sieć obywatelska
WATCHDOG

European
Artificial Intelligence
& Society Fund

KTO, DO CZEGO I JAK KORZYSTA Z AI?

STAN NA NA 8.04.2026

Działamy na rzecz Jawności!

Kto, do czego i jak korzysta z AI?

Autor: Michał Zemełka
Korekta i redakcja: Martyna Bójko
Skład: Autofocus

sieć obywatelska
WATCHDOG

Spis treści

Wprowadzenie

Celem badania było ustalenie, które spośród instytucji realizujących zadania publiczne korzystają z systemów opartych na sztucznej inteligencji (dalej: systemów AI), w jakim celu to robią i na ile są w tym przejrzyste. Chcieliśmy też sprawdzić, czy instytucje samodzielnie przyjmują jakieś wewnętrzne regulacje dotyczące korzystania przez pracowników z ogólnodostępnych narzędzi tego typu, a jeśli tak, to na co zwracają w nich uwagę. W związku z powyższym latem 2025 roku wysłaliśmy wnioski o informację publiczną do blisko pięciu i pół tysiąca różnych podmiotów zobowiązanych.

Wniosek o informację

Na podstawie ustawy o dostępie do informacji publicznej zwracamy się z wnioskiem o udzielenie informacji w zakresie korzystania przez organ z systemów opartych na sztucznej inteligencji (w rozumieniu definicji zawartej w rozporządzeniu Parlamentu Europejskiego i Rady (UE) 2024/1689 z dnia 13 czerwca 2024 r.).

W związku z powyższym wnosimy o udzielenie odpowiedzi na poniższe pytania:

Pytanie 1. Czy w organie lub na jego zlecenie są wykorzystywane **systemy informatyczne**, które **automatycznie** przetwarzają dane wejściowe (np. tekst, obraz, dźwięk, dane liczbowe), aby **generować** raporty, wnioski, przewidywania, rekomendacje lub decyzje, które **mogą wpływać** na działania ludzi lub instytucji?

W przypadku odpowiedzi twierdzącej wnosimy o podanie listy takich systemów z nazwą i funkcją.

Podpowiedź: systemy te mogą obejmować między innymi:

- *automatyczne systemy oceny lub klasyfikacji (np. kandydatów do pracy, wniosków, pacjentów, uczniów),*
- *narzędzia przewidujące ryzyka lub zachowania,*
- *systemy analizujące dane w celu wspomagania decyzji,*
- *narzędzia do rozpoznawania tekstu, mowy, obrazów, filmów (np. analiza nagrań z kamer monitoringu),*
- *chatboty lub wirtualni asystenci,*
- *inne narzędzia wykorzystujące modele „uczenia maszynowego" lub „sztucznej inteligencji".*

Pytanie 2. Czy w organie istnieją **wewnętrzne zasady lub zalecenia** dotyczące **korzystania** przez pracowników z ogólnodostępnych narzędzi **opartych na sztucznej inteligencji** – takich jak czaty i asystenci tekstowi (np. ChatGPT, NotebookLM), generatory filmów (np. Sora), muzyki (np. Suno AI) oraz inne narzędzia wykorzystujące sztuczną inteligencję generatywną?

W przypadku odpowiedzi twierdzącej wnosimy o podanie listy takich regulacji wraz z datą rozpoczęcia obowiązywania.

Kogo pytaliśmy?

Wnioski wysłaliśmy do **5480 instytucji**, w tym:

do wszystkich (na moment badania):

- 2479 gmin (w tym miast na prawach powiatu)
- 314 starostw powiatowych
- 16 marszałków
- 16 wojewodów
- 19 ministerstw
- 16 komend wojewódzkich policji
- 1 komendy stołecznej policji
- 319 sądów rejonowych
- 47 sądów okręgowych
- 11 sądów apelacyjnych
- 16 wojewódzkich sądów administracyjnych
- 46 prokuratur okręgowych
- 16 wojewódzkich inspektorów nadzoru budowlanego
- 16 wojewódzkich inspektoratów ochrony roślin i nasiennictwa
- 16 wojewódzkich inspektoratów ochrony środowiska
- 16 wojewódzkich inspektoratów transportu drogowego
- 16 kuratoriów oświaty
- 16 regionalnych dyrekcji ochrony środowiska
- 17 regionalnych dyrekcji Lasów Państwowych
- 16 regionalnych izb obrachunkowych
- 16 oddziałów Narodowego Funduszu Zdrowia
- 23 parków narodowych
- 341 powiatowych urzędów pracy (w tym czterech miejskich i dwóch grodzkich)
- 305 powiatowych inspektoratów weterynarii
- 11 regionalnych zarządów gospodarki wodnej
- 49 samorządowych kolegiów odwoławczych
- 63 straży gminnych (z czego 3 okazały się zlikwidowane)
- 380 straży miejskich
- 135 uczelni publicznych
- 16 uczelni kościelnych
- 9 ośrodków dla cudzoziemców

i do wybranych:

- 142 organów na szczeblu centralnym
- 14 wojewódzkich urzędów skarbowych
- 16 izb administracji skarbowej
- 203 placówek opieki zdrowotnej (szpitali, lecznic i innych)
- 243 spółek komunalnych
- 4 samorządowych jednostek budżetowych
- 10 zakładów budżetowych
- 70 spółek skarbu państwa

Przejrzystość, czyli jak odpowiadały podmioty zobowiązane?

Ministerstwa

Na początek warto przywołać fragment odpowiedzi Ministerstwa Obrony Narodowej (dalej MON), które deklaruje, że *rozwiązania AI stosowane są w wielu obszarach i systemach, o różnych **klauzulach niejawności***. Oznacza to, że wojsko wykorzystuje technologie oparte na AI, ale nie ujawni szczegółów. Konstrukcja klauzul niejawności w Polsce nie wymaga podawania daty obowiązywania (teoretycznie powinny być rewidowane co 5 lat), a ponadto nie posiadamy niezależnego organu, który mógłby w uproszczonym trybie (bez konieczności korzystania z sądowej ścieżki odwoławczej) potwierdzić zasadność nałożenia danej klauzuli na wnioskowane informacje. Fakt, że MON przyjęło strategię rozwoju sztucznej inteligencji do roku 2039 i jest to dokument publicznie dostępny, niewiele zmienia a wręcz przeciwnie, zapala czerwone lampki, bo napisano tam między innymi, że *wykorzystanie systemów sztucznej inteligencji w działaniach militarnych będzie dotyczyć m.in. **autonomicznych systemów bojowych**, które będą mogły prowadzić operacje bez bezpośredniego zaangażowania człowieka* (patrz str. 11 strategii). I chociaż ryzyka związane z wojskowymi zastosowaniami sztucznej inteligencji wydają się być w tym dokumencie dosyć wyczerpująco zdiagnozowane (patrz str. 16), to brak kontroli w tym obszarze jest niepokojący.

Na nasz wniosek nie odpowiedziały Ministerstwo Kultury i Dziedzictwa Narodowego oraz Ministerstwo Przemysłu (co akurat dziwnym nie jest, bo chwilę po naszym wniosku zostało zlikwidowane). W przypadku kilku innych nie wystarczyło wysłanie wniosku zwykłym mailem, trzeba było dosłać wersję elektronicznie podpisaną. Pozostałe odpowiadały mniej lub bardziej wyczerpująco. Dziesięć z siedemnastu resortów zadeklarowało, że nie korzysta z systemów AI. Odpowiedzi z siedmiu ministerstw były na tyle interesujące, że postanowiliśmy bliżej zbadać temat w kolejnej fazie projektu i szerzej je opisać.

Sądy powszechne

Wiele spośród sądów rejonowych i okręgowych odesłało nas do właściwych dla danego obszaru sądów apelacyjnych, które są odpowiedzialne za ich obsługę informatyczną (zgodnie z art. 31a§1a ustawy – *Prawo o ustroju sądów powszechnych)*. Duża część zdecydowała się jednak wyczerpująco odpowiedzieć na nasze pytania, niekiedy zaznaczając, że mogą nie mieć pełnej wiedzy. W wielu odpowiedziach wskazano jakieś rozwiązania oparte na AI, wśród nich chatbot *Wirtualny asystent* czy transkrypcję mowy na tekst. Jednocześnie sądy apelacyjne gremialnie zaprzeczyły, by *w organie lub na jego zlecenie* wykorzystywano systemy AI w rozumieniu definicji z AI Act. Pozostaje zatem na ten moment kwestią otwartą, czy niektóre sądy rejonowe i okręgowe wdrażają jakieś rozwiązania informatyczne na własną rękę, czy raczej interpretacja systemu AI przez sądy apelacyjne jest wygodnie wąska. A może sposób zadania pytania we wniosku pozwolił sądom apelacyjnym ograniczyć się w odpowiedzi *stricte* do siebie. Warto przy tym zauważyć, że w naszym pytaniu w zasadzie mieściło się wykorzystanie zewnętrznych narzędzi AI przez pracowników (*czy w organie […] **są wykorzystywane** systemy informatyczne…*), a w tym zakresie odsyłanie nas do sądu apelacyjnego nie miało sensu.

Prokuratury

Niemal wszystkie prokuratury okręgowe (40 z 44) zadeklarowały, że nie korzystają z żadnych systemów AI. Wyjątkiem okazały się Prokuratura Okręgowa Warszawa-Praga i Prokuratura Okręgowa w Łodzi, które zasygnalizowały korzystanie z narzędzi wspomaganych AI (przy jednoczesnym braku wewnętrznych regulacji tego obszaru). Z kolei Prokuratura Okręgowa w Radomiu oraz Prokuratura Okręgowa we Włocławku poinformowały, że administratorem sieci komputerowej wszystkich powszechnych jednostek organizacyjnych jest Prokuratura Krajowa (która odżegnuje się od wykorzystywania AI i nie posiada wewnętrznych regulacji).

Ośrodki dla cudzoziemców

Żaden z ośrodków dla cudzoziemców nie odpowiedział na nasz wniosek. Jedyne, co otrzymaliśmy, to automatyczna zwrotka z Ośrodka dla Cudzoziemców w Bezwoli o odczytaniu wiadomości. Ośrodek dla Cudzoziemców w Łukowie usunął wiadomość bez czytania.

Niska responsywność

Jest jeszcze kilka typów instytucji, które odpowiadają na wnioski wyraźnie niechętnie, poniżej średniej (która w przypadku tego konkretnego badania wyniosła trzy podmioty na cztery zapytane). Zazwyczaj ma to jakieś - mniej lub bardziej sensowne - wytłumaczenie, jak w przypadku straży miejskich i gminnych, które w większości przypadków stwierdzały, że odpowiedź w ich imieniu została *implicite* udzielona przez urząd, któremu podlegają, więc nie będą się powtarzać.

Z kolei przypadkiem, którego raczej trudno bronić, są placówki medyczne, odpowiadające w zaledwie jednym przypadku na trzy. Ponadto niektóre spośród instytucji medycznych niebędących szpitalami, jak np. Agencja Oceny Technologii Medycznych i Taryfikacji czy Centrum e-Zdrowia również pozostawiały wniosek bez odpowiedzi. Dla równowagi doceńmy odpowiedzi z Agencji Badań Medycznych i Centrum Medycznego Kształcenia Podyplomowego, które wskazały zewnętrzne narzędzia, takie jak asystent AI.

Statystyka odpowiedzi według typu instytucji

Typ instytucji	Liczba wniosków	Liczba odpowiedzi	Odsetek odpowiedzi
Komendy wojewódzkie i stołeczna policji	17	17	100%
Sądy administracyjne	17	17	100%
Wojewodowie	16	16	100%
Regionalne izby obrachunkowe	16	16	100%
Sądy apelacyjne	11	11	100%
Prokuratury okręgowe	46	44	96%
Wojewódzkie inspektoraty ochrony środowiska	16	15	94%
Regionalne dyrekcje ochrony środowiska	16	15	94%
Ministerstwa	19	17	89%
Kuratoria oświaty	16	14	88%
Parki narodowe	23	20	87%
Miasta na prawach powiatu	66	56	85%
Starostwa powiatowe	314	265	84%
Uczelnie publiczne	135	113	84%
Regionalne Dyrekcje Lasów Państwowych	17	14	82%
Samorządy gminne	2479	2027	82%
Marszałkowie	16	13	81%
Sądy rejonowe	319	256	80%
Samorządowe kolegia odwoławcze	49	39	80%
Powiatowe urzędy pracy	341	268	79%
Organy centralne	142	111	78%
Powiatowe inspektoraty weterynarii	305	233	76%
Wojewódzcy inspektorzy nadzoru budowlanego	16	12	75%
Wojewódzkie inspektoraty ochrony roślin	16	12	75%
Oddziały NFZ	16	12	75%
Sądy okręgowe	47	35	74%

Typ instytucji	Liczba wniosków	Liczba odpowiedzi	Odsetek odpowiedzi
Wojewódzkie inspektoraty transportu drogowego	16	11	69%
Samorządowe jednostki budżetowe	4	2	50%
Spółki komunalne	243	117	48%
Spółki skarbu państwa	70	32	46%
Straże miejskie	380	159	42%
Zakłady budżetowe	10	4	40%
Regionalne zarządy gospodarki wodnej	11	4	36%
Szpitale	203	71	35%
Izby administracji skarbowej	16	5	31%
Straże gminne	60	18	30%
Uczelnie kościelne	16	4	25%
Wojewódzkie urzędy skarbowe	14	1	7%
Otwarte ośrodki dla cudzoziemców	9	0	0%

Odmowy udzielenia informacji

Choć zdecydowana większość odnotowanych braków to bezczynność instytucji (**półtora tysiąca** podmiotów nie zareagowało na wniosek), odnotowaliśmy również decyzje odmowne.

Przykładowo Urząd Gminy Grodziec, przywołując wyrok NSA (I OSK 7/14), stwierdził, że *dane techniczne dotyczące sposobu funkcjonowania systemu komputerowego informacją publiczną nie są. Stanowią jedynie techniczną sferę funkcjonowania narzędzia, jakim organ posługuje się realizując zadania*. Z kolei dyrektor Izby Administracji Skarbowej we Wrocławiu, nie powołując się na żadne przepisy, po prostu arbitralnie stwierdził, że *Systemy informatyczne i dokumenty wewnętrzne tj. zalecenia służą wyłącznie usprawnieniu działalności podmiotu i poprawie funkcjonowania organu. Tym samym nie stanowią źródła informacji o sprawach publicznych, nawet gdy są one wykorzystywane do realizacji celów publicznych*. Wśród instytucji, które odmówiły, podobnie argumentując, znalazł się również Marszałek Województwa Zachodniopomorskiego.

Bardzo ważna jest też korespondencja z Agencją Bezpieczeństwa Wewnętrznego (dalej ABW), która odmówiła udzielenia informacji w zakresie pierwszego pytania wniosku, ponieważ uznała je za niejawne i chronione prawem. Podkreśliła przy tym, że prawo do informacji publicznej nie jest bezwzględne i może być ograniczone ze względu na bezpieczeństwo państwa, a w tym przypadku żądane dane mają charakter informacji niejawnych, do których dostęp mają tylko osoby uprawnione, a wnioskodawca do nich nie należy. Dodatkowo nawet sama odpowiedź – niezależnie od jej treści – mogłaby zdaniem ABW ujawnić ich metody działania oraz zainteresowania operacyjne, co mogłoby zaszkodzić interesom państwa. Organ wskazuje też, że przepisy szczególne wprost zakazują ujawniania informacji o czynnościach operacyjnych, ich metodach i współpracownikach poza wyjątkowymi sytuacjami, które tutaj nie zachodzą. W konsekwencji ABW uznaje, że nie ma podstaw prawnych do udzielenia odpowiedzi. Odmowa ta i argumentacja jest o tyle ważna, że może świadczyć o wykorzystywaniu sztucznej inteligencji przez służby specjalne. Gdyby systemy AI nie były stosowane wcale lub nie byłyby stosowane w związku z działaniami operacyjnymi, to nie byłoby powodu do odmowy udzielenia takiej informacji.

Niektóre instytucje uznawały, że wnioskowana informacja jest informacją przetworzoną, czyli jest to jakościowo nowa informacja publiczna, która nie istnieje w gotowej formie na dzień otrzymania wniosku, a powstaje poprzez analizę, zestawienie lub obróbkę danych źródłowych przez urząd, czyli wymaga zaangażowania dodatkowych sił, środków i pracy intelektualnej. Tak nasz wniosek zakwalifikował m.in. Wojewódzki Szpital Zespolony w Kielcach czy Akademia Sztuk Pięknych im. Władysława Strzemińskiego w Łodzi. W takiej sytuacji podmiot zobowiązany ma prawo zażądać uzasadnienia szczególnego interesu publicznego. Jednak w naszej ocenie złożony przez nas wniosek dotyczył informacji prostej, dlatego nie odpowiadaliśmy na wezwania o wykazanie szczególnie istotnego interesu publicznego. Należy dodać, że w sytuacji, gdy organ kwalifikuje informację jako przetworzoną, spoczywa na nim obowiązek rozstrzygnięcia, czy wnioskodawca spełnia przesłankę szczególnej istotności dla interesu publicznego, która uzasadniałaby konieczność przetworzenia informacji.

Jeszcze inny przypadek stanowią spółki komunalne, które sięgały po tajemnicę przedsiębiorstwa, jako pretekstu do nieudzielenia informacji. Przykładem jest krakowskie Miejskie Przedsiębiorstwo Wodociągów i Kanalizacji SA, według którego żądane przez nas informacje mają charakter strategiczny, a nieprawidłowe ich wykorzystanie mogłoby spowodować niewspółmierną szkodę dla interesów spółki. Podobnie Przedsiębiorstwo Wodociągów i Kanalizacji Sp. z o.o. w Koninie stwierdziło, że żądane informacje wykraczają poza pojęcie informacji publicznej i stanowią tajemnicę przedsiębiorstwa, a ujawnianie lub wykorzystywanie tajemnicy przedsiębiorstwa jest zabronione. Trudno jednak pominąć, że – gdyby informacje te były dostępne – mogłyby mieć znaczenie nie tylko dla tej jednej spółki, lecz również dla podnoszenia standardów świadczenia usług komunalnych w szerszej skali, np. poprzez ich wykorzystanie przez inne podmioty działające w sektorze. Mimo to spółki komunalne, funkcjonujące w warunkach ograniczonej konkurencji i posiadające uprzywilejowaną pozycję na lokalnym rynku, odmawiają udostępnienia informacji, powołu-

jąc się na konieczność ochrony swojej sytuacji rynkowej, w tym także na argumenty dotyczące konkurencji.

Uchylanie się od odpowiedzi

Część instytucji sektora publicznego stoi na stanowisku, że nie jest podmiotem zobowiązanym do udostępniania informacji, o jakie wnioskujemy. Takim przypadkiem w tym badaniu okazał się między innymi Ubezpieczeniowy Fundusz Gwarancyjny (dalej UFG), który - jak sam przyznaje - *jest osobą prawną wykonującą swoje zadania w ścisłych granicach ustawy z dnia 22 maja 2003 r. o ubezpieczeniach obowiązkowych, Ubezpieczeniowym Funduszu Gwarancyjnym i Polskim Biurze Ubezpieczycieli Komunikacyjnych (t.j. Dz. U. z 2025 r. poz. 367), w której m.in. wymieniono zakres gromadzonych przez Fundusz informacji oraz zamknięty katalog podmiotów, którym UFG może - w określonych sytuacjach - je udostępniać (przepisy zawarte w art. 102 i nast. ww. ustawy)*. Dalej stwierdza, że *informację publiczną stanowi informacja (...) wytworzona lub odnoszona do innych podmiotów wykonujących funkcje publiczne w zakresie wykonywania przez nie zadań władzy publicznej i gospodarowania mieniem komunalnym lub mieniem Skarbu Państwa. Tymczasem UFG, realizując swoje ustawowe obowiązki, nie gospodaruje mieniem komunalnym lub mieniem Skarbu Państwa. Mając na uwadze powyższe, uprzejmie informuję, iż na Funduszu nie ciąży obowiązek udostępniania informacji, o jakie Państwo zawnioskowali*. Zdania tego nie podzielił Wojewódzki Sąd Administracyjny w Warszawie, który rozpatrzył naszą skargę na bezczynność w tej sprawie, zobowiązując UFG do rozpoznania naszego wniosku o informację i zwrotu kosztów postępowania. Jednak Ubezpieczeniowy Fundusz Gwarancyjny złożył skargę kasacyjną od tego wyroku, co oznacza, że sprawa przedłuży się o co najmniej dwa lata.

Innym ciekawym przypadkiem okazała się trochę niespodziewanie Wojskowa Akademia Techniczna, która w odpowiedzi przekazała: *proszę o przesłanie wiadomości w formie pisma ułatwi to przekazanie dokumentu do właściwej jednostki organizacyjnej*. Warto przypomnieć, że wysłanie wniosku o informację to odformalizowana procedura, która dopuszcza wysłanie zapytania mailem, który przecież można przesłać do właściwej jednostki organizacyjnej.

Dyrektor Izby Administracji Skarbowej w Lublinie na pytanie pierwsze odpowiedział, że nie jest właścicielem systemów teleinformatycznych wykorzystywanych w Izbie Administracji Skarbowej w Lublinie i podległych jednostkach organizacyjnych i tym samym nie jest właściwy do udzielenia odpowiedzi w powyższym zakresie. Niestety dyrektor zapomniał dodać, kto tym właścicielem jest. Te same instytucje z Krakowa i Olsztyna wskazały na Ministerstwo Finansów i Centrum Informatyki Resortu Finansów. Niestety Ministerstwo okazało się niezbyt skore do dzielenia się wiedzą w zakresie wykorzystania AI w kontekście skarbowym, do czego dalej wrócimy.

Należy tutaj także odnotować odpowiedzi z urzędów miejskich Gdyni i Kalisza, które odesłały nas do swoich centrów informatyki. Oznacza to, że mogą istnieć urzędy z podobną sytuacją, które jednak po prostu zignorowały nasz wniosek, jako błędnie do nich skierowany.

Zaskakujące i oryginalne było wspomniane już stanowisko Agencji Bezpieczeństwa Wewnętrznego, jeśli chodzi o pytanie dotyczące wewnętrznych regulacji dotyczących korzystania z AI: *informacje, których dotyczy pytanie nie mają charakteru informacji publicznych, zawarte są w dokumentach wewnętrznych i nie podlegają udostępnieniu na zasadach określonych w UDIP*. To jeden z wielu przypadków, w których instytucja publiczna powołuje się na pojęcie dokumentu wewnętrznego, mimo że nie występuje ono w ustawie o dostępie do informacji publicznej.

Część instytucji poinformowała nas, że wnioskowane przez nas dane są wyłączone z jawności inną ustawą. Tak m.in. odpisał urząd miasta Konstancin-Jeziorna:

Wnioskowane informacje zostały wyłączone inną ustawą, bo mogą mieć znaczenie z punktu widzenia bezpieczeństwa systemów teleinformatycznych — w szczególności w kontekście ich konfiguracji, funkcjonalności oraz potencjalnych podatności. W związku z powyższym, na podstawie art. 37 ust. ustawy z dnia 5 lipca 2018 r. o krajowym systemie cyberbezpieczeństwa (Dz.U. z 2023 r. poz. 913), szczegóły dotyczące wykorzystywanych narzędzi, ich nazw, funkcji oraz ewentualnego udziału mechanizmów sztucznej inteligencji w procesach organizacyjnych, zostały uznane za informacje wyłączone z dostępu do informacji publicznej.

Urząd nie odpowiedział także na drugie pytanie, choć pozostawało przecież bez związku z podanym uzasadnieniem odmowy. Złożyliśmy skargę na bezczynność i na skutek wyroku Wojewódzkiego Sądu Administracyjnego z 25 lutego 2026 r., sygn. II SAB/Wa 912/25, który zobowiązał urząd do rozpatrzenia wniosku, uzyskaliśmy odpowiedzi na oba pytania. Dane o które wnioskowaliśmy w pierwszym pytaniu nie zawierały w gruncie rzeczy niczego, co by miało ścisły związek z bezpieczeństwem systemów teleinformatycznych - żadne z trzech wymienionych przez urząd narzędzi AI do tego nie służy (Capcut, Adobe Photoshop, Esesja). Co więcej, istnieje wiele narzędzi wspomagających bezpieczeństwo teleinformatyczne z użyciem rozwiązań wykorzystujących sztuczną inteligencję, co wiemy chociażby z odpowiedzi od innych instytucji.

Zdarzało się, że organ nie udostępniał pełnych informacji w odpowiedzi na wniosek i dopiero pojawiały się one w toku dalszej korespondencji. Tak było z Katowicami, które nie od razu wspomniały o systemie Katowickiego Inteligentnego Systemu Monitoringu i Analizy, w którym *działa komponent odpowiedzialny za analizę materiału wideo z kamer miejskiego monitoringu. (...) Platforma w sposób ciągły analizuje strumień wideo i generuje informacje (metadane) opisujące obrazy z kamery. Na podstawie tych metadanych są generowane alerty dot. zdarzeń, które w systemie zostały zdefiniowane jako potencjalnie niebezpieczne lub wymagają zainteresowania operatora monitoringu. Następnie operator na podstawie analizy alertu podejmuje dalsze czynności związane z jego obsługą*. Miasto wyjaśniło, że pierwotny brak wskazania tego systemu wynikał z faktu, że wiedza o nim jest publicznie dostępna i zamieszczona w BIP-ie urzędu. Taka sytuacja nie zwalnia jednak urzędu z obowiązku odniesienia się w odpowiedzi, choćby zdawkowego, do źródeł publicznych.

Rozmywanie definicji

W debacie publicznej AI Act bywa kojarzony przede wszystkim z najnowszą falą sztucznej inteligencji, zwłaszcza z dużymi modelami językowymi i innymi modelami ogólnego przeznaczenia, takimi jak systemy generujące tekst, obrazy czy kod. Jest to skojarzenie zrozumiałe, bo gwałtowny rozwój takich narzędzi wy-

raźnie wpłynął na ostateczny kształt unijnej regulacji. Nie oznacza to jednak, że AI Act dotyczy wyłącznie tego typu rozwiązań.

Sama definicja „systemu AI" przyjęta w AI Act jest szersza. Może obejmować także mniej spektakularne i znane od lat rozwiązania informatyczne, w tym systemy oparte na uczeniu maszynowym, sieciach neuronowych czy modelach predykcyjnych — o ile nie ograniczają się one do prostego, z góry zaprogramowanego przetwarzania danych, lecz na podstawie danych wejściowych generują przewidywania, rekomendacje lub decyzje mogące wpływać na sytuację ludzi. W praktyce oznacza to, że system używany przez administrację do typowania spraw, oceniania ryzyka, wskazywania nieprawidłowości albo wspierania decyzji wobec obywateli może być istotny z perspektywy AI Act nawet wtedy, gdy nie przypomina popularnego chatbota i nie korzysta z dużego modelu językowego.

To rozróżnienie ma znaczenie dla naszego badania. Jeżeli część urzędów utożsamia AI Act głównie z dużymi modelami językowymi lub generatywną AI, mogła nie zgłosić innych narzędzi analitycznych czy predykcyjnych, które — przynajmniej według szerszej interpretacji definicji systemu AI — również mogą mieścić się w zakresie tej regulacji. W rezultacie z badania mogły wypaść systemy mniej widoczne medialnie, ale potencjalnie bardzo istotne z punktu widzenia obywateli, bo wykorzystywane do choćby częściowo zautomatyzowanego wspierania decyzji administracyjnych.

Możemy się tego domyślać między innymi na podstawie odpowiedzi Zakładu Ubezpieczeń Społecznych (dalej ZUS), który w pierwszym odruchu ujawnił korzystanie z narzędzia informatycznego, opartego na działaniu modelu predykcyjnego służącego systemowej analizie danych z wystawionych zaświadczeń lekarskich o czasowej niezdolności do pracy. Jednak na kolejny wniosek, w którym pytaliśmy o więcej szczegółów wykorzystywanych systemów AI, ww. narzędzie się nie pojawiło. Dopiero, kiedy ponowiliśmy pytania, wprost wskazując, że oczekujemy odpowiedzi w kontekście owego narzędzia do analizy zaświadczeń lekarskich, otrzymaliśmy część odpowiedzi, nadal jednak z zastrzeżeniem, że ZUS jednak nie uważa tego systemu za AI w rozumieniu definicji z AI Act.

Innym podobnym przykładem jest zdefiniowany ustawowo (Ordynacja podatkowa i inne ustawy) system STIR, na temat którego informacji nie chce udzielać ani Ministerstwo Finansów, ani zajmująca się nim Krajowa Izba Rozliczeniowa, a tymczasem z uwagi na funkcje, które realizuje (między innymi wyliczanie wskaźnika ryzyka, będącego podstawą do blokowania kont podatników), trudno uwierzyć, że nie korzysta z technologii mieszczących się w definicji AI z AI Act. Jednocześnie nie możemy tego z całą pewnością wykazać, bo nikt nie chce o tym rozmawiać. I koło się zamyka.

Z drugiej strony wiele urzędów podawało w odpowiedzi takie systemy, jak rozwiązania do transkrypcji nagrań z sesji, systemy rozpoznawania tablic rejestracyjnych, czy nawet funkcje OCR stosowane do zeskanowanych dokumentów. Widać więc szeroką rozpiętość w rozumieniu definicji. Ponadto wiele urzędów zdawało się wymieniać po prostu szereg stosowanych w urzędzie systemów informatycznych, nie rozważając, czy zaliczają się do systemów AI, jakby woleli bezpieczniej podać wszystko, zamiast męczyć się z interpretacją i narażać na jej podważanie.

Podsumowanie (nie)przejrzystości

Istotnie niebezpieczne i szkodliwe zastosowania sztucznej inteligencji w obecnym stanie prawnym i praktycznym w Polsce bez trudu można ukryć za zasłoną milczenia lub klauzulą tajności, uzasadniając to pozbawioną jakichkolwiek szczegółów wzmianką o względach bezpieczeństwa, bądź przyjmując wygodnie wąską interpretację systemu AI. Ponadto obecne polskie władze wydają się nie tylko nie mieć żadnych skrupułów przed militarnym użyciem AI, ale nawet wydają się tym szczycić. A przy tym nie ma powodów sądzić, że wobec własnych obywateli stosowane jest odmienne podejście, chociaż w tym przypadku z oczywistych względów nie ma co liczyć na huczne ogłaszanie tego opinii publicznej.

Dlatego tak ważne jest:

- doprecyzowanie zakresu definicji systemu AI,
- powstanie i utrzymywanie publicznego rejestru systemów (właściwie zastosowań) AI w administracji publicznej,
- oraz wymóg oznaczania decyzji i treści wytworzonych z udziałem AI.

Każdy z tych wątków jest obecnie przedmiotem prac i dyskusji zarówno na forum krajowym, jak i europejskim, w tym z naszym i organizacji partnerskich udziałem.

Kto i do czego wykorzystuje AI?

Z 5480 instytucji, do których wysłaliśmy wniosek, odnotowaliśmy odpowiedzi od 4042 (74%). Spośród otrzymanych odpowiedzi 9% zawierało deklarację używania sztucznej inteligencji, przy czym istnieją istotne statystycznie różnice w adopcji tej technologii między różnymi typami instytucji bądź gmin. Przykładowo średnia dla gmin wyniosła 8%, ale jeśli weźmiemy same miasta na prawach powiatu, to tam jest to 34%. Wyraźnie wyższa od średniej adopcja tej technologii wydaje się cechować też między innymi organy centralne (20%), policję na szczeblu wojewódzkim (przynajmniej 18%) czy starostwa powiatowe (14%).

Patrząc całościowo na jednostki samorządu terytorialnego i konfrontując to z danymi z raportu Ministerstwa Cyfryzacji opublikowanego w październiku 2023 roku, gdzie 5% spośród udzielających odpowiedzi w badaniu zadeklarowało używanie sztucznej inteligencji, to - jak się można było spodziewać - widać wzrost, bo w naszym badaniu odsetek zbiorczy takich JST wyniósł 9%. Warto przy tym zauważyć, że część urzędów podeszła drobiazgowo do tego pytania, wymieniając wśród systemów AI również na przykład oprogramowanie typu OCR.

Na drugim biegunie znajduje się rzesza urzędów zdecydowanie odżegnujących się od korzystania z AI i jednocześnie przyznających się do braku wewnętrznych regulacji w tym zakresie (pytanie 2). Zważywszy na dostępność narzędzi opartych na AI i coraz powszechniejsze ich wykorzystywanie przez zwykłych użytkowników, taka postawa wydaje się nieco bagatelizowaniem sytuacji. Potwierdzeniem tego jest raport NASK-u *AI w e-administracji publicznej*, w którym znajdziemy informację, że w celach zawodowych w ciągu ostatnich 6 miesięcy (czyli z grubsza w drugiej połowie roku 2025) 46% urzędników korzystało z generatywnej sztucznej inteligencji.

W szczególności w powyższym kontekście warto przywołać przypadek gminy Wyryki (woj. lubelskie, pow. włodawski), Jarocina (woj. podkarpackie, pow. niżański) i Pysznicy (woj. podkarpackie, pow. stalowowolski), które *unisono* zaprzeczyły korzystaniu z systemów AI, deklarując jednocześnie brak regulacji w tym zakresie lub nie odpowiadając w ogóle na drugie pytanie. Jednocześnie dysponujemy odmownymi decyzjami wydanymi przez te gminy w czasie od kilku do kilkunastu dni od udzielenia przez nie odpowiedzi na nasz wniosek, które powstały najprawdopodobniej z udziałem systemów AI, ponieważ zawierają halucynowane przepisy oraz inne cechy na to wskazujące, jak nieuzupełnione tagi, np. *gmina [nazwa]*.

Wspomniane już Ministerstwo Obrony Narodowej przyznało się do korzystania z systemów wykorzystujących biometrię i technologię rozpoznawania twarzy. Ponadto w Dowództwie Komponentu Wojsk Obrony Cyberprzestrzeni (DKWOC) wykorzystywane są różne modele AI do realizacji zadań, w tym np. do analizy obrazów satelitarnych czy przetwarzania tekstu. A w Wojsku Polskim powstaje Centrum Implementacji Sztucznej Inteligencji, które zostanie utworzone w ramach Dowództwa Komponentu Wojsk Obrony Cyberprzestrzeni.

Ciekawym przykładem jest model predykcyjny wspomniany przez Zakład Ubezpieczeń Społecznych. Tego typu rozwiązania mogą być użyteczne analitycznie, ale wymagają szczególnej ostrożności: z uwagi na sposób wytwarzania model siłą rzeczy powiela wszelkie błędy i uprzedzenia obecne w danych użytych na etapie treningu, a ponadto jego opinia bywa traktowana przez użytkowników z nadmiernym zaufaniem — nie jako pomocnicza prognoza, lecz jako faktyczna rekomendacja działania. Ryzyko nie polega więc wyłącznie na samej automatyzacji, ale także na tym, jak jej wynik zostanie włączony

w proces decyzyjny instytucji, przy jednoczesnym ryzyku trudnych do wykrycia uprzedzeń. Oddajmy głos ZUS-owi:

W Zakładzie Ubezpieczeń Społecznych (ZUS, Zakład) jest wykorzystywane narzędzie informatyczne oparte o działanie modelu predykcyjnego, służące systemowej analizie danych z wystawionych zaświadczeń lekarskich o czasowej niezdolności do pracy. Narzędzie to wspomaga wykonywanie ustawowo określonych zadań związanych z kontrolą orzecznictwa o czasowej niezdolności do pracy.

W procesie kontroli zaświadczeń lekarskich Zakład wykorzystuje systemową analizę ryzyka (tj. analiza rozszerzonego zakresu danych, które mogą wskazywać na prawdopodobieństwo wystąpienia nieprawidłowości w wystawieniu zaświadczenia lekarskiego). Zaświadczenie lekarskie o czasowej niezdolności do pracy wpływające do ZUS poddawane jest takiej analizie i w jej wyniku określone zostaje prawdopodobieństwo (w postaci wartości scoringowej) wystąpienia nieprawidłowości związanych z jego wystawieniem. Wynik analizy systemowej wykonanej przez narzędzie informatyczne jest udostępniany lekarzom wykonującym zadania związane z kontrolą orzekania o czasowej niezdolności do pracy. Jest to jeden z możliwych elementów uwzględnianych w merytorycznej ocenie stanu faktycznego sprawy, odnoszącej się do przesłanek podjęcia kontroli.

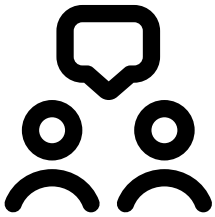
Należy jednak zaznaczyć, że opisywane narzędzie nie typuje zaświadczeń lekarskich do kontroli. Każdorazowo decyzja o podjęciu kontroli zaświadczenia lekarskiego (wytypowanie go do kontroli) podejmowana jest przez głównego lekarza orzecznika po całościowej analizie dostępnych informacji, a w szczególności przesłanek medycznych podjęcia kontroli, w tym także jeśli było to zasadne, po ocenie wskazania narzędzia analitycznego.

Jak wiemy między innymi na podstawie dalszej korespondencji z ZUS-em, toczonej w ramach kolejnej odsłony badania (w związku z czym opiszemy to szerzej w stosownym raporcie), urząd nie uznaje ww. systemu za system AI w rozumieniu AI Act i dlatego nie czuje się zobligowany do stosowania takich mechanizmów zabezpieczających, jak kwalifikacja ryzyka czy ocena skutków dla praw podstawowych (FRIA). Trudniej też w związku z tym uzyskać dostęp do szczegółowych informacji umożliwiających społeczny audyt tego rozwiązania.

Wśród typów instytucji, które zapytaliśmy wybiórczo, znalazły się placówki medyczne. Wnioski wysłaliśmy do 203 z nich, głównie szpitali wojewódzkich. Z tego typu placówek odsetek odpowiedzi był relatywnie niski: odpowiedziało ledwie 71 spośród nich, a więc około 35% (dla porównania w przypadku gmin ten odsetek przekracza zwykle 80%, a w tym badaniu wyniósł 82%). Zaskakująca większość - 66 (93% odpowiedzi) - zanegowała stosowanie systemów sztucznej inteligencji w rozumieniu AI Act, co jest o tyle dziwne, że w zastosowaniach medycznych (np. diagnostyce obrazowej) rozwiązania oparte na sztucznej inteligencji stosowane są od wielu lat i są starsze niż przełom w generatywnej AI. Wydaje się, że może to wynikać z braku świadomości lub niechęci do spełniania ewentualnych nowych obowiązków, np. wynikających z art. 26 i 27 AI Act. Tylko 5 placówek potwierdziło stosowanie AI i były to właśnie obszary związane z diagnostyką obrazową. Dodatkowo jedna z tych odpowiedzi zawierała wzmiankę o voicebocie używanym w procesie rejestracji wizyt pacjentów. W kontekście instytucji medycznych szczególnie wartościowa jest odpowiedź Marszałka Województwa Lubuskiego, który - wykraczając w zasadzie poza ściśle rozumiany zakres wniosku - napisał o realizacji projektu wsparcia diagnostyki i medycyny przez SI (*MedBrain L*), w ramach którego zaimplementowano wytworzone rozwiązania w 3 jednostkach ochrony zdrowia w regionie.

Patrząc zbiorczo na zebrane odpowiedzi, **pomimo badania zakrojonego na tak szeroką skalę, nie udało się zidentyfikować potencjalnie niebezpiecznych zastosowań systemów AI, o których nie wiedzielibyśmy już wcześniej skądinąd** (jak STIR w KIR, czy system typujący podejrzane zwolnienia lekarskie w ZUS). Zamiast tego ujawniły się bariery w dostępie do informacji w tym zakresie, tym silniejsze im potencjalnie bardziej niebezpieczne w kontekście praw człowieka może być zastosowanie AI w danym obszarze (np. militarnym).

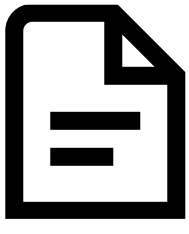
Na podstawie otrzymanych odpowiedzi możemy wyróżnić rozmaite obszary zastosowania AI w polskiej administracji i pozostałych instytucjach realizujących zadania publiczne. Większość z nich wiąże się z wykorzystaniem zewnętrznych narzędzi, często na poziomie zwykłej aplikacji, nie zaś budowaniem własnych rozwiązań, czy stawianiem potężnych systemów mocno opartych na AI. Wyjątki dotyczą organów na szczeblu centralnym, których zasoby pozwalają na sfinansowanie projektów o odpowiednio dużej skali. Oto, jakie obszary zastosowań AI wyłoniły się z otrzymanych odpowiedzi:



Mowa, nagrania i spotkania

Obejmuje wykorzystanie AI do przetwarzania wypowiedzi ustnych i materiałów audio-wideo. Chodzi przede wszystkim o transkrypcję (najczęściej sesji rady), tworzenie napisów, streszczanie spotkań i obrad, przygotowywanie notatek oraz poprawę dostępności nagrań, np. przez syntezę mowy, tłumaczenie czy audiodeskrypcję.

Zastosowania tego typu są stosunkowo bezpieczne, dopóki: dotyczą materiałów jawnych, publicznie i bezwnioskowo dostępnych, są w miarę możliwości weryfikowane i poprawiane przez człowieka (niektóre urzędy to podkreślały), zawierają wyraźne oznaczenie wskazujące na sposób wygenerowania oraz nie służą jako podstawa do podejmowania decyzji.



Dokumenty, tekst i wiedza

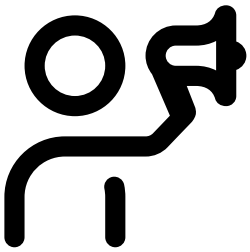
W tej grupie mieszczą się zastosowania związane z opracowywaniem, analizą i porządkowaniem treści pisemnych. AI służy tu do redagowania i streszczania tekstów, wyszukiwania informacji w dokumentach, ekstrakcji danych, anonimizacji, OCR oraz wspomagania pracy z pismami, korespondencją i dokumentacją prawną.

Szczególnie niebezpieczna wydaje się ta ostatnia sytuacja, ze względu na wciąż powszechne halucynacje dużych modeli językowych przy tego typu zadaniach.

Obsługa mieszkańców, klientów i użytkowników

To zastosowania, w których AI pełni funkcję pierwszej linii kontaktu z użytkownikiem. Obejmuje to chatboty i voiceboty udzielające informacji, pomagające znaleźć właściwą usługę, umawiające wizyty, sprawdzające status spraw lub odpowiadajace na pytania mieszkańców, pacjentów, studentów czy przedsiębiorców.

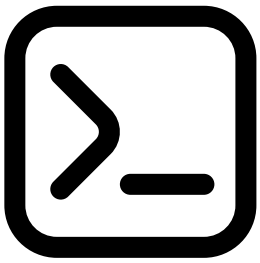
Stosunkowo bezpieczny obszar, w którym korzyści wydają się z nawiązką przewyższać ryzyka. Ważne, żeby użytkownik miał pełną świadomość, z czym ma do czynienia oraz łatwą i zawsze dostępną opcję kontaktu z człowiekiem, bez utraty ciągłości i kontekstu rozmowy.



Komunikacja, promocja i treści kreatywne

Kategoria dotyczy tworzenia i obróbki treści komunikacyjnych oraz promocyjnych. AI jest tu wykorzystywana do generowania grafik, prezentacji, materiałów do mediów społecznościowych, wideo, muzyki i innych form przekazu, a także do poprawy jakości i adaptacji treści do różnych kanałów.

Najpoważniejsze ryzyka w tym obszarze wiążą się z wątpliwościami natury prawnoautorskiej, więc nie dotyczą bezpośrednio osób ani decyzji, chociaż oczywiście nadal powinny być starannie mitygowane.



Programowanie i praca techniczna

AI coraz powszechniej wspiera pracę informatyków i osób technicznych przy tworzeniu oraz utrzymaniu systemów informatycznych. Odnotowane zastosowania obejmują między innymi generowanie i analizę kodu, debugowanie, pracę z dokumentacją techniczną, automatyzację zadań serwisowych, czy wsparcie przy projektowaniu aplikacji.

Pomimo związanych z tym ryzyk, wydaje się, że trend jest bardzo silny i niemożliwy do powstrzymania. Ponadto AI stanowi tu narzędzie podobne do młotka: mniej istotne od samego narzędzia jest to, co udało się (lub nie) zrobić przy jego pomocy, tak i tutaj istnieją stosowne metodyki wytwarzania oprogramowania, których przestrzeganie gwarantuje odpowiednią jakość końcowego produktu.

Analityka i wsparcie decyzyjne

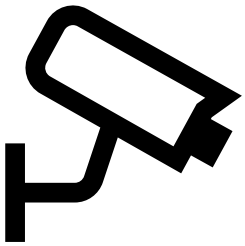
Ta grupa obejmuje użycie AI do analizowania danych i wspierania decyzji operacyjnych, zarządczych lub strategicznych. W odpowiedziach wymieniano takie zadania jak raportowanie, prognozowanie, identyfikowanie trendów, analiza ryzyka, monitorowanie zjawisk oraz predykcyjne utrzymanie infrastruktury i urządzeń.

Wrażliwy obszar z jednej strony, z drugiej w części deklarowanych tutaj zastosowań niekoniecznie wykorzystywano sztuczną inteligencję wykraczającą poza ściśle zdefiniowane reguły. Należy pamiętać, że jest to obszar bardzo delikatny, bo bliski decyzjom wprost wpływającym na ludzi i ich otoczenie. Jednocześnie powszechność narzędzi AI sprawia, że trudno tutaj o realną społeczną kontrolę każdego ich wykorzystania, z konieczności trzeba się skupić na systemowych rozwiązaniach, takich jak typowanie zwolnień lekarskich do kontroli, czyli - na początek - na wymuszeniu ujawniania takich sytuacji. Na dziś wydaje się, że majaczący na horyzoncie kolejny etap wdrożenia AI Act powoduje paradoksalnie większą niechęć do jawności, bo stosowanie tego typu rozwiązań oznacza szereg obowiązków do spełnienia i koniecznych ograniczeń. Wygodnie jest więc z perspektywy urzędu udawać, że dane rozwiązanie nie jest oparte na sztucznej inteligencji w rozumieniu ww. regulacji.

Rekrutacja, edukacja i nauka

W tej kategorii AI wspiera procesy związane z naborem, kształceniem i działalnością badawczą. Obejmuje to m.in. analizę formularzy rekrutacyjnych i CV, tworzenie materiałów dydaktycznych i pytań testowych, wyszukiwanie publikacji, ocenę prac oraz narzędzia do wykrywania plagiatu i użycia AI (przykładem używany przez uczelnie wyższe Jednolity System Antyplagiatowy).

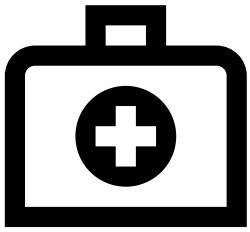
Wiele z tych zastosowań zasługuje na bliższe przyjrzenie się, ponieważ wiążą się z podejmowaniem decyzji w odniesieniu do zwykłych osób. Tak drobne, jak może się wydawać, zastosowanie sztucznej inteligencji, jakim jest chociażby preselekcja CV w rekrutacji, rodzi gigantyczne ryzyka w kontekście oczywistych i szeroko badanych uprzedzeń modeli językowych. Tym bardziej, że osoba, której CV pominięto, raczej nie może liczyć na otrzymanie informacji zwrotnej: *Twoje CV zostało odrzucone przez AI*, co dawałoby szansę na odwołanie się od takiej decyzji, a w szerszym kontekście na ujawnienie uprzedzenia modelu skutkujące masowym dyskryminującym odrzucaniem CV spełniających kryteria rekrutacyjne.



Obraz, monitoring, geodezja i środowisko

Są to zastosowania oparte na analizie obrazu, nagrań lub danych przestrzennych. AI służy tu do rozpoznawania obiektów, tablic i twarzy, analizy monitoringu (wykrywanie określonych zdarzeń), przetwarzania zdjęć satelitarnych, wspierania geodezji, monitorowania przyrody oraz oceny procesów zachodzących w przestrzeni publicznej i środowisku.

O ile zastosowania w odniesieniu do monitorowania środowiska wydają się relatywnie bezpieczne, to już te dotyczące ludzi wymagają dużej ostrożności. Przykładowo, jaką mamy kontrolę nad danymi: czy są przetwarzane lokalnie, czy wysyłane na serwery dostawcy usługi? Czy sposób wykorzystywania tych danych, zakres i czas przechowywania są zgodne z prawem, czy prywatność jest przy tym poszanowana?



Zdrowie i diagnostyka

Ta kategoria dotyczy zastosowań medycznych i okołomedycznych. AI wspiera tu analizę badań obrazowych, wykrywanie nieprawidłowości, ocenę zmian chorobowych i poprawę jakości diagnostyki.

Istnieje wiele wątpliwości związanych z jakością rozwiązań w tym obszarze i pomimo dużego postępu nadal wydaje się niezbędnym angażowanie człowieka w proces każdej jednej diagnozy. Tak jak w wielu podobnych przypadkach, niezbędnym minimum wydaje się skuteczne informowanie osób o zakresie, w jakim zastosowano sztuczną inteligencję do diagnozy ich przypadku, a żaden scenariusz procesu diagnozy w placówce medycznej nie powinien dopuszczać pełnego oparcia na sztucznej inteligencji.

Cyberbezpieczeństwo i bezpieczeństwo IT

W tej grupie AI jest wykorzystywana do ochrony systemów, sieci i danych. Obejmuje to wykrywanie anomalii, analizę incydentów, ochronę poczty, monitorowanie ruchu sieciowego, identyfikację zagrożeń, klasyfikację złośliwych plików oraz wspomaganie reagowania na zdarzenia bezpieczeństwa.

Pod względem praw człowieka są to rozwiązania w zasadzie neutralne, nie wiążą się z decyzjami bezpośrednio dotyczącymi obywateli. Ich jakość i bezpieczeństwo to kwestie odrębne - ważne, ale wykraczające poza zakres naszych kompetencji.

Inne wyspecjalizowane zastosowania administracyjne i sektorowe

To kategoria zbiorcza dla szczególnych, branżowych lub jednostkowych zastosowań, które nie tworzą osobnego bloku tematycznego. Obejmuje ona m.in. wsparcie rozliczeń, podatków, konsultacji społecznych, analiz patentowych, wykrywania ryzyk w ofertach internetowych czy modeli predykcyjnych w specyficznych procesach instytucjonalnych.

Jak instytucje publiczne regulują korzystanie z AI?

Świadomość powszechności narzędzi AI (pośrednio wynikająca z odpowiedzi na pierwsze pytanie wniosku) w gminach nie idzie w parze z tworzeniem wewnętrznych regulacji w tym zakresie (odpowiedzi na drugie pytanie wniosku). W samorządach gminnych jedynie 20 na 2027 to odpowiedzi twierdzące na drugie pytanie. Z tych 20 aż 7 dotyczyło miast na prawach powiatu, zatem w ***zwykłych*** gminach odsetek posiadających własne regulacje wyniósł grubo poniżej 1%. I generalnie wśród zapytanych rozmaitych innych typów instytucji wiele było takich, wśród których ani jeden podmiot nie wypracował własnych zasad korzystania z AI. Spośród placówek medycznych na 71 uzyskanych odpowiedzi twierdząca była tylko jedna: z Wojewódzkiego Szpitala Specjalistycznego Nr 5 im. św. Barbary w Sosnowcu, a wprowadzone przez nich zasady sprowadzają się do zakazu korzystania z ww. narzędzi.

Na drugim biegunie znalazły się ministerstwa, uczelnie i spółki skarbu państwa, gdzie ta potrzeba jest wyraźnie większa. Narodowy Fundusz Zdrowia wraz z oddziałami (a także, sporadycznie, inne instytucje) z kolei twierdził, że dostęp do zewnętrznych narzędzi AI jest u nich całkowicie blokowany na poziomie urządzeń teleinformatycznych, co też można uznać za formę regulacji (jakkolwiek drastyczną, a jednocześnie wątpliwie skuteczną, skoro dostęp do takich narzędzi można bez trudu uzyskać z prywatnego telefonu).

Najbardziej bodaj dojrzałe podejście w tym obszarze wykazywały w momencie badania uczelnie. Wiele spośród nich ustanowiło zasady korzystania z narzędzi AI, zarówno dla pracowników, jak i studentów - przy czym czasem były to osobne dokumenty. Niektóre uczelnie powołały nawet w tym celu specjalne ciała, np. rektor Uniwersytetu Szczecińskiego powołał Komisję ds. stosowania narzędzi SI w Uniwersytecie Szczecińskim, która opracowała zbiór dobrych praktyk dla pracowników i studentów.

Zauważalna grupa instytucji deklarowała, że trwają u nich mniej lub bardziej zaawansowane prace nad wewnętrznymi regulacjami. Jeszcze inne twierdziły, że przepisy związane z ochroną danych osobowych (na przykład RODO, albo przyjęte polityki) czy ogólniej bezpieczeństwem (na przykład wdrożone w wielu miejscach Systemy Zarządzania Bezpieczeństwem Informacji), już regulują wykorzystanie zewnętrznych narzędzi AI. W pewnym sensie i w pewnym zakresie jest to prawda, ale czy to wystarcza? Czy to skutecznie eliminuje wszystkie ryzyka związane z nieumiejętnym wykorzystaniem AI? Wszak nawet jeśli z istniejących regulacji i przepisów prawa dotyczących ochrony prywatności czy bezpieczeństwa danych wynika pośrednio zakaz korzystania z zewnętrznych narzędzi AI (np. w kontekście wprowadzania do nich danych wrażliwych), to nasuwa się pytanie, czy pracownicy danego podmiotu mają tego świadomość? Czy zostali odpowiednio przeszkoleni albo przynajmniej wyraźnie i wprost poinformowani? Jak to się ma do kwestii związanych z halucynacjami modeli językowych i podpieraniem się nimi przy podejmowaniu przez urzędnika rozmaitych decyzji? Te i inne pytania pozostają otwarte. Problemy te dostrzegła między innymi Akade-

mia Muzyczna w Poznaniu, której *pracownicy na okresowych szkoleniach z zakresu RODO informowani są o możliwościach wykorzystywania narzędzi AI w pracy, ale także o towarzyszącym temu zagrożeniu oraz zasadach bezpiecznego ich stosowania, jak również wynikającej z tego odpowiedzialności dla każdego użytkownika*.

Nie rozstrzygamy, czy każda instytucja powinna tworzyć własne regulacje dotyczące wykorzystania AI przez pracowników, bo być może jest to temat do zaopiekowania na wyższym poziomie (a przynajmniej w odniesieniu do wybranych typów instytucji). Dostrzegamy jednak niebezpieczeństwa związane z nieumiejętnym posługiwaniem się zewnętrznymi narzędziami opartymi na sztucznej inteligencji i w związku z tym apelujemy do instytucji publicznych, żeby brały sprawę w swoje ręce i nie odwlekały zajęcia się tematem do momentu, aż pojawi się problem – np. wyciek wrażliwych danych czy wydanie decyzji w oparciu o halucynacje AI. Tym bardziej, że zgodnie ze wspomnianym już wcześniej raportem połowa urzędników przyznaje się do korzystania z AI do celów służbowych, a 59% użyć w tygodniu poprzedzającym omawiane w raporcie badanie było *wyszukiwaniem faktów lub informacji, które pomagają w wykonywaniu pracy*.

Na podstawie analizy regulacji wewnętrznych polskich instytucji publicznych (głównie, jak powiedziano, uczelni wyższych oraz nielicznych przypadków urzędów administracji samorządowej), najczęściej praktykowane w administracji zalecenia dotyczące wykorzystywania narzędzi sztucznej inteligencji można sprowadzić do następujących:

Transparentność i obowiązek ujawniania użycia AI

To najpowszechniejsze zalecenie, wymagające od użytkownika jasnego zadeklarowania, w jakim zakresie i jakie narzędzia zostały wykorzystane w przygotowaniu pracy, raportu czy dokumentu. Często wymaga się zamieszczenia specjalnego oświadczenia we wstępie lub w formie załącznika, a w przypadku prac naukowych – wskazania tego w opisie metodologii.

Pełna odpowiedzialność użytkownika

Instytucje podkreślają, że za ostateczną treść dokumentu przygotowanego z wsparciem AI, jego poprawność merytoryczną oraz ewentualne naruszenia prawa odpowiedzialność ponosi zawsze i wyłącznie człowiek, a nie algorytm. AI jest traktowane jedynie jako narzędzie wspomagające, a nie autor.

Krytyczne podejście i weryfikacja wyników

Ze względu na zjawisko tzw. halucynacji AI (generowania nieprawdziwych informacji) oraz stronniczość modeli, użytkownicy są zobowiązani do każdorazowego sprawdzania faktów, źródeł oraz logicznej spójności wygenerowanych treści. Treści z AI nie mogą być traktowane jako dane źródłowe.

Ochrona danych osobowych i poufnych

Bezwzględnie zakazuje się wprowadzania do publicznie dostępnych narzędzi AI danych osobowych, informacji niepublicznych, finansowych, wrażliwych czy objętych tajemnicą służbową.

Wnioski z badania

Niska przejrzystość i systemowe bariery w dostępie do informacji

Badanie pokazuje, że rzeczywiste wykorzystanie AI w administracji publicznej jest trudne do ustalenia nie dlatego, że ono nie występuje, lecz dlatego, że istnieją liczne mechanizmy jego ukrywania lub nieujawniania. Należą do nich:

- nadużywanie klauzul niejawności i argumentów bezpieczeństwa,
- szerokie stosowanie odmów lub unikanie odpowiedzi,
- uznaniowe wyłączanie informacji jako *dokumentów wewnętrznych*,
- brak realnej, zewnętrznej kontroli zasadności tych decyzji, dziejącej się w sensownym czasie.

W efekcie im bardziej potencjalnie wrażliwe zastosowanie AI (np. militarne, automatyzujące decyzje), tym mniejsza transparentność, a tym samym możliwość kontroli społecznej.

Rozmywanie definicji AI jako kluczowy problem regulacyjny

Instytucje często stosują wąską lub wybiórczą interpretację definicji systemu AI, co pozwala im:

- nie ujawniać stosowanych rozwiązań,
- unikać obowiązków wynikających z regulacji (np. AI Act),
- wyłączać spod kontroli systemy realnie wpływające na obywateli (np. modele predykcyjne).

Jednocześnie inne podmioty stosują definicję bardzo szeroko, co pokazuje brak spójności interpretacyjnej.

Brak systemowego podejścia do jawności zastosowań AI

Z badania wynika, że nie istnieje żaden spójny mechanizm informowania o tym:

- gdzie i w jakim zakresie wykorzystywana jest AI,
- czy AI wpływa na decyzje wobec obywateli,
- jakie są ryzyka i zabezpieczenia.

To znowu ogranicza możliwość społecznej kontroli oraz audytu.

Niedostateczne regulacje wewnętrzne w instytucjach

Większość instytucji – zwłaszcza na poziomie samorządowym – nie posiada żadnych zasad korzystania z AI, mimo że:

- pracownicy faktycznie korzystają z narzędzi AI,
- istnieje ryzyko podejmowania decyzji opartych na błędnych (halucynowanych) treściach,
- istnieje ryzyko przetwarzania danych wrażliwych,

Regulacje – jeśli istnieją – są często fragmentaryczne lub sprowadzają się do wątpliwej skuteczności zakazów.

Rosnące wykorzystanie AI przy jednoczesnym niedoszacowaniu ryzyk

AI jest już wykorzystywana w wielu obszarach (m.in. analiza danych, obsługa obywateli, dokumenty, monitoring), a przy tym często jest stosowana w sposób niekontrolowany, czyli:

- przy użyciu zewnętrznych narzędzi,
- bez pełnej świadomości konsekwencji,
- bez adekwatnych zabezpieczeń.

Niewystarczające oznaczanie wykorzystania AI

W wielu przypadkach odbiorca (obywatel, pacjent, użytkownik) nie ma świadomości, że:

- ma do czynienia z systemem AI,
- treść lub decyzja została wsparta przez AI.

To ogranicza możliwość reakcji, odwołania lub krytycznej oceny.

Potrzeba podejścia systemowego zamiast punktowych działań

Z uwagi na:

- skalę wykorzystania AI,
- trudność w kontroli pojedynczych przypadków,
- asymetrię informacji między instytucjami a obywatelami,

konieczne jest podejście systemowe, a nie reaktywne.

Podsumowanie

Kluczowym problemem nie jest fakt wykorzystywania AI w administracji, lecz **niewystarczająca przejrzystość, brak spójnych zasad oraz realnej kontroli**. Bez wprowadzenia systemowych rozwiązań istnieje ryzyko dalszego rozwoju praktyk, które mogą negatywnie wpływać na prawa obywateli, pozostając jednocześnie poza społecznym nadzorem.

Rekomendacje

Na podstawie wyników i wniosków z badania można wyprowadzić szereg rekomendacji, które będą się różnić w zależności od przekonań i celów osoby je formułującej. Poniżej znajdują się propozycje autora raportu - nie stanowią oficjalnego stanowiska Sieci Obywatelskiej Watchdog Polska i wymagają pogłębionej dyskusji. Ponadto autor dostrzega, że niektóre z wymienionych niżej postulatów są teoretycznie zaadresowane w AI Akcie, jednak odnosi się do praktyki obserwowanej w czasie badania.

Transparentność i dostęp do informacji o wykorzystaniu AI

Konieczne jest wzmocnienie jawności wykorzystania sztucznej inteligencji w administracji publicznej, tak aby obywatele, organizacje społeczne oraz instytucje kontrolne miały wiedzę, gdzie, w jakim celu i z jakim skutkiem stosowane są takie rozwiązania.

Rekomendacje

Wprowadzenie jednoznacznego obowiązku ujawniania informacji o wykorzystaniu systemów AI w administracji publicznej, z ograniczeniem możliwości powoływania się na ogólne przesłanki bezpieczeństwa.

Utworzenie centralnego, publicznego rejestru zastosowań AI w administracji, obejmującego co najmniej informacje o celu systemu, zakresie jego wykorzystania, obszarze zastosowania, podstawowych ryzykach oraz zastosowanych zabezpieczeniach.

Wprowadzenie obowiązku informowania obywateli o tym, że mają do czynienia z systemem AI albo że AI wspiera dany proces, usługę lub decyzję administracyjną.

Zapewnienie, aby informacja o wykorzystaniu AI była przekazywana w sposób zrozumiały, widoczny i adekwatny do kontekstu, w szczególności w pismach, systemach teleinformatycznych i bezpośrednich interakcjach z obywatelami.

Uporządkowanie kwestii *dokumentu wewnętrznego* w taki sposób, aby nie mogła ona służyć blokowaniu dostępu do informacji o systemach wpływających na prawa, obowiązki lub sytuację obywateli.

Powiązanie obowiązków informacyjnych z prawem do uzyskania dodatkowych wyjaśnień oraz możliwością zakwestionowania rozstrzygnięcia lub działania wspieranego przez AI.

Definicje, kwalifikacja i zakres regulacji

Niezbędne jest ograniczenie dowolności interpretacyjnej w zakresie tego, czym jest system AI i które rozwiązania powinny podlegać obowiązkom regulacyjnym. O kwalifikacji systemu powinny decydować jego funkcje, sposób wykorzystania i wpływ na obywateli, a nie wyłącznie deklaracje instytucji.

Rekomendacje

Przyjęcie jednolitej, operacyjnej definicji systemu AI na potrzeby stosowania przepisów w administracji publicznej, ograniczającej możliwość dowolnej interpretacji przez instytucje.

Wprowadzenie zasady, zgodnie z którą o kwalifikacji systemu decyduje jego funkcjonalność, sposób wykorzystania oraz faktyczny wpływ na sytuację obywateli, a nie wyłącznie nazwa, opis lub deklaracja podmiotu go stosującego.

Opracowanie katalogu przykładów systemów podlegających obowiązkom regulacyjnym, w tym systemów predykcyjnych, analitycznych, scoringowych i wspierających decyzje.

Wprowadzenie obowiązku ujawniania zastosowań systemów analitycznych i predykcyjnych niezależnie od tego, czy dana instytucja formalnie uznaje je za AI.

Ustanowienie mechanizmów umożliwiających niezależną ocenę, czy dane rozwiązanie powinno zostać uznane za sys-

tem AI i objęte odpowiednimi obowiązkami.

Powiązanie obowiązków regulacyjnych nie tylko z klasyfikacją techniczną systemu, ale również z poziomem ryzyka oraz wpływem jego działania na prawa i sytuację obywateli.

Zarządzanie ryzykiem i praktyki operacyjne w instytucjach

Wykorzystanie AI w administracji powinno odbywać się na podstawie jasnych zasad wewnętrznych, adekwatnych do poziomu ryzyka. Szczególne znaczenie ma ograniczenie zagrożeń związanych z danymi wrażliwymi, błędami systemów oraz wykorzystywaniem narzędzi zewnętrznych bez właściwej kontroli.

Rekomendacje

Wprowadzenie obowiązku opracowania i wdrożenia wewnętrznych polityk korzystania z AI w każdej instytucji publicznej, dostosowanych do jej specyfiki, zadań i poziomu ryzyka.

Określenie jasnych zasad dotyczących przetwarzania danych przy korzystaniu z narzędzi AI, w tym zakazu wprowadzania danych wrażliwych lub prawnie chronionych do nieautoryzowanych narzędzi zewnętrznych.

Zapewnienie szkoleń dla pracowników w zakresie bezpiecznego, odpowiedzialnego i krytycznego korzystania z AI, obejmujących również rozpoznawanie błędów, ograniczeń i ryzyk związanych z halucynacjami systemów.

Wprowadzenie procedur obowiązkowej weryfikacji treści generowanych przez AI przed ich wykorzystaniem w działaniach urzędowych, analitycznych, komunikacyjnych lub decyzyjnych.

Wprowadzenie obowiązku dokonywania oceny ryzyka przed wdrożeniem systemu AI lub podobnego rozwiązania analitycznego, w szczególności gdy może ono wpływać na prawa obywateli, dane osobowe, rekrutację, monitoring lub decyzje administracyjne.

Określenie minimalnych zabezpieczeń dla zastosowań wysokiego ryzyka, obejmujących m.in. wymóg nadzoru człowieka, dokumentowania działania systemu, testowania poprawności oraz okresowej weryfikacji skutków jego stosowania.

Ustanowienie mechanizmów nadzoru wewnętrznego oraz odpowiedzialności za nieprawidłowe, nieudokumentowane lub ryzykowne wykorzystanie narzędzi AI.

Nadzór, egzekwowanie i koordynacja systemowa

Skuteczność zasad dotyczących AI zależy nie tylko od ich przyjęcia, ale również od istnienia mechanizmów kontroli, egzekwowania i koordynacji działań na poziomie całej administracji publicznej. Potrzebne jest podejście systemowe, a nie wyłącznie reagowanie na pojedyncze przypadki.

Rekomendacje

Ustanowienie niezależnego mechanizmu szybkiej weryfikacji zasadności odmów udostępnienia informacji o wykorzystaniu systemów AI.

Wprowadzenie sankcji lub innych skutecznych mechanizmów egzekwowania obowiązków w przypadku nieuzasadnionego ograniczania dostępu do informacji, nieoznaczania wykorzystania AI lub naruszania przyjętych zasad jego stosowania.

Zapewnienie regularnych audytów systemów AI oraz okresowej aktualizacji informacji o ich działaniu, ryzykach i zabezpieczeniach, w sposób umożliwiający porównywanie danych między instytucjami.

Wzmocnienie niezależnych instytucji nadzorczych oraz zapewnienie im kompetencji, zasobów i narzędzi potrzebnych do skutecznej kontroli wykorzystania AI w administracji publicznej.

Opracowanie i wdrożenie spójnej strategii zarządzania wykorzystaniem AI w administracji publicznej, obejmującej standardy, nadzór, kontrolę oraz zasady odpowiedzialności.

Zapewnienie koordynacji działań na poziomie centralnym w celu ograniczenia fragmentaryczności regulacji i praktyk oraz ujednolicenia standardów postępowania w instytucjach publicznych.

Powiązanie obowiązków informacyjnych i nadzorczych z realną kontrolą społeczną, w tym poprzez dostęp do raportów, mechanizmy zgłaszania zastrzeżeń oraz działania zwiększające dostępność wiedzy o wykorzystaniu AI dla obywateli i organizacji społecznych.

Podsumowanie

Rekomendacje te tworzą cztery wzajemnie powiązane filary odpowiedzialnego wykorzystania AI w administracji publicznej: **transparentność**, **jednoznaczne zasady kwalifikacji systemów**, **zarządzanie ryzykiem w praktyce instytucjonalnej** oraz **skuteczny nadzór i egzekwowanie**. Dopiero łączne wdrożenie tych elementów może ograniczyć ryzyko rozwoju nieprzejrzystych i niesprawdzalnych praktyk, które wpływają na prawa obywateli.

Lista systemów AI wykorzystywanych w administracji publicznej

Zastrzeżenie: Lista zawiera systemy AI, które zostały wymienione w odpowiedzi na wnioski o informację wysłane w ramach niniejszego badania. W obecnej postaci lista ma funkcję wyłącznie poglądową - została przygotowana w sposób łączący ręczną pracę z wykorzystaniem AI i może zawierać błędy. W razie potrzeby odwołania się do poszczególnych przypadków należy zweryfikować informacje podane na liście z odpowiedzią danego urzędu.

Źródła

Korespondencja w sprawach dot. wykorzystywania AI

Arkusz z zebranymi odpowiedziami na wnioski, przygotowany z wykorzystaniem AI

Lista systemów AI wykorzystywanych w administracji publicznej przygotowana na podstawie odpowiedzi

AI w e-administracji publicznej – perspektywa urzędników i instytucji

W drodze ku doskonałości cyfrowej

Resortowa strategia sztucznej inteligencji do roku 2039

AI Act - Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2024/1689 - z dnia 13 czerwca 2024 r.

Podstawowe źródło: Odpowiedzialne wykorzystywanie sztucznej inteligencji

Opracowanie: Sieć Obywatelska Watchdog Polska, ostatnia aktualizacja: 8 kwietnia 2026.

Tekst powstał w ramach projektu finansowanego przez European AI & Society Fund.