

DOI: 10.20535/2522-1078.2022.2(12).270922

УДК 32.328

Надходження до редакції: 25.10.2022

Прийняття до друку: 10.11.2022

Фіялка С. Б.

*канд. наук із соц. комунікацій, доц. кафедри
видавничої справи та редагування
Навчально-наукового видавничо-
поліграфічного інституту, КПІ ім. Ігоря
Сікорського, м. Київ, Україна, fiyalka@i.ua*

ORCID: 0000-0002-1855-7574

Fiialka S.

*PhD of Social Communications, Associate
Professor at the Department of Publishing
and Editing, Publishing and Printing Institute
of Igor Sikorsky Kyiv Polytechnic Institute,
Kyiv, Ukraine, fiyalka@i.ua*

ORCID: 0000-0002-1855-7574

**ПРОБЛЕМАТИКА ДОТРИМАННЯ СТАНДАРТІВ СПІЛЬНОТИ
В СОЦМЕРЕЖІ «ФЕЙСБУК» В УМОВАХ ЗБРОЙНОЇ АГРЕСІЇ
РОСІЙСЬКОЇ ФЕДЕРАЦІЇ****STANDARDS OF THE SOCIAL NETWORK «FACEBOOK»
IN THE CONDITIONS OF ARMED AGGRESSION OF THE RUSSIAN
FEDERATION**

З початком повномасштабної неспровокованої військової агресії Російської Федерації алгоритми та модератори соцмережі «Фейсбук» стали посилено обмежувати проукраїнський контент, посиляючись на стандарти спільноти, в яких декларується захист споживачів мережевого контенту від насильства та злочинної поведінки, неприйняттого змісту, гарантується захист їхньої безпеки та психологічної рівноваги тощо. Мета статті — з'ясувати типові формальні підстави видалення проукраїнського контенту алгоритмами й модераторами мережі «Фейсбук», виокремити стандарти спільноти, які потребують перегляду в умовах війни. Методика дослідження охоплює аналіз попереджень щодо блокування контенту та критичний аналіз Принципів, умов обслуговування та стандартів спільноти «Фейсбук». У повідомленнях користувачам щодо блокування контенту йдеться насамперед про порушення стандартів, прописаних у рубриках «Насильство та підбурювання», «Небезпечні особи та організації», «Мова ворожнечі», «Насильство та чутливий зміст», «Дезінформація». Дедалі частіше алгоритми маркують контент українських користувачів як «чутливий» (наприклад, «бучанська рука», «ізіумська рука», фото тіл цивільних та (або) військових), хоча саме оприлюднення таких матеріалів допомагає донести наслідки військових злочинів рашистського режиму, підвищити обізнаність світової спільноти про

них, сприяючи наданню Україні міжнародної фінансової, гуманітарної, правової та військової допомоги. Поширеною стала практика боротьби з так званим межовим вмістом, коли матеріали не показуються у стрічках новин, але не блокуються. За стандартами спільноти всі користувачі мережі однаково захищені, однак при цьому фактично не враховується історичний контекст та поточні події, які можуть знімати табу на висвітлення контенту. Щоб приймати справедливі рішення, потрібно залучати неупереджених рецензентів з відповідним культурним і політичним досвідом, розумінням контексту. Також вважаємо допустимим частково переглянути вимоги стандартів щодо заборони «мови ворожнечі», яка стає засобом емоційної розрядки всередині спільноти, котра є жертвою масштабної неспровокованої військової агресії Росії.

Ключові слова: стандарти спільноти, контент, модератор, мова ворожнечі, чутливий вміст, блокування, дезінформація, межовий зміст.

With the beginning of the full-scale military aggression of the Russian Federation, the algorithms and moderators of the Facebook social network began to strongly limit pro-Ukrainian content, referring to community standards, which declare the protection of content consumers from violence and criminal behavior, unacceptable content, guarantee the protection of their safety, etc. The purpose of the article is to find out the formal grounds for removing pro-Ukrainian content by algorithms and moderators of the Facebook network and their consistency with Community standards. The research methodology includes an analysis of Facebook's content blocking warnings and a critical analysis of Facebook's Principles, Terms of Service and Community Standards. Notifications regarding blocking of content refer primarily to violations of the standards prescribed in the headings «Violence and incitement», «Dangerous persons and organizations», «Hate speech», «Violence and sensitive content», «Disinformation». More and more often, the algorithms of the social network mark the content of Ukrainian users as «sensitive» (for example, «Bucha's hand», «Izium's hand», photos of the bodies of civilians and (or) soldiers), although it is precisely such materials that have the potential to convey the consequences of the war crimes of the Russian regime, to raise awareness about them, to promote the provision of international humanitarian and military aid to Ukraine. The practice of combating the so-called border content, when materials are not shown in news feeds, but are not blocked, has also become common. According to community standards, all network users are equally protected, but this does not actually consider the historical context and current events that can lift the taboo on content coverage. To make fair decisions, it is necessary to involve impartial reviewers with relevant cultural and political experience, understanding of the context. We also consider it permissible to relax the requirements of the standards for the use of «hate speech», which in

the conditions of war becomes a means of emotional reliving within the community, which is a victim of unprovoked aggression.

Keywords: community standards, content, moderator, hate speech, sensitive content, blocking, misinformation, boundary content.

Постановка проблеми. Процеси модерації контенту соцмереж традиційно були дуже непрозорими [16]. Зокрема, стандарти спільноти «Фейсбук», які окреслювали, який контент є неприпустимим, були поверховими та нечіткими: хоча в них зазначалося, що «контент тероризму» та «мова ворожнечі» заборонені, користувачі ще донедавна не знали, що конкретно належить до цих категорій [12]. Унаслідок занепокоєння громадськості тим, що «Фейсбук» недостатньо підзвітний своїм користувачам та іншим зацікавленим сторонам у країнах і спільнотах, у яких та серед яких він працює, мережа опинилася в епіцентрі розмов про цифрові медіа, технологічну політику та демократію [8]. До того ж дедалі популярнішою ставала думка, що найефективнішим засобом захисту прав людини в мережі можуть бути не окремі країни чи приватні корпорації, і навіть не ООН, а самі користувачі [10].

Очевидно, що модерування було непростим завданням, «особливо коли актори використовували контент-стратегії, які зміщували вікно Овертона (тобто діапазон ідей, прийнятних у публічному дискурсі) праворуч» [5]. Тож у 2020 р. Актом про цифрові сервіси ЄС було запропоновано запровадити базові вимоги щодо діяльності онлайн-платформ, видалення контенту, подання скарг і апеляцій тощо [14].

Фейсбук створив свій Центр прозорості в травні 2021 р. Він також регулярно оновлював свої стандарти спільноти, а наглядова рада Фейсбуку переглядає ці стандарти на основі конкретних кейсів.

Будучи дуже потужним інструментом для формування громадської думки [6], з початком військової агресії Російської Федерації алгоритми соцмережі почали посилено обмежувати контент про російсько-українську війну, блокувати акаунти громадських активістів, що розголошували воєнні злочини російської армії на теренах України [2]. Тобто платформа не стала нейтральним посередником у масовій комунікації і почала дедалі частіше ухвалювати упереджені рішення [13; 17]. Інакше кажучи, практика засвідчила, що управління контентом у мережі «Фейсбук» виводить його з площини стандартів етики та справедливості у площину упередженості.

Блокування патріотичного контенту викликає тривогу в українському сегменті соцмережі, бо при цьому російська пропаганда просуває на західну аудиторію наративи про «екстремізм» та «український фашизм», використовуючи Фейсбук як «поле пропагандистської битви» [9].

Мета статті — з'ясувати типові формальні підстави видалення проукраїнського контенту алгоритмами й модераторами мережі «Фейсбук», виокремити стандарти спільноти, які потребують перегляду в умовах війни.

Методи. Методика дослідження охоплює пошуковий метод, а також аналіз попереджень Фейсбуку щодо блокування контенту та критичний аналіз Принципів, умов обслуговування та стандартів спільноти «Фейсбук» — основних правил платформи, які визначають, як фільтрувати контент і боротися з проявами екстремізму, мови ворожнечі, іншим неприйнятним, відповідно до цих стандартів, контентом. Аналіз цих матеріалів зводиться до відповіді на питання, чому окремі комунікаційні практики користувачів українського сегмента соцмережі «Фейсбук», які формально порушують правила спільноти, можна вважати прийнятними і навіть необхідними в умовах війни.

Результати дослідження. У суспільстві спостерігається соціальний феномен: люди починають масово переміщатися для спілкування в соцмережі саме з міркувань психологічної (особистісно-емоційної) безпеки. Користувачам ненав'язливо пропонується різноманітна інформація на будь-який смак, котру вони охоче сприймають як компенсацію індивідуальних можливостей, яких бракує в реальному житті [1].

Відповідно до стандартів Фейсбуку, користувачі можуть скаржитися на контент, який потенційно порушує правила спільноти, у тому числі на сторінки, групи, профілі, поодинокі дописи й навіть коментарі. Ці стандарти базуються на відгуках користувачів соцмережі та порадах експертів у таких галузях, як технології, громадська безпека, права людини тощо. При цьому зауважується, що в деяких випадках до оприлюднення допускається контент, який потенційно підлягає забороні, проте становить значний суспільний інтерес.

У стандартах декларується захист споживачів інформації від насильства та злочинної поведінки; гарантується захист від неприйнятно-го змісту та забезпечення автентичності й приватності.

У контексті нашого дослідження актуальними є такі підрубрики стандарту, як «Насильство та підбурювання», «Небезпечні особи та організації» (рубрика «Насильство та злочинна поведінка»), «Мова

ворожнечі», «Насильство та чутливий зміст» (рубрика «Неприйнятний зміст»), «Дезінформація» (рубрика «Цілісність та автентичність»). Фейсбук видаляє формулювання, які підбурюють до насильства, а також «не дозволяє організаціям або окремим особам, які проголошують насильницьку місію або залучені до насильства, бути присутніми в соцмережі». Як «мову ворожнечі» Фейсбук визначає насильницьку або дегуманізовану мову, шкідливі стереотипи, заяви про неповноцінність, вираження презирства, огиди чи зневаги, лайку та заклики до виключення чи сегрегації. Адже вважається, що масове розповсюдження «мови ворожнечі» викликає в пересічних користувачів усвідомлення того, що їхні екстремальні погляди рішуче підтримує широка громадськість [11].

Щоб захистити користувачів від травматичних зображень, Фейсбук вилучає вміст особливо жорстокого або відразливого характеру, як-от відео із зображенням розчленування, видимих нутроців або обгорілих тіл, а також вміст, який містить садистські коментарі щодо зображень, що ілюструють страждання людей і тварин. У контексті обговорення важливих питань, таких як порушення прав людини, збройні конфлікти чи акти тероризму, соцмережа, однак, дозволяє окремі специфічні графічні матеріали, «щоб допомогти людям засудити ці ситуації та підвищити обізнаність про них», та додає попередження до деяких графічних чи насильницьких зображень. При цьому обмежується можливість переглядати такий вміст для користувачів віком до 18 років.

Щодо дезінформації, то Фейсбук формулює різні категорії дезінформації та надає вказівки стосовно того, як до неї ставитися. Вилучається дезінформація, яка може безпосередньо спричинити ризик «неминучої фізичної шкоди». У стандартах стверджується, що Фейсбук співпрацює з незалежними експертами, які володіють знаннями та досвідом, щоб оцінити правдивість вмісту та те, чи може він безпосередньо спричинити заподіяння шкоди. Таким чином мережа уповільнює поширення містифікацій і вірусної дезінформації. Сторінки, групи, профілі, які неодноразово розповсюджували дезінформацію, можуть отримати обмеження на поширення контенту, зокрема й рекламного, або бути видалені з платформи [7].

Серед покарань за порушення стандартів спільноти попередження за перший інцидент або тимчасове чи постійне блокування користувача в разі повторних порушень. Якщо користувачі вважають, що видалена інформація не порушує стандартів спільноти, у більшості випадків вони

можуть подати запит на перегляд контенту. При цьому однієї скарги достатньо для блокування невідповідного контенту. Однак якщо він не порушує стандартів спільноти, то його не буде видалено навіть попри велику кількість скарг.

Однією з практик Фейсбуку стало вилучення межового змісту, який не підлягає блокуванню, але може «погіршити якість послуг». За словами М. Цукерберга, внутрішнє дослідження Фейсбуку показує, що межовий вміст привертає більше уваги як людей, яким подобається вміст, так і тих, кому він не подобається. Тому Фейсбук навчив свої системи штучного інтелекту виявляти цей вміст і менше поширювати його — варіант практики, відомої як тіньова заборона (не показувати певний вміст у стрічках новин, але не вилучати його) [18].

В усьому світі понад 100 000 людей працюють модераторами контенту, переважно аутсорсингові працівники з нестабільними контрактами. Незважаючи на перехід до алгоритмічної модерації, люди-модератори все ще є важливим елементом у регулюванні вмісту соцмережі [15].

Українських користувачів у соцмережах з 2014 р. блокують уже не вперше. Наприклад, ще у 2015 р. Фейсбук заблокував відомого українського блогера З. Антипопа за публікацію цитати з вірша Т. Шевченка: «Кохайтесь, чорноброві, та не з москалями» [3]. За слова «москалі», «русня», «кацапи» «рашисти», «свинособаки», «москалі» і навіть «guzznu» тощо українці і зараз можуть потрапити в бан. Фейсбук фіксує це як вияв ненависті до конкретної національності, іноді попри контекст повідомлення.

У липні 2021 р. було заблоковано допис посла України в Чехії Є. Перебийноса з критикою слів президента Росії В. Путіна про «один народ» за нібито «мову ворожнечі»: «Який один народ? Які брати? Судячи з того, як цілеспрямовано і холоднокривно вони нас знищували протягом кількох століть, то ми навіть не родичі. Родичів нормальні люди не вбивають» (коментар Фейсбуку — «Більше ніхто не зможе бачити ваш допис. Ми запровадили ці стандарти, щоб захистити певні групи людей, яких називають неповноцінними порівняно з іншими») (<https://www.facebook.com/yevhen.perebyinis/posts/1043542289515318>). Однак згодом пост повернули.

Подібним чином нібито за пропаганду ненависті та образ було заблоковано відео, на якому члени Львівської академічної чоловічої хорової капели «Дударик» виконували пісню «Батько наш Бандера» (коментар — «Ваш допис порушує наші стандарти спільноти щодо

пропаганди ненависті й образ») (<https://www.facebook.com/Dudaryk.choir/posts/3117448941824355>).

У березні 2022 р. Фейсбук на знак підтримки українців тимчасово вирішив не карати користувачів за висловлення закликів до насильства проти солдатів російських окупаційних військ. Однак уже незабаром відновилося блокування дописів про воєнні злочини Росії [4]. Наприклад, було заблоковано акаунт блогера П. Нека за публікацію фотографій загиблих з Бучі та фотосвідчень систематичних прицільних обстрілів армією Російської Федерації цивільних споруд.

У травні 2022 р. соцмережа Facebook заблокувала допис української письменниці М. Матіос про заступника командира полку «Азов» С. Паламаря, заявивши про «порушення стандартів спільноти щодо небезпечних осіб та організацій». У дописі письменниця розповідала, що дізналася про літературні смаки військового і побачила своє ім'я в переліку авторів, чії твори йому подобаються: «Дай, Боже, всім письменникам таких читачів. Коли з одного інтерв'ю я дізналася, що він читає мої книжки, я заплакала... з вдячності. Бо думати про те, що він — молодий, повний життя і пристрасті — уже належить до когорти сучасних Надлюдей, і одночасно уявляти, що відчуває його мама, ловлячи найменшу згадку про сина і його побратимів, — мені дуже складно. Він — це друг Калина, Святослав Паламар, заступник командира полку АЗОВ».

Письменниця висловила сподівання на зустріч зі С. Паламарем після порятунку захисників Маріуполя з заводу «Азовсталь». Коментар Фейсбуку мав такий вигляд: «Ваша публікація суперечить стандартам нашої спільноти щодо небезпечних осіб та організацій. Ми маємо ці стандарти для запобігання та уникнення офлайн-шкоди. Ми підтвердили, що ваша публікація не відповідає стандартам спільноти. Ми передивилися ваш допис знову, і він не відповідає стандартам нашої спільноти» (<https://www.facebook.com/maria.matios>).

У червні 2022 р. під заборону потрапив пост зі словами «Кременчук. Торговий центр «Амстор». Рашистська ракета. Як же хочеться, щоб ви за все відповіли, тварюки російські» (коментар Фейсбуку — «Більше ніхто не зможе бачити ваш допис. Ми запровадили ці стандарти, щоб захистити певні групи людей, яких називають неповноцінними порівняно з іншими») (<https://www.facebook.com/profile.php?id=100005478302156>)

У вересні 2022 р. було заблоковано пост за фото фрагмента ексгумованого в Ізюмі тіла (коментар Фейсбуку — «Потенційно неприйнятний

контент; це фото може містити сцени насильства чи важкий для сприймання контент»), ілюстрації О. Грекова, який зобразив руку із синьо-жовтим браслетом (коментар Фейсбуку — «Це фото не порушує норми спільноти, але може містити тяжкий для сприймання контент») (<https://www.facebook.com/profile.php?id=100000973248746>), а також допис у групі «Влада Вільних Громадян» про те, що українські вояки відрізняються поведінкою від окупантів, з відео, де військові рятують тварин (коментар Фейсбуку — «Допис містить мову ненависті») (<https://www.facebook.com/profile.php?id=100052975185567>).

У жовтні за пропаганду ненависті було заблоковано допис зі словами «Фейсбук у мене запитує, що у мене на думці? Я вам так скажу. Одним махом, одним ударом знищити всю русню. Бов вони всі в цьому терорі винні» (коментар Фейсбуку — «Ваш допис порушує наші стандарти спільноти щодо пропаганди ненависті. Ми запровадили ці стандарти, щоб обговорення на Фейсбуку велося ввічливо») (<https://www.facebook.com/profile.php?id=100004946356211>) та пост зі словами «Ранок починається не з кави, а з пекучої ненависті до русні» (коментар — Фейсбуку — «Ваш допис порушує наші стандарти спільноти щодо пропаганди ненависті й образ, які натякають на неповноцінність») (<https://www.facebook.com/helen.bohdan.3>).

Цікаво, що тільки-но платформа починає видаляти певний контент, люди починають поширювати його ще більше. Так сталося, наприклад, із цитатою загиблого активіста Р. Ратушного, що стала для українців своєрідним закликком: «Чим більше росіян ми вб'ємо зараз, тим менше росіян доведеться вбивати нашим дітям». Після загибелі героя соцмережах користувачі почали масово публікувати скріншоти із цією цитатою.

Коли алгоритми соцмереж маркують контент українських користувачів як «чутливий», доносити правду світові допомагають ілюстратори, які на своїх малюнках, гіфках, відео тощо рефлексують про все, що відбувається. Серед них О. Греков, А. Данкович, Ю. Журавель, М. Мошковський, А. Немченко, М. Паленко, Н. Тітов та десятки інших. У контексті геополітичної ситуації такий контент стає дієвим способом інформаційної боротьби з російськими загарбниками

Висновки. Фейсбук має власні алгоритми, які автоматично блокують «неприйнятний» контент без участі модератора. Наприклад, це можуть бути невідредаговані фото тіл цивільних або військових. Однак нещодавні блокування можна вважати також цілеспрямованою атакою, до

якої причетні російські інформаційні війська. Рішення урядів західних країн щодо надання допомоги Україні залежить від того, які настрої є в населення цих країн, тому, наприклад, поширення контенту, який розкриває військові злочини Росії, може сприяти наданню такої допомоги, і навпаки. Алгоритми — це теж результат роботи людей, які можуть закладати в них свої упередження та симпатії. Тим більше, що офіційно в стандартах спільноти немає прямої заборони на поширення інформації, відео чи фото злочинів, зокрема й воєнних.

За стандартами спільноти всі користувачі мережі однаково захищені, однак необхідно також враховується історичний контекст та поточні події, які можуть знімати табу на висвітлення чутливого контенту. Для цього потрібні рецензенти з відповідним культурним і політичним досвідом, розумінням контексту та усвідомленням, що спільнота соцмережі, це, з одного боку, одна глобальна спільнота, а з другого — набір різних спільнот, з різним досвідом взаємодії з іншими спільнотами.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Кількість дезінформації в онлайн-медіа зросла майже втричі — Інститут масової інформації. URL: <https://litgazeta.com.ua/news/kilkistdezinformatsii-v-onlajn-media-zroslo-majzhe-vtrychi-instytut-masovoi-informatsii/> (дата звернення: 21.01.2021).
2. Потенційно неприйнятний: як онлайн-платформи обмежують контент українців. URL: <https://ukrainer.net/pravo-na-kontent/> (дата звернення: 20.10.2022).
3. Фейсбук блокував відомого українського блогера Зена Антипопа. Уряд. кур'єр. 2015. 21 січ. URL: <http://ukurier.gov.ua/uk/articles/facebook-ta-youtube-mayut-zminiti-pravila-moderuva/> (дата звернення: 20.10.2022).
4. Чому Фейсбук та Інстаграм блокують дописи про війну від українських користувачів. URL: <https://www.the-village.com.ua/village/business/business-question/331633-chomu-facebook-ta-instagram-blokuyut-dopisi-pro-viynu-vid-ukrayinskih-koristuvachiv-i-scho-z-tsim-roboti> (дата звернення: 20.10.2022).
5. Bouko, C., Van Ostaeyen, P., Voué, P. Facebook's policies against extremism: Ten years of struggle for more transparency. *First Monday*. 2021. Vol. 26(9). <https://doi.org/10.5210/fm.v26i9.11705>.

6. Ciuriak, D. The Role of Social Media in Russia's War on Ukraine. URL: <https://ssrn.com/abstract=4078863> or <http://dx.doi.org/10.2139/ssrn.4078863> (20.10.2022).
7. Facebook community standards [Electronic resource]. URL: <https://transparency.fb.com/policies/community-standards/?source=http%3A%2F%2Fwww.facebook.com%2Fcommunitystandards%2F> (20.10.2022).
8. Garton Ash, T., Gorwa, R., Metaxa, D. GLASNOST! Nine ways Facebook can make itself a better forum for free speech and democracy. In Reuters Institute for the Study of Journalism Report. Reuters Institute for the Study of Journalism [Electronic resource]. URL: <https://reutersinstitute.politics.ox.ac.uk/our-research/glasnost-nine-ways-facebook-can-make-itself-better-forum-free-speech-and-democracy> (20.10.2022).
9. Hauter, J. Forensic conflict studies: Making sense of war in the social media age. *Media, War & Conflict*. 2021. Vol. 1(1). <https://doi.org/10.1177/17506352211037325>.
10. Herrick, K. Breaking Things: Origins and Consequences of Racialized Hate Speech on Facebook. *International Studies Honors Projects*. 2022. Vol. 39. URL: https://digitalcommons.maclester.edu/intlstudies_honors/39 (20.10.2022).
11. Jouglex, P. Hate Speech, Fake News, and the Moderation Problem. In: *Facebook and the (EU) Law. Law, Governance and Technology Series*. 2022. Vol 48. https://doi.org/10.1007/978-3-031-06596-5_7.
12. Myers West, S. Censored, suspended, shadowbanned: User interpretations of content moderation on social media platforms. *New Media & Society*. 2018. Vol. 20(11). Pp. 4366–4383. <https://doi.org/10.1177/1461444818773059>.
13. Papaevangelou, C., Smyrniaos, N. The Case of a Facebook Content Moderation Debacle in Greece. In: Iordanidou, S., Jebril, N., Takas, E. (eds). *Journalism and Digital Content in Emerging Media Markets*. Palgrave Macmillan, Cham., 2022. https://doi.org/10.1007/978-3-031-04552-3_2.
14. Proposal for a regulation of the european parliament and of the council on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC. URL: <https://eur-lex.europa.eu/legal-content/en/TXT/?uri=COM%3A2020%3A825%3AFIN> (20.10.2022).

15. Roberts Sarah, T. *Behind the Screen: Content Moderation in the Shadows of Social Media*. New Haven, CT: Yale University Press, 2019.
16. Roberts, Sarah T. *Social Media's Silent Filter.* 8 March. *The Atlantic*. URL: <https://www.theatlantic.com/technology/archive/2017/03/commercial-content-moderation/518796/> (20.10.2022).
17. Siapera, E., Viejo-Otero, P. *Governing Hate: Facebook and Digital Racism. Television & New Media*. 2021. Vol. 22(2). Pp. 112–130. <https://doi.org/10.1177/1527476420982232>.
18. Zuckerberg, M. *A Blueprint for Content Governance and Enforcement*. URL: <https://www.facebook.com/notes/mark-zuckerberg/a-blueprint-for-content-governance-and-enforcement/10156443129621634/> (20.10.2022).

REFERENCES

1. *Kilkist dezinformatsii v onlain-media zrosla maizhe vtrychi* — Instytut masovoi informatsii (2022) [The amount of disinformation in the online media has almost tripled — Institute of Mass Information]. Retrieved October 21, 2022, from <https://litgazeta.com.ua/news/kilkistdezinformatsii-v-onlajn-media-zrosla-majzhe-vtrychi-institut-masovoi-informatsii/>.
2. *Potentsiino nepryiniatnyi: yak onlain-platformy obmezhuut kontent ukrainsiv* (2022) [Potentially unacceptable: how online platforms limit the content of Ukrainians]. Retrieved October 21, 2022, from <https://ukrainer.net/pravo-na-kontent/>.
3. *Feisbuk blokuvav vidomoho ukrainskoho blohera Zena Antypopa* (2015) [Facebook blocked the famous Ukrainian blogger Zen Antipop]. *Uriadovyi kur'ier*, January 21. Retrieved October 21, 2022, from <http://ukurier.gov.ua/uk/articles/facebook-ta-youtube-mayut-zminiti-pravila-moderuva/>.
4. *Chomu Feisbuk ta Instahram blokuiut dopysy pro viinu vid ukrainskykh korystuvachiv* (2022) [Why Facebook and Instagram block posts about the war from Ukrainian users]. Retrieved October 21, 2022, from <https://www.the-village.com.ua/village/business/business-question/331633-chomu-facebook-ta-instagram-blokuyut-dopisi-pro-viynu-vid-ukrayinskih-koristuvachiv-i-scho-z-tsim-robiti>.
5. Bouko, C., Van Ostaeyen, P., & Voué, P. (2021). Facebook's policies against extremism: Ten years of struggle for more transparency. *First Monday*, 26 (9). <https://doi.org/10.5210/fm.v26i9.11705>.

6. Ciuriak, D. (2022). The Role of Social Media in Russia's War on Ukraine. Retrieved October 21, 2022, from <https://ssrn.com/abstract=4078863>
7. Facebook community standards (2022). Retrieved October 21, 2022, from <https://transparency.fb.com/policies/community-standards/?source=https%3A%2F%2Fwww.facebook.com%2Fcommunitystandards%2F>.
8. Garton Ash, T., Gorwa, R., & Metaxa, D. (2019). GLASNOST! Nine ways Facebook can make itself a better forum for free speech and democracy. In: Reuters Institute for the Study of Journalism Report. Reuters Institute for the Study of Journalism. Retrieved October 21, 2022, from <https://reutersinstitute.politics.ox.ac.uk/our-research/glasnost-nine-ways-facebook-can-make-itself-better-forum-free-speech-and-democracy>.
9. Hauter, J. (2021). Forensic conflict studies: Making sense of war in the social media age. *Media, War & Conflict*, 1(1). <https://doi.org/10.1177/17506352211037325>.
10. Herrick, K. (2022). Breaking Things: Origins and Consequences of Racialized Hate Speech on Facebook. *International Studies Honors Projects*, 39. Retrieved October 21, 2022, from https://digitalcommons.macalester.edu/intlstudies_honors/39.
11. Jougoux, P. (2022). Hate Speech, Fake News, and the Moderation Problem. In: *Facebook and the (EU) Law. Law, Governance and Technology Series*, 48. https://doi.org/10.1007/978-3-031-06596-5_7.
12. Myers West, S. (2018). Censored, suspended, shadowbanned: User interpretations of content moderation on social media platforms. *New Media & Society*, 20 (11), 4366-4383. <https://doi.org/10.1177/1461444818773059>.
13. Papaevangelou, C., & Smyrniaios, N. (2022). The Case of a Facebook Content Moderation Debacle in Greece. In: Iordanidou, S., Jebril, N., Takas, E. (eds). *Journalism and Digital Content in Emerging Media Markets*. Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-031-04552-3_2.
14. Proposal for a regulation of the european parliament and of the council on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC (2020). Retrieved October 21, 2022, from <https://eur-lex.europa.eu/legal-content/en/TXT/?uri=COM%3A2020%3A825%3AFIN>.
15. Roberts Sarah, T. (2019). *Behind the Screen: Content Moderation in the Shadows of Social Media*. New Haven, CT: Yale University Press.

16. Roberts Sarah, T. (2017). Social Media's Silent Filter. 8 March. The Atlantic. Retrieved October 21, 2022, from <https://www.theatlantic.com/technology/archive/2017/03/commercial-content-moderation/518796/>.
17. Siapera, E., & Viejo-Otero, P. (2021). Governing Hate: Facebook and Digital Racism. *Television & New Media*, 22 (2), Pp. 112–130. <https://doi.org/10.1177/1527476420982232>.
18. Zuckerberg, M. (2018). A Blueprint for Content Governance and Enforcement. Retrieved October 21, 2022, from <https://www.facebook.com/notes/mark-zuckerberg/a-blueprint-for-content-governance-and-enforcement/10156443129621634/>.