

Tado Jurić\*

JEL klasifikacija: J1  
Prethodno priopćenje  
<https://doi.org/10.32910/ep.73.2.2>

## FACEBOOK I GOOGLE KAO EMPIRIJSKA OSNOVA ZA RAZVOJ METODE DIGITALNOG PRAĆENJA VANJSKIH MIGRACIJA HRVATSKIH GRAĐANA

*Rad istražuje vanjske migracije hrvatskih građana koristeći inovativni pristup digitalne demografije, odnosno analizu uporabnih vrijednosti podataka koji se mogu dobiti analitičkim aplikacijama Google Trends (GT) i Facebook (FB) API u svrhu migracijskih istraživanja. Cilj rad je promoviranje posrednih izvora podataka kao korisne nadopune „klasičnim“ izvorima podataka u demografskim istraživanjima općenito i u migracijskim istraživanjima specifično, a u svjetlu činjenice da službena hrvatska demografska statistika u mnogobrojnim slučajevima ne omogućava stvarne procjene, što se pogotovo odnosi na recentno iseljavanje. Metoda je od koristi i za ekonomska istraživanja jer su demografski i gospodarski razvoj u kauzalnoj povezanosti i prožetosti. U radu se na primjeru vanjskih migracija hrvatskih građana u Njemačku, Austriju, Švedsku i Irsku prikazuje uporabna vrijednost ovih podataka uz otvaranje rasprave o prednostima i nedostacima istraživačkog pristupa koji se zasniva na ovakvom tipu podataka. Time nastojimo postaviti empirijsku osnovu za razvoj modela digitalnog praćenja migracija hrvatskih građana, koji bi istom osnovom mogao biti upotrijebljen i za šira istraživanja migracija na području Jugoistočne Europe. Rad daje i nove znanstvene spoznaje u području digitalne demografije jer pokazuje novi pristup obrade ovakvog tipa podataka kao i uvide vezane za stavove, namjere i motive migracije hrvatskih građana kao i njihove integracije u pridošlo društvo. Ovo je prva studija ovakvog tipa u Jugoistočnoj Europi, ali i prva*

---

\* T. Jurić, dr. sc., izvanredni profesor, Hrvatsko katoličko sveučilište (e-mail: tado.juric@uni-cath.hr). Rad je primljen u uredništvo 13.01.2021. godine, a prihvaćen je za objavu 01.06.2021. godine.

*koja kombinira izvore dvaju baza podataka Facebooka i Google Trendsa kako bi nadopunila nedostatke koje svaka od ovih baza ima. Unatoč brojnim prednostima ovakvog pristupa razmatramo i specifičnosti i ograničenja sustava ovakvih podataka i upozoravamo u kojim slučajevima može doći do ugroze valjanosti interpretacija i zaključaka.*

*Rezultati istraživanja pokazali su da postoji pozitivna korelacija između rasta pretraživanja određenih pojmova na hrvatskom jeziku u Njemačkoj („misa na hrvatskom“, „katolička misija“ i „hrvatska nastava“) i rasta registriranih useljenih hrvatskih građana u Njemačku. Studija prikazuje i novi način na koji se podaci dobiveni putem GT-a mogu standardizirati u svrhu komparacije sa službenim bazama podataka. Glavna prednost ovog pristupa je pravovremeno identificiranje vanjskih migracija, čime se uvidi u migracijske trendove dobivaju godinu dana prije službenih podataka te se mogu koristiti za modeliranje projekcija i predviđanje daljnjih trendova.*

**Ključne riječi:** *digitalna demografija, Facebook, Google Trends, iseljavanje hrvatskih građana, digitalni tragovi, big data*

## 1. UVOD

Službena hrvatska demografska statistika u određenim slučajevima ne omogućava stvarne procjene, što se pogotovo odnosi na recentno iseljavanje (vidi Jurić, 2021 f). Iz ovog razloga u znanstvenom pristupu pojavljuje se potreba za novim metodama istraživanja, koje smo kreirali nestandardnim postupcima, a čija je namjera da posluže kao nadopuna spoznaja „klasičnih“ demografskih pristupa.

Glavni izvor inovacije u demografskim istraživanjima, naročito u kontekstu migracijskih studija, čini uporaba podataka s weba i društvenih mreža (vidi: Zagheni, Weber, 2012; State i sur. 2013.; Zagheni i sur., 2014, 2017, 2018; Dubois i sur., 2015; Spyrtos i sur., 2018). Kako svake godine raste broj korisnika Interneta tako raste i količina digitalnih tragova na webu, ali i uporaba ovih podataka za različite tipove istraživanja (vidi Dubois i dr., 2018.). Pionirsko istraživanje iz polja demografije u ovom smjeru je pokrenuo UN (UN – Data Revolution), u čijoj je studiji iz 2018. godine dokazano da veliki skupovi podataka (tzv. *big data*) mogu ponuditi korisne uvide u kvantitativne i kvalitativne karakteristike migracija. Unatoč tome, istraživanja o korištenju velikih izvora podataka u području migracija u EU i dalje jedva da postoje – dok takav tip studija na području istočne i jugoistočne Europe do danas nisu učinjene.

Osnovno polazište ovoga rada je da internetski podaci mogu biti transformativni za hrvatsku demografsku znanost, koji će pokušati pokazati da ovakvi pri-

stupi mogu pokrenuti značajna otkrića na području demografije, napose po pitanju vanjskih migracija hrvatskih građana. Uvidima koji se mogu dobiti ovim putem, moguće je otkriti određene obrasce i pravilnosti koje korisnici iskazuju u potrebi za informacijama pri namjeri iseljavanja kao i interese koji odražavaju namjeru migracije. Istraživačima ovakav skup podataka omogućuje uočavanje određenih pravilnosti prema kojima je moguće konstruirati modele predviđanja namjera migracija kao i njihovo pravovremeno uočavanje. Osnovna prednost ovakvog pristupa je omogućavanje pravovremenog uočavanja i utvrđivanja razmjera vanjskih migracija, što je iznimno bitan analitički faktor za razne politike, od tržišta rada do pitanja poduzimanja demografskih revitalizacijskih mjera.

Ovu pretpostavku temeljimo na činjenici da je internetska pretraga često prvo mjesto gdje pojedinci dobivaju odgovore na mnoga pitanja. U situaciji planiranja migracije internetska pretraživanja se koriste uslijed nedostatka ili pak kao nadopuna informacija koje se mogu dobiti od mreže poznanika, rodbine i sl. (vidi Wanner, 2020). Udio korisnika Interneta značajno se povećao od 2009. godine kada je provedena prva studija koja je koristila GT u SAD-u (Ginsberg i sur., 2009), pa su uvidi koji se mogu danas dobiti na taj način puno pouzdaniji. Što se tiče situacije u EU, do sredine 2020. godine 90 % EU populacije bili su korisnici Interneta (Internet World Stats, 2020). U Hrvatskoj udio korisnika iznosi 81% (Eurostat, 2020). Podaci Eurostata pokazuju da je Hrvatska kad je riječ o korištenju Interneta općenito usporediva s prosjekom EU (Jurić, 2021 a). Kako se dob povećava, tako se povećava i postotak građana bez ikakvog iskustva u korištenju Interneta. Odnosno, većina ljudi koji se ne koriste Internetom u Hrvatskoj potječe iz skupine starijih od 65 godina (62 %). Međutim ova činjenica nije ograničavajući čimbenik za našu studiju, budući da je već dokazano da je glavna kohorta hrvatskih iseljenika u dobi od 20 do 40 godina, odnosno da se služi Internetom (vidi: Jurić, 2018.).

Google tražilica najpopularnija je tražilica u Hrvatskoj, koju preferira 97,21% korisnika (StatCounter 2020), slično razini EU (92,92%). Po pitanju FB, ta je društvena platforma kod hrvatskih građana najzastupljenija, što je i objašnjenje zašto smo se u našoj analizi odlučili upravo za ove dvije aplikacije. Ukupan broj aktivnih Facebook korisnika u Hrvatskoj 2019. iznosio je približno 1.900.000 korisnika, s približno jednakom distribucijom korisnika po spolu (vidi: Vuković 2019. i DZS 2019.). Smatramo da je riječ je o izrazito korisnim aplikacijama, koje se, iako su prvenstveno namijenjeni korištenju u marketinške i poslovne svrhe, mogu koristiti i u znanstveno-istraživačke svrhe (vidi Zagheni i sur., 2017).

Ovim radom pokušavamo postaviti empirijsku osnovu za razvoj metode digitalnog praćenja i predviđanja migracija hrvatskih građana, koja se pak može primijeniti i na sve druge populacije Jugoistočne Europe. Uz to nastojimo potaknuti metodološku raspravu o prilikama i izazovima proučavanja digitalnih zapisa u svrhu migracijskih istraživanja na području Jugoistočne Europe.

Osnovna hipoteza rada je da su analitički alati Facebooka i Google Trendsa korisni izvor podataka za utvrđivanje, procjenu i predviđanje vanjskih migracija hrvatskih građana, što ćemo prikazati na primjeru recentnog iseljavanja u Njemačku, Švedsku, Irsku i Austriju. Naglasak rada je na migracijama hrvatskih građana u Njemačku zbog činjenice da se 85 % svih recentnih hrvatskih iseljenika odlučio upravo za tu zemlju (vidi Jurić 2017., 2018.). Smatramo da se glavni doprinos ovog rada osim u popularizaciji ovakvog pristupa odnosi na rezultate dobivene analizom velikih, u probabilističkom smislu nereprezentativnih skupova podataka, kao i prvim predstavljanim uvidima i prognozama o migracijama hrvatskih građana dobivenih iz nestandardnih izvora podataka. U budućnosti će se sve više osjećati potreba za nestandardnim postupcima temeljem kojih će se moći dodatno potvrđivati zakonitosti i spoznaje jer nema naznaka da će u skorijoj budućnosti doći do pomaka u kvaliteti hrvatskog bilježenja vanjskih migracija, dok je isti problem uočljiv i u zemljama Zapadnog Balkana (vidi: Judah, 2019).

Ovdje prikazana metoda bi mogla biti od koristi ne samo za demografska nego i za ekonomska istraživanja jer su demografski i gospodarski razvoj u kauzalnoj povezanosti i prožetosti. Naime, stanovništvo je nositelj gospodarskog razvoja i predstavlja demografski okvir za formiranje radne snage koja pokreće i usmjerava sve djelatnosti u prostoru (Wertheimer-Baletić, 2017). Ovakav pristup omogućuje pravovremeno uočavanje trendova koji ukazuju na smanjenje radne snage i osnovice za njegovu reprodukciju, čime se postavlja podloga za poduzimanje pravovremenih mjera za stabilizaciju hrvatskog tržišta rada te posljedično mirovinskog sustava.

Naime, prirodna i migracijska depopulacija imaju značajne ekonomske posljedice za hrvatsko društvo (Wertheimer-Baletić, 2004). Dok iz sustava odlaze oni koji su neto uplatitelji, u sustavu ostaje sve više onih koji su njegovi korisnici, tako da je danas omjer radnika i umirovljenika 1,24:1 (HZMO, 2021). Starenje stanovništva je pak izravno povezano sa smanjenjem ekonomskih aktivnosti, dok Svjetska banka pokazuje da hrvatske regije u kojima živi starije stanovništvo bilježe znatno niži udio stranih investicija (World Bank, 2020).

Rad je strukturiran na sljedeći način: Nakon što prikažemo relevantne studije iz područja digitalne demografije, prikazat ćemo zasebno metodološki koncept kod pristupa analitičke aplikacije GT te posebno kod analitičke aplikacije FB. U postupku standardizacije podataka prikazujemo kako se indeks dobiven pomoću aplikacije GT može korelirati sa službenim podacima. Mogućnost praćenja korisnika FB prema jeziku i geolokaciji je ključan metodološki dizajn naše studije. Smatramo da su nalazi koje smo dobili indikativni za izradu procjena stanja migracija te su obećavajući alternativni izvor spoznaja o migracijama, ali i njihovih predviđanja što pokazujemo u raspravi rezultata. U posljednjem poglavlju osvrćemo se na ograničenja metodološkog koncepta.

## 2. INOVATIVNI IZVORI PODATAKA U MIGRACIJSKIM STUDIJAMA – DIGITALNA DEMOGRAFIJA

Procjena i utvrđivanje migracijskih tokova je zadatak prepun izazova jer su dostupni administrativni podaci procjene migracijskih tokova unutar EU različiti od države do države te u pravilu koriste neujednačene prakse i metodologije prikupljanja podataka. K tome, budući se podaci o vanjskim migracijskim kretanjima uglavnom oslanjaju na popise stanovništva, koji se pak u većini zemalja održavaju rijetko (najčešće svakih 10 godina), oni često ne bilježe ispravnu distribuciju populacije (vidi: Raymer i sur.; Dubois et all., 2018; Jurić, 2021 f). Nedostatci tradicionalnog prikupljanja podataka su i u tome što su iznimno skupi. Wang i sur. (2015) su pokazali kako se digitalnim alatima mogu prikupiti iznimno vrijedne informacije i to na brži i isplativiji način od tradicionalnih istraživanja. Istraživači su koristili regresijsku analizu i poststratifikaciju na temelju demografskih karakteristika ispitanika u svrhu predviđanje ishoda izbora u SAD-u koristeći kao anketni alat konzolu „Xbox“. Njihova su predviđanja bila gotovo identična rezultatima nacionalno reprezentativnih podataka (Cesare i sur., 2018.).

Osnovni problem hrvatskog bilježenja migracija leži u činjenici da statistiku vodi Odjel za upravne poslove MUP-a, a koju od njega preuzima DZS. Budući se pak veliki broj iseljenika ne odjavljuje pri iseljavanju iz zemlje ta je statistika nedovoljno precizna pa tako hrvatski službeni podaci u slučaju migracija hrvatskih građana u Njemačku pokazuju 40 % odstupanja od podataka EUROSTAT-a. Istovremeno njemački i hrvatski službeni podatci daju potpuno oprečne rezultate koji variraju od godine do godine od 40% do čak 80 % (Pavić, D., Ivanović, I., 2019).

Nakon što su Zagheni i Weber koristili geo-locirane podatke e-pošte *Yahoo!* za procjenu internacionalnih migracijskih tokova, nekoliko se platformi koristilo za razumijevanje mrežne strukture migracija, uključujući FB i *Google+*.<sup>1</sup> Nadalje, geo-podaci korisnika *Twittera* pokazali su se korisnim za proučavanje odnosa unutarne i međunarodne migracije, dok su podaci *LinkedIna* omogućili uvid u globalne migracijske obrasce (Dubois i sur., 2018.). Pri analizi *Yahoo* računa e-pošte u kombinaciji s IP adresama (Zagheni i Weber, 2012) i podataka platforme *Twitter*, (Zagheni i sur., 2014.) istraživački projekti su ukazali na jasne mjerljive podatke u migracijskim obrascima, te su ispravno utvrdili i spol i dob migranata.

I prije ali naročito s pojavom COVID-19 globalno širenje interneta i digitalnih tehnologija radikalno je preobrazio način komunikacije. Sumnjivo obećanje prema kojem je digitalna tehnologija jedini mogući način preživljavanja u vrijeme pandemijske krize COVID-19 svakako će ovaj proces još više ubrzati (vidi Hantrais

---

<sup>1</sup> Platforma je ugašena 2020.

2020). Kao posljedica digitalne revolucije i pandemije, pojedinci su nedvojbeno počeli ostavljati sve veću količinu tragova na društvenim mrežama i Internetu, koji su iznimno zanimljivi društvenim znanostima, napose demografiji.

Podaci dobiveni praćenjem digitalnih tragova osim što su iznimno robusni, lako se prikupljaju i generiraju u stvarnom vremenu te pružaju i vrlo dubok uvid o mišljenjima i interakcijama pojedinaca. U kontekstu brojnih društvenih znanosti, za naglasiti je da ovakav pristup zaobilazi poteškoće u bilježenju stavova i mišljenja oko kontroverznih tema, zbog čega je ovaj pristup u mnogim slučajevima bolji odabir nego tradicionalni upitnici, u kojima ispitanici u pravilu daju socijalno poželjne odgovore (vidi Holbrook i Krosnick, 2010). Štoviše, digitalni tragovi pružaju dokumentaciju o kretanju i aktivnostima, što istraživačima omogućuje zaobilaženje čestih izvora pogrešaka u anketnim podacima, poput nemogućnosti prisjećanja, pristranosti i sl. Konačno, upotreba digitalnih tragova može omogućiti pristup grupama do kojih je teško doći ili koje su općenito nedovoljno zastupljene tradicionalnim tehnikama istraživanja (Cesare, 2018.).

Nekoliko studija koristilo je velike izvore podataka za direktnu analizu pojava povezanih sa migracijama. Veliki doprinos na ovom području dali su Zagheni, Weber i Gumjadi (2017), koji su koristili podatke FB za procjenu stanja međunarodnih migranata u SAD-u. Herdagdelen i sur. (2016), iz znanstvenog tima FB, proučavali su sastav društvenih mreža useljenika u SAD-u koristeći strukturu njihovih prijateljskih veza. Konačno, State i sur. (2014.) te Barslund i Busse (2016.) istraživali su obrasce mobilnosti visokoobrazovanih migranata koristeći podatke LinkedIna.<sup>2</sup>

Zanimljivi uvidi su dobiveni i praćenjem kretanja korisnika mobitela, primjerice uvidom u vremensko trajanje boravka u „roamingu“ (Blumenstock i sur., 2014.). U trenutku dovršetka naše studije, sličnim postupkom dobili smo podatke za hrvatske korisnike prijavljene u „roamingu“ prema kriteriju „boravak izvan zemlje dulje od tri mjeseca“ od teleoperatera *Tele 2* (sada *Telemach*) i *A1*.<sup>3</sup> Prema dostupnim podacima radi se o 91.000 korisnika. Brojne su indicije da se radi o iseljenicima, međutim radi čitavog niza ograničenja ova se stavka ne može precizno izmjeriti, a jednu od prepreka čini i sam GDPR. Stoga smo daljnju analizu vezanu uz ovaj pristup obustavili dok ne usavršimo metodu.

Na kraju ovog poglavlja želimo naglasiti da i podaci prikupljeni alatima digitalne demografije obiluju manjkavostima (vidi: ograničenja) te nam ih nije namjera

---

<sup>2</sup> Za detaljan pregled empirijskih i teorijskih izazova vidi rad „Studying Migrant Assimilation Through Facebook Interests“ (Dubois i sur., 2018.), Cesare i sur. (2018) „Promises and pitfalls of using digital traces for demographic research“ i Jurić (2021 a).

<sup>3</sup> „Hrvatski“ Telekom odbio je sudjelovanje, unatoč svim zajamčenim etičkim postupanjima s podacima.

favorizirati u usporedbi sa službenim podacima. Stava smo da treba učiniti dodatne napore da se osmisle nove metode koje uz postojeće mogu dati dio odgovora na otvorena pitanja te kako bi se pronašla prava mjera i metodologija kombiniranja spoznaja dobivenih iz tzv. „klasičnih“ i „digitalnih“ izvora podataka.

### 3. METODA I PODATCI

Osnovni koncept ovog rada sastoji se u prikupljanju i analizi digitalnih tragova pomoću analitičkih aplikacija GT i FB. Temeljni metodološki koncept kod pristupa metodom GT je praćenje digitalnog traga pretrage po jeziku i geolokaciji. Kod ovakvog pristupa, najvažnije je selektirati riječi koje su prediktivne, specifične i dovoljno uobičajene za upotrebu u predviđanju migracije ali i izabrati pojmove koji se specifično odnose na hrvatske građane. Naime, ovakav pristup otvara pitanje dolaze li pretrage uistinu od građana Hrvatske ili pak populacije drugih državljana koji koristi slične jezike (srpski, bosanski i dr.). Kako bi premostili ovaj problem, tamo gdje je to bilo moguće, izabirali smo skupove pojmova koji se razlikuju na hrvatskom jeziku od istih skupova na srpskom ili bosanskom jeziku.

U postupku standardizacije podataka prikupili smo i obradili podatke za razdoblje od 2010. do 2018., u naznačenim slučajevima i 2019. godine<sup>4</sup> nakon čega smo podijelili indeks pretraživanja za svaku pojedinu ključnu riječ, čime smo dobili indeks učestalosti pretraživanja. Napominjemo da Googleov indeks pretraživanja ne može procijeniti točan broj pretraživanja, niti da svako pretraživanje rezultira migracijom, pa se pomoću ovog alata ne može procijeniti točan broj iseljenika, ali porast trenda može se pratiti vrlo precizno (Jurić, 2021 a). Metodu smo testirali uspoređivanjem rezultata Google indeksa sa službenim njemačkim pokazateljima na osnovu čega smo izračunali korelacije. Ovim postupkom predstavljamo i novi metodološki pristup kako GT indeks komparirati sa službenim podacima. Napominjemo, da se ista metoda može koristiti i za drugačije tipove društvenih istraživanja (primjerice procjenu turističke sezone, natalitet, izbijanje novih žarišta pandemije itd.; Jurić, 2021., b, c, d, e). Nakon analize pojmova, uočili smo da su najprecizniji pojmovi u kontekstu praćenja migracija sljedeći: „misa na hrvatskom“, „katolička misija“ i „hrvatska nastava“. Ove pojmove smo izabrali kao najrelevantnije radi, kako je spomenuto, distinkcije prema ostalim etničkim skupinama s prostora Zapadnog Balkana jer je pretpostavka da pravoslavci i muslimani, odnosno Bošnjaci i Srbi, neće pretraživati ove pojmove.

---

<sup>4</sup> Podatke za 2020. smo izostavili zbog potpuno novih okolnosti vezanih uz COVID-19.

Podaci, koji se mogu dobiti putem FB upravitelja oglasa („Audience insight“) nude iznimno brojne i raznovrsne podatke (koji se mogu besplatno preuzeti), uključujući socio-demografska obilježja o obrazovanju, vrsti zaposlenja te životnim događajima poput rođenja djece, sklapanja braka i mnoge druge. Naime, prilagodba oglašivačke strategije demografskom segmentu, osnovni je cilj FB iz kojeg razloga ova platforma prikuplja podatke o spolu, uzrastu, prebivalištu korisnika, njegovim interesima kao i podacima o podrijetlu (vidi FB ADI). Takva generička segmentacija predstavlja iznimno vrijedne informacije za oglašivače, ali i za demografe. Mogućnost praćenja korisnika FB prema jeziku, u ovom slučaju – hrvatskom jeziku, i geolokaciji je ključan metodološki dizajn i ovog dijela naše studije. Naime, skupno smo pratili sve korisnike koji koriste FB na hrvatskom jeziku u Njemačkoj. Smatramo da su nalazi koje smo dobili indikativni za izradu procjena stanja migracija te su obećavajući alternativni izvor spoznaja o migracijama, ali i njihovih predviđanja što pokazujemo u raspravi rezultata.

Postupak u vezi obrade FB podataka bio je sljedeći: Procjene na državnoj razini hrvatske skupine iseljenika prikupili smo iz baze *Audience insight* i *Facebook Marketing API*, koja podržava povlačenje podataka iz više socio-demografskih kategorija poput obrazovanja, sastava kućanstva, životnih događaja, interesa, itd. (vidi detaljno kod Zagheni i sur., 2017). Iznimno korisnim segmentom pokazali su se podaci o interesima promatrane populacije na temelju „sviđanja“, stranica koje posjećuju i drugih signala (vidi: Dubois i dr., 2018). Taj nam je segment služio kao kontrolni mehanizam, pomoću kojeg smo provjerili radi li se zaista o hrvatskim iseljenicima. Dodatni kontrolni mehanizam nam je bio provjera recenzija koju su korisnici ostavljali, zatim prijavljivanja vlastitih geo-lokacija u Njemačkoj (ali i domovini), posjećenih događaja, recenzija, navođenja osobnih kontakta poput broja telefona koji je bio s njemačkim predznakom i sl.<sup>5</sup> Kako bi testirali metodu i verificirali dobivene podatke, sproveli smo analizu profila FB prijatelja (uz njihovo dopuštenje) koji žive u Njemačkoj. Od 50 sudionika, kod 38 moglo se na osnovu jedne od navedenih stavki ustanoviti da zaista i žive u Njemačkoj, odnosno da su tijekom proteklih godina odselili u Njemačku.

U radu smo u svrhu komparacije serija rezultata koristili „klasične“ demografske podatke dostupne od strane DZS, njemačkih službenih podataka DESTATIS i BAMF, kao i EUROSTAT-a. Nadalje, rezultati su uspoređeni i s rezultatima studije provedene u Njemačkoj na uzorku od 1200 ispitanika kvantitativnom i kvalitativnom metodom (Jurić, 2018). Naime, iako su tradicionalni izvori podataka o migracijama manjkavi, oni su jedini koji mogu verificirati podatke koje dobivamo iz digitalnih izvora pa je ovaj model po prirodi stvari usko naslonjen na „klasične“ izvore podataka.

---

<sup>5</sup> Ovdje još jednom napominjemo da se ne radi o nekoj vrsti špijuniranja, nego isključivo o zbirnim podacima iz kojih se ne može rekonstruirati identitet pojedinačnog korisnika.



### 3.1. Google Trends

Analitička aplikacija Google Trends alat je za pretraživanja trendova koji pokazuje popularnost pojma pri pretraživanju na tražilici *Google*, pri čemu se može vidjeti je li neki određeni trend u porastu ili opadanju (<https://trends.google.com>).

Prva uspješna analiza ovog tipa podataka bila je 2009., a prva studija iz područja migracija ispitivala je tijekom *Migracijske krize 2015.* pretraživanja određenih pojmova na arapskom jeziku u Turskoj i Njemačkoj prema odabranim pojmovima poput „Grčka“ ili „Njemačka“ (Connor, 2016, 2017). Studija *PEW – research* centra pokazala je da digitalni otisci ostavljeni internetskim pretraživanjima mogu pružiti uvid u kretanje migranata. Naime, brojni migranti su za vrijeme svog putovanja 2015. i 2016. koristili pametne telefone koji su pružali pristup informacijama i kartama, kao i savjete o putovanjima putem društvenih medija (Connor, 2017). Tada je nedvojbeno dokazano i da se ovi pokazatelji mogu upotrijebiti za predviđanje migracija.

Pri korištenju ovog alata imali smo na umu da GT ne objavljuje ukupni volumen pretraživanja. Umjesto toga standardizira obujam pretraživanja na skali od 0 do 100 kroz vremensko razdoblje koje se ispituje, s višim vrijednostima koje pokazuju vrijeme kada je opseg pretraživanja bio najveći, čime omogućuju provjerljivu metriku. Pri tumačenju rezultata važno je stoga voditi računa o tome da se mogu dati samo procjene o brzini promjene trenda i usporedbe pojmova, ali se ne može procijeniti ukupan broj pretraživanja, kako je već spomenuto (vidi: Connor, 2017). Ovaj je problem premostiv na način da se detaljnijom analizom trendova u većem vremenskom intervalu rekonstruira broj pretraživanja, a pri tome pomaže i činjenica da GT indeksira samo ona pretraživanja koja prelaze više desetaka tisuća upita.

Treba ipak imati na umu da je svako od tih pretraga provedeno iz vlastitog razloga i ne odgovara na pitanja istraživača, pa tako „guglanje“ primjerice pojma „Njemačka“ nije nužno implikacija da netko želi preseliti u Njemačku, nego se eventualno zanima za turističke informacije ili pak samo traži njemačku *Bundesligu*. Stoga smo iznimno vodili računa o izboru pojmova koji nude najveću moguću prediktivnu razinu. Poseban problem su nam činili pogrešno upisani pojmovi korišteni za pretraživanje jer se pokazalo da (potencijalni) hrvatski iseljenici često neispravno upisuju pojmove. Tako se primjerice pokazalo da se pogrešan pojam „Frankfurt“ češće traži od ispravnog „Frankfurt“ ili pak „Štuttgart“ umjesto „Stuttgart“ itd.

Kod ovakvog pristupa, kako smo već spomenuli, najvažnije je selektirati riječi koje su prediktivne, specifične i dovoljno uobičajene.

*Tablica 1.*

## IZBOR KLJUČNIH RIJEČI I TEMA

Učenje njemačkog jezika; Posao u Njemačkoj; Misa na hrvatskom; Katolička misija; Hrvatska nastava; „Frankfurt“, „Štuttgart“, „Minhen“ (napomena: napisano u izvornom obliku pretraživanja); Frankfurt; München, Stuttgart; Mirovina; Posao u Austriji; Posao u Irskoj; Posao u Švedskoj, smještaj itd.

Jedno od objašnjenja zašto se ovi pojmovi pretražuju na hrvatskom jeziku u Njemačkoj, Austriji, Irskoj i Švedskoj, a ne na jezicima zemlje useljenja, je vrlo vjerojatno povezano s činjenicom postojanja brojnih FB grupa i ostalih društvenih platformi na kojima iseljeni hrvatski građani pokušavaju doći do informacija o životu, radu i sl. u novoj okolini.

### ***3.2. Facebook analitička aplikacija kao metoda za utvrđivanje migracijskih kretanja***

Iako je Facebook najveća društvena platforma, pomoću podataka koji se mogu dobiti njegovom analizom napisano je vrlo malo studija, dok u Hrvatskoj sustavna studija ni ne postoji. Rad na ovom području ograničen je najčešće na projekte koje provode znanstvenici koji su zaposleni pri FB, a rezultati prvenstveno ostaju za internu uporabu. Zaghenni (2017, 2018) je prvi pokazao da FB može biti neprocjenjiv izvor podataka neiskorišten za demografska istraživanja te da nudi skup podataka koji su prema njegovim riječima zapravo neka vrsta neprestano ažuriranog „popisa stanovništva“.

U nastojanju da privuče oglašivače i unaprijedi povrat ulaganja oglašivača, Facebook<sup>6</sup> omogućava oglašivačima da detaljno specificiraju korisnike kojima se njihov oglas treba prikazati. Kako smo spomenuli, socio-demografski podaci uključuju podatke kao što su dob ili spol korisnika, kao i informacije automatski zaključene iz njihovih interakcija na Facebooku, mreže prijatelja i povezanih web mjesta kao i iskazane interese. Kao primjer, FB podržava prikazivanje oglasa isključivo hrvatskim iseljenicima koji su punoljetni i žive primjerice u Ber-

<sup>6</sup> Adverts Manager (FB ADS)

linu, Dublinu ili Beču. Uslijed takve mogućnosti, FB alati<sup>7</sup> nude iznimno korisne uvide i za sama demografska istraživanja.

Za našu studiju naročito je bilo važno vidjeti kako FB prepoznaje hrvatske iseljenike. Osnovni dosad korišten postupak istraživanja migracija na mreži FB bio je u praćenju pojedinaca koji se na vlastitom FB profilu kvalificiraju kao „expats“ („iseljenici“). Međutim, FB pri tome ne daje preciznu definiciju te populacije niti metodologiju kojom se služi u njihovom identificiranju (Zaghenni, i sur., 2017). U FB dokumentaciji kategorijom „expats“ su definirani „ljudi čija je izvorna zemlja prebivališta različita od trenutne zemlje“ (FB API). Na temelju literature koju su objavili istraživači koji interno rade na FB, Zaghenni (2017) je zaključio da su dva čimbenika ključna u Facebookovoj procjeni iseljenih osoba. Riječ je o samoinicijativno upisanom prebivalištu te mjestu rođenja. FB k tome koristi i mrežnu strukturu prijateljstava, pri čemu algoritam procjenjuje je li korisnik iseljenik („expat“). Herdašgdelen i sur. (2017), generirali su procjene iseljenih na temelju niza varijabli pri čemu se pokazalo da postoji čitav niz indicija da se FB zaista koristi spomenutom metodologijom.

#### **4. REZULTATI: GOOGLE TRENDS I FACEBOOK KAO IZVOR PODATAKA O VANJSKIM MIGRACIJAMA HRVATSKIH GRAĐANA**

##### ***4.1. Rezultati analitičkog alata Google Trends***

U nastavku je prikazan pregled geolociranih pretraživanja iz Njemačke, uz određene primjere iz Austrije, Švedske i Irske, čime smo pokušali zaokružiti istraživačku cjelinu i dati naznake ostalim istraživačima za daljnji rad. Ponovno napominjemo da je i u ovom pristupu ključna karakteristika kriterij hrvatski jezik.

Pretraživanje skupine pojmova vezanih uz ključnu riječ „Hrvatska“ od 2014. do 2019. u Njemačkoj prema geolokaciji pokazuje da je ovaj pojam na tražilici Google najviše pretraživan upravo na istim mjestima gdje i službena njemačka statistika<sup>8</sup> potvrđuje da živi najviše hrvatskih građana, odnosno iseljenika, što je na jugu Njemačke.

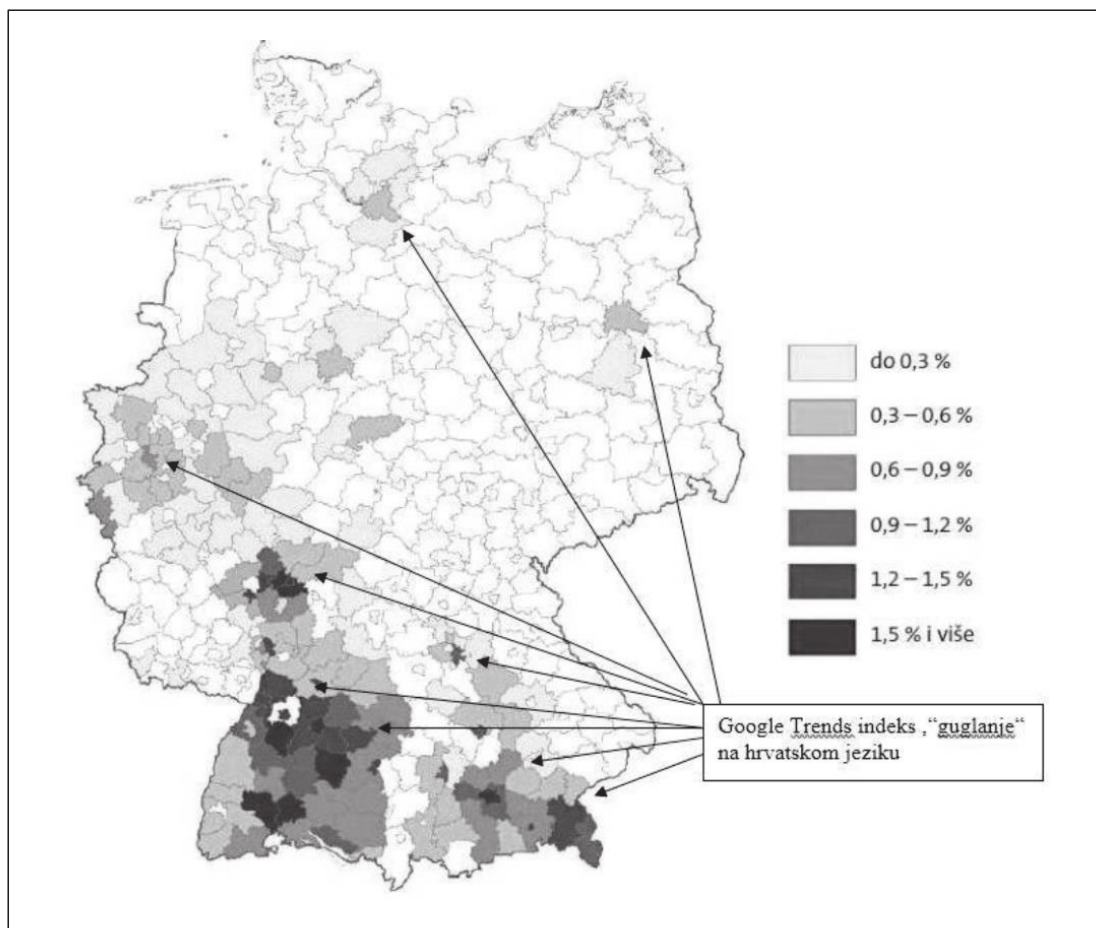
---

<sup>7</sup> Facebook Adverts Manager, Facebook Audiensce Insight te FB API

<sup>8</sup> DESTATIS i BAMF, 2020.

Slika 1.

RASPROSTRANJENOST HRVATSKIH GRAĐANA U NJEMAČKOJ 2017.  
(UDIO PO OKRUZIMA) PREMA SLUŽBENIM POKAZATELJIMA  
I UVIDI DOBIVENI APLIKACIJOM GOOGLE TRENDS (2020.)



Izvor: Autorova izrada prema DESTATIS i BAMF (2018.) te podacima Google Trends (2020.)

Karta prikazuje podudaranje geolokacija u Njemačkoj – na kojima se najčešće *guglaju* pojmovi na hrvatskom jeziku – sa službenom njemačkom statistikom o rasprostranjenosti hrvatskih građana u Njemačkoj. U ovom slučaju vidimo da GT vrlo precizno pokazuje rasprostranjenost hrvatskih građana te da je hrvatskih iseljenika najviše na jugu Njemačke, u pokrajinama Baden-Württemberg i Bavar-ska (što se poklapa sa službenom njemačkom sttistikom). Posebice u gradovima Stuttgartu i Pforzheimu, gdje čine 2,5 % ukupnog broja stanovnika, dok ih je u apsolutnom broju najviše u Münchenu, što ponovno potvrđuju i njemačke statistike

(vidi: DESTATIS, 2017, 2018, 2019.). Analizom Google trendova se nadalje može uočiti porast pretraživanja o uvjetima života i rada u gradu Münchenu, gdje je u posljednjih sedam godina iselilo najviše hrvatskih građana (što se također poklapa s poznatim podacima, vidi: BAMF, 2020; Jurić, 2018).

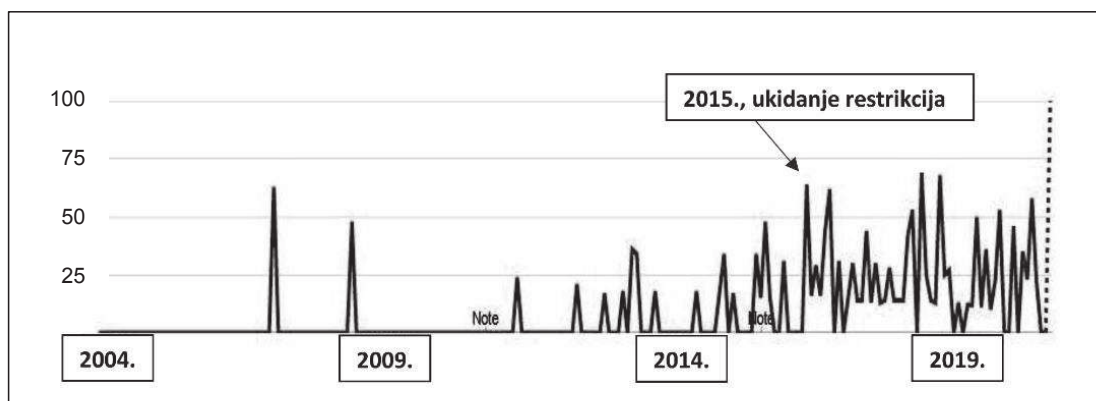
Naša je pretpostavka da je činjenica pretraživanja na hrvatskom jeziku indicija da je riječ o novijim hrvatskim iseljenicima jer bi integrirani hrvatski iseljenici u njemačko društvo ove pojmove pretraživali na njemačkom jeziku. Time smatramo da uklanjamo dio mogućnosti da pretrage dolaze od populacije hrvatskih *gastarbajtera* koji su u Njemačkoj još od vremena 1960-ih godina i njihove (integrirane) djece.

U svrhu testiranja ove metode izabrali smo dva iznimno često pretraživana pojma na hrvatskom jeziku u Njemačkoj: “Hrvatska katolička misija” i “Misa na hrvatskom”. Kako smo naznačili u metodama, ove pojmove smo izabrali kao najrelevantnije iz dva razloga. S jedne strane radi distinkcije s ostalim etničkim skupinama s prostora Zapadnog Balkana, dok s druge strane ova dva pojma odgovaraju i spoznajama iz tzv. „gastarbajterske ere iseljavanja“ kada je uočeno da se hrvatski iseljenici pri dolasku u Njemačku najčešće obraćaju upravo hrvatskim katoličkim misijama za razne vrste informacija kao i zbog potrebe za socijalizacijom i zadovoljenjem religijskih potreba. Ova pretpostavka se naravno ne odnosi na sve skupine novih hrvatskih iseljenika, no budući se 86,28 % hrvatskih građana izjašnjavaju kao katolici (Popis stanovništva, 2011.), smatramo da ova pretpostavka ima smisla. Iako smo uočili da se određeni pojmovi u Njemačkoj pretražuju u znatno većem opsegu od ova dva (primjerice skupina pojmova vezana uz najam stana i potrage za poslom), ne postoji način da se napravi jasna distinkcija između hrvatskog, bosanskog i srpskog jezika u ovom slučaju, stoga ih u svrhu testiranja nismo koristiti.

Već i pri jednostavnoj opservaciji porasta pretrage pojma “katolička misija” uočavamo da se porast te pretrage preklapa s porastom novih hrvatskih doseljenika u Njemačku, a naročito nakon ukidanja restrikcija za rad hrvatskih građana u Njemačkoj od srpnja 2015. godine.

*Slika 2.*

PRETRAŽIVANJE POJMA “KATOLIČKA MISIJA“ NA HRVATSKOM  
JEZIKU U NJEMAČKOJ OD 2004. – 2020.<sup>9</sup>



Dodatno opravdanje za izbor upravo ovog pretraživanja vidimo i u strukturnom uređenju, tradiciji i kohezivnoj snazi hrvatskih katoličkih misija ili župa, koje u Njemačkoj djeluju još od 1960-ih, čime su zasigurno jedna od neprepoznatljivijih institucija u iseljeništvu.

<sup>9</sup> Za potrebe lakše komparacije Google indeksa i službenih podataka, ovdje navodimo podatke o iseljavanju hrvatskih građana u periodu od 2010. do 2020.:

*Tablica 2.*

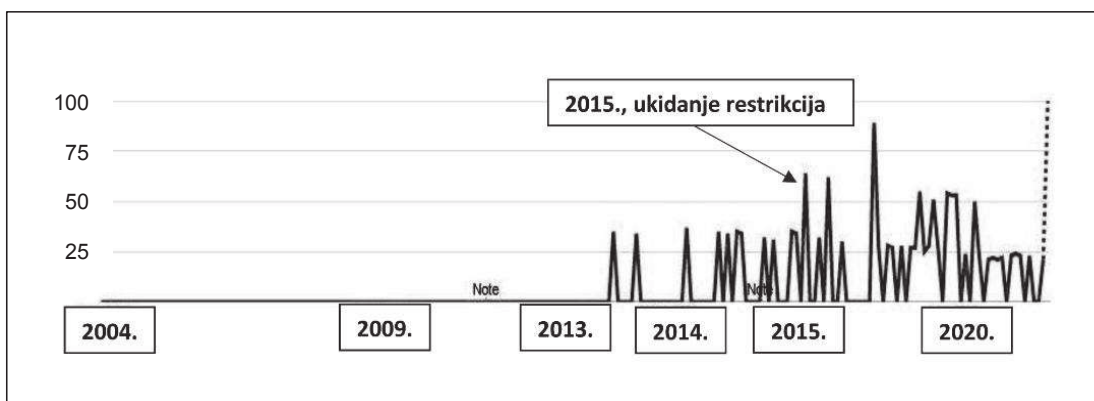
ISELJENI HRVATSKI GRAĐANI U NJEMAČKU 2010. – 2020.

2010.	2011	2012	2013.	2014.	2015.	2016.	2017.	2018.	2019.	2020.
4836	8089	9019	18.633	37.060	50.646	51.163	50.283	48.618	40.151	26.355

Izvor: BAMF, 2017, 2018, 2019, 2020.

Slika 3.

PRETRAŽIVANJE UNOSA „MISA NA HRVATSKOM“ NA HRVATSKOM  
JEZIKU U NJEMAČKOJ OD 2004. – 2020.



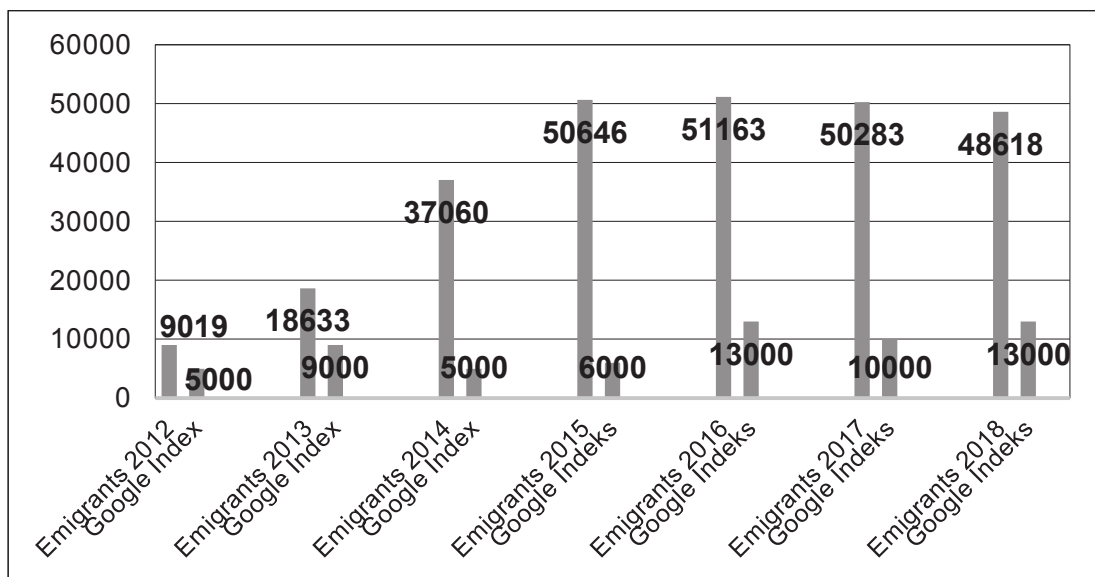
Osim ova dva pojma koji imaju više religijsku konotaciju, dublji uvidi se mogu dobiti i praćenjem trenda pretraživanja unosa „hrvatska nastava“ u Njemačkoj.<sup>10</sup> Od ukupno 328 nastavnih mjesta u cijelome svijetu 167 ih je u Njemačkoj, dok je od ukupnog broja učenika – 6.805 u cijelome svijetu – njih 3.374 u Njemačkoj (Bežen, Bošnjak, 2012; noviji podaci nisu poznati).

U daljnjem postupku standardizacije podataka sakupili smo podatke za razdoblje od 2010. do 2018. i izračunali Google indeks za svaku pojedinu ključnu riječ, čime smo dobili indeks učestalosti pretraživanja. Zatim smo usporedili pretraživanja sa službenim statistikama kako bismo dokazali značaj rezultata.

<sup>10</sup> Hrvatska nastava u inozemstvu – nastava hrvatskoga jezika i kulture, namijenjena je djeci i mladima, hrvatskim državljanima koji privremeno ili stalno žive u drugim državama (vidi: Bežen, Bošnjak, 2012).

*Slika 4.*

USPOREDBA SERIJE PODATAKA IZMEĐU GOOGLE INDEKSA  
ZA PRETRAŽIVANI UNOS/POJAM “HRVATSKA KATOLIČKA MISIJA”  
U SR NJEMAČKOJ NA HRVATSKOM JEZIKU I SLUŽBENE NJEMAČKE  
STATISTIKE ZA GODINE 2010.- 2018.

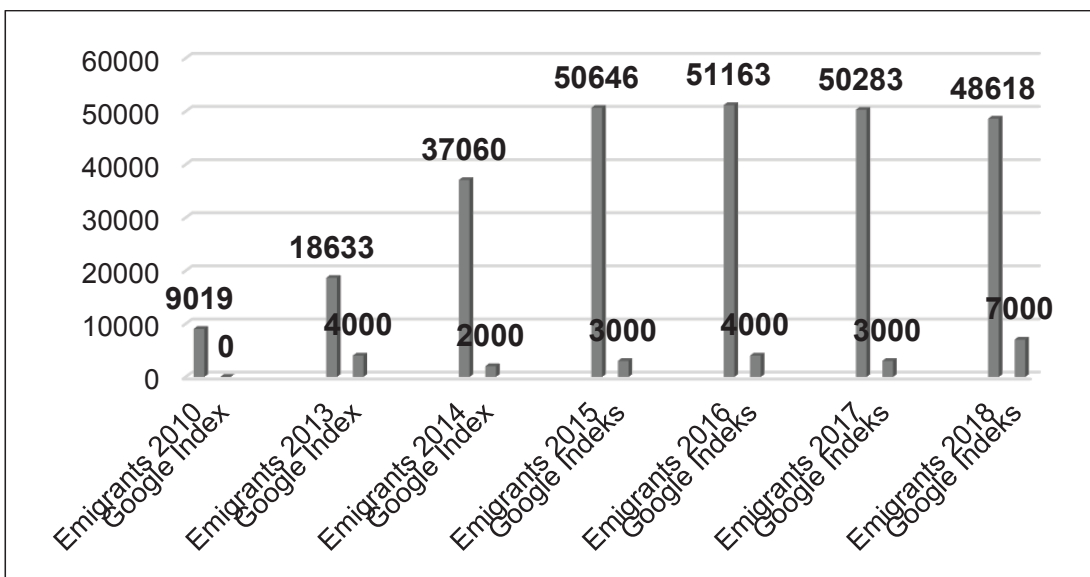


Testirana usporedba serije podataka pokazuje da je porast pretraživanja pojma “Hrvatska katolička misija” povezan s porastom doseljavanja hrvatskih građana u Njemačku.



Slika 5.

USPOREDBA SERIJE PODATAKA IZMEĐU GOOGLE INDEKSA ZA  
PRETRAŽIVANI POJAM „MISA NA HRVATSKOM” U SR NJEMAČKOJ  
NA HRVATSKOM JEZIKU I SLUŽBENE NJEMAČKE STATISTIKE  
ZA GODINE 2010.- 2018.



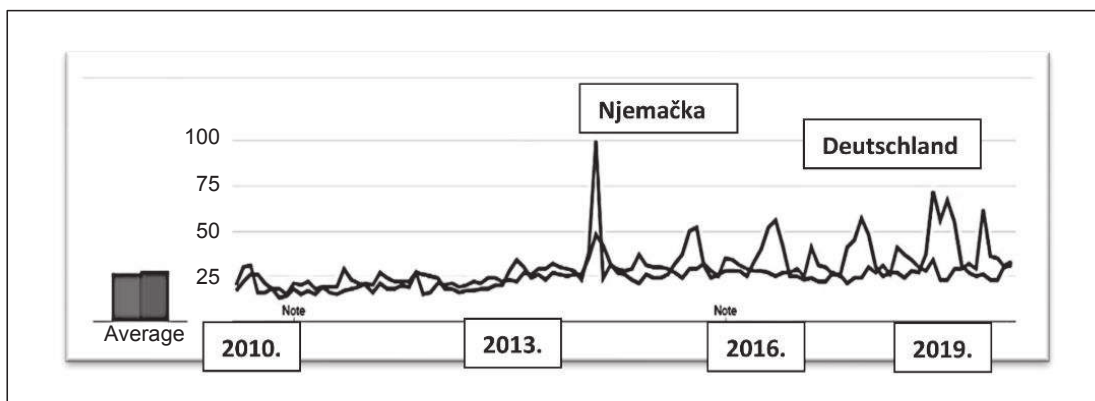
Kod ovog primjera zanimljivo je uočiti da 2010. Google indeks nije uopće indeksirao pojam „misa na hrvatskom“ jer nije bilo dovoljno pretraga tog pojma, da bi od ulaska Hrvatske u EU i nastanka novog vala iseljavanja taj pojam postao jedan od najpretraživanijih pojmova na hrvatskom jeziku u Njemačkoj.

Testirana usporedba serije podataka i kod ovog primjera pokazuju da je porast pretraživanja navedenog pojma povezan s porastom doseljavanja hrvatskih građana u Njemačku.

U nastavku pokazujemo da se provjera može vršiti i obrnutim putem, odnosno iz Hrvatske na njemačkom i hrvatskom jeziku. U dolje navedenom primjeru vidimo kako u Hrvatskoj pretraživanje pojmova vezanih uz tematski kompleks i unose „Njemačka“ i „Deutschland“ naročito raste ukidanjem restrikcija za zapošljavanje hrvatskih radnika u Njemačkoj 2015. godine.

Slika 6.

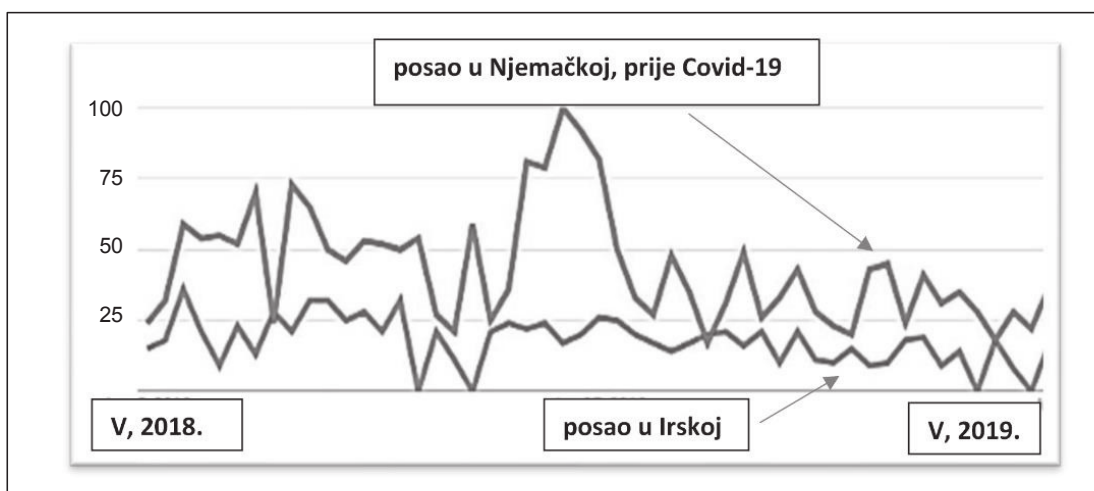
PRIKAZ RASTA TRENTA PRETRAŽIVANJA POJMA „NJEMAČKA“  
I „DEUTSCHLAND“ OD 2010. DO 2020. U HRVATSKOJ



U narednim grafikonima je pak vidljivo da su unosi/pojmovi „posao u Njemačkoj“ kao i s „Posao u Irskoj“ bili redovito pretraživani pojmovi tijekom 2018. i 2019. godine, ali i da je interes za Njemačku bio znatno veći. Ta je činjenica prema našoj metodi trebala rezultirati većim iseljavanjem u Njemačku, koju su pretpostavku naknadno potvrdila službena izvješća Republike Irske (Ured za socijalno osiguranje) i SR Njemačke (BAMF) za 2019. godinu. Nakon izbijanja pandemije COVID-19 smo pak uočili da je interes za Irsku gotovo iščezao, pa se nije očekivao znatan broj novih hrvatskih iseljenika u Irskoj, što su privremena izvješća irskih statističkih podataka iz 2021. ponovno potvrdila (Ured za socijalnu zaštitu Irske, 2021.).

Slika 7.

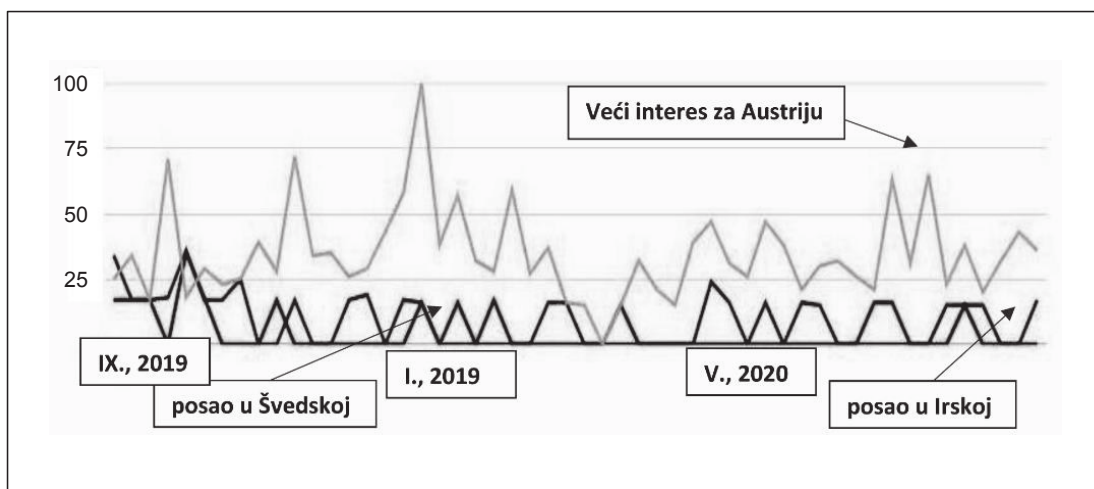
PRETRAŽIVANJE POJMOVA U HRVATSKOJ A) „POSALO U NJEMAČKOJ“,  
B) „POSALO U IRSKOJ“ (V.2018 – V. 2019)



Nadalje se može uočiti da opada interes za iseljavanjem u Švedsku, dok je interes za Austriju u porastu, što pak odgovara službenim podacima austrijske statistike, koje pokazuju da je u 2019. i 2020. porastao broj doseljenih hrvatskih građana u razmjeru od prosječno 3.000 građana više nego u periodima prije 2015. godine (Statistics Austria, 2021.).

Slika 8.

USPOREDBA PRETRAŽIVANJA POJMA „POSAO U AUSTRIJI“ I  
„POS AO U IRSKOJ“, „POS AO U ŠVEDSKOJ“ OD 2019. DO 2020.



Zanimljiv je i uvid da se za Irsku najviše zanima populacija koja živi u Slavoniji. Službeni irski podaci su pokazali da je hrvatskih iseljenika upravo iz Slavonije najviše, čime se uporabna vrijednost ove metode još jednom potvrđuje.

Na primjeru o broju iseljenih iz Hrvatske u Irsku prema podacima DZS i irskim statističkim podacima ponovno se pokazuje da je statistika o vanjskim migracijama Republike Hrvatske iznimno nepouzdana (vidi: Akrap, Strmota, Ivanda, 2017). Prema DZS u Irsku je 2014. godine odselilo samo 128 hrvatskih državljana, te 265 u 2015. godini, dok su irski podaci i podaci FB do 2019. registrirali preko 15 tisuća doseljenih hrvatskih građana, čime se pokazalo da su alati digitalne demografije pouzdaniji izvor stvarnog stanja od službene hrvatske statistike.

I sam „pogled iz Irske“ potvrđuje porast hrvatske populacije u toj zemlji te nalaz da je najviše hrvatskih građana u gradovima Corku i Dublinu (što također potvrđuje i irska službena statistika). Za čitatelja koji se pita u čemu je korist ovih podataka dobivenih pristupima digitalne demografije, budući da te podatke bilježe i vanjski statistički uredi, napominjemo da se s jedne strane ovakvim testiranjem dolazi do potvrde valjanosti podataka dobivenih praćenjem digitalnih tragova, dok s druge takav pristup onda omogućava modeliranje budućih trendova. Glavna uporabna vrijednost se očituje u tome što su podaci GT (kao i FB; vidi u nastavku) pravovremeni, čime se uvidi u migracijska kretanja pristupom digitalne demografije dobivaju godinu dana ranije od službenih podataka, dok su istovremeno u mnogim dokazanim slučajevima precizniji od službenih hrvatskih podataka.

#### 4.2. Facebook analitički alat kao izvor podataka o migracijama

Facebook (2020) registrira između 600 i 630 tisuća korisnika ove aplikacije na hrvatskom jeziku u Njemačkoj, od čega je 70 – 80 tisuća (FB uvijek daje okvirnu procjenu) korisnika koji su na svom profilu izričito naveli da su u zadnjih nekoliko godina doselili u Njemačku (kategorija „expats“). Prednost Facebooka u odnosu na GT je ta da daje preciznije socio-demografske podatke i uvide u brojne interese hrvatskih iseljenika čime se može pratiti njihova integracija. FB, baš kao i službeni njemački podaci, potvrđuje da su najbrojniji iseljenici iz Hrvatske u dobi između 25 i 40 godina. Osim u slučaju Njemačke, FB procjena hrvatskih iseljenika se poklapa i sa demografskim procjenama iz Austrije i Švedske (kao i Irske što je ranije pokazano).

Dok je FB u gradu Münchenu tijekom 2019. godine registrirao između 60 i 70 tisuća Hrvata, 2020. godine je ta brojka porasla za još 10 tisuća (70 – 80 tisuća).<sup>11</sup> Ovdje FB podaci pokazuju preklapanje sa saznanjima iz prethodnih studija (Jurić, 2018.) koje su dokazale da je veći broj hrvatskih građana u Münchenu nego što to pokazuje hrvatska, ali i njemačka statistika.

Provjeru smo izvršili i prema njemačkim pokrajinama, gdje se najviši rezultat korištenja FB na hrvatskom jeziku vidi u pokrajinama Baden-Württemberg, Bavarska i Hessen (između 100.000 i 150.000 u Bavarskoj; između 100.000 i 150.000 u Baden-Württembergu). Usporedivši ove procjene sa službenim njemačkim pokazateljima BAMF i DESTATIS vidimo da podaci po pokrajinama koreliraju sa službenim podacima te da odstupanja gotovo ni nema, odnosno odstupanja iznose oko 5 %, što je i Zagheni u svojim radovima dokazao (Zagheni i sur., 2018.). Analizom smo utvrdili da u najvećem broju slučajeva najniža početna procjena FB odgovara službenim njemačkim pokazateljima. U slučajevima odstupanja, uočili smo da FB pokazuje veći broj korisnika FB na hrvatskom jeziku za 5-10 %, ovisno od grada do grada (primjerice u slučaju grada Kölna i Hamburga), dok je najviše uočeno odstupanje 20 %. Ovdje napominjemo da FB registrira samo punoljetne pojedince pa se i taj podatak treba uzeti u obzir pri konačnom utvrđivanju odstupanja (vidi ograničenja).

Što se tiče Austrije, FB API pokazuje da je broj korisnika FB na hrvatskom jeziku u Austriji čak 200.000 tisuća, što uz sva ograničenja pokazatelja čini znatnu razliku u usporedbi s podacima DZS ali i austrijskih službenih statistika koji tvrde da je u Austriji 89.000 Hrvata (Statistics Austria, 2021.). Podaci pak hrvatskih katoličkih misija u Austriji, baš kao i u Njemačkoj, pokazuju znatno veće brojke.

---

<sup>11</sup> Napomena: Potrebno je voditi vlastitu bazu podataka jer FB ne omogućuje povlačenje podataka za prethodna razdoblja.

Prema tim podacima u Austriji živi najmanje 150.000 Hrvata (samo je u Beču 40.000). Austrijska statistika, naime, Hrvate iz Bosne i Hercegovine bilježi kao građane BiH, bez obzira i na činjenicu ako su živjeli u Hrvatskoj, pa vjerojatno stoga nastaje ovaj brojčani nesrazmjer.

Što se tiče nekih drugih zanimljivih socio-demografskih obilježja koji se mogu uočiti pomoću analitičke aplikacije FB je da je po svemu sudeći jedna trećina svih iseljenih hrvatskih građana u novom migracijskom valu podrijetlom iz BiH. To vidimo po specifičnim interesima koje korisnici pokazuju na svom profilu a vezani su uz specifične karakteristike Hrvata iz BiH (iskazani interesi vezani uz određenu glazbu, tradiciju, mjesta, običaje i sl.; vidi dolje), a isto pokazuju i podaci hrvatskih katoličkim misija.

Testiranje FB podataka nedvojbeno pokazuje preklapanje sa službenim njemačkim podacima ali daje i dodatne zanimljive socio-demografske uvide o iseljenim hrvatskim građanima, koje službena statistika ne bilježi, poput stupnja integracije u njemačko društvo. Naime, pomoću FB se može dosta pouzdano pratiti i stupanj integracije hrvatskih iseljenika u njemačko društvo, naročito putem analize FB interesa (vidi Jurić, 2021 a).

*Tablica 3.*

### NAJČEŠĆI UOČENI INTERESI HRVATSKIH GRAĐANA U NJEMAČKOJ PREMA FACEBOOKU

Učenje njemačkog jezika	FB grupa „Idemo u Njemačku“ i brojne druge tematski slične
informacije i iskustva „zemljaka“ o Njemačkoj	hrvatski portali u Njemačkoj i hrvatski medijski portali u domovini
mogućnost izlazaka u klubove i restorane koje nude „domaću“ glazbu i kuhinju	„ex-yu“ tematski skup
Sv. Ante	mogućnost putovanja u domovinu; informacije na granicama (naročito nakon pojave COVID-19)

## 5. OGRANIČENJA METODOLOŠKOG KONCEPTA

Metoda koju predstavljamo ima važna ograničenja koja želimo istaknuti te ujedno želimo naglasiti potencijal daljnjih istraživanja koja se mogu provesti u ovom području.

Naime, neupitno je da i dalje postoje značajna otvorena metodološka pitanja i upitnost integriteta podataka dobivenih ovim pristupom. Iako podaci dobiveni na društvenim mrežama i webu nude nove mogućnosti za demografska istraživanja i unatoč tome što su to robusni podaci s velikim uzorcima, te da pružaju mnogo-brojnije i raznovrsnije informacije od onih koje se mogu dobiti iz izvješća DZS, EUROSTAT-a i drugih službenih baza, nužno je naglasiti da oni nisu reprezentativni za promatranu populaciju. U znanosti paralelno postoji i općenita rasprava o tome može li se statistički zaključak uopće izvesti iz nereprezentativnih uzoraka. No, na ovo pitanje do danas ne postoji definitivni odgovor. Pitanje reprezentativnosti, međutim, nije svojstveno samo ovom pristupu jer se pristranost može pojaviti i kada se koriste standardni postupci anketiranja, poput telefonskog uzorkovanja (Pettit 2012). U nekim kontekstima istraživanja, nereprezentativnost podataka je pak ključna za dizajn istraživanja, a postoje i teze da analiza novih oblika medija i komunikacije zapravo jako dobro odražavaju postojeće društvene strukture (vidi Cesare et al, 2018).

Glavna ograničenja su sljedeća:

- 1.) Prema onome što znamo, čini se da Facebook klasificira korisnike kao iseljenike prema njihovom rodnom gradu (koji se prijavljuje pri otvaranju računa korisnika) te prema državi prethodnog prebivališta. Budući da se mjesto rođenja korisnika ne podudara uvijek s državom rođenja ili državljanstvom i taj podatak može biti iskrivljen (Spyratos et al, 2018). Nadalje, neki korisnici možda neće objaviti svoje stvarne osobne podatke. Ovaj se problem u slučaju FB može premostiti uporabom dodatnih kriterija za procjenu nalaza, kao što su primjerice prijave geo-lokacija, iskazani interesi i sl.
- 2.) Problem leži i u mogućnosti da korisnici imaju više nepovezanih računa Facebook mreže, što može utjecati na iskrivljavanje podataka. Ono što je ohrabrujuće po ovom pitanju je da je Facebook tijekom 2018. (nakon skandala s *Cambridge Analytica*) izvršio opsežnu kontrolu profila te da je ugasio više milijuna lažnih profila, a naša analiza digitalnih tragova je započela nakon toga. Ovdje je za napomenuti da svakako postoji i određeni broj korisnika FB koji su naša ciljana skupina, ali koriste FB na engleskom ili njemačkom jeziku. Nadalje, određeni korisnici možda nisu hrvatski iseljenici, iako koriste FB na hrvatskom jeziku u Njemačkoj.

Metodom prikazanih osobnih interesa na svojim FB profilima moguće je korigirati ove podatke. Primjerice, takvi iseljenici ne bi pokazivali interes za FB stranicama „sv. Anto“, „Dinamo Zagreb“, hrvatska katolička misija u Njemačkoj i sl. Istovremeno nužno je paziti i na razlikovanje „stare“ i „nove“ hrvatske dijaspore u slučaju Njemačke i Austrije (prikazano prije).

- 3.) Zagheni (2018.) je kontrolnim postupkom pokazao da je odstupanje u njegovom kontrolnom radu bilo 4,7% u usporedbi s podacima koje je prikupio putem FB. To ukazuje da su rezultati dobiveni ovom metodom relativno stabilni (Dubois i sur., 2018). Ipak ta studija je rađena na uzorku u SAD-u gdje je korištenje FB veće nego u hrvatskoj populaciji. Naša kontrola sprovedena je na uzorku od 50 korisnika FB te je pokazala da se podaci FB i naše kontrolne analize preklapaju u 76% slučajeva, odnosno odstupanje je 24%, što je znatno veće od onog kojeg pokazuje Zagheni.
- 4.) Kod FB je posebno problematično što podatke daje samo za aktualan period pa je nužno vršiti stalno periodično praćenje. Problem postoji i u izvozu podataka jer FB ne omogućuje preuzimanje podataka primjerice u Excelu, dok GT to omogućuje. Ovaj se problem može premostiti soft-wareom kao što je primjerice „R.“.
- 5.) Uporaba osobnih podataka iz aplikacija na društvenim medijima izaziva zabrinutost u pogledu otkrivanja osobnih podataka, kao i integriteta i cjelokupnog upravljanja podacima od strane subjekta koji ih prikuplja (vidi: *Cambridge Analytica*). Iz etičkog aspekta, upitno je i koliko je etično koristiti podatke koji nisu prikupljeni uz privolu korisnika – čak i kad se ne znaju osobni podaci kao u slučaju ove studije jer korisnici nisu rabili te aplikacije znajući da će biti predmetom proučavanja.
- 6.) Što se tiče rada s analitičkim alatom Google Trendsa za istaknuti je da se jedan dio porasta trendova pretraživanja može pripisati i generalnom povećanju penetracije Interneta.
- 7.) Poseban problem postoji i u edukaciji samih istraživača koji krenu ovim putem jer moraju biti vješti u programiranju i računskim metodama te vični interdisciplinarnom okruženju.

Ohrabrujuće je da kako GT tako i FB uviđaju određene probleme pri prikupljanju podataka te metodu sve više usavršavaju. Kombinacija tradicionalnih i novih izvora podataka zasigurno je ključna za postizanje napretka u ovom području. Kombinacija vrlo velikog okvira uzorkovanja i razvoj tehnika za prilagođavanje podataka iz nereprezentativnih uzoraka ponudit će neupitno nove mogućnosti za usavršavanje ovog modela (vidi: Zagheni, Weber, Gummadi, 2017).



## 6. ZAKLJUČAK

Službena hrvatska demografska statistika u slučaju recentnog iseljavanja ne omogućava stvarne procjene pa su u znanstvenom pristupu potrebne nove metode istraživanja. U ovoj studiji one su nastale na temelju nestandardnih postupaka uočavanja zakonitosti i spoznaja o migracijama hrvatskih građana temeljenih na metodi digitalne demografije.

Doprinos ovog rada je promoviranje primjene inovativnih izvora podataka u analizi migracija hrvatskih građana i poticanje metodološke rasprave o prilikama i izazovima proučavanja digitalnih zapisa u svrhu migracijskih istraživanja na prostoru Jugoistočne Europe. Međutim, studija ne staje samo na apelu za unapređenje kvalitete podataka o migracijama i popularizaciji novih pristupa u demografiji, nego daje i nove znanstvene spoznaje vezane za stavove, namjere i motive stanovništva vezane uz migracije.

Predstavljeni model doprinosi na način da dokazuje izvedivost mjerenja iseljavanja hrvatskih građana novim nestandardnim pristupom, pomoću analitičkih aplikacija Facebook i Google Trends. Posebnost ovog modela je prvi puta iskušana kombinacija dvaju alata koji se komplementarno nadopunjuju po pitanju praćenja migracija i njihovih prognoza. Ovo je prva studija ovakvog tipa u Jugoistočnoj Europi, ali i prva koja kombinira izvore dvaju baza podataka (FB i GT) kako bi nadopunila određene nedostatke koji karakteriziraju svaku od ovih baza podataka.

Pokazalo se da je FB naročito koristan alat za praćenje socio-demografskih obilježja ali i da može biti koristan u procjeni stupnja integracije hrvatskih iseljenika, dok su alati Google Trendsa prikladniji za predviđanje migracija (vidi detaljno Jurić, 2021 a).

U radu pokazujemo da porast pretraživanja pojmova na hrvatskom jeziku u Njemačkoj: „misa na hrvatskom“, „katolička misija“ i „hrvatska nastava“ korelira s porastom doseljavanja hrvatskih građana u Njemačku. Izvršenim testiranjem uporabne vrijednosti ove metode pokazujemo i novi način i pristup kako se podaci dobiveni putem GT mogu standardizirati u svrhu kompariranja sa službenim bazama podataka.

Ovdje prikazana metoda je od koristi ne samo za demografska nego i za ekonomska istraživanja jer su demografski i gospodarski razvoj u kauzalnoj povezanosti i prožetosti. Osnovna prednost ovakvog pristupa je pravovremeno uočavanje i utvrđivanje vanjskih migracija što predstavlja bitan analitički pokazatelj za razne politike, od tržišta rada do pitanja poduzimanja demografskih revitalizacijskih mjera. Ovakvim pristupom se uvidi u migracijska kretanja dobivaju godinu dana ranije od službenih podataka, čime mogu služiti za modeliranje projekcija i predviđanja daljnjih trendova.

Postoje, međutim, važna ograničenja korištenja izvora velikih skupova podataka i ovakve metode. Stoga svatko tko izabere ovakav pristup treba posvetiti jako puno vremena razmatranju specifičnosti i ograničenja ovakvog pristupa kao i voditi računa o cjelokupnom širem kontekstu istraživanih fenomena.

### *Usklađenost s etičkim standardima:*

U radu koristimo isključivo anonimne, skupne podatke. Svi podaci prikupljeni su u skladu s važećim GDPR-om i etičkim načelima.

### **LITERATURA:**

1. Akrap, A., Strmota, M., Ivanda, K. (2017). Iseljavanje iz Hrvatske od početka 21. stoljeća: uzroci i posljedice, u: Sopta et al: (ur.) Hrvatska izvan domovine II, Zagreb
2. BAMF (2017, 2018, 2019, 2020). Bericht, Nuernberg
3. Barslund, M., Busse, M. (2016). Labour Mobility in the EU. Addressing challenges and ensuring 'fair mobility', No. 139 / July 2016, CEPS Report
4. Bežen, A., Bošnjak, M. (ur.). (2012). Hrvatska nastava u inozemstvu: priručnik za učiteljice i učitelje. MZOS, UFZG. Zagreb
5. Cesare, N., Lee, H., McCormick, T., Spiro, E., Zagheni, E. (2018). Promises and Pitfalls of Using Digital Traces for Demographic Research, Springer, Demography volume 55, p. 1979–1999.
6. Connor, Philip (2017). Can Google Trends Forecast Forced Migration Flows? Perhaps, but Under Certain Conditions, PEW
7. Connor, Philip, (2017). The Digital Footprint of Europe's Refugees, 2017., *PEW RESEARCH*, [https://www.pewresearch.org/\(20.06.2020\)](https://www.pewresearch.org/(20.06.2020))
8. DESTATIS (2019., 2020.), <https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Migration-Integration/Tabellen/auslaendische-bevoelkerung-staatsangehoerigkeit-jahre.html>
9. DZS, *Migracija stanovništva Republike Hrvatske 2011.,; 2016.; 2019.*
10. DZS (2021). Natural change in population, [https://www.dzs.hr/Hrv/publication/FirstRelease/results.asp?pString=Prirodno%20kreta&pSearchString=%20Prirodno%20kreta%20\(01.03.2021\)](https://www.dzs.hr/Hrv/publication/FirstRelease/results.asp?pString=Prirodno%20kreta&pSearchString=%20Prirodno%20kreta%20(01.03.2021))

11. DZS (2019). Primjena informacijskih i komunikacijskih tehnologija (ikt) u kućanstvima i kod pojedinaca u 2019., (12. 11.2020.).
12. Dubois, A., Zagheni E, , Garimella, I. Weber, (2018). Studying Migrant Assimilation Through Facebook Interests, Springer 2018
13. European statistics, [https://ec.europa.eu/info/departments/eurostat-european-statistics\\_hr](https://ec.europa.eu/info/departments/eurostat-european-statistics_hr)
14. EUROSTAT (2020), [https://ec.europa.eu/eurostat/statisticsexplained/index.php?title=Digital\\_economyand\\_society\\_statistics\\_-\\_households\\_and\\_individuals/hr](https://ec.europa.eu/eurostat/statisticsexplained/index.php?title=Digital_economyand_society_statistics_-_households_and_individuals/hr) [accessed 20.02.2021]
15. FB ADI, <https://www.facebook.com/ads/audience-insights/>
16. FB: <https://developers.facebook.com/docs/marketing-api/buying-api/targeting>
17. Google Trends, <https://trends.google.com/trends/?geo=US> (20.08.2020)
18. Herdaždelen, A., Marellib, M. (2016) Social Media and Language Processing: How Facebook and Twitter Provide the Best Frequency Estimates for Studying Word Recognition, <https://onlinelibrary.wiley.com/doi/pdf/10.1111/cogs.12392> (23.07.2020.)
19. HZMO, 2021, <https://www.mirovinsko.hr/hr/statistika/860> (22.05.2021)
20. Internet World Stats, (2020). <https://www.internetworldstats.com/stats.htm> [20.01.2021]
21. Judah, T. (2019). Bye-Bye, Balkans: A region in critical demographic decline, <https://balkaninsight.com/2019/10/14/bye-bye-balkans-a-region-in-critical-demographic-decline/>, 21.10.2020.
22. Jurić, T. (2018). Iseljavanje Hrvata u Njemačku. Gubimo li Hrvatsku? ŠK, Zagreb
23. Jurić, T. (2017). Suvremeno iseljavanje Hrvata u Njemačku: uzroci, motivi i karakteristike, u: Migracijske i etničke teme, Zagreb Vol. 33 No. 3, 337-371.
24. Jurić, T. (2022 a). Forecasting migration and integration trends by using digital demography and big data – the case study of Austria and Germany, Comparative Southeast European Studies, 1-28, <https://doi.org/10.1515/soeu-2021-0090>
25. Jurić, T. (2022 b). Google Trends as a Method to Predict New COVID-19 Cases and Socio-Psychological Consequences of the Pandemic, Athens Journal of Mediterranean Studies, Athens Journal of Mediterranean Studies, Volume 8, Issue 1, January 2022, DOI: 10.30958/ajms/v8i1
26. Jurić, T. (2021 c). Google Trends kao metoda za rano detektiranje pojave novih slučajeva COVID-a 19, Medix, 26/147

27. Jurić, T. (2021 d). Primjena analitičkih alata aplikacija YouTube, Google Photo i Google Web u predviđanju dolazaka turista u Hrvatsku s osvrtom na izazove zdravstvene sigurnosti, *Medix*, 26/147
28. Jurić, T. (2021 e). Posljedice COVID-a 19 na natalitet u Hrvatskoj, *Medix*, 26/147
29. Jurić, T. (2021 f). Razvoj pomoćne metode procjene broja stanovnika: Potrošnja vode kao indikator depopulacije, *Mostariensia* Vol. 25 No. 1.
30. Knežević, A. (2017). Flotantne etničke grupe u demografskim istraživanjima–metodološki problemi, pristupi i primeri, Beograd, DOI: 10.19233/ASHS.2017.31
31. Mellon, J., (2013). Where and When Can We Use Google Trends to Measure Issue Salience? University of Oxford
32. Pavić, D., Ivanović, I., (2019.), Razlike u prikupljanju migracijskih podataka: usporedba Hrvatske i odabranih europskih zemalja, *Migracijske i etničke teme*, 2019, broj 1: 7–32
33. Stocking, G. i Matsa, K. E. (2017). Using Google Trends data for research? Pew Research Center, Apr 27, 2017.
34. Spyrtatos S., Vespe M., Natale F., Weber I., Zagheni E., Rango M., (2018). Migration Data using Social Media a European Perspective, 2018, European Union.
35. Statistics Austria (2021). [https://www.statistik.at/web\\_en/statistics/index.html](https://www.statistik.at/web_en/statistics/index.html)
36. State, Bogdan, Mario Rodriguez, Dirk Helbing and Emilio Zagheni (2019). Migration of Professionals to the U.S. Evidence from LinkedIn data
37. UN (2018). Data Revolution – <http://hdr.undp.org/en/data-revolution> (16.04.2019.)
38. Ured za socijalno osiguranje Republike Irske, <https://www.gov.ie> (16.05.2020)
39. Vuković, I. (2019). Utjecaj društvenih mreža na poslovanje (Thesis), <https://repozitorij.unin.hr/islandora/object/unin%3A3036/datastream/PDF/view> (16.02.2021.)
40. Wang,W., Rothschild, D., Goel, S., & Gelman, A. (2015). Forecasting elections with non-representative polls. *International Journal of Forecasting*, 31, 980–991.
41. Wanner, Ph. (2020), How well can we estimate immigration trends using Google data?, *Quality & Quantity*, <https://doi.org/10.1007/s11135-020-01047-w>
42. Wertheimer-Baletić, A. (2017), Demografska teorija, razvoj stanovništva Hrvatske i populacijska politika, Zagreb
43. World Bank (2020), <https://www.worldbank.org/en/publication/wdr2020>, 11.12.2020.

44. Zagheni, E., Weber, I. (2014). *Inferring International and Internal Migration Patterns from Twitter Data*, Seoul.
45. Zagheni, E., Weber, I., (2017) *Demographic research with non-representative internet data, 2017.*, [www.emeraldinsight.com/0143-7720.htm](http://www.emeraldinsight.com/0143-7720.htm) (14.04.2020.)
46. Zagheni, E., Polimis, M. Alexander, I. Weber, F. C. Billari, (2018). *Combining Social Media Data and Traditional Surveys to Nowcast Migration Stocks*, 2018
47. Zagheni, E., I. Weber, K. Gummadi, (2017). *Leveraging Facebook's advertising platform to monitor stocks of migrants*, *Population and development review* 0(0): 1–14 (xxx 2017)

## FACEBOOK AND GOOGLE AS AN EMPIRICAL BASIS FOR THE DEVELOPMENT OF A METHOD FOR MONITORING EXTERNAL MIGRATION OF CROATIAN CITIZENS

### Summary

Official Croatian demographic statistics in the case of recent emigration do not allow real estimates, so new research methods are needed in the scientific approach. In this study, they were created based on non-standard procedures with the tools of digital demography.

The presented method contributes by proving the feasibility of measuring and forecasting the emigration of Croatian citizens using the analytical applications Facebook and Google Trends. This is the first study of this type in Southeast Europe, but also the first to combine the sources of two databases (FB and GT) to supplement the shortcomings that each of these databases has individually.

The paper shows that the increase in the search for terms in the Croatian language in Germany: "Mass in Croatian", "Catholic mission" and "Croatian teaching" correlates with the increase in the emigration of Croatian citizens to Germany. In this approach, we show a new way in which data obtained through GT can be standardized for the purpose of comparison with official databases.

The main advantage of this approach is the timely detection and identification of external migration, which plays an important role in various policies, from the labour market to the issue of taking demographic revitalization measures. With this approach, insights into migration trends are obtained a year earlier than official data, which can be used to model projections and predict further trends.

Key words: digital demography, Facebook, Google Trends, emigration of Croatian citizens, digital traces, big data