Karlsruhe Institute of Technology

ontotext

SSSC 2011
Linked Data and Services

Barry Norton [1], Denny Vrandecic [2]
[1] Ontotext AD, BG/UK
[2] AIFB, Karlsruhe Institute of Technology, DE

Berkeley, US; August 8, 2011

KARLSRUHE SERVICE RESEARCH INSTITUTE (KSRI)
INSTITUTE OF APPLIED INFORMATICS AND FORMAL DESCRIPTION METHODS (AIFB)

OWLIM GROUP
ONTOTEXT AD

www.aifb.kit.edu

KIT – University of the State of Baden-Württemberg and
National Laboratory of the Helmholtz Association

Ontotext AD,
Sofia, Bulgaria

Outline

- **Background**
  - **Web Technologies**
- Linked Data
  - Linked Data Principles
  - Data Publishing
  - Data Consumption
- Linked Services
  - Linked Services Motivation
  - Linked (Open) Service Principles
  - Linked Service Implementation

HTTP Overview

HTTP, by which all documents on the WWW are served, is a client server protocol.
Every interaction based on:

Request

Method

GET
PUT
POST
PATCH
DELETE (+ OPTIONS, HEADER, TRACE, CONNECT)

URL

Header

[Optional] Body (with POST, PUT, PATCH)

- Reponse
  - Reponse code (integer)
  - Header
  - [Optional] Body

HTTP GET Example

method

Retrieval example:

```
GET /web/Barry_Norton HTTP/1.1
Host: www.aifb.kit.edu
Accept-Language:en
```

header

Can negotiate on
(human) language,
but also...

header

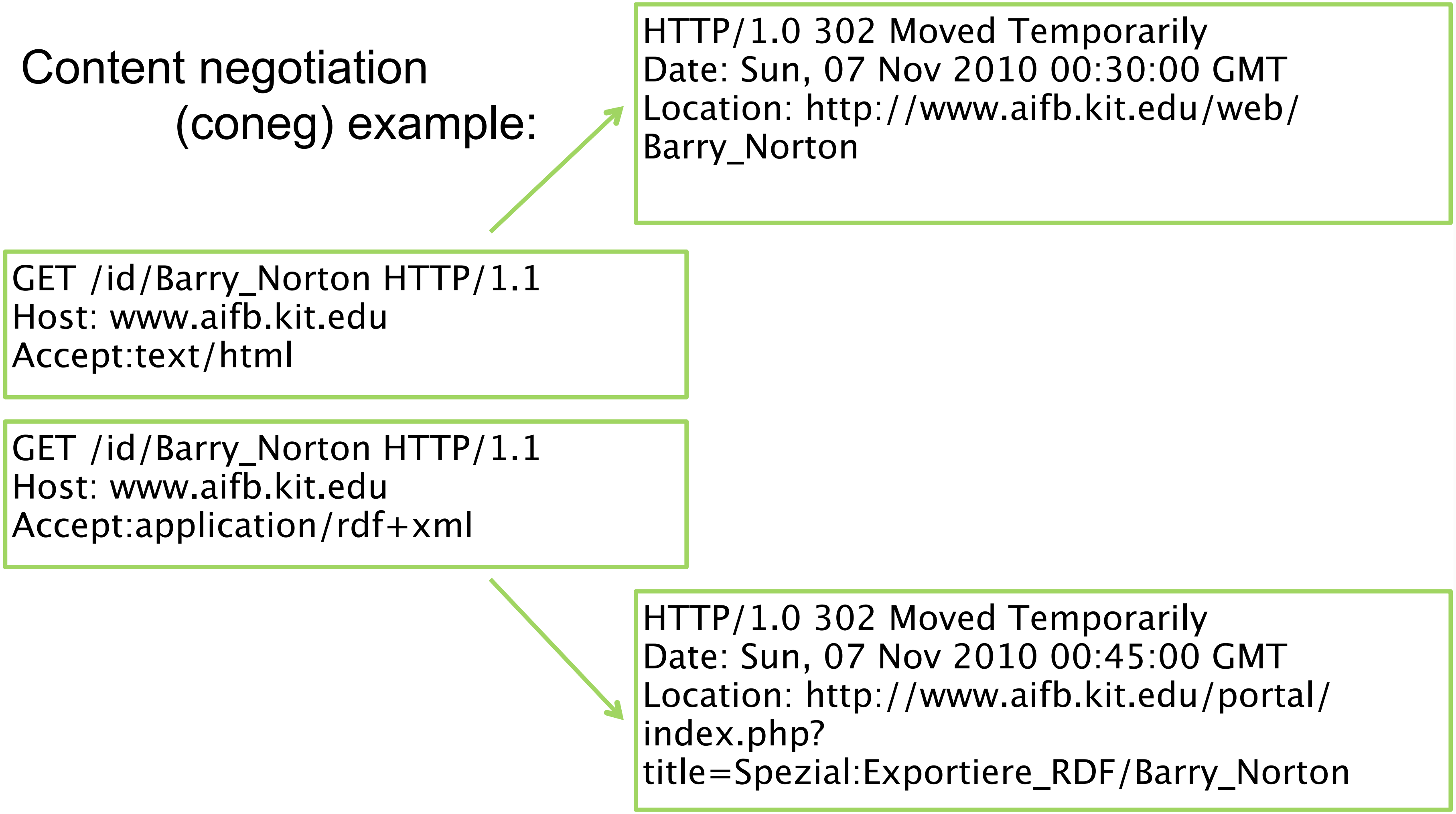
status code

```
HTTP/1.0 200 OK
Date: Sun, 07 Nov 2010 01:00:00 GMT
Content-Type: text/html

<html>
  <head> ...
```

body

HTTP Conneg Example

Content negotiation (coneg) example:



```
GET /id/Barry_Norton HTTP/1.1
Host: www.aifb.kit.edu
Accept:text/html
```

```
HTTP/1.0 302 Moved Temporarily
Date: Sun, 07 Nov 2010 00:30:00 GMT
Location: http://www.aifb.kit.edu/web/
Barry_Norton
```

```
GET /id/Barry_Norton HTTP/1.1
Host: www.aifb.kit.edu
Accept:application/rdf+xml
```

```
HTTP/1.0 302 Moved Temporarily
Date: Sun, 07 Nov 2010 00:45:00 GMT
Location: http://www.aifb.kit.edu/portal/
index.php?
title=Spezial:Exportiere_RDF/Barry_Norton
```

HTTP PUT/PATCH Examples

PUT/PATCH example:

```
PUT /web/Barry_Norton HTTP/1.1
Host: www.aifb.kit.edu
Content-Type: text/html

<html> ...
```

(new resource or complete update)

```
HTTP/1.0 200 OK (or 201 CREATED)
Date: Sun, 07 Nov 2010 00:10:00 GMT
```

```
PATCH /web/Barry_Norton HTTP/1.1
Host: www.aifb.kit.edu
Content-Type: text/html

Change...
```

(partial update)

```
HTTP/1.0 200 OK
Date: Sun, 07 Nov 2010 00:10:00 GMT
```

HTTP POST Examples

POST-compute example:

(input -> computation -> output)

```
POST /web/Barry_Norton HTTP/1.1
Host: www.aifb.kit.edu
Content-Type: something

Input ...
```

```
HTTP/1.0 200 OK
Date: Sun, 07 Nov 2010 00:10:00 GMT
Content-Type: something

Result....
```

POST-append example:

(new related resource)

```
POST /web/Barry_Norton HTTP/1.1
Host: www.aifb.kit.edu
Content-Type: text/html

<html> ...
```

```
HTTP/1.0 201 CREATED
Date: Sun, 07 Nov 2010 00:10:00 GMT
Location:http://www.aifb.kit.edu/...
```

Representational State Transfer

- HTTP, maintained by IETF not W3C, is just one (primary) implementation of an architectural style called:
    - REST = REpresentational State Transfer

REST Principles

1. Application state and functionality is divided into resources
2. Every resource is uniquely addressable
3. All resources share a uniform interface:
    a) A constrained set of well-defined operations
    b) A constrained set of content types

Outline

- Background
  - Web Technologies
- **Linked Data**
  - **Linked Data Principles**
  - Data Publishing
  - Data Consumption
- Linked Services
  - Linked Services Motivation
  - Linked (Open) Service Principles
  - Linked Service Implementation

Linked Data Principles

The defining principles of Linked Data are few and simply stated:

Linked Data Principles

1. Use URIs as names for things
2. Use HTTP URIs so that people can look up those names.
3. When someone looks up a URI, provide useful information, using the standards (RDF*, SPARQL)
4. Include links to other URIs. so that they can discover more things.

These are the latest version of their statement, contained in a W3C Note ‘Linked Data – Design Issues’ by Tim Berners-Lee

http://www.w3.org/DesignIssues/LinkedData.html

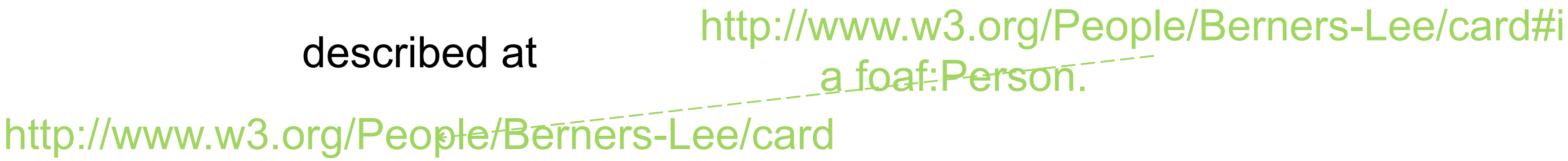
Linked Data Principle 1

Linked Data Principles

1. Use URIs as names for things

…

- A foundational issue in Linked Data was the distinction of URIs for **real-world objects** versus (e.g., RDF) **documents** that might describe them.
- One solution is to include a **Cool URI** with a hash, as follows:



http://www.w3.org/People/Berners-Lee/card#i

a foaf:Person.

described at

http://www.w3.org/People/Berners-Lee/card

•Note that Web browsers already crop hash URIs in this way

http://www.w3.org/TR/cooluris/#semweb

Linked Data Principle 2

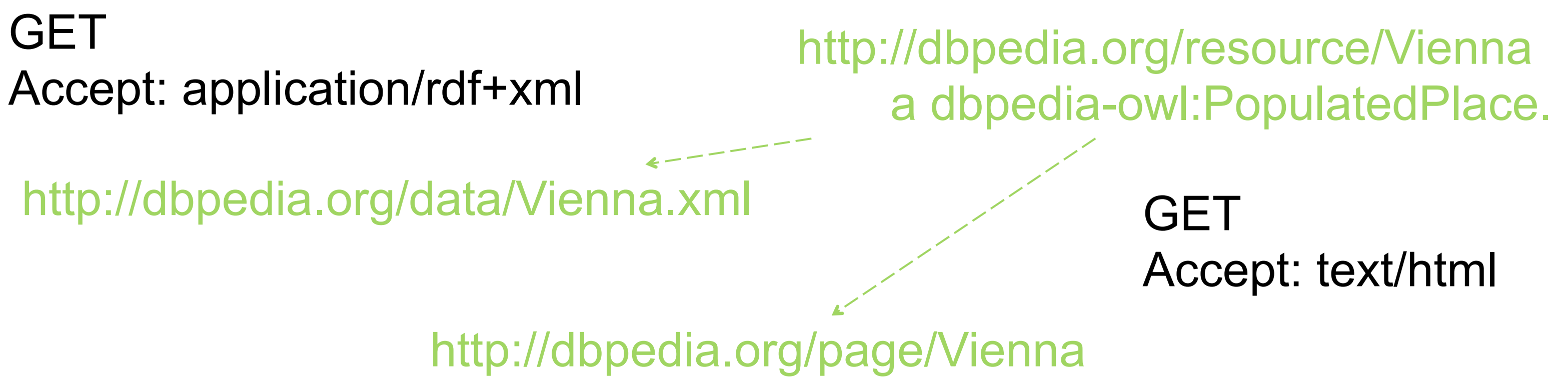
Linked Data Principles

…

2. Use HTTP URIs so that people can look up those names…

…

- HTTP allows a second way to distinguish real-world objects from documents, e.g. In DBPedia:



GET
Accept: application/rdf+xml

http://dbpedia.org/resource/Vienna
a dbpedia-owl:PopulatedPlace.

http://dbpedia.org/data/Vienna.xml

GET
Accept: text/html

http://dbpedia.org/page/Vienna

- Principles say HTTP 303 and Location header should be used

Linked Data Principle 3

Linked Data Principles

…

3. When someone looks up a URI, provide useful information, using the standards (RDF*, SPARQL)

- While RDF/XML should be the default for look-up
  - RDFa annotations in HTML are now also standard
  - it is increasingly encouraged to also offer Turtle:

GET
Accept: text/n3

http://dbpedia.org/resource/Vienna
a dbpedia-owl:PopulatedPlace.

←

http://dbpedia.org/data/Vienna.n3

- A dump of the whole dataset, and a SPARQL endpoint for queries are also encouraged (see Publishing)

Linked Data Principle 4

Linked Data Principles

…

4. Include links to other URIs. so that they can discover more things.

- There are several ways to reuse URIs:
  - At schema level –
    - direct **reuse** of class/property
    - (RDFS) **sub**-class/-property
    - (OWL) **equivalent** class/property
    - SKOS **broad match**

Interlinking

- At instance level –
  - direct **reuse**
  - (RDFS) **seeAlso**
  - (OWL) **sameAs**

There are a number of standard schemas, considered next for publishing, and datasets, considered in the Cloud, that should always be considered.

What to Return for a URI

- **The (immediate) description:** All triples that have the resource's URI as the subject.
- **Backlinks:** All triples that have the resource's URI as the object. This is redundant, but it allows bi-directional traversal.
- **Related descriptions**: Anything about related resources that may be of interest in typical usage scenarios; use prudence.
- **Metadata:** Any metadata such as the author and licensing information.
- **Syntax:** At least RDF descriptions as RDF/XML which is the only official syntax for RDF.
    - As RDF/XML is not very human-readable, the data could additionally be provided in other formats; e.g., for MIME-type application/x-turtle.

How to Publish Linked Data on the Web
Chris Bizer, Richard Cyganiak, Tom Heath

Note that text/n3 and text/rdf+n3
are currently better-supported though
a registered Internet Media type will be established

Outline

- Background
  - Web Technologies
- **Linked Data**
  - Linked Data Principles
  - **Data Publishing**
  - Data Consumption
- Linked Services
  - Linked Services Motivation
  - Linked (Open) Service Principles
  - Linked Service Implementation

Publication Means

There are a number of popular means to publish Linked Data:

- Host RDF/XML as static files behind Web server
- Include RDFa in HTML generated from existing content management system
- Publish direct from triplestore
- Expose relational database via translation

Publishing Static RDF

- The easiest way to create some simple Linked Data content is to host a hand-edited file on a Web server
- For example this is often used to identify a person by hosting a FOAF file (see later)
- There are consequences with respect to best practice:
  - the Web server should be configured to respect (Accept) requests according to the proper Internet Media types
  - it is messy (based on duplication) to offer alternative serialisations
  - SPARQL-based queries have to load whole graph so these should be small
- This approach is only sensible at small scale

Publishing Dynamic RDFa

- Not covered in depth in this module, RDFa is W3C Recommendation for including RDF as annotations to HTML
- It is argued that this can be the minimally invasive way to augment existing Web systems:
    - Web servers need minimal reconfiguration (should recognise application/xhtml+xml media type)
    - Some generic content management systems (CMS), e.g. Drupal, and some eCommerce solutions, already include RDFa support
    - It is argued e.g., for GoodRelations (see later) that annotation of text features can be optional and 'RDFa blocks' included into HTML
- Arguably, though, RDFa is least readable to humans, though both RDF/XML and Turtle can easily be obtained from a 'distiller'

http://www.w3.org/TR/rdfa-syntax/

Publishing Direct from Triple Store

- The most expedient way to follow principles and best practice for large datasets is to store data in RDF form in a triple store
- This has a number of advantages:
  - Most triplestores allow HTTP negotiation of serialisation
  - Allows direct processing of SPARQL queries, and therefore provision of a SPARQL endpoint, over data
  - Implicitly supports per resource deferencing via DESCRIBE queries
- Good solution for hosting new datasets, however in many cases there already exists an infrastructure which cannot be so easily replaced.

Exposing Relational Database via Translation

- Many large datasets are managed using relational DBMSs
- D2R is a popular solution for providing, via translation:
  - Dereferencing of resources
  - SPARQL processing, and
  - complete RDF dumps
- Translation rules expressed in D2RQ, a Turtle encoding of:
  - Mapping between major relational tables and RDFS classes
  - ‘Bridges’ between columns and RDF properties, including
  - Conditions and (programmatic) translations
- The W3C has recently chartered a working group to work on a standard for such mappings

http://www.w3.org/2009/08/rdb2rdf-charter

Vocabulary of Interlinked Datasets

A vocabulary of growing importance in publishing datasets is VoID, the Vocabulary of Interlinked Datasets, defining in RDFS:

•Access metadata:

- Example resources via resolvable URIs and URI patterns – void:exampleResource
- SPARQL endpoints – void:sparqlEndpoint
- Dumps – void:dataDump

- Used vocabularies – void:vocabulary
- Statistics – void:triples, void:distinctObjects
- Interlinkage – void:LinkSet, void:target

VoID also guides the use of existing vocabularies for

- Licensing, via dc:license
- Contact information, via foaf:Person

http://www.w3.org/TR/void/

Outline

- Background
  - Web Technologies
- **Linked Data**
  - Linked Data Principles
  - Data Publishing
  - **Data Consumption**
- Linked Services
  - Linked Services Motivation
  - Linked (Open) Service Principles
  - Linked Service Implementation

Linked Data Consumption

The consumption of Linked Data follows two strategies that are comparable to current Web use, with additional advantages, and one further:

- **Linked Data Browsing** – Linked Data best practice and inter-linkage are exploited to offer an effective human guided traversal of the Web of Data
- **Linked Data Search** – semantics are exploited to improve on existing Web search over the Web of Data
- **Linked Data Mash-ups** – datasets are brought together to create new applications

Linked Data Browsing

Traditional Web Browsers:

- render pre-existing documents
- display these according to their fixed HTML representation
- navigate according to pre-existing hyperlinks (in document).

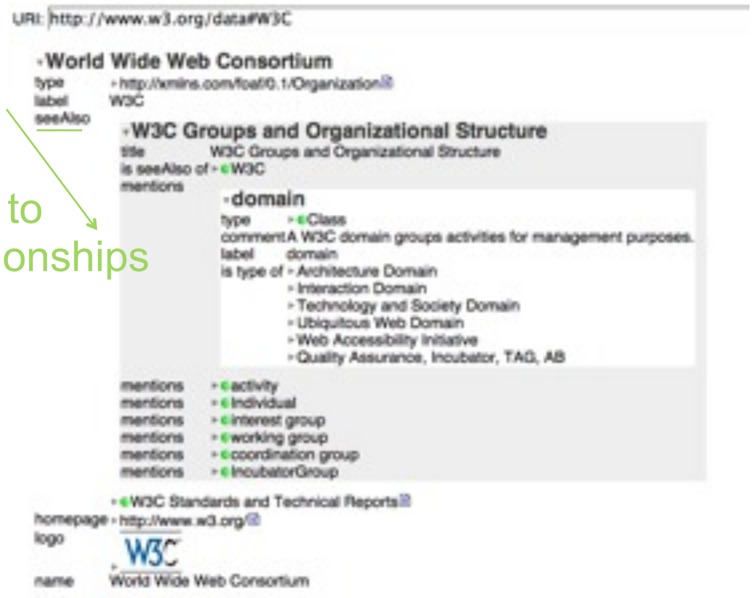
Linked Data Browsers

- render resources according both to pre-existing HTML representations, but also interpretations of RDF terms:
    - rdfs:label implies a label for something, which may have internationalisation via language tagging
    - rdfs:comment implies further information that may be optionally displayed, e.g., as a pop up
    - foaf:depiction implies availability of a picture, etc.
- navigate also according to related datasets, even if not nominated by original publisher of resource of focus

Linked Data Browser - Tabulator

Tabulator is an extension to the Firefox browser with views for tables, maps, calendars, timelines, etc. and mechanism for extension to other views

browsing between resources according to RDF relationships



URI: http://www.w3.org/data#W3C

World Wide Web Consortium

type	http://xmlns.com/foaf/0.1/Organization
label	W3C
seeAlso	

W3C Groups and Organizational Structure

title	W3C Groups and Organizational Structure
is seeAlso of	W3C
mentions	

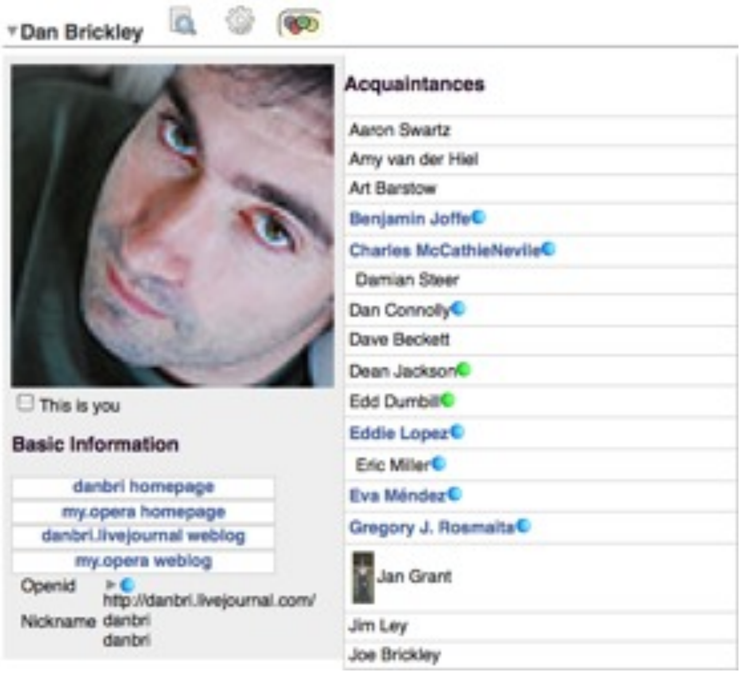
domain

type	Class
comment	A W3C domain groups activities for management purposes.
label	domain
is type of	Architecture Domain Interaction Domain Technology and Society Domain Ubiquitous Web Domain Web Accessibility Initiative Quality Assurance, Incubator, TAG, AB

mentions	activity
mentions	Individual
mentions	interest group
mentions	working group
mentions	coordination group
mentions	IncubatorGroup

	W3C Standards and Technical Reports
homepage	http://www.w3.org/
logo	W3C
name	World Wide Web Consortium

FOAF view



Dan Brickley

□ This is you

Basic Information

danbri homepage

my.opera homepage

danbri.livejournal weblog

my.opera weblog

Openid	http://danbri.livejournal.com/
Nickname	danbri danbri

Acquaintances

- Aaron Swartz
- Amy van der Hiel
- Art Barstow
- Benjamin Joffe
- Charles McCathieNevile
- Damian Steer
- Dan Connolly
- Dave Beckett
- Dean Jackson
- Edd Dumbill
- Eddie Lopez
- Eric Miller
- Eva Méndez
- Gregory J. Rosmaita
- Jan Grant
- Jim Ley
- Joe Brickley

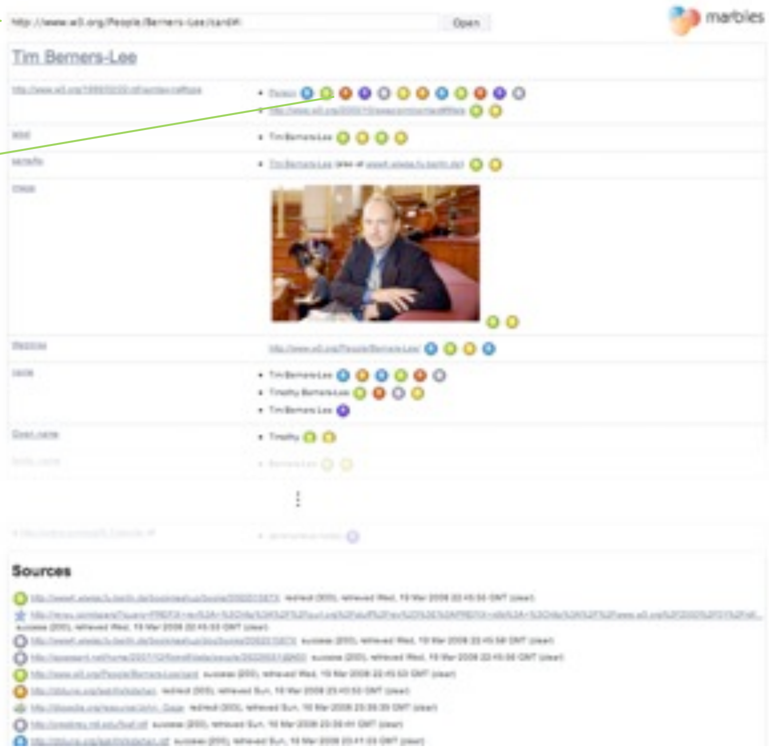
http://www.w3.org/2005/ajar/tab

08.08.2011

Linked Data Browser - Marbles

Tabulator is XHTML server that aggregates Linked Data about nominated resources

It displays this using coloured ‘marbles’ to illustrate where each item came from and to allow navigation



http://marbles.sourceforge.net/

Linked Data Search

Traditional Web search, covered in previous module:

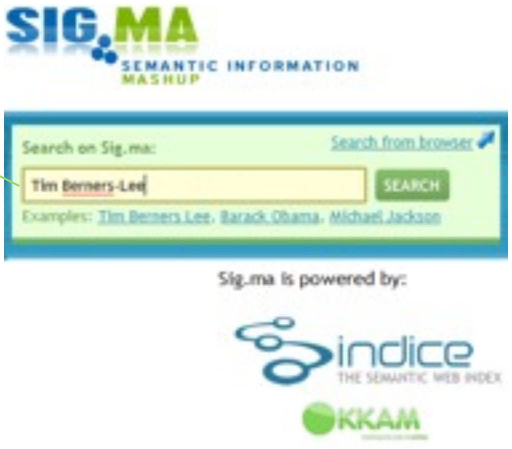
- primarily keyword-based
- crawls documents via hyperlinks
- stores each as ‘bag of words’
- uses (ambiguous) hyperlinks only to judge popularity
- presents results mostly as a set of links

Semantic / Linked Data Search

- crawls given RDF-described inter-linkage structure
- integrates information about unambiguous resources
- interprets query generally, using inference to find matches
- presents results intelligently and in coherent fashion

Linked Data Search – Sig.ma

Sig.ma allows keyword-based search and presents aggregated results, via Sindice, in style of Linked Data Browsing, as well as list of sources



SIG.MA SEMANTIC INFORMATION MASHUP

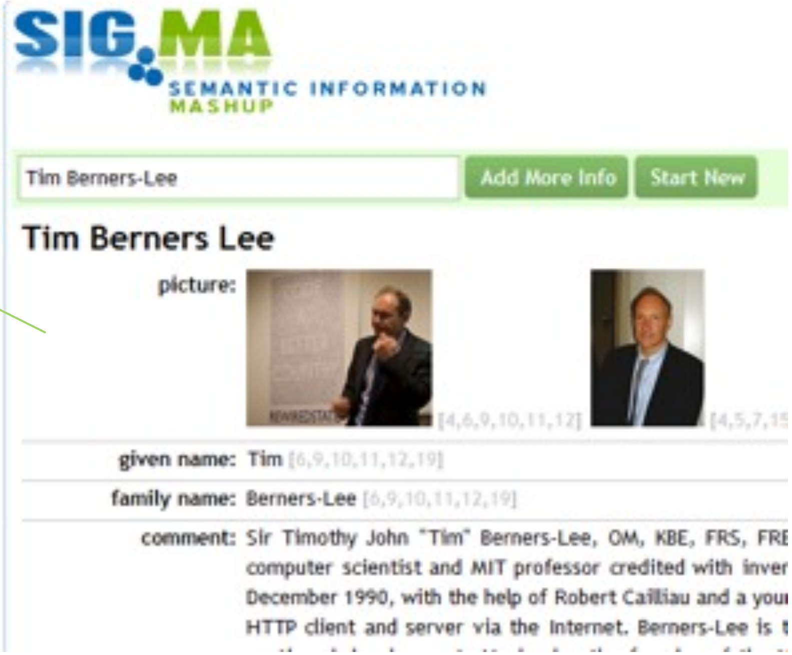
Search on Sig.ma: Search from browser

Tim Berners-Lee

SEARCH

Examples: Tim Berners Lee, Barack Obama, Michael Jackson

Sig.ma is powered by:



SIG.MA SEMANTIC INFORMATION MASHUP

Tim Berners-Lee

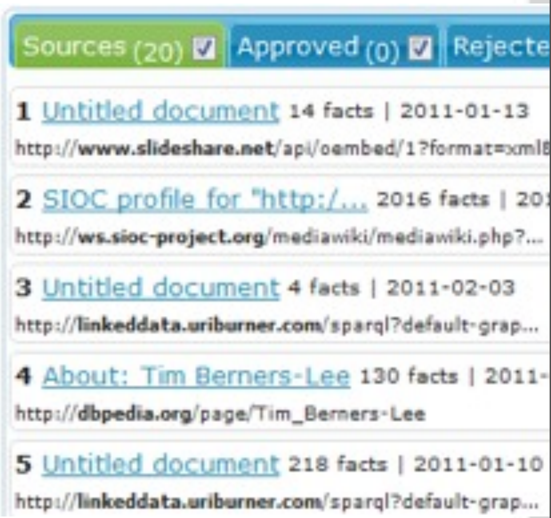
Tim Berners Lee

picture: [4,6,9,10,11,12] [4,5,7,15

given name: Tim [6,9,10,11,12,19]

family name: Berners-Lee [6,9,10,11,12,19]

comment: Sir Timothy John "Tim" Berners-Lee, OM, KBE, FRS, FRE
computer scientist and MIT professor credited with inver
December 1990, with the help of Robert Cailliau and a you
HTTP client and server via the Internet. Berners-Lee is t



Sources (20) ☑ Approved (0) ☑ Rejecte

1 Untitled document 14 facts | 2011-01-13
http://www.slideshare.net/api/oembed/1?format=xml&

2 SIOC profile for "http:/... 2016 facts | 20
http://ws.sioc-project.org/mediawiki/mediawiki.php?...

3 Untitled document 4 facts | 2011-02-03
http://linkeddata.uriburner.com/sparql?default-grap...

4 About: Tim Berners-Lee 130 facts | 2011-
http://dbpedia.org/page/Tim_Berners-Lee

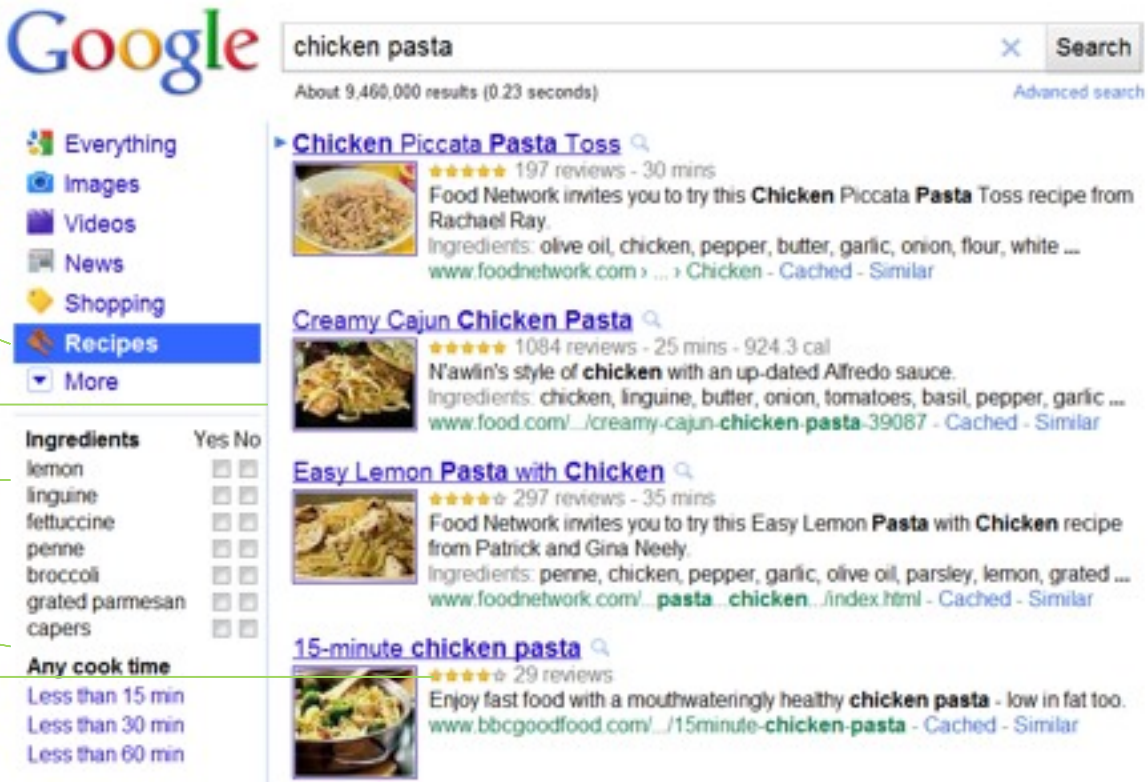
5 Untitled document 218 facts | 2011-01-10
http://linkeddata.uriburner.com/sparql?default-grap...

http://sig.ma/

Linked Data Search – Google Rich Snippets

Google increasingly uses semantic annotations to improve on traditional search results,

for instance
in finding
recipes
& showing
pictures,
ingredients,
timing
& reviews



About 9,460,000 results (0.23 seconds)

Chicken Piccata Pasta Toss
★★★★★ 197 reviews - 30 mins
Food Network invites you to try this **Chicken** Piccata **Pasta** Toss recipe from Rachael Ray.
Ingredients: olive oil, chicken, pepper, butter, garlic, onion, flour, white ...
www.foodnetwork.com › ... › Chicken - Cached - Similar

Creamy Cajun Chicken Pasta
★★★★★ 1084 reviews - 25 mins - 924.3 cal
N'awlin's style of **chicken** with an up-dated Alfredo sauce.
Ingredients: chicken, linguine, butter, onion, tomatoes, basil, pepper, garlic ...
www.food.com/.../creamy-cajun-**chicken**-**pasta**-39087 - Cached - Similar

Easy Lemon Pasta with Chicken
★★★★☆ 297 reviews - 35 mins
Food Network invites you to try this Easy Lemon **Pasta** with **Chicken** recipe from Patrick and Gina Neely.
Ingredients: penne, chicken, pepper, garlic, olive oil, parsley, lemon, grated ...
www.foodnetwork.com/...**pasta**...**chicken**.../index.html - Cached - Similar

15-minute chicken pasta
★★★★☆ 29 reviews
Enjoy fast food with a mouthwateringly healthy **chicken pasta** - low in fat too.
www.bbcgoodfood.com/.../15minute-**chicken**-**pasta** - Cached - Similar

Linked Data Mash-up

Building an application aggregating Linked Data balances two strategies:

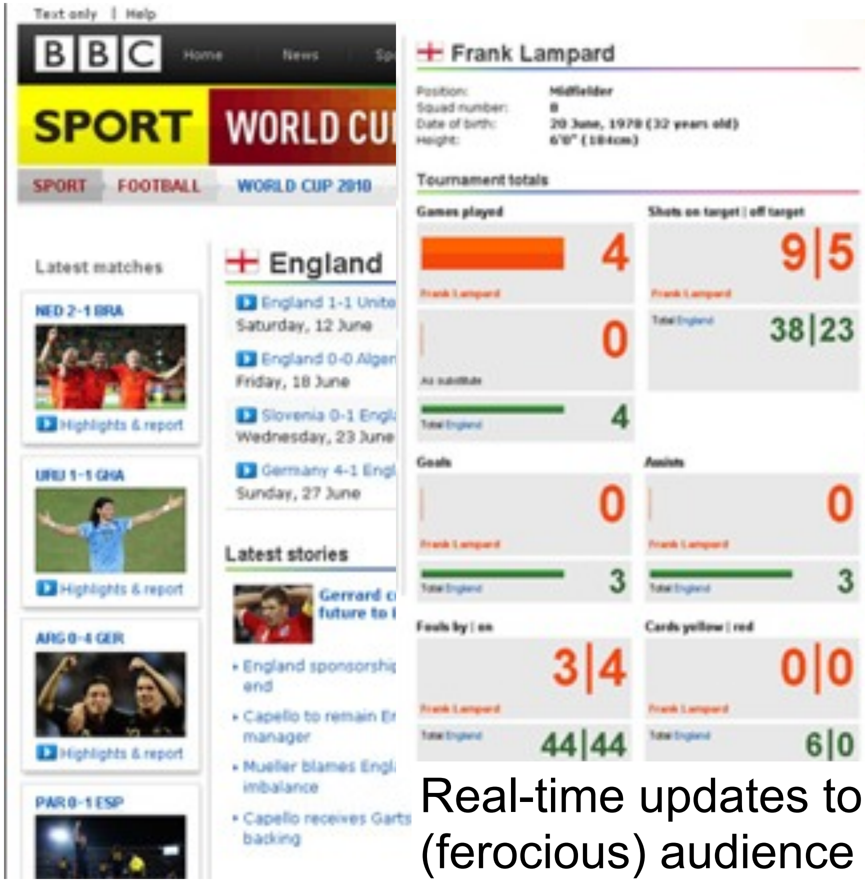
- Utilise a triple store that has application data already available via –
  - **Crawling** ahead of time, using tools such as LDSpider
  - **Loading** of pertinent datastores from dumps
- On-the-fly aggregation, via –
  - **Run-time dereferencing** of RDF resource descriptions
  - **Federation** between SPARQL endpoints

Significant effort is still needed to achieve reusable visualisation of Linked Data; many current mash-ups are forced to transform data to non Semantic Web formats.

Outline

- Background
  - Web Technologies
- Linked Data
  - Linked Data Principles
  - Data Publishing
  - Data Consumption
- **Linked Services**
  - **Linked Services Motivation**
  - Linked (Open) Service Principles
  - Linked Service Implementation

RDF Services at the BBC



Text only | Help

BBC Home News Sp

SPORT WORLD CU

SPORT FOOTBALL WORLD CUP 2010

Latest matches

NED 2-1 BRA
Highlights & report

URU 1-1 GHA
Highlights & report

ARG 0-4 GER
Highlights & report

PAR 0-1 ESP

England

England 1-1 Unite
Saturday, 12 June

England 0-0 Alger
Friday, 18 June

Slovenia 0-1 Engl
Wednesday, 23 June

Germany 4-1 Engl
Sunday, 27 June

Latest stories

Gerrard c
future to I

- England sponsorship end
- Capello to remain En manager
- Mueller blames Engl imbalance
- Capello receives Garts backing

Frank Lampard

Position:	Midfielder
Squad number:	8
Date of birth:	20 June, 1978 (32 years old)
Height:	6'0" (184cm)

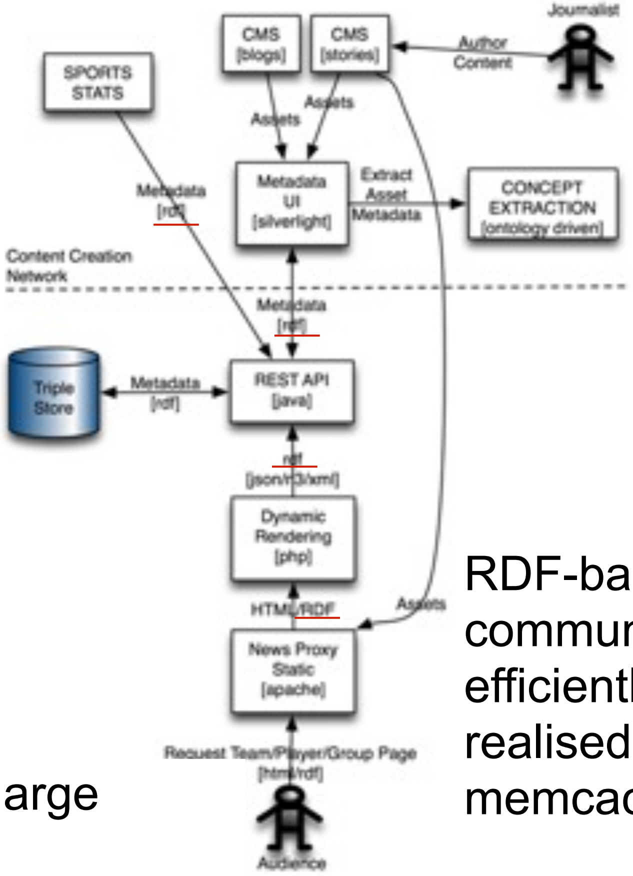
Tournament totals

Games played		Shots on target | off target	
Frank Lampard	4	Frank Lampard	9|5
As substitute	0	Total England	38|23
Total England	4		

Goals		Assists	
Frank Lampard	0	Frank Lampard	0
Total England	3	Total England	3

Fouls by | on		Cards yellow | red	
Frank Lampard	3|4	Frank Lampard	0|0
Total England	44|44	Total England	6|0

Real-time updates to a large (ferocious) audience



RDF-based communication efficiently realised using memcached

Motivating Example Datasets

The GeoNames geographical database is available for download free of charge under a creative commons attribution license. It contains over 10 million geographical names and consists of 7.5 million unique features.

The MusicBrainz music metadata database is available for download under a public domain (CC no rights reserved) and CC Attribution-NonCommercial-ShareAlike 2.0 license. It contains over 10 million track descriptions.

Motivating Example Linked Data

GeoNames

GeoNames categorises geo features in 9 feature classes sub-categorized into 645 feature codes. An SKOS taxonomy reflects this structure and OWL is used to describe features; URIs can be resolved to RDF and a dump is available. Links are made to DBPedia and other LOD sets.

LinkedBrainz

The MusicBrainz NGS schema has been mapped into the Music Ontology. HTML pages describing MusicBrainz entities are now annotated in RDFa. A dump will soon be made available. Links are made to DBPedia and LOD sets, such as BBC Music Reviews, use MB IDs.

Motivating Example APIs

GeoNames

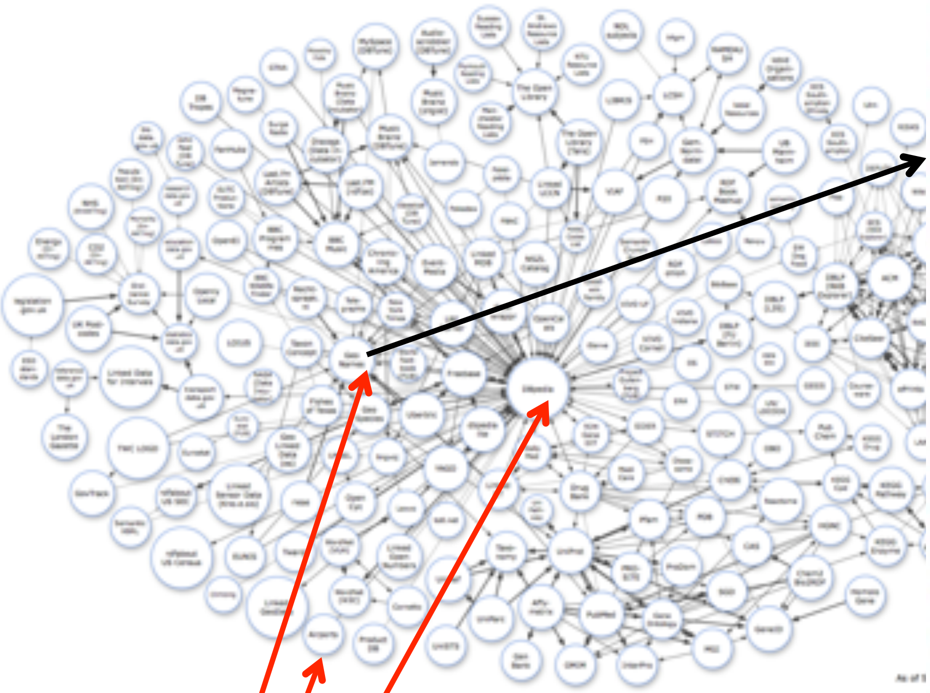
GeoNames offers a number of reverse geolocation (and containment)-based retrievals – the feature is not directly identified, but a set of qualifying resources are computed from a point (/circle/box)

LinkedBrainz

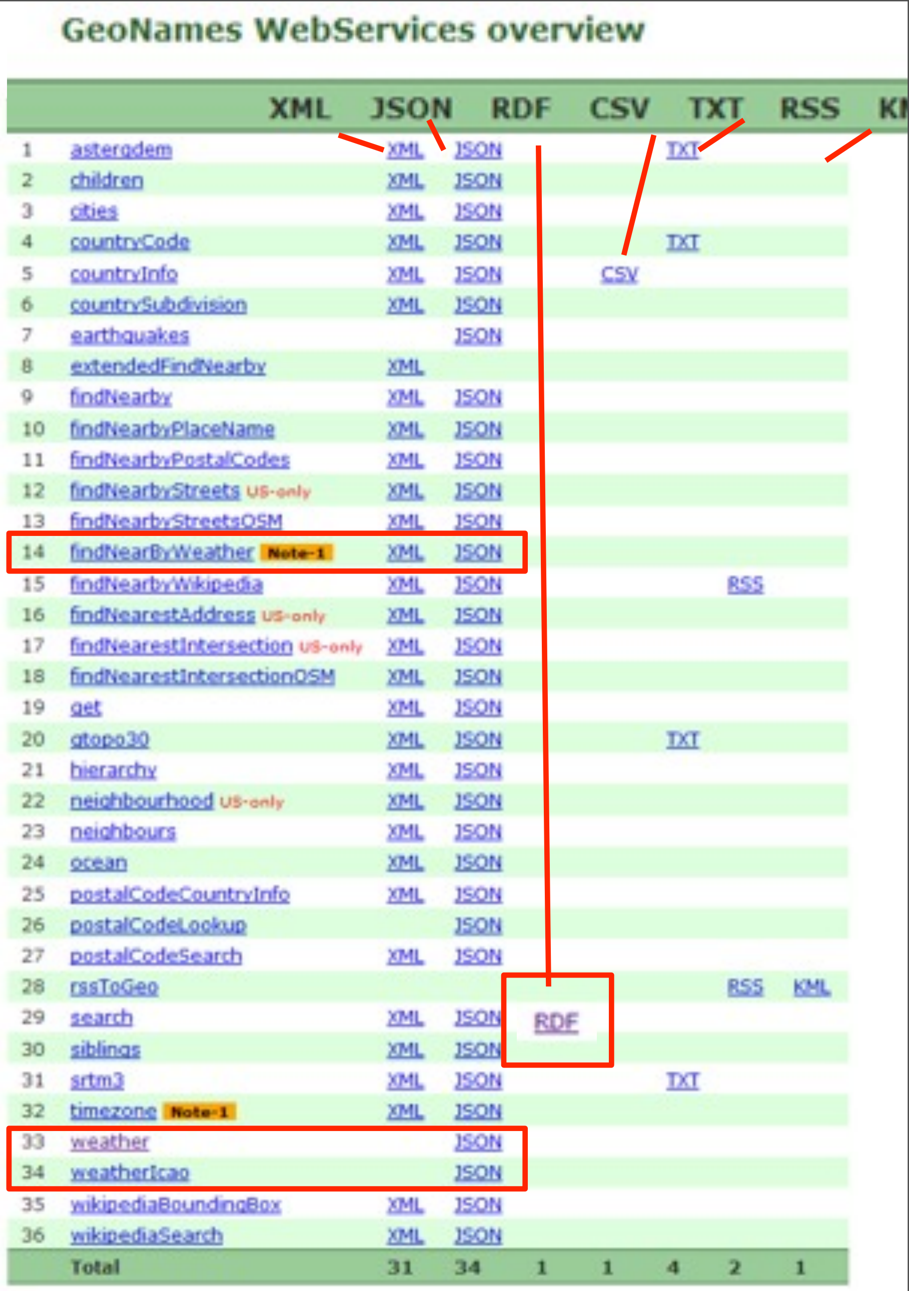
MusicBrainz incorporates PUIDs and uses MusicIP’s MusicDNS audio fingerprinting technologies. When a PUID is submitted they are fuzzily matched to a set of similar tracks.

In both cases **computation** is needed, before retieval, to locate the resource to be included. This needs to happen **near the data**, not be remotely pushed to it

Geonames Services



```
{"weatherObservation":
 {"clouds":"broken clouds",
  "weatherCondition":"drizzle",
  "observation":"LESO 251300Z 03007KT
                340V040 CAVOK 23/15 Q1010",
  "windDirection":30,
  "ICAO":"LESO", ...
```



GeoNames WebServices overview

		XML	JSON	RDF	CSV	TXT	RSS	KI
1	asterdem	XML	JSON			TXT		
2	children	XML	JSON					
3	cities	XML	JSON					
4	countryCode	XML	JSON			TXT		
5	countryInfo	XML	JSON		CSV			
6	countrySubdivision	XML	JSON					
7	earthquakes		JSON					
8	extendedFindNearby	XML						
9	findNearby	XML	JSON					
10	findNearbyPlaceName	XML	JSON					
11	findNearbyPostalCodes	XML	JSON					
12	findNearbyStreets US-only	XML	JSON					
13	findNearbyStreetsOSM	XML	JSON					
14	findNearByWeather Note-1	XML	JSON					
15	findNearbyWikipedia	XML	JSON				RSS	
16	findNearestAddress US-only	XML	JSON					
17	findNearestIntersection US-only	XML	JSON					
18	findNearestIntersectionOSM	XML	JSON					
19	get	XML	JSON					
20	gtopo30	XML	JSON			TXT		
21	hierarchy	XML	JSON					
22	neighbourhood US-only	XML	JSON					
23	neighbours	XML	JSON					
24	ocean	XML	JSON					
25	postalCodeCountryInfo	XML	JSON					
26	postalCodeLookup		JSON					
27	postalCodeSearch	XML	JSON					
28	rssToGeo						RSS	KML
29	search	XML	JSON	RDF				
30	siblings	XML	JSON					
31	srtm3	XML	JSON			TXT		
32	timezone Note-1	XML	JSON					
33	weather		JSON					
34	weatherIcao		JSON					
35	wikipediaBoundingBox	XML	JSON					
36	wikipediaSearch	XML	JSON					
	Total	31	34	1	1	4	2	1

Outline

- Background
  - Web Technologies
- Linked Data
  - Linked Data Principles
  - Data Publishing
  - Data Consumption
- **Linked Services**
  - Linked Services Motivation
  - **Linked (Open) Service Principles**
  - Linked Service Implementation

LOS Principles

REST Principles

1. Application state and functionality is divided into resources
2. Every resource is uniquely addressable
3. All resources share a uniform interface:
   a) A constrained set of well-defined operations
   b) A constrained set of content types

Linked Data Principles

1. Use URIs as names for things
2. Use HTTP URIs so that people can look up those names.
3. When someone looks up a URI, provide useful information, using the standards (RDF*, SPARQL)
4. Include links to other URIs. so that they can discover more things.

Linked Open Service Principles

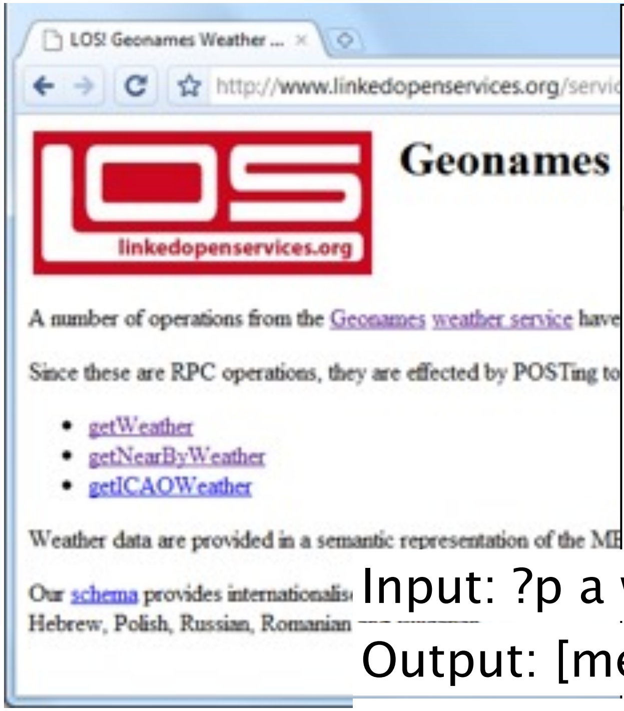
1. Describe services as LOD prosumers with input and output descriptions as **SPARQL graph patterns**
2. **Communicate RDF** by RESTful content negotiation
3. The output should make explicit its **relation** with the input

Linked Open Services Principles

1

Describe services' input and output as **SPARQL graph patterns**

LOS Weather Service



LOS! Geonames Weather ...
http://www.linkedopenservices.org/servic
linkedopenservices.org

Geonames

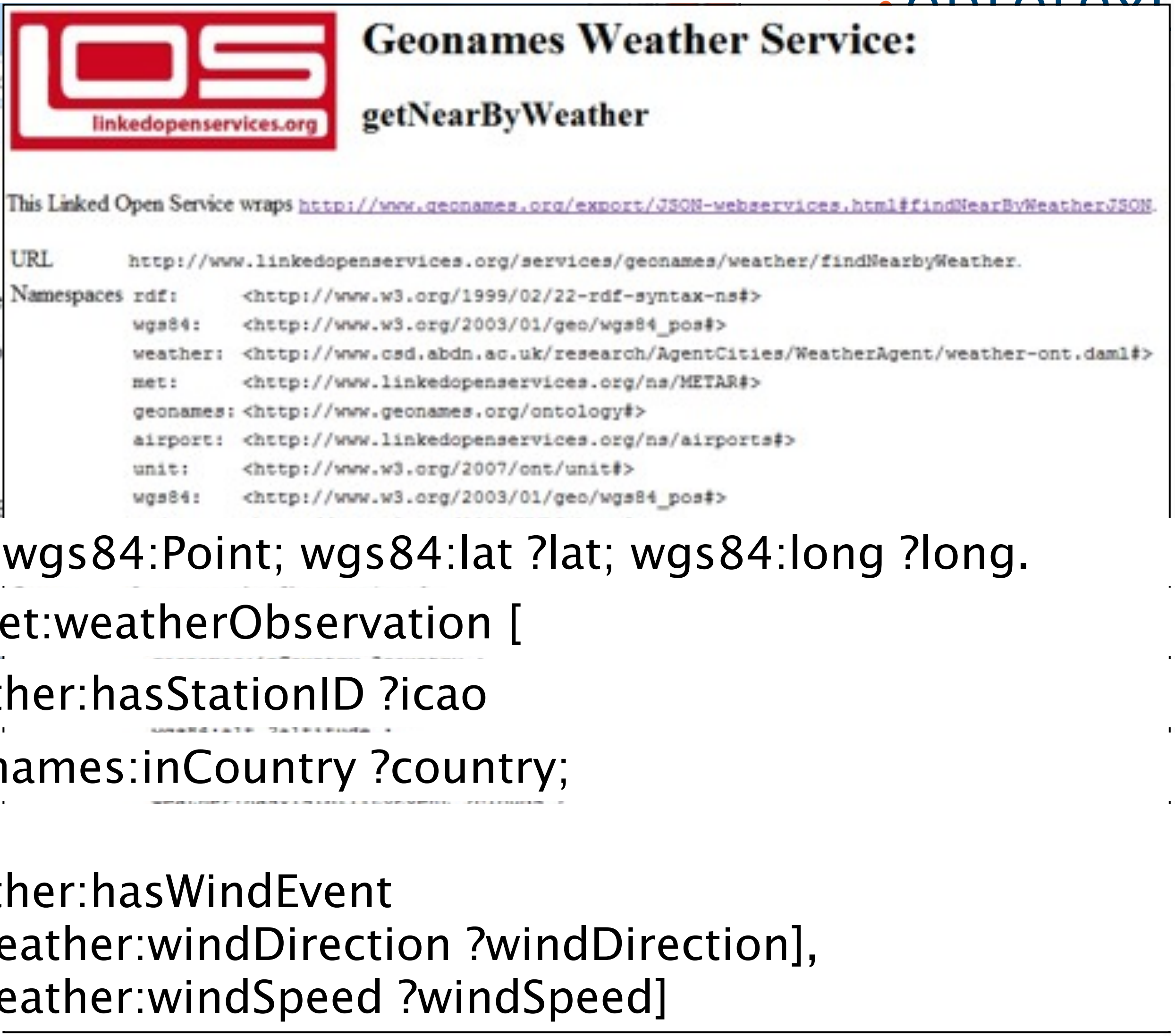
A number of operations from the Geonames weather service have

Since these are RPC operations, they are effected by POSTing to

- getWeather
- getNearByWeather
- getICAOWeather

Weather data are provided in a semantic representation of the ME

Our schema provides internationalis
Hebrew, Polish, Russian, Romanian



Geonames Weather Service:

getNearByWeather

This Linked Open Service wraps http://www.geonames.org/export/JSON-webservices.html#findNearByWeatherJSON.

```
URL        http://www.linkedopenservices.org/services/geonames/weather/findNearbyWeather.
Namespaces rdf:      <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
           wgs84:    <http://www.w3.org/2003/01/geo/wgs84_pos#>
           weather:  <http://www.csd.abdn.ac.uk/research/AgentCities/WeatherAgent/weather-ont.daml#>
           met:      <http://www.linkedopenservices.org/ns/METAR#>
           geonames: <http://www.geonames.org/ontology#>
           airport:  <http://www.linkedopenservices.org/ns/airports#>
           unit:     <http://www.w3.org/2007/ont/unit#>
           wgs84:    <http://www.w3.org/2003/01/geo/wgs84_pos#>
```

```
Input: ?p a wgs84:Point; wgs84:lat ?lat; wgs84:long ?long.
Output: [met:weatherObservation [
         weather:hasStationID ?icao
         geonames:inCountry ?country;
         ...
         weather:hasWindEvent
            [weather:windDirection ?windDirection],
            [weather:windSpeed ?windSpeed]
```

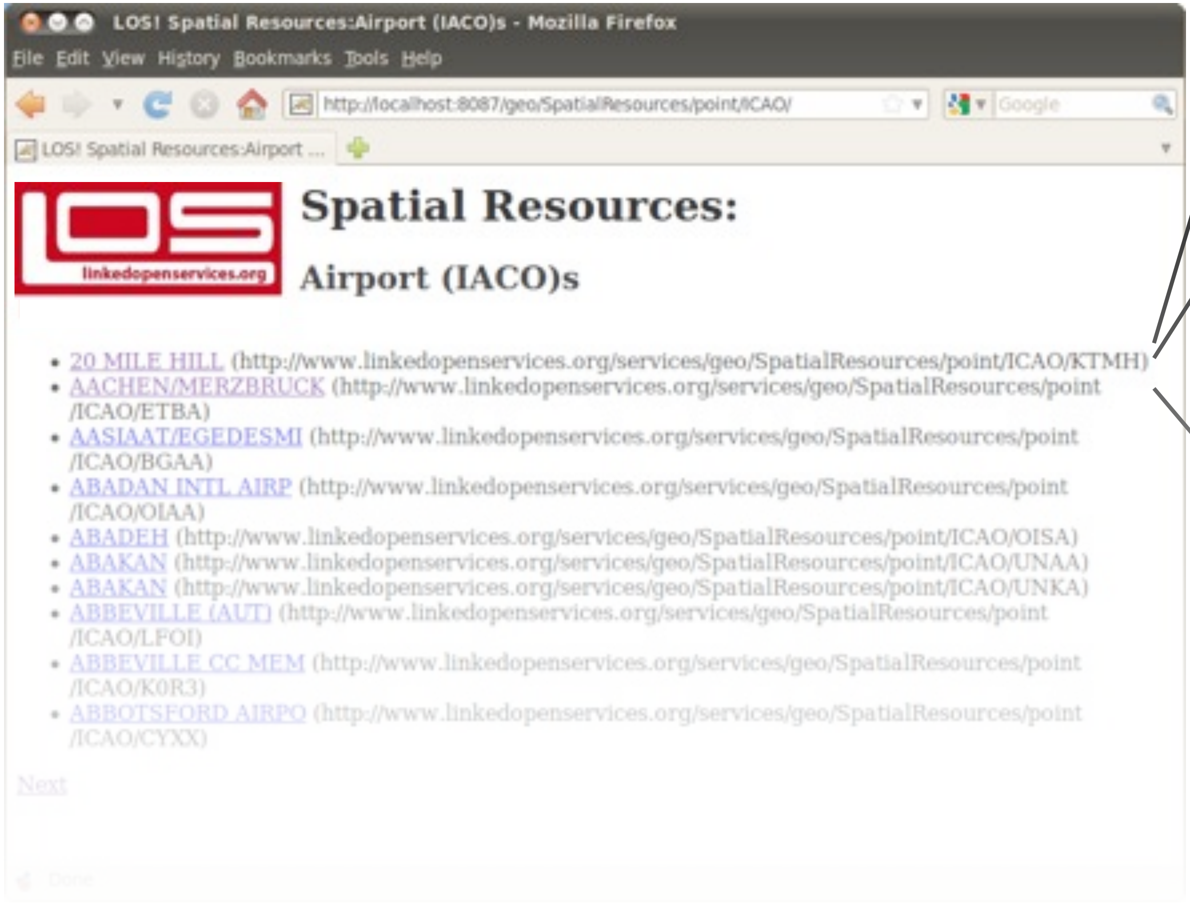
Linked Open Services Principles

2

Communicate RDF

by RESTful content negotiation

LOS Principle 2


LOS! Spatial Resources:Airport (IACO)s - Mozilla Firefox
File Edit View History Bookmarks Tools Help
http://localhost:8087/geo/SpatialResources/point/ICAO/
Google
LOS! Spatial Resources:Airport ...
linkedopenservices.org
Done

Spatial Resources:

Airport (IACO)s

- 20 MILE HILL (http://www.linkedopenservices.org/services/geo/SpatialResources/point/ICAO/KTMH)
- AACHEN/MERZBRUCK (http://www.linkedopenservices.org/services/geo/SpatialResources/point/ICAO/ETBA)
- AASIAAT/EGEDESMI (http://www.linkedopenservices.org/services/geo/SpatialResources/point/ICAO/BGAA)
- ABADAN INTL AIRP (http://www.linkedopenservices.org/services/geo/SpatialResources/point/ICAO/OIAA)
- ABADEH (http://www.linkedopenservices.org/services/geo/SpatialResources/point/ICAO/OISA)
- ABAKAN (http://www.linkedopenservices.org/services/geo/SpatialResources/point/ICAO/UNAA)
- ABAKAN (http://www.linkedopenservices.org/services/geo/SpatialResources/point/ICAO/UNKA)
- ABBEVILLE (AUT) (http://www.linkedopenservices.org/services/geo/SpatialResources/point/ICAO/LFOI)
- ABBEVILLE CC MEM (http://www.linkedopenservices.org/services/geo/SpatialResources/point/ICAO/K0R3)
- ABBOTSFORD AIRPO (http://www.linkedopenservices.org/services/geo/SpatialResources/point/ICAO/CYXX)

Next

GET
Accept: text/html
↓
303 REDIRECT /page

GET
Accept: application/rdf+xml
(or text/n3) ↓

303 REDIRECT /data

Linked Data

GET /weather
Accept: application/rdf+xml
(or text/n3)

200 <rdf:Description>

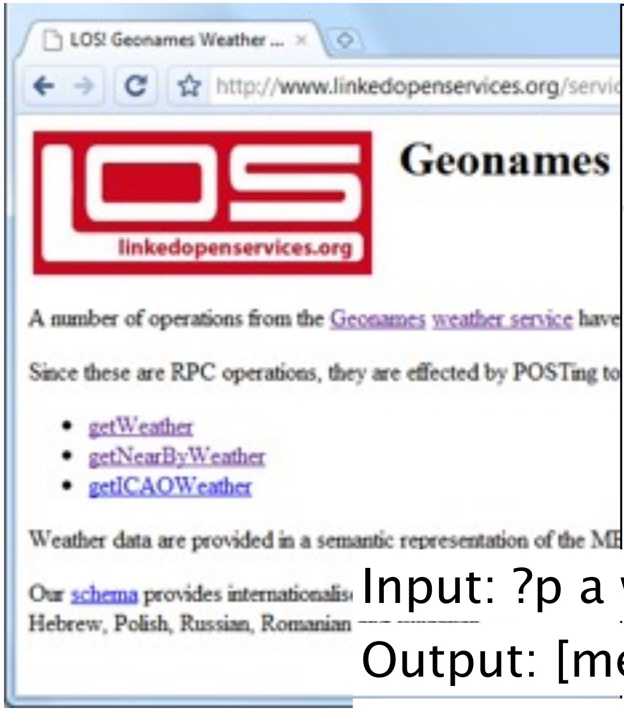
Linked Service

Linked Open Services Principles

3

The **output** should make **explicit** its **relation** with the **input**.

LOS Weather Service


LOS! Geonames Weather ...
http://www.linkedopenservices.org/servic
linkedopenservices.org

Geonames

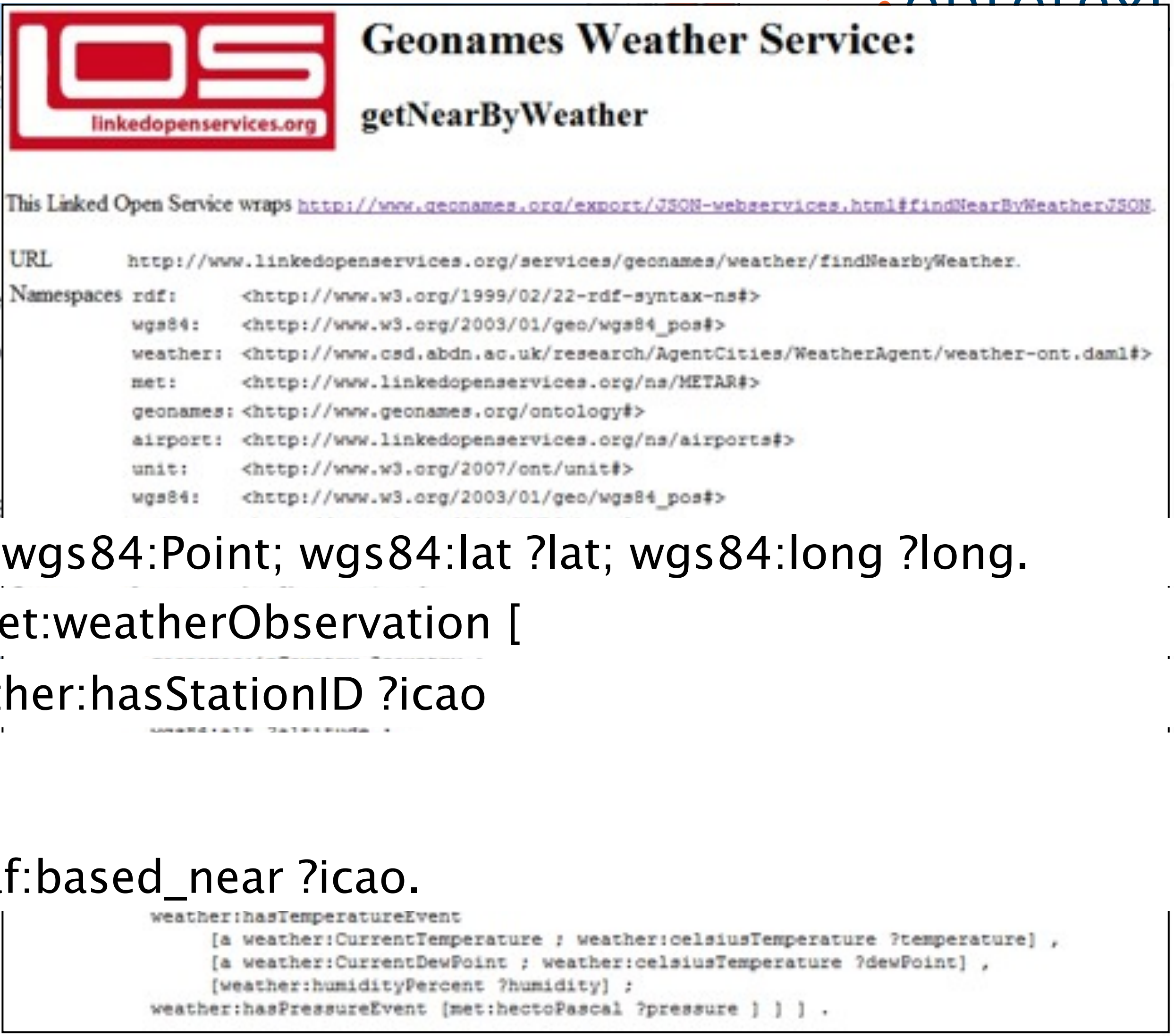
A number of operations from the Geonames weather service have

Since these are RPC operations, they are effected by POSTing to

- getWeather
- getNearByWeather
- getICAOWeather

Weather data are provided in a semantic representation of the ME

Our schema provides internationalis
Hebrew, Polish, Russian, Romanian



Geonames Weather Service:

getNearByWeather

This Linked Open Service wraps http://www.geonames.org/export/JSON-webservices.html#findNearByWeatherJSON.

```
URL        http://www.linkedopenservices.org/services/geonames/weather/findNearbyWeather.
Namespaces rdf:      <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
           wgs84:    <http://www.w3.org/2003/01/geo/wgs84_pos#>
           weather:  <http://www.csd.abdn.ac.uk/research/AgentCities/WeatherAgent/weather-ont.daml#>
           met:      <http://www.linkedopenservices.org/ns/METAR#>
           geonames: <http://www.geonames.org/ontology#>
           airport:  <http://www.linkedopenservices.org/ns/airports#>
           unit:     <http://www.w3.org/2007/ont/unit#>
           wgs84:    <http://www.w3.org/2003/01/geo/wgs84_pos#>
```

```
    weather:hasTemperatureEvent
          [a weather:CurrentTemperature ; weather:celsiusTemperature ?temperature] ,
          [a weather:CurrentDewPoint ; weather:celsiusTemperature ?dewPoint] ,
          [weather:humidityPercent ?humidity] ;
    weather:hasPressureEvent [met:hectoPascal ?pressure ] ] ] .
```

```
Input: ?p a wgs84:Point; wgs84:lat ?lat; wgs84:long ?long.
Output: [met:weatherObservation [
            weather:hasStationID ?icao
            ...]]
            ...
         ?p foaf:based_near ?icao.
```

Linked Open Services Principles

When wrapping non-LOS services:

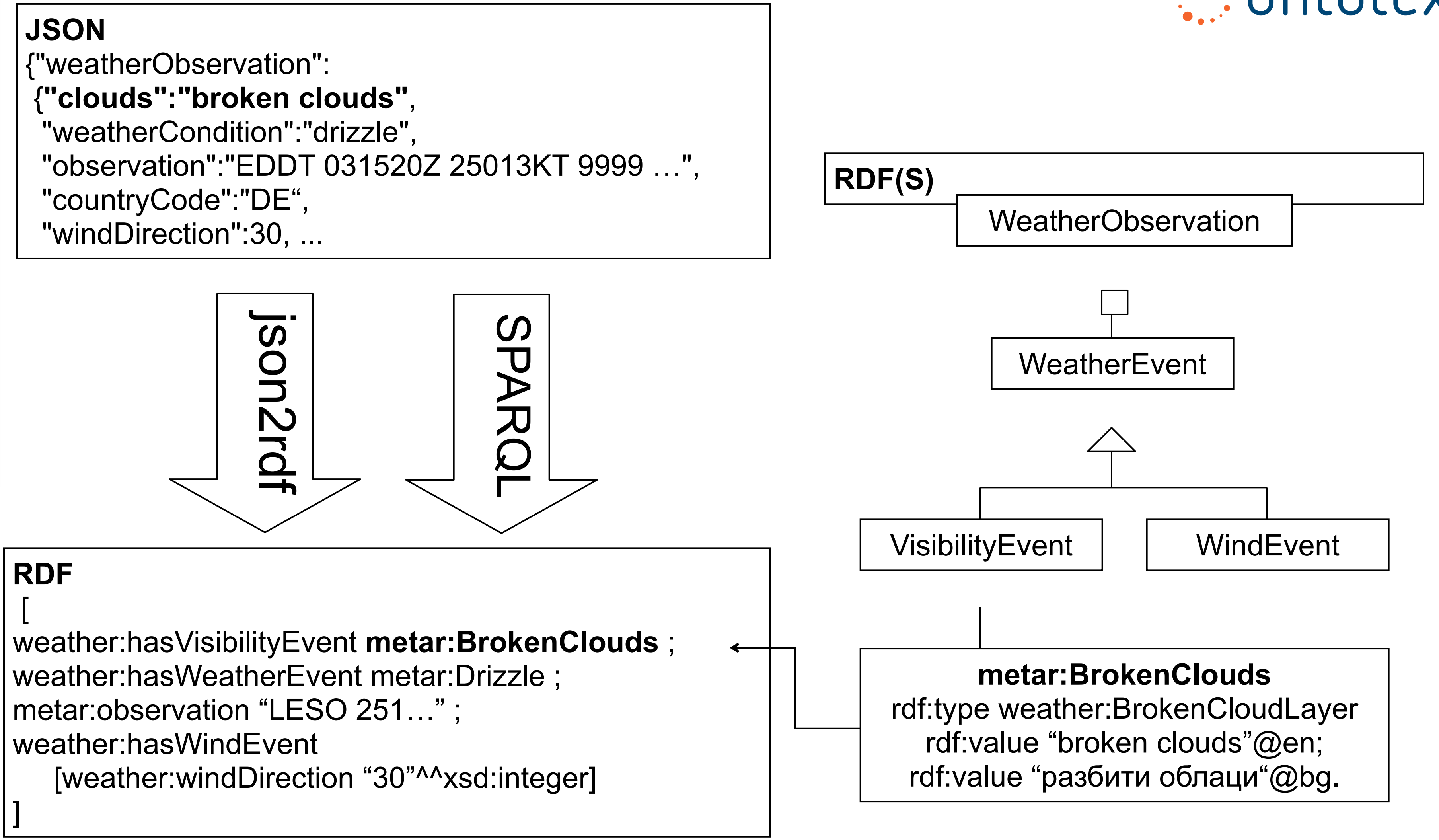
4

Make the **lifting/mapping open** as **SPARQL CONSTRUCT** queries

Outline

- Background
  - Web Technologies
- Linked Data
  - Linked Data Principles
  - Data Publishing
  - Data Consumption
- **Linked Services**
  - Linked Services Motivation
  - Linked (Open) Service Principles
  - **Linked Service Implementation**

LOS Wrapping Overview



JSON

{"weatherObservation":
 {**"clouds":"broken clouds"**,
  "weatherCondition":"drizzle",
  "observation":"EDDT 031520Z 25013KT 9999 …",
  "countryCode":"DE“,
  "windDirection":30, ...

RDF

 [
weather:hasVisibilityEvent **metar:BrokenClouds** ;
weather:hasWeatherEvent metar:Drizzle ;
metar:observation “LESO 251…” ;
weather:hasWindEvent
    [weather:windDirection “30”^^xsd:integer]
]

RDF(S)

metar:BrokenClouds
rdf:type weather:BrokenCloudLayer
rdf:value “broken clouds”@en;
rdf:value “разбити облаци“@bg.

LOS Wrapping Overview

JSON

```
{"weatherObservation":
 {"clouds":"broken clouds",
  "w
  "o
  "c
  "w
```

```
CONSTRUCT {
?bkn weather:hasVisibilityEvent met:BrokenClouds .
?dz weather:hasWeatherEvent met:Drizzle .
...
}
WHERE
{OPTIOI
...
}
```

```
CONSTRUCT {
?obs met:weatherObservation ?f .
?f weather:hasVisibilityEvent ?clouds ;
   weather:hasWindEvent [weather:windDirection ?
windDirection] ;
   ...
} WHERE {
?obs temp:weatherObservation ?f .
    temp:windDirection ?windDirection ;
    OPTIONAL {_:c weather:hasVisibilityEvent ?clouds }
    ...
```

herObservation

t

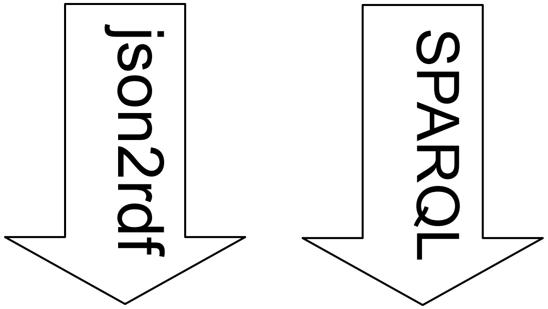
er

RDF

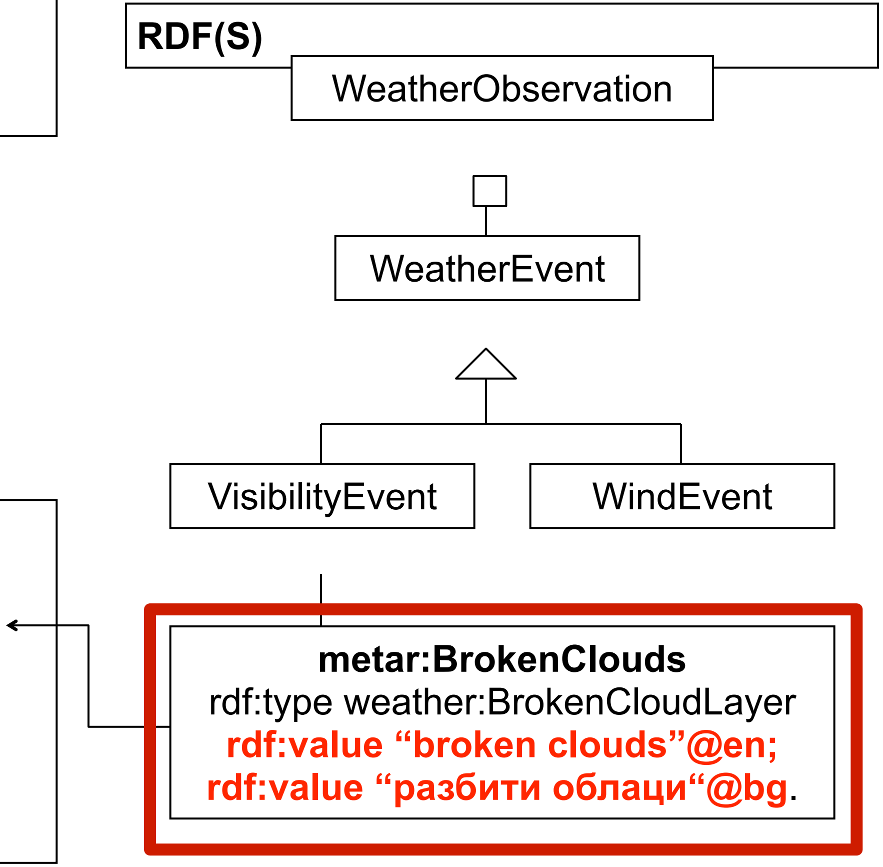
```
 [
weather:hasVisi
weather:hasWe
metar:observati
weather:hasWir
    [weather:win
]
```

LOS Wrapping Overview

JSON
{"weatherObservation":
 {**"clouds":"broken clouds"**,
  "weatherCondition":"drizzle",
  "observation":"EDDT 031520Z 25013KT 9999 …",
  "countryCode":"DE“,
  "windDirection":30, ...



RDF
 [
weather:hasVisibilityEvent **metar:BrokenClouds** ;
weather:hasWeatherEvent metar:Drizzle ;
metar:observation “LESO 251…” ;
weather:hasWindEvent
    [weather:windDirection “30”^^xsd:integer]
]

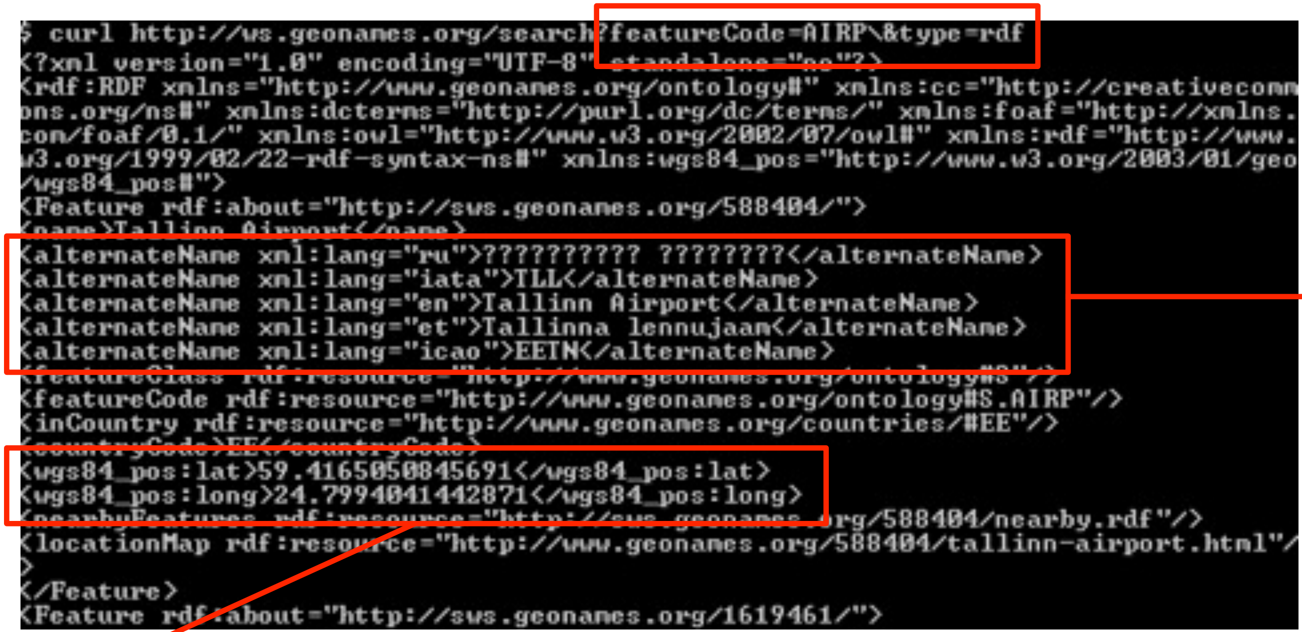


metar:BrokenClouds
rdf:type weather:BrokenCloudLayer
rdf:value “broken clouds”@en;
rdf:value “разбити облаци“@bg.

Linked Data Output

```
_:a1 met:weatherObservation [
    weather:hasStationID
            <http://www.linkedopenservices.org/services/geo/SpatialResources/point/ICAO/
EDDT> ;
 met:stationName "Berlin Tegel" ;
 geonames:inCountry "DE" ;
 wgs84:lat "52.566666"^^xsd:double ;
 wgs84:long "13.316667"^^xsd:double ;
 wgs84:alt "37.0"^^xsd:double ;
 met:datetime "2010-10-27 08:20:00" ;
 met:observation "EDDT 270820Z 18007KT CAVOK 06/02 Q1025 NOSIG" ;
 weather:hasVisibilityEvent metar:BrokenCloud ;
 weather:hasWindEvent [weather:windDirection "180"^^xsd:short] ,
                               [weather:windSpeedKnots "07"^^xsd:short] ;
 weather:hasTemperatureEvent
                [a weather:CurrentTemperature ; weather:celsiusTemperature "6"^^xsd:short] ,
             [a weather:CurrentDewPoint ; weather:celsiusTemperature "2"^^xsd:short] ,
             [weather:humidityPercent "75"^^xsd:short ;
 weather:hasPressureEvent [met:hectoPascal "1025"^^xsd:integer] ] .
```

Geonames Airports

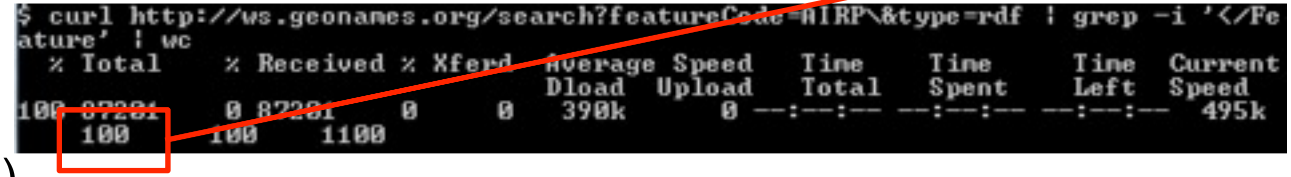


```
$ curl http://ws.geonames.org/search?featureCode=AIRP\&type=rdf
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<rdf:RDF xmlns="http://www.geonames.org/ontology#" xmlns:cc="http://creativecomm
ons.org/ns#" xmlns:dcterms="http://purl.org/dc/terms/" xmlns:foaf="http://xmlns.
com/foaf/0.1/" xmlns:owl="http://www.w3.org/2002/07/owl#" xmlns:rdf="http://www.
w3.org/1999/02/22-rdf-syntax-ns#" xmlns:wgs84_pos="http://www.w3.org/2003/01/geo
/wgs84_pos#">
<Feature rdf:about="http://sws.geonames.org/588404/">
<name>Tallinn Airport</name>
<alternateName xml:lang="ru">?????????? ????????</alternateName>
<alternateName xml:lang="iata">TLL</alternateName>
<alternateName xml:lang="en">Tallinn Airport</alternateName>
<alternateName xml:lang="et">Tallinna lennujaam</alternateName>
<alternateName xml:lang="icao">EETN</alternateName>
<featureClass rdf:resource="http://www.geonames.org/ontology#S"/>
<featureCode rdf:resource="http://www.geonames.org/ontology#S.AIRP"/>
<inCountry rdf:resource="http://www.geonames.org/countries/#EE"/>
<countryCode>EE</countryCode>
<wgs84_pos:lat>59.4165050845691</wgs84_pos:lat>
<wgs84_pos:long>24.7994041442871</wgs84_pos:long>
<nearbyFeatures rdf:resource="http://sws.geonames.org/588404/nearby.rdf"/>
<locationMap rdf:resource="http://www.geonames.org/588404/tallinn-airport.html"/
>
</Feature>
<Feature rdf:about="http://sws.geonames.org/1619461/">
```

‘alternateName’ for both labels and identification schemes (ICAO, IATA)

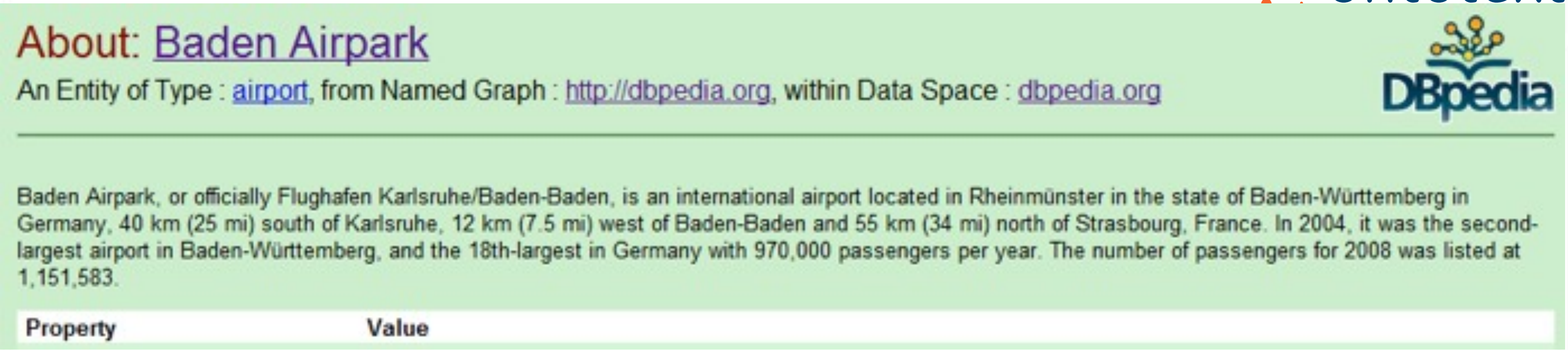
Only ~100 resources

WGS84 geospatial, but only lat/long (not alt)



```
$ curl http://ws.geonames.org/search?featureCode=AIRP\&type=rdf | grep -i '</Fe
ature' | wc
  % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
                                 Dload  Upload   Total   Spent    Left  Speed
100 87201    0 87201    0     0   390k      0 --:--:-- --:--:-- --:--:--  495k
    100     100    1100
```

DBPedia Airports

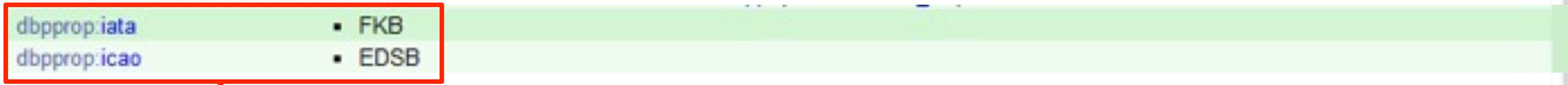


About: Baden Airpark

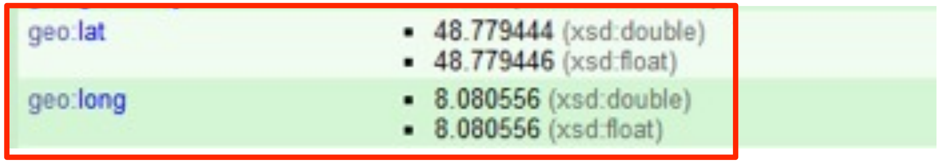
An Entity of Type : airport, from Named Graph : http://dbpedia.org, within Data Space : dbpedia.org

Baden Airpark, or officially Flughafen Karlsruhe/Baden-Baden, is an international airport located in Rheinmünster in the state of Baden-Württemberg in Germany, 40 km (25 mi) south of Karlsruhe, 12 km (7.5 mi) west of Baden-Baden and 55 km (34 mi) north of Strasbourg, France. In 2004, it was the second-largest airport in Baden-Württemberg, and the 18th-largest in Germany with 970,000 passengers per year. The number of passengers for 2008 was listed at 1,151,583.

Property	Value



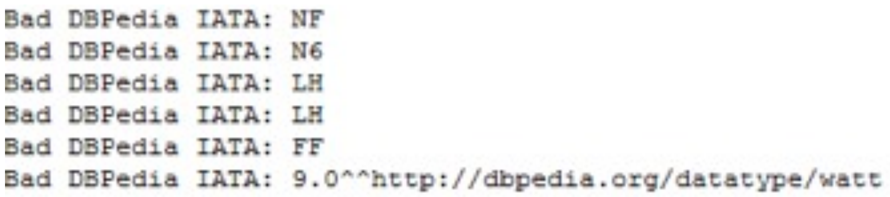
dbpprop:iata	FKB
dbpprop:icao	EDSB



geo:lat	48.779444 (xsd:double) 48.779446 (xsd:float)
geo:long	8.080556 (xsd:double) 8.080556 (xsd:float)

Some ICAO, IATA (with specific properties) but noisy

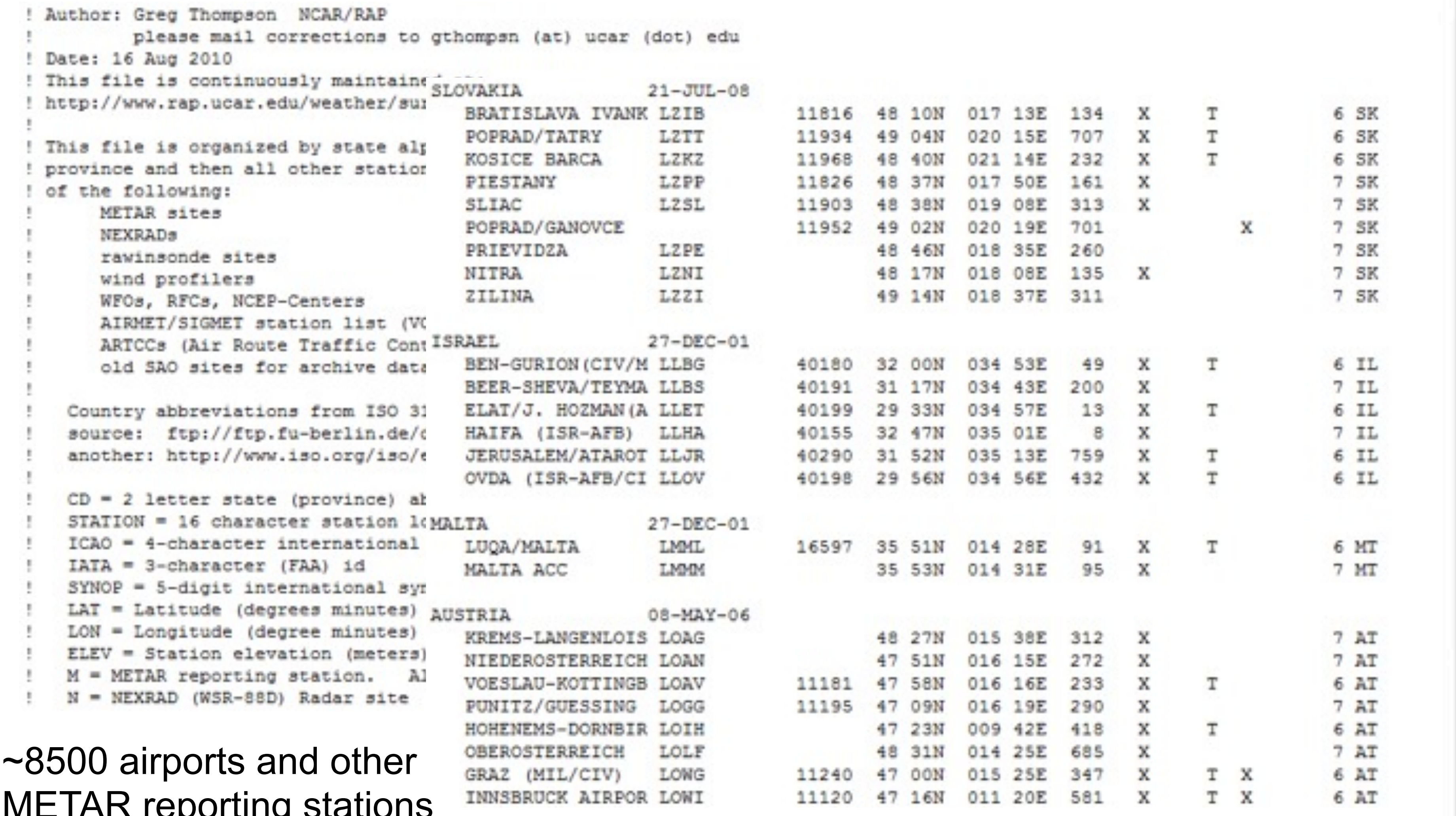
Some WGS84, but not consistently



```
Bad DBPedia IATA: NF
Bad DBPedia IATA: N6
Bad DBPedia IATA: LH
Bad DBPedia IATA: LH
Bad DBPedia IATA: FF
Bad DBPedia IATA: 9.0^^http://dbpedia.org/datatype/watt
```

~3500 verifiable airports (see next)

NCAR METAR Station List

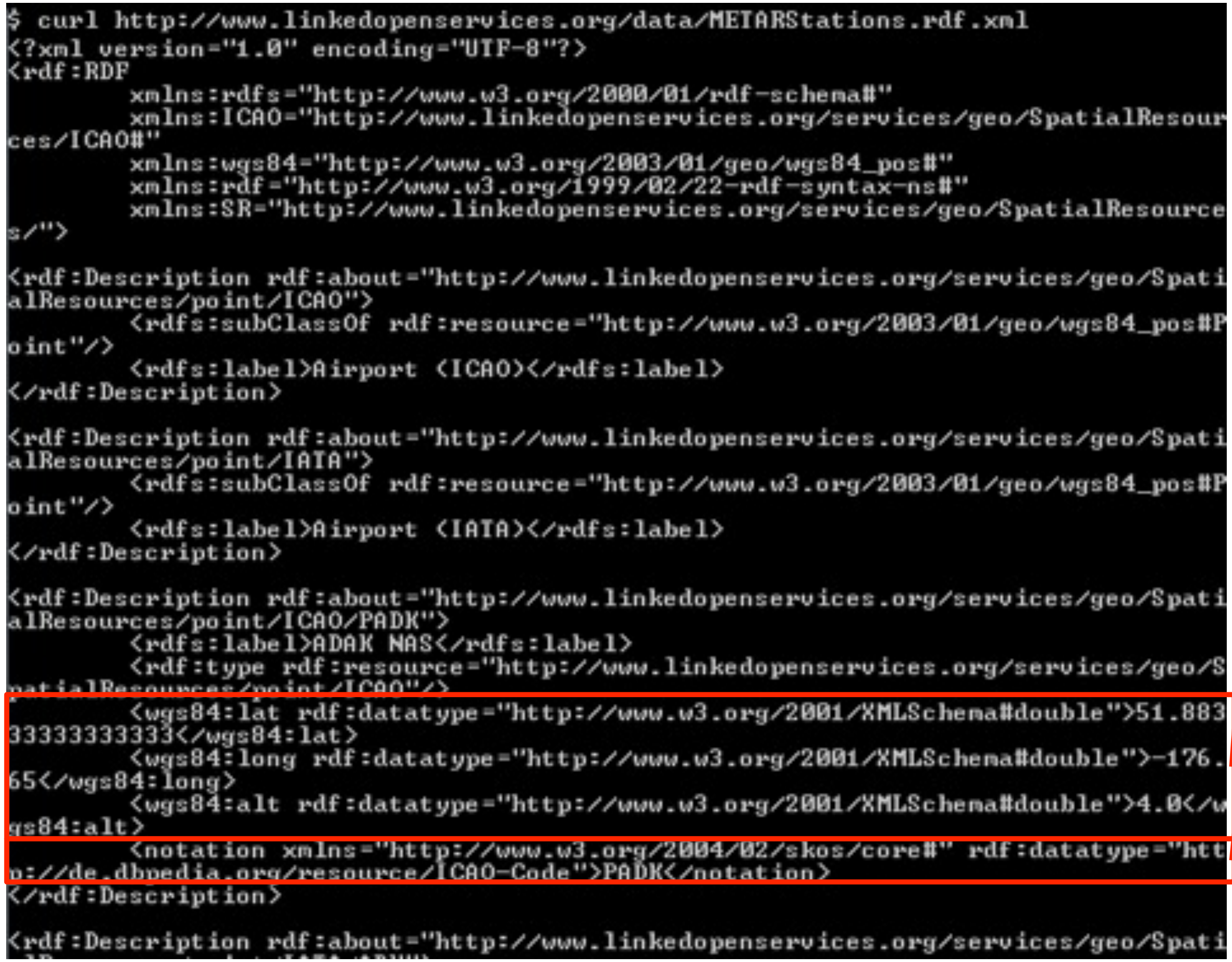


```
! Author: Greg Thompson  NCAR/RAP
!         please mail corrections to gthompsn (at) ucar (dot) edu
! Date: 16 Aug 2010
! This file is continuously maintaine
! http://www.rap.ucar.edu/weather/sur
!
! This file is organized by state alp
! province and then all other statior
! of the following:
!      METAR sites
!      NEXRADs
!      rawinsonde sites
!      wind profilers
!      WFOs, RFCs, NCEP-Centers
!      AIRMET/SIGMET station list (VC
!      ARTCCs (Air Route Traffic Cont
!      old SAO sites for archive data
!
!   Country abbreviations from ISO 31
!   source:  ftp://ftp.fu-berlin.de/c
!   another: http://www.iso.org/iso/e
!
!   CD = 2 letter state (province) ab
!   STATION = 16 character station lo
!   ICAO = 4-character international
!   IATA = 3-character (FAA) id
!   SYNOP = 5-digit international syr
!   LAT = Latitude (degrees minutes)
!   LON = Longitude (degree minutes)
!   ELEV = Station elevation (meters)
!   M = METAR reporting station.   Al
!   N = NEXRAD (WSR-88D) Radar site
```

```
SLOVAKIA             21-JUL-08
   BRATISLAVA IVANK LZIB          11816  48 10N  017 13E  134   X     T              6 SK
   POPRAD/TATRY     LZTT          11934  49 04N  020 15E  707   X     T              6 SK
   KOSICE BARCA     LZKZ          11968  48 40N  021 14E  232   X     T              6 SK
   PIESTANY         LZPP          11826  48 37N  017 50E  161   X                    7 SK
   SLIAC            LZSL          11903  48 38N  019 08E  313   X                    7 SK
   POPRAD/GANOVCE                 11952  49 02N  020 19E  701           X            7 SK
   PRIEVIDZA        LZPE                 48 46N  018 35E  260                        7 SK
   NITRA            LZNI                 48 17N  018 08E  135   X                    7 SK
   ZILINA           LZZI                 49 14N  018 37E  311                        7 SK

ISRAEL               27-DEC-01
   BEN-GURION(CIV/M LLBG          40180  32 00N  034 53E   49   X     T              6 IL
   BEER-SHEVA/TEYMA LLBS          40191  31 17N  034 43E  200   X                    7 IL
   ELAT/J. HOZMAN(A LLET          40199  29 33N  034 57E   13   X     T              6 IL
   HAIFA (ISR-AFB)  LLHA          40155  32 47N  035 01E    8   X                    7 IL
   JERUSALEM/ATAROT LLJR          40290  31 52N  035 13E  759   X     T              6 IL
   OVDA (ISR-AFB/CI LLOV          40198  29 56N  034 56E  432   X     T              6 IL

MALTA                27-DEC-01
   LUQA/MALTA       LMML          16597  35 51N  014 28E   91   X     T              6 MT
   MALTA ACC        LMMM                 35 53N  014 31E   95   X                    7 MT

AUSTRIA              08-MAY-06
   KREMS-LANGENLOIS LOAG                 48 27N  015 38E  312   X                    7 AT
   NIEDEROSTERREICH LOAN                 47 51N  016 15E  272   X                    7 AT
   VOESLAU-KOTTINGB LOAV          11181  47 58N  016 16E  233   X     T              6 AT
   PUNITZ/GUESSING  LOGG          11195  47 09N  016 19E  290   X                    7 AT
   HOHENEMS-DORNBIR LOIH                 47 23N  009 42E  418   X     T              6 AT
   OBEROSTERREICH   LOLF                 48 31N  014 25E  685   X                    7 AT
   GRAZ (MIL/CIV)   LONG          11240  47 00N  015 25E  347   X     T X            6 AT
   INNSBRUCK AIRPOR LOWI          11120  47 16N  011 20E  581   X     T X            6 AT
```

~8500 airports and other METAR reporting stations

LOS METAR Data



```
$ curl http://www.linkedopenservices.org/data/METARStations.rdf.xml
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF
        xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
        xmlns:ICAO="http://www.linkedopenservices.org/services/geo/SpatialResour
ces/ICAO#"
        xmlns:wgs84="http://www.w3.org/2003/01/geo/wgs84_pos#"
        xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
        xmlns:SR="http://www.linkedopenservices.org/services/geo/SpatialResource
s/">

<rdf:Description rdf:about="http://www.linkedopenservices.org/services/geo/Spati
alResources/point/ICAO">
        <rdfs:subClassOf rdf:resource="http://www.w3.org/2003/01/geo/wgs84_pos#P
oint"/>
        <rdfs:label>Airport (ICAO)</rdfs:label>
</rdf:Description>

<rdf:Description rdf:about="http://www.linkedopenservices.org/services/geo/Spati
alResources/point/IATA">
        <rdfs:subClassOf rdf:resource="http://www.w3.org/2003/01/geo/wgs84_pos#P
oint"/>
        <rdfs:label>Airport (IATA)</rdfs:label>
</rdf:Description>

<rdf:Description rdf:about="http://www.linkedopenservices.org/services/geo/Spati
alResources/point/ICAO/PADK">
        <rdfs:label>ADAK NAS</rdfs:label>
        <rdf:type rdf:resource="http://www.linkedopenservices.org/services/geo/S
patialResources/point/ICAO"/>
        <wgs84:lat rdf:datatype="http://www.w3.org/2001/XMLSchema#double">51.883
3333333333</wgs84:lat>
        <wgs84:long rdf:datatype="http://www.w3.org/2001/XMLSchema#double">-176.
65</wgs84:long>
        <wgs84:alt rdf:datatype="http://www.w3.org/2001/XMLSchema#double">4.0</w
gs84:alt>
        <notation xmlns="http://www.w3.org/2004/02/skos/core#" rdf:datatype="htt
p://de.dbpedia.org/resource/ICAO-Code">PADK</notation>
</rdf:Description>

<rdf:Description rdf:about="http://www.linkedopenservices.org/services/geo/Spati
```

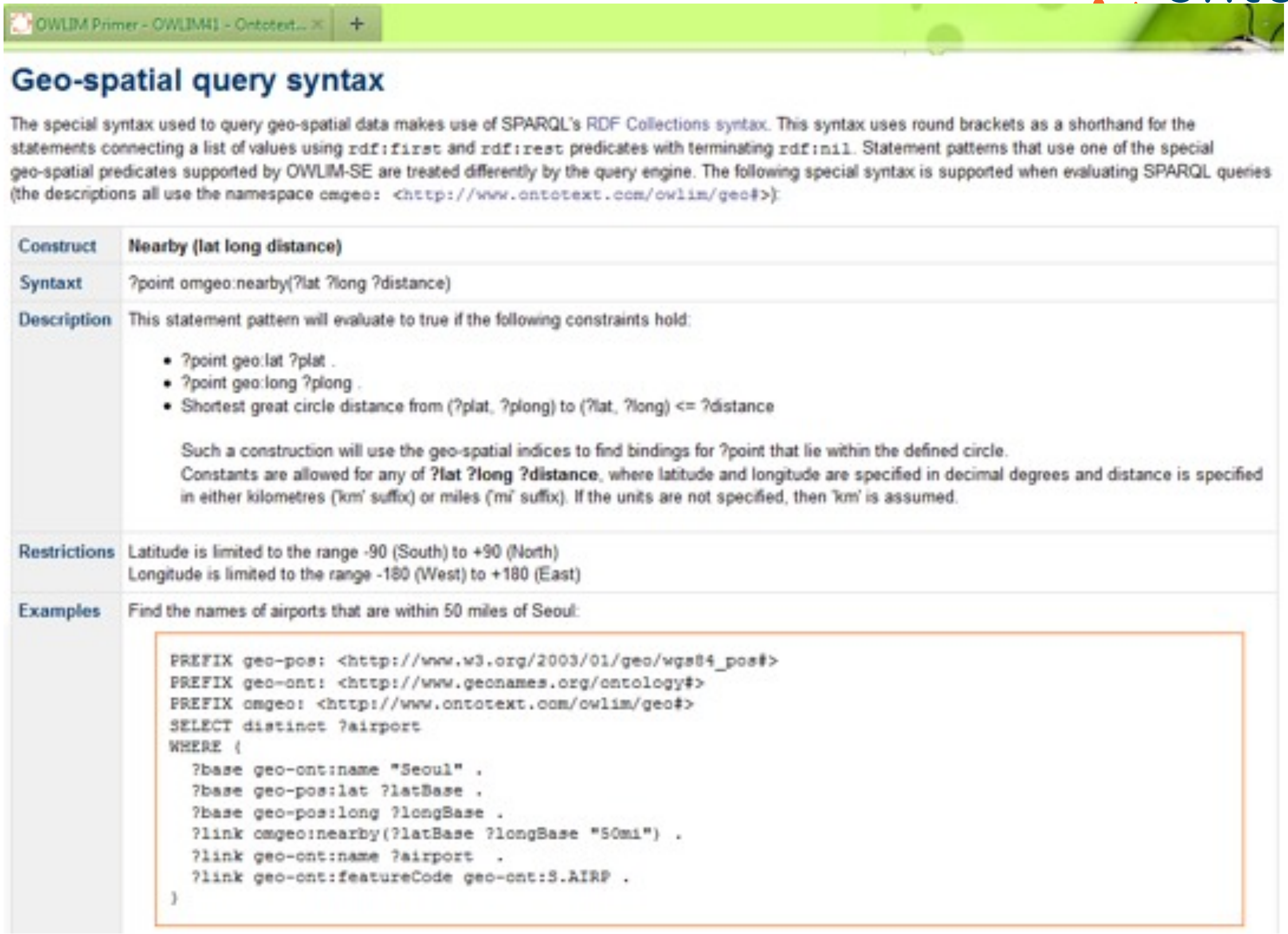
Uniform WGS84 lat/long/alt

ICAO and IATA-based URIs plus skos:notation to originals

owl:sameAs
- between ICAO/IATA
- to Geonames (108)
- to DBPedia (1159)

http://www.linkedopenservices.org/data/METARStations.rdf.xml

OWLIM Geospatial Extensions



Geo-spatial query syntax

The special syntax used to query geo-spatial data makes use of SPARQL's RDF Collections syntax. This syntax uses round brackets as a shorthand for the statements connecting a list of values using `rdf:first` and `rdf:rest` predicates with terminating `rdf:nil`. Statement patterns that use one of the special geo-spatial predicates supported by OWLIM-SE are treated differently by the query engine. The following special syntax is supported when evaluating SPARQL queries (the descriptions all use the namespace `omgeo: <http://www.ontotext.com/owlim/geo#>`):

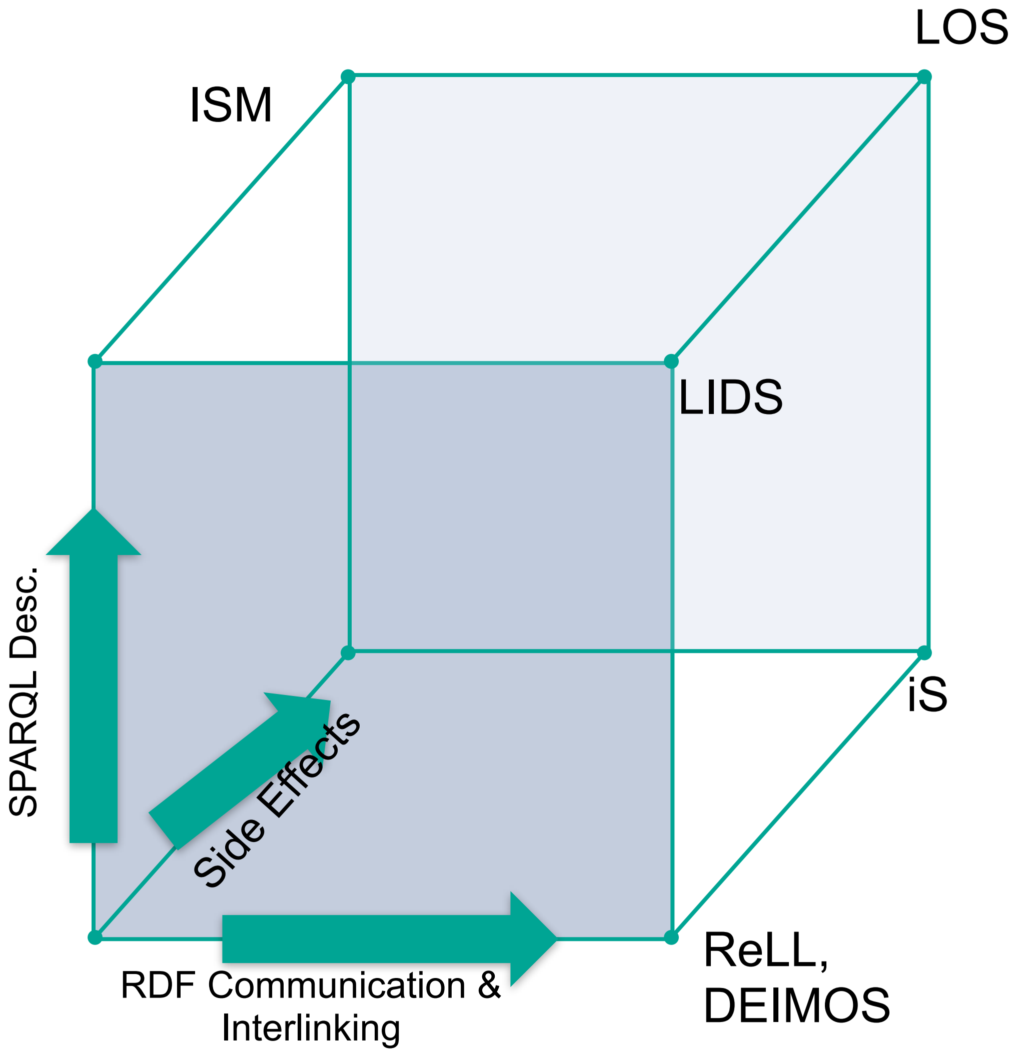
Construct	**Nearby (lat long distance)**
Syntaxt	?point omgeo:nearby(?lat ?long ?distance)
Description	This statement pattern will evaluate to true if the following constraints hold: • ?point geo:lat ?plat . • ?point geo:long ?plong . • Shortest great circle distance from (?plat, ?plong) to (?lat, ?long) <= ?distance Such a construction will use the geo-spatial indices to find bindings for ?point that lie within the defined circle. Constants are allowed for any of **?lat ?long ?distance**, where latitude and longitude are specified in decimal degrees and distance is specified in either kilometres ('km' suffix) or miles ('mi' suffix). If the units are not specified, then 'km' is assumed.
Restrictions	Latitude is limited to the range -90 (South) to +90 (North) Longitude is limited to the range -180 (West) to +180 (East)
Examples	Find the names of airports that are within 50 miles of Seoul:

```
PREFIX geo-pos: <http://www.w3.org/2003/01/geo/wgs84_pos#>
PREFIX geo-ont: <http://www.geonames.org/ontology#>
PREFIX omgeo: <http://www.ontotext.com/owlim/geo#>
SELECT distinct ?airport
WHERE {
  ?base geo-ont:name "Seoul" .
  ?base geo-pos:lat ?latBase .
  ?base geo-pos:long ?longBase .
  ?link omgeo:nearby(?latBase ?longBase "50mi") .
  ?link geo-ont:name ?airport  .
  ?link geo-ont:featureCode geo-ont:S.AIRP .
}
```

Related Activities

- Apache Incubator Clerezza
  - Components (bundles) for building RESTful Semantic Web applications and services leveraging Jax-RS
- Linked Services / iServe
  - A platform for publishing and browsing Semantic Web services as linked data, based on OWLIM
- Linked Data Services (LIDS)
  - An interface spec for data-services, accompanied by a machine-interpretable description of inputs and outputs.
- Iqbal, Sdobio & Moulin
  - Pre and post-conditions, as well as user goals in SPARQL

Related Activities



- ReLL = Resource Linking Language, Wilde (Berkeley)
- DEIMOS, Ambite et al. (USC)
- iS = iServe, Pedrinaci et al. (OU)
- ISM = Iqbal, Sbodio & Moulin (Galway)
- LIDS = Linked Data Services, Speiser & Harth (AIFB)
- LOS = Linked Open Services (Innsbruck -> AIFB)