Plan upravljanja istraživačkim podacima

|  |
| --- |
| Opće informacije |
|  | Ime i prezime predlagatelja  | Tena Perišin |
|  | Matična organizacija | Fakultet političkih znanosti |
|  | Naziv projekta | IP-2019-04-6331 Novinarski istraživački laboratorij (Jourlab): Vjerodostojnost medija kroz kulturu eksperimenata i inovacije u redakcijama |
|  | Upravitelj podacima | Dejan Oblak, dejan.oblak@fpzg.hr |
| 1. | Prikupljanje podataka i dokumentacija |
|  | Koje ćete podatke prikupljati, obrađivati, stvarati ili se ponovno njima koristiti? (navedite format, vrstu i opseg podataka) | S obzirom na opseg projekta, za vrijeme trajanja projekta prikupljat će se anketni podaci, podaci iz analize sadržaja medija, podaci intervjua urednika i novinara te podaci iz fokus skupina prikupljenih na medijskim publikama.1. Anketni podaci prikupljeni od ispitanika na individualnoj razini na nacionalno reprezentativnom više-etapnom stratificiranom slučajnom uzorku od 1009 ispitanika. Podaci su prikupljeni F2F, CAWI metodom u prosincu 2020. Podaci su dostavljeni i pohranjeni u .sav formatu spremnom za obradu u SPSS-u te je za potrebe daljnje analize baza podataka prebačena i u .xslx format. Veličina datoteka ne prelazi 10MB.
2. Podaci analize sadržaja središnjih informativnih televizijskih emisija, naslovnica dnevnih novina i online portala koja obuhvaća analizirani uzorak analiziranih medija tijekom 2020. godine. Podaci obuhvaćaju sadržaj tri televizije s nacionalnom koncesijom – HRT, RTL i Nova, te najtiražnije dnevne novine i njihova online izdanja. Dio uzorka obrađen je koristeći online repozitorije (HRT-i, Pressreader) bez trajnog pohranjivanja od strane istraživačkog tima, dok je dio uzorka (NovaTV, RTL, Jutarnji list) ustupljen istraživačkom timu za potrebe analize podataka. Cijele informativne televizijske emisije komprimirane su od strane istraživačkog tima kako bi zauzimale što manje prostora, no u konačnici veličina video/foto uzorka ne prelazi 100GB. Kvantitativna analiza sadržaja provedena je prema kodnom planu gdje je sadržaj kodiran prema kodovima koji obuhvaćaju sedam tematskih cjelina i u unese u datoteku u Excel formatu čija veličina ne prelazi 10MB
3. Stvaranje multimedijskog sadržaja na temelju prethodnih etapa istraživanja s ciljem ispitivanja istog na fokus grupama. Sadržaj će se stvarati ciljano, kako bi se ispitali formati novinarskih priča (audio/video sadržaj) ili elementi unutar novinarskih priča (grafike, novinarski tekst) na način promjene određenih varijabli unutar iste priče. Sadržaj će biti realiziran (kamere, montaže, računala) postojećim sredstvima Televizije Student pri Fakultetu političkih znanosti. Kreirani sadržaj biti će u .mp4, .wav, .jpg ili .xml formatima te ne očekujemo da će stvoreni sadržaj biti veći od 150GB. Prilikom provođenja fokus grupa, planiramo iste dokumentirati uz pripadajuće audio/video zapise (.wav/.mp4) kako bi se izradili transkripti. Veličina tih podataka ne bi trebala prelaziti 50GB, a samih transkripata i analiza 100MB. Transkripti bi se potom transkribirali i kodirali te bi se transkripti fokus skupina čuvali u tekstualnim dokumentima dok bi se kodirani podaci čuvali u dokumentima proizvedenim u programu za kvalitativnu analizu sadržaja.
4. Podaci vezani za polu-strukturirane intervjue novinara i urednika novinskih redakcija hrvatskih medija (HRT,RTL,NovaTV, N1, Večernji list, Jutarnji list, 24sata) te intervjui s novinarima i urednicima medija koji se bave konstruktivnim novinarstvom (BBC,ZDF,DR). Intervjui će biti snimljeni, transkribirani te potom kodirani sukladno kodnom planu. Audio zapis intervjua pohranjivat će se u .wav formatu te transkripti u .docx formatima. Ukupna veličina podataka ne bi trebala prelaziti 50GB.
 |
|  | Kako će se podaci prikupljati, obrađivati ili stvarati? (ukratko navedite metodologiju i procese osiguranja kvalitete te načine organiziranja podataka) | Različite vrste podataka i sadržaja prikupljati će se, obrađivati i stvarati različitim metodama.1. Podaci su prikupljeni od strane agencije IPSOS F2F istraživanjem i potom su dostavljeni istraživačkom timu od strane agencije u .sav formatu spremnom za statističku obradu u programu SPSS u kojem su korišteni za kvantitativnu statističku analizu. Dodatno, podaci su prebačeni u *Excel* formati te je izvršeno čišćenje i pretvorba podataka iz numeričkih vrijednosti u tekstualne zbog daljnje obrade i vizualizacije podataka. Podaci pripremljeni u *Excelu,* zatim su uneseni u digitalni alat *Flourish****,*** gdje je pomoću predloška za analizu ankete (*Survey 9.3.2)* omogućena interaktivnost podataka te analiza s grafičkim prikazom u realnom vremenu. Sve datoteke iz svih koraka obrade pohranjene su na zajedničkom *Sharepoint-u* Fakulteta političkih znanosti dediciranom za realizaciju projekta kojem imaju pristup članovi istraživačke skupine. Interaktivni rezultati u grafičkom obliku dostupni su na repozitoriju.
2. Za analizu sadržaja medijskih sadržaja istraživačka skupina kreirala je dvije matrice (Televizijske emisije i Naslovnice novina i Online izdanja novina). Svaka matrica sadrži 92 kategorije i 679 indikator-čestica te zaseban dokument – kodnu knjigu. Istraživačka skupina se podijelila u dvije grupe, jedna koja je analizirala televizijske emisije i drugu koja je analizirala naslovnice dnevnih novina i njihovih online izdanja. Nadalje, uzorak unutar svake grupe, podijeljen između istraživačka nasumičnim odabirom datuma i medija, te je napravljeno pilot istraživanje u nekoliko faza izrade matrice. Nakon finalizacije matrice i posljednjeg pilot istraživanja, napravljena je i provjera dosljednosti kodiranja (*intecoder reliability*) između svih članova istraživačke skupine. Nakon završetka analiza, svaki tim objedinio je matricu, pročistio podatke i izvršio provjeru podataka te su podaci analizirani koristeći *SPSS* i *Excel.* Kao i kod analize provedene ankete podaci su prebačeni u *Flourish* koji je omogućio daljnju analizu i vizualizaciju podataka u realnom vremenu. Sve matrice i radni dokumenti pohranjeni su na *Sharepointu* kojem imaju pristup svi članovi istraživačke skupine.
3. Eksperimentalne medijske oblike koji će se testirati na fokus grupama proizvest će članovi istraživačke skupine na temelju prethodnih faza istraživanja, kao i kroz rad sa studentima na kolegiju Istraživački novinarski laboratorij. Tijekom 2 godine kolegija nastali su brojni eksperimentalni novinarski oblici (interaktivne grafike, infografike, video sadržaji) koji za cilj imaju privlačenje mlade publike novinarskim sadržajima. Eksperimentalni medijski oblici redovito su tijekom rada objavljivani na multimedijskim portalima projekta ([www.zagrebnewslab.eu](http://www.zagrebnewslab.eu)) te multimedijskog portala kolegija ([www.novinarz.online](http://www.novinarz.online)). Novinarske oblike izrađene tijekom kolegija istraživački tim podijelit će u kategorije i prema njima napraviti metodološki koncept potreban za testiranje istih na fokus grupama. Novinarski oblici biti će podijeljeni u skupine prema načinu prezentacije informacija (video, tekst, infografika i sl.) te će biti izrađene varijacije odabranog dijela uzorka (primjerice izmjene u količini grafičkog prikaza, izmjene u načinu novinarskog teksta i slično) od strane istraživačkog tima. Nakon definiranja uzorka koji će se testirati pristupiti će se fokus grupama koje će se održati uživo (ili online zbog epidemiološke situacije) na ispitanicima koji pripadaju različitim dobnim skupinama te će se fokus grupe snimati za potrebe izrade transkripata i kodiranja te daljnje analize. Dobiveni podaci pohranjivati će se na *sharepointu* za potrebe projekta, kao i kreirani uzorak na temelju kojeg su provedene fokus grupe.
4. Tematske analize intervjua s novinarima i urednicima domaćih i stranih medija provodit će se tijekom cijelog razdoblja projekta zbog otežanog pristupa redakcijama i medijima tijekom pandemije. Intervjui uživo snimati će se audio snimačem, a intervjui provedeni online biti će snimljeni putem platforme Zoom. Svi intervjui biti će transkribirani i kodirani te analizirani, a podaci iz svih faza će biti pohranjeni na zajedničkom *sharepointu* istraživačke skupine te sortirani prema medijima i tematskim cjelinama. Tematska analiza vršiti će se uz korištenje *NVivo* programa za kvalitativnu analizu sadržaja.
 |
|  | Koju ćete dokumentaciju i metapodatke ustupiti osim podataka? (navedite koje su informacije potrebne korisnicima kako bi mogli čitati i interpretirati podatke u budućnosti te koji će se standardi koristiti pri tumačenju podataka) | Uz sve podatke nastale tijekom projekta istraživačka skupina priložiti će sljedeću dokumentaciju i metapodatke:1. Uz podatke ankete priloženi će biti anketni upitnik kojeg je izradila istraživačka skupina i izvještaj o karakteristikama uzorka. Također na repozitoriju stranica projekta rezultati će biti dostupni u interaktivnom grafičkom obliku gdje korisnik sam bira između svih pitanja odnosno varijabli te je u mogućnosti sam radi analizu koja uključuje jednu ili više varijabli uz detaljno objašnjenje kako pretraživati podatke. Također će podaci biti dostupni za preuzimanje u .CSV i .SAV formatu.
2. Uz rezultate analize sadržaja priložena će biti kodna knjiga s opisom varijabli i oznaka vrijednosti. Na repozitoriju stranica projekta rezultati će biti dostupni u interaktivnom grafičkom obliku gdje korisnik sam bira između svih pitanja odnosno varijabli te je u mogućnosti sam raditi analizu koja uključuje jednu ili više varijabli uz detaljno objašnjenje kako pretraživati podatke. Također će podaci biti dostupni za preuzimanje u .CSV i SAV formatu.
3. Uz podatke priložit će se i kodni plan za kvalitativnu analizu sadržaja s opisom provedene analize. Rezultati istraživanja medijskih sadržaja na fokus grupama biti će dostupni u interaktivnom grafičkom obliku uz mogućnost pregledavanja testiranih sadržaja koje je izradila istraživačka skupina. Demografski podaci sudionika fokus grupa biti će dostupni te će podaci biti dostupni u .CSV obliku dok će transkripti i rezultati kodiranja biti dostupni u .DOC formatu.
4. Podaci vezani uz intervjue biti će dostupni uz opći protokol intervjua te dokumentaciju koja sadrži opis ovog dijela projekta, ciljeve, načinu prikupljanja podataka, vremenskom te prostornom obuhvatu i informaciju o mediju s kojeg intervjuirana osoba dolazi. Transkripti i podaci proizašli iz kodiranja biti će dostupni u .DOC formatu.
 |
| 2. | Pravna i sigurnosna pitanja |
|  | Jeste li ograničeni sporazumom o povjerljivosti? Imate li potrebna dopuštenja za prikupljanje, obradu, čuvanje i dijeljenje podataka? Jesu li osobe čiji se podaci pohranjuju informirani o tome i jesu li dali privolu? Kojim ćete se metodama koristiti u svrhu zaštite osjetljivih podataka (GDPR - posebne kategorije osobnih podataka)? | Pri analizi i prikupljanju podataka svi ispitanici će biti anonimizirani te će bilo koji podatak koji može identificirati identitet ispitanika biti uklonjen i baze podataka. Svi podaci prikupljeni na projektu su izvorni rezultat rada na projektu, i osim ograničenja koja se odnose na zaštitu identiteta ispitanika drugih ograničenja vezanih uz korištenje podataka nema. Svaki ispitanik je dao privolu za korištenje anonimiziranih podataka u analizi. Budući da su osjetljivi podaci vezani jedino uz identitet ispitanika, svi osobni podaci te ostali podaci koji mogu ukazivati na identitet ispitanika će biti uklonjeni iz baza podataka pohranjenih u repozitorijima i dostupnih javnosti.  |
|  | Kako će se regulirati pristup podacima i njihova sigurnost? Koji su potencijalni rizici koje treba uzeti u obzir? Kako ćete osigurati sigurnost pohrane osjetljivih podataka?  | Podaci će se pohraniti u centraliziranom sustavu za pohranu – *OneDrive Sharepoint* kojim upravlja Odjel za informatiku Fakulteta političkih znanosti. Pristup podacima se upravlja preko AAI identiteta, koji je siguran sustav i slijedi najbolje prakse u pogledu upravljanja identitetom. Naš centralni sustav pohranjivanja podataka ima dostatan prostor za pohranu podataka, vrši se zrcaljenje i stalno je nadziran.Podaci će biti dostupni registriranim korisnicima s AAI identitetom te drugim korisnicima iz znanstvene i istraživačke zajednice na temelju upita. |
|  | Kako ćete upravljati zaštitom autorskih prava i intelektualnog vlasništva? Tko će biti vlasnik podataka? Koje će se licencije primjenjivati na podatke? Koja će se ograničenja primjenjivati na ponovnu uporabu osobnih podataka? | Podaci su prikladni za dijeljenje i od interesa zainteresiranoj javnosti i medijima te će kao podaci prikupljeni iz projekta financiranog javnim sredstvima biti dostupni svim pripadnicima znanstvene i istraživačke zajednice uz navođenje JOURLAB projekta kao izvora podataka. Osobni podavi ispitanika neće biti dio baza podatka u niti jednom obliku te neće biti dostupni korisnicima. Naš cilj je omogućiti najširu moguću upotrebu podataka i objavit ćemo ih pod Crative Commons CC0.  |
| 3. | Pohrana i čuvanje podataka |
|  | Kako će podaci biti pohranjeni i kako će biti napravljena sigurnosna kopija podataka (*backup*) tijekom istraživanja? Koji su kapaciteti čuvanja podataka kojim raspolažete? Kojim se procedurama koristite za sigurnosnu kopiju (*backup*)? | Svi podaci pohranjeni su na centraliziranom sustavu za pohranu Fakulteta politčkih znanosti – *Sharepoint* koji ima dostatan prostor za pohranu svih baza. Također osoba za upravljanje podacima izrađuje periodične kopije na vlastitom računalu te jedna sigurnosna kopija radi se na vanjskom disku koji sadrži sve podatke vezane uz projekt veličine 4TB.  |
|  | Koji je vaš plan čuvanja podataka? U kojim će se formatima čuvati? | Podatke ćemo trajno pohraniti u Hrvatskom arhivu podataka za društvenu znanosi (CROSSDA; data.crossda.hr) koji se bavi pripremom skupova podataka za dugoročnu pohranu. Skup podataka čuvat će se u otvorenim formatima za dugoročnu pohranu, uz detaljni metapodatkovni opis. Skup podataka sadržavat će dokumentaciju koja će omogućiti kasniju interpretaciju podataka i mogućnost ponovne upotrebe.Također podatke ćemo čuvati na virtualnom poslužitelju kojeg nam je Fakultet političkih znanosti dodijelio za potrebe projekta – *Sharepoint OneDrive.* Sigurnosna kopija svih podataka biti će pohranjena i na vanjskom tvrdom diskuPodaci će biti javno dostupni na kreiranom repozitoriju na mrežnim stranicama projekta  |
| 4. | Dijeljenje i ponovna uporaba podataka |
|  | Kako i gdje će se podaci dijeliti? Na kojem repozitoriju planirate dijeliti podatke? Kako će potencijalni korisnici doznati za podatke? | Podaci i dokumentacija potrebna za interpretaciju podataka bit će dostupni u Hrvatskom arhivu podataka za društvene znanosti (CROSSDA; data.crossda.hr). Arhiv CROSSDA je nacionalni pružatelj usluga za Konzorcij europskih arhiva podataka za društvene znanosti (CESSDA ERIC). Time se povećava međunarodna vidljivost skupova podataka relevantnoj međunarodnoj zajednici istraživača. CESSDA Data Catalogue (datacatalogue.cessda.eu) uključen je u EOSC portal. Detaljno pretraživanje skupova podataka omogućeno je upotrebom DDI standarda za metapodatke, uz upotrebu kontroliranih rječnika i tezaurusa ELSST (elsst.cessda.eu) za podatke nastale u društvenim istraživanjima. Skupovima podataka dodjeljuje s trajni identifikator iz sustava DOI.Također dio podataka biti će dostupan na repozitoriju stranica projekta ([www.zagrebnewslab.eu](http://www.zagrebnewslab.eu)) u interaktivnom grafičkom obliku te će diseminacijom projekta i njegovih rezultata stranica biti istaknuta i aktivno promovirana. |
|  | Ako postoje podaci koji se ne smiju dijeliti (prijavitelji vezani zakonskim, etičkim, autorskim pravila, povjerljivošću i sl.), pojasnite razloge ograničenja. | Nema ograničenja za dijeljenje podataka.  |
|  | Potvrdite da ćete se koristiti digitalnim repozitorijem koji je u skladu s načelima *FAIR-a*. | Arhiv CROSSDA postupa s podacima prema načelu FAIR. |
|  | Potvrdite da ćete se koristiti digitalnim repozitorijem koji održava neprofitna organizacija (ako ne, objasnite zašto ne možete dijeliti podatke na digitalnom repozitoriju koji nije komercijalan). | Arhiv CROSSDA dio je nacionalne javne istraživačke infrastrukture za područje društvenih znanosti. Osnovan je pri Filozofskom fakultetu Sveučilišta u Zagrebu uz financijsku potporu Ministarstva znanosti i obrazovanja u ulozi pružatelja usluge za Konzorcij europskih arhiva za društvene znanosti (CESSDA ERIC) |

Ref:

[1] Celjak, D., Dorotić Malič, I., Matijević, M., Poljak, Lj., Posavec K. i Turk, I.: „Istraživački podaci - što s njima?“ [Istraživački podaci - što s njima? : priručnik o upravljanju istraživačkim podacima | Digitalni repozitorij Srca (unizg.hr)](https://repozitorij.srce.unizg.hr/islandora/object/srce%3A327)