

ADATKEZELÉSI TERVEK

tájékoztatás a Lendület pályázók részére

Holl András – Virág Gabriella
(MTA KIK)

2025.02.11.

Az adatkezelési terv (Data Management Plan) készítésének célja

- A pályázat kiírójának, a kutatást befogadó intézménynek segít tervezni a szükséges erőforrásokat.
- A pályázat kiírójának, a bírálóknak segít bemutatni, mennyire megalapozott a kutatómódszertan, mennyire lesz verifikálható a kutatás, keletkeznek-e a további kutatásokban felhasználható adatok?
- A kutatóknak segít megtervezni, kutató csoporton belül kommunikálni a kutatási módszertan adatkezelési aspektusait. A kezdetektől felvázolja a kutatási adatkezelés, disszemináció menetét, segít abban, hogy már a kutatás során megfelelően dokumentált adatok kerüljenek rögzítésre. Hozzájárul a nemzetközi szabványok alkalmazásához. (Összegezve segíti a módszeres munkát, a jó gyakorlatok alkalmazását, hozzájárul a kutatás reprodukálhatóságához és növeli a láthatóságot.)
- A DMP (Data Management Plan) alapjául szolgál (esetleg részét képezi) az adatok dokumentációjának.

Milyennek kell lennie az adatkezelési tervnek?

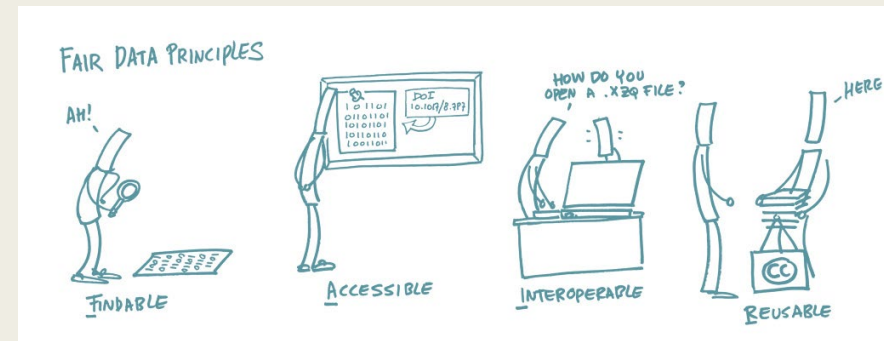
- Az adatkezelési terv élő dokumentum, a kutatás során szükség szerint frissülnie kell.
DMP 1st version/date (Adatkezelési terv 1. változata/dátum)
DMP last update/date (Adatkezelési terv utolsó frissítése/dátum)
- Az adatkezelési terv az adott tudományterület elvárásainak kell megfeleljen, a szakmai bírálók számára érthető kell legyen. Lehet hivatkozni az adott területen elfogadott Domain Data Protocol-ra. (<https://training.ni4os.eu/course/view.php?id=58>)
- Az általános alapelv: „Legyen nyitott, amennyire lehetséges, de legyen zárt, amennyire szükséges.” (“As open as possible, and as closed as necessary.”)
- Az adatkezelést a **FAIR** szempontok szerint kell kifejezni:

Findable – megtalálható

Accessible – hozzáférhető

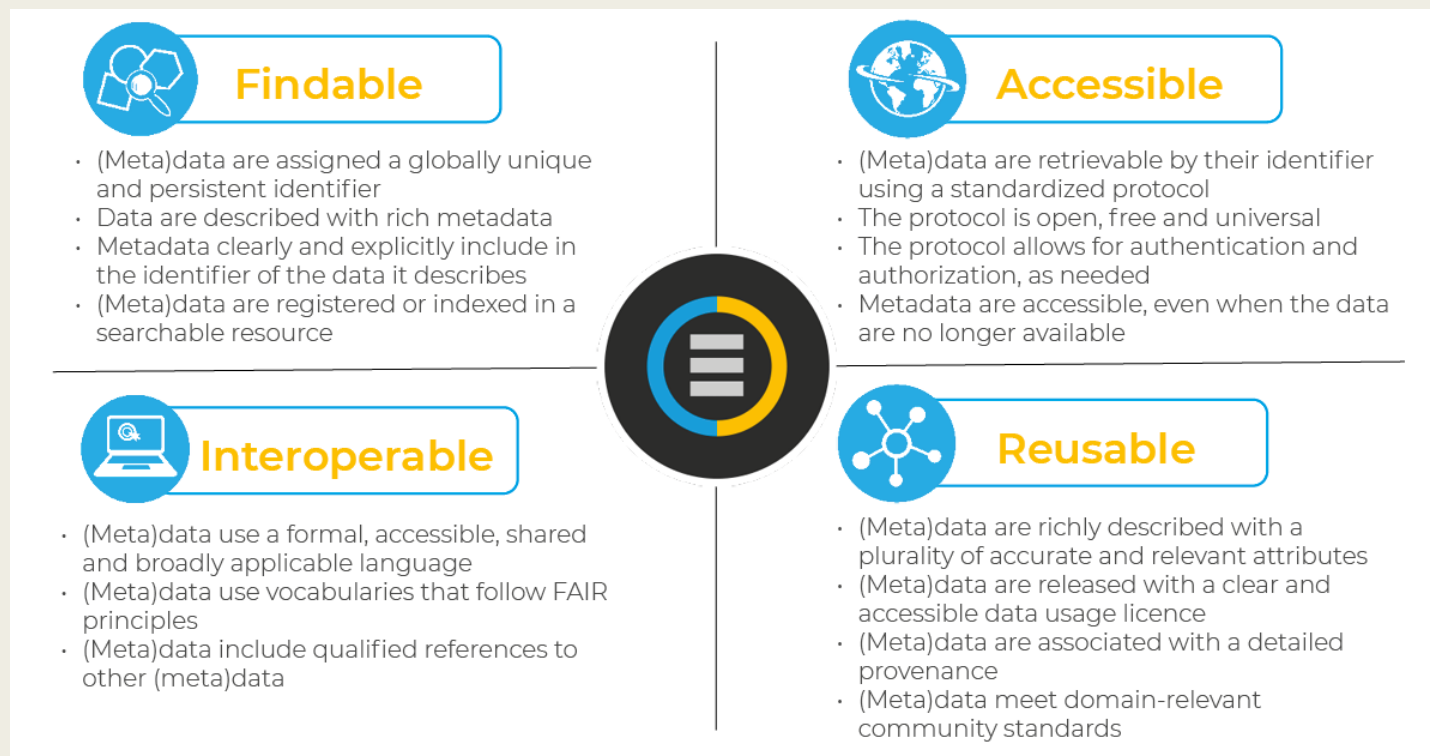
Interoperable – szabványosított

Re-usable – újrafelhasználható



<https://fosteropenscience.eu/>

FAIR adatkezelés



<https://www.ccdc.cam.ac.uk/solutions/about-the-csd/fair-data-principles/>

- NI4OS-Europe tréningplatformján - Minden, amit a FAIR adatokról tudni érdemes címmel - elérhető egy FAIR adatokról szóló dokumentum

https://training.ni4os.eu/pluginfile.php/3638/mod_resource/content/0/HU-NI4OS%20-%20FAIR.pdf

Lendület 2025

Research Data Management Plan (DMP)

szerkezete

- Összefoglaló / Summary
- 1. Az adatok megtalálhatóvá tétele / Making data **F**indable
- 2. Az adatok nyíltan hozzáférhetővé tétele / Making data openly **A**ccessible
- 3. Az adatok szabványosítása / Making data **I**nteroperable
- 4. Az adatok újrafelhasználhatóságának növelése / Increase data **R**e-use
- 5. Erőforrások hozzárendelése és adatbiztonság / Allocation of resources and data security

Összefoglaló / Summary

- az adatgyűjtés/-létrehozás célja
- kapcsolat a projekt célkitűzéseivel
- adatkészlet hivatkozása és neve
- adattípusok és -formátumok
- a létrehozott/gyűjtött adatok eredete és várható mérete
- meglévő adatok újrafelhasználása, ha van ilyen
- az adatok hasznossága: kinek lesz hasznos

Ha a projekt során nem keletkeznek adatok - itt kell kijelenteni.

Mivel a DMP-ben másutt erre nincs hely, itt kell foglalkozni az adatgyűjtés esetleges etikai kockázataival, az esetlegesen szükséges engedélyekkel, a közreadásnál felmerülő anonimizálással vagy zárolással.

1. Az adatok megtalálhatóvá tétele / Making data Findable

- az adatkészlet leírása: metaadatok, tartós és egyedi azonosítók, pl. DOI
- az adatok felkutathatósága (metaadatok biztosítása)
- az adatok azonosíthatósága és általános azonosítási mechanizmusok
- alkalmazott elnevezési szokások
- a keresési kulcsszóval kapcsolatos megközelítés
- az egyértelmű verziószámozásra törekvés
- metaadat-alkotási szabványok, ha vannak ilyenek

Metaadatsémák, PID-ek

- A metaadatséma metaadat elemek meghatározott készlete és az ahhoz kapcsolódó szabályok összessége. Pl.: Dublin Core (DC)

A Research Data Alliance egy metaadat-szabványkönyvtárat biztosít, amelyben kereshetők a szakterület-specifikus szabványok és a kapcsolódó eszközök:

<https://rd-alliance.github.io/metadata-directory/>

Ha egy adott tudományágban nincsenek szabványok, elegendő, ha a kutató/kutatócsoport által használt metaadatokat felsorolják.

- Persistent Identifier (PID)

Handle System

Digital Object Identifier (DOI) - DataCite

Open Researcher and Contributor ID (ORCID)

Research Organization Registry (ROR)

Legtöbb esetben a DOI-t a használandó adatrepozitóriumon keresztül kell igényelni. A DOI-hoz kapcsolható metaadatokról az MTA KIK DOI irodájánál lehet tájékozódni.

2. Az adatok nyíltan hozzáférhetővé tétele / Making data openly Accessible

- mely adatokat teszik nyíltan hozzáférhetővé
 - ha egyes adatkészletek továbbra is zártak maradnak ennek okai (jogi/szerződéses/önkéntes)
- hogyan lehet hozzáférni az adatokhoz
 - (szükség van-e a megfelelő szoftvereszközökre/ módszerekre/megoldásokra?)
- hol helyezik el az adatokat és a kapcsolódó metaadatokat, dokumentációt és kódot (adattár/repozitórium?)
- hogyan biztosítják a hozzáférést, ha korlátozások vannak

A nyílt kutatási adatkezelésben való részvétel nem feltétlenül jelenti az összes kutatási adat megnyitását. Inkább az „as open as possible, as closed as necessary” elvet követi, és a megbízható adatkezelés ösztönzésére összpontosít.

Előnyös, ha az adatok szabványos, nyílt, egyszerű formátumban kerülnek elhelyezésre.

Amennyiben speciális szoftver szükséges az adatok ellenőrzéséhez/megjelenítéséhez, szerencsés, ha ez a szoftver szabadon hozzáférhető.

Érdemes a (végső?) DMP-t elhelyezni az adatrepozitóriumban, a dokumentáció részeként.

Repozitórium választás:

- Szakmai gyakorlat
- Re3data - A Registry of Research Data Repositories hasznos listát kínál azokról a repozitóriumokról, ahol elhelyezhetik kutatási adataikat. <https://www.re3data.org/>
- HUN-REN ARP repozitórium
- Minimál repozitóriumok: pl. zenodo, figshare
- Publikációs repozitóriumok: pl. REAL.

[REAL politika]

POSSIBLE ROLE OF PUBLICATION REPOSITORIES AND CRIS SYSTEMS

Publication repositories are used widely, and services exist for aggregation. The repository of the Hungarian Academy of Sciences - REAL - holds some data entries. We describe our principles for accepting data:

i) There has to be an article associated to the data

We require a publication to be deposited along the data. REAL provides discoverability via the DC metadata of the article.

ii) Size, number and formats: small, few, preferably textual data

We cannot handle huge amounts of data - we lack storage space, and the software we use (EPrints) is not well suited for storing large amounts of files in an item. We accept data in textual format (CSV, TXT, XML, Jason, etc.), and also in widely used image (JPEG, TIFF) and document formats (PDF).

iii) „As is” - curation only at ingestion

We check the submitted data package, and negotiate with the depositing researcher, but no further curation is offered.

We might consider preserving digital data appendices (including those distributed in CD-ROMs) along the digital copies of journals.

In the Hungarian national CRIS system (MTMT) published research data can be recorded. Linking PIDs of data packages to publication entries is considered.



Holl, András (2024) Research Data & Publications - Transparent Journal Papers. In: International Digital Curation Conference (IDCC24), 2024. február 19-21., Edinburgh, Scotland. <https://real.mtak.hu/178535/>

3. Az adatok szabványosítása / Making data Interoperable

- milyen szabványos vagy szakterület-specifikus adat- és metaadatszótárakat és módszereket használnak annak érdekében, hogy lehetővé tegyék az interdiszciplináris adatcserét és újrafelhasználást kutatók, intézmények, szervezetek, országok stb. között
pl. a formátumokra vonatkozó szabványok betartása, a nyílt szoftveralkalmazásoknak való megfelelés, a különböző eredetű, különböző adatkészletekkel való újrakombinálás megkönnyítése

Az esetleges újrafelhasználás egyes tudományterületeken lehet már most elterjedt - ez esetben lehet foglalkozni vele. Egyébként érdemes a reprodukálhatóságra koncentrálni, a felhasznált szótárakról, stb. írni. Célszerű megadni a felhasznált szabványos szótárak nevét és elérhetőségét.

Fontos a gazdag metaadatolás, a metaadatszabványok használata, mert megkönnyíti a gépi olvashatóságot, ezáltal az adatok értelmezhetősége növekszik. Érdemes részletes 'Readme' fájlt készíteni az adatcsomaghoz.

4. Az adatok újrafelhasználhatóságának növelése / Increase data Re-use

- milyen adatok maradnak újrafelhasználhatóak és mennyi ideig, van-e korlátozás (okok, mennyi ideig)
- az adatok engedélyezésének módja
- az adatok minőségbiztosítási eljárásai
- a projektben előállított és/vagy felhasznált adatok harmadik felek számára is felhasználhatók-e különösen a projekt befejezése után

Itt érdemes megadni a konkrét felhasználni szándékozott licenceket

(pl. CC, BSD, Apache, GPL, stb.)


























<https://creativecommons.org/share-your-work/cclicenses/>

Creative Commons

A Creative Commons licencek két jog és négy korlátozó feltétel variációjából állnak össze.

Megosztási jog: a mű szabadon másolható, terjeszthető, bemutatható és előadható legyen

Feldolgozási jog: a mű alapján származékos művek (feldolgozások) hozhatók létre

Creative Commons 4.0 licencek	Jogok		Korlátozó feltételek			A licenc korlátozottságának mértéke
Public Domain (CC0) <i>Közkinccs</i>						szabad licenc
Attribution (CC-BY-4.0) <i>Hivatkozd!</i>						szabad licenc
Attribution - Share Alike (CC-BY-SA-4.0) <i>Hivatkozd! - Ugyanígy oszd meg!</i>						szabad licenc
Attribution - No Derivative Works (CC-BY-ND-4.0) <i>Hivatkozd! - Ne változtasd!</i>						korlátozott licenc
Attribution - Non Commercial (CC-BY-NC-4.0) <i>Hivatkozd! - Ne üzletelj vele!</i>						korlátozott licenc
Attribution - Non Commercial - No Derivative Works (CC-BY-NC-ND-4.0) <i>Hivatkozd! - Ne üzletelj vele! - Ne változtasd!</i>						korlátozott licenc
Attribution - Non Commercial - Share Alike (CC-BY-NC-SA-4.0) <i>Hivatkozd! - Ne üzletelj vele! - Ugyanígy oszd meg!</i>						korlátozott licenc

<https://researchdata.hu/utmutato-kutatasi-adatok-kezelesehez>

5. Erőforrások hozzárendelése és adatbiztonság / Allocation of resources and data security

- a projektadatok nyílt hozzáférésűvé tételének becsült költségei és a hosszútávú adatmegőrzés potenciális költségei
- e költségek fedezésének tervezett módjai
- az adatok biztonsági mentésére és helyreállítására vonatkozó eljárások
- az érzékeny adatok átadása és biztonságos tárolása a repozitóriumokban a hosszútávú megőrzés és gondozás érdekében

Gondolni kell a kutatási adatokhoz kapcsolódó költségek kalkulálására, a finanszírozás megtervezésére. Pl.: kutatás/adatgyűjtés alatti adattárolás felmerülő költségei, repozitóriumi díj egyes repozitóriumok esetében, archiválásra való előkészítésének költségei

**A kutatási adatok kezelési tervének
kitöltéséhez az MTA KIK munkatársai
segítséget nyújtanak**

openaccess@konyvtar.mta.hu

Hasznos linkek

- HRDA webinárium: Az adatkezelési terv – Data Management Plan
2020. március 18. [előadás diák]
Holl András
<https://openaccess.mtak.hu/event/az-adatkezelesi-terv-data-management-plan/>
- ARP Hogyan készítsünk kutatási adatkezelési tervet az új NKFIH pályázatokhoz?
2024. június 7. [video felvétel]
Meiszterics Enikő, Egyed-Gergely Júlia
<https://researchdata.hu/eloadasok/bemutato-hogyan-keszitsunk-kutatasi-adatkezelesi-tervet-az-uj-nkfi-h-palyazatokhoz>
- A kutatási adatok FAIR kezelésének alapjai és az adatkezelési terv - ELKH Cloud / HRDA
2021. Május 20. [előadás diák]
Holl András
<https://science-cloud.hu/sites/default/files/2021-07/HRDA-DMP.pdf>

Hasonló adatkezelési tervek:

- NKFIH OTKA adatkezelési terv:
<https://nkfi.gov.hu/palyazoknak/nkfi-alap/excellence-kutatasi-palyazat-excellence24/palyazati-csomag/research-data-management-plan-nkfp24>
- ERC Data Management Plan:
<https://erc.europa.eu/sites/default/files/document/file/ERC-Data-Management-Plan.docx>
- DMP Tool (publikus adatkezelési tervek és létrehozási segédletek):
<https://dmptool.org/>

Köszönjük a figyelmet!