



Az ELKH kutatók igényeinek felmérése – kérdőíves kutatás és annak eredményei

Egyed-Gergely Júlia
Társadalomtudományi Kutatóközpont
Kutatási Dokumentációs Központ

Az adatfelvételtől

- A felmérés célja:
 - Közös adatrepozitóriummal kapcsolatos kutatói igények megismerése
 - Adatkezeléssel kapcsolatos ismeretek, tapasztalatok, gyakorlatok, felmerülő problémák feltérképezése
- A felmérés módszere:
 - Online kérdőíves felmérés, önkéntes válaszadás, anonim kérdőív
- A kutatás készítői:
 - TK KDK – a SZTAKI és a Wigner FK támogatásával
- Az adatfelvétel időszaka:
 - 2022. február 1-március 31.

Kutatók elérése

- Megkérdezettek:
 - ELKH kutatók
 - Lehetőleg minden ELKH intézményből
- Megkeresés:
 - Intézetigazgatókon, intézeti titkárságokon, a korábbi (2020-as) ELKH-HRDA kutatásban megadott intézeti kapcsolattartókon keresztül
- Elérés:
 - 129 kutató
 - 17 intézmény (2 kitöltő nem kívánta megnevezni intézményét)

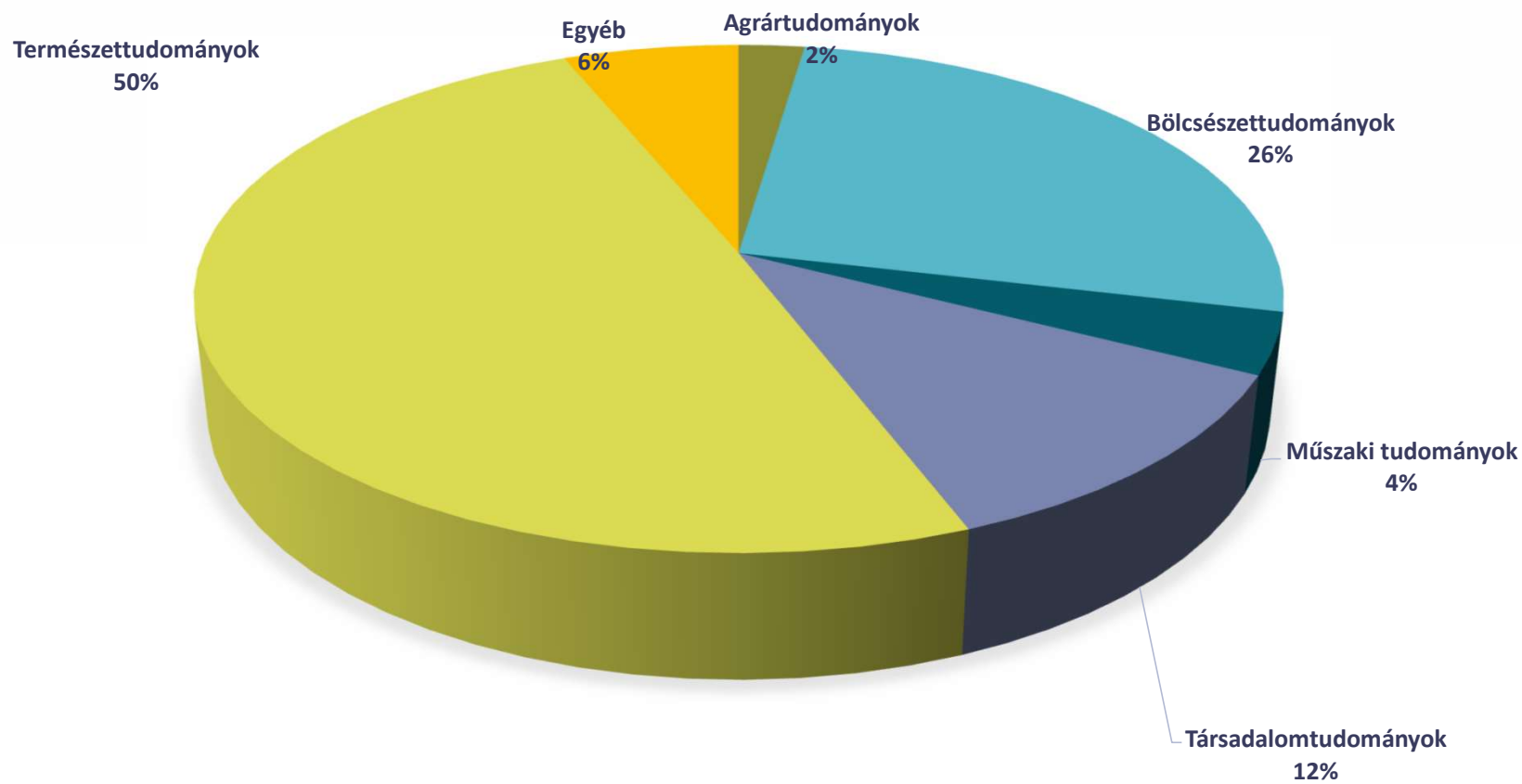
Elért kutatók

- Kitöltők: elsősorban olyan kutatók, kutatócsoportok, akik
 - az adatkezelés, adattárolás területén tájékozottabbak, vagy
 - aktívan szeretnék segíteni, hogy az ARP a felhasználói igényeknek minél megfelelőbb repozitóriumot hozzon létre
- Ezért:
 - A minta nem reprezentatív
 - A feltárt adatok nem általános képet mutatnak, a téma iránt érdeklődő kutatói szegmens adatelőállítási, -kezelési, -megosztási gyakorlatát, ismereteit, felmerülő igényeit, elvárásait segítenek megismerni

Vizsgált kérdéskörök

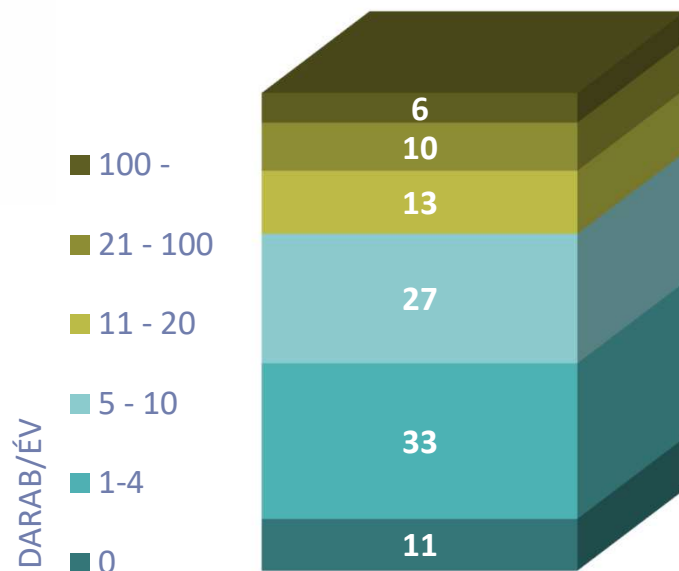
- A kérdőív fő kérdéskörei:
 - Tudományterület (1-5)
 - Kutatási adatok előállítása és tárolása (6-19)
 - Kutatási adatok megosztása (20-28)
 - Adatelérés (29-30)
 - Metaadatok (az adatokat leíró adatok) használata (31-35)
 - A kutatási adatokkal kapcsolatos jogok ismerete (36-42)
 - Kutatási adatok kezelésével kapcsolatos humán infrastruktúra (43-45)
 - FAIR adatkezelési elvek, adatrepozitóriumok, állandó egyedi azonosítók ismertsége (47-60)

Válaszadók tudományterület szerint

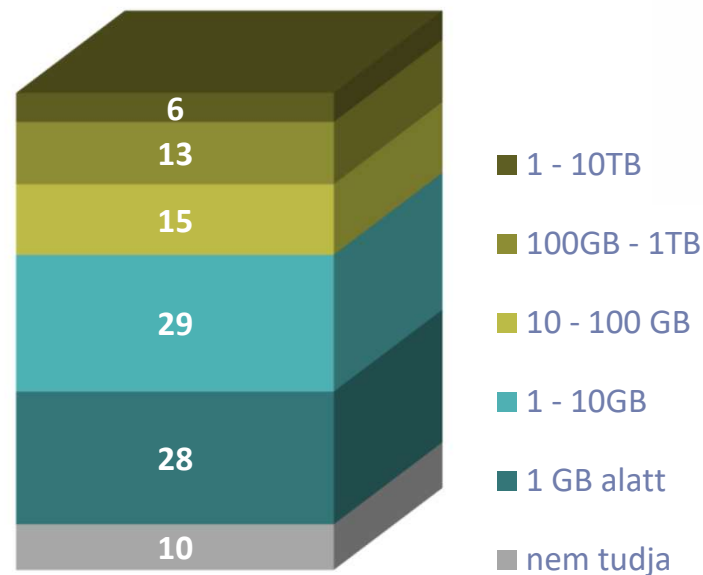


Adatcsomagok* – éves szám és méret

Adatcsomagok száma (%)



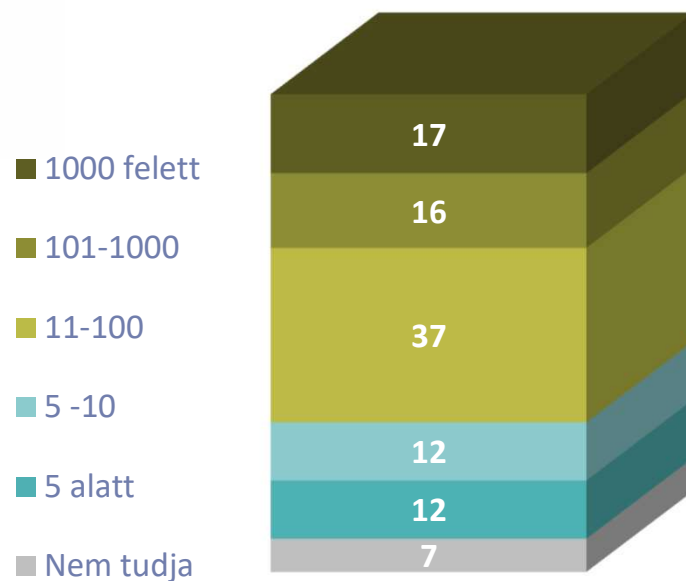
Adatcsomagok mérete (%)



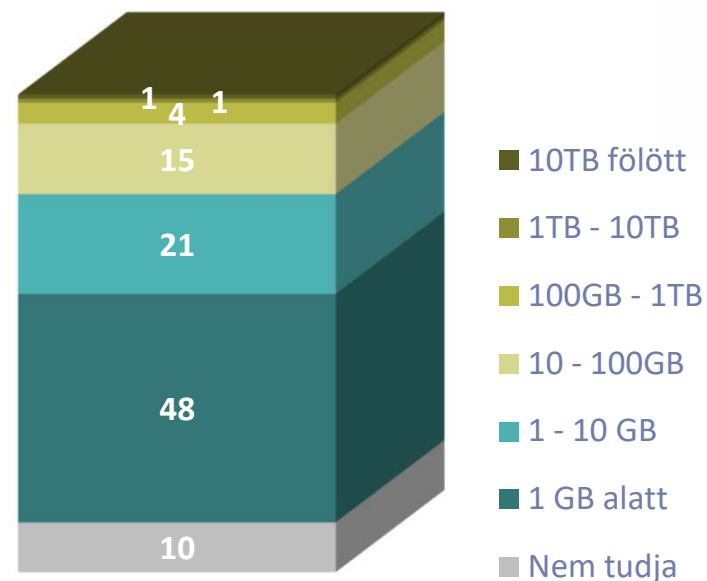
*Adatcsomag = adott kutatásban logikailag egybetartozó és egyben kezelt kutatási adatok, szöveges dokumentumok, mérési eredmények

Adatcsomagok – fájlok száma és mérete

Fájlok maximális száma az adatcsomagban (%)

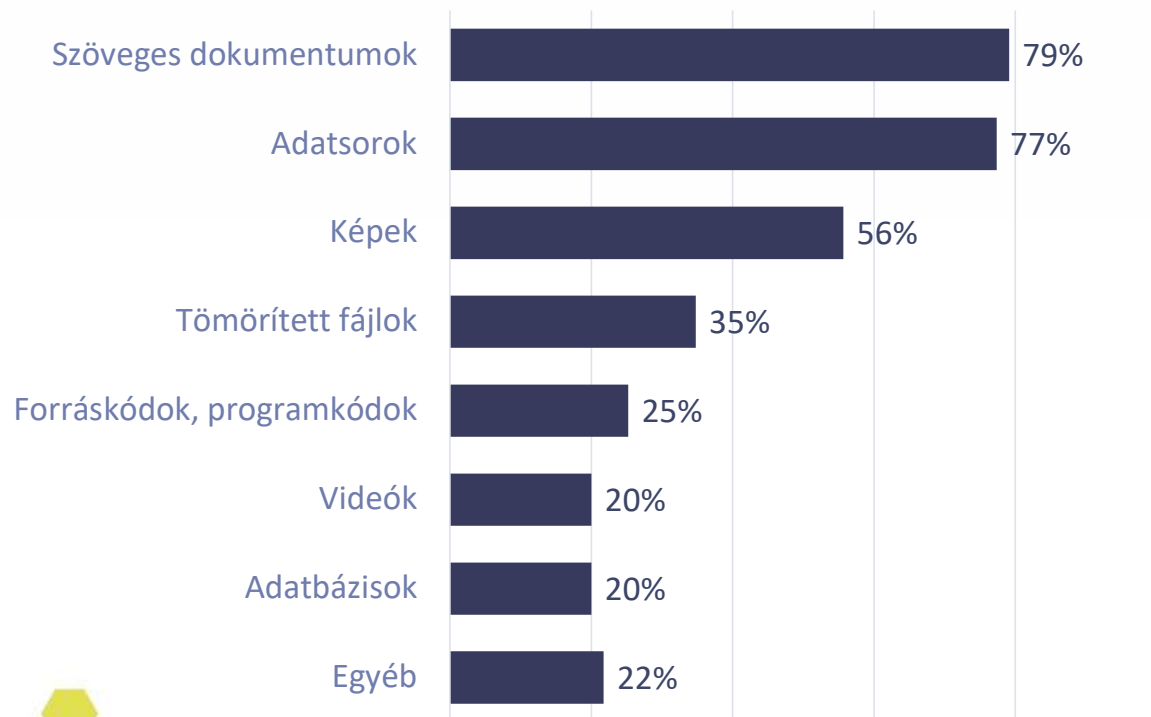


Legnagyobb fájl maximális mérete (%)



Adatcsomagok - fájlok típusa

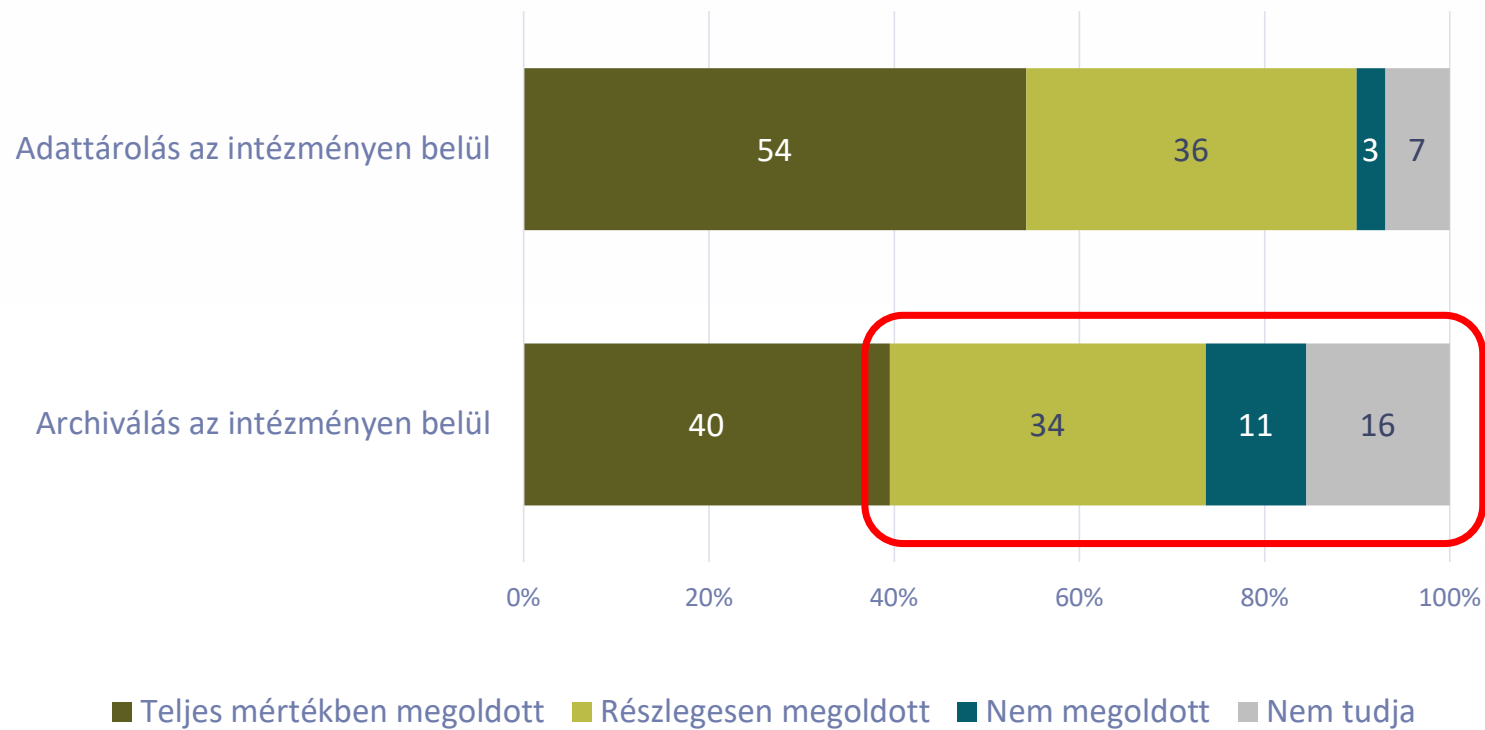
Az adatcsomagban található fájlok típusa



- Szöveges dokumentumok: pl. doc, txt
- Adatsorok: pl. CSV, TSV, Excel
- Adatbázisok: pl. SQL, SPSS
- Tömörített fájlok: pl. zip, tar.gz
- Képek: pl. PNG, JPG
- Videók: pl. MPEG, WMV
- Forráskódok/programkódok
- Egyéb említések: cha, xml, gbk, fasta, shp, kml, kmz, dxf, dwg, psd, ai, axml, spec, mca, spc, fastq, sam, bam, bai, fai, bed, plink, geno, FITS, gexf, TEI XML, HDF5, HEPMC, lmf, RAW, WIFF, prb, prm, Rdata, tex, json, zip, TextGrid, trs, TIF, ND2, AVI, MOV, mgf, dat, baf

Fájltípusok átlagos száma / adatcsomag: 3,3

Adatok tárolása – kutatás közben és után



Adatok tárolása - problémák

Kutatás közben

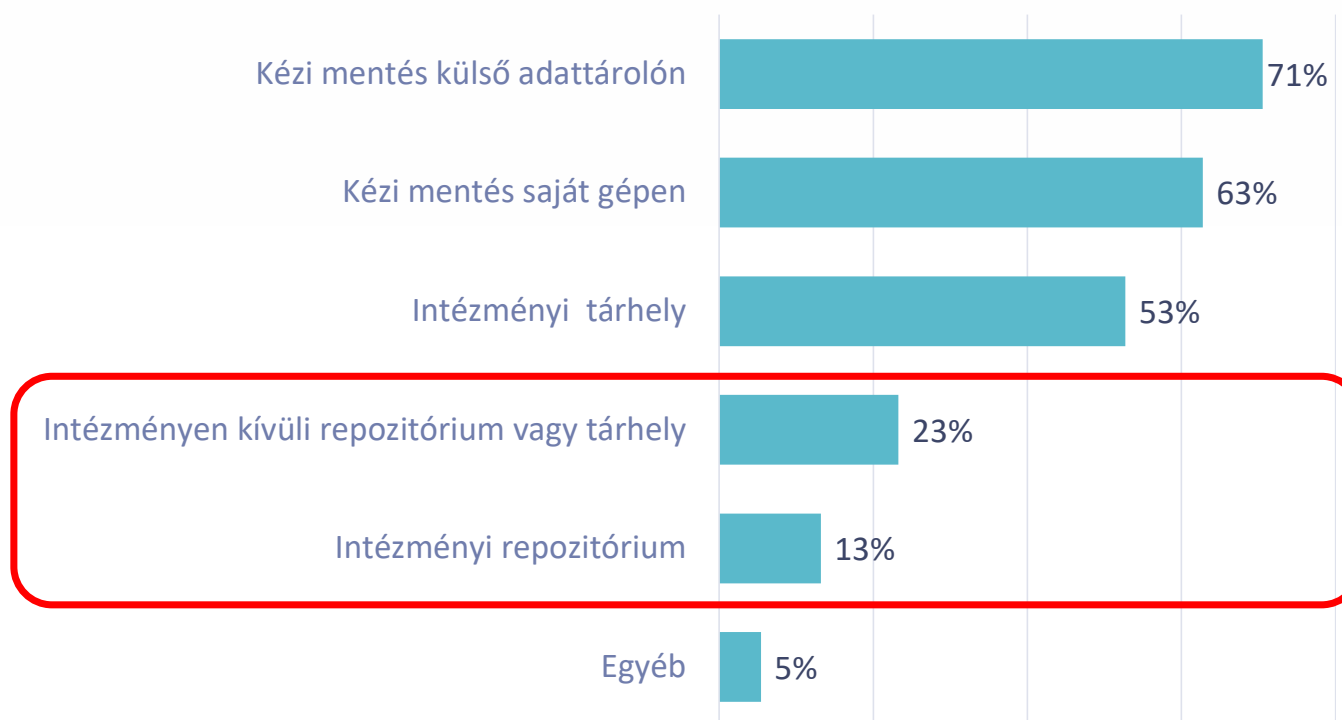
- Kapacitásbeli
*„A külföldi kutatási infrastruktúráknál mért nagy mennyiségű adatokat nem tudjuk hazahozni, mert nem tudja az intézet fogadni”
„A szerver teljesen tele van”
„Folyamatosan szükséges újabb tárhelyek vásárlása”*
- Szervezeti, szervezési
*„Mindenki tárolja valahogy az adatait, de ennek nincs egységes változata”
„A kapott tájékoztatás nem ébresztett bizalmat a szolgáltatás iránt”*

Kutatás után

- Kapacitásbeli
*„tárolóhely hiánya”
„szűkös kapacitás”*
- Szervezeti, szervezési
*„Nincs egységesített gyűjtőhely”
„Van megfelelő tárolóhely, de nincs olyan automatizmus, ami jól bevált”
„Nincs egységes eljárásrend, minden csoport megoldja magának”*
- Humán infrastruktúra hiánya
„az adatok megőrzése megfelelő személyzetet igényelne”

Adatok archiválása a kutatás lezárása után

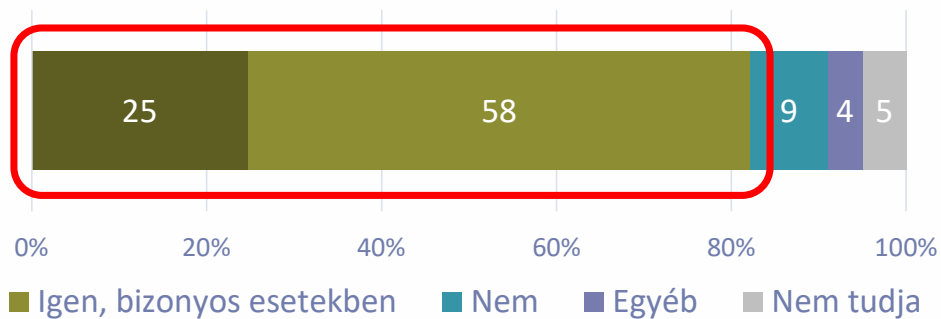
Mentés a kutatás lezárulta után



Mentésfajták átlagos száma: 2,3

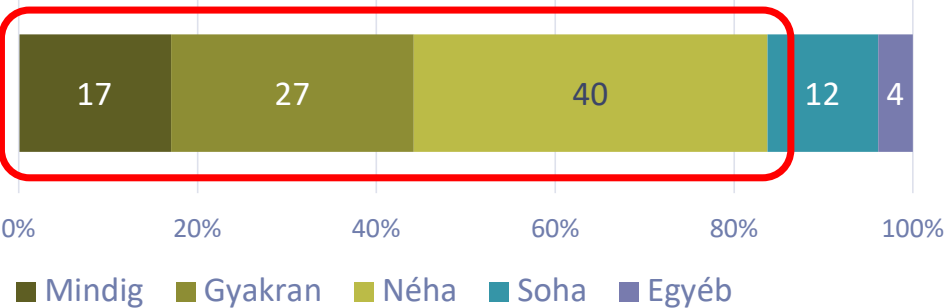
Kutatási adatok közzététele, megosztása

Az adatok közzététele tudományterületi elvárás



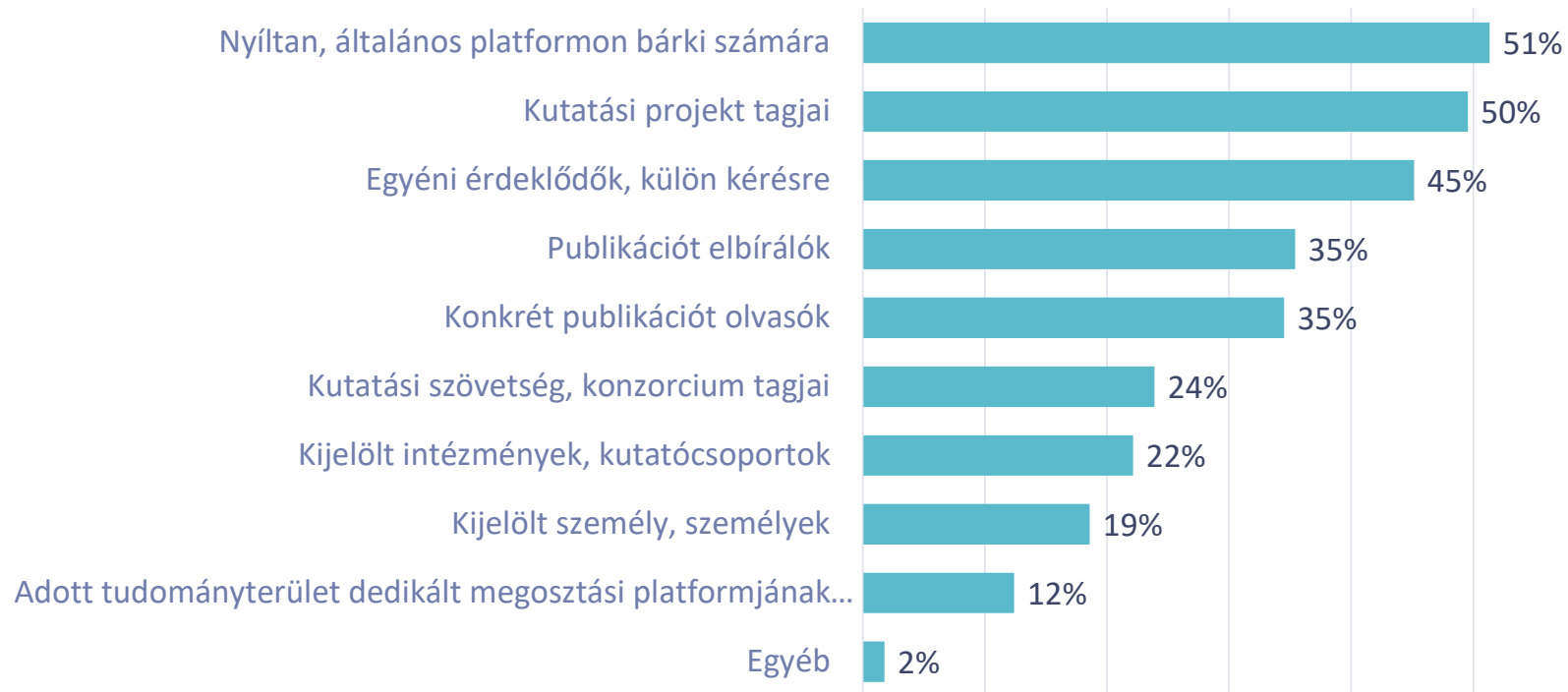
Adatmegosztás a kutatás közben

Adatmegosztás a kutatás lezárását követően



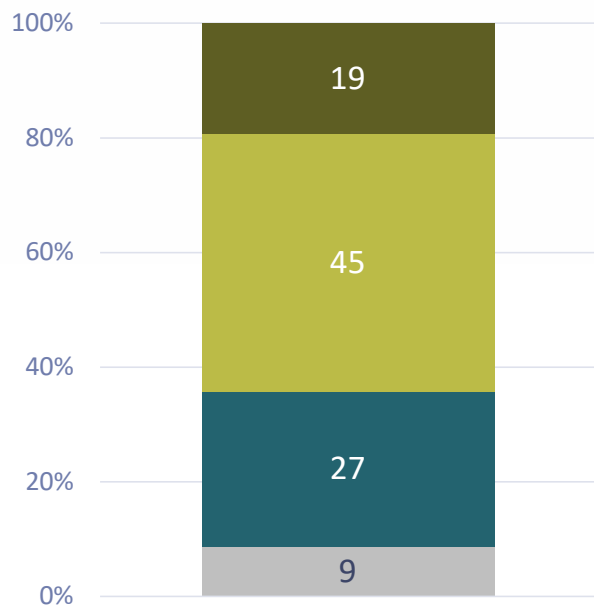
Kutatási adatokhoz hozzáférők köre

Akik számára hozzáférhetővé teszi az adatokat a kutatás után – azok között, akik legalább néha megosztják adataikat



Megosztás utáni igények

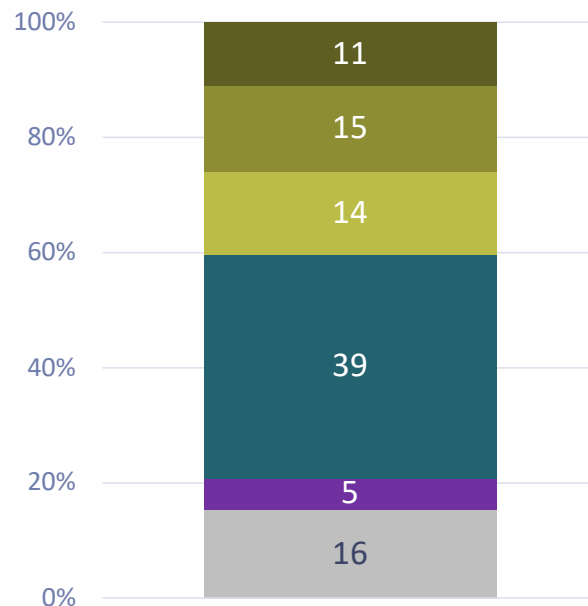
Adatcsomag bővítése megosztás után



- Nem tudja
- Nem
- Alkalmanként
- Gyakran

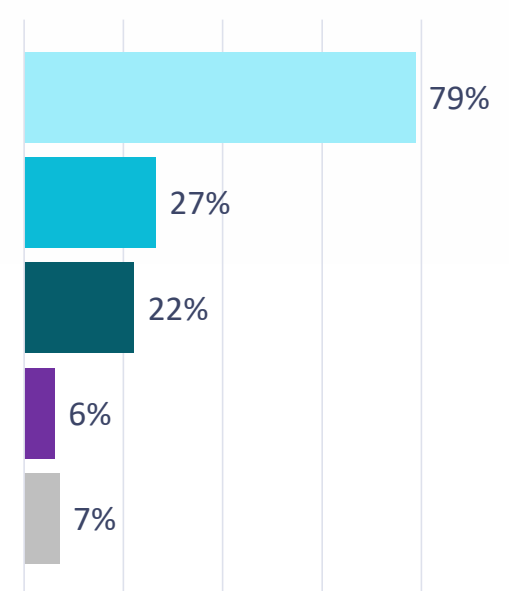
15

Hozzáférési statisztikák elérése



- Minden egyes letöltésről
- Időszakos bontásban
- Összesítő számláló
- Nincs rá szükség
- Egyéb
- Nem tudja

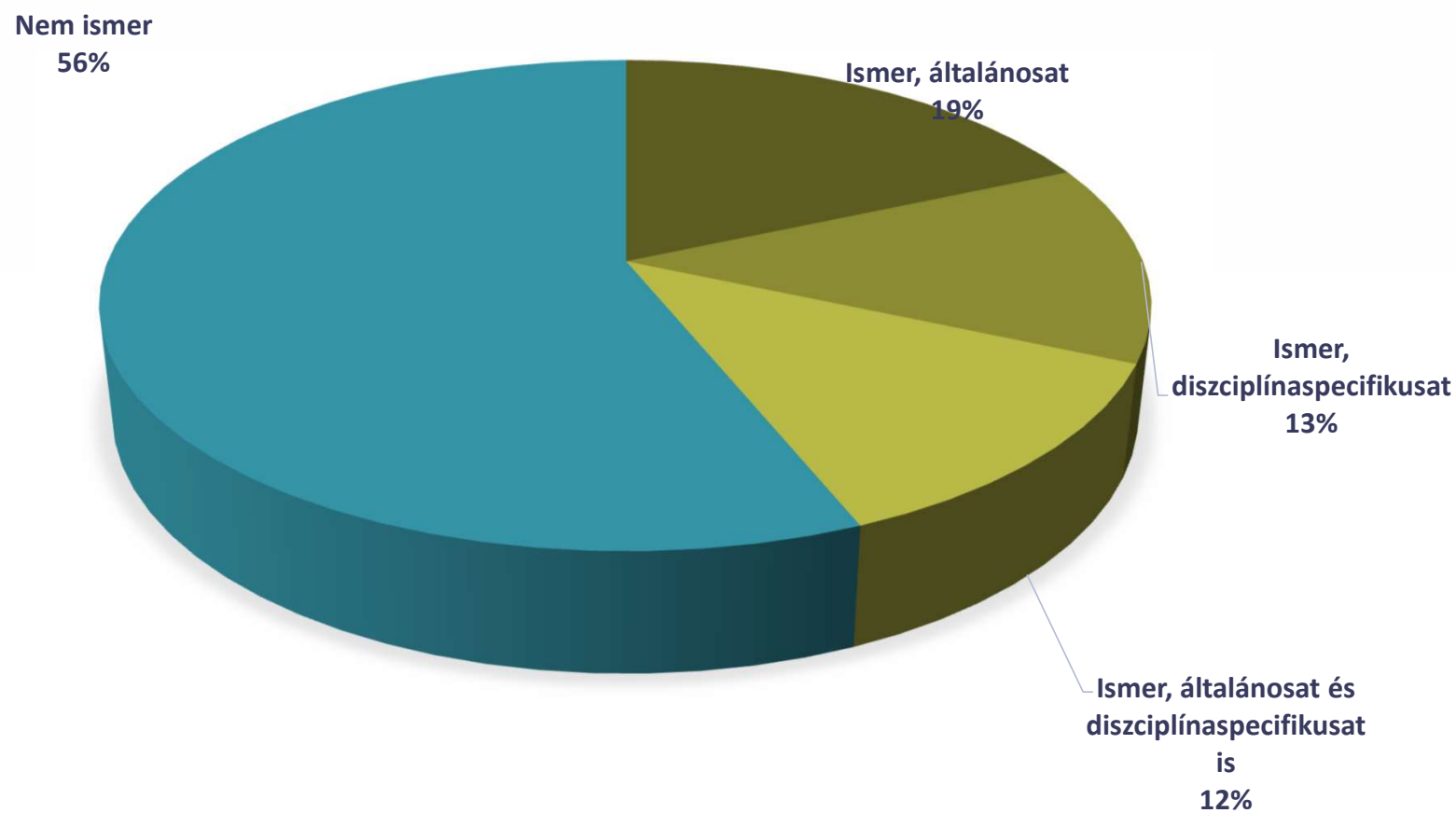
Hozzáférési mód



- Olvásás
- Olvásás és módosítás
- Olvásás, módosítás és kiegészítés
- Egyéb
- Nem tudja



Igénybe vehető repositóriumok ismerete



Repozitóriummal szembeni elvárások

Mennyire fontos egy repositórium esetében? (átlagértékek 1-5 skálán)



Összegzés – Kutatói javaslatok

Repozitóriumhasználat – konkrét igények, javaslatok

- LEGYEN KÖNNYEN KEZELHETŐ:
 - ✓ egyszerűség, egyértelműség, kereshetőség, minél kevesebb extra adminisztráció
- LEGYEN MELLÉ INTÉZMÉNYI TÁMOGATÁS:
 - ✓ az intézmény vezetése támogassa, segítse, jutalmazza a kutatási adatok archiválását és nyílt közzétételét
 - ✓ biztosítson hozzá dedikált személyzetet (intézménykönyvtáros / könyvtárinformatikus / kutatástámogató)
 - ✓ legyenek illeszkedő pályázatok
- LEGYEN HOZZÁ KÖVETHETŐ SZABÁLYOZÁS/AJÁNLÁS:
 - ✓ intézményi iránymutatás



Köszönöm a figyelmet