Krajowa Szkoła Sądownictwa i Prokuratury

Newsletter

PRACOWNI VR/AI

I/26

Grafika - wygenerowana przez AI, Canva

Oddajemy w Państwa ręce kolejny numer newslettera Pracowni VR/AI, poświęcony wykorzystaniu nowoczesnych technologii w wymiarze sprawiedliwości.

Dynamiczny rozwój narzędzi opartych na AI rodzi pytania nie tylko o ich możliwości, lecz także o bezpieczeństwo, odpowiedzialność i ramy prawne ich stosowania. W tym wydaniu przyglądamy się tym zagadnieniom z różnych perspektyw.

Publikujemy rozmowy z ekspertkami – Martą Wilamowską-Juszczyk na temat wyzwań związanych z RODO oraz sędzią Patrycją Dolniak, która opowiada o rozwoju kompetencji cyfrowych w europejskim środowisku prawniczym. Poruszamy również temat deepfake oraz jego wpływu na bezpieczeństwo informacji i wiarygodność materiałów dowodowych. Przygotowaliśmy dla Państwa słownik pojęć AI oraz przegląd harmonogramu wdrażania AI Act. Nowością jest rubryka „Strefa szkoleń AI", w której informujemy o aktualnych wydarzeniach edukacyjnych.

Mamy nadzieję, że przygotowane materiały będą wsparciem w świadomym i odpowiedzialnym korzystaniu z nowych technologii.

Życzymy inspirującej lektury.

Sekcja-Pracownia VR/AI

Przedstawiamy:

DIGITAL AMBASSADOR

nowa rola w erze transformacji cyfrowej

Patrycja DOLNIAK

sędzia, dr nauk prawnych, EJTN Digital Ambassador, tutor w ramach programu HELP Artificial Intelligence and Human Rights, organizowanego przez Radę Europy, wykładowczyni w Krajowej Szkole Sądownictwa i Prokuratury w Krakowie, wieloletnia wykładowczyni akademicka, trenerka z zakresu mediacji, autorka kilkudziesięciu publikacji naukowych, w tym monografii i komentarzy. Absolwentka Uniwersytetu Jagiellońskiego w Krakowie (studia magisterskie), Uniwersytetu Śląskiego w Katowicach (studia doktoranckie), Uniwersytetu Warszawskiego w Warszawie (studia podyplomowe z prawa własności intelektualnej), Akademii Leona Koźmińskiego w Warszawie (studia podyplomowe z prawa sztucznej inteligencji) oraz The George Washington University w Waszyngtonie (Legal writing).

M.KANIA: **Co skłoniło Panią do zaangażowania się w program EJTN Digital Ambassador?**

P.DOLNIAK: Często biorę udział wydarzeniach organizowanych przez EJTN, a EJTN Digital Ambassador jest wyjątkowym przedsięwzięciem. To bowiem nie tylko seria zaawansowanych szkoleń, ale też program, który zakłada dzielenie się swoją wiedzą. Od ponad 15 lat jestem nauczycielem akademickim, prowadzę szkolenia w KSSiP, więc założenia tego programu bardzo mi się spodobały. Sztuczną inteligencją w wymiarze sprawiedliwości zajmuję się już od kilku lat, razem z sędziami Tomaszem Kuźmą, Konradem Wasikiem i prokuratorem Andrzejem Ludwińskim jestem współautorką monografii ***Sztuczna inteligencja w wymiarze sprawiedliwości. Między prawem a algorytmami***, opublikowałam szereg artykułów naukowych poświęconych AI, skończyłam studia podyplomowe z zakresu prawa sztucznej inteligencji. Nie mogłam więc nie wziąć udziału w tej inicjatywie.

Jak - z Pani perspektywy - powinna dziś wyglądać rola Digital Ambassador w polskim wymiarze sprawiedliwości?

Myślę, że my, pracownicy wymiaru sprawiedliwości, potrzebujemy rozmowy o sztucznej inteligencji. AI wzbudza wiele emocji – od dużej fascynacji jej możliwościami, po obawy związane z jej nieuprawnionym użyciem. Chciałabym w związku z powierzoną mi funkcją skupić się na tej rozmowie, przede wszystkim w formie szkoleń adresowanych do sędziów i prokuratorów. Jestem również, jako redaktor naukowy, w trakcie przygotowywania razem z moim wspaniałym zespołem książki ***Od paragrafu do promptu. Podstawy AI dla sędziów i prokuratorów.*** Chcemy w niej oswoić AI, pokazać jej możliwości, ale i zagrożenia, jakie ze sobą niesie. Chciałabym też być słyszalnym dla Ministerstwa głosem w przedmiocie digitalizacji akt.

Jakie kompetencje cyfrowe uważa Pani za kluczowe dla sędziów i pracowników sądów w najbliższych latach?

Myślę, że powinniśmy przede wszystkim skupić się na kwestiach cyberbezpieczeństwa, szczególnie w obliczu coraz częstszych ataków hakerskich. Bardzo się cieszę, że ten temat jest coraz częściej dostrzegany w Ministerstwie Sprawiedliwości i mamy więcej szkoleń w tym zakresie. Uważam jednak, że musimy wzmocnić nasze kompetencje w aspekcie bezpiecznego korzystania z AI. Mam tu w szczególności na myśli duże modele językowe, jak choćby ChatGPT. Uważam, że z takich narzędzi można korzystać nawet w naszej pracy zawodowej, ale ze szczególną ostrożnością i zachowaniem reguł, które pozwolą jak najbardziej zminimalizować potencjalne zagrożenia.

Które doświadczenia lub narzędzia poznane w ramach programu mogą mieć realne, praktyczne zastosowanie w polskich sądach?

Znów powrócę do cyberbezpieczeństwa. W trakcie szkoleń dużo było mowy o tym jak można chronić się przed atakami hakerskimi i jak zabezpieczać dane. Wiele z tych podpowiedzi wdrożyłam w swoje prywatne życie. Chciałabym podzielić się tą wiedzą, tym bardziej, że czasami są to naprawdę proste narzędzia, jak choćby wyłączenie trenowania modelu na naszych danych, co zajmie nam dosłownie kilka sekund. Chciałabym też porozmawiać z sędziami i prokuratorami jak udoskonalić na nasze potrzeby dostępne modele językowe, choćby poprzez stworzenie RAG czy wyłączanie określonego kontentu przy generowaniu przez AI odpowiedzi.

Jak ocenia Pani potencjał technologii AI w edukacji prawniczej i szkoleniu kadr wymiaru sprawiedliwości?

Uważam, że to świetne narzędzie. Zresztą wiem, że Sekcja-Pracownia VR/AI KSSiP już z powodzeniem stosuje nowe technologie w edukacji. Bardzo się cieszę, że będę miała okazję dołożyć do tego swoją cegiełkę, tym bardziej, że zajmuję się też szkoleniem aplikantów aplikacji sędziowskiej. Myślę, że musimy dostosowywać nasze techniki szkoleniowe do potrzeb odbiorcy, tak aby zapewnić jak najefektywniejszy model nauki.

Jakie działania lub inicjatywy planuje Pani podjąć w najbliższym czasie jako Digital Ambassador w Polsce?

Przede wszystkim szkolenia. Razem z KSSiP ruszamy w 2026 r. z serią szkoleń poświęconych AI, w tym zagadnieniu deepfake, który jawi mi się jako jedno z najpoważniejszych zagrożeń dla rzetelnego i sprawiedliwego procesu. Planuję współpracę z sądami, w tym z moim macierzystym sądem – SR Katowice-Wschód w Katowicach i organizowanie szkoleń dla sędziów na miejscu, w sądach. Jak wspomniałam, pracujemy już nad książką – Od paragrafu do promptu. Podstawy AI dla sędziów i prokuratorów, która ma stanowić przystępny przewodnik po świecie AI, zawierający praktyczne porady związane choćby z pisaniem promptów. Bardzo liczę, że będzie dostępna już jesienią tego roku.

Jaką jedną myśl lub rekomendację chciałaby Pani przekazać osobom, które z rezerwą podchodzą do nowych technologii w wymiarze sprawiedliwości?

Sztuczna inteligencja nie jest zła. Złe może być jej wykorzystanie. Korzystajmy więc odpowiedzialnie i bezpiecznie, a możemy być mile zaskoczeni tym, jak może usprawnić naszą pracę.

M.KANIA: **Dziękuję za rozmowę!**

Rozmowy o...

AI W SĄDZIE: POMOC CZY ZAGROŻENIE?

Głos inspektora ochrony danych


Fot. archiwum A.Wilamowska-Juszczyk

MARTA WILAMOWSKA-JUSZCZYK
Inspektor ochrony danych
Sąd Okręgowy w Olsztynie

M.KANIA: **Jak dziś wygląda sytuacja związana z wykorzystaniem sztucznej inteligencji w sądownictwie z perspektywy inspektora ochrony danych?**

M.WILAMOWSKA-JUSZCZYK: W dzisiejszym sądownictwie powszechnym, w kontekście korzystania z narzędzi opartych na sztucznej inteligencji, jedno słowo opisuje sytuację lepiej niż jakiekolwiek analizy czy raporty, a tym słowem jest „niepewność".

To właśnie ona staje się największym wyzwaniem dla administratorów danych, inspektorów ochrony danych i wszystkich osób odpowiedzialnych za bezpieczeństwo informacji w sądach. Przy ogólnodostępnych narzędziach AI nie jesteśmy w stanie z pełnym przekonaniem odpowiedzieć na podstawowe pytania, takie jak: jakie dane pracownik wprowadza do narzędzia, gdzie te dane są przetwarzane, jak długo są przechowywane i do jakich celów wykorzystywane. Największą jednak niepewność budzi pytanie, czy wygenerowana odpowiedź jest poprawna. Mam wrażenie, że zachłyśnięci szybkością generowania rezultatu zapominamy, że AI jest tylko narzędziem technicznym, a nie źródłem wiedzy i to rodzi niestety ogromne zagrożenia.

To nie są drobne wątpliwości, to fundamenty bezpieczeństwa danych, na których opiera się odpowiedzialność administratora.

Największy problem polega na tym, że w praktyce kolejność działań została odwrócona. Najpierw pojawiły się narzędzia AI, a dopiero później zaczęto zastanawiać się nad właściwymi procedurami korzystania z tych narzędzi, obejmującymi zasady bezpieczeństwa, szkolenia pracowników, ocenę ryzyka oraz odpowiedzialność za dane.

Wskazuje Pani, że problemem jest odwrócona kolejność działań: najpierw technologia, potem regulacje. Jak powinien wyglądać właściwy model wdrażania AI w sądach?

Powinno być odwrotnie: najpierw dokładne regulacje przepisami, następnie poszerzanie świadomości pracowników, a dopiero potem technologia. W idealnym modelu narzędzie AI dla sądów powinno być opracowane na szczeblu centralnym oraz skupiać się na zakresie działalności wymiaru sprawiedliwości.

Takie rozwiązanie zapewniłoby odpowiednie zabezpieczenia danych przetwarzanych przez sądy i bezpieczne wykorzystanie szerokiego potencjału sztucznej inteligencji.

Czy są sytuacje, w których AI mogłaby realnie odciążyć sądy?

Paradoks polega na tym, że właśnie tam, gdzie sztuczna inteligencja mogłaby odciążyć sądy, dziś nie możemy z niej skorzystać. Gdy na jednego sędziego przypada obowiązek zapoznania się z setkami tysięcy stron akt, technologia mogłaby stanowić realne wsparcie.

Przykładem jest sprawa prowadzona przez sędziego Olgierda Dąbrowskiego-Żegalskiego, prowadzona w Sądzie Okręgowym w Olsztynie, obejmująca 1600 tomów akt, czyli około 320 tysięcy stron. Zgodnie z obowiązującymi przepisami rozpatrzenie sprawy w całości przypada wyłącznie na jednego sędziego. Mimo poszukiwań rozwiązań technologicznych nie było możliwości ich legalnego zastosowania. A przecież właśnie w takiej sytuacji sztuczna inteligencja mogłaby wpłynąć na takie czynności jakimi są wyszukiwanie kluczowych informacji czy analiza powiązań, to wszystko mogłoby realnie skrócić czas pracy i poprawić jakość orzecznictwa nie tylko w przytoczonej przeze mnie sprawie, ale ogólnie w całym systemie.

Jakie konkretne bariery prawne i organizacyjne uniemożliwiają dziś bezpieczne korzystanie z AI przy analizie akt spraw?

Obecnie litera prawa nie daje takich możliwości, a jednocześnie nie istnieje żaden centralny, bezpieczny system, do którego można byłoby wprowadzić materiały z akt i przetwarzać je przy użyciu AI w sposób zgodny m.in. z RODO. Dopóki taki system nie powstanie, sędzia pozostaje sam z ogromem materiału dowodowego, a AI, choć potencjalnie mogłaby być narzędziem przełomowym, pozostaje poza zasięgiem, bo nie spełnia wymogów prawnych i bezpieczeństwa.


Grafika- źródło: Canva

Z perspektywy administratora danych – jakie są dziś największe ryzyka związane z używaniem ogólnodostępnych narzędzi AI?

W obecnym stanie prawnym i technologicznym administrator danych osobowych, który nie reguluje kwestii korzystania z narzędzi AI, nie ma zapewnionego komfortu, jakim jest pewność co do bezpieczeństwa danych. Nie może mieć gwarancji, że dane nie zostaną wykorzystane w innym celu, niż to określił albo że nie będą przechowywane dłużej, niż to konieczne. Nigdy nie możemy być pewni, że dane, za które odpowiada administrator, nie posłużą do trenowania modeli AI i to są niestety realne ryzyka, a nie teoretyczne rozważania. Ani administrator, ani specjalista IT, ani inspektor ochrony danych, ani pracownik nie mogą zagwarantować, że dane wprowadzone do ogólnodostępnego narzędzia AI nie zostaną wykorzystane w sposób niezgodny z prawem, bo nikt poza dostawcą tego narzędzia nie ma pełnej wiedzy o tym, co dzieje się „po drugiej stronie".

Jaką rolę w tym procesie odgrywa inspektor ochrony danych i jakie działania można podjąć już teraz?

W moim przekonaniu inspektor ochrony danych w sądzie ma na dzień dzisiejszy niestety bardzo ograniczone narzędzia w zakresie ochrony danych przy korzystaniu z narzędzi AI. Jednak podstawą minimalnych środków ostrożności, które powinny być już wdrożone w każdej instytucji, stanowią regulacje wewnętrzne określające konkretne zasady korzystania z ogólnodostępnych narzędzi AI oraz, co jest chyba najważniejsze, edukacja pracowników: przedstawianie sposobu funkcjonowania sztucznej inteligencji oraz obecnych zagrożeń, nauka anonimizacji, czy przypominanie o zasadzie minimalizacji danych.

Dlatego w ubiegłym roku poprosiliśmy o wystąpienie pracowników Sekcji-Pracowni VR/AI Krajowej Szkoły Sądownictwa i Prokuratury, w którym specjaliści dokładnie wyjaśnili sposób działania AI oraz plusy i minusy korzystania z tego typu narzędzi kadrze orzeczniczej i urzędniczej Sądu Okręgowego w Olsztynie. Mogę zapewnić, że nie poprzestaniemy na tym szkoleniu, ponieważ sztuczna inteligencja nie wchodzi na salony polskiego sądownictwa, ona już tu jest i czuje się w tym miejscu jak w domu. Dlatego zadaniem administratorów danych osobowych, jak również inspektorów ochrony danych osobowych jest zapewnienie odpowiedniego poziomu świadomości pracowników w kontekście korzystania ze sztucznej inteligencji.

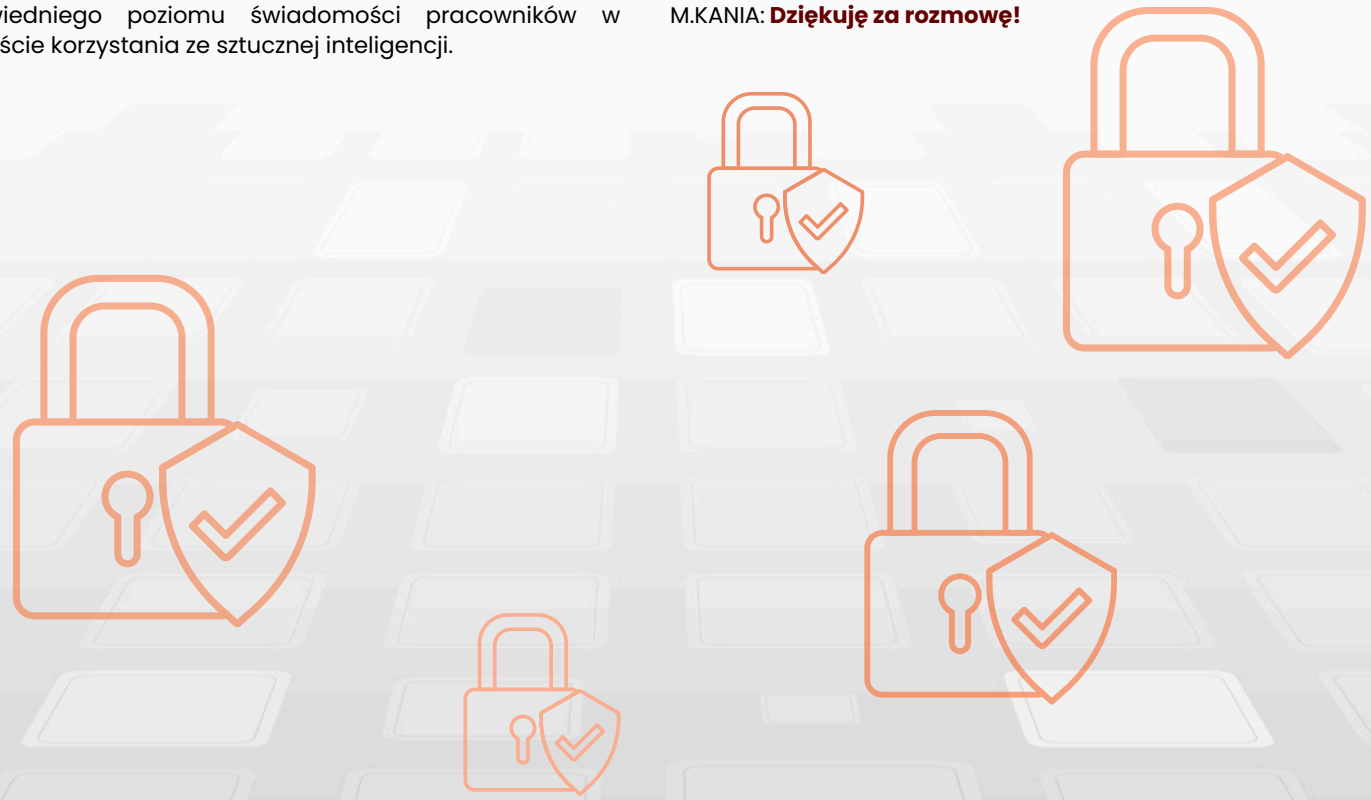
Czy istnieje ryzyko, że narzędzia AI są już wykorzystywane w sądach w sposób niekontrolowany? Jakie mogą być tego konsekwencje?

Najbardziej niepokojący jest fakt, że nie możemy mieć pełnego przekonania, iż w żadnym sądzie nie dochodzi do nieuprawnionego wykorzystywania narzędzi AI w toku pracy. To zjawiska niezwykle trudne do wykrycia, bo mogą się odbywać poza systemami sądowymi, na prywatnych urządzeniach lub w ogólnodostępnych aplikacjach, które nie pozostawiają śladu w infrastrukturze sądu. W praktyce oznacza to, że nie jesteśmy w stanie z całą pewnością wykluczyć, iż np. specjalista OZSS nie wspiera się sztuczną inteligencją przy sporządzaniu opinii bezkrytycznie przyjmując uzyskane odpowiedzi na zadane pytania, asystent sędziego nie wprowadza pełnej umowy do zewnętrznego narzędzia albo sędzia lub pracownik sekretariatu nie prosi AI o streszczenie akt sprawy na podstawie zdjęcia treści. A przecież mówimy o obszarach, w których podejmowane decyzje mogą wpływać na takie aspekty, jak władza rodzicielska, wolność osobista, majątek, czy podstawowe prawa obywatelskie. W tak wrażliwym środowisku nie możemy pozwolić, aby narzędzie, którego działania nie jesteśmy w stanie ani kontrolować, ani zweryfikować, w jakikolwiek sposób oddziaływało na proces orzeczniczy i należy podkreślić, że problem ten nie dotyczy wyłącznie polskiego sądownictwa, także jest to wyzwanie o charakterze globalnym.

Czy mimo tych zagrożeń widzi Pani w sztucznej inteligencji realną szansę dla wymiaru sprawiedliwości?

Sztuczna inteligencja może być ogromnym wsparciem dla sądownictwa. Może przyspieszyć pracę, odciążyć pracowników, poprawić jakość analiz, terminowość rozstrzygania spraw, czy uwolnić sądy od odtwórczych, powtarzalnych czynności. Jednak dopóki nie ma centralnego, bezpiecznego narzędzia, nie ma jasnych procedur wewnętrznych i odpowiednich regulacji prawnych, nie ma obowiązkowych szkoleń, dopóty nie ma kontroli nad przepływem danych, a niepewność będzie największym zagrożeniem dla ochrony danych osobowych w sądach. A w obszarze wymiaru sprawiedliwości niepewność to luksus, na który nie możemy sobie pozwolić.

M.KANIA: **Dziękuję za rozmowę!**



Deepfake

- wyzwania dla praktyki wymiaru sprawiedliwości – cz. I

Wstęp

Dynamiczny rozwój narzędzi generatywnej sztucznej inteligencji sprawił, że coraz częściej spotykamy się z materiałami audio i video wyglądającymi „jak prawdziwe", ukazującymi sytuacje, osoby lub zdarzenia, które w rzeczywistości nigdy nie miały miejsca. Tego typu treści (podmiana twarzy, „mówiąca głowa", klonowanie głosu) określa się mianem deepfake i są one obecnie jednym z najbardziej palących wyzwań dla bezpieczeństwa informacyjnego, zaufania publicznego oraz praktyki dowodowej w postępowaniach sądowych.

W prawie UE pojęcie to doczekało się definicji legalnej. Zgodnie z art. 3 pkt 60 rozporządzenia (UE) 2024/1689 (AI Act) „deepfake" to wygenerowane lub zmanipulowane przez AI obrazy, treści dźwiękowe lub wideo, które przypominają istniejące osoby, przedmioty, miejsca, podmioty, zdarzenia i mogą niesłusznie uchodzić za autentyczne[1].

Z perspektywy praktyki instytucjonalnej ważniejszy niż sama definicja jest jednak art. 50 AI Act dotyczący obowiązków przejrzystości. Wprowadza on m.in. obowiązek ujawnienia, że treść stanowiąca deepfake została sztucznie wygenerowana lub zmanipulowana, z istotnym wyjątkiem dla użyć „autoryzowanych prawem" w obszarze: wykrywania, zapobiegania, ścigania przestępstw, oraz z doprecyzowaniem dla dzieł oczywiście artystycznych i/lub satyrycznych. Dla redakcji oraz odbiorców instytucjonalnych istotny jest harmonogram stosowania przepisów, gdyż wdrażanie AI Act jest stopniowe, a reguły przejrzystości (art. 50) mają zacząć obowiązywać od 2 sierpnia 2026 r. podczas gdy definicje i część przepisów ogólnych już weszły w życie (m.in. od 2 lutego 2025 r.)

Wytwarzanie, wykorzystywanie i wykrywanie

Deepfake może służyć zarówno do wytworzenia fałszywych dowodów (np. rzekome groźby, „przyznanie się", rozmowa telefoniczna, nagranie z rzekomego spotkania), jak i do wywołania efektu „dywidendy kłamcy" (liar's dividend) – czyli strategicznego podważania prawdziwych nagrań poprzez twierdzenie, że „to nie Ja, to na pewno AI". Ten drugi mechanizm jest szczególnie niebezpieczny, gdyż „podnosi próg wiarygodności" oraz wydłuża spory o autentyczność poddając w wątpliwość materiał, którym przykładowo organy ścigania już dysponują.

Niestety prowadzone badania pokazują[2], że ludzie mają ograniczoną i bardzo zróżnicowaną skuteczność w rozpoznawaniu deepfake-ów, a przy tym najczęściej przeceniają własne zdolności i umiejętności ich rzeczywistego rozpoznawania. W cytowanej metaanalizie (wiele badań łącznie) wyników z wielu badań łączna trafność rozpoznawania treści syntetycznych była tylko nieznacznie wyższa od przypadku i z dużą zmiennością zależnie od formatu (wideo/audio/obraz-zdjęcie) oraz warunków odbioru. Autorzy przywołanego wyżej artykułu wskazali m.in., że w badaniach eksperymentalnych przeciętny odbiorca zwykle nie osiąga wysokiej trafności w odróżnianiu materiałów prawdziwych od syntetycznych, a w przeprowadzonej metaanalizie oszacowano zagregowaną trafność na ok. 55% z szerokim przedziałem ufności, który niestety ale w wielu podobszarach obejmuje wynik losowy. Jednocześnie część badań wskazuje na trwały problem nadmiernej pewności siebie („wydaje mi się, że rozpoznam fałszerstwo"), przy niskiej skuteczności i skłonności do błędów w kierunku uznawania fałszu za prawdę[3].

Można bez przesady stwierdzić, że po stronie algorytmów trwa „wyścig zbrojeń". Programy detekcyjne często dobrze działają na znanych już i stosowanych tzw. „laboratoryjnych" typach fałszerstw. Niestety radzą sobie znacznie gorzej w warunkach realnych, tzn. w przypadku wykorzystywania nowych generatorów, kompresji danych, „prania" przez platformy społecznościowe, ponownego kodowania, przycinania, filtrowania oraz innego rodzaju przetwarzania już wygenerowanych danych (audio, video, obrazy-zdjęcia). Wprost podkreślają to programy ewaluacyjne nastawione na odporność i uogólnianie wyników[4]. Wydaje się, że praktycznie istotny wątek to tzw. „nagrania wtórne" (np. telefonem z ekranu), które wprowadzają niepożądane artefakty (m.in. wzory mory[5]) obniżające skuteczność detektorów i zwiększające ryzyko błędnych wniosków[6].

Jak zatem postępować z materiałem „podejrzanym"? Wydaje się, że najbezpieczniejsze podejście to traktowanie materiałów audio-wideo jako dowodów cyfrowych wymagających kontroli integralności, pochodzenia i obiegu. Rekomendacje bazujące na standardowym podejściu do zabezpieczania dowodów cyfrowych akcentują m.in. dokumentowanie źródła pozyskania, utrzymanie łańcucha przekazania, tworzenie kopii roboczych oraz przechowywanie sum kontrolnych i podpisów.

[1] Rozporządzenie (UE) 2024/1689 (AI Act), Dz.U. UE L 1689 z 12.07.2024.

[2] Zob.: Diel, A., Lalgi, T., Schröter, I. C., MacDorman, K. F., Teufel, M., & Bäuerle, A. (2024). Human performance in detecting deepfakes: A systematic review and meta-analysis of 56 papers. Computers in Human Behavior Reports, 16, 100538. https://doi.org/10.1016/j.chbr.2024.100538 (dostęp na dzień: 03.04.2026 r.).

[3] Zob.: Köbis, N. C., Doležalová, B., & Soraperra, I. (2021). Fooled twice: People cannot detect deepfakes but think they can. iScience, 24(11), 103364. https://doi.org/10.1016/j.isci.2021.103364 (dostęp na dzień: 03.04.2026 r.).

[4] Zob. Guan, H., Horan, J., & Zhang, A. (2025). Guardians of forensic evidence: Evaluating analytic systems against AI-generated deepfakes. Forensics@NIST 2024. National Institute of Standards and Technology (NIST). https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=959128 (dostęp na dzień: 03.04.2026 r.).

[5] „Mora – w poligrafii, grafice komputerowej, filmie i fotografii to niepożądany efekt, pojawiający się w postaci regularnych punktów lub wzorów, wskutek krzyżowania się układu co najmniej dwóch regularnych siatek rastrowych lub wzorów podobnego rodzaju. (...) W fotografii efekt mory powoduje powstanie na zdjęciach dziwnie wyglądającego falistego wzoru, którego nie ma na rzeczywistym obiekcie. Powodem występowania efektu mory na zdjęciach wykonanych za pomocą aparatów cyfrowych jest interferencja pomiędzy wzorem na obiekcie a regularnym wzorem pikseli na matrycy, tworząca trzeci wzór, czyli właśnie efekt mory." Cytat zaczerpnięty ze strony: https://fulmanski.pl/zajecia/grafika_2d/zajecia_20152016/wyklad_rastrowa_vs_wektorowa/mora/index.php (dostęp na dzień: 03.04.2026 r.).

[6] Zob.: Tariq, R., Heo, M., Woo, S. S., & Tariq, S. (2024). Beyond the screen: Evaluating deepfake detectors under moire pattern effects. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW 2024). https://openaccess.thecvf.com/content/CVPR2024W/WMF/papers/Tariq_Beyond_the_Screen_Evaluating_Deepfake_Detectors_under_Moire_Pattern_Effects_CVPRW_2024_paper.pdf (dostęp na dzień: 03.04.2026 r.).

W kontekście deepfake istotne znaczenie powinno mieć praktykowanie tzw. „trzech zasad pierwszego kontaktu". Po pierwsze, należy starać się – o ile to realnie możliwe – dążyć do pozyskania pierwotnego pliku (nie z komunikatora jako „zrzutu", lecz z urządzenia lub platformy w wersji źródłowej), bo brak metadanych i re-kompresja ograniczają pole działania i manewru biegłym i narzędziom analitycznych[7]. Po drugie, należy unikać generowania własnych „wtórników" (np. nagrywania ekranu), jeśli intencją jest weryfikacja autentyczności. Takie działania mogą bowiem wprowadzić artefakty, które zaburzają wynik automatycznej detekcji – zarówno w stronę fałszywego: „to jest deepfake", jak i fałszywego: „to jest OK – to nie jest deepfake"[8]. Wreszcie po trzecie, należy traktować detekcję narzędziową jako triage – powiedzmy: wstępne przesiewanie, a nie rozstrzygnięcie. Wynika to z faktu, że ewaluacje systemów wykrywania deepfake kładą nacisk na „ograniczenia uogólniania i odporność na post-processing"[9] – dodatkowo jest to silny argument przemawiający za łączeniem: analizy technicznej, kontekstu źródłowego oraz opinii biegłego.

Podsumowanie

Rozwój technologii i narzędzi umożliwiających tworzenie deepfake-ów stawia przed wymiarem sprawiedliwości złożone wyzwania, które częstokroć wykraczają poza dotychczasowe schematy oceny dowodów. W warunkach rosnącej niepewności co do autentyczności materiałów cyfrowych kluczowe znaczenie zyskuje nie tylko wiedza techniczna i specjalistyczna, ale także odpowiednie procedury i krytyczne podejście do źródeł.

W kolejnej części newsletterowego artykułu przejdziemy od ram teoretycznych do analizy praktycznej, koncentrując się na dostępnych na rynku narzędziach służących zarówno do tworzenia, jak i wykrywania deepfake'ów.

Dariusz Szawurski-Radetz
Monika Kania
Sekcja-Pracownia VR/AI

Grafika- wygenererowana przez AI

[7] Tutaj za: Guttman, B., White, D. R., & Walraven, T. (2022). Digital evidence preservation: Considerations for evidence handlers (NIST Interagency Report 8387). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.IR.8387 (dostęp na dzień: 03.04.2026 r.).
[8] Zob. R. Tariq, M. Heo, S. S. Woo, S. Tariq, Beyond the Screen..., op. cit.
[9] Zob.: Zob. H. Guan, J. Horan, A. Zhang, Guardians of Forensic Evidence..., op. cit.

Słownik wybranych pojęć AI

Deepfake
Technika wykorzystująca sztuczną inteligencję do tworzenia realistycznych, lecz fałszywych nagrań audio, wideo lub obrazów.

Generatywna sztuczna inteligencja
(Generative AI, GenAI)
Rodzaj AI zdolny do tworzenia nowych treści (tekstu, obrazu, dźwięku, wideo) na podstawie wzorców wyuczonych z danych treningowych.

Model generatywny
(generative model)
Model AI, który nie tylko analizuje dane, ale generuje nowe treści – np. twarze, głosy lub całe nagrania.

GAN
(Generative Adversarial Network)
Architektura AI składająca się z dwóch modeli (generatora i dyskryminatora), które „rywalizują" ze sobą, co prowadzi do tworzenia coraz bardziej realistycznych deepfake'ów.

Klonowanie głosu
(voice cloning)
Technologia umożliwiająca odtworzenie głosu konkretnej osoby (głos syntetyczny) na podstawie próbek audio, często wykorzystywana w deepfake'ach dźwiękowych.

Manipulacja multimodalna
(multimodal manipulation)
Jednoczesna manipulacja obrazem, dźwiękiem i tekstem w celu stworzenia spójnego, lecz fałszywego przekazu.

Artefakty cyfrowe
(digital artifacts)
Nienaturalne elementy obrazu lub dźwięku (np. zniekształcenia twarzy, nielogiczne cienie), które mogą wskazywać na użycie deepfake.

Detekcja deepfake
(deepfake detection)
Zestaw metod (technicznych i analitycznych) służących do identyfikacji fałszywych materiałów – np. analiza metadanych, pikseli lub niespójności ruchu.

Autentyczność cyfrowa
(digital authenticity)
Zdolność do potwierdzenia, że dany materiał (np. nagranie) jest oryginalny i nie został zmanipulowany.

Łańcuch dowodowy
(chain of custody)
Udokumentowany proces zabezpieczenia i przechowywania materiału dowodowego, zapewniający jego integralność i wiarygodność – kluczowy przy dowodach cyfrowych, w tym deepfake.

HARMONOGRAM AI ACT

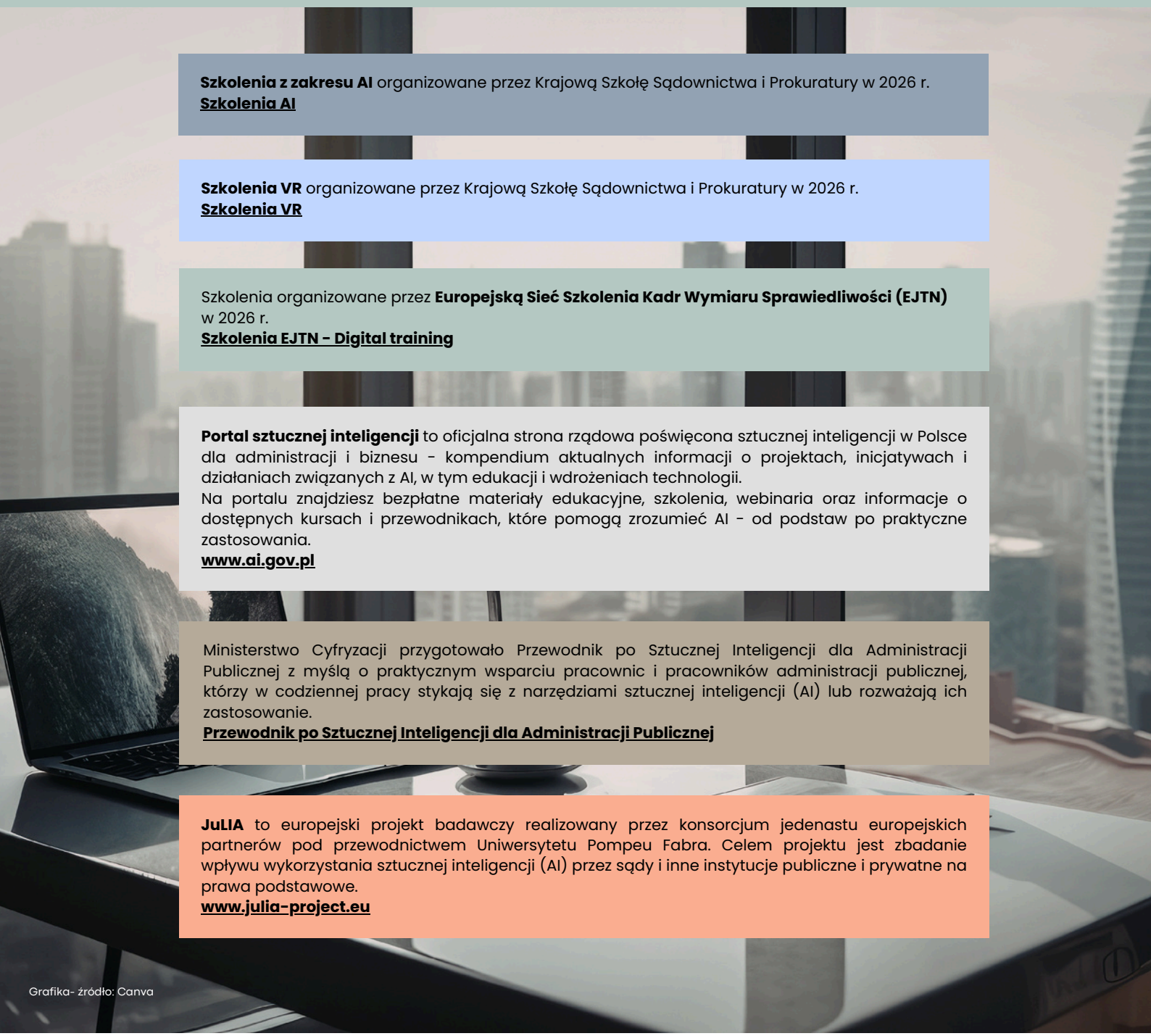
2024	**Fundamenty regulacji**
12 lipca 2024 r.	AI Act opublikowany w Dzienniku Urzędowym UE
1 sierpnia 2024 r.	AI Act wchodzi w życie. Większość zasad będzie wdrażana etapami.
2025	**Pierwsze obowiązki**
2 lutego 2025 r.	Pierwsze przepisy zaczynają obowiązywać: • zakaz stosowania systemów AI o niedopuszczalnym ryzyku (unacceptable risk); • obowiązki związane z AI literacy (podnoszenie świadomości i kompetencji).
2 sierpnia 2025 r.	Kolejne kluczowe przepisy zaczynają obowiązywać: • obowiązki dla dostawców ogólnego przeznaczenia (GPAI); • państwa członkowskie są zobowiązane powołać krajowe organy nadzorcze oraz ustanowić ramy kar.
2026	**Główna faza wejścia w życie**
2 lutego 2026 r.	Termin na wytyczne dotyczące np. planów monitorowania AI (przykładowe wymagania wykonawcze)
2 sierpnia 2026 r.	Większość zasad AI Act zaczyna obowiązywać: • zasady dotyczące systemów wysokiego ryzyka (high-risk); • przepisy o transparencji; • tworzenie co najmniej jednego regulatora sandbox w każdym państwie.
2027	**Dalsze etapy**
2 sierpnia 2027 r.	Obowiązywanie przepisów dotyczących wysokiego ryzyka wbudowanego w regulowane produkty.

Grafika- źródło: Canva

Strefa Szkoleń AI

Chcesz rozwijać kompetencje w obszarze sztucznej inteligencji? W tej rubryce polecamy sprawdzone kursy i szkolenia z zakresu AI - od poziomu podstawowego po bardziej zaawansowane zagadnienia. Wybieramy materiały praktyczne, aktualne i wartościowe, które możesz realnie wykorzystać w pracy oraz w codziennym rozwoju zawodowym.

Otwórz drzwi do nowych kompetencji i możliwości!



Szkolenia z zakresu AI organizowane przez Krajową Szkołę Sądownictwa i Prokuratury w 2026 r.
Szkolenia AI

Szkolenia VR organizowane przez Krajową Szkołę Sądownictwa i Prokuratury w 2026 r.
Szkolenia VR

Szkolenia organizowane przez **Europejską Sieć Szkolenia Kadr Wymiaru Sprawiedliwości (EJTN)** w 2026 r.
Szkolenia EJTN - Digital training

Portal sztucznej inteligencji to oficjalna strona rządowa poświęcona sztucznej inteligencji w Polsce dla administracji i biznesu - kompendium aktualnych informacji o projektach, inicjatywach i działaniach związanych z AI, w tym edukacji i wdrożeniach technologii.
Na portalu znajdziesz bezpłatne materiały edukacyjne, szkolenia, webinaria oraz informacje o dostępnych kursach i przewodnikach, które pomogą zrozumieć AI - od podstaw po praktyczne zastosowania.
www.ai.gov.pl

Ministerstwo Cyfryzacji przygotowało Przewodnik po Sztucznej Inteligencji dla Administracji Publicznej z myślą o praktycznym wsparciu pracownic i pracowników administracji publicznej, którzy w codziennej pracy stykają się z narzędziami sztucznej inteligencji (AI) lub rozważają ich zastosowanie.
Przewodnik po Sztucznej Inteligencji dla Administracji Publicznej

JuLIA to europejski projekt badawczy realizowany przez konsorcjum jedenastu europejskich partnerów pod przewodnictwem Uniwersytetu Pompeu Fabra. Celem projektu jest zbadanie wpływu wykorzystania sztucznej inteligencji (AI) przez sądy i inne instytucje publiczne i prywatne na prawa podstawowe.
www.julia-project.eu

Grafika- źródło: Canva

KONTAKT

81 458 37 81, 83

pracownia@kssip.gov.pl

www.kssip.gov.pl

Newsletter PRACOWNI VR/AI
Opracowanie tekstu i zdjęcia:
Dorota JĘDRASIK, Monika KANIA, Dariusz SZAWURSKI-RADETZ

Projekt graficzny:
Monika KANIA