



Az ELKH kutatók igényeinek felmérése adatrepozitóriumok létrehozására

Társadalomtudományi Kutatóközpont (TK)
Kutatási Dokumentációs Központ (KDK)

Kiinduló helyzet és igényfelmérés

Annak érdekében, hogy a létrehozandó adatrepozitóriumot minél inkább a kutatói igényekre szabhassuk az ELKH kutatóinak adatkezeléssel kapcsolatos ismereteit, tapasztalatait, igényeit egy komplex, az elsődleges és a másodlagos kutatás eszköztárát egyaránt használó felmérés keretében térképezzük/térképeztük fel.

Ennek lépései:

- a 2021. áprilisában végzett ELKH–HRDA online kérdőíves kutatói és intézményi felmérések adatainak másodelemzése (lezárult)
- online kérdőíves felmérés az ELKH intézmények kutatóinak körében (lezárult)



Kiinduló helyzet és igényfelmérés

- interjú adatfelvétel az ELKH intézmények kutatóinak körében (folyamatban)
- az adatfelvételekben megnevezett repozitóriumok, használt sémák, tudományterület specifikus szoftverek, elemzési módszerek kigyűjtése, linkgyűjtemény és információ bázis létrehozása (folyamatos).



Az elsődleges felmérések tapasztalatai

Az elsődleges adatfelvételekről

A cél az volt, hogy az általunk szervezett elsődleges adatfelvételek – online kérdőíves megkérdezés és interjúzás – keretében az ELKH minden intézményének kutatóit, kutatócsoportjait elérjük legalább az egyik, de inkább mind a két módszerrel, ez teljesült.

Az online kérdőívet 17 intézményben töltötték ki, 129 kutató illetve kutatócsoport adatkezelési gyakorlatáról kaptunk információkat.

A már elkészült 48 és további 2 egyeztetett interjú válaszadói 16 intézményt képviselnek.





Online kérdőíves adatfelvétel

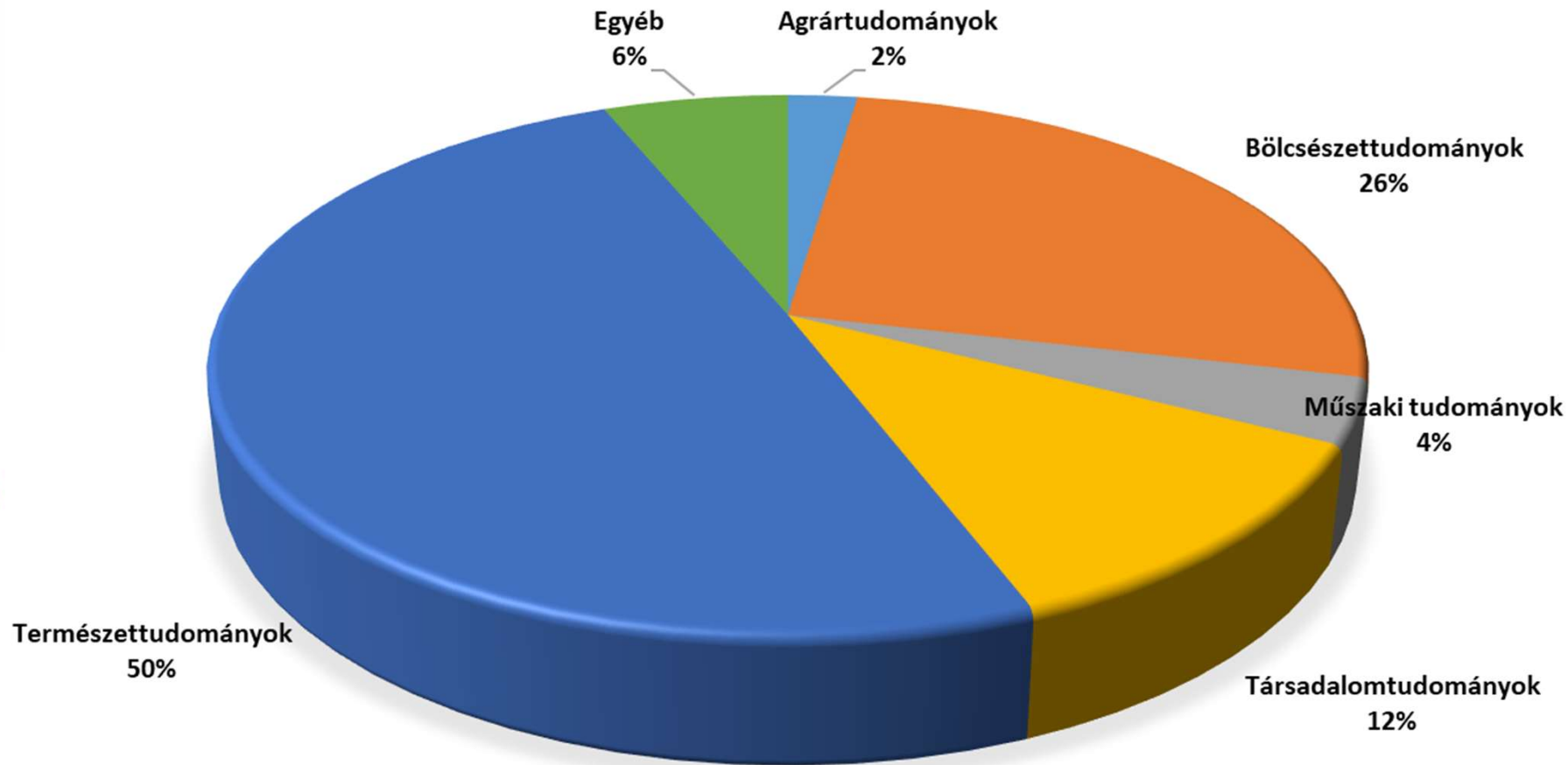
Vizsgált kérdéskörök

A kérdőív a következő kérdésköröket vizsgálta részletesen

- kutatási adatok előállítása és tárolása (6-19)
- kutatási adatok megosztása(20-28)
- adatelérés (29-30)
- metaadatok (az adatokat leíró adatok)(31-35)
- a kutatási adatokkal kapcsolatos jogok (36-42)
- kapcsolódó humán infrastruktúra (43-45)
- FAIR adatkezelési elvek a kutatói gyakorlatban (47-60)

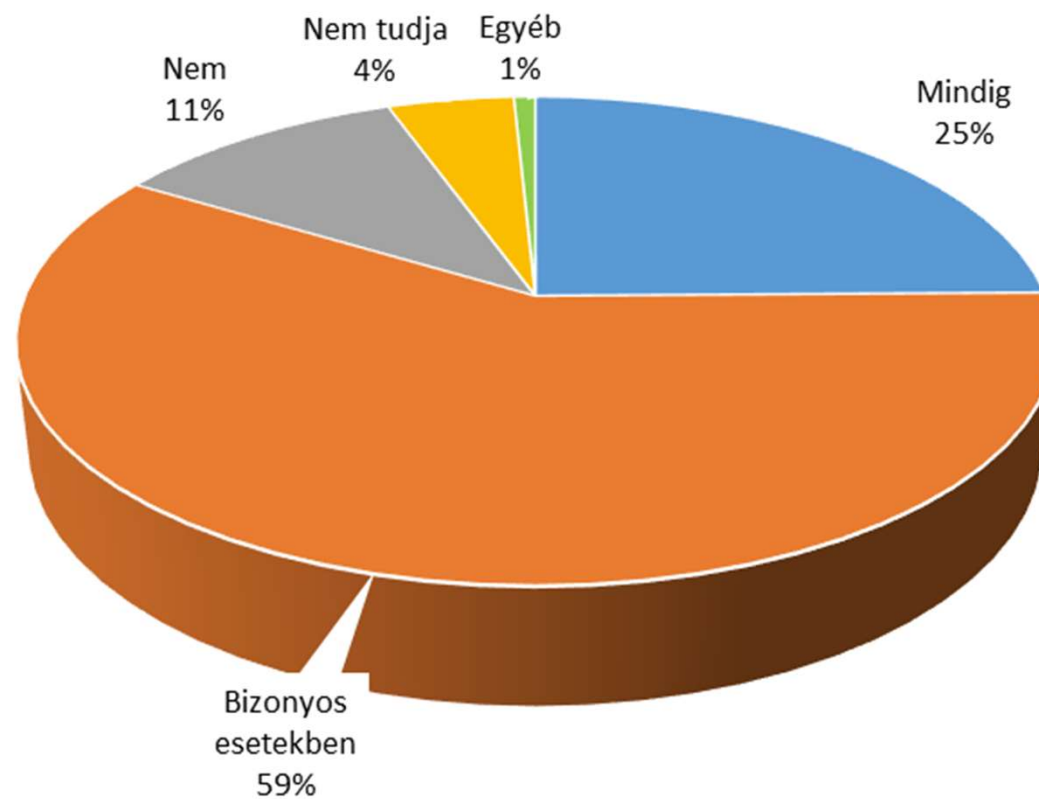


Válaszadók tudományterület szerinti megoszlása

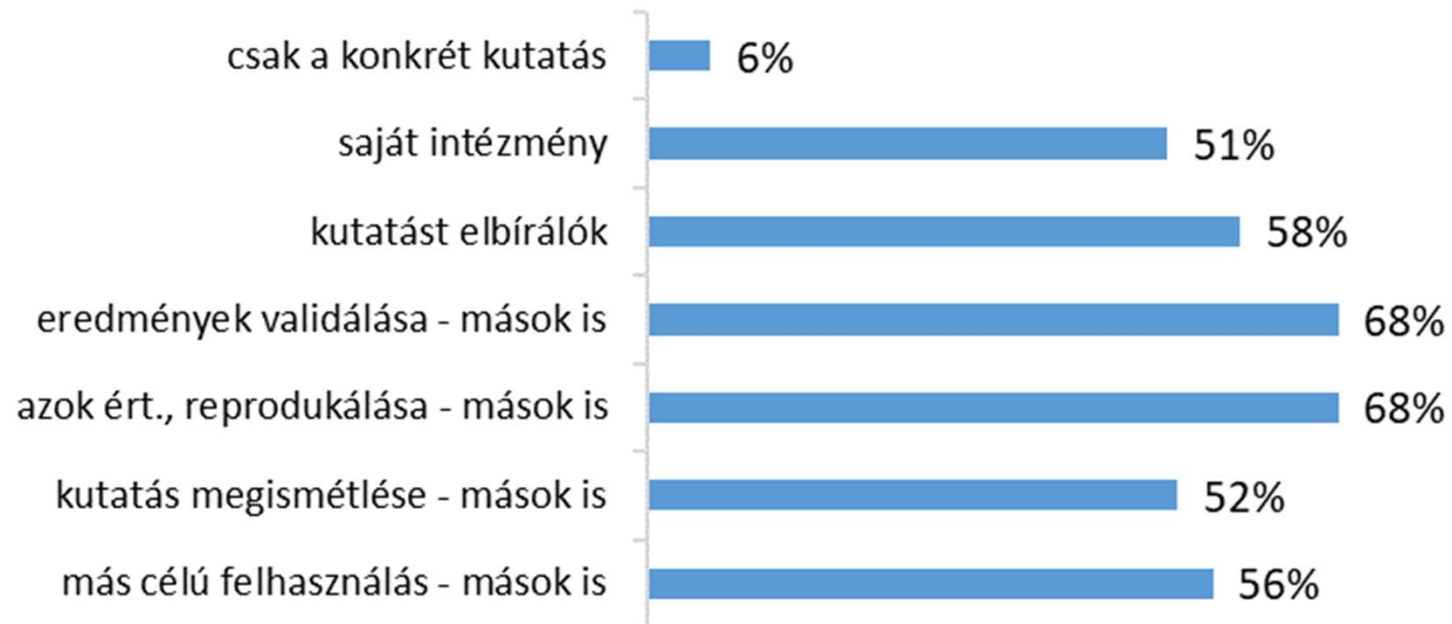


Adatok közzététele

Tudományterületi elvárás az adatok közzététele (%)

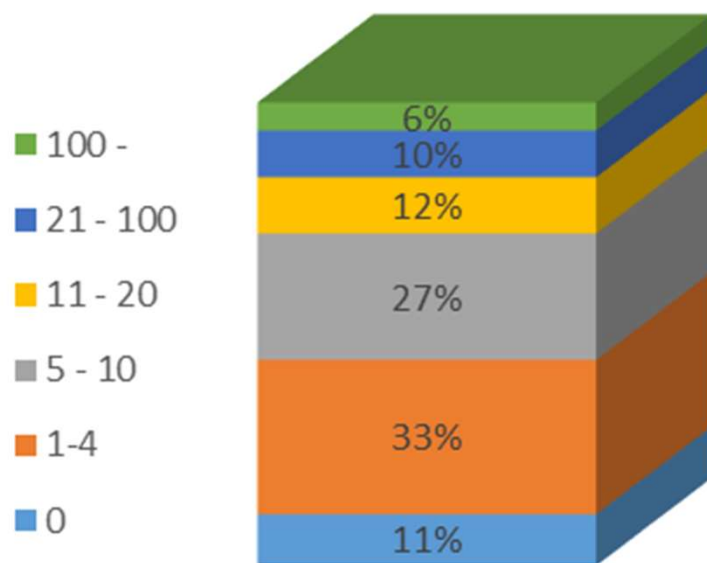


Adatok relevanciája

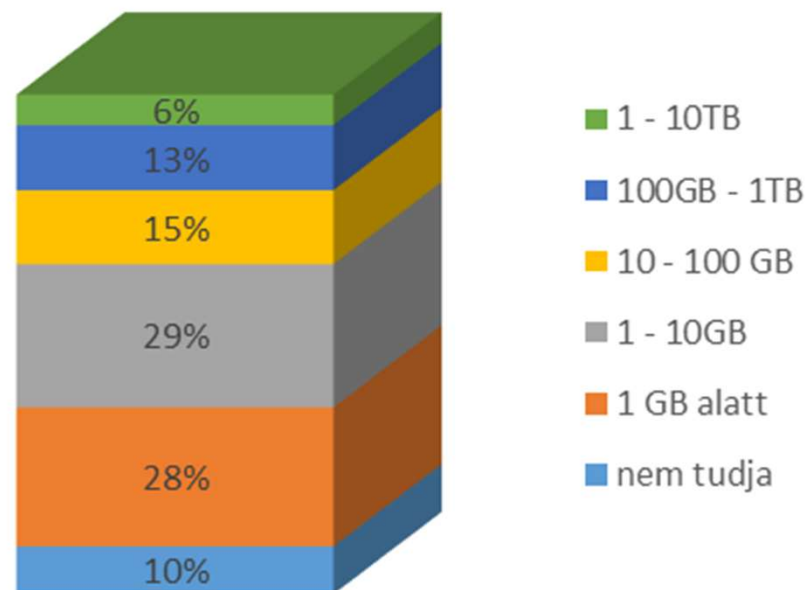


Adatcsomagok - szám és méret

Adatcsomagok száma (db/év)



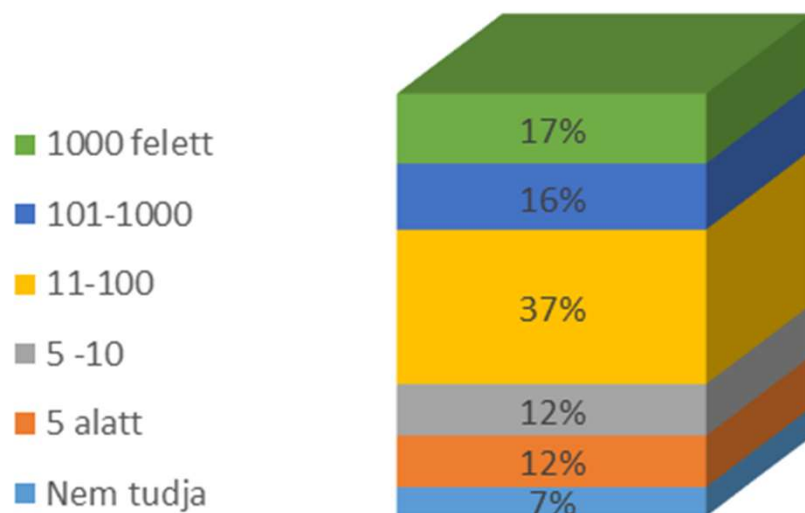
Adatcsomagok mérete



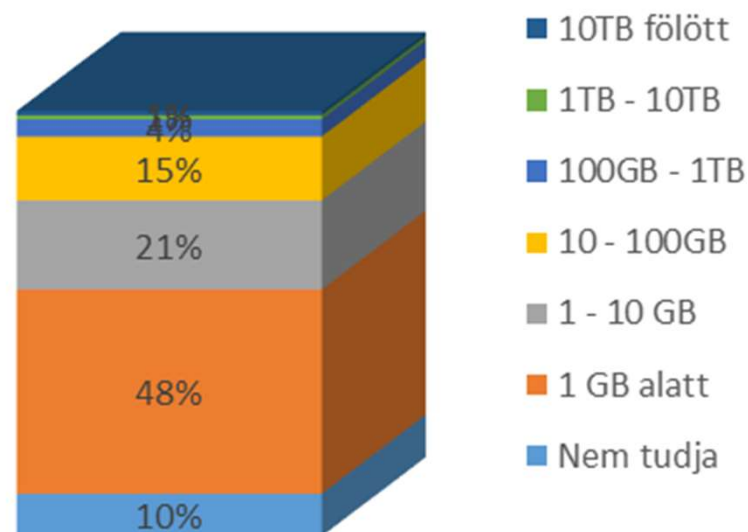
* Adatcsomag = adott kutatásban logikailag egybetartozó és egyben kezelt kutatási adatok, szöveges dokumentumok, mérési eredmények

Adatcsomagok* - fájlok száma és mérete

Fájlok maximális száma az adatcsomagban

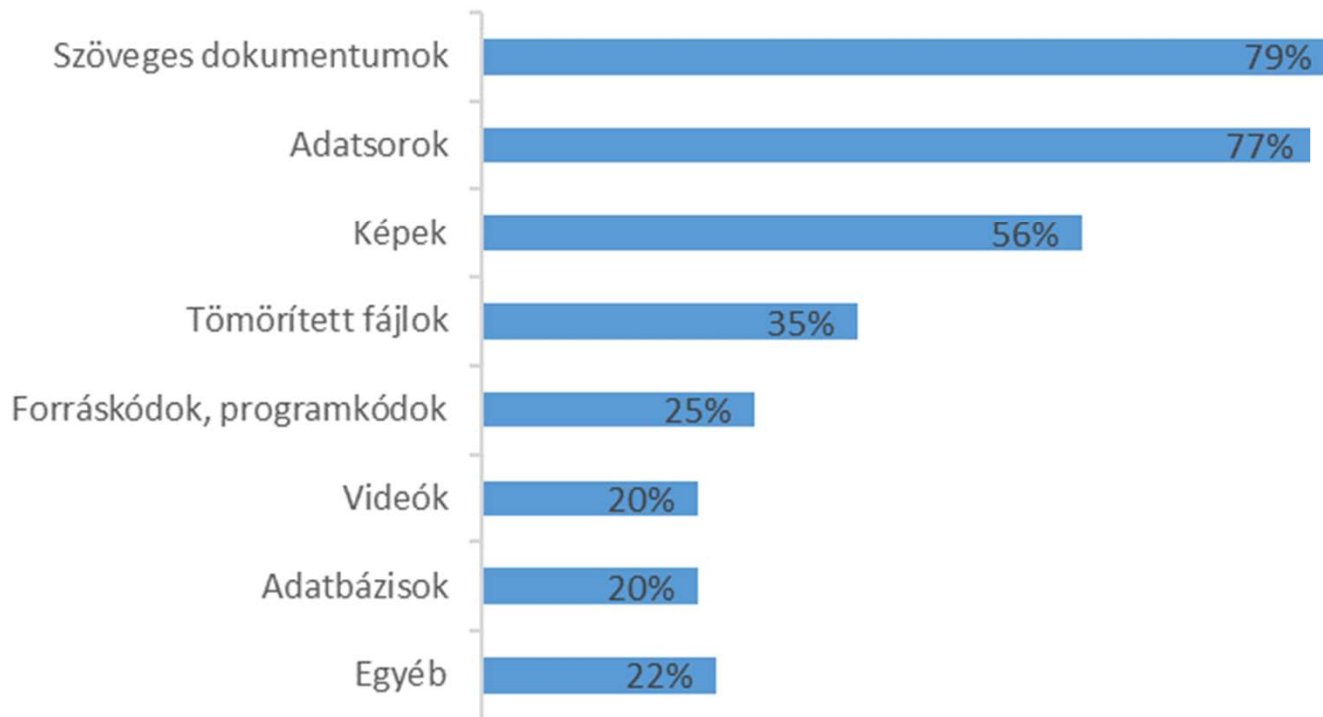


Legnagyobb file maximális mérete

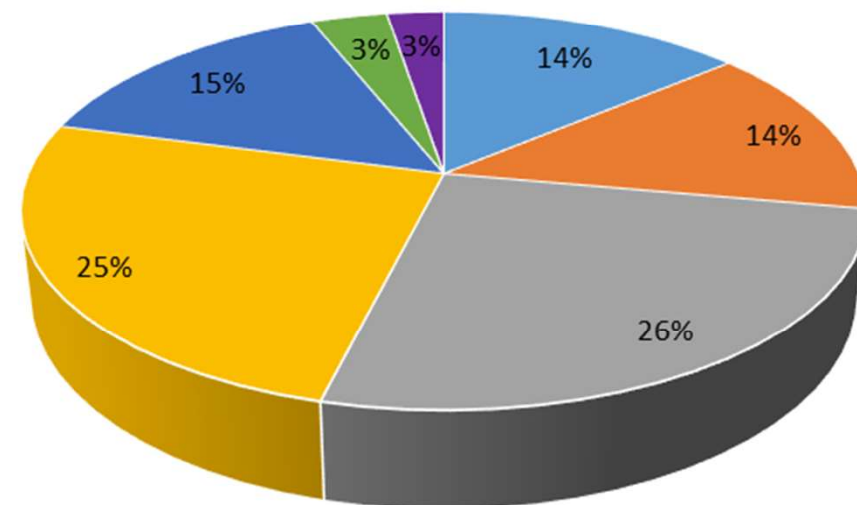


Adatcsomagok - fájlok típusa

Az adatcsomagokban található fájlok típusa

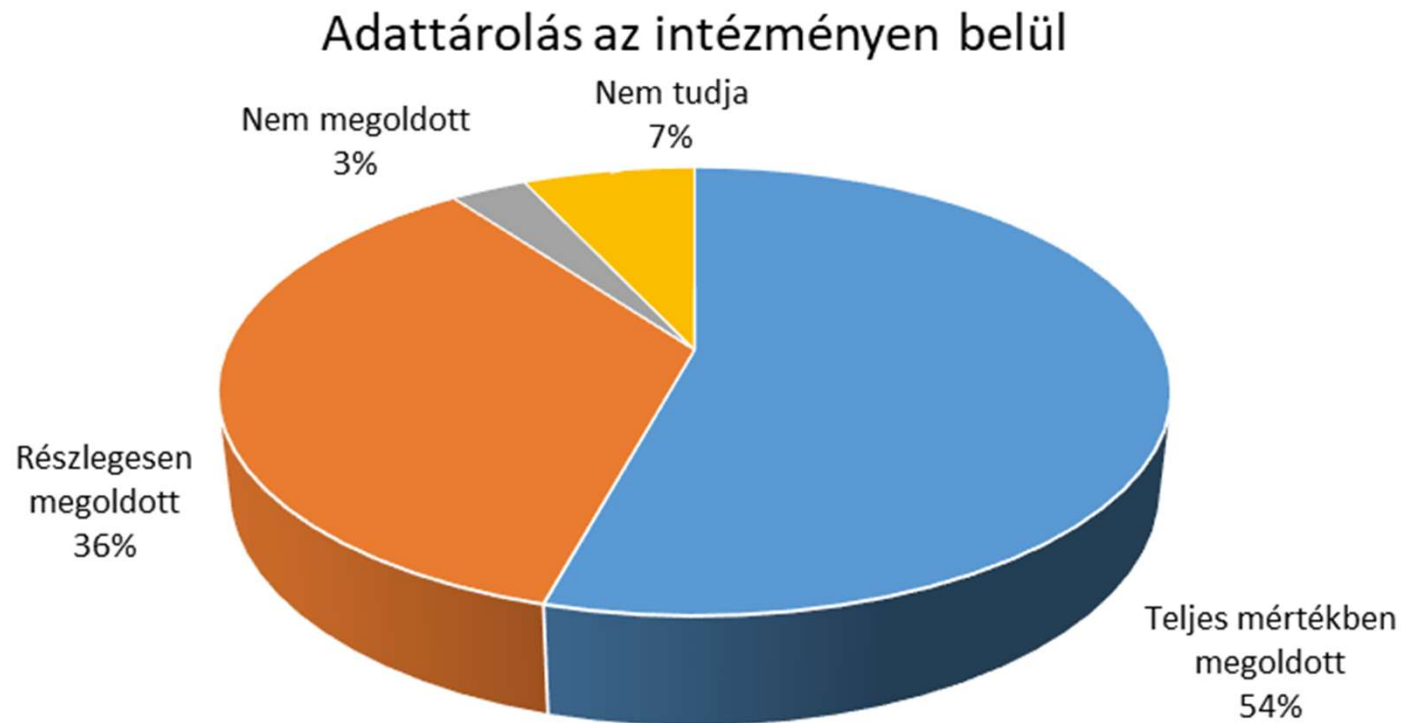


Fájlok típusok száma (átlag =3,3)



■ 1 ■ 2 ■ 3 ■ 4 ■ 5 ■ 6 ■ 7

Adatok tárolása – kutatás közben



Adatok tárolása - problémák

- 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”

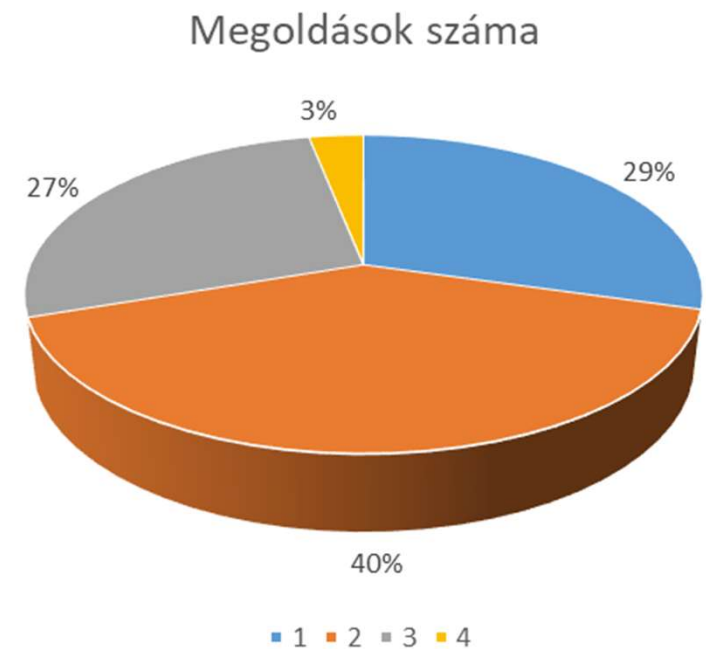
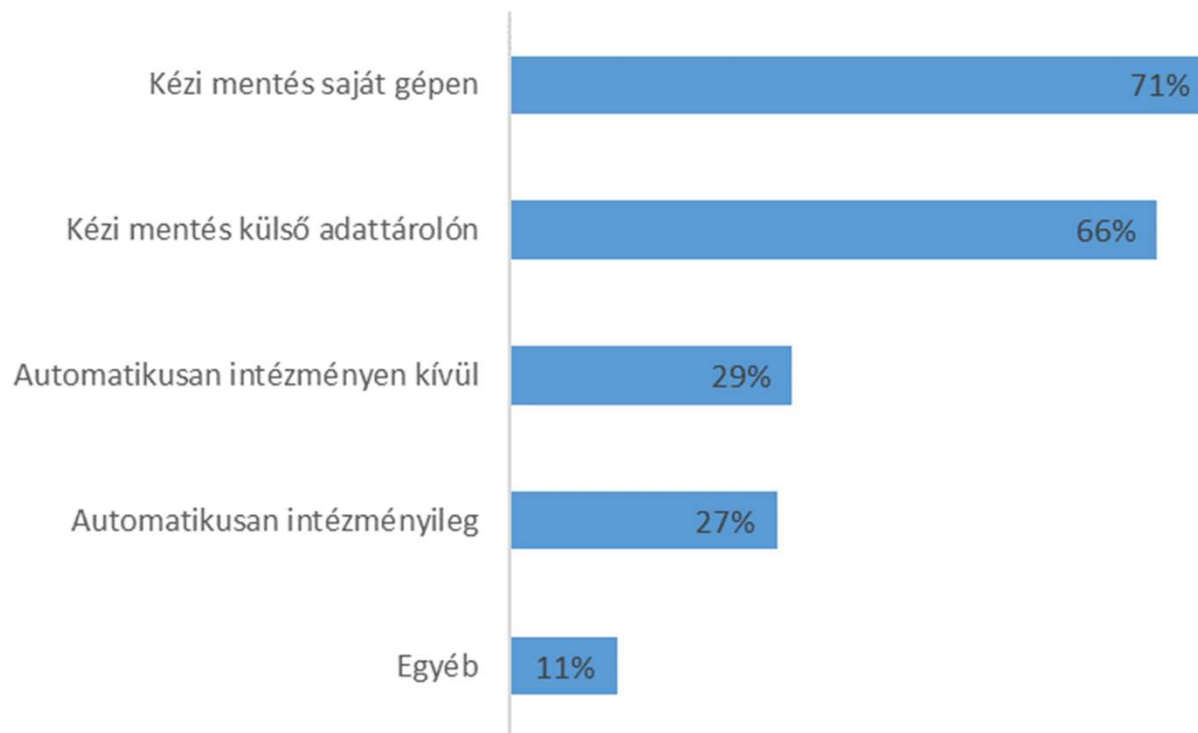
- Szervezeti, szervezési

„Mindenki tárolja valahogy az adatait, de ennek nincs egységes változata

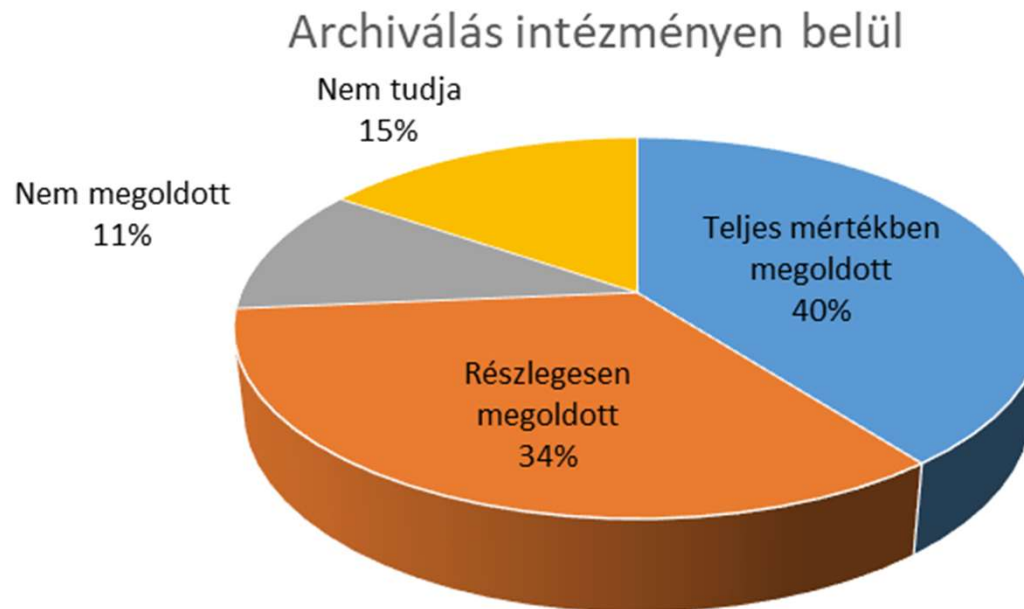
„A kapott tájékoztatás nem ébresztett bizalmat a szolgáltatás iránt”



Adatok mentése – munka közben



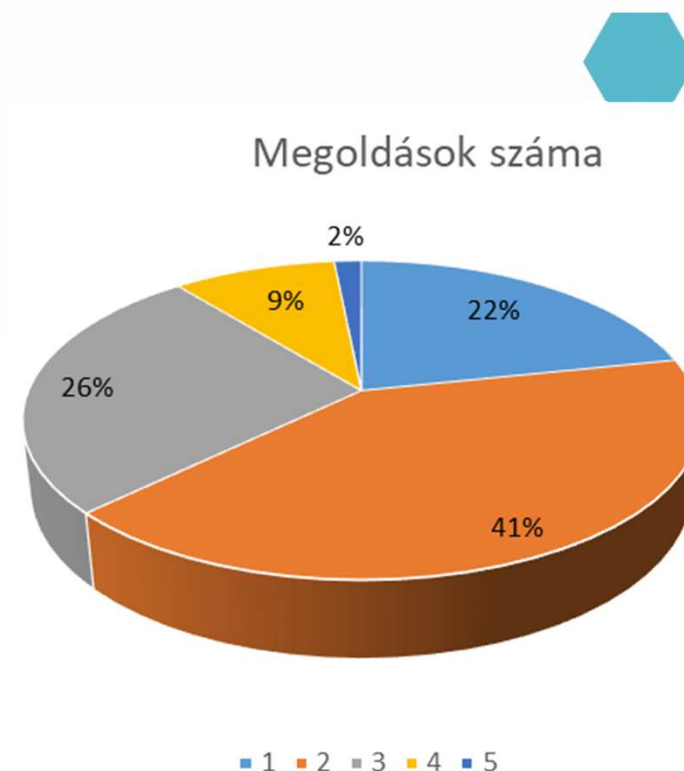
Adatok archiválása – kutatás után



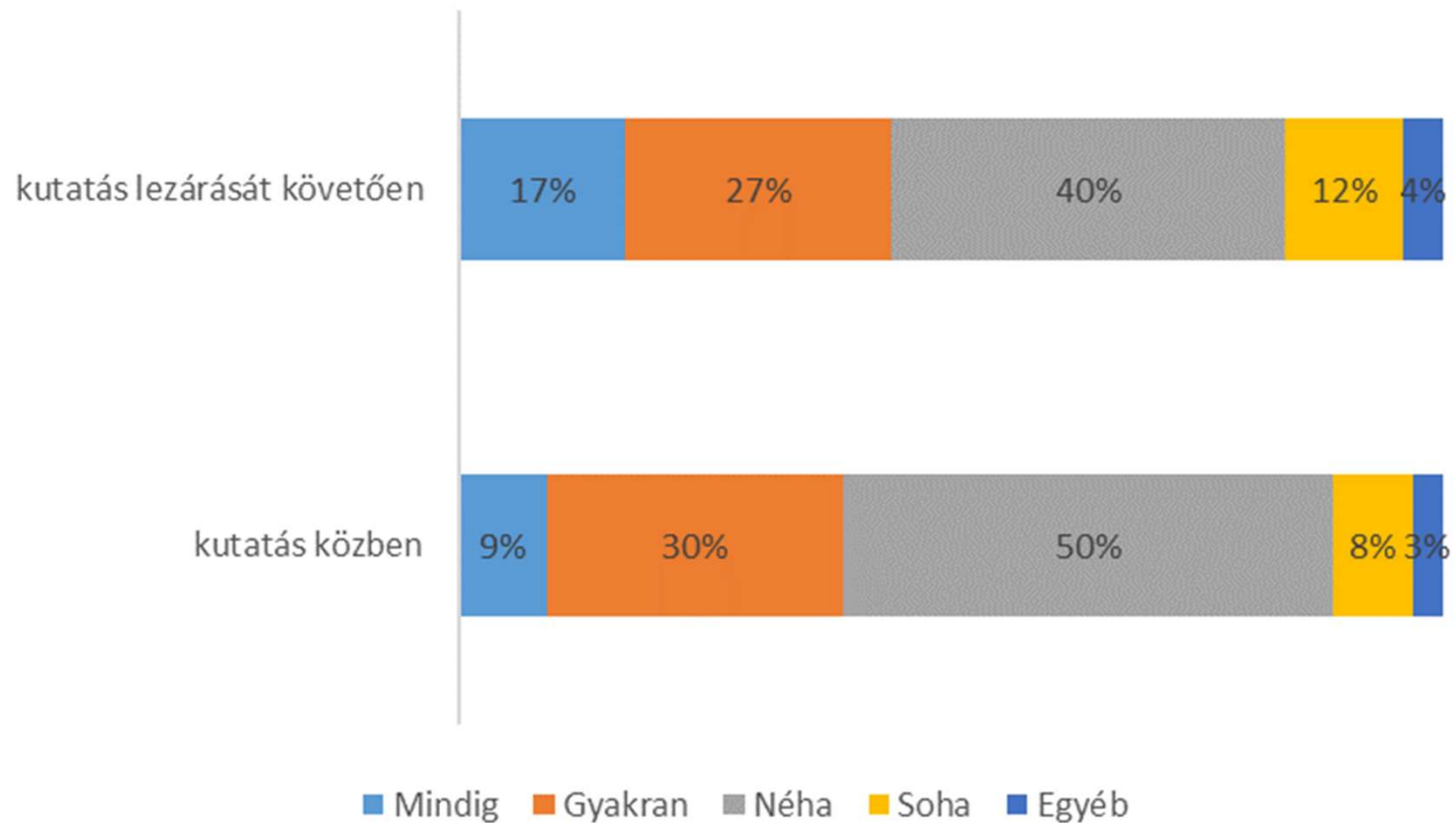
Adatok archiválása-problémák

- 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”.

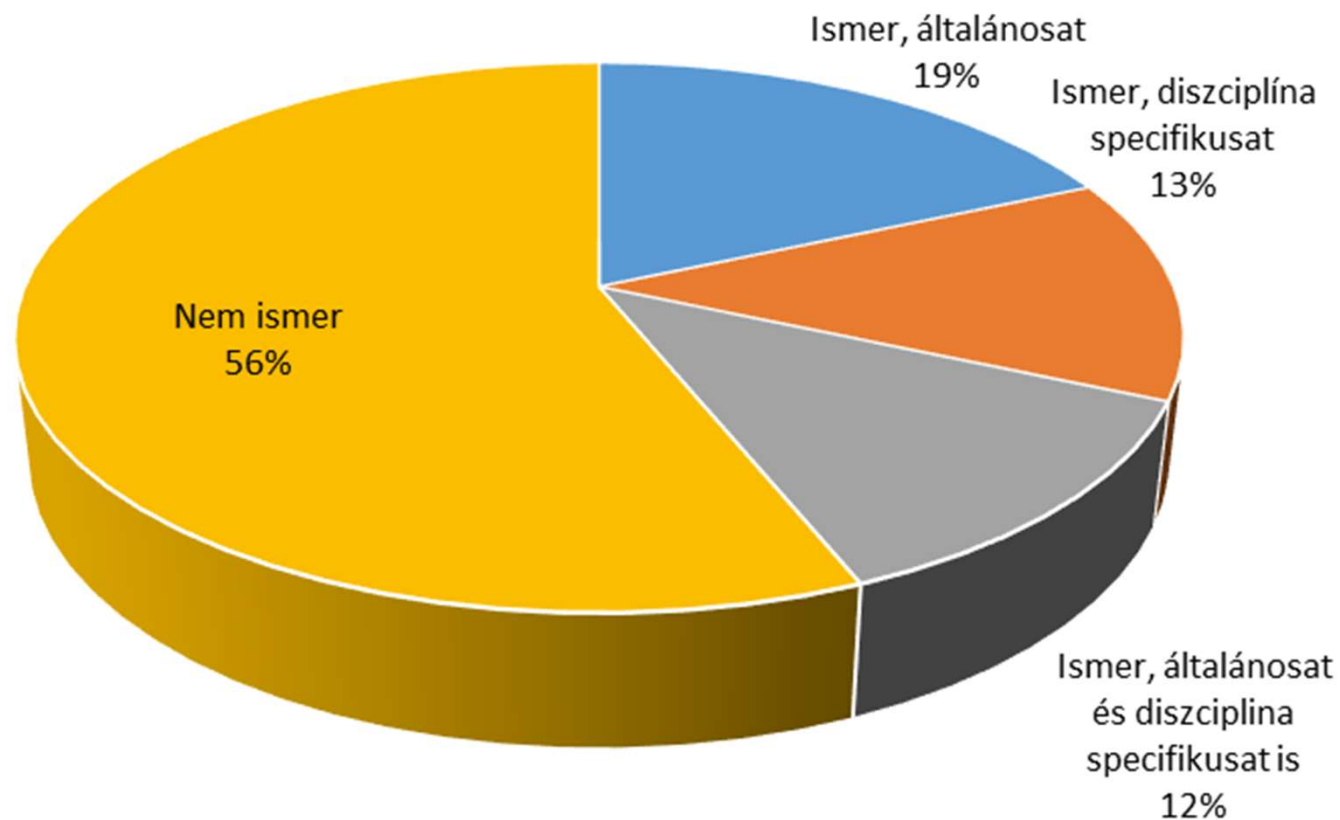
Adatok archiválása – kutatás lezárása után



Kutatási adatok megosztása



Igénybe vehető repositóriumok ismerete



Repozitóriummal szembeni elvárások (ötös skála)





Interjú adatfelvételek

Célcsoportok, témák

Az interjúk célcsoportjai

- intézetigazgatók,
- kutatók,
- Concorda használók.

Az interjúk fő témái

- kutatási terület(ek),
- adat élelciklus,
- esetleges adatkezelési, adattárolási problémák, nehézségek,
- tudományterületi hazai és nemzetközi trendek adatkezelés téren, illetve
- Concorda használóknál a letétbe helyezés során szerzett tapasztalatok.



Általános tapasztalatok



Nem csak az ELKH eltérő tudományterületein, hanem az egyes intézményeken belül, vagy akár egy kutatócsoport különböző projektjeiben meglehetősen különböző jellegű, számú és méretű kutatási adat keletkezik.

A keletkező adatok mennyisége jellemzően nő, miközben több intézményben digitalizálásra és rendezésre várnak korábbi kutatási adatok is.



Általános tapasztalatok

A kutatási adatok sokféleségéhez hasonlóan az adatkezelési, -tárolási, -megosztási és hosszútávú megőrzésre vonatkozó gyakorlatok is igen eltérő mintázatot mutatnak.

- A legtöbb intézményben jelenleg nincs előírás vagy elvárás az adatok rövid- vagy hosszútávú kezelését illetően, ugyanakkor egyre több helyen figyelhető meg erre irányuló törekvés mind az egyes kutatók, mind pedig az intézményvezetés részéről.
- Emellett sok helyen működik olyan kialakult gyakorlat, amely esetenként akár a 20-30 évvel ezelőtti adatok részletes megismerését, értelmezését, feldolgozását (és ilyen módon repositóriumba kerülését) is lehetővé teszi – és vannak ilyen irányú törekvések is.

Concorda használók véleménye

A Concorda-t már használó kutatókkal készült interjúk alapján megállapítható, hogy

- alapvetően felhasználóbarát,
- praktikus, és
- a legtöbb kutatói igénynek megfelel a jelenlegi felület,
- a változtatásokra, fejlesztésekre vonatkozó igényeket felmértük,
- és van, amit már meg is oldottak a repozitóriumot üzemeltető SZTAKI DSD munkatársai.





KÖSZÖNÖM A FIGYELMET!