



HRDA meetup

A kutatási adatok kezelésének FAIR alapelvei

Holl András – MTA KIK

MTA KIK

2020 Ápr. 7. (Máj. 26.)

Élő közvetítés a weben:

<https://mcu.niif.hu/live/play/907-3108-7645>

Az előadás vetített anyaga:

https://openaccess.mtak.hu/dokumentumok/meetup2_FAIR_slides.pdf

Kérdésfeltevés online:

<https://app.sli.do/event/xuzzin3n>

Kérdésfeltevés e-mailben:

openaccess @ konyvtar . mta . hu

Research Data Alliance (RDA), HRDA



<https://www.rd-alliance.org/>

HRDA - az RDA magyar csoportja
(Hungarian National Node of RDA)

<http://hrda.hu/>

HUNOR - KIFÜ - NKFI Hivatal

Kutatási adatok – Open Research Data / FAIR Research Data

A nyílt hozzáférésű tudomány (Open Science) mozgalom eredeti célkitűzései szerint a tudományos közlemények szabad hozzáférhetőségéhez hasonlóan a kutatási adatok is szabadon hozzáférhetővé kell váljanak. Mivel ez sok esetben akadályokba ütközik, paradigmaváltás történt:

„A kutatási adatok legyenek annyira nyíltak, amennyire lehetséges, és annyira zártak, amennyire szükséges.”

„As open as possible, as closed as necessary.”

A kutatási eredmények ellenőrzéséhez szükséges az adatok rendelkezésre állása és a jó dokumentáció. Újrafelhasználásukhoz pedig megtalálhatónak, szabványosnak kell lenniük, és nyilvánvalóak kell legyenek a felhasználás feltételei.

**FAIR – Findable, Accessible,
Interoperable, Reusable
(megtalálható, hozzáférhető,
szabványos, újrafelhasználható)**

2016-ban jelentek meg a FAIR alapelvek.

– Wilkinson, Mark D. et al. “The FAIR Guiding Principles for scientific data management and stewardship.”
Scientific data vol. 3 160018. 15 Mar. 2016,
doi:[10.1038/sdata.2016.18](https://doi.org/10.1038/sdata.2016.18)

– FORCE11 – The FAIR data principles
<https://www.force11.org/group/fairgroup/fairprinciples>

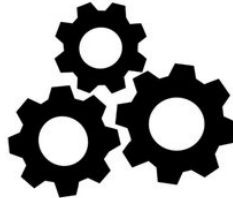
– GoFAIR – FAIR Principles
<https://www.go-fair.org/fair-principles/>

F
Findable

A
Accessible

I
Interoperable

R
Reusable



Rec. 3: A model for FAIR Data Objects

Implementing FAIR requires a model for FAIR Data Objects which by definition have a PID linked to different types of essential metadata, including provenance and licencing. The use of community standards and sharing of code is also fundamental for interoperability and reuse.

FAIR

- **F: Findable** (megtalálható)
 - Egyedi azonosítóval rendelkezik (pl. DOI);
 - gazdagon el van látva leíró adatokkal (metaadatok);
 - be van jegyezve, indexálva van valamilyen nyilvános, kereshető szolgáltatásba.

FAIR

- A: Accessible (hozzáférhető)

- Az egyedi azonosítón keresztül mind a leíró adatok, mind maguk az adatok elérhetőek valamilyen szabványos protokoll segítségével;
- a protokoll nyílt, ingyenes, elterjedt;
- a protokoll lehetővé teszi az azonosítást és a jogosultságkezelést, amennyiben erre szükség van;
- a leíró adatok akkor is hozzáférhetőek, ha az adatok már nem.

FAIR

- I: Interoperable (szabványos)

- Az adatok és a leíró adatok szabványos és értelmezhető formában vannak;

- a leíráshoz, adatrögzítéshez, dokumentációhoz használt szótárak maguk is eleget tesznek a FAIR alapelveknek;

- sok kereszthivatkozás segíti az értelmezést.

FAIR

- R: Reusable (újrafelhasználható)
 - Egyértelmű felhasználási licenc áll rendelkezésre;
 - az adatok származása, keletkezése jól dokumentált;
 - a tudományterületi szabványoknak megfelel.

Hogyan lehet a FAIR követelményeknek megfelelni?

- 1.) DOI azonosítót kell adni az adatoknak. Ebben az MTA KIK segíteni tud.
- 2.) El kell helyezni az adatokat egy (minősített / FAIR) repozitóriumban. Meg lehet keresni a hazai repo-kat - pl. az MTA KIK REAL-t - és megkérdezni, el tudják-e helyezni az adatokat? Vagy nemzetközi, általános vagy szakterületi repozitóriumot kell választani. A kiválasztásban segítenek az intézményi könyvtárak, HUNOR.
- 3.) Példákat kell keresni az adott szakterületről, és megpróbálni annak mintájára rögzíteni és leírni az adatokat.
- 4.) Választani kell egy Creative Commons v. GPL stb. licencet.

Milyen adatokat lehet elhelyezni egy publikációs repozitóriumban?

- Publikációkhoz kapcsolódó adatokat, a publikáció mellett, vagy arra való hivatkozással;
- lehetőleg olyan adatokat, amelyeket a kapcsolódó cikk bírálói ellenőriztek, vagy ellenőrizhetnek;
- egyszerű adatformátumokat lehet használni, melynek kezeléséhez nem szükséges külön szoftver (pl. szöveg, XML, elterjedt, szabványos formátumú kép);
- olyan adatokat, melyek mennyisége (állományok mérete és száma) nem túl nagy.

Ezekben az esetekben a FAIR alapelvek a publikációs kapcsolaton keresztül valósulhatnak meg.

Más esetekben speciális adatrepozitóriumok használata célszerű.

Segítség: [openaccess @ konyvtar . mta . hu](mailto:openaccess@konyvtar.mta.hu)

Információk: <https://openaccess.mtak.hu>
<https://openscience.hu/>

FAIR Data Maturity Model - indicators
<https://docs.google.com/spreadsheets/d/1mkjElFrTBPBH0QViODexNur0xNGhJqau0zkL4w8RRAw/edit#gid=0>

European Commission - Turning FAIR into Reality
https://ec.europa.eu/info/sites/info/files/turning_fair_into_reality_1.pdf

HRDA - MTA KIK kérdőív
https://docs.google.com/forms/d/e/1FAIpQLSdh51_cQhAKk5JUjtzPJHAdk_OqsfFsldlMRN5X0CHo7Q3v-g/viewform

FAIRsFAIR projekt

Fostering Fair Data Practices in Europe

<https://www.fairsfair.eu/>

RDA FAIR Data Maturity Model WG – Indicators

F	F1	F1-01M	Metadata is identified by a persistent identifier	Recommended
	F1	F1-02M	Metadata is identified by a universally unique identifier	Recommended
	F1	F1-01D	Data is identified by a persistent identifier	Mandatory
	F1	F1-02D	Data is identified by a universally unique identifier	Mandatory
	F2	F2-01M	Sufficient metadata is provided to allow discovery, following domain/discipline-specific metadata standard	Recommended
	F2	F2-02M	Metadata is provided for the discovery-related elements defined by the RDA Metadata IG, as much as possible and relevant, if no domain/discipline-specific metadata standard is available	Recommended
	F3	F3-01M	Metadata includes the identifier for the data	Mandatory
	F4	F4-01M	Metadata or landing page is harvested by general search engine	Recommended
	F4	F4-02M	Metadata is harvested by or submitted to domain/discipline-specific portal	Recommended
	F4	F4-03M	Metadata is indexed in institutional repository	Recommended

A	A1	A1-01M	Metadata includes information about access conditions	Optional
	A1	A1-01D	Data is available for manual download	Recommended
	A1	A1-02D	Data is available for automatic download	Recommended
	A1	A1-02M	Metadata identifier resolves to a metadata record	Optional
	A1	A1-03D	Data identifier resolves to a data file	Mandatory
	A1	A1-03M	Metadata is accessed through standardised protocol	Recommended
	A1	A1-04D	Data is accessible through standardised protocol	Recommended
	A1.1	A1.1-01D	Data is accessible through a free access protocol	Mandatory
	A1.1	A1.1-02D	Data is accessible through an open-source access protocol	Recommended
	A1.1	A1.1-01M	Metadata is accessible through a free access protocol	Mandatory
	A1.1	A1.1-02M	Metadata is accessible through an open-source access protocol	Recommended
	A1.2	A1.2-01D	Data is accessible through an access protocol that supports authentication	Recommended
	A1.2	A1.2-02D	Data is accessible through an access protocol that supports authorisation	Recommended
	A1.2	A1.2-01M	Metadata includes information relevant for access control	Mandatory
	A2	A2-01M	Metadata is guaranteed to remain available after data is no longer available	Mandatory

I	I1	I1-01M	Metadata uses knowledge representation expressed in standardised format	Recommended
	I1	I1-02M	Metadata uses machine-understandable knowledge representation	Optional
	I1	I1-03M	Metadata uses self-describing knowledge representation	Optional
	I1	I1-01D	Data uses knowledge representation expressed in standardised format	Recommended
	I1	I1-02D	Data uses machine-understandable knowledge representation	Optional
	I1	I1-03D	Data uses self-describing knowledge representation	Optional
	I2	I2-01M	Metadata uses standard vocabularies	Recommended
	I2	I2-01D	Data uses standard vocabularies	Recommended
	I2	I2-02M	Metadata uses FAIR-compliant vocabularies	Optional
	I2	I2-02D	Data uses FAIR-compliant vocabularies	Optional
	I3	I3-01M	Metadata includes references to other metadata	Recommended
	I3	I3-01D	Data includes references to other data	Recommended
	I3	I3-02M	Metadata includes sufficiently qualified references to other metadata	Recommended
	I3	I3-02D	Data includes sufficiently qualified references to other data	Optional

R	R1	R1-01M	Sufficient metadata is provided to allow reuse, following domain/discipline-specific metadata standard	Recommended
	R1	R1-02M	Metadata is provided for the reuse-related elements defined by the RDA Metadata IG, as much as possible and relevant, if no domain/discipline-specific metadata standard is available	Recommended
	R1.1	R1.1-01M	Metadata includes information about the licence under which the data can be reused	Mandatory
	R1.1	R1.1-02M	Metadata refers to a standard reuse licence	Recommended
	R1.1	R1.1-03M	Metadata includes licence information in the appropriate element of the metadata standard used	Mandatory
	R1.1	R1.1-04M	Metadata refers to a machine-understandable reuse licence	Optional
	R1.1	R1.1-06M	Metadata includes information about consent for reuse (e.g. for personal data)	Recommended
	R1.2	R1.2-01M	Metadata includes provenance information according to community-specific guidelines	Recommended
	R1.2	R1.2-02M	Metadata includes provenance information according to a cross-domain language	Optional
	R1.3	R1.3-01M	Metadata complies with a community standard	Mandatory
	R1.3	R1.3-01D	Data complies with a community standard	Mandatory
	R1.3	R1.3-02M	Metadata is expressed in compliance with a machine-understandable community standard	Optional
	R1.3	R1.3-02D	Data is expressed in compliance with a machine-understandable community standard	Optional

Repository certification - CoreTrustSeal & FAIR

CoreTrustSeal <https://www.coretrustseal.org/>

FAIR követelmények beépítése:

<https://www.coretrustseal.org/why-certification/review-of-requirements/>

MTA KIK - DE RDA grant: repozitórium minősítási ajánlások honosítása