

Kto sprząta Internet podczas pandemii?

13 maja 2020

Nie tylko prywatność może ucierpieć podczas pandemii. I nie tylko z powodu działań podejmowanych przez rządy – uważajmy także na to, co w tym czasie robią platformy społecznościowe i jak może to wpłynąć na naszą wolność wypowiedzi i swobodny dostęp do informacji.

Platformy społecznościowe, takie jak Facebook i YouTube, poinformowały, że z uwagi na zagrożenie koronawirusem część osób odpowiedzialnych za moderację treści w tych serwisach została odesłana z biur do domów. W przypadku Facebooka na miejscu zostali jedynie pełnoetatowi pracownicy, którzy będą koncentrować się na zwalczaniu treści powodujących bezpośrednie zagrożenie dla życia lub zdrowia (czyli dotyczących np. samobójstw, samookaleczenia, wykorzystywania seksualnego dzieci czy terroryzmu). Większość ludzi zatrudnionych do „sprzątania” niechcianych treści nie pracuje jednak bezpośrednio dla Facebooka, tylko dla współpracujących z nim firm zewnętrznych. Platforma przyznała, że nie jest w stanie zagwarantować im bezpiecznych warunków pracy zdalnej (np. wsparcia psychologicznego), których niewątpliwie wymaga ta brudna robota. Warto przypomnieć, że w przeszłości poziom odpowiedniej ochrony zapewnianej moderatorom – eksponowanym m.in. na sceny przemocy i inne drastyczne treści – wzbudzał poważne wątpliwości, co dobrze pokazuje film dokumentalny „Czyściciele Internetu” oraz reportaż portalu „The Verge”. Zdalna praca moderatorów miałaby ponadto stwarzać ryzyko dla prywatności użytkowników.

Kto zatem będzie czyścił Internet pod ich nieobecność? Platformy przekazały to zadanie w ręce algorytmów. Fakt, że biorą one udział w procesie moderacji, to oczywiście żadna nowość. Teraz jednak automatyczne filtry mają zastąpić czynnik

ludzki w znacznie większym stopniu niż dotychczas, a wiele decyzji o usunięciu treści będzie podejmowanych bez udziału człowieka. W praktyce oznacza to śledzenie i skanowanie wszystkich materiałów publikowanych przez użytkowników, ale także ogromne ryzyko nadmiernego usuwania treści. Także takich, które wcale nie są szkodliwe, a zostały tylko źle ocenione przez algorytm, który popełnił błąd w analizie obrazu albo nie wychwycił satyrycznego kontekstu. Jakiś czas temu informowaliśmy na przykład o usunięciu posta organizacji kampanijnej WeMove, która udostępniła na swoim profilu krytyczny artykuł portalu Euroactiv na temat polityki migracyjnej Unii Europejskiej. Artykuł był zilustrowany zdjęciem wyrzuconych na plażę ciał migrantów z Afryki, którzy utonęli w drodze do Europy. Facebook usunął go z uwagi na naruszenie „zasad dotyczących nagości i czynności seksualnych”. Jak widać, automatycznym filtrem daleko do perfekcji.

Same platformy nie kryją zresztą, że większe zaangażowanie algorytmów może spowodować – jak ujął to YouTube – „przypadkowe usunięcia treści, które nie naruszają ich zasad”. Tymczasem przypomnijmy, że media społecznościowe to dla wielu osób, firm, organizacji ważne narzędzie pracy, komunikacji, organizowania pomocy dla innych czy czerpania wiedzy o bieżących wydarzeniach. Jak podaje Pew Research Center, w 2018 r. w Polsce 30 proc. badanych pozyskiwało wiadomości z Facebooka codziennie, a w grupie wiekowej 18–29 odsetek ten wynosił aż 72 proc. Nie trzeba chyba nikogo przekonywać, że w trakcie aktualnego kryzysu, kiedy jeszcze więcej aspektów naszego życia zostało zepchniętych do sfery online, a wiele osób bardziej niż zwykle łaknie dostępu do bieżących informacji o sytuacji zdrowotnej, politycznej czy gospodarczej, znaczenie platform dodatkowo wzrosło (a wraz z nim ich zyski).

Filtr ochronny czy ochrona przed filtrem?

Zakładając, że decyzja o odesłaniu moderatorów do domów była w obecnych okolicznościach nieunikniona oraz że jakaś moderacja treści dodawanych przez użytkowników jest konieczna (jako że wyzwań związanych z niebezpiecznymi treściami w czasie pandemii nie ubyło, a wręcz przeciwnie) – czy można zrobić cokolwiek, aby ochronić nas przed arbitralną oceną i usuwaniem treści przez automatyczne filtry? Są dwa rozwiązania, które (łącznie) pomogłyby zrównoważyć to zagrożenie. Po pierwsze: wprowadzenie skutecznej procedury odwoławczej od decyzji o usunięciu treści. Po drugie: wprowadzenie gwarancji przejrzystości stosowanych algorytmów.

Co do pierwszej kwestii – Facebook, ogłaszając zmiany w systemie moderacji, poinformował, że... zawiesza możliwość wniesienia odwołania od decyzji o usunięciu treści. A dokładnie: będzie można wyrazić swój sprzeciw wobec zasadności blokady, z którym firma się zapozna, może wykorzystać do tego, aby usprawnić swój mechanizm odwoławczy w przyszłości, ale na razie raczej nie będzie przywracać usuniętych materiałów. Można się jedynie pocieszyć, że w praktyce ten proces odwoławczy i tak mocno odbiegał od ideału skuteczności. Dobrą ilustracją jest tutaj sprawa SIN vs Facebook, w której walczymy o ochronę praw osób, których wolność słowa została niesłusznie ograniczona poprzez arbitralną blokadę nałożoną przez platformy internetowe. Przypomnijmy, że strona i grupy Społecznej Inicjatywy Narkopolityki (SIN) zostały bez ostrzeżenia i przejrzystych wyjaśnień usunięte przez Facebook jako „niezgodne ze Standardami społeczności”. Proces odwoławczy sprowadzał się do wykorzystania przycisku „odwołaj się” w komunikacie o usunięciu treści, bez możliwości przedstawienia jakiejkolwiek argumentacji. Z kolei na ten ruch platforma już nie zareagowała. Członkowie SIN do tej pory nie wiedzą, które z publikowanych przez nich materiałów i z

jakiego powodu Facebook uznał za naruszenie swoich zasad. Proces jest w toku, mamy nadzieję, że usunięte strony uda się przywrócić na mocy decyzji sądu. Na razie sąd nakazał Facebookowi tymczasowe zabezpieczenie zablokowanych materiałów i zakazał usuwania treści publikowanych aktualnie przez SIN na ich nowym profilu.

YouTube co prawda nie zawiesił możliwości składania odwołań, ale poinformował, że mogą być one rozpoznawane „z opóźnieniem”. Na użytkowników nie będą natomiast przynajmniej nakładane tzw. ostrzeżenia, chyba że... pracownicy serwisu będą mieli „silne przekonanie”, że doszło do naruszenia jego zasad.

Jeśli chodzi o drugą kwestię, tj. transparentność algorytmów – przed pandemią zablokowany użytkownik nie wiedział, czy sporne treści zostały wyłapane przez automatyczny filtr (czy też np. zgłoszone przez innego użytkownika jako szkodliwe) ani czy decyzję o ich usunięciu podjął człowiek, czy automat. Nie wiedzieliśmy też, jak często algorytmy na platformach się mylą, w oparciu o jakie dane działają i czy przypadkiem nie są zbudowane tak, że dyskryminują pewne treści. Teraz niestety niewiele się zmieniło. Dowiedzieliśmy się jedynie, że algorytmy będą odgrywały większą rolę, ale szczegóły ich funkcjonowania wciąż pozostają owiane tajemnicą.

Wnioski (i czy jest dla nas nadzieja?)

Jakie możemy wyciągnąć wnioski z tej sytuacji? Reakcja platform na kryzys wywołany przez koronawirusa potwierdza, że realne gwarancje dla użytkowników platform nie mogą wynikać z ich wewnętrznych regulacji, bo te firmy internetowe mogą w każdej chwili dowolnie zmienić. Nie ma też żadnej pewności, że po pandemii platformy zrezygnują z tych tymczasowych rozwiązań. Problemu tego nie rozwiążą więc nawet najlepsze, przyjmowane dobrowolnie „standardy społeczności” ani rada ekspertów utworzona właśnie przez Facebooka do rozstrzygania

sporów dotyczących treści. Nie oznacza to, że nie należy wspierać dążenia platform do tworzenia dla siebie jak najbardziej przejrzystych i sprawiedliwych zasad funkcjonowania. One jednak nie wystarczą. Od jakiegoś czasu stało się jasne, że potrzebujemy bardziej stabilnego zabezpieczenia dla naszych praw – w postaci regulacji prawnych, które z jednej strony narzuca platformom standardy, poniżej których nie będą mogły schodzić, a z drugiej strony wprowadzą skuteczny mechanizm ich egzekwowania.

Być może szansą na przyjęcie takich regulacji będzie kodeks usług cyfrowych (Digital Service Act), nad którym prace rozpoczęła niedawno Komisja Europejska. Kodeks może dać przestrzeń, aby w końcu uregulować wiele palących kwestii związanych z funkcjonowaniem platform internetowych i wymusić na nich większe poszanowanie praw swoich użytkowników. Jednym z tych zagadnień jest niewątpliwie zarządzanie treścią w mediach społecznościowych w taki sposób, żeby skutecznie walczyć z tym, co szkodliwe, ale jednocześnie ukrócić tzw. prywatną cenzurę. W tym zakresie kluczowe jest wprowadzenie obowiązku stworzenia sprawiedliwej procedury stosowanej przez same platformy przy okazji rozstrzygania sporów o treści, a także zewnętrznej, niezależnej kontroli nad ich decyzjami. Taką kontrolę powinny sprawować sądy, tradycyjnie posiadające kompetencję do oceny tego, czy dana wypowiedź mieści się w granicach wolności słowa. Musiałyby jednak zostać wyposażone w odpowiednie zasoby, które gwarantowałyby ich realną dostępność i efektywność w kontekście spodziewanej dużej liczby takich spraw. Jak bardzo jest to ważne, widać na przykładzie wspomnianej wyżej sprawy SIN vs Facebook, która mimo upływu roku od złożenia pozwu utknęła na kwestiach proceduralnych i w efekcie jest wciąż na wczesnym etapie. Pokazuje to, że kontrola sądowa w obecnym kształcie nie stwarza wystarczająco dobrej ochrony. Jest to kolejny argument, który przemawia za pilnością interwencji regulacyjnej oraz za tym, aby na poważnie rozważyć przedstawioną przez koalicję European Digital Rights – EDRi (której członkiem jest Panoptikon)

propozycję wsparcia sądów poprzez utworzenie specjalnych organów quasi-sądowych, które rozstrzygałyby spory dotyczące treści na platformach znacznie szybciej i taniej.

Z uwagi na szczególnie duże ryzyko arbitralności sprawiedliwa procedura i niezależny nadzór są szczególnie istotne w przypadku wykorzystywania algorytmów do oceny i usuwania treści. Aby dodatkowo zminimalizować to ryzyko, kodeks usług cyfrowych powinien narzucić platformom odpowiednie standardy przejrzystości stosowania takich algorytmów, pozwalające użytkownikom zrozumieć, kiedy i na jakiej zasadzie one działają, jak można kwestionować ich decyzje czy w jakich sprawach dotyczących się myliły. To jednak wciąż nie wystarczy – EDRi sugeruje ponadto powołanie specjalnego organu kontrolnego na poziomie UE, który miałby m.in. prawo przeprowadzać niezależny audyt takich narzędzi.

Do wprowadzenia powyższych rozwiązań jeszcze długa i pewnie wyboista droga, wypełniona trudnymi negocjacjami w Brukseli. Pamiętajmy jednak, że bez tych gwarancji swobodny przepływ informacji w sieci będzie zagrożony – i to nie tylko w czasie pandemii. Jeśli nie chcemy, żeby platformy dowolnie decydowały o tym, co możemy powiedzieć czy przeczytać w sieci, teraz jest czas, żeby o to zawałczyć i przekonać do tego europejskich decydentów.

Autorstwo: Dorota Głowacka

Źródło: [Panoptikon.org](https://panoptikon.org)

Bibliografia

1.

<https://about.fb.com/news/2020/05/coronavirus/#content-review>

2.

https://support.google.com/youtube/answer/9777243?p=covid19_updates&visit_id=637194868739116880-4122912225&rd=1

3.

<https://panoptikon.org/wiadomosc/kto-czysci-dla-nas-internet-i-dlaczego-nadal-jest-tak-brudno>

4.

<https://www.theverge.com/2019/2/25/18229714/cognizant-facebook-content-moderator-interviews-trauma-working-conditions-arizona>

5.

https://support.google.com/youtube/answer/9777243?p=covid19_updates&visit_id=637194868739116880-4122912225&rd=1

6. <http://www.pewglobal.org/interactives/media-habits-table>

7. <https://panoptikon.org/gafa-zniwa-w-pandemii>

8. <https://panoptikon.org/sinvsfacebook>

9.

<https://panoptikon.org/wiadomosc/pierwsza-decyzja-sadu-w-sprawie-sin-vs-facebook>

10. <https://www.oversightboard.com/>

11. <https://panoptikon.org/digital-services-act>

12.

<https://panoptikon.org/wiadomosc/czy-sprawa-przeciwko-facebookowi-w-koncu-przyspieszy>

13.

https://edri.org/wp-content/uploads/2020/04/DSA_EDRiPositionPaper.pdf