

Marko Poljak, Jelena Hadžić i Maša Martinić

Alma Mater Europea, Hrvatska
marko.poljak@diktiranje.hr

Fakultet hrvatskih studija Sveučilišta u Zagrebu, Presscut d.o.o., Hrvatska
Jelena.hadzic@presscut.hr

NEWTON Technologies Adria, Hrvatska
masa.martinic@diktiranje.hr

Govor mržnje u hrvatskom medijskom prostoru

Sažetak

Autori u svom radu razmatraju aktualni problem zakonski reguliranog neprihvatljivog ponašanja u Hrvatskom medijskom prostoru, uključujući i društvene mreže, koncentrirajući se na problem govora mržnje. Temeljno je pitanje je li prisutnost govora mržnje u porastu. Primjenom metode analize sadržaja, a na obuhvatu kompletnog medijskog prostora Hrvatske, autori zaključuju da je porast prisutnosti neprihvatljivog ponašanja i govora mržnje hrvatska svakodnevnica. Osnovom dobivenih rezultata istraživački napori autora usmjeravaju se prema detekciji vrsta govora mržnje i na njihove primjere pojavnosti pri čemu detekcija pojavnosti neprihvatljivog govora a i govora mržnje pripada vrsti usmjerenoj prema etnicitetu, religiji i izbjeglicama. U primjeni detekcije, obrade i analize su računalne tehnologije na čijem razvoju i testiranju rezultata autori nastavljaju istraživanje. Treći istraživački napor autora usmjeren je prema događajima koji su prisutni u medijskom prostoru, a odnose se na neprihvatljivi govor, te kvalitativnom i kvantitativnom analizom sadržaja daju kontekstualni okvir najvažnijih događaja.

Ključne riječi: govor mržnje, društvene mreže, medijski sadržaj, računalne tehnologije.

UVOD U PROBLEM ISTRAŽIVANJA

U fokus interesa znanstvene i stručne javnosti dolazi tematika govora mržnje, lažnih vijesti i dezinformacija. Dijelom je to produkt namjere EU da se spriječi pojava i širenje istih.

Predmet našeg rada je govor mržnje. Pretpostavljali smo da se u hrvatskom medijskom prostoru najčešće pojavljuje govor mržnje usmjeren prema etnicitetu, migrantima i religiji pa smo, shodno tome, prikupili recentne teorijske radove na ove teme.

PREGLED RADOVA O GOVORU MRŽNJE

Interesirala su nas temeljna dva aspekta i to teorijska polazišta radova i metodološki i računalni aspekti načina obrade teme. Da se poslužimo sintagmom Aslana (Aslan, A. 2016), govor mržnje je postojao i prije nego je stekao digitalnu dimenziju. S digitalnom dimenzijom stekla se mogućnost da gotovo svatko preko društvenih mreža može biti nositelj a umreženost mu uz sveprisutnost daje potencijalnu moć. Iz tog razloga smo pored klasičnih medija masovne komunikacije u naš istraživački korpus uvrstili i društvene mreže. Naizgled logičan pristup jer se na društvenim mrežama najviše pojavljuju oblici neprihvatljivog ponašanja poput uvrede, klevete i govora mržnje ali se pritom, čini nam se, neopravdano izostavljaju mainstream mediji (klasični mediji masovnog komuniciranja), za koje se u pravilu ne može očekivati prisutnost govora mržnje i poticanje na nasilje zbog čvrsto postavljenih uredničkih pravila i kodeksa ali su važan faktor u davanju značaja i osude svih oblika neprihvatljivog ponašanja.

Temeljna karakteristika prikupljenih recentnih teorijskih radova je **inter i multidisciplinarnost u pristupu proučavanja** te izbor analitičkog, istraživačkog korpusa sa društvenih mreža.

Carney (2014) u svome radu analizira i naglašava **lingvističku** važnost prilikom zakonske analize govora mržnje na primjeru sudskog slučaja iz Južnoafričke republike. Carney polemizira s presudom suda po kojoj se percepcija slušatelja mora smatrati jedinom valjanom u donošenju presude. Autor smatra da je sud u procesu trebao uključiti forenzičkog lingvistu jer razloge problematičnosti oslanjanja samo na percepciju slušatelja odnosno primatelja nalazi u principu pragmatičnosti – nije važno što kažete, već kako; ton govora i govor tijela; u kulturalnoj pozadini izgovorenog; u komunikacijskim vještinama; te u nesporazumu u razumijevanju netaktično prezentiranih osjećaja i stavova.

Ostali autori koji su dotakli tematiku govora mržnje slažu se po pitanju naglaska na lingvistički element u navedenom govoru, poput Ousidhoum i drugih (2019) koji prezentiraju višejezičnu bazu podataka o govoru mržnje vezano uz engleske, francuske i arapske komentare na Twitteru koji prikazuju potencijalnu problematiku ovog velikog projekta – prvenstveno lingvističku stranu priče. Također, Baumgarten i drugi (2019) prikupili su i analizirali podatke sa Facebooka i Twittera te komparirali lingvističke instancije govora mržnje u danskom i njemačkom jeziku. Zatim, Sap i drugi (2019) prikazuju kako neosjetljivost antagonista na razlike u dijalektu može dovesti do rasne

pristranosti u modelima automatskog otkrivanja govora mržnje, što potencijalno povećava štetu za manjinsko stanovništvo. Analiza afroameričkog dijalekta pokazuje kako se navedeni dijalekt nalazi u dva puta većoj šansi da bude proglašen kao uvredljiv za razliku od drugih. Kada su sugovornici eksplicitno upozoreni kako se radi o Afroamerikancima, značajno pada mogućnost karakteriziranja *tvitova* kao uvredljivih.

Većina radova ističe nepreciznost definicije govora mržnje i razlika istog u odnosu na uvredljiv jezik i/ili klevetu.

Tako na primjer autori Malmasi i Zampieri (2017) predlažu slijedeću kategorizaciju:

a) komentar sadrži govor mržnje (HATE), b) komentar sadrži uvredljiv jezik, ali ne i govor mržnje (OFFENSIVE), c) komentar uopće ne sadrži uvredljiv jezik (OK) i napominju kako se bez detaljnijih lingvističkih analiza ne može lako (s velikom točnošću) utvrditi što je točno govor mržnje. U svome radu navode definicije „bulinga“ (zastrašivanje određenog pojedinca) i „govora mržnje“ (omalovažavanje grupe ljudi) zaključujući kako je govor mržnje sam po sebi nejasan pojam.

Kao primjer retoričko komunikološkog pristupa obradi predmetne tematike ističemo rad Arofaha (2018) koji ističe da govor mržnje kao žanr retorike uvijek sadrži tri elementa: etos, patos i logos (na primjeru Indonezije, odnosno kandidaturi za guvernera Jakarte). Ovo istraživanje pokazuje da je govor mržnje uglavnom zanemario aspekt etosa i logosa. Radije naglašava i oslanja se na patos aspekt koji izaziva osjećaje publike.

Bosco i drugi (2017) bave se alatima i resursima za detektiranje mržnje i predrasuda prema imigrantima na društvenim mrežama. Pristup slijedi **holističku i višejezičnu perspektivu**, koja uključuje različite izvore znanja radi smislenijeg prikupljanja podataka. Premda je studija usmjerena na pokrajinu Piemont u Italiji, može se upotrijebiti i za druge (veće) lokalitete i jezike. Također, talijanskim slučajem bave se i Comandini i Patti (2019) koji analiziraju ulogu populističkih tema i retorika na talijanskom Twitteru vezano uz govor mržnje usmjeren na imigrante. Istom tematikom samo za nešto sjeverniju državu bavi se i Hrdina (2016) koja navodi rezultate istraživanja govora mržnje u Češkoj, 2015. godine, koji pokazuju kako je val govora mržnje prema migrantima pooštren i namjerno i slučajno sjedinjenim snagama različitih korisnika Facebooka, propagande ekstremističkih skupina, novinskih medija i dizajna same društvene mreže. Što se tiče društvene pozadine učestalih proizvođača govora mržnje, postojala je snažna prevalencija muškaraca srednje dobi i srednje klase i značajno podzastupljenost i starijih i mladih korisnika Facebooka. **Komunikacijske aktivnosti** većih, dobro organiziranih **populističkih skupina**, političkih stranaka ili zajednica bile su vidljivo prisutne, ali nisu imale značajnu ulogu u samoj proizvodnji govora mržnje - iako se njihovo moguće sudjelovanje u postavljanju dnevnog reda ne može podcijeniti. Svi skupovi podataka pokazali su da je velika većina govora mržnje u navedenom vremenskom razdoblju bila usmjerena bilo protiv migranata općenito ili protiv muslimana, dok su se te dvije skupine često preklapale. Istim podnebljem bavio se i Pejchal (2018) analizom izbjegličke krize iz 2015. godine na primjeru Češke i Slovačke. U nedostatku univerzalne definicije govora mržnje, poticanje na mržnju tumači na tri

različita, ali komplementarna načina: poticanje na nasilje; poticanje na diskriminaciju; i poticanje na uskraćivanje ljudskog dostojanstva. Autori Jaki i De Smedt (2018) analiziraju govor mržnje njemačkih desničara na Twitteru te naglašavaju nove izazove, poput točnog definiranja što je to slobodni govor i što je nelegalno u specifičnoj državi te saznanja o tome koje su točno lingvističke karakteristike govora mržnje.

Autori Laaksonen i ostali (2020) su na razvijenoj tehničkoj infrastrukturi za automatsko nadgledanje izjava izbornih kandidata na društvenim mrežama zaključili da je računalno potpomognut pristup zahtijevao izrazitu simplifikaciju jer se model oslanjao na pristup temeljen na riječima dok je realnost suprotna. Govor mržnje je, zaključuju autori lingvistički i društveni fenomen s različitim tonovima i formama. Određene poteškoće kod analize govora mržnje navode i Enarsson i Lindgren (2018) koji su prikupili materijal iz švedskih tvitova o romskom stanovništvu kako bi mapirali različite diskurse. Složenost i činjenica da bilo koji tvit ili drugi izrazi na mreži moraju biti promatrani i procjenjivani pojedinačno kako bi se mogao uzeti u obzir okolni kontekst i diskurs, doprinose teškoći davanja jasnih smjernica o tome koji su izrazi dopušteni, a koji ne. To je izazov za sudove i agencije za provođenje zakona koji su suočeni sa suprotstavljanjem govoru mržnje.

Pretežito politološko - sociološki radovi odnose se većinom na raspravu od **odnosu slobode govora i zakonske regulative zabrane govora mržnje**.

Baider (2019) navodi kako u međunarodnom zakonu o ljudskim pravima ne postoji formalna definicija govora mržnje, ali EU temelji svoje preporuke za suočavanje s govorom mržnje (2008) na sljedećoj definiciji: „svi oblici i izrazi koji javno potiču na nasilje ili mržnju a usmjereni su protiv skupine ljudi ili člana takve skupine definirani s obzirom na rasu, boju, religiju, nacionalno ili etničko podrijetlo“¹¹⁵. Autor istražuje stav prema diskriminatornom govoru među uzorkom mladih ciparskih Grka i ima za cilj identificirati faktore koji bi mogli proizvesti stavove mržnje. Spominje „teške varijable“ koje sadrže klasu, etničku pripadnost, spol, dob, vjeru i izbjeglički status. Ovi parametri moduliraju vjerovanja i stavove, a to je vrlo teško promijeniti jer su unutarne varijable. „Mekše“ i više „iskustvene varijable“ uključuju više društvenih i iskustvenih parametara poput obrazovnih iskustava, građanske uključenosti i kontakta i izloženosti kulturnim „drugima“ u obrazovnim programima, kao i dijeljenja traumatičnih iskustava rata i nasilja. Promicanje ovih iskustava i njihovo dijeljenje može stvoriti otvaranje „nacionalnog načina razmišljanja“, a što može donijeti promjenu stava i možda čak i promjenu vjerovanja (o „drugom“).

Kazneni zakon Republike Hrvatske¹¹⁶ definira sljedeće elemente ključne za definiranje govora mržnje: poticanje na nasilje ili mržnju, nasilje ili mržnja su usmjereni na skupine ljudi ili na pojedince temeljem pripadnosti nekoj skupini; poticanje na nasilje ili mržnju je javno.

¹¹⁵ Citirana je Okvirna odluka o suzbijanju određenih oblika i načina izražavanja rasizma i ksenofobije kazneno - pravnim sredstvima, 2008/913/PUP koju je 2008. godine donijelo Vijeće EU.

¹¹⁶ Kazneni zakon RH, članak 325, pročišćeni tekst, na snazi od 1. 1. 2020. <https://www.zakon.hr/z/98/Kazneni-zakon>.

Demaske (2019) u radu predlaže omanju promjenu paradigme vezanu uz slobodu govora, odnosno govor mržnje - pomak naglašavanja s negativne zaštite prava pojedinca prema pozitivnijoj potpori socijalne jednakosti. Koristeći teoriju prepoznavanja Axela Honnetha (kao mogući način premošćivanja jaza između današnje regulacije i razvoja strategije slobodnog govora za budućnost), autor razvija okvir za analizu slobode govora koji bi promovirao strategiju za borbu protiv govora mržnje u dvadeset prvom stoljeću. Navodeći razne primjere kroz sudske procese te naglašavajući postojanje američkog Prvog amandmana o slobodi govora, autor smatra nužnim promjenu bavljenja problemom „govora mržnje“. Tradicionalistički pristup branjenja slobode govora, smatra autor, nije dovoljno fleksibilan za moderno društvo. Ako je cjelokupni razlog zaštite govora promicanje i poticanje slobodnog demokratskog društva, tada se zaštita koja je zakonski propisana mora razmotriti u odnosu na stvarne ishode. Autor predlaže da se umjesto prethodnih pravnih doktrina primijenjenih u slučajevima govora mržnje uspostavlja novi analitički okvir koji će omogućiti sudovima da pronađu određeno zakonodavstvo protiv govora mržnje kako bi izdržalo kontrolu Prvog amandmana.

Gould (2019) navodi kako je najuvjerljivije argumente protiv zabrane mržnje iznio Ronald Dworkin, koji je upozorio na prijetnju političkom legitimitetu koji postavlja zakone koji onima koji im podliježu uskraćuje odgovarajuću priliku za neslaganje. Autor navodi kako se Dworkinovoj argumentaciji suprotstavlja Jeremy Waldron, koji smatra kako se oponenti zakona o govoru mržnje bave problemom tako generalizirano da nisu uspjeli identificirati specifični izazov demokratskom legitimitetu koji predstavlja taj govor mržnje. Reid (2019) kritizira verziju argumentacije kako regulacija govora mržnje od strane države narušava njezin demokratski legitimitet, što zastupaju Ronald Dworkin i James Weinstein. Autor smatra kako u nekim slučajevima štetni učinci govora mržnje na demokratski proces nadmašuju ograničenja. Ovaj rad svrstao se uglavnom na stranu onih koji tvrde kako se ograničenjem govora mržnje zapravo može povećati demokratski legitimitet. Na sličnom tragu je i rad autora Cohen - Almagor (2019) koji analizira govor mržnje u sklopu kritike Thomasa Scanlona i njegove teorije ugovornosti kojom promiče “zaštitu interesa” različitih dionika. Scanlon govorniku pripisuje snažnu zaštitu, ponekad na štetu zaštite osnovnih prava ciljane skupine kojoj je govor namijenjen. Autor smatra kako je navedeni liberalni stav previše otvoren za razvoj govora mržnje. U demokraciji ljudi moraju uživati apsolutnu slobodu zagovaranja i rasprave ideja, ali to je toliko dugo dok se suzdržavaju od zlouporabe te slobode da bi napali prava drugih ili njihov status u društvu kao ljudskih bića i ravnopravnih članova zajednice. Autor smatra kako se demokracija temelji na dva osnovna načela: poštivanje drugih, i ne nanošenje štete drugima. Sloboda izražavanja treba se uravnotežiti s ne manje važnim načelima, poput zaštite ranjivih manjina. Sličnog mišljenja je i Mondal (2018) koji navodi kako liberalna teorija slobodnog govora konceptualizira slobodu kao glatku, kontinuiranu, homogenu, nedjeljivu i proširivu bez prekida dok ne dosegne vanjske granice. Autor predstavlja alternativnu „tekuću“ teorije slobodnog govora koja objašnjava „oblikovanje“ slobode od strane društvenih snaga, kulture i institucionalnih praksi. Također, O`Hehir (2019), pozivajući se na Johna Milтона i Johna Stuarta Milla, naglašava kako je

sloboda izražavanja važna za razvoj i održavanje svakog demokratskog društva. Mišljenja je kako pod krinkom prava na slobodu izražavanja ekstremna desnica koristi situaciju, što je negativno utjecalo na ljudska prava izbjeglica u Europi. Stoga se može tvrditi da sloboda govora gubi moć uzvratanja tlačitelju, najvažnijeg elementa ovog prava.

Chandrasekharan i drugi (2017) prezentiraju efikasnost rješavanja govora mržnje na platformi Reddit u 2015. godini. Sukladno Redditoj politici protiv uznemiravanja, platforma je zatvorila, odnosno ugasila svoje dvije podplatforme. Zabrana može umanjiti ponašanje mržnje ili može premjestiti takvo ponašanje u različite dijelove web mjesta. Autori smatraju kako je zabrana bila uspješna za Reddit. Više korisničkih računa nego što se očekivalo prekinulo je korištenje web stranice; oni koji su ostali drastično su smanjili svoju upotrebu mržnje - za najmanje 80 %. Premda su mnoge podplatforme doživjele priljev „migranata“ s ugašene dvije podplatforme, nove lokacije nisu doživjele značajne promjene u govoru mržnje. Drugim riječima, druge podplatforme nisu naslijedile navedeni problem. Korisnici koji su bili aktivni na ugašene dvije podplatforme drastično su smanjili korištenje govora mržnje. Iako su filozofska pitanja koja se tiču umjerenosti (i posebno zabrane) složena, autori su nastojali razviti raspravu vezanu uz rezultate o učinkovitosti zabrane devijantnih skupina mržnje s internetskih platformi. S druge strane, autori Carlson i Rouselle (2020) na primjeru Facebooka prikazuju kako se jedna druga društvena mreža odnosi prema govoru mržnje. Autori smatraju kako su Facebookova pravila o ponašanju na toj društvenoj mreži iz 2018. godine više načelna nego provedena u praksi te su mišljenja kako je Facebooku pri zaobilaženju govora mržnje na mreži važnija financijska motivacija. Autori Doring i Mohseni (2019) na primjeru još jedne platforme - YouTubea, pri analizi komentara i videa vezanih uz tzv. „fail video“, prikazuju kako su žene znatno više nego muškarci prikazane kao objektivizirane i seksularizirane odnosno kako su žene sklonije biti mete govoru mržnje. Istom temom bavili su se i autori Fuchs i Schafer (2020) koji predstavljaju rezultate istraživačke analize slučajeva mizoginističkog ili seksističkog govora mržnje i jezika zlostavljanja prema političarkama na Twitteru, primjenjujući računske korpuslingvističke alate i metode, dopunjene kvalitativnom dubinskom studijom verbalnog zlostavljanja četiri istaknute japanske političarke.

METODE ISTRAŽIVANJA I PRIMJENA RAČUNALNIH METODA – PREGLED RADOVA

Većina radova upotrebljava analizu društvenih mreža (SNA), (Blanco - Herrero, D. i Calderon, C., 2019) proučavaju kako se sadržaji šire u društvenim medijima; a eksperimentalnom anketom promatraju kako se ti sadržaji primaju i u kojoj mjeri im građani vjeruju. Namjera autora je prikazati cijeli komunikacijski proces od - produkcije, prijenosa i primanja poruka

Cilj projekta koji predstavljaju autori (Bosco, C. i Patti, V. i Bogetti, M. i Conoscenti, M. i Ruffo, G. i Schifanella, R. i Stranisci, M., 2017) je izgradnja interaktivnog mrežnog alata za prikupljanje, analizu i vizualizaciju diskursa o imigrantima i reakcija zajednice. Autori upotrebljavaju znanja računalne lingvistike, semantičke i diskurzivne analize.

Lazaridis i Veikou (2016), autori koje smo već spomenuli, provodili su klasičnu analizu sadržaja nakon koje je primijenjen strukturirani intervju sa nevladinim udrugama i pučkim pravobraniteljem, što je primjer kombinacije kvantitativnog i kvalitativnog istraživanja.

Autori Blanco-Herrero i Calderon (2019) prezentiraju rad vezan uz španjolski nacionalni kontekst, kombinirajući lažne vijesti i govor mržnje prema migrantima i izbjeglicama. Navedene dvije teme predstavljaju veliku prijetnju za društvo u cjelini, a posebno za migrante i pojedince koji traže azil. U sklopu projekta u kojem sudjeluju najavljuju provedbu **eksperimenta** u kojemu namjeravaju uzorak građana izložiti sadržajima o migrantima i izbjeglicama prikupljenim s društvenih medija i tražiti povezanost između sadržaja u koji vjeruju i faktora poput njihove političke ideologije, stava prema imigraciji, spola, dobi ili stupnja obrazovanja.

Florio i ostali (2019) **kombinirajući analizu statističkih podataka i analizu sadržaja** navode kako nekoliko studija ističe opasnost društvenih medija kao polazišta za govor mržnje (ili cyberhate). Pronašli su korelacije koje sugeriraju međusobnu povezanost ekonomskih i kulturnih čimbenika i izražavanja mržnje na internetu.

Također, Baumgarten i drugi (2019) prikupili su i analizirali podatke sa Facebooka i Twittera te **komparirali lingvističke instancije govora mržnje u danskom i njemačkom jeziku.**

Multidisciplinarni teorijski pristupi ukazuju na pretežito zastupljenu kognitivnu lingvistiku, komunikološku perspektivu te sociološko - politološku dimenziju radova na temu govora mržnje.

U metodološkom smislu najzastupljenija je analiza sadržaja potpomognuta računalnim alatima jer je prikupljanje i selekcija podataka ogromnih baza podataka nemoguća bez računalnih alata.

U pogledu računalne obrade podataka najzastupljenije je područje obrade prirodnog jezika (natural language processing) koja je disciplina na razmeđu računarske znanosti i inženjerstva, lingvistike i statistike a primjenjuje metode strojnog učenja (machine learning) i umjetnih neuronskih mreža (artificial neural networks).

U znanstvenoj zajednici koja se bavi obradom prirodnog jezika okušane metode koje se mogu rabiti za rješavanje podzadataka detekcije govora mržnje uključuju detekciju stava (engl. stance detection), zaključivanje u prirodnom jeziku (engl. natural language inference) i izračun semantičke sličnosti tekstova (engl. semantic textual similarity). Osim nabrojanih, ovdje su također i prepoznavanje i klasifikacija imenovanih entiteta (engl. named entity recognition and classification), klasifikacija tema (engl. text categorization) te ekstrakcija izvora i citata (engl. source extraction, quote extraction).

Alati kognitivnog prepoznavanja neverbalnih komunikacijskih procesa uz prepoznavanje emocionalnog sentimenta verbalnog i neverbalnog izričaja omogućiti će brže prikupljanje i selekciju ogromne količine sadržaja s društvenih mreža.

Shodno izloženom, mi smo se u svom radu opredijelili na komunikološku perspektivu izučavanja govora mržnje u hrvatskim uvjetima i to u segmentima etniciteta, migranata i religije primjenjujući NLP tehnologije u prikupljanju, selekciji i djelomičnoj obradi podataka prikupljenih sa društvenih mreža.

PROBLEM I CILJEVI ISTRAŽIVANJA

Temeljni problem ovog istraživanja je govor mržnje u medijskom prostoru Republike Hrvatske. Pritom se pod medijskim prostorom podrazumijevaju internetski portali, tiskovine, radijske i televizijske informativne emisije i društvene mreže.

DEFINIRANJE TEMELJNIH POJMOVA

Već pri samoj definiciji pojma 'govor mržnje' nailazimo na mnoge probleme. Prvi problem je taj što ne postoji jednoznačna definicija pojma. Drugi problem prilikom definiranja pojma govora mržnje je primjenjivost definicije. Pregled problematike govora mržnje prikazan je u pregledu literature.

Pod pojmom **medijska objava** podrazumijevaju se: članci tiskovina, članci news portala objavljeni na news portalu, prilozi informativnih radijskih i televizijskih emisija emitirani putem radija ili televizije.

Negativan događaj je afera, eksces, konflikt, direktno nasilje ili neki slični negativni događaji u kojem je neki pojedinac ili skupina govorio mržnju ili djelovao u skladu s postupcima govora mržnje, te su mediji prenijeli ovakav događaj karakterizirajući ga kao govor mržnje. Generalno prenošenje statistika i podataka o količini govora mržnje nije konkretan negativni događaj i nije kodirano kao negativan događaj.

Da bi izbjegli raspravu o definiranju pojma Internet medija, društvenih medija, društvenih mreža, kanala komuniciranja, mi smo se u radu opredijelili za termin **društvene mreže**.

Neprihvatljiv govor je pojam koji u sebi obuhvaća: psovke i vulgarijeme, uvrede, klevete, prijetnje i govor mržnje. Neprihvatljiv govor može biti samo jedna riječ ili cijela rečenica.

Istraživači ne ulaze u raspravu radi li se o uvredi, kleveti, govoru mržnje ili nekom drugom obliku neprihvatljivog ponašanja.

CILJEVI ISTRAŽIVANJA

(za analizu sadržaja medijskih objava tiska, radija, televizije i news portala)

Glavni cilj:

Koliko je i kakav sadržaj o govoru mržnje prisutan u hrvatskom medijskom prostoru

Sporedni ciljevi:

Kvantifikacija medijskih objava koje se odnose na govor mržnje i promjena kroz vrijeme

Ekstrakcija ključnih medijskih događaja (ili tema)

Kategorizacija i kvantifikacija medijskih objava na one koje govore o konkretnom negativnom događaju i one koje ne govore o konkretnom negativnom događaju

Istražiti u kojem se omjeru, s obzirom na pojam govora mržnje, mediji osvrću na hrvatske događaje, a koliko se odnose na događaje van Republike Hrvatske

Istražiti je li govor mržnje dopijeva u medije kao usputni spomen ili kao samostalna cjelina

(za analizu sadržaja komentara sa društvenih mreža)

Glavni cilj:

Prema kojim skupinama u komentarima s društvenih mreža ima najviše neprihvatljivog govora.

Sporedni ciljevi:

Zastupljenost komentara s neprihvatljivim sadržajem u odnosu na ukupan broj komentara i pojedinu skupinu

Koja skupina generira najviše neprihvatljivih komentara na društvenim mrežama

Upotreba jezičnih tehnologija u prikupljanju, obradi i analizi medijskog sadržaja i komentara na društvenim mrežama i ocjena upotrebljivosti tehnologija.

Prema kojoj skupini je usmjereno najviše neprihvatljivih komentara

METODOLOGIJA

U radu je primijenjena kvantitativna i kvalitativna analiza sadržaja medijskih objava (tiska, radija, televizije i news portala) i komentara sa društvenih mreža.

Baza medijskih objava u periodu od 1. 1. 2013. pa do 30. 4. 2019. sadržava 14. 086. 565 medijskih objava iz tiskovina, informativnih radijskih i televizijskih emisija i news portala.¹¹⁷ Iz baze medijskih objava su izdvojene sve one medijske objave koje u sebi sadržavaju pojam govora mržnje. Tijekom pretrage smo pazili da obuhvatimo sve morfološke oblike sintakse govora mržnje. Populacija sadrži 50. 724 medijskih objava koje u naslovu ili sadržaju sadržavaju pojam govora mržnje.¹¹⁸

Kvantitativna analiza sadržaja medijskih objava temelji se na populaciji.

Kvalitativna analiza sadržaja medijskih objava napravljena je na uzorku. U uzorak su izabrane sve objave sa najposjećenijih¹¹⁹ hrvatskih news portala: jutarnji.hr, dnevno.hr, tportal.hr, index.hr, 24sata.hr, telegram.hr, vecernji.hr. Uzorak se odnosi na period od 1. 5. 2018. do 30. 4. 2019. (n=878).

Jedinica kvalitativne analize su objave hrvatskih portala. Analiziran je tekst objave, ali ne i grafička oprema teksta (poput fotografija, infografika i slično).

Prilikom **analize sadržaja društvenih mreža** korišten je namjerni uzorak. Pretražene društvene mreže su: Youtube, Facebook, Reddit, forumi, gdje forumi obuhvaćaju komentare ispod medijskih objava news portala. Promatrani period je 03. 04. 2019. do 02. 07. 2019. U uzorak su izabrani svi komentari objavljeni pod naslovima u kojima se spominje izabrana ključna riječ. Pod naslovima se u ovom radu podrazumijeva: za forume koji su dio news portala naslovi novinarskih članaka odnosno medijskih objava, za Reddit su naslovi topics (odnosno tema rasprave), za Facebook su to objave (postovi), za Youtube su to nazivi objavljenih videa. Pretraga naslova i ekstrakcija komentara je napravljena uz pomoć alata za pretraživanje društvenih mreža¹²⁰. Izabrane ključne riječi¹²¹ se odnose na etničke i vjerske skupine te izbjeglice. Veličina uzorka je 71.914 komentara.

117 Kriteriji praćenja odnosno ulaska medijskog sadržaja u medijsku bazu objava su definirani na sljedeći način: tiskovine koje imaju nakladu 3000 ili više primjeraka, barem 500 news portala koji su najposjećeniji prema Alexa ili Gemius listi najposjećenijih portala, informativne emisije svih nacionalnih i najgledanije/najslušanije informativne emisije regionalnih, županijskih i lokalnih radijskih i televizijskih postaja sa terestrijalnim signalom.

118 Pretraga je tekstualna. Kako smo pretražili tekstualno radijske i televizijske priloge, pitate se? Sav govor zvučnih i video zapisa je uz pomoć tehnologije pretvaranja govora u tekst tekstualiziran. Sav tiskovni sadržaj prolazi kroz OCR i tako sadržaj članka iz slike pretvara u tekst. Točnost pretvaranja govora u tekst je 92 %, a točnost OCRa je 97 %, što ostavlja prostor za pogrešno uključivanje ili isključivanje objave iz populacije.

119 Najposjećeniji prema Gemiusovoj listi portala, prema prosjeku posjećenosti kroz promatrani period.

120 Alat Metricom je korišten za pretragu društvenih mreža i uključuje samo javne objave.

121 Izabrane ključne riječi su: Srbija, Srbi, srpski, Bosna, Bošnjak, Bosanac, bosanski, bošnjački, Slovenija, Slovenac, slovenski, Romi, romski, Cigan, ciganski, Albanija, Albanac, albanski, Šiptar, šiptarski, Italija, Talijan, talijanski, Židov, židovski, katolici, katolički, katoličanstvo, pravoslavlje, pravoslavac, pravoslavan, islam, musliman, muslimanski, izbjeglica, migrant. Rezultati pretrage po ključnim riječima obuhvaćaju sve morfološke oblike izabrane ključne riječi (a ne samo nominativ jednine).

Istraživači su odabrali skupine prema kojima su očekivali pojavu neprihvatljivog govora na društvenim mrežama. Od skupina koje hrvatski Kazneni zakon¹²² navodi, u ovom radu su zastupljene vjerske i etničke skupine. Skupina izbjeglice,¹²³ treća skupina zastupljena u ovom radu, ne može se svrstati u neku od navedenih kategorija već spada u „kakve druge osobine“ prema Kaznenom zakonu. Etničke skupine zastupaju: Albanci, Bosanci,¹²⁴ Romi, Slovenci, Srbi i Talijani. Vjerske skupine zastupaju: katolici, muslimani, pravoslavci i Židovi.¹²⁵

Jedinica analize sadržaja društvenih mreža je komentar. Prilikom analize komentara u obzir je uzet samo tekst, dok slike, emotikoni ili linkovi nisu analizirani.

Analiziralo se postojanje ili nepostojanje neprihvatljivog govora u komentaru. Za svaki komentar sa neprihvatljivim govorom određena je skupina na koju je neprimjeren govor bio usmjeren.

Neprihvatljiv govor, za potrebe ove analize, definiran je kao negativan izričaj kojim se vrijeđa, kleveće, izražava mržnja ili prijetnja s obzirom na pripadnost nekoj skupini ili na neku skupinu u cijelosti.

REZULTATI ISTRAŽIVANJA

PODACI O POPULACIJI MEDIJSKIH OBJAVA KOJE SADRŽAVAJU POJAM GOVORA MRŽNJE

Udio medijskih objava koji sadržavaju pojam govora mržnje se kroz godine povećava u hrvatskom medijskom prostoru

Tablica: udio medijskih objava govora mržnje prema vrsti medija¹²⁶ u odnosu na ukupan broj objava, po godinama

	2013	2014	2015	2016	2017	2018	1.4. 2019	ukupno
tisak	0,34 %	0,23 %	0,32 %	0,49 %	0,36 %	0,37 %	0,50 %	0,36 %
radio	0,03 %	0,02 %	0,03 %	0,06 %	0,05 %	0,05 %	0,11 %	0,04 %
televizija	0,11 %	0,10 %	0,16 %	0,24 %	0,20 %	0,22 %	0,44 %	0,19 %
portali	0,42 %	0,39 %	0,60 %	0,73 %	0,51 %	0,59 %	0,83 %	0,58 %
ukupno	0,22 %	0,20 %	0,33 %	0,49 %	0,35 %	0,41 %	0,60 %	0,36 %

122 Kazneni zakon RH, članak 325, stavak 1., <https://www.zakon.hr/z/98/Kazneni-zakon>.

123 Istraživači priznaju da postoje razlike u značenju riječi izbjeglica i migrant, kako u konotaciji tako i pravne. Za potrebe ovog istraživanja se nije radila razlika između migranta i izbjeglice već su ta dva pojma uzimana kao istovjetna.

124 Za potrebe ovog istraživanja nije napravljena razlika između Bosanaca i Bošnjaka i oba pojma su promatrana kao istoznačna; etničku pripadnost bošnjačkoj nacionalnosti. Jedan od razloga sjedinjavanja ovih dvaju pojmova se nalazi u samom materijalu koji je obrađivan: komentatori na društvenim mrežama ne slijede službenu podjelu pojmova već se pojmovi često brkaju i mijenjaju.

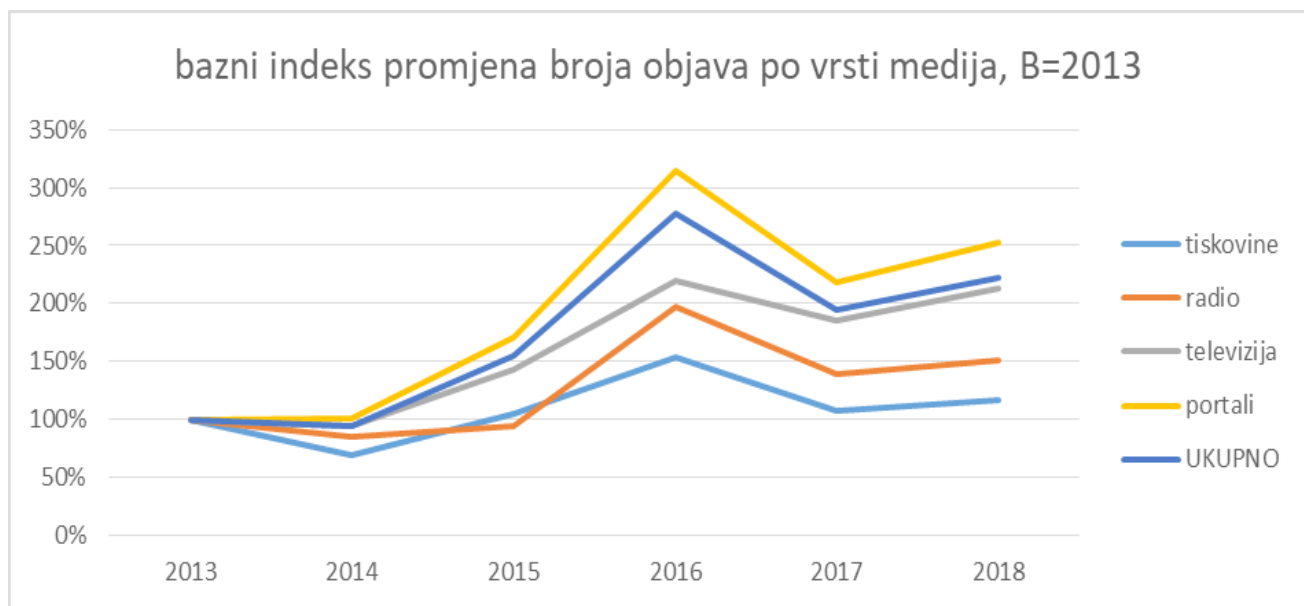
125 Istraživači priznaju da se skupinu Židovi može podjednako promatrati kao etničku ili kao vjersku skupinu, no za potrebe ovog rada je skupina Židovi grupirana u vjersku a ne etničku skupinu.

126 Vrste medija su: tiskovine, news portali, radio i televizija.

Ukupan broj medijskih objava u razdoblju od 1. 1. 2013. do 30. 4. 2019. je 14. 086. 565. Broj medijskih objava koje sadržavaju pojam govora mržnje za isto razdoblje je 50.724.

Tablica pokazuje postotak učešća medijskih objava koje u sebi sadržavaju pojam govora mržnje u odnosu na ukupan broj objava u promatranom razdoblju.

Broj objava koje sadržavaju pojam govor mržnje niti jedne godine ne prelazi 1 % ukupnog broja medijskih objava. Udio objava koje sadržavaju pojam govora mržnje, u odnosu na ukupan broj medijskih objava po godinama, se povećava.



Promatrajući zadnje razdoblje (4 mjeseca 2019. godine) u odnosu na prvo promatrano razdoblje (cijela 2013. godina) vidljivo je kako je broj objava s pojmom govora mržnje u samo 4 mjeseca 2019. godine premašio ukupan broj objava iz 2013. (omjer broja objava u 2019. u odnosu na 2013., za sve medije iznosi 118 %). Ukoliko upotrijebimo verižni indeks umjesto baznog i promatramo samo 4 mjeseca 2019. u odnosu na cjelokupnu 2018. godinu, broj objava u samo 4 mjeseca 2019. godine iznosi 53 % ukupnog broja objava iz cijele 2018. godine.

Razlozi povećanja i smanjivanja broja medijskih objava svakako ovise o događajima vezanim uz govor mržnje i objašnjenja za povećanje broja objava trebalo bi detaljnije analizirati kroz sam sadržaj koji se medijima prenosi. Tako bi se posebno velik broj objava s pojmom govora mržnje koji je pronađen 2016. godine trebalo povezati sa nekim medijski značajnim događajima koji su se zbili te godine.

Ipak, vidljivo je povećanje broja objava vezanih uz govor mržnje po godinama.

Broj dana bez medijskih objava koje sadržavaju pojam govora mržnje

Kroz promatrano razdoblje smanjuje se broj dana u kojima nije bilo niti jedne objave o govoru mržnje. Ovaj podatak je drugi način dokazivanja hipoteze da je tema govora mržnje kroz godine sve zastupljenija u medijskom prostoru.

Tablica: broj dana po mjesecu i godini u kojima nije pronađena niti jedna medijska objava s pojmom govora mržnje

godina/ mjesec	2013	2014	2015	2016	2017	2018	1-4 2019	ukupno
1	1							1
2	2	1						3
3	3							3
4		1						1
5	1	1					X	2
6	1	1					X	2
7	4	3	1				X	8
8	6			1	1		X	8
9	1						X	1
10	3				1		X	4
11							X	0
12	1					1	X	2
ukupno	23	7	1	1	2	1	0	35

Dnevni prosjek broja medijskih objava sa pojmom govora mržnje raste

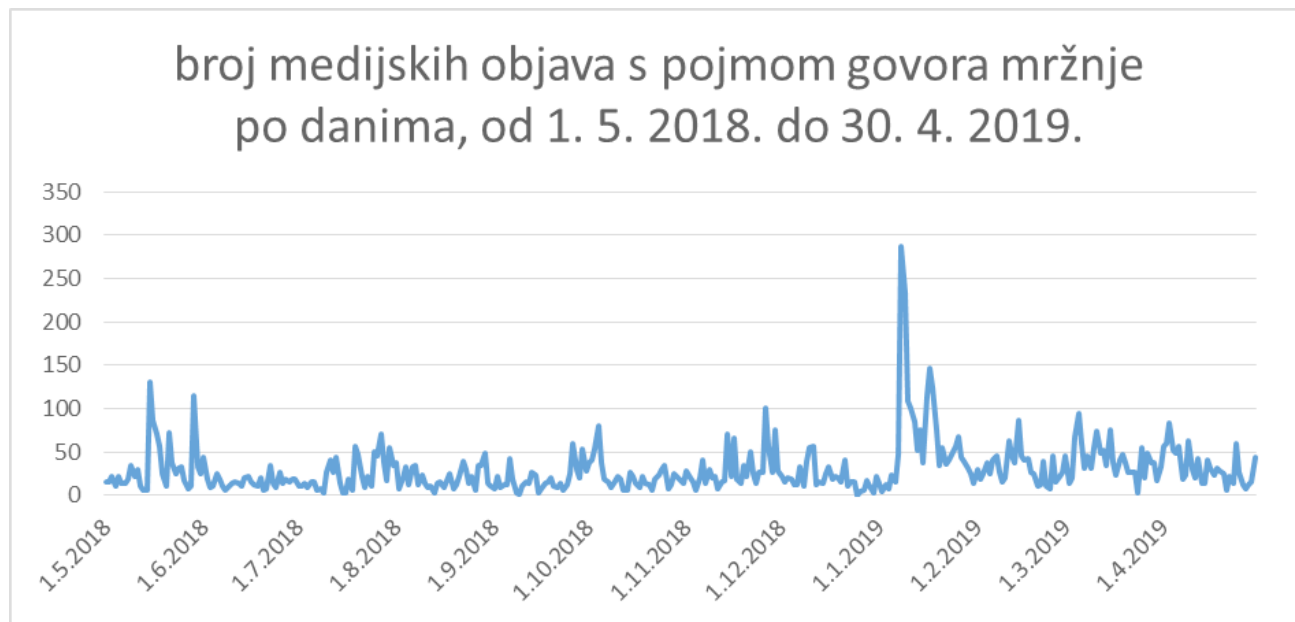
U promatranom periodu je 2311 dana, prosjek dnevnog broja objava za sve vrste medija je 21,9 objava sa standardnom devijacijom od 13,7 objava. Radi jednostavnosti, prikazane su tablice i grafovi gdje je dnevni prosjek broja objava dan na mjesečnoj razini.

prosječan dnevni broj medijskih objava							
mjesec/godina	2013	2014	2015	2016	2017	2018	2019
1	16,0	13,8	32,3	60,8	17,4	56,3	66,2
2	6,6	11,0	26,6	47,8	38,3	30,7	31,6
3	26,0	9,8	16,0	69,2	38,1	20,1	40,6
4	16,2	9,9	13,6	37,7	18,3	28,8	32,5
5	10,3	11,4	23,1	31,7	27,2	33,0	
6	7,2	5,6	22,3	18,8	19,6	16,2	
7	3,1	5,4	10,0	23,7	15,6	24,1	
8	2,6	5,3	14,1	20,5	19,1	20,0	
9	5,7	8,1	14,0	23,9	24,8	17,9	
10	8,8	7,2	16,1	19,5	15,2	21,9	
11	24,7	29,9	12,7	23,7	17,4	31,0	
12	16,1	18,2	21,3	19,9	29,0	19,6	
godišnji prosjek	12,0	11,3	18,5	33,1	23,3	26,6	43,1



Jedini dan bez govora mržnje je Božić

Od ukupno 365 dana u razdoblju od 1. 5. 2018. pa do 30. 4. 2019, samo jedan dan nije bilo objava na news portalima koji bi u sebi sadržali riječi govora mržnje, i taj dan je 25. 12. 2018., Božić. U prosjeku je objavljivano 29,6 objava dnevno, sa standardnom devijacijom od 28,5 objava po danu. 9. 1. 2019. objavljeno je najviše objava, čak 287. U tih 365 dana ukupno je pronađeno 10. 769 medijskih objava s pojmom govora mržnje, od čega 84,5 % objava su pronađene na news portalima.



Osim 2018. godine, u promatranom razdoblju samo je Božić 2013. bio bez objava koje su u sebi uključivale pojam govora mržnje.

PODACI O UZORKU MEDIJSKIH OBJAVA NEWS PORTALA

Pregled ključnih događaja koji se odnose na govor mržnje

Podatak o ključnim događajima vezanim uz govor mržnje temelji se na tehnologiji. Upotrijebljen je računalni softver¹²⁷ koji grupira medijske objave u događaje obzirom na sličnost sadržaja (i neke druge kriterije). Prikaz ključnih događaja dan je za razdoblje: 1. 5. 2018 – 30. 4. 2019. Kao kriterij za važnost pojedinog događaja uzet je broj medijskih objava. Drugim riječima, ključni medijski događaj je onaj koji je brojem objava izazvao najveću medijsku pozornost. Izdvojeni su oni događaji s preko 100 objava. Događaji su uzeti iz populacije medijskih objava, gdje populaciju čine sve one objave tiska, radija, televizije ili news portala koje u sebi sadržavaju pojam govora mržnje.

¹²⁷ Računalni softver upotrijebljen za generiranje događaja se zove event detection.

Naziv ili opis događaja (teme), preuzet iz medijskog sadržaja	Datum prve objave	Broj medijskih objava
Ivan Đakić, sin HDZ-ovog saborskog zastupnika Josipa Đakića čestitao pravoslavni Božić slikom na kojoj neki ustaša u ruci drži glavu mrtvog četnika. Sliku je objavio uz popratni tekst u kojem piše: „Svim prijateljima ‘srbicima’ sretan božić”.	6. 1. 2019.	523
Ivan Penava, gradonačelnik Vukovara, prozvan za izazivanje “spirale govora mržnje” te se skoro potukao u Saboru	15. 1. 2019.	235
Milorad Pupovac, predsjednik SDSSa: ‘Spaljivanje krnje s mojim likom, to je govor mržnje’	4. 3. 2019.	155
Izvešće Europske komisije o porastu govora mržnje u Hrvatskoj	14. 5. 2018.	154
Osuđivanje srpske političarke Vjerice Radeta za tweet povodom smrti Hatidže Mehmedović, predsjednice Udruženja “Majke Srebrenice”. Tweet glasi: “Čitam umrla Hatidža Mehmedović iz udruženja biznismenki Srebrenice. Ko li će je sahraniti. Muž ili sinovi?!”	24. 7. 2018.	130
Pokolj u džamiji na Novom Zelandu	14. 3. 2019.	113
Svim televizijama koje su emitirale (ili prenosile) emisiju Bujica na temu migranata ukinuta koncesija na 24 sata	22. 11. 2018.	107
Nenad Stazić, saborski zastupnik SDP-a, komentirao na Facebooku: “Izgleda da u svibnju 1945. posao nije obavljen temeljito. Kakva šlampavost pobjednika”	27. 5. 2018.	102

Promatrani medijski sadržaj se ne odnosi samo na negativne događaje proglašene govorom mržnje

Prva asocijacija na pojam govora mržnje su afere, ekscesi, konflikti, direktno nasilje ili neki slični negativni događaji u kojima je neki pojedinac ili skupina govorio mržnju ili djelovao u skladu sa postupcima govora mržnje, te su mediji prenijeli ovakav događaj karakterizirajući ga kao govor mržnje. Rezultati istraživanja ukazuje na pojavnost pojma govora mržnje unutar medijskih objava koje se ne tiču nekog konkretnog negativnog događaja.

Od ukupno 878 analiziranih objava iz uzorka, 357 ili 40,6 % se ne odnose na medijsko prenošenje konkretnog negativnog događaja karakteriziranog kao govor mržnje.

Postojanje konkretnog negativnog događaja	Zastupljenost govora mržnje u sadržaju objave		
	Glavna tema	Sporedna tema	ukupno
Da	371	150	521
Ne	74	283	357
ukupno	445	433	878

Od ukupno 445 medijskih objava koje se punim sadržajem odnose na govor mržnje, svega 16,6% objava nisu medijsko prenošenje nekog ekscesa ili negativnog događaja. Svaka čast tim iznimkama, ali ovi podaci pokazuju kako je medijsko tematiziranje pojma govora mržnje uglavnom vezano za prenošenje ekscesa, nasilja, optužbi za govor mržnje i drugih negativnih događaja.

Od 74 medijske objave news portala iz uzorka koje se punim sadržajem odnose na govor mržnje ali istovremeno ne govore o nekom konkretnom negativnom događaju, istraživači su sastavili pregled tema ili događaja. Pregled je rađen uz pomoć ljudske (a ne strojne) klasifikacije. Izabrani su oni događaji koji imaju puno objava ili oni događaji koji su svojim sadržajem vrlo različiti od ostatka medijskog sadržaja iz uzorka.

Opis ili naziv događaja (preuzet iz članka)	mjesec i godina događaja	društveno područje
Izvešće Europske komisije protiv rasizma i nesnošljivosti	5/2018.	europska politika
Motovunski filmski festival prikazao dokumentarni film 'Srbenka' u produkciji Restarta – rad koji prati pripreme i probe Frlićeve predstave i tematizira ksenofobiju	7/2018.	kultura
Potpredsjednik Hrvatskoga sabora, Željko Reiner, govori na misi u Ludbregu i poziva na rad, ljubav i toleranciju	9/2018.	politika, religija
Međunarodna konferencija u Zagrebu pod nazivom "Regulacija govora mržnje u medijima – uloga regulatornih tijela i pravosuđa"	11/2018.	znanost
Tribina u Zagrebu pod nazivom "Zašto se mrzimo?"	11 i 12/2018.	znanost
Mišljenja poznatih Hrvatica o negativnim komentarima na društvenim mrežama	12/2018.	estrada
Istraživanje Fakulteta političkih znanosti o zabrani ustaških i komunističkih simbola	1/2019.	znanost
Ministarstvo kulture predlaže novi zakon o medijima kojim će se regulirati nedopušteno ponašanje na internetu	1/2019.	politika i zakoni
Biskupska konferencija osudila govor mržnje i totalitarne ideologije	2/2019.	religija
2. Dani medijske pismenosti - rasprava i o govoru mržnje na društvenim mrežama	4/2019.	znanost mediji
Studenti u studentskom domu na Savi objesili zastavu duginih boja kako bi ukazali na diskriminaciju i govor mržnje	6/2018.	studenti
Okrugli stol o govoru mržnje u Hrvatskoj održan na Fakultetu političkih znanosti u Zagrebu	12/2018.	znanost
Engleski niželigaški nogometni klub je igrao utakmicu u dresovima s duginim bojama, kao izraz podrške LGBT zajednici	2/2019.	nogomet

Facebook nakratko uvodi mogućnost da svaki korisnik označi <i>post</i> kao govor mržnje	5/2018.	društvene mreže
Facebook objavljuje podatke o količini uklonjenog sadržaja, jedna od kategorija je i govor mržnje	5/2018.	statistike o govoru mržnje, društvene mreže
Facebook ne uklanja objave u kojima se negira holokaust, dok takav sadržaj ne prijeđe granicu i postane nasilan ili mrzilački	7/2018.	društvene mreže
Europska komisija najavljuje stroža pravila za regulaciju sadržaja na društvenim mrežama	8/2018.	društvene mreže, europsko zakonodavstvo
Reguliranje i samoreguliranje Facebooka	8/2018. 3 i 4/2019.	društvene mreže
Donald Trump nezadovoljan rezultatima pretrage (o sebi) na Googleu	8/2018.	
“Nije bilo govora mržnje ili netrpeljivosti” na dočeku nogometne reprezentacije u Zagrebu, izjavljuje predsjednica Kolinda Grabar Kitarović	7/2018.	nogomet
Pismo čitateljice u kojem se pita zašto moramo vrijeđati jedni druge umjesto da uživamo u plasmanu na nogometnom prvenstvu	7/2018.	pisma čitatelja
Tumblr ažurirao službena pravila ponašanja te time nanio izravni udar govoru mržnje	9/2018.	društvene mreže
U Europi je na djelu “stalni porast” ksenofobnog populizma i govora mržnje, upozorila je stručna skupina Vijeća Europe	6/2018.	europska politika
Europski parlament raspravljao o jačanju ekstremizma	10/2018.	europska politika
Europska komisija je zadovoljna provedbom Kodeksa postupanja za borbu protiv nezakonitog govora mržnje na internetu	2/2019.	europska politika
Francuski predsjednik Emmanuel Macron mađarskom premijeru Viktoru Orbanu i ministru unutarnjih poslova Matteou Salviniju “Ništa neću prepustiti nacionalistima i onima koji se zalažu za govor mržnje”	8/2018.	europska politika

Iz popisa tema/događaja razvidno je kako inicijativa za poboljšanje stanja vezano uz govor mržnje dolazi iz znanstvenih, političkih i civilnih krugova, gdje se kao nositelji civilnih inicijativa pojavljuju studenti, (pisma) čitatelja i nogometni klub.

Osim inicijativa, ovakve medijske objave zorno prikazuju medijsko prihvaćanje promicanja javne svijesti o problematici govora mržnje.

Medijski sadržaj koji potiče javnu svijest o problemu govora mržnje ne odnosi se samo na Hrvatsku. Od ukupno 74 objave koje se punim sadržajem odnose na govor mržnje i koje se ne odnose na neki konkretni negativni događaj, njih 33 (44,5 %) se ne odnose na događaje unutar teritorija Hrvatske.

Sadržaj hrvatskih portala o govoru mržnje uglavnom se odnosi na govor mržnje unutar teritorija Hrvatske.

Raspodjela medijskih objava obzirom na zastupljenost govora mržnje u sadržaju i postojanje negativnog događaja, za objave koje se odnose na događaje van teritorija Hrvatske			
Postojanje konkretnog negativnog događaja	Zastupljenost govora mržnje u sadržaju objave		
	Glavna tema	Sporedna tema	ukupno
Da	109	37	146
Ne	33	57	90
ukupno	142	94	236

Četvrtina (26,8 % ili 236) objava iz uzorka se odnosi na događaje medijski označene kao govor mržnje koji se nisu odvijali unutar teritorija Hrvatske.

Od ukupno 236 objava koje se svojim sadržajem odnose na govor mržnje na teritoriju neke druge zemlje (a ne Hrvatske), najzastupljenije zemlje su: Srbija (33 objave), SAD (31 objava), Internet (26 objava), Europska unija (24 objave), Bosna i Hercegovina (23 objave), Poljska (19 objava), Slovenija (14 objava) i Novi Zeland (11 objava).

Kategorija 'zemlje Europske Unije' i 'Internet' zahtijevaju objašnjenje. Ukoliko se u objavi moglo odrediti u teritoriju koje zemlje se odvija govor mržnje, ta je zemlja i kodirana ovom kategorijom. Kategorija Europske Unije se odnosi na one objave koje u svojem sadržaju govor mržnje lociraju na cijelu Europsku Uniju ili na više od dvije zemlje Europske Unije istovremeno. Kategorija Interneta se odnosi na objave čiji sadržaj o govoru mržnje se ne odnosi na neku određenu zemlju nego Internet. Objave označene kao Internet su, na primjer, one objave koje govore o ukupnom broju objava koje Facebook uklanja sa svoje platforme zbog detekcije govora mržnje.

Osim Sjedinjenih Američkih Država (SAD), koje kao neeuropska zemlja svakako prednjače brojem objava vezanim uz govor mržnje, na listi se nalaze i: Novi Zeland, Mjanamar, Brazil, Papua Nova Gvineja, Šri Lanka, Venezuela, Izrael i Saudijska Arabija. Sve skupa je pronađena 51 objava koja se odnosi na govor mržnje koji se zbiva van Europe što čini 5,8 % ukupnog broja objava iz uzorka.

Europske zemlje (ne zemlje Europske Unije, već zemlje koje su barem dijelom u Europi) čine 151 objavu, što čini 17,2 % ukupnog broja objava iz uzorka.

Zanimljivo je za primijetiti kako su neke objave toliko generalizirane s obzirom na lokaciju izvora iz kojeg je potekla medijska objava na koju se odnose da je bilo potrebno u kategoriju „zemlja u kojoj se radi o govoru mržnje uvesti opciju „svijet“. Takvih objava ima u ovom uzorku 6¹²⁸.

128 U kategoriji „svijet“ se ne nalaze objave označene sa „Internet“. Primjeri: statistike govora mržnje na svjetskoj razini, opasnosti od razvoja tehnologija poglavito umjetne inteligencije.

Govor mržnje se podjednako pojavljuje kao glavna i kao sporedna tema objave

Ovom hipotezom se pokušava pokazati je li tema govora mržnje u medijima zastupljena kao tema koja se provlači kroz neke druge teme ili je to tema koja u medije dopijeva uglavnom kao samostalna tema. Time se istovremeno nastoji pokazati je li tematika govora mržnje nadišla svoj milje i počinje se širiti i na druge medijske teme.

Usporedbom kategorija glavna ili sporedna tema s podacima o tome radi li se o prenošenju nekog ekscesa vezanog uz govor mržnje nastoji se pokazati jesu li medijske objave uglavnom vezane za prenošenje ekscesa i u kojoj mjeri se medijski sadržaj odnosi na pokušaje promicanja svijesti o govoru mržnje ili poticanju onog suprotnog od govora mržnje što su tolerancija i razumijevanje ili jednostavno ljudskost i neosuđivanje.

Glavna i sporedna tema su uvriježene analitičke kategorije. Ukoliko se objava punim svojim sadržajem odnosi na govor mržnje, označen je kao glavna tema. Ukoliko se sadržaj objave samo dotiče govora mržnje ili je govor mržnje samo spomenut u objavi, govor mržnje označen je kao sporedna tema objave.

Od ukupno 878 objava, njih 445 imaju govor mržnje kao glavnu temu što čini 50,68 % broja objava iz uzorka.

Iako nije kodirano, iz analiziranog korpusa se primjećuje kako su članci o govoru mržnje učestalo vezani uz temu slobode govora. Također je primjetljivo kako je pojam govora mržnje učestalo vezan uz političke optužbe jedne političke opcije spram druge političke opcije. O prosudbi ispravnosti osude za govor mržnje u ovoj se analizi nije prosuđivalo, no svakako je ispravni daljnji smjer prosuditi koriste li se javne optužbe za govor mržnje ispravno ili su to samo prazne neutemeljene optužbe.

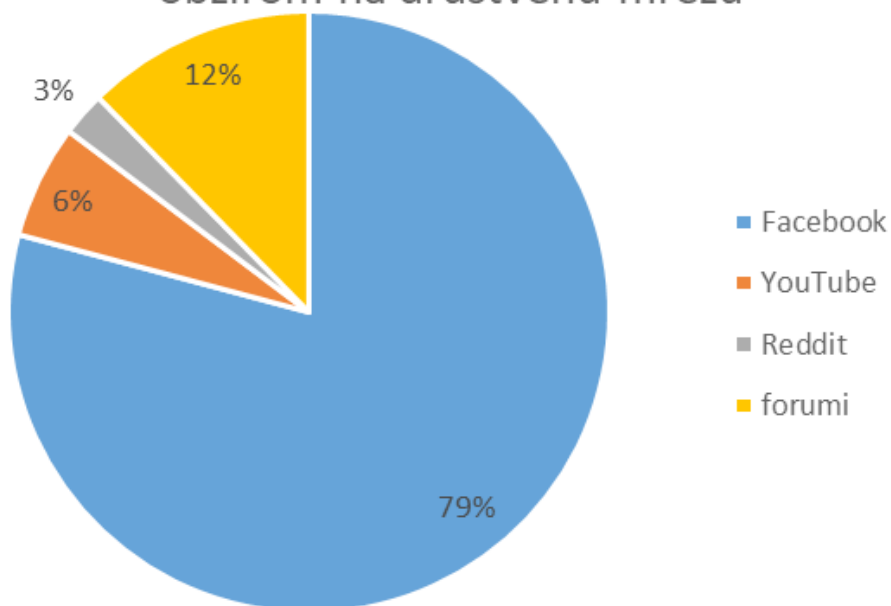
Podaci o komentarima na društvenim mrežama

Od ukupno 72.001 prikupljenih komentara s društvenih mreža, neprihvatljiv govor pronađen je u 1012 komentara što čini 1,4 % komentara.

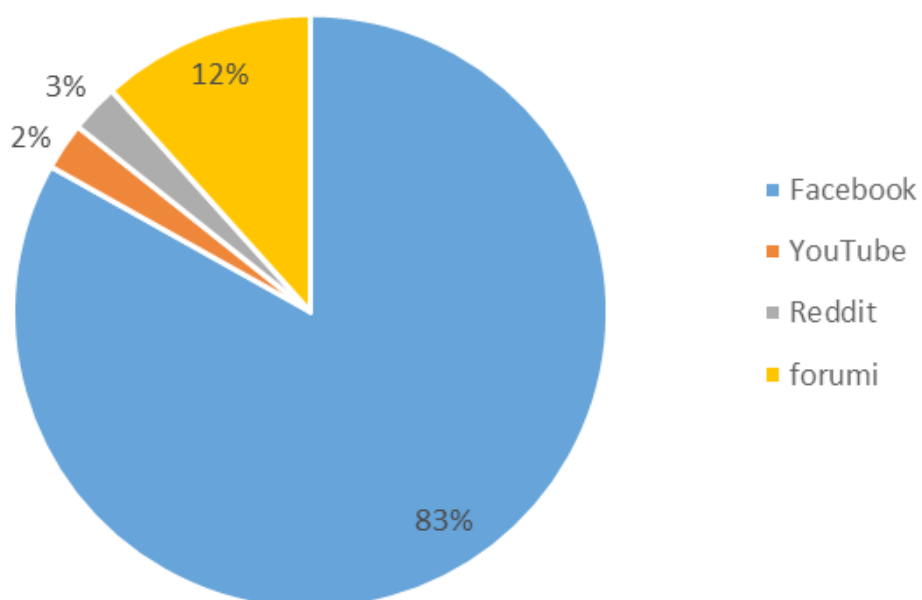
Najviše komentara je prikupljeno s društvene mreže Facebook. Najviše neprihvatljivih komentara je pronađeno na društvenoj mreži Facebook. Najveći udio neprihvatljivih komentara u odnosu na ukupan broj komentara pronađen je na društvenoj mreži YouTube.

	Facebook	YouTube	Reddit	forumi	ukupno
ukupno komentara	59.847	1.861	1.906	8.387	72.001
neprihvatljivih komentara	801	62	24	125	1.012
udio neprihvatljivih komentara	1,34 %	3,33 %	1,26 %	1,49 %	1,41 %

Udio neprihvatljivih komentara s obzirom na društvenu mrežu



Ukupan broj prikupljenih komentara s obzirom na društvenu mrežu



Prema kome je usmjereno najviše neprihvatljivih komentara na društvenim mrežama

Neprihvatljiv govor u komentaru nije uvijek usmjeren prema skupini na koju se naslov odnosi.

Tablica: broj komentara sa neprihvatljivim govorom s obzirom na skupinu prema kojoj je govor usmjeren i s obzirom na to unutar koje pretrage je neprihvatljivi govor pronađen.

Skupine prema kojima su pretraživani naslovi	Skupina prema kojoj je usmjeren neprihvatljivi govor																			Ukupno
	Etnicitet														Migranti	Vjera				
	Albanci	Amerikanci	Hrvati	Srbi	Bosanci	Slovinci	Romi	Nijemci	Kinezi	Talijani	Englezi	Francuzi	Iranci	Poljaci		Katolici	Pravoslavci	Muslimani	Židovi	
Albanci	5	1	1	30											2					39
Bosanci		1	11	18	106										22			17		175
Migranti			3		2		5	2							51			2	1	66
Katolici		1	1	7	3				2	1						44	81	1		141
Muslimani		2	1		23				1		1	1	1		8	11		59	5	113
Pravoslavci			1	16												13	20			50
Romi				1			10													11
Slovinci			8	5		86	2			1					1		1			104
Srbi	8		12	113			4		1	2					1		2	1		144
Talijani			1	4		4				75					3	5		5		97
Židovi			1	1												11			58	72
Ukupno	13	5	40	195	134	90	21	2	4	79	1	1	1	1	88	84	104	85	64	1012

U tablici su crvenim brojevima označene one skupine prema kojima je bilo najviše komentara s neprimjerenim govorom.

Skupina za koju je pronađeno najviše komentara s neprihvatljivim govorom su Srbi. Broj komentara s neprihvatljivim govorom je 195. od toga je svega 58 % neprihvatljivih komentara unutar naslova koji govore o Srbima. Neprihvatljivi govor prema Srbima može se pronaći i u komentarima vezanim uz naslove o Albancima, Bosancima, katolicima, pravoslavcima, Slovincima, Romima, Talijanima i Židovima. Jedine teme (naslovi) unutar kojih nisu pronađeni neprihvatljivi komentari prema Srbima su migranti i muslimani.

Nakon Srba, druga skupina prema kojoj ima najviše neprihvatljivog govora su Bosanci (134 komentara sa neprihvatljivim govorom). 106 ili 79 % neprihvatljivih komentara je pronađeno na temu Bosanaca, dok se neprihvatljivi govor pojavio i unutar tema o migrantima, katolicima i muslimanima.

Treća skupina prema kojoj je pronađeno najviše neprihvatljivih govora su pravoslavci. Najviše neprihvatljivog govora usmjerenog prema pravoslavcima pronađeno je između komentara s naslovima o katolicima (78 %).

Skupine (naslovi) koje imaju najveći broj neprihvatljivih komentara (bez obzira prema kome je usmjeren) su Bosanci. Ukupno 175 neprihvatljivih komentara je pronađeno unutar naslova koji se odnose na Bosance. Iako je većina pronađenog neprihvatljivog govora unutar naslova o Bosancima usmjerena prema Bosancima (60 %), preostalih 69 komentara sa neprihvatljivim govorom su usmjereni prema: migrantima, Srbima, muslimanima, Hrvatima i Amerikancima. Redoslijed skupina je od najviše prema najmanje neprihvatljivih komentara.

Druga skupina prema kojoj je usmjereno najviše neprihvatljivog govora su Srbi. Ukupno 144 neprihvatljiva komentara je pronađeno unutar naslova koji se odnose na Srbe. Iako je većina pronađenog neprihvatljivog govora unutar naslova o Srbima usmjerena prema Srbima (78 %), preostalih 31 komentara sa neprihvatljivim govorom su usmjereni prema: Hrvatima, Albancima, Romima, Talijanima, pravoslavicima, muslimanima, migrantima i Kinezima. Redoslijed skupina je od najviše prema najmanje neprihvatljivih komentara.

Od izabranih skupina, najmanje neprihvatljivog govora je usmjereno prema Albancima i prema Romima.

Skupine prema kojima je usmjeren neprihvatljiv govor, a koje nisu bile predmetom istraživanja (odnosno nisu bile pretraživane) su: Amerikanci, Kinezi, Nijemci, Englezi, Francuzi, Iranci i Poljaci.

Unutar korpusa o Albancima i katolicima, najveći broj neprihvatljivih komentara je usmjeren na neke druge skupine (Srbe i pravoslavce). Sve ostale pretraživane skupine imaju najveći broj neprihvatljivog govora usmjerenog prema onoj skupini prema kojoj je pretraga rađena.

Nije pronađena niti jedna skupina prema kojoj se pretraživalo a da nema neprihvatljivog govora prema toj skupini.

Komparirajući etnicitet, vjeru i migrante, najviše neprihvatljivih komentara ima temeljem etniciteta (587 komentara), zatim temeljem vjere (337) te su na ovoj listi zadnji migranti (88). Ovaj podatak svakako treba promatrati kroz činjenicu da je pretraživano 6 različitih etničkih skupina i 4 vjerske skupine dok migrante predstavlja samo jedna skupina.

Gledajući naslove s obzirom na etnicitet, vjeru i migrante, najviše neprihvatljivih komentara ima na naslove koji se odnose na etnicitet (570), zatim na religiju (376) te najmanje unutar naslova o migrantima (66).

Što se najviše komentira i gdje ima najviše neprihvatljivog govora

Skupine prema kojima su pretraživani naslovi	ukupan broj komentara	ukupan broj komentara sa neprihvatljivim govorom	udio neprihvatljivog govora
Muslimani	5013	113	2,25 %
Židovi	3238	72	2,22 %
Albanci	1971	39	1,98 %
Talijani	4949	97	1,96 %
Pravoslavci	2959	50	1,69 %
Bosanci	10755	175	1,63 %
Katolici	8892	141	1,59 %
Slovenci	7652	104	1,36 %
Srbi	12837	144	1,12 %
Romi	1448	11	0,76 %
Migranti	12287	66	0,54 %
Ukupno	72001	1012	1,41 %

Najviše komentara prikupljeno je unutar naslova koji se odnose na Srbe. Najmanje naslova je prikupljeno za naslove koji se odnose na Rome.

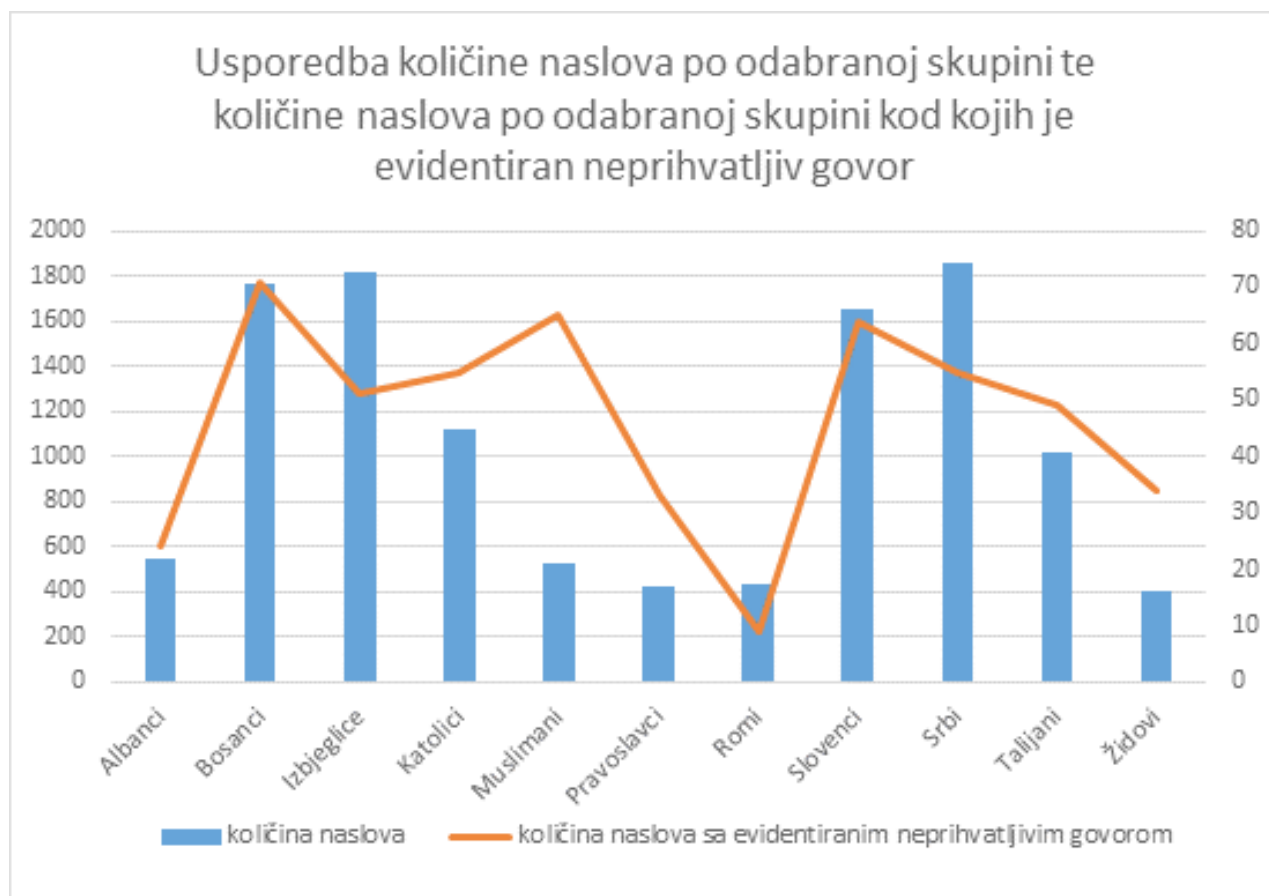
Promatrajući udjele neprihvatljivih komentara u odnosu na ukupan broj komentara, najčešće se neprihvatljivi komentari nalaze na naslove koji se odnose na muslimane.

Očigledno je kako redoslijed skupina prema broju komentara, broju neprihvatljivih komentara i udjela neprihvatljivih komentara nisu usklađeni. Vidljivo je kako skupine koje nemaju najviše komentara niti najviše neprihvatljivih komentara u apsolutnom iznosu, su na prvim mjestima prema udjelu neprihvatljivih komentara.

Najmanji udio neprihvatljivih komentara u ukupnom broju prikupljenih komentara je za skupinu migranti.

Iznalaženje odgovora na pitanje zašto pojedine skupine imaju manji ili veći udio neprihvatljivih komentara nije predmet ovog istraživanja. Dobivanje odgovora na to pitanje zahtijevalo bi dublju analizu samih naslova, događaja, konteksta a možda i komentatora (pošiljatelja).

Skupine sa najviše naslova i tko ima najviše neprihvatljivih komentara po naslovu



Sam način izbora sadržaja u uzorak se svodi na pronalaženje naslova koji se odnose na pretraživane skupine. Povrh broja prikupljenih komentara, dajemo i prikaz broja pronađenih naslova s obzirom na izabrane skupine za pretraživanje sadržaja društvenih mreža.

Na grafu je prikazan ukupan broj prikupljenih naslova i ukupan broj naslova u kojima je pronađen barem jedan neprihvatljivi komentar. Udio naslova koji imaju neprihvatljive komentare nije sukladan ukupnom broju prikupljenih naslova na pojedinu pretraživanu skupinu i taj broj se vidi na desnoj ordinati (u apsolutnom iznosu).

U prosjeku svaki 25. naslov ima barem jedan komentar sa neprihvatljivim sadržajem (4,4 %). 12,4 % naslova skupine muslimani imaju barem jedan komentar sa neprihvatljivim sadržajem. Najmanji udio naslova sa neprihvatljivim sadržajem je pronađen prilikom pretraživanja skupine Romi (2 %).

Slijedom ovog zaključivanje naslovi na koje se najčešće pojavljuju komentari koji su neprihvatljivi su: muslimani (12,4 %), Židovi (8,5 %), pravoslavci (7,8 %). Skupine čiji naslovi najrjeđe imaju neprihvatljive komentare su: Romi (2 %), izbjeglice (2,8 %) i Srbi (2,9 %).

Očekivali smo da ćemo od primjene oblaka ključnih riječi koji u sebi sadrži mogućnost posebnog prikaza glagola, pridjeva i imenica dobiti određene analitičke parametre koji će nas upućivati prema zaključku. Međutim, do toga nismo došli iz razloga što se u neprihvatljivom izričaju prema bilo kojem od predmeta naše analize (religija, etnicitet i migranti) upotrebljavaju metafore, ironija, sarkazam a većina izgovorenog ima neke kontekstualne i semantičke specifične značajke.

Poučeni ovim dijelom istraživanja shvatili smo da trebamo napraviti bazu upotrjebljenih izraza prema različitim skupinama koja je u postupku izrade i za koju se nadamo da će nam omogućiti dublji analitički uvid u sadržaj neprihvatljivog govora. Rezultate ćemo prikazati u narednom radu i zbog toga ovaj rad ima karakter prethodnog priopćenja.

ZAKLJUČAK

Problem govora mržnje kroz godine dobiva sve više medijskog prostora.

Mediji vrše funkciju podizanja javne svijesti o problemu govora mržnje kroz prenošenje sadržaja i događaja vezanih uz govor mržnje. Zastupljenost medijskog sadržaja s pojmom govora mržnje se kroz vrijeme povećava. Osim nasilnog i neprimjerenog ponašanja, medijski sadržaj odnosi se i na promicanja problematike govora mržnje kroz različite znanstvene i stručne skupove, edukacije, kulturna događanja koja tematiziraju govor mržnje, općenite ili statističke podatke o govoru mržnje te nastojanja vlasti da prilagode zakonodavni okvir.¹²⁹

Medijski sadržaj o govoru mržnje većinom je vezan uz događaje koji su se zbili na teritoriju Republike Hrvatske, no četvrtina svih promatranih medijskih objava odnosi se na govor mržnje u nekim drugim zemljama, pretežno susjednim zemljama. Mediji portretiraju problematiku govora mržnje kao sveopćeg društvenog problema koji nadilazi granice Hrvatske države. Istovremeno se svrstava Hrvatsku kao zemlju u trendu sa svjetskim nastojanjima u borbi s ovim problemom.

Objave koje u sebi sadrže pojam govora mržnje: u podjednakom omjeru su objave koje se punim svojim sadržajem odnose na govor mržnje kao i objave koje se manjim dijelom odnose na govor mržnje ili je govor mržnje samo spomenut u objavi. Govor mržnje nije nužno samostalna medijska tema već se isprepliće sa drugim temama i sadržajima.

Analizom sadržaja komentara sa društvenih mreža utvrđena je pojavnost neprihvatljivog govora na razini od 1,4 % prikupljenih i analiziranih komentara.

Obzirom na naslove (odnosno tematiku) pod kojima su komentari skupljeni, najviše se neprihvatljivih komentara pojavljuje uz naslove vezane uz Bosance. Skupina prema kojoj je usmjereno najviše neprihvatljivih komentara su Srbi. Neprihvatljivi komentari nisu uvijek usmjereni prema onoj skupini na koju se odnosi naslov.

¹²⁹ Pitanje je je li samo prenošenje pozitivnih primjera podizanje javne svijesti ili se i negativni događaji mogu tumačiti kao podizanje javne svijesti o problemu govora mržnje.

U ovom radu smo primijenili jezične tehnologije i to:

1. Klasifikator - Za prikupljanje podataka iz baze podataka, gdje je isti pokazao točnost od 95 %
2. Simularnosti (iste vijesti) – radi detekcije istovjetnosti medijskog sadržaja i brzine širenja istog kroz medijski prostor. Razina točnosti je 92 %
3. Event detection – za grupiranje medijskog sadržaja u događaje. Razina točnosti je 86 %
4. Metricom za prikupljanje komentara i naslova sa društvenih mreža

Većina upotrjebljenih i novih tehnologija primijenit će se u sljedećoj fazi istraživačkog procesa koji ima za cilj napraviti bazu najčešće upotrebljivanih riječi govora mržnje prema vrsti, značenjske odrednice upotrjebljenih izraza i najčešćih metafora te najčešće izraze, što ovaj rad svrstava u kategoriju prethodnih priopćenja.

DISKUSIJA

Medijski prostor zaista zastupa tematiku govora mržnje, i to ne samo kao prenošenje događaja već se kroz medijski sadržaj proteže jedna (izravna ili između redova) osuda govora mržnje. Svakoliki je dojam istraživača da se status osude za govor mržnje tretira kao nešto zaista grozno. Svakako je stav medija koji osuđuje (izravno ili neizravno) govor mržnje pozitivan smjer, poglavito imajući u vidu ponekad nedovoljno naglašenu odgojnu funkciju medija. Pritom, kad se kaže odgojna, smatramo to da sadržaj medija masovnog komuniciranja može utjecati na promjene mišljenja i stavova pojedinaca ili društva i promijeniti javno mnijenje o pojedinoj temi. Svakako bi ovakva razmišljanja trebalo provjeriti nekim daljnjim istraživanjima, bilo kroz analizu sadržaja (da se utvrde stavovi spram govora mržnje) bilo kao ispitivanje javnog mnijenja ili neki drugi oblik istraživanja.

Zanimljivo bi bilo istražiti prema kojoj skupini je najviše usmjeren govor mržnje te koje su to učestale fraze kojima se „časti“ pojedinu skupinu. Koji su to najčešći kanali putem kojih se govori mržnja?

Koje su to medijske teme uz koje je govor mržnje vezan? Iako ovo pitanje nije bilo u fokusu istraživanja, prilikom istraživanja „isplivale“ su teme: slobode govora, reguliranja društvenih mreža, politike i političkih govora i djelovanja, nasilja, sporta. Ovakva klasifikacija tema svakako ne predstavlja znanstveno utemeljenu klasifikaciju, već samo opservaciju istraživača koja bi mogla poslužiti nekom drugom.

Gledajući nekoliko stoljeća unazad, veliki je društveni pomak da se govor mržnje postavio kao društveni problem. Ovo je logičan pomak u modernom humanističkom društvu koje je na temelju slobode i jednakosti priznalo neprihvatljivost nasilja spram bilo kojeg pojedinca temeljem pripadnosti nekoj skupini i iznjedrilo posebnu klasifikaciju: nasilje iz mržnje. Problem govora mržnje je, prema mišljenju ovih istraživača, samo logičan daljnji korak u nastojanju popravaka društva na način da se nasilje iz mržnje prevenira još u ranom stadiju a što govor mržnje jest.

Ništa ne pogodi čovjeka kao spoznaja da je nasilje vezano uz govor mržnje prisutno u tvom susjedstvu. Čitati o ubijanju muslimana na Novom Zelandu, ubojstvu nožem u Poljskoj i smrtonosnoj pucnjava u sinagogi u SAD - u nije isto kao i čitati o žigosanju ustaškog „U“ u Zadarskoj srednjoj školi.

Zašto pojedini događaji dobivaju više medijske pozornosti od drugih? Radi li se o usklađivanju medijskog sadržaja po principu „ako drugi o tome pišu, moram i ja“ ili će više medijske pozornosti dobiti onaj događaj koji se odnosi na skandal vezan uz neku javnu osobu (poput Đakićeve Facebook „čestitke“ Božića „srbićima“)? Postoji li možda neka druga varijabla koja čini pojedini sadržaj o govoru mržnje medijski više ili manje zastupljenim?

Postoji li (i u kojoj mjeri) svojevršno razbacivanje osudom za govor mržnje? Postoji li opasnost da će osuda za govor mržnje postati još jedna izlizana politička fraza? U kojoj se mjeri različite medijski eksponirane osobe brane pozivajući se na govor mržnje prema njima?

Prati li broj događaja koji se odnose na imenovani negativni događaj u kojem je točno određeni pojedinac prozvan za govor mržnje broj sudskih procesa? Ili prozivanje u medijima za govor mržnje nikad ni ne dospije do prijave ili optužnice? Uostalom, je li veći broj prijava i / ili optužnica za govor mržnje od medijskih događaja u kojima se nekoga proziva za govor mržnje? I koji udio prijava i optužnica za govor mržnje uopće dospije u medije?

Onome tko prvi put zaluta u tematiku govora mržnje stanje se može činiti banalnim i nevrijednim znanstvenog interesa. No, kao i uvijek, kad dobiješ odgovor na jedno pitanje pojavi se stotina novih. To je valjda tako, jer da nema novih pitanja ne bi bilo niti znanosti.

Zahvale

Zahvaljujemo poduzeću Presscut i Medianet na pristupu bazi medijskih objava bez koje ovaj rad ne bi bilo moguće provesti.

Literatura:

- Alaburić, V. (2003). Ograničavanje „govora mržnje“ u demokratskome društvu – Teorijsko, zakonodovani i pratkični aspekti. *Hrvatska pravna revija: časopis za promicanje pravne teorije i prakse*, 3 (2), 80-90.
- Arofah, K. (2018) Rhetorical Analysis of Hate Speech - Case Study of Hate Speech Related to Ahok's Religion Blasphemy Case, *Journal Komunikasi*, 11(1), str. 91-105.
- Aslan, A. (2017) Online Hate Discourse - A Study on Hatred Speech Directed Against Syrian Refugees on YouTube, *Journal of Media Critiques*, 12(3), str. 227-256, DOI: 10.17349/jmc117413.
- Baider, F. (2019) Double speech act - Negotiating inter-cultural beliefs and intra-cultural hate speech, *Journal of Pragmatics*, 151, str. 155-166, DOI: 10.1016/j.pragma.2019.05.006.
- Bajt, V. (2015) Anti-Immigration Hate Speech in Slovenia, u *Razor Wired: Reflections on Migration Movements through Slovenia in 2015*, str. 51-61.
- Banks, J. (2010). Regulating hate speech online, *International Review of Law, Computers & Technology*, 24:3, 233-239
- Baumgarten, N. & Bick, E. & Geyer, K. & Iversen, D. & Kleene, A. & Lindø, A. & Jana Neitsch, Niebuhr, O. & Nielsen, R. & Nedenskov Petersen, E. (2019) Towards Balance and Boundaries in Public Discourse: Expressing and Perceiving Online Hate Speech (XPEROHS), *RASK: International Journal of Language and Communication*, 50(Autumn 2019), str. 87-108.
- Ben-David, A. & Matamoros-Fernandez, A. (2016) Hate Speech and Covert Discrimination on Social Media - Monitoring the Facebook Pages of Extreme-Right Political Parties in Spain, *International Journal of Communication*, 10, str. 1167–1193.
- Benkler, Y., Faris, R., Roberts, H., Zuckerman, E., (2017). Study: Breitbart-led right-wing media ecosystem altered broader media agenda. *Columbia Journalism Review* [online] 3. 3. 2017. <https://www.cjr.org/analysis/breitbart-media-trump-harvard-study.php> (pristupljeno 28. 6. 2019.).
- Bentzen, N. (2017). Parliamentary briefing: 'Fake news' and the EU's response. *European Parliament Think Tank* [online] 31.3.2017. [http://www.europarl.europa.eu/thinktank/en/document.html?reference=EPRS_ATA\(2017\)599384](http://www.europarl.europa.eu/thinktank/en/document.html?reference=EPRS_ATA(2017)599384) (pristupljeno 28. 6. 2019.).
- Blanco-Herrero, D. & Calderon, C. (2019) Spread and reception of fake news promoting hate speech against migrants and refugees in social media, u *TEEM'19: Proceedings of the Seventh International Conference on Technological Ecosystems for Enhancing Multiculturality* (Leon, Španjolska, 16.- 18. 10. 2019.), str. 949–955, DOI: 10.1145/3362789.3362842.
- Blazak, R. (2009). Toward a working definition of hate groups. In B. Perry, B. Levin, P. Iganski, R. Blazak, & F. Lawrence (Eds.), *Hate crimes* (133–148). Westport, CT: Greenwood Publishing Group.

- Bosco, C. & Patti, V. & Bogetti, M. & Conoscenti, M. & Ruffo, G. & Schifanella, R. & Stranisci, M. (2019) Tools and resources for detecting hate and prejudice against immigrants in social media, u *Symposium Iii. Social Interactions In Complex Intelligent Systems (SICIS)* (Falmouth, Ujedinjeno Kraljevstvo, 16. - 18. 4. 2019), str. 79-84.
- Britannica. Hate speech [online] <https://www.britannica.com/search?query=hate+speech> (preuzeto 28. 6. 2019.).
- Brown, A. (2018). What is so special about online (as compared to offline) hate speech? *Ethnicities*, 18(3), 297–326.
- Busolo, D. & Ngigi, S. (2018) Understanding Hate Speech in Kenya, *New Media and Mass Communication*, 70, str. 43-49.
- Carlson, C. & Rousselle, H. (2020) Report and repeat - Investigating Facebook's hate speech removal process, *First Monday*, 25(2), DOI: 10.5210/fm.v25i2.10288.
- Carney, T. (2014) Being (im)polite: A forensic linguistic approach to interpreting a hate speech case, *Language Matters: Studies in the Languages of Africa*, 45(3), str. 325-341, DOI: 10.1080/10228195.2014.959545.
- Chandrasekharan, E. & Pavalanathan, U. & Srinivasan, A. & Glynn, A. & Eisenstein, J. & Gilbert, E. (2017) You Can't Stay Here - The Efficacy of Reddit's 2015 Ban Examined Through Hate Speech, *PACM on Human-Computer Interaction*, 1(CSCW), str. 1-22, DOI: 10.1145/3134666.
- Chetty, N., Alathur, S. (2018). Hate speech review in the context of online social networks. *Aggression and Violent Behavior*, Volume 40, 108-118
- Cohen-Almagor, R. (2019) Racism and hate speech – A critique of Scanlon's Contractual Theory, *First Amendment Studies*, 53(1-2), str. 41 - 66, DOI: 10.1080/21689725.2019.1601579.
- Comandini, G. & Patti, V. (2019) An Impossible Dialogue! Nominal Utterances and Populist Rhetoric in an Italian Twitter Corpus of Hate Speech against Immigrants, u *Proceedings of the Third Workshop on Abusive Language Online* (Firenza, Italija, 1. 8. 2019.), str. 163–171.
- Costello, M., Hawdon, J., Ratliff, T. N. (2017). Confronting Online Extremism: The Effect of Self-Help, Collective Efficacy, and Guardianship on Being a Target for Hate Speech. *Social Science Computer Review*, 35(5), 587–605.
- Culpeper, J. (2011). *Impoliteness. Using Language to Cause Offence*. Cambridge: Cambridge University Press.
- Čerkez, I. (2009). Osnovna obilježja medijske komunikacije u demokratskoj kulturi. *Socijalna ekologija*. 18(1), 28-45.
- Davids, N. (2018) On the (in)tolerance of hate speech - does it have legitimacy in a democracy, *Ethics and Education*, 13(3), str. 296-308, DOI: 10.1080/17449642.2018.1477036.
- Davidson, T., Warmsley, D., Macy, M., Weber, I. (2017). Automated hate speech detection and the problem of offensive language. *arXiv Preprint arXiv:1703.04009*.

- Demaske, C. (2019) Social Justice, Recognition Theory and the First Amendment - A New Approach to Hate Speech Restriction, *Communication Law and Policy*, 24(3), str. 347-401, DOI: 10.1080/10811680.2019.1627800.
- Doring, N. & Mohseni, M. (2019) Fail videos and related video comments on YouTube - a case of sexualization of women and gendered hate speech, *Communication Research Reports*, 36, str. 1-11, DOI: 10.1080/08824096.2019.1634533.
- Enarsson, T. & Lindgren, S. (2018) Free speech or hate speech - A legal analysis of the discourse about Roma on Twitter, *Information & Communications Technology Law*, 28(1), str. 1-18, DOI: 10.1080/13600834.2018.1494415.
- Ezeibe, C. & Ikeanyibe, O. (2017) Ethnic Politics, Hate Speech, and Access to Political Power in Nigeria, *Africa Today*, 63(4), str. 64-83, DOI: 10.2979/africatoday.63. 4. 04.
- Florio, K. & Basile, V. & Lai, M. & Patti, V. (2019) Leveraging Hate Speech Detection to Investigate Immigration-related Phenomena in Italy, u *8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)* (Cambridge, Ujedinjeno Kraljevstvo, 3.- 6. 9. 2019.), str. 219-225.
- Fortuna, P., Nunes, S. (2018). A Survey on Automatic Detection of Hate Speech in Text. *ACM Comput. Surv.* 51, 4, Article 85.
- Fuchs, C. (2014). *Social Media: A Critical Introduction*. London: SAGE.
- Fuchs, T. & Schafer, F. (2020) Normalizing misogyny - hate speech and verbal abuse of female politicians on Japanese Twitter, *Japan Forum*, str. 1-27, DOI: 10.1080/09555803.2019.1687564.
- Gelber, K. (2019): Differentiating hate speech: a systemic discrimination approach, *Critical Review of International Social and Political Philosophy*, str. 1-22, DOI: 10.1080/13698230.2019.1576006
- Gould, R. (2019) Is the 'hate' in hate speech the 'hate' in hate crime? Waldron and Dworkin on political legitimacy, *Jurisprudence - An International Journal of Legal and Political Thought*, 10(2), str. 171-187, DOI: 10.1080/20403313.2018.1552468.
- Hawdon, J., Oksanen, A., Rasanen, P. (2016). Exposure to online hate in four nations: A cross-national consideration. *Deviant Behavior*, 38:3, 254-266
- Hebrang, B. (2006). *Informacijske (novinske) agencije*. Zagreb: Jarbol d.o.o.
- Hemker, K. (2018) *Data Augmentation and Deep Learning for Hate Speech Detection*. MSc. Imperial College London.
- Hrdina, M. (2016) Identity, activism and hatred - Hate speech against migrants on facebook in the Czech Republic in 2015, *Naše společnost*, 14(1), str. 38-45, DOI: 10.13060/1214438X.2016.1.14.265.
- Jaki, S. & De Smedt, T. (2018) Right-wing German Hate Speech on Twitter - Analysis and Automatic Detection, *arXiv - Computation and Language*, arXiv:1910.07518 [cs.CL].

- Kazneni zakon RH, zadnja izmjena 1. 1. 2020., NN 125/11, 144/12, 56/15, 61/15, 101/17, 118/18, 126/19
- Kimotho, S. & Nyaga, R. (2016) Digitized Ethnic Hate Speech - Understanding Effects of Digital Media Hate Speech on Citizen Journalism in Kenya, *Advances in Language and Literary Studies*, 7(3), str. 189-200, DOI: 10.7575/aiac.all.v.7n.3p.189.
- Kulenović, E. (2016) *Govor mržnje u Hrvatskoj*. Zagreb, Fakultet političkih znanosti Sveučilišta u Zagrebu
- Labaš, D. i Grmuša, T. (2011). Istinitost i objektivnost u informaciji i društveno štetne komunikacijske forme. *Kroatologija*. 2(12), 87-122.
- Laaksonen, S-M. & Haapoja, J. & Kinnunen, T. & Nelimarkka, M., & Pöyhtäri, R. (2020). The Datafication of Hate: Expectations and Challenges in Automated Hate Speech Monitoring, *Frontiers in Big Data*, 3(3), str. 1-16, DOI:10.3389/fdata.2020.00003.
- Lazaridis, G. & Veikou, M. (2016) The rise of the far right in Greece and opposition to 'othering', hate speech, and crime by civil and civic organizations, *Journal of Civil Society*, 13(1), str. 1 - 17, DOI: 10.1080/17448689.2016.1251032.
- Leets, L. (2002). Experiencing Hate Speech: Perceptions and Responses to Anti-Semitism and Antigay Speech. *V Journal of Social Issues*. 58 (2): 341 – 361.
- Macquarie rječnik (2016) *Riječ godine*. [online] 23. 11. 2016. <https://www.macquariedictionary.com.au/resources/view/word/of/the/year/2016> (pristupljeno 28. 6. 2019.).
- Malmasi, S. & Zampieri, M. (2017) Challenges in discriminating profanity from hate speech, *Journal of Experimental & Theoretical Artificial Intelligence*, 30(4), str. 1-16. DOI: 10.1080/0952813X.2017.1409284.
- Marwick, A., Lewis, R. (2017). Media Manipulation and Disinformation Online. *Data and Society* [online] 15. 5. 2017. <https://datasociety.net/output/media-manipulation-and-disinfo-online/> (pristupljeno 28. 6. 2019.).
- Mattsson, C. (2019) Lost in translation – A case study of a public debate on freedom of expression and a neo-Nazi rally, *Social Identities*, 26(1), str. 92-108, DOI: 10.1080/13504630.2019.1679620.
- Meibauer, J. (2013). *Hassrede/Hate Speech. Interdisziplinäre Beiträge zu einer aktuellen Diskussion*. Gießener Elektronische Bibliothek.
- Mondal, A. (2018) The shape of free speech: rethinking liberal free speech theory, *Continuum*, 32(4), str. 503-517, DOI: 10.1080/10304312.2018.1480463.
- Munivrana V., Šurina, M., Marton, A. (2016). Gdje prestaju granice slobode izražavanja, a počinje govor mržnje? Analiza hrvatskog zakonodavstva i prakse u svjetlu europskih pravnih standarada. *Hrvatski ljetopis za kaznene znanosti i praksu* vol. 23/2, str. 435-467.

- O’Hehir, R. (2019) ‘Europe’s cockroaches’ - freedom of expression, hate speech and their impact on refugee human rights, *Aristotle University of Thessaloniki - European Master’s Degree in Human Rights and Democratisation*, A.Y. 2018/2019.
- Ousidhoum, N. & Lin, Z. & Zhang, H. & Song, Y. & Yeung, D. (2019) Multilingual and Multi-Aspect Hate Speech Analysis, *arXiv - Computation and Language*, arXiv:1908.11049 [cs.CL].
- Oxfordov rječnik (2016) *Riječ godine*. [online] <https://en.oxforddictionaries.com/word-of-the-year/word-of-the-year-2016> (pristupljeno 28. 6. 2019.)
- Pejchal, V. (2018) Hate Speech Regulation in Post-Communist Countries - Migrant Crises in the Czech and Slovak Republics, *International Journal for Crime, Justice and Social Democracy*, 7(2), str. 58-74, DOI: 10.5204/ijcjsd.v7i2.500.
- Reid, A. (2019) Does Regulating Hate Speech Undermine Democratic Legitimacy - A Cautious ‘No’, *Res Publica*, str. 1-19, DOI: 10.1007/s11158-019-09431-6.
- Sap, M. & Card, D. & Gabriel, S. & Choi, Y. & Smith, N. (2019) The Risk of Racial Bias in Hate Speech Detection, u *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages (Firenza, Italija, 28.7.-2.8.2019.), str. 1668–1678.
- Sharma, H.K., Singh, T.P., Kshitiz, K.K., Singh, H., Kukreja, P. (2017). Detecting Hate Speech and Insults on Social Commentary using NLP and Machine Learning. *International Journal of Engineering Technology Science and Research* 4(12), 279-285.
- Tsesis A (2001) Hate in cyberspace: Regulating hate speech on the Internet. *San Diego Law Review* 38, 817–874.
- Vigna, F.D., Cimino, A., Dell’Orletta, F., Petrocchi, M., Tesconi, M. (2017). Hate Me, Hate Me Not: Hate Speech Detection on Facebook. In: *Proceedings of the First Italian Conference on Cybersecurity (ITASEC17)*.
- Vilović, G. (2011). Politološki pojmovnik: Govor mržnje. *Političke analize* 2(6). 68-70.
- Vijeće EU, Okvirna odluka o suzbijanju određenih oblika i načina izražavanja rasizma i ksenofobije kazneno-pravnim sredstvima, 2008/913/PUP, 2008. godina
- Vozab, D. (2017). Pristrani i neprijateljski mediji te polarizacija u novom medijskom okolišu. *Političke analize*. 8(30), 3-9.
- Zhang, Z., Robinson, D., Tepper, J. (2018). Detecting hate speech on Twitter using a convolution-GRU based deep neural network. In: Gangemi, A., Navigli, R., Vidal, M.E., Hitzler, P., Troncy, R., Hollink, L., Tordai, A. and Alam, M., (eds.) *ESWC 2018: The semantic web*. ESWC 2018, 03-07 Jun 2018, Heraklion, Greece. Lecture Notes in Computer Science, 10843. Springer Verlag, 745-760.

Hate speech in Croatian Media Space

Abstract

In this paper, authors are considering current problem of legally regulated unacceptable behavior in Croatian media space, social networks included, by focusing on hate speech. Fundamental question is whether hate speech presence is on the rise. Through content analysis of the entire Croatian media content, authors find an increase of unacceptable behavior and hate speech. Based on results, research is focused on detection of types of hate speech and examples of their occurrence whereas this type targets ethnicity, religion and migrants. Third research is focused on events present in media space, targeting unacceptable behavior and provide contextual frame of important events using qualitative and quantitative content analysis and NLP technologies. Research continues with development of NLP technologies for detecting, processing and analyzing media content.

Key words: *hate speech, social networks, media content, NLP technologies.*



This journal is open access and this work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.