
МЕДИИ И КОМУНИКАЦИИ НА 21. ВЕК
21ST CENTURY MEDIA AND COMMUNICATIONS
3/2019

Богомил КАЛИНОВ

Софийски университет „Св. Климент Охридски“, София, България
b.kalinov@nvictus.eu

**МОРАЛНО НЕСЪГЛАСИЕ ОНЛАЙН. ПРОТИВОРЕЧЕНИЕТО ИЛИ ПРЕМАХВАНЕТО
НА СЪДЪРЖАНИЕ В СЛУЧАИ НА ОНЛАЙН ОМРАЗА Е ПО-ЕФЕКТИВНО?
СИСТЕМАТИЧНО ЛИТЕРАТУРНО РЕВЮ¹**

Bogomil KALINOV

Sofia University „St. Kliment Ohridski“, Sofia, Bulgaria
b.kalinov@nvictus.eu

**WHEN TRYING TO RESOLVE AGGRESSIVE CONFLICTS ONLINE,
IS IT MORE EFFECTIVE TO EMPLOY COUNTER-SPEECH,
OR IS IT MORE EFFECTIVE TO REMOVE THE AGGRESSIVE CONTENT?**

Abstract: The terms and conditions for using social platforms and the Internet grow more complex as governments struggle to provide instruments that avoid employing censorship techniques to revoke free speech, but at the same time provide online users with a safe environment and shelter from experiencing abuse or aggressive behavior. This ethical dilemma has apparently split Europe into two contradicting schools of thought. In July 2017 a search to identify relevant publications on the topic of online moral disagreement was conducted on the “Communication and Mass Media Complete” electronic database. The objective of the search was to find articles related to the topic, published in English between 2015 and 2018 with access to the full text. This systematic literature review is the final product of the materials found.

Key words: online moral disagreement, counter-speech, hate speech, systematic review.

1. Увод

Условията за използване на социалните платформи стават все по-сложни, тъй като правителствата се борят да предоставят инструменти, които избягват цензуриращи техники за отмяна на свободата на словото, но в същото време осигуряват подслон и безопасна среда за онлайн потребителите от злоупотреба, обиди или агресивно поведение. Тази етична дилема очевидно е разделила Европа в две противоречиви морални школи. През 2017 г. норвежкият министър на децата, равенството и социалното включване – Солвейг Хорне, бе домакин на „Северна конференция за речта на омразата“². Във въстъпителното си изказване тя изрази своята подкрепата за младите хора, които искат да се обръщат към речта на омразата и по този начин насърчи метода на контрареч³ и диалог (Horne 2017). Този похват е подкрепен и от ЮНЕСКО, които публикуват наръчник за борба с речта на омразата (Gagliardoneq, Gal, Alves, Martinez, 2015). Междувременно Европейската комисия обяви,

¹ Настоящата статия е част от проекта „КоМ: Инсист“, финансиран от Фонд „Научни изследвания“ по договор № КП-06-М25/3.

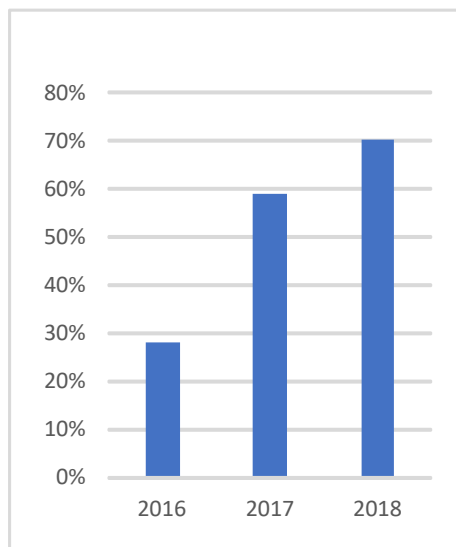
² Nordic Conference on Hate Speech.

³ Counter-speech.

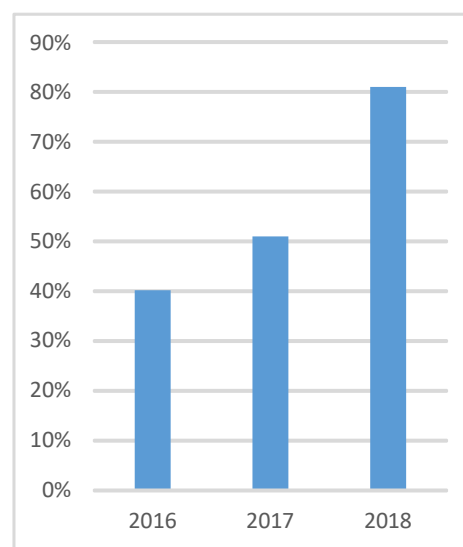
че една година след приемането на „Кодекса за поведение“⁴ – инициативата, одобрена от четирите основни социални медийни платформи (Facebook, Twitter, YouTube и Microsoft) за противодействието на незаконната реч и омраза в интернет, компаниите успешно са укрепили своите системи за докладване на агресия. Те са улеснили сигнализирането за обидни публикации, което е довело до повишаване на броя преразгледани съобщения в рамките на 24 часа и са удвоили случаите с премахването на съдържанието (Countering online hate speech 2017; Countering illegal hate speech online 2018). Подобно поведение избягва взаимодействията в полза на комфорта и липсата на спорове.

Резултатите от прилагането на инициативата „Кодекс за поведение“

% Известията, прегледани в рамките на 24 часа



Отговорено с премахване на съдържанието



Фиг. 1. Резултатите от прилагането на инициативата „Кодекс за поведение“

Източник: European Commission

Съществуването на тези двата подхода повдига въпроса коя техника е по-ефективна за борба с речта на омразата⁵? – **Известията, прегледани в рамките на 24 часа или Отговорено с премахване на съдържанието.**

От една страна, има непреодолими аргументи, че обсъждането е от съществено значение за демократичния процес, защото позволява на всеки да участва в колективното вземане на решения, вместо да ограничава тази дискусия до избрана група от хора (Masullo Chen 2017: 2). В допълнение проучвания сочат, че чувствителността у модерните публики се е повишила в сравнение с предходни години и поради тази причина потребителите са склонни да възприемат по-широк набор от съдържание като обидно за тях (Калинов 2017). Тези идеи пораждаат съмнение дали данните на Комисията отразяват положително въздействие за решаване на проблема с речта на омразата, или инициативата е имала разрушителен ефект върху социалния диалог и единственият ѝ принос е да скрие речта на омраза далеч от общественото ползрение, без действително да се бори с нея? В същото време обаче има проучвания, които предлагат тезата, че е възможно противоречието да не е винаги ефективно и някои типове контрареч биха могли дори да бъдат контрапродуктивни (Bartlett 2015: 5).

Въпросът как да се противопоставим на речта на омразата, ще се оказва все по-труден за решаване с напредване на времето поради постоянното нарастване броя на интернет потребителите⁶, повишаването на достъпността до мобилна интернет връзка и новите устройства (Cindy 2015), както

⁴ Code of Conduct.

⁵ Hate speech.

⁶ Виж Internet World Stats.

и включването на по-млади аудитории⁷, които фактори ще усложняват разговора занапред. Но въпреки това тези са същите аргументи защо трябва своевременно да се въведе надеждно регулативно законодателство за решаване на проблема. Изглежда, че Комисията възнамерява да продължи с прилагането на „Кодекса за поведение“, който достига до тревожните 70% на премахнато съдържание от всички сигнали. Ако бъдат разпознати вградени недостатъци в толкова ранен етап на прилагането на инициативата, ще бъде необходимо въвеждането на нови инструменти, които да коригират въпросните пропуски, преди да са настъпили опустошителни социални последици.

2. Методология (процедура при търсенето на литература)

През 2018 г. в електронната база данни „Комуникация и пълномащабни средства за масово осведомяване“⁸ беше проведено търсене на подходящи публикации по темата за морално несъгласие онлайн. Целта на търсенето бе да се намерят релевантни статии, публикувани на английски език между 2015 г. и 2018 г., с достъп до пълен текст. Бяха избрани четиринадесет различни комбинации от ключови думи, в резултат на които беше изграден набор от общо двеста тридесет и една статии. Петдесет и две от тях бяха отстранени поради дублиране в заглавията. Сто шестдесет и седем бяха премахнати, тъй като не са свързани с изследователския въпрос. Останалите дванадесет статии бяха използвани за текущия преглед на литературата.

3. Състояние на научните изследвания

В своя наръчник „Контриране на онлайн речта на омразата“ ЮНЕСКО накратко обобщава международните стандарти за определения за видовете онлайн реч на омразата и как те могат да бъдат преборени:

- За омразата, свързана с расата, Международната конвенция за премахване на всички форми на расова дискриминация изисква забрана за изразяване на идеи за превъзходство или малоченност на група хората, категоризирани от „раса“.
- За омразата, основана на националност или религия, се води като криминална проява по смисъла на член 20 от Международния пакт за граждански и политически права (МПГПП) – но с квалификацията, че единствено изразите, които използват застъпничество, което представлява подбуждане към дискриминация, враждебност или насилие, могат да бъдат сведени под този знаменател.
- За омразата, основана на пола на хората, сексуалната ориентация или други подобни характеристики, наръчникът допуска, че може да съществува и би могла да бъде ограничена от гледна точка на МПГПП (член 19) в интерес на зачитането на правата или репутацията на другите, но не е поставено като задължително условие (Gagliardoneq, Gal, Alves, Martinez, 2015: 6).

От представените определения става ясно, че рамката е непълна в зададените стандарти и начините за решаване на подобни казуси. Тези въпросни липси са логично свързани към същността на моралното несъгласие. Когато разглеждаме проблема за речта на омразата онлайн и начините за борба с нея, сме поставени пред две основни препятствия. Макар съдебно-лингвистичните изследвания за заплахата и насилствен език постепенно да се увеличават, липсват изследвания за лингвистичната агресия в интернет (Hardaker 2015: 1). Това твърдение се доказва и от ниския брой резултати при търсенето на литература по темата в пълномащабна база данни. Подобна липса в знанията по отношение на една толкова важна тема като контраречта и онлайн агресията разкрива колко небрежно е общественото отношение към назряващия проблем на интернет комуникацията. Вследствие на подобно negliжиране адекватното разбиране и разрешаване на проблема са невъзможни поради недостиг на актуална информация и ограничен опит в конкретното поле. Въпреки липсата на източници по темата изследването се възползва от свързани тематични проучвания, като: поведение онлайн; емоционална компетентност; взаимодействие с конфликти; демократично обсъждане; групова динамика; социална и онлайн цензура, за които са провеждани завиден брой изследвания.

Второто предизвикателство е представено при прецизното оценяване на качеството на събитие, базирано на намеренията на другите участници в него. Определенията на това, което се счита за

⁷ Виж Child Trends DataBank (2015)

⁸ Communication and Mass Media Complete.

обидно, силно зависят от социално-политическия контекст (Hintz 2016: 5), който постоянно се мени и еволюира. В допълнение, тъй като не може да знаем намеренията на друга личност, се създава една средна област, с която всеки провокатор би могъл да злоупотреби – „неизвестна зона“, в която може той да внушава искреност, докато възнамерява злодеяние; и тъй като е невъзможно да се „докаже“ какво наистина е предназначено, винаги остава място за съмнение и липса на доверие (Hardaker 2015: 4). Това е основен аспект на всяко изследване, което се занимава с онлайн морално несъгласие. Ефектът на Гигес, или с други думи, начините, по които анонимността може да повлияе на поведението (власт, която предоставя такава корумпираща сила, че дори и най-моралният не може да устои) (Hardaker 2015: 3), се превръща в задълбочен проблем, който трябва винаги да се взема предвид, но в същото време изглежда, че не може да бъде измерен. Този парадокс формулира същността на дебата между контраречта и премахването на съдържанието. Наистина ли се борим с онлайн агресията, когато използваме контрареч, или просто обгрижваме нуждата на един интернет трол и дали наистина ограничаваме трола от възможността да тормози другите потребители, или потискаме свободата на своеволен потребител?

4. Премахване на съдържание

В случаите, в които се прилага премахване на съдържание, един от основните аргументи срещу действието е, че лишава обществения диалог от жизненоважен елемент на демократичното общество – шанс за несъгласие. За да бъдат политически, решенията се вземат чрез несъгласие и убеждаване (Arendt в Masullo 2017: 2), но тези възможности отсъстват при отстраненото съдържание. При липса на несъгласие не е необходимо аргументи да се предлагат или обмислят (Esterling, Fung, Lee, 2015: 2). Като се има предвид това, изглежда, че съществува и положителна страна, която често е пренебрегната.

4.1. Позитивни страни

Пространството произвежда комуникация и комуникацията произвежда пространство (Mishra 2015: 5), така че, когато виртуалното пространство е изградено, създателите имат това, което Хинц нарича „социална отговорност“, за да определят границите на комуникацията, която ще се използва в него. Освен това Чандлър твърди, че *надстройката* (вярванията на групата) може да повлияе на *основата* (отношенията между индивидите) и обратното (Chandler 2018). От това става ясно, че представянето на набор от правила и ограничения при общуването може да подобри комуникацията в рамките на общността. Ако разговорът не може да се осъществи в рамките на установените правила, то споменатото несъгласие може да бъде по-разрушително, отколкото конструктивно и следователно трябва да бъде премахнато изцяло от платформата за взаимна изгода на всички участници. Чрез прилагането на правилата за поведение общността придобива чувство за ред и дисциплина (Mishra 2015: 6). В допълнение предходните изследвания показват, че индивидите не са добре пригодени да се справят с разногласията, които непременно съпътстват демократичното обсъждане (Esterling 2015: 15). Според Берковиц (1983) конфликтът неизбежно включва проявление на несъвместима дейност, която с течение на времето засилва враждебността, а моделът на Инфанте⁹ подкрепя тези констатации, като уточнява, че хората, които не могат да се аргументират, прибегват до съобщения, които атакуват отсрещната страна (Aloia, Solomon 2015: 17). Всъщност няма гаранции, че взетите решения (чрез обсъждане) ще бъдат добре информирани, мъдри или справедливи (Maia 2007: 1). След като се вземат предвид всички изброени фактори, възниква въпросът дали премахването на съдържанието е наистина вредно за процеса на обсъждане, или е от полза чрез прилагане на ред и следователно норми за поведение? За да отговорим на този въпрос, трябва да разгледаме другия край на спектъра.

4.2. Негативни страни

Изглежда, че има много критики по отношение на премахването на съдържание, които се простират отвъд ограничението или избягването на разногласия.

Основното притеснение е свързано с факта, че много от ключовите компоненти на интернет са създадени „без много държавен или друг надзор“ (Cerf в Hintz 2004), поради този факт самото

⁹ Infante's Argumentativeness Skill Deficiency Model.

задаване на правила в общност, формирана в така нареченото беззаконово измерение, не само противоречи на себе си, но и на самата природа на средата, в която то е създадено.

Въпреки че е обществено прието в обсъжданията за гражданството и демократичен идеал да присъства задължително оценка за качеството на информацията, циркулираща в едно общество (Tewksbury, Riles 2015: 1), изследванията показват, че селективното предлагане на информация води до екстремни нагласи и поляризация на мнението (Song 2017 и Tewksbury 2015). Информацията, която прелива в моралните, религиозните или политическите ограничения, поставени от правителствата, е все по-често блокирана (Hintz 2016: 4) и степента на „кръстосано“¹⁰ излагане на информация се превръща в основна грижа за учените по комуникация (Song 2017: 2). Поради това е важно да се насърчи участието на всички членове на групата, за да се поддържа приобщаващият характер на интернет като комуникационна платформа (Uysal 2016: 14), вместо да бъде ограничаван. Хинц определя социалните медии и други цифрови платформи като ключови места, в които напрежението между свободната комуникация и възникващата реалност на ограничаването и цензурата се разиграва (Hintz 2016: 12). Той изразява своята загриженост по отношение на промяната в управлението, която възлага на частните посредници по-ключова роля в прилагането и формулирането на правила и разпоредби и в изкушението да използват тези правила за по-широк кръг от скрити цели, като допълнително подкрепя своята теория с приложни примери. Тези доказателства подсказват, че съществува смущаващ модел на злоупотреба с регулаторната система, която надскача тона на разговора, и изглежда подобно поведение бързо се разпространява и нараства. Хинц предполага, че такива широки тълкувания са възможни поради достатъчно неясните регламенти, които са така оформени, че да могат да включват всякакъв брой политически или икономически проблеми, чието състояние може да доведе до преход от право на изразяване на мнение към рамките на често променящия се и непрозрачен режим на кодекса за поведение (Hintz 2016: 6). Въпреки че разпространението на новините в интернет е довело до драматично увеличаване броя на опциите за съдържание (Tewksbury 2015: 3) чрез предлагането на информационни и нематериални блага и превръщането им в пазари, социалните платформи определят като ключов икономически ресурс контрола върху идеите и знанията, като по този начин засягат съдържанието, достъпно в социалните медии и самото използване на тези платформи (Hintz 2016: 7). В допълнение Хинц твърди, че социалните медии са обвързани от собствената си търговска логика и политически убеждения, а не от граждански права и върховенство на закона.

Освен политическото въздействие трябва да се вземе предвид и ефектът върху индивида. Личният стил често се основава на груповия, тъй като културните очаквания формират фона, върху който се развиват личните стилове (Chioricki 2017: 1), а експерименталните проучвания показват, че хората изпитват заплахата и емоционална болка, ако са изключени от взаимодействията (Masullo Chen 2017: 3). Нивото на сравнение на алтернативите предполага, че мотивацията на дадено лице да се присъедини, или да напусне групата, се влияе от стойността на други алтернативни групи (Forsyth 2006), така че чрез ограничаване на комуникацията между индивида и групата не само може да бъде вредно, но и пагубно за индивида до точката, в която той може да чувства, че сегашната му социална група е единствената му възможност и ще бъде демотивиран да се ангажира с по-нататъшна комуникация или да промени мнението си.

4.3. Контрареч

За демократичното управление институциите не могат да функционират без несъгласие и те не разрешават несъгласието с никакво правило или формула, а вместо това делиберационният идеал предизвиква хората да слушат други, които не споделят техните убеждения, ценности или интереси, и да се ангажират в конструктивен диалог (Esterling 2015: 1). Такава идея може да се представи като по-демократична в сравнение с алтернативата за премахване на съдържание, но и тя има своя цена.

4.3.1. Позитивни страни

Емпиричните доказателства до голяма степен съответстват на твърдението, че комуникационните умения предотвратяват ескалацията на конфликти (Aloia, Solomon 2015: 4). Проучванията показват, че участниците изпитват по-голямо удовлетворение от обсъждането под умерена идеологическа разлика, отколкото когато са в хомогенни или силно разнородни групи и ако бъдат помолени

да обсъдят с други, които не споделят мненията им, участниците биха могли да намерят обогатяване и удовлетворение от изживяването, което удовлетворение от процеса и резултатите от обсъждането има немонотонна връзка с несъгласието (Esterling 2015: 15). Друго изследване показва, че няма значителни различия между група, изложена на нецивилизовани коментари, и група, която не е, и твърди, че коментарите в неприлично несъгласие няма да доведат до по-голям негативен ефект от разговора. Тези констатации показват, че нецивилизованите и цивилизованите коментари не се различават по въздействието им върху агресивни намерения или отрицателни емоции (Masullo Chen 2017: 14). Умереното несъгласие предизвиква оптимално обмисляне в съответствие с нормативните очаквания и емпирично демонстрира съвещателното качество на това събитие (Esterling 2015: 1). Всички изброени доказателства обрисуват един много благоприятен портрет на тази техника, докато не разгледаме контрааргументите.

4.3.2. Негативни страни

Идеалът на обсъждането изисква гражданите да участват в разумни дискусии въпреки разногласията си (Esterling 2015: 1), което изглежда като нереалистична ситуация. Изследването, което предлага, че нецивилизованите и цивилизованите коментари не се различават по своите последици, също така заявява, че и двата вида несъгласие могат да доведат до определени нива на негативни емоции и агресивни намерения и може да имат смразяващ ефект върху обществения дискурс, жизненоважен за делиберационната демокрация (Masullo Chen 2017: 14). Моделът на дефицит на аргументирани умения заключава, че лицата, които нямат способност да се включат в ефективни аргументи, вероятно ще прибегнат до словесна агресия (Aloia, Solomon 2015: 1). В този контекст, в който не може да сме напълно сигурни в способностите за дискусия на всички участници, насърчаването за дебатиране е равнозначно на насърчаване на конфликтите. От своя страна, агресивни устни забележки ескалират конфликти и увеличават вероятността от физическо насилие (Infante, Chandler, Rudd 1989 в Aloia 2015). Според теорията на афектираната интелигентност срещането на разногласия може да предизвика и негативни емоции (Parson 2010). Това е два пъти по-валидно в случаите, в които системата за наблюдение е подразнена от нова или заплашителна информация, която се различава от убежденията на субекта (Masullo Chen 2017: 2). Като алтернатива участниците в дискусията могат да отговорят, търсейки разногласия, и по този начин да целят конфликт заради самата идея за пререкание, с предпочитание към словесен сблъсък пред възможността за конструктивен диалог, а с твърде много разногласие дори добрите аргументи попадат в глухи уши (Esterling 2015: 2). В допълнение контраречта може да предразположи към взаимно подсилващи се процеси на спирален модел, без да се разглеждат междуличностните фактори за дискусия, които биха довели до потенциално погрешни или поне непълни интерпретации. Интерперсоналното дискусийно мрежово (не)съгласие може значително да усилва влиянието на поляризацията на отношението (Song 2017: 14). Трябва да се вземе под внимание и фактът, че хората с висока емоционална компетентност са по-уверени в себе си и по-правилно се справят с конфликтите (Rahim, Psenicka 2005 в Aloia), но въпреки това всеки може и ще се включи в обратна реч независимо от своя стил на общуване (Chiopicki 2017), емоционална компетентност (Aloia 2015) или когнитивни способности. Един активен участник в противоречие, у когото липсват или недостигат едно или повече от споменатите комуникационни умения, би могъл да има пагубен ефект над диалога. В такъв случай премахването на съдържание би било много по-полезно за социума, отколкото допускането да се проведе подобна „вредна“ дискусия.

4.4. Умереност

След като и двата похвата за борба с речта на омразата бяха внимателно разгледани, става ясно, че няма един напълно верен отговор, който да предлага абсолютно решение. Следователно както контраречта, така и премахването на съдържание трябва да се прилагат умерено, за да се осигури най-добрата възможна среда за комуникация и гражданско несъгласие. Но остава въпросът как да определим кога да приложим едната техника над другата, ако не можем да оценим намеренията на участниците? Изследванията показват, че характерът на предишни постъпки служи като един от най-силните показатели за бъдещото отношение (Krosnick, Petty 1995 в Song: 7). Следвайки тази

¹⁰ Cross-cutting.

логика, става ясно, че ако намеренията не могат да бъдат оценени, а поведението е по-скоро константна величина, оказваща влияние върху комуникацията, то може да бъде използвано като по-ефективна единица при измерване на качеството за бъдещи дискусии. Вече има някои модели на поведение, които оценяват способността на човека да реагира и да се справя с конфликтите и несъгласията въз основа на личните си черти.

- **Модел на дефицит на аргументативни умения** – уменията за общуване определят дали субектите предпочитат да предотвратят враждебност по време на конфликтни взаимодействия, или да ескалират ситуацията (Aloia 2015).
- **Модел на комуникационните стилове** – група от аспекти на поведението на разговорен език, които колективно определят модела на културната комуникация (Chiopicki 2017).
- **Измерение на контекстния модел** – класифицира поведението на континуума в зависимост от степента на тяхната зависимост от контекста (Chiopicki 2017).

Чрез възприемането на един от гореспоменатите или подобни модели лесно може да се определи дали разговорът или несъгласието ще бъдат цивилизовани и конструктивни, или не, което предлага възможност да бъдат взети необходимите действия за коригиране на поведението спрямо ситуацията и участниците в нея.

5. Дискусия

Настоящото изследване, подобно на всяко друго такова, подлежи на свои собствени ограничения. Поради недостига на ресурси като време и финансиране то разглежда единствено набор от публикации в една база данни. Липсата на предишни проучвания по темата също се отрази на обхвата и броя на използваните материали. Заслужава да се отбележи обаче, че независимо от размера на извадката изследването осигурява актуален систематичен преглед по разгледаният въпрос в избраната база данни. Документът идентифицира предизвикателствата, свързани с моралното несъгласие в интернет по структуриран начин, който предлага методи и модели за противодействие при случаи на онлайн омраза. Той също така спомага за развитието на не толкова популярна тема, като същевременно осигурява нови отговори на това, което, изглежда, е многопластов проблем. Предложените насрещни мерки могат да бъдат тествани в реални времеви условия и ако са успешни, могат да бъдат приложени в практиката. Както е очевидно, тяхното прилагане би довело до подобрене в регламентите за онлайн комуникациите и социалните платформи.

Създадена е нова изследователска основа за анализ на ефективността на предложените модели в контекста на борбата с речта на омразата или алтернативно допълнителни изследвания могат да бъдат разширени, за да включват нови бази данни, за да се осигури по-пълнен систематичен преглед по изследователския въпрос.

6. Заключение

След сравнителен анализ на две различни техники за борба с онлайн агресията въз основа на предишни изследвания статията излага както силните, така и слабите страни на двата подхода и задава най-ефективния начин за противодействие на онлайн агресията според наличните изследвания по темата до момента. Никоя отделна техника не се оказа напълно устойчива, за да бъде достатъчна сама по себе си, което предполага необходимостта от измерване на модерацията.

ЛИТЕРАТУРА // BIBLIOGRAPHY

Калинов 2018: Калинов, Б. Опасността в безопасното интернет пространство. Дигиталната революция в културните и социалните процеси, 364. Благоевград, България: Югозападен университет „Неофит Рилски“. // **Kalinov 2018:** Kalinov, B. Opasnostta v bezopasnoto internet prostranstvo. Digitalnata revolyutsia v kulturnite i sotsialnite protsesi, 364. Blagoevgrad, Bulgaria: Southwest University “Neofit Rilski”.

Aloia, Solomon 2015: Aloia, L. S., Solomon, D. H. The Physiology of Argumentative Skill Deficiency: Cognitive Ability, Emotional Competence, Communication Qualities, and Responses to Conflict. Communication Monographs, 82(3), 315–338. doi:10.1080/03637751.2014.989868.

- Arendt 1958:** Arendt, H. *The human condition*. Chicago, IL: University of Chicago Press.
- Berkowitz 1983:** Berkowitz, L. The experience of anger as a parallel process in the display of impulsive, “angry” aggression. In R. G. Geen & E. I. Donnerstein (Eds.), *Aggression: Theoretical and empirical reviews* (Vol. 1, pp. 103–133). New York, NY: Academic Press.
- Bartlett, Jones 2015, October:** Bartlett, J., A. K., Jones. Counter-speech examining content that challenges extremism online [PDF]. London: Demos. <https://www.demos.co.uk/wp-content/uploads/2015/10/Counter-speech-1.pdf>. (16 май 2018)
- Cerf 2004:** Cerf, V. G. First, do no harm. In D. MacLean (Eds.), *Internet governance: A grand collaboration* (pp. 13–15). New York, NY: United Nations ICT Task Force.
- Chandler 2000:** Chandler, D. *Marxist Media Theory*. U.K:
- Child Trends DataBank 2015:** Child Trends DataBank. Home computer access and internet use. or <https://www.childtrends.org/?indicators=home-computer-access>. (16 май 2018)
- Chłopicki 2017:** Chłopicki, W. *Communication Styles – An Overview*. Styles Of Communication. Kraków, Poland: Jagiellonian University.
- Cindy 2015, August:** Cindy, L. [Worldwide Internet and Mobile Users. eMarketer’s Updated Estimates for 2015]. https://insights.ap.org/uploads/images/eMarketer_Estimates_2015.pdf. (16 май 2018)
- Countering illegal hate speech online** – Commission initiative shows continued improvement, further platforms join. (2018, January 19). European Commission [Press release]. http://europa.eu/rapid/press-release_IP-18-261_en.htm. (16 май 2018)
- Countering online hate speech** – Commission initiative with social media platforms and civil society shows progress. (2017, June 1). European Commission [Press release]. http://europa.eu/rapid/press-release_IP-17-1471_en.htm. (16 май 2018)
- Deibert 2009:** Deibert, R. The geopolitics of Internet control: Censorship, sovereignty, and cyberspace. In A. Chadwick & P. Howard (Eds.), *The Routledge handbook of Internet politics* (pp. 323–336). London: Routledge.
- Esterling, Fung, Lee 2015:** Esterling, K. M., Fung, A., Lee, T. How Much Disagreement is Good for Democratic Deliberation?. *Political Communication*, 32(4), 529-551. doi:10.1080/10584609.2014.969466
- Forsyth 2014:** Forsyth, D. R. *Group dynamics* (6th ed.). Belmont, CA: Wadsworth.
- Gagliardoneq, Gal, Alves Martinez 2015:** Gagliardoneq, I., Gal, D., Alves, T., Martinez, G. *Countering online hate speech*. UNESCO, ISBN 978-92-3-100105-5.
- Hardaker 2015:** Hardaker, C. ‘I refuse to respond to this obvious troll’: an overview of responses to (perceived) trolling. *Corpora*, 10(2), 201–229. doi:10.3366/cor.2015.0074.
- Horne 2017, June 22:** Horne, S. Opening Speech. In *Speech: Countering Hate – Nordic Conference on Hate Speech*. <https://www.regjeringen.no/en/aktuelt/speech-countering-hate---nordic-conference-on-hate-speech/id2558269/>. (16 май 2018)
- Hintz 2016:** Hintz, A. Restricting digital sites of dissent: commercial social media and free expression. *Critical Discourse Studies*, 13(3), 325-340. doi:10.1080/17405904.2016.1141695.
- Infante, Chandler, Rudd 1989:** Infante, D. A., Chandler, T. A., Rudd, J. E. Test of an argumentative skill deficiency model of interspousal violence. *Communication Monographs*, 56, 163–177. doi:10.1080/03637758909390257.
- Infante 1987:** Infante, D. A. Aggressiveness. In J. C. McCroskey and J. A. Daly (Eds.), *Personality and interpersonal communication* (pp. 157–192). Newbury Park, CA: Sage.
- Internet World Stats 2018:** [INTERNET USAGE STATISTICS. The Internet Big Picture]. <https://www.internetworldstats.com/stats.htm>. (16 май 2018)
- Krosnick, Petty 1995:** Krosnick, J. A., Petty, R. E. Attitude strength: An overview. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 1–24). Hillsdale, NJ: Erlbaum.
- Maia 2007:** Maia, R. C. Deliberative Democracy and Public Sphere Typology. *Estudos Em Comunicazgo*. <http://ec.ubi.pt/ec/01/pdfs/maia-rousiley-deliberative-democracy.pdf>. (16 май 2018)
- Masullo Chen, Lu 2017:** Masullo Chen, G., Lu, S. Online Political Discourse: Exploring Differences in Effects of Civil and Uncivil Disagreement in News Website Comments. *Journal Of Broadcasting & Electronic Media*, 61(1), 108–125. doi:10.1080/08838151.2016.1273922.
- Mishra 2015:** Mishra, N. Space, communication and spatial violence: pace, communication and spatial violence: situating the mechanism of spatial exclusion and inclusion. *Journal For Communication & Culture*, 4(1), 71–85.
- Parsons 2010:** Parsons, B. M. Social networks and the affective impact of political disagreement. *Political Behavior*, 32, 181–204. Springer US: doi.org/10.1007/s11109-009-9100-6
- Rahim, Psenicka 2005:** Rahim, R. M., Psenicka, C. Relationship between emotional intelligence and effectiveness of leader role: A dyadic study in four countries. *International Journal of Organizational Analysis*, 13, 327–342. doi:10.1108/eb029010.
- Song, Boomgaarden 2017:** Song, H., Boomgaarden, H. G. Dynamic Spirals Put to Test: An Agent-Based Model of Reinforcing Spirals Between Selective Exposure, Interpersonal Networks, and Attitude Polarization. *Journal Of Communication*, 67(2), 256–281. doi:10.1111/jcom.12288.

Tewksbury, Riles 2015: Tewksbury, D., Riles, J. M. Polarization as a Function of Citizen Predispositions and Exposure to News on the Internet. *Journal Of Broadcasting & Electronic Media*, 59(3), 381-398. doi:10.1080/08838151.2015.1054996.

Uysal 2016: Uysal, N. Social Collaboration in Intranets: The Impact of Social Exchange and Group Norms on Internal Communication. *International Journal Of Business Communication*, 53(2), 181–199. doi:10.1177/2329488415627270