

Dbpedia:

Glue for all Wikipedias and a Use Case for Multilingualism

Marco Fossati



Mariano Rico



Martin Brümmer



Dbpedia:

Extracting knowledge from Wikipedia

Martin Brümmer

bruemmer@informatik.uni-leipzig.de

UNIVERSITÄT LEIPZIG

How?

**Mapping wikipedia data to
Linked data**

why?

**Turn documents into data to granularly
use and query it**

Result:

Multilingual data with a common structure

Organized in chapters

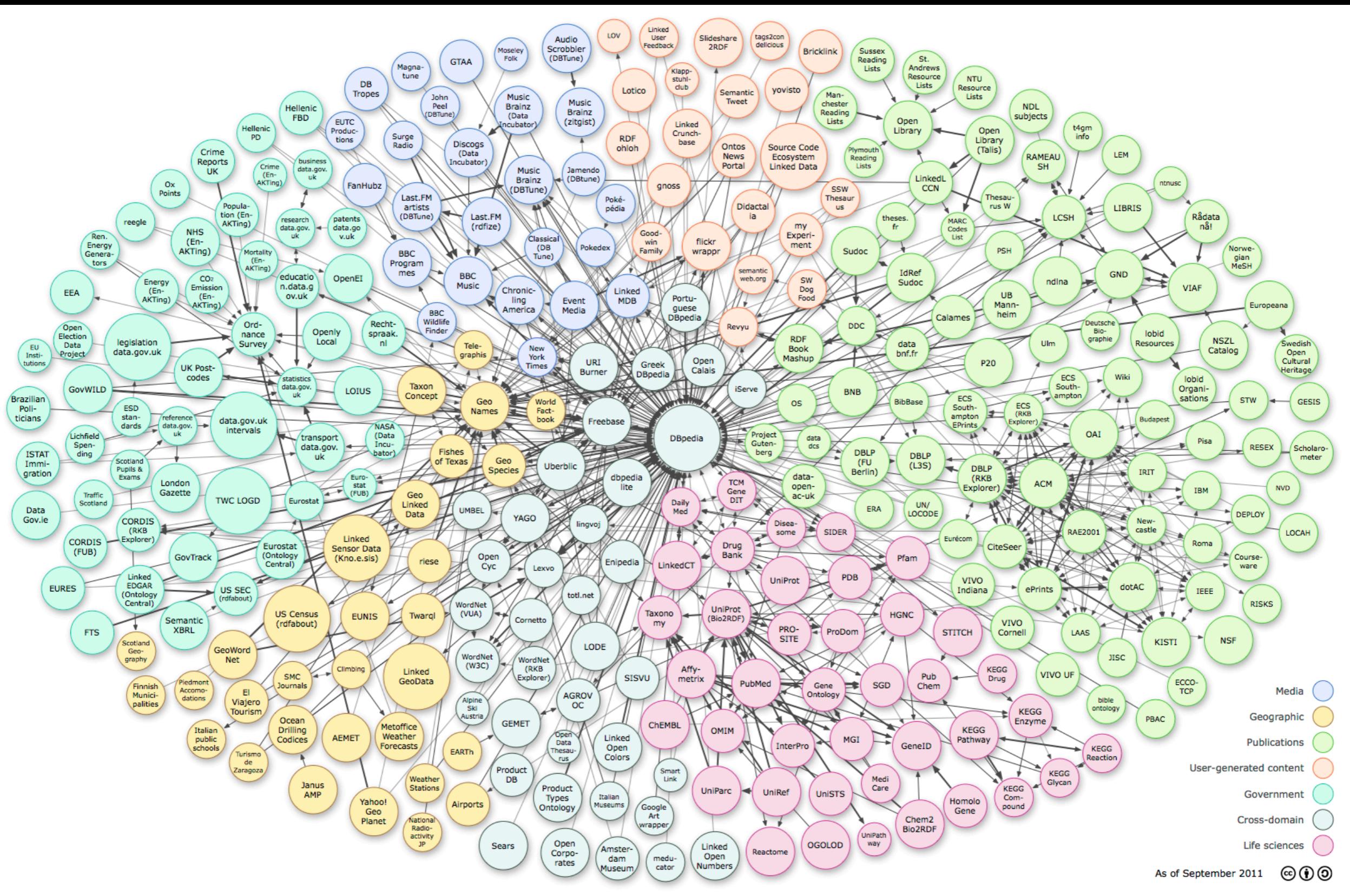
**14 Language communities maintaining their
language dbpedias**

Multilingual community

**Guarantee data quality and coverage
beyond language borders**

Supported by DBpedia association

**Opening the research project for long-term
sponsoring**



The center of the Iod cloud

Internationalization

Extracting multilingual knowledge

English
Mapping

Dbpedia.org



Wiki Loves Earth in focus during May and June
Discover nature, take photos, help Wikipedia and win!

Coordinates: 40°24'N 3°41'W

Madrid

From Wikipedia, the free encyclopedia

This article is about the capital city of Spain. For the autonomous community, see [Community of Madrid](#). For other uses, see [Madrid \(disambiguation\)](#).

Madrid (English /məˈdrɪd/, Spanish: [maˈðɾið]) is the capital of Spain and its largest city. The population of the city is roughly 3.3 million^[4] and the entire population of the Madrid metropolitan area is calculated to be around 6.5 million. It is the third-largest city in the European Union, after London and Berlin, and its metropolitan area is the third-largest in the European Union after London and Paris.^{[5][6][7][8]} The city spans a total of 604.3 km² (233.3 sq mi).^[9]

The city is located on the Manzanares River in the centre of both the country and the Community of Madrid (which comprises the city of Madrid, its conurbation and extended suburbs and villages); this community is bordered by the autonomous communities of Castile and León and Castile-La Mancha. As the capital city of Spain, seat of government, and residence of the Spanish monarch, Madrid is also the political, economic and cultural centre of Spain.^[10] The current mayor is Ana Botella from the People's Party (PP).

The Madrid urban agglomeration has the third-largest GDP^[11] in the European Union and its influences in politics, education, entertainment, environment, media, fashion, science, culture, and the arts all contribute to its status as one of the world's major global cities.^{[12][13]} Due to its economic output, high standard of living, and market size, Madrid is considered the major financial centre of Southern Europe^{[14][15]} and the Iberian Peninsula; it hosts the head offices of the vast majority of the major Spanish companies, such as Telefónica, Iberia or Repsol. Madrid is the 10th most livable city in the world according to Monocle magazine, in its 2010 index.^{[16][17]}

Madrid houses the headquarters of the World Tourism Organization (WTO), belonging to the United Nations Organization (UN), the SEGIB, the Organization of Ibero-American States (OEI), and the Public Interest Oversight Board (PIOB). It also hosts major international institutions regulators of Spanish: the Standing Committee of the Association of Spanish Language Academies, headquarters of the Royal Spanish Academy (RAE), the Cervantes Institute and the Foundation of



Internationalization

Extracting multilingual knowledge

Wiki Loves Earth in focus during May and June
Koordinaten: 40° 25' N, 3° 42' W |

Wiki Loves Earth:
Natur für die Wikipedia fotografieren und gewinnen!

From Madrid

Madrid of the calcu and E Paris

The c Madr com Manc Madr Botel The I influe the a outpu Soutl majo in the Madr Natic Publi Span head

Diese Artikel behandelt die Hauptstadt von Spanien. Zu weiteren Bedeutungen siehe [Madrid \(Begriffsklärung\)](#).

Madrid (dt. [maˈdʁɪt], span. [maˈðɾið]) ist die [Hauptstadt Spaniens](#) und der [Autonome Gemeinschaft Madrid](#).

Der Großraum Madrid zählt mit etwa sieben Millionen Einwohnern zu den größten [Metropolen Europas](#). Madrid ist (ohne Vororte) mit rund 3,2 Millionen Einwohnern nach London und Berlin die [drittgrößte Stadt der Europäischen Union](#) und die [größte Stadt Südeuropas](#).

Madrid ist seit Jahrhunderten der geographische, politische und kulturelle Mittelpunkt Spaniens (siehe [Kastilien](#)) und der Sitz der spanischen Regierung. Hier residieren auch der König, ein katholischer [Erzbischof](#) sowie wichtige Verwaltungs- und Militärbehörden. Als Handels- und [Finanzzentrum](#) hat die Stadt nationale und internationale Bedeutung.

In Madrid befinden sich sechs öffentliche [Universitäten](#) sowie verschiedene andere Hochschulen, Theater, Museen und Kultureinrichtungen.

Die Einwohner Madrads nennt man „[Madrilenen](#)“ (span. *madrileños*) oder „[Madriden](#)“.

Inhaltsverzeichnis [[Verbergen](#)]

- [1 Geographie und Klima](#)
 - [1.1 Geographische Lage](#)
 - [1.2 Klima](#)
- [2 Geschichte](#)

Madrid	
Wappen	Karte von Spanien
Basisdaten	
Autonome Gemeinschaft:	Madrid
Provinz:	Madrid
Comarca:	Metropolitana de Madrid
Koordinaten	40° 25' N, 3° 42' W
Höhe:	667 msnm
Fläche:	605,770 km²
Einwohner:	3.207.247 (1. Jan. 2013) ^[1]
Bevölkerungsdichte:	5.294,5 Einw./km²
Postleitzahl:	28001–28080
Gemeindenummer (INE):	28079
Nächster Flughafen:	Flughafen Madrid-Barajas

Internationalization

Extracting multilingual knowledge

The screenshot shows the Italian Wikipedia page for Madrid. The page title is "Madrid". The sidebar on the left lists various language versions of the city, such as English, German, French, Spanish, and others. The main content area features a large administrative index (Indice) with several sections:

- 1 Generalità
- 2 Clima
- 3 Divisione amministrativa
- 4 Origine del nome
- 5 Storia
 - 5.1 Medioevo ed Età Moderna
 - 5.2 L'"Ensanche" e l'Era Industriale
 - 5.3 Seconda Repubblica e Guerra Civile
 - 5.4 Dittatura di Franco
 - 5.5 Democrazia
 - 5.5.1 Terrorismo
- 6 Popolazione
 - 6.1 Popolazione straniera
- 7 Luoghi d'interesse
- 8 Turismo a Madrid
 - 8.1 Luoghi da vedere
 - 8.2 Edifici e monumenti
 - 8.3 Musei e gallerie
 - 8.4 Zone commerciali
 - 8.5 Parchi e giardini

Internationalization

Extracting multilingual knowledge

Wiki Loves Earth in focus during May and June

Koordinaten: 40° 25' N, 3° 42' W |

From

Madrid

Para otros usos de este término, véase [Madrid \(desambiguación\)](#).

Madrid es un municipio y una ciudad de España. La localidad, con categoría histórica de villa,⁴ es la capital de España⁵ y de la Comunidad de Madrid. También conocida como la *Villa y Corte*, es la más poblada del estado, con 3 207 247 de habitantes en el año 2013,^{6 7} mientras que, con la inclusión de su área metropolitana⁸ la cifra de población asciende a 6.543.031 habitantes⁸, siendo por ello la cuarta área metropolitana —por detrás de las de París, Londres y Región del Ruhr— y la tercera ciudad más poblada de la Unión Europea —por detrás de Berlín y Londres—.^{9 10 11 12}

Como capital del Estado, Madrid alberga las sedes del gobierno, Cortes Generales, ministerios, instituciones y organismos asociados, así como la residencia oficial de los Reyes de España¹³ y del Presidente del Gobierno. En el plano económico, es la cuarta ciudad más rica de Europa, tras Londres, París y Moscú.¹⁴ Para el 2009, el 50,1 % de los ingresos de las 5000 principales empresas españolas son generados por sociedades con sede social en Madrid, las cuales representan el 31,8 % de ellas.¹⁵ Es sede del 3^{er} mayor mercado de valores de Europa,¹⁶ del 2º de ámbito latinoamericano (Latibex) y de varias de las más grandes corporaciones del mundo.^{17 18} Es la 8.^a ciudad del mundo con mayor presencia de multinacionales, tras Pekín y por delante de Dubái, París y Nueva York.^{19 20}

En el plano internacional, acoge la sede central de la Organización Mundial del Turismo (OMT), perteneciente a la ONU, la sede de la Organización Internacional de Comisiones de Valores (OICV), la sede de la Secretaría General Iberoamericana (SEGIB), la sede de la Organización de Estados Iberoamericanos para la Educación, la Ciencia y la Cultura (OEI), la Organización Iberoamericana de Juventud (OLJ), y la sede de Public Interest Oversight Board (PIOB).²¹ También alberga las principales instituciones internacionales reguladoras y difusoras del idioma español: la Comisión Permanente de la Asociación de Academias de la Lengua Española,²² y sedes centrales de la Real Academia Española (RAE), del Instituto Cervantes y de la Fundación del Español Urgente (Fundéu). Madrid organiza ferias como FITUR, Madrid Fusión, ARCO, SIMO TCI, el Salón del Automóvil y la Cibeles Madrid Fashion Week.

Es un influyente centro cultural y cuenta con museos de referencia internacional, entre los que destacan el Museo del Prado, el Museo Nacional Centro de Arte Reina Sofía, CaixaForum Madrid y el Thyssen-Bornemisza, que ocupan, respectivamente, el 11º, 15º, 24º y 48º puesto entre los museos más visitados del mundo.²³

Los orígenes de la ciudad son objeto de revisión tras recientes hallazgos de enterramientos visigodos así como de restos que se remontan a los carpetanos o periodo prerromano. Las excavaciones arqueológicas también arrojan restos que se atribuyen al Madrid romano. Estos hallazgos de época visigoda han venido a confirmar que el posterior asentamiento fortificado musulmán de Maṣrīt (del siglo IX) se había asentado sobre un *vicus* visigodo del siglo VII llamado *Matrice* o matriz, arroyo. (AFI [maðsrit])^{24 25}

No sería hasta el siglo XI que Madrid es incorporado a la Corona de Castilla, tras su conquista por Alfonso VI de León y Castilla en 1083. Es designada como sede de la Corte por Felipe II en 1561, convirtiéndose en la primera capital

Madrid

Municipio de Madrid

Bandera

Internationalization

Extracting multilingual knowledge

The screenshot shows a Wikipedia article for "Madrid". The main content discusses Madrid as a municipality and city in Spain, its status as the capital, and its role as a global economic center. It mentions the city's international influence, including its role as a major tourism hub and its presence in various international organizations. The sidebar on the left lists categories such as "From", "T", "Madrid", "Da Wi", "Madi", "Auto", "Madri", "abitar", "Der C", "größl", "Einw", "Ruhr", "Unio", "Madr", "Botel", "Madr", "The I", "Mitte", "Regi", "wichti", "4 Ori", "5 Sto", "In Ma", "ande", "Madr", "Natic", "„Mad", "Die E", "6 Poj", "Inh", "1 G", and "2 G". On the right, there is a "Wikimapia" section with a map of Madrid, coordinates (40° 25' N, 3° 42' W), and a "Wikivoyage" link. Below the map, there is a "Bandera" section with the Madrid coat of arms, and a series of five small images showing landmarks like the Royal Palace and the Metropolis building.

Mapped by
chapters

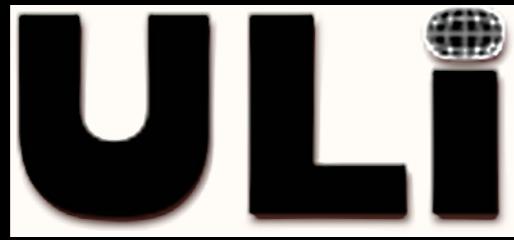
\$lang.Dbpedia.org

dbpedia internationalization

USE CASES

industrial use case

Abbrev. base



LIST OF ABBREVIATIONS

absol.	Absolutely.	v.	Verb.
acc.	Accusative.	v. trans.	Verb transitive.
act.	Active.	v. intrans.	Verb intransitive.
adj.	Adjective.	voc.	Vocative.
adv.	Adverb.	Xen.	Xenophon.
Æsch.	Æschylus.	TITLES OF ÆSCHYLUS' PLAYS.	
Andoc.	Andocides.	Ag.	Agamemnon.
aor.	Aorist.	Choe.	Choephoree.
Ar.	Aristophanes.	Eum.	Eumenides.
Arist.	Aristotle.	Pers.	Persae.
conj.	Conjunction.	P.V.	Prometheus Vinctus.
cp.	Compare.	Supp.	Supplices.
dat.	Dative.	Theb.	Septem Contra Thebas.
Dem.	Demosthenes.	TITLES OF ARISTOPHANES' PLAYS.	
Eur.	Euripides.	Ach.	Acharnians.
fem.	Feminine.	Av.	Aves.
frag.	Fragment.	Eccl.	Ecclesiazusae.
fut.	Future.	Eq.	Equites.
gen.	Genitive.	Lys.	Lysistrata.
Hdt.	Herodotus.	Nub.	Nubes.
imperf.	Imperfect.	Pl.	Plutus.
indic.	Indicative.	Ran.	Ranae.
infin.	Infinitive.	Thesm.	Thesmophoriazusae.
interj.	Interjection.	Vesp.	Vespae.
intrans.	Intransitive.	TITLES OF EURIPIDES' PLAYS.	
Isae.	Isaeus.	Alc.	Alcestis.
Isoc.	Isocrates.	And.	Andromache.
lit.	Literally.	Bacc.	Bacchae.
Lys.	Lysias.	Cycl.	Cyclops.
masc.	Masculine.	El.	Electra.
met.	Metaphorically.	Hec.	Hecuba.
mid.	Middle.	Hel.	Helen.
neut.	Neuter.	Heracl.	Heraclidae.
opt.	Optative.	H.F.	Hercules Furens.
P.	Prose.	Hipp.	Hippolytus.
part.	Participle.	I.A.	Iphigenia in Aulis.
perf.	Perfect.	I.T.	Iphigenia in Tauris.
pl.	Plural.	Med.	Medea.
Plat.	Plato.	Or.	Orestes.
prep.	Preposition.	Phoen.	Phoenissae.
pres.	Present.	Rhes.	Rhesus.
pro.	Pronoun.	Supp.	Supplices.
sing.	Singular.	Tro.	Troades.
Soph.	Sophocles.		
subj.	Subjunctive.		
subs.	Substantive.		
Thuc.	Thucydides.		
V.	Verse.		

what?

Multilingual Knowledge base of abbreviations

why?

Help in text segmentation in the form of exceptions to segmentation rules

how?

Extract words that look like sentence boundaries, model via **Lemon**

DBPEDIA INTERNATIONALIZATION THE ITALIAN JOB

Marco Fossati
fossati@fbk.eu



**GAZETTEER - INDEX
of
THE WORLD**

**INDUSTRIAL USE
CASE
BUILDING
HUGE
GAZETTEERS**



SPAZIODATI
ALL YOU NEED IS DATA

Country	Page	Index	Capital	Area In Sq. Miles	Population	Country	Page	Index	Capital	Area In Sq. Miles	Population	
Abyssinia, see Ethiopia.	49	H-5		73	48,338	Dominican Republic	16	H-6	Ciudad Trujillo (Santo Domingo)	19,325	1,616,361	
Aden	32	E-7	Aden	42,000	100,000	Dutch East Indies, see Netherlands Indies	38	E-6				
Aden Protectorate	32	E-7	Aden	1,038	140,848	East Indies	38	E-6				
Angra Is.	29	F-7		280,000	12,000,000	Ecuador	17	H-4	Quito	25,000	3,200,000	
Afghanistan	32	I-3	Kabul			Egypt	49	G-3	Cairo	383,000	13,820,213	
Africa	40			11,378,730	135,000,000	Eire	19	H-4	Baile Atha Cliath (Dublin)	26,650	2,968,429	
Alabama	11	I-4	Montgomery	31,279	2,812,961	England	19	E-4	London	59,874	37,794,008	
Alaska	12		Junior	59,270	1,200	Erica	40	E-4	Asmara	68,000	1,690,000	
Albania	25	G-5	Tirana	10,523	1,050,000	Estonia	26	H-4	Tallinn (Reval)	18,358	1,127,885	
Alberta	13	H-3	Edmonton	253,283	769,009	Ethiopia	40	H-5	Addis Ababa	250,000	12,000,000	
Alexand Is.	12				Europe, Southern	20-21						
Algeria	40	K-3	Algiers (Algiers)	847,870	6,910,000	Europe, Northern	22-23					
Andorra	33	G-5	Port Blair	2,508	14,272	Europe, Axis Dominated	41					
Angola (Port. West Africa)	40	K-7	Andorra	191	5,251	3,813,291						409,466,408
Anguilla	28	C-2	Chortiaw	957,500	6,342,477	Faroe Is., The	22	H-1				
Antilles	16					Falkland Is.	17	I-14				
Antigua	40	G-6	St. John	198	38,527	Federated Malay States	23	H-5				
Aracaju	32	D-5		1,000,000	10,000,000	Fernando Po	40	E-5	Santa Isabel	771	20,870	
Argentina	17	I-10	Nova Lisboa (Humbo)	487,788	3,098,281	Fiji Is.	37	G-6	Fiji Is.	7,083	210,630	
Arizona	21	I-4	Hue	30,738	5,122,249	Finland	26	H-3	Helsinki (Helsingfors)	118,386	3,863,733	
Arkansas	21	H-4			Florida	11	J-3	Tallahassee	54,851	24,200		
Ascension Is.	40	G-6			Foomoo, see Taiwan	24	H-7					
Ashanti	40	D-5	Georgetown	34	158	France	27		Paris and Vicinity	212,070	41,040,000	
Asia	31		Kumasi	24,379	278,702	French Equatorial Africa	40	E-5	Brazzaville	912,048	3,325,629	
Astr. and Saudi Arabia	32	D-6			French Indo-China	28	C-3	Hanoi	295,422	23,833,300		
Australia (Including Tasmania)	39				French Oceania	37	C-6	Papeete	1,520	40,392		
Azores Is.	18	F-4	Canberra	2,974,581	6,929,091	French Sudan	40	D-4	Bamako	383,357	3,649,883	
Bahama Is.	16	E-3	Nomaoi	4,404	68,905	French West Africa	40	D-4	Dakar	1,702,220	14,875,978	
Bahrain	32	F-4	Mamuna	213	120,000	Galapagos Is.	9	J-5	Nova Zem.	2,868	2,047	
Baleno Is.	20	G-8	Palma	1,938	276,735	Gambia	40	C-4	Portsmouth	4,033	199,520	
Baluchistan	23	A-3	Quetta	124,000	380,517	Georgia	11	J-4	Atlanta	38,725	3,133,723	
Barbados	16	N-8	Bridgetown	166	295,945	Germany (incl. Austria)	20	L-5	Berlin	214,107	79,364,408	
Barbuda Is.	16	M-5			Gibraltar	20	E-8	Gibraltar	176	21,847		
Basutoland	40	G-8	Maseru	11,716	270,000	Galapagos Is.	9	J-5	Hanau	200	34,337	
Bechuanaland	40	G-8	Mafeking	275,000	161,000	Gambia	40	C-4	Havana	4,033	199,520	
Belgian Congo	40	G-6	Leopoldville	902,082	30,326,191	Greenland	19	O-2	Helsinki	295,218	17,000	
Belgium	20	G-5	Brussels (Brussels)	11,775	8,386,553	Guinea	16	M-8	Hong Kong	3,320	84,888	
Bermuda Is.	18	B-4	Hamilton	19,3	31,051	Guinea-Bissau	16	M-6	Buenos Aires	688	267,407	
Bhutan	51	K-6	Furnaka	18,000	300,000	Guanacaste	37	D-3	Agana	225	22,843	
Bismarck Archipelago	27	F-5	Ratulai	12,600	129,072	Guanacast.	19	D-3	Guatemala	45,452	3,284,269	
Bohemian-Moravian	21	J-6	Hangang	20,000	10,522,010	Guanica, British	17	L-3	Georgetown	89,480	341,237	
Bolivia	17	H-7	Sucre & La Paz	395,732	3,226,296	Guanica, Dutch						
Borneo, Dutch	28	D-6			Guanica, Spanish	17	L-3	Cayenne	34,740	22,159		
Brazil	17	J-5	Rio de Janeiro	2,375,510	2,168,951	Guanica, French	17	J-3	Conakry	89,436	2,044,010	
British Columbia	13	D-3	Victoria	372,639	46,115,826	Guanica, Portuguese	19	C-4	Rolama	13,944	364,929	
British Honduras	15	C-2	Belize	8,308	38,759	Guinea, Spanish	40	E-5	Port au Prince	10,294	3,000,000	
British Isles	19			123,600	50,715,000	Guinea	18	J-5	Hainan	22,359	801,090	
British Malaya	38	B-5	Singapore	51,001	4,413,530	Hungary	19	C-2	Honolulu	6,407	423,330	
British North Borneo	28	E-4	Jesson	31,006	2,905,000	Iceland	32	C-4	Mexico	180,000	1,860,000	
Bruce	28	D-5	Bromet	2,300	32,071	India	11	D-2	London	94,294	46,225,444	
Bulgaria	29	K-4	Sofya (Sofia)	43,825	6,500,000	Indonesia	22	H-5	Athens (Athina)	50,447	7,108,000	
Burma	32	H-3	Hango	261,619	14,867,146	Ireland	29	D-3	London	94,294	46,225,444	
California	11	C-8	Santamento	163,612	6,267,287	Iceland	19	C-2	London	94,294	46,225,444	
Camodia	38	H-3	Puerto-Peñ	67,559	2,905,000	Iran (Persia)	18	J-2	London	94,294	46,225,444	
Camerons, British	40	F-5		34,559	585,157	Iceland	19	C-2	London	94,294	46,225,444	
Camerons, French	40	F-5	Tacouide	165,459	2,200,000	Iran Free State, see Elze	19	B-4	London	94,294	46,225,444	
Canada	18	E-6	Ottawa	3,694,853	11,200,000	Iceland	19	D-3	London	94,294	46,225,444	
Canal Zone	15	H-6	Balbo Heights	569	81,527	Indonesia	22	H-5	London	94,294	46,225,444	
Capri Is.	40	C-3	Stra. Cras de Tenerife	2,907	599,183	India	11	D-2	London	94,294	46,225,444	
Cape Verde Is.	18	F-6	Cape Town	270,945	3,520,000	India	11	D-2	London	94,294	46,225,444	
Caroline Is.	37	D-4	Fins	1,507	163,182	India	11	D-2	London	94,294	46,225,444	
Cayman Is.	16	O-5	Georgetown	104	6,009	India	11	D-2	London	94,294	46,225,444	
Cebbes	38	N-5	Mandal	72,000	251,006	India	11	D-2	London	94,294	46,225,444	
Central America	15		Colombia	29,812	5,022,000	India	11	D-2	London	94,294	46,225,444	
Ceylon	33	E-4		75	93,256	India	11	D-2	London	94,294	46,225,444	
Chilean Is.	19	I-2	Hartford	4,820	1,700,242	International Zone	40	D-2	Tangier	225	41,000	
China, Proper	34	E-6		3	1,142	Iceland	11	J-2	Des Moines	55,586	2,518,268	
Chosen (Korea)	35	F-5	Bogota	438,997	8,701,816	Iceland	11	J-2	Tokyo	23,071	9,160,232	
Christian Is.	33	C-8	Denver	103,058	1,123,296	Iceland	11	J-2	Baghdad	136,600	3,560,498	
Cochin China	33	C-4		780	133,962	Iceland	11	J-2	Bahia (Athia Clath)	26,601	3,033,000	
Corse Is.	9	A-6			Kazakhstan	19	D-3	Douglas	221	48,398		
Colombia	17	F-3			Kazakhstan	19	D-3					
Comoro Is.	11	K-3			Kazakhstan	19	D-3					
Connecticut	40	I-7			Kazakhstan	19	D-3					
Cook Is.	37	J-7			Kazakhstan	19	D-3					
Corsia	27	I-7			Kazakhstan	19	D-3					
Costa Rica	15	K-5	Ajedza	3,367	297,255	Kazakhstan	19	D-3				
Cote d'Ivoire, see Ivory Coast	29	I-8	San Jose	25,000	639,197	Kazakhstan	19	D-3				
Croatia	29	A-3	Zagreb	40	1,021	Kazakhstan	19	D-3				

WHAT?

linguistic resource

language, domain-specific

WHY?

natural language understanding

HOW?

the simplest query

USERS THE OPEN DATA LANDSCAPE



open government

OPENCOESIONE.GOV.IT

OPENCOESIONE

Verso un migliore uso delle risorse: scopri, segui, sollecita.



Dipartimento per lo Sviluppo
e la Coesione Economica

Ex art. 10 del decreto legge n.101/2013
convertito, con modificazioni, dalla legge
n.125/2013 (in G.U. 30/10/2013, n.255)

Progetti ▾ Soggetti ▾ Risorse ▾

Progetti ▾

Cerca tra i progetti



FINANZIAMENTI MONITORATI

**74,9
MILIARDI
DI EURO**

PAGAMENTI MONITORATI

**29,7
MILIARDI
DI EURO**

PROGETTI MONITORATI

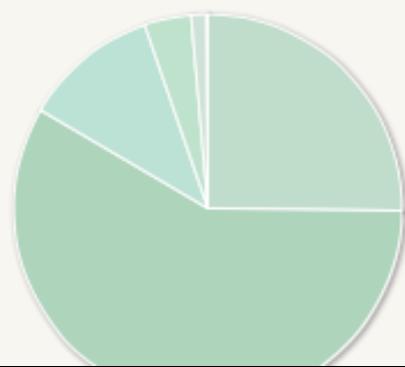
749.112

OpenCoesione è il primo portale sull'attuazione degli investimenti programmati nel ciclo 2007-2013 da Regioni e amministrazioni centrali dello Stato con le risorse per la coesione. I dati sono pubblicati perché i cittadini possano valutare se i progetti corrispondono ai loro bisogni e se le risorse vengono impegnate in modo efficace. [Leggi qui](#) per saperne di più.
I dati pubblicati sono aggiornati al 31/12/2013.

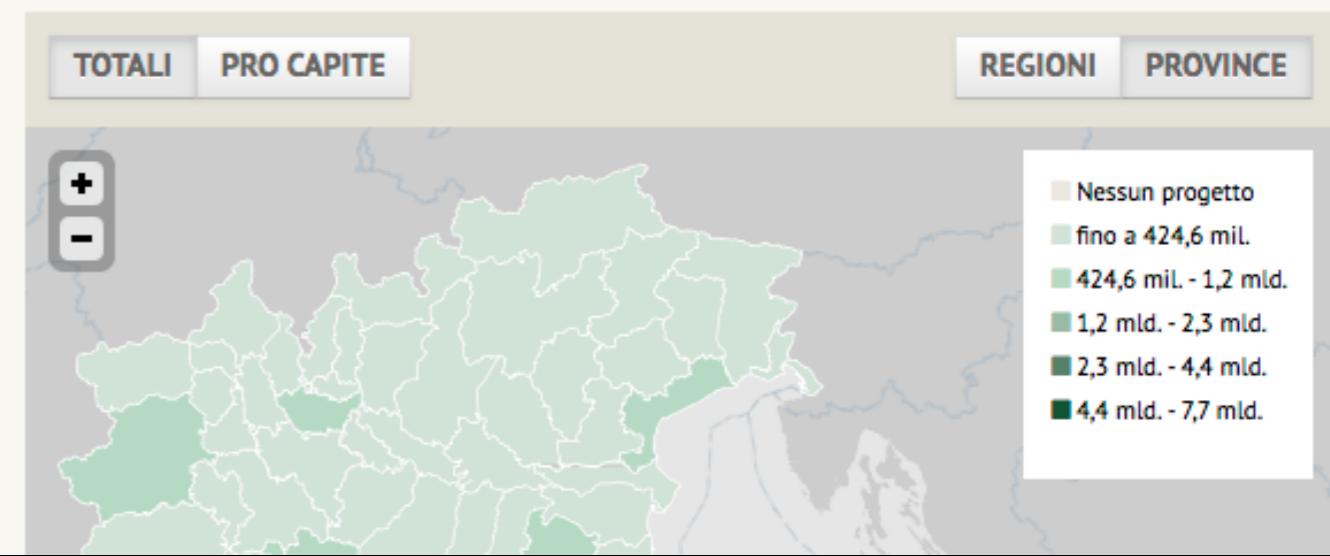
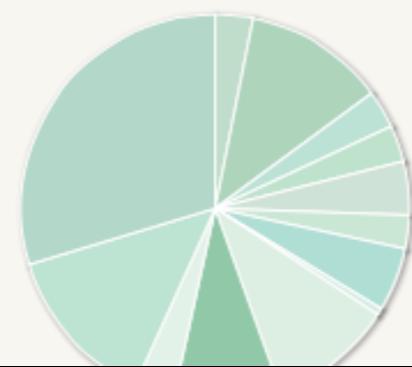
RISORSE TOTALI 2007-2013

**99,286
MILIARDI
DI EURO**

Natura dell'investimento
Cosa si fa con i progetti?



Temi
In quali settori si interviene?



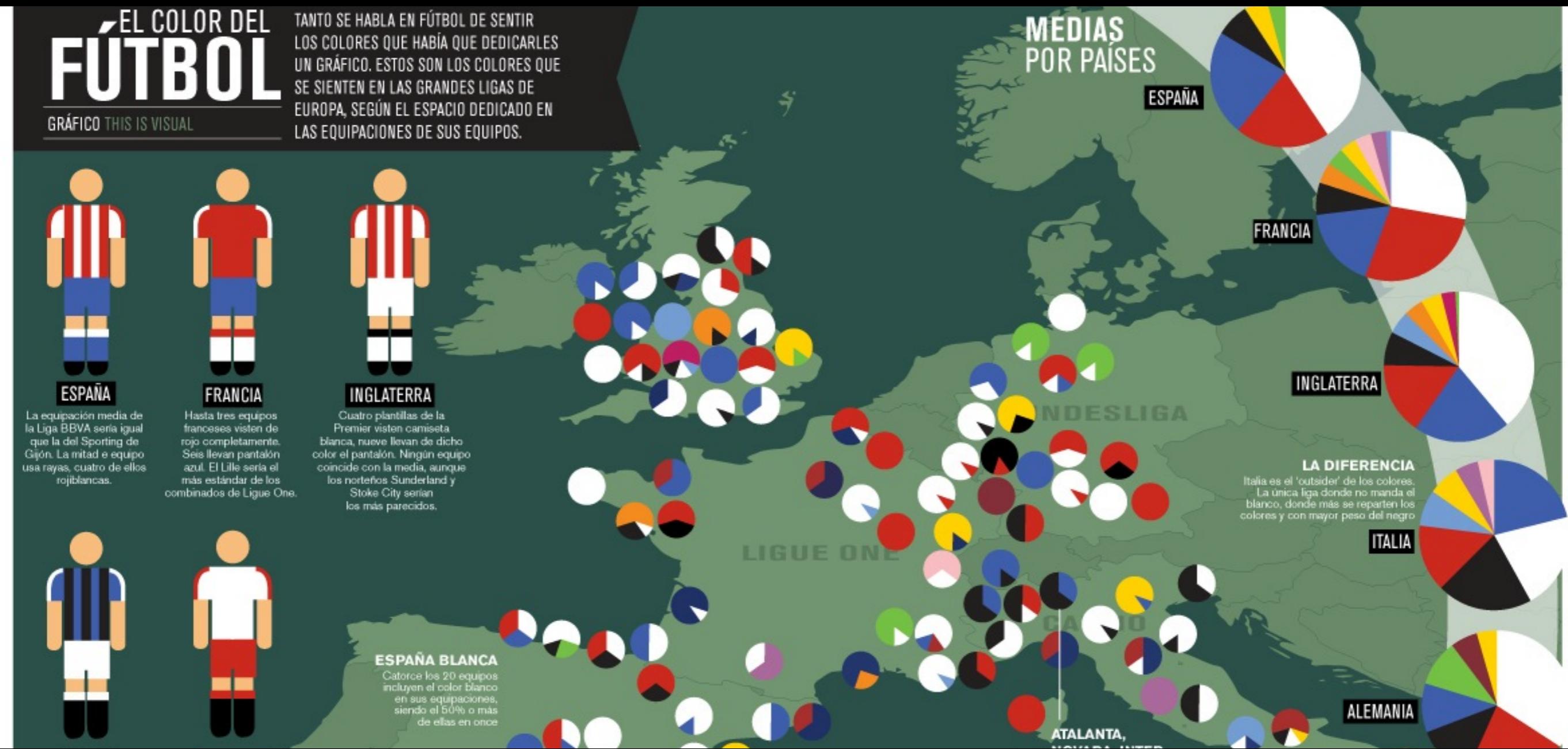
digital libraries

FLORENCE NATIONAL LIBRARY



data-driven journalism

INFOGRAPHICS



THE FIRST ITALIAN DBPEDIA MAPPING SPRINT STUDENTS LEARN HOW TO TRANSLATE A CULTURE



WHAT?

mapping **italian** data to the dbpedia ontology

WHY?

High quality, **multilingual** data

HOW?

hackathon in a high school

AND NOW...

THE SPANISH APARTMENT

Marco Fossati
fossati@fbk.eu



DBPEDIA I18N

THE SPANISH JOB



Mariano.Rico@upm.es

DBPEDIA I18N

THE SPANISH JOB

MEXICAN

ARGENTINIAN

COLOMBIAN

... (UP TO 22)



Mariano.Rico@upm.es

WIKIPEDIA LANGUAGES

Ranking:

(As of 29th Jan. 2014)

- 1.- English (4.4 M)
- 2.- German (1.7 M)
- 3.- French (1.5 M)
- 4.- Italian (1.1M)
- Russian (1.1M)
- Spanish (1.1M)
- Polish (1.1M)
- 5.- Japanese (0.9)
- 6.- Portuguese (0.8M)
- 7.- Chinese (0.8M)

WIKIPEDIA

English

The Free Encyclopedia

4 433 000+ articles

日本語

フリー百科事典

892 000+ 記事

Deutsch

Die freie Enzyklopädie

1 680 000+ Artikel

Português

A enciclopédia livre

817 000+ artigos

Polski

Wolna encyklopedia

1 025 000+ haset

Español

La enciclopedia libre

1 075 000+ artículos

Русский

Свободная энциклопедия

1 083 000+ статей

Français

L'encyclopédie libre

1 470 000+ articles

Italiano

L'enciclopedia libera

1 094 000+ voci

中文

自由的百科全書

747 000+ 條目



1 052 000+ 頁面

Wolna encyklopedia

پاکیزہ

811 000+ 文章

1 411 000+ 頁目

自由的百科全書

中文

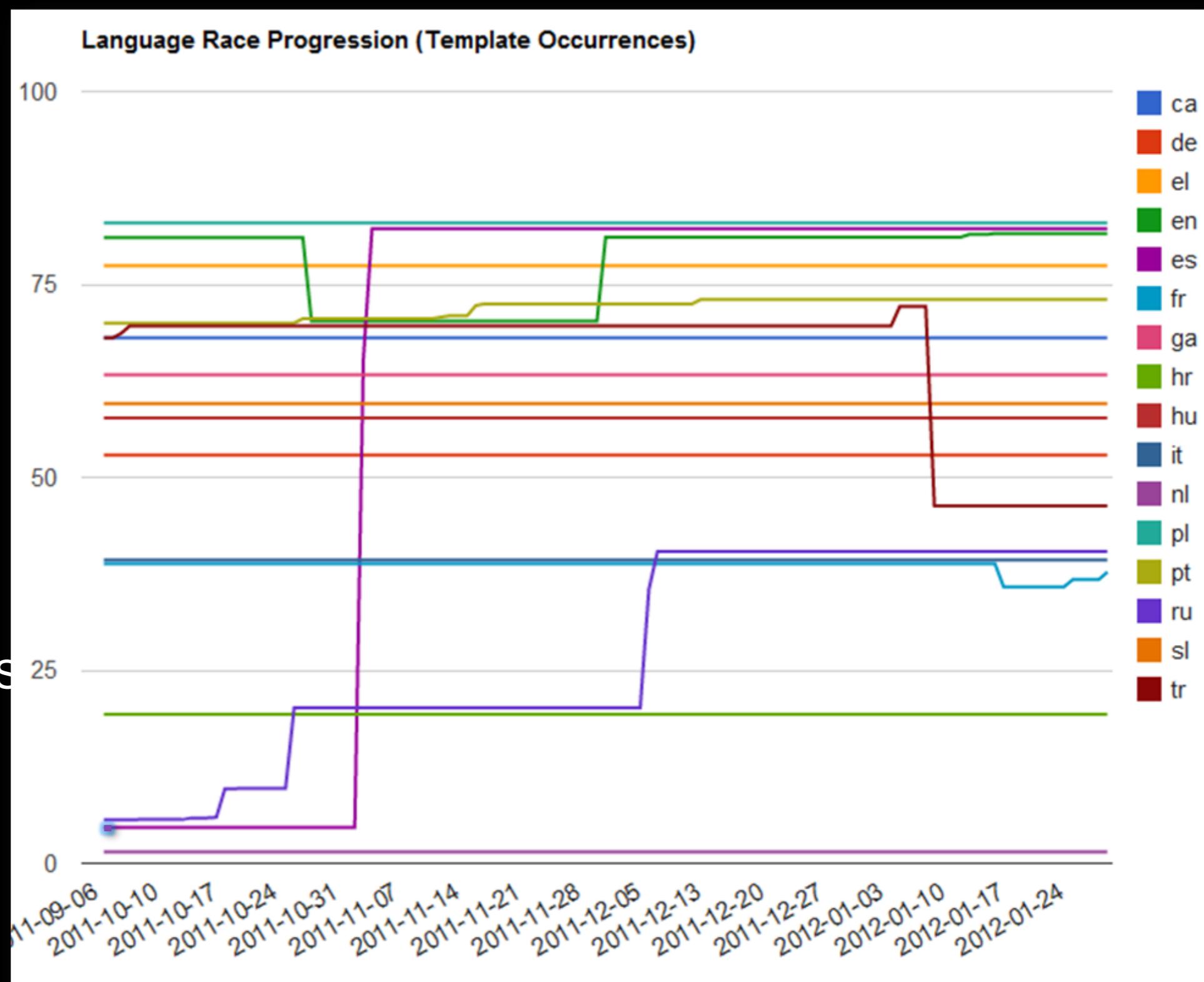
1 084 000+ 文章

MAPPING RACE

2011

ESDBPEDIA HACKATON (NOV. 2011)

15 PEOPLE
4H 4H
101 CLASSES
MAPPED
80% INSTANCES
MAPPED



ESDBPEDIA: THE WEB SITE

ENGLISH HUMANS

[Tweet 109](#) [Follow @esDBpedia 155 followers](#) [g+ 11](#) [Like](#) [Share](#) You, Freddy Priyatna and 30 others like this.

[FOLLOW US ON FACEBOOK](#)

Welcome

Your trail:

[View](#) [Attach](#) [Info](#) [Edit](#) [More... ▾](#)

[Versión en español](#)

Welcome! #

You are at the Spanish chapter of DBpedia.

In a hurry? This is a query over the Spanish dbpedia: "[What bullfighters were married to copla singers?](#)" ([copla](#) is a traditional Spanish style, existing as a category in the Spanish Wikipedia).

Do you want to see it slowly? Copy and paste in the Spanish [SPARQL endpoint](#) the query text you see in the following box.

PREFIX dcterms: <http://purl.org/dc/terms/>
SELECT ?torero ?cantante WHERE{
 ?torero rdf:type dbpedia-owl:BullFighter .
 ?torero dbpedia-owl:spouse ?cantante .
 ?cantante dcterms:subject <http://es.dbpedia.org/resource/Categoría:Cantantes_de_coplas>
}

Click the "Run Query" button. As you can see, this is the answer:

torero	cantante
http://es.dbpedia.org/resource/José_Ortega_Cano	http://es.dbpedia.org/resource/Rocío_Jurado
http://es.dbpedia.org/resource/Curro_Romero	http://es.dbpedia.org/resource/Concha_Márquez_Piquer
http://es.dbpedia.org/resource/Francisco_Rivera	http://es.dbpedia.org/resource/Isabel_Pantoja

[More examples](#) and [Open Questions](#) #

Quick Navigation [Search](#)

G'day, [Mariano Rico](#) (not logged in) [Log in](#) [My Prefs](#)

DBpedia español

UNIVERSIDAD AUTONOMA DE MADRID

Ontology Engineering Group

ESDBPEDIA: THE WEB SITE

SPANISH HUMANS

[Tweet 8](#) [Follow @esDBpedia 155 followers](#) [g+1 3](#) [FOLLOW US ON FACEBOOK](#)

[Like](#) [Share](#) 33 people like this. Be the first of your friends.

Bienvenido

Your trail: [Welcome](#)

[View](#) [Attach \(4\)](#) [Info](#) [Edit](#) [More...](#)

[English version](#)

iBienvenido!

Gracias a todos los que habéis enviado cartas de apoyo para la [propuesta de nueva infraestructura](#)

Para evitar spammers, sólo los usuarios registrados pueden editar las páginas

Este es la sección del idioma español de la DBpedia.

¿Tienes prisa?. Esta es una consulta sobre la dbpedia del español: ["¿Qué toreros se casaron con cantantes de copla?"](#).

Lo puedes ver más despacio?

1) Selecciona y copia el texto de la consulta que aparece en el siguiente recuadro (expresada en lenguaje SPARQL):

```
PREFIX dcterms: <http://purl.org/dc/terms/>
SELECT ?torero ?cantante WHERE{
  ?torero rdf:type dbpedia-owl:BullFighter .
  ?torero dbpedia-owl:spouse ?cantante .
  ?cantante dcterms:subject <http://es.dbpedia.org/resource/Categoría:Cantantes_de_coplas>
}
```

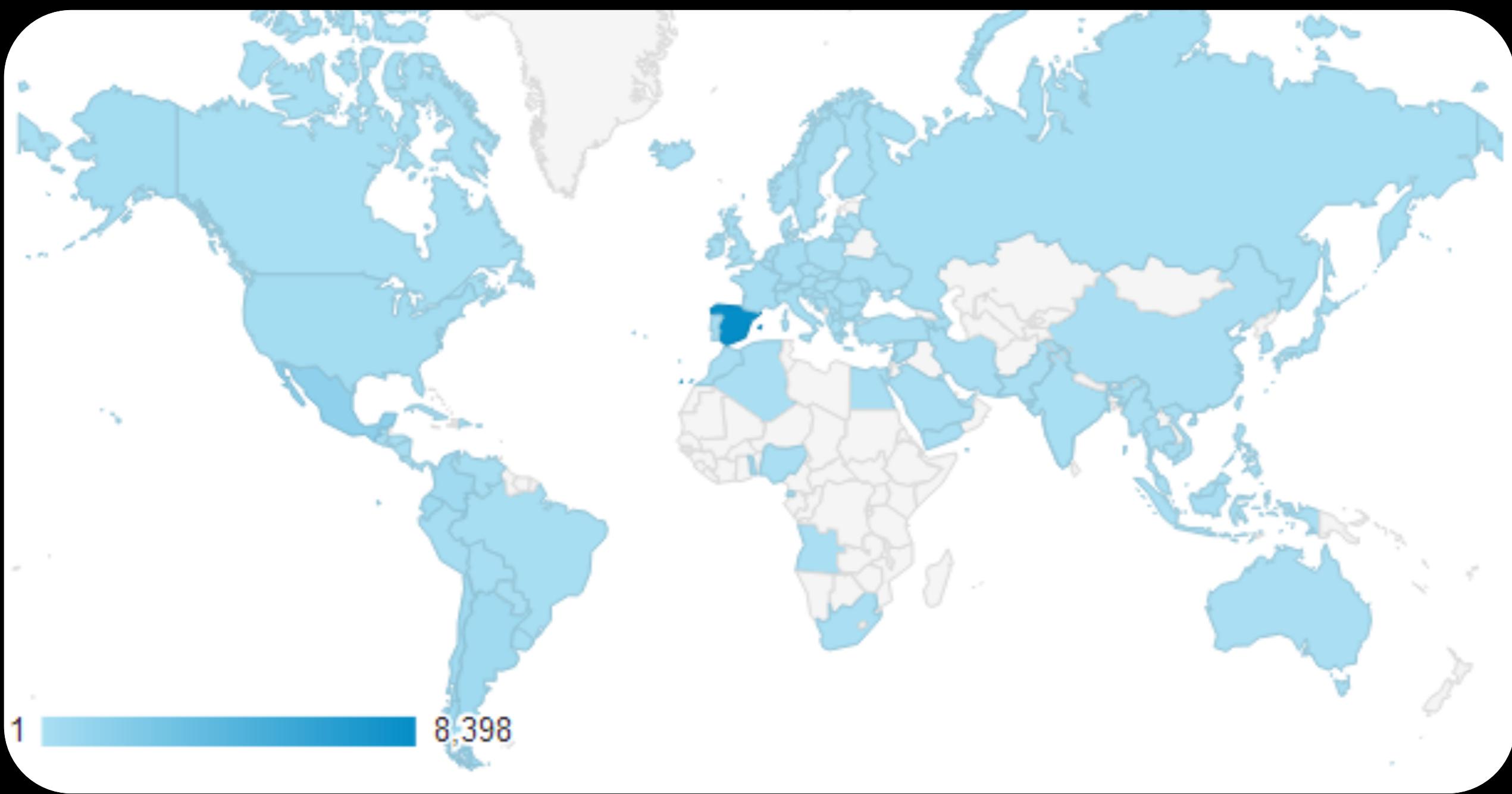
2) Pégalo en el [SPARQL endpoint](#) y pulsa sobre el botón "Run Query".

Como has visto, esta es la respuesta:

torero	cantante
http://es.dbpedia.org/resource/José_Ortega_Cano	http://es.dbpedia.org/resource/Rocío_Jurado
http://es.dbpedia.org/resource/Curro_Romero	http://es.dbpedia.org/resource/Concha_Márquez_Piquer
http://es.dbpedia.org/resource/Francisco_Rivera	http://es.dbpedia.org/resource/Isabel_Pantoja

ESDBPEDIA: THE WEB SITE

LOCATIONS



ESDBPEDIA: THE WEB SITE

LOCATIONS

Spanish
(browser)

users:
78%
(10048 in 12686)

English
(browser)

users:
16%
(2091 in 12686)

No es|
No en
(browser)
users:
5%

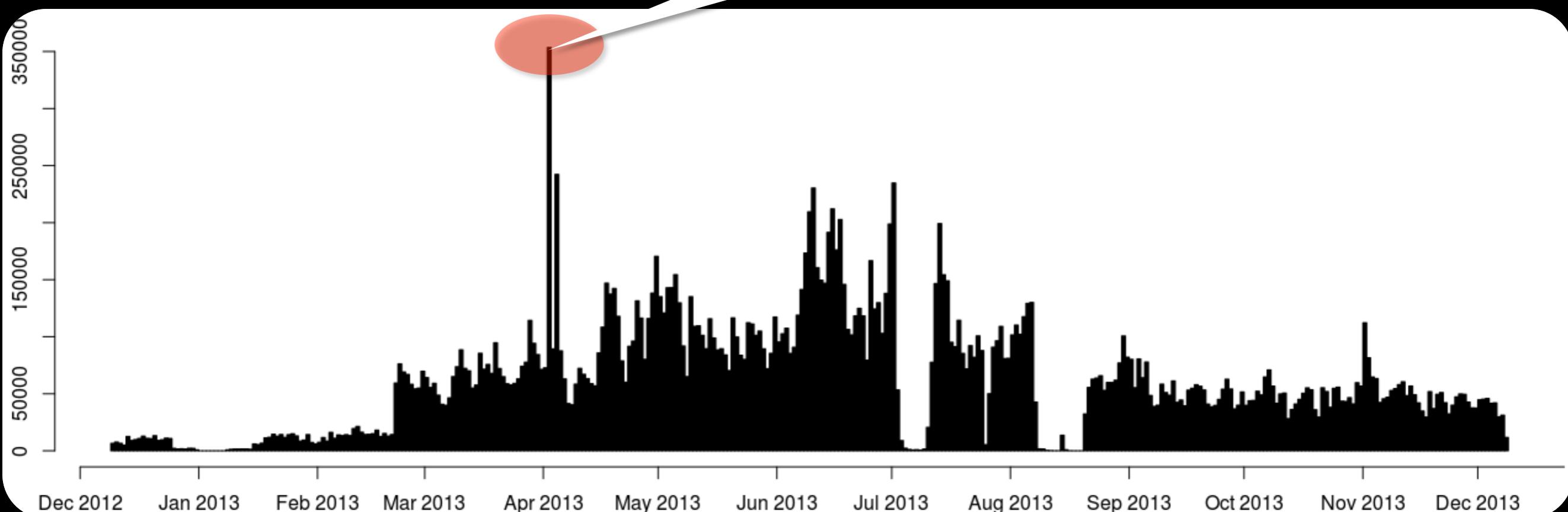
Language	Acquisition			Behavior			
	Visits	% New Visits	New Visits	Bounce Rate	Pages / Visit	Avg. Visit Duration	
	12,686	68.75% % of Total: 100.00% (12,686)	8,721	51.96% % of Total: 100.22% (8,702)	3.17	00:03:21	
1. es	5,452	69.19%	3,772	49.58%	3.31	00:03:28	
2. es-es	3,698	67.47%	2,495	50.22%	3.19	00:03:11	
3. en-us	1,762	65.83%	1,160	57.78%	3.20	00:03:43	
4. es-419	394	66.50%	262	46.70%	3.17	00:04:31	
5. es-mx	270	52.22%	141	56.30%	3.09	00:05:18	
6. en	147	70.07%	103	68.71%	2.27	00:01:31	
7. es-ar	143	70.63%	101	45.45%	2.59	00:03:16	
8. ca	109	85.32%	93	42.20%	3.44	00:03:12	
9. fr	91	82.42%	75	63.74%	1.99	00:01:14	
10. en-gb	73	82.19%	60	53.42%	2.77	00:02:58	
10. eu-0p	13	85.10%	80	23.45%	5.11	00:05:28	

ESDBPEDIA: THE SPARQL ENDPOINT

SPARQL QUERIES

22M SPARQL queries
FROM 2200 IPs

Up to 350,000
sparql queries
per day



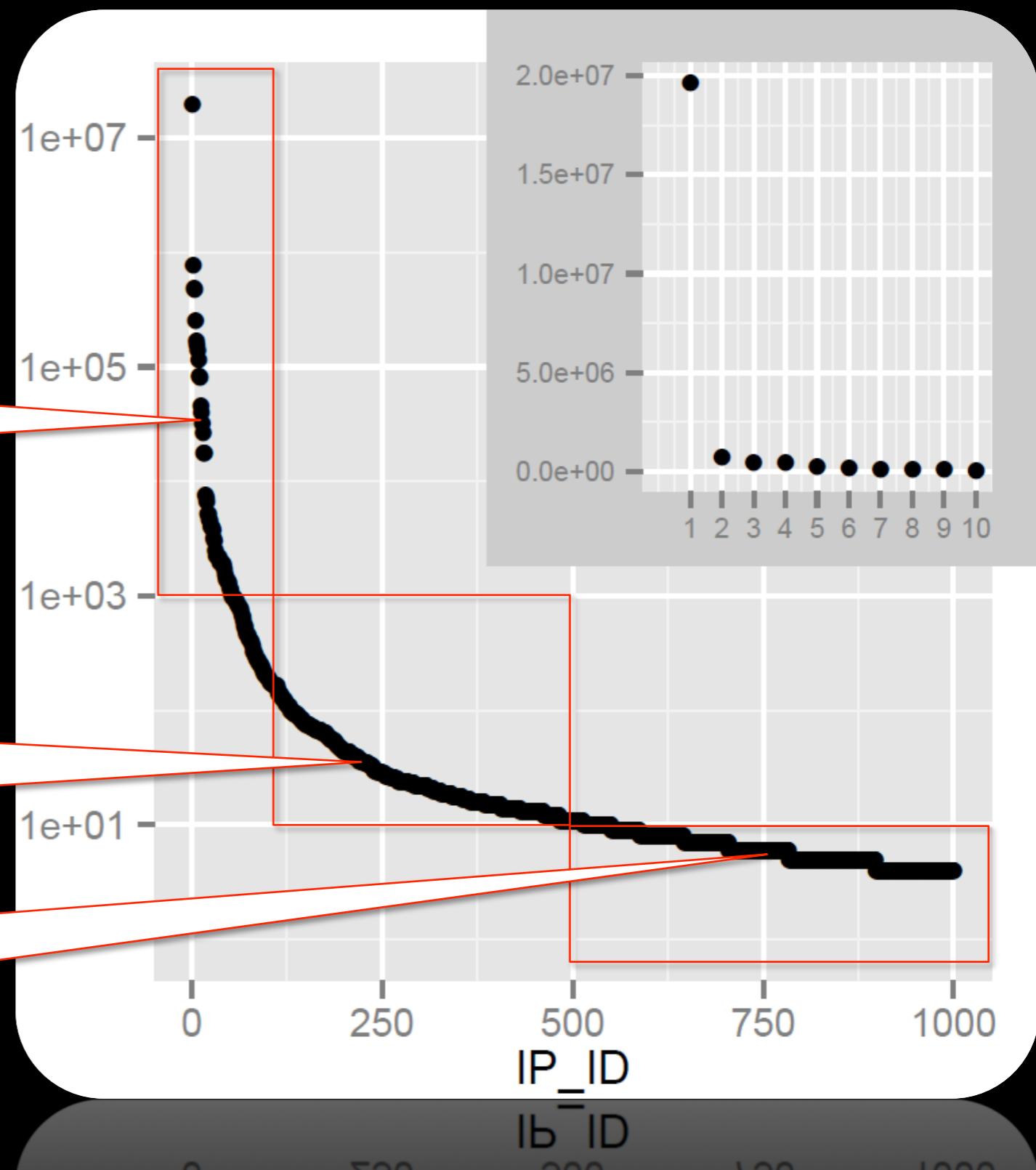
SPARQL QUERIES

22M SPARQL queries
FROM 2200 IP

IPs with more than 10^3
requests: 60

IPs with requests between
 10^3 and 10: 440

IPs with <less than 10
requests: 1700

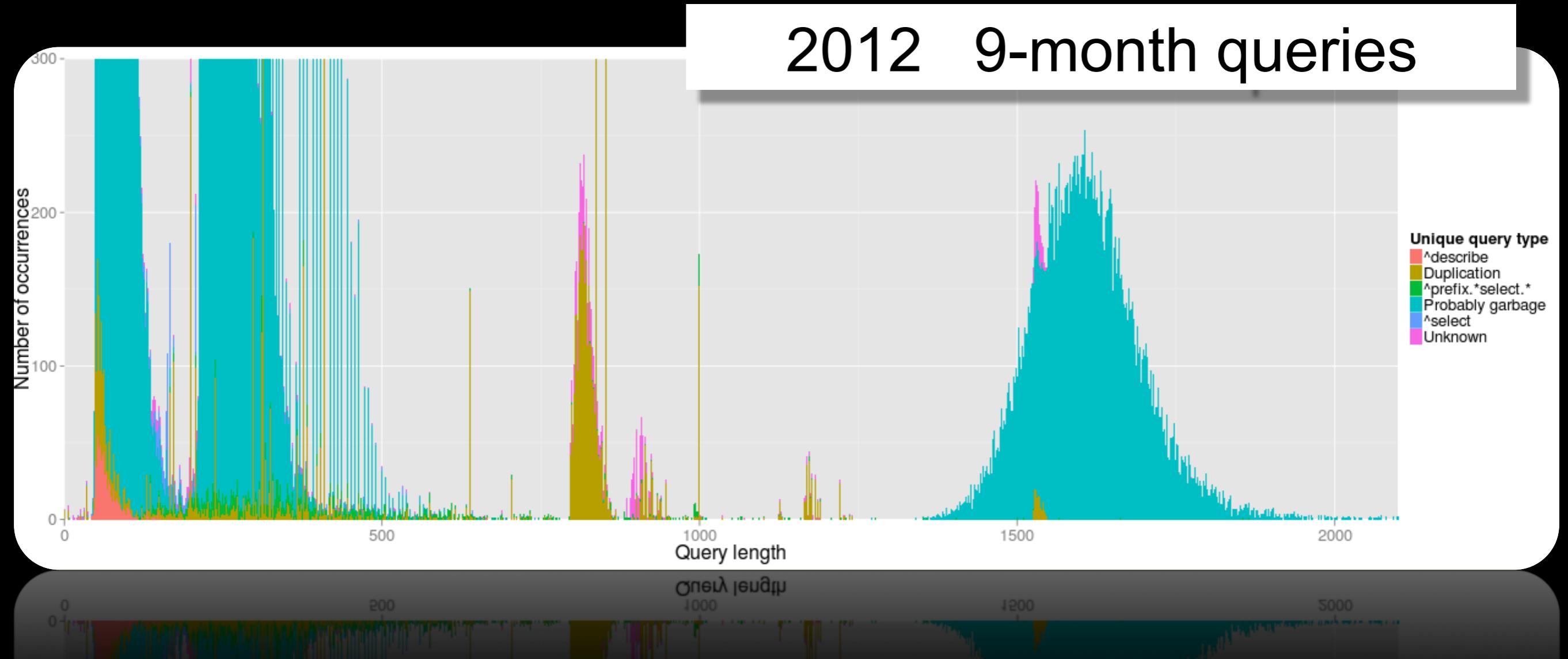


ESDBPEDIA: THE SPARQL ENDPOINT

NOISE GENERATORS

22M SPARQL queries
FROM 2200 IP

2012 9-month queries



LESSONS LE

Lesson 1

Take care of
IP monsters



LESSONS LEARNT

Lesson 2

Take care of
NOISE
GENERATORS



Thanks for your attention!

[HTTP://DBPEDIA.ORG](http://dbpedia.org)



fossati@fbk.eu

Mariano.Rico@upm.es



UNIVERSITÄT LEIPZIG

bruemmer@informatik.uni-leipzig.de