



Horyzonty Polityki
2024, Vol. 15, N° 53



ANNA LESZCZUK-FIEDZIUKIEWICZ

<http://orcid.org/0000-0002-8800-2486>

Uniwersytet w Białymstoku
a.fiedziukiewicz@uwb.edu.pl

DOI: 10.35765/HP.2632

Polityka mediów społecznościowych wobec „niewygodnych treści”

Streszczenie

CEL NAUKOWY: Refleksja nad rolą mediów społecznościowych w kształtowaniu polityki wobec treści wyrażających pogardę i nienawiść oraz wyjaśnienie kontrowersyjnej zgody koncernu Meta na zamieszczanie wpisów życzących śmierci W. Putinowi i A. Łukaszence, usprawiedliwianej jako „tymczasowy wyjątek”.

PROBLEM I METODY BADAWCZE: Problemem jest postawa Mety dot. mowy nienawiści po agresji Rosji na Ukrainę w 2022 r. Metodami badawczymi były krytyczna analiza dyskursu, teoria ramowania oraz analiza danych zastanych.

PROCES WYWODU: Rozwinięto trzy potencjalne odpowiedzi na pytania związane z decyzjami Mety, nazwane interpretacjami. Pierwsza dotyczyła możliwości strukturalnych firmy, zdefiniowanych jako rozumienie potrzeb użytkowników („wyrażanie wściekłości”). Druga interpretacja wiązała się z użyciem sztucznej inteligencji. Eksperymenty inżynierów były normą, a podczas wojny hybrydowej niekonwencjonalne metody zarządzania treścią stały się przydatne w tworzeniu nowych algorytmów. Trzecie wyjaśnienie koncentrowało się na potrzebie retuszu skompromitowanej reputacji Facebooka wskutek skandalu z Cambridge Analytica i narracji F. Haugen o manipulacjach firmy.

WYNIKI ANALIZY NAUKOWEJ: Wyniki analiz pokazały, że pozwolenie na życzenie śmierci Putinowi i Łukaszence było swoistym eksperymentem

Sugerowane cytowanie: Leszczuk-Fiedziukiewicz, A. (2024). Polityka mediów społecznościowych wobec „niewygodnych treści”. *Horyzonty Polityki*, 15(53), 181–204. DOI: 10.35765/HP.2632.

napędzającym ruch online i karmiącym sztuczną inteligencję. Facebook przez lata realizował pomysły mające ograniczyć mowę nienawiści, a brak precyzji w określaniu regulaminu tłumaczył niuansami kulturowymi lub problemami z moderatorami. Analiza obnażyła ciemną stronę hegemonia komunikacji (3 mld użytkowników), a zwłaszcza mechanizmy zarządzania treściami, które prowadziły do polaryzacji poglądów i elektronicznych wojen, również w kontekście wojny hybrydowej z Rosją.

WNIOSKI, INNOWACJE I REKOMENDACJE: Media społecznościowe stały się narzędziem masowego hejtingu i trollingu, rosną więc społeczne oczekiwania, by ich właściciele reagowali na nie, łącząc paradoksy: bezpieczeństwo w internecie z wolnością słowa. Publikacja jest pretekstem do debaty publicznej o roli inżynierów w moderowaniu treści pojawiających się w internecie. Z uwagi na dynamiczność i skomplikowanie problemu zarządzanie mową nienawiści wymaga pogłębionej refleksji wielu dyscyplin nauki oraz zaangażowania wielu aktorów społecznych na poziomie prawnym, edukacyjnym i technologicznym.

SŁOWA KLUCZOWE:

mowa nienawiści, trolling, polaryzacja, media społecznościowe, zagrożenia internetowe

Abstract

SOCIAL MEDIA POLICY TOWARDS
“INCONVENIENT CONTENT”

SCIENTIFIC OBJECTIVE: To reflect on the role of social media in shaping policy towards content expressing contempt and hatred, and to explain the controversial approval of the Meta concern for posts wishing death to V. Putin and A. Lukashenko, justified as a ‘temporary exception’

THE RESEARCH PROBLEM AND METHODS: The problem is Meta's attitude towards hate speech after Russia's aggression against Ukraine in 2022. The research methods were critical discourse analysis, framing theory and analysis of found data.

PROCESS OF ARGUMENTATION: Three potential answers to questions related to Meta's decisions, called interpretations, were developed. The first related to the company's structural capabilities, marketing-defined as understanding user needs (‘expressing rage’). The second interpretation involved the use of artificial intelligence and new algorithms. The third explanation focused on the need to retouch Facebook's discredited reputation.

RESEARCH RESULTS: The results of the analysis showed that allowing hate speech was a kind of experiment to drive online traffic and feed artificial intelligence. Facebook has for years pursued ideas to curb hate speech, and explained its lack of precision in defining its rules and regulations by cultural nuances or problems with moderators. The analysis exposed the dark side of the communications hegemon (3 billion users), especially content management mechanisms that led to polarised views and electronic warfare, including in the context of the hybrid war.

CONCLUSIONS, INNOVATIONS AND AND RECOMMENDATIONS: As social media has become a tool for mass trolling, there is a growing public expectation for its owners to respond by combining paradoxes: online safety and freedom of expression. The publication is a pretext for a public debate about the role of engineers in moderating content appearing on the internet. Given the dynamism and complexity of the problem, the management of hate speech requires in-depth reflection from multiple scientific disciplines and the involvement of multiple social actors at the legal, educational and technological levels.

KEYWORDS:

hate speech, trolling, polarisation, social media, online threats

1. WSTĘP, CZYLI MOWA NIENAWIŚCI TO „TYMCZASOWY WYJĄTEK”

Agencja informacyjna Reuters w marcu 2022 r. poinformowała, że szef koncernu Meta Mark Zuckerberg, będący właścicielem Facebooka i innych mediów społecznościowych (takich jak WhatsApp, Instagram), przesłał pozwolenie na niereagowanie na mowę nienawiści na profilach swoich użytkowników, zgadzając się na nawoływanie do zabójstwa Władimira Putina i Aleksandra Łukaszenki oraz na groźby pod adresem rosyjskich żołnierzy. Pod dwoma warunkami: po pierwsze, posty nie będą zawierały szczegółów i metody zabójstwa, po drugie, nie będą dotyczyły rosyjskich cywilów. Andy Stone, szef komunikacji Meta, tłumaczył, że jego firma nie toleruje rusofobii ani żadnego rodzaju dyskryminacji i prześladowania, ale:

W świetle trwającej inwazji na Ukrainę zrobiliśmy tymczasowy wyjątek dla osób dotkniętych wojną, aby mogli wyrazić gwałtowne uczucia wobec atakujących sił zbrojnych (...) Są to tymczasowe

środki, których celem jest wyrażenie ekspresji przez ludzi, którzy stoją w obliczu inwazji.

Pozwolenie to obejmowało mieszkańców wybranych krajów: Ukrainy, Rosji, Polski, Węgier, Słowacji, Armenii, Azerbejdżanu, Litwy, Łotwy, Estonii, Gruzji. Logiki geografii wyboru nie wyjaśniono. Większość dziennikarzy śledzących kroki technologicznego giganta elektryzowało nagłówkami: „Facebook pozwala na mowę nienawiści wobec Putina” (Newsweek), „Nie życz nikomu śmierci, chyba że Putinowi” (Telepolis) itp. Podobnie kuriozalne było wycofanie się koncernu z tej decyzji. Kilka dni potem inny reprezentant firmy wyjaśniał, że nowa zasada będzie dotyczyć tylko mieszkańców Ukrainy. Zmianę Nick Clegg, dyrektor Mety ds. globalnych, tłumaczył: „zasady moderacji treści na platformie nie powinny być interpretowane jako «generalna akceptacja przemocy wobec Rosjan»”, a „wezwania do śmierci głowy państwa są zakazane”. Pisano o kolejnej rysie na wizerunku, niekonsekwencji i hipokryzji.

2. PROBLEM I METODY BADAWCZE

Celem artykułu, poza zbadaniem powodów kontrowersyjnych decyzji firmy Meta, jest zastanowienie się nad kształtowaniem odpowiedzialności mediów społecznościowych za ich politykę walki z mową nienawiści. Pytania badawcze skupiają się na etyczności działań i rozdźwięku między deklaracjami a działaniami, bo od kilku lat zauważa się nieprzejrzystość działań i proces monetyzacji treści umieszczonych na platformach społecznościowych. Czy koncern Meta pozwolił na mowę nienawiści z powodów komercyjnych, chcąc przyciągnąć reklamodawców, czy z powodów wizerunkowych, chcąc odbudować reputację? Czy zgoda na elektroniczny hejting była próbą odświeżenia relacji z użytkownikami? A może próbą użycia sztucznej inteligencji, która ma go eliminować?

Jednym z istotnych pytań badawczych tego artykułu jest pytanie o zgodę na mowę nienawiści wobec przywódców Federacji Rosyjskiej i Białorusi. Pierwsza hipoteza dotyczy odpowiedzi na potrzeby użytkowników oburzonych atakiem na Ukrainę w 2022 r. Powodem rezygnacji z decyzji mogły być możliwości strukturalne i nienadążanie moderatorów za detekcją materiałów. Druga hipoteza wiąże

się ze strategią firmy, polegającą na sprawdzeniu działania nowych algorytmów sztucznej inteligencji. Eksperymentowanie było w firmie normą, a w trakcie wojny hybrydowej w 2022 niekonwencjonalne metody zarządzania treścią mogły być przydatne. Trzecia hipoteza podpowiada potrzebę retuszu reputacji skompromitowanej za sprawą afer z Cambridge Analytica i wypowiedziom Frances Haugen o manipulacyjnych algorytmach Facebooka.

Metodami badawczymi były: krytyczna analiza dyskursu, teoria ramowania (*framing theory*) oraz analiza danych zastanych. W konstrukcji pojęć teoretycznych zastosowano socjologiczne i medioznawcze koncepcje mowy nienawiści, trollingu i flamingu. Istotny był też dyskurs dziennikarski widoczny na portalach związanych z branżą technologiczną (IT).

Krytyczna analiza dyskursu zakładała analizę wypowiedzi oraz kontekstów, w jakich one powstawały (Kawka, 2011, s. 216). Koncentracja na dyskursie technologicznym wynikała ze specyfiki działania firmy Meta. Jedną z metod była analiza treści publikacji dotyczących przemiany Facebooka po 2016 r., ponieważ Facebook był miejscem eksperymentów z algorytmami, przyciągał uwagę opinii publicznej, ponadto – jak pokaże casus wojny na Ukrainie – również zaangażuje się w politykę. Gatunkowo publikacje analizowane w artykule były różnorodne (reportaż, książka popularnonaukowa, praca naukowa), a łączącą ich klamrę stanowiła ewolucja polityki platformy. Scott Galloway, jeden ze znawców rynku usług informatycznych, w książce *Wielka czwórka. Four. Ukryte DNA: Amazon, Apple, Facebook i Google* (2018, rok wydania podany w bibliografii dotyczy tłumaczenia na język polski) skoncentrował się na specyfice kultury organizacyjnej Doliny Krzemowej, która promuje niestandardowe zachowania, wychodzące poza granice norm prawnych i społeczno-etycznych. Inna pozycja *Nabici w Facebooka. Przestroga przed katastrofą* (2019) oprócz tytułu intryguje tym, że napisał ją doradca Marka Zuckerberga, Roger McNamee, początkowo inwestor, potem krytyk firmy przestrzegający przed jej manipulacjami. Sheera Frenkel i Cecilia Kang wydały książkę *Brzydka prawda. Kulisy walki Facebooka o dominację* (2021) o przyczynach procesów sądowych wytoczonych firmie Zuckerberga. Śledztwo prowadzone przez dziennikarki New York Timesa polegało na nieformalnym dojściu do dokumentów Facebooka i przeprowadzeniu rozmów z jego pracownikami. Publikacja *Facebook*.

A miało być tak pięknie. Za kulisami imperium z Doliny Krzemowej Stevena Levy'ego (2022) to efekt wieloletniej reporterskiej pracy, polegającej na analizie dyskursu medialnego nt. Facebooka i przeprowadzeniu wywiadów z pracownikami firmy, a także ekspertami zajmującymi się nowymi technologiami. Autor demaskuje początkowe postulaty Zuckerberga, konfrontując je z działaniami serwisu, który manipulował bezpieczeństwem danych jego użytkowników.

Patrząc krytycznie, należy stwierdzić, że większość wspomnianych pozycji nie ma charakteru naukowego, dominuje w nich retoryka dziennikarska, nastawiona na obnażenie sekretów technologicznego giganta. Wypowiedzi mają charakter wtórny i nieoficjalny, nie pochodzą od samego Zuckerberga (jedynie McNamee i Levy przeprowadzili z nim wywiady). Naukowy charakter mają jedynie dwie pozycje: *Antisocial Media. Jak Facebook oddala nas od siebie i zagraża demokracji* (2018), której autor, Siva Vaidhyanathan, medioznawca, profesor Uniwersytetu Virginia, jest bezlitosny dla manipulacji i nieuczciwości platformy, oraz *Władza platform* Jana Krefta, która stanowi naukowe podejście do globalnej polityki technologicznych gigantów i jej konsekwencji społecznych. W wymienionych publikacjach do kategorii analitycznych zaliczono następujące problemy: 1) bezpieczeństwo danych użytkowników, 2) wprowadzanie przez Facebook nowych algorytmów, 3) skandale dotyczące kampanii wyborczej w USA i Cambridge Analytica, 4) polityka Facebooka wobec mowy nienawiści.

3. CZĘŚĆ ZASADNICZA

3.1. Hejting i trolling w definicjach i koncepcjach

Zjawisko hejtingu początkowo opisywały takie synonimy jak: „bluzg”, „agresywny wpis”, „nienawistny mem” (Juza, 2015, s. 30) lub „agresja elektroniczna” (Pyżalski, 2015). Mówienie o nienawiści jako zwerbalizowanej emocji jest hiperbolą. Definicje są semantycznie wyolbrzymione, nie chodzi w nich o „wywrotową negatywność”, bardziej o „uczucie wrogości” (Żmuda, 2016, s. 42) lub „narzędzie służące wyrządzeniu krzywdy” (Olczyk, 2017, s. 35). Obecne w dyskursie naukowym pojęcia nie uwzględniają filozoficznych odniesień,

które nienawiść definiują jako wzdarcę dla słabych, chorych i biednych lub trujący od środka resentment (Żardecka-Nowak, 2016, s. 59). Eksperti od retoryki i politologii, opisując agresję językową w debacie publicznej, piszą o „retoryce nienawiści”, która przejawia się w ośmieszaniu, degradowaniu i dyskredytacji przeciwnika (Głowiński, 2007, s. 24) oraz przypisywaniu osobom lub grupom szczególnie negatywnych cech lub uczynków (Nijakowski, 2008, s. 113). Pionierskie badania nad hejtingiem są niepełne i pokazują skrajne stanowiska. Językoznawcy określają ją wulgarnie zwroty, służące obrażaniu konkretnych grup, a do jej przyczyn zaliczają chęć deprecjacji, frustrację, potrzebę zauważenia (Naruszewicz-Duchlińska, 2015, s. 35) oraz patologicznie rozumianą wolność słowa (Osowiecka, 2015). W opinii socjologów źródłami mowy nienawiści są pogarda i zepsucie życia politycznego w Polsce (Bilewicz, 2019, s. 27) oraz kulturowy zanik prywatności i medialny ekshibicjonizm (Frącz, 2019, s. 30). Diagnozuje się pojawienie się wielu odmian hejtu, np. hejtu swojskiego, będącego konsekwencją wojny polsko-polskiej prowadzonej przez polityków (Leszczuk-Fiedziukiewicz, 2020, s. 102). Zdaniem badaczy hejting jest inspirowany przez właścicieli platform, ponieważ nienawistne komentarze napędzają dyskusję i zwiększają zyski (Juza, 2015, s. 27).

Trolling jest nowym zjawiskiem, bliskim hejtingowi, definiowanym jako:

intencjonalne publikowanie w Sieci wrogich, obraźliwych, kontrowersyjnych wiadomości, wywołujących gwałtowną reakcję innych użytkowników, przybierającą formę serii wzajemnych reprimend, powodującą eskalującą kłótnie wojnę (*flame war*) i wymykającą się spod kontroli (Kamińska, 2008, s. 288).

Zdaniem badaczy retoryczna strategia trolla polega na stworzeniu kilku fikcyjnych osobowości i wygłaszaniu kontrowersyjnych poglądów. Owe „marionetkowe osoby” (*sock puppetry*) są aktywne w internecie, używając kłamstw, gróźb, sarkazmu, demonstracyjnie łamiąc zasady netykiety, innych internautów oskarżając o to samo (Kamińska, 2010, s. 644). Najbardziej niebezpieczną formą jest trolling zorganizowany, który „ma szerzyć ideologie destrukcyjne i działać na negatywne emocje społeczne”, a jego celem jest „uzyskiwanie korzyści ideologicznych, pozyskiwanie zwolenników lub dyskredytowanie

oponentów” (Warchala, 2020, s. 240). Znaczący mechanizmów rosyjskiej propagandy twierdzą, że trolle działali w świecie online długo przed wybuchem wojny na Ukrainie. Według jednego z nich trolle

[w]ypełniają internet zaprzeczeniami, budują chaos plotkami i dezinformacją, zaostrzają konflikt wewnętrzny (przez podkreślanie niekompetencji i sporów politycznych), stosują groźby, koncentrują się na podżeganiu narodowej, etnicznej, religijnej nienawiści i na budowaniu teorii spiskowych (Wasiuta, 2017, s. 178).

Trolling staje się użyteczny w niszczeniu reputacji osób i organizacji i „każdy może tanio stworzyć własną armię botów lub założyć własną farmę prywatnych fałszywych kont na Facebooku i Twitterze” (Aro, 2020, s. 18). Stosowały go rosyjskie służby wywiadowcze już w 2013 r. Federalna Agencja Informacyjna (FAN), nazywana farmą trolli, nadzorowała 12 firm, które tworzą nieprawdziwe informacje. Zatrudnieni w nich menadżerowie, blogerzy, tworzyli treści propagandowe na potrzeby Kremla i rozsyłali w świecie online. Relacja fińskiej dziennikarki i aktywistki Jessikki Aro, badającej rosyjską propagandę, udowodniła, że trolle mogą poprzez zmasowane dyskredytacje zniszczyć komuś życie, zmuszając do ucieczki z własnego kraju. Destrukcyjny potencjał internetu, obecny w trollingu, zauważa już wielu badaczy, pisząc o medium, które sprzyja ekshibicjonizmowi i narcyzmowi (Levinson, 2010, s. 98), promuje kulturę uczestnictwa (Ogonowska, 2014, s. 33), ułatwia zachowania dewiacyjne i patologiczne (Leszczuk-Fiedziukiewicz, 2018, s. 113), będąc ponadto źródłem ludzkich cierpień (Vaidhyanathan, 2018, 331). Co z tym wszystkim mają wspólne Facebook i wojna na Ukrainie?

3.2. Rozumienie gniewu i wściekłości użytkowników – hipoteza pierwsza

Po ataku na Ukrainę uśmiercenie rosyjskiego dyktatora było obecne w mediach na całym świecie, w pytaniach dziennikarzy i odpowiedziach ekspertów symulujących rzeczywistość polityczną po skończeniu wojny (kto będzie rządził po Putinie? czy są sposoby jego eliminacji?). Służyła temu tabloidyzacja newsów o chorobie dyktatora, wzbogacana pseudorelacjami lekarzy o nowotworze i złych

rokowaniach na wyzdrowienie. Dziennikarze definiowali Putina: morderca, potwór, ludobójca. Niemal na całym świecie stał się globalnym generatorem negatywnych emocji, takich jak: złość, wściekłość, chęć zemsty.

Zgoda na mowę nienawiści była być może psychologiczną odpowiedzią na potrzeby użytkowników, która zawierała obietnicę: damy wam miejsce na emocje, zawiesimy nasze ostre zasady, tj. Standardy społeczności, na potrzeby waszej ekspresji. Żądanie śmierci może być desperacką reakcją na śmierć niewinnych ludzi i jednoczesnym oczekiwaniem na sprawiedliwość. Może być próbą dyskredytacji i dewaluacji wroga, który w sposób mitomański informował świat o hegemonicznych aspiracjach. Życzenie śmierci, na które zgadzał się Facebook, mogło być propozycją sposobu na odreagowanie negatywnych emocji przez ludzi wstrząśniętych makabrycznymi obrazami wojny (np. masakry w Buczy) lub formą elektronicznej mobilizacji zbiorowej przeciwko wrogowi, co zwykle pełni funkcję wspólnototwórczą. Poza tym wojna zawsze była źródłem zysków dla różnych grup społecznych. Czy Facebook chciał wtedy zarabiać na wojnie? Mógł mieć profity z werbalizowanej nienawiści: tak wizerunkowe, jak ekonomiczne. Silne emocje internautów przydałyby się Facebookowi, bo emocje powodują klikalność i dzielenie się postami, a więc przyływ reklamodawców, z czym wiąże się model biznesowy Facebooka (darmowy dostęp do kanału komunikacji, filtrowanie i moderowanie treści algorytmami sprzyjającymi reklamodawcom). O roli emocji pisano:

jeśli użytkownik poczuje się wściekły, czy przestraszony, zwiększy to szanse na jego zaangażowanie, (...) wściekli ludzie częściej udostępniają więcej informacji, tak, by inni podzielili ich wściekłość. A przestraszeni lub wściekli użytkownicy dają się uwieść emocjonalnie i jeszcze bardziej reagują na naładowanie emocjami informacje (McNamee, 2019, s. 94).

Pozwolenie na hejting online mogło przynieść symboliczny efekt *katharsis*, ale mogło być eksperymentem działania algorytmu, o którym mówiła sygnalistka Haugen. Doświadczenie z manipulacją emocjami Facebook miał duże już w 2013 r. dzięki News Feedowi, nie informując użytkowników o zmianie ustawień kont, by wywołać określone zachowania. Swoją decyzją próbował wykreować wrażenie,

że jeśli pozwala na mowę nienawiści, to ma powód do zarządzania emocjami i głowami ludzi. Przyznawał sobie hegemoniczną funkcję w społeczeństwie uzależnionym od technologii i zastosował manipulację, monetyzując wściekłość wobec agresora zostawianą na stronach platformy.

3.3. Facebook na wojnie informacyjnej – hipoteza druga

Facebook miał w zamierzeniach jego pomysłodawcy łączyć ludzi, jednak nie potrafił być neutralny wobec globalnych wydarzeń. Jednym z przykładów zaangażowania w wojnę na Ukrainie jest blokada profili (na prośbę ukraińskiego rządu) rosyjskich mediów rządowych, takich jak Sputnik, Russia Today w Europie, Wielkiej Brytanii i na Ukrainie. Stygmatyzacja propagandowych mediów agresora była dowodem na opowiedzenie się po stronie ofiar. Blokowanie treści było bezprecedensowym zachowaniem organizacji, która nie chce nazywać się mediami, a platformą internetową, więc wojna zmieniła normy komunikacyjne. Dziennikarz BBC nazwał Facebooka „bronią w wojnie informacyjnej” między Federacją Rosyjską a Ukrainą. Tablice Facebooka były wykorzystywane do oglądania wojny na żywo, śledzenia jej przebiegu, wojna hybrydowa działała się na oczach internautów całego świata, jednak strona rosyjska również korzystała z tego kanału, umieszczając propagandowe antyukraińskie treści, które eksperci po jakimś czasie diagnozowali jako materiały dezinformacyjne.

Decyzja Mety o zgodzie na mowę nienawiści podczas wojny mogła oznaczać z jednej strony przyznanie się do niepanowania nad chaosem, który jest na stronach platformy, gdy miliony ludzi reaguje na wojnę. Sama moderacja treści bywa dwuetapowa: robi to najpierw robot, a potem moderator. Algorytmy nie odczytują wszystkich obraźliwych słów lub zdjęć, mają problem z przemówieniami polityków w formacie wideo. Zapanować nad zmasowaną komunikacją nie jest prosto, nikt nie wie, czy w ogóle jest to możliwe. Nieskuteczność kasowania wpisów jest poza tym wzmacniana przez wiralową naturę samego internetu, gdy jedne strony są kasowane, inne powstają. Choć liderzy firmy zapewniali w 2022, że zatrudniają więcej fact-checkerów znających język rosyjski i ukraiński w celu

ograniczenia dezinformacji. Wraca pytanie, czy Facebook wycofał się z decyzji o pozwoleniu na mowę nienawiści dlatego, że miał problem z liczbą moderatorów. Podpowiedzią są medialne newsy z branży IT z 2022 r., mówiące o planach Facebooka co do zwolnienia 11 tysięcy pracowników (w pandemii miał 87 tysięcy). Może to był czytelny przekaz: firma dobrze sobie radzi bez ludzi, bo używa sztucznej inteligencji. Wycofanie się z akceptacji nienawistnych wpisów mogło być odejściem od odpowiedzialności firmy matki, bo wtedy musiałyby reagować krajowe oddziały Facebooka, a nie amerykańska centrala. Można też je tłumaczyć obawą o bezpieczeństwo i wolność internautów, bo według rosyjskiego prawa mogli otrzymać kary więzienia za pisanie prawdy. A może druga decyzja Facebooka była związana z jakimś szantażem rosyjskich trolli i prowokacją, którą trzeba było zdementować? Ktoś mógł wkraść się na konto, opublikować oświadczenie, a firma nie chciała przyznać się do braku zabezpieczeń. O rosyjskiej propagandzie inspirowanej służbami wywiadowczymi zaczęto wówczas mówić bardzo głośno. Amerykanie nagłaśniali wykrycie siatki botów i trolli, uprawiających prorosyjską propagandę. Podobnie polska agencja rządowa, Centrum Przeciwdziałania Dezinformacji, wydała oświadczenie o propagandzie i dezinformacji będących „jedną z najważniejszych składowych wojny hybrydowej”, które zostały zauważone w takich serwisach jak Telegram, YouTube, Facebook i Twitter (Tak Rosjanie...). Być może te absurdalnie brzmiące przypuszczenia nie wykluczają się.

Próbie innej odpowiedzi na zmianę stanowiska Mety może przynieść rama interpretacyjna (*framing*), związana z wojną informacyjną. Roskomnadzor, państwowy regulator komunikacyjny (nazywany w Moskwie watchdogiem), wprowadził karę do 15 lat łagru za nazywanie konfliktu wojną, za relacjonowanie ataków, podawanie informacji nt. ostrzałów ukraińskich miast i relacjonowanie strat rosyjskiej armii. W konsekwencji wiele mediów zawiesiło aktywność, chroniąc swoich pracowników, ponieważ za publikację „kłamliwych informacji” kilka redakcji zostało zablokowanych (radio Echa Moskwy i telewizja Dożd´). W międzyczasie Meta zablokowała dostęp do rosyjskich stron Facebooka, a Roskomnadzor ograniczył szybkość działania Twittera na komputerach, by po kilku dniach go zawiesić. Do tego Rosyjskie Ministerstwo Spraw Zagranicznych oskarżyło Meta o „podżeganie do wojny”, a redakcja Al Jazeera poinformowała na

angielskojęzycznych stronach o nazwaniu Mety organizacją ekstremistyczną i procesie oskarżającym ją o tolerowanie rusofobii. Zdaniem dziennikarza BBC News, bycie na tej liście równało się zamrożeniu pieniędzy na koncie korporacji. Warto odnotować komentarz wiceministra polskiego rządu, pełnomocnika rządu ds. cyberbezpieczeństwa, Janusza Cieszyńskiego, który o zablokowaniu Facebooka i Twittera dosadnie powiedział: „Meta i Twitter przekonały się, jak to jest cnotę stracić, a rubelka nie zarobić. Tak się wychodzi na zginiętych kompromisach z Rosją” (jd/pap, 2022). Podobne antyrosyjskie działania zastosowali inni liderzy na rynku usług technologicznych, jak chociażby Google Alphabet zablokował narzędzia Google Maps, które dostarczając na żywo informacji o warunkach drogowych i lokalizacji różnych miejsc (pap/jd, 2022), tłumacząc to bezpieczeństwem. Z kolei WhatsApp wprowadził opcję dla Ukraińców, w której wiadomość zniknęła po odczytaniu przez odbiorcę.

3.4. Lifting zniszczonej reputacji – hipoteza trzecia

Przyczyn kryzysu wizerunku Facebooka było kilka: skandal z mową nienawiści sztabu Donalda Trumpa (2015), rosyjscy trolle (2016), ujawnienie istoty algorytmów promujących gniew i nienawiść (2019) i Cambridge Analytica (2021). Pierwsza rysa na wizerunku wiązała się z brakiem reakcji Facebooka na mowę nienawiści w bogatym w antymigrancką retorykę przemówieniu Trumpa z 2015 r., umieszczonym w formie nagrania wideo, które szybko zyskało 100 tysięcy polubień. Facebook nie wiedział, jak postąpić z polityką szokowania, podał więc argumentację wolności słowa, według której obywatele powinni mieć dostęp do niecenzurowanych wypowiedzi polityków, aby wyrobić swoje zdanie (Frenkel, 2022, s. 24). Na spotkaniu z personelem Zuckerberg mówił, że Amerykanie powinni sami wyciągnąć wnioski z przemówienia kandydata. Potem w wywiadach powtarzał, że jest za nieograniczoną wolnością słowa i nie chce decydować o tym, co jest prawdą.

Reputację Facebooka zniszczyły w 2016 r. oskarżenia angażowania się w wybory prezydenckie w USA, co nazwano przełomem w pozorowanej neutralności politycznej firmy. Sprzyjanie dezinformacji miało polegać na nieblokowaniu fałszywych kont i niezdejmowaniu reklam

opłacanych przez Rosję. W artykule brytyjskiego „Guardiana” opisano proceder Rosjan, którzy już w 2014 r. siali niezgodę między Ameryką a Europą Zachodnią. Śledztwo dziennikarzy pokazało, że większość protrupowskich materiałów pochodziła z Agencji ds. Badań Internetowych (Internet Research Agency) z siedzibą w Petersburgu, określanej „farmą trolli”. Podczas procesu w Waszyngtonie Zuckerberg oświadczył, że nie wiedział o dezinformacji, jednak przeprosiny za incydent opublikował dopiero po roku. Przyznał się do sprzedaży reklam zagranicznym klientom i ujawnił, że 126 milionów kont było narażonych na ingerencję Rosji. Oficjalnie długo nie mówiono o tym, że firma została zaatakowana przez Rosjan, używając argumentu bezpieczeństwa USA (Levy, 2022, s. 382). Skutek niedopatrzania okazał się druzgocący, spadło morale pracowników, a rewitalizację zaufania firma zaczęła od powołania dyrektora ds. bezpieczeństwa i poszerzenia struktury o działy moderacji treści. Wywołany skandal był powodem odejścia z serwisu ogromnej liczby użytkowników.

Kolejny kryzys wywołała w 2021 r. pracownica Facebooka, Frances Haugen. Wyniosła z pracy tysiące wewnętrznych mejli, udzieliła wielu wywiadów, a jej oskarżenia zostały sprawdzone przez dziennikarzy śledczych. Mówiła o algorytmie, który został przez firmę wymyślony tak, by prowokować ruch na stronie tylko tych informacji, które zawierają prowokacyjne treści. System w konsekwencji izolował użytkowników, zamykając ich w bańkach informacyjnych (filtrujących i preferencyjnych), a bycie częścią tzw. plemienia zamykało na inne treści w myśl zasady: dzielimy się z przyjaciółmi naszymi treściami i czytamy ich treści, bo im ufamy. Haugen mówiła: „Facebook świadomie szkodzi, algorytmy celowo zaogniają podziały społeczne w USA, nie są apolityczne”. W jej opinii platforma stała się „miejscem bardziej gniewnym”, bo algorytmem kształtowała opinię publiczną, stosując manipulację. Lars Backstrom, wiceprezes ds. inżynierii, odpowiedział na zarzuty sygnalistki, tłumacząc, że każdy algorytm ryzykuje promowanie treści, które są szkodliwe dla niektórych użytkowników. Badacze obserwujący manipulacje algorytmów komentowali, że media społecznościowe nie są zainteresowane weryfikacją treści, ale prowokowaniem do interakcji, które przynoszą ruch, bo ten przyciąga reklamodawców (Morales-Sánchez & Martin-Villarreal, 2019, s. 4).

Kolejny skandal dotyczył wycieku danych. Firma Cambridge Analytica zajmująca się analizą danych i konsultacją polityczną wykorzystała dane milionów Amerykanów wydobyte z Facebooka, bo chciała stworzyć psychograficzny profil wyborców i zastosować namierzenie behawioralne, wcześniej opracowane przez inżynierów Facebooka, żeby określonym wyborcom wysyłać określone treści (Galloway, 2021, s. 135). Firma została ukarana przez Federalną Komisję Handlu (FTC) rekordową w USA grzywną 5 mld dolarów.

W przestrzeni publicznej w USA coraz głośniej było słycać głosy wyrażające oczekiwanie wzięcia odpowiedzialności za treści umieszczane na stronach Facebooka. Znaczący pragmatycznych zachowań Zuckerberga podkreślali biznesowe podejście zorientowane na powiększenie liczby użytkowników. Dostrzegali też niekonsekwentne decyzje co do szanowania ich prywatności. Powtarzalną przez lata formą reakcji Zuckerberga na nieuczciwości serwisu były przeprosiny w formie wpisu na oficjalnej stronie platformy. O uruchomieniu aplikacji Beacon (udostępniającej reklamodawcom aktywność zakupową użytkowników bez ich zgody) mówił ogólnikowo: „Nie udało się nam zachować właściwej równowagi” (Frenkel, 2022, s. 78), obiecując zmianę ustawień, przeprosił „za zaskoczenie nowym narzędziem reklamowym”. Roger McNamee komentował:

Facebook ciągle eksperymentuje, coś zmienia i przekracza granice, (...) niektóre rzeczy działają inaczej niż zakładano, co zmusza firmę do przeprosin i kolejnych eksperymentów, (...) do mistrzostwa opracował sztukę spadania na cztery łapy (McNamee, 2019, s. 146).

Zdaniem krytyków Facebooka jego zmiany są prowizoryczne i nie dość, że wymuszone przez rząd lub dziennikarzy, realizowane są z opóźnieniem, bo po 2016 r. Facebook sporządził „listę dawnych błędów, które stara się wyeliminować, ale nie sięgnął korzeni zła, modelu biznesowego i algorytmu” (McNamee, 2019, s. 271). Po ujawnieniu skandalu w wyciekiem danych do Cambridge Analytica Facebook po długim okresie milczenia przeprosił konsumentów, jednak w kuluarach mówiło się, że nie wszystkie dane z 50 mln profili zostały skasowane (Levy, 2022, s. 443). Zuckerberg tłumaczył się niewiedzą o incydencie, potem nadmiernym zaufaniem wobec firmy zbierającej dane. Taki akt skruchy można potraktować jako sposób

na zachowanie dobrego wizerunku i dobrych relacji z otoczeniem (Sobczak, 2020, s. 107).

4. POLITYKA KASOWANIA „NIEWYGODNYCH TREŚCI” NA FACEBOOKU – OD HAKERSTWA DO SZTUCZNEJ INTELIGENCJI

Facebook wielokrotnie zmieniał podejście do mowy nienawiści, podejmując różne działania, w zależności od prawa danego państwa, gdzie wystąpiły naruszenia Standardów społeczności (Frenkel, 2022, s. 26). Brak precyzji w określeniu kryteriów usprawiedliwiał przyczynami kulturowymi. Bardzo długo nie posiadał żadnej polityki reagowania na „szkodliwe treści”. To, w jaki sposób uczył się usuwania wpisów, zostało opisane w książce *Facebook. A miało być tak pięknie*. Pionierska metoda usuwania materiałów miała cechy typowe dla hackerstwa. Pracownicy centrali Facebooka, mając sygnał o niewłaściwym poście, włamywali się na konto osoby zgłaszającej w celu sprawdzenia wiarygodności nadawcy. Potem włamywali się na konto osoby zgłaszanej i ręcznie kasowali wpis. Podstawową zasadą było unikanie pornografii i zdjęć obnażających ciało. Regulamin był niejasny i wiele razy zmieniany. Przy tworzeniu nowych zasad pojawiły się szersze kryteria dotyczące słów lub wyrażen, przedstawiane na wewnętrznych szkoleniach. Dopiero w 2009 r. zatrudniono ludzi do ingerencji w treści użytkowników, tworząc zagraniczny oddział w Dublinie. W roku 2012 firma założyła następne centra moderacji w Manili i w Indiach, zatrudniając przez firmy outsourcingowe tysiące pracowników. Ludzie ci:

stali się odpowiednikami cyfrowych dozorców, czyszczących News Feeda naszym niewidzialni sprzątacze, którym płaci się za oglądanie takich treści jak sceny gwałtów, nielegalne operacje medyczne, ludzkie genitalia (...), a praca rzuca im się na mózg (Levy, 2022, s. 465–466).

Wskutek kontaktów ze wstrząsającymi treściami Facebook od 2016 r. zaczął zapewniać moderatorom pomoc psychoterapeutów i cotygodniowe sesje. Facebook był monitorowany całodobowo, a w danym

momencie nad zamieszczanymi treściami czuwało naraz 400 moderatorów – każdy z nich miał ok. 40 sekund na decyzję, czy dany post lub nagranie należy usunąć. Wyścig polegał na interpretacji zasad z dokumentu nazwanego Standardami społeczności (na początku jednostronicowego, potem rozbudowanego do 27 stron). Monika Bickert, koordynatorka działań firmy w zakresie wolności słowa, przyznała w wywiadzie: „polityka w sprawie nienawiści jest naszym najtwardszym orzechem do zgryzienia, ponieważ brakuje nam kontekstu” (Levy, 2022, s. 469). W 2019 r. powiększono liczbę moderatorów trzykrotnie, do 15 tysięcy, a Bickert tłumaczyła to wcześniejszymi zaniedbaniami (Levy, 2022, s. 464).

Pomysłem Zuckerberga na przesianie mowy nienawiści było korzystanie z pomocy zespołu doradców, którzy pochylając się nad postami, mieli oceniać, co jest obelgą, a co nie. Inny pomysł, tzw. Sąd Najwyższy, składał się z 40 osób, prawników, dziennikarzy, działaczy na rzecz praw człowieka, którzy spotykali się w Singapurze, Berlinie i Nowym Jorku, gdzie 150 ekspertów miało dać rekomendacje co do polityki wobec mowy nienawiści. Paradoksalnie założyciel Facebooka podkreślał często, że nie chce być arbitrem, bo woli zajmować się biznesem, nie moralnością, jednak skargi użytkowników wymuszały wymyślenie systemu reagowania. Zdaniem znawców sama struktura platformy pozwala na praktyki plemienne, sianie niezgody i organizację kampanii nienawiści (McNamee, 2019, s. 246). Kryteria decydowania, co należy usunąć, są i będą płynne i subiektywne. Kto je stworzy? Może robot?

Problem Facebooka z definiowaniem mowy nienawiści ma rozwiązać sztuczna inteligencja (*artificial intelligence*, AI). Od 2017 r. pracuje nad nią francuski naukowiec Yann LeCun w stworzonej przez Zuckerberga Pracowni Sztucznej Inteligencji (FAIR), współpracując z zespołem Stosowanego Uczenia Maszynowego (AML). Celem tych działań jest nie tylko identyfikacja fałszywych kont, ale także reagowanie, szybka identyfikacja treści, zanim ktokolwiek je zobaczy i zgłosi (w Facebooku określano to „działaniem proaktywnym”). Pierwsze próby maszynowego odkrycia treści terrorystycznych okazały się pozytywne, jednak problemem stały się treści w językach obcych. Co do fałszywych kont, od stycznia do marca 2019 r. robot zablokował 2 miliardy prób otwarcia nieprawdziwych kont (Levy, 2022, s. 476). Pod koniec 2020 r. firma opublikowała raport

o usunięciu aż 26,9 mln postów (pojawiających się od października do końca grudnia, w ciągu 3 miesięcy), deklarując, że ciągle usprawnia wykorzystanie sztucznej inteligencji do rozpoznawania i blokowania „szkodliwych treści” (ps, 2021).

5. WYNIKI ANALIZY NAUKOWEJ – DIAGNOZA PYCHY I KANIBALIZMU FACEBOOKA I POSTULAT JEGO KONTROLOWANIA

Warto pamiętać, że firma do 2016 r. nie podlegała żadnej kontroli i nie dokonywała samoregulacji w zakresie monopolu usług. Kongresmeni USA oczekiwali, że start-upy będą naprawiały się same. Komentatorzy podkreślali patologiczne cechy kultury organizacyjnej Doliny Krzemowej, która „przypomina klub towarzyski i nawet dyrektorom trudno występować przeciwko platformom internetowym” (McNamee, 2019, s. 266) i w której powstała popularna idea niekwestionowania założyciela, będącego „niczym król czczony przez inwestorów”. Pisano o specyfice firm złożonych z informatyków z dyplomami prestiżowych uczelni, którzy marzą o pracy przy nadzwyczajnych projektach, pod presją, by z innowacją być pierwszym na rynku. Za kilka lat obserwatorzy działań Facebooka postawią katastroficzną diagnozę: „pycha i wiara w rozwój zapędziła Zuckerberga w przepaść”. Facebooka cechowała pozorna neutralność polityczna, bo w strukturach firmy zatrudniony był człowiek odpowiedzialny za globalną politykę i lobbing w Waszyngtonie. Kaplan, były wiceszef sztabu wyborczego G. Busha, miał być człowiekiem kształtującym dobrą opinię platformy wśród republikanów. Znawcy kuluarów przypominają nieetyczność działań Zuckerberga, który wycofał się z porozumień nt. regulacji związanych z monopolem władzy i złamał obietnice z 2012 r. dane założycielom Instagrama i WhatsAppa o zachowaniu przez nich odrębności (Frenkel, 2022, s. 257–258). Zaczął też ingerować w technologiczne decyzje Instagrama w obawie o „kanibalizację aplikacji Facebooka”. Za kilka lat obserwatorzy napiszą, że ludzie Facebooka padli ofiarą „efektu Midasa”, wierząc w to, że „cokolwiek zrobią, zawsze będą najlepsi i będą służyć ludzkości. Porzucono pokorę” (McNamee, 2019, s. 85).

Rozdźwięk między deklaracjami a realizacjami nazywa się w etyce hipokryzją i to ona przykleiła się do wizerunku koncernu na lata. Nawet w Polsce pojawiły się komentarze ekspertów z portalu Spiders' Web typu: „to, co na początku jest siłą tych firm, później staje się ich słabością. Brak regulacji wpływa na to, że sytuacja wymyka im się z rąk” (Koprz, 2022).

Analiza badawcza odkryła ciemną stronę Facebooka i mechanizmy zarządzania treścią w celu uzyskania silniejszych interakcji, prowadzących do elektronicznych wojen i silnej polaryzacji społeczeństwa. Nie jest tajemnicą, że Facebook do krajów rozwijających się wysyłał pracowników w celu przeszkolenia, jak obsługiwać się platformą. Pomagał Hindusom, Filipińczykom w „łączeniu się ze światem”. Krytycy punktowali go za destrukcyjne działania wykorzystane przez dyktatorów, opisując armie trolli w Indiach (2014), na Filipinach (2016), w Kambodży (2017), w Mjanma (dawnej Birnie) (2018). Zdaniem badaczy „autorytarni przywódcy, religijni fanatycy oraz etniczni nacjonaliści używają Facebooka do siania dezinformacji” (Vaidhyanathan, 2018, s. 312). Działania armii trolli były podobne: zakładano fałszywe konta, produkowano fake newsy i rozpowszechniano je przez prawdziwe i fałszywe konta na Facebooku. Dystrybuowana propaganda dominowała polityczny dyskurs i dyskredytowała przeciwników. Podobnie wojna na Ukrainie w 2022 przyniosła zmasowaną kampanię dezinformacyjną po obu stronach konfliktu, wobec której platformy internetowe stały się bezradne.

Zgoda koncernu na mowę nienawiści w 2022 r. mogła być eksperymentem zarządzania emocjami w wojnie hybrydowej z Rosją. Facebook wrócił do zakazu, by nie doprowadzić do „pożarów nienawiści” (McNamee, 2019, s. 228). Do dzisiaj w Europie i w Polsce jest duże zamieszanie z decyzjami krajowych oddziałów Facebooka, które reagują niekonsekwentnie, jedne posty kasując, bo szkodzą społeczności, zostawiając inne, etycznie dwuznaczne i obraźliwe. Sztuczna inteligencja, chwalona przez Zuckerberga i Muska, nie radzi sobie z diagnozowaniem mowy nienawiści. Mimo zachwyków nad jej potencjałem zarzuca się jej brak ludzkiej empatii i błędy w odczytaniu znaczeń (Frenkel, 2022, s. 236). Czy Facebook jest w stanie wyuczyć AI wyrażen związanych z wojną, polityką i konfliktami społecznymi, jakie dzieją się na całym świecie? Niepokój słychać w tezie eksperta: „narzędzia, które służą do dzielenia się ideami, są

wspaniałe w idealnym świecie, ale przy zaangażowaniu sztucznej inteligencji i automatyzacji stają się podatne na manipulację” (McNamee, 2019, s. 235). Amerykański profesor medioznawstwa napisał, że dzięki Facebookowi „jesteśmy dla siebie okrutniejsi”, podkreślając, że informacje na stronach platformy służą prześladowaniu i upokarzaniu, a „chwalebne zamiary Marka Zuckerberga zamieniły się w koszmar” (Vaidhyanathan, 2018, s. 90–91). Pesymizm widać też w innej diagnozie:

Facebook jest zbyt duży, a różnorodność form ludzkiej nikczemności zbyt wielka, żeby firma mogła zatrudnić odpowiednio dużo ludzi lub wykorzystać kod komputerowy do przewidywania i korygowania zachowań milionów użytkowników, działając na skalę globalną (Vaidhyanathan, 2018, s. 326).

Nie napawa optymizmem jednorazowa aktywność prokuratorska koncernu Meta w 2022 r., który udawał stróża porządku publicznego, ujawniając fałszywe siatki kont kreowanych przez boty szerzących prorosyjską dezinformację (Siwik, 2022).

6. WNIOSKI – GASZENIE „POŻARÓW NIENAWIŚCI” W ELEKTRONICZNYM ŚWIECIE

Od kilku lat dyskutuje się temat przeciwdziałania mowie nienawiści, tak w środowisku technologicznym, biznesowym (hejting dotyka wielu przedsiębiorców), jak i politycznym, państwowym (mieliśmy w polityce kilka afer z farmami trolli). Wiele głosów mówi o moderacjach w cyberprzestrzeni i blokadzie obraźliwych treści (zaangażowaniu platform), inne wspominają o penalizacji hejterów (przepisach prawnych dotyczących zarówno sprawców, jak i właścicieli platform typu Facebook).

Wyzwaniem jest nałożenie odpowiedzialności na właścicieli platform, która idzie w poprzek ich biznesowych zysków. Presja organizacji pozarządowych (NGOs) w USA przeniosła na europejski grunt debatę o sankcjach za nieprzestrzeganie zasad dotyczących pracy algorytmów, bezpieczeństwa danych internautów i ich poczucia bezpieczeństwa. Unia Europejska wprowadziła system wysokich kar

za brak reakcji firm technologicznych działających na europejskim rynku. Obserwatorzy twierdzą, że firmy zajmą się bezpieczeństwem głównie z powodów unikania kar finansowych.

Problematyczna jest uznaniowość w rozumieniu przepisów prawnych związanych z definiowaniem mowy nienawiści. W Polsce obowiązują przepisy Kodeksu karnego, sankcjonujące przemoc słowną, zapisaną w art. 119, 256, 257, jak i przepisy dotyczące znieważenia głowy państwa (art. 135, par. 2) i znieważenia głowy obcego państwa (art. 136, par. 3). Przesłębstwa mające charakter nawoływania do nienawiści i dyskryminacji są ścigane w trybie publicznoskargowym. Dochodzi do tego społeczna krytyka decyzji prokuratorów decydujących o tym, czy casus obrażania, zniesławiania lub znieważania jest przestępstwem. Powstają obecnie pierwsze prace w polskim Sejmie dotyczące ustawy o mowie nienawiści, w innych krajach Europy zjawisko ma bardziej dynamiczny charakter. Powtarzającym się postulatem jest edukacja społeczeństwa nt. odpowiedzialności za słowa w przestrzeni online, gdzie anonimowość i bezkarność mają złudny charakter. Jednak oprócz licznych kampanii społecznych wyzwaniem jest uporządkowanie pojęć prawnych i konsekwencja w ich stosowaniu, a także współpraca wielu podmiotów publicznych w ochronie przed agresją słowną i dezinformacją.

Nie ma jednej skutecznej metody na eliminację hejtu z internetu tworzonego przez farmy trolli lub boty, tworzących siatki zasypujące internet kłamstwami i teoriami spiskowymi. W opinii krytyków nadzieją stać się może samoświadomość użytkowników nt. maszynowo napędzanych treści lub stworzenie ruchu obywatelskiego wymuszającego zmiany platform w myśl rady: „trzeba opuścić cyfrowe kokony i zacząć współdziałać” (Levy, 2022, s. 260). Grupa 150 naukowców z USA opłacanych przez fundację Chan Zuckerberg Initiative wysłała do Marka Zuckerberga list, krytykując firmę za „rozpowszechnianie dezinformacji i podziałów”, co pokazuje siłę społeczną obywateli (Dailyweb, 2020). Potrzebne jest również uporządkowanie przepisów prawa dotyczących mowy nienawiści online, zarówno w Polsce, jak i w Unii Europejskiej. Media społecznościowe przestały odgrywać rolę sprzed 20 lat, gdy powstawały, zaczynają zaś być partnerami w konfliktach o charakterze globalnym. Z uwagi na dynamiczność i skomplikowanie problemu zarządzanie mową nienawiści wymaga

pogłębionej refleksji wielu dyscyplin nauki oraz współpracy i zaangażowania wielu aktorów społecznych.

BIBLIOGRAFIA

- Aro, J. (2020). *Trolle Putina*. Kraków: Wydawnictwo SQN.
- BBC News (2022, 11 października). *Russia confirms meta's designation as extremist*. <https://www.bbc.com/news/technology-63218095> (dostęp: 03.11.2022).
- Bilewicz, M., & Lis, D. (2019). Więcej niż polityka. *Znak*, 4(27).
- Cyberdefence24 (2022, 11 marca). *Facebook pozwoli na posty wzywające do przemocy wobec rosyjskich żołnierzy oraz z życzeniami śmierci dla Putina i Łukaszenki*. <https://cyberdefence24.pl/social-media/facebook-pozwoli-na-posty-wzywajace-do-przemocy-wobec-rosyjskich-zolnierzy-oraz-z-zyczeniami-smierci-dla-putina-i-lukaszenki> (dostęp: 13.11.2022).
- Dailyweb (2020, 8 czerwca). *To już agonia Facebooka? Eksperci opłacani przez Zuckerberga krytykują serwis*. <https://dailyweb.pl/to-juz-agonia-facebook-a-eksperci-oplacani-przez-zuckerberga-krytykuja-serwis/> (dostęp: 13.11.2022).
- Do Rzeczy (2022, 11 marca). *Wojna na Ukrainie. Facebook robi wyjątek od zasad dotyczących brutalnych treści*. <https://dorzeczy.pl/kraj/274366/wojna-facebook-robi-wyjatek-dotyczacy-brutalnych-tresci.html> (dostęp: 04.11.2022).
- Dynkowska, J. (Red.). (2017). *Hejterstwo. Nowa praktyka kulturowa? Geneza, przypadki, diagnozy*. Łódź: Wydawnictwo Uniwersytetu Łódzkiego
- Flores, S. (2014). *Sfejsowani. Jak media społecznościowe wpływają na nasze życie, emocje i relacje z innymi*. Warszawa: MUZA S.A.
- Frenkel, S., & Kang, C. (2022). *Brzydka prawda. Kulisy walki Facebooka o dominację*. Poznań: Dom Wydawniczy REBIS.
- Galloway, S. (2021). *Wielka czwórka. Four. Ukryte DNA: Amazon, Apple, Facebook i Google*. Poznań: Dom Wydawniczy REBIS.
- Goczał, B (2021, 19 września). *Algorytm Facebooka nagradza nienawiść. Wirtualne Media*. <https://www.wirtualnemedial.pl/arttykul/algorytm-facebook-nagradza-nienawisc-mark-zuckerberg-nie-chce-zmian-boi-sie-odwrotu-uzytkownikow> (dostęp: 10.10.2022).
- Jaroszewski, D. (2022, 11 marca). *Facebook: nie życz nikomu śmierci, chyba że Putinowi*. Telepolis. <https://www.telepolis.pl/wiadomosci/wydarzenia/facebook-mowa-nienawisci-putin-rosja> (dostęp: 13.11.2022).

- jd/pap (2022, 5 marca). *Facebook i Twitter zablokowane w Rosji*. Wirtualne Media. <https://www.wirtualnemedi.pl/artykul/rosja-blokada-facebook-twitter> (dostęp: 03.11.2022).
- Juza, M. (2015). Hejterstwo w komunikacji internetowej: charakterystyka zjawiska, przyczyny i sposoby przeciwdziałania. *Profilaktyka Społeczna i Resocjalizacja*, 25, 27–50.
- Kamińska, M. (2008). Flaming i trolling – kulturotwórcza forma konfliktu we wspólnocie wirtualnej. W M. Wawrzak-Chodaczek (Red.), *Komunikacja społeczna w świecie wirtualnym* (s. 280–296). Toruń: Wydawnictwo Adam Marszałek.
- Kamińska, M. (2011). *Niecie memy. Dwanaście wykładów o kulturze internetu*. Poznań: Galeria Miejska Arsenał.
- Kawka, M. (2011). Dyskurs medialny w świetle współczesnych analiz dyskursu. *Studia Logopaedica*, IV, 212–220.
- Koprz (2022, 11 listopada). *Mark Zuckerberg w tarapatach. Ekspert: Zawiodł jako wizjoner, powinien się zwolnić*. TOK FM. <https://www.tokfm.pl/Tokfm/7,103086,29126249,marc-zuckerberg-w-tarapatach-ekspert-zawiodl-jako-wizjoner.html> (dostęp: 13.11.2022).
- Leszczuk-Fiedziukiewicz, A. (2018). Czy hejt i mowa nienawiści staną się normą? Społeczne uwarunkowania zachowań dewiacyjnych w Internecie. *Media – Kultura – Komunikacja Społeczna*, 14(3), 99–117.
- Leszczuk-Fiedziukiewicz, A. (2020). Inicjatywy antyjejtowe. Wybrane przykłady zaangażowania społecznego w ograniczanie mowy nienawiści. W E. Horyń, B. Skowronek, A. Walecka-Rynduch (Red.), *Język a media. Perspektywy i zagrożenia we współczesnych mediach*. Kraków: Collegium Columbinum.
- Levy, S. (2022). *Facebook. A miało być tak pięknie. Za kulisami imperium z Doliny Krzemowej*. Białystok: MOVA.
- McNamee, R. (2019). *Nabici w Facebooka. Przestroga przed katastrofą*. Poznań: Media Rodzina.
- Milmo, D. (2022, 14 marca). *Facebook and Instagram users not allowed to call for death of Putin*. The Guardian. <https://www.theguardian.com/technology/2022/mar/14/facebook-instagram-death-putin-meta> (dostęp: 13.11.2022).
- Morales-Sánchez, I., & Martin-Villarreal, J.P. (2019). Double-click Rhetoric: Rhetorical Strategies of Communication in the Digital Context. *Res Rhetorica*, 6(1). DOI: 10.29107/rr2019.1.1.
- Naruszewicz-Duchlińska, A. (2015). *Nienawiść w czasach internetu*. Gdynia: Novae Res.
- Nijakowski, L. (2008). Mowa nienawiści w świetle teorii dyskursu. W A. Horolets, *Analiza dyskursu w socjologii i dla socjologii*. Toruń: Wydawnictwo Adam Marszałek.

- Ogonowska, A. (2014). *Uzależnienia medialne, czyli o patologicznym wykorzystaniu mediów*. Kraków: Wydawnictwo Edukacyjne.
- Osowiecka, M. (2015). *Nienawistny język w internecie*, „Ja. My. Oni. Poradnik Psychologiczny Polityki”. T. 18: *Psychologia i polityka*, 126–127.
- pap/jd (2022, 27 lutego). *Google tymczasowo wyłącza dane o ruchu na żywo w mapach Google w Ukrainie*. Wirtualne Media. <https://www.wirtualnemedi.pl/artykul/google-mapy-ograniczone-w-ukrainie> (dostęp: 03.11.2022).
- ps (2021, 15 lutego). *Facebook na wojnie z hejtem. Prawie 27 mln usuniętych postów*. Wirtualne Media. <https://www.wirtualnemedi.pl/artykul/facebook-hejt-27-mln-usunietych-postow-dlaczego> (02.12.2024).
- Pyżalski, J. (2012). *Agresja elektroniczna i cyberbullying jako nowe ryzykowne zachowania młodzieży*. Kraków: Oficyna Wydawnicza Impuls.
- Siwik, K. (2022, 28 września). *Meta zdemaskowała Rosję. Największa akcja dezinformacyjna od początku wojny*. Dobre Programy. <https://www.dobreprogramy.pl/meta-zdemaskowala-rosje-najwieksza-akcja-dezinformacyjna-od-poczatku-wojny,6817280331340608a> (dostęp: 02.12.2024).
- Sobczak, B. (2020). *Publiczne przeprosiny – akt skruchy czy chwyt perswazyjny? Res Rhetorica. Retoryka (nie)grzeczności*, 7(3). DOI: 10.29107/rr2020.3.7.
- Stawiany J. (2024, 11 lutego). *20 lat Facebooka. „To już nie jest medium społecznościowe, ale informacyjno-reklamowe”*. Wirtualne Media. <https://www.wirtualnemedi.pl/artykul/facebook-kiedy-powstal-krytyka-metawersum-polaryzacja> (dostęp: 02.12.2024).
- Vaidhyanathan, S. (2018). *Antisocial Media. Jak Facebook oddala nas od siebie i zagraża demokracji*. Warszawa: Grupa Wydawnicza Foksal.
- Vengattil, M., Culliford, E. (2022, 11 marca). *Facebook allows war post urging violence against russian invaders*. Reuters. <https://www.reuters.com/world/europe/exclusive-facebook-instagram-temporarily-allow-w-calls-violence-against-russians-2022-03-10/> (dostęp: 03.11.2022).
- War in Ukraine: Facebook to allow calls for violence against Putin*. BBC News. 11.03.2022. <https://www.bbc.com/news/technology-60702619> (dostęp: 03.11.2022).
- Wasiuta, O., & Wasiuta, S. (2017). *Wojna hybrydowa Rosji przeciwko Ukrainie*. Kraków: Arcana.
- zew, qm (2022, 4 marca). *Facebook. Twitter: Władze Rosji blokują dostęp do mediów społecznościowych*. Rzeczpospolita. <https://www.rp.pl/konflikty-zbrojne/art35809981-facebook-twitter-youtube-instagram-wladze-rosji-blokuja-dostep-do-mediow-spolesnosciovych-i-ograniczaja-dzialanie-vpn> (dostęp: 13.11.2022).

Żardecka-Nowak, M. (2016). Resentymnt jako postać nienawiści. W W. Zuziak, & J. Synowiec (Red.), *Nienawiść w życiu publicznym* (s. 57–68). Kraków: Wydawnictwo Naukowe Uniwersytetu Papieskiego Jana Pawła II.

Żmuda, B. (2016). Czy może istnieć nienawiść usprawiedliwiona? W W. Zuziak, & J. Synowiec (Red.), *Nienawiść w życiu publicznym* (s. 41–53). Kraków: Wydawnictwo Naukowe Uniwersytetu Papieskiego Jana Pawła II.

Copyright and License



This article is published under the terms of the Creative Commons Attribution – NoDerivs (CC BY- ND 4.0) License
<http://creativecommons.org/licenses/by-nd/4.0/>