



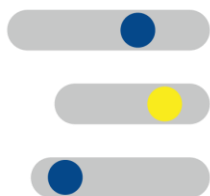
FAIR principi za istraživačke podatke i za softver

Dr Nadica Miljković, vanredni profesor, Rukovodilac Co-creation projekta

Elektrotehnički fakultet Univerziteta u Beogradu, kabinet 68

e-mail: nadica.miljkovic@etf.rs

Serbi.RDM



EOSCsecretariat.eu

Setup and management of the EOSC Secretariat supporting the EOSC Governance

FAIR principi

- Na radionici “*Jointly Designing a Data Fairport*” (Lajden, Holandija, 2014) definisani su FAIR principi.
 - U definisanju FAIR principa učestvovali su predstavnici akademije i industrije.
 - Motivacija je bio zajednički interes da se iskoriste sredstva uložena u istraživanja.
- Da bi se skupovi podataka mogli što bolje iskoristiti neophodno je da podležu FAIR principima:
 - *Findable* (pretraživi)
 - *Accessible* (dostupni)
 - *Interoperable* (interoperabilni)
 - *Reusable* (ponovo upotrebljivi)
- FAIR važi i za ljude i za mašine.
- Skupovi podataka koji zadovoljavaju ove principe imaju velike šanse da budu pronađeni, citirani i upotrebljeni.
- FAIR ≠ otvoreno
- Izvor (Serbia.RDM vodič): <https://rdm.open.ac.rs/index.php/8-rdm-basics/3-rdm-data?highlight=WyJmYWlyll0=> i Zenodo: <https://zenodo.org/communities/ords>

ANNEX 1: Horizon 2020 FAIR Data Management Plan (DMP) Template

INTRODUCTION

This Horizon 2020 FAIR DMP template has been designed to be applicable to any Horizon 2020 project that produces, collects or processes research data. You should **develop a single DMP for your project** to cover its overall approach. However, where there are specific issues for individual datasets (e.g. regarding openness), you should clearly spell this out.

FAIR data management

In general terms, your research data should be 'FAIR', that is findable, accessible, interoperable and re-usable. These principles precede implementation choices and do not necessarily suggest any specific technology, standard, or implementation-solution.

This template is not intended as a strict technical implementation of the FAIR principles, it is rather inspired by FAIR as a general concept.

More information about FAIR:

[FAIR data principles \(FORCE11 discussion forum\)](#)

[FAIR principles \(article in Nature\)](#)

PrtSc: https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm#A1-template

Od 2017. godine svaki novi projekat mora da dostavi DMP u prvih šest meseci od dobijanja sredstava.

Izvor: <https://rdm.open.ac.rs/index.php/9-dmp/9-dmp-funders?highlight=WyJmYWlyllo=>

FAIR

Findable

(pristupačnost, pretraživost)

- jedinstveni i perzistentni identifikatori;
- opisani detaljnim metapodacima;
- (meta)podaci se indeksiraju u pretraživom repozitorijumu.

Accessible

(dostupnost)

- (meta)podaci se preuzimaju putem standardizovanih komunikacionih protokola;
- metapodaci su uvek dostupni.

Interoperable

(interoperabilnost)

- metapodaci koriste rečnik koji prati FAIR principe;
- metapodaci obuhvataju kvalifikovane reference na druge metapodatke.

Reusable

(ponovna upotrebljivost)

- metapodaci imaju jasnu i pristupačnu licencu za korišćenje;
- detaljne informacije o poreklu podataka (eng. *provenance*).

Iz: Vučkovic, Obrad. (2021, February). Plan upravljanja istraživačkim podacima kao primer dobre prakse (Version 1). Zenodo. <http://doi.org/10.5281/zenodo.4560394>

Metapodaci?

- Podaci o podacima (na primeru rada):
 - Naslov rada
 - Ključne reči
 - Autor(i)
 - Naziv časopisa
 - Informacija o projektima
 - Trajni identifikatori (DOI)
 - Apstrakt
 - Identifikatori (URL, WoS, Scopus, ...)
 - Godina
 - Strane
 - Informacije o pravima i uslovima korišćenja
 - ...
- Najčešće promenljivi u repozitorijumima
- Mogu biti formirani po opštoj ([Dublin Core](#)) ili specifičnoj ([AVM](#) od eng. *Astronomy Visualization Metadata*) šemi
- Više na <https://rdm.open.ac.rs/index.php/docs-metadata/metadata?lang=sr>



GDE DELITI?

Izbor repozitorijuma i FAIR principi

- Izbor repozitorijuma igra ključnu ulogu u uređenju podataka po FAIR principima.
- Pri izboru repozitorijuma potrebno je obratiti pažnju da pruža mogućnost da podaci budu FAIR, odnosno da:
 - nudi standardizovane metapodatke;
 - tipovi podataka koji mogu da se pohranjuju budu jasno naznačeni;
 - poseduje svu potrebnu dokumentaciju;
 - dodeljuju trajne identifikatore (DOI, HANDLE);
 - ...
- Pri izboru repozitorijuma može Vam koristiti registar re3data.org (eng. *Registry of Research Data Repositories*) u kojem su detaljno opisani svi repozitorijumi podataka.
- Izvor: <https://rdm.open.ac.rs/index.php/13-deljenje/27-data-repositoties?highlight=WyJmYWlyllo=>



FAIR PRINCIPI ZA SOFTVER

FAIR za istraživački softver?

- Trenutno postoji više inicijativa za re-definisanjem FAIR principa tako da budu primenljivi i na softver, a ne samo na podatke.
 - Izvor: <https://rdm.open.ac.rs/index.php/sharing-publishing/o-softverskim-licencama?highlight=WyJmYWlyl10=>
- Lamprecht, Anna-Lena, et al. "Towards FAIR principles for research software." *Data Science* 3.1 (2020): 37-59, <https://doi.org/10.3233/DS-190026>:
 - Softver nije podatak (FAIR je definisan za podatke).
 - Softverski kod se može pokretati (eng. *executable*) i softver je složene prirode (može biti zavistan od drugog softverskog koda). Takođe, softver može proizvoditi podatke.
 - Preporuka je da se prošire FAIR principi.
- Ovo je tema [RDA](#) (eng. *Research Data Alliance*) grupe [FAIR4RS](#) (eng. *FAIR for Research Software*)
- Ima i kritika FAIR principa za softver, red je da ih spomenem, <https://peertube.social/videos/watch/d99c9e7c-892c-4248-beac-3e463920c808>, "Free Software for Open Science - 36c3" predavanje koje je održao Purine Bitter iz [Fondacije za slobodan softver](#)
 - Softver se može unaprediti. Dograditi. Podaci ne mogu.

FAIR za softver, ali ne za CR?

- CR – *Computational Reproducibility* (srp. računarska ponovljivost)
- Ovo je tema RDA grupe [CURE-FAIR WG](#) (eng. *Working Group*).
 - Još uvek je u izradi (eng. *draft report*) sledeći izveštaj "CURE-FAIR: Challenges" koji je pripremila RDA CURE-FAIR WG podgrupa 3, a autori su: Limor Peer, Florio Arguillas, Sonia Barbosa, Andrew Davison, Tom Honeyman, Nadica Miljković, Wolmar Nyberg Åkerström, Karsten Peters-von Gehlen
 - Ako želite da budete obavešteni o tome kada radna verzija bude zamenjena finalni, dokumentom, možete se priključiti CURE-FAIR WG na <https://www.rd-alliance.org/node/66744/members>

WG

CURE-FAIR WG

Taxonomy:

Posts

Wiki

Events

Repository

Outputs

Case Statements

Plenaries

Members

create new content

Group Status:  WGs Producing deliverables (~6-12 months after RDA endorsement)

 **Join Group**

> **status:** Recognised & Endorsed
Chair (s): Limor Peer, Florio Arguillas, Thu-Mai Christian, Tom Honeyman
Group Email: [group_email]
Secretariat Liaison: Stefanie Kethers

Computational reproducibility CR

- “An article about computational science in a scientific publication is not the scholarship itself, it is merely advertising of the scholarship. The actual scholarship is the complete software development environment and the complete set of instructions which generated the figures.” ([J. B. Buckheit and D. L. Donoho *Wavelets Statist.* 103, 55–81; 1995](#))
 - Više u Perkel, Jeffrey M. "A toolkit for data transparency takes shape." *Nature* 560.7718 (2018): 513-516, <https://doi.org/10.1038/d41586-018-05990-5>.
- CR može biti delimična ili potpuna
 - Nivo deljenja (eng. *white or black box*)
 - Portabilnost (isti OS ili sličan)
 - Pokrivenost (koliko se može ponoviti)
 - Više u Freire, Juliana, Philippe Bonnet, and Dennis Shasha. "Computational reproducibility: state-of-the-art, challenges, and database research opportunities." *Proceedings of the 2012 ACM SIGMOD international conference on management of data*. 2012, <https://doi.org/10.1145/2213836.2213908>

Trenutna rešenja za deljenje koda?

- Da li je [GitHub](#) dobro rešenje?
 - Odličan za razvoj softvera (npr. verzionisanje, saradnja)
 - Problem je što nema odgovarajućih metapodataka koji su ključni za FAIR principe
 - Na GitHub-u nema trajnih identifikatora (eng. *Persistent Identifier*), ali ima uputstvo za kreiranje DOI preko Zenoda, <https://guides.github.com/activities/citable-code/> - VAŽNO za citiranje softverskog koda
- A [Software Heritage](#)?
 - Preporučuje se (Lamprecht, Anna-Lena, et al. "Towards FAIR principles for research software." *Data Science* 3.1 (2020): 37-59, <https://doi.org/10.3233/DS-190026>)
 - Deo su i [FAIRsFAIR](#) Evropskog projekta (eng. *Fostering Fair Data Practices in Europe*)

Loše prakse?

- FAIR nije isto što i otvoreni podaci.
- FAIR nije isto što i slobodan softver (eng. [Free and Open Source Software](#), FOSS).
- Marwick, Ben. "Computational reproducibility in archaeological research: Basic principles and a case study of their implementation." *Journal of Archaeological Method and Theory* 24.2 (2017): 424-450, <https://doi.org/10.1007/s10816-015-9272-9>
 - Ponovljiva istraživanja su moguća ako se koristi FOSS
 - "Mouse-operated point-and-click interface with commercial software... cannot be available for inspection due to... proprietary code of the software"
- Iako FAIR podrazumeva dobre prakse, nije garancija ni kvaliteta softverskog koda ni računarske ponovljivosti.
- Videćemo šta će razni projekti i inicijative tek da ponude u budućnosti.

Buduća rešenja? Na primer:

- Rešenja zasnovana na primeni računarstva u oblaku (eng. *cloud-based platform*) za
 - organizaciju podatka,
 - saradnju na razvoju i
 - računarsku ponovljivost.
- Uz [Docker](#) alate za kontejnersku virtuelizaciju okruženja u kome kod radi postiže se računarska ponovljivost.
- Uz integraciju ovih alata [Code Ocean](#), [Whole Tale](#), [Renku](#) sa npr. [Dataverse](#) repozitorijumom za podatke (ili sa [Zenodom](#)) ispunjeni su i FAIR principi.
- Više u Trisovic, Ana, et al. "Advancing computational reproducibility in the Dataverse data repository platform." *Proceedings of the 3rd International Workshop on Practical Reproducible Evaluation of Computer Systems*. 2020, <https://doi.org/10.1145/3391800.3398173>

Naučna publikacija u budućnosti

- Nije jednostavno čitati publikaciju koja sadrži tekst, podatke i softver. Zahtevaju se dodatne veštine od autora i čitalaca.
 - Više u: Nüst, D., Boettiger, C., & Marwick, B. (2018). How to Read a Research Compendium. *arXiv preprint arXiv:1806.09525*
- Kako čitati takve radove?
- Kako organizovati recenzije?
- Generalno
 - Kako nagraditi/motivisati istraživače?
 - Da li treba usloviti autore da dele podatke i kod?



ZA KRAJ



FAIR AWARE upitnik

- <https://fairaware.dans.knaw.nl/>
- U formi kratkog kviza
- Odlično da proverite svoje prakse i/ili svoja znanja
- Razvio ga je [FAIRsFAIR](#) tim ([DANS](#), [DCC](#), [UniHB](#))
- Lako se pamti.



Data Archiving and Networked Services



Let's assume you have research data almost ready for uploading to a repository: do you already know how you and the repository can work together to make the data as findable, accessible, interoperable and reusable (FAIR) as possible? By guiding you through the assessment process, the FAIR-Aware tool can help you to better understand the FAIR Principles [↗](#) and how making data FAIR can increase the potential value and impact of your data.

FAIR-Aware is an disciplinary-agnostic online tool developed by the FAIRsFAIR [↗](#) project. Different scientific communities can adapt it to their own use. You should, however, have a target dataset in mind to be able to answer the questions and complete the assessment.

FAIR institucije?

- Taco de Bruin, Sarah Coombs, Jutta de Jong, Irene Haslinger, Henk van den Hoogen, Frans Huigen, ... Jacquelijnn Ringersma. (2020, November 3). Do I-PASS for FAIR. A self assessment tool to measure the FAIR-ness of an organization (Version 1). Zenodo. <http://doi.org/10.5281/zenodo.4080867>
- Da proverite prakse u Vašim institucijama? Samovrednovanje?
- Razvijaju se i FAIR sertifikati za repozitorijume, <https://www.fairsfair.eu/fair-certification>

**DO
I-PASS
FOR FAIR?**



Self assessment tool to
measure the FAIR-ness
of an organization

BEGINNER

INTERMEDIATE

ADVANCED

Korisni linkovi

- https://en.wikipedia.org/wiki/FAIR_data
- <https://www.go-fair.org/fair-principles/>
- <https://www.openaire.eu/how-to-make-your-data-fair>
- <https://www.force11.org/group/fairgroup/fairprinciples>
- <https://fair-software.nl/>
- <https://www.fairsoftware.co.uk/>
- <https://librarycarpentry.org/Top-10-FAIR/>
- <https://www.fairsfair.eu/articles-publications/decoding-fair-principles-are-they-relevant-software>
- <https://www.uu.nl/en/research/open-science/tracks/fair-data-and-software>
- <https://www.software.ac.uk/tags/fair-software>
- <https://www.library.universiteitleiden.nl/researchers/data-management/fair-data>
- Gruenpeter, Morane, Di Cosmo, Roberto, Koers, Hylke, Herterich, Patricia, Hooft, Rob, Parland-von Essen, Jessica, ... Jones, Sarah. (2020). M2.15 Assessment report on 'FAIRness of software' (Version 1.1). Zenodo, <https://zenodo.org/record/4095092>.
- Wilkinson, Mark D., et al. "The FAIR Guiding Principles for scientific data management and stewardship." *Scientific data* 3.1 (2016): 1-9, <https://doi.org/10.1038/sdata.2016.18>.