
Szemantikus web

(Wikidata és az EU Open Data Portal)



Hajdu Kálmán

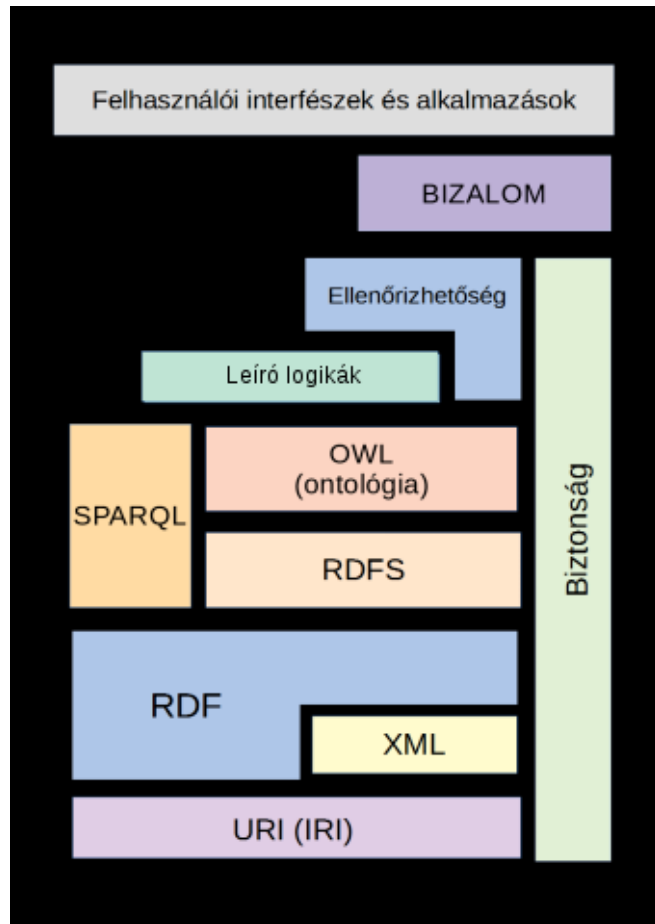
Szemantikus web

Az elmélet



Hajdu Kálmán

Szemantikus web



2007, Tim Berners-Lee

- ❑ <http://www.linkeddata.tools.com/semantic-web-basics>
- ❑ <http://wiki.dbpedia.org>
- ❑ <http://www.geonames.org/ontology/documentation.html>



Hajdu Kálmán

RDF

- Ez egy erőforrást leíró formátum
- Információ tárolás alapegysége a szemantikus triplet
- Angol terminológia szerint SPO (Subject, Predicate, Object) azaz: alany, állítmány, tárgy



RDF

- Subject: URI / Üres csomópont
- Predicate: URI
- Object: URI / Üres csomópont / karakter sor

- Példa:

Hankiss Elemér – foglalkozása – szociológus

[wd:Q957435](#) – [wdt:P106](#) – [wd:Q2306091](#)



RDF adatbázis

- A relációs adatbázisok indexelik az adatokat és index segítségével döntenek el, hogy melyik adatot kell a memóriába beolvasni és mikor.
- Speciális megoldások a triplestor-ok pl: 3store, OpenLink Virtuoso, Wikibase.



RDFS & OWL

Szemantikus metaadat szabványok (W3C-szabványok)

- RDFS (RDF Schema) Az RDF szabályok kiterjesztése
- OWL (Web Ontology Language)

Osztályok, alosztályok, osztálykiterjesztések (egyedek) és az osztályok, alosztályok stb. közötti kapcsolatok leírására szolgálnak.

„Friend of a Friend” <http://www.foaf-project.org/>

„Dublin Core” <http://dublincore.org/>



SPARQL

- ❑ A SPARQL egy RDF-hez tartozó triplestore-ra optimalizált szabványos lekérdező nyelv (query language)
- ❑ Nagyon hasonlít az SQL-re, de a logikája a szemantikus triplettekre épül
- ❑ Az adatbázisok eléréséhez szükséges egy SPARQL végpont



Szemantikus web

A gyakorlat



Hajdu Kálmán

Európai Unió adatportáljai



Az Európai Adatportál egyablakos hozzáférést biztosít az EU 24 hivatalos nyelvén mindazokhoz az adatokhoz, amelyeket Európa-szerte a különböző szintű kormányzati szervek tettek közzé. Az Európai Bizottság által finanszírozott Európai Adatportál tartalma kiegészíti a nyílt hozzáférésű adatok európai uniós portálján található információkat.

<https://www.europeandataportal.eu/hu/>



A **nyílt hozzáférésű adatok** EU-portálja
Az Európai Unió nyílt hozzáférésű adatainak lelőhelye

A nyílt hozzáférésű adatok európai uniós portálja hozzáférést biztosít az Európai Unió intézményei és szervei által közzétett adatokhoz. Az ebben a katalógusban található összes adat – akár kereskedelmi, akár más célból – szabadon felhasználható.

<https://data.europa.eu/euodp/hu/linked-data>



Hajdu Kálmán

SPARQL

Ha a nyílt hozzáférésű adatok európai uniós portálján tárolt metaadatok között kíván keresni, használja az alábbi lekérdező interfészt (SPARQL-végpont).

Névterek *

```
PREFIX dcat: <http://www.w3.org/ns/dcat#>
PREFIX odp: <http://data.europa.eu/euodp/ontologies/ec-odp#>
PREFIX dc: <http://purl.org/dc/terms/>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
```

SPARQL-keresés *

```
select distinct ?g ?o where { graph ?g {?s dc:title ?o. filter regex(?o, 'Hungary', 'i') filter regex(?o, 'court', 'i')} } LIMIT 10
```

Formátum *

HTML ▾

A találatok maximális száma *

10 ▾

Lekérdezés indítása

A találatokat tartalmazó URL lekérése

Metaadatszótár

Létrehoztunk egy metaadatszótárt, amely a [DCAT](#) adatkatalógus-szótár és a Dublin Core ([DC](#)) szótár terminológiáján alapul. A szótár [munkalap](#) és [ontológia](#) formájában is rendelkezésre áll, és úgy lett kialakítva, hogy igazodjon az állományleíró metaadatok sémájához ([ADMS](#)).



Hajdu

A lekérdezés eredménye

g	o
http://data.europa.eu/euodp/data/dataset/judicial-systems-in-member-states-specialised-courts	"Specialised courts - Hungary"@en
http://data.europa.eu/euodp/data/dataset/information-on-judicial-systems-in-member-states-ordinary-courts	"Ordinary courts - Hungary"@en

Vissza

MEGTALÁLTA, AMIT KERESETT?

Alkalmazások: <https://data.europa.eu/euodp/hu/apps>

Fejlesztési támogatás: <https://data.europa.eu/euodp/hu/developerscorner>



Hajdu Kálmán

Wikidata?

- ❑ A Wikidata egy **ténytár**, nyelvektől független a **szemantikus web** eszközeit használó adatbázis.
- ❑ A Wikipédiával a kapcsolatot speciális API-k segítik.
- ❑ Az adattárolást RDF szabvány szerint történik
- ❑ A lekérdezésekhez SPARQL nyelvet használjuk



Wikidata statisztika

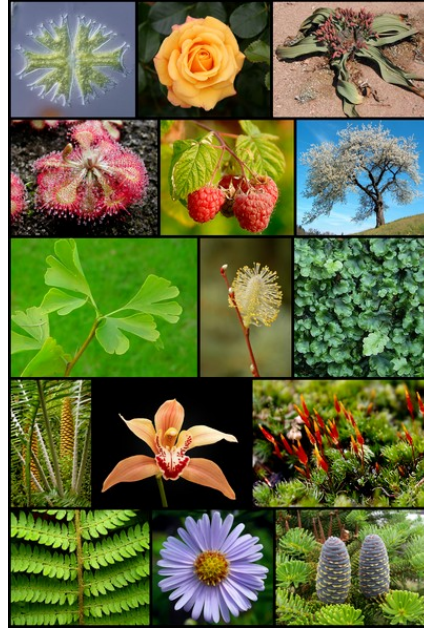
- 42 678 193 elem
- 4699 tulajdonság (predikátum vagy property)



Hajdu Kálmán

Wikidata statisztika

~2 427 000
taxon



~4 170 000
ember

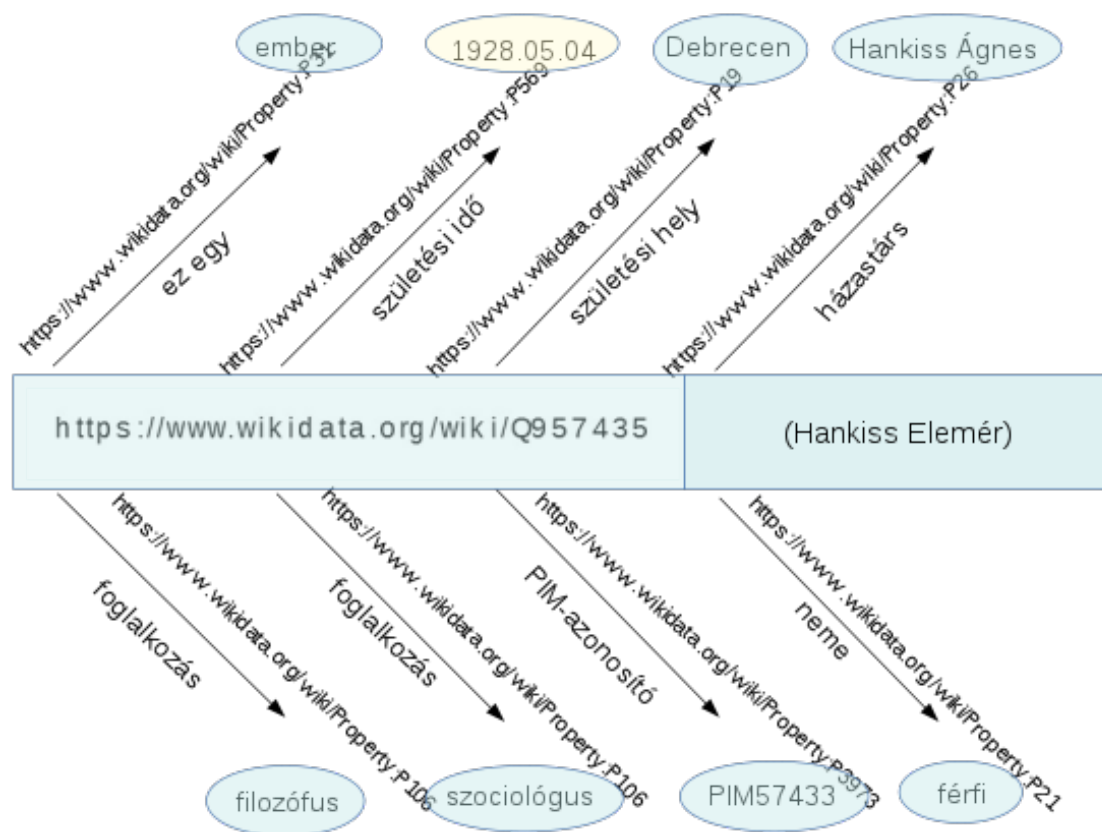


~5 820 000
Hely (koordináta)



Hajdu Kálmán

RDF formátum



Wikidata elem (részlet)

Hankiss Elemér (Q957435)

magyar szociológus, író

▼ További nyelveken












Nyelv	Címke	Leírás	Más néven
magyar	Hankiss Elemér	magyar szociológus, író	
német	Elemér Hankiss	Nincs leírás	
angol	Elemér Hankiss	Hungarian sociologist	Elemer Hankiss
lett	Nincs címke	Nincs leírás	
orosz	Nincs címke	Nincs leírás	
arab	Nincs címke		عالم اجتماع مجري
bangla	Nincs címke	হাজেরীয় সমাজবিজ্ঞানী	
katalán	Nincs címke	sociòleg hongarès	
walesi	Elemér Hankiss	Nincs leírás	
kanadai angol	Nincs címke	Hungarian sociologist	
brit angol	Nincs címke	Hungarian sociologist	
eszperantó	Elemér Hankiss	Nincs leírás	
spanyol	Elemér Hankiss	sociólogo húngaro	Elemer Hankiss



Hajdu Kálmán

Wikidata elem (részlet)

Állítások

osztály, amelynek példánya	 ember ▶ 2 hivatkozás	 szerkesztés + érték hozzáadása
kép	 Hankiss Elemér closeup.jpg  ▼ 0 hivatkozás	 szerkesztés + forrás hozzáadása + érték hozzáadása
nem	 férfi ▶ 2 hivatkozás	 szerkesztés + érték hozzáadása
állampolgárság	 Magyarország ▼ 0 hivatkozás	 szerkesztés + forrás hozzáadása + érték hozzáadása
utónév	 Elemér	 szerkesztés



Hajdu Kálmán

SPARQL lekérdezés

```
PREFIX entity: <http://www.wikidata.org/entity/>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX hint: <http://www.bigdata.com/queryHints#>
#Data of Elemér Hankiss
SELECT ?propUrl ?propLabel ?valUrl ?valLabel ?picture WHERE {
  hint:Query hint:optimizer "None".
  {
    BIND(entity:Q957435 AS ?valUrl)
    BIND("N/A" AS ?propUrl)
    BIND("identity"@en AS ?propLabel)
  }
  UNION
  {
    entity:Q957435 ?propUrl ?valUrl.
    ?property ?ref ?propUrl.
    ?property rdf:type wikibase:Property.
    ?property rdfs:label ?propLabel.
  }
  ?valUrl rdfs:label ?valLabel.
  OPTIONAL { ?valUrl wdt:P18 ?picture. }
  FILTER((LANG(?valLabel)) = "hu")
  FILTER((LANG(?propLabel)) = "hu")
}
ORDER BY ?propUrl ?valUrl
LIMIT 200
```

property	value
foglalkozás	szociológus
foglalkozás	filozófus
díj	Alföld-díj
születési hely	Debrecen
halálozási hely	Budapest
halálozási hely	Budapest
nem	férfi
házastárs	Hankiss Ágnes
állampolgárság	Magyarország
ez egy	ember
ez egy	ember
beszélt nyelvek	magyar
utónév	Elemér



Személy infobox kifejtés

```
{{Személy infobox
|típus =
|név = Hankiss Elemér
|kép =
|képméret =
|képaláírás = Hankiss Elemér a 2012-es
Ünnepi Könyvhéten
|művésznév =
|álnév =
|becenév =
|nemzetiség = [[Magyarok|magyar]]
|házastárs = [[Hankiss Ágnes]]<br />
Zétényi Lili
|gyermekai = Hankiss Katalin<br
/>Hankiss Ádám
|szakma =
```







```
{{Személy infobox
|típus=
|név=
|kép= <!-- Wikidata: p18; A
WP:NEMSZABAD szerint nem helyezhető
el jogvédett kép (filmkocka, poszter stb)
élő személy esetén -->
|képméret =
|képaláírás =
|születési dátum= <!-- Wikidata: p569 -->
|születési hely= <!-- Wikidata: p19 -->
|halál dátuma= <!-- Wikidata: p570 -->
|halál helye= <!-- Wikidata: p20 -->
|halál oka= <!-- Wikidata: p509 -->
|sírhely= <!-- Wikidata: p119 -->
|születési név= <!-- Wikidata: p513 -->
|állampolgárság= <!-- Wikidata: p27 -->
|házastárs = <!-- Wikidata: p26 -->
|élettárs = <!-- Wikidata: p451 -->
|gyermekai = <!-- Wikidata: p40 -->
|szakma = <!-- Wikidata: p106 -->
```



Hajdu Kálmán

Wikidata elem (részlet)

Azonosítók

VIAF-azonosító	 24643569	 szerkesztés
	▶ 1 hivatkozás	+ érték hozzáadása
ISNI	 0000 0001 1608 8674	 szerkesztés
	▶ 1 hivatkozás	+ érték hozzáadása
SUDOC-azonosító	 030318599	 szerkesztés
	▶ 1 hivatkozás	+ érték hozzáadása
NTA-azonosító	 087450941	 szerkesztés
	▶ 1 hivatkozás	+ érték hozzáadása



Hajdu Kálmán

Wikipédia Infobox

{{Személy infobox

|típus =

|név = Hankiss Elemér

|kép = Hankiss Elemér closeup.jpg

|képméret =

|képaláírás = Hankiss Elemér a 2012-es Ünnepi Könyvhéten

|művésznév =

|álnev =

|becenév =

|nemzetiség = [[Magyarok|magyar]]

|házastárs = [[Hankiss Ágnes]]
 Zétényi Lili

|gyermekai = Hankiss Katalin
Hankiss Ádám

|szakma =

Nemzetközi katalógusok

[WorldCat](#) · [VIAF: 24643569](#) · [PIM: PIM57433](#) ·

[LCCN: n83204808](#) · [ISNI: 0000 0001 1608 8674](#) ·

[SUDOC: 030318599](#) · [BNF: cb121759013](#)



Hajdu Kálmán

Hankiss Elemér



Hankiss Elemér a 2012-es Ünnepi Könyvhéten

Született	1928. május 4. ^[1] Debrecen
Elhunyt	2015. január 10. (86 évesen) ^[1] Budapest ^[2]
Állampolgársága	magyar
Nemzetisége	magyar
Házastársa	Hankiss Ágnes Zétényi Lili
Gyermekei	Hankiss Katalin Hankiss Ádám
Foglalkozása	szociológus,

Nemzetközi katalógusok

- ❑ **VIAF ID (P214)** VIAF-azonosító azonosító a nemzetközi virtuális katalógustárban
- ❑ **ISNI (P213)** nemzetközi standard névazonosító
- ❑ **BnF ID (P268)** a francia nemzeti könyvtár által kiadott azonosítószám
- ❑ **PIM authority ID (P3973)** Petőfi Irodalmi Múzeum névazonosítója

Nemzetközi katalógusok

[WorldCat](#) · [VIAF: 22263856](#) · [PIM: PIM57994](#) · [LCCN: n81142432](#) · [ISNI: 0000 0001 1873 8204](#) · [GND: 11673759X](#) · [SUDOC: 121977366](#) · [BNF: cb12667481b](#)



Wikidata Query Service

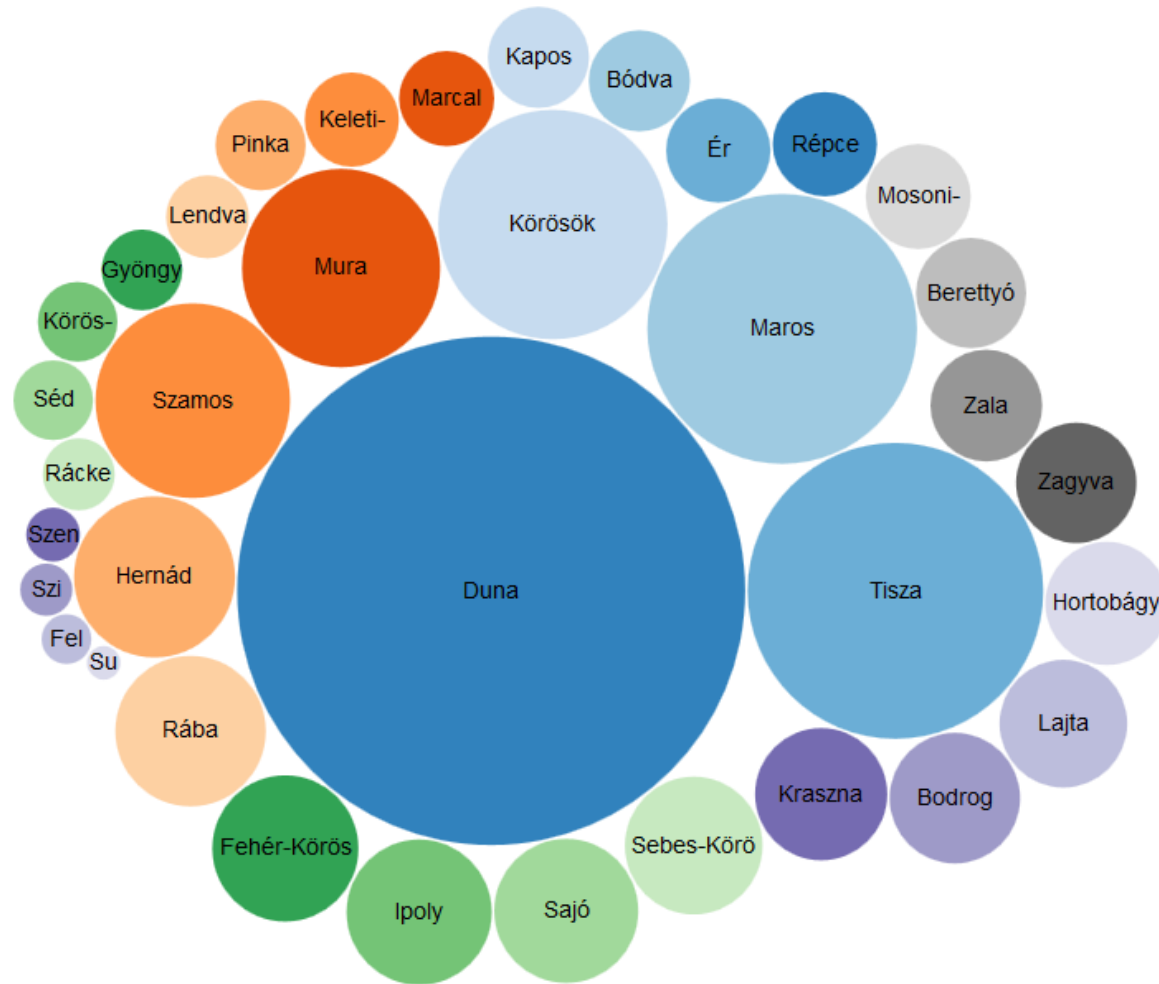
<https://query.wikidata.org/>

```
1 #Magyarország folyói
2 #added before 2017-11
3 #defaultView:BubbleChart
4 SELECT ?item ?itemLabel ?length ?pic ?location WHERE {
5   ?item (wdt:P31/wdt:P279*) wd:Q4022. #természetes vízfolyások
6   ?item wdt:P2043 ?length. #vízfolyás hossza
7   ?item wdt:P18 ?pic. #kép a folyóról
8   #?item wdt:P625 ?location.
9   SERVICE wikibase:label { bd:serviceParam wikibase:language "hu". }
10  ?item wdt:P17 wd:Q28. #ország amely területén folyik (Magyarország)
11 }
12 ORDER BY DESC(?length) ?itemLabel
13 LIMIT 50
```



Hajdu Kálmán

Wikidata Query Service



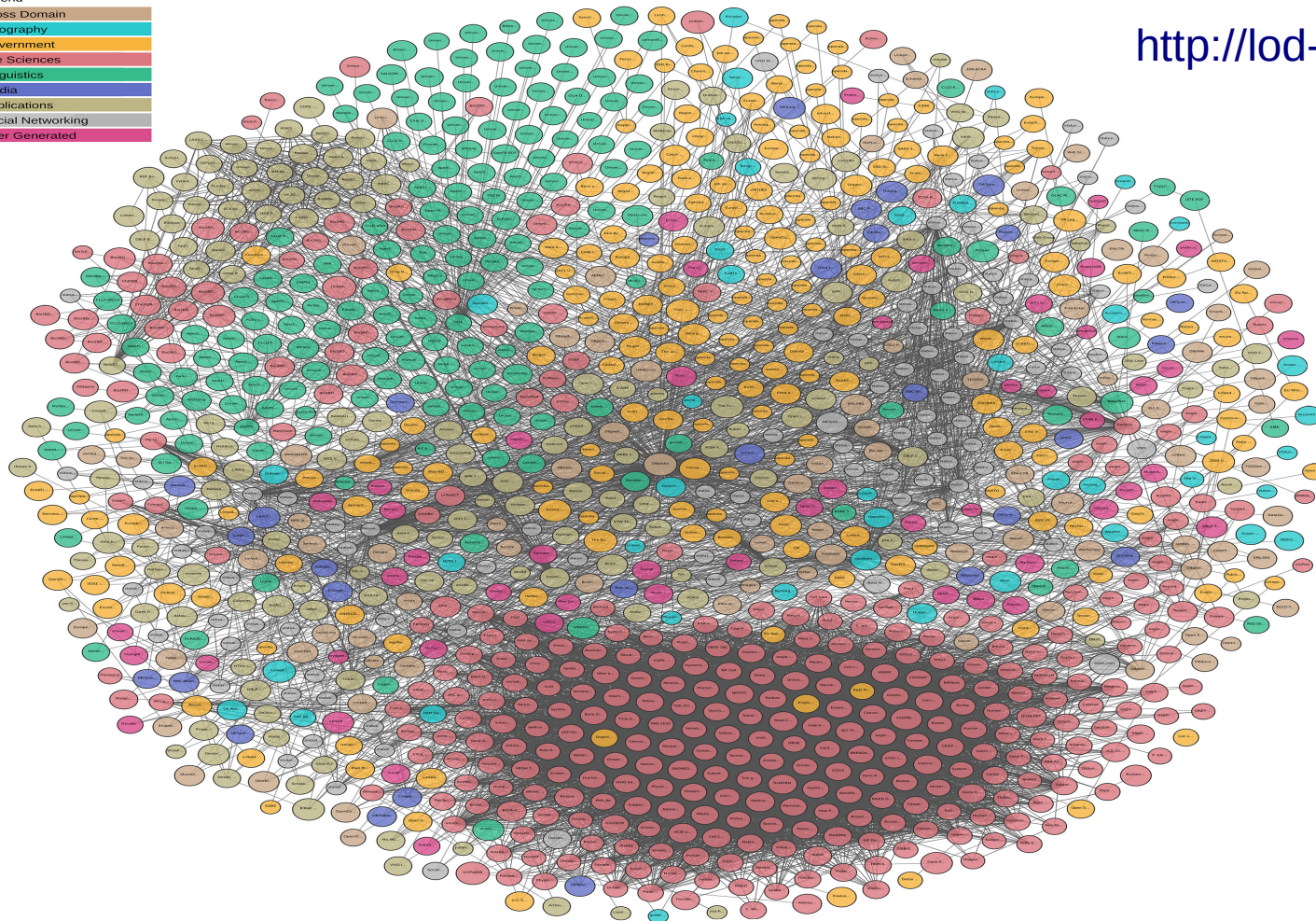
Hajdu Kálmán

LOD - a Linked Open Data Cloud

Legend

Cross Domain
Geography
Government
Life Sciences
Linguistics
Media
Publications
Social Networking
User Generated

<http://lod-cloud.net/>



The Linked Open Data Cloud from lod-cloud.net



Hajdu Kálmán

Nyílt hozzáférésű adatbázisok Magyarországon

- ❑ Szépművészeti Múzeum nyílt adatbázisa (336145 triplet) → <http://data.szepmuveszeti.hu/>
- ❑ Országos Széchényi Könyvtár (Torzó, elvileg 19 300 000 triplet)
http://nektar.oszk.hu/wiki/Szemantikus_web
- ❑ SZTAKI DSD kísérleti szolgáltatása (11 millió triplet) <http://lod.sztaki.hu/>

