

Grzegorz Tylec

Katedra Języka, Retoryki i Prawa Mediów

Instytut Dziennikarstwa i Zarządzania KUL JP II

grzegorztylec@kul.pl

ORCID: <https://orcid.org/0000-0003-2016-4523>

Czy prawo autorskie może hamować rozwój sztucznej inteligencji?

Streszczenie: W artykule opisano problemy prawne związane z interpretacją art. 4 Dyrektywy Parlamentu Europejskiego i Rady (UE) 2019/790 z dnia 17 kwietnia 2019 r. w sprawie prawa autorskiego i praw pokrewnych na jednolitym rynku cyfrowym oraz zmiany dyrektyw 96/9/WE i 2001/29/WE dotyczących nieodpłatnego wykorzystania utworów oraz przedmiotów praw pokrewnych do szkolenia wykorzystywanych komercyjnie algorytmów generatywnej sztucznej inteligencji. Przybliżono uregulowania w zakresie prawa autorskiego zawarte w Akcie w sprawie sztucznej inteligencji. Omówiono różnicę wynikającą z definicji pojęcia „eksploracja tekstów i danych” zawartego w tekście wymienionej dyrektywy oraz definicji terminu „system sztucznej inteligencji” zawartego w tekście wymienionego aktu, a także ich konsekwencje prawne w sferze prawa autorskiego.

Słowa kluczowe: prawo autorskie, sztuczna inteligencja, eksploracja tekstów i danych, akt w sprawie sztucznej inteligencji

Współcześnie mamy do czynienia z bardzo szybkim rozwojem technologii sztucznej inteligencji (ang. *artificial intelligence*, dalej: AI), która dotyczy już niemal każdej sfery życia, a jej dalszy rozwój i ekspansja wydają się nieuniknione. Jak wskazano w literaturze prawniczej, jedynym czynnikiem, który może znacząco spowolnić jej rozwój jest prawo autorskie. Trzeba jasno powiedzieć, że mimo funkcjonowania na rynku ogromnej ilości rozwiązań informatycznych opartych na AI, to legalność ich działania w wielu przypadkach może być kwestionowana.

Problem naruszeń praw autorskich przez algorytmy AI powinien być rozważany z perspektywy rozmaitych systemów prawnych obowiązujących w różnych państwach. Jak wiemy jednak, w praktyce najistotniejsza jest perspektywa prawa USA oraz prawa UE. Większość państw demokratycznych podąża za rozwiązaniami w zakresie nowych technologii wypracowanymi w tych dwóch systemach prawnych. Jeśli chodzi o sytuację prawną w USA, to warto wskazać, że toczy się tam obecnie wiele głośnych procesów sądowych, w których organizacje reprezentujące twórców lub sami twórcy, w tym m.in. *New York Times*, wskazują na nielegalne wykorzystywanie ich utworów przez AI. Stroną pozwaną w tych procesach są kluczowi na rynku deweloperzy AI jak np.: *OpenAI*, *Midjourney* czy *Stable Diffusion*.

Jeśli chodzi zaś o regulacje prawne UE, to zdaniem części prawników – także moim – należy stwierdzić, że wykorzystywanie utworów przez AI odbywa się z naruszeniem praw autorskich podmiotów uprawnionych. Naruszenia te mogą dotyczyć zarówno danych wejściowych, jak i danych wyjściowych. W zakresie danych wejściowych w pierwszej kolejności należy zwrócić uwagę na działania związane z procesem trenowania AI. Kwestią, która w tej materii budzi najwięcej wątpliwości jest odpowiedź na pytanie, czy zapisy zawarte w artykule 4 *Dyrektywy Parlamentu Euro-*

pejskiego i Rady (UE) 2019/790 z dnia 17 kwietnia 2019 r. w sprawie prawa autorskiego i praw pokrewnych na jednolitym rynku cyfrowym oraz zmiany dyrektyw 96/9/WE i 2001/29/WE (Dz. Urz. UE, L 130/92) (dalej: dyrektywa), regulujące kwestie legalności eksploracji tekstu i danych (ang. *text and data mining*, TDM) do celów komercyjnych, pozwalają deweloperom AI na nieodpłatne wykorzystywanie dzieł chronionych prawem autorskim do szkolenia algorytmów AI.

Gdy Parlament Europejski uchwalił w 2019 r. wskazaną wyżej dyrektywę, zawarte w niej przepisy dotyczące eksploracji tekstu i danych mało kogo interesowały. Wyobrażano sobie, że dotyczą one jedynie sytuacji, kiedy zainteresowany jakimś tematem naukowiec chce przeanalizować tysiące pdf-ów zawierających teksty naukowe, aby zautomatyzować analizę informacji w nich zawartych do celów swych badań naukowych. W owym czasie temat TDM miał charakter niszowy i nie budził większych dyskusji ani jakichkolwiek zastrzeżeń. Pytanie o legalność tego typu działania pojawiła się, gdy twórcy oraz podmioty uprawnione z tytułu praw pokrewnych zorientowali się, że deweloperzy AI do szkolenia swych algorytmów wykorzystują bez ich zgody oraz bez zapłaty ogromne ilości chronionych prawem utworów. W odpowiedzi na tak postawione pytanie deweloperzy AI, a wraz z nimi wysocy urzędnicy UE przypomnieli, że sprawa jest wyraźnie uregulowana w art. 4 dyrektywy i nie można mówić tu o jakimkolwiek naruszeniu prawa autorskiego.

Warto zauważyć, że kilka tygodni temu w Polsce problem ten był przedmiotem ożywionej debaty, gdyż polskie Ministerstwo Kultury i Dziedzictwa Narodowego jako ostatnie w Europie, po dwóch i pół roku opóźnienia w implementacji dyrektywy, przedstawiło swój zupełnie odmienny projekt jej wdrożenia do prawa polskiego. Nowatorskie podejście ministerstwa, inne niż przyjęte dotychczas we wszystkich krajach UE, dotyczyło m.in. opisanego wyżej komercyjnego TDM i przewidywało, że szkolenie algorytmów AI powinno odbywać się po uiszczeniu należnego twórcom wynagrodzenia za korzystanie z ich utworów. Ostatecznie polskie ministerstwo wycofało się ze swego pierwotnego pomysłu, jednak walka o uznanie, że deweloperzy AI korzystają w sposób nielegalny z chronionych prawem utworów do szkolenia AI w Polsce oraz całej Europie wciąż trwa. Prawdopodobnie temat ten powróci niebawem, gdy projekt nowelizacji ustawy o prawie autorskim i prawach pokrewnych implementujący przepisy dyrektywy będzie przedmiotem debaty polskiego Sejmu.

Warto przyjrzeć się argumentom, które są podnoszone na dowód tego, że obecna interpretacja art. 4 dyrektywy uznająca, że wolno deweloperom korzystać z utworów do szkolenia AI za darmo jest błędna:

1. Rozwój technologiczny AI następuje w bardzo szybkim tempie i uchwalając dyrektywę w 2019 r., prawodawca nie miał świadomości, jak będzie wyglądała sfera eksploatacji utworów w ramach TDM.
2. Działalność firm tworzących rozwiązania generatywnej AI oparta jest na wykorzystaniu utworów i przedmiotów praw pokrewnych z pokrzywdzeniem podmiotów uprawnionych, bez jakiegokolwiek rekompensaty.
3. Wytwory kreowane przez generatywną AI stanowią produkty konkurencyjne dla utworów i przedmiotów praw pokrewnych, na których bazują.
4. Przewidziana w art. 4 ust. 3 dyrektywy klauzula „opt out” pozwalająca każdemu twórcy nie zgodzić się na wykorzystywanie jego utworów do szkolenia AI faktycznie nie funkcjonuje, pozostając martwym zapisem. W praktyce uprawnieni z tytułu praw autorskich nie mają możliwości dokonania jakiegokolwiek sprzeciwu co do tego, aby ich utwory nie były wykorzystywane w procesie szkolenia AI.

5. Pojawia się coraz więcej wypowiedzi przedstawicieli nauki oraz praktyków, że realizacja uprawnienia z art. 4 ust. 3. dyrektywy (czyli złożenia zastrzeżenia w ramach klauzuli „opt out”) jest w praktyce niemożliwa i nie będzie mogła objąć swym zakresem dzieł opublikowanych przed datą uchwalenia dyrektywy, te bowiem w większości funkcjonują w wersji analogowej.
6. Pojawiły się przekonujące argumenty przemawiające za tym, że regulacja z art. 4 dyrektywy w obecnej wersji narusza prawa podstawowe twórców, m.in. prawo do kultury i nauki, wolności słowa oraz wolność wypowiedzi artystycznej.
7. Podnosi się, że dokonany przez państwa UE sposób implementacji przepisów dotyczących TDM jest sprzeczny z art. 13 *Porozumienia w sprawie handlowych aspektów praw własności intelektualnej TRIPs* (2017) oraz art. 9 ust. 2 *Konwencji berneńskiej o ochronie dzieł literackich i artystycznych* (1935), gdyż wyjątek ten nie spełnia tzw. trójstopniowego testu zawartego w konwencji.
8. Brytyjska Komisja Kultury Mediów i Sportu w dniu 30 sierpnia 2023 r. wydała oświadczenie, że pierwotny plan rządu mający na celu wyłączenie TDM przez sztuczną inteligencję z ochrony praw autorskich grozi ograniczeniem sztuki i produkcji kulturalnej. W oświadczeniu stwierdzono, że zamiast realizować plany dotyczące szerokiego wyłączenia TDM spod ochrony praw autorskich, rząd powinien aktywnie wspierać możliwość nabywania licencji do korzystania z utworów przez niewielkie przedsiębiorstwa tworzące algorytmy sztucznej inteligencji.

Dodatkowo wskazać należy, że uchwalony przez Parlament Europejski 13 marca 2024 r. *Akt w sprawie sztucznej inteligencji* (Artificial Intelligence Act, 2024) nie przesądził jednoznacznie analizowanych kwestii. Regulacje tego aktu normatywnego w opisywanym obszarze nie tworzą nowych zasad w obrębie prawa autorskiego ani nie mogą być wykorzystywane do interpretacji treści dyrektywy, gdyż sprowadzają się one jedynie do:

- a) obowiązku przestrzegania przez deweloperów AI obowiązującego prawa autorskiego, w tym art. 4 dyrektywy (por. art. 53 ust. 1 pkt c aktu) oraz
- b) sporządzenia i udostępnienia do publicznej wiadomości odpowiednio szczegółowego podsumowania treści wykorzystywanych do szkolenia modelu AI ogólnego przeznaczenia, według wzoru udostępnionego przez Urząd ds. AI (por. art. 53 ust. 1 pkt d aktu).

Wskazane regulacje zawarte są w rozdziale V sekcji II *Obowiązki dostawców modeli sztucznej inteligencji ogólnego przeznaczenia*, a nie części poświęconej prawu autorskiemu, gdyż takowa w tym dokumencie nie istnieje. Kluczowym argumentem przemawiającym za tym, że przepisy *Aktu w sprawie sztucznej inteligencji* nie mogą być wykorzystywane w dyskusji nad interpretacją przepisów dyrektywy, jest stwierdzenie zawarte w ostatnim zdaniu motywu aktu (recital nr 108), gdzie zapisano: *Urząd ds. AI powinien monitorować, czy dostawca wypełnił obowiązki nałożone na dostawców modeli AI ogólnego przeznaczenia dotyczące wprowadzenia polityki w celu uzyskania zgodności z unijnym prawem autorskim i podania do wiadomości publicznej streszczenia na temat treści wykorzystywanych do trenowania, jednak nie powinien weryfikować danych treningowych pod kątem przestrzegania praw autorskich ani przystępować do oceny tych danych w podziale na poszczególne utwory. Niniejsze rozporządzenie nie wpływa na egzekwowanie przepisów dotyczących praw autorskich przewidzianych w prawie Unii.*

Kluczowym argumentem przemawiającym za tym, że procesy wykorzystania utworów na potrzeby eksploracji tekstu i danych oraz szkolenia generatywnej AI powinny być od siebie wyraźnie oddzielone i odrębnie uregulowane prawnie, jest analiza i porównanie definicji pojęcia TDM zawartej w art. 2 dyrektywy oraz definicji pojęcia „system sztucznej inteligencji”, przyjętej w art. 3 *Aktu w sprawie sztucznej inteligencji* (Tylec i in., 2024).

TDM, według dyrektywy, to termin definiujący szereg metod i technik wydobywania przydatnych informacji ze zbiorów dokumentów poprzez identyfikację i wykorzystywanie wzorców z nieustrukturyzowanych zbiorów danych tekstowych. Z formalnego punktu widzenia, TDM ogranicza się do tworzenia nowych informacji w oparciu o bazę dokumentów, które można poddać takim operacjom, jak kategoryzacja tekstu, grupowanie tekstów, ekstrakcja pojęć, analiza sentymentów czy modelowanie relacji między elementami. Ideą zastosowania TDM jest zmniejszenie wysiłku związanego z przeszukiwaniem ogromnych zbiorów dowolnego rodzaju dokumentów spełniających wymagania Big Data. Zestawiając tę definicję z treścią pojęcia „system sztucznej inteligencji” zawartego w *Akcie w sprawie sztucznej inteligencji*, widać, że TDM to jedynie zautomatyzowana technika analizy tekstów i danych cyfrowych w celu generowania informacji. Koncepcja sztucznej inteligencji jest natomiast znacznie szersza i wykracza daleko poza cel TDM, którym jest jedynie generowanie informacji zawierających wzorce, trendy i korelacje. Cechą systemów AI odróżniającą je od TDM jest bowiem możliwość tworzenia nowych treści, czyli tekstów, obrazów czy innych mediów, które wcześniej nie istniały. Można powiedzieć więc, że systemy AI mogą nadawać wydobytym informacjom indywidualną formę wyrażenia, prowadząc do zupełnie nowych kreacji. Różnica polega na tym, że samo TDM nie prowadzi do wykreowania nowych wytworów, a jedynie danych (informacji), które dopiero mogą stać się punktem wyjścia do tworzenia na ich bazie nowych tekstów, filmów czy innych mediów.

To spostrzeżenie jest dowodem na to, że regulacja prawna przyjęta przez UE w zakresie TDM nie rozwiązuje problemu legalności uczenia maszynowego w oparciu o utwory chronione prawem autorskim. Dotyczy bowiem jedynie procesu TDM, nie zaś zagadnienia generatywnej sztucznej inteligencji, co jak wykazano wyżej jest czymś zgoła innym.

Na zakończenie pragnę zaznaczyć swoje krytycznie podejście do argumentu sugerującego, że nałożenie na deweloperów AI obowiązku płacenia twórcom wynagrodzenia zahamuje rozwój nowych technologii i spowolni procesy innowacyjne. Moim zdaniem, jest dokładnie odwrotnie. Tylko zrównoważona, uczciwa i ustalona na rozsądnym poziomie opłata za korzystanie z utworów przyczyni się do realnego wzrostu innowacyjności UE zarówno w branży deweloperów AI, jak i twórców. Korzyścią dla deweloperów będzie to, że zniknie ryzyko prawne związane z procesami sądowymi i zapłatą odszkodowań za naruszanie praw autorskich. Natomiast twórcy nadal będą tworzyć контент, który będzie mógł być rozwijany przez AI. Brak stosownych rozwiązań w tej materii w niedługiej przyszłości może doprowadzić do zapaści działów kreatywnych zatrudniających twórców. Wizja oparcia kultury i sztuki jedynie na wytworach AI może budzić głębokie obawy. Od dawna wiadomo, że prawo autorskie jest instrumentem polityki kulturalnej państwa i jako takie powinno równoważyć i godzić uzasadnione interesy twórców i konsumentów dzieł. W mojej opinii przedstawione w artykule, obecnie obowiązujące na terenie UE, regulacje prawne dotyczące prawa autorskiego w zakresie korzystania z utworów do szkolenia AI nie realizują tego postulatu.

Bibliografia:

1. Akt w sprawie sztucznej inteligencji – posłowie przyjmują przełomowe przepisy (2024). W: *Parlament Europejski* [online]. [Dostęp 16.04.2024]. Dostępny w: <https://www.europarl.europa.eu/news/pl/press-room/20240308IPR19015/akt-w-sprawie-sztucznej-inteligencji-poslowie-przyjmujaja-przelomowe-przepisy>.
2. Artificial Intelligence Act (2024). W: *The EU Artificial Intelligence Act* [online]. [Dostęp 16.04.2024]. Dostępny w: <https://artificialintelligenceact.eu/the-act/>.
3. *Dyrektywa Parlamentu Europejskiego i Rady (UE) 2019/790 z dnia 17 kwietnia 2019 r. w sprawie prawa autorskiego i praw pokrewnych na jednolitym rynku cyfrowym oraz zmiany dyrektyw 96/9/WE i 2001/29/WE (2019)*. Dz. Urz. UE, L 130/92. W: *EUR-Lex* [online]. [Dostęp 16.04.2024]. Dostępny w: <https://eur-lex.europa.eu/legal-content/PL/TXT/?uri=CELEX:32019L0790>.
4. *Konwencja berneńska o ochronie dzieł literackich i artystycznych z dnia 9 września 1886 r., przejrzana w Berlinie dnia 13 listopada 1908 r. i w Rzymie dnia 2 czerwca 1928 r. (ratyfikowana zgodnie z ustawą z dnia 5 marca 1934 r.)*. Dz. U. 1935, nr 84, poz. 515. W: *ISAP – Internetowy System Aktów Prawnych* [online]. [Dostęp 16.04.2024]. Dostępny w: <https://isap.sejm.gov.pl/isap.nsf/DocDetails.xsp?id=wdu19350840515>.
5. TYLEC, G., KWIECIEN, S., SAROWSKI, Ł., SADOWSKI, A. (2024). *Is it Possible to License Works Used in the Learning Process of Artificial Intelligence Algorithms?* [online]. SSRN. [Dostęp 16.04.2024]. Dostępny w: <https://ssrn.com/abstract=4729495>.
6. WTO: porozumienie w sprawie handlowych aspektów praw własności intelektualnej (2017). W: *EUR-Lex* [online]. [Dostęp 16.04.2024]. Dostępny w: <https://eur-lex.europa.eu/PL/legal-content/summary/wto-agreement-on-trade-related-aspects-of-intellectual-property.html>.