Elmélkedések az adatrepozitóriumokról és a hosszú távú fenntarthatóságról

Holl András

MTA KIK

ARP konferencia, Budapest – 2026 10.8.

A kutatási adatok megőrzéséről és elérhetővé tételéről folyó modern hazai diskurzus több, mint tíz éve tart. Talán érdemes elővenni néhány alapvető kérdést, amelyek egy részére az alkalmazásban már valamilyen válasz született, más részük pedig még továbbra is megoldatlan.

Vizsgáljuk az adatrepozitálás célját, tárgyát, finanszírozását, időtávját, az adatok beágyazottságát, megőrzésének végét, és szót ejtünk az OAIS szabvány értelmezéséről is.

Repozitóriumok:

Kutatói játszadozás: arXiv elődje Paul Ginsparg személyi számítógépén



Könyvek nyilvántartása

1997. évi CXL. törvény (Kulttv.),
3/1975. (VIII. 17.) KM–PM együttes rendelet

MTA KIK eszköz selejtezési szabályzat

Magyar Tudományos Akadémia Könyvtár és Információs Központ
1051 Budapest, Arany János utca 1.

Felesleges vagyontárgyak hasznosításának és a hasznosíthatatlan eszközök selejtezésének szabályzata

Érvényes: 2013. január 1. napjától

Jóváhagyta:……..

Dr. Monok István
főigazgató

Toldy, Ferenc (2017) Utasítás a Magyar Tudományos Akadémia Könyvtára tisztviselői számára. Lichniae, MTA KIK, Budapest. https://real-eod.mtak.hu/7634/



Adatrepozitóriumok:

– Adatrepozitóriumok működtetésének alapelvei
Policy-making for Research Data in Repositories: A Guide

https://repositorium.uminho.pt/server/api/core/bitstreams/4aa34c31-9ec7-4c7d-b453-1da83e6a7044/content

– Mi az adat? (publikációnál: "supporting material": mindaz, ami a publikációban megfogalmazott állítások alátámasztásához szükséges

– cél: reprodukálhatóság/ellenőrizhetőség; újrafelhasználhatóság/későbbi publikáció; kutatásértékelés; szellemi vagyon nyilvántartás/megőrzés; kutatásmódszertan-támogatás/egységesítés

Reprodukálhatóság

https://opusproject.eu/open-science-news/he-reproducibility-crisis-how-open-science-can-save-research/

https://www.ipr.northwestern.edu/news/2024/an-existential-crisis-for-science.html

-- NVO demo (https://spacenews.com/virtual-observatory-prototype-provides-surprise-discovery-early-demo-project-identifies-new-brown-dwarf/)

-- későbbi publikáció (hosszú távú adatgyűjtés) pl. Konkoly Obs., változócsillag fotografikus fotometria

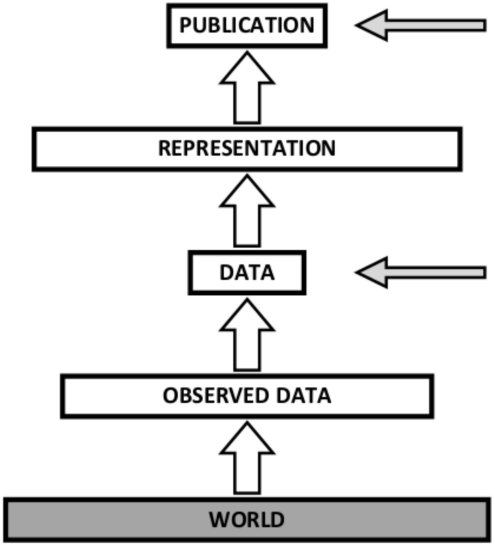
– intézményi háttér (működtetés / informatikai fejlesztés / kapcsolattartás a kutatói közösséggel)

-- külföldi példák: DANS, DCC, CDS (szakmai háttér / konkrét adatkezelés) Archaeology Data System (UK)

- finanszírozás (lehet vegyes)

-- központi finanszírozás: pl. OTKA kiosztott pénzek X százaléka évente (intézmény/infrastruktúra fenntartás)
-- intézményi projekt finanszírozás
-- kutatási projektek költségeibe épülő finanszírozás (működési költségek - a futó projekteknek kell eltartani az archív adatok kezelését. Exponenciálisan növekvő adatmennyiség: költségek X szerese)
- mit kell archiválni?

-- Holl - Small Data Archives and Libraries, LISA VI.



https://real.mtak.hu/3879/

másik megközelítés:

nyers adatok -> script -> publikáció

csak a nyers adatot és kalibrációs adatokat tárolni

– meddig kell archiválni? mikor törölhető?
törölhető, ha nem hivatkozik rá semmi(?)

-- Konkoly Obszervatórium: fotografikus változócsillag észlelések (lemezek) esete

- adatvagyon gazdálkodás, apasztás, sírkő

- adatok értéke (?) (létrehozás költsége? újra-létrehozás költsége? archiválás/tárolás költsége?)

- mentés vagy redundancia?

- láthatóság – DOI szerepe

- kapcsolódási hálózat – publikációk / CRIS / EOSC

– OAIS – reflexiók GD NWS cikkére

Görög, Dániel (2026) OAIS a gyakorlatban In: A tudomány, az oktatás és a közgyűjteményi kiszolgálás új informatikai szinergiái : NETWORKSHOP 2026 : https://real.mtak.hu/243906/

– János vitéz: az adatmegőrzés csak az értelmezhető metaadatok biztosítását vállalhatja fel, a beágyazó világ archiválását nem.

Köszönöm a figyelmet!