



Seminar za raziskovalce

PRIPRAVA RAZISKOVALNIH PODATKOV ZA ODPRTI DOSTOP
Priprava, izročitev, spravilo

10. december 2014, Arhiv družboslovnih podatkov, Fakulteta za družbene vede, Univerza
v Ljubljani



PRAKTIČNI DEL: PREIZKUS ORODJA IN SEZNANJANJE Z NAVODILI

*Sonja Bezjak in Irena Vipavc Brvar, Arhiv družboslovnih
podatkov*



Kaj je NRRP?

„NAČRT ZA RAVNANJE S PODATKI“: je dokument/specifikacija o tem, kako se bo ravnalo s podatki med izvajanjem projekta in po tem, ko je projekt končan.

Namen načrtovanja ravnanja: podpreti **življenjski krog vseh podatkov**, ki bodo zbrani, ustvarjeni ali pridobljeni v projektu.

Življenjski krog podatkov

- ✓ distribucija podatkov
- ✓ kontrola dostopa
- ✓ avtorske pravice
- ✓ promocija podatkov

- ✓ načrtovanje raziskovanja
- ✓ načrt ravnanja s podatki

- ✓ interpretiranje podatkov
- ✓ produciranje rezultatov / publikacije

- ✓ hramba podatkov v priporočenih formatih
- ✓ na priporočenih medijih
- ✓ priprava varnostne kopije
- ✓ priprava metapodatkov in dokumentacije
- ✓ arhiviranje podatkov

- ✓ zbiranje
- ✓ zajem in priprava metapodatkov

- ✓ vnos podatkov, digitalizacija, transkript
- ✓ preverjanje veljavnosti, čiščenje podatkov
- ✓ anonimizacija
- ✓ opis podatkov



Področja vključena v Pilot (2014-2015)

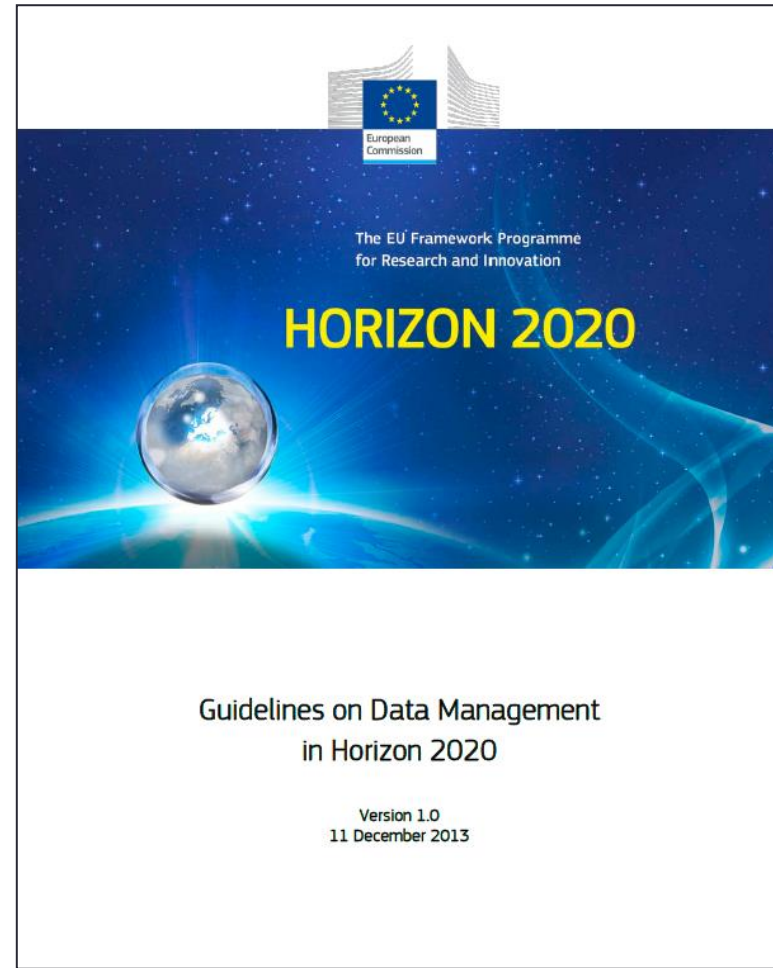
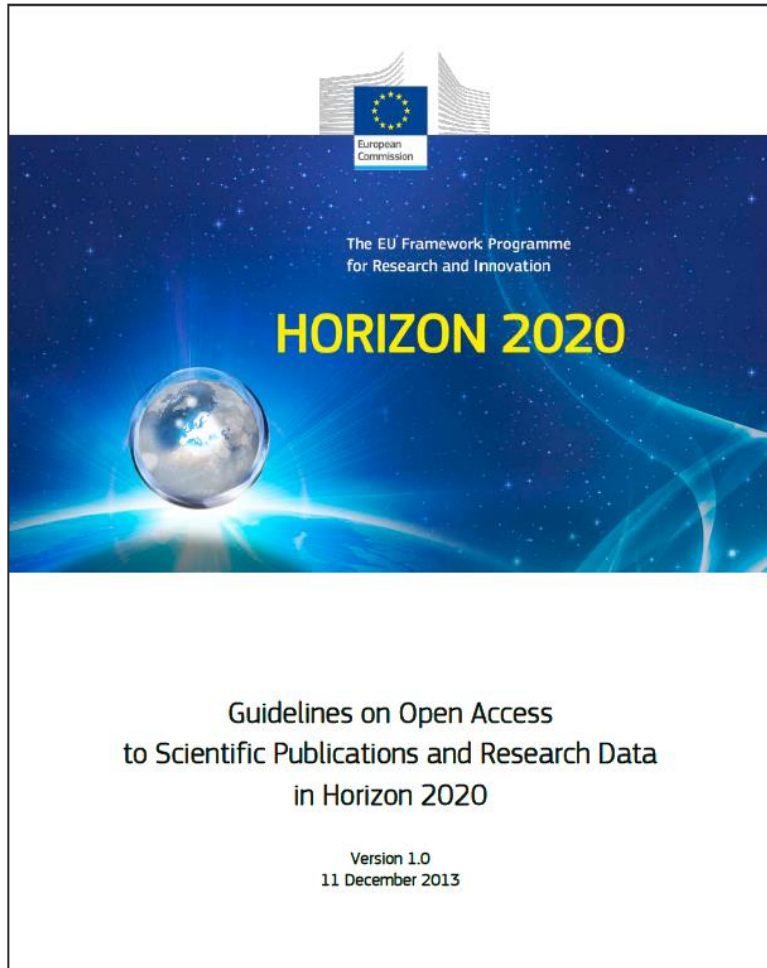
- Future and Emerging Technologies
- Research infrastructures - part e-Infrastructures
- Leadership in enabling and industrial technologies - Information and Communication Technologies
- Societal Challenge: Secure, Clean and Efficient Energy - part Smart cities and communities
- Societal Challenge: Climate Action, Environment, Resource Efficiency and Raw materials - with the exception of topics in the area of raw materials
- Societal Challenge: Europe in a changing world - inclusive, innovative and reflective Societies
- Science with and for Society

Za projekte z drugih področij **možna prostovoljna udeležba**:

Prostovoljna udeležba pomeni, da **bodo v pogodbi sprejeli še to obvezo**, da bo financer spremljal tudi to področje in da bodo s tem v zvezi **deležni pomoči**, kot ostali obvezno vključeni v pilot.

([Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020](#), 2013 str. 10)

Viri



Obzorje 2020: kdaj in kako začeti?

V **prijavi projekta** opredeliti, kako boste ravnali s podatki:

- Kakšen tip podatkov bo nastal v okviru projekta?
- Kakšni standardi bodo uporabljeni?
- Kako bodo nastali podatki dostopni za preverjanje in ponovno rabo? In če ne, zakaj ne?
- Kako bodo podatki hranjeni in vzdrževani?

(Guidelines on Data Management, 2013, str. 2-3)

Obzorje 2020: kako naprej?

Prvo verzijo NRRP **pripraviti v 6 mesecih** po začetku projekta.

NRRP naj vključuje:

1. oznako in naslov podatkov;
2. opis podatkov: izvor, tip, uporabnost, zanimivost, podobnost z drugimi;
3. standardi in metapodatki: uporaba področnih, utemeljitev, če drugače;
4. deljenje podatkov: pogoji dostopa in izjeme, obdobje embarga, odprti dostop ali omejen, **izbrano mesto dostopa** (npr. podatkovni arhiv);
5. arhiviranje in dolgotrajno digitalno ohranjanje: predviden čas, obseg, stroški.

(Annex 1: Data Management Plan (DMP) template)

Priporočila za kasnejše verzije NRRP

Med projektom je NRRP **mogoče po potrebi popravljati** in ga oddati v kasnejših fazah.

Obzorje 2020:

Kasnejša verzija NRRP naj odgovori na naslednja vprašanja:

1. Ali so podatki ustvarjeni tako, da jih je mogoče odkriti/najti? (npr. DOI)
2. Ali so podatki dostopni in na kakšne načine? (npr. licence, embargo, komercialna raba)
3. Ali so podatki pripravljene tako, da jih lahko tretji vrednotijo in razumejo? (npr. recenzenti)
4. Ali so podatki tudi dolgo po nastanku uporabni za tretje? (npr. so varno spravljene v pooblaščenem repozitoriju)
5. Ali so povezljivi z drugimi: omogočena izmenjava med raziskovalci, ustanovami, organizacijami, državami?

(Annex 2: Additional guidance for Data Management Plans)

Za prevzem gradiva v ADP potrebujemo

- 1) Osnovne podatke o raziskavi: naslov raziskave, leto izvedbe, avtorji, serija raziskav
- 2) Vsebinska področja: kratek povzetek raziskave in poglavja
- 3) Osnovne podatke o metodologiji: enota za analizo, populacija, vrsta podatkov, način zbiranja podatkov, vzorčenje, stopnja odgovora, geografska pokritost

<http://www.adp.fdv.uni-lj.si/evidentiranje/>

- 4) Izjavo o izročitvi: zagotovilo, zaupnosti, avtorske pravice, pogoji dostopa (dostopna: http://www.adp.fdv.uni-lj.si/za_dajalce/izjava_o_izrocitvi/)

Vnos nove raziskave

Polja označena z * so obvezna polja. Prosimo, da jih izpolnite. Izpolniti morate tudi vsaj eno izmed polj nosilec raziskave ali ustanova.

Izvorni naslov raziskave*:	<input type="text"/>
Izvorni naslov raziskave v angleškem jeziku:	<input type="text"/>
Nosilec raziskave:	<input type="text"/>
Izdelala ustanova:	<input type="text"/>
Leto izdelave raziskave:	Izberi <input type="text"/>
Kraj izdelave raziskave:	<input type="text"/>
Podatki v ADP ?:	<input checked="" type="radio"/> Ni podatka <input type="radio"/> Podatki so že v ADP <input type="radio"/> Podatke bi verjetno lahko izročili v ADP <input type="radio"/> Možnosti za izročitev podatkov niso gotove
Izvorno mesto dostopa do podatkov (v kolikor različno od ustanove izdelave):	<input type="text"/>
URL naslov izvornega opisa raziskave/mesta dostopa do podatkov:	<input type="text"/>
Naziv serije:	<input type="text"/>
Vsebinsko področje (1):	IZOBRAŽEVANJE <input type="text"/>
Vsebinsko podpodročje (1):	Izberi podkategorijo <input type="text"/>
Vsebinsko področje (2):	Izberi dodatno poglavje <input type="text"/>
Vsebinsko podpodročje (2):	Izberi podkategorijo <input type="text"/>
Povzetek (do 255 znakov):	<input type="text"/>
Država:	<input type="text"/>
Populacija:	<input type="text"/>

Kriteriji za sprejem v ADP

Dolgotrajno shranjevanje podatkov zahteva **dodatne napore in stroške** za pripravo v obliki, ki omogoča njihovo nadaljnjo uporabo. Ti so **upravičeni ob prihranku**, ki ga predstavlja večkratna uporaba podatkov.

- **vsebinsko bogastvo** podatkov v smislu ustreznosti konceptualizacije in tematskega dopolnjevanja zbirke ADP,
- izpopolnjenost uporabljenih metod, celovitost in ustreznost **podatkov in dokumentacije** za nadaljnje analize,
- **dajalec avtorsko razpolaga s podatki** in je pripravljen podatke izročiti arhivu za razširjanje.

(http://www.adp.fdv.uni-lj.si/za_uporabnike/izrocanje_podatkov/)

Izbor na podlagi kakovosti

Prednostno se v arhiv uvrstijo:

(1) najvišji rang, primerjalne ali kontinuirane raziskave, pomembne populacije, metodološko ustrezne; (2) teoretsko ali praktično pomembne raziskave, ki **zapolnjujejo raziskovalno praznino ali imajo široko uporabnost** za več praktičnih problemov, imajo dolgoročno vrednost za znanost; (3) **teoretsko ali praktično pomembne raziskave**, omejene na manj splošno populacijo, manj vplivne.

Prevzeti podatki štejejo kot [znanstvena objava po merilih ARRS.](#)

3. Samostojni znanstveni sestavek ali poglavje v monografski publikaciji (Cobiss tip 1.16 v 2.01, 2.02, 2.06, 2.16, 2.17, 2.20)

- | | |
|---|---------|
| A. Obsežni znanstveni sestavek ali poglavje (nad 50 strani) v obsežni znanstveni monografiji (2.01A nad 500 strani) | 60 točk |
| B. Samostojni znanstveni sestavek ali poglavje v znanstveni monografiji (2.01A) | 30 točk |
| C. Samostojni znanstveni sestavek ali poglavje v znanstveni monografiji (2.01B) | |
| ○ izdana pri tuji založbi | 25 točk |
| ○ izdana pri domači založbi | 20 točk |
| D. Samostojni znanstveni sestavek ali poglavje v drugi monografski publikaciji | 10 točk |
| E. Zaključena znanstvena zbirka podatkov (2.20) iz seznama agencije | 30 točk |

Samostojni znanstveni sestavek ali poglavje je praviloma objavljen v znanstveni monografiji (2.01), izjemoma pa tudi v strokovni monografiji (2.02), slovarju ali leksikonu (2.06), kot spremna študija k umetniškemu delu (2.16) ali kot razprava v katalogu razstave (2.17). Obsegati mora najmanj 4 strani (izjema so znanstveni zemljevidi).

Skupno število točk enega avtorja v eni monografski publikaciji ne sme preseči 100 točk.



Vnos nove raziskave

Preko obrazca podam prvo informacijo o svoji raziskavi.

Polja označena z * so obvezna polja. Prosimo, da jih izpolnite. Izpolniti morate tudi vsaj eno izmed polj nosilec raziskave ali ustanova.

Izvorni naslov raziskave*:

 ?

Izvorni naslov raziskave v angleškem jeziku:

 ?

Nosilec raziskave:

 ?

Izdelala ustanova:

 ?

Leto izdelave raziskave:

 ?

Kraj izdelave raziskave:

 ?

Podatki v ADP ?:

- Ni podatka
- Podatki so že v ADP
- Podatke bi verjetno lahko izročili v ADP
- Možnosti za izročitev podatkov niso gotove

Izvorno mesto dostopa do podatkov (v kolikor različno od ustanove izdelave):

 ?

URL naslov izvornega opisa raziskave/mesta dostopa do podatkov:

 ?

Naziv serije:

 ?

Vsebinsko področje (1):

Vsebinsko podpodročje (1):

Vsebinsko področje (2):

Vsebinsko podpodročje (2):

Povzetek (do 255 znakov):

Država:

 ?

Populacija:

 ?

Vrsta podatkov:

 ?

Časovna opredelitev ?:

- Ni podatka
- Enkratna raziskava
- Ponavljajoča se raziskava
- Panelne in kohortne raziskave
- Časovne serije
- Drugo

Tip vzorca ?:

- Ni podatka
- Brez vzorčenja - zajeta celotna populacija
- Slučajno (verjetnostno) vzorčenje
- Ne-slučajno vzorčenje

Način zbiranja podatkov:

 ?

Omejitve pri uporabi ?:

- Ni podatka
- Brez omejitev
- Omejeno za akademske namene
- Omejeno za potrebe naročnika ali raziskovalne skupine
- Informacije niso razpoložljive

Ostalo gradivo o raziskavi:

 ?

Število enot ?:

- Ni podatka
- manj kot 100

Sledi še informacija o
Številu spremenljivk in
Ocena urejenosti
podatkov

Urejevalnik



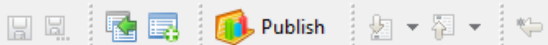
- Odprto kodni program.
- V pomoč raziskovalcem pri pripravi NRRP za prijavo projekta.
- V pomoč pri pripravi in nadgrajevanju NRRP v procesu samega izvajanja raziskovanja.

- Načrt se lahko izvozi v DDI XML in objavi v PDF, HTML, RTF in drugo.

- Poročila se lahko uporabljajo za ravnanje podatki znotraj raziskovalne skupine (organizacije), kakor tudi za predajo v ADP in za potrebe poročanja financerju.

- Trenutne podlage - National Science Foundation (NSF), National Institutes for Health (NIH) iz ZDA in the Digital Curation Center iz VB.

http://www.dcc.ac.uk/sites/default/files/documents/resource/DMP_Checklist_2013.pdf



- ▶ Data Management Plans
 - Moja raziskava
- ▶ Data Management Definitions
 - DCC_Full
 - DCC_Simple
 - NIH General DMP
 - NSF-AGS DMP
 - NSF-AST DMP
 - NSF-BIO DMP
 - NSF-CHE DMP
 - NSF-CISE DMP
 - NSF-DMR DMP
 - NSF-EAR DMP
 - NSF-EFRI DMP
 - NSF-EHR DMP
 - NSF-ENG DMP
 - NSF General DMP
 - NSF-PHY DMP
 - NSF-SBE DMP

Data Documentaiton and Metadata

Are the datasets which you will be capturing/creating self-explanatory, or understandable in isolation?

If you answered No to DCC 2.5.1, what contextual details are needed to make the data you capture or collect meaningful?

How will you create or capture these metadata?

What form will the metadata take?

Why have you chosen particular standards and approaches for metadata and contextual documentation?

Ethics and Intellectual Property

Ethical and Privacy Issues

Are there ethical and privacy issues that may prohibit sharing some or all of the dataset(s)?

If you answered Yes to DCC 3.1.1, how will these be resolved?

Is the data that you will be capturing/creating "personal data" in terms of the Data Protection Act (1998) or equivalent legislation if outside the UK?

What action will you take to comply with your obligation under the Data Protection Act (1998) or equivalent legislation outside the UK?

Intellectual Property Rights

Will the dataset(s) be covered by copyright or the Database Right? If so give details in DCC 3.2.2 below

Is you answered Yes to DCC 3.2.1, who owns the copyright and other intellectual property?

If you answered Yes to DCC 3.2.1, How will the dataset be licensed?

For multi-partner projects, what is the dispute resolution process/mechanism for mediation?

Access, Data Sharing and Reuse

Access and Data Sharing

Are you under obligation or do you have plans to share all or part of the data you create/capture?

If you answered No to DCC 4.1.1, why will you not share your data?

An outline is not available.

Bean Information

No identifiable object is advertised.

Data Management Plans

*Moja raziskava

Data Management Definitions

DCC_Full

DCC_Simple

NIH General DMP

NSF-AGS DMP

NSF-AST DMP

NSF-BIO DMP

NSF-CHE DMP

NSF-CISE DMP

NSF-DMR DMP

NSF-EAR DMP

NSF-EFRI DMP

NSF-EHR DMP

NSF-ENG DMP

NSF General DMP

NSF-PHY DMP

NSF-SBE DMP

Basic Plan Information

Plan Name: Moja raziskava

Definition Name: DCC_Full

Plan Documentation

Normal Arial 2





Tukaj |

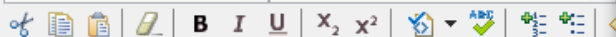
Rich Text HTML

Introduction and Context (1/13)

Basic Project Information (0/6)

Short Description of the project's fundamental aims and purpose

Normal Arial





Zapišemo kratek opis raziskovalnega projekta

Rich Text

Related Policies (0/3)

Funding body requirements relating to the creation of a data management plan

Introduction and Context (1/13)

Basic Project Information (0/6)

Short Description of the project's fundamental aims and purpose

Related Policies (0/3)

Funding body requirements

Institutional or Research

Other policy-related dependencies

Basic Data Management Plan

Data Types, Formats, Standards

Give a short description of the data

Existing Data (0/3)

New Data (1/4)

Why do you need to capture the data?

Describe the process by which you will capture the data

Which file formats will you use?

What criteria will you use to select data?

Relationship between old and new data

Data Documentaiton and Metadata

Are the datasets which you are capturing unique?

If you answered No to the question above, how will you ensure that the data is unique?

How will you create or capture the data?

What form will the metadata take?

Why have you chosen particular metadata standards?

Ethics and Intellectual Property

Information summarised from the main body of your research proposal will help potential re-users understand the purposes your data has been collected or created for, and they are unlikely to have access to your proposal. Briefly summarise what you set out to discover and how that is likely to affect the kind of data you collect or create and how.

Dodatne povezave v pomoč

- [Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020](#)
- [Guidelines on Data Management in Horizon 2020](#)
- [Obrazec Evidentiraj raziskavo pri ADP](#)
- [Obrazec Izjava o izročitvi v ADP](#)

- [DCC Data Management Plans](#)
- [UK Data Archive Data Management Checklist](#)
- [University of Leeds Data Management Planning](#)