

Countering Online Antigypsyism and Cyberhate (COACH)
Project: 101084694 — COACH — CERV-2022-EQUAL

Реч, съдържаща обиди, предразсъдъци и омраза спрямо ромите в социалната медия Фейсбук

Боян Захариев,
Фондация "Институт Отворено общество - София"

януари, 2024 г.



Funded by the European Union

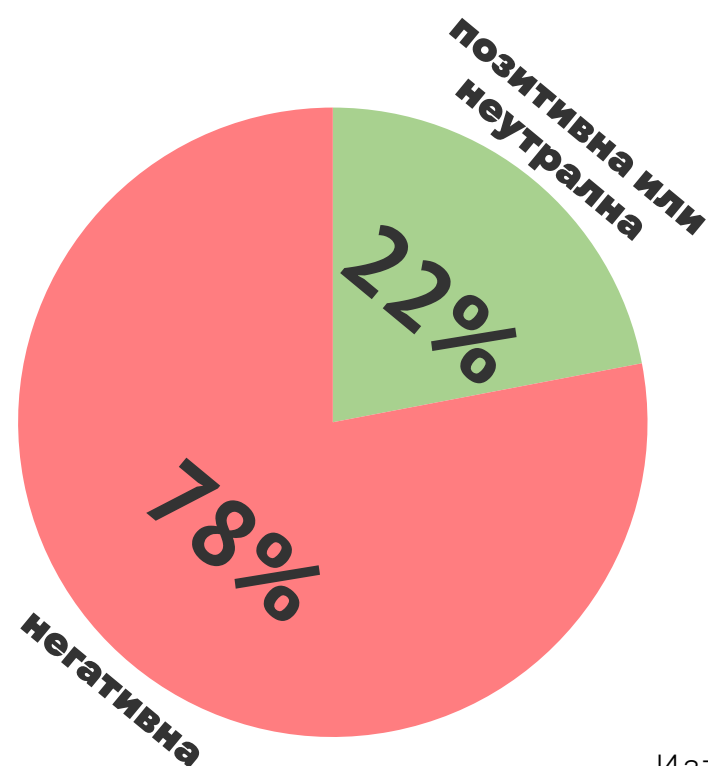
За изследването

Материалите са събрани между 1-ви септември и 15 октомври 2023 г. От 1252 текста, в които под някаква форма се споменават ромите, поне около 80% имат обидно съдържание по отношение на ромите, но ако не броим публикациите в чисто ромски групи в социалните медии или от ромски организации, всъщност делът на негативните публикации би надхвърлил 90%.

Търсенето във Фейсбук е осъществено автоматично с помощта на ключови думи, предварително подбрани от експерти. Публикациите са анализирани експериментално с помощта на изкуствен интелект.

Методологията и използваните алгоритми за автоматичен анализ са разработени от консорциум от организации, основали с помощта на финансиране от Европейския съюз Европейската обсерватория за омразата онлайн (European Observatory of Online Hate) - EOOO с водещ партньор спиноф компанията за езикови технологии Textgain, създадена от Антверпенския университет. Партньор за събирането и обработката на данните на български език е ЦМДТ „Амалипе“.

Дял на публикациите, съдържащи обиди, предразсъдъци или омраза спрямо ромите във Фейсбук за периода 01.09-15.10.2023 г.



Източник: EOOO и ЦМЕДТ „Амалипе“

По най-консервативни оценки, поне 80% от публикациите имат обидно съдържание по отношение на ромите. Този резултат се получава ако броим и публикациите от ромски Фейсбук групи и ромски организации, които допринасят за позитивната статистика и ако считаме неголемия брой двусмислени (неясни) публикации за неутрални или позитивни. Огромната част от материалите на български език, засягащи ромите са от Фейсбук. Материалите по тази тема, намерени в други социални медии са единични.

От общо 2641 записа са извадени 1805 уникални записа, от които 1252 се отнасят до ромите, а останалите 553 са на различна тематика, макар че част от тях съдържат реч на омразата.

От 1252 текста, в които под някаква форма се споменават ромите, след внимателен преглед за подробен анализ са оставени 1096 уникални публикации (постове). Записите на различна тематика, дори да съдържат реч на омразата не трябва да влизат в статистиката за честотата на поява на реч на омразата срещу ромите в определена медия, нито в оценката на корпуса от текстове, насочени към ромите.

Под уникален запис се има предвид появата на даден текст в определен профил в социалната медия. Споделянето на същия текст в друг профил е отделен текст в корпуса, тъй като всяко отделно споделяне поражда различни по брой и съдържание коментари. Освен това от статистическа гледна точка появата на даден текст в повече от един профил увеличава вероятността той да бъде прочетен от повече потребители, да порожи повече реакции и да има като цяло по-голям ефект.



Видове публикации във Фейсбук, съдържащи обида, омраза или предрасъдъци към ромите

В настоящия анализ са представени данни от вторична полуавтоматизирана обработка, която разчита на критичен преглед на работата на алгоритмите. Целта е от една страна да бъде представена по-достоверна информация за обидната реч спрямо ромите в социалните медии, а от друга страна – да бъдат открити слабите места и типичните пропуски/грешки в използваните за автоматична обработка алгоритми.

След анализ на корпуса от публикации обособихме **9 основни категории**, които частично се припокриват, но всяка публикация е поставена само в една категория, която най-добре съответства на съдържанието ѝ:

Категории

1. Публикации, съдържащи пряка обида.

2. Публикации, съдържащи ирония и сарказъм.

3. Публикации, съдържащи обща обида.

4. Вицове.

5. Публикации, разказващи някаква (истинска) история.

6. Публикации, разказващи история с участието на животни.

Този вид истории са обособени в отделна категория, поради специфичните послания, различната аудитория и голямата им токсичност.

7. Публикации, съдържащи предрасъдъци.

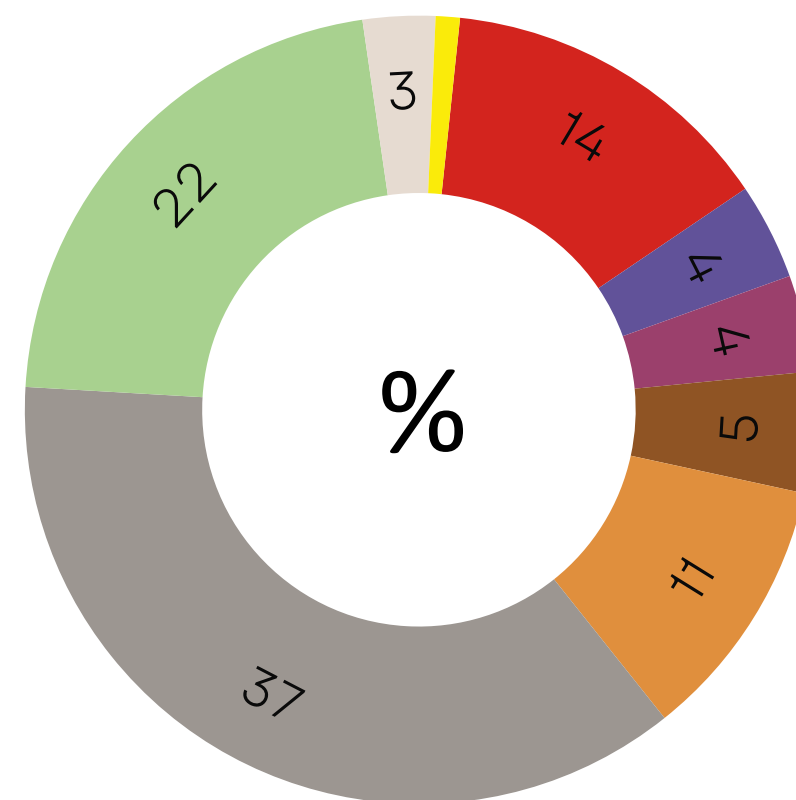
8. Обяви.

9. Публикации с позитивно или неутрално съдържание.

Публикациите с позитивно или неутрално съдържание са обособени от останалите 8 категории, които не са напълно взаимноизключващи се и понякога се припокриват. Отделни текстове могат да бъдат подходящи за повече от една категория. Например някои истории носят анекдотичен характер и се доближават до вицовете, макар че са разказани като реални случки. За обобщената статистика всяка публикация е поставена само в една категория, която най-добре описва съдържанието ѝ.

Тематично разпределение на публикациите във Фейсбук за периода 1.09-15.10.2023 г., съдържащи обида, предрасъдъци или омраза спрямо ромите

- пряка обида
- сарказъм
- обща обида
- виц
- история
- предрасъдък
- позитивна или неутрална
- обяви
- история с животни



Източник: ЕООО и ЦМЕДТ „Амалипе“

Публикации, съдържащи пряка обида и израз на омраза към ромите като група или към отделни хора от ромски произход.

Това са едни от най-токсичните публикации, които почти сигурно попадат в категорията на публикации, подлежащи на премахване според стандартите на Фейсбук. Тези публикации използват директни оскърбителни наименования на ромите, съпроводени с нецензурна реч и обидни епитети. Тези публикации формират до 15% от корпуса с текстове (152 публикации), като не винаги е лесно да бъдат различени от общите обиди.

Публикации, съдържащи ирония или сарказъм.

Това са между 4% и 10% от публикациите. Тази категория се застъпва с вицовете. Отделни саркастични забележки и коментари се съдържат и в постове с преки обиди и стереотипи и с историите. Саркастичните бележки са насочени към различни стереотипни характеристики, приписвани на ромите, сред които са начин на говорене и изразяване. Използват се и игра на думи, омоними и каламбури. Публикациите съдържащи сарказъм попадат в гранична зона, за която не е сигурно дали нарушава стандартите на Фейсбук, но част от тях подлежат на премахване ако бъдат докладвани.

Публикации, съдържащи обща обида, която се основава на предразсъдъци и омраза към ромите.

Тези публикации използват обиди, свързани с ромите, но ги прилагат към всеки, който се е превърнал в обект на омраза, независимо от етническия му произход. По този начин предполагаемата принадлежност към ромския етнос се третира като обидна сама по себе си и така предразсъдъците и омразата се затвърждават и добиват статут на общоприета практика. Това прави въпросните публикации изключително токсични. Публикациите от тази категория би трябвало да бъдат премахнати ако бъдат докладвани. Има постове, в които ромите не са пряко споменати, но в тях се развиват откровено расистки теории, споменаващи директно Адолф Хитлер.

Публикации, съдържащи вицове с участието на ромите.

Голяма част от вицовете с участието на ромите са с нецензурно съдържание. На ромите от всякаква възраст, включително на жените се приписва нецензурна лексика. Вицовете осмиват образи на ромите, които се държат и говорят по начин, приписван на тях от стереотипите. Рядко, във вицовете в стереотипизирани роли например на полицаи, участват и персонажи, които не са ромите.

Публикации, разказващи история с участието на ромите.

Участието на ромите в различни истории е свързано с някои от най-обидните публикации. Като истории могат да бъдат класифицирани малко повече от 11% от публикациите. Някои вицове на пръв поглед приличат на истории. Историите са значително по-дълги текстове в сравнение с всички останали категории, включително вицовете. Макар, че по-често не съдържат преки обиди, историите съдържат многобройни силно негативни описания на ромите и вероятно действат като мощен двигател за затвърждаване на стереотипи и създаване на омраза към ромите. Ромите са обрисувани като хора без задръжки, готови на измами, склонни да нарушават закона за лична изгода. Особена черта на историите е създаването на чувство на морално превъзходство, което от своя страна е основен инструмент за дехуманизацията на определени групи. В историите не рядко присъства контрастен образ на основен герой (протагонист) – самият разказвач или друг, който е надарен с ценни човешки качества. Това е обичайно образ на човек, когото предполагаемата аудитория, посещаваща определен профил в социалните мрежи ценят. Именно този протагонист описва ромите като аморални по подраждане и това засилва ефекта на отхвърляне.

Публикации, разказващи истории, в които става дума за отношението към животните.

Особена категория истории са историите, в които става дума за отношението на хората към домашните животни и домашните любимци. Тези публикации са силно токсични. В тях ромите обичайно са описани като хора, склонни към необяснима жестокост. В немалко от историите от този вид авторите по презумпция допускат, че някое изоставено или пострадало животно е било жертва на лошо отношение или наранено именно от ромите. Съдейки по контекста и реакциите, читателите на подобни материали са най-често хора, които обичат животните и се интересуват предимно от домашни любимци, а не толкова от политика или отношенията между хората. Това показва степента на проникване на предразсъдъците спрямо ромите в различни обществени групи. Начините, по които работят предразсъдъците се разкриват например в историите с животни, където ромите са всъщност жертва, но авторите им отново намират начин да ги поставят от обратната страна. Например, когато санитарните власти отнемат от ромите техни домашни животни, за да ги евтаназират заради предполагаема зараза, фактът че потърпевшите са ромите не се коментира. Вместо това акцентът е изместен към необосновани и маловажни твърдения, че част от санитарния екип вероятно са ромите – съответно цялата негативна енергия се насочва към тях, макар че за читателите е ясно, че нито санитарните правила, нито решенията за тяхното прилагане са прерогатив на ромите.

Публикации, основани на предрасъдъци.

В тази категория поставяме обичайно кратки текстове, които нямат сюжет (не са истории или вицове). Най-често става дума за кратки коментари и реакции на новини за политически или други събития или на други публикации в социалните медии. Публикациите, основани на предрасъдъци не използват директно обидни реплики или нецензурна лексика. Вместо това, нерядко завоалирано, подкрепят утвърдени стереотипи спрямо ромите. Публикации, основани на предрасъдъци се появяват понякога и в отговор на силно позитивни новини, например за постиженията на отделни роми в спорта, изкуствата, работата или образованието. Типичната реакция в подобни случаи се свежда до опит да се обезцени постижението („има толкова български деца с още по-впечатляващи резултати, а за тях никога не говори“) или да се препотвърди утвърдения стереотип („изключението потвърждава правилото“). Любопитното в случая е, че чисто статистически позитивни публикации за постиженията на роми всъщност се появяват изключително рядко и достигат до ограничена аудитория, за разлика например от информацията за постиженията на деца от училища, където не се обучават роми, която често се появява в някои от най-популярните електронни и печатни медии. Една от характеристиките на стереотипите е именно, че те не се подлагат на емпирична проверка, защото имат статута на общоизвестни и общопризнати истини.

Обяви.

Обявите са публикации, отнасящи се основно до продажбата и отдаването под наем на недвижими имоти. В корпуса с текстове има обяви както за продажба и отдаване под наем, така и обяви от хора търсещи да наемат жилище. Това, което обединява обявите е, че в тях се посочва недвусмислено, че ромите са нежелани съседи и нежелани наематели. При предлагането на имоти по селата обявите специално подчертават, че на мястото не живеят роми. Подразбира се, че това повишава стойността и привлекателността на имота. Наемодателите пишат в обявите си, че не желаят с тях да се свързват роми. Това са прояви на неприкрит дискриминация спрямо ромите.

Публикации с позитивно или неутрално съдържание.

Това са публикации, в които ромите се споменават с неутрално или позитивно отношение. При най-благоприятна оценка тези публикации формират около 20-22% от общия корпус. Този резултат се получава, когато публикации с двусмислено или не съвсем ясно значение бъдат определени като неутрални или позитивни. В тази категория са включени и материалите от Фейсбук група, в която участниците са само роми, както и няколко съобщения и становища на ромски организации или групи. Материалите от ромската Фейсбук група и сходните на нея материали по всяка вероятност не достигат до потребители извън

ромските общности и нямат влияние върху общото отношение към ромите. В този смисъл те биха могли дори да бъдат изключени от обобщената статистика на позитивната информация. Другата част от неутралните и позитивните публикации се дължат на неправителствените организации – те съдържат съобщения за проведени събития, покани, обяви за конкурси, информация за проекти и друга подобна информация. Извън кръга на тези материали има малък брой позитивни публикации (в нашата извадка те са общо 3), които обаче са споделени много пъти в социалните медии и са добили сериозна популярност. Към тях спадат две отделни съобщения за младеж и девойка от ромски произход с големи успехи в образованието, както и видео на ромска танцова група, която играе хоро. Както беше споменато в раздела за стереотипите дори тези публикации, на фона на преобладаващото одобрение, успяват да породят и негативни реакции.

Анализ и коментар

Наименования на ромите

В 54% от текстовете е употребено под някаква форма самоназванието на етническата група – роми, в 37% е употребено външното название – цигани. Обичайно употребата спрямо дадена етническа група на нейното самоназвание е начин да се засвидетелства уважение и е проява на любезност и зачитане. Външните наименования нерядко съдържат в самата си етимология обиден подтекст. Случаят с понятието „цигани“ според известните етимологии не е такъв, но в резултат от продължителната дискриминация и отхвърляне, понятието е придобило много обидни конотации. Употребата на понятието „роми“ съвсем не означава, че съответните публикации/постове са позитивни. Често се срещат словосъчетания с пейоративно значение като „братята роми“. Всъщност понятието „роми“ по-често се ползва за саркастични забележки, докато цигани се среща по-често в директни и неприкрити обиди към ромите. Думи с пряко обидно значение като м* се ползват доста по-рядко – в 2-3% от публикациите.

Както вече споменахме, преките обиди с използване на различни наименования на ромската етническа група не са непременно насочени към хора, които се самоопределят като роми. Понякога те са насочени към хора, които самите обиждащи не считат сериозно за роми, просто съответните думи се използват като обиди с общо предназначение, подобно на нецензурната реч. Те се употребяват по отношение на политически опоненти, но и към други мразени или нехаресвани лица, за да изразят чувство на превъзходство, презрение и други крайни форми на отхвърляне. Това беше причината да обособим тази форма на обидна реч в отделна категория.

Ромите и местните избори

Периодът на събиране на данните частично съвпада с предизборната кампания за местните избори през 2023 г. Затова не е учудващо, че 120 (около 10%) от материалите се занимават с темата за гласуването на местните избори, а в 76 се споменават други теми, отнасящи се до проблемите на общините, които са станали предмет на предизборната кампания. Материалите, посветени на изборите попадат в почти всички категории – от отделни предразсъдъчни коментари и реакции, до истории със сложен сюжет.

През месец септември 2023 г., в периода преди началото на предизборната кампания, темите свързани с местните избори вече заемат важно място в медиите и в обществения дебат. Влиянието на атмосферата, създадена от местните избори се чувства и в други публикации, споменаващи ромите, което не се занимават пряко нито със самите избори, нито с проблемите на общините.

В контекста на изборите ромите почти изключително се описват като лишени от политическа субектност, манипулируеми хора, които са готови да продадат гласа си. Друга част от ромите са описани като участници в електорални измами и посредници при купуването на гласове. Неутралните публикации са относително редки и се срещат в публикации на части от предизборни платформи на някои политически партии, които предлагат мерки за приобщаване на ромите. В други предизборни платформи резюмирани в социалните медии ромите се описват като заплахата, която общността, която в съвкупността си не спазва законите и срещу която трябва да се предприемат наказателни мерки. Позитивните публикации за ромите в контекста на изборите се свеждат до единични примери и се дължат на ромски активисти, които се обръщат към съграждани от своите общности с различни призови за по-голяма гражданска активност и участие в демократичния процес.

Немалко текстове са посветени на кметската надпревара в столицата, където освен в контекста на електоралните измами и социално-битовите проблеми в ромските махали, ромите присъстват както в позитивен така и в силно негативен план и в коментарите за етническия произход на единия от основните кандидати за кмет. Масивът от текстове, посветени на общинските избори като цяло съдържа многобройни примери за обидна реч спрямо ромите като група и спрямо отделни личности. Немалко са примерите и за използване на отделни наименования на ромите като директни обиди спрямо хора, които никога не са декларирали да са от ромски произход, включително такива, които никога не смята за роми.

Автоматично търсене и обработка на публикации, съдържащи предразсъдъци и омраза спрямо ромите

Алгоритмите, разработени от ЕООО изчисляват и индекс на токсичност/отровност на материалите. Понятието „токсичност“ обичайно се използва в анализа на медийно съдържание, за да се оцени като цяло степента на обидност,

интензивността на проявената омраза и потенциала на съответния материал за пораждаване на вражда и конфликт. Няма общоприето разбиране или универсален подход за измерване на токсичността на медийно съдържание. Алгоритмите на ЕООО, както и повечето съществуващи алгоритми разчитат за оценка на токсичността на списък от ключови думи, определени и оценени от хора. Този подход има сериозни ограничения тъй като участниците в социалните медии са изключително изменчиви и иновативни, използват разнообразни подходи за генериране на мултимедийно съдържание, разчитат на метафори, езикови игри и богат споделен контекст, който е трудно разпознаваем дори за по-сложни алгоритми.

Голяма част от погрешно попадналите в извадката текстове са подбрани от алгоритъма поради неразпознатата омонимия с алкохолната напитка „ром“. От друга страна въпросната омонимия е инструмент за изключително отровни обиди към ромите, както съвсем директни, така и силно завоалирани. Към директните обиди спадат коментарите на етнонима „ром“, използван в единствено число. Авторите на подобни коментари твърдят например, че става дума за „(обикновен) циганин“ докато „ром“ е наименование на алкохолна напитка. Едно от завоалираните послания, използващо тази омонимия се съдържа в поста на известна личност – общественик и музикант. Авторът твърди, че обича „отлежал ром“. Под твърдението има снимка на бутилка с ром, а под нея снимка на затворническа килия. Постът използва две снимки, които могат да бъдат изтълкувани само в съчетание и двойна игра на думи – омонимията на „ром“ и двойното значение на „отлежал“ – оставена да отлежи напитка и човек лежал в затвора. Този пост ни препраща завоалирано към системата от възгледи, която определя ромите като (по природа) склонни към престъпления, недисциплинирани и неуправляеми и в допълнение предлага наказанията като основен или единствен подход спрямо тях.

Дори безспорно позитивни съобщения могат да станат обект на коментари, заредени с омраза и предразсъдъци, които трябва да бъдат определени като силно токсични. В разглеждания корпус с особена популярност се ползват два подобни примера, които са споделени в много профили. Единият пример е за българка от ромски произход, която с високите си постижения е станала лице на престижна гимназия в чужбина. Това съобщение поражда множество естествени реакции на одобрение и поздравления с постигнатия успех. Срещат се обаче и предразсъдъчни коментари, които развиват тезата, че успехът на ромите в чужбина се дължи на строгите закони там, които принуждават иначе недисциплинираните (по природа) роми да ги спазват. Това е причината в чужбина ромите да не рискуват да не си плащат тока. Авторът на този пост косвено препоръчва използването спрямо ромите в България само на наказателни политики, защото те са единственото възможно обяснение за успеха на хора от ромски произход. Авторът на поста не привежда никакви доказателства за тази причинно-следствена връзка, не предлага и обяснение защо много други хора, които си плащат тока не са станали първенци в престижни гимназии – законопослушанието може да се нарежда сред тривиалните необходими условия, но очевидно не е достатъчно условие за такова постижение. Авторът не разглежда и възможността успехът в чужбина да се дължи на други фактори – например на липсата на дискриминация или на подходяща подкрепа.

Премахване на неприемливо съдържание от Фейсбук

В материалите, засягащи ромите, попаднали в изследвания корпус има такива, които могат да бъдат принудително премахнати или поради нецензурното съдържание с еротичен характер или поради наличието на признаци за реч на омраза.

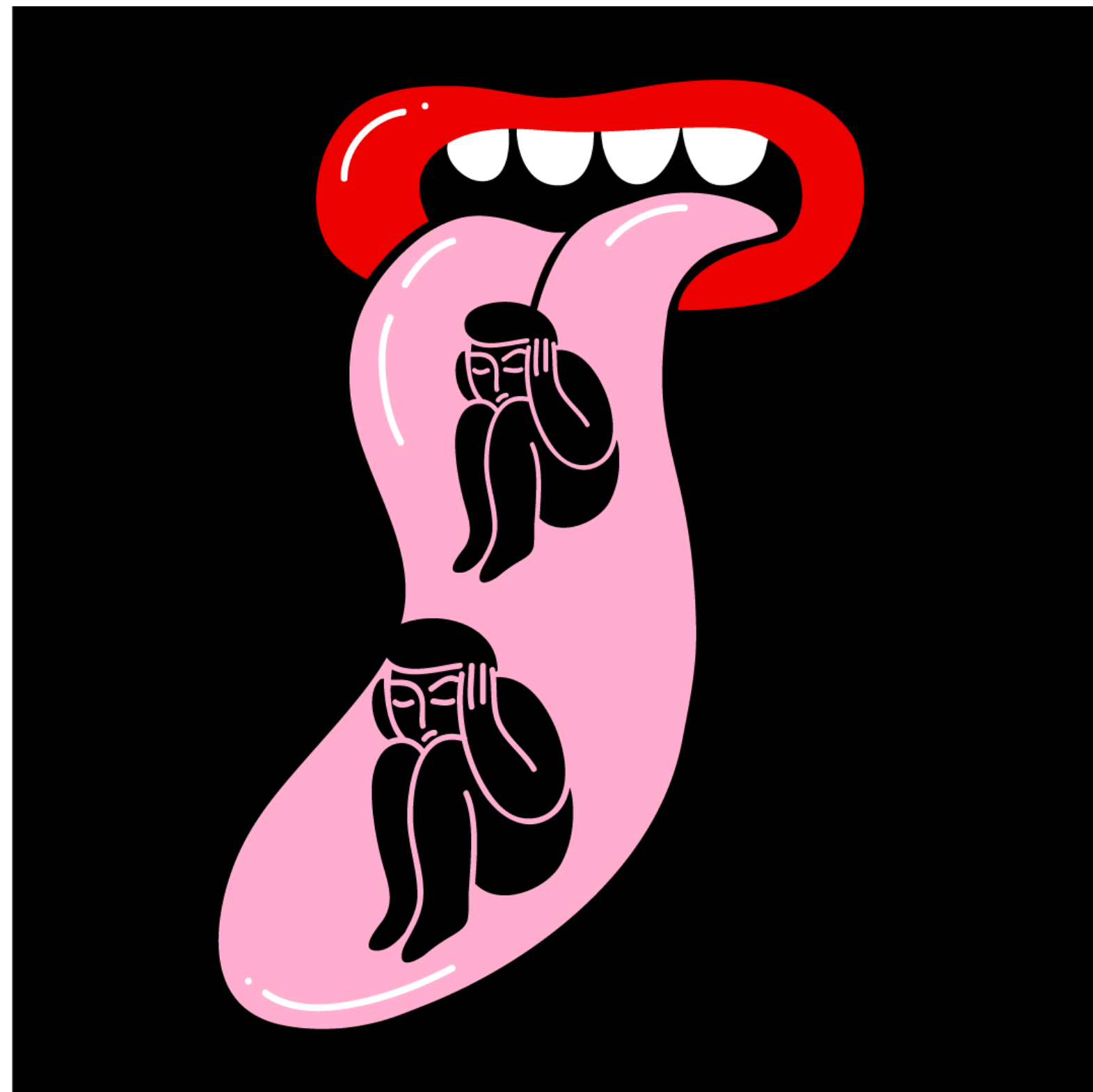
Обидните вицове, съдържащи нецензурна реч вероятно няма да попаднат сред забранените във Фейсбук материали. Политиката на Фейсбук относно съдържанието със сексуално-еротичен характер прави специално изключение за материали със сатирично съдържание, където „елементите, нарушаващи политиката относно съдържанието са със сатиричен характер или са приписвани на нещо или някого, за да бъдат осмислени или критикувани¹“. По всичко личи, че разказването на нецензурни вицове представлява изключение от материалите с откровено сексуално съдържание, които могат да бъдат премахнати. Всички вицове с участието на ромите, открити от алгоритъма през септември и октомври 2023 могат да бъдат открити на постоянните си линкове към 08.01.2024 г. Част от тези вицове биха могли да подлежат на премахване от мрежата заради съдържащите се в тях предразсъдъци и обиди спрямо ромите.

Сред останалите материали, извън вицовете и тези със сатирично съдържание определено има много такива, които според определенията, публикувани от Фейсбук би трябвало да нарушават стандартите на мрежата. Сред публикациите обаче не намерихме такива, за които със сигурност да може да се каже, че са били премахнати в резултат от нарушения на политиката за съдържание на Фейсбук. Нямаме информация дали някои от публикациите са били докладвани, но съдейки по групите, в които са публикувани това е малко вероятно, тъй като тези групи едва ли са посещавани от хора, които биха докладвали за наличието на подобни материали, освен ако не става дума за хора, които са специално ангажирани с откриването и докладването на обиди и реч на омразата, нарушаващи стандартите на Фейсбук.

Огромната част от материалите, за които са свалени постоянни линкове продължават да са достъпни към началото на м. януари 2024 г. Сред материалите, които вече не са достъпни има само няколко, които потенциално биха могли да бъдат отстранени принудително заради нарушаване на стандартите на Фейсбук за съдържание, но нямаме основание да смятаме, че това е причината да не са достъпни. Става дума за директно използване на обидни изрази на етническа основа, понякога съпроводени с нецензурна лексика. Възможно е самите потребители да са решили да ги премахнат или да променят настройките кой може да вижда съдържанието. Обичайно вместо недостъпните вече постове се появява съобщение което обяснява, че вероятно авторите им са ги направили достъпни само за определена група от потребители или са ги изтрили.

Освен това някои от въпросните материали са споделени и в други профили, от които не са били премахнати.

¹ “Adult Sexual Solicitation and Sexually Explicit Language | Transparency Center.” Retrieved January 7, 2024 (<https://transparency.fb.com/policies/community-standards/sexual-solicitation/>).



София, януари 2024 г.

Автор: Боян Захариев, Фондация "Институт Отворено общество – София

Графичен дизайн: Елица Маркова, Фондация "Институт Отворено общество – София

Снимки: AdobeStock, Canva.Илюстрация стр. 12: Nikoleta Nosovska for Fine Acts.