

Meta chciała pozwolić na grożenie śmiercią

19 marca 2022

Firma Meta, właściciel m.in. Facebooka i Instagrama, w zeszłym tygodniu poinformowała swoich moderatorów, że w kilku krajach (m.in. w Polsce) mają nie usuwać nawoływania do śmierci rosyjskich żołnierzy oraz Władimira Putina i Aleksandra Łukaszenki. To nie był dobry pomysł (zresztą Meta po chwili zmodyfikowała tę decyzję, ograniczając jej zasięg do samej Ukrainy). Póki platformy internetowe funkcjonują w obecnym modelu, przyzwolenie na hejt (nawet wobec zbrodniarza wojennego) doprowadziłoby jedynie do wyparcia z naszych strumieni aktualności wszelkich wartościowych informacji na rzecz „zalegalizowanej” mowy nienawiści.

Chociaż na pewnym poziomie decyzja Mety wydaje się zrozumiała ze względu na wyjątkowe okoliczności i brutalny kontekst wojny, to jej konsekwencje mogłyby być fatalne. Po pierwsze, groziła ona narastaniem w sieci fali nienawiści, która najpewniej rozlałaby się także na zwykłych Rosjan. Po drugie, doprowadziłaby do jeszcze większego zanieczyszczenia platformy społecznościowej przez negatywny przekaz, który – podbijany przez algorytmy preferujące angażujący content – wyparłby bardziej wartościowe, a za to mniej angażujące wiadomości. W rezultacie jeszcze trudniej byłoby trafić na wiarygodne i istotne informacje: czy to o akcjach pomocowych na rzecz osób dotkniętych wojną, czy to o samym konflikcie. Pamiętajmy, że media społecznościowe dla wielu osób są dziś głównym, jeśli nie jedynym, kanałem informacyjnym.

Takie decyzje przypominają, że ustalając i egzekwując swoją politykę moderacji treści, Meta ma większy wpływ na wyznaczanie faktycznych granic debaty publicznej niż politycy, sądy i inne demokratyczne instytucje. W dodatku to nie jedyne narzędzie kontroli nad globalnym obiegiem treści i sterowania

naszym postrzeganiem świata, którym dysponuje ta firma. Wspomniane algorytmy, profilujące użytkowników i personalizujące treści (zarówno sponsorowane, jak i organiczne), odpowiadają za to, co ostatecznie zobaczysz i przeczytasz na „swoim” Facebooku czy Instagramie, jak również za to, do kogo dotrze to, co opublikujesz. Jako użytkownicy nie mamy nad tym żadnej kontroli.

Instytucje Unii Europejskiej właśnie negocjują długo oczekiwany akt o usługach cyfrowych (Digital Services Act, DSA). Ta regulacja prawna wprowadzie nie wymusi radykalnej przebudowy modelu funkcjonowania wielkich platform (takie zmiany proponowały organizacje społeczne, w tym Panoptikon), ale obiecuje lepszą ochronę użytkowników przed jego negatywnymi skutkami.

Digital Services Act ma wzmocnić gwarancje dla ofiar nienawistnych treści, wprowadzając m.in. instytucję zaufanych podmiotów zgłaszających tego rodzaju materiały. Na plus zmienia się też zasady rozliczania platform z tego, jak moderują treści: DSA wymusi większą przejrzystość i ograniczy arbitralność takich decyzji. Użytkownicy będą szczegółowo informowani o przyczynach zablokowania konkretnej treści lub całego konta oraz o tym, w oparciu o jaki mechanizm taka blokada została nałożona. Dostaną możliwość przedstawienia swoich argumentów w sporze oraz gwarancję, że ostatecznych decyzji nie będą podejmowały algorytmy. DSA ma też wprowadzić system niezależnego nadzoru nad rozstrzygnięciami samych platform.

Jeśli chodzi o regulację algorytmów wykorzystywanych do personalizowania treści, liczymy na to, że w rozporządzeniu znajdą się:

- obowiązkowa ocena ryzyka związanego z funkcjonowaniem algorytmów odpowiedzialnych za rekomendowanie treści na platformach (polegającego np. na regularnym podbijaniu zasięgów treści zawierających hejt czy dezinformację);

- większa kontrola użytkowników nad wyborem logiki tych systemów;
- zakaz wykorzystywania tzw. danych wywnioskowanych (które platformy generują często bez wiedzy użytkowników i użytkowników, na podstawie ich aktywności w sieci) do celów reklamowych, w szczególności danych ujawniających wrażliwe cechy, takie jak pochodzenie etniczne, religia czy stan zdrowia.

Digital Services Act wyposażony w te wszystkie elementy ograniczyłby manipulację bazującą na wykorzystaniu ludzkich słabości i inne szkodliwe praktyki stosowane przez platformy internetowe. To, czy uda się je przeforsować w finalnym tekście rozporządzenia, zależy jednak od wyniku negocjacji między Komisją Europejską, Radą i Parlamentem, które właśnie teraz wkroczyły w decydującą fazę. A kluczowy głos będzie należeć do Rady UE – czyli państw członkowskich, a więc m.in. polskiego rządu.

Autorstwo: Dorota Głowacka, Anna Obem, Katarzyna Szymielewicz
Źródło: [Panoptykon.org](https://panoptykon.org)