



Kapcsolódási pontok a bölcsészettudományok területén adatgazdász szemmel

Szabó Bálint // szabo.balint@htk.elte.hu

ARP adatkonferencia // 2025.10.08.

HUN-REN Rényi Alfréd Matematikai Kutatóintézet

ÁTTEKINTÉS

- Kulcsterületek – kulcs metaadatok
- Földrajzi lefedettség
- Földrajzi név és a valóság
- Módszertan és eredmények



KULCSTERÜLETEK – KULCS METAADATOK

Kulcsterületek – kulcs metaadatok

Access points fall into six broad categories:

- Names
- Places
- Subjects
- Documentary forms
- Occupations
- Functions

Forrás: Describing Archives: A Content Standard

https://saa-ts-dacs.github.io/dacs/05_overview_of_archival_description.html

E1 CRM Entity

Superclass of:

E2 Temporal Entity

E52 Time-Span

E53 Place

E54 Dimension

E59 Primitive Value

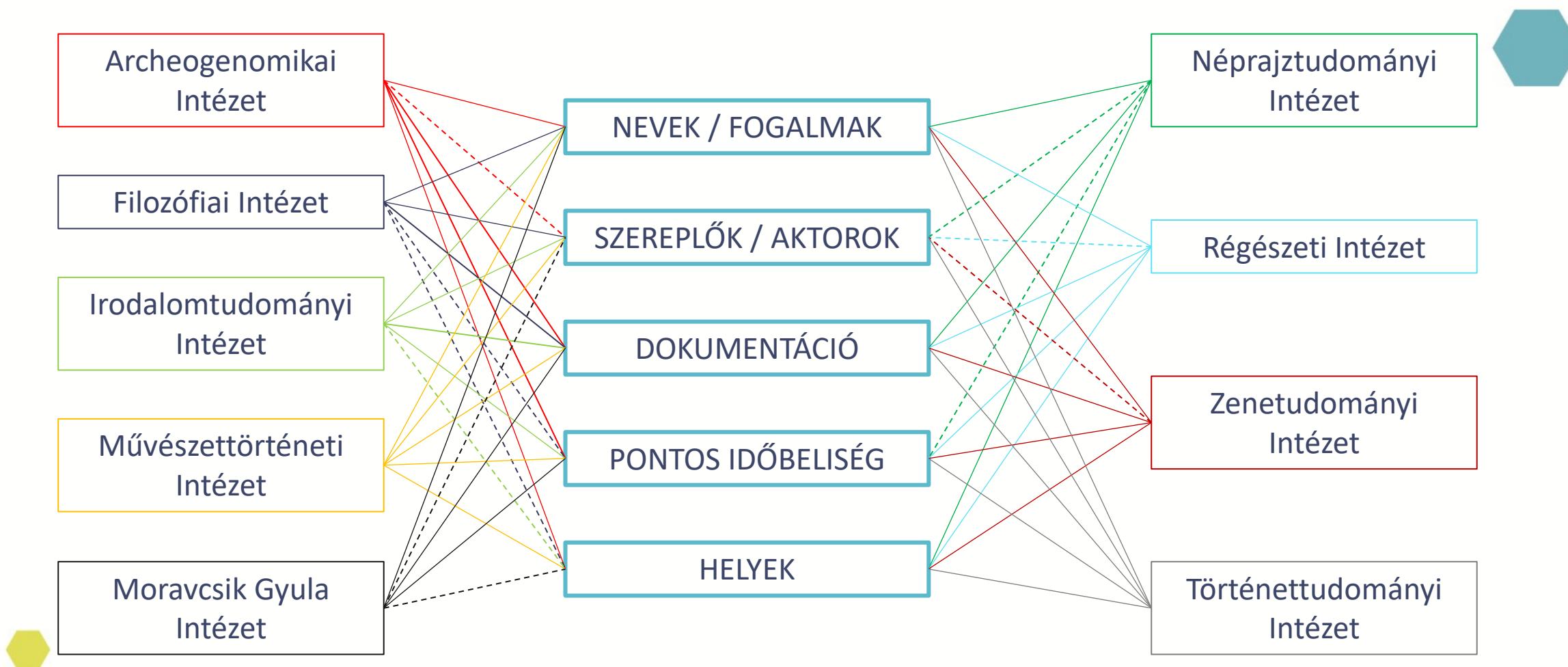
E77 Persistent Item

E92 Spacetime Volume

Forrás: ICOM CIDOC CRM

https://cidoc-crm.org/sites/default/files/cidoc_crm_version_7.3.1_1_0.pdf

Intézeti kapcsolatok



Kulcsterületek – kulcs metaadatok

Access points fall into six broad categories:

- Names
- **Places**
- Subjects
- Documentary forms
- Occupations
- Functions

Forrás: Describing Archives: A Content Standard

https://saa-ts-dacs.github.io/dacs/05_overview_of_archival_description.html

E1 CRM Entity

Superclass of:

E2 Temporal Entity

E52 Time-Span

E53 **Place**

E54 Dimension

E59 Primitive Value

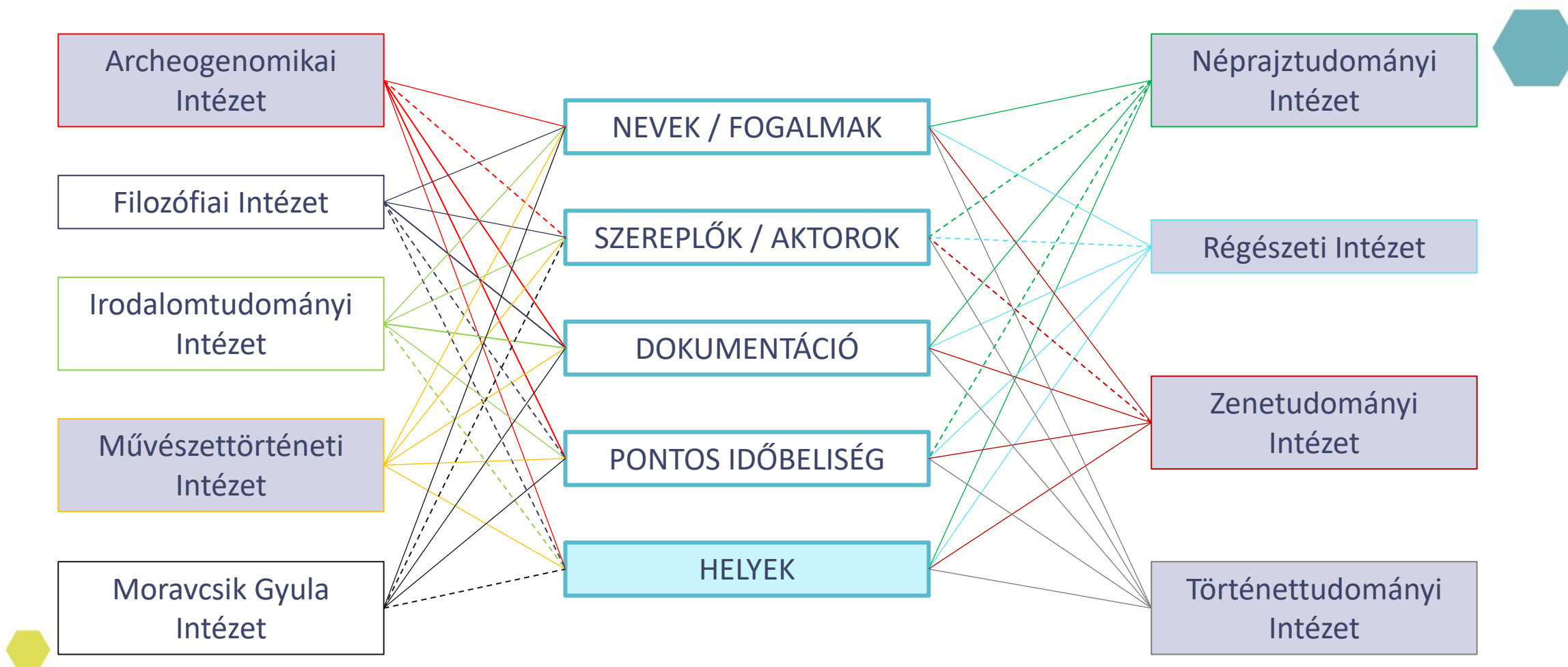
E77 Persistent Item

E92 Spacetime Volume

Forrás: ICOM CIDOC CRM

https://cidoc-crm.org/sites/default/files/cidoc_crm_version_7.3.1_1_0.pdf

Intézeti kapcsolatok





FÖLDRAJZI LEFEDETTSÉG

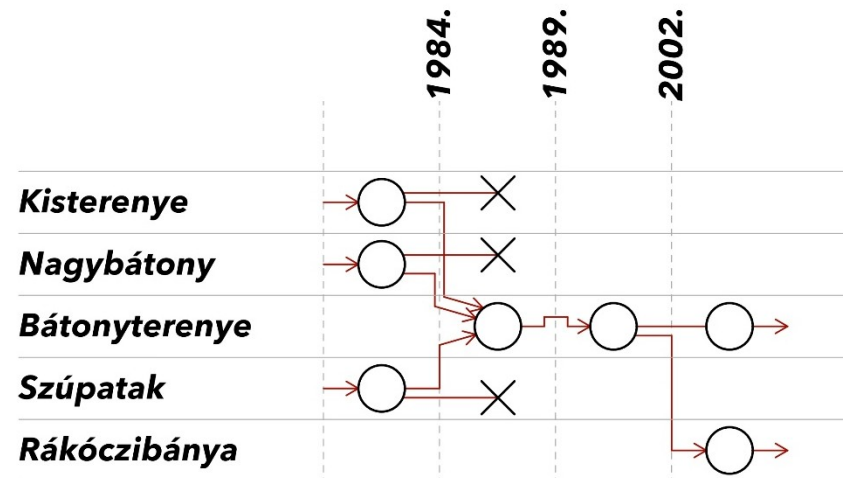


Földrajzi nevek problematikái

„A földrajzi név az egyetlen olyan téradat, ami nem mérhető eszközökkel, sem a Földről, sem repülőből, sem az űrből.”

/Márton Mátyás – térképész/

- A metaadat metaadatai
 - Időbeliség
 - Kulturális kontextus



Forrás: saját ábra

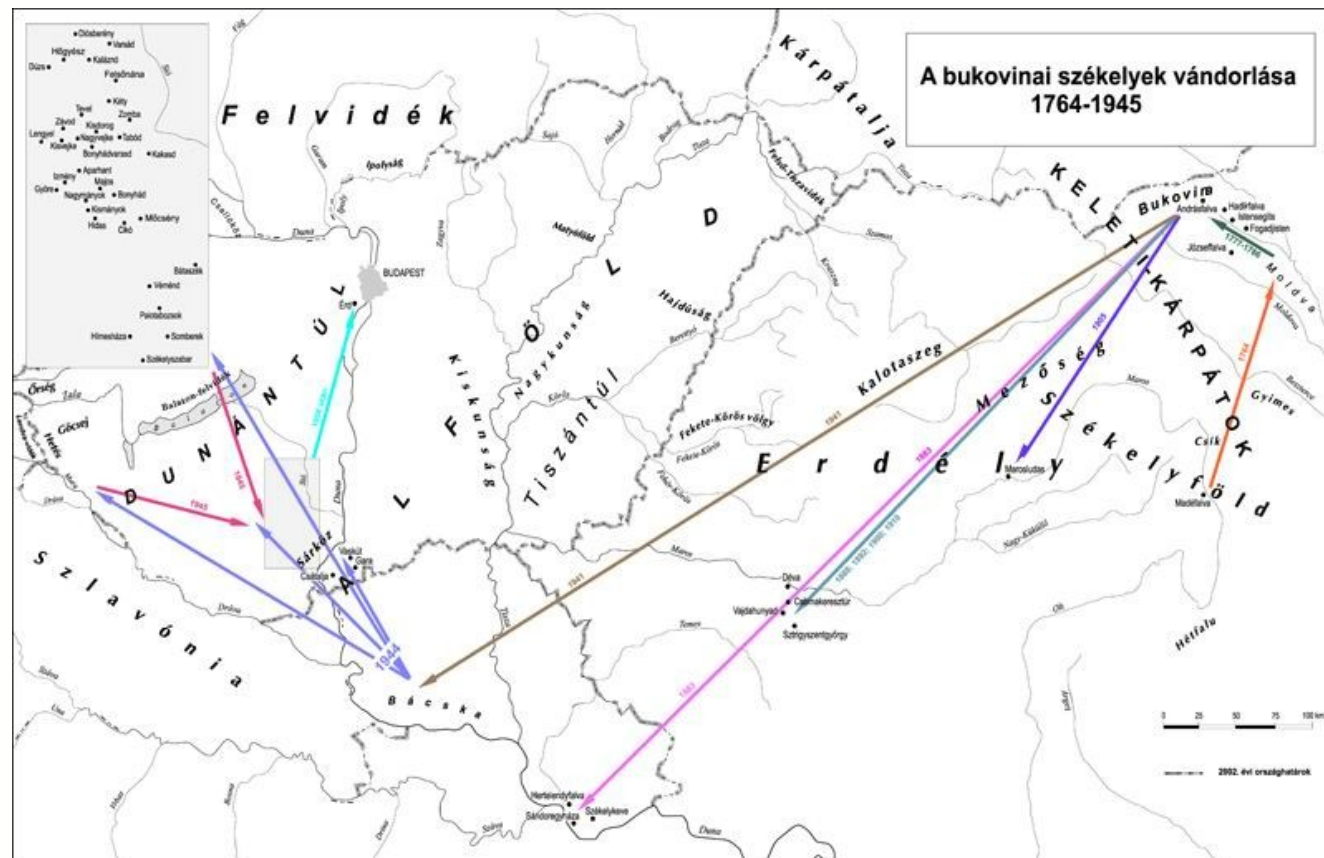
Gyűjtési hely ÉS érvényességi terület

- A gyűjtések helyfüggetlensége

- Elköltözött, elvándorolt népesség
- Kitelepítések
- Kisebbségek lokális hatásai

gyűjtési hely

érvényességi terület





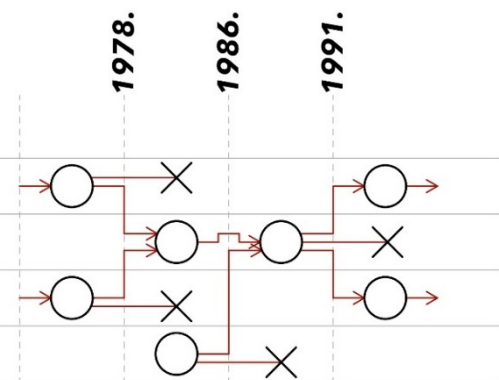
FÖLDRAJZI NÉV ÉS A VALÓSÁG



Időbeliség

- Átnevezések

- helyesírási változások
- átnevezések
- összevonások
- beolvadások
- kiválások



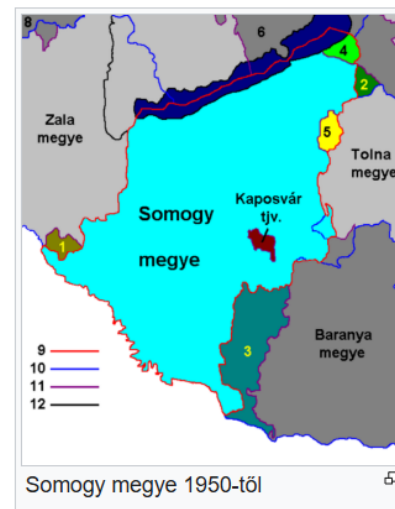
dezés
sok

Somogy megye

Somogy megye új területe 1950. február 1-jétől az alábbiak szerint alakult:

- Somogy megye addigi területe, kivéve
 - a Zala megyéhez csatolt Belezna, Nemespátró és Surd községeket (a térképen: 1),
 - a Fejér megyéhez csatolt Szabadhidvég községet (a térképen: 2),
 - a Barcsi járásból Baranya megyéhez csatolt Felsőszentmárton községet és Magyarlad, Németlad és Patosfa községek kivételével a Szigetvári járást (a térképen: 3),
 - Veszprém megyéből Balatonszabadi, Siófok és Siómaros községek (a térképen: 4),
 - Tolna megyéből Bedegkér, Kánya és Tengőd községek (a térképen: 5).

Kaposvár 1950. június 15. napjától lett a megye része, amikor a törvényhatósági jogú városi jogállás az egész országban megszűnt.



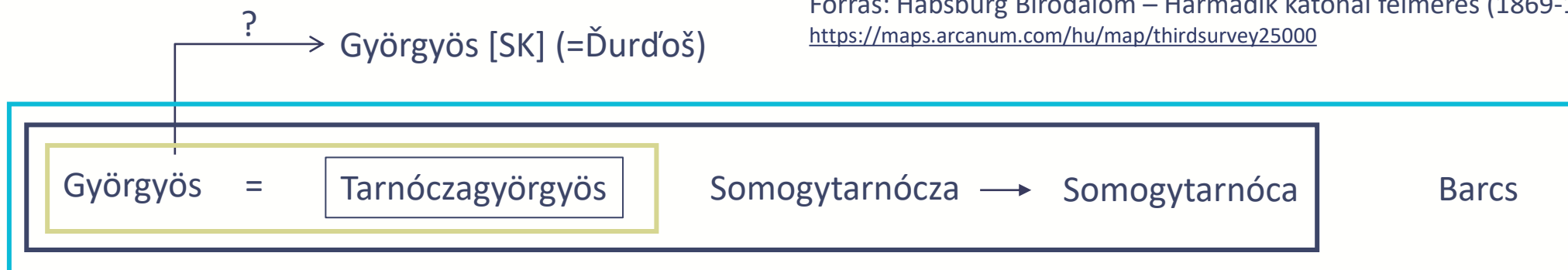
Forrás: Wikipedia – 1950-es megyerendezés
https://hu.wikipedia.org/wiki/1950-es_megyerendez%C3%A9s

Kulturális kontextus

- Változatokban élő nevek
- Beszélt nyelvi különbségek
- Lokális tudás – *pl.: Hol van a Gyöngyösi csárda?*



Forrás: Habsburg Birodalom – Harmadik katonai felmérés (1869-1887)
<https://maps.arcanum.com/hu/map/thirdsurvey25000>

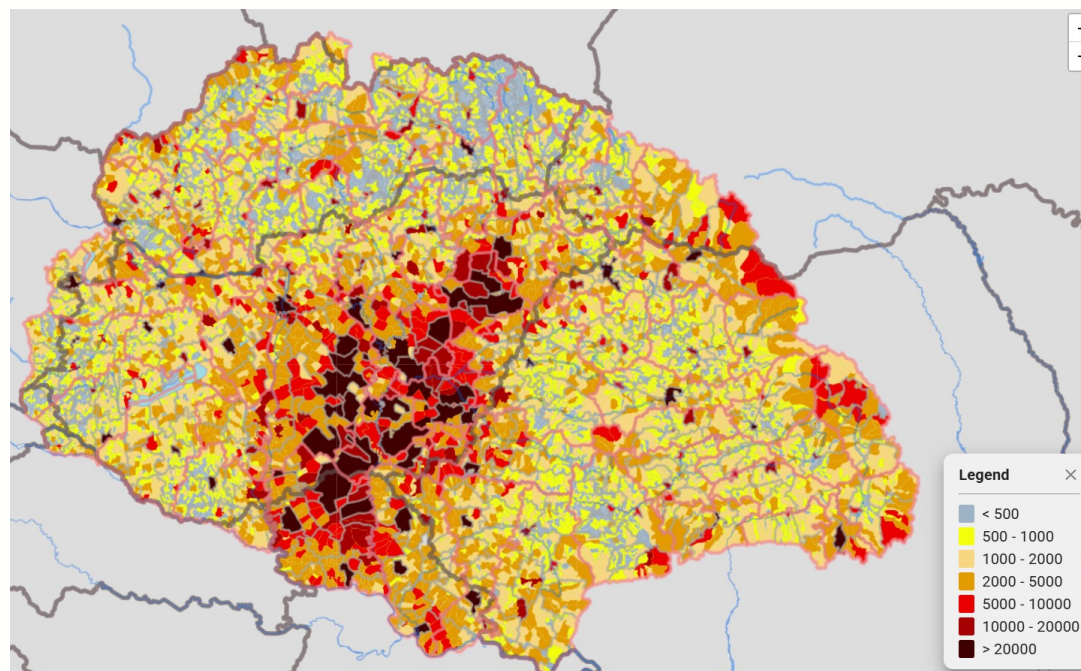




MÓDSZERTAN ÉS EREDMÉNYEK

Megvalósítási keretek

- HUN-REN Tech pályázat
- GISTA Hungarorum platform
<http://gistahungarorum.abtk.hu/>



Forrás: <http://gistahungarorum.abtk.hu/>

TelKPID	Telepulesnev_1912	TelepulesnevHNT1913	JarasNev	MegyeNev	GISCO_ID	LAU_NAME
1155	Pécs	Pécs	Pécsi	Baranya	HU_19415	Pécs
904	Albertfalu	Albertfalu	Baranyavári	Baranya	HR_00647	Čeminac
918	Baranyabán	Baranyabán	Baranyavári	Baranya	HR_03468	Popovac

Forrás: Referencia adatbázis a térképes felületen megjeleníthető helynevekről

Érintett adatbázisok

- Zenetudományi Intézet Népzene Archívumának leíró táblázatai
- Zenetudományi Intézet Néptánc Archívumának leíró táblázatai
 - Táncfolyamatok táblázat [22.405 sor]
 - Fotótár leíró táblázata [45.408 sor]
 - Kézirattár leíró táblázata [1.898 sor]
 - Tánckataszter összesítő kartonok adatbázisa [7.219 sor]

+1: Referencia táblázat a térképes felületen megjeleníthető helynevekről [13.036 sor]

> A táblázatok elemi gyűjtési egységekre való hivatkozást tesznek lehetővé



Vizualizálási célok

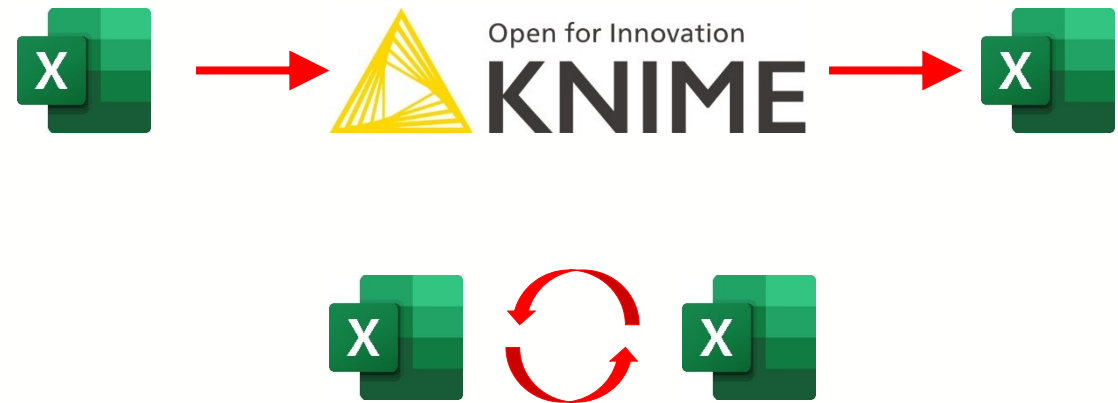
- táncfolyamatok száma / település [darab]
- rögzített táncfelvételek hossza / település [hh:mm:ss]
- releváns kéziratok oldalszáma / település [darab]
- releváns tánckataszteri kartonok száma / település [darab]
- nemzetiségi táncos adatok / település / nemzetiség [I/N]
- Kiemelten fontos táncgyűjtők tevékenysége / település [darab]

> mindez idő intervallumra szűrhető módon



Feldolgozás

- Cél: a kutatók által létrehozott struktúra továbbra is használható legyen
- Alakítási eszközök:
 - MS Excel
 - KNIME



Folyamatok

- Adatgazdagítás
- Helynevek egyértelmű beazonosítása
 - Helyes alak
 - Elnevezések vizsgálata több időpontban
 - Járási / megyei egyezés
- Helykód hozzárendelés
- Adatbázisban nem szereplő elemek beazonosítása
 - nincs benne (pl.: Moldvai terület)
 - benne van, de az említett település nem volt önálló abban a három időpontban, ami a mi táblázatunkban jelenleg kiemelt



Perspektíva

- Térképes adatvizualizációs eszköz
- Interdiszciplináris kutatási platform
- Adatfeldolgozási mintaprojekt
- KNIME szoftver integráns használata
- Gyűjtemények metaadatainak repozitálása
 - fél éves frissítésekkel a folyamatos változások miatt



Eddigi eredmények

- Nagy mennyiségű adatgazdagítás és pontosítás
- Általános érvényű tapasztalatok fájlkezelésben
 - digitalizálás
 - nevezéktan
 - csoportos átnevezések módszertanai
- Kezdetleges KNIME workflow-k az Excelben feljavított adatok gyors kimentéséhez
- Településszintű névtér gazdagításának módszertana





KÖSZÖNÖM A FIGYELMET!

Szabó Bálint // szabo.balint@htk.elte.hu // adatgazdasz@htk.elte.hu