Ile kontroli musi zachować urzędnik nad zautomatyzowanymi decyzjami?

Posted on 2026-11-02

W sierpniowym numerze „Europejskiego Przeglądu Sądowego” ukazał się artykuł Prezesa UODO Mirosława Wróblewskiego zatytułowany „Interwencja człowieka jako granica automatyzacji decyzji administracyjnych – relacje między art. 22 RODO a art. 14 i art. 86 aktu w sprawie sztucznej inteligencji”. Poniżej przedstawiamy główne tezy tego tekstu wraz z kilkoma cytatami (za pozwoleniem redakcji EPS – polecamy jednak Czytelnikom zapoznać się z oryginałem).

Artykuł skupia się na analizie relacji między art. 22 ogólnego rozporządzenia o ochronie danych osobowych (RODO) a art. 14 i art. 86 AI Act. Autor uzasadnia, że systemowy nadzór człowieka nie może zastąpić rzeczywistej i znaczącej interwencji ludzkiej w konkretnych sprawach. Z przytoczonych przepisów wyłania się bowiem model odpowiedzialności złożony z trzech elementów: projektowego (za działanie systemu AI), operacyjnego (za sposób korzystania z niego) i decyzyjnego (za konkretne rozstrzygnięcia).

Pozwala to wskazać granicę automatyzacji, której przekroczenie prowadziłoby do niedopuszczalnej depersonalizacji aktów stosowania prawa.

Spór sprowadza się zatem do tego, czy w demokratycznym państwie prawnym można akceptować decyzję wydawaną automatycznie, bez realnego udziału urzędnika; taką której treść tworzy AI, a człowiek jedynie ją zatwierdza. Nie jest to więc wyłącznie spór o metodę działania i o to, czy algorytm działa sprawniej od człowieka. Zarówno polski Trybunał Konstytucyjny, jak i Trybunał Sprawiedliwości UE w swoim orzecznictwie wskazują, że taki czysto formalny nadzór człowieka jest niewystarczający.

Zakaz podlegania automatycznym decyzjom

W tym kontekście kluczowe znaczenie ma art. 22 RODO, który przyznaje każdemu prawo by nie podlegać decyzjom opartym wyłącznie na zautomatyzowanym przetwarzaniu danych (w tym profilowaniu), jeżeli decyzja ta wywołuje dla tej osoby skutki prawne lub podobnie na nią wpływa.

W ustępie 3 tego przepisu przewidziano odpowiednie gwarancje praw osób podlegających takiemu automatycznemu przetwarzaniu, jeśli odbywa się ono na podstawie umowy lub zgody – minimalnym zabezpieczeniem jest tu właśnie możliwość interwencji ze strony człowieka. Jest to więc w gruncie rzeczy prawo do odwołania od decyzji administratora podjętej przez algorytm.

„W odniesieniu do administracji publicznej art. 22 RODO pełni funkcję dalej idącą: wyznacza granicę dopuszczalnego przeniesienia procesu stosowania prawa na system automatyczny. W tym sensie jest nie tylko instrumentem ochrony danych osobowych, lecz także normą chroniącą konstytucyjne minimum odpowiedzialności człowieka za wykonywanie władzy publicznej. Dlatego punktem wyjścia niniejszej analizy jest teza, że art. 22 RODO pozostaje centralną normą europejskiego prawa ochrony jednostki i europejskiego prawa

konstytucyjnego w kontekście i w warunkach postępującej automatyzacji decyzji administracyjnych” – pisze Mirosław Wróblewski.

Tymczasem art. 14 i art. 86 AI Act przenosi tę zasadę z momentu podejmowania decyzji już na wcześniejszy etap projektowania systemu. Dlatego wszystkie te trzy przepisy należy rozpatrywać łącznie jako elementy tej samej normy prawnej.

Automatyczne decyzje jako problem konstytucyjny

W dalszej części autor czyni rozróżnienie między informatyzacją administracji, czyli usprawnieniem wykonywanych przez nią czynności technicznych (obieg dokumentów, prowadzenie rejestrów, komunikacja czy analiza dużych zbiorów danych) a automatyzacją kompetencji władczych, czyli udziału systemu informatycznego w akcie stosowania prawa (klasyfikacja obywatela, ocena ryzyka, scoring lub rekomendacja rozstrzygnięcia).

Ten drugi przypadek rodzi problem z punktu widzenia konstytucyjnych praw obywatela, gdyż błąd algorytmu może skutkować błędnymi decyzjami co do konkretnych osób, nieraz na skalę masową. Ponadto, powoduje to rozmycie odpowiedzialności za decyzję, zwłaszcza w kontekście zjawiska black box (czyli braku pełnej wiedzy o tym jak AI dochodzi do konkretnego rozstrzygnięcia). Można bowiem powoływać się na działanie systemu, producenta lub dane treningowe jako przyczyny błędnej decyzji. Osłabia to także gwarancje proceduralne i utrudnia realną kontrolę sądową.

Sytuacja taka może naruszać zarówno Kartę Praw Podstawowych Unii Europejskiej, w szczególności jej art. 41 (prawo do dobrej administracji) oraz art. 47 (prawo do skutecznej ochrony sądowej), a także samą zasadę demokratycznego państwa prawnego z art. 2

Konstytucji RP. By tego uniknąć, nie wystarczy zapewnić samej obecności człowieka przy podejmowaniu decyzji. Obecność ta musi być jeszcze znacząca i wpisana w szerszą konstrukcję rozliczalności administratora.

Art. 22 RODO jako zasada państwa prawa

Mirosław Wróblewski zwraca uwagę na szczególny charakter art. 22 RODO, który chroni znacznie więcej niż tylko prywatność: daje obywatelowi sprawczość w procesie decyzyjnym. Odejście od tej zasady przez administrację publiczną jest możliwe tylko jeśli decyzja taka jest dozwolona prawem UE lub państwa członkowskiego, które przewiduje odpowiednie środki ochrony praw, wolności i uzasadnionych interesów jednostki (art. 22 ust. 2 lit. b RODO).

Nie wystarczy tu więc ogólna kompetencja do rozpoznania spraw danego rodzaju, ani podstawa prawna do przetwarzania danych. Przepis musi dawać podstawę do podjęcia zautomatyzowanej decyzji. W przeciwnym wypadku organ sam przenosiłby swoje kompetencje na system AI, co naruszałoby zasadę legalizmu.

„Należy podkreślić, że motyw 71 RODO potwierdza tę wykładnię. Wskazuje on, że zabezpieczenia powinny obejmować co najmniej prawo do interwencji człowieka, prawo do przedstawienia własnego stanowiska oraz prawo do zakwestionowania decyzji. Wbrew niekiedy spotykanemu uproszczeniu nie są to trzy luźne gwarancje proceduralne. Tworzą one jedną strukturę: obywatel musi wiedzieć, że decyzja została zautomatyzowana, musi móc przedstawić argumenty nieuwzględnione przez system, musi mieć dostęp do człowieka zdolnego do rzeczywistej rewizji wyniku oraz musi móc zakwestionować decyzję przed organem lub sądem. Dlatego art. 22 RODO pozostaje fundamentalną normą państwa prawa

także po wejściu w życie AIA” – pisze Prezes UODO.

Orzecznictwo TSUE w kwestii automatyzacji decyzji

W kolejnej części tekstu autor omawia podejście Trybunału Sprawiedliwości Unii Europejskiej do analizowanego problemu. Przede wszystkim przywołuje więc wyrok C-634/21, SCHUFA Holding, w którym trybunał uznał, że wyliczenie przez agencję informacji kredytowej prawdopodobieństwa danej osoby do spłaty zobowiązań już może stanowić formę „zautomatyzowanego podejmowania decyzji”, jeśli podmiot udzielający kredytu w znacznym stopniu opiera się na tej wartości przy zawieraniu umowy. W ten sposób TSUE odchodzi od formalnego rozumienia decyzji – jeśli jej rezultat zależy od wyniku wygenerowanego automatycznie (nawet jedynie w formie rekomendacji), to wynik ten realnie determinował decyzję podjętą przez inny podmiot.

Innym ważnym orzeczeniem jest wyrok TSUE z 27.02.2025 r., C-203/22, CK przeciwko Magistrat der Stadt Wien, przy udziale: Dun & Bradstreet Austria GmbH30. Zdefiniowano w nim bowiem pojęcie „znaczących informacji o zasadach podejmowania decyzji” których może domagać się osoba, której dane dotyczą. Obejmują one wyjaśnienie – w zwięzłej i przejrzystej formie – procedury i zasad zastosowanych do przetwarzania jej danych w celu osiągnięcia określonego rezultatu. Wykracza to więc poza obowiązki informacyjne wynikające z RODO – wyjaśnienie procesu decyzyjnego, w interpretacji TSUE, staje się warunkiem zachowania gwarancji praw jednostki. Organ administracji nie może więc stwierdzić jedynie, że decyzja została podjęta na podstawie wskazań systemu AI – musi jeszcze wskazać, jakie konkretnie dane, kryteria i mechanizmy doprowadziły do takiego, a nie innego rozstrzygnięcia.

„Oba te wyroki przesunęły ciężar wykładni art. 22 RODO z pytania „czy człowiek formalnie uczestniczył w procesie?" na pytanie „kto rzeczywiście ukształtował decyzję?". To przesunięcie ma fundamentalne znaczenie. Udział człowieka w procesie decyzyjnym nabiera w tym świetle znaczenia dopiero wtedy, gdy rozumie on wynik działania systemu AI, może go zakwestionować i ma rzeczywistą kompetencję do podjęcia odmiennego rozstrzygnięcia. W przeciwnym razie mamy do czynienia jedynie z administracyjnym odpowiednikiem „rubber stamping", fasadowością i iluzorycznością rzeczywistej decyzyjności, nieodpowiadającej standardom prawa ochrony danych osobowych ani wymaganiom konstytucyjnym państwa prawa" – podsumowuje tę część Mirosław Wróblewski.

Rzeczywista interwencja człowieka – czyli jaka?

Mirosław Wróblewski podkreśla, że kluczowe nie jest to, czy człowiek formalnie bierze udział w podejmowaniu decyzji, ale czy ten udział ma istotny wpływ na kształt tej decyzji. W przeciwnym radzie zatwierdzenie czy podpisanie decyzji przez urzędnika miałoby charakter czysto ceremonialny, co pozbawiałoby art. 22 RODO realnego znaczenia. „Znacząca interwencja człowieka" (*meaningful human intervention*) wyznaczająca właśnie różnicę między decyzją faktycznie podjętą przez urzędnika (przy wsparciu AI), a decyzją jedynie przez człowieka poświadczoną. Nie ma przy tym znaczenia, czy udział ten ma miejsce w trakcie działania algorytmu, czy już po nim.

Autor wyróżnia cztery minimalne warunki które należy spełnić, by można było mówić, że interwencja człowieka była rzeczywista.

Po pierwsze, osoba interweniująca musi mieć odpowiednie kompetencje prawne i organizacyjne do zmiany wyniku wygenerowanego przez system. Nie będzie więc taką

interwencja, w której człowiek może potwierdzić rekomendację algorytmu, ale już nie kwestionować ją, uchylić lub zastąpić swoją. Drugi warunek to kompetencja poznawcza, czyli zdolność do zrozumienia jakie dane, kryteria i założenia doprowadziły do określonego wyniku. Nie oznacza to rozumienia kodu źródłowego, lecz możliwość weryfikacji, czy rezultat odpowiada stanowi faktycznemu, normie prawnej i celowi postępowania.

Element trzeci to czasowa relewantność interwencji. Jeśli bowiem człowiek wkracza dopiero, gdy wywoła już ona skutki prawne, to główna gwarancja z art. 22 RODO traci na sile. Wreszcie po czwarte, interwencja taka musi być udokumentowana, gdyż tylko wówczas można ustalić, czy człowiek rzeczywiście brał udział w ocenianiu sprawy czy jedynie zatwierdził wynik systemu. To właśnie ten ostatni aspekt jest szczególnie ważny w administracji publicznej, ze względu na konieczność uzasadniania decyzji. Uzasadnienie takie bowiem nie tylko informuje o motywach rozstrzygnięcia, ale też pozwala na kontrolę instancyjną i sądową. Organ, który korzysta z narzędzia AI, musi dokumentować interwencję człowieka jako odpowiednik uzasadnienia w przypadku decyzji podejmowanych bez AI. Bez takiej dokumentacji nie da się odróżnić aktu wydanego przez urzędnika od takiego, który jest przez niego jedynie formalnie podpisany.

„W tym miejscu ujawnia się słabość rozwiązań opartych wyłącznie na założeniu, że człowiek „zawsze może” zmienić wynik działania systemu. Sama możliwość ingerencji nie przesądza o jej realności. Literatura dotycząca automation bias pokazuje, że osoby korzystające z systemów automatycznych mają skłonność do nadmiernego polegania na rekomendacjach algorytmicznych, zwłaszcza gdy system przedstawiany jest jako obiektywny, naukowy albo statystycznie bardziej trafny od oceny ludzkiej. (…) Zwraca się uwagę, że sam proces interwencji powinien być projektowany z uwzględnieniem automatycznego zaufania do systemu, rutyny, obciążenia pracą i kompetencji osoby dokonującej oceny. W administracji publicznej problem ten jest szczególnie poważny, ponieważ jeżeli urzędnik rozpatruje masowo tysiące spraw, ma on w oczywisty sposób ograniczony czas na analizę każdej z nich, nie zna dokładnie założeń modelu i wie, że odejście od rekomendacji wymaga dodatkowego

uzasadnienia, to formalna prawna kompetencja do zmiany decyzji nie gwarantuje rzeczywistej kontroli” – zwraca uwagę Mirosław Wróblewski.

Następnie autor omawia rozwiązanie przyjęte w tej sprawie we Francji, które jest zgodne także z prawem ochrony danych osobowych. Artykuł L311-3-1 Code des relations entre le public et l’administration wymaga bowiem, aby indywidualna decyzja podjęta na podstawie przetwarzania algorytmicznego zawierała wyraźną informację o stosowaniu algorytmów, a na żądanie zainteresowanego należy udostępnić mu reguły definiujące przetwarzanie danych i główne cechy wdrożenia systemu. Zasady ochronione w tym przepisie są podobne do tych z art. 22 RODO: obywatel nie może pozostawać bez wiedzy, że technologia współkształtowała jego sytuację prawną.

Co na to AI Act – omówienie art. 14 i art. 86

Przytoczony wyżej Artykuł 14 AI Act wprowadza wymóg, żeby systemy AI wysokiego ryzyka projektować i rozwijać w sposób, który umożliwi ich skuteczne nadzorowanie przez człowieka (w tym np. poprzez przystępny interfejs). Ma to na celu zapobieganie ryzykom dla zdrowia, bezpieczeństwa lub praw podstawowych albo przynajmniej ich minimalizowanie. Przepis ten nie ustanawia jednak prawa dla jednostki i nie koncentruje się na używaniu AI wobec konkretnej osoby. Autor podkreśla jednak, że między tym przepisem a art. 22 RODO istnieje relacja funkcjonalna.

„Dobrze zaprojektowany nadzór człowieka może stworzyć warunki umożliwiające rzeczywistą interwencję człowieka w rozumieniu art. 22 RODO. Może zapewnić szkolenie personelu, interfejs pozwalający zrozumieć wynik systemu, możliwość wstrzymania działania modelu AI w razie takiej potrzeby. Nie przesądza jednak sam przez się, że w konkretnej

sprawie doszło do meaningful human intervention. (…) Nadzór człowieka z art. 14 AI Act można porównać do stworzenia infrastruktury odpowiedzialności. Organ administracji publicznej ma dysponować systemem, personelem i procedurami, które pozwalają człowiekowi rozumieć, monitorować i w razie potrzeby zakwestionować działanie systemu AI. Jednak decyzja administracyjna nadal wymaga przyjęcia odpowiedzialności decyzyjnej przez konkretnego decydenta" – czytamy w tekście.

A odpowiedzialność ta powstaje dopiero wtedy, gdy osoba działająca z upoważnienia organu rzeczywiście dokonuje własnej oceny i przyjmuje odpowiedzialność za wynik. Dlatego AI Act nie może być interpretowany jako przejście od modelu zakazu z art. 22 RODO do modelu dopuszczalności decyzji algorytmicznych pod warunkiem nadzoru. Lepsze wydaje się ujęcie warstwowe: akt o sztucznej inteligencji reguluje bezpieczeństwo i organizację systemu algorytmicznego, a RODO – granice poddania jednostki decyzji automatycznej.

Z kolei art. 86 AI Act przewiduje prawo osoby dotkniętej decyzją podjętą na podstawie wyniku systemu AI wysokiego ryzyka wymienionego w załączniku III do tego aktu, jeżeli wywołuje ona skutki prawne lub podobnie istotnie wpływa na tę osobę – do uzyskania jasnego i znaczącego wyjaśnienia roli systemu AI w procedurze decyzyjnej oraz głównych elementów podjętej decyzji. W pewnym sensie przepis ten stanowi więc pomost pomiędzy art. 22 RODO a art. 14 AI Act. Z jednej strony nawiązuje bowiem nawiązuje do systemowego obowiązku przejrzystości i nadzoru nad sztuczną inteligencją, a z drugiej umożliwia jednostce zakwestionowanie decyzji i uruchomienie odpowiedzialności człowieka.

Jak zwraca uwagę Mirosław Wróblewski, holenderski organ ochrony danych wprost zaś wskazał, że wyjaśnienie ma umożliwić sprawdzenie decyzji i jej zakwestionowanie – skorzystanie z tego prawa jest zaś trudne bez odpowiednich informacji. Autor podkreśla przy tym, że w administracji publicznej prawo do wyjaśnienia ma dodatkowy wymiar, a uzasadnienie decyzji administracyjnej jest klasycznym instrumentem kontroli legalności działania organu władzy publicznej.

Ile kontroli musi zachować urzędnik nad zautomatyzowanymi decyzjami?

„Wyjaśnienie nie powinno być bowiem ani czysto techniczne, ani czysto formalne. Powinno być prawniczo użyteczne: umożliwiać adresatowi decyzji i sądowi ocenę, czy organ pozostał rzeczywistym decydentem. W tym sensie art. 86 AI Act jest normą łączącą transparentność systemu AI z odpowiedzialnością organu administracji publicznej. Bez art. 86 AI Act nadzór z art. 14 AI Act mógłby pozostać wewnętrzną procedurą niewidoczną dla obywatela. Bez art. 22 RODO wyjaśnienie z art. 86 AI Act mogłoby natomiast zostać zredukowane do informacji o technologii, bez realnej możliwości żądania ludzkiej rewizji rozstrzygnięcia. Dopiero łączne zastosowanie tych przepisów obu unijnych rozporządzeń tworzy pełny konstytucyjny model ochrony prawnej godny demokratycznego państwa prawa: system AI musi być nadzorowalny, decyzja musi być wyjaśnialna, a odpowiedzialność za jej treść musi pozostać *człowiecza*" – podsumowuje Mirosław Wróblewski.