



Pohrana i objava podataka u arhivu CROSSDA

2. prosinca 2022.

Radionica za zaposlenike Instituta za razvoj i međunarodne odnose

Marijana Glavica

CROSSDA

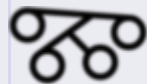
Hrvatski arhiv podataka za društvene znanosti

Nacionalni javni servis za istraživačke podatke u području društvenih znanosti.

Pruža podršku istraživačima oko upravljanja podacima tijekom cijelog istraživačkog procesa, kako bi se očuvao integritet i autentičnost podataka zajedno sa svim informacijama potrebnim za interpretaciju podataka.

Omogućuje dugoročno očuvanje podataka i smanjuje prepreke ponovnom korištenju podataka.

<https://crossda.hr>



Filozofski fakultet
Sveučilišta u Zagrebu



REPUBLIKA HRVATSKA
Ministarstvo znanosti i
obrazovanja

CESSDA ERIC

Konzorcij europskih arhiva podataka za društvene znanosti

Distribuirana istraživačka infrastruktura. Članice konzorcija su države. Svaka država imenuje jednog pružatelja usluga. Za Hrvatsku to je CROSSDA.

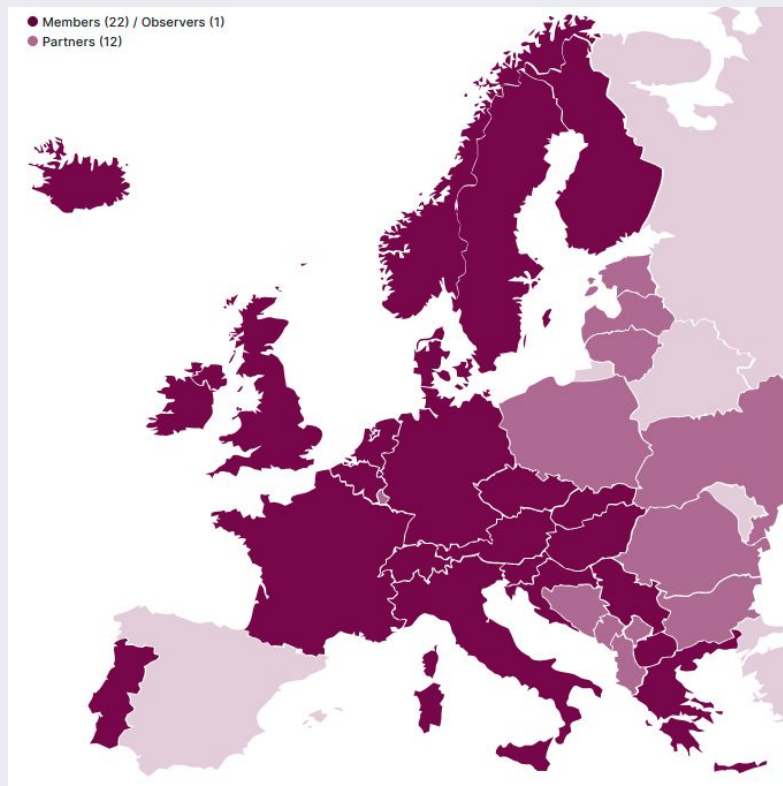
CESSDA razvija alate i servise za arhiviranje, objavu i ponovno korištenje istraživačkih podataka u društvenim znanostima.

Vizija:

Osiguravanje pristupa podacima i metapodacima u području društvenih znanosti od ključne je važnosti za razvoj znanosti i društva.



<https://cessda.eu>





Upravljanje podacima

Što je upravljanje podacima?

- odnosi se na to kako rukujete, organizirate i strukturirate svoje istraživačke podatke tijekom cijelog istraživačkog procesa
- obuhvaća razmišljanje o pitanjima koja su bitna za korištenje postojećih ili prikupljanje novih podataka
- uključuje postupke kojima se postiže integritet podataka, smanjivanje rizika od gubljenja podatka zbog tehničkih problema, osiguravanje pristupa podacima
- olakšava stvaranje detaljne i strukturirane dokumentacije koja je potrebna da bi se podaci mogli ponovo koristiti kroz nekoliko godina ili nekoliko desetljeća



Image courtesy of <http://aukeherrema.nl> CC-BY

Zašto je važno upravljati podacima?

- snalaženje u podacima, za vrijeme i nakon projekta
- suradnja na projektu
- ponovna upotreba podataka
 - podaci koje ste vi prikupili, mogu biti baš ona puzzla koja nedostaje nekom drugom (ili vama samima u budućnosti!) da odgovorite na važna pitanja
- odgovornost, vjerodostojnost, autentičnost



Plan upravljanja podacima (PUP)

- alat za strukturiranje upravljanja podacima koji nastaju ili se koriste unutar istraživačkog projekta
- izrađuje se na početku projekta, ažurira tijekom cijelog projekta
- EK već nekoliko godina zahtijeva od istraživača izradu PUP-a (Data Management Plan, DMP) tijekom prve faze provedbe projekta
- za projekte HRZZ-a obavezno je izraditi PUP prilikom prijave projekta
- struktura PUP-a može se razlikovati od financijera do financijera
- preporuča se i razvoj institucijskih politika o upravljanju podacima



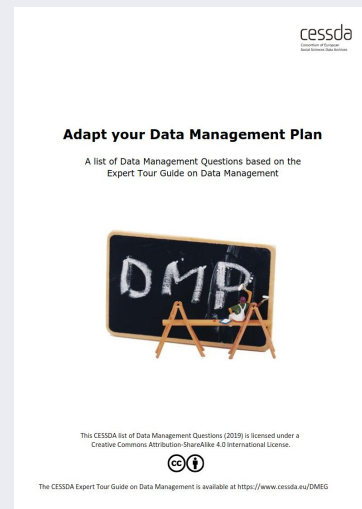
Vodiči za izradu PUP-a

- **CESSDA**

- Popis pitanja o upravljanju podacima baziran na [CESSDA Data Management Expert Guide-u](#)
[PDF](#) verzija
[Docx](#) verzija za uređivanje

- **Hrvatska zaklada za znanost**

- [Upute za izradu PUP-a](#)
- [Primjer PUP-a za društvene znanosti](#)



Plan upravljanja istraživačkim podacima

Opće informacije	
Ime predlagatelja i broj projekta	Upišite ime glavnog istraživača
Matična organizacija	
Naziv i sifra projekta	
Upravitelj podacima	upišite ime prezime te e-adresu osobe koja je odgovorna za upravljanje podacima i <i>Planom upravljanja istraživačkih podataka</i>
I Prikupljanje podataka i dokumentacija	
Koje ćete podatke prikupljati, obrađivati, stvarati ili se ponovno njima koristiti? (navedite format, vrstu i opseg podataka)	Primjer: Podaci koji će se prikupljati su odgovori na upitnik kreiran u LimeSurvey aplikaciji. Podaci koji se prikupljaju su: dob, spol, županja u kojoj je osoba boravila većinu vremena u proteklih 6 mjeseci, te odgovori na IPIP-50 upitnik. Podaci će biti pohranjeni u CSV datoteci. 5 obziram na to da planiramo prikupiti podatke od oko 1000 sudionika, veličina datoteke trebala bi iznositi maksimalno 10 MB.
Kako će se podaci prikupljati, obrađivati ili stvarati? (ukratko navedite metodologiju i procese osiguranja kvalitete te načine organiziranja podataka)	Primjer: Podaci će se prikupljati vijanim slanjem upitnika zainteresiranim skupinama sudionika. Nakon preuzimanja podataka s LimeSurvey servisa, sirovi podaci će se staviti u poseban direktorij, te će se s njih ukloniti pravo pisanja (eng. <i>write permission</i>). Podaci će se čistiti koristeći programski jezik R, a pročišćeni podaci će se spremiti u posebnu datoteku i zaseban direktorij. Skripta korištena za čišćenje podataka bit će pohranjena zajedno s podacima. Nakon čišćenja, podaci će se validirati kako bi se utvrdilo jesu li unosi valjani (npr. jesu li vrijednosti dobi valjan broj manji od 100, odgovaraju li rasponi vrijednosti na česticama <i>IPIP</i> upitnika uporišnim točkama koje su bile zadane u aplikaciji LimeSurvey). Osim toga, provest će se ručna usporedba vrijednosti u datoteci s pročišćenim podacima i vrijednosti u datoteci sa sirovim podacima za 50 nasumično odabranih unosa. Istraživači će voditi <i>README</i> datoteku u kojoj će ukratko biti opisana struktura direktorija unutar projekta, te sadržaj pojedinih datoteka ili skupova datoteka.
Koji ćete dokumentaciju i metapodatke izraditi 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)	Primjer: Dokumentacija uz podatke bit će upitnik korišten za prikupljanje podataka, pozivna pisma poslana sudionicima, upute sudionicima te upute anketarima. Dokumentacija će sadržavati opis projekta, ciljeve, vijanu skupinu te informacije o uzorkovanju, jedinici analize, načinu prikupljanja podataka, stupnju odaziva, vremenskom te prostornom obuhvatu. Izradit ćemo kodnu knjigu s opisom varijabli i oznaka vrijednosti. Opisat ćemo upotrebu varijabli otežavanja te izvedenih varijabli nastalih nakon procesa prikupljanja podataka.



PUP - Opće informacije

- Naslov projekta
- Datum i verzija plana
- Opis projekta
- Porijeklo podataka
- Glavni istraživač(i)
- Suradnici
- Financijer
- Ustanova odgovorna za projekt/podatke
- Kontakt osoba za podatke
- **Vlasnici podataka**
- **Uloge**
- Troškovi i resursi

Dileme su česte oko toga tko je vlasnik podataka, tako da to treba razriješiti odmah kod planiranja projekta.

Preporučljivo je zadužiti jednu osobu koja će brinuti o upravljanju podacima, izraditi plan i ažurirati ga (*data manager*)

Overview

Title of the project/study

Date of this plan

Description of the project

- What is the nature of the project?
- What is the research question?
- What is the project time line?

Origin of Data

- What kind of data will be used during the project?
- If you are reusing existing data: What is the scope, volume and format? How are different data sources integrated?
- If you are collecting new data can you clarify why this is necessary?

Principal researchers

- Who are the main researchers involved?
- What are their contact details?

Collaborating researchers (if applicable)

- What are their contact details and their roles in the project?

Funder (if applicable)

- If funding is granted, what is the reference number of the funding granted?
- What is the project's title in the funding contract?

Data producer

- Which organisation has the administrative responsibility for the data?

Project data contact

- Who can be contacted about the project during and after it has finished?

Data owner(s)

- Which organisation(s) own(s) the data?
- If several organisations are involved, which organisation owns what data?

Roles

- Who is responsible for updating the DMP and making sure that it's followed?
- Do project participants have any specific roles?
- What is the project time line?

Costs and Resources

- Are there costs you need to consider to buy specific software or hardware?
- Are there costs you need to consider for storage and backup?
- Are potential expenses and resources for (preparing the data for) archiving covered?
- What resources will be dedicated to data management ensuring that data will be FAIR?

Upravljanje podacima

osiguravanje kvalitete:
organiziranje, dokumentiranje, procesiranje

Što je kvaliteta podataka?

- "*fit for purpose*" - kvaliteta ovisi o kontekstu u kojem se podaci primjenjuju
- naglasak je na mogućnosti ponovnog korištenja podataka - podaci trebaju biti razumljivi i upotrebljivi bez da se kontaktiraju osobe koje su prikupile te podatke
- iz perspektive upravljanja podacima: **organiziranje, dokumentiranje i procesiranje** podataka ima velik utjecaj na kvalitetu podataka

Dizajniranje strukture datoteke s podacima

- ovisno o jedinici analize, mogućim analitičkim ciljevima i metodama analize, softveru koji koristite za prikupljanje i/ili analizu podataka
- misliti na mogućnost povezivanja s drugim podacima - uključiti jedinstvene identifikatore
- moguća tehnička ograničenja (npr. velike datoteke zahtijevaju veće računalne kapacitete za obradu)
- varijable trebaju biti u logičnom redoslijedu (varijable koje se odnose na isti fenomen trebaju biti jedna pored druge; izvorne i derivirane varijable; identifikatore staviti na početak)

primjeri struktura za kvantitativne podatke:

hijerarhijska struktura, flat datoteka, baza podataka

Imenovanje varijabli

- počnite sa slovom; nemojte započinjati brojem, upitnikom, uskličnikom ili posebnim znakom kao što su #, &, \$, @ (često su rezervirani za posebne svrhe u softverskim aplikacijama);
- imena varijabli ne mogu sadržavati razmake;
- nazivi varijabli također se koriste kao pozivni kodovi u softverskim operacijama - zato nazivi trebaju biti kratki i poštovati uobičajene zahtjeve standardnog softvera; standard je da imena varijabli ne budu duža od osam znakova;
- ne koristite dijakritičke znakove ili nacionalno specifična slova;
- nazivi neka budu smisleni (radi lakše orijentacije)

više informacija

- [Organisation of variables](#) / CESSDA DMEG

Imenovanje datoteka i struktura direktorija

- dobra imena datoteka pružaju korisne informacije o sadržaju, statusu i verziji datoteke, jedinstveno identificiraju datoteku i pomažu u klasifikaciji i sortiranju datoteka.
- nazivi datoteka koji odražavaju sadržaj datoteke olakšavaju pronalaženje datoteka (npr. sortiranjem)
- strategija imenovanja datoteka ne bi se trebala mijenjati kroz vrijeme svi suradnici bi ju trebali slijediti
- svaki projekt treba dobiti vlastiti folder
- napraviti posebne foldere za različite tipove datoteka
 - podaci (data)
 - tekstualne datoteke (rukopisi, dokumentacije) (doc)
 - datoteke vezane uz analize (stats)

više informacija:

- [File naming and folder structure](#) / CESSDA DMEG
- [How to name files](#) / Jennifer (Jenny) Bryan

```
megahrt-zadar/  
├── analyses  
│   ├── data  
│   │   ├── psycholing-estimates  
│   │   │   ├── age-of-acquisition.csv  
│   │   │   ├── concreteness.csv  
│   │   │   ├── raw-estimates  
│   │   │   │   ├── dob.txt  
│   │   │   │   ├── konkretnost.txt  
│   │   │   │   └── subjektivna.txt  
│   │   │   └── subjective-frequency.csv  
│   │   ├── reaction-time  
│   │   │   ├── clean-non-printable.sh  
│   │   │   ├── megahrt-part1-clean.csv  
│   │   │   ├── megahrt-part1.csv  
│   │   │   ├── megahrt-part2-clean.csv  
│   │   │   └── megahrt-part2.csv  
│   │   └── words  
│   │       ├── words-wave-1.csv  
│   │       └── words-wave-2.csv  
│   ├── helpers  
│   │   ├── 01-psycholing-evaluations.R  
│   │   └── 02-plot-helpers.R  
│   ├── plots  
│   │   ├── 01-plot-ppc-overall.R  
│   │   ├── 02-plot-ppc-participant.R  
│   │   ├── 03-effects-plots.R  
│   │   ├── effects-plot.png  
│   │   ├── ppc-overall.png  
│   │   └── ppc-per-participant.png  
│   ├── stats  
│   │   ├── 01-prepare-data-rt-analysis.R  
│   │   ├── 02-rt-model.stan  
│   │   ├── 03-model-fitting.R  
│   │   └── 04-ppc-sampling.R  
│   └── wrangling  
│       ├── 01-rt-data-anonimization.R  
│       ├── ...  
│       └── 05-combine-datasets.R  
├── presentation  
│   ├── hrzz-logo.jpg  
│   ├── megahr-logo.png  
│   ├── megahr-zadar.pdf  
│   └── megahr-zadar.tex  
├── README.md  
└── reference.bib
```

Dokumentacija

- mnogo dokumentacije nastaje kroz dobru istraživačku praksu - vi već radite veliki dio toga!
- počnite rano - već pri izradi PUP-a; to štedi vrijeme i trud - na početku se uloži nešto više vremena, ali se to isplati kasnije u projektu;
- razmislite o informacijama koje su potrebne za razumijevanje podataka; što će drugi istraživači (ili vi sami kroz godinu, dvije, tri) trebati da razumiju vaše podatke?
- napravite posebnu datoteku za dokumentaciju i zapišite osnovne informacije o svakom datasetu
- dobra imena varijabli su također dokumentacija; sve varijable moraju biti i dodatno opisane (labels) kodovi za vrijednosti varijabli također moraju biti dokumentirani (ako radite u SPSS-u ili Stati to je moguće u samom programu; ako radite u Excelu ili koristite .csv datoteke, obavezno opisati varijable u zasebnoj datoteci)
- koristite engleski jezik za dokumentaciju; to povećava šanse da se vaši podaci razumiju i ponovno koriste

više informacija:

- [Documentation and metadata](#) / CESSDA DMEG

Metapodaci

- dokumentacija u strojno čitljivom formatu
- standard za društvene znanosti: DDI

Study title

Data Type

Authors

Abstract

Keywords

Topic Classification

Series

Distributor

Access

Data Collector

Collection Dates

Nation

Geographical Coverage

Analysis/Observation Unit Type

Universe

Time Method

Sampling Procedure

Collection Mode

Research Instrument

Data Processing

Data Files

Data File Language

Data Version

Citation Requirement

Bibliographic Citation

Deposit Requirement

Special Terms and Conditions for Access

Disclaimer

Related Publications

Provjera podataka

- tradicionalni postupak: prikupljanje podataka, unos podataka i uređivanje/provjera podataka se danas malo kad primijenjuje - veća automatizacija procesa općenito sprječava neke vrste pogrešaka, ali istodobno proizvodi druge vrste pogrešaka! npr. pogreške u skriptama tijekom računalno potpomognutog anketiranja mogu uzrokovati sistematske greške, pomake u podacima i sl.
- provjerite sadrže li vaše podatkovne datoteke točan broj zapisa, broj varijabli ili duljinu zapisa itd.
- npr. 5–10% slučajno odabranih zapisa trebalo bi podvrgnuti detaljnijoj, dubinskoj provjeri kako bi se provjerili postupci i identificirale moguće sustavne pogreške
- provjere logičnosti i dosljednosti - raspon vrijednosti (npr. ispitanik stariji od 100 godina nije vjerojatan); najniže i najviše vrijednosti i ekstreme; odnosi između povezanih varijabli (npr. obrazovna postignuća trebaju odgovarati minimalnoj dobi, ukupan broj sati provedenih u različitim aktivnostima ne smije prelaziti 100% raspoloživog vremena); usporedite svoje podatke s povijesnim podacima (npr. provjerite broj članova kućanstva s prethodnim valom panel ankete).

Više informacija: [Data entry and integrity](#) / CESSDA DMEG

Formati datoteka i konverzije

- Izbor formata datoteke ovisi o fazi istraživanja
- kratkoročno (za vrijeme trajanja projekta) mogu se koristiti zatvoreni (*proprietary*) formati
 - Rich text format (*.rtf), MP3, MPEG, JPG, MS Excel (*.xls), SPSS (*.sav, *.por), STATA (*.dta) naširoko se koriste i možete pretpostaviti da će biti čitljivi neko razumno vrijeme
- za dugoročnu pohranu - otvoreni formati
 - PDF/A, CSV, TIFF, ASCII, Open Document Format (ODF), XML, Office Open XML, JPEG 2000, PNG, SVG, HTML, XHTML, RSS, CSS

Kontrola verzija - očuvanje autentičnosti

- mnoge su promjene moguće u podacima: čišćenje podataka; ispravljanje grešaka; konstruiranje novih varijabli; dodavanje novih informacija iz vanjskih izvora; konverzija formata; mijenjanje strukture podataka

Moguće metode kontrole verzija:

- jedinstveno identificiranje različitih verzija datoteka korištenjem konvencije sustavnog imenovanja, kao što je korištenje brojeva verzija ili datuma (format datuma korisno je pisati prema ISO standardu: GGGG-MM-DD)
- korištenje mogućnosti verzioniranja unutar softvera koji koristite;
- korištenje softvera za praćenje verzija kao što je *git*;
- Google Drive, Dropbox, i sl. imaju ugrađene kontrole verzija (ali i nametnuta pravila!)
- dizajniranje i korištenje **tablice za kontrolu verzija** - u svakom slučaju je dobro da takva povijest bude uključena u datoteku; tako možete pratiti verzije i opis o promjena koje su napravljene

Upravljanje podacima

etička i pravna pitanja;
sigurna pohrana podataka tijekom projekta

Istraživačka etika i zaštita osobnih podataka

- Prilikom prikupljanja, korištenja i dijeljenja istraživačkih podataka, etička pitanja i pravne obveze su važne teme
- Tri važna principa:
 - **pridržavanje znanstvenih standarda** - uzeti u obzir sve relevantne dokaze i prezentirati svoje istraživanje bez izostavljanja, pogrešnog predstavljanja ili prijevare; pri formuliranju istraživačkih pitanja, dizajnu ankete, upitnike ili intervju, istraživači ne prejudiciraju ishod svojim izborom pitanja ili postupaka
 - **usklađenost sa zakonom** - poznavanje relevantnih nacionalnih i međunarodnih zakona; kod međunarodnih projekata, to može uključivati različite zakone; od posebne važnosti su oni koji se odnose na zaštitu podataka i intelektualno vlasništvo
 - **izbjegavanje društvene i osobne štete** - istraživački projekt treba biti osmišljen odgovorno prema sudionicima; npr. sudjelovanje u istraživačkom projektu treba biti dobrovoljno i kontinuirano na temelju potpuno informiranog pristanka.

Etička provjera

- potrebna kada se prikupljaju (osjetljivi) osobni podaci
- zaštititi sigurnost, prava i dobrobit sudionika istraživanja te promicati etički ispravno istraživanje
- to uključuje osiguravanje da je istraživanje u skladu s nacionalnim i međunarodnim zakonima o zaštiti podataka koji se odnose na korištenje osobnih podataka prikupljenih u istraživanju
- povjerenstva za etiku istraživanja

Autorska prava

- tko je vlasnik podataka? tko ima pravo dati dozvolu za pohranu i pružiti podatke na korištenje?
 - autor
 - financijer?
 - ustanova pri kojoj je autor zaposlen?
- kako financijeri ili ustanove stječu pravo raspolaganja podacima?
 - isključivo pisanim putem!
 - ugovor o financiranju projekta
 - ugovor o radu
 - izjava o prijenosu prava
- korištenje sekundarnih podataka nije moguća bez dozvole vlasnika, odn. nositelja prava

Načela obrade osobnih podataka (GDPR)

- **zakonitost, poštenost i transparentnost** - sudionik je obaviješten o tome što će se raditi s podacima
- **izvorna svrha** - podatke treba prikupljati u određene, eksplicitne i legitimne svrhe i ne smiju se dalje obrađivati na način koji je nespojiv s tim svrhama; daljnja obrada u svrhe arhiviranja u javnom interesu, u svrhe znanstvenog ili povijesnog istraživanja ili u statističke svrhe, ne smatra se neusklađenom s prvotnim svrhama
- **smanjenje količine podataka** - osobni podaci koji se prikupljaju trebaju biti primjereni, relevantni i ograničeni na ono što je potrebno.
- **točnost** - osobni podaci trebaju biti točni i, prema potrebi, ažurirani;
- **ograničenje pohrane** - osobni podaci trebaju se čuvati u obliku koji omogućuje identifikaciju nositelja podataka ne dulje nego što je potrebo za svrhe za koje se osobni podaci obrađuju.
- **cjelovitost i povjerljivost** - osigurati odgovarajuću sigurnost osobnih podataka, uključujući zaštitu od neovlaštene ili nezakonite obrade i od slučajnog gubitka, ništenje ili oštećenje, korištenjem odgovarajućih tehničkih ili organizacijskih mjera

Pravni temelji za obradu osobnih podataka (GDPR)

- ispitanik je dao **privolu** za obradu svojih osobnih podataka u jednu ili više posebnih svrha;
- obrada je nužna za izvršavanje ugovora u kojem je ispitanik stranka ili kako bi se poduzele radnje na zahtjev ispitanika prije sklapanja ugovora;
- obrada je nužna radi poštovanja pravnih obveza voditelja obrade;
- obrada je nužna kako bi se zaštitili ključni interesi ispitanika ili druge fizičke osobe;
- obrada je nužna za izvršavanje **zadaje od javnog interesa** ili pri izvršavanju službene ovlasti voditelja obrade;
- obrada je nužna za potrebe legitimnih interesa voditelja obrade ili treće strane, osim kada su od tih interesa jači interesi ili temeljna prava i slobode ispitanika koji zahtijevaju zaštitu osobnih podataka, osobito ako je ispitanik dijete

Za znanstvena istraživanja mogući temelji su: privola i zadaća od javnog interesa. Hrvatska implementacija GDPR-a ne regulira dodatno korištenje podataka u znanstvene svrhe, tako da je privola sigurniji temelj.

Informirani pristanak

- sudioniku istraživanja pružiti informacije o istraživanju, tako da sudionik može dobrovoljno i informirano odlučiti hoće li prihvatiti ili odbiti suradnju
- ako se prikupljaju osobni podaci - potrebna je **privola (GDPR)**
 - obavezno u pisanom obliku ili kao zvučni zapis
 - formulacije trebaju biti nedvosmislene i specifične
 - privola se daje jasnom potvrdnom akcijom, najbolje granularno
- ako se ne prikupljaju osobni podaci (anonimne ankete) - **etički pristanak**
 - sudionik pristaje ili ne pristaje sudjelovati u istraživanju na temelju informacija koje mu je istraživač pružio

Informirani pristanak - primjeri

Istraživanje “Izdavačke i uredničke prakse hrvatskih znanstvenih časopisa u humanističkom području i uloga javnih znanstvenih politika”

- [informativno pismo](#)
- [informirani pristanak](#)

Obrazac za informacijsko pismo i privolu koji su izradili kolege u [Norwegian Centre for Research Data](#)

- [template for information letter and consent](#)

Resursi dostupni u [UK Data Service](#)

- [Consent for sharing](#)
- [Model consent form](#)

Informiranje o korištenju podataka (1)

Sudionike je potrebno informirati o tome tko i kako će koristiti podatke - pritom ne isključiti mogućnost da podatke koriste i osobe (ili samo istraživači) izvan istraživačkog tima koji je podatke prikupio.

Razmislimo o ovim formulacijama:

"Prikupljeni podaci objavit će se samo u sumarnom obliku u izvještajima i znanstvenim radovima."

"Vaši će odgovori biti dostupni isključivo niže potpisanim istraživačima za potrebe znanstvenih analiza."

Kako ste vi u svojem posljednjem istraživanju informirali sudionike o korištenju podataka? Zašto vam je bitna takva formulacija?

Informiranje o korištenju podataka (2)

- istraživačima je bitno poručiti sudionicima da će se s njihovim **osobnim** podacima postupati odgovorno i tako zadobiti povjerenje sudionika
- treba napraviti jasnu **razliku između osobnih podataka i anonimiziranih podataka!**
- za osobne podatke želimo ograničiti upotrebu samo na usku skupinu članova istraživačkog tima
 - osobne podatke možemo dijeliti s drugima samo ako sudionik na to pristane!
 - treba uzeti u obzir i etičku dimenziju - ponekad ispitanici nisu svjesni na što pristaju, tako da javno dijeljenje podataka u kojima je moguće identificirati pojedince treba izbjegavati u svakom slučaju - za takve podatke često treba ograničiti pristup samo za istraživačke svrhe
 - kod dijeljenja podataka uvijek se uklanjanju direktni identifikatori, no to nije dovoljno za zaštititi identitet sudionika!
- u svrhu ponovnog korištenja, pristup je najčešće dovoljno omogućiti samo anonimiziranim podacima;
 - svedeno, sudionike treba obavijestiti da će anonimizirani podaci biti dostupni drugim istraživačima

Formulacije u pristanku koje omogućuju dijeljenje osobnih podataka

Za istraživanja u kojima podatke nije moguće anonimizirati tako da se ne izgubi njihova informacijska vrijednost (najčešće za kvalitativna istraživanja):

"Razumijem da će moji osobni podaci poput imena, adrese e-pošte i broja telefona biti dostupni samo osobama koje provode ovaj projekt"

- direktne identifikatore najčešće ne želimo podijeliti s drugima! treba ih izbrisati kada više nisu potrebni

"Slažem se da moji odgovori budu pohranjeni u Hrvatskom arhivu podataka za društvene znanosti."

"Razumijem da će drugi istraživači imati pristup ovim podacima samo ako se obvežu da će čuvati povjerljivost informacija sadržanih u podacima."

Informacije sudionicima o dijeljenju podataka

Dijeljenje anonimiziranih podataka ne zahtijeva izričit pristanak, ali informiranje sudionika o tome što će se dogoditi s njihovim odgovorima.

”Anonimizirani podaci bit će trajno pohranjeni u Hrvatskom arhivu podataka za društvene znanosti kako bi bili dostupni u znanstvene svrhe ili za korištenje u podučavanju studenata.”

”Vaši podaci su anonimizirani, što znači da nije moguće otkriti koja je osoba dala koji odgovor. Takvi podaci bit će trajno pohranjeni u Hrvatskom arhivu podataka za društvene znanosti i javno dostupni kako bi se povećala iskoristivost i odjek informacija koje ste nam u ovom upitniku pružili.”

”Svi osobni podaci koji bi vas mogli identificirati bit će uklonjeni ili promijenjeni prije nego što se datoteke podijele s drugim istraživačima”

Pohrana podataka za vrijeme trajanja projekta

Prilikom odabira prikladnog rješenja za pohranu koje odgovara potrebama vašeg projekta potrebno je odgovoriti na mnoga pitanja. Na primjer:

- Koliko prostora za pohranu trebam?
- Tko treba pristup podacima?
- Koje mjere opreza trebam poduzeti da zaštitim svoje podatke od gubitka?
- Koja su rješenja za pohranu prikladna za osobne podatke?

Laptop, tablet, vanjski disk, USB stick, CD?

- koristiti samo za privremenu, kratkotrajnu pohranu neosjetljivih podataka
- **ako radite s osjetljivim podacima - obavezna enkripcija i jaka zaštitna lozinka**
- nemojte za glavne kopije datoteka - samo za radne verzije
- nakon završene određene faze rada obavezno obrisati podatke
 - želite spriječiti rizik od ostavljanja prijenosnih uređaja u kafiću, tramvaju.... šansa da se to dogodi je mala, ali **potencijalna šteta je ogromna!**

Podaci u oblaku

- korisno za zajednički, udaljeni pristup podacima i drugim datotekama
- Google Drive, OneDrive, Dropbox, OSF, ... ?
 - obavezno pročitajte uvjete servisa; usredotočite se na prava korištenja sadržaja koje dajete pružatelju usluge
 - izbjegavajte korištenje osobnih korisničkih računa - politike privatnosti obično su puno manje za osobne korisničke račune u odnosu na organizacijske
 - neka ovo ne bude vaše jedino rješenje za pohranu i izradu sigurnosnih kopija
- ako prikupljate osobne podatke:
 - ispitanici moraju biti informirani o tome gdje čuvate podatke
 - obavezno koristite servise čiji serveri su u Europi
 - preporuča se servis Srca <https://www.srce.unizg.hr/puh>
- **osjetljive osobne podatke obavezno je enkriptirati!**





Hvala na pažnji.

arhiv.podataka@ffzg.hr
<https://crossda.hr>
@crossda_data