KRAJOWA IZBA
RADCÓW PRAWNYCH

OŚRODEK BADAŃ, STUDIÓW I LEGISLACJI KRAJOWEJ RADY RADCÓW PRAWNYCH

KRAJOWA IZBA RADCÓW PRAWNYCH | ul. Powązkowska 15, 01-797 Warszawa | tel.: 22 300 86 40 | kirp@kirp.pl | NIP 5261043011

Opinia Ośrodka Badań, Studiów i Legislacji Krajowej Rady Radców Prawnych do projektu Polityki Rozwoju Sztucznej Inteligencji w Polsce do 2030 r. z dnia 17 maja 2026 r.

I. Wykaz skrótów

Skrót	Rozwinięcie
AI	Sztuczna inteligencja
AI Act	Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2024/1689 z dnia 13 czerwca 2024 r. w sprawie ustanowienia zharmonizowanych przepisów dotyczących sztucznej inteligencji oraz zmiany rozporządzeń (WE) nr 300/2008, (UE) nr 167/2013, (UE) nr 168/2013, (UE) 2018/858, (UE) 2018/1139 i (UE) 2019/2144 oraz dyrektyw 2014/90/UE, (UE) 2016/797 i (UE) 2020/1828 (akt w sprawie sztucznej inteligencji)
DSA	Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2022/2065 z dnia 19 października 2022 r. w sprawie jednolitego rynku usług cyfrowych oraz zmiany dyrektywy 2000/31/WE (akt o usługach cyfrowych)
KNF	Komisja Nadzoru Finansowego
Konstytucja RP	Konstytucja Rzeczypospolitej Polskiej Dz.U. 1997 Nr 78, poz. 483
Opinia	Ta opinia
Polityka	Projekt Polityki Rozwoju Sztucznej Inteligencji w Polsce do 2030 r. – załącznik do uchwały Rady Ministrów
PUODO	Prezes Urzędu Ochrony Danych Osobowych
RODO	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)

RPD	Rzecznik Praw Dziecka
RPO	Rzecznik Praw Obywatelskich
UE	Unia Europejska
UKE	Urząd Komunikacji Elektronicznych
UOKiK	Urząd Ochrony Konkurencji i Konsumentów
Wytyczne HLEG	„Wytyczne w zakresie etyki dotyczące godnej zaufania sztucznej inteligencji", opracowane przez HLEG, opublikowane w dniu 8 kwietnia 2019r. , https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai, dostęp: 15 .05.2026 r.

II. Podsumowanie Opinii

Zdaniem autorek i autora Opinii Polityka zasadniczo jest zgodna z regulacjami unijnymi (w tym AI Act). Autorki i autor Opinii uważają, że pomimo jej ogólności (co jest zasadniczą istotą tworzenia dokumentów w rodzaju polityk i strategii), niektóre obszary Polityki wymagają doprecyzowania.

III. Generalne uwagi do Polityki

Niezależnie od omawianego fragmentu Polityki, należy wskazać kilka generalnych uwag, które należy uwzględnić w Polityce.

1) Przede wszystkim, aby Polityka mogła zostać sprawnie zrealizowana w Polsce, konieczne jest sprawne przeprowadzanie przez polskiego ustawodawcę procesów legislacyjnych, szczególnie w zakresie tworzenia ustaw implementujących akty prawne UE do polskiego porządku prawnego. Na dzień sporządzenia Opinii projekt polskiej ustawy o systemach sztucznej inteligencji (wdrażający do Polski poszczególne obszary AI Act) jest po I czytaniu w Sejmie, stanowi on wyłącznie jeden z aktów, na jaki należy zwrócić uwagę przy właściwym ustawodawczym zabezpieczeniu wątków związanych z AI Act, a które regulują szeroko pojęte prawo nowych technologii. Dla przykładu, w dalszym ciągu nie uchwalono polskiej ustawy wdrażającej DSA, a jest to akt prawny niezwykle istotny z perspektywy funkcjonowania Internetu, a jego postanowienia oddziałują również na regulacje AI Act (np. w ramach przeciwdziałania deepfake). Analogicznie konieczne jest jak najszybsze uchwalenie ustawy o sprawiedliwym dostępie do danych i ich wykorzystywaniu (wdrażającej Data Act). Jeżeli polska nie

będzie posiadała uchwalonych i przejrzystych regulacji, może to powodować niepewność prawną dla przedsiębiorców, a tym samym może zniechęcać z jednej strony zagranicznych przedsiębiorców od inwestowania w Polsce, a z drugiej krajowe podmioty przed wdrażaniem bardziej kosztowych rozwiązań w obszarze AI.

2) W zakresie prac legislacyjnych w Polityce warto uwzględnić potrzebę bieżącą współpracę ze środowiskami, nie tylko m.in. z akademickimi i biznesowymi podmiotami (co uwzględniono), ale także samorządami zawodów zaufania publicznego (w tym samorządem radców prawnych) – ujęcie takiego wątku w Polityce będzie wskazywało, że troska o rozwój AI w Polsce spoczywa również na licznych grupach zawodowych, które na co dzień korzystają z rozwiązań tej technologii.
3) Do rozważenia byłoby dodanie jako załącznika do Polityki zakresu polskich aktów prawnych, jakie wymagają nowelizacji w związku z potrzebą realizacji Polityki – dzięki temu zainteresowane środowiska (w tym przedsiębiorcy) będą poinformowane, jakie obszary w niedalekiej przyszłości będą zmieniane, co pozwoli im lepiej przygotować się, w tym zabezpieczyć odpowiednie zasoby.

IV. Szczegółowe uwagi do Polityki

Poniżej zamieszczono konkretne fragmenty Polityki, które zdaniem autorek i autora Opinii zasługują na szczególną analizę.

1. Otwarte dane. Modele open-source

Odnosząc się do podrozdziału 3.4. „Otwarte dane. Modele open-source" należy w pierwszej kolejnosci podzielić zdecydowanie pozytywny obraz realizowanej w Polsce strategii otwierania danych. Jednocześnie jednak wydaje się, że w dokumencie stanowiącym Politykę rozwoju sztucznej inteligencji w Polsce do 2030 r. więcej uwagi powinno być poświęcone temu, w jaki sposób wykorzytany ma być potencjał otwartych danych oraz ponownego wykorzytywania informacji sektora publicznego na potrzeby rozwoju sztucznej inteligencji. Bez watpienia „[o]twarte dane umożliwiają tanie trenowanie algorytmów AI, wspierając itd. rozwiązania dla smart city, transportu, zdrowia czy ochrony środowiska czy innych obszarów. Ułatwiają też współdzielenie danych w ramach ekosystemu i rozwój innowacji branżowych" (s. 27). Od Polityki oczekiwać raczej należałoby nie tyle formułowania tego typu aksjomatów, co wyraźnego i możliwie skonkretyzowanego nakreślenia celów, zadań, kierunków korzystania z otwartych danych na potrzeby rozwoju AI. Nie chodzi tu bowiem o dokonywanie opisu

okoliczności faktycznych (nawet najbardziej sprzyjających i pozytywnych), ale o wyznaczanie celów i zadań priorytetowych.
Podobnie rzecz się ma z częścią poświęconą modelom OPEN-SOURCE, ponieważ w Polityce nie tyle chodzi o ich charakterystykę istniejącego stanu rzeczy, ale o wytyczenie kierunków rozwoju, założeń, standardów działania, zadań do realizacji.

Wątpliwe konstrukcyjnie jest również zawarcie w treści tej części Polityki rozwoju sztucznej inteligencji w Polsce informacji bieżących i działaniach podejmowanych przez Komisję Europejską i jej planach uruchomienia Inicjatywy na rzecz Frontier AI, która połączy wiodące ośrodki naukowe i przemysłowe w Europie wraz z wyjaśnieniem, że dzięki temu Europa wzmocni swoją pozycję jako lider w bezpiecznej i etycznej sztucznej inteligencji. To treść zdecydowanie zbędna w polskiej Polityce rozwoju AI.

Na aprobatę i uwagę zasługują natomiast „Główne kierunki działań" wyznaczone w tym podrozdziale. Merytorycznie są one poprawne i odpowiadają oczekiwaniom, jakie stawia się tego rodzaju dokumentom.

2. Sprawne państwo (pkt 4 Polityki)

Polityka rozwoju sztucznej inteligencji w Polsce do 2030 r. obejmuje rozdział 4 zatytułowany „Sprawne państwo", w obrębie którego wyróżnione zostały cztery wiodące zagadnienia tematyczne określone jako: 4.1 Cel i działania dla rozwoju AI w administracji; 4.2. Wdrożenia AI w administracji publicznej (AI HUB Poland); 4.3. AI w usługach publicznych (ochrona zdrowia, wymiar sprawiedliwości); 4.4. Bezpieczeństwo i infrastruktura krytyczna.

Wyraźne wyodrębnienie w ten sposób w projektowanej Polityce płaszczyzny zastosowania sztucznej inteligencji w sektorze publicznym uznać należy za rozwiązanie nie tylko uzasadnione merytorycznie, ale również zasługujące na zaakcentowanie i pochwałę. W powszechnie prowadzonym dyskursie oraz w licznych opracowaniach dotyczących sztucznej inteligencji szczególną uwagę zwraca się przede wszystkim na wykorzystanie tej technologii w biznesie i gospodarce oraz najczęściej wyraźnie osadza się tę problematykę wśród zagadnień istotnych dla funkcjonowania sektora prywatnego. Tymczasem korzystanie z rozwiązań AI w szeroko pojętym sektorze publicznym jest zagadnieniem zdecydowanie godnym wyodrębnienia i bez wątpienia zasługującym na osobną uwagę i pogłębioną analizę ze względu na specyfikę i sektora publicznego, jego niejednorodność, ograniczenia i uwarunkowania w

zakresie sprawowania władztwa publicznego, ogromny zakres oddziaływania w sferze wolności i praw, istotne zaangażowanie środków publicznych, liczne ograniczenia wynikające z reżimu regulacyjnego kształtowanego już na gruncie przepisów konstytucyjnych, jak i uwarunkowań prawnych wynikających z prawa UE, w tym przepisów mających bezpośrednie zastosowanie do rozwiązań opartych na sztucznej inteligencji[1].

Wdrożenie sztucznej inteligencji w sektorze publicznym niewątpliwie niesie za sobą ogromny potencjał – od automatyzacji rutynowych zadań po lepsze dopasowanie usług do potrzeb obywateli[2]. Założenie to trafnie odzwierciedla zamysł, by w opiniowanej Polityce osobno potraktować problematykę sprawnego państwa wykorzystującego rozwiązania sztucznej inteligencji. Biorąc pod uwagę treść poszczególnych jednostek redakcyjnych rozdziału 4, pewien niedosyt budzi jednak stopień konkretyzacji zadań i działań, jakie w świetle Polityki mają zmierzać do osiągnięcia założonych w niej rezultatów. Rzecz jasna specyfika dokumentów o charakterze strategicznym pozwala na przyjmowanie w retoryki dość ogólnej i kierunkowej, jednak zważywszy na to, że w przypadku ustaleń dotyczących zadań zmierzających do efektywnego i zgodnego z prawem wdrożenia rozwiązań opartych na sztucznej inteligencji w sektorze tak istotnie osadzonym w rygorystycznym reżimie prawnym wynikającym z konstytucyjnej zasady legalizmu (itd. 7 Konstytucji RP) i normatywnych pochodnych tej zasady, uzasadnionym byłoby przy formułowaniu kierunkowych zadań i wytycznych w sposób zdecydowanie bardziej konkretny określać cele i zadania z uwzględnieniem regulacji prawnych i pewnych systemowych wymogów kształtowanych zarówno przez prawo UE, jak i prawo krajowe. Tytułem przykładu – od strony prakseologicznej trudno dyskutować z tezą postawioną na wstępie podrozdziału 4.1. – w ramach przybliżenia „Celu 3", iż „[r]ozwój i wdrażanie sztucznej inteligencji w administracji publicznej (rządowej i samorządowej) pozwoli na zautomatyzowanie rutynowych czynności, usprawni procesy decyzyjne, komunikację z obywatelami i między urzędami, a usługi publiczne będą świadczone w sposób łatwy, skuteczny, terminowy i dostosowany do potrzeb". Nakreślono w ten sposób standard obsługi, który ma być osiągnięty w stosunkowo niedługiej perspektywie czasowej, bo

[1] Szerzej zob. E. Feret, *Decentralizacja władzy publicznej – relacje czy zależności w aspekcie finansowego działania jednostek samorządu terytorialnego w dobie AI*, FK 2026, nr 2, s. 7-18; H. Guz, M. L. Żabiński, *Emergentne AI. Jak samorząd lokalny adaptuje się do rewolucji AI*, ST 2025, nr 4, s. 56-71; P. Dolniak, T. Kuźma, A. Ludwiński, K. Wasik, *Rozdział 3 AI w służbie publicznej* [w:] P. Dolniak, T. Kuźma, A. Ludwiński, K. Wasik, *Sztuczna inteligencja w wymiarze sprawiedliwości. Między prawem a algorytmami*, Warszawa 2024, s. 73 i n.

[2] Szerzej zob. M. Sakowska-Baryła, Sztuczna inteligencja w sektorze publicznym, [w:] *Prawo sztucznej inteligencji i nowych technologii. 3*, B. Fischer, A Pązik, M. Świerczyński, Warszawa 2023, s. 21 i n.

zaledwie do 2030 r. jest zdecydowanie kuszący. Od strony prawnej jednak, co w praktyce przełoży się na prakseologię, wiele wątpliwości wiąże się z tym, jak ów stan osiągnąć. Nie wdając się w tym miejscu w bardziej szczegółowe ustalenia, już na tle tak ogólnie (acz kusząco) określonego rezultatu, można podawać w wątpliwość, czy sposób wyznaczania w Polityce celów i narracja przybliżająca metodę ich realizacji jest adekwatna.

Zasadne jest przyjmować, że wyznaczanie zadań w dokumencie rangi Polityki strategicznej dla rozwoju państwa powinno następować w sposób, który w pozwoli dekodować konkretne zadania podmiotów, które tę Politykę będą realizować. Należą do nich w pierwszym rzędzie organy władzy publicznej każdej z doktrynalnie wyodrębnianych kategorii – od władzy ustawodawczej po wykonawczą i sądowniczą, których piastuni na podstawie treści Polityki powinny móc w sposób przybliżony chociaż przewidzieć, jakie działania, w jakiej kolejności, w jakim czasie będą obowiązane podjąć, aby osiągnąć rezultat opisany w Polityce jako wyznacznik tytułowego „Sprawnego państwa", o którym mowa w rozdziale 4.

Biorąc pod uwagę powyższe, podstawową rekomendacją dotycząca treści tego rozdziału, byłoby skupienie się w nim nie tyle na nakreślaniu kuszącej wizji tego, by sztuczną inteligencję „pozycjonować jako strategiczny atut zintegrowany z instytucjami i usługami" (s. 39), ale na bliższym określeniu w jaki sposób i za sprawą czyich działań ten stan osiągnąć. Wskazać bowiem trzeba, że w sektorze publicznym korzystanie z AI nie jest wyłącznie kwestią efektywności, ale jest to zagadnienie o charakterze ustrojowym dotykające: zasady legalizmu, godności ludzkiej, prawa do sądu, prawa do prywatności, prawa do ochrony danych, zasady równości, niedyskryminacji, transparentności działania państwa. Korzystanie z systemów sztucznej inteligencji w sektorze publicznym wymaga zatem nie tyle wyznaczenia standardu obsługi interesanta (sprawnie, szybko, trafnie, terminowo, itd.) i analizy technologicznej, ale przede wszystkim prawnej, ustrojowej i etycznej[3].

Poniżej zostaną poczynione uwagi odnoszące się do poszczególnych jednostek redakcyjnych rozdziału 4 opiniowanej Polityki, ale zważywszy na dokonane już kierunkowe spostrzeżenia dotyczące tego rozdziału, uzasadnione wydaje się, by w tym miejscu sformułować kilka

[3] Szerzej zob. M. Szpyrka, *Uczenie maszynowe a prawa podstawowe. Czyli o obowiązkach organów administracji publicznej, które wdrażają systemy AI wysokiego ryzyka*, [w:] *Wpływ aktu w sprawie sztucznej inteligencji na funkcjonowanie administracji publicznej*, G. Sibiga (red.), Warszawa 2025, s. 79 i n.

rekomendacji merytorycznych dotyczących koniecznych obszarów zainteresowania twórców Polityki, które powinny być odzwierciedlone w konkretnych ustaleniach zawartych w jej treści:

a) określenie w Polityce celów i zadań powinno nastąpić z poszanowaniem zasady legalizmu (itd. 7 Konstytucji RP), a więc fundamentalnego założenia, że organy władzy publicznej mogą działać wyłącznie na podstawie i w granicach prawa; wdrożenie systemu AI wymaga zatem: identyfikacji wyraźnej podstawy prawnej przetwarzania danych, oceny, czy regulacja ustawowa dopuszcza automatyzację procesu, ustalenia, czy decyzje mogą być podejmowane w sposób częściowo lub w pełni zautomatyzowany – w kontekście ochrony danych osobowych kluczowe znaczenie mają uwarunkowania określone w RODO zwłaszcza w zakresie podstaw przetwarzania oraz zakazu podejmowania wyłącznie zautomatyzowanych decyzji w indywidualnych przypadkach;
b) zapewnienie mechanizmów ochrony godności ludzkiej, ochrony przed dyskryminacją i nierównościami, co ma istotne znaczenie zważywszy na związane ze stosowaniem systemów AI ryzyko dyskryminacji *biasu* algorytmicznego, a w konsekwencji utrwalania nierówności społecznych, systemowego wykluczania określonych grup, wzmacniania uprzedzeń;
c) kierunkowe wytyczne i założenia dotyczące zautomatyzowanego podejmowania decyzji i profilowania, w tym dotyczące dopuszczalności wydawania decyzji administracyjnych przy wsparciu AI, realizacji prawa jednostki do uzyskania wyjaśnień i możliwości skutecznego odwołania;
d) kierunkowe wytyczne dotyczące zapewnienia odpowiednich podstaw prawnych przetwarzania danych osobowych, ze szczególnym uwzględnieniem danych szczególnych kategorii oraz problematyki zmiany wtórnego przetwarzania danych itd. na potrzeby tworzenia modeli AI wykorzystywanych w sektorze publicznym[4];
e) kierunkowe wytyczne dotyczące ochrony prywatności;
f) uwzględnienie okoliczności, że w sektorze publicznym wiele systemów AI będzie z gruntu kwalifikowanych jako systemy wysokiego ryzyka, a co za tym idzie wzięcie pod uwagę konieczności realizacji szczególnych obowiązków z tym związanych, których niezbędnie trzeba będzie dopełnić w obszarze administracji publicznej, edukacji, zatrudnienia, świadczeń socjalnych;

[4] Szerzej zob. M. Sakowska-Baryła, *RODO i AI Act – przegląd zagadnień na styku współstosowanych regulacji*, [w:] *Sztuczna inteligencja a RODO – wyzwania prawne i praktyczne*, M. Sakowska-Baryła (red.), Wrocław 2026, s. 15 i n.

g) wytyczne dotyczące zapewnienia transparentności i prawa do informacji, w tym wprowadzenia rozwiązań dotyczących jawności algorytmów stosowanych przez administrację, dostępu do informacji publicznej o zasadach działania systemów, łagodzenia konfliktu między wartościami chronionymi konstytucyjnej i ustawowo a zasadą transparentności[5];

h) przybliżenie granic dopuszczalnej automatyzacji w demokratycznym państwie prawnym, ponieważ w kontekście korzystania z AI zwłaszcza przez organy władzy publicznej podstawową kwestią jest to, czy każdą czynność można zautomatyzować bez naruszenia istoty sprawowania władzy oraz istoty prawa do dobrej administracji; kierunkowe wskazówki powinny dotyczyć rozstrzygania takich zagadnień, jak automatyzacja decyzji uznaniowych, stosowanie AI w postępowaniach o charakterze represyjnym, wykorzystywanie systemów predykcyjnych w nadzorze nad obywatelami.

Odzwierciedlenie w Polityce realizacji przywołanych tu założeń dotyczących stosowania sztucznej inteligencji pozwoliłoby na skonkretyzowanie „treści tego dokumentu, a jednocześnie pozwoliło podmiotom potencjalnie zainteresowanym realizacją poszczególnych zadań (czy zobowiązanym też do ich wykonania) na odpowiednie przygotowanie i podjęcie skonkretyzowanych działań.

Cel i działania dla rozwoju AI w administracji (Ad. 4.1. Polityki)

Treść podrozdziału 4.1. ma charakter wyraźnie postulatywny i jak wynika z pierwszych akapitów tego podrozdziału, zawarty w nim opis zmierza do określenia działań urzeczywistniających celu 3 – „Sprawne Państwo wykorzystujące rozwiązania sztucznej inteligencji". O ile kierunkowo uzasadnione jest, że w tym miejscu Polityka nawiązuje do „Apply AI Strategy" ogłoszonej przez Komisje Europejską, o tyle nie wydaje się celowe, aby w priorytetowym dokumencie dotyczącym rozwoju AI w Polsce, jakim przecież jest przedmiotowa Polityka, przywoływać działania Komisji, jej zamierzenia i cele wynikające z rzeczonej strategii. Zdecydowanie bardziej uzasadnione byłoby, gdyby w tym miejscu sprecyzowane zostały zadania po stronie polskich organów władzy publicznej i instytucji służące wykorzystaniu potencjału AI dla zwiększenia efektywności administracji publicznej i szeroko pojętego publicznego sektora.

[5] Szerzej zob. G. Bar, *Przejrzystość, w tym wyjaśnialność, jako wymóg prawny dla systemów Sztucznej Inteligencji* (dodatek MoP 20/2020), Monitor Prawniczy 2020, nr 20, s. 75–79.

Choć pierwsza część podrozdziału poświęcona jest przywołaniu projektowanych działań Komisji oraz dość ogólnym postulatom polepszenia działania administracji publicznej dzięki zastosowaniu AI, to zdecydowanie na uznanie zasługuje zawarta tu deklaracja, że „[w]drażane będą inicjatywy podnoszące świadomość pracowników w zakresie ochrony danych osobowych oraz przestrzegania przepisów dotyczących prywatności, zgodnie z oczekiwaniami obywateli", a także zapowiedź wdrażania AI w administracji publicznej w sposób gwarantujący suwerenność infrastruktury informacyjnej Państwa (za sprawą projektu PLLuM) oraz zapowiedź, że „Planowane jest także stworzenie jednolitej, dostępnej publicznie listy systemów AI używanych w administracji publicznej wraz z opisem funkcji oraz podstawowych parametrów technicznych dostępnych dla obywateli w celu zwiększenia transparentności administracji publicznej i zaufania społecznego do właściwie zaprojektowanych systemów AI", jak również dostrzeżenie nieodzowności opracowania i egzekwowania zasad zamawiania i stosowania AI w administracji publicznej i istotnej roli samorządów. W przypadku tych ostatnich jednak zdecydowanie rekomendowane byłoby sformułowanie bliższych wytycznych dotyczących założeń i realizacji programów pilotażowych wykorzystywania AI.

Za godne pochwały i prawidłowo określone należy uznać „Główne kierunki działań" wieńczące podrozdział 4.1.

Wdrożenia AI w administracji publicznej (AI HUB Poland) (Ad. 4.2. Polityki)

Ideą zasługującą na aprobatę jest stworzenie AI HUB Poland jako centralnej platformy zarządzania, rozwijania i wdrażania technologii AI w sektorze publicznym w Polsce, która to instytucja będzie pełnić funkcję platformy koordynującej wdrożenia AI w administracji oraz wspierającej projekty pilotażowe. Prawidłowo określone zostały również rezultaty zaplanowanych działań w ramach AI HUB Poland i główne kierunki prac. Pewien niedosyt budzi jednak brak wskazania przybliżonego choćby terminarza realizacji zadań oraz wytycznych dla podmiotów, które potencjalnie będą uczestniczyć w tym przedsięwzięciu.

Na tle ustaleń zawartych w treści podrozdziału 4.2. pojawia się wątpliwość dotycząca tego, jaka jest relacja pomiędzy utworzeniem i działaniem AI HUB Poland a opisanymi w podrozdziale 4.1. pilotażowych programów wdrożeń AI. Uzasadnione byłoby choćby punktowe wskazanie tego, jak mają się do siebie ww. pilotaże do działania tej centralnej platformy zarządzania, rozwijania i wdrażania technologii AI w sektorze publicznym w Polsce. Pomiędzy tymi inicjatywami powinna być przecież spójność i choć szczątkowa koordynacja. Doświadczenia ostatnich lat cyfryzacji i wdrożeń różnego rodzaju systemów informatycznych pokazują

bowiem, że brak spójnej koncepcji wdrożeń IT skutkuje wieloma niedogodnościami i problemami o charakterze technicznym, organizacyjnym i finansowym. Chociażby problem - ogólnie rzecz ujmując – niekompatybilności systemów i zróżnicowanych formatów gromadzenia danych stanowi istotną przeszkodę w otwieraniu danych i zapewnianiu możliwości ich ponownego wykorzystywania. Podobnie rzecz się będzie miała w obszarze wdrożeń systemów AI. Nieskoordynowana różnorodność wdrażanych rozwiązań może być przeszkodą w zapewnieniu tytułowego dla analizowanego rozdziału sprawnego państwa w dziedzinie stosowania AI.

AI w usługach publicznych (ochrona zdrowia, wymiar sprawiedliwości) (Ad. 4.3. Polityki)

Ze względu na uwarunkowania ustrojowe, proceduralne, organizacyjne i techniczne nie wydaje się rozwiązaniem adekwatnym i uzasadnionym, by stosowanie systemów AI w ochronie zdrowia i wymiarze sprawiedliwości w dokumencie rangi strategicznej państwowej polityki AI ujmować wspólnie jako usługi publiczne. To dwa obszary funkcjonowania państwa o tak istotnie odmiennym charakterze i odrębnych uwarunkowaniach ustrojowych, że już z tego powodu powinno się im poświęcić uwagę odrębnie. Inne cele i specyfika charakteryzuje też systemy AI wykorzystywane odpowiednio w wymiarze sprawiedliwości[6]. Powoduje to, że polityki, kierunkowe wytyczne i strategie w zakresie działania tych dziedzin powinny znaleźć odzwierciedlenie w odrębnych podrozdziałach. Dodatkowo niepokój budzi to, że w analizowanej polityce brak osobnego podrozdziału dotyczącego obszaru sprawowania władztwa (głównie w formie decyzji administracyjnych) przez organy administracji publicznej. To również odrębna problematyka, której nie sposób uplasować ani jako ochrona zdrowia, ani jako wymiar sprawiedliwości. Być może wedle zamysłu projektodawcy tego dokumentu kwestia sprawowania władztwa przez organy administracji publicznej z wykorzystaniem sztucznej inteligencji miała się zamknąć w ogólnych wywodach dotyczących tego, jak sprawnie działać będzie administracja publiczna wsparta systemami AI (wprowadzenie podrozdziału 4.1). Zważywszy jednak na różnorodność działań podejmowanych przez administrację, niezwykle szoki zakres merytoryczny stosowania prawa przez organy administracji publicznej i związany z tym istotny stopień ingerencji w sferę wolności i praw osób fizycznych,

[6] Szerzej zob. K. Kiejnich-Kruk, *Społeczne uwarunkowania wykorzystania sztucznej inteligencji w sądownictwie*, PS 2023, 68-86; P. Dolniak, T. Kuźma, A. Ludwiński, K. Wasik, *Rozdział 7. Sędzia i AI – sojusz czy starcie* [w:] P. Dolniak, T. Kuźma, A. Ludwiński, K. Wasik, *Sztuczna inteligencja w wymiarze sprawiedliwości. Między prawem a algorytmami*, Warszawa 2024, s. 221 i n.; K. Maciejewska, *Etyczne aspekty związane z wykorzystaniem AI jako narzędzia stanowiącego wsparcie usług cyfrowych w wymiarze sprawiedliwości*, [w:] *Prawo sztucznej inteligencji i nowych technologii 2*, B. Fischer, A. Pązik, M. Świerczyński (red.), Warszawa 2022, s. 263 i n.

uzasadnione wydaje się szersze potraktowanie w Polityce rozwoju AI korzystanie z systemów sztucznej inteligencji przez te właśnie podmioty. Nie da się bowiem ani zaliczyć tej sfery działalności sektora publicznego do ochrony zdrowia, ani do wymiaru sprawiedliwości, ani też do ogólnie ujętych „usług publicznych", bo często nie o usługi tu chodzi, ale o władztwo nacechowane znacząco działaniami o charakterze sankcyjnym czy represyjnym, bo przecież i w tym zakresie organy administracji publicznej mogą być potencjalnie wspierane przez AI. Co więcej – jak wynika z treści podrozdziału 4.1. – oczekiwać należy, że w istocie organy administracji publicznej tak właśnie wspierane będą. To wszystko prowadzi do rekomendowania, by w Polityce odrębnie określić zasady, priorytety i założenia dotyczące stosowania systemów AI w administracji publicznej ze szczególnym uwzględnieniem praktyki orzeczniczej. Jednocześnie – jak się wydaje – ta część Polityki byłaby właściwym miejscem na sformułowanie odpowiednich założeń dotyczących funkcjonowania piaskownic regulacyjnych, o których mowa w AI Act, jako rozwiązaniom, dzięki którym w bezpiecznych, kontrolowanych warunkach można by, testować rozwiązania AI służące wsparciu administracji publicznej. Takie kierunkowe ustalenia w dokumencie rangi strategicznej dla rozwoju sztucznej inteligencji w państwie byłyby zdecydowanie pożądane.

Odnosząc się do sekcji „OCHRONA ZDROWIA" wyodrębnionej w podrozdziale 4.3. niedosyt budzi sposób sformułowania treści Polityki w tym względzie. Siedem pierwszych akapitów zawartego tam wywodu to ogólne luźno powiązane ze sobą spostrzeżenia i prognozy dotyczące stosowania AI w sektorze zdrowia. O ile zastosowanie systemów AI w medycynie trudno przecenić, bo notoryjnie znaną okolicznością jest to, że sztuczna inteligencja jest szeroko stosowana zwłaszcza w diagnostyce, o tyle w analizowany tekście trudno dopatrzyć się treści, jakich oczekiwać należałoby w przypadku dokumentu będącego państwową polityką rozwoju AI[7]. Zwykle terminem „polityka" określa się zbiór zasad, priorytetów i kierunków działania wyznaczony przez wprowadzający ja podmiot. Z polityki zwykle wynika co i dlaczego chce się osiągnąć, co jest istotne, jakie działania są oczekiwane, wokół jakich celów skupiony jest wysiłek jej adresatów. Tymczasem analizowany dokument niestety nie spełnia tego standardu. Obejmuje umiarkowanie spójny merytorycznie opis faktów, przypuszczeń i postulatów, choć należałoby oczekiwać, że w akcie tej rangi zostanie wskazane co, jak i kiedy sektor medyczny w obszarze stosowania AI ma osiągnąć. Większym stopniem konkretyzacji charakteryzuje się

[7] Szerzej zob. M. Jankowska, M. Pawełczyk, M. Sakowska-Baryła, *Sztuczna inteligencja i wybrane aspekty cyfrowej transformacji w systemie medycznym*, [w:] *Prawo sztucznej inteligencji i nowych technologii 2*, B. Fischer, A. Pązik, M. Świerczyński (red.), Warszawa 2022, s. 218 i n.

część tekstu określona jako „Główne kierunki działań". Cześć ta jest również merytorycznie poprawna. Dodatkowe zagadnienie, jakie zasadnie powinno być ujęte w tej części polityki, stanowi problematyka stosowania rozporządzenia Parlamentu Europejskiego i Rady (UE) 2025/327 z dnia 11 lutego 2025 r. w sprawie europejskiej przestrzeni danych dotyczących zdrowia oraz zmiany dyrektywy 2011/24/UE i rozporządzenia (UE) 2024/2847 (Dz. U. UE. L. z 2025 r. poz. 327), które z pewnością będzie miało istotne znaczenie w kontekście przetwarzania danych medycznych w systemach AI i na potrzeby ich tworzenia.

Odnosząc się do wyodrębnionej w podrozdziale 4.3.sekcji „WYMIAR SPRAWIEDLIWOŚCI", podobnie, jak w przypadku ww. sekcji poświęconej ochronie zdrowia, pierwsza część wywodu poświęcona jest ogólnej (raczej negatywnej) diagnozie stanu faktycznego. W pierwszej części tej sekcji nie tyle wyznacza się kierunki rozwoju zastosowania AI w wymiarze sprawiedliwości, co przede wszystkim utyskuje się nad licznymi bolączkami tego sektora. Ów zabieg retoryczny należy uznać za chybiony, bowiem zwykle w politykach wyznacza się cele, kierunki zmian, kluczowe założenia, pożądane działania, czego zdecydowanie brak w tej części podrozdziału 4.3. Pewną nadzieję na to, że projektodawca Polityki rzeczywiście zechciał nakreślić jakieś perspektywy działań niesie dopiero część „Główne kierunki działań". Zasadniczo są one określone poprawnie, jednak nader wąsko i niejednoznacznie – zważywszy na to, jak krytycznie stan wymiaru sprawiedliwości został nakreślony na wstępie tej samej sekcji i jak wiele wyzwań związanych jest z wdrożeniem systemów AI w sądownictwie. Pamiętać trzeba jednak, że korzystanie z AI w wymiarze sprawiedliwości to nie tylko postępowania sądowe, ale także wspomagane przez AI działania prokuratorów[8], biegłych, mediatorów itd., ale na ten temat brak wzmianek w analizowanej Polityce.

Bezpieczeństwo i infrastruktura krytyczna (Ad. 4.4. Polityki)

Treść podrozdziału 4.4. skupiona została na zagadnieniu bezpieczeństwa i infrastruktury krytycznej. Współcześnie to zagadnienia o zdecydowanie fundamentalnym charakterze, więc uzasadnione jest, by dokument strategiczny, jakim jest analizowana Polityka, określał cele, zadania i priorytety w tej dziedzinie. Zaprojektowana treść Polityki budzi jednak istotne wątpliwości. Po pierwsze dlatego, że skupiona jest wyłącznie wokół zagadnień dotyczących infrastruktury krytycznej, tymczasem istotnych, kierunkowych wytycznych zdecydowanie

[8] Szerzej zob. P. Drembkowski, *Granice prawne stosowania sztucznej inteligencji i Big Data przez prokuraturę – analiza prawna, etyczna i technologiczna*, Prok.i Pr. 2025, s. 10-35.

wymaga już samo bezpieczeństwo (cyberbezpieczeństwo), więc uzasadnione byłoby – zgodnie z zapowiedzią zawartą w tytule podrozdziału – poświęcić mu choć kilka wskazówek i sformułować jakieś bazowe złożenia. Po drugie, zagadnienie infrastruktury krytycznej potraktowane zostało mało przejrzyście . Właściwie pewien walor informacyjny ma dopiero akapit, w którym stwierdzono, że „[z]godnie z podejściem przedstawionym w strategii „Apply AI Strategy", bezpieczeństwo i odporność systemów sztucznej inteligencji stanowią kluczowy filar ich wdrażania, zwłaszcza w sektorach o znaczeniu strategicznym i w infrastrukturze krytycznej. Strategia podkreśla konieczność budowy „trustworthy and safe AI", czyli rozwiązań zapewniających niezawodność technologiczną, cyberbezpieczeństwo oraz zgodność z zasadami prawa i etyki". Merytorycznie poprawne wydają się natomiast „Główne kierunki działań" i należy je przyjąć z aprobatą.

3. **Monitorowanie wpływu AI na gospodarkę, rynek pracy i społeczeństwo (pkt 5.5. Polityki)**

Należy zwrócić uwagę na wątek Polityki związanej z „Wpływem AI na rynek pracy". Do docenia jest świadomość twórcy Polityki, że rozwiązania AI w coraz większym stopniu będą miały wpływ na rynek pracy.

Bezpośrednio Polityka nie porusza jednak tematu **wpływu AI na warunki pracy i prawa pracownicze**, poza stwierdzeniem, że trzeba **kształcić kadry** (by nie było niedoboru specjalistów ICT) oraz i że programy *reskilling/upskilling* będą wspierać dostosowanie rynku pracy do AI.

Polityka nie wspomina natomiast nic o **regulacji monitorowania pracowników przez AI** czy **ochronie zatrudnionych przed zautomatyzowanym zwalnianiem**.

Warto zwrócić w Polityce uwagę na Dyrektywę Parlamentu Europejskiego i Rady (UE) 2024/2831 z dnia 13 września 2024 r. w sprawie poprawy warunków pracy za pośrednictwem platform cyfrowych oraz zmiany dyrektywy 2019/1152 i 2019/1158 (tzw. dyrektywa o platformach) oraz polskie prace legislacyjne w tym zakresie.

Warto również byłoby poruszyć w Polityce wątek wykorzystywania przez pracodawców rozwiązaniami AI w celu zarówno prowadzenia rekrutacji, jak również zarządzania

zatrudnionymi pracownikami – z roku na rok coraz więcej pracodawców będzie powierzać tego typu zadania algorytmom, co będzie miało znaczący wpływ na prawa i wolności pracowników.

4. Sztuczna inteligencja godna zaufania (pkt 6 Polityki)

4.1. UWAGI OGÓLNE

W rozdziale 6. Opiniowanego projektu „Polityki rozwoju sztucznej inteligencji w Polsce do 2030r." zatytułowanym „Sztuczna inteligencja godna zaufania" przyjęto szeroką perspektywę aksjologiczną i regulacyjną. Projekt trafnie lokuje rozwój sztucznej inteligencji w kontekście ochrony praw człowieka, prawa do prywatności i ochrony danych osobowych, ochrony konkurencji i konsumentów, praw twórców, a także potrzeb grup szczególnie wrażliwych. Takie ujęcie pozostaje zasadniczo zgodne z podejściem prezentowanym w Wytycznych HLEG[9] dotyczących godnej zaufania AI (**Wytyczne HLEG**), zgodnie z którym systemy AI powinny być nie tylko zgodne z prawem, lecz także etyczne oraz solidne z technicznego i społecznego punktu widzenia.

Na aprobatę zasługuje również dostrzeżenie w projekcie szczególnej sytuacji dzieci, seniorów, osób z niepełnosprawnościami oraz osób zagrożonych wykluczeniem. Uwzględnienie tych grup jest istotne z perspektywy zasady zorientowania AI na człowieka (*human-centric AI*), a także potrzeby przeciwdziałania pogłębianiu istniejących nierówności społecznych i informacyjnych. Podejście to koresponduje z Wytycznymi HLEG, które akcentują konieczność szczególnej ostrożności w odniesieniu do grup wrażliwych oraz sytuacji asymetrii władzy lub informacji, w szczególności w relacjach pracodawca – pracownik, przedsiębiorca – konsument oraz organ publiczny – obywatel.

Projekt właściwie identyfikuje istotne obszary ryzyka i interwencji, takie jak dezinformacja, ochrona danych osobowych, prawa autorskie, transparentność systemów AI oraz piaskownice regulacyjne. W całej Polityce wskazano również bardziej konkretne cele odnoszące się do infrastruktury, wsparcia MŚP, piaskownic regulacyjnych oraz wdrożeń w sektorze publicznym, w tym cel zakładający testowanie rozwiązań przez 50 firm rocznie w środowisku piaskownic regulacyjnych oraz wzrost poziomu zaufania społecznego do rozwiązań AI o co najmniej 20%.

[9] High-Level Expert Group on Artificial Intelligence (HLEG AI) - niezależna grupa ekspertów wysokiego szczebla ds. sztucznej inteligencji powołana przez Komisję Europejską w 2018 r. w celu doradzania Komisji w zakresie strategii dotyczącej sztucznej inteligencji, w tym opracowania wytycznych etycznych oraz rekomendacji dotyczących polityki i inwestycji w obszarze AI.

Jednocześnie należy wskazać, że powyższe elementy nie zostały w wystarczającym stopniu przełożone na treść samego rozdziału dotyczącego sztucznej inteligencji godnej zaufania. Rozdział ten wymaga silniejszego powiązania z konkretnymi instrumentami realizacyjnymi przewidzianymi w innych częściach Polityki.

4.2. GŁÓWNA UWAGA SYSTEMOWA

Zasadnicza uwaga do rozdziału „Sztuczna inteligencja godna zaufania" dotyczy **nieuwzględnienia modelu godnej zaufania AI wypracowanego w ramach Wytycznych HLEG**. Wytyczne te opierają się na czytelnej strukturze, zgodnie z którą godna zaufania AI powinna łącznie spełniać trzy cechy: **zgodność z prawem, zgodność z zasadami etycznymi oraz solidność techniczną i społeczną**. Cechy te są następnie rozwijane poprzez **cztery zasady etyczne:**

1) poszanowanie autonomii człowieka,
2) zapobieganie szkodom,
3) sprawiedliwość oraz
4) możliwość wyjaśnienia

oraz **siedem wymogów operacyjnych**, obejmujących:

1) przewodnią i nadzorczą rolę człowieka,
2) solidność techniczną i bezpieczeństwo,
3) ochronę prywatności i zarządzanie danymi,
4) przejrzystość,
5) różnorodność, niedyskryminację i sprawiedliwość,
6) dobrostan społeczny i środowiskowy oraz
7) odpowiedzialność.

Analizowany rozdział nie wykorzystuje tej struktury jako podstawowej **matrycy porządkującej**. Zamiast tego koncentruje się na wybranych zagadnieniach, takich jak ochrona danych osobowych, ochrona konsumentów, prawa autorskie i prawa twórców, bez równorzędnego uwzględnienia wszystkich kluczowych wymogów godnej zaufania AI. W szczególności niedostatecznie rozwinięte pozostają kwestie nadzoru człowieka nad systemami AI, przeciwdziałania dyskryminacji algorytmicznej, dobrostanu społecznego i środowiskowego, odpowiedzialności, audytowalności oraz technicznych i organizacyjnych środków zapewniania zgodności, bezpieczeństwa i kontroli nad systemami AI.

Zastosowanie struktury opartej na Wytycznych HLEG umożliwiłoby systematyczne powiązanie zasad etycznych z konkretnymi instrumentami wdrożeniowymi, podmiotami odpowiedzialnymi, harmonogramem oraz miernikami realizacji, a tym samym zwiększyłoby spójność Polityki z europejskim podejściem do AI oraz jej praktyczną użyteczność.

Należy przy tym podkreślić, że model zaproponowany w wytycznych HLEG ma charakter **horyzontalny** i został opracowany jako ogólne ramy oceny oraz wdrażania godnej zaufania AI niezależnie od sektora jej zastosowania. Nie ogranicza się zatem do wybranych obszarów, takich jak ochrona danych, konsumenci czy własność intelektualna, lecz obejmuje pełny cykl życia systemów AI oraz ich wpływ na jednostki, społeczeństwo, instytucje publiczne i rynek. Jest to podejście **szeroko wykorzystywane w europejskiej debacie regulacyjnej i strategicznej dotyczącej AI**, a jego zastosowanie w Polityce pozwoliłoby uniknąć fragmentarycznego i niespójnego ujęcia omawianej problematyki.

4.3. UWAGI I REKOMENDACJE SZCZEGÓŁOWE

4.3.1. Brak odniesienia do agentów AI i systemów agentowych

Rozdział odnosi się do modeli AI, systemów AI, chatbotów, personalizacji oraz wykorzystania AI w administracji publicznej, jednak nie uwzględnia w wystarczającym stopniu rozwoju systemów agentowych, w tym tzw. *agentic AI*. Jest to istotna luka, zwłaszcza że Polityka ma wyznaczać kierunki działań do 2030 r., a systemy tego rodzaju będą odgrywać coraz większą rolę zarówno w sektorze prywatnym, jak i publicznym.

Systemy agentowe mogą być zdolne do planowania, wykonywania sekwencji działań, korzystania z narzędzi zewnętrznych, delegowania zadań, działania w środowisku cyfrowym oraz podejmowania czynności w imieniu użytkownika. Wiąże się to z odrębnymi kategoriami ryzyka, w szczególności z ryzykiem **autonomicznego działania poza intencją użytkownika, eskalacji uprawnień, niezamierzonych skutków działań, trudności w przypisaniu odpowiedzialności oraz braku efektywnego nadzoru człowieka.**

Rekomendowane jest zatem uzupełnienie Polityki o wyraźne odniesienie do systemów agentowych. W szczególności Polityka powinna przewidywać zasady nadzoru człowieka nad takimi systemami, obowiązek rejestrowania działań agenta, ograniczania jego uprawnień, testowania w środowiskach kontrolowanych, określania dopuszczalnego zakresu autonomii oraz stosowania procedur awaryjnego zatrzymania lub wyłączenia systemu.

4.3.2. Wyjaśnialność systemów AI

Wątpliwości budzi sformułowanie, zgodnie z którym „zasady wsparcia urzędników przez AI będą przejrzyste i wyjaśnialne". Deklaracja ta jest zbyt kategoryczna i wymaga doprecyzowania. W przypadku wielu systemów AI, zwłaszcza opartych na sztucznych sieciach neuronowych i modelach głębokiego uczenia maszynowego, **pełna techniczna wyjaśnialność działania modelu może być ograniczona albo niemożliwa do osiągnięcia w praktyce.**

Nie oznacza to jednak, że systemy takie nie mogą być wykorzystywane w sposób kontrolowany i zgodny z wymogami przejrzystości. Konieczne jest rozróżnienie między pełną wyjaśnialnością techniczną modelu a szerszą **przejrzystością procesu jego wykorzystania**. W administracji publicznej szczególne znaczenie powinny mieć: dokumentacja systemu, identyfikowalność danych i działań, ścieżka audytu, nadzór człowieka, informowanie obywatela o wykorzystaniu AI oraz możliwość zakwestionowania wyniku lub decyzji podjętej z wykorzystaniem AI.

4.3.3. Brak odniesienia się do kryteriów sprawiedliwości i niedyskryminacji jako samodzielnego filaru Polityki

Omawiany rozdział wspomina o ryzyku uprzedzeń algorytmicznych oraz o potrzebach osób wykluczonych, jednak nie czyni z przeciwdziałania dyskryminacji algorytmicznej odrębnego i systemowo ujętego obszaru działań. Jest to istotny brak, ponieważ **różnorodność, niedyskryminacja i sprawiedliwość stanowią jeden z kluczowych wymogów godnej zaufania AI w ujęciu HLEG.**

Ryzyka dyskryminacyjne mogą występować w wielu obszarach stosowania AI, w tym w administracji publicznej, zatrudnieniu, edukacji, ochronie zdrowia, usługach finansowych, ubezpieczeniach, usługach cyfrowych oraz wymiarze sprawiedliwości. Mogą one wynikać zarówno z jakości danych treningowych, jak i z błędnych założeń projektowych, niewłaściwej walidacji, braku reprezentatywności danych lub niewystarczającego testowania wpływu systemu na różne grupy społeczne. Rekomendowane jest zatem dodanie opisu odrębnego kierunku działań dotyczącego przeciwdziałania dyskryminacji algorytmicznej.

4.3.4. Brak dostatecznego uwzględnienia dobrostanu człowieka i środowiska

Rozdział 6. Projektu „Polityki…" jedynie w ograniczonym zakresie odnosi się do wpływu AI na dobrostan człowieka i środowisko. Tymczasem dobrostan społeczny i środowiskowy jest jednym z wymogów godnej zaufania AI. Obejmuje on nie tylko wpływ systemów AI na środowisko naturalne, ale także ich oddziaływanie na jednostki i społeczeństwo.

Polityka powinna zatem przewidywać **ocenę wpływu systemów AI na zdrowie psychiczne, relacje społeczne, autonomię informacyjną, warunki pracy oraz środowisko, w tym zużycie energii, wody i zasobów obliczeniowych przez systemy AI**. Wymóg ten powinien być uwzględniany w zamówieniach publicznych, audytach oraz projektach finansowanych ze środków publicznych.

4.3.5. Niedostateczne ujęcie krajowego systemu nadzoru nad AI

W kontekście godnej zaufania AI warto byłoby w omawianym rozdziale skupić się na stworzeniu skutecznego, skoordynowanego i niezależnego systemu nadzoru nad AI, obejmującego zarówno **organ właściwy do spraw AI, jak i mechanizmy współpracy z innymi organami publicznymi**, w szczególności Prezesem Urzędu Ochrony Danych Osobowych (PUODO), Urzędem Ochrony Konkurencji i Konsumentów (UOKiK), Rzecznikiem Praw Obywatelskich (RPO), Rzecznikiem Praw Dziecka (RPD), Komisją Nadzoru Finansowego (KNF), Urzędem Komunikacji Elektronicznej (UKE) oraz Polskim Centrum Akredytacji.

Ma to szczególne znaczenie ze względu na horyzontalny charakter AI oraz nakładanie się różnych reżimów prawnych: AI Act, RODO, prawa konsumenckiego, prawa antydyskryminacyjnego, prawa pracy, przepisów sektorowych i regulacji dotyczących cyberbezpieczeństwa.

4.3.6. Nieprecyzyjne ujęcie kwestii ochrony danych osobowych i „legalizowania" działań organów publicznych

Fragment dotyczący ochrony danych osobowych jest zasadny kierunkowo, jednak wymaga doprecyzowania. W szczególności wątpliwości budzi sformułowanie odnoszące się do przyjęcia ram regulacyjnych „legalizujących działania organów publicznych stosujących rozwiązania sztucznej inteligencji". Może ono sugerować, że ram prawnych dla AI w sektorze publicznym nie ma, tymczasem podstawy prawne wykorzystania AI zarówno w sektorze

prywatnym jak i publicznym wynikają z obowiązujących przepisów prawa, w tym z RODO i AI Act. W szerszym kontekście zakres uprawnień i obowiązków w tym zakresie określa także Konstytucja RP, Karta praw podstawowych UE oraz właściwe ustawy.

Dokument strategiczny może wskazywać potrzebę doprecyzowania podstaw prawnych w określonych sektorach, jednak nie powinien sugerować konieczności blankietowej „legalizacji" stosowania AI przez administrację publiczną. Bardziej adekwatne byłoby wskazanie, że celem Polityki jest zapewnienie praktycznych mechanizmów stosowania istniejących i projektowanych ram prawnych.

Ochrona danych osobowych powinna być traktowana jako wymóg horyzontalny przy projektowaniu, wdrażaniu i wykorzystywaniu systemów AI we wszystkich obszarach Polityki. Nie chodzi o tworzenie odrębnego reżimu wobec RODO, lecz o **zapewnienie praktycznych mechanizmów jego stosowania w projektach AI, w szczególności przez wytyczne, wzory oceny skutków dla ochrony danych, standardy dokumentowania danych, zasady anonimizacji i pseudonimizacji, mechanizmy realizacji praw osób, audytowalność oraz współpracę organu nadzoru nad AI z PUODO.**

4.3.7. Zbyt wąskie ujęcie ochrony konsumentów i relacji z DSA

Uwagi dotyczące ochrony konsumentów są zasadne, jednak zostały ujęte zbyt wąsko. Wskaźnik dotyczący „transparentności algorytmów personalizujących oferty zakupowe" koncentruje się przede wszystkim na e-commerce, podczas gdy ryzyka konsumenckie związane z AI są znacznie szersze.

Ponadto część obowiązków dotyczących przejrzystości systemów rekomendacyjnych, reklam i działania platform internetowych wynika już z DSA. Polityka powinna zatem unikać prostego dublowania istniejących obowiązków regulacyjnych i wskazać wartość dodaną działań krajowych, w szczególności działań UOKiK, UKE i innych właściwych organów.

Zamiast ograniczać działania do transparentności algorytmów personalizujących oferty zakupowe, **Polityka powinna objąć szersze ryzyka konsumenckie związane ze stosowaniem AI**, w tym *dark patterns*, manipulację behawioralną, *dynamic pricing*, profilowanie podatności konsumenta, syntetyczne recenzje, *deepfake marketing*, chatboty podszywające się pod ludzi oraz automatyczne systemy obsługi reklamacji. Działania krajowe

powinny skupiać się na praktycznej i skutecznej realizacji obowiązków już wynikających z DSA, AI Act i prawa ochrony konsumentów.

4.3.8. Zbyt kategoryczne ujęcie kwestii twórczości z udziałem AI

Fragment dotyczący ochrony praw autorskich i praw twórców wymaga ostrożniejszego ujęcia. Stwierdzenie, że treści wygenerowane wyłącznie przez AI nie mogą być uznane za utwory, odpowiada dominującemu podejściu, zgodnie z którym ochrona prawnoautorska wymaga twórczego i indywidualnego wkładu człowieka. Polityka nie powinna jednak przesądzać abstrakcyjnie, czy określone kategorie treści powstałych z udziałem AI są utworami. Zasadne jest przyjęcie **neutralnego technologicznie podejścia**, zgodnie z którym ochronie prawnoautorskiej może podlegać twórczy i indywidualny wkład człowieka, także wtedy, gdy AI była narzędziem wspierającym proces twórczy.

Konieczne jest natomiast **wdrożenie praktycznych narzędzi** w zakresie transparentności, dokumentowania wkładu człowieka, poszanowania praw do materiałów treningowych oraz efektywnych mechanizmów opt-out i dochodzenia roszczeń.

4.3.9. Niedostateczne wskaźniki realizacji celu „godnej zaufania AI”

Wskaźniki zaproponowane w rozdziale nie mierzą w wystarczającym stopniu rzeczywistego poziomu wdrożenia godnej zaufania AI. Liczba audytów, liczba wytycznych czy odsetek przeszkolonych osób są przede wszystkim wskaźnikami aktywności, a nie rezultatów. Nie pozwalają one jednoznacznie ocenić, czy systemy AI są faktycznie bardziej przejrzyste, bezpieczne, niedyskryminacyjne, audytowalne i poddane skutecznemu nadzorowi człowieka. Rekomendowane jest uzupełnienie wskaźników o mierniki odnoszące się bezpośrednio do wymogów godnej zaufania AI.

Obszar	**Proponowany wskaźnik**
Nadzór człowieka	Odsetek systemów AI stosowanych w administracji publicznej, dla których określono tryb nadzoru człowieka
Przejrzystość	Odsetek publicznych systemów AI posiadających publicznie dostępną kartę/instrukcję systemu AI
Niedyskryminacja	Liczba przeprowadzonych ocen wpływu algorytmicznego na prawa podstawowe i ryzyko dyskryminacji

Odpowiedzialność	Liczba systemów AI wyposażonych w procedurę skargi, odwołania lub zakwestionowania wyniku przez obywatela
Bezpieczeństwo	Odsetek systemów wysokiego ryzyka objętych testami odporności, procedurą zarządzania incydentami i regularnym audytem
Dane osobowe	Liczba sektorowych ocen skutków dla ochrony danych (DPIA) dla projektów AI skonsultowanych z PUODO
Środowisko	Obowiązek raportowania śladu energetycznego i środowiskowego publicznych projektów AI
Zamówienia publiczne	Odsetek zamówień publicznych na systemy AI zawierających klauzule dotyczące godnej zaufania AI

Takie wskaźniki lepiej odpowiadałyby celowi opiniowanego dokumentu, ponieważ pozwalałyby ocenić nie tylko liczbę podjętych działań, lecz także ich realny wpływ na bezpieczeństwo, przejrzystość, odpowiedzialność i ochronę praw jednostki.

4.4. PODSUMOWANIE

Rozdział „Sztuczna inteligencja godna zaufania” należy ocenić pozytywnie co do przyjętego kierunku aksjologicznego. Projekt trafnie wskazuje, że rozwój AI powinien być powiązany z ochroną praw człowieka, prywatności i danych osobowych, interesów konsumentów, konkurencji, praw twórców oraz potrzeb grup szczególnie wrażliwych. Na aprobatę zasługuje również dostrzeżenie znaczenia dezinformacji, transparentności, piaskownic regulacyjnych oraz edukacji w zakresie odpowiedzialnego korzystania z AI.

Jednocześnie rozdział ten wymaga istotnego uporządkowania i doprecyzowania. W obecnym brzmieniu ma on w znacznej mierze charakter deklaratywny i nie przekłada w wystarczającym stopniu pojęcia „sztucznej inteligencji godnej zaufania” na konkretne mechanizmy wdrożeniowe.

Rekomendowane jest oparcie rozdziału na spójnej, horyzontalnej matrycy wymogów godnej zaufania AI, zgodnej z Wytycznymi HLEG, obejmującej zgodność z prawem, etykę i solidność systemów AI oraz siedem wymogów operacyjnych: nadzór człowieka, bezpieczeństwo techniczne, ochronę prywatności i zarządzanie danymi, przejrzystość,

niedyskryminację i sprawiedliwość, dobrostan społeczny i środowiskowy oraz odpowiedzialność.

W szczególności opiniowany rozdział powinien zostać uzupełniony o kwestie obecnie pominięte albo potraktowane zbyt ogólnie: systemy agentowe i agentic AI, realne granice wyjaśnialności modeli AI, przeciwdziałanie dyskryminacji algorytmicznej, wpływ AI na dobrostan człowieka i środowisko, krajowy system nadzoru nad AI, praktyczne mechanizmy stosowania RODO i AI Act, szersze ujęcie ryzyk konsumenckich oraz ostrożniejsze, neutralne technologicznie podejście do twórczości powstającej z udziałem AI.

Za konieczne należy uznać także wzmocnienie warstwy operacyjnej rozdziału. Wydaje się, że opiniowany rozdział powinien zostać przebudowany tak, aby nie tylko deklarował rozwój AI godnej zaufania, ale wskazywał, **jak państwo zamierza ten cel osiągnąć, kto będzie za to odpowiedzialny i według jakich kryteriów realizacja celu będzie oceniana**. Dopiero takie ujęcie pozwoli nadać pojęciu „godnej zaufania AI" praktyczne znaczenie i zwiększy spójność Polityki z europejskim podejściem regulacyjnym oraz strategicznym.

dr hab. Marlena Sakowska-Baryła, prof. UŁ, radczyni prawna
dr Gabriela Bar, radczyni prawna
dr Wojciech Lamik, radca prawny