„Krytyka Prawa", tom 18, nr 1/2026, s. 115–130, ISSN 2080-1084, e-ISSN 2450-7938, © 2026 Author.

MONIKA GÓRAL[1]

Problematyka „wytrenowania" generatywnej sztucznej inteligencji w świetle przepisów o ochronie danych osobowych[2]

Wpłynął: 10.07.2024. Akceptacja: 15.12.2025

Streszczenie

W artykule poddano badaniu aspekty prawne dotyczące „wytrenowania" generatywnej sztucznej inteligencji w kontekście obowiązujących przepisów ochrony danych osobowych. Zidentyfikowano problemy występujące między zapewnieniem prywatności i zgodności regulacyjnej a wymogami technologicznymi z uwzględnieniem samego procesu technologicznego. Zwrócono uwagę na analizę ryzyka prawnego, obowiązujące normy prawne oraz na identyfikację wyzwań związanych z ochroną danych osobowych podczas „trenowania" generatywnej sztucznej inteligencji.

Słowa kluczowe: dane osobowe, sztuczna inteligencja.

[1] Monika Góral – Akademia Leona Koźmińskiego (Polska), e-mail: monika.goral6@gmail.com; ORCID: 0009-0007-7403-3218.

[2] Badania wykorzystane w artykule nie zostały sfinansowane przez żadną instytucję.

 DOI: 10.7206/kp.2080-1084.838

„Krytyka Prawa", tom 18, nr 1/2026, s. 115–130, ISSN 2080-1084, e-ISSN 2450-7938, © 2026 Author.

MONIKA GÓRAL

The issue of "training" generative artificial intelligence in light of personal data protection regulations[3]

Abstract

The article examines the legal aspects of 'training' generative artificial intelligence in the context of current data protection legislation. It identifies the issues that arise between ensuring data protection and regulatory compliance and the technological requirements of the technological process itself. Attention was paid to the analysis of legal risks, applicable legal standards and the identification of data protection challenges when "training" generative artificial intelligence.

Keywords: personal data, artificial intelligence.

[3] The research in this article has not been supported financially by any institution.

Wprowadzenie

Generatywna sztuczna inteligencja oznacza system sztucznej inteligencji zdolny do generowania tekstu, obrazów lub innych mediów w odpowiedzi na komunikaty. Modele generatywne bazują na wzorcach i strukturze danych wejściowych, a następnie generują nową zawartość podobną do danych treningowych, ale z pewnym stopniem nowości[4].

Dane osobowe stanowią cenny kapitał dla systemów generatywnej sztucznej inteligencji, a użycie ich w celu trenowania danego systemu może przynieść wiele korzyści dla optymalizacji oraz kompatybilności systemu. Skuteczność modeli sztucznej inteligencji w dużej mierze zależy od jakości otrzymywanych danych, co sprawia, że ochrona danych jest integralnym aspektem ich projektowania[5]. Warto jednak zaznaczyć, że wykorzystywanie danych osobowych podczas trenowania generatywnej sztucznej inteligencji wiąże się z licznymi ryzykami zarówno dla twórców modeli, jak i dla użytkowników. Dlatego twórcy modeli powinni zwracać szczególną uwagę na przepisy dotyczące prywatności i ochrony danych osobowych w celu zapewnienia zgodności regulacyjnej[6]. Dbałość o prywatność użytkowników oraz uczciwe i etyczne wykorzystanie danych osobowych powinny być zawsze priorytetem przy projektowaniu i wdrażaniu systemów opartych na sztucznej inteligencji.

Mając powyższe na uwadze, w niniejszym artykule poddano analizie zgodność procesu „trenowania" generatywnej sztucznej inteligencji z wykorzystaniem zwykłej kategorii danych osobowych z regulacją UE[7].

---

4 C. Stryker, M. Scapicchio, *What is generative AI?* Pozyskano z: https://www.ibm.com/topics/generative-ai (dostęp: 16.08.2025).

5 A. Aldoseri, K. N. Al-Khalifa, A. M. Hamouda, *Re-Thinking Data Strategy and Integration for Artificial Intelligence: Concepts, Opportunities, and Challenges*, s. 2. Pozyskano z: https://www.mdpi.com/2076-3417/13/12/7082 (dostęp: 16.08.2025).

6 https://lexdigital.pl/sztuczna-inteligencja (dostęp: 20.08.2025).

7 RODO dzieli dane osobowe na dwie kategorie: dane zwykłe oraz dane sensytywne. Dane sensytywne stanowią szczególną kategorię danych osobowych oraz wymagają wyższego poziomu ochrony ze względu na ich potencjalnie większy wpływ na prawa i wolności osób fizycznych.

 DOI: 10.7206/kp.2080-1084.838

Proces trenowania generatywnej sztucznej inteligencji dane osobowe

Proces trenowania generatywnej sztucznej inteligencji zaczyna się zazwyczaj od pozyskiwania danych treningowych. Dane te stanowią podstawę, na której model uczy się generować nowe, realistyczne dane[8]. Proces zbierania danych jest jednym z najbardziej krytycznych etapów, gdyż dane pochodzą z różnych źródeł oraz mogą na dodatek zawierać dane osobowe.

Modele generatywnej sztucznej inteligencji są zazwyczaj trenowane na ogromnych ilościach danych czerpanych z różnorodnych źródeł[9]. Te duże zestawy danych odgrywają kluczową rolę w generatywnej sztucznej inteligencji, ponieważ umożliwiają modelom wychwytywanie szerokiej gamy wzorców i różnorodności w danych, co skutkuje bardziej zróżnicowanymi i twórczymi wynikami[10]. Do trenowania są wykorzystywane m.in. dane publicznie dostępne w internecie, bazy danych i inne zasoby cyfrowe. Na tym etapie istnieje duże ryzyko pozyskania danych osobowych, które mogą być ukryte w treściach tekstowych, obrazach lub innych formach. Niesie to ze sobą liczne kontrowersje z zakresu ochrony prywatności oraz zgodności z przepisami ochrony danych osobowych.

Metody pozyskiwania danych do trenowania sztucznej inteligencji obejmują m.in.: *scraping*, *web crawling*, *crowdsourcing*, *synthetic data generation*, wykorzystanie publicznych baz danych oraz użycie zgromadzonych danych klientów[11]. Do najbardziej kontrowersyjnych metod pozyskiwania danych do trenowania sztucznej inteligencji należy zaliczyć *scraping* oraz *web crawling*.

Metoda *scraping* stanowi zautomatyzowany proces zbierania oraz przechowywania informacji, które następnie są poddawane analizie oraz ekstrakcji[12]. W przypadku użycia tej metody zostaje wstępnie określona lista adresów URL, z której następnie pozyskiwane są dane.

Z kolei *web crawling* jest w znacznej mierze używane w celu indeksowania stron na podstawie ich zawartości oraz tworzenia replik wszystkich odwiedzanych stron, które są następnie przetwarzane przez wyszukiwarkę. Sieć internetowa

---

8 S.H. Bach, B. He, A. Ratner, C. Ré, *Learning the Structure of Generative Models without Labeled Data*. Pozyskano z: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6417840/ (dostęp: 19.08.2025).

9 T. Davenport, M. Alavi, *How to Train Generative AI Using Your Company's Data*. Pozyskano z: https://hbr.org/2023/07/how-to-train-generative-ai-using-your-companys-data (dostęp: 19.08.2025).

10 https://medium.com/@rickspair/introduction-understanding-the-importance-of-data-in-generative-ai-9603780ab256 (dostęp: 19.08.2025).

11 C. Dilmegani, *6 Risks of Generative AI & How to Mitigate Them in 2025*. Pozyskano z: https://research.aimultiple.com/generative-ai-data/ (dostęp: 19.08.2025).

12 B. Zhao, *Web Scraping*. Pozyskano z: https://www.researchgate.net/publication/317177787_Web_Scraping (dostęp: 19.08.2025).

z użyciem niniejszej metody jest przeglądana w sposób zautomatyzowany, a zawartość poszczególnych stron internetowych jest pobierana i umieszczana w lokalnym repozytorium[13].

Po zebraniu danych zazwyczaj następuje ich przygotowanie do treningu. Proces ten powinien obejmować kilka podetapów, spośród których szczególnie istotne, zwłaszcza w kontekście ochrony danych osobowych, są czyszczenie danych oraz anonimizacja. Czyszczenie danych polega na usunięciu szumów, duplikatów i błędów, co jest kluczowe dla zapewnienia jakości danych treningowych[14]. Anonimizacja danych to proces, w którym informacje osobowe są nieodwracalnie zmieniane, aby uniemożliwić identyfikację osób, których one dotyczą. Jest to kluczowy etap dla zgodności z przepisami o ochronie danych osobowych, aby zapewnić zgodność regulacyjną.

Następny etap stanowi trening modelu generatywnej sztucznej inteligencji, który wymaga zastosowania odpowiednich algorytmów oraz architektur. Proces treningu obejmuje dostarczanie modelowi dużych ilości danych treningowych, na podstawie których model uczy się rozpoznawać wzorce i generować nowe dane. Na etapie treningu dane osobowe mogą znaleźć się w części zbioru treningowego. Ważne jest zatem, aby dane osobowe były odpowiednio zanonimizowane przed ich użyciem w procesie treningu, aby zminimalizować ryzyko naruszenia prywatności, jeżeli z jakichś przyczyn nie zostały poddane wcześniejszej anonimizacji.

Kolejne etapy powinny obejmować optymalizację oraz wdrożenie na rynek. Nie będą one jednak przedmiotem szerszej analizy w niniejszym artykule.

Z powyższego wynika, że na każdym etapie trenowania generatywnej sztucznej inteligencji mogą pojawić się dane osobowe, które niosą ze sobą ryzyko naruszenia prywatności oraz wymogi zgodności z przepisami prawnymi.

Zarówno w Rzeczypospolitej Polskiej, jak i w całej Unii Europejskiej kluczową regulacją w tym zakresie jest rozporządzenie Parlamentu Europejskiego i Rady (UE) 2016/679 z dnia 27 kwietnia 2016 r. w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych i w sprawie swobodnego przepływu takich danych oraz uchylenia dyrektywy 95/46/WE (ogólne rozporządzenie o ochronie danych) (dalej: RODO). Zawiera ona w sobie szereg unormowań, które przedsiębiorstwa europejskie muszą spełnić, aby z poszanowaniem prawa przetwarzać dane osobowe. Należy w tym kontekście również zaznaczyć, że RODO ma także zastosowanie do sytuacji, gdy przetwarzanie danych osobowych jest związane z działalnością prowadzoną przez jednostkę organizacyjną administratora

[13] M.A. Kausar, V.S. Dhaka, *Web Crawler: A Review*. Pozyskano z: https://www.researchgate.net/publication/258789938_Web_Crawler_A_Review (dostęp: 19.08.2025).

[14] https://www.puttingdatatowork.com/post/is-your-data-ready-for-generative-ai-a-comprehensive-guide (dostęp: 19.08.2025).

 DOI: 10.7206/kp.2080-1084.838

bądź podmiotu przetwarzającego w Unii, niezależnie od tego, czy przetwarzanie odbywa się w Unii[15].

Ponadto wyżej wymienione rozporządzenie będzie miało zastosowanie zgodnie z art. 3 ust. 2 RODO, m.in. „do przetwarzania danych osobowych osób, których dane dotyczą, przebywających w Unii przez administratora lub podmiot przetwarzający niemających jednostek organizacyjnych w Unii, jeżeli czynności przetwarzania wiążą się z:

a) oferowaniem towarów lub usług takim osobom, których dane dotyczą, w Unii – niezależnie od tego, czy wymaga się od tych osób zapłaty; lub
b) monitorowaniem ich zachowania, o ile do zachowania tego dochodzi w Unii".

Trenowanie sztucznej inteligencji a zasady RODO

Proces trenowania generatywnej sztucznej inteligencji, obejmujący gromadzenie, selekcję, przetwarzanie i wykorzystywanie danych, w sposób bezpośredni styka się z obszarem regulacji przewidzianym w ogólnym rozporządzeniu o ochronie danych (RODO). Szczególne znaczenie ma art. 5 RODO, ustanawiający podstawowe zasady przetwarzania danych osobowych, które mają charakter uniwersalny i znajdują zastosowanie niezależnie od rodzaju technologii czy modelu biznesowego. W doktrynie przyjmuje się, że zasady te pełnią funkcję normatywnego „kręgosłupa" całego systemu ochrony danych, determinując cele przetwarzania danych osobowych.

Trenowanie modeli generatywnych w szczególności narażone jest na zarzut naruszenia art. 5 RODO, ponieważ w naturalny sposób wiąże się z pozyskiwaniem ogromnych, często nieustrukturyzowanych zbiorów danych, wśród których mogą znajdować się informacje o osobach fizycznych. Wykorzystywanie metod takich jak *scraping* czy *web crawling* powoduje, że dane te trafiają do zbiorów treningowych bez wiedzy i zgody osób, których dotyczą, często w postaci niekompletnej, zdezaktualizowanej czy obarczonej błędem. Z perspektywy prawnej i etycznej rodzi to pytania nie tylko o legalność przetwarzania, ale również o rzetelność i przejrzystość całego procesu. Przetwarzanie niedokładnych danych osobowych może mieć realne konsekwencje dla osób, których dane dotyczą. Z racji tego dostawcy modeli generatywnej sztucznej inteligencji, np. OpenAI, ostrzegają użytkowników o niebezpieczeństwach z tym związanych i o tym, że nie można całkowicie ufać dokładności

[15] Art. 3 ust.1 RODO.

pobieranych danych[16]. Punktem wyjścia dla oceny zgodności prawnej każdego procesu przetwarzania danych osobowych powinno być zatem przestrzeganie nw. zasad RODO[17].

Zasada legalności, rzetelności i przejrzystości (art. 5 ust. 1 lit. a RODO) wymaga, by każde przetwarzanie opierało się na przesłance wskazanej w art. 6 ust. 1 RODO, a więc m.in. na zgodzie, wykonaniu umowy lub prawnie uzasadnionym interesie administratora. W praktyce uzyskanie zgody od szerokiego i anonimowego kręgu podmiotów jest niewykonalne, a powiązanie przetwarzania z umową – trudne do udowodnienia. Dlatego najczęściej rozważaną przesłanką staje się art. 6 ust. 1 lit. f wymagający przeprowadzenia testu równowagi interesów. Niezależnie jednak od wyboru podstawy prawnej obowiązek zapewnienia przejrzystości nakazuje przekazywanie osobom, których dane dotyczą, jasnych informacji o sposobie i celu przetwarzania (art. 12–14 RODO). W tym aspekcie należy zaznaczyć, że RODO normuje obowiązki informacyjne, które zostały określone w artykułach 12–15 RODO. Artykuł 14 RODO jest szczególnie istotny w kontekście generatywnej sztucznej inteligencji, ponieważ dotyczy on sytuacji, w których dane osobowe nie są pozyskiwane bezpośrednio od osoby, której dotyczą. Warto tutaj zwrócić szczególną uwagę na art. 14 ust. 5 lit. b RODO, który przewiduje wyjątek od obowiązku informacyjnego. Oznacza on, że w kontekście przetwarzania danych na potrzeby trenowania sztucznej inteligencji, jeżeli poinformowanie każdej osoby, której dane dotyczą, byłoby niewspółmiernie trudne lub niemożliwe, administrator może być zwolniony z tego obowiązku, pod warunkiem podjęcia odpowiednich środków ochrony. Jednakże nawet w takich sytuacjach administrator powinien dążyć do maksymalnej przejrzystości i starać się informować osoby, których dane dotyczą, w sposób pośredni, np. poprzez publiczne udostępnianie informacji na swojej stronie internetowej. Dzięki temu przepisowi administratorzy mogą skupić się na rozwijaniu technologii sztucznej inteligencji, pod warunkiem, że podejmą odpowiednie środki ochrony praw i wolności osób, których dane są przetwarzane. Jednakże należy w tym aspekcie podkreślić, że zastosowanie w celach komercyjnych wyjątku, jaki określa art. 14 ust. 5 pkt b RODO, jest wysoce sporne, zwłaszcza jeśli pozyskanie danych osobowych odbywa się za pomocą metody *scraping*[18].

Zasada ograniczenia celu (art. 5 ust. 1 lit. b RODO) zasadniczo zabrania wykorzystywania danych w celach innych niż pierwotne. W przypadku SI zagrożenie

---

16 https://openai.com/policies/eu-terms-of-use/ (dostęp: 16.08.2025).

17 S. Kowalski, *Sztuczna inteligencja a ochrona danych osobowych*, [w:] *Metaświat prawne i techniczne aspekty przełomowych technologii*, red. R. Bieda, Z. Okoń, Warszawa 2022, s. 351.

18 C. Novelli, F. Casolari, P. Hacker, G. Spedicato, L. Floridi, *Generative AI in EU Law: Liability, Privacy, Intellectual Property, and Cybersecurity*. Pozyskano z: https://arxiv.org/pdf/2401.07348, s. 11 (dostęp: 16.08.2025).

 DOI: 10.7206/kp.2080-1084.838

naruszenia tej zasady jest szczególnie duże, ponieważ dane pozyskane np. do obsługi klienta mogą zostać użyte do trenowania algorytmu marketingowego, co w świetle prawa wymagałoby dodatkowej podstawy przetwarzania.

Zasada minimalizacji danych (art. 5 ust. 1 lit. c RODO) zobowiązuje administratora do ograniczania danych do zakresu niezbędnego. W przypadku SI wymaga to stosowania wstępnej selekcji danych, anonimizacji, pseudonimizacji czy metod takich jak *federated learning*, które pozwalają ograniczyć centralne gromadzenie danych. Brak takich działań prowadzi do gromadzenia nadmiarowych informacji, co zwiększa ryzyko naruszeń. Motyw 28 RODO wprost wskazuje na potrzebę stosowania pseudonimizacji jako środka ograniczającego ryzyko.

Zasada prawidłowości (art. 5 ust. 1 lit. d RODO) nakłada obowiązek zapewnienia, by dane były aktualne i poprawne. W praktyce oznacza to konieczność walidacji zbiorów przed treningiem, ponieważ wykorzystanie niezweryfikowanych lub przestarzałych danych może skutkować generowaniem przez model treści błędnych, krzywdzących czy dyskryminujących. Jest to szczególnie istotne w sektorach wrażliwych, takich jak ochrona zdrowia czy wymiar sprawiedliwości, gdzie skutki decyzji opartych na wadliwych danych mogą być nieodwracalne.

Zasada ograniczenia przechowywania (art. 5 ust. 1 lit. e RODO) nakazuje usuwanie lub anonimizację danych, gdy przestają być niezbędne do celu, dla którego je zebrano. W trenowaniu SI pojawia się pokusa ich długotrwałego przechowywania w celu ponownego użycia przy modyfikacji modelu, co bez jasnej i zgodnej z prawem polityki retencji może stanowić naruszenie.

Zasada integralności i poufności (art. 5 ust. 1 lit. f RODO), w powiązaniu z art. 32 RODO, wymaga stosowania adekwatnych zabezpieczeń technicznych i organizacyjnych. W kontekście SI oznacza to m.in. ochronę przed atakami pozwalającymi na odtworzenie danych treningowych (np. *model inversion*) oraz wdrożenie kontroli dostępu i mechanizmów szyfrowania.

Z kolei zasada rozliczalności (art. 5 ust. 2 RODO) zobowiązuje administratora do wykazania zgodności wszystkich etapów trenowania z prawem, co w praktyce wymaga prowadzenia rejestrów czynności przetwarzania (art. 30 RODO), dokumentacji DPIA (art. 35 RODO) oraz archiwizacji decyzji dotyczących metod przetwarzania i zabezpieczeń.

W rezultacie konieczność stosowania art. 5 RODO do trenowania generatywnej sztucznej inteligencji wymaga ścisłej integracji wiedzy prawnej i technologicznej. Samo formalne powołanie się na zasady nie gwarantuje zgodności, ponieważ konieczne jest ich rzeczywiste przełożenie na procedury operacyjne i architekturę systemu. Zaniedbanie któregokolwiek z elementów może skutkować nie tylko naruszeniem prawa, ale także utratą zaufania publicznego, co w przypadku technologii o rosnącym wpływie społecznym ma znaczenie fundamentalne.

Właściwa podstawa prawna przetwarzania

Wybór odpowiedniej podstawy prawnej przetwarzania danych zgodnie z RODO jest kluczowym elementem procesu „trenowania" sztucznej inteligencji. Podstawy te normuje art. 6 RODO, który warunkuje zapewnienie zgodności z prawem poprzez spełnienie co najmniej jednej z nich. Poniżej poddano analizie trzy najbardziej „pozornie" odpowiednie podstawy prawne, które mogłyby mieć zastosowanie przy trenowaniu generatywnej sztucznej inteligencji oraz zaakcentowano główne wyzwania z nimi związane.

Pierwszą z nich jest zgoda osoby, której dane dotyczą. Uzyskanie takiej zgody jest jednym z najtrudniejszych wyzwań w kontekście trenowania modeli sztucznej inteligencji. Ponadto zgoda zgodnie z RODO musi być dobrowolna, konkretna, świadoma i jednoznaczna, co w praktyce oznacza konieczność dokładnego informowania osób o celu przetwarzania ich danych. Uzyskanie zgody od dużej liczby osób może być logistycznie i finansowo nieopłacalne[19]. W przypadku dużych zbiorów danych, w tym danych osobowych pochodzących od ogromnej grupy osób nieznanych wcześniej deweloperom i podmiotom tworzącym i rozwijającym modele sztucznej inteligencji, uzyskanie ważnej zgody od każdej osoby zazwyczaj nie jest możliwe ze względu na wygórowane koszty transakcji[20]. Ponadto, osoby wyrażające zgodę muszą w pełni rozumieć, jak ich dane będą wykorzystywane, co może być skomplikowane w przypadku zaawansowanych technologii wykorzystywanych w sztucznej inteligencji. Poza tym takie osoby mają również prawo w każdej chwili wycofać zgodę, co może skomplikować proces przetwarzania danych i wpłynąć na integralność całego modelu. Co do zasady niniejsza przesłanka ma bardzo ograniczone zastosowanie w szkoleniu generatywnych systemów sztucznej inteligencji w obecnym kształcie, co wynika w znacznej mierze ze sposobu pozyskiwania danych za pomocą takiej metody jak ww. *scraping*.

Kolejną poddaną analizie podstawą jest niezbędność przetwarzania do wykonania umowy. Zastosowanie tej podstawy prawnej wymaga m.in., aby przetwarzanie danych osobowych było bezpośrednio związane z wykonaniem umowy, której stroną jest osoba, której dane dotyczą. W kontekście trenowania sztucznej inteligencji często trudno jest powiązać przetwarzanie danych z konkretną umową. Podstawa ta mogłaby mieć zastosowanie, gdy korzystanie z systemu sztucznej inteligencji byłoby przedmiotem umowy zawartej między operatorem sztucznej inteligencji a użytkownikiem i gdy nie ma innego sposobu wykonania tej umowy

[19] C. Novelli, F. Casolari, P. Hacker, G. Spedicato, L. Floridi, *op. cit*.

[20] J. Barugahare, *Bioethical reflexivity and requirements of valid consent: conceptual tools*. Pozyskano z: https://bmcme dethics.biomedcentral.com/articles/10.1186/s12910-019-0385-7, s. 1 (dostęp: 19.08.2025).

 DOI: 10.7206/kp.2080-1084.838

niż szkolenie sztucznej inteligencji z wykorzystaniem danych użytkownika. Ponadto w odniesieniu do drugiej gałęzi tej podstawy prawnej, tj. kroków poprzedzających zawarcie umowy, jej zastosowanie wymagałoby wykazania, że osoba, której dane dotyczą, złożyła wniosek w kontekście potencjalnego zawarcia umowy i że nie ma innego sposobu na spełnienie jej żądań niż wyszkolenie (a nie tylko wykorzystanie raz wyszkolonej) sztucznej inteligencji. Jedną z przeszkód dla tej podstawy jest brak bezpośredniej relacji twórców z użytkownikami, co sprawia, że trenowanie modeli często obejmuje dane osób, które nie są bezpośrednimi stronami żadnej umowy z twórcami modeli[21]. Ponadto trudności budzi również zdefiniowanie samego zakresu umowy, co powoduje trudność w jednoznacznym określeniu, jakie przetwarzanie danych jest niezbędne do wykonania danej umowy.

Najczęściej przywoływaną możliwą podstawą prawną jest natomiast prawnie uzasadniony interes administratora. Do zastosowania niniejszej podstawy prawnej konieczne jest spełnienie trzech kumulatywnych przesłanek przed rozpoczęciem przetwarzania danych osobowych:

a) istnienie prawnie uzasadnionego interesu,
b) niezbędność przetwarzania do celów wynikających z prawnie uzasadnionych interesów,
c) wobec wskazanych interesów nadrzędnego charakteru nie mają interesy lub podstawowe prawa oraz wolności osoby, której dane dotyczą, wymagające ochrony danych osobowych.

Zgodnie z ww. lit. c należałoby przeprowadzić test równowagi, który ma na celu weryfikację, czy podstawowe prawa, wolności oraz interesy osób, których dane dotyczą są nadrzędne względem prawnie uzasadnionego interesu administratora[22].

Ponadto, zgodnie z wytycznymi opublikowanymi przez Holenderski Urząd Ochrony Danych (AP), tylko prawnie chronione interesy kwalifikują się jako uzasadnione interesy, co oznacza, że interes administratora musi wynikać z przepisów prawnych. Jednakże może to również być niepisana reguła prawna lub zasada prawna. Poza tym zostało też podkreślone, że prawnie uzasadniony interes nie może wynikać tylko z interesów handlowych danego podmiotu.

Z perspektywy prawnej i ekonomicznej szkolenia związane ze sztuczną inteligencją często muszą opierać się na teście balansującym określonym w art. 6 ust. 1 lit. f RODO. Aby stwierdzić, czy tzw. test balansujący (test równowagi interesów,

[21] https://cedpo.eu/wp-content/uploads/generative-ai-the-data-protection-implications-16-10-2023.pdf, s. 8 (dostęp:19.08.2025).

[22] *Ibidem*, s. 9–10.

o którym mowa w art. 6 ust. 1 lit. f RODO) stanowi właściwą podstawę prawną, konieczna jest indywidualna analiza każdego przypadku. Przykłady zastosowań korzystnych społecznie mogą wzmacniać argumentację na korzyść deweloperów. Jednak takie kryterium rzadko jest spełniane. Strategie zwiększające prywatność, takie jak przejrzystość, pseudonimizacja czy szyfrowanie brane pod uwagę w ramach testu balansującego mają istotne znaczenie dla legalności trenowania sztucznej inteligencji[23].

Warto jednak podkreślić, iż uzasadniony interes nie stanowi solidnej podstawy prawnej dla przetwarzania danych w celu wytrenowania generatywnej sztucznej inteligencji, biorąc pod uwagę potrzebę przeprowadzenia oceny uzasadnionych interesów, a także fakt, że osoby, których dane dotyczą, mogą w dowolnym momencie sprzeciwić się takiemu przetwarzaniu. Ponadto, w doktrynie istnieje znaczna wątpliwość, czy niniejsza podstawa prawna może być zastosowana w celach komercyjnych.

Możliwe rozwiązania aby zapewnić zgodność z RODO

Nieustanny postęp technologiczny stwarza nowe wyzwania regulacyjne, szczególnie w obszarze ochrony danych osobowych i rozwoju technologii generatywnej sztucznej inteligencji. W miarę jak polskie przedsiębiorstwa dążą do zapewnienia konkurencyjności na globalnym rynku, stają przed koniecznością balansowania między innowacyjnością a zgodnością z obowiązującymi przepisami. Poniżej przedstawiono najważniejsze, zdaniem autorki, innowacyjne podejścia oraz propozycje zapewnienia zgodności z RODO podczas trenowania generatywnej sztucznej inteligencji[24].

Jednocześnie należy pamiętać, że dynamiczny rozwój SI nie ogranicza się wyłącznie do kwestii prawnych – towarzyszą mu również wyzwania praktyczne, takie jak konieczność ochrony danych osobowych, zapewnienie cyberbezpieczeństwa czy przeciwdziałanie wykorzystaniu algorytmów do szerzenia dezinformacji[25].

Zgodność z RODO nie jest stanem jednorazowo osiągniętym, lecz procesem ciągłym, wymagającym regularnego przeglądu stosowanych środków technicznych i organizacyjnych (art. 24 i 25 RODO). Zgodnie z motywem 78 preambuły

---

23 C. Novelli, F. Casolari, P. Hacker, G. Spedicato, L. Floridi, *op. cit*., s. 9.

24 Jednakże trzeba w tym aspekcie zaznaczyć, że sam proces zbierania danych osobowych, jeżeli mają one być następnie wykorzystane z użyciem opisanych technik wymaga odpowiedniej podstawy prawnej na gruncie RODO.

25 M. Frączek, K. Spaliński, *Bezpieczeństwo informacji w dobie sztucznej inteligencji: analiza ryzyka i wyzwania związane z ChatGPT*, „Zeszyty Naukowe Pro Publico Bono" 2023, 1(1), s. 293.

 DOI: 10.7206/kp.2080-1084.838

środki te powinny być projektowane w sposób uwzględniający ryzyko naruszenia praw i wolności osób fizycznych, a także charakter, zakres, kontekst i cele przetwarzania. W przypadku trenowania modeli SI szczególnego znaczenia nabiera połączenie mechanizmów prawnych z rozwiązaniami technicznymi, które ograniczają zakres i skutki ewentualnych naruszeń. Nie można jednak zapominać, że równie istotnym elementem pozostaje etyka – zarówno na etapie tworzenia i wdrażania systemów sztucznej inteligencji, jak i podczas ich praktycznego wykorzystania. To człowiek ponosi odpowiedzialność za to, by standardy etyczne były respektowane od fazy badań i rozwoju, przez implementację rozwiązań, aż po bieżące monitorowanie i nadzór nad ich funkcjonowaniem[26].

Pierwszym z nich jest *federated learning*. Stanowi ono podejście do uczenia maszynowego, w którym wiele podmiotów współpracuje w celu treningu modelu maszynowego pod nadzorem centralnego serwera, jednocześnie utrzymując dane treningowe zdecentralizowane[27]. Oznacza to, że dane surowe pozostają na urządzeniach lokalnych i nie są wymieniane ani transferowane, a zamiast tego przesyłane są jedynie zaktualizowane informacje, które są następnie agregowane, aby osiągnąć cel uczenia. Metoda *federated learning* jest kluczowa w kontekście trenowania generatywnej sztucznej inteligencji na danych osobowych z powodu m.in. ochrony prywatności oraz redukcji ryzyka bezpieczeństwa. W tradycyjnych metodach uczenia maszynowego dane osobowe muszą być przesyłane do centralnego serwera w celu przetwarzania, co wiąże się z ryzykiem naruszenia prywatności. *Federated learning* eliminuje potrzebę przesyłania surowych danych do centralnego miejsca, minimalizując ryzyko wycieków danych i zapewniając, że dane osobowe pozostają na urządzeniu użytkownika[28]. Należy zauważyć, że rozwiązanie to w sposób bezpośredni realizuje zasadę minimalizacji danych (art. 5 ust. 1 lit. c RODO), a także wspiera zasadę integralności i poufności (art. 5 ust. 1 lit. f RODO). Może być ono traktowane jako implementacja koncepcji *privacy by design* i *privacy by default* (art. 25 RODO). W świetle wytycznych Europejskiej Rady Ochrony Danych z 2020 r. dotyczących nowych technologii zastosowanie *federated learning* powinno być jednak uzupełnione o dodatkowe środki, takie jak uwierzytelnianie wieloskładnikowe, szyfrowanie parametrów modelu podczas transmisji oraz regularne testy penetracyjne, w celu wykrycia i neutralizacji ewentualnych luk bezpieczeństwa.

Niemniej jednak metoda ta wymaga do zapewniania pełnej zgodności regulacyjnej dodatkowych środków ochrony, takich jak: anonimizacja i szyfrowanie. Mimo licznych zalet *federated learning* napotyka także pewne wyzwania i ograniczenia.

---

[26] W. Ilnicka, *Sztuczna inteligencja –korzyści i zagrożenia*, „Zeszyty Naukowe Wyższej Szkoły Bankowej w Poznaniu" 2024, 105(2), s. 41.

[27] https://arxiv.org/pdf/1912.04977, s. 4–5 (dostęp: 19.08.2025).

[28] *Ibidem*, s. 37–40.

Wśród najczęstszych ograniczeń tej metody należy wymienić heterogeniczność danych, która polega na tym, iż dane na różnych urządzeniach mogą być niejednorodne, co może utrudniać trenowanie modeli. Poza tym mimo że dane nie są centralizowane, istnieje ryzyko ataków na urządzenia końcowe, co może zagrozić integralności modeli[29]. Jednak mimo tych wyzwań *federated learning* ma ogromny potencjał i już teraz znajduje zastosowanie w wielu dziedzinach.

Kolejnym innowacyjnym rozwiązaniem jest *differential privacy*, która polega na dodawaniu kontrolowanego szumu do danych w taki sposób, aby zapewnić prywatność indywidualnych danych jednostek w ogólnym zbiorze danych[30]. W kontekście przepisów o ochronie danych osobowych *differential privacy* pomaga w minimalizacji ryzyka identyfikacji osób na podstawie danych przetwarzanych przez generatywne modele sztucznej inteligencji[31]. Pod względem prawnym *differential privacy* wpisuje się zasadniczo w obowiązki z art. 32 RODO dotyczące bezpieczeństwa przetwarzania oraz art. 89 ust. 1 RODO, regulujący przetwarzanie w celach badawczych i statystycznych. Metoda ta może również wspierać realizację zasady rozliczalności (art. 5 ust. 2 RODO), ponieważ pozwala na wykazanie, że administrator wdrożył mierzalne i zweryfikowane środki ochrony. W doktrynie podkreśla się, że przewagą tego podejścia jest możliwość formalnego oszacowania poziomu prywatności przy użyciu parametru epsilon[32]. Jednak w świetle RODO dobór tego parametru musi wynikać z analizy ryzyka przeprowadzonej w ramach oceny skutków dla ochrony danych (art. 35 RODO), co zapewnia zgodność z wymogami motywu 76 preambuły, zgodnie z którym prawdopodobieństwo i powaga ryzyka naruszenia praw lub wolności osoby, której dane dotyczą, powinny być określane poprzez odniesienie do charakteru, zakresu, kontekstu i celów przetwarzania. Ryzyko należy szacować na podstawie obiektywnej oceny, pozwalającej stwierdzić, czy operacje przetwarzania wiążą się z ryzykiem lub wysokim ryzykiem.

Jednym z istotnych aspektów prawnych, które *differential privacy* pomaga spełnić, jest zgodność z zasadą rozliczalności określoną w RODO. *Differential privacy*, jako metoda matematycznie udowodniona i oparta na solidnych podstawach teoretycznych, pozwala na dokumentowanie i wykazywanie zgodności z przepisami. W kontekście odpowiedzialności prawnej metoda ta oferuje dodatkowy poziom ochrony przed potencjalnymi roszczeniami. Nawet jeśli dane są wykorzystywane

[29] *Ibidem*, s. 11–14, 62–88.

[30] C. Dwork, A. Roth, The Algorithmic Foundations of „differential privacy", "Foundations and Trends in Theoretical Computer Science" 2014, 9(3–4) (2014) 211–407.

[31] Opinion 05/2014 on Anonymisation Techniques. Pozyskano z: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf, s. 15 (dostęp: 19.08.2025).

[32] Zob. szerzej, C. Dwork, A. Roth, *The Algorithmic Foundations…*

 DOI: 10.7206/kp.2080-1084.838

przez strony trzecie, poziom dodanego szumu sprawia, że indywidualne informacje są ukryte, co redukuje ryzyko związane z naruszeniem prywatności[33]. Jednakże implementacja *differential privacy* nie jest pozbawiona wyzwań prawnych. Jednym z głównych problemów jest określenie odpowiedniego poziomu szumu (epsilon), który zapewni zgodność z przepisami, a jednocześnie pozwoli na uzyskanie użytecznych wyników z analizy danych[34]. W obliczu rosnących wymagań regulacyjnych dotyczących ochrony danych osobowych *differential privacy* staje się nieodzownym elementem strategii ochrony danych w nowoczesnych systemach analizy i przetwarzania danych.

Skuteczne zapewnienie zgodności trenowania generatywnej sztucznej inteligencji z RODO wymaga łączenia metod takich jak *federated learning* i *differential privacy* z kompleksowym systemem zarządzania ochroną danych. Powinno to obejmować m.in. jasne procedury retencji danych, kontrolę dostępu opartą na rolach, mechanizmy monitorowania incydentów oraz regularne audyty zgodności. Tylko takie holistyczne podejście pozwala na rzeczywiste, a nie jedynie deklaratywne, wypełnienie obowiązków wynikających m.in. z art. 5, 24, 25 i 32 RODO.

Podsumowanie

Dane, wśród których mogą znajdować się dane osobowe, mają kluczowe znaczenie dla sukcesu rozwoju narzędzi generatywnej sztucznej inteligencji. Problematyka wytrenowania modeli sztucznej inteligencji na takich danych w kontekście zgodności z RODO stawia przed nami szereg wyzwań. Podnoszony niekiedy pogląd iż „problem wysokich standardów RODO, które mogą ograniczać innowacyjność w zakresie SI, wydaje się jednak niezbędnym i proporcjonalnym kosztem, którego poniesienie umożliwia pozyskiwanie zaufania użytkowników SI, a dzięki temu sprawne wprowadzanie nowych rozwiązań" jest moim zdaniem w perspektywie rozwoju technologii problematyczny[35].

Przeprowadzone analizy prowadzą do wniosku, że obecne regulacje prawne mogą w istotny sposób ograniczać rozwój technologii sztucznej inteligencji, w szcze-

---

33 R. Cummings, D. Desfontaines, D. Evans, R. Geambasu, Y. Huang, M. Jagielski, P. Kairouz, G. Kamath, S. Oh, O. Ohrimenko, N. Papernot, R. Rogers, M. Shen, S. Song, W. Su, A. Terzis, A. Thakurta, S. Vassilvitskii, Y. Wang, L. Xiong, S. Yekhanin, D. Yu, H. Zhang, W. Zhang, *Advancing „differential privacy": Where We Are Now and Future Directions for Real-World Deployment*. Pozyskano z: https://hdsr.mitpress.mit.edu/pub/sl9we8gh/release/3 (dostęp: 19.08.2025).

34 https://utrechtuniversity.github.io/dataprivacyhandbook/differential-privacy.html (dostęp: 19.08.2025).

35 M. Wróblewski, *Sztuczna inteligencja a ochrona praw człowieka*, [w:] *Metaświat prawne i techniczne aspekty przełomowych technologii*, red. R. Bieda, Z. Okoń, Warszawa 2022, s. 334.

gólności w kontekście przetwarzania danych osobowych. Jednak brak pewnej i jednoznacznej podstawy prawnej może prowadzić do ograniczenia wdrożenia innowacyjnych systemów lub nawet do zaniechania wprowadzenia ich na rynek europejski[36].

RODO, zgodnie z zamiarem legislatora, jest neutralne technologicznie. Liczne obowiązki nałożone przez ten akt stanowią jednak pewną barierę, zwłaszcza dla przedsiębiorstw działających na rynku europejskim.

Istotne jest zatem wyważenie praw jednostek oraz potrzeb innowacji technologicznych, aby nie doprowadzić do sytuacji, w której przedsiębiorstwa naruszające prawo lub znajdujące się na jego granicy pod względem ochrony danych osobowych mają bardziej konkurencyjną pozycję na rynku europejskim w porównaniu z tymi, które starają się w jak największym stopniu zapewnić zgodność z restrykcyjnymi wymogami RODO.

Należy jednak zauważyć, że choć wysokie standardy RODO mogą ograniczać rozwój generatywnej sztucznej inteligencji, stanowią one niezbędny mechanizm ochrony praw jednostek. Aby pogodzić te dwa cele, konieczne jest dalsze doprecyzowanie przepisów RODO oraz rozwijanie technologii w sposób ich nienaruszający, co pozwoli na tworzenie innowacyjnych rozwiązań, które będą zarazem bezpieczne i respektujące prywatność użytkowników.

Bibliografia

Aldoseri A., Al-Khalifa K.N., Hamouda A. M, *Re-Thinking Data Strategy and Integration for Artificial Intelligence: Concepts, Opportunities, and Challenges*. Pozyskano z: https://www.mdpi.com/2076-3417/13/12/7082 (dostęp: 19.08.2025).

Bach S.H., He B., Ratner A., Ré C., *Learning the Structure of Generative Models without Labeled Data*. Pozyskano z: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6417840/ (dostęp: 19.08.2025).

Barugahare J., *Bioethical reflexivity and requirements of valid consent: conceptual tools*. Pozyskano z: https://bmcmedethics.biomedcentral.com/articles/10.1186/s12910-019-0385-7 (dostęp: 19.08.2025).

Cummings R., Desfontaines D., Evans D., Geambasu R., Huang Y., Jagielski M., Kairouz P., Kamath G., Oh S., Ohrimenko O., Papernot N., Rogers R., Shen M., Song S., Su W., Terzis A., Thakurta A., Vassilvitskii S., Wang Y., Xiong L., Yekhanin S., Yu D., Zhang H., Zhang W., *Advancing „differential privacy": Where We Are Now and Future*

[36] Zostało tutaj założone, że w danych treningowych na skutek zawężenia parametrów pozyskiwania danych nie znajdą się dane Europejczyków.

Directions for Real-World Deployment. Pozyskano z: https://hdsr.mitpress.mit.edu/pub/sl9we8gh/release/3

Davenport T., Alavi M., *How to Train Generative AI Using Your Company's Data*, https://hbr.org/2023/07/how-to-train-generative-ai-using-your-companys-data (dostęp: 19.08.2025).

Dilmegani C., *6 Risks of Generative AI & How to Mitigate Them in 2025*. Pozyskano z: https://research.aimultiple.com/generative-ai-data/ (dostęp: 19.08.2025).

Dwork C., Roth A., *The Algorithmic Foundations of „differential privacy"*, "Foundations and Trends in Theoretical Computer Science" 2014, 9(3–4), s. 211–407. https://doi.org/10.1561/0400000042

Frączek M., Spaliński K., *Bezpieczeństwo informacji w dobie sztucznej inteligencji: analiza ryzyka i wyzwania związane z ChatGPT*, „Zeszyty Naukowe Pro Publico Bono" 2023, 1(1). https://doi.org/10.5604/01.3001.0054.1726

Ilnicka W., *Sztuczna inteligencja – korzyści i zagrożenia*, „Zeszyty Naukowe Wyższej Szkoły Bankowej w Poznaniu" 2024, 105(20. https://doi.org/10.58683/dnswsb.2013

Kausar M. A., V.S. Dhaka, *Web Crawler: A Review*. Pozyskano z: https://www.researchgate.net/publication/258789938_Web_Crawler_A_Review (dostęp: 19.08.2025).

Kowalski S., *Sztuczna inteligencja a ochrona danych osobowych*, [w:] *Metaświat prawne i techniczne aspekty przełomowych technologii*, red. R. Bieda, Z. Okoń, Warszawa 2022.

Novelli C., Casolari F., Hacker P., Spedicato G., Floridi L., *Generative AI in EU Law: Liability, Privacy, Intellectual Property, and Cybersecurity*. Pozyskano z: https://arxiv.org/pdf/2401.07348 (dostęp: 16.08.2025).

Stryker C., Scapicchio M., *What is generative AI?* Pozyskano z: https://www.ibm.com/topics/generative-ai (dostęp: 16.08.2025).

Wróblewski M., *Sztuczna inteligencja a ochrona praw człowieka*, [w:] *Metaświat prawne i techniczne aspekty przełomowych technologii*, red. R. Bieda, Z. Okoń, Warszawa 2022.

Zhao B., *Web Scraping*. Pozyskano z: https://www.researchgate.net/publication/317177787_Web_Scraping (dostęp: 19.08.2025).