

Tematikus összeállítás

AZ EÖTVÖS LORÁND KUTATÁSI HÁLÓZAT ADATREPOZITÓRIUM PLATFORM

THE ELKH ARP RESEARCH DATA REPOSITORY

VENDÉGSZERKESZTŐ: GÁRDOS JUDIT

BEVEZETŐ

INTRODUCTION

Kovács László¹, Gárdos Judit², Kacsuk Péter³

¹műszaki doktor, osztályvezető

ELKH Számítástechnikai és Automatizálási Kutatóintézet Elosztott Rendszerek Osztály, Budapest

laszlo.kovacs@sztaki.hu

²PhD, Társadalomtudományi Kutatóközpont Kutatási Dokumentációs Központ, Budapest

³az MTA doktora, tudományos tanácsadó, tudományos laborvezető

ELKH Számítástechnikai és Automatizálási Kutatóintézet, Budapest

Gyakran mondják, hogy az adat az új arany – és valóban, a tudományos kutatók során és azok lezárultával is egyre nagyobb figyelmet kap a kutatási adatok kezelése, gondozása, megosztása. Ezt ismerte fel több évvel ezelőtt az Európai Bizottság, és kezdte jelentős összegekkel támogatni az adatorientált kutatások feltételeinek megteremtését. Az Open Science mozgalom szerves részévé vált, hogy nemcsak a tudományos publikációk, hanem a kutatási adatok is minél szélesebb körben váljanak hozzáférhetővé.

Két évvel ezelőtt az Eötvös Loránd Kutatási Hálózat (ELKH) vezetése úgy döntött, hogy a kutatási adatkezelés és adatnyilvánosság olyan fontos terület, amelyet érdemes anyagi forrásokkal támogatni. Ennek jegyében először 2020-ban adott 50 millió Ft támogatást az RDA (Research Data Alliance) nemzetközi szervezet magyarországi alszervezetének, a HRDA-nak (Hungarian Research Data Alliance), hogy kezdje el szervezni az ilyen irányú magyar törekvéseket. A támogatásból miniprojektek valósultak meg, amelyek azokat a kutatói törekvéseket voltak

hivatottak ösztönözni, amelyek FAIR-elven¹ osztják meg a kutatási adatokat (tehát úgy, hogy azok megtalálhatók, hozzáférhetők, együttműködők, interoperábilisak – digitális platformok között átjárhatók –, és végül újrafelhasználhatók legyenek). Ezen miniprojektek többsége a SZTAKI-ban (Számítástechnikai és Automatizálási Kutatóintézet) 2019-ben kifejlesztett CONCORDA- (Concentrated Cooperation on Research Data) adatrepozitóriumot használta.

2021-ben a három legaktívabb, ELKH-hoz tartozó HRDA-tagintézmény, a SZTAKI, a Társadalomtudományi Kutatóközpont és a Wigner Fizikai Kutatóközpont az ELKH-től anyagi támogatást kapott egy olyan projekt megvalósítására, amely egyrészt megteremti az ELKH intézmények számára az adatorientált kutatás infrastrukturális környezetét, másrészt aktívan járul hozzá a FAIR-elvű adatorientált kutatás kultúrájának kialakításához és elterjesztéséhez. 2021. november 1-jén mintegy 500 millió forintos ELKH-támogatással indult el az ELKH Adatrepozitórium Platform (ELKH ARP) létrehozását lehetővé tevő projekt.

Az ELKH ARP projekt az intézményhálózat kutatási adatkezelését támogató, új hálózati szolgáltatásrendszer alapjainak megteremtését célozza. A kutatási adatok kezelésének alapvető feltételeit, az adattárolás és adatmegosztás országos szervezeti, humán és műszaki infrastruktúráját hozza létre, felismerve, hogy a (digitális) kutatási adatok tárolása és az adattárolókra, repozitóriumokra épülő adatszolgáltatások kapcsán bőven van fejleszteni valónk Magyarországon.

A projekt célja, hogy lehetővé váljék a kutatási adatok mint értékkel bíró adatvagyron biztonságos digitális tárolása, azok használata, újrafelhasználása, másodlagos és többedleges hasznosítása a kutatási és egyéb területek, például a gazdasági szféra számára. A kutatási adatok értékének tudatosítása, az értéknek megfelelő adekvát adatkezelési és tárolási erőforrás-használat fontos és felvállalt küldetése a projektnek. A projekt gondoskodik a kutatási adatrepozitóriumok felhasználásával kapcsolatos elméleti ismeretek, gyakorlati képességek terjesztéséről, a szükséges, ELKH- és intézményi szintű adatkezelési *policyk* és szabályozások létrehozásáról, a hazai és nemzetközi adat- és metaadat-kezelési, -tárolási szabványok és ajánlások, jó gyakorlatok hazai elterjesztéséről.

A kutatási adatok tárolásának, kezelésének fejlesztése, a szükséges *know-how* létrehozása és a tervezett szakmai képzés megteremti az ELKH kutatóintézet-hálózatban keletkező kutatási adatok jobb és hatékonyabb felhasználhatóságát. A hazai kutatási adatrepozitóriumok létrehozása és az ehhez szükséges fejlesztési és felhasználói kompetenciák kialakítása reményeink szerint javítja a kutatók (európai) K+F-pályázatokban várható sikerességét.

A hazai kutatási adattárolási infrastruktúrák megteremtése lehetőséget biztosít a hasonló célú európai kezdeményezésekbe való bekapcsolódásra mind intézményes, mind pedig technológiai és közvetlen adatkapcsolati szinteken is. Ezen

¹ FAIR – Findable, Accessible, Interoperable, Reusable

túlmenően, a projekt létrehozza a magyar K+F-szférában élenjáró kutatóintézetek olyan együttműködését, amely elegendő hazai és nemzetközi szakmai nyomaték-kal bír e szakmai terület valóságos előmozdításához, és a nemzetközi, valamint hazai intézményi partnerkapcsolatok szélesebb körű kiterjesztéséhez.

Ebben a blokkban az ELKH ARP legfontosabb célkitűzéseiről számolunk be, amelyek a következők:

1. Olyan infrastruktúra tervezése és megvalósítása, amely lehetővé teszi nagy mennyiségű kutatási adat tárolását, archiválását, gyors feldolgozását, kereshetőségét, felhasználását és újrafelhasználását magyar és nemzetközi kutatási projektek számára. Kacsuk Péter és szerzőtársai cikkükben az ennek a célnak a megvalósítása érdekében végzett munkákról és tervekről számolnak be.
2. Az ELKH ARP-adatrepozitórium létrehozása (a már működő, CONCORDA-adatrepozitórium továbbfejlesztésével²), és ezáltal egy FAIR-elvű adatrepozitórium megteremtése. A fejlesztés alatt lévő infrastruktúráról Kovács László cikkéből tájékozódhatunk.
3. A korszerű adatrepozitórium-technikák és technológiák megismertetése a hazai kutatókkal és általában, a FAIR-elvű adatkezelési kultúra terjesztése és meghonosítása az ELKH kutatói körében. Egyed-Gergely Júlia és szerzőtársai cikkében olvashatunk az ELKH ARP keretében történt kutatói felmérés legfontosabb tapasztalatairól.
4. Az, hogy minél könnyebben hozzáférhetővé váljanak a kutatási eredmények (kutatási adatok és az azokból írott publikációk), az EU egyik fontos prioritása. Többek között emiatt is egyre erősödnek Európában az adatpublikációs elvárások. A kutatóhálózatban használt legismertebb folyóiratok adatpublikációs elvárásairól és a gyakorlat néhány nemzetközi példájáról Kutnyánszky Anikó és szerzőtársai írásából tájékozódhatunk.

² A CONCORDA repozitóriumkezelő szoftverrendszernek jelenleg 51 felhasználója van, akik összesen 2,5 TB tárhelyet használnak, kb. 7300 fájlal. Adatcsomagonként egy-egy DOI/handle került alkalmazásra. Jelenleg 1199 adatcsomagot használnak, ebből 1180 nyilvános.