycji dotyczących obszaru sądownictwa, chociaż określono minimalne wymogi dla użytkowania systemów służących do zdalnej identyfikacji biometrycznej osób fizycznych.

Wymóg dokładności, solidności i cyberbezpieczeństwa zmierza do zapewnienia ochrony przed wpływami z zewnątrz, w tym ataków hakerskich i prób wpłynięcia na wyniki pracy systemu poprzez ingerowanie w zbiory danych wejściowych. Szczególnie istotne jest w tym zakresie postulowane w propozycji AIA wprowadzenie środków pozwalających na weryfikację integralności zbiorów danych treningowych i wejściowych. Wejście w życie przepisów AIA, w obecnym ich brzmieniu, może pozwolić na skuteczniejsze zapobieganie potencjalnym naruszeniom wymogów niezawisłości i bezstronności sędziowskiej, które mogą wystąpić jako efekt zewnętrznego wpływu na dane wprowadzane do systemów SI lub algorytmy, według których takie systemy funkcjonują. Sędzia w wykonywaniu swoich zadań musi być całkowicie niezależny od czynników zewnętrznych, a także wolny od uprzedzeń i stronniczości. Systemy SI wykorzystywane przez sędziów stanowią niewątpliwie czynnik zewnętrzny wobec nich. Gwarancja bezpieczeństwa tych systemów, w tym brak ukrytych celów oraz odporność na ataki z zewnątrz, jest konieczna do zapewnienia, że system pozostaje narzędziem i nie wpływa na sędziego w sposób, który może budzić uzasadnione wątpliwości co do jego niezawisłości. Przepis art. 15 AIA stanowi, że dostawcy systemów SI powinni wprowadzać nadmierne zabezpieczenia i rozwiązania techniczne mające służyć solidności tych systemów. Wymogi dokładności, solidności i cyberbezpieczeństwa są jednymi z nielicznych wymogów zawartych w AIA, które prawdopodobnie nie wymagają szczególnego dostosowania lub doszczegółowienia dla systemów SI wykorzystywanych w sądownictwie. Zagrożenia naruszenia zasady skutecznej ochrony sądowej w powyższej sferze, którym stara się przeciwdziałać ten przepis, są bardzo poważne, ale ich bezpośrednią przyczyną są kwestie czysto techniczne: bezpieczeństwo systemu i odporność na ataki z zewnątrz.

Czysto techniczne jest zatem także rozwiązanie tego problemu, ale prawo powinno przewidywać obowiązek spełnienia konkretnych kryteriów.

Oprócz zapewnienia zgodności systemów SI z wymogami sformułowanymi w AIA na dostawcach systemów spoczywają również inne obowiązki. W większości są to jednak obowiązki proceduralne związane z właściwym certyfikowaniem systemów SI oraz przeprowadzaniem kontroli zgodności z wymogami. W kontekście ochrony praw podstawowych, w tym zasady skutecznej ochrony sądowej, znaczenie mają trzy spośród pozostałych obowiązków dostawców. Przede wszystkim dostawca ma obowiązek sporządzać dokumentację techniczną swoich systemów SI. Taka dokumentacja jest przekazywana użytkownikowi celem spełnienia wymogu przejrzystości, a możliwość zaznajomienia się z nią jest jednym z elementów umożliwiających użytkownikowi lepsze zrozumienie funkcjonowania systemu SI. Podobną funkcję pełnią rejestry zdarzeń generowane z użyciem funkcji rejestrowania zdarzeń. Do obowiązków dostawcy należy przechowywanie tych rejestrów, jeżeli znajdują się one pod kontrolą dostawcy. Im lepiej użytkownik (sędzia) rozumie działanie systemu SI, tym lepiej może wyjaśnić, w jaki sposób system osiąga konkretne wyniki. Wzmacnia to gwarancje praw wynikających z zasady skutecznej ochrony sądowej, w szczególności prawa do uzyskania uzasadnienia orzeczenia. Trzecim obowiązkiem dostawców jest podejmowanie niezbędnych działań naprawczych, jeżeli system SI nie spełnia któregośkolwiek z wymogów zawartych w AIA. Nie ma to bezpośredniego związku z zasadą skutecznej ochrony sądowej, ale służy podtrzymaniu i umocnieniu ochrony wynikającej z wymogów proponowanych przez Komisję, w tym także przepisów, które w odniesieniu do systemów sztucznej inteligencji stosowanych w sądownictwie służą ochronie sprawnego funkcjonowania sądów⁵³.

Projekt AIA nakłada obowiązki również na użytkowników systemów sztucznej inteligencji wysokiego ryzyka. Wśród tych obowiązków najdonioślejsze znaczenie w kontekście zapobiegania naruszeniom zasady skutecznej ochrony sądowej ma zwłaszcza obowiązek zapewnienia adekwatności danych wejściowych, zgodnie z przeznaczeniem systemu, w granicach, w jakich użytkownik sprawuje kontrolę nad tymi danymi⁵⁴. Danymi wejściowymi w przypadku systemów SI wykorzystywanych w sądownictwie mogą być m.in. akta sprawy. Nadzór nad sposobem wprowadzenia tych danych do systemu, ich integralnością i autentycznością służy zagwarantowaniu prawa do sprawiedliwego rozpatrzenia sprawy, poprzez zapewnienie poprawności pracy systemów SI, a co za tym idzie także rzetelności całości postępowania. Dodatkowo, każdorazowa kontrola wprowadzanych danych służy również minimalizacji ryzyka powielania potencjalnej tendencyjności lub innych usterek w tych danych, uzupełniając ochronę wymogu bezstronności sędziowskiej.

⁵³ Zob. art. 16-24 AIA.

⁵⁴ Zob. art. 29 AIA.

7. Podsumowanie

Proponowane w AIA wymogi i obowiązki, w kontekście opisanych w niniejszym artykule problemów, pozwalają zapobiegać w szczególności zagrożeniom naruszenia zasady skutecznej ochrony sądowej związanym z zabezpieczeniem systemów sztucznej inteligencji przed wpływem zewnętrznym, w tym również atakiem hakerskim, a także przed potencjalną manipulacją danymi wejściowymi. Również postulat spełnienia szczególnych wymogów przez dane wykorzystywane podczas projektowania i testowania systemu sztucznej inteligencji pozwala liczyć na wyższy stopień wczesnego wykrywania problemów wynikających z potencjalnego braku obiektywizmu w zbiorach danych, co służy m.in. minimalizacji ryzyka naruszenia wymogu bezstronności sędziowskiej. Niestety, zgłaszane przez organizacje pozarządowe trudności związane z osiągnięciem wymaganego przepisami AIA poziomu doskonałości danych budzą wątpliwości co do rzeczywistej możliwości realizacji tych propozycji.

Zagrożenie braku transparentności procesów decyzyjnych podejmowanych przez systemy sztucznej inteligencji, którego efektem może być naruszenie prawa do uzasadnionego orzeczenia, minimalizowane jest poprzez wprowadzenie wymogów tworzenia szczegółowej specyfikacji technicznej oraz funkcji rejestrowania zdarzeń. Do ich spełnienia konieczne będzie podniesienie poziomu wiedzy użytkowników systemów SI (sędziów) na temat działania tych systemów. Efektem zwiększonej wiedzy powinna być umiejętność adekwatnego wyjaśnienia efektów pracy systemów SI, co zmniejszy zagrożenie wystąpienia „efektu czarnej skrzynki”. Nie jest to jednak rozwiązanie dopasowane do wykorzystania systemów sztucznej inteligencji w sądownictwie.

Powyższe uzasadnia stwierdzenie, że w pracach Komisji nad utworzeniem ram prawnych regulujących sztuczną inteligencję zauważalna jest tendencja do tworzenia przepisów mających możliwie uniwersalny charakter. Nie można wykluczyć, że takie przepisy w połączeniu z aktami prawa Unii Europejskiej harmonizującymi konkretne sektory rynku wewnętrznego będą wystarczające do zapewnienia skutecznego zapobiegania zagrożeniom dla życia, zdrowia i praw podstawowych w tych sektorach. Sądownictwo jest jednak obszarem odmiennym, ponieważ użycie jakiegokolwiek systemu do administrowania pracą sądów, wspierania procesu rozstrzygania spraw przez sędziów, a nawet potencjalnie do rozstrzygania spraw zamiast sędziów musi być zgodne z szeregiem praw i wymogów, których rolą jest zapewnienie skutecznego funkcjonowania całego wymiaru sprawiedliwości – wykonywania władzy sądowniczej. W mojej ocenie postulowane przez CCBE i zgodne ze stanowiskiem CEPEJ rozwiązanie sektorowe, zaprojektowane specyficznie dla potrzeb wykorzystania systemów sztucznej inteligencji w sądownictwie, jest konieczne do osiągnięcia celów sformułowanych w *Wytocznych* i umożliwienia pełnego wykorzystania potencjału technologii sztucznej inteligencji,

przy jednoczesnym zagwarantowaniu efektywnej ochrony praw podstawowych, czego nieodłącznym elementem jest zgodność z zasadą skutecznej ochrony sądowej.

Bibliografia

Akty prawne

- Karta praw podstawowych Unii Europejskiej, Dz. Urz. UE C 326/391 z 26.10.2012 r.
Konwencja o ochronie praw człowieka i podstawowych wolności, Dz. U. z 1993 r. Nr 61, poz. 284.
Traktat o funkcjonowaniu Unii Europejskiej, Dz. Urz. UE C 326/47 z 26.10.2012 r.

Orzecznictwo sądowe

- Wyrok ETPC z dnia 28 czerwca 1984 r. w sprawie *Campbell i Fell przeciwko Zjednoczonemu Królestwu*, skargi nr 7819/77 i 7878/77.
Wyrok ETPC z dnia 22 października 1984 r. w sprawie *Sramek przeciwko Austrii*, skarga nr 8790/79.
Wyrok ETPC z dnia 30 listopada 1987 r. w sprawie *H. przeciwko Belgii*, skarga nr 8950/80.
Wyrok Trybunału z dnia 14 grudnia 1971 r. w sprawie C-43/71 *Politi przeciwko Ministero delle finanze*, ECLI:EU:C:1971:122.
Wyrok Trybunału z dnia 13 marca 2007 r. w sprawie C-432/05 *Unibet (London) Ltd i Unibet (International) Ltd przeciwko Justitiekanslern*, ECLI:EU:C:2007:163.
Wyrok Trybunału z dnia 6 września 2012 r. w sprawie C-619/10 *Trade Agency Ltd przeciwko Seramico Investments Ltd*, ECLI:EU:C:2012:531.
Wyrok Sądu Najwyższego Stanu Wisconsin z 13 lipca 2016 r. w sprawie *State przeciwko Loomis*, sygn. 881 N.W.2d 749 (2016).

Literatura

- Aletras N., Tsarapatsanis D., Preotiuc-Pietro D., Lampos V., *Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective*, "PeerJ Computer Science" 2016, t. 93, nr 2.
Angwin J., Larson J., Mattu S., Kirchner L., *Machine Bias. There's a software used across the country to predict future criminals. And it's biased against blacks*, "ProPublica", 23 maja 2016 r., <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.
Dede G., Hamon R., Junklewitz H., Naydenov R., Malatras A., Sanchez M., *Cybersecurity challenges in the uptake of Artificial Intelligence in Autonomous Driving*, Publications Office of the European Union, Luxembourg 2021.
Flores A., Lowenkamp C., Bechtel K., *False Positives, False Negatives, and False Analyses: A Rejoinder to "Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And it's Biased Against Blacks"*, "Federal Probation Journal" 2016, t. 80, nr 2.
Garlicki L. (red.), *Konwencja o Ochronie Praw Człowieka i Podstawowych Wolności. Tom I. Komentarz do artykułów 1-18*, Warszawa 2010.
Letteron R., *Digital access to the law*, "Digital Issues of Les Annales des Mines" 2018, nr 3.

- Metz C., *The Sadness and Beauty of Watching Google's AI Play Go*, "Wired", 3 listopada 2016 r., <https://www.wired.com/2016/03/sadness-beauty-watching-googles-ai-play-go/>.
- Morison J., Harkens A., *Re-engineering justice? Robot judges, computerised courts and (semi) automated legal decision-making*, Queen's University Belfast, "Legal Studies" 2019, t. 39, nr 4.
- Nakad-Weststrate H., Jongbloed T., van der Herik H., Salem A., *Digitally Produced Judgements in Modern Court Proceedings*, "International Journal of Digital Society" 2015, nr 4.
- Rahman F., *COMPAS Case Study: Fairness of a Machine Learning Model*, "Towards Data Science", 7 września 2020 r., <https://towardsdatascience.com/compas-case-study-fairness-of-a-machine-learning-model-f0f804108751>.
- Russel S., Norvig P., *Artificial Intelligence: a modern approach*, Berkeley 1994.
- Safjan M., *Prawo do skutecznej ochrony sądowej – refleksje dotyczące wyroku TSUE z 19.11.2019 r. w sprawach połączonych C-585/18, C-624/18, C 625/18*, „Palestra” 2020, nr 5.
- Thiebes S., Lins S., Sunyaev A., *Trustworthy artificial intelligence*, "Electronic Markets" 2021, t. 31, nr 2.

Inne

- AlgorithmWatch, *Automating Society Report 2019*, 1 stycznia 2019 r., <https://algorithmwatch.org/en/automating-society/>.
- Europejska Komisja na rzecz skutecznego wymiaru sprawiedliwości, *European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment*, przyjęta w Strasbourgu 4 grudnia 2018 r.
- Grupa Ekspertów Wysokiego Szczebla ds. Sztucznej Inteligencji powołana przez Komisję Europejską w czerwcu 2018 r., *Wytoczne w zakresie etyki dotyczące godnej zaufania sztucznej inteligencji*, opublikowane 10 kwietnia 2019 r.
- Komisja Europejska, *Biała Księga w sprawie sztucznej inteligencji – Europejskie podejście do doskonałości i zaufania*, Bruksela 19 lutego 2020 r., COM/2020/65 final.
- Opinia nr 11 Rady Konsultacyjnej Sędziów Europejskich (CCJE) w sprawie jakości orzecznictwa, 18 grudnia 2008 r., <https://www.coe.int/en/web/ccje/opinion-n-11-on-the-quality-of-judicial-decisions>.
- Predictive justice: when algorithms pervade the law*, "Paris Innovation Review" 2017, nr 9, <http://parisinnovationreview.com/articles-en/predictive-justice-when-algorithms-pervade-the-law>.
- Rada Adwokatur i Stowarzyszeń Prawniczych Europy, *CCBE Response to the consultation on the European Commission's White Paper on Artificial Intelligence*, 5 czerwca 2020 r., <https://www.ccbe.eu/documents/position-papers/>.
- Reinhold F., Muller A., *AlgorithmWatch's response to the European Commission's proposed regulation on Artificial Intelligence – A major steps with major gaps*, 22 kwietnia 2021 r., <https://algorithmwatch.org/en/response-to-eu-ai-regulation-proposal-2021/>.
- Swedsoft, *Comments Regarding the European Commission's proposal for an Artificial Intelligence Act*, 24 czerwca 2021 r., <https://www.swedsoft.se/en/2021/06/29/comments-regarding-the-artificial-intelligence-act/>.
- Wniosek Komisji Europejskiej dotyczący rozporządzenia Parlamentu Europejskiego i Rady ustanawiającego zharmonizowane przepisy dotyczące sztucznej inteligencji (akt w sprawie sztucznej inteligencji) i zmieniającego niektóre akty ustawodawcze Unii, Bruksela 21 kwietnia 2021 r., COM/2021/206 final.

Źródła internetowe

Dane na temat „e-demokracji” i internetowego głosowania w wyborach powszechnych w Estonii, <https://e-estonia.com/solutions/e-governance/i-voting/>.

Skan formularza z pytaniami używanego do oceny ryzyka przez program COMPAS, <https://www.documentcloud.org/documents/2702103-Sample-Risk-Assessment-COMPAS-CORE.html>.

Różne definicje sztucznej inteligencji, <https://www.sztuczna inteligencja.org.pl/definicja/sztuczna-inteligencja/>.