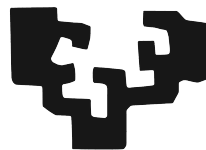


eman ta zabal zazu



UPV EHU



Workflows para 'five-star Open Data' usando infraestructuras libres

David Lindemann david.lindemann@ehu.eu

**I Workshop CLARIAH-CM:
Humanidades Digitales y Tecnologías del Lenguaje**

30 y 31 de mayo de 2024

Contenidos

- ▼ Collections as data
 - ▼ Las cinco estrellas de Berners-Lee
 - ▼ Linked Open Data
 - ▼ Wikidata, Wikibase
- ▼ Workflows (ejemplos)
 - ▼ datos bibliográficos
 - ▼ datos léxicos
- ▼ Dinámica práctica (Open Refine)



<https://etherpad.wikimedia.org/p/31-mayo-2024>

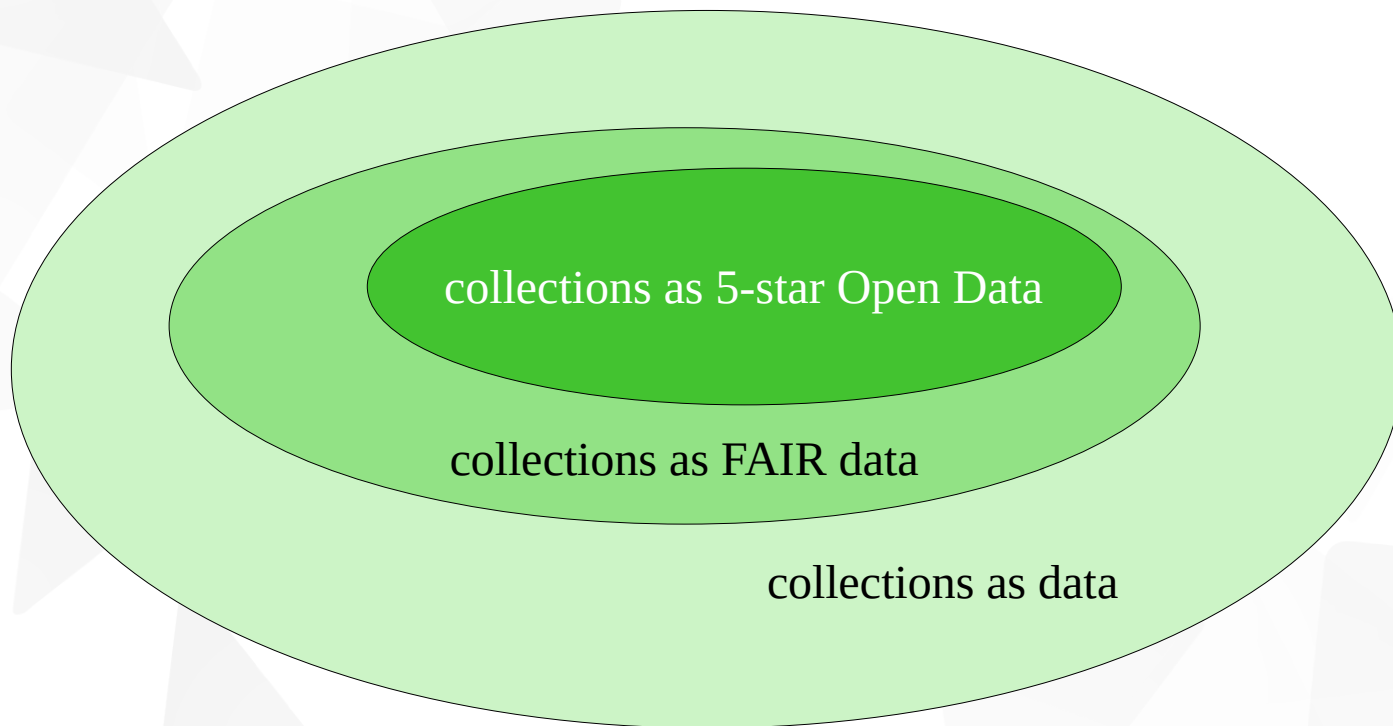
Antes de empezar

- ▼ Descargar una copia de esta presentación
- ▼ Instalar OpenRefine

<https://etherpad.wikimedia.org/p/31-mayo-2024>

- ▼ Sobre mí
 - ▼ Ciencias Políticas > Traducción e Interpretación > Filología Vasca > Lexicografía
 - ▼ Ayudante Doctor en EHU, activo en DARIAH-EU, CLARIAH-EUS
 - ▼ Rol en el Grupo de Investigación: apoyo en metodología HHDD
 - ▼ Wikimedia Basque Country > [Workshop Wikidata \(euskera\)](#) > [Workshop Wikibase \(inglés\)](#)
 - ▼ Implicación en una [serie de instancias de Wikibase](#)

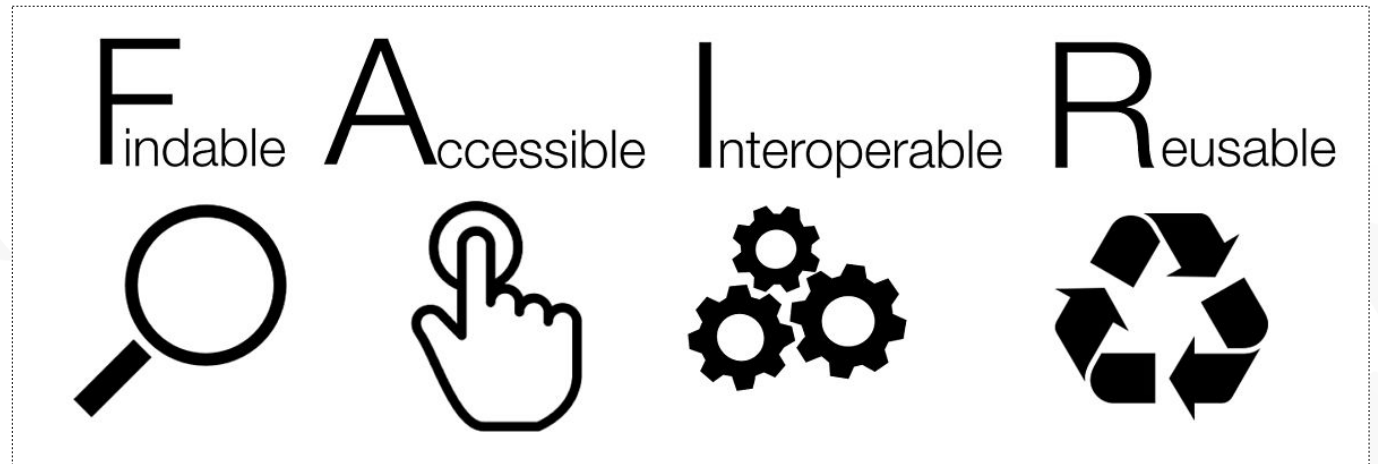
Collections as data (→ seminario de Gustavo Candela)



FAIR data (→ seminario de Gustavo Candela)

- ▼ Findable
- ▼ Accessible
- ▼ Interoperable
- ▼ Re-usable

<https://www.go-fair.org/fair-principles/>



Open Science: FAIR Principles



Findable

Data and materials enriched with metadata assigned with a unique identifier



Accessible

Data and metadata stored in a trusted repository with an open and free protocol. Accessible by machines and humans



Interoperable

Using vocabularies and public domain ontologies the metadata can be referenced and linked



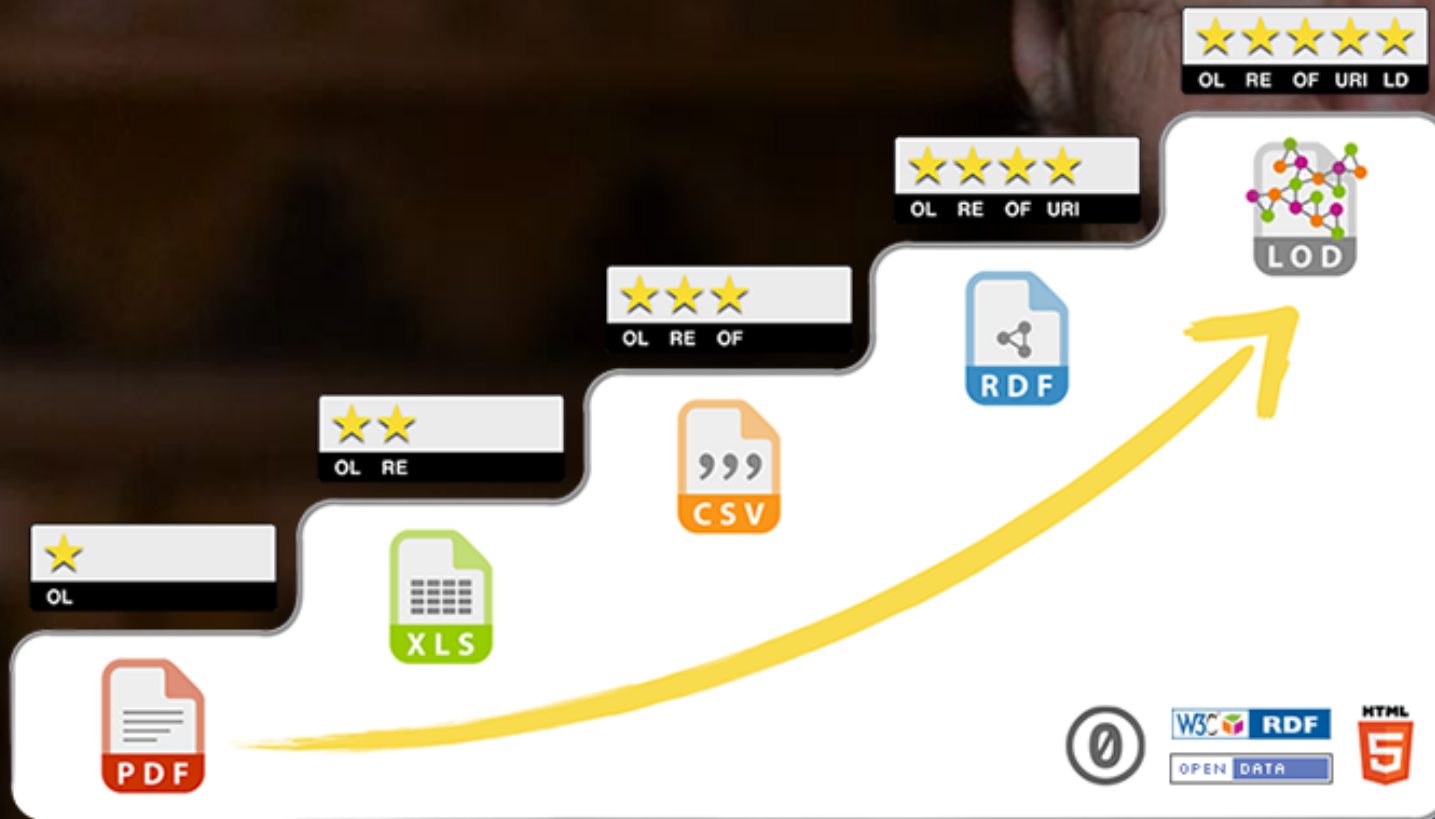
Reusable

Additional documentation and protocols describing the acquisition of the data, licensed with a detailed provenance

5 ★ OPEN DATA

<https://5stardata.info>

Tim Berners-Lee, the inventor of the Web and Linked Data initiator, suggested a 5-star deployment scheme for Open Data. Here, we give examples for each step of the stars and explain costs and benefits that come along with it.



1 estrella: un documento digital



BIBLIOGRAFÍA

OBSERVACIONES: Las reseñas se consignan a continuación del libro (o artículo) reseñado. Cuando éste se ha descrito en una entrega anterior de la Bibliografía, se pone entre paréntesis el número de ficha correspondiente. También se hacen referencias a ediciones anteriores de una obra, a tomos anteriores de una obra en curso de publicación, etc. En estos casos, la cifra que precede al número de referencia indica el volumen de la NRFH en que figura la ficha en cuestión. Así, el "núm. 39-39695" corresponde a la ficha 39695, la cual está en el volumen 39 de la NRFH. Cuando no hay cifra inicial, ha de entenderse que el número remite a una ficha incluida en el presente fascículo.

LINGÜÍSTICA HISPÁNICA

OBRAS GENERALES

- 45-1363. ALARCOS LLORACH, EMILIO – "Apostillas leves a las ponencias". – *EAG*, 61 (1994), 81-86.
- 45-1364. ARIZA VIGUERA, MANUEL – "África y América: dos ejes de la investigación del profesor Germán de Granda". – *ALM*, 34 (1996), 273-290.
- 45-1365. BENIERS, ELIZABETH – "Bibliografía de trabajos descriptivos del español de México". – *ALM*, 34 (1996), 293-349.
- 45-1366. BRUMME, JENNY – "El español moderno y el siglo XIX, en especial, como objeto de estudio en la historia de la lengua. (Balance, lagunas y tareas)". – *Estudis de lingüística i filologia oferts a Antoni M. Badia i Margarit* (Barcelona, 1996-97), t. 1, pp. 131-140.
- 45-1367. BRUNO, LIDIA, y MIRTA SERAFINI – "Lenguaje, idiosincrasia y educación entre los mapuches". – *S&S*, 1995, núm. 4, 223-235.
- 45-1368. CABRÉ, M. TERESA – "Sobre manlleus i diccionaris". – *Estudis de lingüística*

ca i filologia oferts a Antoni M. Badia i Margarit (Barcelona, 1996-97), t. 3, pp. 39-61.

- 45-1369. CORTÉS RODRÍGUEZ, LUIS – *Tendencias actuales en el estudio del español hablado*. – Universidad, Almería, 1994. 153 pp. (*Monografías Humanidades*, 2). || *Verba*, 22 (1995), 677-680 (Casánovas).
- 45-1370. GÓMEZ TORREGO, L. – *Manual del español correcto*. 4ª ed. corr. y aum. – Arco Libros, Madrid, 1994. 2 ts. || *EAG*, 60 (1993), 109 (Albalá).
- 45-1371. *Historia de la alfabetización y la educación para adultos en México*. – SEP-El Colegio de México, México, 1994. xxxi + 772 pp. || *Ciencia y Desarrollo*, México, 1996, núm. 130, 90-92 (Barriga Villanueva).
- 45-1372. LEBSANFT, FRANZ – *Spanische Sprachkultur*. – Max Niemeyer, Tübingen, 1997. xv + 322 pp. (*BZRP*, 282).
- 45-1373. LÓPEZ MORALES, H. – *Métodos de investigación lingüística* (N. 44-22). || *RFE*, 75 (1995), 379-381 (Hernández).
- 45-1374. MACRÌ, ORESTE – *Studi Ispanici*. T. 2: *Critici*. Ed. de Laura Dolfi. – Liguori Editore, Napoli, 1996. 463 pp.

NRFH, XLV (1997), núm. 2, 541-608

arranparrila. *iz.* (1912; *arranbarrilla* *1883, 1894; *aramarrilla* *1905). Paraleloki ezarritako burdin hariz osatutako tresna giderduna, janariak sutan prestatzeko erabiltzen dena. Ik. **grise-la matxarro** Zartagia eta arranparrilak. Okela arranparrilan errea.

arrantxo. *iz.* (*XVIII, 1930). *Heg.* (Beh.). [esp]. Soldaduen jantzia, jatena. *Arrantxo ederra egin dut hementxe jarrita, diru gabe nintzena zeuden igarrita.*

1 arrantza. *iz.* (1627; *arrainza* 1571). Arrainak harrapatzeko jarduna eta lanbidea. Ik. **arrantzu**. *Arrantzara joan. Ur handitako, itsas zabaleko, itsas bazterreko arrantza. Urrutietako arrantza eta merkataritza. Balearen edo bakailaoaren arrantza. Arrantzatik bizi direnak. Arrantzan ikasteko. Arrantzan egin. Ik. arrantzatu* Arrantzan ari direnak. Hamabi marinela, arrantzan usaturik. Bere lagunekin itsas arrantzan aurkitzen zen batez. *Arrantza ontzia. Arrantza handia: itsas zabaleko arrantza. Arrantza txikia: itsasbazterreko arrantza.*

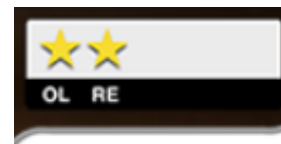
2 arrantza. *iz.* (*1746, 1800). **1.** Astoaren orroa. *Astoaren arrantza min gabea* (esr. zah.). **2.** Negar zaratatsua. *Seme bat hil eta amak aterako ditu buruko ileak, deiadarraz eta arrantzaz.* – ARRANTZA EGIN. (1842). *Astoa arrantza egiten. II Kantatu beharrean egiten arrantza.*

arrantzaka. *nola adlag.* (1842). Arrantza eginez; oihuka. *Astoa arrantzaka hasten denean ez belarri eman. Tramana eta Brix, bata txilioka eta arrantzaka bestea.*

arrantzale. *iz.* (1571; *arrantzale* *XVII ea., ~1660). Arrainak harrapatzen dituen, berez., lanbidetzat arrantza duen pertsona. *Arrantzaleak sareak garbitzen ari ziren. Arrantzale Anaidia. II Giza arrantzale egingo zaituztet.*

- ▶ Si el PDF tiene una capa de texto, admite búsquedas en el texto completo
- ▶ Elementos hipertextuales (referencias cruzadas, enlaces http)

2 estrellas: una tabla de Excel

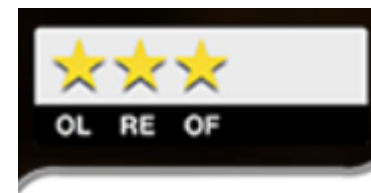


	A	B	C
1	<i>Author</i>	<i>Year</i>	<i>Title</i>
2	Michelena Elissalt, Luis	1953	Apellidos vascos
3	Michelena, Luis	1964	Textos arcaicos vascos
4	Michelena Elissalt, Luis	1970	Estudio sobre las fuentes del Diccionario de Azkue
5	Mitxelena, Koldo; Sarasola, Ibon	1986	Diccionario general vasco - Orotariko euskal hiztegia

	A	B	C
1	<i>Basque lemma</i>	<i>Part of Speech</i>	<i>Spanish equivalent</i>
2	sagar	iz.	manzana
3	ibili	ad.	andar
4	txiki	adj.	pequeño, -a
5	gaur	adb.	hoy

- Los datos tienen una estructura (explicable a la máquina)
- El formato tiene límites
 - Formato propietario
 - No todo es explicable a la máquina (aquí: colores)

3 estrellas: una tabla CSV

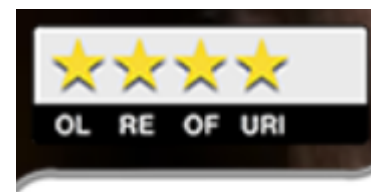


	A	B	C	D
1	<i>Author</i>	<i>Year</i>	<i>Title</i>	<i>status</i>
2	Michelena Elissalt, Luis	1953	Apellidos vascos	green
3	Michelena, Luis	1964	Textos arcaicos vascos	green
4	Michelena Elissalt, Luis	1970	Estudio sobre las fuentes del Diccionario de Azkue	red
5	Mitxelena, Koldo; Sarasola, Ibon	1986	Diccionario general vasco - Orotariko euskal hiztegia	green

```
1 "Author","Year","Title","status"  
2 "Michelena Elissalt, Luis",1953,"Apellidos vascos","green"  
3 "Michelena, Luis",1964,"Textos arcaicos vascos","green"  
4 "Michelena Elissalt, Luis",1970,"Estudio sobre las fuentes del Diccionario de Azkue","red"  
5 "Mitxelena, Koldo; Sarasola, Ibon",1986,"Diccionario general vasco - Orotariko euskal hiztegia","green"
```

- ▼ El formato es universal (no depende de apps concretas)
- ▼ Todo el contenido es interpretable por una máquina

4 estrellas: RDF



	A	B	C
1	<i>Author</i>	<i>Year</i>	<i>Title</i>
2	Michelena Elissalt, Luis	1953	Apellidos vascos
3	Michelena, Luis	1964	Textos arcaicos vascos
4	Michelena Elissalt, Luis	1970	Estudio sobre las fuentes del Diccionario de Azkue
5	Mitxelena, Koldo; Sarasola, Ibon	1986	Diccionario general vasco - Orotariko euskal hiztegia

PREFIX example: <<http://example.org/entity/>>

example:work1 example:author "Michelena Elissalt, Luis"

example:work1 example:publicationYear "1953"

example:work1 example:title "Apellidos vascos"

example:work5 example:author "Mitxelena, Koldo"

example:work5 example:author "Sarasola, Ibon"

example:work5 example:publicationYear "1986"

example:work5 example:title "Diccionario general vasco"@es

example:work5 example:title "Orotariko euskal hiztegia"@eu

▼ Triples semánticos:
sujeto – predicado – objeto

▼ sujeto y predicado están
expresados por
identificadores unívocos

5 estrellas: LOD (Linked Open Data)



```
example:work1 example:author "Michelena Elissalt, Luis"  
example:work1 example:publicationYear "1953"  
example:work1 example:title "Apellidos vascos"  
  
example:work5 example:author "Mitxelena, Koldo"  
example:work5 example:author "Sarasola, Ibon"  
example:work5 example:publicationYear "1986"  
example:work5 example:title "Diccionario general vasco"@es  
example:work5 example:title "Orotariko euskal hiztegia"@eu
```

<http://purl.org/dc/elements/1.1/creator>

PREFIX dc: <<http://purl.org/dc/elements/1.1/>>

```
example:work1 dc:creator "Michelena Elissalt, Luis"  
example:work1 dc:date "1953"  
example:work1 dc:title "Apellidos vascos"  
  
example:work5 dc:creator "Mitxelena, Koldo"  
example:work5 dc:creator "Sarasola, Ibon"  
example:work5 dc:date "1986"  
example:work5 dc:title "Diccionario general vasco"@es  
example:work5 dc:title "Orotariko euskal hiztegia"@eu
```

Datos bibliográficos para consumo humano

Aldizkari Zientifikoen Ataria

Anuario del Seminario de Filología Vasca "Julio de Urquijo"

ERREGISTRATU

ALDIZKARIAREN EDUKIA

Hizkuntza

- > English
- > Español
- > Euskara

Make a Submission

Make a Submission

Informazioa

> Irakurleentzat

CAMPIONEN "GRAMÁTICA" ETA BESTE

PDF

Argitaratua 1992-01-15

DOI <https://doi.org/10.1387/asju.9357>

Koldo Mitxelena

Laburpena

Nola aipatu

Mitxelena, Koldo. 1992. «Campionen "gramática" Eta Beste». *Anuario Del Seminario De Filología Vasca "Julio De Urquijo"*, urtarrilak, /10.1387/asju.9357.

Abstract 50 | PDF Downloads 88

Larramendi, M., 1728, *De la antigüedad y universalidad y ventajas sobre otras muchas Lenguas, Demostrada en la Lengua Vasca*, Salamanca. Ediciones Vascas, Bilbo, 1979.
_____, 1729, *El Imposible Vencido. Arte de la Lengua Vasca*, Hordago, Donostia, 1979.
_____, 1745, *Diccionario Trilingüe del Castellano, Francés y Vasco*, Montero, San Sebastián. *Diccionario Trilingüe de la Lengua Vasca*, San Sebastián, 1853. Lehenengo arg. 1984.
_____, c. 1764, *Corografía de Guipúzcoa*. J.I. Tellechea de Ediciones y Publicaciones, Donostia, 1973.
_____, *Autobiografía y otros escritos*, J.I. Tellechea de Ediciones y Publicaciones, Donostia, 1973.
Lázaro Carreter, F., 1981, *Estudios de Lingüística*, I. Etxepare, Donostia.
Malkiel, Y., 1975, "A Typological Classification of the Languages of the Pyrenean Area", in Housholder & Saporta (arg.), *Typological Studies in Language*, 10: 1-10.
Mitxelena, K., 1970, *Estudio sobre las fuentes del castellano de Vizcaya*, Bilbo.
_____, 1978, "Para una historia de la ortografía vasca", *FLV* XXX, 389-413.

Entrada en repositorio „Open Journal Systems“

▼ Bibliografía en PDF

▼ Layout features sí

▼ Estructura no

Metadatos de publicación en triples semánticos

Aldizkari Zientifikoek Ataria

B

Anuario del Seminario de Filología Vasca "Julio de Urquijo"

ERREGISTRATU HASI SAIOA UNEKOA ONLINE FIRST ARTXIBOAK GUR

ALDIZK

Hizkunt

> English

> Español

> Euskara

Make

Informa

> Irakur

```
<link rel="schema.DC" href="http://purl.org/dc/elements/1.1/" />
<meta name="DC.Creator.PersonalName" content="Koldo Mitxelena"/>
<meta name="DC.Date.created" scheme="ISO8601" content="1992-01-15"/>
<meta name="DC.Date.dateSubmitted" scheme="ISO8601" content="2013-06-03"/>
<meta name="DC.Date.issued" scheme="ISO8601" content="1992-01-15"/>
<meta name="DC.Date.modified" scheme="ISO8601" content="2013-06-03"/>
<meta name="DC.Description" xml:lang="eu" content="-"/>
<meta name="DC.Format" scheme="IMT" content="application/pdf"/>
<meta name="DC.Identifier" content="9357"/>
<meta name="DC.Identifier.pageNumber" content="495-501"/>
<meta name="DC.Identifier.DOI" content="10.1387/asju.9357"/>
<meta name="DC.Identifier.URI" content="https://ojs.ehu.eus/index.php/ASJU/article/view/9357"/>
<meta name="DC.Language" scheme="ISO639-1" content="eu"/>
<meta name="DC.Rights" content="Copyright (c) 2015 Anuario del Seminario de Filología Vasca &quot;Julio de Urquijo&quot;"/>
<meta name="DC.Rights" content=""/>
<meta name="DC.Source" content="Anuario del Seminario de Filología Vasca &quot;Julio de Urquijo&quot;"/>
<meta name="DC.Source.ISSN" content="2444-2992"/>
<meta name="DC.Source.URI" content="https://ojs.ehu.eus/index.php/ASJU"/>
<meta name="DC.Title" content="Campionen &quot;gramática&quot; eta beste"/>
<meta name="DC.Type" content="Text.Serial.Journal"/>
<meta name="DC.Type.articleType" content="Artikuluak"/>
<meta name="gs_meta_revision" content="1.1"/>
<meta name="citation_journal_title" content="Anuario del Seminario de Filología Vasca &quot;Julio de Urquijo&quot;"/>
<meta name="citation_journal_abbrev" content="ASJU"/>
<meta name="citation_issn" content="2444-2992"/>
<meta name="citation_author" content="Koldo Mitxelena"/>
<meta name="citation_title" content="Campionen &quot;gramática&quot; eta beste"/>
<meta name="citation_language" content="eu"/>
<meta name="citation_date" content="1992/01/15"/>
<meta name="citation_firstpage" content="495"/>
<meta name="citation_lastpage" content="501"/>
<meta name="citation_doi" content="10.1387/asju.9357"/>
```

Metadatos con semántica completa: entidades enlazadas

- ▶ Triples semánticos:
sujeto – predicado – objeto
- ▶ Los tres son entidades

```
<link rel="schema.DC" href="http://purl.org/dc/elements/1.1/" />
<meta name="DC.Creator.PersonalName" content="Koldo Mitxelena" />
<meta name="DC.Date.created" scheme="ISO8601" content="1992-01-15"/>
<meta name="DC.Date.dateSubmitted" scheme="ISO8601" content="2013-06-03"/>
<meta name="DC.Date.issued" scheme="ISO8601" content="1992-01-15"/>
<meta name="DC.Date.modified" scheme="ISO8601" content="2013-06-03"/>
<meta name="DC.Description" xml:lang="eu" content="-"/>
<meta name="DC.Format" scheme="IMT" content="application/pdf"/>
<meta name="DC.Identifier" content="9357"/>
<meta name="DC.Identifier.pageNumber" content="495-501"/>
<meta name="DC.Identifier.DOI" content="10.1387/asju.9357"/>
<meta name="DC.Identifier.URI" content="https://ojs.ehu.eus/index.php/ASJU/article/view/9357"/>
<meta name="DC.Language" scheme="ISO639-1" content="eu"/>
<meta name="DC.Rights" content="Copyright (c) 2015 Anuario del Seminario de Filología Vasca &quot;Julio de Urquijo&quot;"/>
<meta name="DC.Rights" content=""/>
<meta name="DC.Source" content="Anuario del Seminario de Filología Vasca &quot;Julio de Urquijo&quot;"/>
<meta name="DC.Source.ISSN" content="2444-2992"/>
<meta name="DC.Source.URI" content="https://ojs.ehu.eus/index.php/ASJU"/>
<meta name="DC.Title" content="Campionen &quot;gramática&quot; eta beste"/>
<meta name="DC.Type" content="Text.Serial.Journal"/>
<meta name="DC.Type.articleType" content="Artikuluak"/>
<meta name="gs_meta_revision" content="1.1"/>
<meta name="citation_journal_title" content="Anuario del Seminario de Filología Vasca &quot;Julio de Urquijo&quot;"/>
<meta name="citation_journal_abbrev" content="ASJU"/>
<meta name="citation_issn" content="2444-2992"/>
<meta name="citation_author" content="Koldo Mitxelena"/>
<meta name="citation_title" content="Campionen &quot;gramática&quot; eta beste"/>
<meta name="citation_language" content="eu"/>
<meta name="citation_date" content="1992/01/15"/>
<meta name="citation_firstpage" content="495"/>
<meta name="citation_lastpage" content="501"/>
<meta name="citation_doi" content="10.1387/asju.9357"/>
```



Item Discussion Read View history ☆ More ▾

Koldo Mitxelena (Q2457804)

Basque writer

Koldo Mitxelena Elissac Luis Michelena

▼ In more languages

Language	Label	Description
English	Koldo Mitxelena	Basque writer

"Koldo Mitxelena" en Wikidata

"Koldo Mitxelena" como grafo

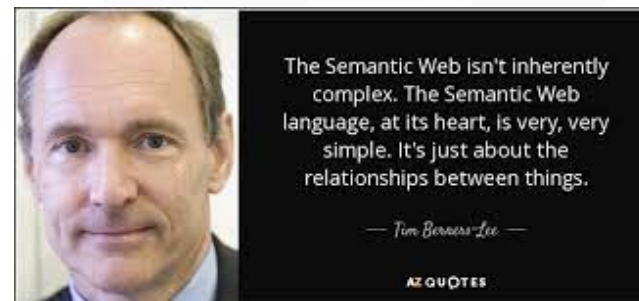
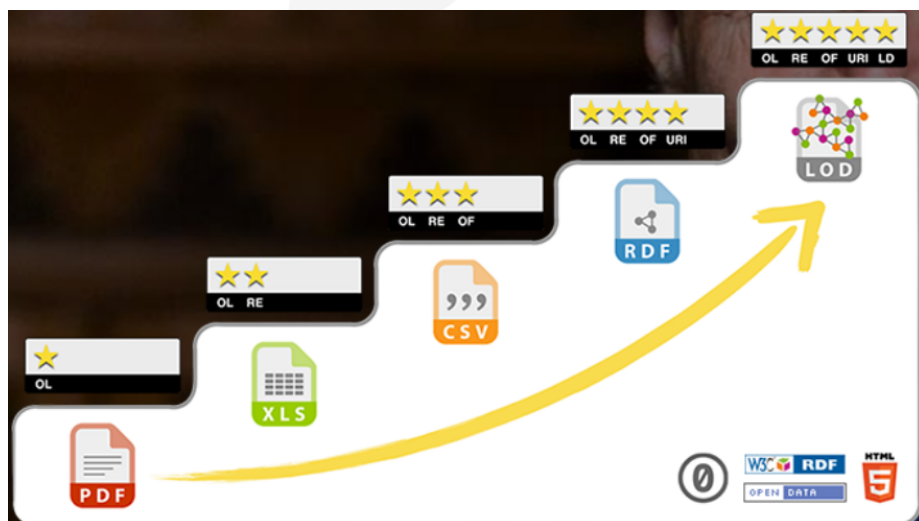
Recapitulando y encaminándonos hacia LOD

▼ Colecciones de datos según Berners-Lee

- ▼ PDF ★
- ▼ XLS ★★
- ▼ CSV ★★★
- ▼ RDF ★★★★★
- ▼ LOD ★★★★★★

▼ LOD ★★★★★: "Semantic Web"

- ▼ Término propuesto en 1999
- ▼ También llamado "Web 3.0"
- ▼ LOD: Un modelo para contenidos web
- ▼ Interpretable por máquinas
- ▼ Enlaza diferentes conjuntos de datos entre si
- ▼ Entidades representadas por URI:
Sujeto – Predicado – Objeto
- ▼ URI de clase y propiedad descritos en vocabularios estándar (W3C)



URL y URI

Web 1.0

https://es.wikipedia.org/wiki/Tim_Berners-Lee

URL

<https://es.wikipedia.org/wiki/Londres>

URL

Web 3.0

<http://www.wikidata.org/entity/Q80>

URI

<http://www.wikidata.org/entity/Q84>

URI

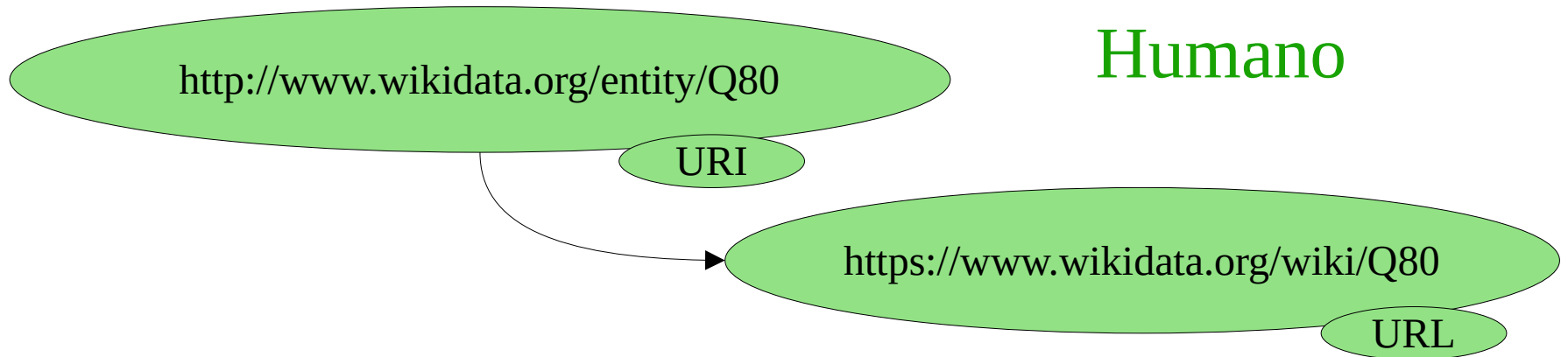
"Tim Berners-Lee"@en

"London"@en

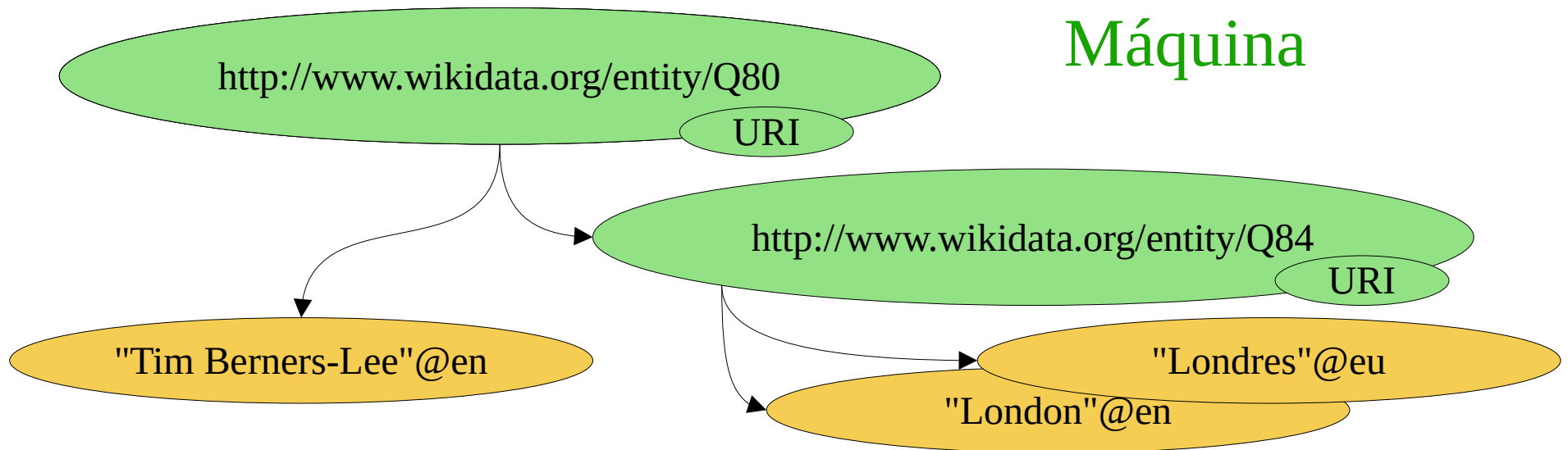
"Londres"@es

URI, content negotiation

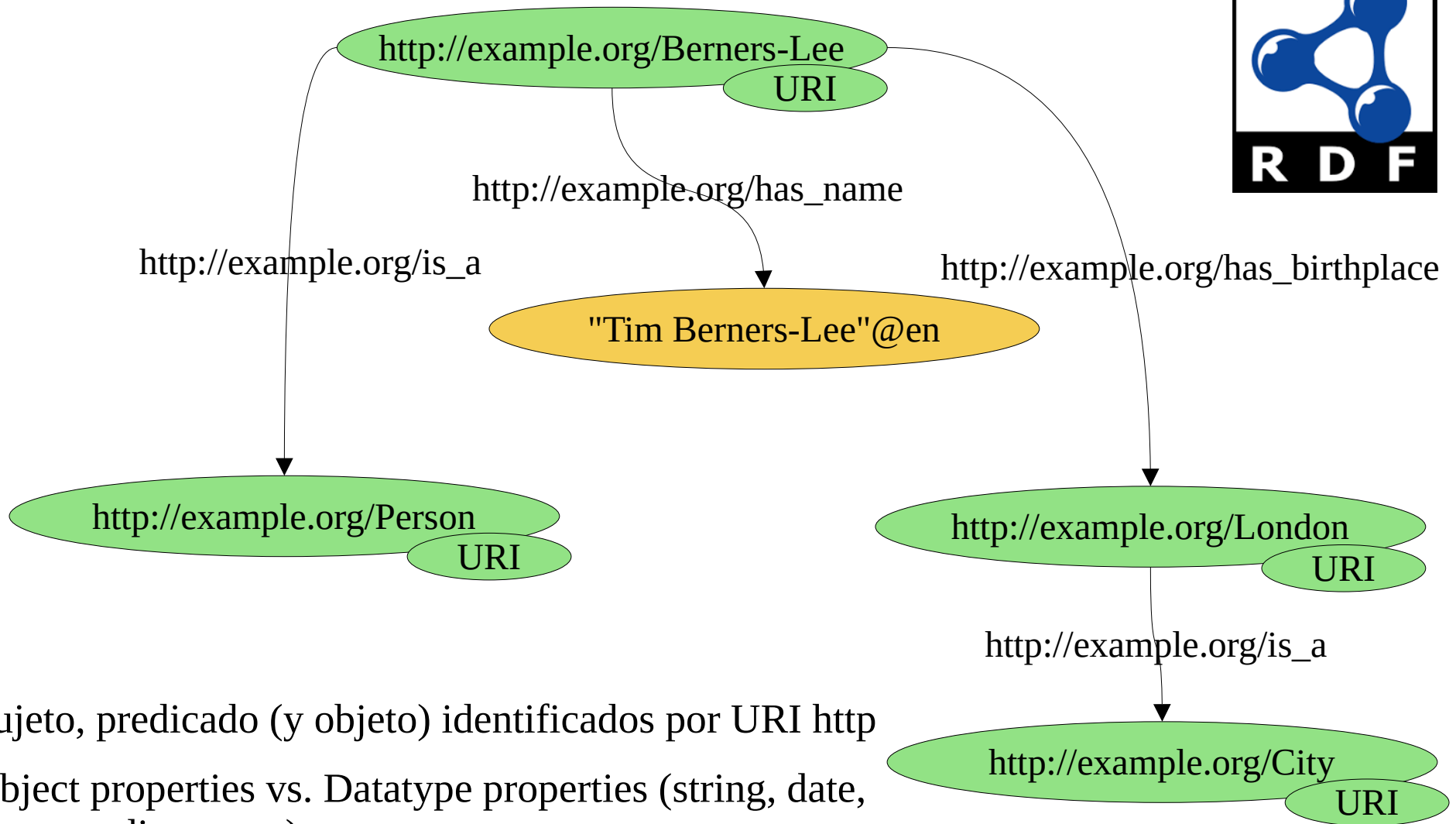
Humano



Máquina



Semántica para la máquina: Relaciones tipificadas

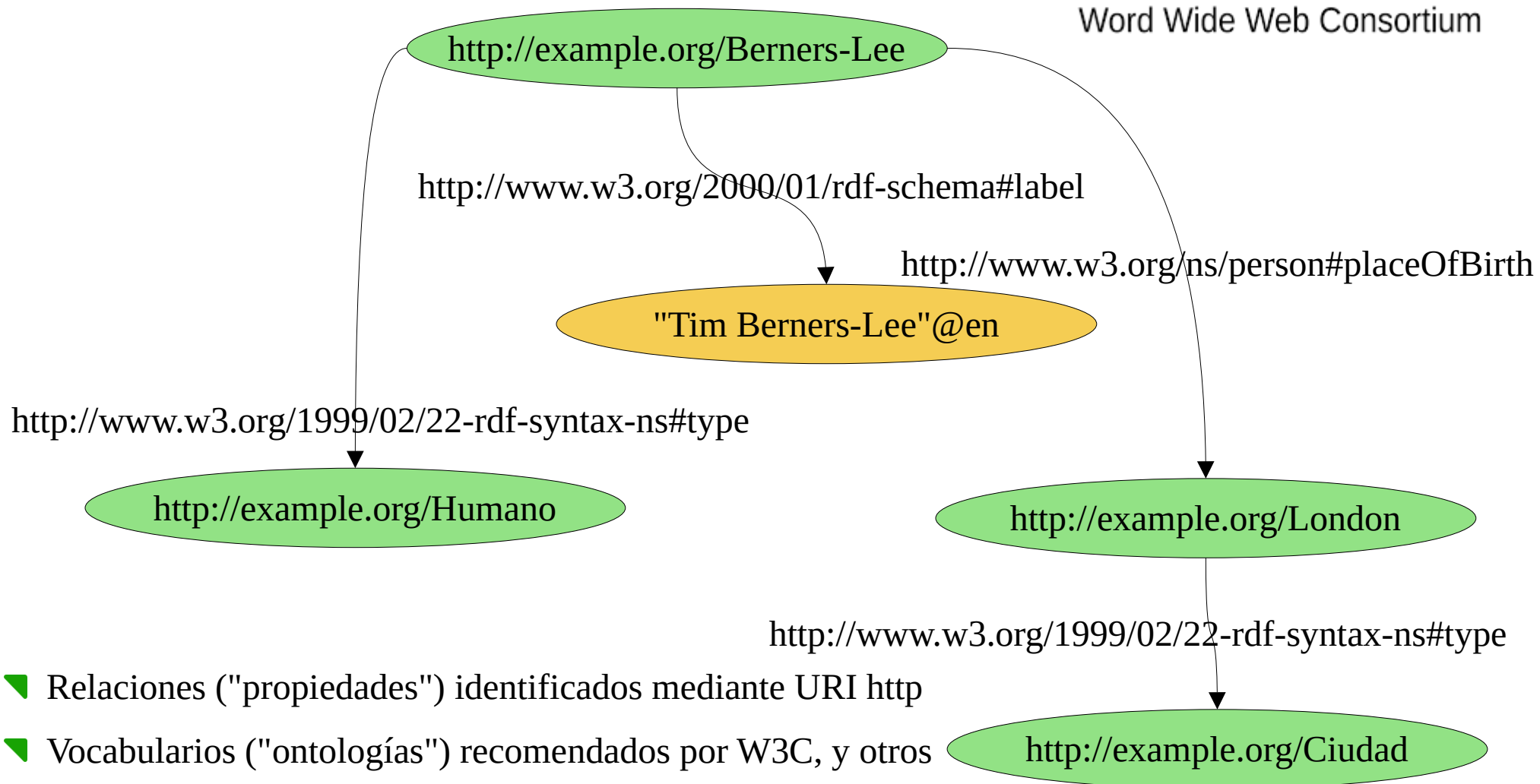


- ▼ Sujeto, predicado (y objeto) identificados por URI http
- ▼ Object properties vs. Datatype properties (string, date, geo-coordinates, ...)

Relaciones tipificadas: vocabularios estándar



Word Wide Web Consortium



- ▼ Relaciones ("propiedades") identificados mediante URI http
- ▼ Vocabularios ("ontologías") recomendados por W3C, y otros
- ▼ *Love Linked Data*: buscador de términos (bilatu "label")
- ▼ Si está bien hecho, un URI lleva al usuario humano a una descripción textual: [person:placeOfBirth](http://www.w3.org/ns/person#placeOfBirth)

Espacios de nombres (namespaces)

Namespaces used

prefix	namespace
dcterms	http://purl.org/dc/terms/
foaf	http://xmlns.com/foaf/0.1/
rdfs	http://www.w3.org/2000/01/rdf-schema#
schema	http://schema.org/
person	http://www.w3.org/ns/person#

Namespaces used in this schema

▼ Ejemplos:

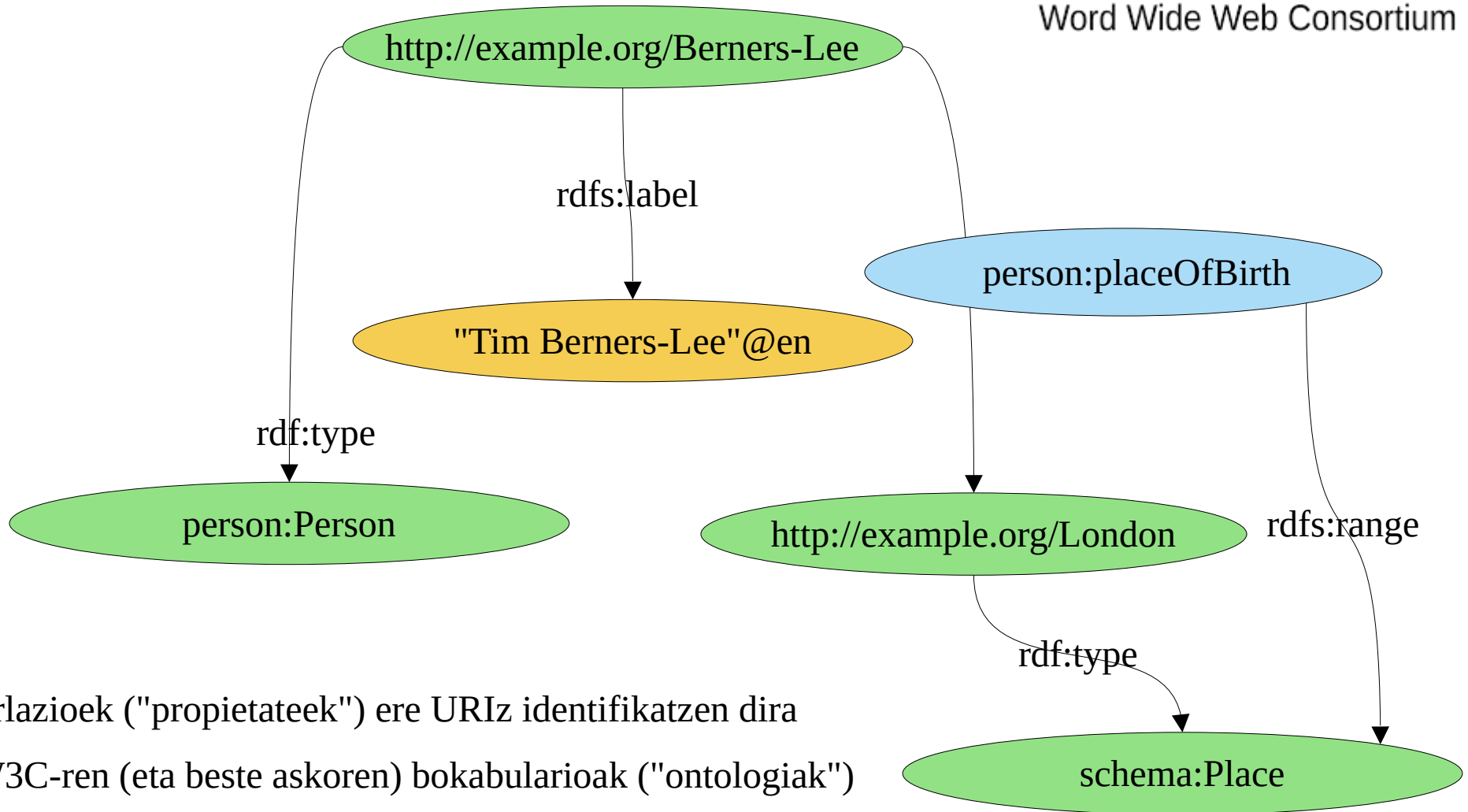
- ▼ `dcterms:title`
- ▼ `foaf:knows`
- ▼ `rdfs:label`
- ▼ `schema:birthPlace`
- ▼ `person:hasBirthPlace`

Summary of terms

The Person Core Vocabulary defines one the following terms.

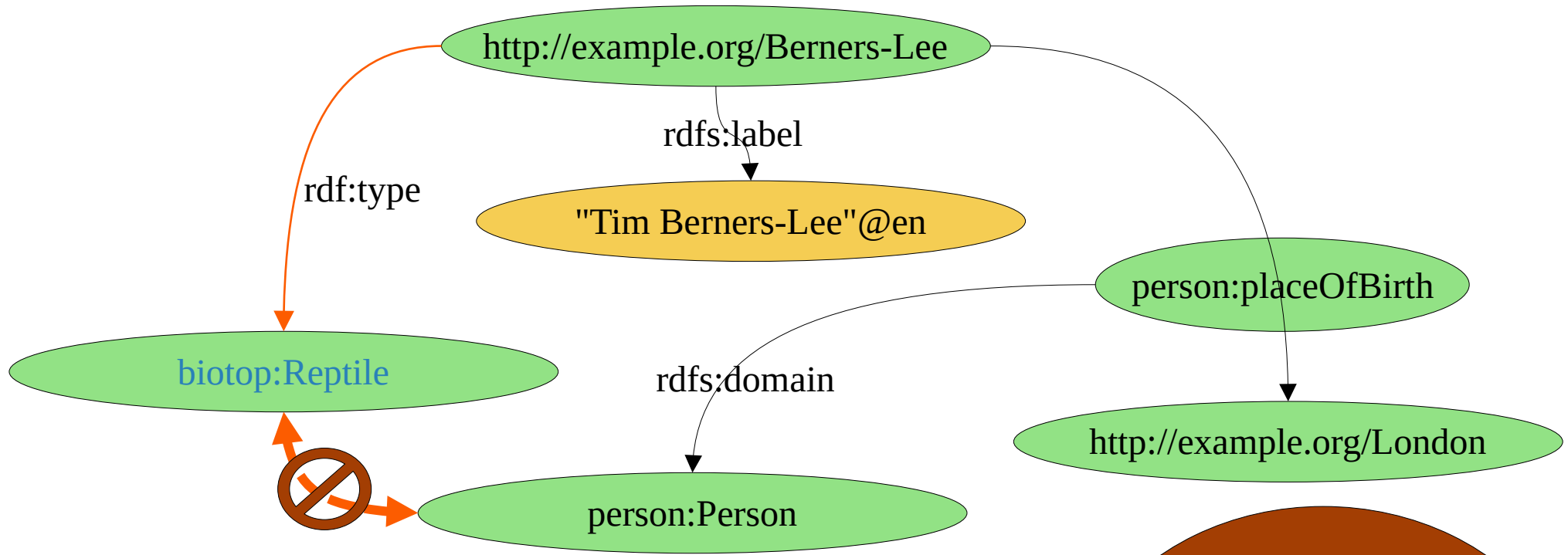
Object property	<u>person:placeOfBirth</u>
label	place of birth
comment	A person's place of birth.
domain	<u>person:Person</u>
range	<u>dcterms:Location</u> , <u>schema:Place</u>

Propiedades y Clases



- ▼ Erlazioek ("propietateek") ere URIz identifikatzen dira
- ▼ W3C-ren (eta beste askoren) bokabularioak ("ontologiak")
- ▼ Love Linked Data: Termino-bilatzailea (bilatu "label")
- ▼ URI batzuek content negotiation bitartez, guk irakurtzeko URL batera daramate: [placeOfBirth](#)

RDF eta OWL: Ontologia (mundu ikuskera) eraikitzeko

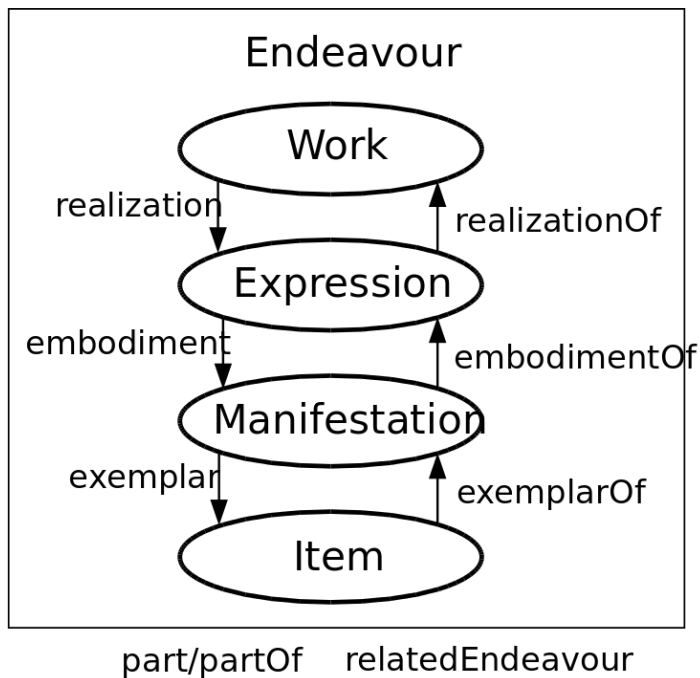


- ▼ `person`: "The Person Core Vocabulary"
- ▼ `biotop`: "Upper-Level ontology for Biology and Medicine"
- ▼ **Berners-Lee can't be a reptile**



Vocabularios RDF para (casi) todo; ejemplo bibliografía

- ▼ Dublin Core DCMI Metadata Terms
- ▼ BIBO (bibliographic ontology)
- ▼ FRBR (orain: LRM, Library Reference Model)



The screenshot shows the Dublin Core website with the following metadata for the DCMI Metadata Terms:

Title:	DCMI Metadata Terms
Creator:	DCMI Usage Board
Identifier:	http://dublincore.org/specifications/dublin-core/dcmi-terms/2020-01-20/
Date Issued:	2020-01-20
Latest Version:	https://www.dublincore.org/specifications/dublin-core/dcmi-terms/

<http://purl.org/ontology/bibo/>

The Bibliographic Ontology describes bibliographic things on the semantic Web in RDF. This ontology can be used as a citation ontology, as a document classification ontology, or simply as a way to describe any kind of document in RDF. It has been inspired by many existing document description metadata formats, and can be used as a common ground for converting other bibliographic data sources.

Turtle: Una serialización de RDF (también para SPARQL)

▼ <https://felixlohmeier.github.io/turtle-web-editor/>

Turtle Web Editor

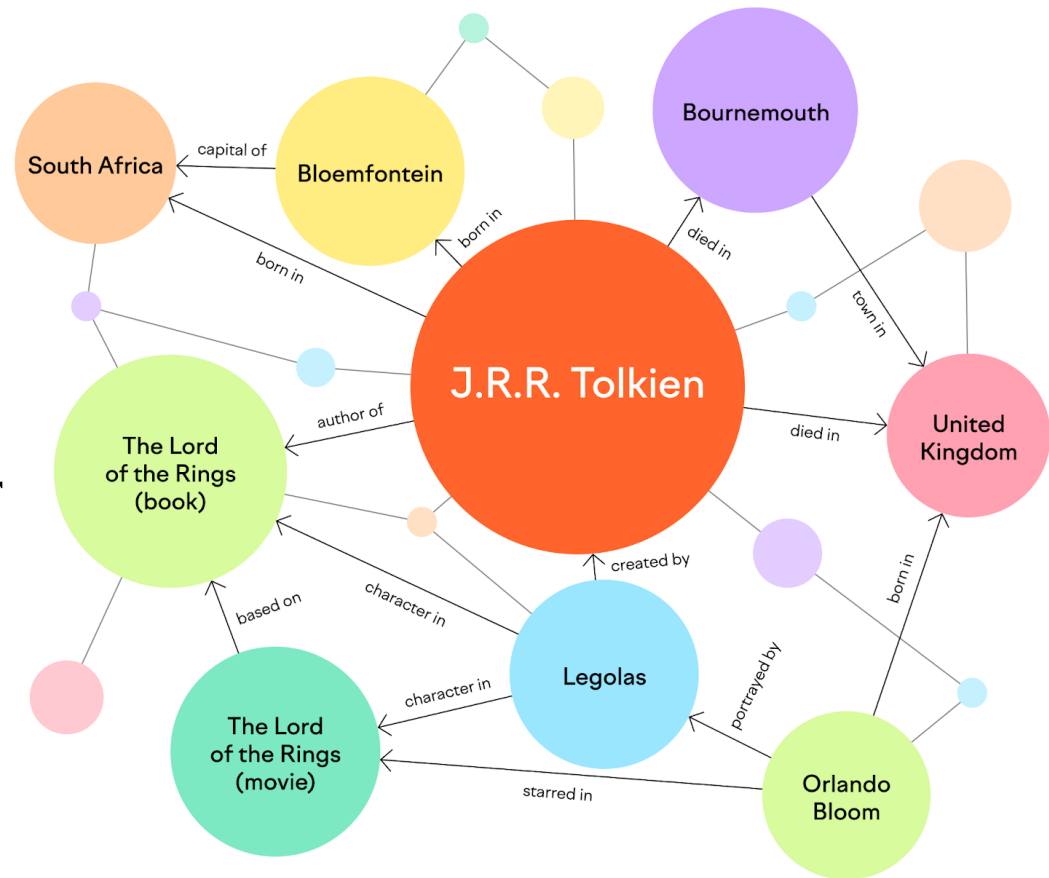
Online text editor with Turtle syntax highlighting and validation. This is a remix of IDLab Turtle Validator using Codemirror.

```
1 @prefix gurea: <http://gurea.eus/>.
2 @prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>.
3 @prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>.
4 @prefix person: <http://www.w3.org/ns/person#>.
5 @prefix schema: <https://schema.org/>.
6
7 gurea:Berners-Lee rdf:type person:Person.
8 gurea:Berners-Lee rdfs:label "Tim Berners-Lee"@en.
9 gurea:Berners-Lee person:birthPlace gurea:Londres.
10
11 gurea:Londres rdf:type gurea:City;
12               rdf:type schema:Place;
13               rdfs:label "Londres"@eu, "London"@en.
14
15
```

▼ Tres serializaciones: Turtle, JSON-LD, RDF/XML

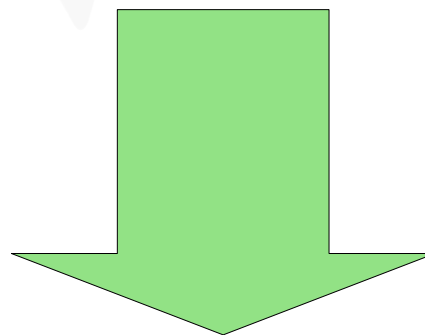
Los triples semánticos forman un grafo

- ▼ Los triples se almacenan en una base de datos de tipo grafo („triple store“)
 - ▼ Ontotext GraphDB
 - ▼ Fuseki
 - ▼ Python RDFlib
 - ▼ VocBench
 - ▼ Protégé
- ▼ SPARQL: un lenguaje para obtener resultados en tablas
 - ▼ Por interfaz gráfico
 - ▼ programáticamente



Turtle y SPARQL

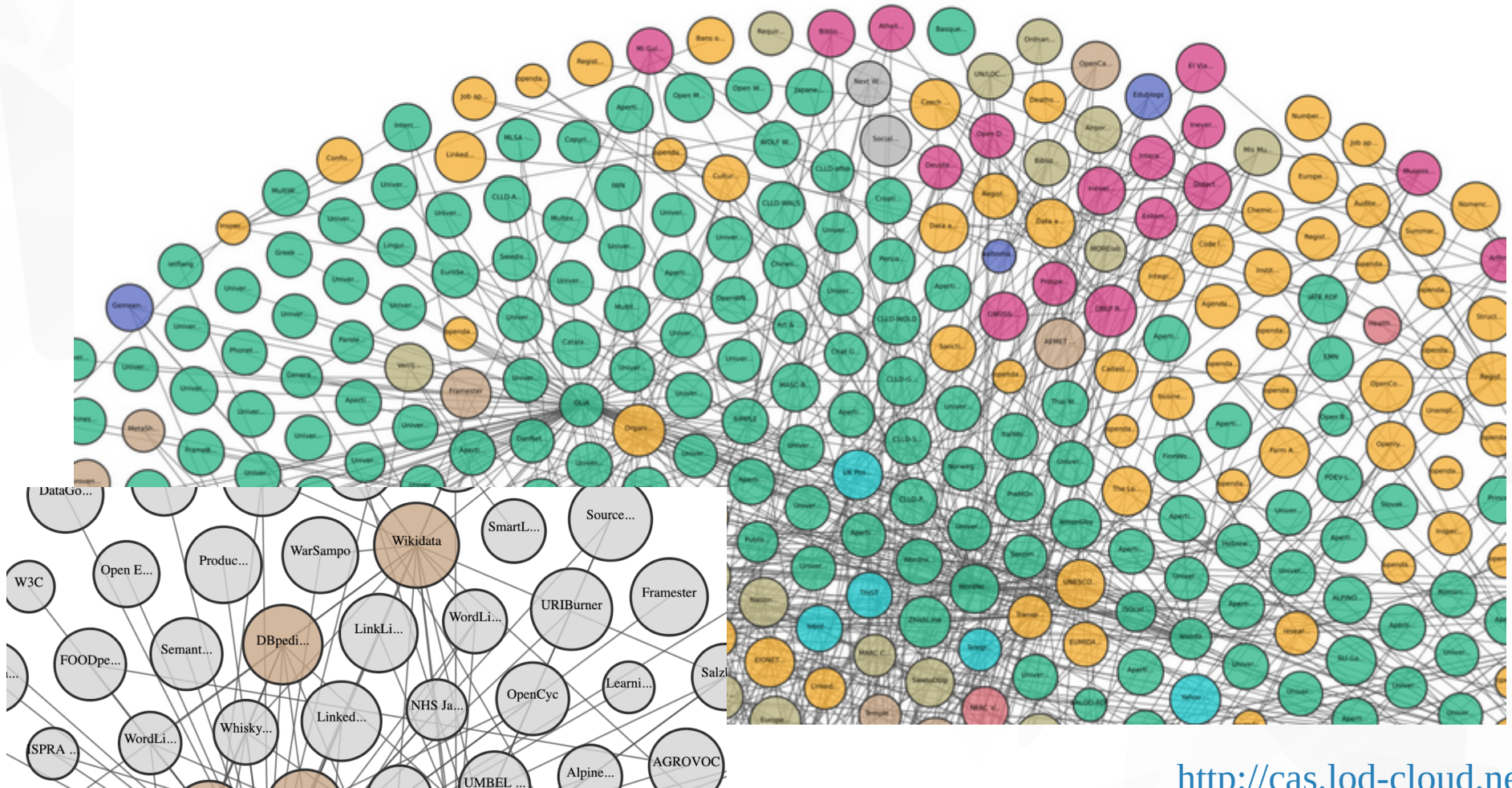
```
example:BernersLee rdf:type person:Person.  
example:BernersLee rdfs:label "Tim Berners-Lee"@en.  
example:BernersLee person:birthPlace example:London.
```



```
select ?quien  
where {  
  ?quien rdf:type person:Person.  
  ?quien person:birthPlace example:London.  
}
```

Linked Open Data

The **Linked Open Data Cloud**



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

Wikidata in the LOD cloud

The Linked Open Data Cloud


Browse Submit a dataset Diagram Subclouds About

Wikidata (Edit)

★★★★★

About this dataset

free knowledge database project hosted by Wikimedia and edited by volunteers



License: <http://www.opendefinition.org/licenses/cc-zero>

[user_generated](#) [wikimedia](#) [wikipedia](#)

Contact Details

Contact Point: [Lucas Werkmeister](#)
Website: <https://www.wikidata.org/>

Download Links

Full Downloads

- [Latest full dump, bzip2 compressed](#) (The latest full dump of all Wikidata entity data, in Turtle format, compressed with bzip2. Dumps are generated on a weekly basis.)
- [Latest full dump, gzip compressed](#) (The latest full dump of all Wikidata entity data, in Turtle format, compressed with gzip. Dumps are generated on a weekly basis.)
- [Latest truthy dump, bzip2 compressed](#) (The latest "truthy" dump of all Wikidata entity data, in N-Triples format, compressed with bzip2. This contains only "truthy" or "best" statements, without qualifiers or references. Dumps are generated on a weekly basis.)
- [Latest truthy dump, gzip compressed](#) (The latest "truthy" dump of all Wikidata entity data, in N-Triples format, compressed with gzip. This contains only "truthy" or "best" statements, without qualifiers or references. Dumps are generated on a weekly basis.)

SPARQL Endpoints

- [Wikidata Query Service](#) (The Wikidata Query Service generally contains the full data of Wikidata, modulo a slight delay in updating (usually no more than a few seconds), a handful of spurious synchronization errors (requests are routed to a randomly selected server from the pool, and the different servers may be slightly out of sync), and a few differences between the RDF dump format and the WDQS version (see https://www.mediawiki.org/wiki/Wikibase/Indexing/RDF_Dump_Format#WDQS_data_differences.)

<http://cas.lod-cloud.net/dataset/wikidata>

artículo en inglés muy bueno



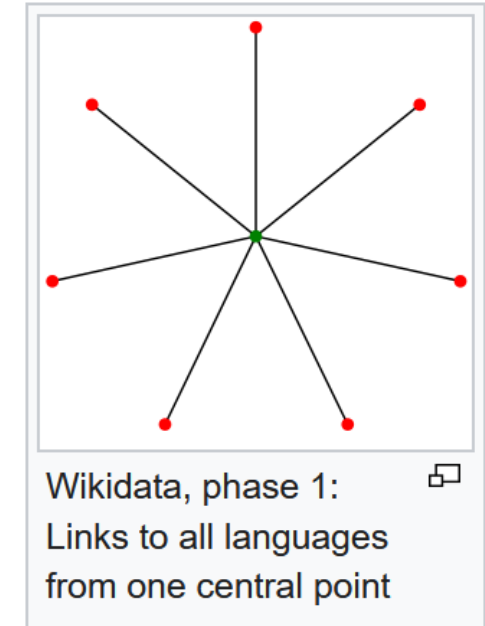
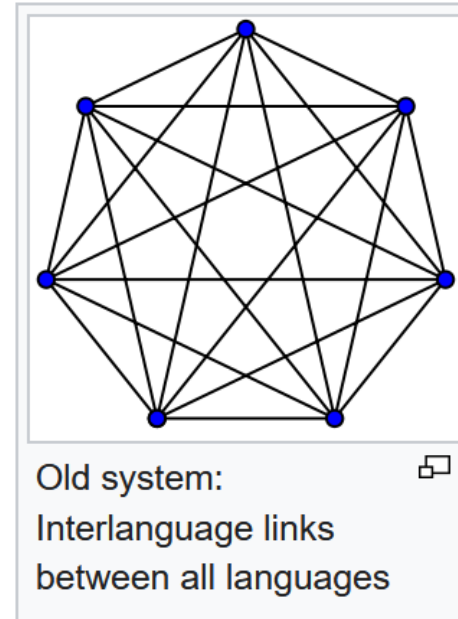
Wikidata: Five-Star-Wikipedia

“Until Wikidata can give me a list of all movies *shot in the 60s in Spain, in which a black female horse is stolen by a left-handed actor playing a Portuguese orphan, directed by a colorblind German who liked sailing, and written by a dog-owning [woman] from Helsinki*, we have more work to do.”

—User:Tobias1984, Wikidata project chat, December 2013

Qué es Wikidata? Un poco de historia

- ▼ 2012: Q1, la primera entidad
- ▼ 2013: interlanguage links
- ▼ 2013: statements
- ▼ 2015: Wikidata Query Service
- ▼ 2018: L1, el primer lexema
- ▼ 2023: 100M+ entidades ([estatistikak](#))



<https://meta.wikimedia.org/wiki/Wikidata/Help>

Wikidata, para qué?

- Wikipedia utiliza datos de Wikidata, por ejemplo en las *infobox*

Item Discussion

Norwegian Lundehund (Q38895)

dog breed edit

In more languages

Language	Label	Description	Also known as
English	Norwegian Lundehund	dog breed	
German	Norwegischer Lundehund	Hunderasse	Lundehund Lunde-Hund Norwegischer Vogelhund
Spanish	Lundehund	raza de perro originaria de Noruega	
Catalan	No label defined	No description defined	

All entered languages

Statements

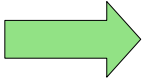
instance of dog breed edit

1 reference

+ add value

image edit

Lundehund-2003.jpg
716 x 526; 121 KB



Norwegian Lundehund

A Norwegian Lundehund

Other names Norsk Lundehund
Norwegian Puffin Dog

Nicknames Lundehund

Country of origin Norway

Traits [hide]

Weight 6–7 kilograms (13–15 lb)

Height 30–40 centimetres (12–16 in)

Classification / standards [hide]

FCI Group 5 Section 2 #265 standard edit

AKC Non-Sporting standard edit

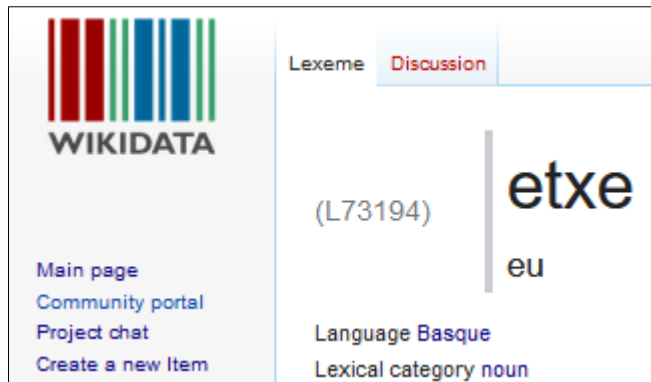
UKC Northern Breed standard edit

Dog (*Canis lupus familiaris*)

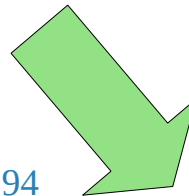
Example of an infobox edit

Wikidata, para qué?


- Un script en el Wiktionary del euskera que pregunta por acepciones de lexemas en Wikidata



<https://www.wikidata.org/wiki/Lexeme:L73194>



<https://eu.wiktionary.org/wiki/etxe>

 **Wikidata Adierak** [editatu]

Ikusi eta editatu Wikidatan

- [1](#): gizakia bizitzeko prestatzen den edozein eraikin
- [2](#): etxebizitza
- [3](#): erabilera jakin bat duten zenbait eraikinen izena
- [4](#): familia; leinua, dinastia
- [5](#): (Teknol.) hitz-elkarte baten bigarren osagai gisa, pieza edo tresna bati eusteko erabiltzen den elementu edo egitura

Wikidata, un grafo del conocimiento libre

WIKIDATA

Orrialdea **Eztabaida**

<http://www.wikidata.org/entity/Q65216433>

Larramendiren Hiztegi Hirukoitza (Q65216433)

Larramendik XVIII. mendean idatzitako hiztegia (Euskara-Gaztelania-Latina)
Larramendi hiztegia | Larramendiren hiztegia

Azala
Komunitatearen ataria

Statements

instance of	version, edition, or translation
	▼ 0 references
dictionary	
	▼ 0 referen

author	Manuel de Larramendi
	▼ 0 references
place of publication	Donostia-San Sebastián
	▼ 0 references

WIKIDATA

Lexeme **Eztabaida**

(L49165) **bilaukeria**

eu

<http://www.wikidata.org/entity/L49165>

Language euskara
Lexical category substantibo

Adierazpenak

non jasoa	Larramendiren Hiztegi Hirukoitza
honela jasoa	billauqueria
dagokion URLa	https://eu.wikisource.org/wiki/Hiztegi_Hirukoitza/V#villanila

Lexicographical data
Lexema berria sortu

L49165-F14 | bilaukeriaz
eu

Grammatical features **instrumental, singular**

L49165-S1 | euskara

bilaua denaren nolakotasuna, gaitzesgarritzat hartua; bilauari dagokion eintza makurra

RDF >> Wikidata

- ▶ Punto de partida: [Berners-Lee et al. 2001: Semantic Web / "Web 3.0"](#)
- ▶ Esquema fundamental: Sujeto - Predicado - Objeto ('Semantic Triple')
- ▶ Todas las entidades tienen su **URI**, un identificador unívoco
- ▶ Wikidata implementa los criterios Semantic Web

Sujeto	http://www.wikidata.org/entity/L49165 'bilaukeria' (lexema)
Predicado	http://www.wikidata.org/entity/P5323 'non jasoa' (propiedad)
Objeto	http://www.wikidata.org/entity/Q65216433 'Hiztegi Hirukoitza' (un item de la clase 'diccionario')

Datos Enlazados en Wikidata: Entidades del tipo L

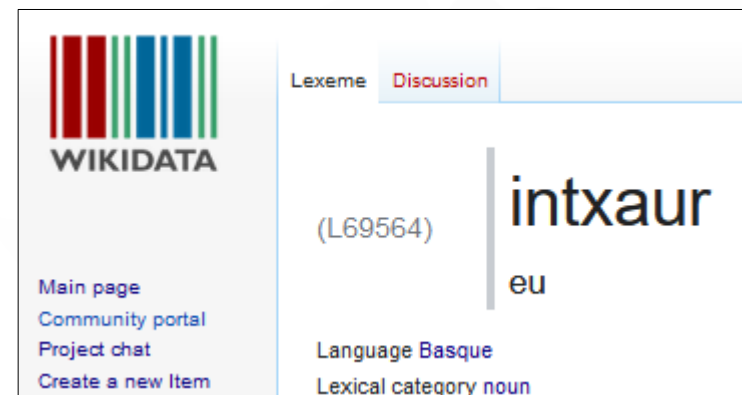
- ▼ Entidades L: lexemas (entradas de diccionario, con lengua, lema, cat. gramatical)
- ▼ Estructura fundamental: lexema - acepción - forma (vocabulario Ontolex RDF W3C)
- ▼ Lista de lexemas en Wikidata
- ▼ Documentación modelo de datos: lexemas

<https://www.wikidata.org/wiki/Lexeme:L184867>



The screenshot shows the Wikidata page for Lexeme L184867. The page has a header with the Wikidata logo and navigation tabs for 'Lexeme' and 'Discussion'. The main content area displays the lexeme '(L184867) entzun' in Basque, with the language 'eu' and the lexical category 'verb'.

<https://www.wikidata.org/wiki/Lexeme:L69564>



The screenshot shows the Wikidata page for Lexeme L69564. The page has a header with the Wikidata logo and navigation tabs for 'Lexeme' and 'Discussion'. The main content area displays the lexeme '(L69564) intxaur' in Basque, with the language 'eu' and the lexical category 'noun'.

Datos Enlazados en Wikidata: entidades P

- ▼ Entidades P: propiedades (relaciones)
 - ▼ *Datatype*:
 - ▼ Cada propiedad apunta a un solo tipo de datos
 - ▼ Ejemplos
 - ▼ P31 "Instance of [this class]": Item
 - ▼ P50 "Author": Item
 - ▼ P2799 "BVMC person ID": external ID
 - ▼ P569 "Date of birth": point in time
 - ▼ P1476 "Title": monolingual text

Built-in data types

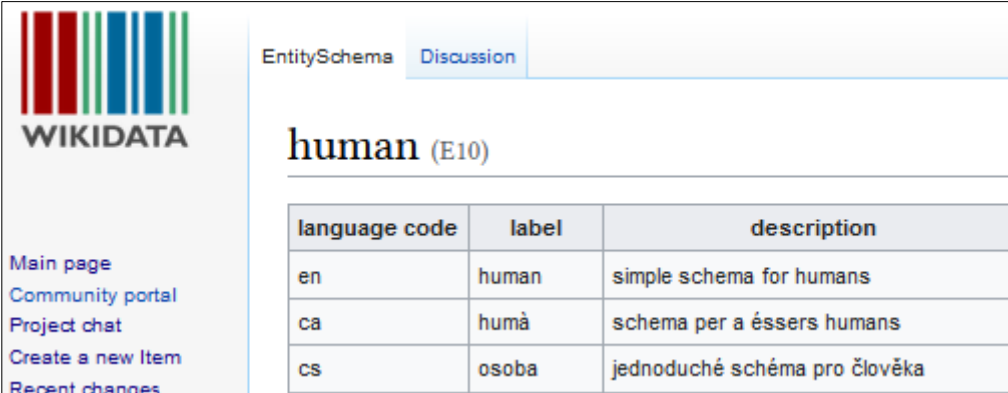
Data type	Number of properties
External identifier	8,091
Item	1,598
Quantity	644
String	320
URL	97
Commons media file	78
Point in time	63
Monolingual text	60
Property	21
Geographic coordinates	11
Tabular data	6
Geographic shape	3

Extra data types

Data type	Number of properties
Mathematical expression	36
Lexeme	16
Sense	16
Form	7
Musical Notation	6

Wikidata: otros tipos de entidades

- ▼ E: Entity Schemes
 - ▼ ejemplo: [esquema del sustantivo en euskera](#), [esquema de persona](#)
- ▼ M: Media (en Wikimedia commons)
- ▼ S: Lexeme Sense (acepción de lexema) > parte de la descripción de un lexema
- ▼ F: Lexeme Form (forma de lexema) > parte de la descripción de un lexema



EntitySchema Discussion

human (E10)

language code	label	description
en	human	simple schema for humans
ca	humà	schema per a éssers humans
cs	osoba	jednoduché schéma pro člověka

Labels, Description



- Main page
- Community portal
- Project chat
- Create a new Item
- Recent changes
- Random Item
- Query Service
- Nearby
- Help
- Donate
- Lexicographical data
- Create a new Lexeme
- Recent changes
- Random Lexeme
- Tools
- What links here
- Related changes
- Special pages
- Permanent link
- Page information
- Concept URI
- Cite this page

Item [Discussion](#)

Bob Marley (Q409)

Jamaican reggae musician (1945–1981)

[Donald Marley](#) | [Nesta Robert Marley](#) | [Tuff Gong](#) | [Bab Maali](#) | [Rabat Nesta Maali](#) | [Robert Nesta "Bob" Marley](#) | [Robert Nesta Marley](#)

edit

▼ In more languages

Language	Label	Description	Also known as
English	Bob Marley	Jamaican reggae musician (1945–1981)	Donald Marley Nesta Robert Marley Tuff Gong Bab Maali Rabat Nesta Maali Robert Nesta "Bob" Marley Robert Nesta Marley
German	Bob Marley	jamaikanischer Sänger, Gitarrist und Songwriter (1945-1981)	Robert Nesta Marley Tuff Gong Donald Marley Nesta Robert Marley
Spanish	Bob Marley	músico, guitarrista y compositor jamaicano	Donald Marley Nesta Robert Marley Tuff Gong Robert Nesta Marley
Basque	Bob Marley	No description defined	
Latin	Robertus Marley	No description defined	Bob Marley

consulta

- [wd:Q409](#) `rdfs:label` “Bob Marley”@en , “Боб Марли”@ru .
- [wd:Q409](#) `skos:altLabel` “Robert Nesta Marley”@en, “Nesta Robert Marley”@en .
- [wd:Q409](#) `schema:description` “Jamaican reggae musician (1945-1981)”@en .

Statements

<https://www.wikidata.org/wiki/Q409>

The screenshot shows the Wikidata page for Bob Marley (Q409). The main content area displays the name 'Bob Marley (Q409)' and his description as a 'Jamaican reggae musician (1945–1981)'. Below this, there are two statements:

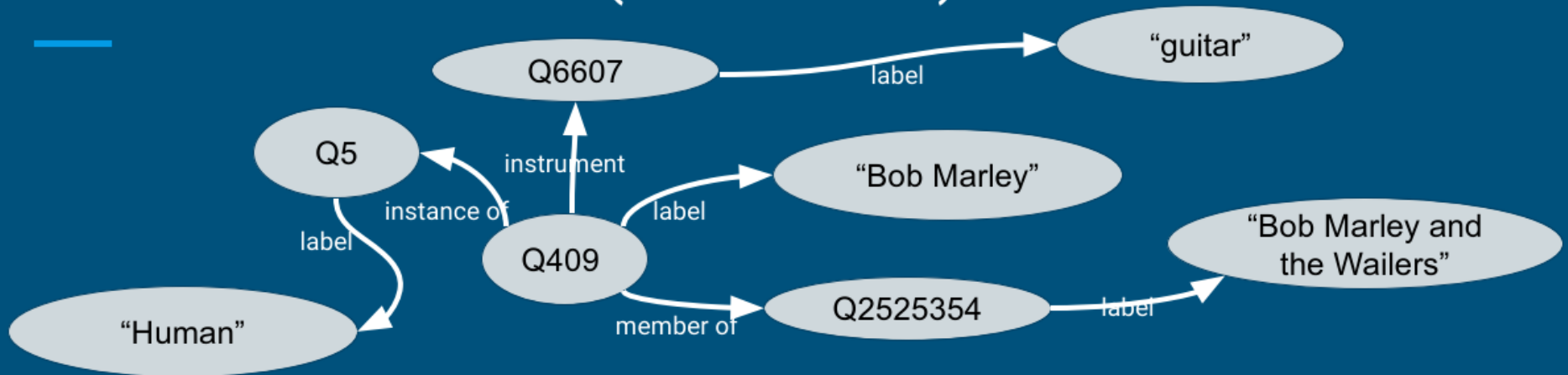
- member of** (Property:P463) with value **The Upsetters** (0 references).
- religion or worldview** (Property:P140) with value **Latin Church** (1945–1981, 0 references) and **Rastafari movement** (1966–1980, 0 references).

The 'Rastafari movement' statement is highlighted with a tooltip showing its ID: Q101462.

- [Q409](#) (Bob Marley) > [P463](#) (member of) > [Q2525354](#) (Bob Marley and the Wailers)
- [Q409](#) (Bob Marley) > [P140](#) (religion or worldview) > [Q101462](#) (Rastafari movement)

Labels y Statements, vistos como grafo

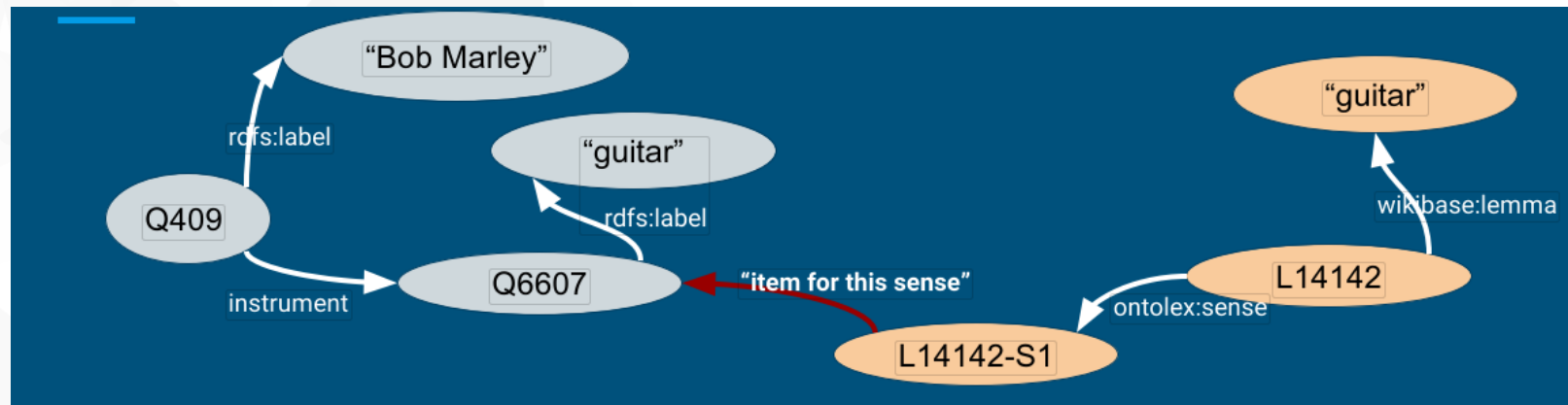
Wikibase “items” (Q-entities)



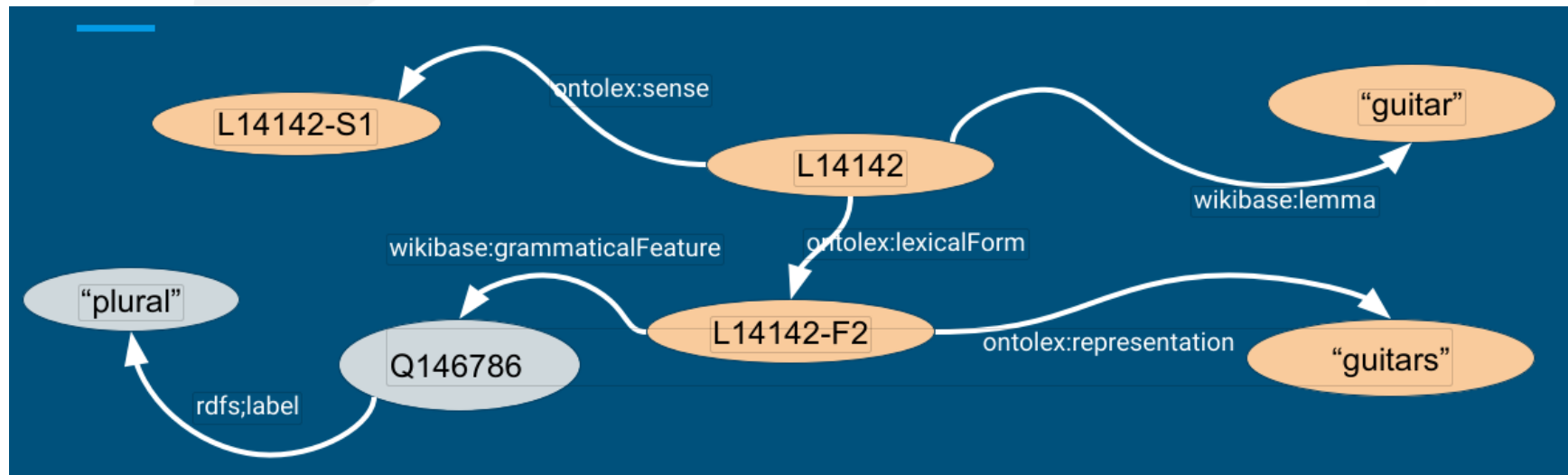
- [wd:Q409](#) rdfs:label “Bob Marley”@en , “Боб Марли”@ru .
- [wd:Q409](#) wdt:P31 (“instance of”) [wd:Q5](#) .
- [wd:Q409](#) wdt:P463 (“member of”) [wd:Q2525354](#) .

<https://w.wiki/6yJL>

Entidades Q y L, enlazados entre si



`wd:L14142-S1 wdt:P5137 ("item for this sense") wd:Q6607 .`



ORDIA text to lexemes

Search after lemma or forr **Ordia** Languages Aspects Tools Statistics

Text to lexemes

Hendrix destrozaba guitarras.

Language:

Español - es - Spanish

Submit

<https://ordia.toolforge.org/text-to-lexemes>

Statements y calificadores ("qualifiers")

Qualifiers

Help:Qualifiers

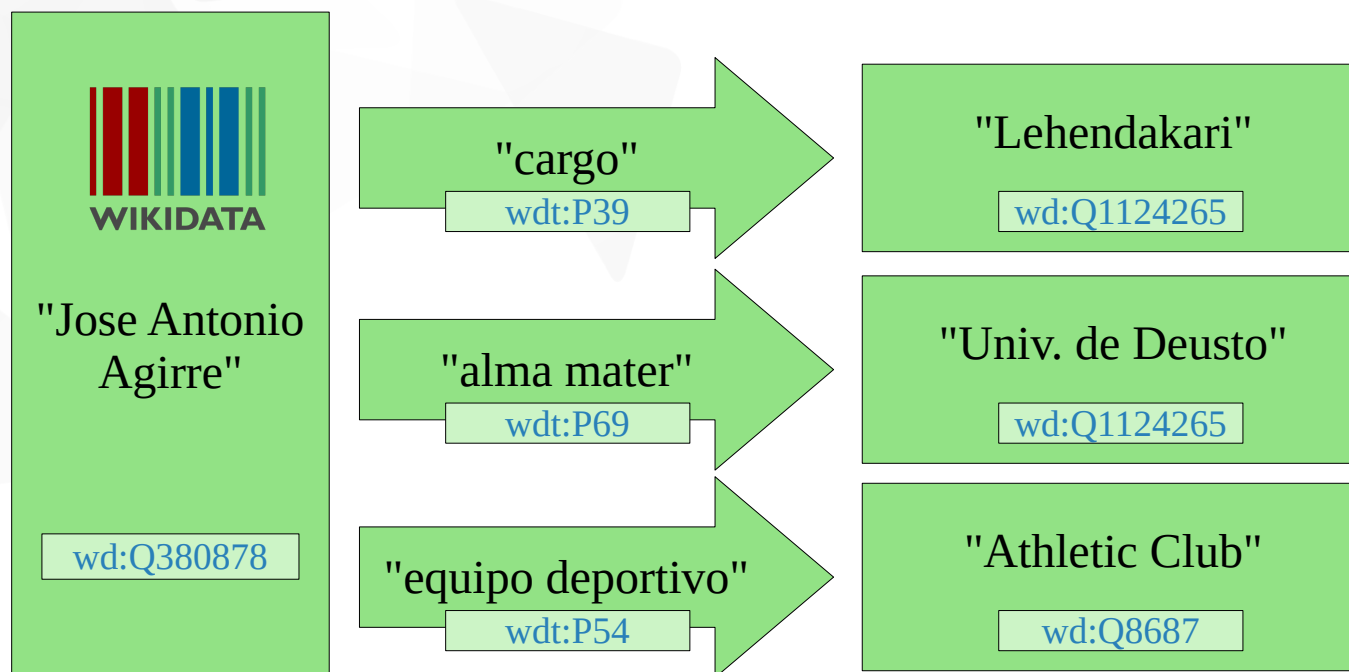
[wd:Q409](#) (Bob Marley) [wdt:P140](#) (religion or worldview) [wd:Q101462](#) (Rastafari movement)

[wd:Q409](#) p:[140](#) [ps:[P140](#) [wd:Q101462](#) ; pq:[P580](#) {time object}; pq:[P582](#) {time object}].

```
PREFIX wd: <http://www.wikidata.org/entity/>
PREFIX wdt: <http://www.wikidata.org/prop/direct/>
PREFIX p: <http://www.wikidata.org/prop/>
PREFIX ps: <http://www.wikidata.org/prop/statement/>
PREFIX pq: <http://www.wikidata.org/prop/qualifier/>
```

- La misma entidad P con diferentes usos, en diferentes namespaces (wdt: / p: / ps: / pq:)

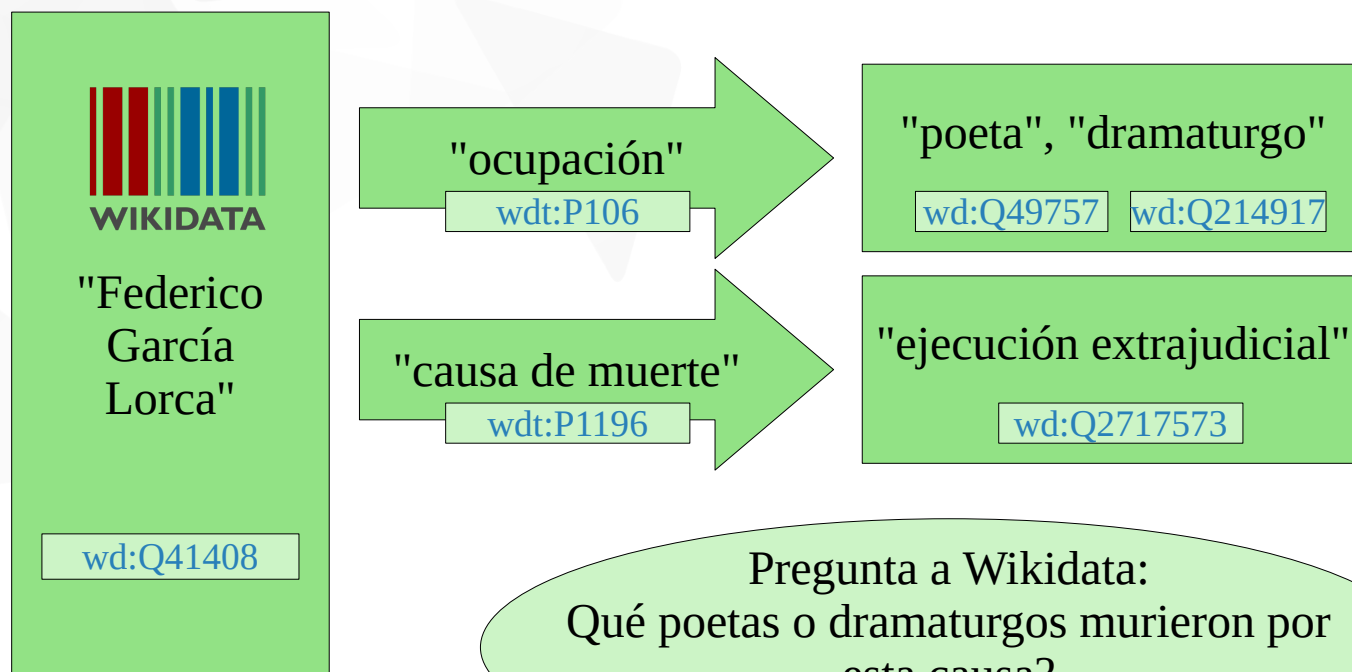
Wikidata: Five-Star-Wikipedia



Pregunta a Wikidata:
Dónde estudiaron los lehendakaris?
Demo: <https://w.wiki/M3L>

Pregunta a Wikidata:
Qué jefe de gobierno fue futbolista?
Demo: [https://w.wiki/572\\$](https://w.wiki/572$)

Wikidata: Five-Star-Wikipedia



Pregunta a Wikidata:
Qué poetas o dramaturgos murieron por esta causa?

<https://w.wiki/AE8C>

Pregunta a Wikidata:
Qué artistas murieron por muerte violenta?

<https://w.wiki/AE8B>

Consultas federadas: Wikidata y BVMC

- ▶ Obras de Lorca y obras sobre Lorca: <https://w.wiki/AASW>
- ▶ Lorca, y otras personas muertas por ejecución extrajudicial y con entrada en BVMC: <https://w.wiki/AAU6>
- ▶ Lorca en Wikidocumentaries

wd:Q41408

WIKIDATA

Item Discussion

Federico García Lorca (Q41408)

Spanish poet, dramatist and the director (1898-1936)

García Lorca | García Lorca, Federico | G. F. Lorca | Federi
| Federico García Lorca | Phenteriko Gkarthia Lorka | Lor
| Federico García Lorca | Lorca | Federico García Lorca |
| Federico García Lorca | F. García Lorca | Federico del Sa

Main page
Community portal
Project chat
Create a new Item
Recent changes

BVMC person ID 273



Inicio / Datos enlazados / García Lorca, Federico, 1898-1936

García Lorca, Federico, 1898-1936

escritor español

Información extraída de [Wikipedia](#) / [\(CC BY-SA 3.0\)](#)

Identificador: 273

Nombre: García Lorca, Federico

Fecha de nacimiento: [1898](#)

Fecha de fallecimiento: [1936](#)

VIAF: <https://viaf.org/viaf/34456883>

ISNI: <https://isni-url.oclc.nl/isni/0000000122781811>

Variantes de nombre:

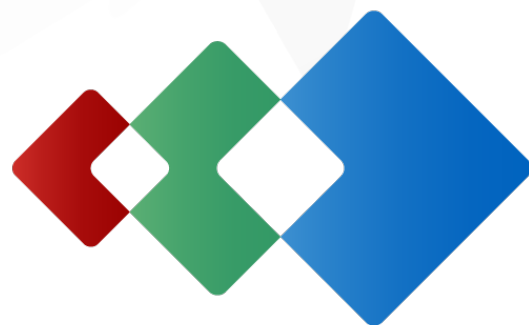
- Federico García Lorca
- Federico García Lorca

EXPORTAR:

RDF JSON

Wikimedia Commons

Wikibase



WIKIBASE

Wikibase, Wikibase Cloud



<https://wikiba.se>



<https://wikibase.cloud>

Por qué Wikibase?

▼ Características del software

- ▼ Licencia libre, código abierto
- ▼ Comunidad
- ▼ Self-hosted: Versión para Docker
- ▼ Wikibase Cloud
- ▼ 3 bases de datos:
 - ▼ mariaDB datos de entidad [en JSON](#)
 - ▼ elastic search
 - ▼ blazegraph > SPARQL
- ▼ [Modelo de datos](#) fundamental 100% compatible con Wikidata

▼ Proceso de edición de datos

- ▼ Exhibición de datos en páginas de entidad
- ▼ User y user rights management
- ▼ Historiales de entidad y de usuario reversibles

▼ Acceso programático

- ▼ mediawiki API
- ▼ SPARQL endpoint

Wikidata API, Turtle entity data, elastic search

▼ API

▼ <https://www.wikidata.org/wiki/Special:ApiSandbox?action=wbgetclaims>

▼ Adibidea: Tim Berners-Lee buruzkoak

▼ Documentación del API

▼ Entity data dump

▼ Ejemplo: Berners-Lee (Q80.ttl)

▼ Elastic Search

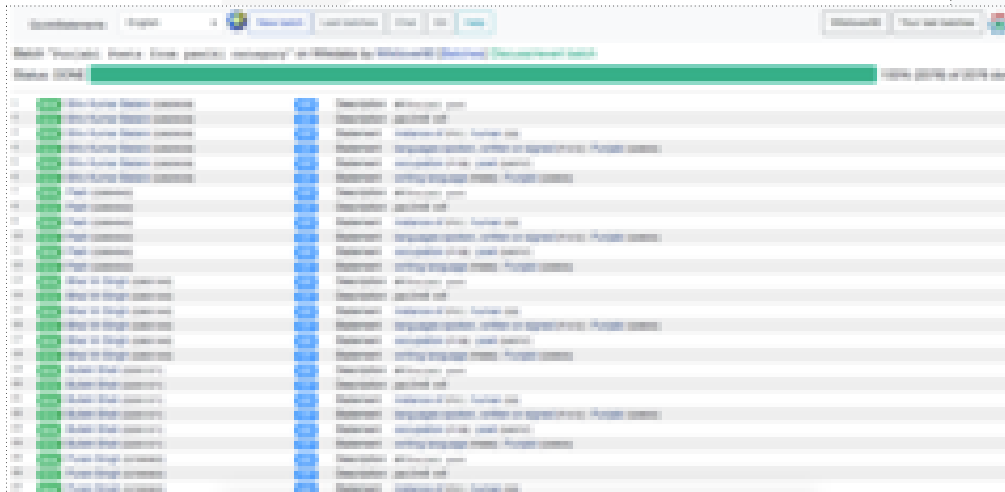
▼ ejemplo de búsqueda: 'Lorca'

```
{
  "claims": {
    "P734": [
      {
        "mainsnak": {
          "snaktype": "value",
          "property": "P734",
          "hash": "211c4757a2945ca46fc2bbce95ac1a029fb247fc",
          "datavalue": {
            "value": {
              "entity-type": "item",
              "numeric-id": 18375238,
              "id": "Q18375238"
            },
            "type": "wikibase-entityid"
          },
          "datatype": "wikibase-item"
        },
        "type": "statement",
        "id": "Q80$f415fcf7-4fec-59f7-7793-e71aa5874323",
        "rank": "normal",
        "references": [
          {
            "hash": "c0d60ea6dd8479b916e20c9aef43446852006f64",
            "snaks": {
              "P248": [
                {
                  "snaktype": "value",
                  "property": "P248",
                  "hash": "019a50b7de741e0068bde41c9d9955b22a5de47b",
                  "datavalue": {
                    "value": {
                      "entity-type": "item",
                      "numeric-id": 36578,
                      "id": "Q36578"
                    },
                    "type": "wikibase-entityid"
                  }
                }
              ]
            }
          }
        ]
      }
    ]
  }
}
```

Cómo subir datos a una Wikibase?

- ▼ API
 - ▼ Acepta representaciones JSON de las entidades a crear o modificar
 - ▼ python `mwclient`
 - ▼ python `wikibaseintegrator`
- ▼ Manualmente en la página de la entidad
- ▼ Cradle (ejemplo)
- ▼ Quickstatements: importar CSV

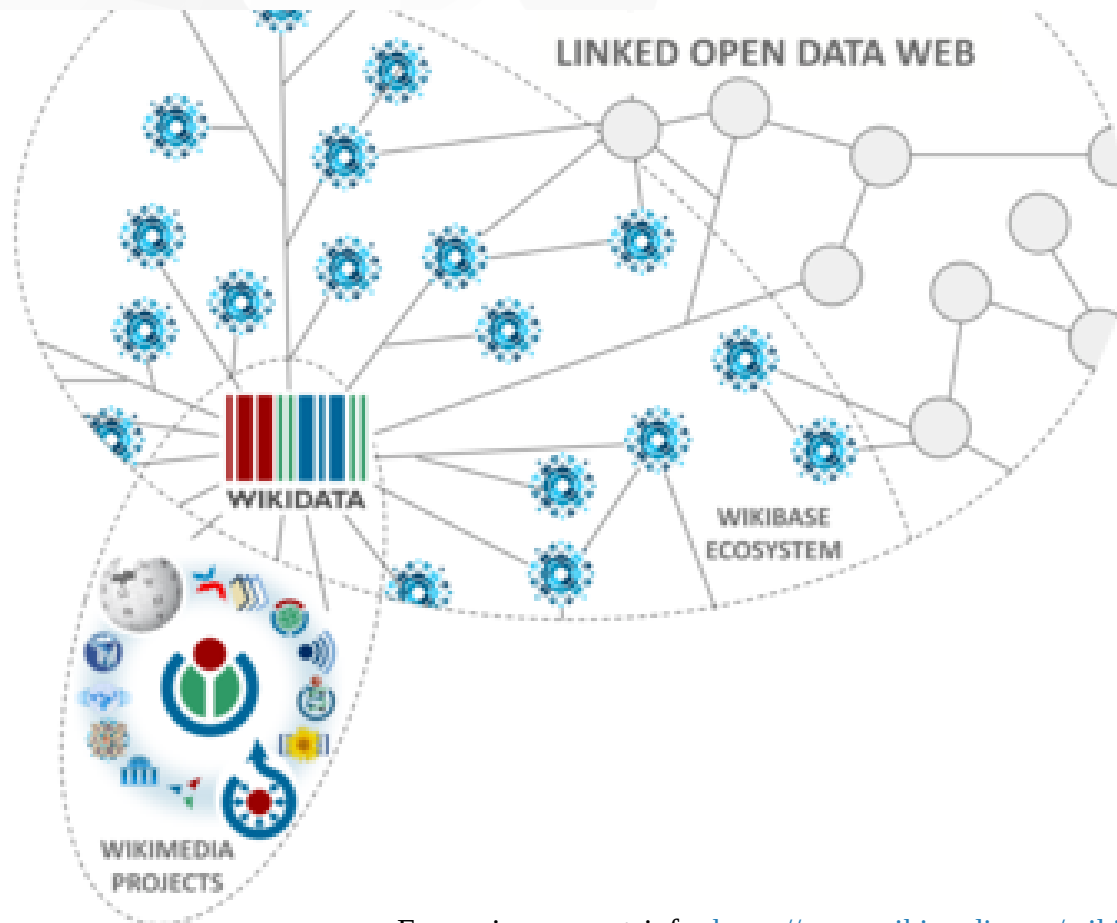
```
{
  "claims": {
    "P734": [
      {
        "mainsnak": {
          "snaktype": "value",
          "property": "P734",
          "hash": "211c4757a2945ca46fc2bbce95ac1a029fb247fc",
          "datavalue": {
            "value": {
              "entity-type": "item",
              "numeric-id": 18375238,
              "id": "Q18375238"
            }
          },
          "type": "wikibase-entityid"
        },
        "datatype": "wikibase-item"
      },
      {
        "type": "statement",
        "id": "Q805f415fcf7-4fec-59f7-7793-e71aa5874323",
        "rank": "normal",
        "references": [
          {
            "hash": "c0d60ea6dd8479b916e20c9aef43446852006f64",
            "snaks": {
              "P248": [
                {
                  "snaktype": "value",
                  "property": "P248",
                  "hash": "019a50b7de741e0068bde41c9d9955b22a5de47b",
                  "datavalue": {
                    "value": {
                      "entity-type": "item",
                      "numeric-id": 36578,
                      "id": "Q36578"
                    }
                  },
                  "type": "wikibase-entityid"
                }
              ]
            }
          }
        ]
      }
    ]
  }
}
```



The screenshot shows a web browser window displaying the Wikibase QuickStatements interface. The interface is a table with multiple columns, including 'Statement ID', 'Property', 'Value', and 'Rank'. The table contains several rows of data, with some cells highlighted in green and blue. The interface is used for manually entering or editing Wikibase statements.

Wikibase

- ▼ "Un Wikidata vacío" para cualquiera y cualquier uso
- ▼ Federación con Wikidata ("Wikibase Ecosystem")



Una consulta federada:
Wikibase-Wikidata

Fuente imagen y + info: https://meta.wikimedia.org/wiki/LinkedOpenData/Strategy2021/Joint_Vision

DL: Proyectos / experimentos con Wikibase: Bibliodata



▼ LexBib Wikibase

- ▼ Metadatos de publicación
 - ▼ Diccionarios
 - ▼ Artículos metalexigráficos
- ▼ Personas, congresos, organizaciones, lugares, lenguas...
- ▼ Objetivos:
 - ▼ Elexifinder
 - ▼ CLARIN VLO
 - ▼ Wikidata (trans. / fed.)



▼ CLB-LOD Wikibase

- ▼ Czech Literary Bibliography
- ▼ MARC-XML to Wikibase



▼ ENEOLI Wikibase

- ▼ Bibliografía de la Neología
- ▼ Terminología de la Neología



▼ Inguma Wikibase

- ▼ Metadatos de publicación
- ▼ Personas, congresos, organizaciones, lugares, lenguas...
- ▼ Objetivos:
 - ▼ Convertir todo a LOD
 - ▼ Transferir a o federar con Wikidata

Workflow para una bibliografía digital LOD y indexada

- Metadatos de publicación y textos completos en [Zotero](#) ([ejemplo](#))
- Volcado a Wikibase, desambiguación de autores usando [ZotWb](#)
- Procesamiento de textos completos
 - [GROBID](#): PDF > TEI-XML
 - [SpaCy](#): lematización
 - [Flashtext](#): encontrar términos en textos completos
- Resultados
 - [Ejemplo ENEOLI](#) (en curso)
 - [Ejemplo LexBib](#)

The screenshot shows the LexBib Wikibase interface. At the top, there is a search bar and navigation options. The main content area displays the title 'La lexicografía occitana entre recuperación del idioma y tesoro sentimental' (Q13706). Below the title, there are links for 'Item', 'Discussion', and 'History'. The item is identified as a 'Publication by Naro Rouquette (2014)'. A table below shows the item's representation in different languages:

Language	Label	Description	Also known as
English	La lexicografía occitana entre recuperación del idioma y tesoro sentimental	Publication by Naro Rouquette (2014)	
German	No label defined	No description defined	

Below the table, there are sections for 'Statements'. The first statement is 'instance of' with the value 'BibItem'. The second statement is 'LexBib Zotero item' with the value '45XIA79X'. Both statements have options to 'add reference' and 'add value'.

paper1 paper2

Inguma

- Comunicaciones científicas en euskera
 - Autores, Publicaciones Revistas, organizaciones, lugares
- Conversión desde **INGUMA** tablas SQL
- Una vez transferidos a Wikidata...
 - ejemplo: SCHOLIA

<http://wikibase.inguma.eus>

INGUMA WIKIBASE

Inguma Wikibase wikian bilatu

Konbertasen rola tumorearen garapenean eta metastasian

Item Eztabaida ☆

Ez da deskribapenik zehaztu

Hizkuntza gehiagotan Konfiguratu

Hizkuntza	Etiketa	Deskribapena
euskara	Konbertasen rola tumorearen garapenean eta metastasian	Ez da deskribapenik zehaztu
ingeles	Konbertasen rola tumorearen garapenean eta metastasian	Ez da deskribapenik zehaztu
aleman	Ez da etiketarik zehaztu	Ez da deskribapenik zehaztu

Adierazpenak

author ^{ingeles}	Patricia Garcia Gallastegui	serie hurrenkera	1
		0 erreferentzia	
	Olatz Crende Arruabarrena	serie hurrenkera	2

Nabigazioa

- Azala
- Aldaketa berriak
- Ausazko orria
- MediaWikiri buruzko laguntza

Wikibase

- New Item
- New Property
- New Schema
- All Properties
- Query Service
- Cradle
- QuickStatements

Wiki-tresnak

SCHOLIA Author Work Organization Location Event

author / Q212730

Steven Pinker (Q212730)

Steven Arthur Pinker is a Canadian-American cognitive psychologist, evolutionary psychologist and the computational theory of mind.

Related: Paul Greengard · Robert Darnton · Alfred Sturtevant · Michael Thomas Nagel · Paul Berg

<https://orcid.org/0000-0002-2319-4085> | [@sapinker](https://twitter.com/sapinker)

Table of Contents

- List of publications
 - Number of publications per year
 - Number of pages per year
- Topics

<https://scholia.toolforge.org/author/Q212730>

DL: Datos léxicos



▼ Qichwabase

- ▼ Lexemas en Quechua
- ▼ Fuente: diccionario CSV
- ▼ Objetivos:
 - ▼ Crear una comunidad de hablantes de quechua en esta Wikibase / en Wikidata
 - ▼ Enriquecer las descripciones de los lexemas
 - ▼ Federación / migración con / a Wikidata



▼ Kurdî Wikibase

- ▼ Lexemas kurdos
 - ▼ Diferentes dialectos
- ▼ Fuente: Ontolex TTL
- ▼ Objetivos:
 - ▼ Crear una comunidad de hablantes de kurdo en esta Wikibase en Wikidata
 - ▼ Enriquecer las descripciones de los lexemas
 - ▼ Federación / migración con / a Wikidata



▼ Monumenta Linguae Vasconum (MLV) Wikibase

- ▼ Un modelo para conectar corpus y diccionario
 - ▼ Textos en Wikisource
 - ▼ Corpus token: entidad en Wikibase
 - ▼ Datos léxicos enlazados
- ▼ Primer experimento:
 - ▼ Un manuscrito de 1737
 - ▼ [En Wikisource](#)
 - ▼ [En Wikibase](#)

Qichwabase

- ▼ Fuente: [Runasimi Dictionary CSV](#)
- ▼ Equivalentes multilingües
- ▼ Información dialectal

ver online

The screenshot shows the Qichwabase interface for the entry 'achka t'asla'. The page includes a navigation sidebar with links like 'Main page', 'Recent changes', and 'Tools'. The main content area displays the word 'achka t'asla' with its language 'Quechua' and category 'Lexical category noun'. It lists statements such as 'interim ID 112' and 'described in Runa Simi Dictionary'. A 'dialectal lemma variant' section lists 'achka tasla' and 'ashka t'asla' with their respective dialects and source references.

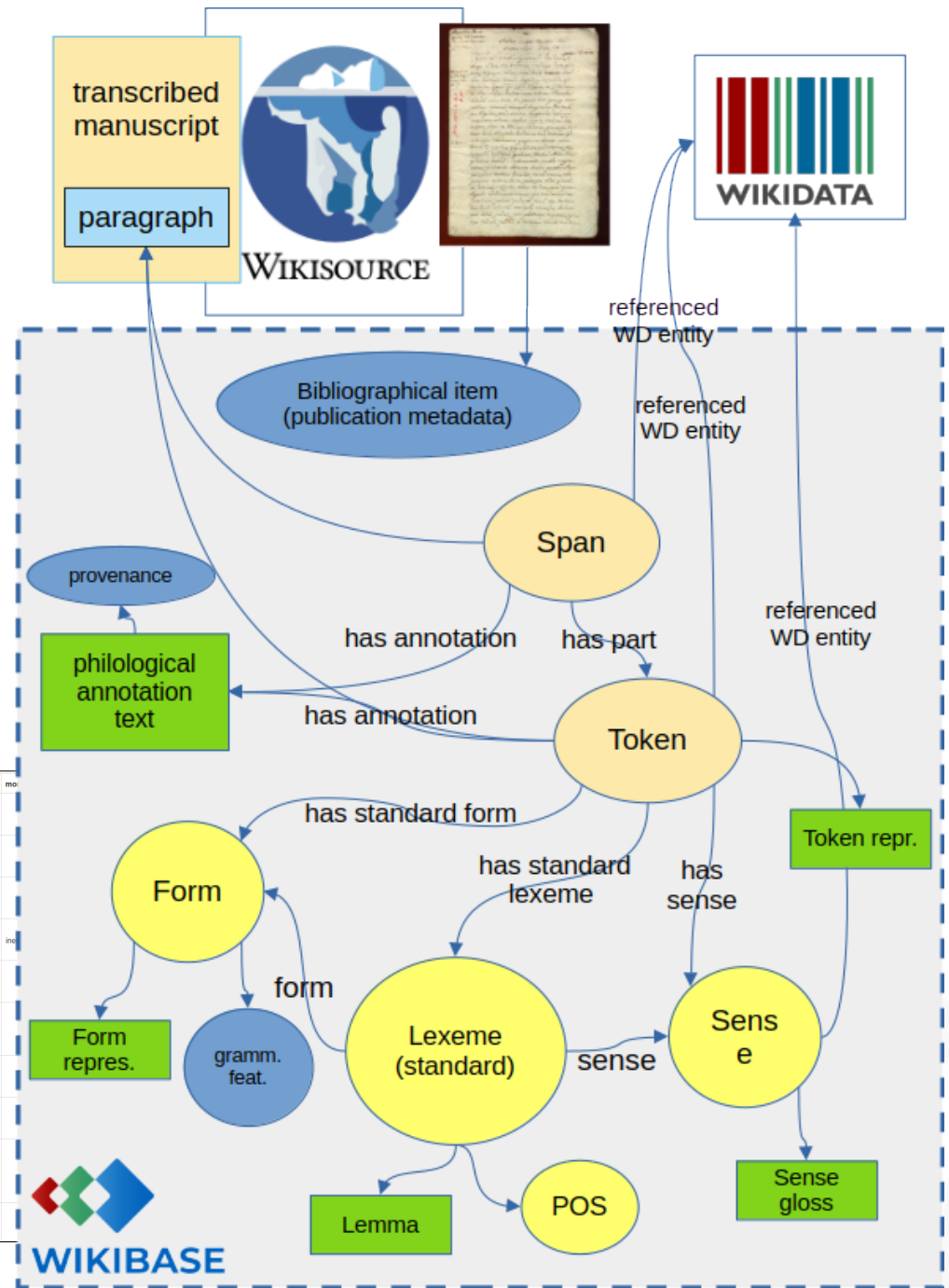
51	sich schmücken	to adorn	adornarse; ataviarse; embellecerse	decorarsi; abbellirsi; farsi bello; farsi bella
73	leuchten; scheinen; strahlen	to shine; to light	alumbrarse	illuminarsi
96	ja, gewiss; gewiss; sicherlich; sehr wahrscheinlich	yes indeed; certainly; surely; very probably	sí pues; seguramente; posiblemente; muy probable	si certo; sicuramente; probabilmente; molto probabilmente
100	viel; viele	a lot; much; many; much	mucho; mucha; muchos; muchas; harto; bastante; abundante; harto; mucho; varios; demasiado; bastante; mucho; bastante; gran cantidad; cuantioso; numeroso; gran porción; profusamente; tanto	molto; molta; molti; molte; troppo; sufficiente; abbondante; vari; grande quantità; numeroso
102	vergrößern; vermehren	to increase	aumentar; acrecentar	aumentare; accrescere

MLV Wikibase

- Fuente: Wikisource (ejemplo)
- Corpus anotado en Wikibase
 - anotaciones morfosintácticas
 - anotaciones filológicas
 - "corpus-dictionary interface"

ver online

token	token_zbk	token_forma	mlv_lexema	wikidata_sense	wd_pos_label	wikisource_paragraph	lemma	sense	forma	mo
Q->https://monumenta.wikibase.cloud/entity/Q1273	363.1	euangelioan	Q->https://monumenta.wikibase.cloud/entity/L18191>			->https://eu.wikisource.org/wiki/Azkoitko_Sermoiak4<	ebanjelo			
Q->https://monumenta.wikibase.cloud/entity/Q1274	402.1	ezi				->https://eu.wikisource.org/wiki/Azkoitko_Sermoiak4<				
Q->https://monumenta.wikibase.cloud/entity/Q454	1	Azcoitico	Q->https://monumenta.wikibase.cloud/entity/L1>			->https://eu.wikisource.org/wiki/Azkoitko_Sermoiak1<	Azkoita			
Q->https://monumenta.wikibase.cloud/entity/Q455	2	Parroquian	Q->https://monumenta.wikibase.cloud/entity/L94131>	wd:L72187-S3	substantibo	->https://eu.wikisource.org/wiki/Azkoitko_Sermoiak1<	parrokia	(Eri.) delako barmuko eliza nagusia, elkartearren elizkizunetarako biloki ofiziala dena	parrokan	ine
Q->https://monumenta.wikibase.cloud/entity/Q808	3	1737.				->https://eu.wikisource.org/wiki/Azkoitko_Sermoiak1<				
Q->https://monumenta.wikibase.cloud/entity/Q809	4	urtean	Q->https://monumenta.wikibase.cloud/entity/L79782>	wd:L8818-S1	substantibo	->https://eu.wikisource.org/wiki/Azkoitko_Sermoiak1<	urte	hamabi hilaletako aldia. Lurrak Eguzkiaren inguruan bira osoa emateko behar duen denborari dagokiona; eskuarki, urtarilaren leiki abenduaren 31ra bitariko denbora-bitartea		
Q->https://monumenta.wikibase.cloud/entity/Q816	5	.				->https://eu.wikisource.org/wiki/Azkoitko_Sermoiak1<				
Q->https://monumenta.wikibase.cloud/entity/Q811	6	Virgiharen	Q->https://monumenta.wikibase.cloud/entity/L13381>			->https://eu.wikisource.org/wiki/Azkoitko_Sermoiak1<	birjina			
Q->https://monumenta.wikibase.cloud/entity/Q812	7	icnaren	Q->https://monumenta.wikibase.cloud/entity/L38502>	wd:L61410-S1	substantibo	->https://eu.wikisource.org/wiki/Azkoitko_Sermoiak1<	izen	persona, animala edo gauza bat aipatzen edo adierazten duen hitz; banako bat aipatzeko eta kolektiboak bereizteko erabiltzen den hitz edo hitz-multzoa; kategoria edo mota bereko izakiak edo gauzak adierazteko erabiltzen den hitza		
Q->https://monumenta.wikibase.cloud/entity/Q813	8	egunean	Q->https://monumenta.wikibase.cloud/entity/L19077>	wd:L50616-S1	substantibo	->https://eu.wikisource.org/wiki/Azkoitko_Sermoiak1<	egun	Lurrak bere ardatzaren inguruan bira osoa egiten ematen duen denbora; 24 orduko denbora-bitartea, gaurerik gaurdera		



Ejercicio Open Refine



Extensions Wikibase

« first < previous 1 next > last »

publisher

- Facet
- Text filter
- Edit cells
- Edit column
- Transpose
- Sort...
- View
- Reconcile

Start reconciling...

Reconcile text in this column with items in another data source

Actions

- Copy reconciliation data...
- Use values as identifiers...
- Add entity identifiers column...

- ▶ Crear un proyecto, con un CSV de entrada
- ▶ "Reconciliar": identificar valores literales con entidades de Wikidata
- ▶ Exportar resultados

publisher	Wikidata
University of Southern Brittany Choose new match	Q1125958
University of Deusto Choose new match	Q1468481
EliPhi. Éditions de linguistique et de philologie Choose new match	Q108267805
Euskaltzaindia Choose new match	Q202042
Euskaltzaindia Choose new match	Q202042
Auñamendi (argitaletxea) Choose new match	Q3579236
Royal Galician Academy Choose new match	Q2728155
Government of Navarra Choose new match	Q578870
Isaac López Mendizábal Choose new match	Q3772654
Imprenta de Fortanet Choose new match	Q19491061

Match this cell Match all identical cells Cancel

WIKIDATA

EliPhi. Éditions de linguistique et de philologie (Q108267805)
publisher in France

impr. de F. de P. Diaz (Sevilla)
 Create new item
Search for match

University of Southern Brittany
Choose new match

University of Deusto
Choose new match

Éditions de Linguistique et de Philologie
EliPhi. Éditions de linguistique et de philologie (92)
[vø.rta.r]. Mélanges de linguistique, de philologie et d'histoire ancienne offerts à Rudolf Wachter (52)
Create new item
or match

zaindia
new match

zaindia
new match

Diccionario vasco-español

Auñamendi

- Pic d'Anie (100)
- Auñamendi Eusko Entziklopedia (100)
- Auñamendi (100)
- Auñamendiko lorea (69)
- Auñamendiko lorea (69)
- Auñamendi (argitaletxea) (67)
- Auñamendi (mitologia) (64)
- Bernardo Estornés Lasa (58)
- Enciclopedia General Ilustrada del País Vasco (56)
- Sir Juan de Arzadun (48)

<https://etherpad.wikimedia.org/p/31-mayo-2024>



Eskerrik asko

david.lindemann@ehu.eus

<https://www.wikidata.org/wiki/Q57694630>