

VODIČ ZA ANONIMIZACIJO V KVALITATIVNEM RAZISKOVANJU: BESEDILNI PODATKI

Različica 2.0

Sonja Bezjak, Sergeja Masten, Irena Vipavc Brvar



Vodič za anonimizacijo v kvalitativnem raziskovanju: besedilni podatki

Elektronska izdaja, različica 1.0

Sonja Bezjak, Sergeja Masten, Irena Vipavc Brvar

Založila in izdala: Fakulteta za družbene vede, Založba FDV

Za založbo: Iztok Prezelj, dekan

Jezikovni pregled: Tomaž Petek

Ljubljana, 2025

To delo je ponujeno pod licenco Creative Commons Priznanje avtorstva – Nekomercialno. Mednarodna licenca (CC BY-NC 4.0). Če si želite ogledati kopijo tega dovoljenja, obiščite: <https://creativecommons.org/licenses/by-nc/4.0/>.

Publikacija je v digitalni obliki prostodostopna prek <https://ebooks.uni-lj.si>.

Katalogni zapis o publikaciji (CIP) pripravili v Narodni in univerzitetni knjižnici v Ljubljani

[COBISS.SI](#)-ID **236061699**

ISBN 978-961-295-117-7 (PDF)

DOI: <https://doi.org/10.51936/9789612951177>

KAZALO

UVODNIK.....	4
OSEBNI PODATKI IN ZAKONODAJA.....	4
RAZISKOVALNA ETIKA.....	6
ŽIVLJENJSKI KROG RAZISKOVANJA.....	6
ANONIMIZACIJA V KVALITATIVNEM RAZISKOVANJU.....	7
NAČRTOVANJE ANONIMIZACIJE.....	8
PRED ZAČETKOM ANONIMIZACIJE.....	9
POSTOPEK ANONIMIZACIJE.....	9
KAJ ANONIMIZIRATI?.....	10
POSREDNI IN NEPOSREDNI IDENTIFIKATORJI	10
KAKO ANONIMIZIRATI?	11
PSEVDONIMIZACIJA.....	14
ODSTRANITEV PODATKOV	14
PREGLEDNICA: IDENTIFIKATORJI IN METODE	15
DOSTOP POD POSEBNIMI POGOJI	19
ARHIVIRANJE IN OBJAVA PODATKOV V ADP	20
DODATNI VIRI	21
UPORABLJENI VIRI	22
Priloga: Seznam pogostih identifikatorjev in primeri anonimizacije.....	23

UVODNIK

Osnovno poslanstvo **Arhiva družboslovnih podatkov (ADP)** je prevzem, izbor in trajno skrbništvo kakovostnih raziskovalnih podatkov, zanimivih za družboslovne analize, s poudarkom na problemih, povezanih s slovensko družbo ali sicer pomembnih za slovensko družbo in družboslovje ne glede na geografske meje ([Poslanstvo ADP](#)). Pri uresničevanju poslanstva se pogosto srečujemo s primeri, ko podatki še niso zadostno urejeni, da bi jih lahko prevzeli in ustrezno obdelali. V takih primerih raziskovalkam in raziskovalcem svetujemo glede nadaljnje obdelave, kot sta npr. čiščenje in anonimizacija.

Poseben izziv predstavlja varstvo osebnih podatkov, vključno s posebnimi vrstami osebnih podatkov. Podatki o spolu, starosti, verski pripadnosti, o političnih prepričanjih so v družboslovnih raziskavah stalnica, ki omogočajo razumevanje sprememb na ravni posameznika, skupnosti in družbe.

Da bi lahko ohranjali in delili raziskovalne podatke, vključno z osebnimi podatki, morajo biti izpolnjene pravne podlage za zbiranje in obdelavo osebnih podatkov ter po potrebi izbrani zaščitni ukrepi, kot je npr. anonimizacija ali omejen dostop do podatkov. V ADP opažamo, da je anonimizacija boljše uveljavljena v kvantitativnih kot kvalitativnih raziskavah. Zato se na področju kvalitativnega raziskovanja pogosto srečujemo z več zadržki in negotovostmi glede deljenja podatkov.

Na podlagi izkušenj smo pripravili vodič za anonimizacijo v kvalitativnem raziskovanju, s poudarkom na besedilnih podatkih, ki običajno nastanejo z metodo intervjujev ali fokusnih skupin. V vodiču smo osvetlili najpogostejše izzive, s katerimi se v praksi srečujemo pri izvedbi postopka anonimizacije.

OSEBNI PODATKI IN ZAKONODAJA

Osební podatki so vse informacije, ki se nanašajo na določeno ali določljivo fizično osebo. Določljiv posameznik je tisti, ki ga je mogoče neposredno ali posredno identificirati, zlasti z navedbo identifikatorja, kot so: ime, identifikacijska številka, podatki o lokaciji, spletni identifikator ali z navedbo enega ali več dejavnikov, ki so značilni za fizično, fiziološko, genetsko, duševno, gospodarsko, kulturno ali za družbeno identiteto tega posameznika.

Ta definicija izhaja iz **Splošne uredbe o varstvu podatkov (GDPR)** in je prenesena v **Zakon o varstvu osebnih podatkov (ZVOP-2)**, ki velja v Sloveniji. Z [Uredbo](#) EU o varstvu osebnih podatkov (2018) in Zakonom o varstvu osebnih podatkov ([ZVOP-2](#), 2023) smo v raziskovalni skupnosti zakonsko zavezani k odgovorni in ozaveščeni obdelavi osebnih podatkov.

Obdelava osebnih podatkov po določbi 2. točke 4. člena GDPR pomeni vsako dejanje ali niz dejanj, ki se izvaja v povezavi z osebnimi podatki ali nizi osebnih podatkov, tj. z avtomatiziranimi sredstvi ali brez njih. Sem spadajo: zbiranje, beleženje, urejanje, strukturiranje, shranjevanje, prilagajanje ali spreminjanje, priklic, vpogled, uporaba, razkritje s posredovanjem, razširjanje ali drugačno omogočanje dostopa, prilagajanje ali kombiniranje, omejevanje, izbris ali uničenje.

Ta široka definicija je namenjena zagotavljanju celovite zaščite osebnih podatkov v vseh fazah njihovega življenjskega kroga.

Po GDPR in ZVOP-2 so raziskovalci in raziskovalke dolžni pred začetkom zbiranja podatkov udeležence in udeleženske ustrezno seznanimi o vsebini in poteku raziskave ter o načrtovani obdelavi njihovih osebnih podatkov. To vključuje pridobitev njihovega prostovoljnega, informiranega in nedvoumnega soglasja za sodelovanje v raziskavi ter za obdelavo osebnih podatkov skladno z določbami 6. člena GDPR.¹

Če raziskovalec ali raziskovalka ne pridobi ustreznega soglasja za obdelavo osebnih podatkov, obdelava teh podatkov ni dovoljena, razen če obstaja druga pravna podlaga, kot je določena v 6. členu GDPR ali 69. členu ZVOP-2.

Namig

Razmislimo o možnosti, da udeležencem in udeleženkam omogočimo izbiro, ali želijo v intervjuju ohraniti svojo identiteto. To pomeni, da jih bomo v objavah, kot so članki, monografije ali podatkovne objave, navajali s polnim imenom in priimkom. Praksa kaže, da se v določenih situacijah, na primer, ko pogovor teče o osebnih izkušnjah ali spominih na pomembne zgodovinske dogodke, udeleženci odločijo ostati vidni, saj je s tem prepoznan in priznan njihov prispevek. Pri tem je pomembno, da ne pozabimo na zaščito zasebnosti tretjih oseb – identifikatorje, po katerih bi jih prepoznali, moramo anonimizirati, saj nismo prejeli njihovega soglasja za obdelavo njihovih osebnih podatkov.

Zakonodaja s področja varstva osebnih podatkov ne velja za:

1. **obdelavo podatkov, ki so anonimizirani do te mere, da posameznik ni več določljiv;**
2. **podatke, ki so del javne komunikacije ali zadevajo javne osebnosti.**

Za podrobnejše informacije in razumevanje specifičnih primerov je priporočljivo posvetovanje z osebo, pooblaščenca za varstvo osebnih podatkov (DPO) na raziskovalni organizaciji, ali z informacijskim pooblaščencom. Za podrobnejši vpogled v tematiko priporočamo *Priročnik o načrtovanju ravnanja z raziskovalnimi podatki* (2025), poglavje [Zaščita podatkov](#).

¹ V družboslovju je soglasje pogostejše uporabljena pravna podlaga, zakonodaja pa dopušča tudi druge možnosti. Podrobnejše informacije o informiranem soglasju in drugih pravnih podlagah so zbrane v Priročniku o načrtovanju ravnanja z raziskovalnimi podatki, poglavje [Zaščita](#)). Več o oblikovanju informiranega soglasja v poglavju 5. Protect [Informed Consent](#) učbenika CESSDA DMEG.

RAZISKOVALNA ETIKA

V raziskovalnem delu smo zavezani tudi k spoštovanju raziskovalne etike; ta vključuje tudi raziskave na ljudeh in o ljudeh. Tako npr. Evropski kodeks ravnanja za raziskovalno integriteto določa:

»Raziskovalci ravnaajo z udeleženci in s subjekti raziskav (človeškimi, živalskimi, kulturnimi, biološkimi, okoljskimi ali fizikalnimi) ter z njimi povezanimi podatki spoštljivo in skrbno ter skladno z zakonskimi določbami in etičnimi načeli« (ALLEA, 2023).²

Na raziskovalnih organizacijah večinoma delujejo komisije za etična vprašanja, ki pomagajo pri presoji tveganj in svetujejo glede primernih praks, pripravljajo smernice in pravilnike, ki se nanašajo na zagotavljanje etičnih praks.³ Za več priporočamo Priročnik o načrtovanju ravnanja z raziskovalnimi podatki, [Komisije za etična vprašanja](#).

V povezavi s tem pa velja premisliti tudi o načelih CARE (C – kolektivna korist, A – pravica nadzora, R – odgovornost, E – etičnost), ki so nastala v staroselskih skupnostih in so priporočena, kadar sodelujemo z manjšinskimi skupnostmi ([Načela CARE](#)).

ŽIVLJENJSKI KROG RAZISKOVANJA

Preden se poglobimo v anonimizacijo besedilnih podatkov v družboslovju, opozorimo še na pomen skrbnega in pravočasnega načrtovanja ravnanja z raziskovalnimi podatki. K temu nas zavezujejo načela odprte znanosti, s ciljem, da bi bili podatki, financirani z javnimi sredstvi, najdljivi, dostopni, interoperabilni in uporabljivi oz. na kratko FAIR, tudi po koncu raziskave. Ko ravnamo z osebnimi podatki, zaradi varstva in zaščite pogosto ni mogoče zagotavljati »odprtih podatkov«, takrat se ravnamo po načelu »odprti, kolikor je mogoče, in zaprti, kolikor je nujno«.

Pri uresničevanju teh načel nam bo v pomoč [Načrt ravnanja z raziskovalnimi podatki](#), ki nas s premišljenimi vprašanji strukturirano vodi skozi življenjski krog podatkov in nam pomaga vnaprej razmisliti o tem, kako podatke zbrati in obdelati, da jih bomo lahko v primerni obliki arhivirali in objavili. V praksi to pomeni, da v zasnovi raziskave predvidimo delo z osebnimi podatki, ocenimo, katere podatke nujno potrebujemo ([Načelo najmanjšega obsega podatkov](#), Informacijski pooblaščenec), kako bomo poskrbeli za njihovo varnost ter kako bomo zagotovili transparentnost in objavo podatkov kljub omejitvam. V branje priporočamo učbenik CESSDA Data Management Expert Guide, v katerem so teme in rešitve za družboslovje obsežneje predstavljene (CESSDA DMEG, poglavje [Protect](#)).

² ALLEA (2023) Evropski kodeks ravnanja za raziskovalno integriteto – revidirana izdaja 2023 (slovenščina). Berlin. DOI 10.26356/ECOC-Slovenian.

³ Npr. [Etika in integriteta v raziskovanju](#), Univerza v Ljubljani.

ANONIMIZACIJA V KVALITATIVNEM RAZISKOVANJU

Anonimizacija je »obdelava podatkov, s katero se odstranijo ali spremenijo tisti podatki, na podlagi katerih je mogoče določiti konkretnega posameznika ali skupino posameznikov« ([Slovar odprte znanosti](#)). Gre torej za obdelavo osebnih podatkov na način, da se z različnimi tehnikami zakrije informacije, po katerih bi bilo mogoče prepoznati udeležence in udeleženke raziskave.

Pri odločitvi za anonimizacijo moramo upoštevati več ključnih vidikov:

- **Izražena volja udeležencev in udeleženk**, kot je navedena v soglasju ali drugih oblikah privolitve glede obdelave njihovih osebnih podatkov.
- **Ocena tveganja**, da bi raziskovalni podatki razkrili značilne in edinstvene informacije o posamezniku, kar bi mu lahko povzročilo škodo.
- **Vložena sredstva in napor v zbiranje podatkov** ter ocena možnosti in stroškov morebitnega ponovnega zbiranja. Redki in težko dostopni podatki imajo lahko visoko dodano vrednost.
- **Ohranitev vsebinske bogatosti podatkov**, ki je ključna za zagotavljanje znanstvene preverljivosti in uporabnosti rezultatov.

Če anonimizacijo izberemo kot ukrep za zaščito podatkov, moramo poskrbeti, da zakrijemo edinstvene informacije, po katerih bi bilo mogoče identificirati udeležence. Obenem pa moramo paziti, da z anonimizacijo ne odstranimo preveč podatkov, saj bi to lahko bistveno zmanjšalo njihovo uporabnost – za lastno analizo in morebitne nadaljnje raziskave.

Kadar bi anonimizacija vodila v izgubo dragocenih informacij ali znatno okrnila uporabnost podatkov, je smiselno že v fazi načrtovanja raziskave razmisliti o alternativnih rešitvah, kot je npr. **omejen dostop do podatkov**.

NAČRTOVANJE ANONIMIZACIJE

Že v fazi načrtovanja raziskave je pomembno, da razmislimo o vseh korakih v življenjskem krogu podatkov, ki lahko vplivajo na ravnanje z osebnimi podatki. Pripravimo načrt dela, ki vključuje naslednje vidike:

- **Določitev odgovornih oseb in časovnice:** kdo in kdaj bo obdeloval podatke.
- **Uporabljena programska oprema:** izbira orodij za zbiranje, obdelavo in za hrambo podatkov.
- **Varnostni ukrepi za hrambo podatkov:** kako bomo zagotovili zaščito osebnih podatkov pred nepooblaščenim dostopom.
- **Ocena občutljivosti raziskovalne teme:** določimo morebitna tveganja, povezana z vsebino raziskave.
- **Število udeleženk in udeležencev:** večje število sodelujočih lahko zmanjša tveganje za prepoznavo posameznika.
- **Status udeležencev:** upoštevamo demografske in družbene značilnosti, kot so: starost, spol, verska pripadnost, pripadnost manjšinam, poklici ali življenjske okoliščine, ki lahko povečajo tveganje za razkritje identitete.
- **Časovni okvir raziskave:** starejši podatki običajno predstavljajo manjše tveganje za razkritje, zlasti kadar se nanašajo na zgodovinsko oddaljene dogodke.
- **Druge posebnosti:** upoštevamo dodatne dejavnike, ki bi lahko vplivali na varnost in anonimizacijo podatkov.⁴

⁴ Povzeto po finskem arhivu FSD: <https://www.fsd.uta.fi/aineistonhallinta/en/anonymisation-and-identifiers.html>

PRED ZAČETKOM ANONIMIZACIJE

Pripravimo smernice, navodila in druge protokole, ki bodo v pomoč pri izvajanju postopkov anonimizacije. Ti vključujejo naslednje:

- **Navodila za izvedbo anonimizacije** in sprotno spremljanje poteka dela, s čimer zagotovimo sistematično načrtovanje sprememb in njihovo dosledno beleženje skozi celoten postopek.
- **Priprava poimenskega seznama sodelavk in sodelavcev**, ki sodelujejo pri anonimizaciji, ter **vodenje evidence o izvedbi postopkov**, kar omogoča sledljivost in odgovornost.
- **Načrt za varen prenos in hrambo dokumentov**, ki vsebujejo osebne podatke oziroma drugo gradivo, ki bi lahko razkrilo identiteto udeležencev.⁵
- **Načrt za delo z matično datoteko**: ustvarimo kopijo izvirne (matične) datoteke transkripcije, na kateri se izvaja anonimizacija. Izvirnik hranimo na varnem in ločenem mestu, da ga po potrebi lahko uporabimo za morebitne poznejše preverbe.
- **Načrt za uničenje datotek z osebnimi podatki**, ki ga pripravimo skladno s pravnimi podlagami oziroma z dogovori z udeleženci (npr. soglasje). Na primer: zvočni posnetek izbrišemo po pripravi transkripta; transkripcijo z osebnimi podatki izbrišemo po pripravi anonimizirane različice besedilne datoteke.

POSTOPEK ANONIMIZACIJE

Informacije, ki bi lahko razkrile identiteto posameznika ali posameznice, lahko zaščitimo z različnimi postopki, kot so:

- nadomestitev;
- posplošitev;
- združevanje;
- popačenje ali
- odstranitev podatkov.

⁵ Več o varni hrambi in ravnanju s podatki iz delavnice *Pravni in etični vidiki ravnanja z raziskovalnimi podatki 2018*: [Storing and Securely Handling Research Data, Copyright, Access Controls and Licensing Data, Gaining Informed Consent for Data Sharing, Research Ethics and Data Protection Legislation](#)

V praksi pogosto uporabimo kombinacijo več teh postopkov. Pri tem sledimo načelu, da želimo ohraniti čim višjo informativno vrednost podatkov, obenem pa jih zaščititi le v tolikšni meri, kot je nujno potrebno.

KAJ ANONIMIZIRATI?

Anonimiziramo lahko podatke, zbrane z različnimi metodami in napravami ali ročno, na primer z anketami, zvočnimi posnetki, s slikami, podatki z družbenih omrežij, sodnimi spisi in z drugimi viri. V tem vodiču so predstavljena priporočila za anonimizacijo podatkov, pridobljenih z intervjuji v obliki besedilnih datotek. Takšne datoteke so pogoste v družboslovju pa tudi v drugih disciplinah, ki pri raziskovanju uporabljajo pogovore z ljudmi.

Pred začetkom anonimizacije **opredelimo neposredne in posredne identifikatorje** – to so vse informacije, na podlagi katerih bi bilo mogoče razkriti identiteto sodelujočih v raziskavi. V praksi se pogosto zgodi, da dodatne identifikatorje, ki jih sprva nismo predvideli, odkrijemo šele med postopkom anonimizacije.

Bodimo pozorni tudi na primere, ko respondenti in respondentke v pogovoru omenjajo **osebne podatke tretjih oseb** (npr. imena, kraje, poklice in druge prepoznavne podatke), za katere nimamo dovoljenja za deljenje. Takšne informacije lahko dodatno povečajo tveganje za razkritje identitete sodelujočih, obenem pa nimamo dovoljenja, da bi delili informacije o tretjih osebah.

Prav tako preverimo, ali so bili za isto osebo v pogovoru uporabljeni **različni načini poimenovanja** (npr. vzdevek, ime, priimek). V tem primeru moramo dosledno anonimizirati vse pojavitve, ne glede na obliko.

POSREDNI IN NEPOSREDNI IDENTIFIKATORJI

Večinoma najprej pomislimo na **neposredne identifikatorje**, kot so: ime in priimek, naslov prebivališča ali e-poštni naslov, vendar je pri anonimizaciji treba razmisliti tudi o **manj očitnih, t. i. posrednih identifikatorjih**. To so značilnosti, po katerih udeleženec ali udeleženka izstopa iz preučevane populacije, na primer večje število otrok, etnična, narodna ali druga pripadnost, zaposlitveni sektor, prostoračasne dejavnosti itn.

Posredni identifikatorji so lahko tudi takšni podatki, ki **vsak zase ne razkrivajo identitete**, njihova kombinacija pa lahko pripelje do prepoznavnosti posameznika.

O tem, **kateri identifikatorji predstavljajo tveganje za razkritje**, razmislimo že v fazi načrtovanja raziskave, ko opredeljujemo cilje, namen in ciljno populacijo.

Primer: ocenimo, ali bi bil posameznik iz manjšega kraja lahko prepoznaven, če v raziskavi navajamo, da živi v gospodinjstvu s petimi otroki.

Namig

Pri iskanju identifikatorjev si lahko pomagamo z orodji »**Najdi (Find)**« in »**Zamenjaj (Replace)**«, če jih omogoča naš urejevalnik besedila. Tako bomo lažje in hitreje izvedli spremembe na več mestih hkrati. Bodimo pri tem **temeljiti in dosledni**.

KAKO ANONIMIZIRATI?

Pri anonimizaciji na mestu identifikatorja uporabimo **besedno zvezo v oglatem oklepaju**, in sicer v obliki **[zakrita vrednost + opis]**. S tem ohranimo sledljivost sprememb – za lastno dokumentacijo in za potrebe morebitnih poznejših raziskav.

Med postopkom anonimizacije uporabljamo **šifrant**, s katerim sistematično beležimo zamenjane izraze. Tako zmanjšamo tveganje napak in nedoslednosti. Šifrant je treba **varno hraniti** in ga po končanem postopku **uničiti ali ustrezno anonimizirati**, če vsebuje osebne podatke.

NADOMESTITEV

Zamenjava osebnih imen

Namesto resničnih imen uporabimo **izmišljena imena**, pri čemer:

- se držimo kulturne in jezikovne skladnosti (npr. slovenskih imen ne prevajamo v angleška in nasprotno),
- po potrebi ime zamenjamo z vlogo ali drugo opisno lastnostjo.

Primer 1:

Stanka Novak je izdala beli recept.

☐ *[zakrita ime in priimek: Metka Jazbec] je izdala beli recept.*

Primer 2:

Stanka Novak je izdala beli recept.

☐ *[zakrita ime in priimek: Zdravnica] je izdala beli recept.*

Ime in priimek lahko nadomestimo tudi z **vlogo v odnosu**, na primer:

- *[zakrita ime in priimek: prijateljica]*
- *[zakrita ime in priimek: sošolka]*
- *[zakrita ime in priimek: sosed]*

Takšna poimenovanja omogočajo razumevanje konteksta, ne da bi razkrili identiteto.

POSPLOŠITEV

Posplošitev pomeni, da s kategorizacijo ali klasifikacijo oblikujemo skupine, razrede ali kategorije, ki opisujejo lastnosti posameznika ali posameznice na način, da ohranimo čim višjo informativno vrednost originalnih podatkov. Kategorije določimo na podlagi vrednosti oziroma vsebine podatkov in jih dosledno uporabljamo skozi celoten proces.

Primer 3:

Intervjuvali smo 44-letno profesorico sociologije na Fakulteti za družbene vede v Ljubljani, ki je poročena in ima dva otroke, stara 9 in 11 let (primer je izmišljen). Posplošimo informacije na naslednji način:

- **Spol:** ženski
- **Starost:** 40–50 let
- **Delovno mesto:** Univerza v Ljubljani
- **Poklic:** profesorica družboslovnih ved
- **Status:** poročena, vsaj en otrok

Pri tem moramo biti pozorni, da posplošitev ne razkrije preveč informacij. Na primer, če bi bile ženske v preučevanem prostoru slabše zastopane, bi že podatek o spolu lahko predstavljal tveganje za razkritje identitete.

Posplošitev krajevnih in državnih imen

Krajevna imena posplošimo na višjo geografsko enoto, stopnja posplošitve pa je odvisna od zaščite podatkov in pomena geografske enote za razumevanje podatkov.

Primer 4:

Naša sedemčlanska družina prihaja iz Cvetličnega Dola.

- ☐ *Naša sedemčlanska družina prihaja iz [zakrito ime kraja: občine Mozirje].*

Primer 5:

V soseski je veliko priseljencev, najmanj moteča sta iz Španije.

- ☐ *V soseski je veliko priseljencev, najmanj moteča sta iz [zakrito ime države: evropske države].*

Posplošitev imen delovnih mest in izobraževalnih institucij

Namesto natančnega opisa delovnega mesta ali šole zapišemo bolj splošno ime.

Primer 6:

Moja sestra je že 15 let zaposlena v Mercatorju v Spodnjem Kašlju.

☐ *Moja sestra je že 15 let zaposlena v [zakrito ime podjetja: enoti verige živilskih trgovin] v [zakrito ime kraja: v manjšem kraju v osrednjeslovenski regiji].*

Primer 7:

Na šoli v Spodnjem Kašlju je tudi otrok medicinske sestre, ki je povedal vse podrobnosti o tragičnem družinskem dogodku.

☐ *Na šoli [zakrito ime šole: v osrednjeslovenski regiji] je tudi otrok [zakrit poklic in spol: zdravstvenega delavca/delavke], ki je povedal vse podrobnosti o tragičnem družinskem dogodku.*

Posplošitev je učinkovit način zaščite podatkov, vendar mora biti dosledna in premišljena, da ne zmanjša uporabnosti podatkov za raziskovanje.

PSEVDONIMIZACIJA

Po **Uredbi o varstvu osebnih podatkov** ([GDPR](#), 4. člen, 5. odstavek):
»**Psevdonimizacija** pomeni obdelavo osebnih podatkov na tak način, da osebnih podatkov brez dodatnih informacij ni več mogoče pripisati specifičnemu posamezniku, na katerega se ti podatki nanašajo, če so te dodatne informacije shranjene ločeno in so za njih sprejeti tehnični in organizacijski ukrepi, ki zagotavljajo, da osebni podatki ne bodo pripisani določenemu ali določljivemu posamezniku.«

Po **Zakonu o varstvu osebnih podatkov** ([ZVOP-2](#), 69. člen, 7. odstavek):
»Rezultati raziskave se lahko objavijo tudi v psevdonimizirani obliki, če objava podatkov v anonimizirani obliki iz tehničnih razlogov ali zaradi ciljev raziskave ni mogoča.«

Ob **psevdonimizaciji** se »ključ« za povezovanje psevdonimiziranih podatkov in identitete posameznikov ločeno hrani. Ta postopek je običajen pri longitudinalnih raziskavah, pri katerih želimo v različnih časovnih točkah ponovno vzpostaviti stik z istimi udeleženci in udeleženkami. V takih primerih moramo sprejeti ustrezne tehnične in organizacijske ukrepe, da zaščitimo informacije, ki bi nepooblaščenim osebam lahko razkrile identiteto udeležencev in udeleženek ter povzročile morebitno škodo.

ODSTRANITEV PODATKOV

Odstranitev podatkov naj bo skrajna odločitev, saj je cilj, da podatkovno datoteko čim manj okrnimo. Podatke odstranimo oziroma izbrišemo, le če nadomestitev ali posplošitev nista primerna rešitev za zaščito osebnih podatkov.

PREGLEDNICA: IDENTIFIKATORJI IN METODE

V preglednici so navedene različne vrste identifikatorjev, kot so jih opredelili v **Finskem arhivu družboslovnih podatkov**,⁶ dodatno pa so vključene tudi vrste identifikatorjev, ki so pogostejše v slovenskem prostoru.

Z zvezdico (*) so označene informacije, ki spadajo v posebne kategorije osebnih podatkov, kot jih določa **GDPR**.

V zadnjem stolpcu so navedene priporočene metode anonimizacije. Odgovornost raziskovalcev je, da za vsak identifikator ocenijo tveganje za razkritje in izberejo najprimernejšo metodo anonimizacije.

Vrsta identifikatorja	direktni identifikator	močen posredni identifikator	posredni identifikator	metoda anonimizacije
EMŠO	x			odstranitev
Davčna številka	x			odstranitev
Ime in priimek	x			odstranitev/zamenjava (psevdonim)
E-naslov	x	x		odstranitev
Telefonska številka		x		odstranitev
Poštna številka			x	odstranitev/kategorizacija
Uporabniško ime v spletnem okolju ali drug spletni identifikator	x	x		odstranitev/zamenjava (psevdonim)

⁶ Glejte: Identifier type table: <https://www.fsd.tuni.fi/en/services/data-management-guidelines/anonymisation-and-identifiers/#what-kind-of-information-constitutes-identifiable-data>

Vrsta identifikatorja	direktni identifikator	močen posredni identifikator	posredni identifikator	metoda anonimizacije
Območje prebivanja/del mesta, občina, regija			x	kategorizacija
Drugi podatki o lokaciji			x	kategorizacija
Informacije o ozadju, na podlagi katerih je v kombinaciji mogoče identificirati osebo			x	odstranitev/kategorizacija
Osebe na zvočnem posnetku	x			odstranitev/zakritje
Osebe na videoposnetku	x			odstranitev/zakritje
Osebe na fotografiji	x			odstranitev/zakritje
Datum rojstva		x		kategorizacija
Starost			x	kategorizacija
Spol			x	kategorizacija
Zakonski stan			x	kategorizacija
Sestava gospodinjstva			x	kategorizacija
Poklic		x	x	kategorizacija

Vrsta identifikatorja	direktni identifikator	močen posredni identifikator	posredni identifikator	metoda anonimizacije
Zaposlitveni sektor			x	kategorizacija
Zaposlitveni status			x	kategorizacija
Stopnja izobrazbe			x	kategorizacija
Področje izobrazbe			x	kategorizacija
Materni jezik			x	kategorizacija
Nacionalnost				(kategorizacija)
Delovno mesto/Zaposlovalec		(x)	x	odstranitev
Registrska številka vozila		x		odstranitev
Naslov publikacije		x		odstranitev
Naslov spletne strani		(x)	x	odstranitev
ID-številka študenta, zaposlenega		x		odstranitev
Številka zdravstvenega zavarovanja		x		odstranitev
Številka bančnega računa		x		odstranitev

Vrsta identifikatorja	direktni identifikator	močen posredni identifikator	posredni identifikator	metoda anonimizacije
IP-naslov		x		odstranitev
Zdravstveni podatki*		x	x	kategorizacija/odstranitev
Genetske informacije*			x	kategorizacija/odstranitev
Biometrični podatki*			x	kategorizacija/odstranitev
Etnična pripadnost*		x	x	kategorizacija/odstranitev
Podatki o kaznivih dejanjih			x	odstranitev/kategorizacija
Članstvo v sindikatu*			x	kategorizacija
Politična ali verska pripadnost*			x	kategorizacija
Položaj v drugem združenju ali članstvo*		(x)	x	odstranitev/kategorizacija
Prejemniki socialnih transferjev*			x	kategorizacija/odstranitev
Spolna usmerjenost*			x	odstranitev
Politična, verska, filozofska prepričanja				kategorizacija

DOSTOP POD POSEBNIMI POGOJI

Če ocenimo, da bi z anonimizacijo preveč posegli v vsebino in izgubili dragocene informacije, je dodaten ukrep, da **delno anonimizirane** ali **neanonimizirane podatke** predamo v **repozitorij**, v katerem bodo dostopni **omejenim skupinam raziskovalk in raziskovalcev**, ob zagotovitvi **nadzorovanega dostopa**.

ADP zagotavlja skrbništvo nad anonimiziranimi in neanonimiziranimi podatkovnimi datotekami ter omogoča različne vrste dostopa, s katerimi nadzira uporabo občutljivih podatkov.

Dostop brez registracije:

- Odprti dostop brez registracije do **neobčutljivih** ali **anonimiziranih mikropodatkov** za vse uporabnike **ADP**.

Dostop ob registraciji:

- Dostop do **polno zaščiteneh** ali **anonimiziranih mikropodatkov** za registrirane uporabnike (največkrat v obliki t. i. **PUF** – Public Use Files).
- Dostop do **manj zaščiteneh** (delno anonimiziranih) mikropodatkov (t. i. **SUF** – Scientific Use Files) za registrirane uporabnike, ki morajo izpolniti obrazec za dostop do gradiva na zahtevo. Za namen analize lahko podatke prenesejo na svoj računalnik.
- Dostop do **nezaščiteneh občutljivih mikropodatkov** (t. i. **ScUF** – Secure Use Files) za registrirane uporabnike s statusom raziskovalca oziroma raziskovalke (s šifro ARIS), ki so oddali izpolnjeno **Vlogo za dostop do gradiva na zahtevo**. Komisija za zaščito zaupnosti vloge odobri, uporabniki pa morajo podpisati **Pogodbo o dostopu**. Ti podatki se lahko analizirajo le v varni sobi **ADP**.

Več informacij o različnih vrstah dostopa je na voljo [na spletni strani ADP](#).

ARHIVIRANJE IN OBJAVA PODATKOV V ADP

V **ADP** skladno s poslanstvom izvajamo trajno skrbništvo raziskovalnih podatkov, da bodo ti na voljo za različne namene tudi v prihodnosti. Za manj pomembne raziskave je na voljo tudi možnost samoshranitve v začasno skrbništvo.⁷

Ko raziskovalka ali raziskovalec predaja gradiva iz raziskave v ADP, ki vključuje osebne podatke, poskrbi za zaščito udeležencev in udeleženk skladno s pravnimi podlagami, ki so bile osnova za izvedbo raziskave. Po potrebi lahko skupaj ocenimo tveganja za razkritje identitete udeleženk in udeležencev raziskave, pri tem pa je ključno, da raziskovalka in raziskovalec v **ADP** predstavit:

- pravne podlage za obdelavo osebnih podatkov, npr. **informirano soglasje**;
- **vrste osebnih podatkov**, vključenih v gradivo:
- **tematiko raziskave**.

Namig

Skrbno in pravočasno načrtovanje ravnanja z raziskovalnimi podatki prepreči morebitne ovire za arhiviranje in objavo gradiva v **ADP** ob sami predaji.

Skladno z dogovorom se na tej točki po potrebi pripravi več različic mikropodatkov, ki imajo različne stopnje zaščite, odvisno od vrste dostopa do podatkov.

Za več o tej temi priporočamo v branje podstran [Predaj in objavi](#) na spletni strani ADP.

⁷ [Dva sistema skrbništva v ADP](#)

DODATNI VIRI

Raziskovalcem in raziskovalkam priporočamo, da več praktičnih primerov in napotkov za ustrezno anonimizacijo (in pripravo kvalitativnih podatkov) poiščejo še v:

- spletnem učbeniku o ravnanju z raziskovalnimi podatki Data Management Expert Guide konzorcija CESSDA, v poglavju 5. Protect, <https://dmeg.cessda.eu/Data-Management-Expert-Guide/5.-Protect/Anonymisation>
- Timescapes Archive, spletna stran Depositing Your Data. General guidelines for anonymisation, <https://timescapes-archive.leeds.ac.uk/depositing/>
- Mannheimer, S., Pienta, A., Kirilova, D., Elman, C., & Wutich, A. (2019). Qualitative Data Sharing: Data Repositories and Academic Libraries as Key Partners in Addressing Challenges. American Behavioral Scientist, 63(5), 643–664, <https://doi.org/10.1177/0002764218784991>
- UK Anonymisation Network (UKAN). Anonymisation Decision Making Framework (ADF), spletna stran z uporabnimi smernicami in orodji, <https://ukanon.net/framework/>
 - ADF 2nd Edition: The Full Book
 - ADF Tools and Templates:
 - Data Situation Evaluation Template
 - Data Features Template
 - Data Situation Evaluation Tool
 - Drugo

UPORABLJENI VIRI

CESSDA (b. d.). Data Management Expert Guide. Dostopno prek <https://www.CESSDA.eu/Training/Training-Resources/Library/Data-Management-Expert-Guide>

Finnish Social Science Data Archive. (b.d.). Anonymisation and Personal Data. Dostopno prek <http://www.fsd.uta.fi/aineistohallinta/en/anonymisation-and-identifiers.html>

Timescapes (2008). Timescapes Anonymisation Guidelines. Dostopno prek <https://timescapes-archive.leeds.ac.uk/wp-content/uploads/sites/47/2018/04/Timescapes-Anonymisation-Guidelines-18Aug08-in-use.pdf>

Finnish Social Science Data Archive, spletna stran Anonymisation and Personal Data: <https://www.fsd.tuni.fi/en/services/data-management-guidelines/anonymisation-and-identifiers/>

Priloga: Seznam pogostih identifikatorjev in primeri anonimizacije

Tip identifikatorja	Primer v izvirnem besedilu	Predlog za anonimizacijo
Ime in priimek	Matej Kovač mi je povedal ...	[zakrita ime in priimek: Rok Novak] ali [zakrita ime in priimek: prijatelj]
Naslov	Stanujem na Dunajski 48.	[zakrit naslov: ulica v Ljubljani]
Geografska lokacija	Živim v Kanalu ob Soči.	[zakrit kraj: manjše mesto na zahodu Slovenije]
Delovno mesto ali poklic	Delam kot psihiater na Onkološkem inštitutu.	[zakrit poklic in delovno mesto: zdravnik v javnem zavodu]
Ime podjetja/ustanove	Zaposlena sem v Gorenju.	[zakrito ime podjetja: večje slovensko podjetje]
Ime šole/fakultete	Obiskujem Filozofsko fakulteto v Mariboru.	[zakrito ime ustanove: fakulteta v večjem mestu]
Redke značilnosti	Imava šest otrok.	[zakrita družinska značilnost: veliko otrok]
Verska/Etnična pripadnost	Kot pripadnik bahaistov ...	[zakrita pripadnost: verska manjšina]
Prostočasne dejavnosti	Ukvarjam se z restavriranjem starin.	[zakrito: specifičen konjiček]
Ime tretje osebe	To mi je rekla moja sodelavka Mojca.	[zakrita ime in vloga tretje osebe: Marija]
Kombinacija značilnosti	Sem zdravnica v Mali vasi.	[zakrito: zdravstvena delavka v manjšem kraju]