



Institut za nuklearne
nauke Vinča

Šta je upravljanje istraživačkim podacima? Zašto je to uopšte važno?

Obrad Vučkovic

Institut za nuklearne nauke "Vinča" - Biblioteka
Univerzitet u Beogradu



Šta je upravljanje istraživačkim podacima?

Upravljanje istraživačkim podacima se odnosi na niz aktivnosti kako bi se na najefikasniji način prikupili i organizovali podaci, i pripremili za buduću upotrebu i deljenje sa drugima.



ZAŠTO UPRAVLJATI PODACIMA?

01

SMANJITI PONAVLJANJE

02

ZAHTEVI FINANSIJERA

03

SPREČITI GUBITAK PODATAKA

04

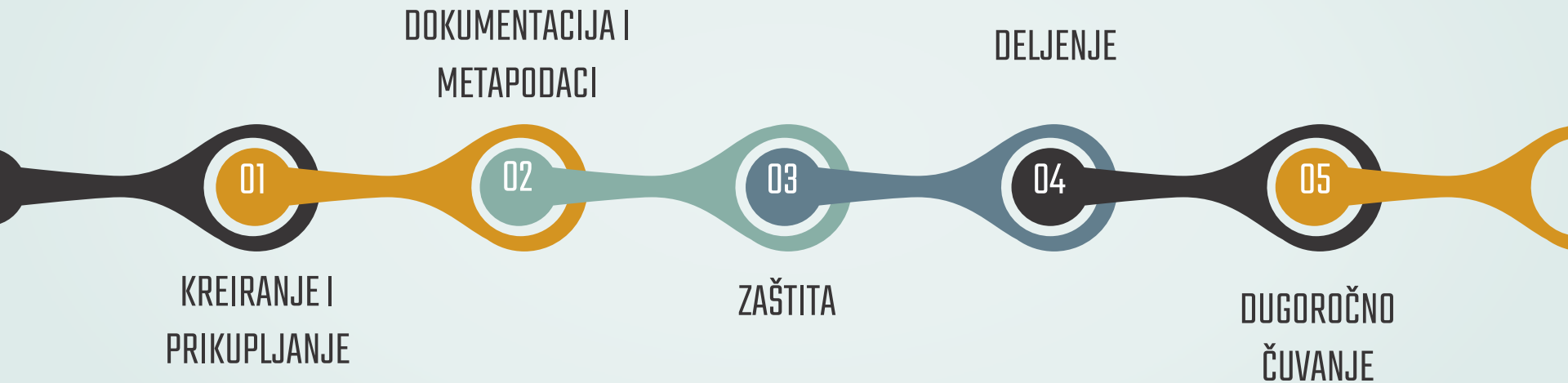
TRANSPARENTNOST

05

BOLJA SARADNJA



RESEARCH LIFECYCLE



Ali, prvo...

F T 1 P

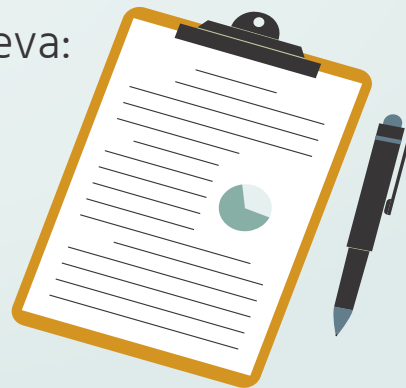
(fali ti jedan papir)

DATA MANAGEMENT PLAN

Dokument koji služi kao okvir svim učesnicima na projektu kako da upravljaju podacima i na koji način da ih čuvaju.

Donosi se na početku projekta i svi učesnici bi trebali da su upoznati sa njim.

Odlični alati za pomoć pri pisanju DMP-eva:
[DMPOnline](#) & [DMPTool](#)





KREIRANJE I PRIKUPLJANJE

Kreiranje i prikupljanje podataka



Odabrati odgovarajući format

- otvoreni (open) formati i *loseless* formati održivi i bolji za dugoročno čuvanje
na primer: CSV, PDF/A, TIFF, WAV, XML
- vlasnički (proprietary) i kompresovani formati često u širokoj upotrebi, izbegavati kao izbor za dugoročno čuvanje
na primer: DOCX, XLSX, JPG, MP3

Možete koristiti vlasnički format za analiziranje, ali odaberite otvoreni format za dugoročno čuvanje.

Kreiranje i prikupljanje podataka

Organizovanje po folderima

- Imenovati folder po oblasti rada, a ne po istraživaču
- Hijerarhijska struktura
- Odvajati radne od urađenih (gotovih) foldera

Imenovanje fajlova

- Koristiti iste termine
- Odvajanje reči: velika slova, crtice (_ ili -); izbegavati dijakritičke znake i prazninu
- Dogovoriti format datuma – dobra praksa: ISO 8601 format YYYY-MM-DD (20200708)
- Red reči: fajlovi sa istom temom su grupisani i lako se nalaze

Verzije

Svaka bitnija promena na fajlu treba da dobije novu verziju (npr. v2 ili v3_2)

Poslednji fajl dobija verziju „**final**“.

DOKUMENTACIJA I METAPODACI

02

An abstract graphic design on the right side of the page. It features several organic, teardrop-like shapes in shades of orange, olive green, and dark grey. A prominent orange shape in the foreground contains a white circle with the number '02' inside. Other shapes are scattered around, including a dark grey circle with a teal center, a solid black circle, and a white teardrop shape. The background is a light, pale blue.

Dokumentacija i metapodaci

```
1 /*
2  * This line basically imports the "stdio" header file, part of
3  * the standard library. It provides input and output functionality
4  * to the program.
5  */
6 #include <stdio.h>
7
8 /*
9  * Function (method) declaration. This outputs "Hello, world\n" to
10 * standard output when invoked.
11 */
12 void sayHello(void) {
13     // printf() in C outputs the specified text (with optional
14     // formatting options) when invoked.
15     printf("Hello, world!\n");
16 }
17
18 /*
19 * This is a "main function". The compiled program will run the code
20 * defined
21 */
22 int main(void)
23 {
24     // Inv
25     sayHel
26     return
27 }
```



Image by TeroVesalainen from Pixabay

Dokumentacija

sva prpratna građa koja omogućava pravilnu upotrebu podataka

- upitnici i pitanja za intervju,
- laboratorijske sveske,
- programski kod,
- informacije o podešavanjima i kalibraciji uređaja,
- *codebooks*,
- dokumentacija na nivou fajla (npr. kolone u spreadsheet fajlovima)

Dokumentacija i metapodaci

Metapodaci

- Standardizovani
- Strukturisani
- Čitljivi za ljude i mašine

Najpoznatiji standardi za metapodatke za biološke nauke:

- Darwin Core – namenjen za opisivanje biološkog diverziteta
- EML (Ecological Metadata Language) – specijalno razvijeni standard za nauku o ekologiji
- OME-XML (Open Microscopy Environment XML) – standard za metapodatke i format za opisivanje bioloških mikroskopskih podataka
- Pojedini repozitorijumi razvili su sopstvene standarde kako bi zadovoljili potrebe svojih korisnika

ZAŠTITA PODATAKA

03

The graphic features several organic, teardrop-shaped elements in shades of orange, olive green, and teal. These shapes are interconnected by thin lines, creating a network-like structure. Scattered throughout the light blue background are small, solid-colored dots in teal, black, and white. The overall aesthetic is clean, modern, and tech-oriented.

ZAŠTITA PODATAKA I
REZERVNE KOPIJE (BACKUP)



Image by John from Wikimedia Commons

Zaštita podataka

Rezervne kopije (*backup*)

- Laptop, fleš drajv (mogu da se pokvare ili budu ukradeni/izgubljeni)
- Mrežni server
- *Cloud storage* (Dropbox, Google Drive, etc.)

Pravilo 3-2-1 - 3 kopije na 2 različita medijuma i 1 *offsite* (server ili cloud)

Backup i dugoročno čuvanje nije isto!!!



Zaštita podataka

Kontrola pristupa

Lozinke

Osetljivi podaci, podaci o ličnosti, poslovne tajne, ...

LastPass – brauzer dodatak (*add-on*) za čuvanje lozinki

Enkripcija

Proces promene podataka u nečitljivi kod.

- BitLocker (Windows), File Vault (Mac) – enkripcija hard-diska
- Encrypto, AES Encrypt (enkripcija na nivou fajlova i foldera)



Image by Pete Linforth from Pixabay



04

DELJENJE PODATAKA

Deljenje podatoka

Otvorena nauka => Otvoreni podaci













Potrebno je obratiti pažnju na:


- zaštitu podatoka o ličnosti
- intelektualnu svojino, zaštitu konkurentnosti, ...


“as open as possible, as closed as necessary” (*Horizon 2020*)


Finansijeri često dozvoljavaju embargo period za objavljivanje podatoka.


Deljenje podataka


CREATIVE COMMONS LICENSES		 COPY & PUBLISH	 ATTRIBUTION REQUIRED	 COMMERCIAL USE	 MODIFY & ADAPT	 CHANGE LICENSE
	PUBLIC DOMAIN	✓	✗	✓	✓	✓
	CC BY	✓	✓	✓	✓	✓
	CC BY-SA	✓	✓	✓	✓	✗
	CC BY-ND	✓	✓	✓	✗	✓
	CC BY-NC	✓	✓	✗	✓	✓
	CC BY-NC-SA	✓	✓	✗	✓	✗
	CC BY-NC-ND	✓	✓	✗	✗	✓

 You can redistribute (copy, publish, display, communicate, etc.)

 You have to attribute the original work

 You can use the work commercially

 You can modify and adapt the original work

 You can choose license type for your adaptations of the work.

Modified image from foter.com (CC BY-SA)



05

DUGOROČNO ČUVANJE

Dugoročno čuvanje podataka

Prilikom procene i izbora podataka za dugoročno čuvanje obratiti pažnju na:

1. značaj za cilj istraživanja
2. naučna ili istorijska vrednost
3. jedinstvenost
4. neponovljivost
5. ekonomska računica
6. puna dokumentacija

Dugoročno čuvanje podataka

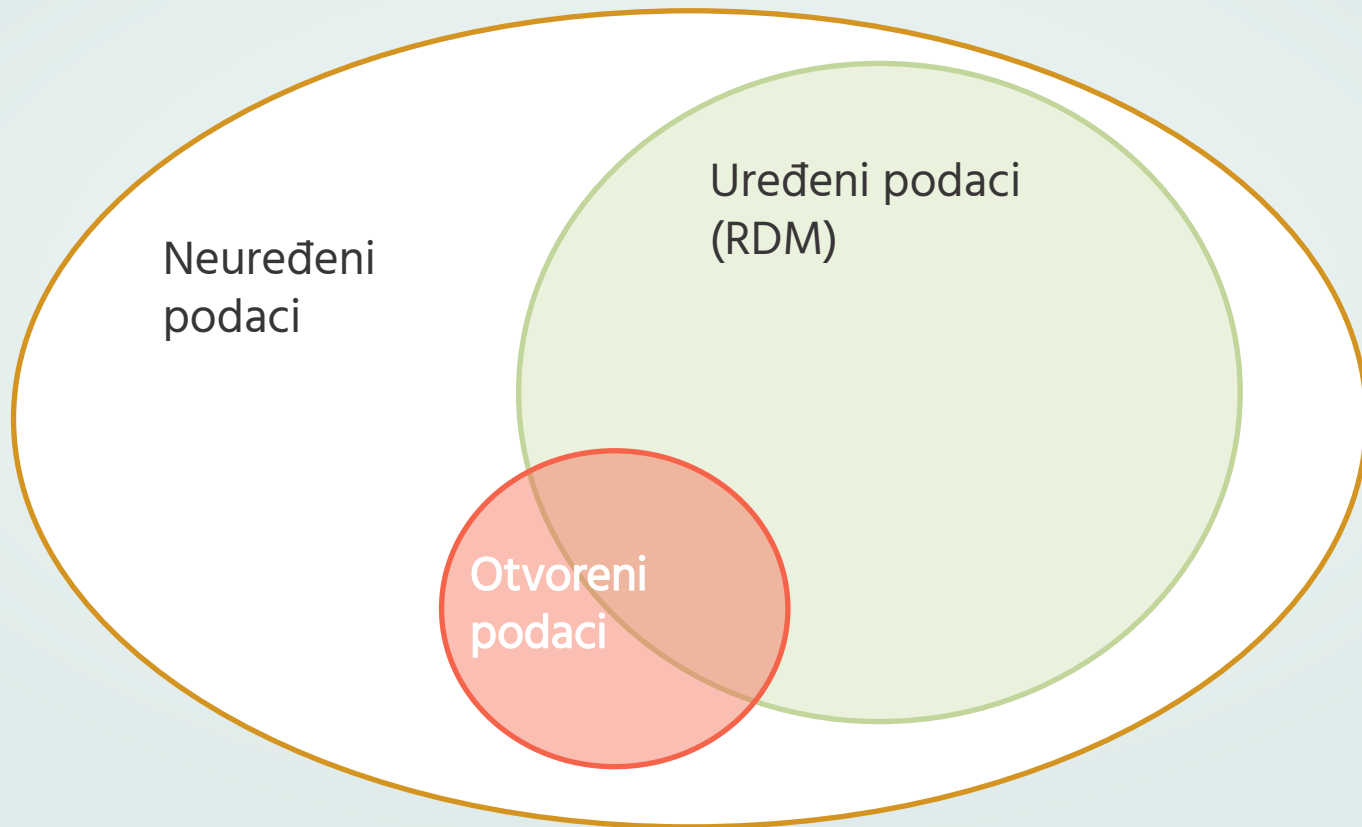
Kriterijumi za izbor repozitorijuma:

- Domain specific > general repository
- Koje formate prihvataju,
Open/Restricted Access
- Perzistentni i jedinstveni identifikatori
- Sertifikati (DSA – Data Seal of Approval, ISO 16363)

re3data.org
REGISTRY OF RESEARCH DATA REPOSITORIES

Image by re3data.org from Wikimedia Commons

Otvoreni i uređeni podaci



FAIR Data Principles



FINDABLE

ACCESSABLE



INTEROPERABLE

REUSABLE



FAIR Data Principles



Findable (Pristupačnost)

- (Meta)podaci imaju jedinstven i perzistentni identifikator
- Podaci se opisuju detaljnim metapodacima
- Metapodaci jasno uključuju identifikator podataka koje opisuju
- (Meta)podaci se indeksiraju u pretraživom repozitorijumu

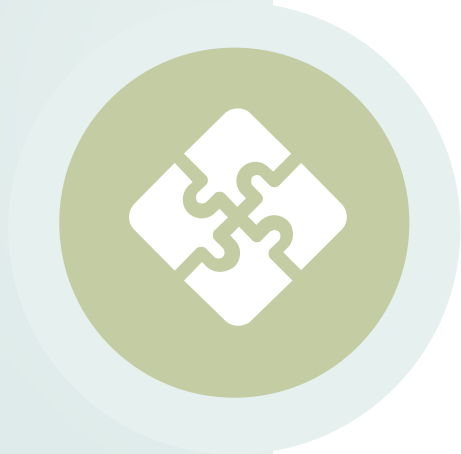
FAIR Data Principles

Accessible (Dostupnost)

- (Meta)podaci se mogu preuzimati putem standardizovanih komunikacijskih protokola, koji su otvoreni, besplatni i omogućavaju autorizaciju
- Metapodaci su dostupni, čak i kada podaci nisu



FAIR Data Principles



Interoperable (Interoperabilnost)

- Metapodaci koriste formalan, pristupačan, deljen i široko primenjiv jezik za predstavljanje informacija
- Metapodaci koriste rečnik koji prati FAIR principe
- Metapodaci obuhvataju kvalifikovane reference na druge metapodatke

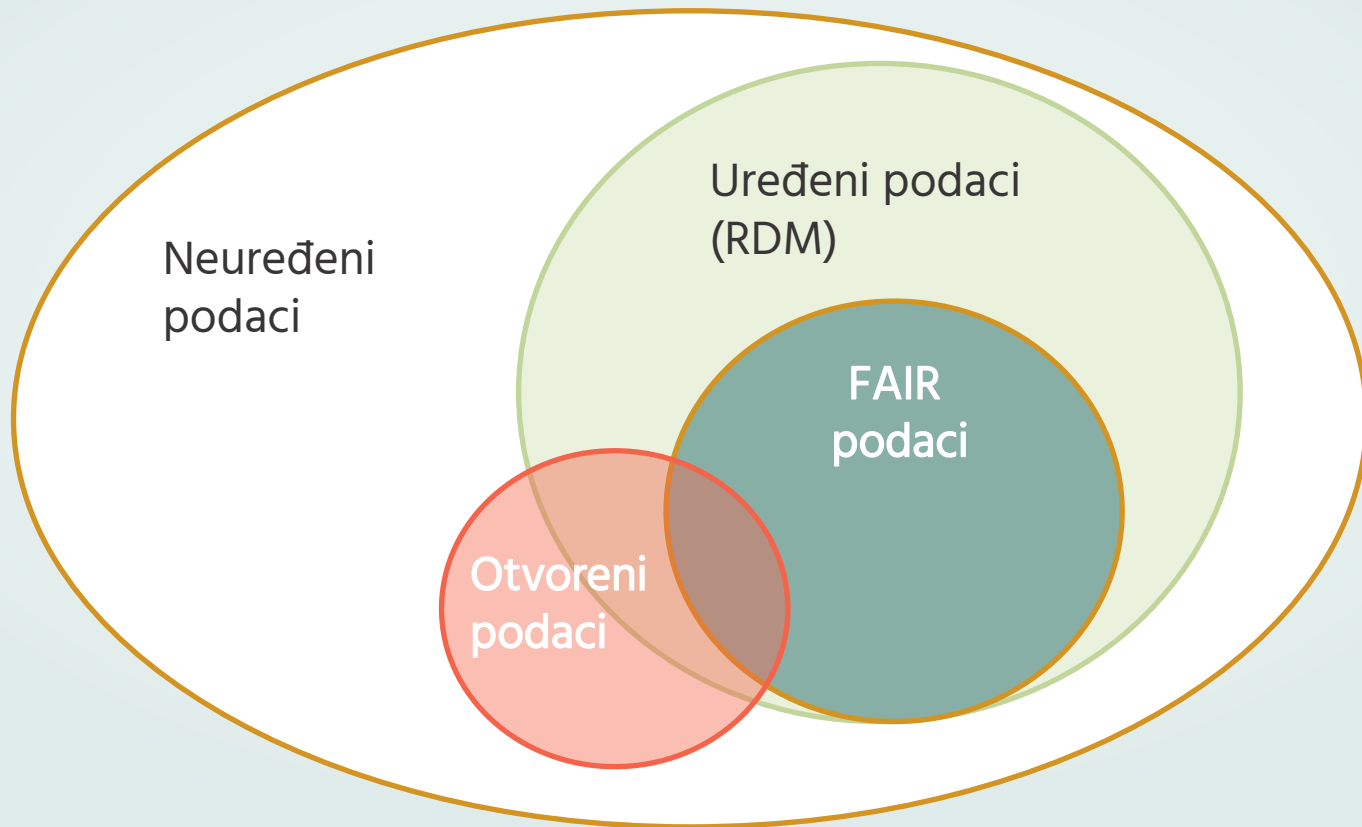
FAIR Data Principles

Reusable (Ponovna upotrebljivost)

- Metapodaci su detaljno opisani tačnim i relevantnim atributima, kao što su:
 - jasna i pristupačna licenca za korišćenje
 - detaljne informacije o izvorima iz kojih potiču (*provenance*)
 - ispunjavaju relevantne standarde u određenoj naučnoj oblasti




Otvoreni i uređeni podaci





Dodatni izvori

- DMPOnline - <https://dmponline.dcc.ac.uk/>
 - DMPTool - <https://dmptool.org/>
 - Michener WK (2015) Ten Simple Rules for Creating a Good Data Management Plan. PLoS Comput Biol 11(10): e1004525. <https://doi.org/10.1371/journal.pcbi.1004525>
 - Whyte, A. & Wilson, A. (2010). "How to Appraise and Select Research Data for Curation". DCC How-to Guides. Edinburgh: Digital Curation Centre. Available online: <http://www.dcc.ac.uk/resources/how-guides>
 - MANTRA – Research Data Management Training. 2019. University of Edinburgh. <https://mantra.edina.ac.uk/>
 - Research Data Management and Sharing (offered by the University of North Carolina at Chapel Hill & the University of Edinburgh at Coursera). <https://www.coursera.org/learn/data-management?>
 - OpenAIRE Guides for researchers – How to select data repository? <https://www.openaire.eu/opendatapilot-repository-guide>
 - Vodič *Upravljanje istraživačkim podacima*. <http://rdm.open.ac.rs/>
- 

HVALA

Imate li pitanja?

obrad.vuckovac@vin.bg.ac.rs

Institut za nuklearne nauke “Vinča”
Univerzitet u Beogradu



CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, and infographics & images by **Freepik**

Please keep this slide for attribution

