



Cyberűr: Adatok birodalma

Az adatgazdászok vezetik a Digitális Univerzumot

Farkas Henrietta
HUN-REN Wigner Adatközpont
farkas.henrietta@wigner.hun-ren.hu

Meiszterics Enikő
HUN-REN TK KDK
Meiszterics.Eniko@tk.hun-ren.hu

A digitális világ adatmennyisége és növekedése

- **Üzleti élet:** Web, e-kereskedelem, pénzügyi tranzakciók, tőzsde
- **Közösségi és mindennapos élet:** Facebook, Instagram, hírek, Youtube, digitális kamerák, okostelefonok és eszközök
- **Tudomány:** Távérzékelő rendszerek, bioinformatika, tudományos szimulációk
- **Automatikus adatgyűjtő mérőeszközök, adatbázisrendszerek, közösségi hálók, számítógépes ügyfélszolgálatok**
- **Mesterséges intelligencia és gépi tanulás**



Mi is az a Cyberűr?

- A "**Cyberűr**" egy képzelt kifejezés, amely az adatok virtuális világát és az abban rejlő végtelen lehetőségeket jelöli.
- A "**Cyberűr**" összekapcsolja a "**cyber**" (kibertér) és az "**űr**" (univerzum) szavakat, hangsúlyozva a digitális tér végtelen és felfedezésre váró jellegét.



Hát az „Adatok Birodalma”?

- Az **Adatok Birodalma** egy olyan metafora, amely rámutat arra, hogy az adatok terjedelmes és összetett világa kiterjed az emberi gondolkodás határain túlra, és számtalan lehetőséget rejt magában a felfedezésre és az innovációra.
- A "**Cyberűr: Adatok Birodalma**" című prezentáció ezt a képzelt világot és annak jelentőségét hivatott bemutatni az adatgazdászok szerepének és tevékenységének kontextusában.





Adatkezelési kihívások és megoldások

Növekvő adat mennyiség, növekvő nyílt tudományos elvárások

Egyre több és egyre gyorsabban növekvő mennyiségű – döntően digitális – adat keletkezik a tudományos kutatások során is, ezeket a kutatás alatt és után is kezelni kell.

A kutatási adatok mennyiségének növekedésével párhuzamosan mind kutatásfinanszírozói mind kiadói oldalról egyaránt növekszik az igény a publikációk mellett a kutatások eredményeit alátámasztó kutatási adatok megtalálhatóvá és elérhetővé tételére is.

Tudatos, FAIR adatkezelés – de hogyan?

A kutatási-adat kezelés az adatokkal kapcsolatos döntések és tevékenységek összessége. A kutatóknak válaszolni kell azokra a kérdésekre, hogy

- a kutatás során
 - hol és hogy tárolják az adatokat?
 - kivel, hol és milyen feltételekkel osszák meg az azokat?
 - milyen metaadatokat társítsanak a kutatási adatokhoz?
- a kutatás lezárása után
 - mely adatokat archiválják?
 - milyen fájlformátumokat használjanak az archiváláshoz?
 - hol tárolják az archivált adatokat?
 - kivel és hol és milyen feltételekkel osszák meg azokat?



Akadályok a megoldásban

A tudatos, FAIR elveknek megfelelő adatkezelés megvalósítását

- kapcsolódó ismeretek
- az iránymutatások, szabályozások,
- az erre fordítható plusz idő és
- a feladatok elvégzését támogató munkatárs hiánya egyaránt akadályozhatja.

Ezeknek a hiányoknak a csökkentésre, megszüntetésére megoldás lehet a kutatókat az adatkezelés területén támogató adatgazdász, adatgazdászok alkalmazása.





Cyberkalandok: Az adatgazdászok a Kutatás tengerén

Navigálás a kutatási adatok tengerén

A főbb kikötők:

- Adatkezelés és tárolás
- Adatbiztonság és adatvédelem
- Adatminőség és adattisztítás
- Adatkezelési irányelvek betartása
- Adatelérhetőség és -megosztás



Az adatgazdászok útja

Adatkezelés és tárolás

- A kutatócsoportok adatkezelési szokásának felmérése:
 - Adatgyűjtési gyakorlatok
 - Adattárolási módszerek
 - Adatelérés
 - Adatkezelési irányelvek és szabályozások betartása
 - Adatmegosztási szokások
- Adatkezelési stratégia kidolgozása
 - Munkafolyamatok folyamatos dokumentálása
 - Mappastruktúrák, névkonvenciók kialakítása
 - Metaadatok összegyűjtése, metaadat sémák
 - Adatok előkészítése archiváláshoz (HUN-REN ARP megismertetése a kutatókkal)
- Oktatások, segédanyagok készítése



Az adatgazdászok útja

Adatbiztonság és adatvédelem

- Jogosultságok kezelése
- Adatok titkosítása, anonimizálása
- Biztonsági mentések és visszaállítási folyamatok
- Felhasználói oktatás és tudatosság növelése
 - Aktuális fenyegetések
 - Adatvédelmi tréning



Az adatgazdászok útja

Adatminőség és adattisztítás

- Adatellenőrzés: Adatok strukturáltságának ellenőrzése, az érvényes értékek ellenőrzése, valamint a hiányzó vagy hibás adatok azonosítása és korrigálása
- Adat átalakítás és normalizálás: Egységes formátum és szerkezet kialakítása, kategóriák és címkék normalizálása
- Adattisztítás: Hibás vagy nem reprezentatív értékek kezelésére javaslattétel
- Adatforrások integrációja: duplikátumok azonosítását és kezelését



Az adatgazdászok útja

Adatkezelési irányelvek betartása

- Iránymutatások kidolgozása (ARP portálon)
- Képzés és tudatosság növelése:
 - Adatkezeléssel kapcsolatos legfrissebb információk megosztása
 - Az adatbiztonsági gyakorlatok bemutatása
- Támogatás és tanácsadás
 - Adatkezelési eljárások javítására vonatkozó javaslatok és gyakorlati alkalmazásának támogatása



Az adatgazdászok útja

Adatelérhetőség és -megosztás

- Adatmegosztási irányelvek kidolgozása
- Adatmegosztási folyamatok támogatása
 - Adatok feltöltése, letöltése
 - Adatmegosztási formátumok
 - Repozitórium használatával kapcsolatos ismeretek terjesztése
- Hozzáférési jogosultságok kezelése
 - Platformon a hozzáférési jogosultságok beállítása
- Adatmegosztásra vonatkozó politikák és szabályok betartása
 - Az adatvédelmi és személyes adatokra vonatkozó irányelvek





Jövőbeli kilátások és lehetőségek

Fő trendek

Növekedés



dreamstime.com

ID 112740655 © Savano

Hálózatosodás



Specializáció



dreamstime.com

ID 281360897 © Adona1969

Növekedés

Növekszik az adatok mennyisége.

Növekednek a finanszírozók és a kiadók nyílt adat publikációs elvárásai.

Nagyobb lesz az adatkezelésre, megosztásra fordítható összeg.

Nő az adatgazdászok száma, egyre több intézményben lesz adatgazdász.

Több elérhető képzési lehetőség lesz, kutatóknak, adatgazdászoknak egyaránt.

Nő a témához kapcsolódó formális képzések száma is.



Hálózatosodás

A kapcsolódó képzések és az adatgazdászok számának növekedése, az adatgazdászok mint szakma elfogadása ezt eleve magával hozza.

A Horizone Europe által támogatott Skills4EOSC projektben 2022. szeptember 1-óta dolgoznak azon, hogy kompetenciaközpontok páneurópai hálózatát hozzák létre.

Nemzeti kompetencia központokat szeretnének létrehozni, céljuk az új adatgazdász hálózatok harmonizálása és létrehozásának ösztönzése. 18 országban 44 partnerük van jelenleg.



Specializáció

Az adatgazdász funkciók elválnak, specializálódnak az adatgazdászok.

Az oktatás is új lehetőségeket nyújt, több és többféle képzés lesz, ez is támogatja a specializációt.





Köszönjük a figyelmet!

Elérhetőségeink:

farkas.henrietta@wigner.hun-ren.hu

Meiszterics.Eniko@tk.hun-ren.hu