

Filozofski fakultet Sveučilišta u Zagrebu

Plan osnivanja nacionalnog arhiva podataka za društvene znanosti

Veljača, 2019.

Sadržaj

Uvod	2
Osnovna obilježja arhiva podataka	4
Definicija organizacije	4
Misija	4
Obuhvat i vrste podataka	5
Usluge i aktivnosti	5
Korisnici	6
Upravljanje	6
Financiranje	7
Postojeća iskustva u Hrvatskoj	7
Potrebni resursi i organizacija	9
Sjedište arhiva podataka	10
Osoblje	10
Partnerstvo i suradnja	11
Tehnička infrastruktura	13

Uvod

Empirijska istraživanja ključna su komponenta rada u društvenim znanostima. Primjenom širokog spektra metoda, modela i teorijskih pretpostavki, cilj je doseći spoznaje o društvenoj zbilji. Međutim, takva istraživanja su, pogotovo ukoliko teže reprezentativnosti i analizi složenih fenomena, izuzetno opsežni te organizacijski i financijski zahtjevni poduhvati. Osim izvještaja i znanstvenih radova, iza svakog takvog istraživanja ostaje i baza prikupljenih podataka. No, za razliku od izvještaja i radova, koji su dostupni u knjižnicama, sustav koji omogućuje pohranu i korištenje podataka iz prethodnih istraživanja u području društvenih znanosti u Hrvatskoj još ne postoji.

Stoga predlažemo **osnivanje arhiva podataka koji će pohranjivati i objavljivati istraživačke podatke u području društvenih znanosti, pružati ih na korištenje i brinuti o njihovom dugoročnom očuvanju** i tako služiti istraživačima, nastavnicima, studentima i ostalim zainteresiranima za podatke o društvenim pojavama i procesima.

U ovome dokumentu najprije ćemo obrazložiti bitnu motivaciju na kojoj su utemeljene sve dosadašnje i planirane aktivnosti oko uspostave arhiva podataka za društvene znanosti. Nakon toga, iznijet ćemo zamišljeni koncept arhiva podataka kroz opis njegovih osnovnih obilježja poput: usluga koje pruža, aktivnosti kojima se bavi, zbirki podataka o kojima brine, korisnika kojima služi, načina na koji se treba upravljati organizacijom te načina financiranja. Taj koncept osnova je za detaljno planiranje aktivnosti prilikom uspostave arhiva podataka. U narednom poglavlju opisat ćemo postojeće resurse, odnosno iskustva i rezultate prethodnih projekata u kojima je Filozofski fakultet Sveučilišta u Zagrebu (FFZG) sudjelovao kao partner zajedno s članovima CESSDA ERIC konzorcija. Na kraju ćemo razmotriti resurse i druge organizacijske elemente koje je potrebno osigurati u procesu osnivanja arhiva podataka kao nacionalnog javnog servisa za društvene znanosti, te ih staviti u kontekst onoga što već postoji u nacionalnom okruženju, kako bi se osigurao kontinuitet aktivnosti u koje se ulagalo posljednjih nekoliko godina.

Motivacija

Motivacija za uspostavljanjem arhiva podataka proizlazi s jedne strane iz vrijednosti prikupljene građe i njenog potencijala za ponovno korištenje te same prirode znanstvenog rada, a s druge strane iz svijesti o potrebi da podaci budu uredni i jasno dokumentirani te dugoročno očuvani kako bi bili upotrebljivi i u dalekoj budućnosti. Osim toga, važno je i uključivanje u međunarodne mreže sličnih organizacija kako bi se potaknula međunarodna razmjena znanja.

Vrijednost materijala za sekundarnu analizu

Materijal prikupljen empirijskim istraživačkim radom u pravilu je bogat sadržajem tako da izvorne analize jednostavno ne mogu konceptualizirati niti testirati sve hipoteze i modele koji se mogu izraditi na temelju prikupljenih podataka, kao niti obratiti detaljnu pozornost na sve prikupljene varijable. Dostupnost izvornih podataka prethodnih istraživanja omogućuje sekundarnu analizu kroz nove tematske, teorijske i analitičke pristupe i predstavlja vrlo vrijedan izvor informacija za znanstvenike u društvenim znanostima te vrlo bitnu točku osnaživanja društvenih znanosti. Tim više što, uzevši ograničenost resursa kao datost, ovakav istraživački proces ne zahtjeva ponavljanje izuzetno skupog i zahtjevnog terenskog prikupljanja podataka.

Osnivanje arhiva podataka olakšat će izradu empirijski utemeljenih analiza, kako kod znanstvenika-istraživača čija bi se akademska produktivnost tako povećala, tako i kod studenata preddiplomskog,

diplomskog i poslijediplomskog studija koji za svoje završne radove rijetko mogu prikupiti visokokvalitetne izvorne podatke.

Kumulativna priroda znanstvenog rada: konzistentnost, nadogradnja i verifikacija

U većini okolnosti, znanost je kolektivni poduhvat u kojemu se spoznaje gomilaju doprinosima znanstvenika u obliku prethodnih radova. Dostupnost tih prethodnih radova nužan je preduvjet za daljnji razvoj i utemeljenu kritiku postignuća na bilo kojem području. To se podrazumijeva kad su u pitanju izvori poput recenziranih članaka, knjiga, izvještaja i drugih publikacija, no dostupnost nije ništa manje bitna niti kada su u pitanju sami podaci na kojima se istraživanja temelje.

S jedne strane, sekundarna analiza izvornih podataka omogućava nadogradnju izvornih zaključaka, testiranje alternativnih hipoteza na temelju istih podataka, pa i verifikaciju izvornih analiza. Svi ovi procesi mogu samo doprinijeti kvaliteti znanstvene produkcije.

S druge strane, dizajn novih empirijskih istraživanja može biti značajno konzistentniji i kvalitetniji ukoliko se zasniva na temeljitom uvidu u konstrukciju i karakteristike instrumenata prethodnih istraživanja (što može predstavljati svojevrсно pred-testiranje). Na taj se način mogu izbjeći mnoge prethodno učinjene greške u dizajnu, a rezultirajuće serije rezultata u različitim sredinama ili vremenskim razdobljima mogu biti u većoj mjeri usporedive, što povećava upotrebljivost i starih i novih istraživanja te omogućuje meta-analizama temeljenima na agregiranju više baza podataka.

Sustavno i pouzdano očuvanje podataka radi ponovnog korištenja

Zbog nepostojanja institucije koja sustavno pohranjuje podatke, mnoga opsežna istraživanja iz proteklih nekoliko desetljeća potpuno su izgubljena za naknadne analize, dio postoji samo u fragmentima, a u potrazi za ostatkom mora se potrošiti velika količina vremena. Spašavanje podataka od "zuba vremena" te očuvanje (socijalno) povijesno vrijedne građe bitna je uloga arhiva istraživačkih podataka. Arhivirani podaci imaju ključno mjesto u analizama prethodnih društvenih konfiguracija, jer se naknadno prikupljeni podaci mogu tek u vrlo ograničenoj mjeri koristiti za analizu prethodnog stanja. Dakle, arhiv je nužan kako društvene znanosti ne bi postojale u stanju permanentne amnezije.

Ustanovama koje se bave financiranjem i vrednovanjem znanstvenog rada, arhiv podataka treba biti partner u ostvarivanju ciljeva koji se odnose na unapređenje znanstvene informacijske infrastrukture, u podršci otvorenom pristupu znanstvenim informacijama i općenito u osiguravanju povoljnih uvjeta rada istraživačima. Financijeri istraživanja i uredništva časopisa sve češće zahtijevaju od istraživača da pohranjuju podatke radi njihovog ponovnog korištenja, a postojanje arhiva podataka omogućuje ostvarenje takvih zahtjeva. Budući da se sustavnim arhiviranjem osigurava upotrebljivost već prikupljenih podataka u novim istraživanjima, ulaganje u razvoj arhiva podataka je i ekonomski isplativo¹ jer korištenje postojećih podataka može znatno smanjiti cijenu istraživačkog procesa, a pažnja se može više usmjeriti na nove spoznaje i znanstvene doprinose.

Uključivanje u paneuropske istraživačke infrastrukture i svjetsku "obitelj" arhiva podataka

U većini europskih, pa i tranzicijskih zemalja (izuzetak su npr. Albanija, BiH, Bugarska, Crna Gora, Kosovo, Latvija, Makedonija i Hrvatska), osnovani su arhivi podataka za društvenih znanosti. Na europskoj razini, takvi arhivi podataka okupljaju se u **CESSDA ERIC - Konzorcij europskih arhiva podataka za društvene znanosti**. CESSDA je dio paneuropske istraživačke infrastrukture, a ESFRI

¹ Beagrie N., Houghton, J., Palaiologk, A., and Williams, P. (2012). *Economic Impact Evaluation of the Economic and Social Data Service*. <http://www.esrc.ac.uk/files/research/research-and-impact-evaluation/economic-impact-evaluation-of-the-economic-and-social-data-service/>

Roadmap iz 2018². ju identificirana kao "*landmark*" što znači da se radi o potpuno implementiranoj infrastrukturi koja već pruža usluge korisnicima i čiji rad Europska komisija podupire kroz infrastrukturne projekte namijenjene upravo razvoju takvih istraživačkih infrastruktura. Članstvo u CESSDA ERIC ostvaruje se na nacionalnoj razini, preko Ministarstva znanosti i obrazovanja, a uvjet za članstvo je imenovanje nacionalnog pružatelja usluga koji je sposoban zadovoljiti odgovarajuće profesionalne i organizacijske kriterije. Članstvo u CESSDA ERIC je za Republiku Hrvatsku važno jer se na taj način omogućuje razvoj usluga arhiva podataka uz podršku međunarodne zajednice s velikim iskustvom u ovome poslu, što je od posebnog značaja jer za osoblje koje treba biti zaposleno u arhivu podataka ne postoji u ovome trenutku adekvatno formalno obrazovanje. CESSDA svake godine koordinira i prijavljuje nekoliko EU projekata u kojima predstavnici zemalja članica mogu direktno sudjelovati, što je način za dodatno financiranje rada arhiva podataka. Općenito, članstvo u ESFRI projektima i ERIC konzorcijima je jedan od načina za ostvarivanje integracije u europski istraživački prostor, a kroz članstvo u CESSDA ERIC moguće je ostvariti veću vidljivost hrvatskih znanstvenika u tom prostoru, ali i samim istraživačima omogućiti pristup nekim setovima podataka kojima sada ne mogu pristupiti.

Osnovna obilježja arhiva podataka

U ovome poglavlju opisat ćemo zamišljeni koncept arhiva podataka koji nam kasnije može pomoći za konkretnije planiranje prilikom uspostavljanja njegove organizacije.

Definicija organizacije

Hrvatski arhiv podataka za društvene znanosti je javni servis u nacionalnoj znanstveno-istraživačkoj infrastrukturi koji se bavi prikupljanjem, pohranjivanjem, distribucijom i dugoročnim očuvanjem izvornih podataka stvorenih u istraživanjima u području društvenih znanosti i sličnih podataka u drugim područjima. Svojim korisnicima - istraživačima, nastavnicima, studentima i ostalim zainteresiranima, arhiv podataka osigurava vjerodostojan izvor istraživačkih podataka za primjenu u vlastitim istraživanjima ili u nastavnom procesu.

Misija

Misija arhiva podataka je pružanje usluga temeljne podrške istraživanjima u društvenim znanostima tijekom cijelog životnog ciklusa istraživačkog projekta, od nastanka ideje i pripreme projekata za financiranje, za vrijeme prikupljanja podataka (iz originalnih i sekundarnih izvora) te njihove obrade i analize, a podrška se nastavlja i nakon istraživačkog projekta kako bi se osigurale pretpostavke za ponovno korištenje podataka u bližoj i daljoj budućnosti. Arhiv nastoji smanjiti prepreke ponovnom korištenju podataka te podupire otvoreni pristup podacima gdje god je to moguće.

Svoju misiju arhiv ispunjava kroz kontinuiranu kontrolu digitalnih objekata koji sadrže podatke i njihov opis s ciljem da se podaci dugoročno očuvaju, da budu dostupni u otvorenim formatima i dobro dokumentirani te da ih je uz primjenu informacijskih tehnologija što lakše koristiti. U svom radu arhiv podataka slijedi međunarodne profesionalne standarde te surađuje sa sličnim organizacijama širom Europe i ostatka svijeta.

² <http://roadmap2018.esfri.eu/>

Obuhvat i vrste podataka

Primarni fokus arhiva podataka je na kvantitativnim i kvalitativnim podacima u znanstvenim poljima sociologija, politologija, psihologija, pedagogija, informacijske i komunikacijske znanosti te ekonomija, ali prihvaćaju se i podaci iz drugih područja, ako su nastali metodologijom društvenih znanosti i ako pomažu u razumijevanju društva. Kvalitativni podaci, zbog njihove kompleksnije prirode, prihvaćaju se u skladu s raspoloživim resursima za njihovu obradu i očuvanje. Arhiv može prihvatiti i administrativne podatke koje proizvode vladini uredi i agencije tijekom svojih svakodnevnih aktivnosti. Prikupljaju se podaci nastali u suvremenim istraživanjima, ali i podaci nastali u starijim istraživanjima ukoliko su značajni za povijesne preglede i komparativna istraživanja. Radi zaštite privatnosti i osobnih podataka, svi podaci pohranjuju se u anonimiziranom obliku kako temeljem arhiviranih podataka ne bi bilo moguće utvrditi identitet pojedinih sudionika istraživanja.

Usluge i aktivnosti

Osnovne usluge arhiva podataka odnose se na objavljivanje podataka nastalih u istraživanjima u području društvenih znanosti te arhiviranje i dugoročno očuvanje tih podataka. S tim u skladu, arhiv treba obavljati sljedeće poslove:

- identifikacija i selekcija podataka kroz praćenje istraživačkih projekata i komunikaciju sa stvarateljima podataka;
- preuzimanje podataka od stvaratelja zajedno s punim skupom opisnih podataka o provedbi, metodologiji, osobama i institucijama uključenim u provođenje istraživanja (osiguravanje sigurne okoline za prihvrat podataka, zaštita osobnih podataka, dogovaranje prava pristupa);
- kontrola podataka i popratne dokumentacije (konzistentnost podataka, formati datoteka, oznake za varijable i njihove vrijednosti, standardni metapodatkovni opisi)
- pohranjivanje podataka, standardnog metapodatkovnog opisa i dokumentacije u repozitorij;
- objavljivanje podataka (dodjela identifikacijskih oznaka, administracija korisnika i omogućavanje pristupa podacima u skladu s uvjetima dogovorenim sa stvarateljima podataka);
- implementacija informacijskih pomagala za pronalaženje podataka (katalozi) te drugih alata i protokola kojima se postiže veći doseg informacija o dostupnim podacima u lokalnoj i međunarodnoj zajednici;
- uključivanje metapodataka iz lokalnih zbirki u svjetske mreže repozitorija s podacima, naročito u CESSDA skupni katalog;
- kontinuirana kontrola i održavanje zbirki digitalnih objekata radi njihovog dugoročnog očuvanja.

Kako bi svoje zadaće obavljao što bolje, što se odražava ne samo u količini već i u kvaliteti arhiviranih setova podataka, arhiv podataka treba pružati i usluge podrške korisnicima. Stoga treba organizirati i izrađivati različite edukacijske aktivnosti i materijale namijenjene:

- stvarateljima podataka pri upravljanju podacima i planiranju (npr. podrška u izradi plana upravljanja podacima);
- korisnicima podataka oko pronalaženja i korištenja podataka.

Dodatne aktivnosti, u skladu s misijom arhiva podataka i raspoloživim resursima, mogu biti:

- organizacija skupova i promotivnih aktivnosti radi edukacije i podizanja svijesti o važnosti arhiviranja podataka i prednostima ponovnog korištenja;
- zagovaranje i promocija otvorenog pristupa podacima;
- implementacija alata za online pregled varijabli i njihovih vrijednosti (pitanja i odgovora) te za analizu podataka uz kreiranje tablica, grafova i drugih analitičkih rezultata;

- osiguravanje sigurnog pristupa osjetljivim podacima (sigurna soba).

Radni procesi koji se odvijaju u arhivu podataka trebaju osigurati autentičnost podataka i njihovu iskoristivost u dalekoj budućnosti, što se postiže poštivanjem postojećih međunarodnih standarda i najboljih praksi i certifikaciju kvalitete usluga čime se daje garancija pouzdanosti arhiva podataka i gradi povjerenje korisnika arhiva podataka.

Arhiv podataka aktivno sudjeluje u mreži ustanova koje pružaju slične ili povezane usluge (knjižnice, računski centri, nacionalni arhivi, Državni statistički zavod i druge) nudeći toj zajednici ekspertizu o podacima u području društvenih znanosti i razmjenjujući s njima stručna znanja o upravljanju podacima i njihovom očuvanju, uz zajedničko nastojanje da se usluge i njihova kvaliteta usklade i koordiniraju na nacionalnoj razini. Kroz sudjelovanje u međunarodnim projektima i skupovima te kroz suradnju s CESSDA ERIC konzorcijem i ostalim sličnim ustanovama i organizacijama, arhiv podataka kontinuirano razvija svoje radne procese i usluge u skladu s međunarodnim standardima, a njeguje i regionalnu suradnju, posebno sa zemljama s kojima Hrvatska dijeli dio povijesti zabilježene u zajedničkim istraživanjima.

Korisnici

Arhiv podataka na raspolaganju je pojedincima i organizacijama koje žele pohraniti svoje podatke ili koristiti podatke nastale u drugim projektima. Primarni korisnici usluga Arhiva podataka su istraživači na sveučilištima, nastavnici i studenti, te istraživači u javnim znanstvenim institutima i vladinim organizacijama. Druga kategorija korisnika su donosioci znanstvenih politika i financijeri istraživanja koji među kriterijima za evaluaciju istraživačkih projekata imaju i one koji se odnose na upravljanje i arhiviranje istraživačkih podataka. Njima arhiv podataka može osigurati temeljnu infrastrukturu potrebnu istraživačima za ispunjavanje takvih kriterija. Usluge mogu koristiti i istraživači iz nevladinih organizacija i privatnog sektora, srednjoškolski nastavnici i učenici, novinari, tvrtke, strukovne organizacije i drugi. Usluge su besplatne za svako nekomercijalno korištenje, a za komercijalne svrhe usluga se može naplatiti.

Upravljanje

Budući da se radi o nacionalnom servisu, važno je da u osnivanju i u kasnijem upravljanju arhivom podataka sudjeluju organizacije koje proizvode podatke ili podatke trebaju u svakodnevnom radu. Ključnu ulogu u razvoju i potpori nacionalnim infrastrukturama ima Ministarstvo znanosti i obrazovanja. Arhiv podataka moguće je osnovati i kao konzorcij više ustanova.

Arhivom podataka rukovodi ustanova sjedište i voditelj/ica arhiva koji osiguravaju poslovni okvir za provođenje usluga i aktivnosti i odgovorni su za poslovanje arhiva. Ukoliko se radi o konzorciju, u provođenju aktivnosti arhiva podataka trebaju sudjelovati sve članice konzorcija.

Radi osiguravanja usklađenosti s nacionalnim i institucionalnim strategijama te oblikovanja razvojnih planova, misije i ciljeva arhiva podataka u nacionalnom kontekstu, potrebno je osnovati nacionalno savjetodavno tijelo koje će pratiti rad arhiva te odobravati godišnje planove i izvještaje. Članovi ovoga tijela su institucije koje se bave donošenjem nacionalnih znanstvenih politika te financiranjem znanosti i znanstvene infrastrukture (Ministarstvo znanosti i obrazovanja, Hrvatska zaklada za znanost), zajedno sa predstavnicima znanstvenih instituta, fakulteta i drugih znanstvenih organizacija u području društvenih znanosti, Državnog zavoda za statistiku i drugih organizacija koje imaju potrebu pohraniti vlastite i koristiti postojeće podatke te koje mogu donositi vlastite politike o upravljanju istraživačkim podacima. Osobe koje su izabrane u ovo tijelo ujedno su i komunikacijska spona sa svojim matičnom

ustanovom tako što podižu svijest o uslugama, aktivnostima i inicijativama arhiva podataka te daju povratne informacije o vlastitim potrebama prema kojima se treba ravnati razvoj usluga arhiva podataka.

Potrebno je još jedno savjetodavno tijelo koje će raspravljati o stručnim pitanjima koja se odnose na metodologiju znanstvenih istraživanja i upravljanje istraživačkim podacima u društvenim znanostima te o procedurama dugoročnog očuvanja podataka kroz upotrebu standardnih formata, metapodatkovnog opisa i drugo. Ovo tijelo treba okupiti istraživače iz raznih polja unutar društvenih znanosti, stručnjake za podatke i knjižničare specijalizirane za podatke, IT stručnjake te predstavnike CESSDA ERIC konzorcija.

Financiranje

Financiranje arhiva podataka može se gledati kroz dva stadija razvoja institucije: početni stadij uspostave arhiva i operativni stadij. Početni stadij uključuje aktivnosti poput zapošljavanja osoblja, njihovog osposobljavanja za rad te uspostavljanja pravila i procedura prema kojima djeluje arhiv podataka. Troškovi, u smislu vremena koje osoblje ulaže, u početnoj fazi mogu biti vrlo značajni, dok će operativna faza biti učinkovitija kada se radni procesi normaliziraju. Pregled financijskih zahtjeva može se temeljiti na KRDS modelu aktivnosti kojeg su izradili Beagrie, Ch. i sur.³ Taj model opisuje cijeli niz aktivnosti potrebnih za osiguravanje dugoročnog očuvanja istraživačkih podataka i objašnjava raspodjelu troškova kroz te aktivnosti. Završni izračuni troškova ovisit će o sljedećim ključnim varijablama: osoblje (na temelju definiranih uloga i odgovornosti); softver, hardver, poslužitelji i ostala oprema; te ekonomska prilagodba kao što su stope inflacije.

U idealnom slučaju, budući da se radi o javnom nacionalnom servisu, rad arhiva podataka financirat će se iz državnog proračuna kao redovna aktivnost, sa stalnim radnim mjestima i godišnjim proračunom. Razvoj i funkcioniranje usluga arhiva podataka može se djelomično financirati i kroz višegodišnje projekte, putem nacionalnih poziva kojima se ostvaruju ciljevi nacionalne strategije za znanost, istraživanje i obrazovanje i koji se odnose na ulaganje u razvoj istraživačkih infrastruktura u europskom kontekstu. Arhiv podataka treba sudjelovati i u međunarodnim i regionalnim projektima (EC pozivi i drugi) zajedno s drugim europskim arhivima, kako bi se omogućio daljnji razvoj, obrazovanje i suradnja s međunarodnim zajednicama. Doprinos ustanove koja je sjedište arhiva podataka sastoji se od pružanja administrativnih resursa, prostora i opreme za rad.

Postojeća iskustva u Hrvatskoj

Prva inicijativa i završeni projekti uspostavljanja arhiva podataka

Prva inicijativa za osnivanje arhiva podataka u području društvenih znanosti krenula je još 2006. godine kada su se okupili istraživači i informacijski stručnjaci iz različitih ustanova⁴. Djelovanje inicijative prepoznato je kroz *CESSDA Preparatory Phase Project* (2008) kroz koji se CESSDA pripremala za dobivanje ERIC statusa i kroz koji su uspostavljeni prvi kontakti s međunarodnom zajednicom. Slijedom toga dobiven je prvi poziv za sudjelovanje u FP7 projektu **SERSCIDA - Support for**

³ <https://beagrie.com/krds/>

⁴ Henjak, A., Potočnik, D., Matković, T. i Vodopijevac, A. (2007). Zasnivanje arhiva podataka društvenih i humanističkih znanosti u RH. *Korisnik u fokusu: 9. CARNetova korisnička konferencija*. https://cuc.carnet.hr/cuc2007/program/radovi/prezentacije/b2_7_pr_vodopijevac.pdf

*Establishment of National/Regional Data Archives*⁵ (2012-2014) koji se bavio uspostavljanjem arhiva podataka za društvene znanosti u Hrvatskoj i regiji i u kojem je Filozofski fakultet Sveučilišta u Zagrebu (FFZG) sudjelovao kao partner. Rad se nastavlja kroz dva nova projekta na FFZG: *SEEDS - South-Eastern European Data Services*⁶ (2015-2017) financiran kroz SCOPES program kojeg je podupirala Swiss National Science Foundation i Swiss Agency for Cooperation and Development te *CESSDA Strengthening and Widening* (2015-2017) financiran u okviru Obzor 2020 programa. U svim ovim projektima partneri su bili iskusni arhivi podataka za društvene znanosti okupljeni oko CESSDA konzorcija.

Kontakti s relevantnim organizacijama i pojedincima

Tijekom razdoblja provođenja ovih projekata, u više su navrata kontaktirani istraživači i ustanove u području društvenih znanosti. Informacije o aktivnostima oko uspostavljanja arhiva podataka poslane su i institucijama koje se bave oblikovanjem znanstvenih politika, financiranjem znanosti i vrednovanjem (MZO, HRZZ, AZVO, Područno znanstveno vijeće za društvene znanosti, Hrvatska akademija znanosti i umjetnosti). S nekolicinom ustanova obavljani su intervjui čiji rezultati su poslužili za analizu postojećih potencijala za osnivanje arhiva podataka u Hrvatskoj⁷ Od Ministarstva znanosti i obrazovanja dobivena je podrška projektu SERSCIDA⁸, a Filozofski fakultet u Sveučilišta u Zagrebu prepoznat kao partner Ministarstva u izgradnji podatkovne infrastrukture. Organizirano je nekoliko posjeta europskim arhivima podataka⁹ (slovenski ADP, britanski UKDA, švedski NSD) kojima su, osim osoba neposredno uključenih u projekte, prisustvovali i istraživači, informacijski i informatički stručnjaci te predstavnice Hrvatske zaklade za znanost. Održana su i dva okrugla stola o arhiviranju podataka u društvenim znanostima. Zadnji sastanak u Ministarstvu znanosti i obrazovanja s pomoćnikom za znanost i predstavnicom RH u ESFRI održan je u travnju 2017. godine povodom aktivnosti na projektu CESSDA SaW koje su se odnosile na izradu nacionalnog plana za osnivanje arhiva podataka u za društvene znanosti.

Edukacija

Kroz sva tri navedena projekta, ali i kroz redovne aktivnosti i programe CESSDA konzorcija, održane su brojne radionice i prezentacije o arhiviranju istraživačkih podataka. Sudionici su upoznali radne procese koji se odvijaju u arhivu podataka i kojima se osigurava autentičnost podataka i njihova iskoristivost u dalekoj budućnosti, naučili kako se izrađuje metapodatkovni zapis prema DDI standardu, koja dokumentacija je potrebna uz podatke, na koji način i u kojem obliku se podaci pohranjuju u repozitorij, s kojim pravnim i etičkim pitanjima se arhiv podataka treba baviti, kakva tehnička infrastruktura je potrebna, kako arhiv podataka može certificirati kvalitetu svojih usluga i drugo.

Prototip arhiva podataka, pravila, procedure i planovi razvoja

Tijekom SERSCIDA projekta izrađen je prvi prototip web stranice arhiva i repozitorija podataka, a kroz projekt SEEDS izrađena je nova verzija tog prototipa¹⁰ korištenjem Dataverse platforme te je

⁵ <http://serscida.eu/hr>

⁶ <http://seedsproject.ch>

⁷ SERSCIDA (2012). *Analysis of existing potentials for the establishment of a social sciences digital data base archive in Croatia*.

http://www.serscida.eu/images/deliverables/SERSCIDA_D_2_2_Mapping_Report_Croatia.pdf

⁸ <https://www.serscida.eu/data/index.php/s/mr2esd2LuFp7NAO>

⁹ SERSCIDA (2015). *Report on Training Activities*.

http://www.serscida.eu/images/deliverables/SERSCIDA_D_4_3_Report_on_Training_Activities.pdf

¹⁰ Prototip web stranice arhiva podataka: <https://hr.seedsproject.ffzg.unizg.hr/>

Prototip repozitorija podataka: <https://dataverse.ffzg.unizg.hr>

prikupljeno i pripremljeno za objavu deset setova podataka. Kroz CESSDA SaW projekt elaboriran je plan razvoja nacionalnog arhiva podataka¹¹, dokument koji je, zajedno s dokumentima nastalim za vrijeme prve inicijative i prethodnih projekata, poslužio kao osnova ovoga dokumenta. Razvijeni su dokumenti koji opisuju pravila i procedure rada arhiva podataka koji služe kao temelj i okvir za razvoj i upravljanje uslugama i aktivnosti arhiva podataka¹².

Anketno istraživanje o korištenju podataka među istraživačima u društvenim znanostima

Online anketnim istraživanjem, koje je provedeno 2012. godine u okviru SERSCIDA projekta¹³ i koje je obuhvatilo znanstvenike iz područja društvenih znanosti u Hrvatskoj, identificirano je kako se većina bavi empirijskim istraživanjima, u prosjeku izrađujući jedno istraživanje godišnje. Oni gotovo bez iznimke i čuvaju podatke tih istraživanja. Međutim, ti su podaci uglavnom izuzetno slabo dokumentirani i čuvaju se na osobnim računalima i medijima istraživača. Praksa pristupa vlastitim podacima uglavnom je ograničena na izvorni istraživački tim, ponekad isključivo voditelja projekta. To je uža razina pristupa od one koju većina sudionika vidi poželjnom; čak polovica sudionika istraživanja smatra da bi podaci istraživanja u kojima su sudjelovali trebali biti dostupni čitavoj akademskoj zajednici ili svima koji iskazuju interes. U pravilu istraživači iskazuju spremnost za arhiviranje svojih podataka u sigurnom i reguliranom repozitoriju. Također, dominantno je gledište sudionika kako je pristup istraživačkim podacima drugih istraživača, kako domaćih tako i inozemnih vrlo bitno u njihovim disciplinama, te kako vodi unapređenju vlastitog znanstvenog i nastavnog rada. Međutim, u praksi sekundarnim analizama se pristupa rijetko (posebno kvalitativnih podataka), a istraživači često niti ne znaju koristi li netko drugi njihove podatke. Oni koji trenutno primjenjuju sekundarne analize, najčešće do podataka dolaze kroz kolegijalne kanale, projektne mrežne stranice i DZS-a. Kao najveća trenutna prepreka sekundarnim analizama prepoznaje se slaba dostupnost postojećih istraživanja, te u nešto manjoj mjeri nesklonost dijeljenju podataka u domaćoj istraživačkoj kulturi. Potonje je vidljivo i iz procjene spremnosti dijeljenja podataka, koju sudionici vide kao vrlo visoku kod sebe i svojih suradnika, ali često kao upitnu kod drugih kolega. U svjetlu ovih nalaza, ali i temeljem eksplicitnog iskaza sudionika, potreba za formiranjem nacionalnog arhiva čini se se akutnom i široko je prepoznata među istraživačima.

Potrebni resursi i organizacija

Ovdje ćemo razmotriti neke elemente organizacijske strukture, naročito resurse koje je potrebno osigurati u procesu osnivanja arhiva podataka kao nacionalnog javnog servisa za društvene znanosti, te

SEEDS (2017). *Report on tool evaluation and selection*. https://seedsproject.ch/wp-content/uploads/2015/06/D9_FINAL.pdf

SEEDS (2017). *Report on integration of technical systems*. <https://seedsproject.ch/wp-content/uploads/2015/06/CroatiaD6-Reportonintegrationoftechnicalsystemfinal.pdf>

¹¹ CESSDA SaW (2017). *National development plan: Croatia*.

https://www.cessda.eu/content/download/3881/35232/file/NDP_Croatia.pdf

¹² SERSCIDA (2014). *Documents and Materials for Social Science Digital Data Archives*.

http://www.serscida.eu/images/deliverables/D5.1_FINAL.pdf

¹³ SERSCIDA (2012). *Analysis of existing potentials for the establishment of a social sciences digital data base archive in Croatia*.

http://www.serscida.eu/images/deliverables/SERSCIDA_D_2_2_Mapping_Report_Croatia.pdf

ih staviti u kontekst onoga što već postoji u nacionalnom okruženju, kako bismo nove servise gradili na postojećim iskustvima i povezali ih s drugim elementima znanstvene infrastrukture.

Sjedište arhiva podataka

Sjedište arhiva podataka treba biti javna znanstvena organizacija u kojoj se njeguje empirijski pristup istraživanjima u području društvenih znanosti i koja strateški podupire sudjelovanje u razvoju istraživačkih infrastrukture u međunarodnom kontekstu. Ta znanstvena organizacija treba imati iskustvo u projektima i aktivnostima Konzorcija europskih arhiva podataka za društvene znanosti (CESSDA ERIC) i mora biti sposobna zadovoljiti uvjete koji su potrebni za članstvo Republike Hrvatske u CESSDA ERIC, kako bi mogla biti imenovana predstavnicom RH u tom konzorciju tj. dobiti status nacionalnog pružatelja usluga.

Filozofski fakultet Sveučilišta u Zagrebu (FFZG) prepoznat je kao moguće sjedište budućeg arhiva podataka za društvene znanosti još 2006. godine, za vrijeme prve inicijative prema uspostavljanju arhiva podataka. Upravo slijedom te identifikacije, FFZG je postao partner u projektima opisanim u ovome dokumentu, u poglavlju o prethodnim iskustvima. FFZG je najveća istraživačka i obrazovna ustanova u Hrvatskoj u području humanistike i društvenih znanosti na kojoj se izvodi više od 100 preddiplomskih i diplomskih programa te 19 doktorskih programa. S preko 500 istraživača i oko 7000 studenata, FFZG je dobro integriran u nacionalnu istraživačku mrežu te ima najveću visokoškolsku knjižnicu u Hrvatskoj za područje humanistike i društvenih znanosti čije su usluge i građa na raspolaganju cijeloj akademskoj zajednici i šire.

FFZG sudjeluje u predlaganju dva nova projekta u suradnji s CESSDA ERIC-om. Jedan projekt priprema se za poziv *INFRADEV-03-2019: Individual support to ESFRI and other world-class research infrastructures (RIA)* u okviru programa Obzor 2020, a koordinirat će ga CESSDA. Svrha ovoga projekta je jačanje i širenje konzorcija kroz aktivnosti usmjerene na izgradnju infrastrukture i održivost. Drugi projekt priprema se za pozivi u okviru *Interreg - Danube Transnational Programme*, koordinirat će ga AUSSDA - Austrijski arhiv podataka za društvene znanosti, a CESSDA će sudjelovati u njegovom provođenju

FFZG je trenutno uključen u provođenje projekta "Monitoring i evaluacija projekta Rijeka 2020 Europska prijestolnica kulture" unutar kojeg će se izraditi plan upravljanja podacima, osigurati sigurno mjesto za razmjenu datoteka među istraživačima tijekom izvođenja projekta te arhivirati prikupljene podatke prema međunarodnim standardima s ciljem njihove ponovne upotrebe.

Osim toga, FFZG sudjeluje i u mentorskom programu za 2019. godinu kojeg provodi CESSDA kao svoju internu aktivnosti. Cilj ovog programa je ponuditi direktnu podršku novoosnovanim arhivima podataka i ustanovama koje uspostavljaju arhive podataka u definiranju i postizanju njihovih aktualnih kratkoročnih ciljeva. ADP - Arhiv družboslovnih podatkov iz Slovenije bit će mentor FFZG-u.

Osoblje

Za normalno funkcioniranje arhiva podataka potrebne su minimalno četiri različite uloge koje treba preuzeti osoblje zaposleno u arhivu.

Jedna je uloga voditelja arhiva podataka, osobe koja je odgovorna za svakodnevno funkcioniranje arhiva te za komunikaciju s dionicima, regionalnim i međunarodnim partnerima. To mora biti osoba koja dobro razumije radne procese, standarde, pravna pitanja i dobre prakse arhiviranja i objavljivanja istraživačkih podataka te njihovog dugoročnog očuvanja. Ona mora poznavati okolinu, uvjete i resurse potrebne za rad arhiva, uključujući i razumijevanje tehničke infrastrukture za pohranu podataka i

softverskih alata za upravljanje podacima. Treba imati dobre komunikacijske i upravljačke vještine te biti uključena u nacionalne i međunarodne mreže i organizacije, arhive podataka i udruženja koja se bave istraživačkim podacima. Voditelj/ica treba imati završen najmanje diplomski studij u području društvenih znanosti, doktorat bi bio prednost.

Arhivu su potrebna i dva stručnjaka za podatke. Uloge stručnjaka za podatke mogu se podijeliti na niže i više položaje, na temelju prethodnog radnog iskustva. Viši stručnjak za podatke trebao bi pokazati znanje u području metodologije društvenih znanosti, statističke obrade podataka i upravljanja podacima. Poželjan je završen diplomski studij u području društvenih znanosti i dodatni studij informacijskih znanosti, doktorat ili upisan doktorski studij može biti prednost. Mlađi stručnjak za podatke trebao bi imati završen najmanje preddiplomski studij iz područja društvenih znanosti. Ta pozicija ne zahtijeva prethodno veliko iskustvo u upravljanju istraživačkim podacima, ali je potrebno dovoljno znanja o metodologiji društvenih znanosti i statističkoj obradi podataka.

Četvrtu ulogu trebao bi preuzeti IT stručnjak. Za poziciju IT stručnjaka u arhivu podataka potrebne su vještine administracije poslužitelja, programiranja i upravljanja bazama podataka. Položaj također zahtijeva opće razumijevanje protokola za razmjenu podataka u umreženom okruženju i pitanjima tehničke sigurnosti koja su s njom povezana. Ta osoba mora imati iskustvo u arhitekturi informacijskih sustava te razvoju i održavanju softverskih aplikacija. Potrebna je diploma iz područja informatike/računarstva ili ekvivalentno iskustvo, a poželjno je i prethodno iskustvo s informacijskim sustavima u znanosti.

U početku, minimalno dvije osobe trebaju biti zaposlene u arhivu kako bi se pokrenuli procesi oko uspostave normalnog rada arhiva. Obje osobe trebaju biti stručnjaci za podatke i trebaju biti u stanju preuzeti različite uloge. Jedna osoba treba preuzeti ulogu voditelja arhiva. Barem jedna osoba treba dobro razumjeti IT potrebe i biti u stanju raditi na implementaciji tehničkih rješenja u suradnji s drugim partnerima (primarno Srce i CESSDA).

Budući da se ne postoji formalno obrazovanje stručnjaka za podatke, obrazovanje zaposlenika treba se oslanjati na zajednicu iskusnih stručnjaka kakva je dostupna preko CESSDA konzorcija u čijoj organizaciji se svake godine odvija cijeli niz aktivnosti usmjerenih na potporu zaposlenika.

Partnerstvo i suradnja

Izuzetno je važno da arhiv podataka bude prepoznat kao vrijedan nacionalni servis u svim institucijama u kojima se podaci proizvode i koriste, ali i u institucijama koje se bave razvojem i implementacijom znanstvenih politika i financiraju znanstveni rad. Osim toga, suradnja je potrebna i s ostalim institucijama koje se bave znanstvenom infrastrukturom. Važnu ulogu u radu arhiva podataka trebaju imati ove organizacije:

Ministarstvo znanosti i obrazovanja (MZO) kao glavna ustanova za donošenje znanstvenih politika i financiranje znanstvene infrastrukture. Usluge arhiva podataka moraju biti prepoznate kao jedan od načina ostvarivanja nacionalnih strateških ciljeva za znanost i uvrštene u akcijske planove. Osnivanje i buduće funkcioniranje arhiva podataka kao nacionalnog servisa nije moguće bez da se MZO obveže da će takvim aktivnostima pružiti potporu. Članstvo RH u CESSDA ERIC ostvaruje se preko MZO, pri čemu MZO treba imenovati pružatelja usluge koji je u stanju zadovoljiti potrebne stručne kriterije.

Hrvatska zaklada za znanost (HRZZ) kao glavna ustanova koja osigurava potporu znanstvenih, visoko obrazovnih i tehnologijskih programa i projekata. Zaklada može u svojim internim dokumentima, strategijama i pravilnicima razviti kriterije za evaluaciju projektnih prijedloga i rezultata projekata koji zahtijevaju izradu plana upravljanja istraživačkim podacima i obvezuju istraživače da

pohranjuju istraživačke podatke radi njihovog ponovnog korištenja. Zaklada je mogući izvor financiranja nekih aktivnosti arhiva podataka, a također bi mogla unutar istraživačkih projekata odobriti financiranje aktivnosti oko upravljanja istraživačkim podacima i na taj način promovirati važnost dobrih praksi upravljanja istraživačkim podacima.

Sve **znanstvene organizacije u području društvenih znanosti**, posebno sveučilišta, fakulteti i javni istraživački instituti, trebaju biti upoznati s uslugama i aktivnostima arhiva, a neki i neposredno uključeni u njegov rad kroz partnerstvo i ostale oblike suradnje. Istraživači zaposleni pri takvim ustanovama glavni su stvaratelji podataka, a također i primarni korisnici podataka. Te ustanove mogu donositi svoje vlastite politike o upravljanju istraživačkim podacima i objavljivanju tih podataka pri čemu im arhiv podataka može pružiti informacije i primjere. Također, u takvim ustanovama često su zaposleni knjižničari i drugi informacijski stručnjaci koji nisu samo posrednici između arhiva podataka i istraživača, već ponekad i sami mogu preuzeti dio poslova oko upravljanja istraživačkim podacima.

Vladini uredi, agencije i zavodi koji prikupljaju podatke i provode istraživanja u svrhu potpore procesima oblikovanja socijalnih i gospodarskih politika, planiranja i provođenja mjera za postizanje zadanih ciljeva te planiranje i provođenje projekata kojima se unapređuju usluge koje država obavlja za građane.

Državni zavod za statistiku (DZS) proizvodi nacionalne statističke podatke koje istraživači u području društvenih znanosti često trebaju. Formalna suradnja između DZS-a i arhiva podataka u budućnosti može se ostvariti u domeni mikropodataka jer su takvi podaci za istraživače bogatiji izvor u odnosu na agregirane podatke koje standardno objavljuje DZS. Arhiv podataka može pružiti ekspertizu oko standardnog metapodatkovnog opisa i dokumentacije tih podataka te pomoći u njihovoj diseminaciji do ciljane publike.

Sveučilišni računski centar Sveučilišta u Zagrebu (Srce) kao pružatelj ključne tehničke infrastrukture za dugoročno očuvanje digitalnih objekata koji se prikupljaju i uređuju u arhivu podataka. Osoblje Srca bilo je uključeno u aktivnosti projekata SERSCIDA, SEEDS i CESSDA-SaW. Srce implementira i održava tehničke komponente Dabar sustava repozitorija, kakav je potreban i za istraživačke podatke. Sustav AAI@EduHr kojeg razvija i održava Srce, također je važna komponenta u arhivu podataka.

Institut Ruđer Bošković (IRB) kao institucija s velikim iskustvom u razvoju informacijskih sustava za znanstvene informacije. IRB-ov **Centar za znanstvene informacije** aktivni je partner u OpenAIRE projektima te i kroz brojne druge aktivnosti promovira i podupire otvoreni pristup znanstvenim informacijama. Djelatnici Centra također su sudjelovali u projektima koji su se bavili uspostavljanjem arhiva podataka za društvene znanosti (SERSCIDA, SEEDS i CESSDA-SaW). Druge relevantne aktivnosti su suradnji na razvoju i održavanju sustava Dabar i partnerstvo EOSC-hub projektu.

Nacionalna i sveučilišna knjižnica (NSK) kao ustanova koja se bavi dugoročnim očuvanjem hrvatske baštine koja se digitalizira ili nastaje u digitalnom obliku. Iskustvo i ekspertiza u primjeni standardnih metapodatkovnih opisa, u razvoju tezaurusa i drugih sustava za predmetno označavanje, te u razvoju informacijskih sustava za pronalaženje, čini temelj suradnje NSK i arhiva istraživačkih podataka. NSK je također jedna od potpisnica Sporazuma o suradnji na razvoju i održavanju sustava Dabar.

Hrvatski državni arhiv kao ustanova koja se također bavi dugoročnim očuvanjem hrvatske baštine susreće se sa sličnim stručnim i tehničkim izazovima kao i arhiv istraživačkih podataka.

Regionalna i međunarodna suradnja je za arhiv podataka neophodna kako bi povećao svoju relevantnost i razvio potrebne kompetencije. Sudjelovanje u regionalnim projektima i nastavak suradnje s partnerima u regiji zanimljivo je zbog jezičnih i kulturalnih sličnosti, kao i zbog dijela zajedničke

povijesti. Međunarodna suradnja postiže se primarno kroz članstvo u CESSDA ERIC koja predstavlja integriranu infrastrukturu za društvene znanosti, ali i kroz članstva u međunarodnim organizacijama poput IASIST i RDA.

Suradnja je neophodna i s ostalim **paneuropskim istraživačkim infrastrukturama** u području humanistike i društvenih znanosti koje su aktivne ili su bile aktivne u Hrvatskoj (DARIAH-HR, CLARIN-HR, ESS, SHARE, ISPP, CSES). Među tim infrastrukturama postoje dodirne točke koje se i na europskoj razini pokušavaju dovesti u sinergiju.

Tehnička infrastruktura

Arhivu podataka potrebna je pouzdana tehnička infrastruktura u kojoj se podaci mogu dugoročno pohraniti u sigurnoj okolini te koja može omogućiti nesmetan pristup podacima i potporu ostalim uslugama arhiva. Ona uključuje i različite softverske alate i pomagala koji se koriste u svakodnevnom radu arhiva (prihvata podataka, komunikacija s korisnicima, konverzija podataka, kontrola konzistentnosti podataka) ili koje istraživači koriste za pristupanje podacima te za pregledavanje i analiziranje podataka.

Prototip repozitorija podataka¹⁴ izrađen na Dataverse platformi, kao jedan od elemenata potrebne tehničke infrastrukture, omogućuje pohranu, objavu i pronalaženje istraživačkih podataka te reguliranje pristupa podacima, trenutno je smješten na infrastrukturi FFZG-a koja se koristi za potrebe Knjižnice filozofskog fakulteta u Zagrebu. Knjižnični informacijski sustav u jednom dijelu je sličan informacijskom sustavu za arhiviranje podataka te se postojeća konfiguracija poslužiteljskih računala i sustava za pohranu podataka može koristiti prilikom osnivanja arhiva podataka i za obavljanje njegovih svakodnevnih aktivnosti, zajedno s ostalim elementima infrastrukture kojima raspolaže cijeli fakultet.

Dataverse sustav razvijen je u University of Harvard, a sve je popularniji i u Europi, naročito za novoosnovane arhive podataka (npr. austrijska AUSSDA¹⁵). Unutar svojih aktivnosti koje se odnose na širenje i jačanje organizacije te kroz projekte koji se prijavljuju na natječaje Europske komisije, CESSDA razvija sustav podrške arhivima podataka koji teže zadovoljiti uvjete za članstvo. Trenutno je u tijeku predlaganje projekta, u kojem FFZG sudjeluje kao partner, kroz koji CESSDA planira implementirati DataverseEU¹⁶ i razvijati softversku podršku novim arhivima bez razvijene vlastite infrastrukture, ali i već postojećim i iskusnim arhivima podataka za razvoj novih usluga poput samoarhiviranja. Osim kao repozitorij podataka, Dataverse se može koristiti i kao platforma za razmjenu podataka među istraživačima tijekom projekta, prije objave podataka.

Sveučilišni računski centar Sveučilišta u Zagrebu (Srce) neizostavan je u području suradnje oko tehničke podrške radu arhiva podataka. Srce osigurava e-infrastrukturu potrebnu za obavljanje djelatnosti arhiva podataka naročito u onim elementima koji su nužni kao sigurna okolina za pohranu i dugoročno očuvanje podataka, kakva se može ostvariti kroz **Dabar** sustav repozitorija u čijem radu FFZG već sudjeluje. Suradnja arhiva podataka s Dabrom važna je zbog dugoročne održivosti samih podataka, pa time i usluga koje arhiv pruža korisnicima. Drugi važan sustav kojeg razvija i održava Srce je **AAI@EduHr** koji omogućuje autorizaciju i autentikaciju korisnika te se može koristiti pri administriranju prava pristupa podacima. Trenutno je u razvoju i sustav **CroRIS** - Sustav informacija o hrvatskoj znanstvenoj djelatnosti, putem kojeg će se omogućiti unos, pohranjivanje, upravljanje,

¹⁴ <http://dataverse.ffzg.unizg.hr>

¹⁵ <https://aussda.at/>

¹⁶ https://www.cessda.eu/belgrade2018/presentations/13.CESSDA_Widening_2018_2_Wittenberg.pdf

pristup i diseminacija svih relevantnih informacija o hrvatskoj znanstvenoj djelatnosti i njezinim temeljnim sastavnicama.

Rad na povezivanju Dataverse sustava i Islandora sustava (kojeg Srce koristi za repozitorije u Dabru) već je započeo u projektu CESSDA SaW kada su se kroz pilot aktivnost ispitivale mogućnosti korištenja Dataverse platforme za prikupljanje podataka od istraživača u području društvenih znanosti i Islandora softvera (na kojem je izgrađen Dabar) u kojem bi se mogle dugoročno očuvati arhivske verzije digitalnih objekata¹⁷.

¹⁷ <http://cessdasaw.eu/content/uploads/2017/11/SAW-D4.5.pdf>