

Dozwolona kradzież? Sąd w Kalifornii staje po stronie Big Techów. Ale nie do końca
wyborcza.pl, (04.07.2025)
Wojciech Szot

Dozwolona kradzież? Sąd w Kalifornii staje po stronie Big Techów. Ale nie do końca
Wojciech Szot

- 'The Atlantic' ujawnił, że Mark Zuckerberg, dyrektor generalny Mety, zatwierdził korzystanie z LibGen (Fot. Alex Wong/Getty Images)

Czy można "karmić" AI treściami z książek ściągniętych z sieci? Kalifornijski sędzia wydał kontrowersyjny wyrok, a pisarze protestują przeciwko sztucznej inteligencji.

Ta historia zaczęła się w 2023 roku, gdy Meta, firma do której należy m.in. platforma społecznościowa Facebook, zaczęła prace nad budową modelu językowego Llama. To odpowiednik słynnego Chata GPT, którego Meta zamierza używać zarówno do komunikacji na swoich platformach, jak i generowania treści.

Już w połowie 2023 roku okazało się, że zarówno Meta, jak i OpenAI, czyli twórcy ChataGPT, do trenowania swoich narzędzi użyli książek znanych autorów i autorek.

Twórcy pytali, jak to możliwe, że ChatGPT generuje streszczenia książek, które trzeba kupić w księgarni, i pozwali firmę.

Nie poznaliśmy wówczas odpowiedzi na jeszcze jedno pytanie: jak firmy technologiczne weszły w posiadanie tych książek? Nie kupowały przecież pojedynczych kopii e-booków.

"Największa kradzież dzieł sztuki w historii"?

Odpowiedź poznaliśmy w 2025 roku – po prostu pobrano książki z internetu. Wykorzystano do tego bibliotekę zawierającą pliki lub odnośniki do stron, z których można ściągnąć e-booki.

"The Atlantic" ujawnił, że Mark Zuckerberg, dyrektor generalny Mety, zatwierdził korzystanie z LibGen (skrót od "Library Genesis") – nielegalnego repozytorium plików – jako źródła treści do trenowania modelu AI rozwijanego przez firmę. Magazyn opublikował również wyszukiwarkę, dzięki której autorzy i autorki mogli sprawdzić, czy ich książki znalazły się w tej bibliotece.

Wybuchł skandal. Jedni pisali, że to "największa kradzież w historii", inni relatywizowali działania Big Techów. – Naukowcy masowo korzystają z pirackich baz książek – przyznawał w "Wyborczej" Mateusz Kostecki, neurobiolog w Uniwersytecie w Heidelbergu.

Artur Wojciechowski, dyrektor Centrum Cyberbezpieczeństwa Wyższej Szkoły Informatyki i Zarządzania w Rzeszowie zauważył, że skoro takie firmy kradną w imię tak zwanych „wyższych celów”, to "dlaczego nie mają tego robić inni, zwykli użytkownicy sieci? Wielcy monopolisci kształtujący opinię społeczną – pokazują, że można i że dzieje się to bezkarnie".

Teraz kalifornijski sędzia okręgowy, Vince Chhabria uznał, że Meta może trenować swoje modele językowe wykorzystując do tego książki bestsellerowych pisarzy. Bez ich zgody.

Pisarze vs Meta

Firmę pozwało 13 pisarzy, wśród nich komiczka Sarah Silverman, autor kultowego eseju "Między światem a mną" Ta-Nehisi Coates i Junot Diaz, autor "Krótkiego i niezwykłego

życiu Oscara Wao", powieści uważanej za jedną z najwybitniejszych amerykańskich książek XXI wieku.

Decyzja Chhabrii zapadła dwa dni po tym, jak inny sędzia w Kalifornii – William Alsup – orzekł, że firma Anthropic PBC, twórca innego modelu językowego, mogła wykorzystywać przy trenowaniu modeli AI książki, do których nie posiadała licencji.

Obaj sędziowie uznali, że firmy technologiczne używają książek w ramach tzw. dozwolonego użytku (fair use).

Co to jest dozwolony użytek?

Dozwolony użytek to specyficzna instytucja prawa autorskiego, która – zależnie od kraju – mieści różne definicje. Polskie prawo mówi o "dozwolonym użytku", amerykańskie o "uczciwym użytku".

Zasadniczo przyjmuje się, że w ramach dozwolonego użytku zawieszona jest możliwość dochodzenia roszczeń przysługujących twórcom z tytułu naruszenia praw autorskich.

Kiedy korzystanie z książek jest dozwolone i nie trzeba za nie płacić? Na przykład wtedy, gdy kopiujemy książkę wyłącznie na własny użytek, zwłaszcza w celach edukacyjnych. Możemy stworzyć jej kopię, by czytać ją samodzielnie – o ile nie udostępniamy jej innym osobom ani nie rozpowszechniamy dalej.

W ramach dozwolonego użytku możemy też używać fragmentów książek jako materiałów edukacyjnych, czy pożyczać książkę znajomym.

Pisarze nie przedstawili dowodów

Czy trenowanie modeli językowych, które obecnie lub w przyszłości będą używane komercyjnie, jest uczciwe?

Kalifornijscy sędziowie uznali, że tak. Sędzia Chhabria zwrócił uwagę twórcom, że nie przedstawili dowodów na to, że działania Mety miały jakikolwiek wpływ na rynek książki. Innymi słowy – Meta wykorzystując ich książki nie przyczyniła się do ich rozpowszechnienia, a więc nie zmniejszyła potencjalnych dochodów pisarzy i pisarek.

Sędzia zauważył również, że nie ma twardych dowodów, że Llama – model tworzony przez Metę – "jest w stanie odtworzyć małe fragmenty" ich książek.

To nie koniec potyczek sądowych

W kalifornijskim sądzie wciąż pozostaje bez rozstrzygnięcia pytanie, czy Meta naruszyła prawa autorskie twórców, pobierając miliony pirackich książek za pośrednictwem sieci P2P, czyli – jak mówił "Wyborczej" ekspert – sieci komputerów, której użytkownicy m.in. "za pomocą stron torrentowych, oprogramowania klienckiego i zgromadzonych na swoich dyskach zasobów udostępniają i wymieniają się plikami".

- Czy Meta naruszyła prawa autorskie twórców, pobierając miliony pirackich książek za pośrednictwem sieci P2P? Fot. REUTERS/Peter Dasilva

Wyrok w tej sprawie ma zapaść w najbliższych tygodniach. Komentatorzy zwracają uwagę, że duży wpływ na decyzję sądu może mieć rozstrzygnięcie w sprawie przeciwko Anthropic PBC, w której sędzia William Alsup uznał, że firma mogła korzystać z nielegalnych źródeł, ponieważ rezultat jej działań miał charakter „transformacyjny” – czyli przyczyniał się do postępu i tworzenia nowej jakości.

"Technologia, o której mowa, jest jedną z najbardziej transformacyjnych, jakie wielu z nas zobaczy w swoim życiu" – napisał Alsup w swojej decyzji.

Co ciekawe, sędzia Chhabria już teraz był bardzo krytyczny wobec twórców "dużych modeli językowych".

Podkreślił, że jego orzeczenie "nie wspiera tezy, że wykorzystanie przez Meta materiałów chronionych prawem autorskim do trenowania swoich modeli językowych jest zgodne z prawem. Potwierdza jedynie tezę, że powodowie przedstawili błędne argumenty i nie stworzyli dokumentacji na poparcie właściwego".

Meta chwali decyzję sądu

"Modele sztucznej inteligencji typu open source napędzają innowacje, produktywność i kreatywność osób fizycznych i firm, a wykorzystanie materiałów chronionych prawem autorskim w ramach dozwolonego użytku jest kluczowe dla budowy tej transformacyjnej technologii" – napisał rzecznik Meta w oświadczeniu przesłanym mediom.

I choć decyzja sędziego może zaskakiwać, to w uzasadnieniu wyroku znalazły się bardzo ostre słowa krytyki wobec Big Techów. Chhabria wyraził zaniepokojenie faktem, że rynek książek i muzyki może zostać zalany materiałami generowanymi przez sztuczną inteligencję, w efekcie czego AI może "dramatycznie podważyć rynek tych dzieł".

Kalifornijski sędzia zauważył, że firmy muszą znaleźć sposób na to, jak legalnie wchodzić w posiadanie treści wykorzystywanych do trenowania dużych modeli językowych.

"Jeśli korzystanie z dzieł chronionych prawem autorskim do trenowania modeli jest tak konieczne, jak twierdzą firmy, będą musiały znaleźć sposób, aby zrekompensować to właścicielom praw autorskich", powiedział Chhabria.

"Stoimy nad przepaścią"

Tak o przyszłości literatury piszą w liście do największych amerykańskich wydawnictw znani pisarze i pisarki.

Autorzy zwrócili się do wydawców z prośbą, by nie publikowali książek napisanych przy użyciu narzędzi sztucznej inteligencji, nie tworzyli fikcyjnych autorów, ani nie pozwalali prawdziwym autorom na używanie pseudonimów tak, by mogli publikować dzieła napisane przy pomocy AI.

"Najprościej rzecz ujmując, naszym zadaniem jako artystów jest reagowanie na ludzkie doświadczenie. Ale sztuka, którą tworzymy, jest towarem, a nasz świat chce rzeczy szybko, tanio i na żądanie. Pędzimy ku przyszłości, w której nasze powieści, biografie, wiersze i wspomnienia – nasze zapisy ludzkich doświadczeń – są "pisane" przez modele sztucznej inteligencji, które z definicji nie mogą wiedzieć, co to znaczy być człowiekiem. Krwawić, głodować lub kochać".

Czy wydawcy będą chcieli być strażnikami przyszłości?

Czy możliwe do spełnienia będą postulaty takie jak niezastępowanie pracowników narzędziami AI czy zatrudnianie wyłącznie ludzkich narratorów do nagrań audiobooków, zamiast głosów generowanych przez sztuczną inteligencję? Sądząc po tym, że niemal wszyscy wydawcy już dziś korzystają z takich technologii, nie wszystkie postulaty zawarte w liście pisarzy znajdą zrozumienie wśród nich zrozumienie.

"Wzywamy wydawców, by stanęli w obronie autorów – przeciwko kradzieży naszej twórczości i deprecjonowania naszej pracy przez sztuczną inteligencję, która czerpie zyski z tej kradzieży"

– piszą autorzy i wzywają wydawców, by "byli strażnikami przyszłości naszej pracy i pracy przyszłych pokoleń".

Pod listem podpisało się kilkuset autorów, między innymi Lauren Groff, Colleen Hoover, R.F. Kuang, Jodi Picoult, czy autorka bestsellerowych "Lekcji chemii" Bonnie Garmus. Pełna lista jest dostępna pod tym adresem.

Redagowała Magdalena Birska

Wojciech Szot

Dziennikarz działu kultura "Wyborczej", pisze o literaturze. Autor książki "Panna doktor Sadowska" (wyd. Dowody na istnienie), prowadzi blog Zdaniem Szota.

Teraz możesz słuchać naszych artykułów. Wersja audio została stworzona przy pomocy AI. Dowiedz się więcej.