UNIVERSIDADE FEDERAL
DE SANTA CATARINA

Linked Open Data to Enrich the World

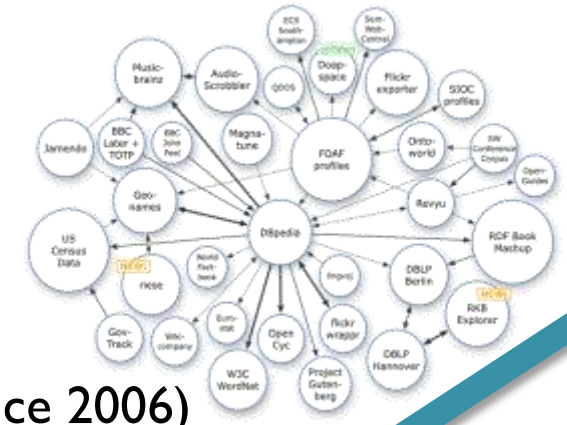
Renato Fileto
Cleto May

ine INFORMÁTICA E ESTATÍSTICA
CTC • UFSC
GBD
INE-UFSC

Topics

1. **Introduction**
   - **Motivation**
2. Linked Open Data (LOD)
   - Ontologies
   - Semantic Web Standards
   - The Web of Data
   - LOD Collections and Tools
3. Semantic Enrichment of Movement Data
   - Moving Object Trajectories
   - The Baquara Ontology
   - Trajectories *X* Social Media Trails
   - Fusion of Trajectories with Social Media Trails (Ricardo)
   - Connecting Annotated Movement Data with LOD (Cleto)
4. Conclusions and Future Work

From the Document Web to the *Linked Open Data Web* (and beyond)



Data Web (since 2006)
- URI de-referencability
- RDF serializations

Semantic Web
(Vision 1998, starting ???)
- Reasoning
- Logic, Rules
- Trust

Social Web (since 2003)
- Folksonomies/Tagging
- Reputation, sharing
- Groups, relationships

Web (since 1992)
- HTTP
- HTML/CSS/JavaScript

Web 1.0

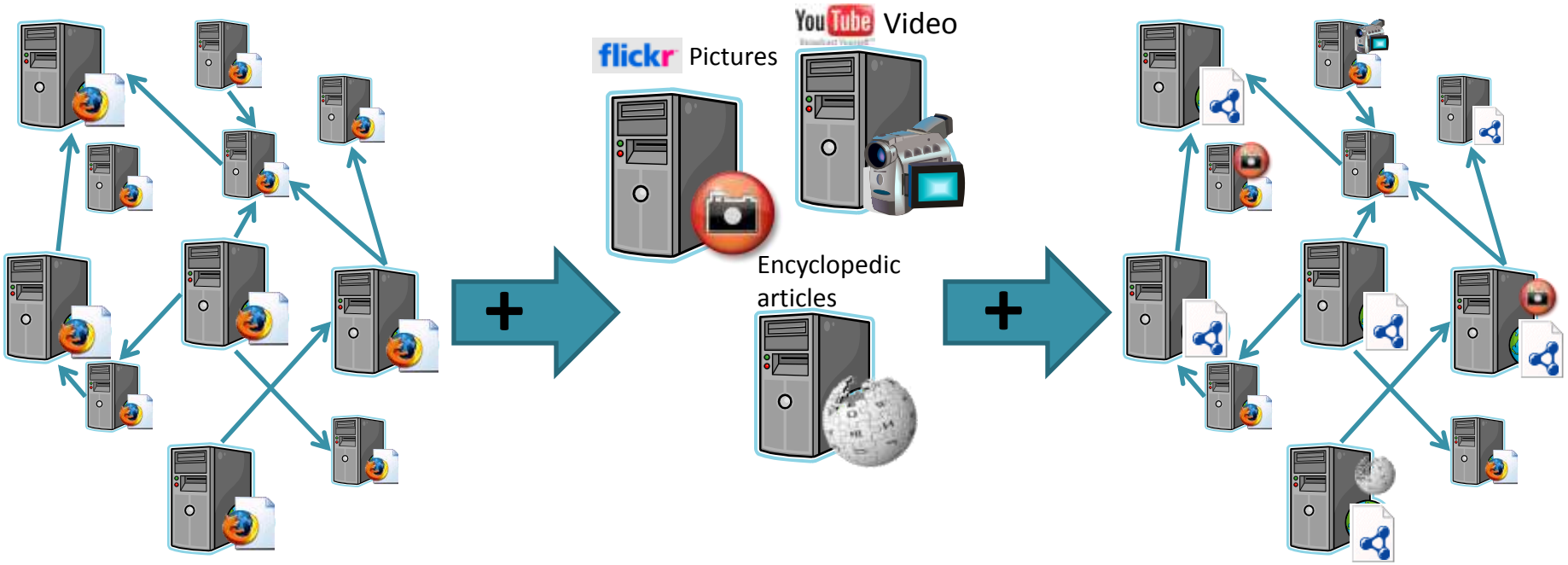
Many Web sites containing unstructured contents

Web 2.0

Few large Web sites specialized on specific content types

Web 3.0

Semantically syndicating arbitrarily structured content



Conceptual Level
Data Access and Integration

Data Integration

Enterprise Information Integration
sets of heterogeneous data sources appear as a single, homogeneous data source

Data Warehousing
- Based on extract, transform load (ETL)
- Global-As-View (GAV)

Research
Mediators
Ontology-based
P2P
Web service-based

Data Web
- URIs as entity identifiers
- HTTP as data access protocol
- Local-As-View (LAV)

Data Access

Object-relational mappings (ORM)
- NeXT's EOF / WebObjects
- ADO.NET Entity Framework
- Hibernate

Procedural APIs
- ODBC
- JDBC

Linked Data
- de-referencable URIs
- RDF serialization formats

Query Languages
- Datalog, SQL
- **SPARQL**
- XPATH/XQuery

Data Models

RDBMS
- Organize data in relations, rows, cells
- Oracle, DB2, MS-SQL

Column-oriented DBMS
- Collocates column values rather than row values
- Vertica, C-Store, MonetDB

Triple/Quad Stores
- RDF data model
- Virtuoso, Oracle, Sesame

Entity-attribute-value (EAV)
- HELP medical record system, TrialDB

Others
- XML, hierachical, tree, graph-oriented DBMS

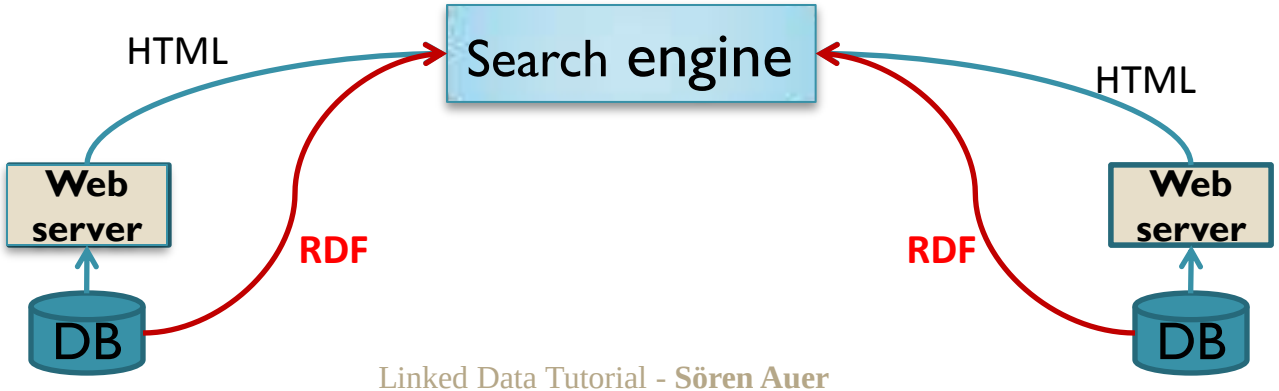
Why Do We Need Another Web?

Try to search for these things on the current Web:

- **Researchers working on LOD and social media in Floripa.**
- **Moving objects that stop at Touristic Hot Spots.**

Information to answer such queries **is available** on the Web, **but opaque to current Web search**.

Semantic Data Web allows to intelligently combine and integrate such structured information from different sources:



Topics

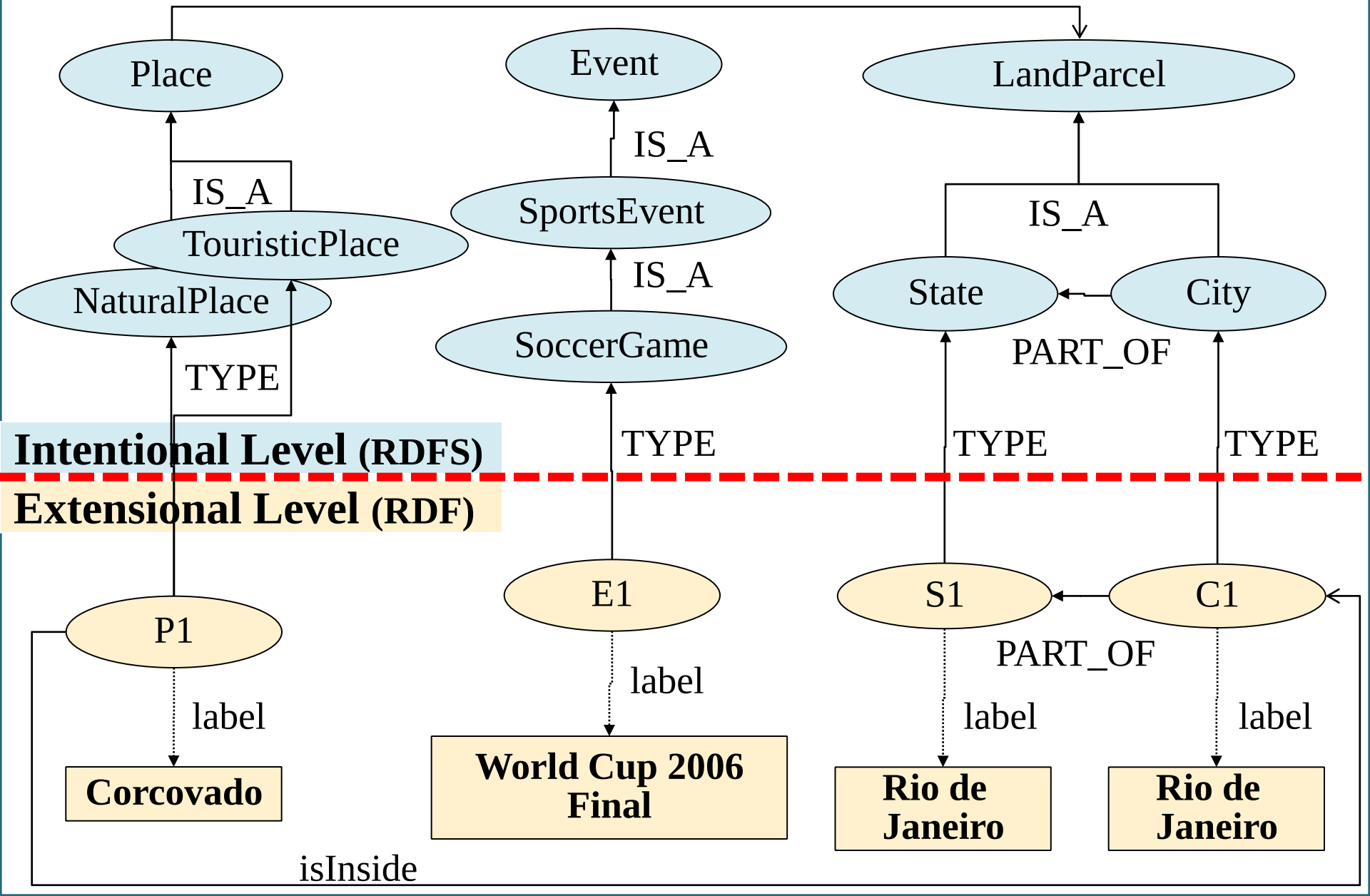
1. Introduction
   - Motivation
2. **Linked Open Data (LOD)**
   - **Ontologies**
   - **Semantic Web Standards – RDF/RDFS**
   - **The Web of Data**
   - **LOD Collections and Tools**
3. Semantic Enrichment of Movement Data
   - Moving Object Trajectories
   - The Baquara Ontology
   - Trajectories *X* Social Media Trails
   - Fusion of Trajectories with Social Media Trails (Ricardo)
   - Connecting Annotated Movement Data with LOD (Cleto)
4. Conclusions and Future Work

Ontology

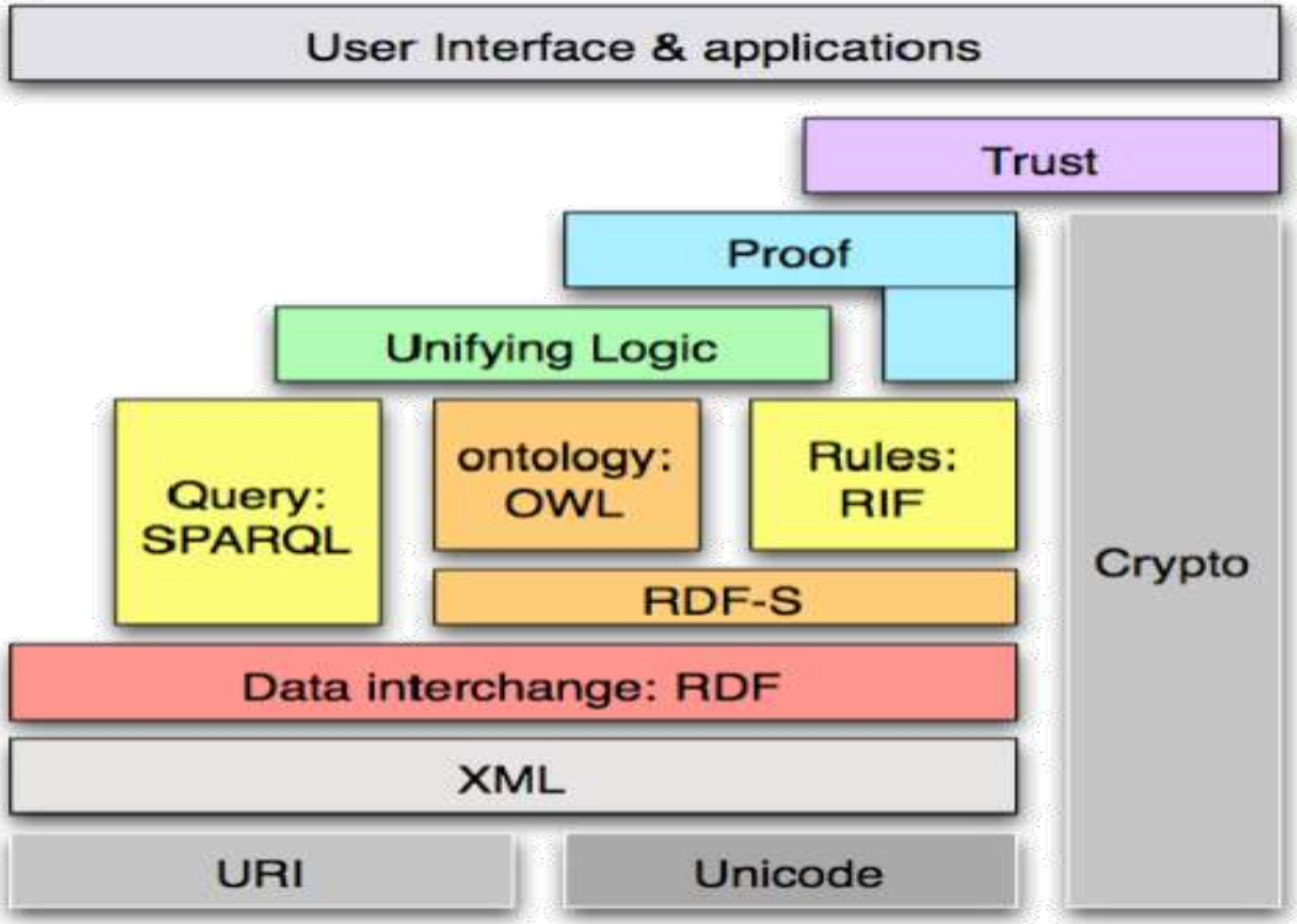
- **Shared conceptualization**
- **Unified view for a universe of discourse**

- **Related stuff**
  - **Taxonomy**
  - **Thesaurus**
  - **Class Diagram**
  - **Knowledge base**
    - Classes, properties and their relationships
    - Instances of classes

An Ontology Example



Semantic Web standards



RDF/RDFS

Standard language & data model to represent information in the Semantic Web.

An **RDF statement** is a triple of the form:

- **Resource**: anything that has a URL
- **Property**: any property of the resource
- **Value**: a literal or another resource

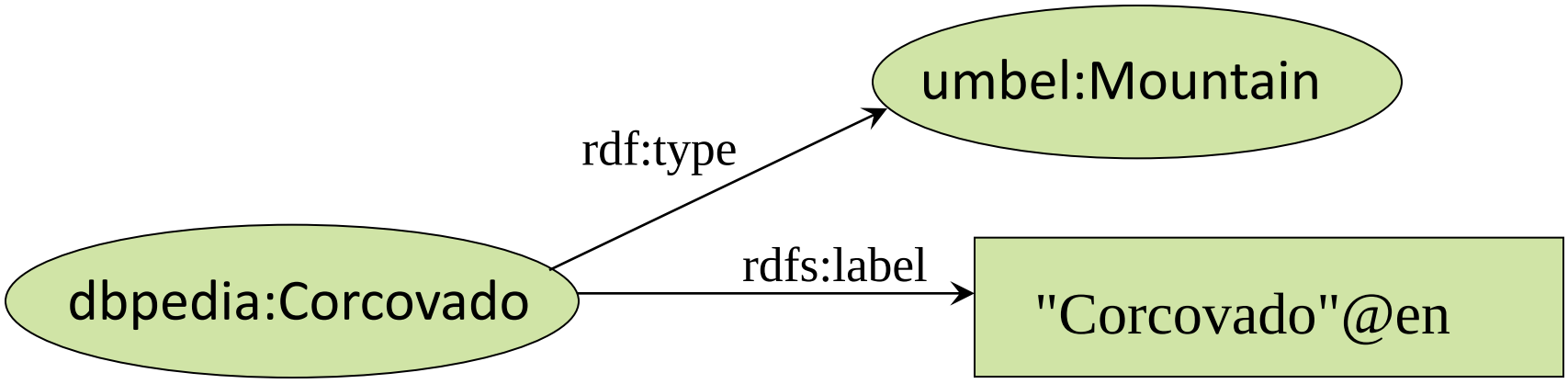
RDFS (RDF-Schema) defines the classes of resources, the possible properties for each class, and the possible values for these properties.

RDF represents information according to a conceptual model expressed in RDF.

RDF Statement / Triple Paradigm

RDF/N3:

http://**dbpedia**.org/resource/**Corcovado**
rdf:type http://**umbel**.org/umbel/rc/**Mountain** ;
rdfs:label **"Corcovado"@en** ;
rdfs:label "Corcovado (Brésil)"@fr ;
http://www.w3.org/2003/01/geo/wgs84_pos#geometry
"POINT(-43.2117 -22.9524)"

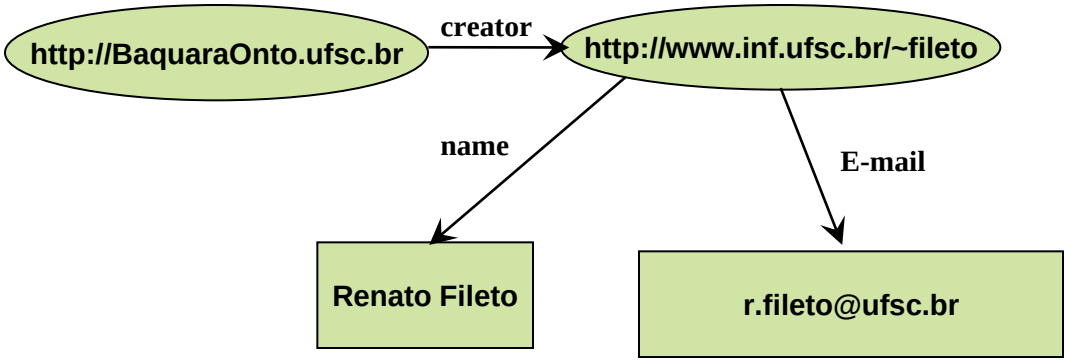


Subject (Resource)

Predicate (Resource)

Object (Resource/Literal)

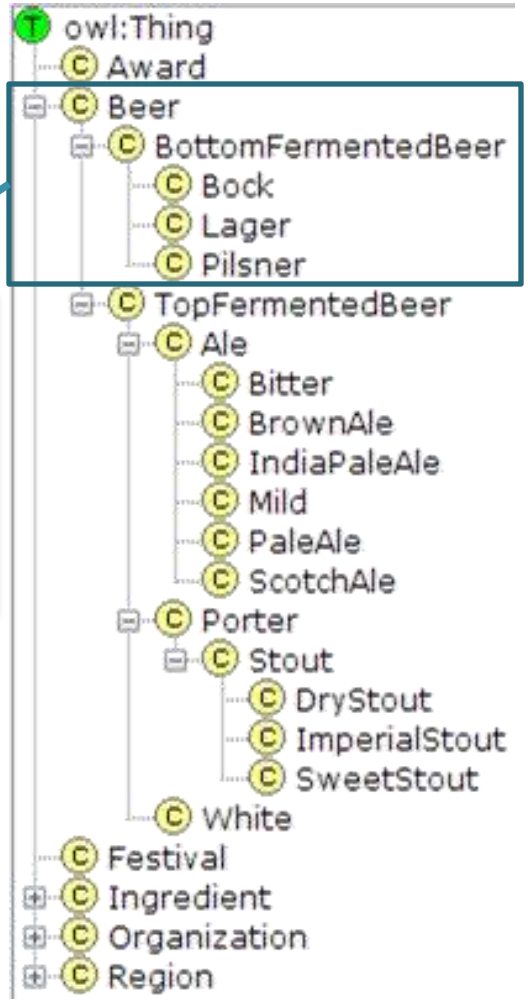
RDF Serialization




```
<?xml version="1.0"?>
<rdf:RDF
    xmlns="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:dc="http://purl.org/metadata/dublin_core#">
  <rdf:Description about="http://BaquaraOnto.ufsc.br ">
    <dc:Creator>
      <rdf:Description>
      <rdf:Description about="http://www.inf.ufsc.br/~fileto">
        <dc:Name>Renato Fileto</dc:Name>
        <dc:Email>r.fileto@ufsc.br</dc:Email>
      </rdf:Description>
    </dc:Creator>
  </rdf:Description>
</rdf:RDF>
```

http://BaquaraOnto.ufsc.br	http://purl.org/metadata/dublin_core#Creator	http://www.inf.ufsc.br/~fileto
http://www.inf.ufsc.br/~fileto	http://purl.org/metadata/dublin_core#Name	“Renato Fileto"
http://www.inf.ufsc.br/~fileto	http://purl.org/metadata/dublin_core#Email	r.fileto@ufsc.br

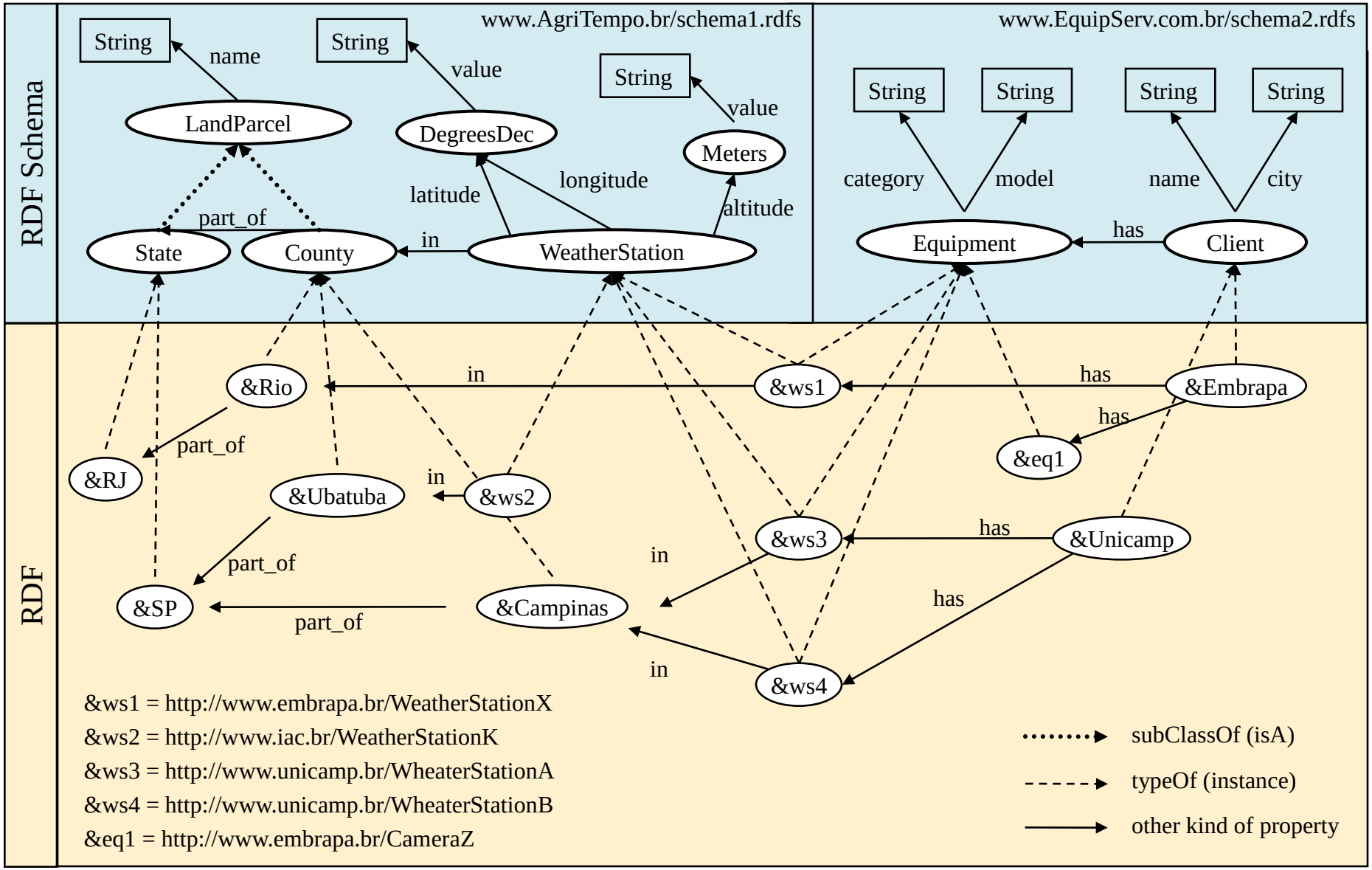
RDF-S Class & Property Hierarchies



Beer	rdf:type	rdfs:Class
BottomFermentedBeer	rdfs:subClassOf	Beer
Bock	rdfs:subClassOf	BottomFermentedBeer
Lager	rdfs:subClassOf	BottomFermentedBeer
Pilsner	rdfs:subClassOf	BottomFermentedBeer

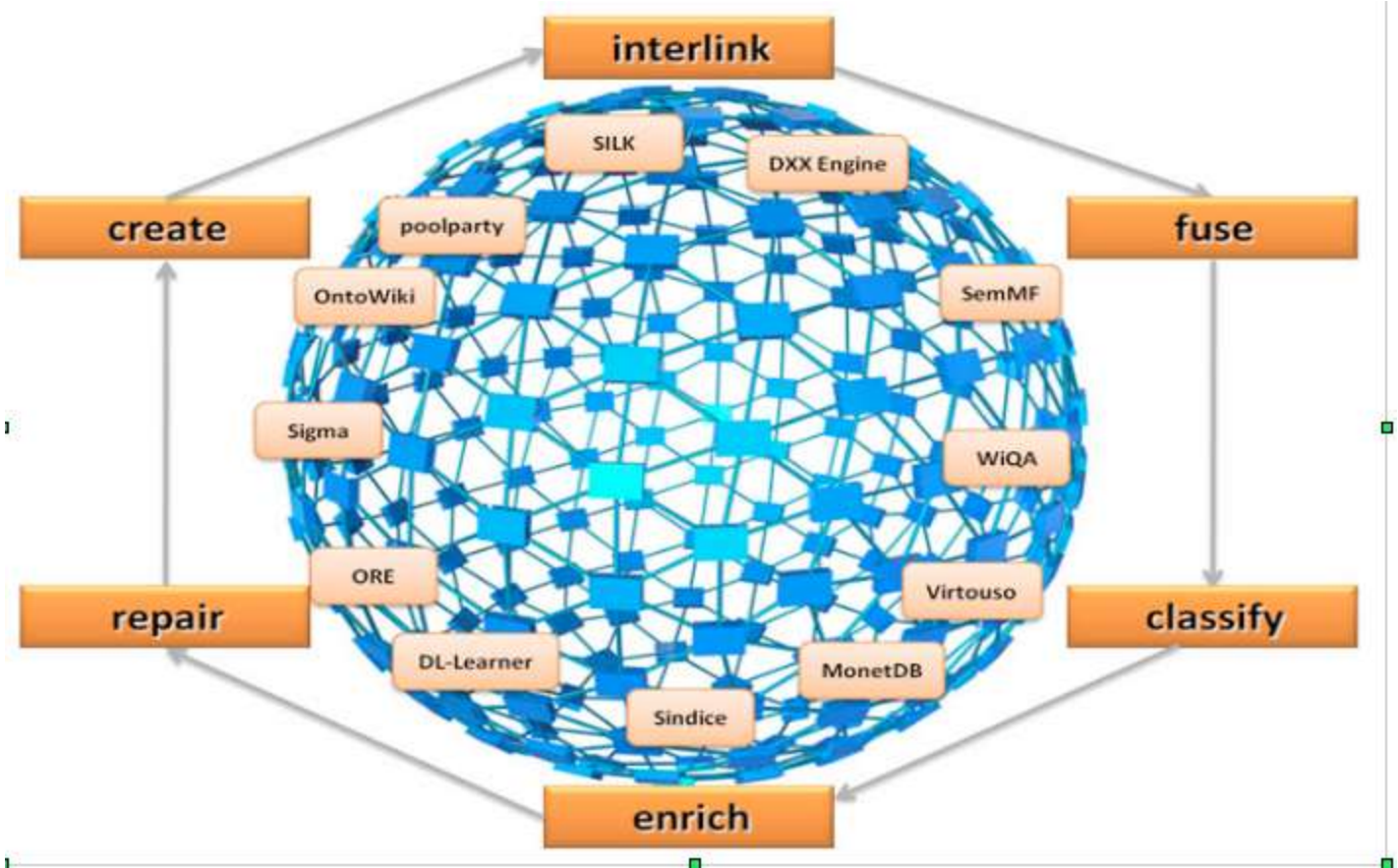
hasContent	rdf:type	rdfs:Property
hasAlcoholicContent	rdfs:subPropertyOf	Beer
hasOriginalWortContent	rdfs:subClassOf	BottomFermentedBeer

Each RDF instance can be associated to several classes

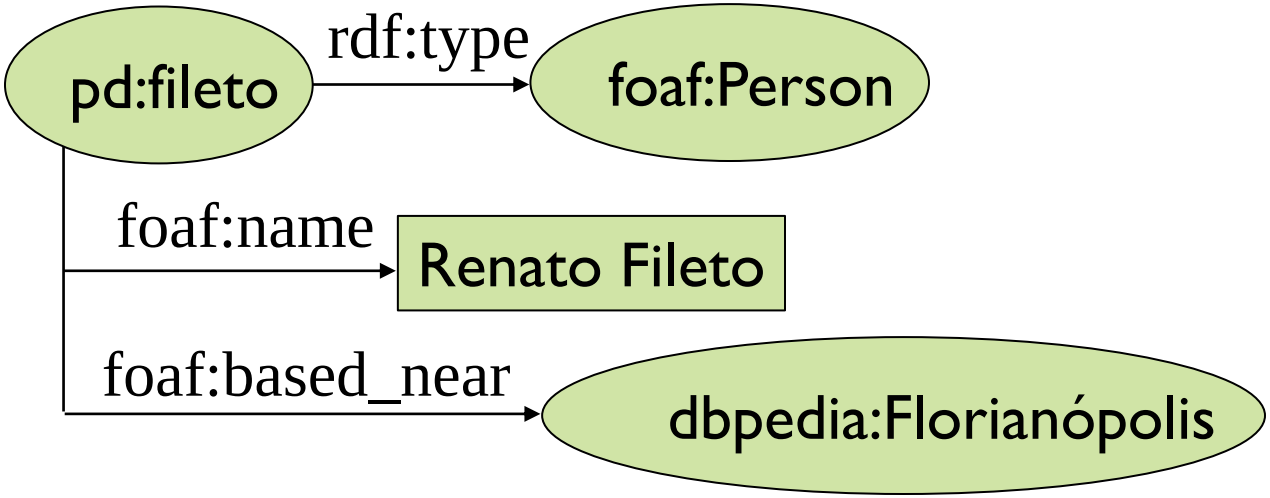


Linked Open Data

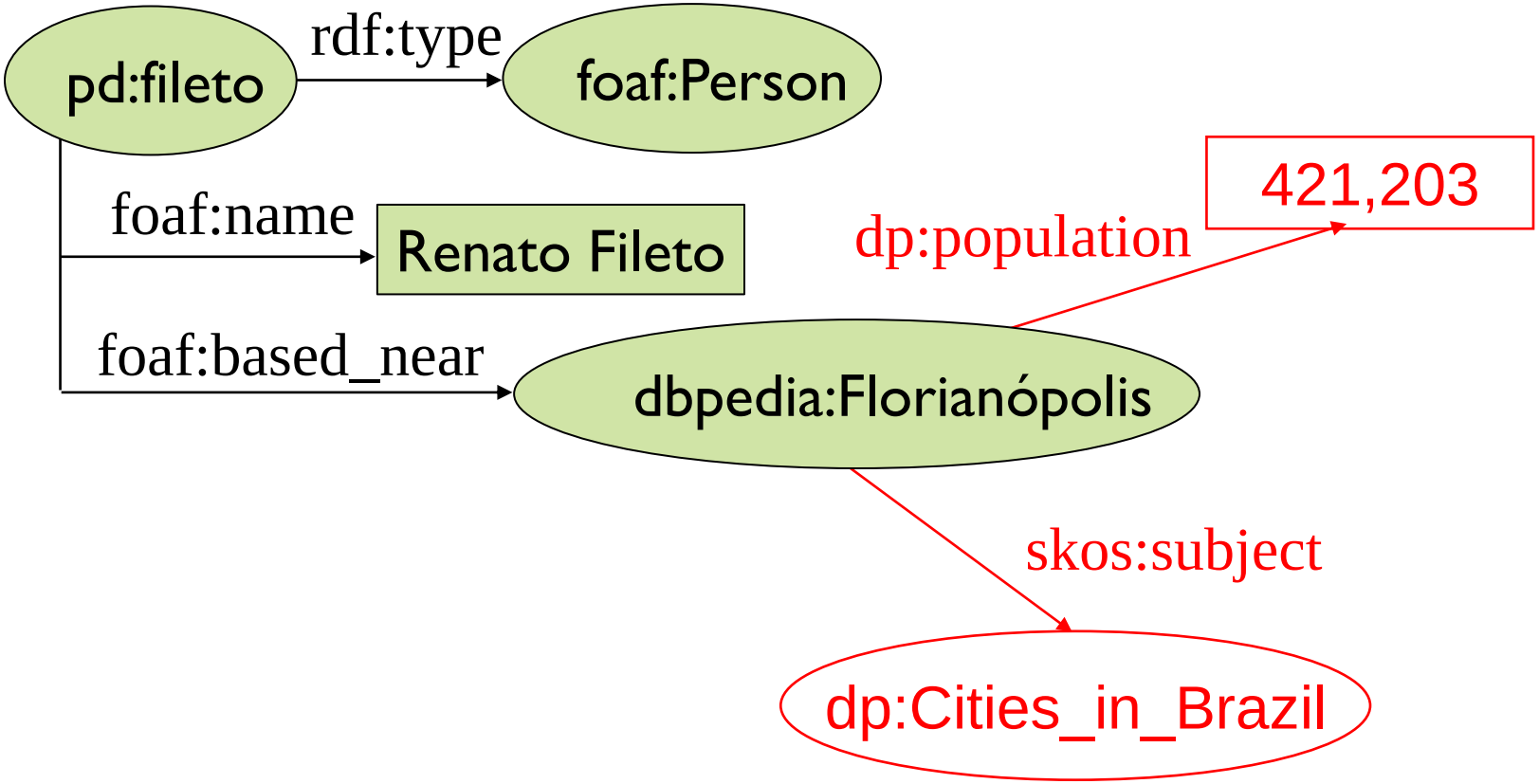
- collections of data from different sources;
- represented in accordance to standards (RDF/RDFS + specific vocabularies/schemas) and guidelines;
- enable interlinking, querying, and reuse.



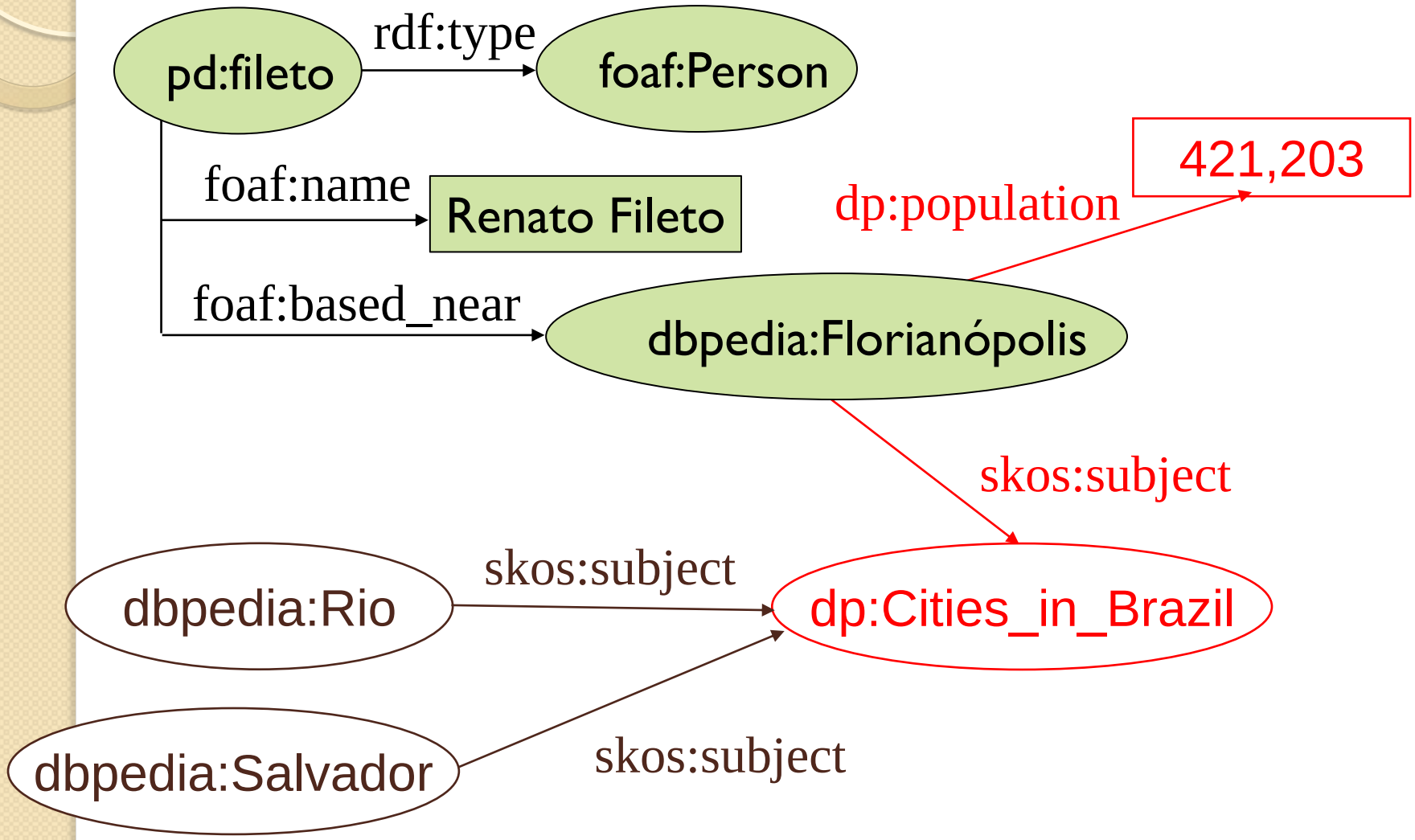
The RDF Data Model

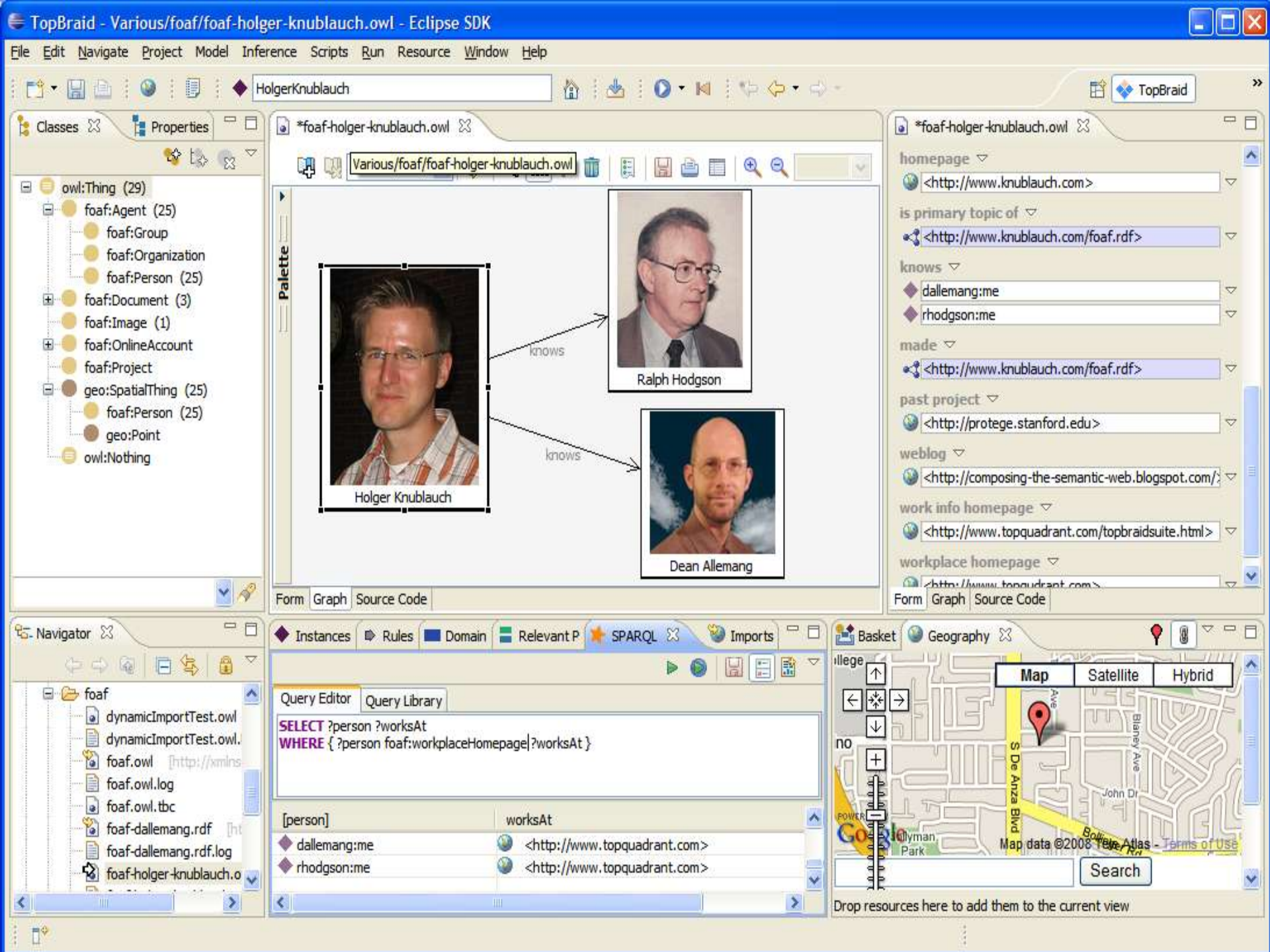


Dereferencing URIs over the Web



Dereferencing URIs over the Web




TopBraid - Various/foaf/foaf-holger-knublauch.owl - Eclipse SDK
File Edit Navigate Project Model Inference Scripts Run Resource Window Help
HolgerKnublauch
TopBraid
Classes
Properties
owl:Thing (29)
foaf:Agent (25)
foaf:Group
foaf:Organization
foaf:Person (25)
foaf:Document (3)
foaf:Image (1)
foaf:OnlineAccount
foaf:Project
geo:SpatialThing (25)
foaf:Person (25)
geo:Point
owl:Nothing
*foaf-holger-knublauch.owl
Various/foaf/foaf-holger-knublauch.owl
Palette
knows
knows
Ralph Hodgson
Holger Knublauch
Dean Allemang
Form Graph Source Code
*foaf-holger-knublauch.owl
homepage
<http://www.knublauch.com>
is primary topic of
<http://www.knublauch.com/foaf.rdf>
knows
dallemang:me
rhodgson:me
made
<http://www.knublauch.com/foaf.rdf>
past project
<http://protege.stanford.edu>
weblog
<http://composing-the-semantic-web.blogspot.com/>
work info homepage
<http://www.topquadrant.com/topbraidsuite.html>
workplace homepage
Form Graph Source Code
Navigator
foaf
dynamicImportTest.owl
dynamicImportTest.owl.
foaf.owl
foaf.owl.log
foaf.owl.tbc
foaf-dallemang.rdf
foaf-dallemang.rdf.log
foaf-holger-knublauch.o
Instances
Rules
Domain
Relevant P
SPARQL
Imports
Query Editor
Query Library
SELECT ?person ?worksAt
WHERE { ?person foaf:workplaceHomepage ?worksAt }
[person]
worksAt
dallemang:me
<http://www.topquadrant.com>
rhodgson:me
<http://www.topquadrant.com>
Basket
Geography
Map
Satellite
Hybrid
Search
Map data ©2008
Terms of Use
Drop resources here to add them to the current view

Guidelines to publish LOD (Berners-Lee)

- Use URIs as identifiers for things
- Use HTTP URIs so that people can look up those names.
- When someone looks up a URI, provide useful information.
- Include links to other URIs, so that more information can be discovered.

Example RDF Links

- RDF links from DBpedia to other data sources

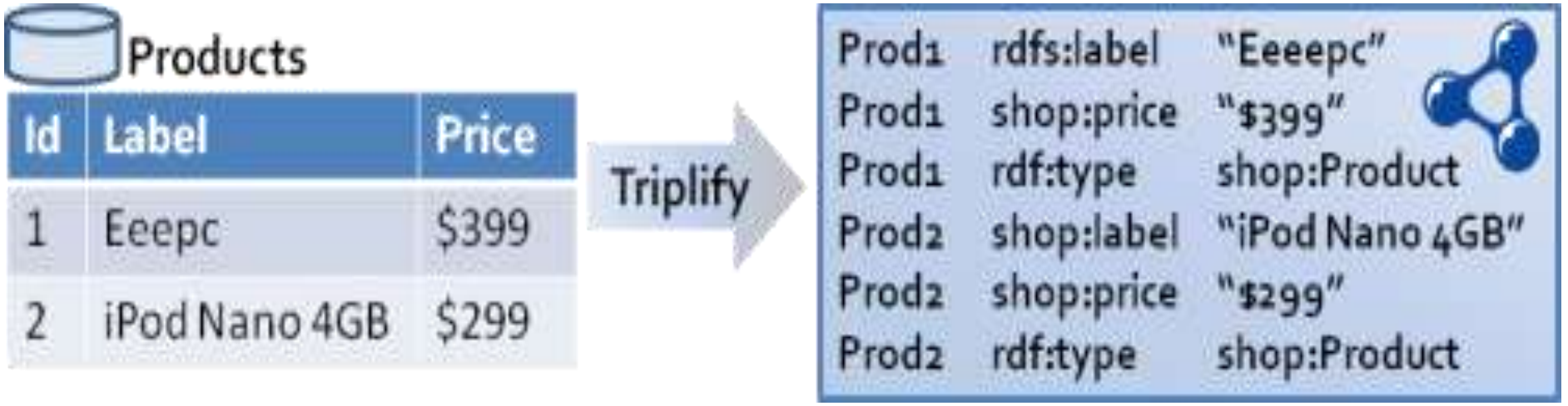
<http://dbpedia.org/resource/Berlin> **owl:sameAs** <http://sws.geonames.org/2950159> .

<http://dbpedia.org/resource/Tim_Berners-Lee> **owl:sameAs** <http://www4.wiwiss.fu-berlin.de/dblp/resource/person/100007> .

- RDF link from a FOAF profile to DBpedia

<http://richard.cyganiak.de/foaf.rdf#cygri> **foaf:topic_interest** <http://dbpedia.org/resource/Semantic_Web> .

Publishing linked data with Triplify



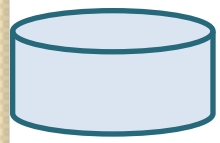
RDB2RDF tool comparison

Tool	Triplify	R2DQ	Virtuoso RDF Views
Technology	Scripting languages (PHP)	Java	Whole middleware solution
SPARQL endpoint	-	X	X
Mapping language	SQL	RDF based	RDF based
Mapping generation	Manual	Semi-automatic	Manual
Scalability	Medium-high (but no SPARQL)	medium	High

More at: http://esw.w3.org/topic/Rdb2RdfXG/StateOfTheArt

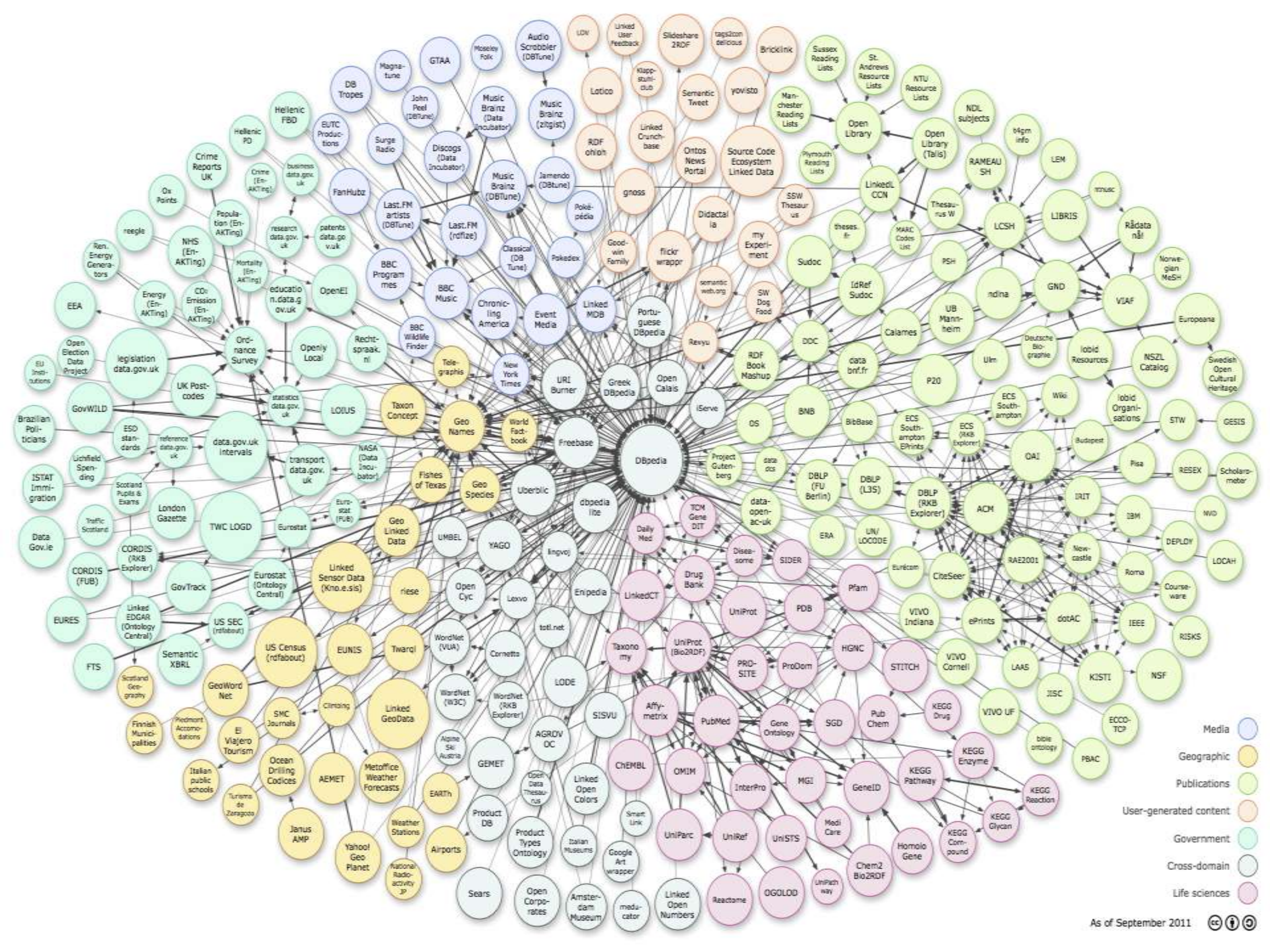
Marrying DBs with RDF & Ontologies

	Relational Databases	RDF & Ontologies
Data Model	**Relational** (tables, columns, rows)	**Triples** (subject, predicate, object)
Schema and data separation	☑	☐
Implicit information	☐	☑
Scalability	☑	☐
Schema flexibility	☐	☑
Web data integration readiness	☐	☑



Using DBs for **storage and querying** of RDF & ontologies

Publishing DB content as RDF


DBpedia
Freebase
URI Burner
Greek DBpedia
Open Calais
iServe
Uberblic
dbpedia lite
YAGO
UMBEL
lingvoj
Open Cyc
Lexvo
Enipedia
totl.net
WordNet (VUA)
Cornetto
WordNet (W3C)
WordNet (RKB Explorer)
LODE
SISVU
AGROV OC
GEMET
Alpine Ski Austria
Open Data Thesaurus
Linked Open Colors
Product DB
Product Types Ontology
Italian Museums
Google Art wrapper
Smart Link
Sears
Open Corporates
Amsterdam Museum
meducator
Linked Open Numbers
Portuguese DBpedia
Geo Names
World Factbook
Taxon Concept
Telegraphis
Fishes of Texas
Geo Species
Geo Linked Data
Linked Sensor Data (Kno.e.sis)
riese
US Census (rdfabout)
EUNIS
Twarql
Linked GeoData
Climbing
SMC Journals
GeoWord Net
El Viajero Tourism
Ocean Drilling Codices
AEMET
Metoffice Weather Forecasts
EARTh
Weather Stations
Airports
Janus AMP
Yahoo! Geo Planet
National Radioactivity JP
Scotland Geography
Finnish Municipalities
Piedmont Accomodations
Italian public schools
Turismo de Zaragoza
Audio Scrobbler (DBTune)
Music Brainz (zitgist)
Music Brainz (Data Incubator)
Music Brainz (DBTune)
Moseley Folk
GTAA
Magnatune
DB Tropes
John Peel (DBTune)
Surge Radio
EUTC Productions
Discogs (Data Incubator)
FanHubz
Last.FM artists (DBTune)
Last.FM (rdfize)
Jamendo (DBtune)
Pokédex
Pokepédia
Classical (DB Tune)
BBC Programmes
BBC Music
BBC Wildlife Finder
Chronicling America
Event Media
Linked MDB
New York Times
LOV
Linked User Feedback
Slideshare 2RDF
tags2con delicious
Bricklink
Klappstuhlclub
Lotico
Semantic Tweet
yovisto
RDF ohloh
Linked Crunchbase
Ontos News Portal
Source Code Ecosystem Linked Data
gnoss
SSW Thesaurus
Didactalia
my Experiment
flickr wrappr
Good-win Family
semantic web.org
SW Dog Food
Revyu
Sussex Reading Lists
St. Andrews Resource Lists
NTU Resource Lists
Manchester Reading Lists
Open Library
Open Library (Talis)
Plymouth Reading Lists
NDL subjects
t4gm info
LEM
RAMEAU SH
LinkedL CCN
Thesaurus W
MARC Codes List
LCSH
LIBRIS
ntnusc
Rådata nå!
theses.fr
Sudoc
PSH
IdRef Sudoc
ndlna
GND
VIAF
Norwegian MeSH
UB Mannheim
Calames
DDC
Deutsche Biographie
lobid Resources
NSZL Catalog
Europeana
Swedish Open Cultural Heritage
RDF Book Mashup
data bnf.fr
P20
Ulm
lobid Organisations
ECS Southampton
Wiki
OS
BNB
BibBase
ECS Southampton EPrints
ECS (RKB Explorer)
STW
GESIS
Budapest
OAI
Pisa
RESEX
Scholarometer
Project Gutenberg
data dcs
DBLP (FU Berlin)
DBLP (L3S)
DBLP (RKB Explorer)
ACM
IRIT
IBM
NVD
data-open-ac-uk
ERA
UN/LOCODE
DEPLOY
LOCAH
Eurécom
CiteSeer
RAE2001
Newcastle
Roma
Courseware
VIVO Indiana
ePrints
dotAC
IEEE
RISKS
VIVO Cornell
LAAS
KISTI
NSF
JISC
VIVO UF
ECCO-TCP
bible ontology
PBAC
Hellenic FBD
Hellenic PD
Crime Reports UK
Crime (EnAKTing)
business data.gov.uk
Ox Points
Population (EnAKTing)
research data.gov.uk
patents data.gov.uk
reegle
NHS (EnAKTing)
Mortality (EnAKTing)
Ren. Energy Generators
Energy (EnAKTing)
CO2 Emission (EnAKTing)
education.data.gov.uk
OpenEI
EEA
Ordnance Survey
Openly Local
Rechtspraak.nl
Open Election Data Project
EU Institutions
legislation data.gov.uk
UK Postcodes
statistics data.gov.uk
LOIUS
GovWILD
Brazilian Politicians
ESD standards
reference data.gov.uk
data.gov.uk intervals
transport data.gov.uk
NASA (Data Incubator)
Lichfield Spending
ISTAT Immigration
Scotland Pupils & Exams
London Gazette
TWC LOGD
Eurostat
Eurostat (FUB)
Traffic Scotland
Data Gov.ie
CORDIS (RKB Explorer)
CORDIS (FUB)
GovTrack
Eurostat (Ontology Central)
EURES
Linked EDGAR (Ontology Central)
US SEC (rdfabout)
FTS
Semantic XBRL
Daily Med
TCM Gene DIT
Diseasome
SIDER
Drug Bank
LinkedCT
Pfam
UniProt
PDB
Taxonomy
UniProt (Bio2RDF)
HGNC
PRO-SITE
ProDom
STITCH
Affymetrix
PubMed
Gene Ontology
SGD
Pub Chem
KEGG Drug
KEGG Enzyme
ChEMBL
OMIM
InterPro
MGI
GeneID
KEGG Pathway
KEGG Reaction
KEGG Glycan
KEGG Compound
Medi Care
UniParc
UniRef
UniSTS
Homolo Gene
Chem2 Bio2RDF
Reactome
OGOLOD
UniPathway
Media
Geographic
Publications
User-generated content
Government
Cross-domain
Life sciences
As of September 2011

DBPedia (OWL onto and several datasets)

- 97 languages
- dataset describes > 3.6 million things
- ~ 2 million things classified in an ontology
  - 416,000 persons, 526,000 places, 106,000 music albums, 60,000 films, 17,500 video games, 169,000 organizations, 183,000 species, …
- Several properties between things, usually taken from Info Boxes, but some not standardized yet

DBpedia SPARQL Endpoint (2)

```
SELECT ?name ?birth ?description ?person WHERE {
    ?person  dbp:birthPlace  dbp:Berlin .
    ?person  skos:subject    dbp:Cat:German_musicians .
    ?person  dbp:birth       ?birth .
    ?person  foaf:name       ?name .
    ?person  rdfs:comment    ?description .
  FILTER (LANG(?description) = 'en') .
} ORDER BY ?name
```

name	birth	description	person
"Moser, Edda"@de	"1938-10-27"^^xsd:date	"The German soprano Edda Moser was born on October 27, 1938 in Berlin, Germany. She is the daughter of the musicologist Hans Joachim Moser."@en	:Edda_Moser
"Möbius, Ralph Christian""@de	"1950-01-09"^^xsd:date	"Rio Reiser (January 9, 1950 - August 20, 1996), was a German rock musician and singer of the famous rock group Ton Steine Scherben. He was born Ralph Christian Möbius in Berlin and died at the age of 46 in the little German town of Fresenhagen. Rio Reiser was politically active during his whole life. In the early 70ies he participated in the squatter scene, for which he wrote the famous Rauchhaussong.""@en	:Rio_Reiser
"Straube, Karl"@de	"1873-01-06"^^xsd:date	"Montgomery Rufus Karl/Carl Siegfried Straube (January 6, 1873, Berlin - April 27, 1950, Leipzig) was a German church musician , organist, and choral conductor, famous above all for championing the abundant organ music of Max Reger."@en	:Karl_Straube
"Tricht, Käte van""@de	"1909-10-22"^^xsd:date	"Käte van Tricht (October 22, 1909–July 13, 1996), was a German organist, pianist, harpsichordist, and pedagogue.""@en	:K%C3%A4te_van_Tricht
"Urlaub, Farin"@de	"1963-10-27"^^xsd:date	"Jan Ulrich Max Vetter, better known as Farin Urlaub (like German "Fahr in Urlaub!" ("Go on holiday!"), after his love of travelling) was born on October 27, 1963 in what was then West Berlin, Germany. He is best known as the guitarist/vocalist for the German punk rock band Die Ärzte.""@en	:Farin_Urlaub
"Voormann, Klaus"@de	"1938-04-29"^^xsd:date	"Klaus Voormann (born 29 April 1938) is a German artist, musician, and record producer who was associated with the early days of The Beatles in Hamburg and later designed the cover of their album Revolver."@en	:Klaus_Voormann

GeoCodes (XML, JSON, RDF, TXT, RSS, CSV, KML)

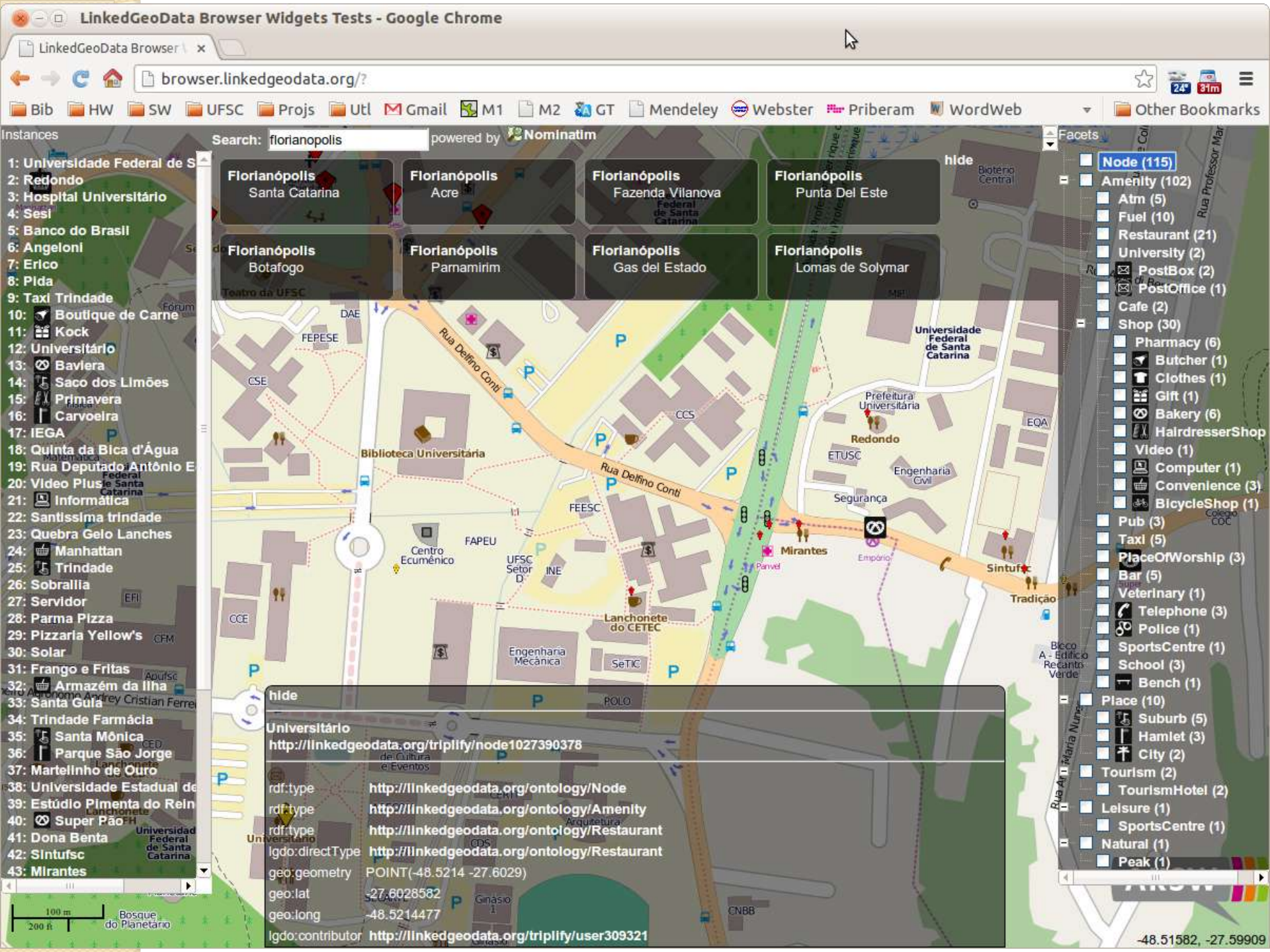
- ~10 million geographical names
- ~ 8 million unique features
- ~ 2.8 million are populated places
- ~ 5.5 million alternate names.
- All features are categorized into one out of nine feature classes
  - Administrative Boundary, Hydrographic, Area, Populated Place, Road / Railroad, Spot, Hypsographic, Undersea, Vegetation
- and further subcategorized into one out of 645 feature codes

LinkedGeoData (nt format)

Information collected from OpenStreetMap.

Up to 2013-Apr-29 :

Dataset	#Triples	Download Size	Uncompressed
Ontology	8K	50KB	1MB
RelevantNodes	66Mio	323MB	10GB
RelevantWays	65Mio	661MB	10GB
RelevantWayNodes	74Mio	329MB	11GB
RelevantNodePositions	60Mio	414MB	11GB
DBpedia Interlinks	101K	917KB	14MB
GeoNames Interlinks	487K	3MB	60MB


LinkedGeoData Browser Widgets Tests - Google Chrome
LinkedGeoData Browser
browser.linkedgeodata.org/?
Bib HW SW UFSC Projs Utl Gmail M1 M2 GT Mendeley Webster Priberam WordWeb Other Bookmarks
Instances
1: Universidade Federal de S
2: Redondo
3: Hospital Universitário
4: Sesi
5: Banco do Brasil
6: Angeloni
7: Erico
8: Pida
9: Taxi Trindade
10: Boutique de Carne
11: Kock
12: Universitário
13: Baviera
14: Saco dos Limões
15: Primavera
16: Carvoeira
17: IEGA
18: Quinta da Bica d'Água
19: Rua Deputado Antônio E
20: Video Plus
21: Informática
22: Santissima trindade
23: Quebra Gelo Lanches
24: Manhattan
25: Trindade
26: Sobrallia
27: Servidor
28: Parma Pizza
29: Pizzaria Yellow's
30: Solar
31: Frango e Fritas
32: Armazém da Ilha
33: Santa Gula
34: Trindade Farmácia
35: Santa Mônica
36: Parque São Jorge
37: Martelinho de Ouro
38: Universidade Estadual de
39: Estúdio Pimenta do Rein
40: Super Pão
41: Dona Benta
42: Sintufsc
43: Mirantes
Search: florianopolis
powered by Nominatim
Florianópolis Santa Catarina
Florianópolis Acre
Florianópolis Fazenda Vilanova
Florianópolis Punta Del Este
Florianópolis Botafogo
Florianópolis Parnamirim
Florianópolis Gas del Estado
Florianópolis Lomas de Solymar
hide
Facets
Node (115)
Amenity (102)
Atm (5)
Fuel (10)
Restaurant (21)
University (2)
PostBox (2)
PostOffice (1)
Cafe (2)
Shop (30)
Pharmacy (6)
Butcher (1)
Clothes (1)
Gift (1)
Bakery (6)
HairdresserShop
Video (1)
Computer (1)
Convenience (3)
BicycleShop (1)
Pub (3)
Taxi (5)
PlaceOfWorship (3)
Bar (5)
Veterinary (1)
Telephone (3)
Police (1)
SportsCentre (1)
School (3)
Bench (1)
Place (10)
Suburb (5)
Hamlet (3)
City (2)
Tourism (2)
TourismHotel (2)
Leisure (1)
SportsCentre (1)
Natural (1)
Peak (1)
Rua Delfino Conti
Biblioteca Universitária
Universidade Federal de Santa Catarina
Prefeitura Universitária
Redondo
Mirantes
Sintufsc
Tradição
Lanchonete do CETEC
Centro Ecumênico
FAPEU
UFSC Setor D
INE
FEESC
CCS
ETUSC
Engenharia Civil
Segurança
Engenharia Mecânica
SeTIC
CCE
FEPESE
DAE
CSE
EQA
Teatro da UFSC
Bosque do Planetário
hide
Universitário
http://linkedgeodata.org/triplify/node1027390378
rdf:type http://linkedgeodata.org/ontology/Node
rdf:type http://linkedgeodata.org/ontology/Amenity
rdf:type http://linkedgeodata.org/ontology/Restaurant
lgdo:directType http://linkedgeodata.org/ontology/Restaurant
geo:geometry POINT(-48.5214 -27.6029)
geo:lat -27.6028582
geo:long -48.5214477
lgdo:contributor http://linkedgeodata.org/triplify/user309321
-48.51582, -27.59909

Linked Data Mashups

- Domain-specific applications using LOD

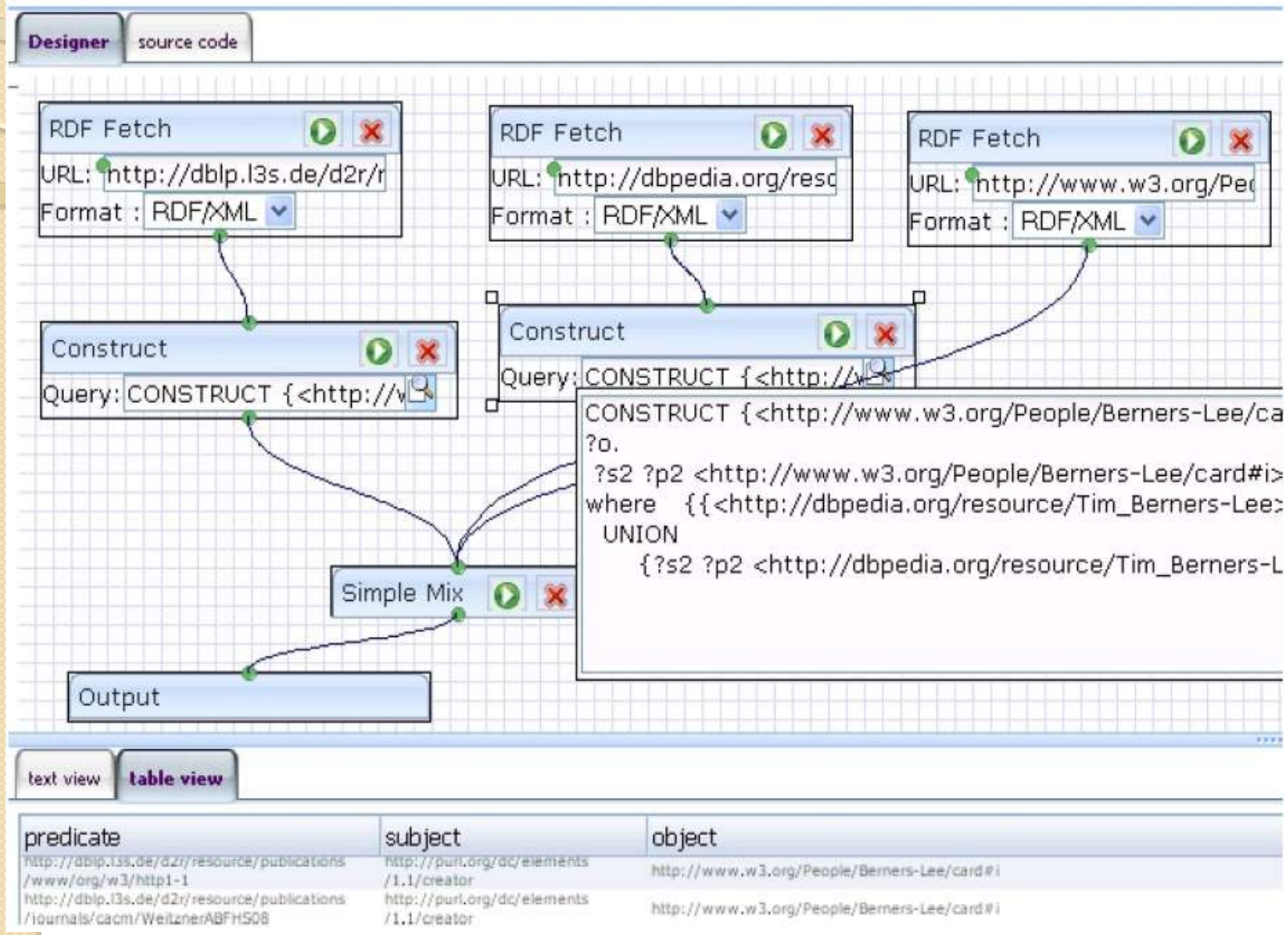
DBtune Slashfacet

- Visualizes music-related Linked Data
- Uses LastFM, MySpace, and BBC data

DBpedia Mobile

- Geospatial entry point into the Web of Data
- Starts with data from
  - DBpedia,
  - Revyu, and
  - Flickr

DERI Semantic Web Pipes



Topics

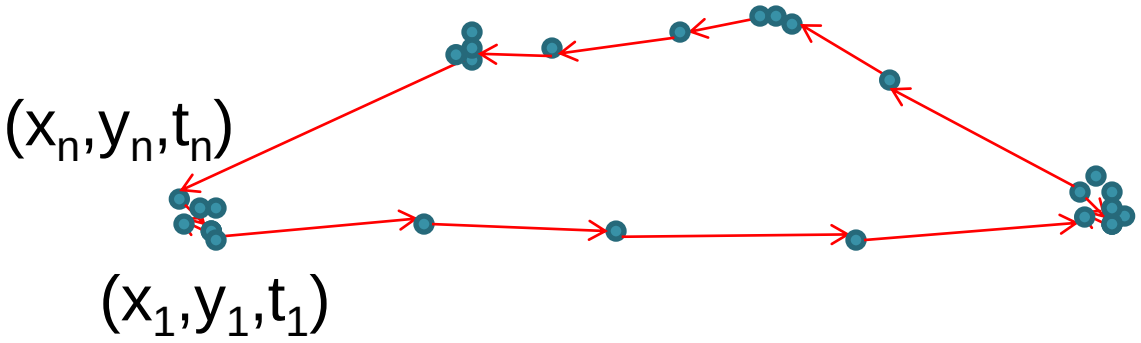
1. Introduction
   - Motivation
2. Linked Open Data (LOD)
   - Ontologies
   - Semantic Web Standards
   - The Web of Data
   - LOD Collections and Tools
3. **Semantic Enrichment of Movement Data**
   - **Moving Object Trajectories**
   - **The Baquara Ontology**
   - **Trajectories *X* Social Media Trails**
   - **Fusion of Trajectories with Social Media Trails (Ricardo)**
   - **Connecting Annotated Movement Data with LOD (Cleto)**
4. Conclusions and Future Work

Raw Trajectory

- Time ordered sequence of spatial-temporal positions of a moving object

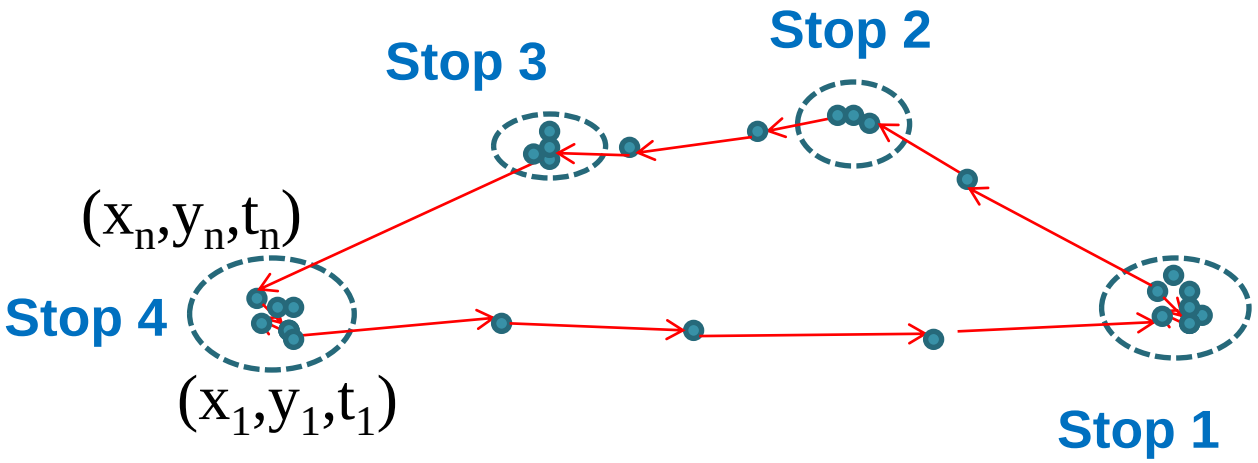
$$(x_1,y_1,t_1)\ (x_2,y_2,t_2)\ \ldots\ (x_n,y_n,t_n)$$

- Sampled by using sensors (GPS, GSM, RFID, câmeras, etc.)



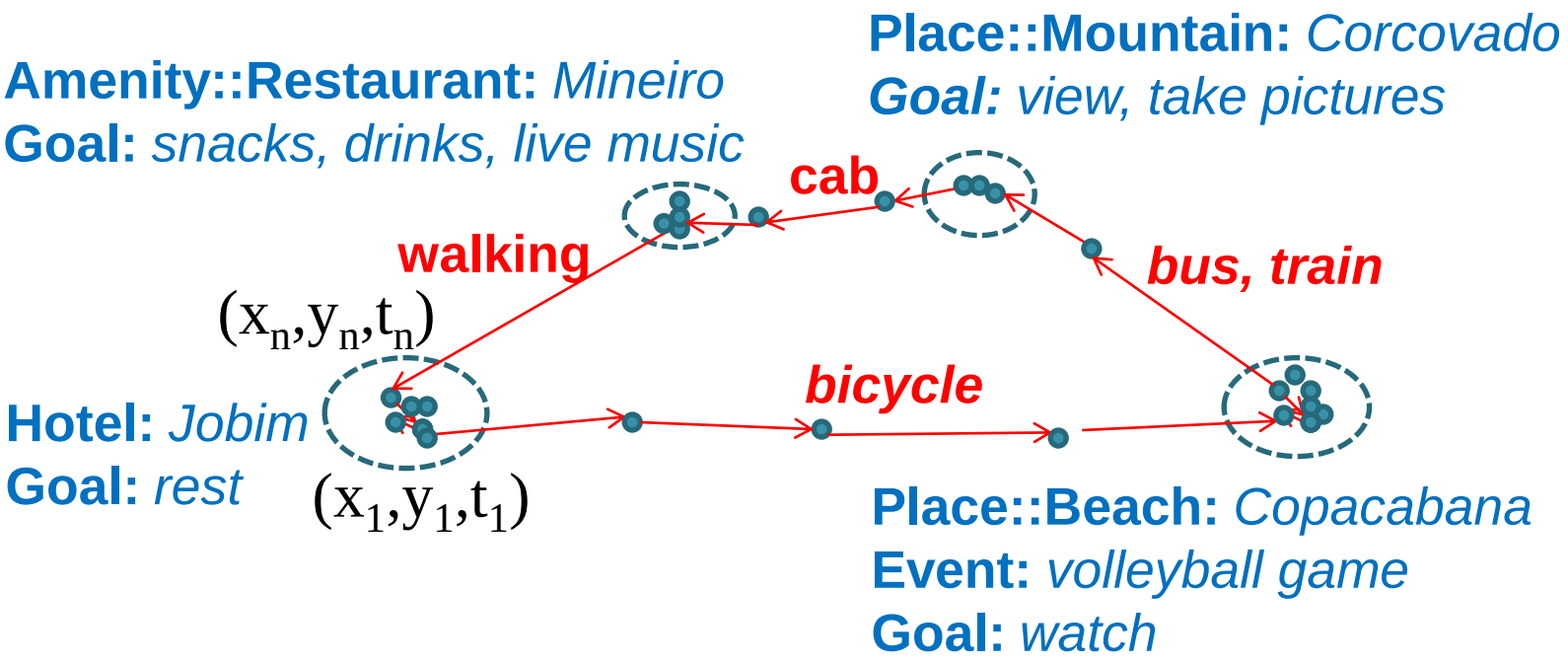
Structured Trajectory

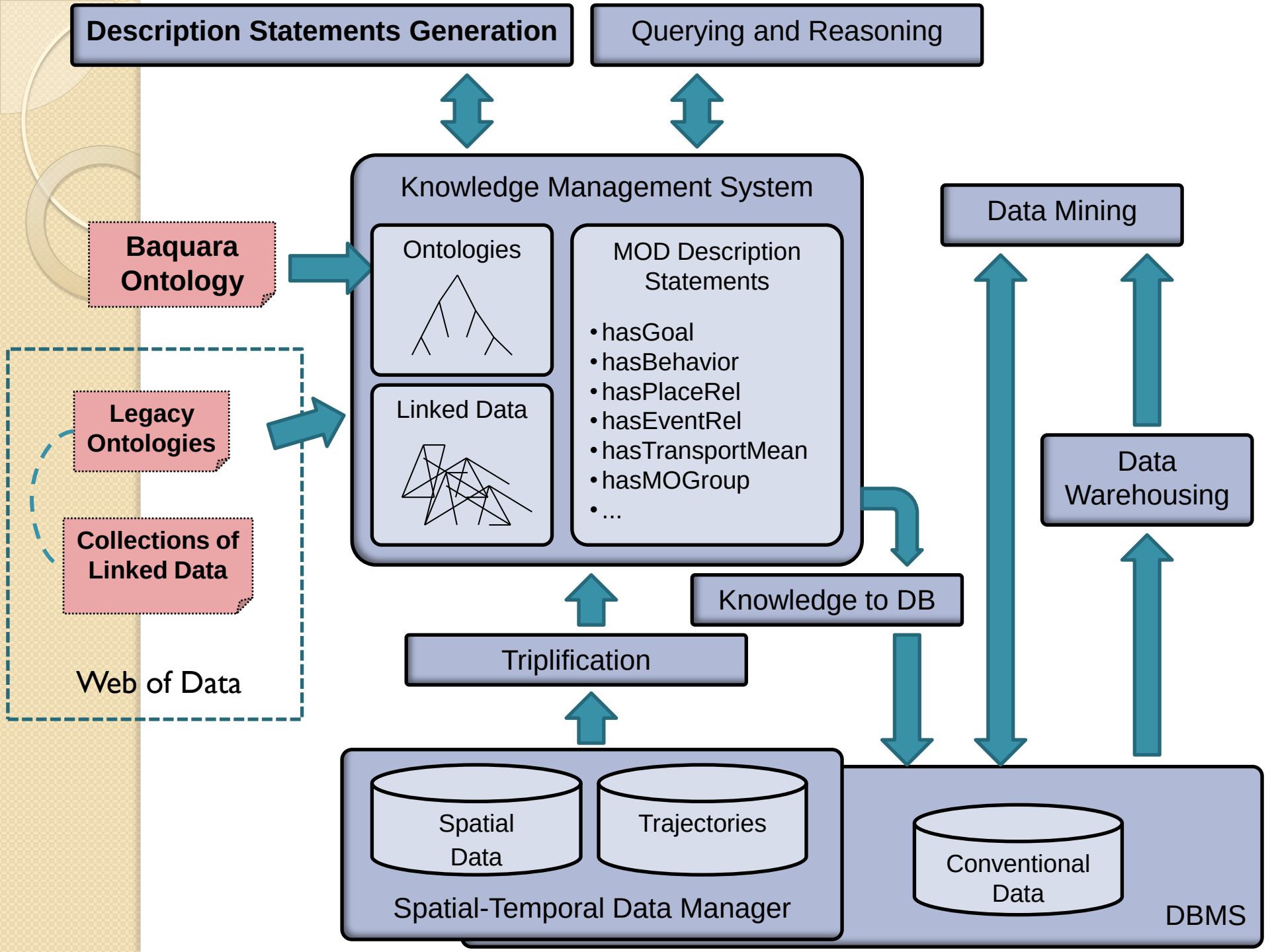
- Time ordered sequence of **episodes**.
- An **episode** (e.g., stop, move) is a maximal continuous segment of a trajectory in which a predicate holds. For example:
  - the MO speed is below a threshold;
  - the MO is at a place for a certain time.

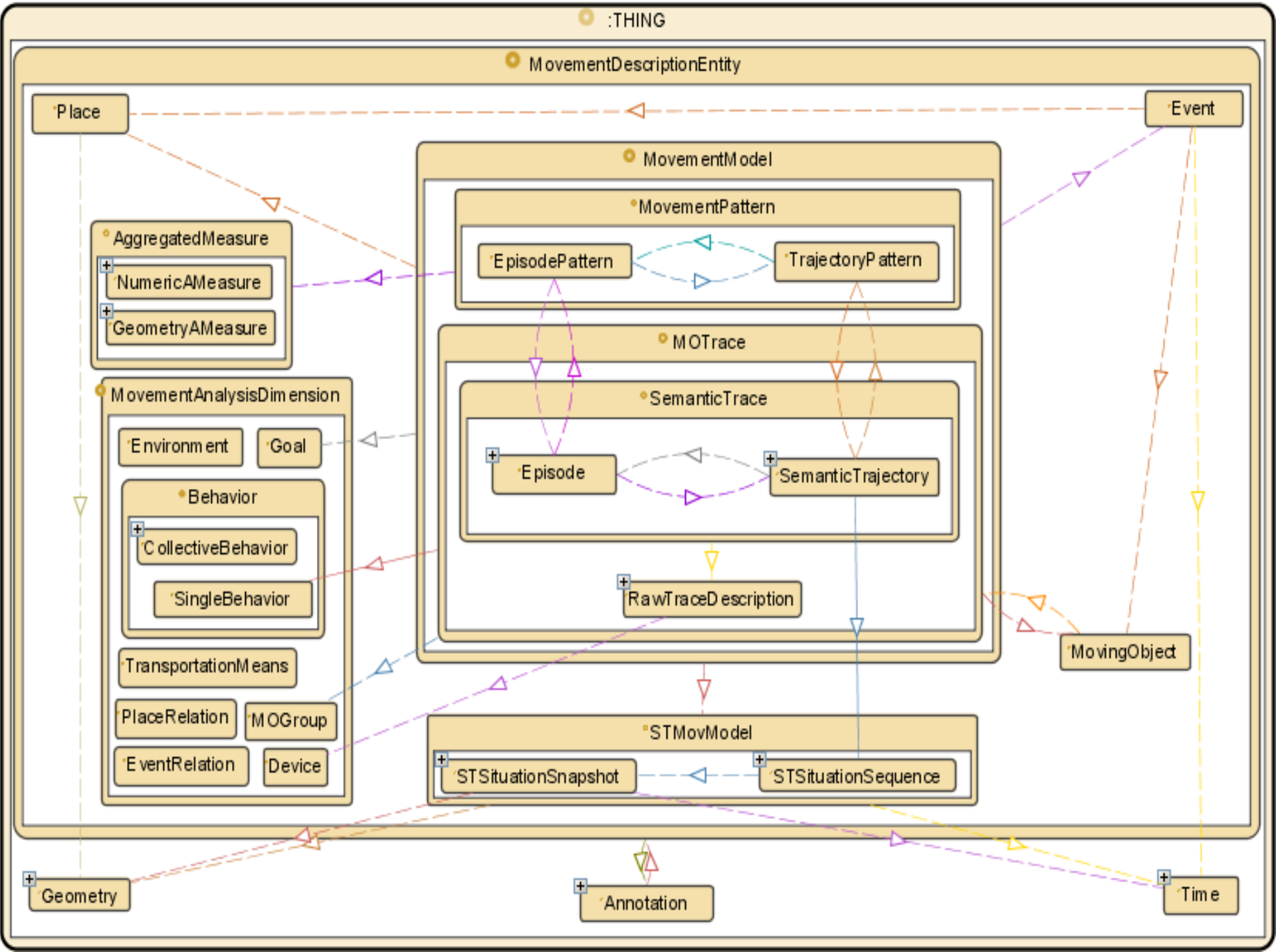


Semantic Trajectory

- Structured trajectory whose components episodes are associated with information that has well-defined semantics.




Description Statements Generation
Querying and Reasoning
Knowledge Management System
Ontologies
Linked Data
MOD Description Statements
• hasGoal
• hasBehavior
• hasPlaceRel
• hasEventRel
• hasTransportMean
• hasMOGroup
• ...
Baquara Ontology
Legacy Ontologies
Collections of Linked Data
Web of Data
Data Mining
Data Warehousing
Knowledge to DB
Triplification
Spatial Data
Trajectories
Spatial-Temporal Data Manager
Conventional Data
DBMS


:THING
MovementDescriptionEntity
Place
Event
MovementModel
MovementPattern
EpisodePattern
TrajectoryPattern
AggregatedMeasure
NumericAMeasure
GeometryAMeasure
MOTrace
SemanticTrace
MovementAnalysisDimension
Environment
Goal
Episode
SemanticTrajectory
Behavior
CollectiveBehavior
SingleBehavior
RawTraceDescription
TransportationMeans
MovingObject
PlaceRelation
MOGroup
STMovModel
EventRelation
Device
STSituationSnapshot
STSituationSequence
Geometry
Annotation
Time

General Algorithm for Semantic Enrichment

INPUT: MovData(id,MOid,T,S,$A_1 \ldots A_n$) % Movement data

LD % Baquara ontology and collected linked data

OUTPUT: DS(id,property,value) % Description statements

1. DS = ∅
2. FOR EACH r IN MovData DO
3. Matches = **FindMatches(r,LD)**;
4. FOR EACH v IN Matches DO
5. CandidateProperties = **ChooseProperties(r,v)**;
6. FOR EACH p IN CandidateProperties DO
7. DS = DS + (r.id,p,v);
8. RETURN DS

Baquara enabled query 1

```
SELECT   ?trajectory
WHERE {
  ?trajectory   a   bq:SemanticTrajectory;
                bq:hasEpisode   ?episode.
  ?episode   bq:hasAnnotation   ?a;
             geo:geometry   ?eGeo.
  ?a   bq:hasValue   ?v.
  ?v   a   <http://dbpedia.org/ontology/Mountain>;
       rdfs:label   "Corcovado"@pt;
       geo:geometry   ?corGeo.
  ?rio   a   <http://dbpedia.org/ontology/City>;
         rdfs:label "Rio de Janeiro"@pt;
         geo:geometry ?rioGeo.
  FILTER(bif:st_intersects (?corGeo,?rioGeo,20) &&
         bif:st_intersects (?eGeo,?corGeo,50))
}
```

Baquara enabled query 2

```
SELECT ?trajectory
WHERE {
  ?trajectory   a   bq:SemanticTrajectory ;
         bq:hasEpisode   ?s1;
         bq:hasEpisode   ?s2;
         bq:hasEpisode   ?s3.
  ?s1   a   bq:Stop; bq:occursIn   ?p1.
  ?s2   a   bq:Stop; bq:hasAnnotation   ?a2.
  ?s3   a   bq:Stop; bq:hasAnnotation   ?a3.
  ?a2   bq:hasValue   ?v2.
  ?a3   bq:hasValue   ?v3.
  ?p1 a <http://linkedgeodata.org/ontology/Amenity>.
  ?v2 a <http://dbpedia.org/resource/SportsEvent>.
  FILTER(regex(?v3,"Beach"))
}
```

Experimental Setup - Flickr user's trails

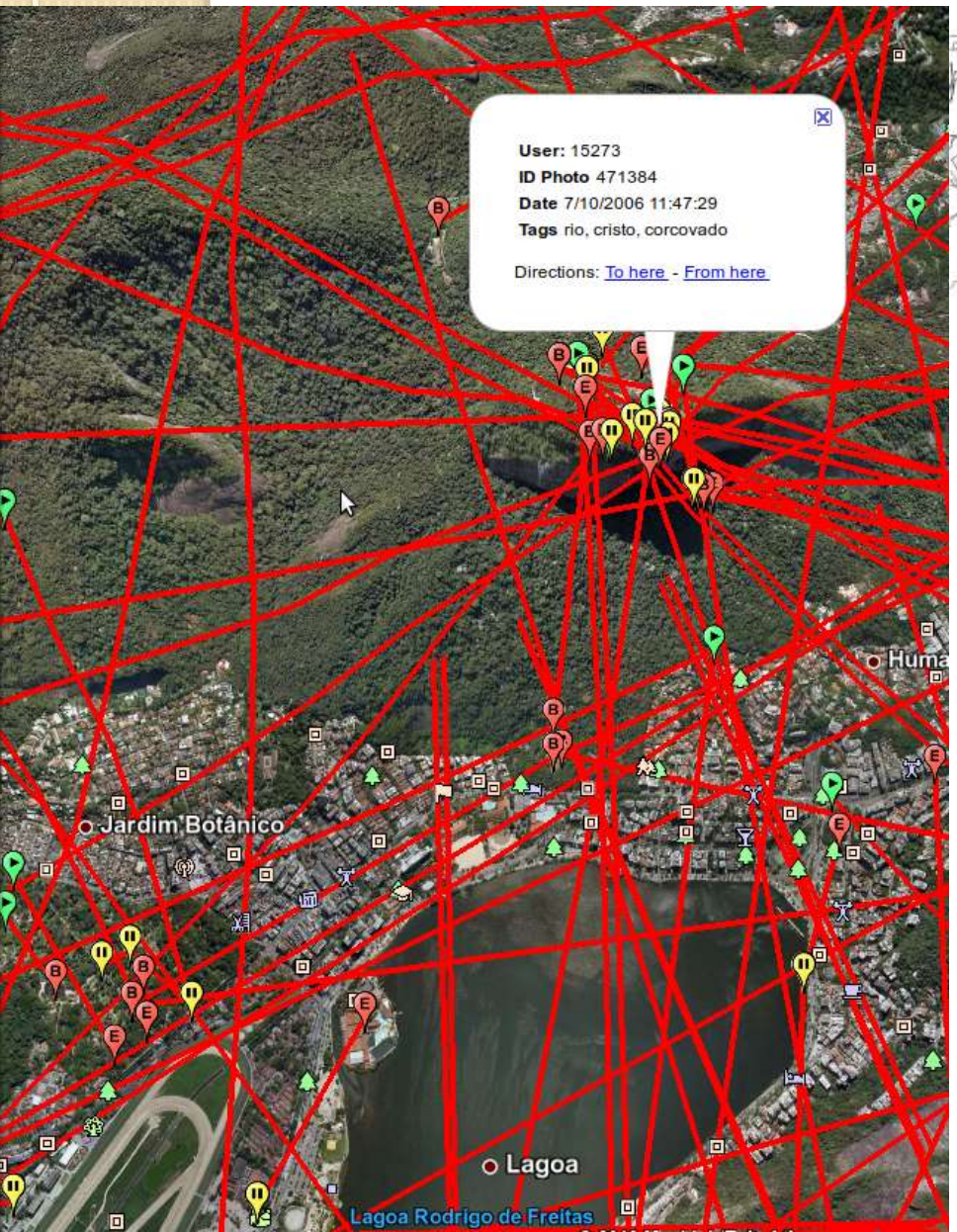
MovementData::

FlickT-BR(Uid,Instant,Lat,Long,Tags)

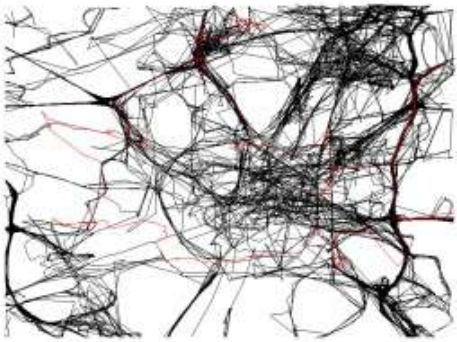
- After data filtering:
  - 2143 trajectories (each trajectory a day),
  - 564 distinct users,
  - 14504 sampled spatial-temporal positions,
  - 12443 distinct tags,
  - 117146 point-tag pairs.

Flickr Trails in Brazil (CoPhIR)

Social Media Trails *vs.* Trajectories



(a) The 60 known trajectories for Athens.



(b) The 90 known trajectories for Milan.

(c) Athens, Success-rate 0.60

(d) Milan, Success-rate 0.60

(e) Athens, Success-rate 0.85

(f) Milan, Success-rate 0.85

Social Media Trails (sequences of media footprints)	**Trajectories** (sequences of spatio-temporal points or episodes)
Sparse samples of positions	Usually **dense**, allowing detailed spatio-temporal representation
Asynchronous sampling (when the user decides to send/post something)	**Sampling intervals** with respect to **time** or **distance** travelled are common
Imprecise and frequently **uncertain** spatio-temporal data	**Precise and reliable**
Map matching is not possible for data from several sources (Flickr, Instagram, Twitter, Facebook)	**Map matching as a pre-processing step** for many datasets and applications
Textual data available	**Annotated trajectories are hard to obtain**

Proposal

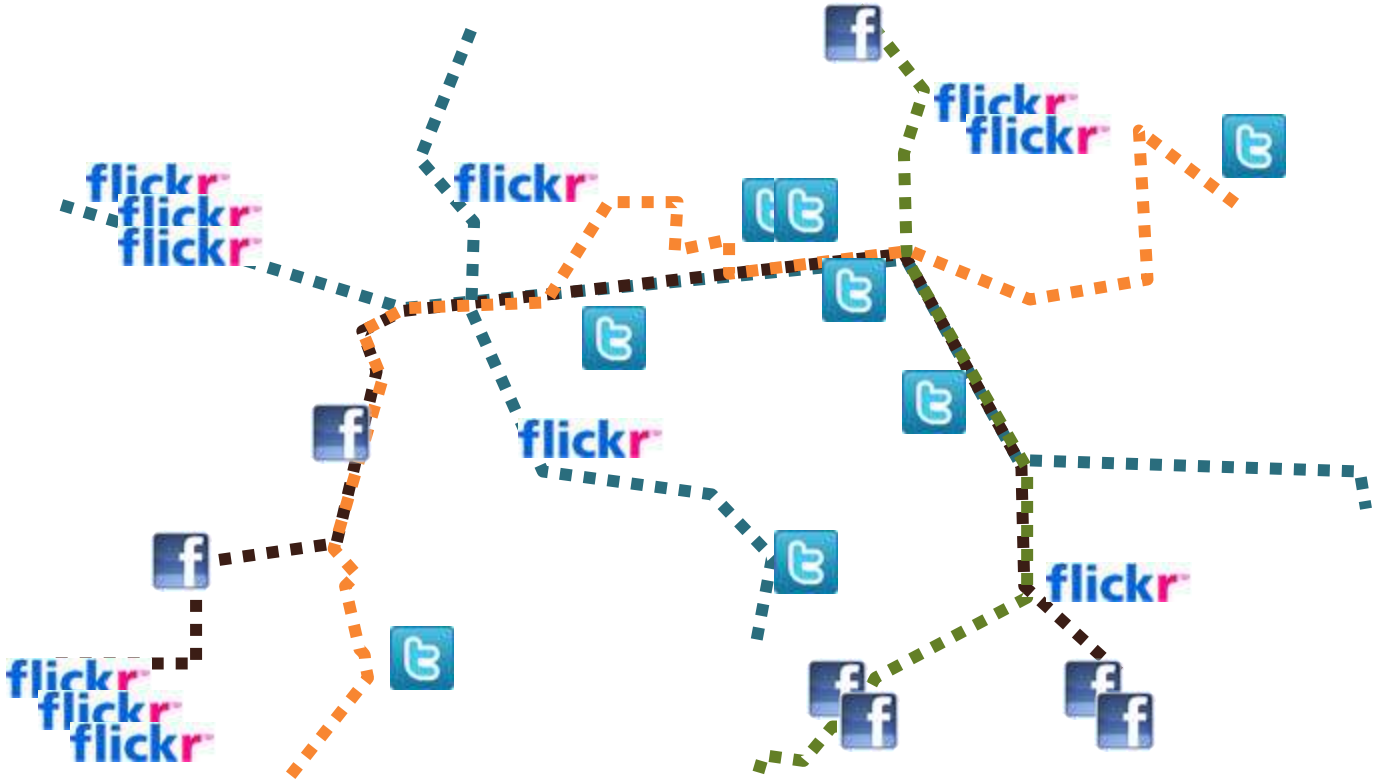
Semantic enrichment of moving object trajectories in 2 steps:

1. fuse trajectories with social media footprints (Ricardo)
2. connect the resulting annotated trajectories to LOD (Cleto)

Fusing Trajectories with SMTs

"Similarity-join" based on:

1. Spatio-temporal proximity/compatibility
2. Moving objects´ and users´ profiles and contexts

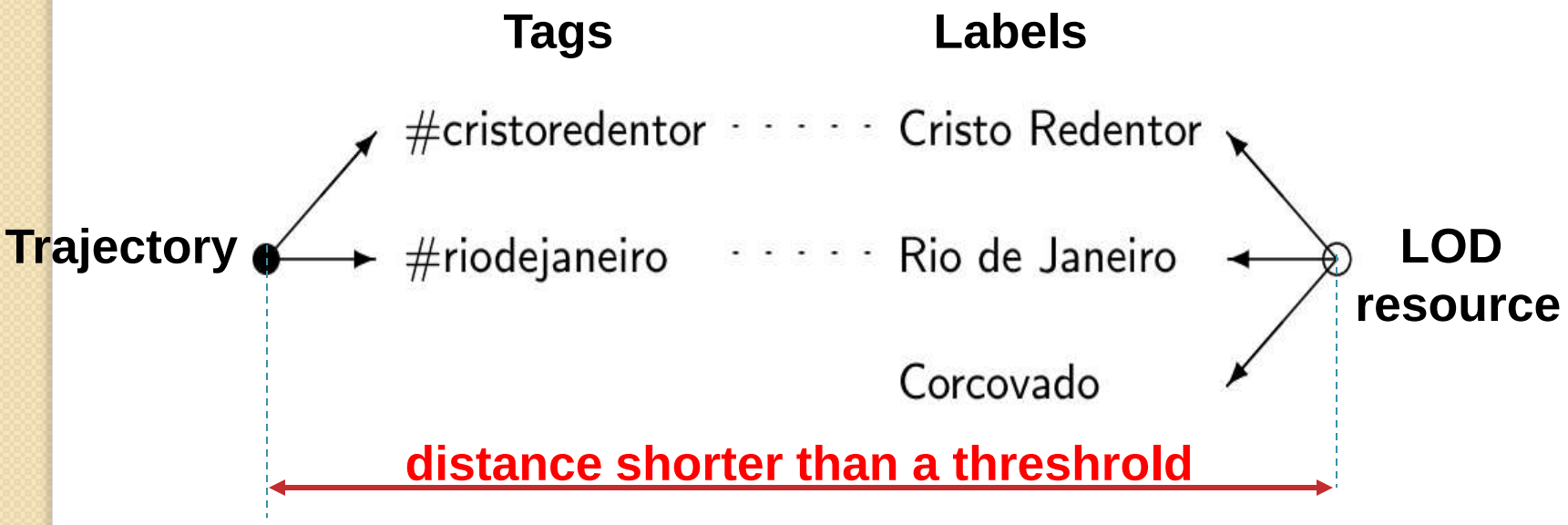


Connecting Annotated Trajectories to LOD

"Similarity-join" based on:

1. Spatio-temporal proximity/compatibility
2. Textual similarity (e.g., traj.e.tag = trail.fp.label)
3. Moving objects´ profiles and contexts

Example: connections SMTs to LOD



Topics

1. Introduction
   - Motivation
2. Linked Open Data (LOD)
   - Ontologies
   - Semantic Web Standards
   - The Web of Data
   - LOD Collections and Tools
3. Semantic Enrichment of Movement Data
   - Moving Object Trajectories
   - The Baquara Ontology
   - Trajectories *X* Social Media Trails
   - Fusion of Trajectories with Social Media Trails (Ricardo)
   - Connecting Annotated Movement Data with LOD (Cleto)
4. **Conclusions and Future Work**

Conclusions

- Semantic enrichment and analysis of trajectories require continuously updated information about the reality in which the trajectories occur.
- Linked data collections are rich sources of integrated information, with well-defined and widely agreed semantics.
- Major problems:
  1. Obtain annotated trajectories, by crowdsourcing or fusing trajectories with social media footprints
  2. Find connections between movement data and linked data to build precise description statements.

Future Work

1. Develop suitable strategies and efficient algorithms to solve sub-problems.
2. Evaluate the proposed approach with trajectories, social media trails and linked data from distinct sources and application domains.
3. Employ contextual information and profiles to better connect trajectories with linked data.
4. Investigate the use of linked data for trajectories warehousing and mining.

Questions?