GP Partners Gawroński, Biernatowski Sp.K.
al. Jana Pawła II 12, 00-124 Warszawa | +48 22 243 4953 | info@gppartners.pl

Warszawa, 27 listopada 2023 r.

Skarżący:
Łukasz Olejnik

reprezentowany przez:
r. pr. Macieja Gawrońskiego
GP Partners Gawroński, Biernatowski Sp.K.
al. Jana Pawła II 12, 00-124 Warszawa

Prezes Urzędu Ochrony Danych Osobowych
ul. Stawki 2
00-193 Warszawa

Znak sprawy: DS.523.4800.2023.PR.SPI.

PISMO
w sprawie uzupełnienia stanowiska Pana Łukasza Olejnika zawartego w skardze z dnia 29 sierpnia 2023 r.

Działając w imieniu skarżącego, Pana Łukasza Olejnika przedstawiamy dalsze stanowisko w przedmiocie zagadnień prawnych związanych z ochroną danych osobowych przetwarzanych przez systemy generatywnej sztucznej inteligencji, w tym narzędzie ChatGPT udostępniane przez OpenAI.

Niniejsze pismo kierujemy w nawiązaniu do komunikatu Urzędu Ochrony Danych Osobowych z dnia 20.09.2023 („*Technologia musi być zgodna z RODO*") w przedmiocie rozpatrywania skargi Pana Łukasza Olejnika. Komunikat UODO przyjmujemy z entuzjazmem jako sygnał uznania przez Urząd zagadnień prawnych poruszanych w skardze za istotne i mające fundamentalne, uniwersalne znaczenie dla praw podmiotów danych w związku z dalszym rozwojem narzędzi sztucznej inteligencji. Dokładnie w ten sposób przedmiot sprawy postrzega Pan Łukasz Olejnik. Skarżący, bardziej niż stwierdzeniem dotychczasowych naruszeń, zainteresowany jest oceną modelu działania OpenAI w celu poprawy poziomu ochrony danych osobowych przetwarzanych w ramach ChatGPT w przyszłości.

Kluczowe zagadnienia prawne. Urząd w komunikacie z dnia 20.09.2023 r. zwraca uwagę, że *rozwój nowych technologii musi odbywać się z poszanowaniem praw osób fizycznych wynikających m.in. z RODO*, a sprawa dotyczy systemowego podejścia OpenAI do europejskich zasad ochrony danych osobowych.

W związku z tymi spostrzeżeniami Urzędu, Pan Łukasz Olejnik chciałby uwypuklić zagadnienia, które w jego ocenie odgrywają kluczową rolę w ocenie systemowego podejścia OpenAI do zgodności przetwarzania danych z obowiązującymi regulacjami. Te zagadnienia to:

1) realizacja obowiązku zapewnienia ochrony danych osobowych w fazie projektowania (**data protection by design**),
2) zapewnienie wykonywania **prawa do sprostowania danych** (w odniesieniu do danych wykorzystywanych w celu trenowania modeli językowych),
3) realizacja **obowiązków informacyjnych** odnoszących się do procesów przetwarzania danych w celu trenowania modeli językowych,

Wszystkie powyższe zagadnienia są przejawem podstawowej z perspektywy zgodności z RODO zasady przetwarzania danych osobowych **zgodnie z prawem, rzetelnie i w sposób przejrzysty** dla osoby, której dane dotyczą, tj. zasady zapisanej w art. 5 ust. 1 lit. a RODO.

Wybrane kwestie wymagające zbadania. Aby możliwa była analiza zgodności przetwarzania danych osobowych w ramach ChatGPT, OpenAI powinna przekazać Urzędowi informacje w szczególności w poniższych obszarach:

- czy dla narzędzia ChatGPT została przeprowadzona ocena skutków dla przetwarzania danych osobowych (DPIA) i czy ocena ta uwzględnia szczególne ryzyka związane z narzędziem, w tym ryzyko generowania treści nieprawidłowych (tzw. halucynacje),
- w jaki sposób OpenAI wdrożyło mechanizmy mające na celu obsługę praw jednostki, w szczególności prawo do informacji (art. 13 i 14 RODO), prawo dostępu (art. 15 RODO), prawo do sprostowania danych (art. 16 RODO),
- czy OpenAI poddało analizie w jaki sposób można wykonać prawo do sprostowania danych, czy podjęło próbę opracowania rozwiązania technicznego pozwalającego na wykonanie tego prawa,
- czy OpenAI poddało analizie możliwość wprowadzenia rozwiązania, modułu pozwalającego na korygowanie nieprawidłowych danych, mechanizmy filtrujące generowane treści, mechanizmy oduczania maszynowego (*machine unlearning*),

  Taki moduł został zaproponowany w przypadku narzędzia generatywnej sztucznej inteligencji Bard udostępnianego przez Google[1]. Narzędzie umożliwia dodatkową funkcjonalność weryfikacji „na drugą rękę" treści generowanych przez sztuczną inteligencję. Dodatkowy moduł na żądanie użytkownika oznacza, w oparciu o informacje w przeglądarce Google, w jakim stopniu można opierać się o wygenerowaną treść co do jej wiarygodności.
- czy w przypadku oceny, że nie jest możliwe dokonanie sprostowania danych osobowych, OpenAI wprowadziło zasady informowania podmiotów danych o tym, kiedy i dlaczego sprostowanie danych nie jest możliwe,
- jakie procedury stosuje OpenAI w przypadku powzięcia wiedzy (np. wskutek zgłoszenia), że określone treści generowane przez narzędzie ChatGPT obejmujące dane osobowe są nieprawdziwe,
- w jaki sposób OpenAI zapewniło, że narzędzie ChatGPT zostało opracowane i udostępnione w zgodzie z zasadą privacy by design, jaka dokumentacja pozwala na zapewnienie zasady rozliczalności w tym zakresie.

Ocena powyższych zagadnień wymaga opisania przez OpenAI modelu i oprogramowania, na którym opiera się narzędzie ChatGPT i ich analizy w kontekście właściwych przepisami regulującymi ochronę danych osobowych.

Gdyby Urząd ustalił, że OpenAI prowadziło analizę dotyczącą w szczególności wykonania prawa do sprostowania danych, uzasadnione będzie pozyskanie od OpenAI dokumentacji z tej analizy i jej uwzględnienie przy rozpoznaniu skargi.

Celowe dla rozstrzygnięcia sprawy, ze względu na wagę zagadnień, byłoby także spotkanie i wymiana korespondencji Urzędu z przedstawicielami OpenAI, przy udziale Pana Łukasza Olejnika jako skarżącego i jego pełnomocnika. Pozwoliłoby to zrozumieć jakie wyzwania dotyczące ochrony danych wiążą się z rozwojem narzędzi generatywnej sztucznej inteligencji i w jaki sposób można te wyzwania zaadresować, dążąc do wysokiego poziomu ochrony praw i wolności osób, których dane dotyczą.

[1] https://support.google.com/bard/answer/14143489?hl=en&co=GENIE.Platform%3DAndroid (dostęp: 13.11.2023).

Pod rozwagę Urzędu chcielibyśmy poddać także zwrócenie się do Urzędu Ochrony Konkurencji i Konsumentów w celu zasięgnięcia stanowiska tego urzędu w sprawie konsumenckiego wymiaru zagadnień, których dotyczy skarga i ryzyk, jakie narzędzia sztucznej inteligencji generują dla praw konsumentów. Zagadnieniami konsumenckimi w ramach narzędzi generatywnej sztucznej inteligencji interesują się także zagraniczne organy nadzorujące ochronę praw konsumentów[2].

Najnowsze stanowiska organów. Poniżej wskazujemy dodatkowe stanowiska i wypowiedzi organów, które pojawiły się w ostatnim czasie i odnoszą się do zgodności procesów przetwarzania danych przy użyciu narzędzi sztucznej inteligencji.

- Niemieckie organy nadzoru nad ochroną danych osobowych od dłuższego czasu interesują się OpenAI i udostępnianym przez OpenAI narzędziem ChatGPT. W ostatnim czasie **Heski komisarz ds. ochrony danych i wolności informacji** (HBDI) poinformował, że do OpenAI skierował kolejne, szczegółowe **pytania** w sprawie warunków przetwarzania danych z wykorzystaniem narzędzia ChatGPT[3].

  Jak wynika z informacji prasowej organu kolejne 79 pytań dotyczy m.in. kwestii zgodności z prawem masowego gromadzenia danych w celu trenowania modeli językowych oraz możliwości wykonania przez podmioty danych praw do sprostowania danych, ich usunięcia lub uzyskania informacji dotyczących przetwarzania. Organ wskazuje, że zachodzi wątpliwość czy prawa te mogą w ogóle być wykonywane zgodnie z prawem w kontekście przetwarzania w ChatGPT.

  Heski komisarz ds. ochrony danych sygnalizuje, że podejmuje działania nadzorcze w stosunku do OpenAI we współpracy z innymi krajowymi organami nadzorczymi, jak też organami unijnymi.

- W skardze, w pkt 1.4. (str. 5 skargi) opisywaliśmy zapowiedź francuskiego CNIL zintensyfikowania aktywności w obszarze ochrony danych osobowych w związku z narzędziami generatywnej AI. Zgodnie z zapowiedzią, organ opublikował w dniu 16 października 2023 r. **siedem arkuszy instruktażowych dotyczących rozwijania systemów sztucznej inteligencji**[4]. Arkusze mają służyć jako wsparcie w projektowaniu systemów sztucznej inteligencji zgodnie z przepisami o ochronie danych osobowych.

  Z arkuszy CNIL wynika m.in., że projektując systemy AI należy uwzględnić następujące zagadnienia:

  (i) potrzeba przeprowadzenia **oceny skutków dla przetwarzania** (DPIA). Ocena powinna obejmować m.in. kwestie ryzyk dla praw podmiotów danych, środków pozwalających na wykonywanie praw podmiotów danych, poziomu przejrzystości przetwarzania.

  CNIL wskazuje, że jednym z istotnych ryzyk związanych z systemami AI jest ryzyko tworzenia fałszywych treści na temat prawdziwej osoby, co jest szczególnie ważne w przypadku generatywnych systemów sztucznej inteligencji i może mieć konsekwencje dla jej reputacji.

---

[2] Brytyjski Urząd ds. Konkurencji i Rynków (CMA), będący krajowym organem nadzoru regulacyjnego opracował a następnie w dniu 18 września 2023 r. przedstawił 7 zasad dotyczących odpowiedzialnego rozwijania oraz wykorzystywania narzędzi sztucznej inteligencji (Proposed principles to guide competitive AI markets and protect consumers, https://www.gov.uk/government/news/proposed-principles-to-guide-competitive-ai-markets-and-protect-consumers, dostęp: 13.11.2023).

[3] Pytania kierowano wcześniej (w kwietniu br.) i OpenAI udzieliło na te pytania odpowiedzi. Organ z Hesji po analizie odpowiedzi uznał, że zachodzi potrzeba zadania dodatkowych, szczegółowych pytań, aby dogłębnie zbadać zgodność przetwarzania danych osobowych w ramach narzędzia ChatGPT (*Hessischer Datenschutzbeauftragter fordert erneut Antworten zu ChatGPT, 24.10.2023,* dostęp: 13.11.2023)

https://datenschutz.hessen.de/presse/hessischer-datenschutzbeauftragter-fordert-erneut-antworten-zu-chatgpt

[4] CNIL, *AI how-to sheets, 16.10.2023*, dostęp: 13.11.2023 (https://www.cnil.fr/en/ai-how-sheets).

Podmioty projektujące systemy sztucznej inteligencji w oparciu o DPIA powinny zdecydować o potrzebie wdrożenia szczególnych środków technicznych takich jak (i) minimalizacja danych oraz (ii) mechanizmy ułatwiające wykonywanie praw lub innych środków prawnych, takie jak techniki oduczania maszynowego (*machine unlearning*) lub środki służące wyjaśnianiu i śledzeniu wyników systemów AI.

(ii) potrzeba uwzględnienia ochrony danych w **projektowaniu architektury systemów sztucznej inteligencji** – CNIL wskazuje, że należy wybrać taką architekturę, która zapewnia najwyższy poziom ochrony praw i wolności podmiotów danych, tak aby przetwarzanie było zgodne z zasadą minimalizacji danych[5].

Arkusze opublikowane przez CNIL są wyrazem tego, jak dużą wagę organ przywiązuje do procesu zapewnienia zgodności systemów sztucznej inteligencji na etapie ich projektowania. Stanowisko organu podkreśla fundamentalność zasady *data protection by design* (*privacy by design*) na tle regulacji ochrony danych osobowych.

Arkusze są obecnie poddane konsultacjom społecznym, ich ostateczna wersja ma zostać opublikowana na początku roku 2024.

- W dniu 20 października 2023 r. **Global Privacy Assembly** (GPA)[6] przyjęło rezolucję w sprawie generatywnych systemów sztucznej inteligencji. GPA wskazuje w rezolucji:

  (i) systemy generatywnej sztucznej inteligencji muszą być projektowane, rozwijane i wdrażane w oparciu o zasadę ochrony danych osobowych, a ochrona danych powinna być zapisana w koncepcji i projekcie systemu. GPA podkreśla tym samym rolę jaką odgrywa zasada **privacy by design** oraz potrzebę przeprowadzenia oceny skutków dla ochrony danych,

  (ii) istotnym elementem zgodności z przepisami o ochronie danych osobowych jest **zapewnienie prawidłowości danych osobowych** wykorzystywanych do trenowania modeli. GPA wskazuje, że jednym z ryzyk wiążących się z systemami GAI jest generowanie treści nieprawdziwych zawierających dane osobowe (tzw. halucynacji). Aby zmitygować m.in. to ryzyko, zdaniem GPA należy wdrożyć dodatkowe środki techniczne – np. wykorzystanie filtrów dla danych wejściowych i wyjściowych.

  (iii) w zakresie **transparentności** GPA wskazuje m.in. że osoby, których dane dotyczą powinny być w stanie uzyskać informację jak, kiedy i dlaczego dane osobowe są używane w procesie trenowania systemów generatywnej sztucznej inteligencji. Zasada transparentności oznacza, że przejrzysta powinna być informacja z jakiego źródła pochodzi zbór danych wykorzystywanych do trenowania modeli.

  (iv) GPA jako jeden z kluczowych aspektów zgodności systemów GAI wskazuje obowiązek **zapewnienia podmiotom danych możliwości wykonania ich praw**, w tym poprzez wdrożenie odpowiednich środków technicznych i organizacyjnych. Osoby, których dane dotyczą powinny mieć zapewnione m.in. prawo dostępu do danych osobowych jak też prawo do skorygowania nieprawidłowych danych osobowych.

[5] CNIL, Taking data protection into account in the system design choices, 16.10.2023, https://www.cnil.fr/en/taking-data-protection-account-system-design-choices, dostęp: 13.11.2023)

[6] Global Privacy Assembly to światowe stowarzyszenie ponad 130 organów nadzoru w obszarze ochrony danych osobowych i prywatności.

- Prezydent Stanów Zjednoczonych Joe Biden 30 października 2023 r. wydał rozporządzenie wykonawcze w sprawie „bezpiecznej i godnej zaufania sztucznej inteligencji". W rozporządzeniu wzywa Kongres do przyjęcia przepisów dotyczących prywatności danych w celu ochrony obywateli Stanów Zjednoczonych, a zwłaszcza dzieci. W rozporządzeniu zwraca uwagę m.in. na konieczność wsparcia dla przyspieszenia rozwoju i wykorzystywania technik ochrony prywatności, opracowania wytycznych w celu oceny skuteczności ochrony prywatności, czy też oceny sposobu, w jaki podmioty gromadzą i wykorzystują informacje dostępne na rynku[7].

Wnioski. Powyższe stanowiska potwierdzają, że zagadnienia poruszane w skardze są aktualne i relewantne, a obszary, w których Pan Łukasz Olejnik identyfikuje nieprawidłowości zbiegają się z obszarami, w których wątpliwości i rekomendacje formułują organy nadzoru.

Istotę sprawy i perspektywę Pana Łukasza Olejnika szczególnie trafnie opisuje następujący fragment rezolucji GPA, w której GPA wzywa podmioty udostępniające narzędzia generatywnej sztucznej inteligencji do:

> *uznania ochrony danych i prywatności jako podstawowego prawa człowieka oraz budowania odpowiedzialnych i godnych zaufania technologii generatywnej sztucznej inteligencji, które chronią ochronę danych, prywatność, godność ludzką i inne podstawowe prawa i wolności[8].*

W załączeniu przekazujemy pomocniczo tłumaczenie maszynowe niniejszego pisma na język angielski.

_____

Maciej Gawroński
radca prawny

Załączniki:

1. Tłumaczenie maszynowe pisma

[7] FACT SHEET: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/

[8] Tumaczenie własne.