

---

## Рекомендації щодо запобігання поширення нелегального контенту в українському сегменті Facebook та утвердження принципів свободи слова в Інтернеті

Ці рекомендації зібрані в процесі спостережень та обговорень найбільших викликів, пов'язаних з питанням балансу свободи слова та захисту від дезінформації та інформаційних спецоперацій.

### Преамбула

1. Соціальні мережі протягом останніх 15 років стали невід'ємним майданчиком для Р2Р-комунікації, обміном інформацією за професійними та іншими інтересами та для поширення новин й іншої соціально значущої інформації. Після запровадження легкого і доступного поширення фото та відео, соціальна мережа Facebook отримала значний приріст користувачів ще в кінці першого десятиліття, а після встановлення санкцій щодо компаній, що походять з держави-агресора (aggressor State)<sup>1</sup> вийшов, разом з Youtube в лідери платформ спільного доступу та обміну інформацією.
2. Попри глобальність самої соцмережі і можливість комунікувати з 3 мільярдами користувачів Facebook, існує частина соцмережі, яку умовно можна назвати "український FB", що частково може бути пов'язано з перебудовування держави в умовах воєнного конфлікту. Це пов'язано з низкою факторів, як от мова (українська, російська), потреба в отриманні швидкої інформації про події, що відбуваються навколо України, її регіонів та місць проживання тих чи інших користувачів. Особливою проблемою є так звані "інформаційні бульбашки", в яких опиняються користувачі, очікувано оточуючи себе певним колом контактів та інтересів, що створює потенційні ризики для таргетування з метою впливів на політичні уподобання або для поширення дезінформації.
3. Окрім цього, існує контентна відмінність українського сегменту FB від західноєвропейських країн. Найбільш виразно вона проявилася під час Революції Гідності, коли FB разом з YouTube стали платформою для

---

<sup>1</sup> Поняття «victim State» та «aggressor State» вжиті для позначення сторін збройного конфлікту відповідно до термінології міжнародного гуманітарного права (див. зокрема, рішення Міжнародного суду ООН у справі Нікарагуа проти Сполучених Штатів Америки): <https://www.icj-cij.org/en/case/70/judgments>.



найшвидшого і оперативного поширення інформації про Майдан і атаки на нього. До, а надто після Революції гідності FB став майданчиком для активної комунікації щодо соціально важливих питань, спровокувавши феномен, коли платформа стає інструментом комунікації для ключових лідерів думок, в тому числі й топ-посадовців в країні, а в 2019 році продукти FB стали активним елементом президентських і парламентських виборів. Це значно відрізняється від сімейних фото, “котиків”, мандрівок користувачів із країн західної Європи.

4. Основним документом Facebook є [Стандарти](#) спільноти<sup>2</sup> (аналогічно у Youtube існує поняття Правил спільноти<sup>3</sup> - прим), що визначає правила поведінки на цій платформі, обмеження щодо поширення певних типів контенту і, водночас підстави для накладення санкцій, як от вилучення певних публікацій або блокування користувача. Стандарти спільноти є динамічним документом і кожен користувач може побачити їх оновлення, як от наприклад [останні](#) оновлення<sup>4</sup> щодо мови ворожнечі в контексті анімалізації осіб та використання позначок тварин у комбінації з певними етносами або віруючими. Такі оновлення є надзвичайно важливими, адже дозволяють вчасно реагувати на нові форми ненавистницьких висловів та інших проявів загрозової поведінки (див., наприклад, Замість слів: ненависть, булінг та домагання мовою емодзі<sup>5</sup>). Разом з тим, існує низка проблем, які потребують особливої уваги з боку FB та інших соцмереж, з огляду на специфіку та особливу ситуацію, в якій перебуває Україна, і, власне, що є практичною моделлю для вирішення аналогічних конфліктів що виникають або можуть виникати в світі.

## Проблема 1. Війна і комунікація “навколовоєнних” питань

5. Україна з листопада 2013 року опинилася під потужним інформаційним тиском з боку Російської федерації. Від початку маніпулятивних сюжетів<sup>6</sup>

<sup>2</sup> <https://www.facebook.com/communitystandards/> .

<sup>3</sup> <https://www.youtube.com/about/policies/#community-guidelines> .

<sup>4</sup> [https://www.facebook.com/communitystandards/recentupdates/hate\\_speech/](https://www.facebook.com/communitystandards/recentupdates/hate_speech/) .

<sup>5</sup> <https://old.cedem.org.ua/analytics/emoji/> .

<sup>6</sup> <https://www.presscouncil.ru/index.php/praktika/rassmotrennye-zhaloby/3007-zhaloba-na-vesti-nedeli-s-dmitriem-kisilevym-iz-za-osveshcheniya-evromajdana?showall=&start=9> .



Дмитрія Кисельова про Майдан до низки інформаційних спецоперацій щодо збитого рейсу МН-17, “участі військ НАТО” тощо. Слід зауважити, що командам Facebook вдається боротися з фейками або понижує їх рейтинг, або видаляючи їх з платформи, якщо це є елемент інформаційних спецоперацій (координувана неавтентична поведінка). Разом з тим, більш проблемним є взаємодія інформаційних сфер країн, що перебувають у стані проголошеної чи непроголошеної війни.

6. Очевидно дещо спрощеним є підхід і інструменти до ситуації, коли країни є у стані конфлікту, а відтак населення має значно вищий поріг чутливості і рівень агресії намагаються застосувати механізми загального характеру.
7. Загальновідомим фактом є те, що Російська федерація окупувала частину України — Крим, а також бере активну роль у збройному конфлікті на території Донецької і Луганської областей. Аналогічна ситуація існує на окупованій частині Грузії та Молдови. Російська федерація, а надто її силові структури успадкували техніки і методики своїх попередників. Одним з важливих елементів ведення війни є застосування інформаційних військ, про створення яких в РФ заявлялося<sup>7</sup> ще у 2017 році. Потужна внутрішня пропаганда, а також інформаційні впливи назовні покликані підтримати ворожість<sup>8</sup> до демократичної України та цінностей свободи, але, водночас демонізувати Україну назовні через чутливі і надто складні теми минулого через просування теми “нацизму”, “бандеризму” тощо. Така обробка має наслідком ворожість з боку рядових росіян та використання таких штампів як “майданутые”, “а вы все скачете” “хохлы” “нацисты на майдане” тощо. Крім того, використовуються штампи “ваши хозяева”, “продались Америке”, “марионетковое правительство”, “failed state”, що зводиться до заперечення суб’єктності України і її уряду. Також окремою темою є заперечення вини Росії за збройний конфлікт на Донбасі, де українці звинувачуються у “расстреле женщин и детей на Донбассе”.

<sup>7</sup> <https://tyzhden.ua/World/186503>

<sup>8</sup> <https://www.dw.com/uk/%D1%80%D0%BE%D1%81%D1%96%D1%8F%D0%BD%D0%B8-%D0%B2%D0%B2%D0%B0%D0%B6%D0%B0%D1%8E%D1%82%D1%8C-%D1%81%D1%88%D0%B0-%D1%82%D0%B0-%D1%83%D0%BA%D1%80%D0%B0%D1%97%D0%BD%D1%83-%D0%BD%D0%B0%D0%B9%D0%B1%D1%96%D0%BB%D1%8C%D1%88-%D0%B2%D0%BE%D1%80%D0%BE%D0%B6%D0%B8%D0%BC%D0%B8-%D0%BA%D1%80%D0%B0%D1%97%D0%BD%D0%B0%D0%BC%D0%B8-%D0%BE%D0%BF%D0%B8%D1%82%D1%83%D0%B2%D0%B0%D0%BD%D0%BD%D1%8F/a-51013726>



8. *Vise versa* українські користувачі обурені російською військовою агресією, російською пропагандою, а також маючи так чи інакше загиблих на Майдані та в зоні проведення ООС не обмежують себе в епітетах починаючи з відомої кричалки на адресу російського лідера і закінчуючи такими виразами як “кацапи”, “русня”, “свинособаки”, “трихромосомні”<sup>9</sup>.
9. Варто зауважити, що Facebook у 2017 році [визначив](#)<sup>10</sup>, що “хохол” та “москаль” є неприйнятними на платформі, разом з тим таке рішення не до кінця враховує історичне та ономастичне значення слова “москаль”, яке означає як жителя історичної Московії, солдат, що служили в царській армії, а також низку прізвищ — Москаль, Москалець, Москаленко тощо. Найбільш відомим є цитата Тараса Шевченка “Кохайтеся чорнобриві, та не з москалями”, що первісно означало засторогу про трагічність стосунків з чоловіками, яких забирали в царську армію, а нині — засторогу про дружбу з представниками країни-агресора. Натомість “хохли” є більш виражене дискримінуюче, адже має чітку прив’язку до українців, часто використовується у виразах на кшталт “тупые хохлы” або у зв’язці “хохол” - “сало”, таким чином виставляючи українців як неосвічених недалеких людей. Дзеркальним відповідником є слово “кацап”. Також варто звернути увагу на контекст використання слова “русня”, що використовується найперше щодо російських військових, що беруть участь у військових діях на території Донецької та Луганської областей.
10. Зважаючи на ситуацію відкритих чи прихованих конфліктів базові засади Стандартів спільноти, як от Достовірність, Безпека та Гідність мають містити додаткові запобіжники та обмеження для того, щоб користувачі могли вільно почуватися на платформі. Такі обмеження цілком відповідають обмеженням, передбаченим ч.2. ст.10 Європейської конвенції захисту прав людини та основоположних свобод.
11. Слід зауважити, що країна-агресор та її інформаційні сателіти не можуть за замовчуванням підпадати під критерії Достовірності, Безпеки і Гідності, адже наміри таких суб’єктів полягають в агресивних діях щодо іншої держави або окремих груп, а відтак і у поширенні дезінформації. Ба більше, для Facebook слід ретельніше вивчити догмати так званого “русского міра”, які передбачають дискримінацію різного роду меншин. З огляду на це варто

<sup>9</sup> [https://www.bbc.com/russian/rolling\\_news/2013/01/130121\\_rn\\_medinsky\\_usa\\_interview](https://www.bbc.com/russian/rolling_news/2013/01/130121_rn_medinsky_usa_interview) .

<sup>10</sup> <https://about.fb.com/news/2017/06/hard-questions-hate-speech/>.



розглянути кілька механізмів, які мають нівелювати діяльність країни-агресора в інформаційному полі країни-жертви:

- 1) Чітка ідентифікація, що це продукт, спільнота, контент походять з країни-агресора;  
*Такий інструмент надає можливість користувачу, надто необізнаному, розуміти, що він споживає контент, що походить з країни агресора.*
- 2) Додавання в механізмі скарг опцій “контент пов’язаний зі збройною агресією” в розділі “мова ворожнечі” та “неправдиві новини”.  
*Такий механізм дозволяє підсвітити модератору, що той чи інший контент може використовуватися як елемент інформаційної агресії, відтак потребує ретельнішої перевірки.*
- 3) Обмеження діяльності суб’єктів з країни агресора, надто таргетування реклами, заборону на участь в групах і спільнотах, що направлені на користувачів з України (загалом - країни-жертви).
- 4) При модерації скарг зважати на пост одночасно і в його текстуальній і візуальній частині, адже саме поєднання нейтрального коментаря з агресивним зображенням або vice versa нейтрального зображення з агресивним коментарем використовується для поширення ворожнечі.
- 5) В контексті мови ворожнечі додати можливість позначати, що це не лише мова ворожнечі, що пов’язана з етнічним чи расовим походженням, але також зазначити що це мова ворожнечі, що поширюється представниками країни-агресора.  
*В ідеалі, варто розглянути можливість максимального зниження будь-яких контактів користувачів країни-агресора і країни-жертви, адже сама насильницька природа породжує справедливий гнів в реакції з боку держави-жертви (Victim state)<sup>11</sup>.*

12. Варто зауважити про протилежний ефект щодо блокування користувачів чи контенту за різкі і агресивні коментарі на адресу країни агресора лише посилює відчуття правоти користувача, а сама платформа починає сприйматися як така, яка сприяє діяльності країни-агресору, та посилює низку міфів, на кшталт “московський офіс FB”.

---

<sup>11</sup> Поняття «victim State» та «aggressor State» вжиті для позначення сторін збройного конфлікту відповідно до термінології міжнародного гуманітарного права (див. зокрема, рішення Міжнародного суду ООН у справі Нікарагуа проти Сполучених Штатів Америки): <https://www.icj-cij.org/en/case/70/judgments> .



## Проблема 2. Безпека журналістів, медіа, громадських активістів

13. Журналісти, громадські активісти та медіа в Україні є активним користувачами FB та Instagram. Часто кількість їх підписників може сягати десятків і сотень тисяч. Формат стрічки соціальної мережі дозволяє поширювати короткі швидкі повідомлення, інформуючи суспільство і локальні спільноти про важливі події, чи, у випадку активістів — важливі ініціативи та заходи. Разом з тим, діяльність таких суб'єктів пов'язана з низкою ризиків — протидія з боку суб'єктів, на яких направлена діяльність (наприклад корпціонери проти активістів чи журналістів-розслідувачів, шахрайство, інтернет-помста, намагання перехоплення аудиторії тощо). Яскравими випадками стали спроба рейдерського захоплення сторінки одеського видання “Думская”, що, як видається, стало помстою за низку журналістських розслідувань, атака, у тому числі за допомогою купівлі реклами на інтернет-сторінку “Судовий репортер” з метою рейдерства останнього, блокування постів журналістів.
14. Ми розуміємо, що створення “білих списків” журналістів чи медіа є неможливим ані з точки зору адміністрування, ані з точки зору гарантій, що медіа будуть дотримуватися стандартів постійно. Тут яскравим прикладом є купівля телеканалу “ZIK” і тотальна зміна редакційної політики на угоду новим власникам. Разом з тим, ми переконані, що платформі слід забезпечити простий механізм верифікації автентичності профілю медіа, журналіста чи активіста, що має знизити ризик створення клонів та захоплення сторінок. Зокрема, одним із ефективних способів ідентифікації медіа, якщо точніше – новинних сторінок – можна згадати News Page Index<sup>12</sup>, який уже трохи менше двох років використовується як інструмент, що допомагає платформі розпізнати сторінку як видавця новин. У цій сфері платформі потрібно працювати над досягненням розуміння власниками сторінок, що News Page Index відрізняється від процедури верифікації та за зовнішніми ознаками ніяк не вирізняє новинну сторінку від усіх інших. При цьому, така процедура полегшує розпізнавання платформою такої сторінки як новинної, а також згодом включатиме додаткові опції для таких сторінок у

<sup>12</sup> <https://www.facebook.com/journalismproject/indexing-facebook-news-pages-ad-archive>.



чат-маркетингу (поява опції підписки у Messenger)<sup>13</sup>. Також, одним із можливих елементів полегшення для новинних сторінок процесу подачі заявки, є забезпечення україномовного інтерфейсу у відповідних налаштуваннях (зараз доступно англійською, німецькою, французькою, нідерландською, португальською, італійською, хінді, бенгальською, тамільською, маратською та сингальською).

15. Водночас варто ретельніше перевіряти будь-які скарги на медіа, як такі, що можуть бути елементом тиску чи неправдивого використання скарг для ведення в оману модераторів FB. Особливо ретельно слід ухвалювати рішення щодо блокування медіа і користувачів і намагатися консультуватися з незалежними експертами щодо пропорційності шкоди/блокування. Також варто передбачити, що у випадку застосування санкцій щодо профілів журналістів, активістів, медіа має діяти ефективний і швидкий механізм оскаржень.
16. Варто також розглянути можливість для підвищеної уваги і захисту вимагати від медіа, громадських організацій розкривати інформацію про учасників, керівний склад тощо. Таким чином платформа убезпечує себе від захисту сайтів-одноденок або “зливних бачків”, а споживачі інформації мають отримати інформацію про те, хто відповідає та впливає на редакційну політику чи на діяльність громадської організації.

## **Проблема 2. Зловживання механізмами захисту авторського права**

17. Одним з найбільш ризикових сфер для свободи слова є зловживання механізмами захисту авторських прав. Основна відмінність українського законодавства полягає в презумпції авторства, де від автора не вимагається наявність певного ідентифікатора для контенту. Разом з тим платформа діє за правилами DMCA, відповідних Директив ЄС, що регулюють питання авторських прав в цифровому середовищі. Механізм “notice and take down” є дійсно ефективним для контенту, що має ідентифікатор, разом з тим до “дрібного” контенту, як от фото чи текст, такі механізми можуть стати елементом цензури, коли заява від “правовласника” може спричинити охолоджуючий ефект (chilling effect) або навіть блокувати медіа. Відтак, платформі слід зважити на а) законодавство України щодо захисту

<sup>13</sup> <https://developers.facebook.com/docs/messenger-platform/policy/policy-overview/>.



авторських прав, особливо, коли мова йде про скарги від одного резидента України до іншого; б) на значний ризик охолоджуючого ефекту для медіа; в) презумпцію злого умислу, якщо скарги надходять з країни-агресора на українські медіа.

### Проблема 3. Дифамація

18. Окремою проблемою є використання платформи для поширення дезінформації та чорного піару (в українських реаліях найбільш відповідає розумінню терміну malinformation). Природа такої діяльності може мати як політичний аспект, надто під час виборчих кампаній та в міжвиборчий період, а також в комерційних питаннях і як протидія діяльності громадських активістів і журналістів-розслідувачів. Однією з найбільш конфліктних сфер є питання забудов, яке історично завжди пов'язано з протистоянням “збудувник vs. активісти” або “збудувник vs. місцеві жителі”. При цьому ресурсно забудовники мають більші ресурси, зокрема, і фінансові. Це все відбувається, для прикладу, шляхом закупівлі реклами з посиланням на статті, які дискредитують діяльність захисників історичної спадщини чи осіб, які виступають лідерами відповідних груп. Іншою ситуацією є публікації з боку політиків, політтехнологів, де вони атакують ту чи іншу особу. При цьому особа не має жодного ефективного захисту від такої атаки. Ба більше, публікація спростування на стіні особи, що є під атакою, не може вважатися реалізацією права на відповідь, адже аудиторія такої особи відрізнятиметься від аудиторії, на яку було поширено первинне повідомлення.
19. Платформа не здійснює редакційної політики, адже дійсно, неможливо відслідкувати кожен конкретний кейс — це має залишатися за судовими інстанціями. Проте, швидкість і шкода, яку може нанести поширення дифамації на платформі є значною в силу самої природи потоку інформації. Саме тому важливо надати інструментарій, коли особа може опублікувати право на відповідь і ця відповідь буде помітна аудиторії первинного поширення, а у випадку використання рекламних інструментів -встановити вимоги призупинення або зняття реклами, що має на меті нанести шкоду.





#### Проблема 4. Сторінки-одноденки і реклама

20. Зважаючи на природу платформи і швидкість поширення інформації, інструменти платформи стають магнітом для осіб, що мають на меті шахрайство, поширення дифамації під час виборів, маніпулювання громадською думкою тощо. Одним зі способів таких дій є створення сторінок, на яких є мінімальна кількість контенту, які водночас запускають низку рекламних постів з метою просування посту або посилань на зовнішні ресурси. Очевидно, що платформі слід більш уважно ставитися до реклами з боку сторінок з малою кількістю контенту.
21. Крім цього, в контексті розміщення реклами на платформі актуальними залишаються питання таргетингу та збору приватних даних. Так, це проблема неодноразово обговорювалася на слуханнях у американському сенаті<sup>14</sup> і супроводжувалося вимогами контролювати збір приватної інформації, суб'єктів, які мають до неї доступ, а також чітко визначати способи та цілі її використання. Наразі алгоритми платформи щодо розміщення реклами є частково прихованими від користувачів, втім при реєстрації на ресурсі та внесенні до умов користування сервісом платформа має надавати чіткий перелік даних, які вона збирає, а також цілей використання таких даних. Зокрема, це може допомогти вирішити контраверсійне питання щодо збирання інформації з третіх веб-сайтів для подальшого використання в цілях рекламного таргетингу. Ще одним важливим елементом є належне повідомлення користувачів про їхні права та обов'язки. По-перше, вони мають бути обізнані з будь-якими новаціями на платформі (зокрема, своєчасне повідомлення про імплементацію нових алгоритмів, появу нових функцій, встановлення нових контентних заборон тощо); по-друге, ознайомлення з можливістю відмовитися від поширення певних типів реклами через блокування конкретних рекламних суб'єктів чи тематик на платформі. Основними питаннями, що потребують вирішення в цій площині є: переклад усіх роз'яснень Facebook щодо умов користування та особливостей функціонування мовою держави, де платформа функціонує, а також використання зрозумілих та стислих формулювань, що дозволять користувачу краще засвоювати подану інформацію.

<sup>14</sup> <https://www.washingtonpost.com/news/the-switch/wp/2018/04/10/transcript-of-mark-zuckerbergs-senate-hearing/>.



## Проблема 5. Модерація

22. Однією з найбільш поширених проблем є неефективність оскаржень на порушення чи блокування. Обираючи скаргу, особі випадає меню, де вона вказує модераторам на сферу порушень. При цьому, особа не має можливості коротко описати свою проблему і єдиним додатковим механізмом мережі нині є лише опція “надіслати на повторну перевірку”. Для більш ефективного захисту або для більш якісної протидії мові ворожнечі, платформі слід на апеляційному рівні запровадити більш ґрунтовний механізм оскарження, де оскаржувач має обрати відповідний пункт стандартів спільноти та надати коротке пояснення в чому саме він вбачає порушення Стандартів. Адже попри намагання Facebook залучати носіїв мови, знання контексту може бути вузькопрофільним і невідомим модератору, разом з тим досить зрозумілим особам про яких чи на яких спрямована мова ворожнечі.
23. У питаннях поліпшення функціонування платформи дієвим механізмом є співрегулювання. По-перше, в багатьох випадках це дозволяє належним чином і в короткі строки скорегувати умови користування платформою, привівши їх у відповідність із національним законодавством. По-друге, це допомагає розділити відповідальність платформи за формування власних політик, уникнути фундаментальних помилок у формуванні підходів. Варто відзначити, що механізм співрегулювання вже довів свою ефективність. наприклад, з листопада 2018 року Facebook посилено співпрацює з французькими державними органами для обмеження поширення мови ворожнечі<sup>15</sup>, а з червня 2019 – погодився надавати інформацію на судові запити у таких справах<sup>16</sup>. В результаті, платформа розробила ефективніший механізм реагування на скарги, а також спромоглася сформувати більш локальне керівництво для оцінки контенту.

<sup>15</sup> <https://www.reuters.com/article/us-france-facebook-macron/france-to-embed-regulators-at-facebook-to-combat-hate-speech-idUSKCN1NH1UK>.

<sup>16</sup> <https://www.reuters.com/article/us-france-tech-exclusive/exclusive-facebook-to-give-data-on-hate-speech-suspects-to-french-courts-minister-idUSKCN1TQ1TJ>.

