



PROCENA KVALITETA DIGITALNIH ARHIVA PRIMARNIH PODATAKA U NAUCI

Aleksandra Bradić-
Martinović^{1, *},
Jelena Banović¹,
Larisa Malić²

¹Institut ekonomskih nauka,
Beograd, Srbija

²Institut za književnost i umetnost,
Beograd, Srbija

Rezime:

Ponovna upotreba istraživačkih podataka zavisi od stepena do kog se podaci mogu iskoristiti bez ograničenja za akademsku zajednicu. Diverzitet ciljnih grupa takođe je jedan od pokazatelja kvaliteta upotrebe podataka. Otvorena i slobodna upotreba istraživačkih podataka odvija se u unutar relevantnih zakonskih okvira i političkih odluka, određenih od strane nadležnih vlasti. Da bi arhivi imali jasne smernice razvoja, a korisnici garanciju do kog nivoa je rad arhiva usklađen sa najboljom praksom, uvedeni su nivoi sertifikacije. U ovom radu detaljno je opisan CoreTrustSeal okvir sertifikacije, na primeru dva operativna arhiva u društvenim naukama, Data Centra Srbija u društvenim naukama, koji funkcioniše kao nacionalni arhiv podataka i Arhiva družboslovnih podataka iz Slovenije. Cilj rada je upoznavanje šire naučne zajednice sa kompleksnom temom digitalnog arhiviranja i diseminacije podataka u nauci.

Ključne reči:

digitalni arhivi, sertifikacija, istraživački podaci, Data centar Srbija za društvene nauke.

1. UVOD

Kako bi se obezbedilo kvalitetno i pravilno čuvanje i diseminacija primarnih mikro podataka u nauci, dugoročno pohranjivanje u digitalnim arhivima postalo je nezaobilazan korak svakog istraživačkog ciklusa. Arhiviranje podataka nastalih prilikom naučnih istraživanja može biti od izrazitog nacionalnog značaja, jer njihovo skladištenje omogućuje ponovnu upotrebu u budućnosti, čime se produžuje životni vek podataka i omogućuju longitudinalne i komparativne analize. Osim toga, plan upravljanja istraživačkim podacima sve češće postaje zahtev finansijera projekata, a samim tim i obaveza istraživača. Konačno, svetski renomirani časopisi pokrenuli su inicijativu koja podrazumeva da istraživač uz rad priloži i podatke koje je koristio za analizu. Kao rezultat globalnih kretanja pojavili su se otvoreni (često komercijalni), kao i nacionalni, tematski i institucionalni repozitorijumi - digitalni arhivi podataka. Pojedini, poput *UK Data Service*, imaju tradiciju dužu od pedeset godina.

Revizija i sertifikacija digitalnih arhiva na međunarodnom nivou predstavljaju najbolji pokazatelj (samo)procene i garancije pouzdanosti ovih infrastruktura.

Odgovorno lice:

Aleksandra Bradić-Martinović

e-pošta:

abmartinovic@ien.bg.ac.rs



Većina država koje finansiraju nacionalne digitalne arhive imaju i zakonske okvire koji utiču na njihov rad, kao i na kvalitet čuvanja i diseminacije podataka. Ovi zakonski okviri obuhvataju raspon od statutarnih (koji su u funkciji zaštite privatnosti ispitanika obuhvaćenim istraživanjem), do kodeksa ponašanja koji uključuju etiku i moral, kao dva najvažnija načela naučnog istraživanja. Situacija se posebno komplikuje u slučaju podataka koji su rezultat istraživanja u društvenim naukama, kao posledica činjenice da su subjekti istraživanja najčešće fizička lica, čija privatnost ne sme biti ugrožena. Iako je oblast digitalnog očuvanja još uvek u povoju, veliki napredak se postiže dobrom infrastrukturom, razvojem politika, preporuka i standarda za sertifikaciju [1]. U slučaju digitalnih arhiva podataka u društvenim naukama sertifikacija predstavlja važnu referencu garancije kvaliteta. Cilj ovog rada je da ukaže istraživačkoj zajednici na činjenicu da je za čuvanje vrednih podataka, koji imaju veliki potencijal za ponovnu upotrebu i za koje se proceni da je opravdano dugoročno očuvanje, potrebno obezbediti kvalitetne i stabilne uslove rada, solidnu infrastrukturu i ekspertska znanja. U protivnom, sačuvani podaci se mogu oštetiti, izgubiti ili se ispitanici obuhvaćeni istraživanjem izložiti riziku ugrožavanja privatnosti.

2. EVROPSKI OKVIR ZA REVIZIJU I SERTIFIKACIJU DIGITALNIH ARHIVA

U cilju efikasne koordinacije revizije i sertifikacije 2010. godine potpisan je memorandum „Evropski okvir za reviziju i sertifikaciju digitalnih arhiva“ (*European Framework for Audit and Certification of Digital Repositories*). Okosnicu ovog memoranduma objedinjuju tri standarda ključna za sticanje pouzdanosti:

- ♦ *Data Seal of Approval* i *The World Data System of the International Science Council* (DSA 2009 + WDS 2013), sada objedinjeni kao *CoreTrustSeal* standard; dobijanjem ovog sertifikata digitalni arhiv stiče **osnovni** vid sertifikacije [2];
- ♦ *Audit and certification of trustworthy digital repositories – ISO 16363, 2012*: **prošireni** vid sertifikacije i
- ♦ *Criteria for Trustworthy Digital Archives DIN 31644, 2012*, **formalni** vid sertifikacije. [3]

Navedeni standardi konstruišu kompletan okvir sertifikacije digitalnih arhiva na svetskom nivou. Nivoi sertifikacije obezbeđuju različite nivoe poverenja i pouzdanosti, odnosno ispunjenja traženih kriterijuma i smernica.

CoreTrustSeal standard pre svega obuhvata osnovni vid sertifikacije za pouzdane digitalne arhive. Njegov katalog smernica i zahteva zasnovan je na principu transparentnosti i otvorenosti, kao i određivanja balansa između jednostavnosti i uloženog truda [2].

Proširenu sertifikaciju mogu steći samo arhivi koju su prethodno ispunili sve uslove *CoreTrustSeal*-a i izvršili krajnju javnu samoprocenu zasnovanu na ili ISO 16363 ili DIN 31644 standardima. ISO standard sadrži više od osamdeset kriterijuma koji se odnose na:

- ♦ organizacionu infrastrukturu,
- ♦ digitalno upravljanje objektima i
- ♦ upravljanje rizikom bezbednosti.

Formalna sertifikacija podrazumeva sertifikaciju koja, pored stečene osnovne, sadrži i spoljnu reviziju koja garantuje ispunjenje svih uslova, kriterijuma i smernica koje propisuju ISO 16363 ili DIN 31644. [3]

U ovom radu fokus je postavljen na *CoreTrustSeal* sertifikaciju, jer je ona usklađena sa zahtevima CESS-DA ERIC za arhive u društvenim naukama u Evropi. Kriterijumi za kvalitet i pouzdanost objedinjeni u *CoreTrustSeal* standardu su od velikog interesa i značaja za sve stejkholdere - istraživače i institucije koje proizvode istraživačke podatke, zatim za organizacije koje arhiviraju te podatke, kao i za njihove krajnje korisnike. Kriterijumi su sastavljeni sa ciljem da se podaci sačuvaju, da se osigura visok kvalitet njihovog skladištenja i da se obezbedi njihova pravilna distribucija. *CoreTrustSeal* obezbeđuje istraživačima sigurnost da će njihovo istraživanje biti sačuvano i distribuirano na odgovarajući način i isključivo u naučnoistraživačke svrhe. Ovaj standard baziran je na šesnaest smernica koje služe za proveru i potvrdu kvaliteta i pouzdanosti jednog digitalnog arhiva. Smernice i uputstva se u najvećoj meri odnose na kreiranje, skladištenje, upotrebu i distribuiranje digitalnih podataka iz oblasti naučnog istraživanja. Homogenost svih smernica obeležena je principima koji zajedno definišu da li se podaci mogu smatrati održivim na digitalnom nivou i ti principi podrazumevaju sledeće: da podaci mogu biti pronađeni na Internetu, da se podacima može pristupiti uz određene uslove i odgovarajućim nalogom sa ličnim podacima i svesti o intelektualnoj svojini, da su podaci skladišteni u odgovarajućim i upotrebljivim formatima i da su podaci relevantni.

Svaki kriterijum praćen je smernicama, kao i numeričkom ocenom koja pokazuje do kog je nivoa uspostavljena usaglašenost sa standardom. Sistem sertifikacije podrazumeva da arhiv vrši samo-evaluaciju i u skladu sa stepenom usklađenosti sa zahtevom određuje jedan od sledećih nivoa:



- 0 – Nije primenljiv
- 1 – Digitalni arhiv još uvek nije razmotrio ovaj zahtev
- 2 – Digitalni arhiv ima teorijski koncept primene ovog zahteva
- 3 – Digitalni arhiv je u fazi implementacije
- 4 – Zahtev je u potpunosti primenjen. [4]

U narednom delu rada predstavice ćemo zahteve koje propisuje *CoreTrustSeal* standard.

3. ZAHTEVI CORETRUSTSEAL STANDARDA

Prilikom ocene kvaliteta digitalnog arhiva podataka polazi se ona sledećih zahteva:

R1. Misija digitalnog arhiva - Digitalni arhiv ima javnu i eksplicitnu misiju u okviru oblasti u kojoj vrše digitalno arhiviranje.

R2. Licence - Digitalni arhiv primenjuje sve adekvatne licence koje se odnose na pristup i upotrebu podataka.

R3. Kontinuitet pristupa, u dugoročnom okviru - Digitalni arhiv ima dugoročni plan za skladištenje svog digitalnog fonda.

R4. Poverljivost - Digitalni arhiv osigurava da su svi podaci proizvedeni, sakupljeni, organizovani i sačuvani u skladu sa svim etičkim normama svojstvenih konkretnoj naučnoj disciplini.

R5. Organizaciona struktura - Arhiv ima odgovarajuće i održivo finansiranje, kao i dovoljan broj stručnog i kvalifikovanog osoblja koje obavlja aktivnosti kroz jasno definisan sistem upravljanja sa ciljem ispunjenja postavljene misije.

R6. Stručno usavršavanje - Arhiv usvaja mehanizme za neprekidno stručno usavršavanje, kao i za pružanje povratnih informacija (*feedback*), na svim nivoima - unutar samog arhiva, kao i konsultacije sa drugim spoljnim ili nadležnim organizacijama.

R7. Integritet i autentičnost podataka - Digitalni arhiv osigurava integritet i autentičnost digitalnih objekata i meta-podataka. Ovo se najpre odnosi na stepen pouzdanosti i približnosti originalnim istraživačkim podacima.

R8. Procena kvaliteta podataka - Repozitorijum prihvata samo one podatke i meta-podatke koji su kreirani u skladu sa definisanim kriterijumima, kako bi osigurao relevantnost i čitljivost za krajnje korisnike.

R9. Dokumentovane procedure za skladištenje - Arhiv poseduje procedure kojima se vodi pri upravljanju skladištenjem podataka. Kada podaci stignu u arhiv, vrši se njihova detaljna provera koja podrazumeva poštovanje

standarda o (meta)podacima, obavlja se unifikacija svih fajlova i njihovih formata uz proveru dodatnih tehničkih specifikacija. Skladištenje podrazumeva različite sigurnosne mehanizme koji se sprovode radi utvrđivanja kvaliteta i održivosti podataka, proveravaju se potencijalne greške, vrši se evaluacija materijala. Takođe, arhiv je u obavezi da za sve materijale omogući odgovarajući tip pristupa, kao i odgovarajući nivo zaštite.

R10. Plan očuvanja - Arhiv preuzima odgovornost za dugoročno očuvanje podataka na unapred predviđen i dokumentovan način.

R11. Kvalitet podataka - Arhiv ima odgovarajuća ekspertna znanja koja se odnose na tehnički kvalitet podataka i meta-podataka, a koje omogućuje da se obezbede dovoljne informacije krajnjih korisnicima kako bi bili u stanju da izvrše procenu kvaliteta podataka. U ovom procesu arhiv treba da sarađuje sa provajderima podataka i da ima uvid u sve informacije o podacima koji će se eventualno čuvati u arhivu. Te informacije će poslužiti i arhivu i potencijalnim korisnicima podataka prilikom procene kvaliteta raspoloživih podataka. Arhiv ima pravo, za one skupove podataka za koje ne postoji minimalan broj opisnih informacija, odlučiti da ih ne pohranjuje. Zbog toga, deponenti moraju da obezbede dokumenta u kojima se vidi kako je istraživanje sprovedeno, na kom uzorku, šta je bio osnovni cilj, da priloži metodologiju rada, spisak članova istraživačkog tima i ostale relevantne informacije.

R12. Tok aktivnosti - Arhiviranje se odvija u skladu sa definisanim tokom, od preuzimanja podataka od deponenta, do diseminacije ka istraživačima koji ih ponovo koriste.

R13. Dostupnost i identifikacija podataka - Arhiv omogućuje korisnicima da pronađu potrebne podatke i da ih citiraju na jednoznačan način. Mogućnost lakog pristupa podacima je ključ za dobro deljenje podataka, a većina arhiva obezbeđuje pretražive kataloge u kojima se pored pristupa podacima, obezbeđuju i meta-podaci, kako bi potencijalni korisnik utvrdio da li podaci odgovaraju njegovim potrebama.

R14. Ponovna upotreba podataka - Arhiv je dužan da omogući da se podaci mogu ponovo koristiti, obezbeđujući da odgovarajući meta-podaci budu dostupni kako bi podržali razumevanje i upotrebu podataka.

R15. Tehnička infrastruktura - Arhiv koristi dobro implementirane operativne sisteme i hardverske i softverske tehnologije koje odgovaraju uslugama koje arhiv isporučuje zajednici. Infrastruktura arhiva mora maksimizirati dostupnost usluga krajnjem korisniku.



R16. Bezbednost - Tehnička infrastruktura arhiva mora da obezbedi zaštitu svih digitalnih objekata koji su u njemu pohranjeni, kao i svih elementa koji su uključeni u proces – podataka, proizvoda, usluga i korisnika. Arhiv mora da vrši redovne analize potencijalnih pretnji i rizika, ali i da kreira dosledan sistem zaštite. Takođe, neophodno je da vrši procene potencijalnih šteta koje bi eventualno nastale usled napada malicioznih softvera, zbog ljudskog faktora ili nekog drugog tehničkog problema koji može oštetiti bilo koji entitet u arhivu. Neophodno je meriti verovatnoću i uticaj ovih šteta na arhiv, kao i odlučiti koje mere treba preduzeti kako bi se rizici i pretnje izbegle, a rad arhiva se nastavio neometano. Za ovaj zahtev, neophodno je da arhiv ispuni sledeće zahteve, da opiše procedure i *backup* opcije ukoliko dođe do kolapsa i da opiše sistem zaštite, plan ukoliko dođe do kolapsa, plan za kontinuirani nastavak rada, podelu poslova svih zaposlenih ukoliko do kolapsa dođe. [5]

Do sada je 143 repozitorijuma dobilo *CoreTrustSeal* sertifikat [6]. Da bismo čitaocima približili zahteve i proces sertifikacije odabrali smo dva digitalna arhiva koji skladište primarne podatke u društvenim naukama, Data Centar Srbija za društvene nauke (DCS), čije je sedište u Beogradu i Arhiv družboslovnih podataka (ADP), čije je sedište u Ljubljani.

4. ANALIZA USKLAĐENOSTI SA ZAHTEVIMA CORETRUSTSEAL STANDARDA NA PRIMERU ADP I DCS

Imajući u vidu kompleksnost teme i oblasti smatramo da je najbolji način za razumevanje i približavanje materije široj naučnoj zajednici i svim zainteresovanim stranama uporedna analiza ispunjenosti zahteva odabranog standarda, na primeru ADP-a i DCS-a. Potrebno je istaći da je ADP dobio *CoreTrustSeal* sertifikat 23. januara 2018. godine, a da DCS ima za cilj da dostigne kvalitet koji će mu omogućiti sertifikaciju. [7]

Data centar Srbija za društvene nauke (<http://data-centarserbia.com/>) osnovan je 2014. godine u okviru Instituta ekonomskih nauka u Beogradu. Nastao je kao rezultat FP7 projekta SERSCIDA (*Support for Establishment of National/Regional Social Sciences Data Archives*). Njegov dalji razvoj podržala je Švajcarska nacionalna fondacija za nauku, u okviru SCOPES projekta SEEDS (*South-Eastern European Data Services*) i EC u okviru projekta H2020 - CESSDA SaW (*Strengthening and Widening*). Na osnovu ostvarenih rezultata, Ministarstvo prosvete, nauke i tehnološkog razvoja Republike Srbije prepoznalo je značaj novoosnovane

istraživačke infrastrukture i podržalo rad DCS-a kroz učlanjenje Republike Srbije u CESSDA ERIC, u novembru 2018. godine, pri čemu je DCS postao nacionalni pružalac usluga (*Service Provider*). CESSDA je konzorcijum koji okuplja države članice koje u okviru Evropskog istraživačkog prostora imaju operativne digitalne arhive u društvenim naukama [8]. Trenutno DCS čuva i distribuira i kvantitativne i kvalitativne podatke preko portala podataka SEEDSbase, čiji je host švajcarski nacionalni digitalni arhiv FORS i preko ePrints platforme, koja je hostovana na internim resursima.

Osnivanjem i radom DCS-a istraživači u društvenim naukama dobili su mogućnost da čuvaju i preuzimaju mikro podatke prikupljene u primarnim istraživanjima, u skladu sa odredbama „Platforme za otvorenu nauku“.

Arhiv družboslovnih podataka (ADP) je osnovan 1997. godine, kao organizaciona jedinica Instituta za društvene nauke na Fakultetu društvenih nauke, Univerziteta u Ljubljani. ADP je nacionalna istraživačka infrastruktura za društvene nauke, čija je glavna misija upravljanje podacima i uslugama podataka kako bi se podržalo istraživanje, obrazovanje i opšte blagostanje. ADP je, takođe pružalac usluga u CESSDA ERIC. Ministarstvo prosvete, nauke i sporta Republike Slovenije finansira rad ove institucije od osnivanja. Od 2004. godine finansiranje rada ADP-a omogućeno je kroz program finansiranja infrastruktura koje funkcionišu u okviru Univerziteta u Ljubljani – MRIC UL. Arhiv u svom radu odgovara Savetu ADP-a i Slovenačkoj istraživačkoj agenciji. [9]

Zahtev 1: Misija

Misija ADP-ja je „promocija čuvanja podataka iz istraživanja sprovedenih u istraživačkim centrima u Sloveniji“. ADP je uspostavljen kao deo nacionalne infrastrukture i kroz svoju misiju potencira čuvanje, deponovanje i promovisanje podataka, kao i mogućnost njihovog kasnijeg korišćenja u naučne, obrazovne i druge [10].

Misija DCS je da „obezbedi podršku istraživačima u procesu upravljanja podacima kroz razvoj pouzdane infrastrukture i kroz uvećanje i deljenje znanja u ovoj oblasti“. [11]

Nivo usklađenosti sa zahtevom 1: ADP - nivo 4; DCS – nivo 4

Zahtev 2: Licence

ADP funkcioniše u skladu sa zakonskom regulativom koja se odnosi na pristup i upotrebu digitalnih



objekata. Takođe, vodi računa da zajednica korisnika (*designated community*) bude upoznata sa ovim okvirom i svim zahtevima koji se postavljaju pred njih. Osnovni dokumenti koji se odnose na ovaj zahtev su „Politika digitalnog očuvanja“ (www.adp.fdv.uni-lj.si/publikacije_adp/publikacija/248/) i „Zakonski osnov, etičke norme i odgovornosti“ [12].

U praktičnom radu, ADP sa provajderima podataka (deponentima) potpisuje se „Ugovor o skladištenju podataka“, koji se bazira na poštovanju nacionalnih pravnih okvira o zaštiti podataka i citiranju podataka. Pored toga, Ugovorom se definiše način upotrebe i nivo zaštite Podataka. Sa korisnicima podataka, takođe potpisuje „Ugovor o upotrebi podataka“, po kome se korisnik se obavezuje da će koristiti podatke u skladu sa zakonskim i etičkim normama, kao i da anonimizirane podatke koristi kao takve. [13].

Prilikom zaštite autorskih prava, ADP koristi *Creative Commons* licence (CC0 - upotreba bez ograničenja; CCBY – obaveza navođenja izvora (autora) i CCBYNC – obaveza navođenja izvora (autora), uz obavezu upotrebe u nekomercijalne svrhe) [13].

DCS takođe funkcioniše u skladu sa zakonskim i etičkim standardima koji se odnose na privatne podatke i specifičnost njihovog deljenja u skladu sa Opštom uredbom o zaštiti podataka o ličnosti (GDPR), kao i etičkim standardima svojstvenim društvenim naukama. Deponovanje i ponovna upotreba podataka regulisana je odgovarajućim ugovorima, kao i u ADP-u. Konačno, isti sistem zaštite autorskih prava primenjen je i u radu DCS-a.

Nivo usklađenosti sa zahtevom 2:

ADP - nivo 4; DCS – nivo 4

Zahtev 3: Kontinuitet pristupa, u dugoročnom okviru

Kontinuitet rada ADP-a obezbeđen je kroz „Pravila organizacije i implementacije istraživačkih arhiva Fakulteta za društvene nauke Univerziteta u Ljubljani“ (2017), kao i kroz „Strategiju očuvanja“ i „Digitalno očuvanje podataka“ – dva interna dokumenta. ADP radi u skladu sa odredbama Otvorenog pristupa.

DCS, u ovoj fazi razvoja, nema razvijenu pravnu podršku koja bi mu omogućila kontinuitet rada i razvoja, ali postoji svest o potrebi zadovoljenja ovog zahteva i čine se koraci za njegovo ispunjenje, u saradnji sa resornim ministarstvom.

Nivo usklađenosti sa zahtevom 3:

ADP - nivo 4; DCS – nivo 2

Zahtev 4: Poverljivost

ADP je u potpunosti uskladio rad sa zahtevima relevantnih dokumenata koji uslovljavaju zaštitu anonimnosti ispitanika. Za svaku studiju koja se deponuje proverava se da li poštuje sve etičke i pravne norme. To znači da deponent mora da osigura da su svi podaci koji se tiču učesnika studije zaštićeni od zloupotreba, kao i da poseduje potpisani pristanak za deljenje podataka od strane ispitanika, u slučaju kvalitativnih podataka. Deponent je u obavezi da sam izvrši anonimizaciju podataka, ali ADP stoji na raspolaganju za svaku vrstu pomoći. U ekstremnim slučajevima, arhiv vrši anonimizaciju podataka, prateći unapred utvrđene procedure, ali isključivo u slučajevima kada se radi o jako osetljivim podacima koje deponent nije u stanju sam da anonimizira.

Na isti način i DCS zadovoljava ovaj zahtev.

Nivo usklađenosti sa zahtevom 4:

ADP - nivo 4; DCS – nivo 4

Zahtev 5: Organizaciona infrastruktura

ADP je organizaciona jedinica Instituta za društvene nauke u Ljubljani, koji funkcioniše u okviru Fakulteta za društvene nauke, Univerziteta u Ljubljani. Rad arhiva se bazira na pravnim aktima potpisanim između ovih institucija. Takođe, arhiv ima podršku Ministarstva, koje obezbeđuje sredstava za njegovo funkcionisanje [14].

DCS je organizacioni deo Instituta ekonomskih nauka iz Beograda. U Centru rade stručni saradnici za oblasti upravljanja podacima, a obezbeđena je i adekvatna IT podrška. Na žalost, još uvek ne postoji adekvatna i održiva šema finansiranja centra, te se njegov rad i razvoj i dalje oslanja na podršku partnera iz inostranstva, pre svega CESSDA ERIC i njenih članica, što nije dovoljno za ostvarenje postavljene [15].

Nivo usklađenosti sa zahtevom 5:

ADP - nivo 4; DCS – nivo 3

Zahtev 6: Stručno usavršavanje

ADP je dobro integrisan u naučnu zajednicu Slovenije, a i aktivni je član međunarodne istraživačke infrastrukture, što omogućuje praćenje i uključivanje u



globalne inicijative, koje implementira u skladu sa svojim mogućnostima. Vršu stručnu konsultaciju sa partnerima iz zemlje i inostranstva, promovišu svoje podatke putem različitih okruglih stolova, seminara i radionica namenjenih učesnicima naučno-istraživačke zajednice.

DCS, takođe ima jedinstvenu priliku stručnog usavršavanja učešćem u radu CESSDA ERIC, kroz aktivni rad u radnim grupama i radionicama. Jačanje pozicije DCS-a u lokalnom naučnom okruženju je cilj koji će biti realizovan tokom 2019. godine, kroz organizaciju radionica i okruglih stolova.

Nivo usklađenosti sa zahtevom 6:
ADP - nivo 4; DCS – nivo 3

Zahtev 7: Integritet i autentičnost podataka

ADP u svojoj dugoročnoj misiji očuvanja za cilj ima da blisko isprati principe očuvanja integriteta i autentičnosti podataka. Materijali koji stignu u arhiv, podvrgnuti su detaljnim ispitivanjima, gde se fokus stavlja na kompletnost dokumentacije, značaj studije, ispitivanje anonimnosti materijala, usklađenost dostavljenih formata sa preporučenim formatima. Kreirane su brojne procedure koja na formalan način propisuju tok podataka i zahteve koji u tom procesu moraju biti ispunjeni [7].

DCS ima ključni dokument „Politike i procedure“ u kome su definisane odredbe o integritetu i autentičnosti podataka, ali još uvek nema razrađene sve korake pojedinačno.

Nivo usklađenosti sa zahtevom 7:
ADP - nivo 4; DCS – nivo 3

Zahtev 8: Procena podataka

Politika rada ADP-a propisuju da se za skladištenje i dugoročno očuvanje prihvataju oni podaci za koje se proceni da imaju veliki potencijal ponovne upotrebe, da su zanimljivi za slovenačku naučnu zajednicu, da ističu određeni istraživački i naučni problem, kao i da je značajan za slovenačko društvo uopšte. Zadovoljenje ovog zahteva proizilazi iz potrebe da se ograničeni materijalni i ljudski resursi iskoriste na optimalan način [7].

DCS, takođe ima propisane uslove za podatke koje prikuplja, mada u praksi, koja je uslovljena niskim nivoom svesti istraživača u ovoj oblasti, uslovi se ne postavljaju rigorozno. Zbog toga, prihvataju se svi raspoloživi podaci.

Nivo usklađenosti sa zahtevom 8:
ADP - nivo 4; DCS – nivo 3

Zahtev 9: Dokumentovane procedure za skladištenje

U ADP-ju se arhiviranje vrši prateći unapred određeni tok aktivnosti. Sam postupak skladištenja je detaljno opisan u usvojenim internim dokumentima arhiva.

DCS ima teorijski koncept primene ovog zahteva.

Nivo usklađenosti sa zahtevom 9:
ADP - nivo 4; DCS – nivo 2

Zahtev 10: Plan očuvanja

ADP je razvio kompletan sistem digitalnog očuvanja podataka, uključujući i jasno definisan tok aktivnosti sa pravilima za njihovo sprovođenje. Tok aktivnosti usklađen je sa OAIS (*Open Archival Information System*) modelom. Definisani su formati podataka koji se prihvataju od provajdera podataka, kao i formati u kojima se vrši diseminacija [7].

DCS je započeo da razvija dugoročni plan očuvanja i upravljanja podacima, koji će obezbediti održivost arhiva. Takođe, ovim planom DCS obezbediće kvalitetno očuvanje podataka od prijema u arhiv do kasnije mogućnosti za ponovnu upotrebu. Aktivnosti u arhivu će, takođe, biti usklađene se OAIS referentnim modelom. Dokumentom „Politike i procedure“ definisani su formati podataka koji se primaju za arhiviranje i koji se koriste u diseminaciji.

Nivo usklađenosti sa zahtevom 10:
ADP - nivo 4; DCS – nivo 2

Zahtev 11: Kvalitet podataka

Kao što smo već naveli, ADP prati zahteve OAIS modela kroz obezbeđenje dovoljne količine informacija o svakoj studiji (istraživanju) na nivou meta-podataka, tako da su krajnji korisnici u stanju da procene u kojoj meri podaci zadovoljavaju njihove istraživačke potrebe. Šema i zahtevi za meta-podacima raspoloživi su na njihovoj veb prezentaciji.

Ovaj zahtev je na isti način postavljen i u radu DCS-a, s tim što postoji prostor za unapređenje.



**Nivo usklađenosti sa zahtevom 11:
ADP - nivo 4; DCS – nivo 3**

Zahtev 12: Tok aktivnosti

ADP ima jasno definisan tok aktivnosti koje su jasno opisane usvojenim procedurama, usklađenim sa OAIS referentnim modelom. Postoji obaveza praćenja svih transformacija koje se obave na originalno dostavljenim mikro podacima (registruje se svaka izmena), a proces se sprovodi iz saglasnost deponenta, koji zadržava autorsko pravo nad podacima. ADP u najvećem broju slučajeva skladišti kvantitativne podatke koje je lakše anonimizirati i koji su najčešće u SPSS formatu, ali postoje i jasno definisane procedure obrade kvalitativnih podataka. Poštuje se DDI šema meta-podataka, a u upotrebi je ELLST tezaurus za tematsku klasifikaciju [7].

DCS ima za cilj da unapredi svoje politike i procedure koje će jasno definisati tok rada, od prijema do mogućnosti za ponovnu upotrebu podataka.

**Nivo usklađenosti sa zahtevom 12:
ADP - nivo 4; DCS – nivo 2**

Zahtev 13: Dostupnost i identifikacija podataka

Dostupnost podataka pohranjenih u ADP-u moguća je kroz njihov katalog [16] i *Nesstar* platformu. Obe opcije poseduju mogućnost pretraživanja (*search*), a postoje i uputstva za istraživače. Fajlovi sa podacima su u najčešće korišćenim formatima (SPSS (*.sav), SPSS Portable (*.por), Statistica, Stata v.8, Stata v.7. DIF, Dbase, SAS). Korisnicima je stavljena na raspolaganje i opcija direktnog preuzimanja podataka. Svi podaci imaju DOI brojeve.

DCS je još uvek u fazi implementacije platforme, a najveći broj podataka se nalazi na SEEDSbase platformi [17], u kojoj su sačuvani podaci prikupljeni u regionu Zapadnog Balkana, prikupljeni u istraživanjima u društvenim naukama, a u okviru podrške RRPP. U periodu april-december 2019. godine biće izvršna migracija podataka u lokalni repozitorijum [18]. U budućnosti se predviđa uvođenje platforme *Dataverse-EU*. Oba arhiva su aktivno uključena razvoj CESSDA kataloga podataka i usavršavanje pomenute platforme.

**Nivo usklađenosti sa zahtevom 13:
ADP - nivo 3; DCS – nivo 2**

Zahtev 14: Ponovna upotreba podataka

U ADP-u dokument „Opis studije“ i „Ugovor o deponovanju“ definišu koji su obavezni meta-podaci (deskriptivni, strukturni i administrativni). Ukoliko nisu obezbeđeni svi obavezni meta-podaci, studija neće biti objavljena u katalogu i podaci neće biti javno dostupni. U upotrebi je DDI 2.0 šema meta-podataka. Dugoročno arhiviranje vrši se u formatima „otpornim na promene“, kao što su TXT, HTML, XML i sl. [7]

Na isti način organizovani su zahtevi za meta-podacima u DCS-a, uz otvoren prostor za napredak.

**Nivo usklađenosti sa zahtevom 14:
ADP - nivo 3; DCS – nivo 2**

Zahtev 15: Tehnička infrastruktura

ADP implementira politiku tehnološkog razvoja i bezbednosne zaštite u procesu digitalne prezervacije, a koja je kreirana na osnovu liste međunarodnih smernica i procedura. U tehničkom smislu, ADP koristi resurse Slovenačke akademske i istraživačke mreže (ARNES). [7]

U tehničkom smislu, DCS se oslanja na interne resurse Instituta ekonomskih nauka, što bi u jednom trenutku razvoja moglo da postane ograničenje. Takođe, DCS trenutno testira najbolju mogućnost za implementaciju arhiva i kataloga, kao posledica činjenice da je CESSDA prestala da održava *Nesstar* platformu, a još uvek nije završila unapređenje i prilagođavanje *Dataverse-EU* platforme.

**Nivo usklađenosti sa zahtevom 15:
ADP - nivo 3; DCS – nivo 2**

Zahtev 16: Bezbednost

U okviru politike o digitalnoj prezervaciji, ADP je propisao i poštuje procedure za zaštitu osetljivih podataka, prvo kroz fizičku (virtuelnu) zaštitu pristupa, a dalje kroz upravljanje nivoima pristupa podacima. Pristup arhivskom materijalu dozvoljen je samo ovlašćenim osobama ADP-a. U slučaju da je procenjeno da su podaci veoma osetljivi, postoji mogućnost uvođenja dodatnih restrikcija, uprkos opredeljenju za Otvoreni pristup. Redovno se vrše procedure bekapa. [7]

DCS na isti način reguliše svoje bezbednosne aspekte, uz prostor za unapređenje u budućnosti, u skladu sa količinom i vrstom podataka koji će se naći u sistemu.



**Nivo usklađenosti sa zahtevom 16:
ADP - nivo 3; DCS – nivo 2**

5. ZAKLJUČAK

Dugoročno čuvanje mikro-podataka prikupljenih u naučnim istraživanjima ima veliku vrednost. To posebno dolazi do izražaja ako su u pitanju istraživanja u društvenim naukama, jer se presek društva u jednom trenutku nikada više ne može replikovati. Jedan od načina da se omogući kreiranje naučne zaostavštine je deponovanje podataka u digitalne repozitorijume/arhive, a zatim i obezbeđenje lakog i bezbednog pristupa tim podacima. Velike kompanije su prepoznale značaj čuvanja podataka i reagovala tako što su tržištu ponudile otvorene, javne, najčešće komercijalne platforme, kao što su *Figshare*, *Harvard Dataverse*, *Open Science Framework*, *Zenodo*, *Mendeley Data* i druge. Države su, takođe procenile obim oportunitetnih troškova koji se generišu gubitkom skupova podataka koji mogu da opišu važne društvene fenomene i zbog toga su spremne da ulože u kreiranje i kontinuiran rad nacionalnih digitalnih arhiva podataka. Konačno, istraživačka zajednica se nalazi pod pritiskom javnih politika, a u Srbiji je sredinom jula 2018. godine usvojena „Platforma za otvorenu nauku“ prema kojoj je propisana obaveza otvorenog pristupa naučnim publikacijama, ali i preporuka otvorenog pristupa naučnim podacima, kao prvi korak ka uvođenju obaveze. Međutim, istraživači se u najvećem broju slučajeva nalaze pred dilemom gde da deponuju svoje podatke. Jedan od navedenih, javnih arhiva je opcija, ali ukoliko procene da su njihovi podaci veoma vredni za naučnu zajednicu i za buduća istraživanja, ukoliko žele da budu sigurni da će njihovi podaci biti arhivirani za dugoročno očuvanje, da će biti promovisani na pravi način, što će im omogućiti da izvrše širi uticaj, a obezbedi im i korektno citiranje, kao i da će biti zaštićeni i oni i njihovi ispitanici od zloupotrebe, važno je da shvate značaj kvaliteta rada i organizacije jednog arhiva. Standardi za procenu kvaliteta su jedan od najsigurnijih načina da donesu pravu odluku o izboru virtuelnog prostora na kome će čuvati svoje podatke.

Republika Srbija spada u grupu zemalja koje imaju solidnu infrastrukturu, koja se razvija. Angažovani u Data Centru Srbija za društvene nauke poseduju ekspertsku znanja iz ove oblasti i stoje na raspolaganju istraživačima u društvenim naukama.

ZAHVALNOST

Ovaj rad je deo istraživačkih projekata pod šiframa 179015 (Izazovi i perspektive strukturnih promena u Srbiji: Strateški pravci ekonomskog razvoja i usklađivanje sa zahtevima EU) I 47009 (Evropske integracije i društveno-ekonomske promene privrede Srbije na putu ka EU), finansiranih od strane Ministarstva prosvete, nauke i tehnološkog razvoja Republike Srbije.

LITERATURA

- [1] Xie, K. Matusaik, „Digital preservation“. Discover Digital libraries: theory and practice. Elsevier, Amsterdam, pp. 255-279, 2016.
- [2] [CTS] CoreTrustSeal <https://www.coretrustseal.org/about/> posećeno 27.3.2019.
- [3] Trusted Digital Repository, <http://www.trusteddigitalrepository.eu/Trusted%20Digital%20Repository.html>. Posećeno 26.3.2019.
- [4] [Core Trustworthy Data Repositories Requirements] https://www.coretrustseal.org/wp-content/uploads/2017/01/Core_Trustworthy_Data_Repositories_Requirements_01_00.pdf. Posećeno 22.3.2019.
- [5] Core Trustworthy Data Repositories Extended Guidance <https://www.coretrustseal.org/wp-content/uploads/2017/01/20180629-CTS-Extended-Guidance-v1.1.pdf>. Posećeno 22.3.2019.
- [6] [CTS], www.coretrustseal.org/why-certification/search-repositories/. Posećeno 27.3.2019.
- [7] Implementation of the CoreTrustSeal, <https://www.coretrustseal.org/wp-content/uploads/2018/01/ADP-Social-Science-Data-Archives.pdf>. Posećeno 12.3.2019.
- [8] [CESSDA] Consortium of European Social Science Data Archives, <https://www.CESSDA.eu/>. Posećeno 15.3.2019.
- [9] [ADP] Arhiv društvenih podataka, www.adp.fdv.uni-lj.si/. Posećeno 13.3.2019.
- [10] [ADP], Our mission, <https://www.adp.fdv.uni-lj.si/eng/spoznaj/adp/poslanstvo/>. Posećeno 20.3.2019.
- [11] [DCS] Data centar Srbija <http://datacentarserbia.com/svrha/>. Posećeno 20.3.2019.
- [12] [ADP] www.adp.fdv.uni-lj.si/eng/spoznaj/politika/infrastruktura/. Posećeno 20.3.2019.
- [13] Bradić-Martinović, Aleksandra, Zdravković, Aleksandar. Zaštita privatnosti – anonimizacija podataka. Naučni skup „Usson“, Fakultet zdravstvenih, pravih i poslovnih studija u Valjevu, Univerzitet Singidunum, 2013.



- [14] [ADP] <https://www.adp.fdv.uni-lj.si/eng/kontakt/>. Posećeno 20.3.2019.
- [15] [DCS] <http://datacentarserbia.com/osoblje/>. Posećeno 20.3.2019.
- [16] [ADP] Katalog ADP, www.adp.fdv.uni-lj.si/opisi/. Posećeno 21.3.2019.
- [17] SEEDSbase, <https://seedsdata.unil.ch/>. Posećeno 15.3.2019.
- [18] Data centar Srbija u društvenim naukama, dcs.iien.bg.ac.rs. Posećeno 21.3.2019