

RDF

Mikel Egaña Aranguren

mikel-egana-aranguren.github.io

mikel.egana@ehu.eus



RDF

<https://github.com/mikel-egana-aranguren/DBA>



Web Semantikoa

Web sare berri baterako proiektu utopikoa: baliabide bakoitzak bere datuak argitaratzen ditu eta agente adimendunek modu autonomoan kontsumitzen dituzte

Neurri batean iadanik inplementatua (Adib. [JSON-LD](#) + [Schema](#): [Google Knowledge Graph](#))

LLM-ak?

Web Semantikoa

Teknologiak:

- [URI](#): identifikatzaileak
- [RDF](#): datuak
- [SPARQL](#): RDF kontsultak
- [SHACL](#): RDF balioztatzea
- [OWL](#): RDF datuak deskribatzeko hiztegiak (ontologiak)

RDF

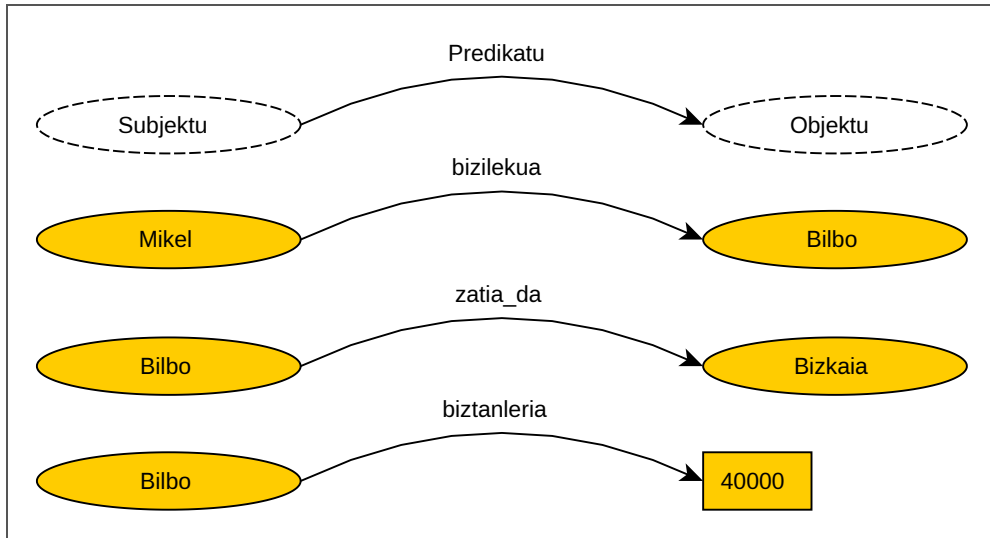
RDF: Resource Description Framework

[W3C-ren estandar ofiziala](#) baliabideak web sarean argitaratzeko

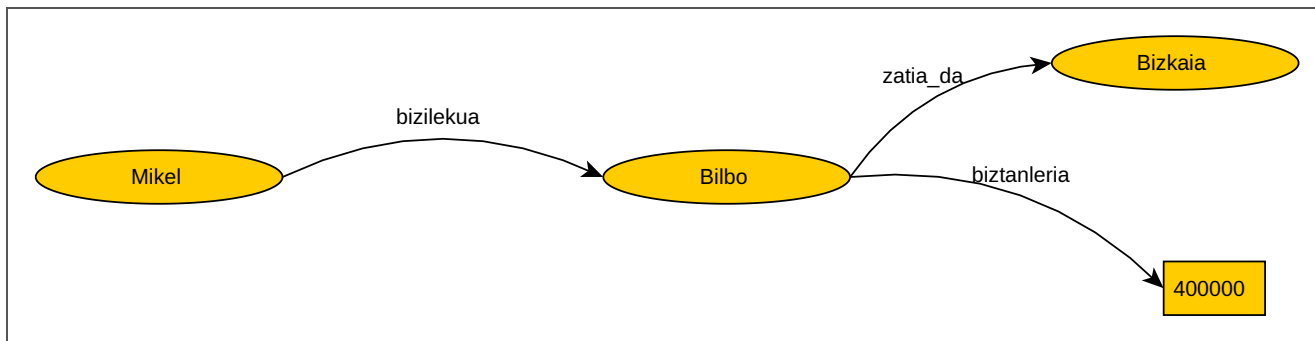
Gaur egun bere erabilera hazten ari da enpresa eremuan, datuak Knowledge Graphs (KGs) bidez gordetzeko, [datuetan-zentratutako arkitekturak](#) inplementatzeko

RDF triplea

RDF tripleak subjektu-predikatu-objektu eredua erabiltzen du, Euskarak ez bezala (subjektu-objektu-predikatu)



RDF grafoa



RDF grafoa

Grafoaren banako guztiak URLen bidez identifikatzen dira

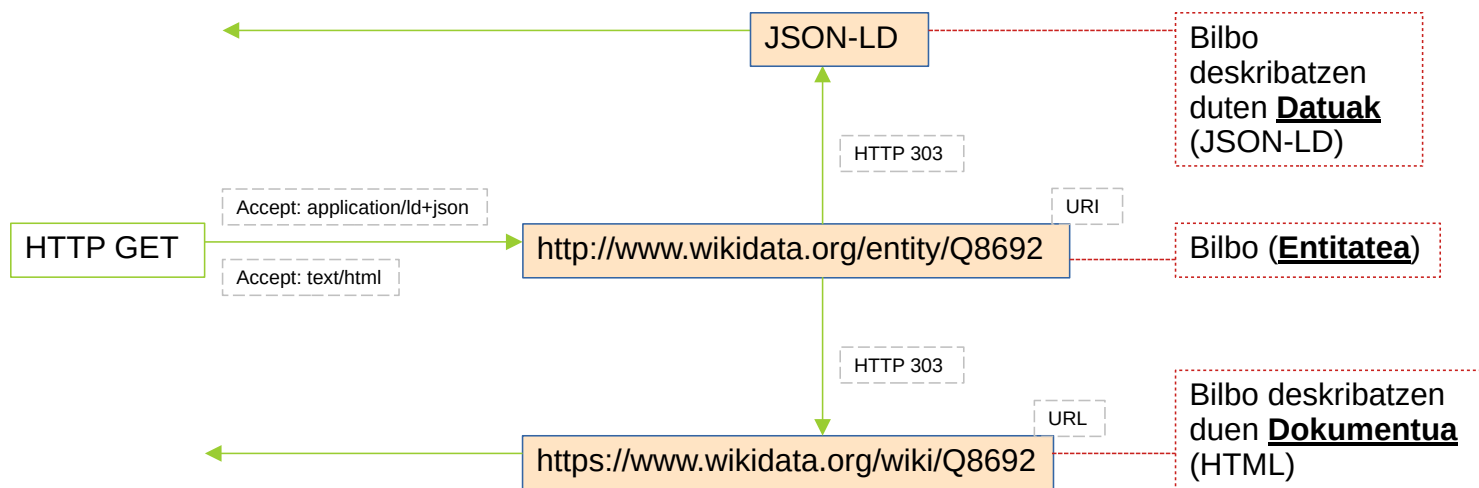
URI: Uniform Resource Identifier (\neq URL!). Baliabideak identifikatzen ditu:

<http://www.wikidata.org/entity/Q8692>

URL: Uniform Resource Locator. Sarean (Baliabide bat deskribatzen duen) dokumentu baten kokapen fisikoa adierazten duen URI-a:

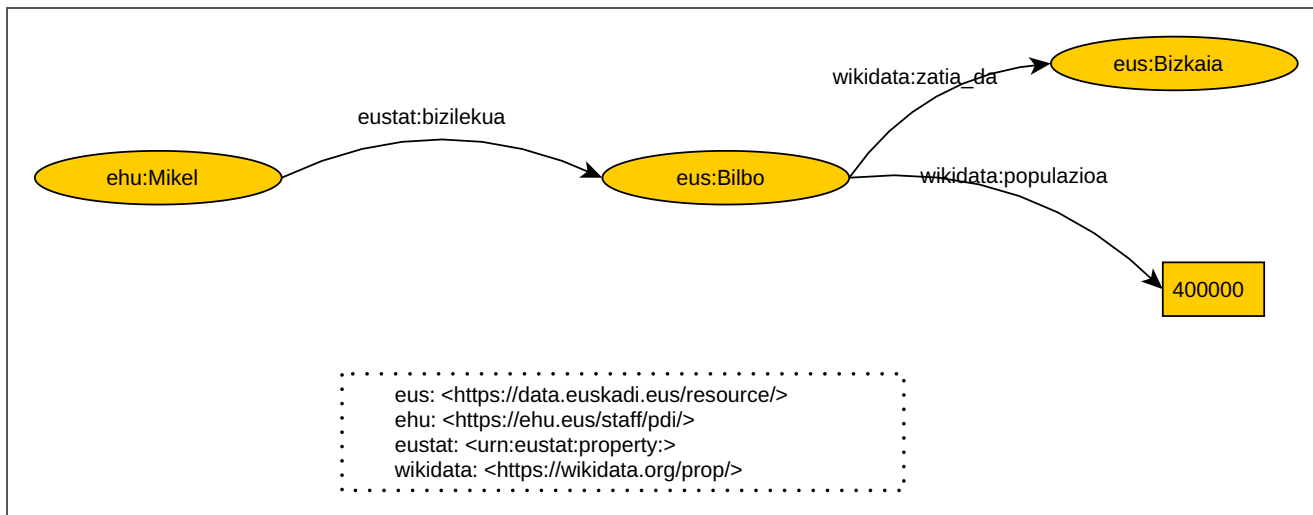
<https://www.wikidata.org/wiki/Q8692>

RDF grafoa: URI-ak



(Adibidea exekutatu)

RDF grafoa: URI-ak



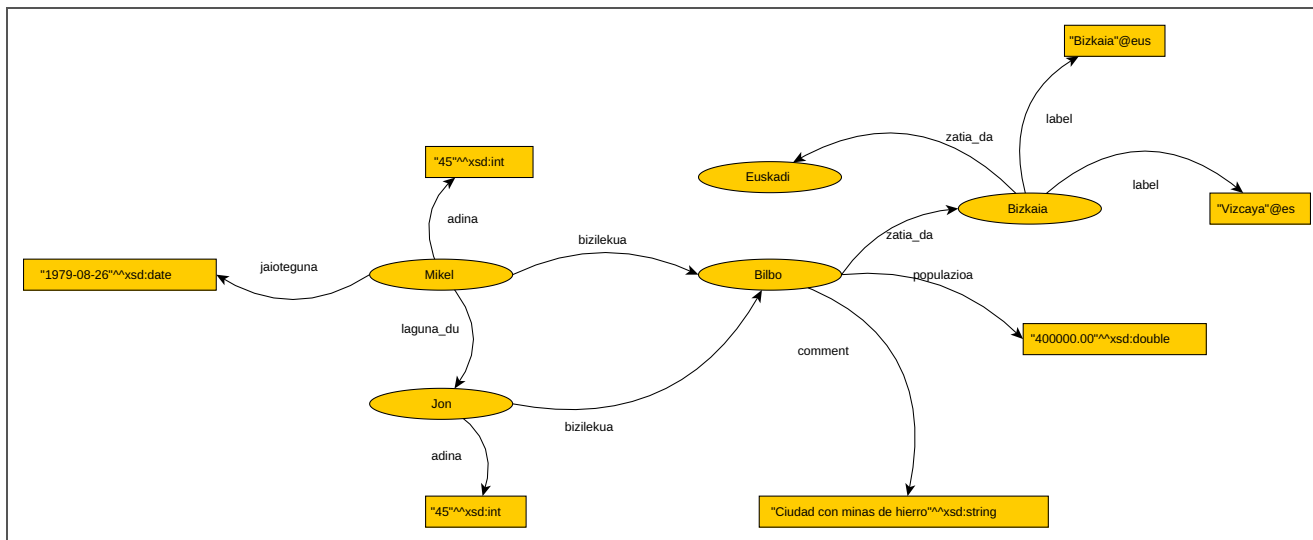
RDF grafoa

Subjektuak eta predikatuak baliabideak soilik izan daitezke (URLak)

Zenbait objektu balio literalak izan daitezke (Karaktere kateak)

Balio literalek datatype-a izan dezakete ([XML Schema Datatypes](#))

RDF grafoa: datatype-ak



RDF: eredua eta sintaxia

RDF datuak kodifikatzeko eredua da

Eredu abstraktu hori sintaxi desberdinen bidez irudika daiteke: Fitxategi batean "Serializatu" (idatzi)

Sintaxi horietako bat RDF/XML da

Ez nahastu eredua sintaxiarekin: RDF XML fitxategi bat baino askoz gehiago da!

RDF serializatu

RDF/XML (<http://www.w3.org/TR/rdf-syntax-grammar/>)

RDFa (<http://www.w3.org/TR/rdfa-core/>)

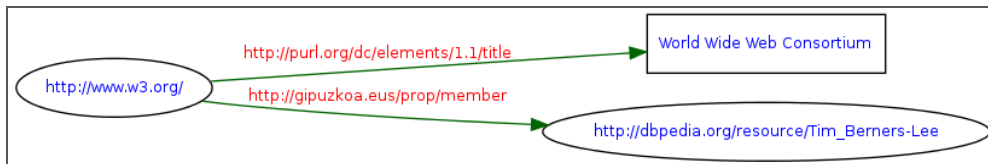
Turtle (<http://www.w3.org/TR/turtle/>)

N3 (<http://www.w3.org/DesignIssues/Notation3.html>)

JSON-LD (<https://www.w3.org/TR/json-ld11/>)

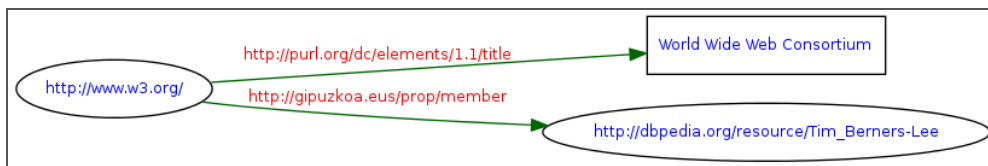
RDF serializatu: RDF/XML

```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:gip="http://gipuzkoa.eus/prop/">
  <rdf:Description rdf:about="http://www.w3.org/">
    <dc:title>World Wide Web Consortium</dc:title>
    <gip:member rdf:resource="http://dbpedia.org/resource/Tim_Berners-Lee"/>
  </rdf:Description>
</rdf:RDF>
```



RDF serializatu: Turtle

```
@prefix dc11: <http://purl.org/dc/elements/1.1/> .  
@prefix ns0: <http://gipuzkoa.eus/prop/> .  
  
<http://www.w3.org/>  
  dc11:title "World Wide Web Consortium" ;  
  ns0:member <http://dbpedia.org/resource/Tim_Berners-Lee> .
```



Tresna erabilgarriak

<http://www.w3.org/RDF/Validator/>

<http://rdf-translator.appspot.com/>

<https://www.easyrdf.org/converter>

Triple Store-ak

RDF W3C estandar bat da (=HTML, !=SQL), RDF fitxategi batetik abiatuz:

- Datu eredua: Triple Store guztiek berdina kargatzen dute
- Sintaxia (RDF/XML, Turtle, ...): Triple Store guztiek berdina kargatzen dute

Triple Store-ak internet existitzen zenean sortu ziren (NoSQL!):

- Guztiek HTTP(S) SPARQL bidez funtzionalitate bera eskaintzen dute
- Federazioa

Triple Stores

[Amazon Neptune:](#)

- Amazon AWS
- Property Graph-ak biltegitratzeko balio du ere

Triple Store-ak

[Virtuoso](#)

[Stardog](#)

[GraphDB](#)

[RDFox](#)

GraphDB

- Arrazoiketa automatikoa
- Lucene edo Kafka bezalako zerbitzuetarako konektoreak
- Datu Base erlazionalen birtualizazioa (Ontop)
- ...
- [1-1, 1-2 adibideak exekutatu]

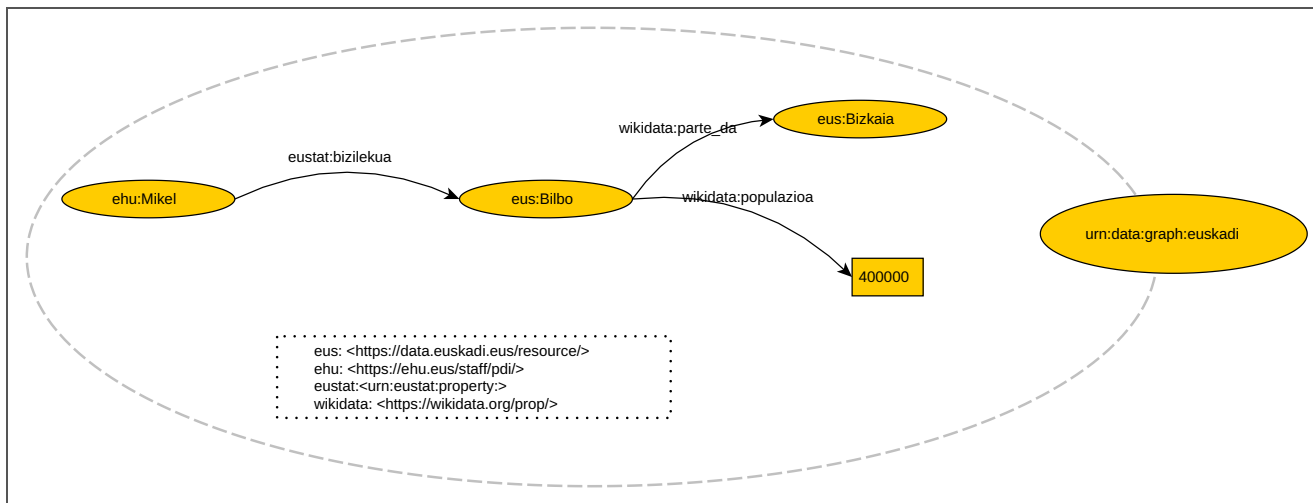
RDF grafoak

Tripleak URI batek identifikatutako "poltsa" batean biltzea (datuekiko desberdina): Named Graphs

Implementazioaren araberakoa: Triple Store guztiek Default Graph bat dute, non triple guztiak biltzen diren

[Adibidea exekutatu]

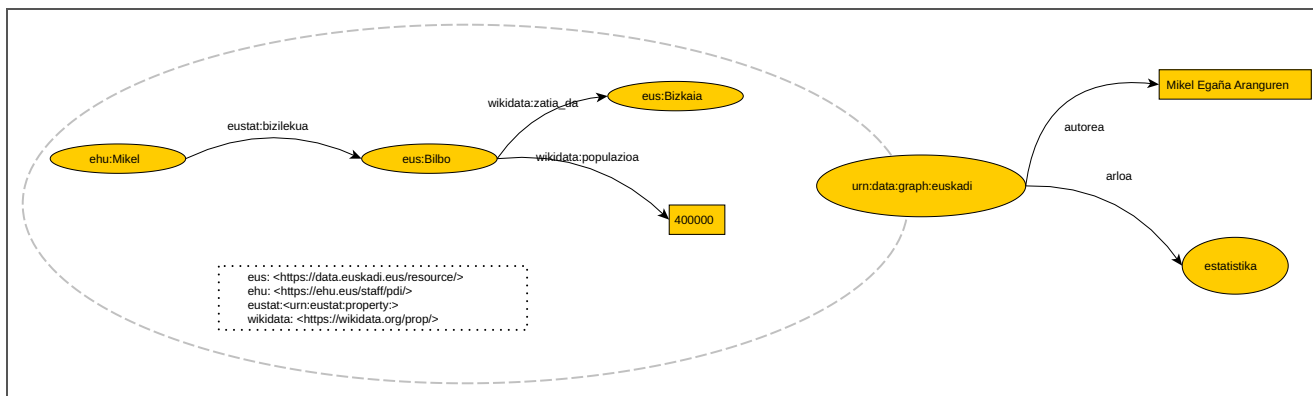
RDF grafoak



[Adibidea exekutatu]

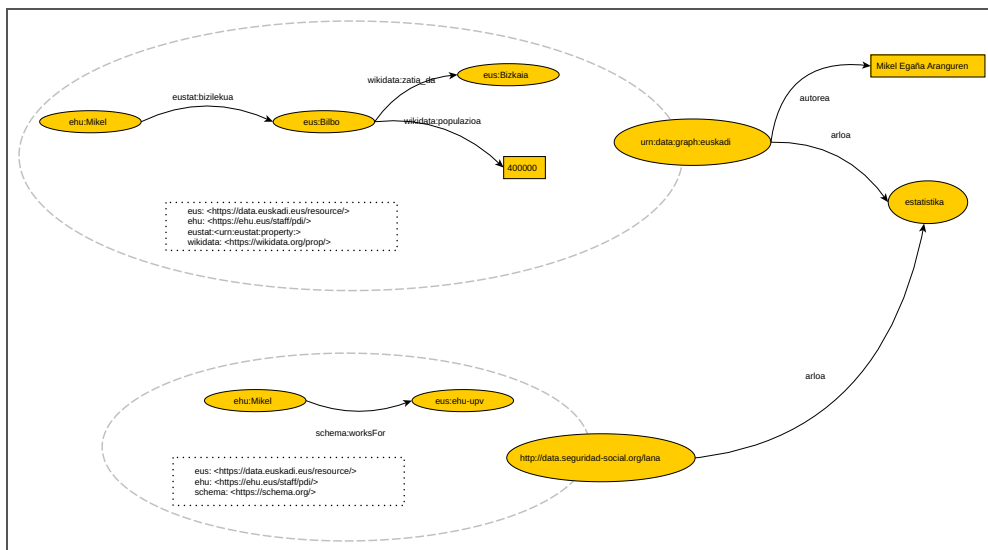
RDF grafoak

Datuak eta metadatuak hizkuntza berdinean (NoSQL! RDF!)



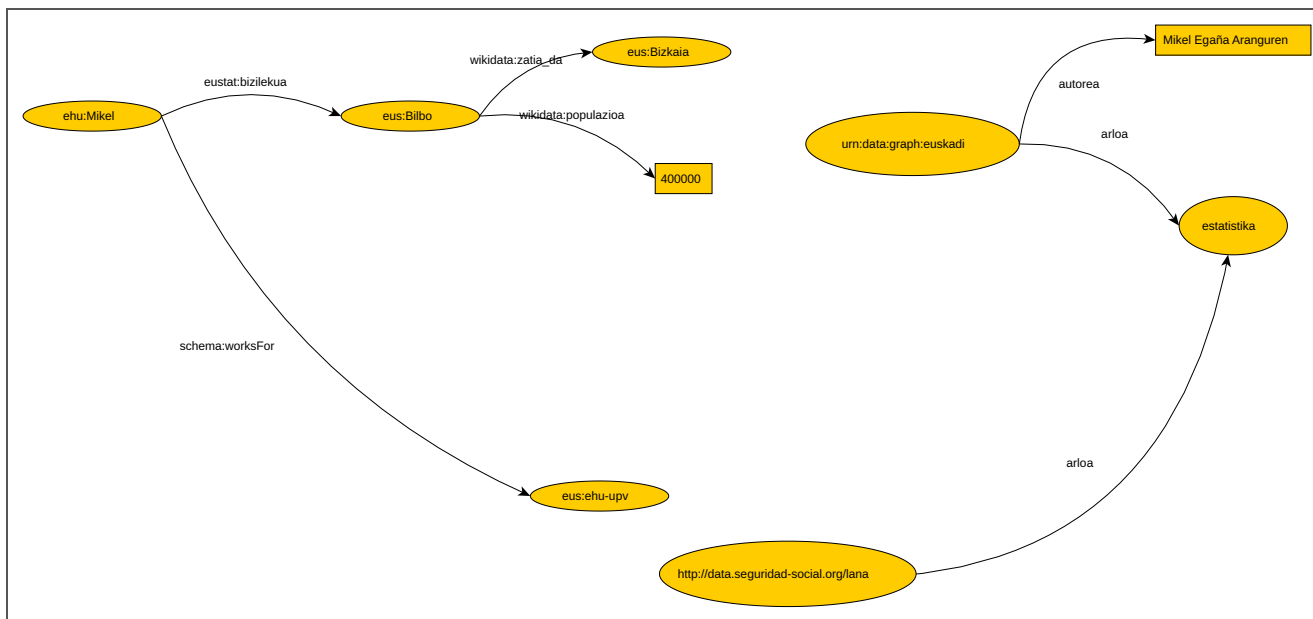
RDF grafoak

Datuak eta metadatuak hizkuntza berdinean (NoSQL! RDF!)



RDF grafoak

Default Graph

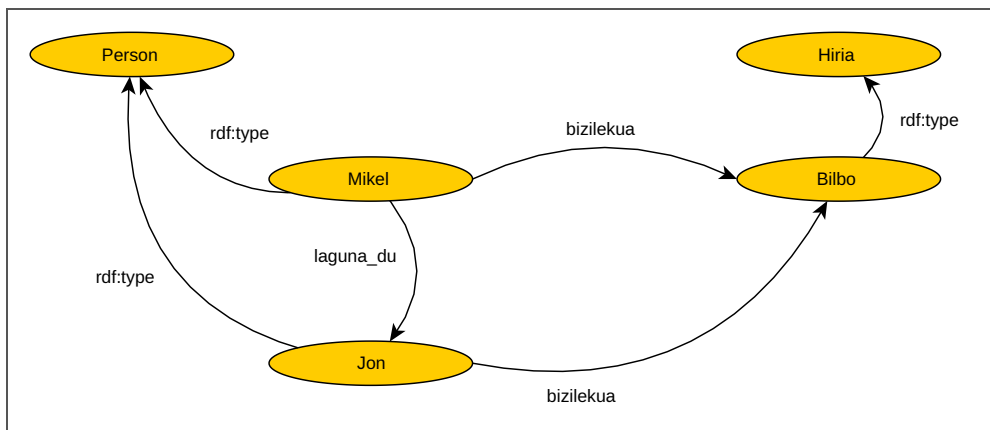


RDF grafoak

[Adibideak exekutatu]

RDF grafoak

rdf:type prdikatua: baliabideak klaseetan taldekatu



[Adibidea exekutatu - Class hierarchy]

Hiztegi erresebatuak (W3C estandarrak - NoSQL!)

Hizkuntza bera definitzen dute

RDF: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> (Adib. `rdf:type = http://www.w3.org/1999/02/22-rdf-syntax-ns#type`), RDFek definitua (Bere burua definitzen du), RDFSek eta OWLek

RDFS: <http://www.w3.org/2000/01/rdf-schema#> RDFek, RDFSek (Bere burua definitzen du) eta OWLek definitua

Hiztegi erreserbatuak (W3C estandarrak - NoSQL!)

Hizkuntza bera definitzen dute

SHACL: <https://www.w3.org/ns/shacl.ttl> RDFek, RDFSek eta OWLek definitua

OWL: <http://www.w3.org/2002/07/owl#> RDFek, RDFSek eta OWLek definitua
(Bere burua definitzen du)

Domeinu hiztegiak (Ontologiak)

Domeinu zehatza deskribatzen dute: geografia, biologia, merkataritza, ...

Argitaratu nahi ditugun datuen propietate orokorrak definitzen dituzte

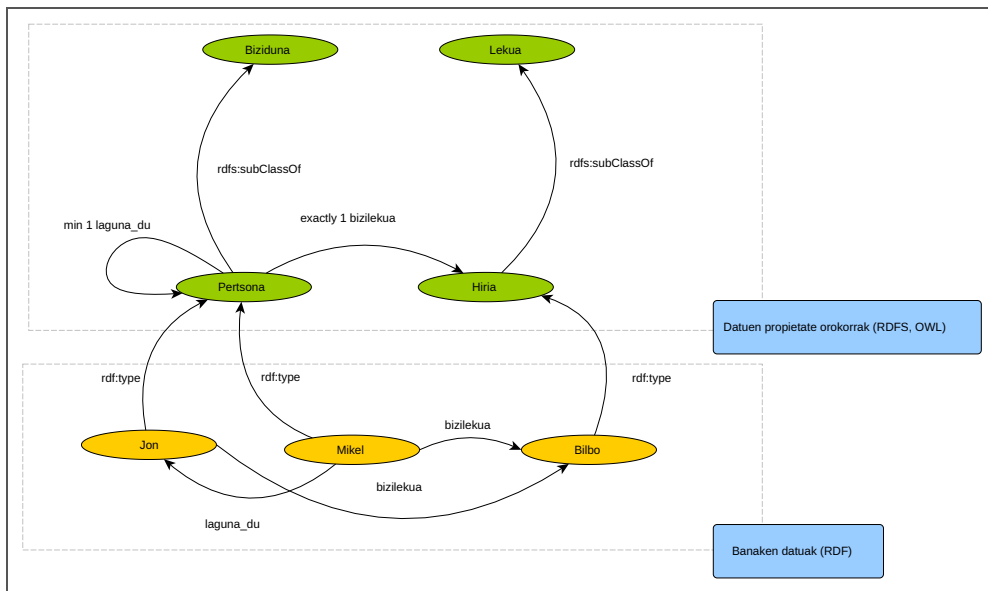
foaf:person, dbpedia-ont:city, dc:book, schema:person, ...

Ontologiak

RDF: datu espezifikoak (Mikel, Jon, Bilbo, ...)

Ontologiak (RDFS, OWL): datu horien propietate orokorrak (Pertsona, Hiria, ...)

Ontologiak



Tresna erabilgarriak

[Prefix.cc](https://prefix.cc)

[Linked Open Vocabularies](#)