

# Upravljanje istraživačkim podacima

Obrad Vučkovic

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



**Upravljanje  
istraživačkim  
podacima**



**Research Data  
Lifecycle  
&  
Data Management  
Plan**



**FAIR principi**



Upravljanje istraživačkim podacima se odnosi na niz aktivnosti koje istraživači međusobno dogovaraju i usvajaju kako bi na najefikasniji način prikupili i organizovali svoje skupove podataka, kao i da bi ih sačuvali za buduću upotrebu koja može uključiti i deljenje sa drugim istraživačima.

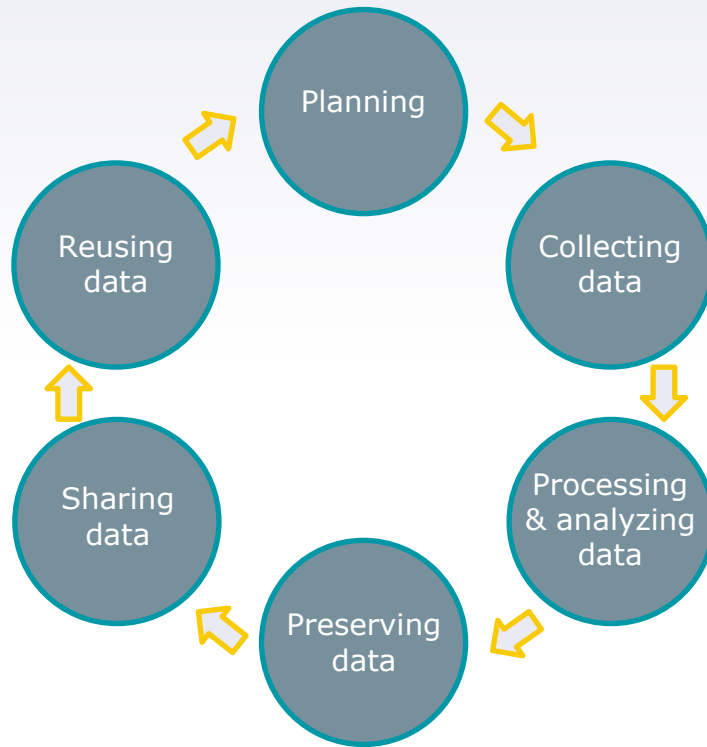
*[rdm.open.ac.rs](http://rdm.open.ac.rs)*

# Zašto RDM?

Kriza replikacije u nauci (*Replication crisis*)

Neke prednosti RDM-a:

- ▶ bolja naučna praksa
  - ▷ transparentnost naučnog istraživanja;
  - ▷ (trajno) čuvanje;
  - ▷ veća citiranost;
  - ▷ mogućnost saradnje
- ▶ ekonomski razlozi
  - ▷ manje ponavljanja eksperimenata (finansijski i etički razlozi)
- ▶ korišćenje podataka u edukativne svrhe



## Research Data Lifecycle

# Data Management Plan

Plan upravljanja podacima

(eng. Data Management Plan, DMP)

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.

DMP se može revidirati i uređivati u toku projekta.

Institucije i finansijeri (uglavnom) omogućavaju obrasce za izradu DMP-ova.

# Alati za DMP: Argos, DMPOnline, DMP tool, DSW

### 3.2 How will data security and protection of sensitive data be taken care of during the research?

- Not applicable (no sensitive data)
- Default security measures of the institution networked research storage
- Additional security measures (please specify)

**Additional Information**

**B** *I*

There will be no sensitive data

**Save**

Answered just now by [obrad.vuckovac@gmail.com](mailto:obrad.vuckovac@gmail.com)

Guidance	Comments
<b>NWO</b>	Consider data protection, particularly if your data is sensitive – for example, containing personal data, politically sensitive information or information relating to religion and health, trade secrets or national security information. Describe the main risks and how these will be managed. Inquire with your institution's research support staff whether your intended storage solution meets your institution's data security policy if your research involves sensitive data. If you are using offline storage, describe how data will be recovered in the

# Alati za DMP: Argos, DMPOnline, DMP tool, DSW

Horizon 2020 Dataset Description

Template: Horizon 2020

External References

- Data Repositories
  - Zenodo, GitHub
- External Datasets
- Registries
- Services

Dataset Description

1 Data Summary

1.1 What is the purpose of the data collection/generation and its relation to the objectives of the project?

The data is collected to validate the novel approach and demonstrate the new coating technology. The construction details, protocols, process parameters, and analytical measurements on the produced coatings all aim to fulfil the three objectives. The raw data on plasma-treated and plasma-coated wood substrates might further be helpful for readers of our scientific articles, that are to be published, allowing them to verify our findings. Thus, publishing all data allows to achieve a more complete transparency and reproducibility. Furthermore, the data may help to form a more complete view on the effects of different plasma treatments on wood surfaces, and thus might enable to generate a general model covering all different plasma treatments. The three main objectives of the action are: (I) Building an integrated device, (II) optimizing the parameters of PMMA deposition for exterior use, thereby further improving the understanding of the processes, and (III) demonstrating the technique's capability and priming the industrial implementation. The created data will therefore include: (I) construction details and computer-aided design (CAD) assisted drawings, (II) coating deposition protocols, plasma diagnostic data, and data for the characterisation of the deposited coatings, as well as (III) aging, weathering and adhesion tests of the coatings, amongst other measurements, that indicate the industrial usability. However, variables and types of the data required to fulfil these three objectives are too complex to be stated in one paragraph. These will be explored in more detail later within this

2 FAIR Data

2.1 Making data findable, including provisions for metadata

2.1.1 Will you use metadata to describe the data?

No

*Comment: All data will be uploaded together with the relating metadata, including project context and labbook entries.*

2.1.3 Will your metadata use standardised vocabularies?

No

*Comment: No field-specific recommendations are known to the PI at the point of the DMP creation. As metadata, we will thus provide: Publication date, Title, Authors including contact information, Description, Version, Language, Keywords, Grant acknowledgement, and References to all publications referring to the dataset. Further, we will include complete lab notebook excerpts, as well as protocols for measurements and analysis within the dataset. This is fully in line with the repository's policy. Further, the repository stores all metadata in JSON-format according to a defined JSON schema. Metadata is exported in several standard formats such as MARCXML, Dublin Core, and DataCite Metadata Schema in accordance with the OpenAIRE Guidelines.*

2.1.7 Will you use naming conventions for your data?

Yes

*Comment: We ascertain that the data will be easily recognized and correlated to experiments via the following naming conventions: Raw data: YYMMDD\_[experiment]\_[technique]\_XXX.\* Processed results: YYMMDD\_[experiment]\_[technique]\_XXX\_analysis\_ZZZ.\* Herein, symbols represent the following: YYMMDD - the inverted date of the day the experiment was conducted [experiment] - a short title for the experimental series [technique] - a unique denominator for each technique, such as CLSM, SEM XXX - a running number for individual measurements ZZZ - a running number for individual data points*



## Alati za DMP: Argos, DMPOnline, DMP tool, DSW

```
{
  "dmp" : {
    "contact" : {
      "contact_id" : {
        "identifier" : "c22450b2-9999-4896-9ec6-f7c0af5bfa37",
        "type" : "other"
      },
      "mbox" : "obrad.vuckovac@gmail.com",
      "name" : "Obrad Vuckovac"
    },
    "contributor" : [ {
      "contributor_id" : {
        "identifier" : "http://orcid.org/0000-0001-5616-2680",
        "type" : "orcid"
      },
      "name" : "Obrad Vučkovac"
    } ],
    "cost" : [ ],
    "created" : "2020-08-06T18:19:38Z",
    "dataset" : [ {
      "dataset_id" : {
        "identifier" : "62c5029c-2322-4eb7-ba52-bf808de1c615",
        "type" : "other"
      },
    },
  ],
}
```

# Zaštita podataka

## Zaštita tokom istraživanja

- Rezervne kopije (*backups*)
  - pravilo 3-2-1
  - "cloud storage"
- Kontrola pristupa
  - zaštita lozinkom
  - enkripcija



Image by [John](#) from [Wikimedia Commons](#) under [CC-BY](#) license

# Zaštita podataka

Dugoročno čuvanje (nakon istraživanja):

- ▶ Procena i izbor podataka
- ▶ Otvoreni i održivi formati
  - ▷ npr. CSV – a ne XLSX ili OPJ;
  - ▷ [lista preporučenih formata \(DANS\)](#)
  - ▷ (vlasnički formati ako su prihvaćeni u naučnoj zajednici)
- ▶ Digitalni repozitorijumi za dugoročno čuvanje podataka



Image by [John](#) from [Wikimedia Commons](#) under [CC-BY](#) license



## *metadata*

data that provides  
information about other data

(Merriam-Webster dictionary)

# Metapodaci


База података ▾		Ново претраживање	🔍
Аутор - особа	Андрић, Иво, 1892-1975 = Andrić, Ivo, 1892-1975		
Наслов	На Дрини ћуприја / Иво Андрић		
Врста грађе	роман		
Језик	српски		
Година	2022		
Издање	2. изд.		
Издавање и производња	Београд : Vulkan izdavaštvo, 2022 (Београд : Вулкан штампарија)		
Физички опис	351 стр. : 22 cm		
Други аутори - особа	Ђукић-Перишић, Жанета, 1956- = Đukić-Perišić, Žaneta, 1956-		
ISBN	978-86-10-03422-6 (картон са омотом)		
Напомене	Тираж 1.500 Стр. 5-20: На Дрини ћуприја: борба против пролазности / Жанета Ђу Речник турцизама, провинцијализама и неких мање познатих изра На омоту белешка о аутору с његовом сликом.		
Предметне одреднице	Андрић, Иво, 1892-1975 -- "На Дрини ћуприја"		
УДК	821.163.41-31 821.163.41-31.09 Andrić I.		
COBISS.SR-ID	57547273		

# Metapodaci

Texas Data Repository > Texas A&M University Dataverse Repository > Forensic Science Study Dataverse >

## Forensic Science Information Seeking 2017 Data

Version 2.2



Sare, Laura; Bankston, Sarah, 2019, "Forensic Science Information Seeking 2017 Data", <https://doi.org/10.18738/T8/2BAQE0>, Texas Data Repository, V2

Cite Dataset ▾ [Learn about Data Citation Standards.](#)

**Description** ⓘ  
Response data from survey posted to the American Academy of Forensic Sciences in 2017 in fulfillment of NIJ Grant. (2019-06-29)

**Subject** ⓘ  
Social Sciences; Other

**Keyword** ⓘ

Access Dataset ▾

Contact Owner | Share

Dataset Metrics ⓘ

448 Views ⓘ

14 Downloads ⓘ

0 Citations ⓘ


Perzistentni identifikator (DOI)

Statistika korišćenja

Verzija

# Metapodaci

**License/Data Use Agreement**

 CC0 1.0

Uslovi korišćenja i pristupa


Files Metadata Terms Versions

Search this dataset...

Filter by  
File Type: All Access: All File Tag: All

Sort

1 to 2 of 2 Files Download

 **American Academy of Forensic Sciences Survey Data.csv**  
Comma Separated Values - 137.4 KB  
Published Jun 29, 2019  
13 Downloads  
MD5: e8a...c34  
Survey Data in CSV format  
Data

Fajlovi (format, veličina)

# Metapodaci

```
▼ <resource xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns="http://datacite.org/schema/kernel-4" xsi:schemaLoc
  <identifier identifierType="DOI"> 10.5072/DataCollector_dateCollected_geoLocationBox </identifier>
  <creators>
    <creator>
      <creatorName nameType="Personal"> Peach, A. </creatorName>
    </creator>
  </creators>
  <titles>
    <title xml:lang="en"> Temperature and Humidity in School Classrooms, Ponhook Lake, N.S., 1961-1962 </title>
  </titles>
  <publisher xml:lang="en"> National Research Council Canada </publisher>
  <publicationYear> 1963 </publicationYear>
  ▼ <subjects>
    <subject xml:lang="en" schemeURI="https://lccn.loc.gov/sh85062931" subjectScheme="LCCN"> Temperature </subject>
    <subject xml:lang="en" schemeURI="https://lccn.loc.gov/sh85133712" subjectScheme="LCCN"> Humidity </subject>
    <subject xml:lang="en" schemeURI="https://lccn.loc.gov/sh88003345" subjectScheme="LCCN"> Classrooms </subject>
    <subject xml:lang="en"> Ponhook Lake (N.S.) </subject>
  </subjects>
  <contributors>
    <contributor contributorType="DataCollector">
      <contributorName nameType="Personal"> Pomegranate, B. </contributorName>
    </contributor>
  </contributors>
```



# ► Dokumentacija

Sve druge informacije koje su neophodne za bolje shvatanje sadržaja i konteksta podataka.

- ▶ (elektronske) laboratorijske sveske
- ▶ *codebooks*
- ▶ softver i kod
- ▶ informacije o uređajima, kalibraciji i sl.
- ▶ `readme.txt`
- ▶ i sl.

# ► Čuvanje i deljenje

Deponujte podatke u **proverene digitalne repozitorijume** za dugoročno čuvanje.

Tipovi repozitorijuma za podatke:

- ▶ tematski
- ▶ opšti
- ▶ institucionalni

Skupovi podataka se mogu publikovati i u: *Data papers*.

# ► Čuvanje i deljenje

## Tematski repozitorijumi

- ▶ najbolja vidljivost;
- ▶ metapodaci odgovaraju naučnoj disciplini;
- ▶ nedostatak: deponovanje može da košta;
- ▶ registar repozitorijuma:  
[re3data.org](https://re3data.org)



# Čuvanje i deljenje

## Opšti repozitorijumi

- ▶ Zotero, Figshare, OSF i sl.;
- ▶ besplatni PID (DOI);
- ▶ podržavaju razne tipove podataka;
- ▶ ne koštaju (osim za dodatne usluge)
- ▶ nedostatak: metapodaci su opšteg tipa (najčešće DataCite ili Dublin Core);

The screenshot displays the Zenodo interface. At the top, there is a search bar, an 'Upload' button, and a 'Communities' dropdown menu. The main content area shows the title 'Open Research Data in Serbia' and a 'Recent uploads' section. A search result is highlighted with a white box, showing the following details:

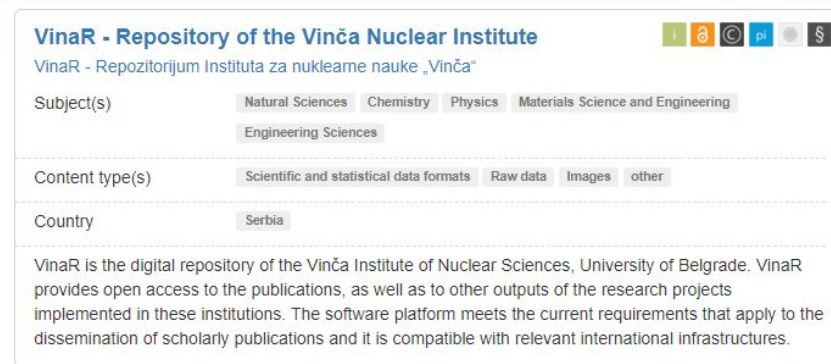
- Publication date:** January 23, 2023
- DOI:** DOI 10.5281/zenodo.7559361
- Keyword(s):** COVID-19, coronavirus, scientometrics, bibliometrics
- Related identifiers:**
  - Cites: <https://pages.semanticscholar.org/coronavirus-research> (Dataset)
  - <https://github.com/diwis/PaperRanking> (Software)
- Supplement to:** [www.biorxiv.org/content/10.1101/2020.04.11.037093v2](http://www.biorxiv.org/content/10.1101/2020.04.11.037093v2) (Preprint)
- Communities:** Coronavirus Disease Research Community - COVID-19, Zenodo
- License (for files):** Creative Commons Attribution 4.0 International


The background of the screenshot shows the 'Serbia.RDM' community page, which includes a 'New upload' button and a description of the community's focus on open research data in Serbia.

# ► Čuvanje i deljenje

## Institucionalni repozitorijumi

- ▶ VinaR je registrovan u re3data.org;
  - ▷ (uskoro i u EOSC-u)
- ▶ sigurno dugoročno čuvanje;
- ▶ nedostatak: slabija vidljivost;



**VinaR - Repository of the Vinča Nuclear Institute** 

VinaR - Repozitorijum Instituta za nuklearne nauke „Vinča“

Subject(s)  Natural Sciences  Chemistry  Physics  Materials Science and Engineering  
 Engineering Sciences

Content type(s)  Scientific and statistical data formats  Raw data  Images  other

Country  Serbia

VinaR is the digital repository of the Vinča Institute of Nuclear Sciences, University of Belgrade. VinaR provides open access to the publications, as well as to other outputs of the research projects implemented in these institutions. The software platform meets the current requirements that apply to the dissemination of scholarly publications and it is compatible with relevant international infrastructures.

# FAIR principi

Findable

Accessible

Interoperable

Reusable





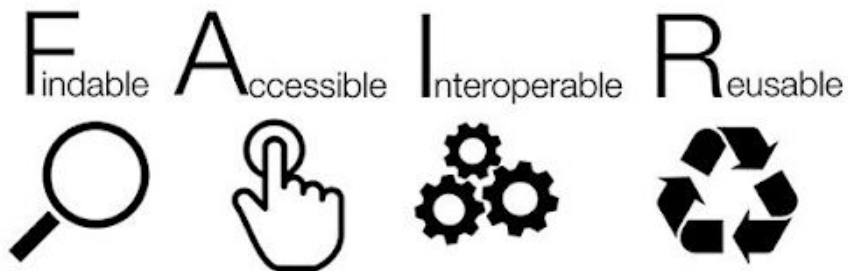
## Mandatory Open Science practices:

*responsible management of research data in line with the **FAIR principles** of 'Findability', 'Accessibility', 'Interoperability' and 'Reusability', notably through the generalised use of data management plans, and open access to research data under the principle 'as open as possible, as closed as necessary', under the conditions required by the grant agreement*

*Horizon Europe Programme Guide*

# FAIR

Istraživački rezultati treba da budu organizovani kako bi bili dostupni, razumljivi, i da bi se lakše delili i ponovo upotrebljavali, kako od strane drugih istraživača tako i od strane **mašina**.





# Findable

F1. (Meta)data are assigned a **globally unique** and **persistent identifier** (DOI, Handle, ARK, etc.)

F2. Data are described with **rich metadata**

F3. Metadata clearly and explicitly include the identifier of the data they describe

F4. (Meta)data are registered or indexed in a **searchable resource**



# Accessible

A1. (Meta)data are retrievable by their identifier using a **standardised communications protocol**

A1.1 The protocol is open, free, and universally implementable

A1.2 The protocol allows for an authentication and authorisation procedure, where necessary

A2. **Metadata are accessible**, even when the data are no longer available



# Interoperable

11. (Meta)data use a formal, accessible, shared, and broadly applicable **language for knowledge representation**.
12. (Meta)data use **vocabularies** that follow FAIR principles
13. (Meta)data **include qualified references** to other (meta)data



# Reusable

R1. (Meta)data are **richly described** with a plurality of accurate and relevant attributes

R1.1. (Meta)data are released with a clear and accessible **data usage license**

R1.2. (Meta)data are associated with detailed **provenance**

R1.3. (Meta)data meet domain-relevant **community standards**



# Open Data

“as open as possible,  
as closed as necessary”



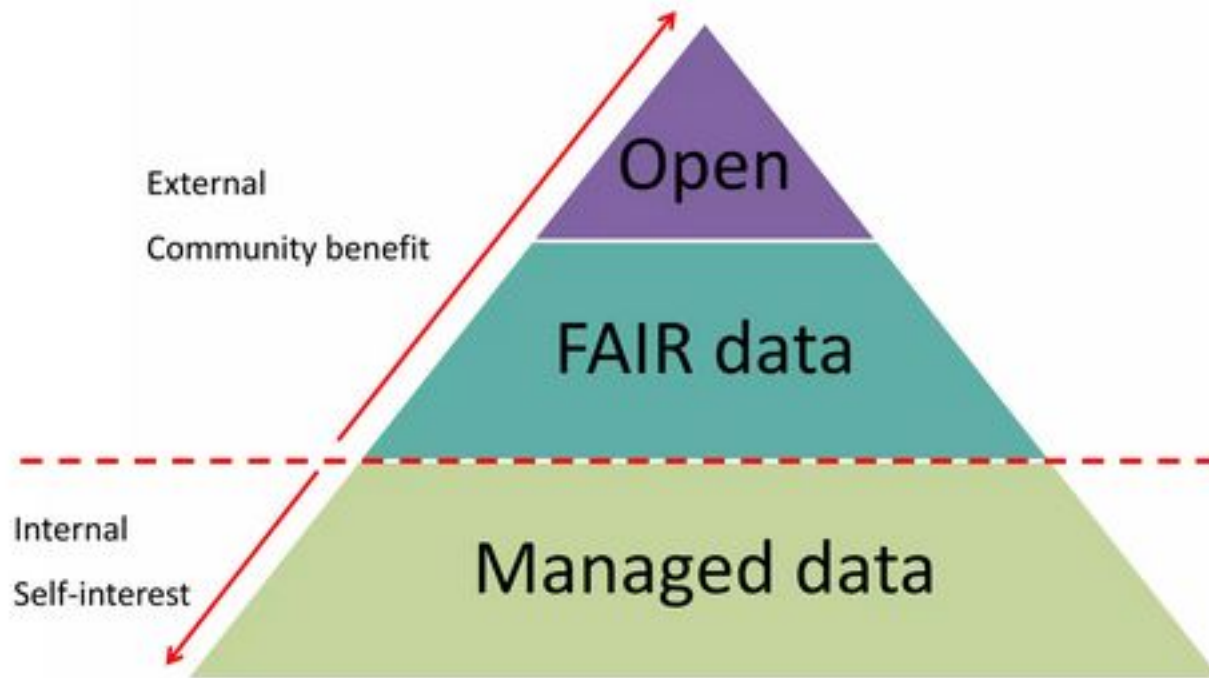
## European Commission - Horizon Europe

- “as soon as possible and within the deadlines set out in the DMP, **ensure open access** – via the repository – **to the deposited data**, under the latest available version of the Creative Commons Attribution International Public License (CC BY) or Creative Commons Public Domain Dedication (CC0) or a licence with equivalent rights, following the principle ‘as open as possible as closed as necessary’, unless providing open access would in particular:
  - be against the beneficiary’s legitimate interests, including regarding commercial exploitation, or
  - be contrary to any other constraints, in particular the EU competitive interests or the beneficiary’s obligations under this Agreement; if open access is not provided (to some or all data), this must be justified in the DMP
- “Metadata of deposited data must be open under a CC Public Domain (CC 0) or equivalent, in line with the FAIR principles (in particular machine-actionable).”

*EU Grants - Model Grant Agreement*



As open as possible,  
as closed as necessary



Jones, S. Open, FAIR data and RDM. 2018.  
<https://www.slideshare.net/sjDCC/open-fair-data-and-rdm>.  
Available under [Creative Commons Attribution](#) License.



# ► Zaključak

Upravljajte podacima na odgovoran način i u skladu sa FAIR principima.

- ▶ pripremite Data Management Plan (DMP);
- ▶ deponujte podatke u repozitorijum;
- ▶ omogućite informacije o svim drugim rezultatima, alatima, instrumentima i sl. preko tog istog repozitorijuma.

# ► Hvala na pažnji

Obrad Vučkovic

INN "Vinča" - Biblioteka

Univerzitet u Beogradu

Prezentacija: [biblioteka.vin.bg.ac.rs](http://biblioteka.vin.bg.ac.rs)



Except as otherwise noted, this presentation is licensed under the Creative Commons Attribution 4.0 International Licence. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

# Credits

Special thanks to all the people who made and released these awesome resources for free:

- ▶ Presentation template by [SlidesCarnival](#)
- ▶ Illustrations by [Sergei Tikhonov](#)
- ▶ Photographs by [Unsplash](#)