

Semantic Web Technologies

Topic: RDF and Linked Data

Olaf Hartig

olaf.hartig@liu.se

Overview

- Basic concepts of RDF
- Formats for serializing RDF data
- Advanced features of RDF
- What is Linked Data?



RDF: Basic Concepts

RDF in General

- Resource Description Framework
- A **resource** may basically be everything
 - e.g. persons, places, Web documents, abstract concepts
- **Descriptions** of resources
 - Attributes
 - Relationships
- The **framework** contains:
 - A data model, and
 - Languages and syntaxes

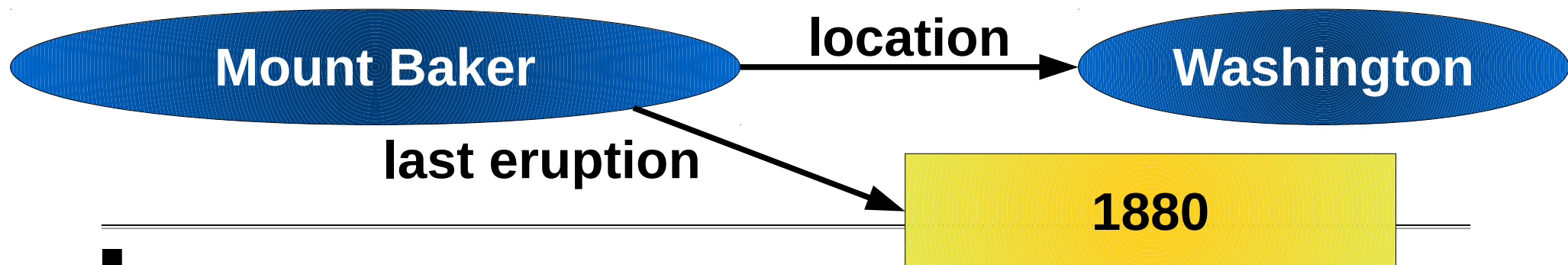


RDF Data Model

- Data comes as a set of **triples** (subject, predicate, object)
- **Subject**: resources
- **Predicate**: properties
- **Object**: literals or resources
- **Examples**:
 - (Mount Baker , last eruption , 1880)
 - (Mount Baker , location , Washington)

RDF Data Model (cont'd)

- RDF based data may be understood as a **graph**:
 - Triples as directed edges
 - Subjects and objects as vertices
 - Edges labeled by predicate
- **Example:**
 - (Mount Baker , last eruption , 1880)
 - (Mount Baker , location , Washington)

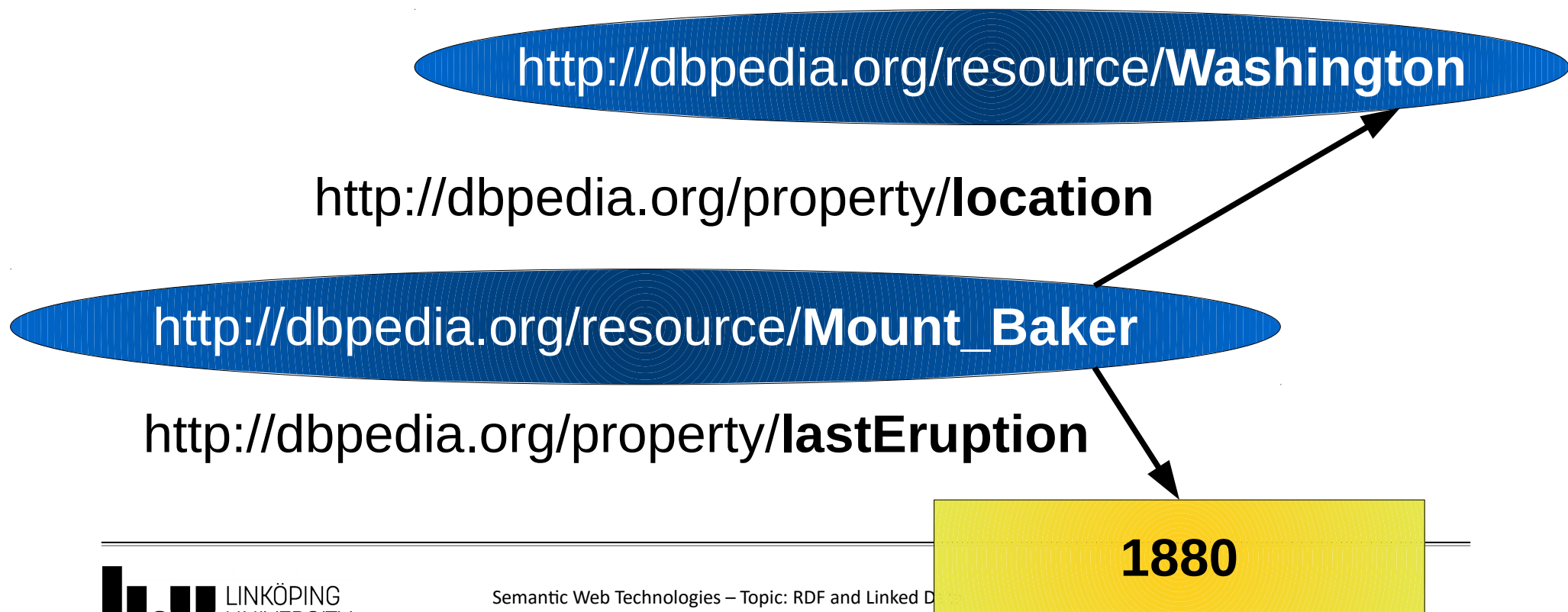


Uniform Resource Identifier (URI)

- URIs extend the concept of URLs
 - Globally **unique identifier** for resources
 - URL of a Web document usually used as its URI
 - Attention: URIs identify not only Web documents
- Example:
 - Me: <http://olafhartig.de/foaf.rdf#olaf>
 - RDF document about me:
<http://olafhartig.de/foaf.rdf>
 - HTML document about me:
<http://olafhartig.de/index.html>

Example Revisited

- (http://dbpedia.org/resource/Mount_Baker,
<http://dbpedia.org/property/lastEruption>, 1880)
- (http://dbpedia.org/resource/Mount_Baker,
<http://dbpedia.org/property/location>,
<http://dbpedia.org/resource/Washington>)

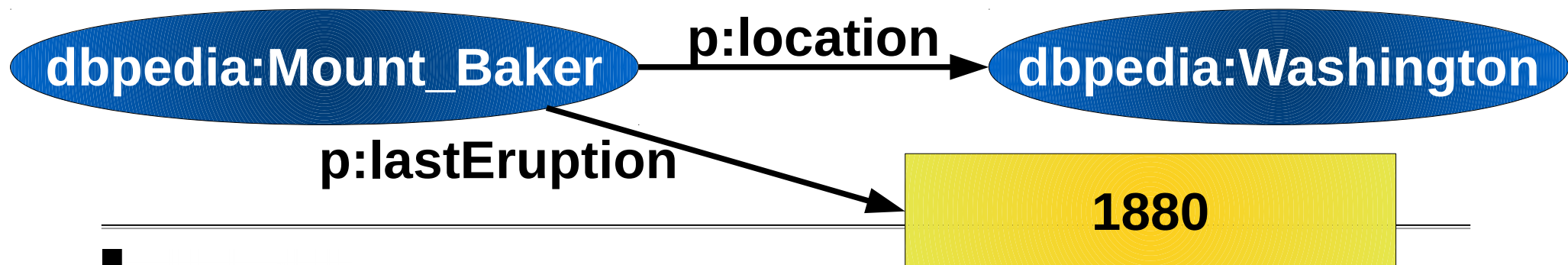


Compact URI (CURIE)

- Abbreviated Notation for URIs
- Syntax:
 - Prefix name (references a prefix of the URI)
 - Colon character (":")
 - Reference part
- URI by **concatenating** the prefix and the reference part
- Examples:
 - `dbpedia:Mount_Baker` for
`http://dbpedia.org/resource/Mount_Baker`
 - `myfoaf:olaf` for
`http://olafhartig.de/foaf.rdf#olaf`

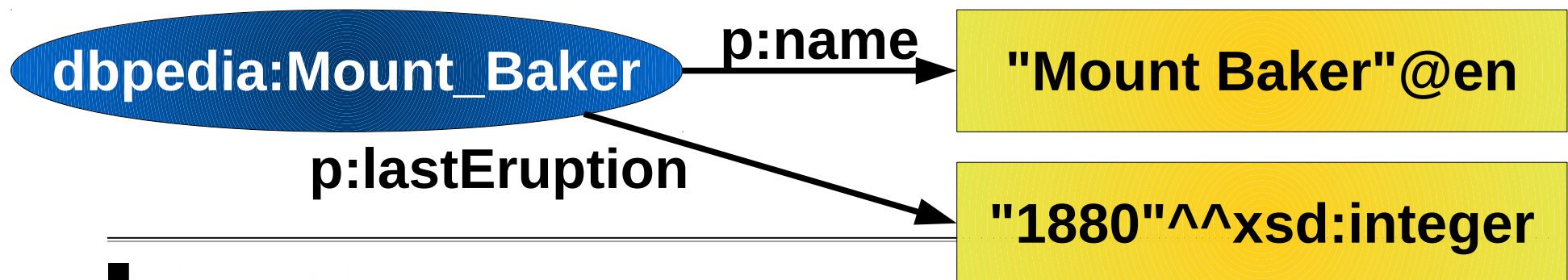
Example with CURIEs

- Using
 - *dbpedia* for prefix *http://dbpedia.org/resource/*
 - *p* for prefix *http://dbpedia.org/property/*
- we have
 - (dbpedia:Mount_Baker, p:lastEruption, 1880)
 - (dbpedia:Mount_Baker, p:location, dbpedia:Washington)



Literals

- Literals may occur in the **object position** of RDF triples
- Represented by strings
- Literal strings interpreted by **datatypes**
 - Datatype identified by a URI
 - Common to use the XML Schema datatypes
 - If no datatype, then interpreted as xsd:string
- Untyped literals may have **language tags** (e.g. @de)



RDF Turtle

Turtle: A Human-Readable Syntax for RDF

- Simple, human-readable notation to list RDF triples:
 - Triples separated by a period (“.”) character
 - Example:

```
<http://dbpedia.org/resource/Mount_Baker>  
    <http://dbpedia.org/property/lastEruption>  
        "1880"^^xsd:integer .  
  
<http://dbpedia.org/resource/Mount_Baker>  
    <http://dbpedia.org/property/location>  
        <http://dbpedia.org/resource/Washington> .
```

CURIEs in Turtle

- `@prefix` directive binds a prefix to a namespace URI

```
@prefix dbpedia : <http://dbpedia.org/resource/> .
```

```
@prefix p : <http://dbpedia.org/property/> .
```

```
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
```

```
dbpedia:Mount_Baker p:lastEruption "1880"^^xsd:integer .
```

```
dbpedia:Mount_Baker p:location dbpedia:Washington .
```

```
dbpedia:Washington p:borderingstates dbpedia:Oregon .
```

```
dbpedia:Washington p:borderingstates dbpedia:Idaho .
```

Syntactic Sugar in Turtle

- Property lists separated by a semicolon (“;”) character
- Object lists separated by a comma (“,”) character

```
@prefix dbpedia : <http://dbpedia.org/resource/> .
```

```
@prefix p : <http://dbpedia.org/property/> .
```

```
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
```

```
dbpedia:Mount_Baker p:lastEruption "1880"^^xsd:integer ;  
                    p:location      dbpedia:Washington .
```

```
dbpedia:Washington p:borderingstates dbpedia:Oregon ,  
                    dbpedia:Idaho .
```

Some More Syntactic Sugar

- Shortcuts for number literals

```
dbpedia:Mount_Baker p:lastEruption "1880"^^xsd:integer ;  
                    geo:lat "48.777222"^^xsd:float ;  
                    geo:long "-121.813332"^^xsd:float .
```

Equivalent:

```
dbpedia:Mount_Baker p:lastEruption 1880 ;  
                    geo:lat 48.777222 ;  
                    geo:long -121.813332 .
```

Other Serialization Formats

N-Triples

```
@prefix dbpedia: <http://dbpedia.org/resource/> .  
@prefix p: <http://dbpedia.org/property/> .  
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .  
  
dbpedia:Mount_Baker p:lastEruption "1880"^^xsd:integer .  
dbpedia:Mount_Baker p:location dbpedia:Washington .
```

Turtle

```
<http://dbpedia.org/resource/Mount_Baker> <http://dbpedia.org/property/lastEruption> "1880"^^<http://www.w3.org/2001/XMLSchema#integer> .  
<http://dbpedia.org/resource/Mount_Baker> <http://dbpedia.org/property/location> <http://dbpedia.org/resource/Washington> .
```

N-Triples

- One line per triple

RDF/XML: An XML Syntax for RDF

```
@prefix dbpedia: <http://dbpedia.org/resource/> .
@prefix p: <http://dbpedia.org/property/> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .

dbpedia:Mount_Baker p:lastEruption "1880"^^xsd:integer .
dbpedia:Mount_Baker p:location      dbpedia:Washington .
```

Turtle

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:p="http://dbpedia.org/property/">
  <rdf:Description rdf:about="http://dbpedia.org/resource/Mount_Baker">
    <p:lastEruption ▶
      rdf:datatype="http://www.w3.org/2001/XMLSchema#integer" ▶
      >1880</p:lastEruption>
    <p:location rdf:resource="http://dbpedia.org/resource/Washington"/>
  </rdf:Description>
</rdf:RDF>
```

RDF/XML

RDFa: RDF inside HTML Documents

```
@prefix dbpedia: <http://dbpedia.org/resource/> .
@prefix p: <http://dbpedia.org/property/> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .

dbpedia:Mount_Baker p:lastEruption "1880"^^xsd:integer .
dbpedia:Mount_Baker p:location      dbpedia:Washington .
```

Turtle

```
<div xmlns="http://www.w3.org/1999/xhtml"
  prefix=" p: http://dbpedia.org/property/
         rdf: http://www.w3.org/1999/02/22-rdf-syntax-ns#
         xsd: http://www.w3.org/2001/XMLSchema#">
  <div about="http://dbpedia.org/resource/Mount_Baker">
    <div property="p:lastEruption" datatype="xsd:integer">1880</div>
    <div rel="p:location" resource="http://dbpedia.org/resource/Washington">
      </div>
    </div>
  </div>
```

HTML with RDFa Markup

JSON-LD

```
@prefix dbpedia: <http://dbpedia.org/resource/> .
@prefix p: <http://dbpedia.org/property/> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .

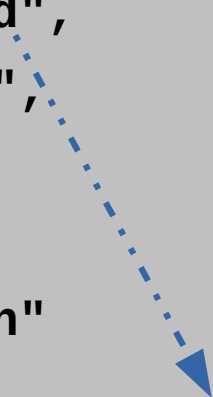
dbpedia:Mount_Baker p:lastEruption "1880"^^xsd:integer .
dbpedia:Mount_Baker p:location      dbpedia:Washington .
```

Turtle

```
{
  "@context": "http://example.org/mycontext.jsonld",
  "@id": "http://dbpedia.org/resource/Mount_Baker",
  "lastEruption": 1880,
  "location": {
    "@id": "http://dbpedia.org/resource/Washington"
  }
}
```

JSON-LD

```
{ "@context": {
  "lastEruption": "http://dbpedia.org/property/lastEruption",
  "location": "http://dbpedia.org/property/location" }
}
```

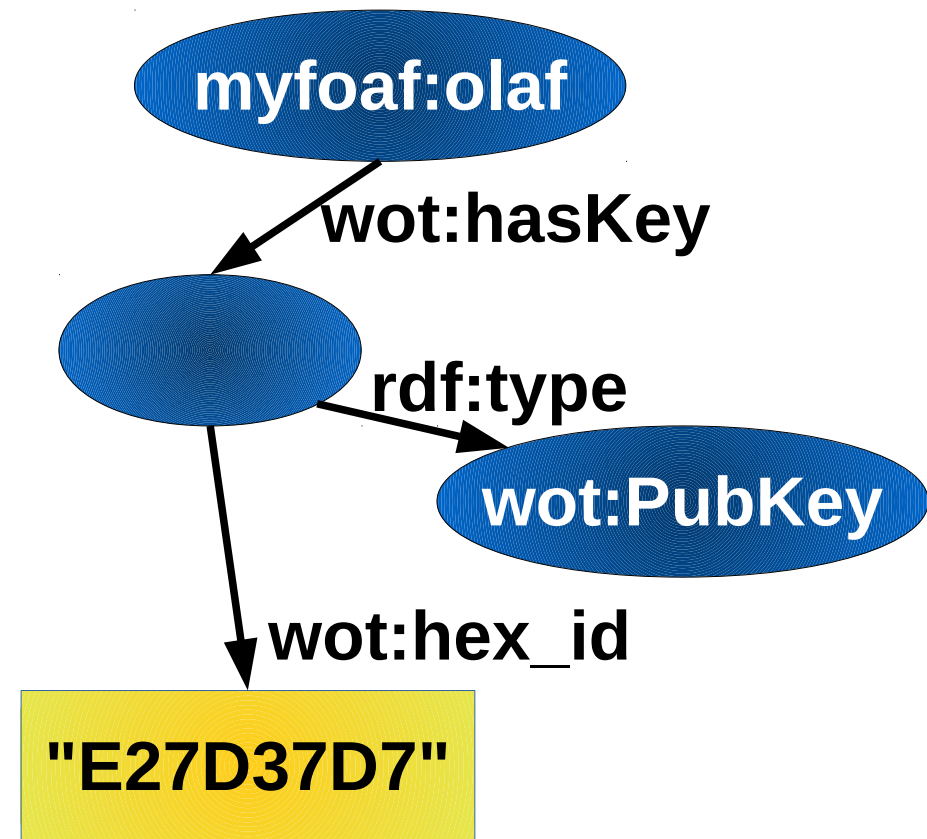


RDF: Advanced Features

Blank Nodes

Blank Nodes

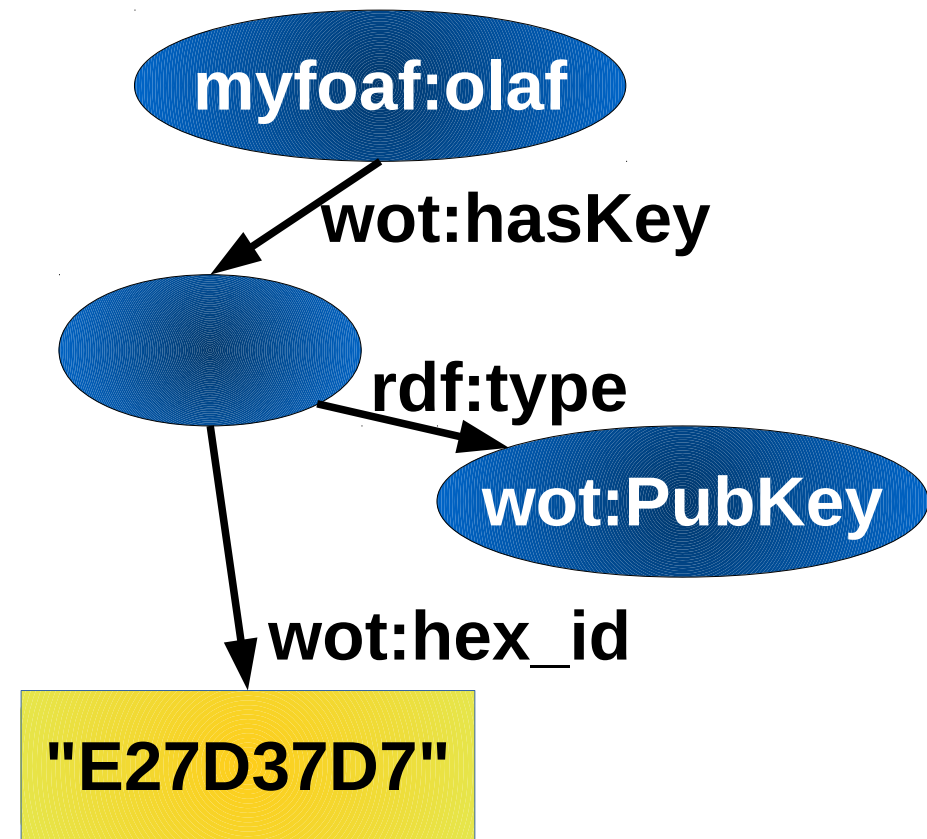
- Unnamed, anonymous nodes
- Resources for which we do not have a URI as identifier, because
 - We may not know the URI
 - There is no point in minting a URI for them



Blank Node Identifiers

- Needed to denote blank nodes when writing RDF to a file
- Typical form: `_:xyz`
- Example in Turtle:

```
myfoaf:olaf wot:hasKey _:x .  
_:x rdf:type    wot:PubKey ;  
    wot:hex_id  "E27D37D7" .
```

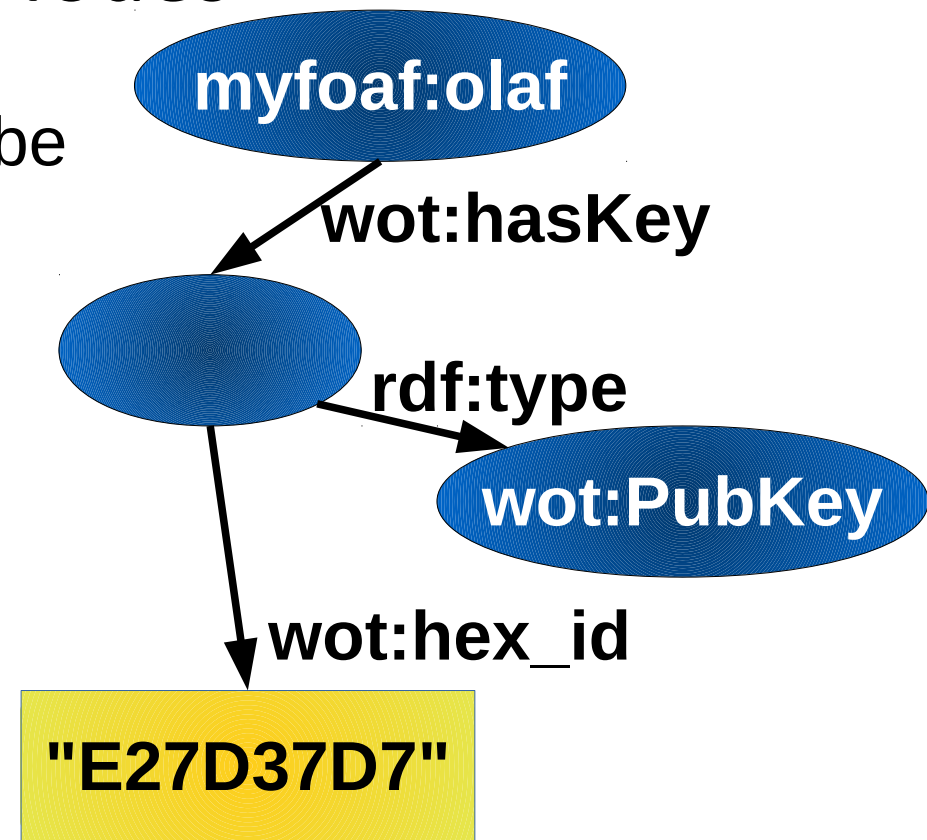


- Scope only within the serialization of a single RDF graph
- i.e., the same bnode identifier in two different serializations cannot be assumed to denote the same thing

Abbreviated Syntax for BNodes

- In some cases, blank nodes may be represented in Turtle without an explicit blank node identifier

```
myfoaf:olaf wot:hasKey _:x .  
_:x rdf:type    wot:PubKey ;  
    wot:hex_id  "E27D37D7" .
```



```
myfoaf:olaf wot:hasKey [ rdf:type    wot:PubKey ;  
                        wot:hex_id  "E27D37D7" ]
```

RDF: Advanced Features

Containers and Collections

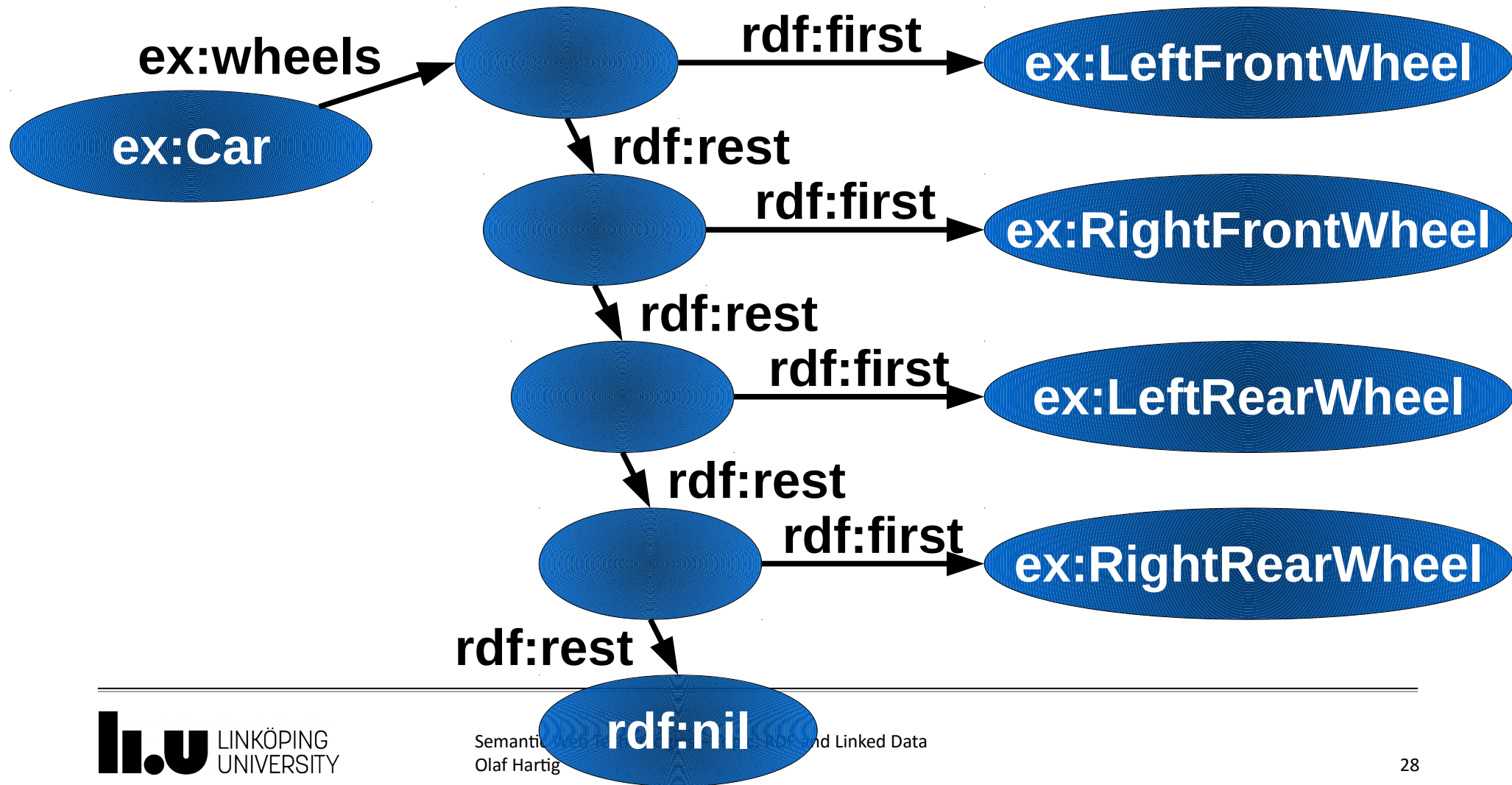
Containers

- Containers are an open group
- Contain resource or literals, possibly duplicates
 - `rdf:Seq` – ordered
 - `rdf:Bag` – unordered
 - `rdf:Alt` – for alternatives (typical processing will be to select one of the members)

```
dbpedia:Mount_Etna ex:eruptions [ rdf:type rdf:Bag ;  
                                   rdf:_1 "1669" ;  
                                   rdf:_2 "1949" ;  
                                   rdf:_3 "1971" ;  
                                   rdf:_4 "2001" ] .
```

Collections

- Closed list of resources or literals, possibly duplicates



Collections in Turtle

- Option 1:

```
ex:Car ex:wheels _:a ;  
_:a rdf:first ex:LeftFrontWheel ; rdf:rest _:b .  
_:b rdf:first ex:RightFrontWheel ; rdf:rest _:c .  
_:c rdf:first ex:LeftRearWheel ; rdf:rest _:d .  
_:d rdf:first ex:RightRearWheel ; rdf:rest rdf:nil .
```

- Option 2 (special support for collections in Turtle):

```
ex:Car ex:wheels ( ex:LeftFrontWheel ex:RightFrontWheel  
                  ex:LeftRearWheel ex:RightRearWheel ) .
```

RDF: Advanced Features

Named Graphs

Definitions

- **Named graph**: an RDF graph that is associated with a URI
- **RDF dataset**: one default graph + zero or more named graphs

```
dbpedia:Mount_Etna rdfs:seeAlso <http://example.org/d1> .  
dbpedia:Mount_Baker rdfs:seeAlso <http://example.org/d2> .
```

Default
Graph

```
dbpedia:Mount_Etna http://example.org/d1  
rdf:type umbel-sc:Volcano ;  
rdfs:label "Etna" .
```

```
dbpedia:Mount_Baker http://example.org/d2  
rdf:type umbel-sc:Volcano .
```

```
dbpedia:Beerenberg http://example.org/d3  
rdf:type umbel-sc:Volcano ;  
rdfs:label "Beerenberg"@en .
```

Use Cases

- Maintaining data from different sources separately
- Graph-level metadata (e.g., provenance information)
- Versioning
- Access control
- etc



Serialization Formats

- **TriG**: an extension of Turtle

```
dbpedia:Mount_Etna rdfs:seeAlso <http://example.org/d1>.  
dbpedia:Mount_Baker rdfs:seeAlso <http://example.org/d2>.
```

```
GRAPH <http://example.org/d1>  
{ dbpedia:Mount_Etna rdf:type umbel-sc:Volcano ;  
  rdfs:label "Etna" . }
```

```
GRAPH <http://example.org/d2>  
{ dbpedia:Mount_Baker rdf:type umbel-sc:Volcano . }
```

```
GRAPH <http://example.org/d3>  
{ dbpedia:Beerenberg rdf:type umbel-sc:Volcano ;  
  rdfs:label "Beerenberg"@en . }
```

Serialization Formats (cont'd)

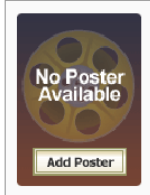
- **TriG**: an extension of Turtle
- **N-Quads**: like N-Triples, with a fourth element added
- **JSON-LD**: @graph keyword

Linked Data

Data on the Web



IMDb > War Child (1999) (TV)



War Child (1999) (TV) [More at IMDbPro »](#)

Overview

User Rating: ★★★★★★ (awaiting 5 votes)

Director: [Michael Davie](#)

Release Date: 12 July 1999 (USA) [more▶](#)

Genre: [Documentary](#) | [War](#) [more▶](#)

Tagline: Young people affected by the war in Kosovo

Plot Keywords: Spoiler alert! R

Additional Details

Parents Guide: [Add content advisory for parents](#)

Runtime: USA: 25 min

Country: [USA](#)

Language: [English](#)

Quicklinks

[main details](#)

Top Links

[Country List](#) | [World Factbook Home](#)



Location: Southeastern Europe, bordering the Adriatic Sea and Ionian Sea, between Greece in the south and Montenegro and Kosovo to the north

Geographic coordinates: 41 00 N, 20 00 E

Map references: Europe

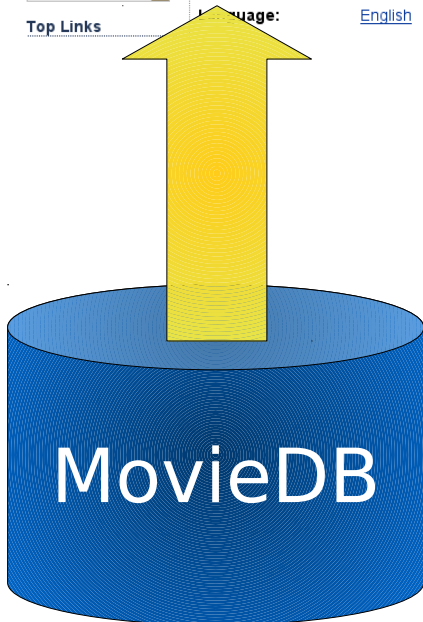
Area: total: 28,748 sq km
land: 27,398 sq km
water: 1,350 sq km

Area - comparative: slightly smaller than Maryland

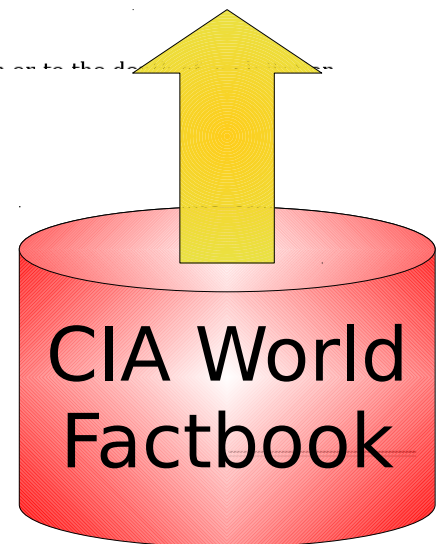
Land boundaries: total: 717 km
border countries: Greece 282 km, Macedonia 151 km, Montenegro 172 km, Kosovo 112 km

Coastline: 362 km

Maritime claims: territorial sea: 12 nm
continental shelf: 200 m depth or to the deep-sea

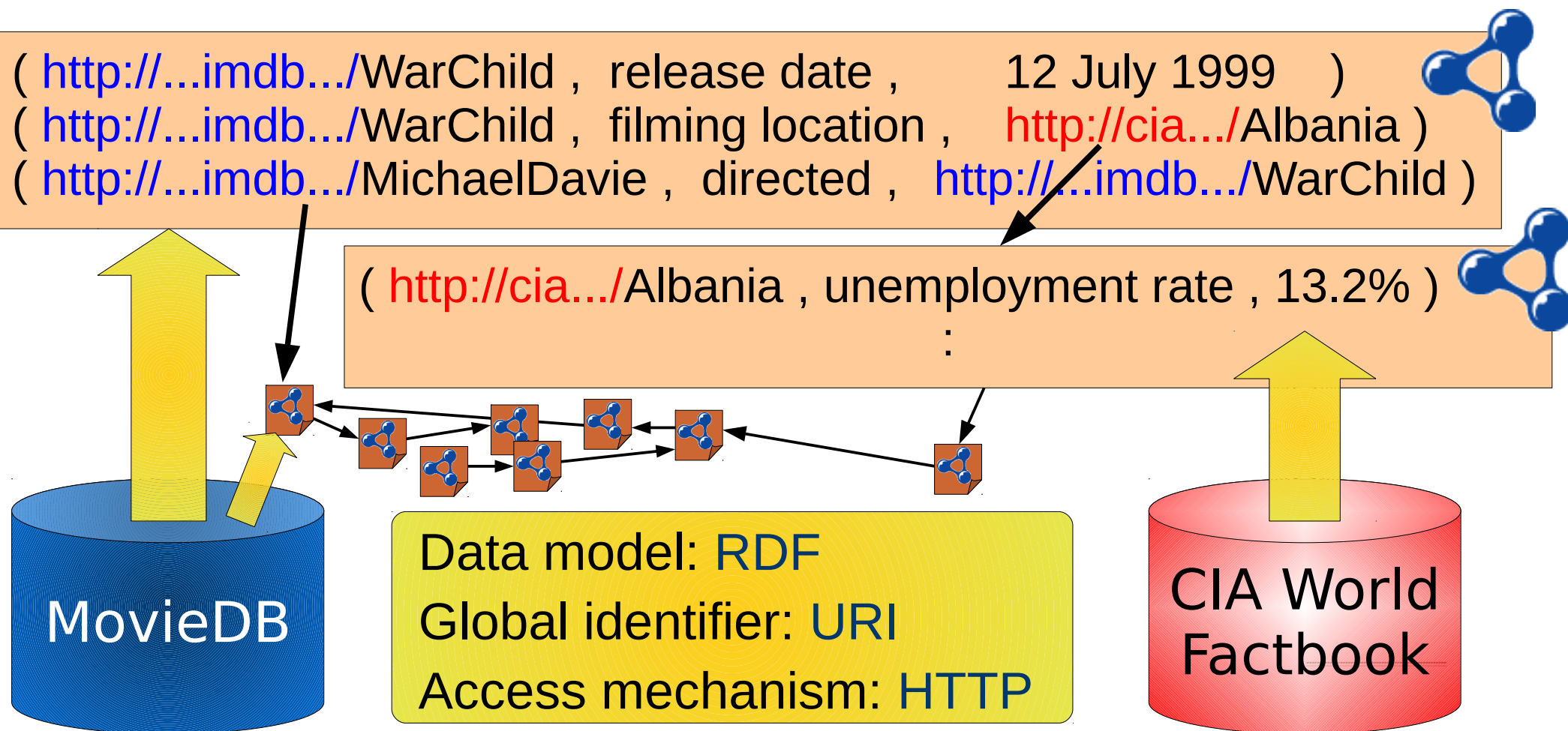


Traditionally, published in HTML documents that are designed for human consumption



Linked Data Publishing Principles

- Goal: publishing structured data on the WWW in a standardized, machine-readable manner



Adoption

- Started as a grassroots community effort in 2007
- Publish existing, open license datasets as Linked Data
- Interlink things between different data sources
- Prominent publishers joined the effort
 - e.g., BBC, NY Times, Library of Congress, Thomson Reuters, Springer, Nature, Best Buy, Sears, Renault, UK Government
- Numbers
 - BTC 2014 Crawl (February – June 2014):
ca. 4.1B triples in ca. 44M docs from ca. 48K sites
 - lodstats (as of September 24, 2017):
ca. 149B triples in ca. 3.0K datasets



www.liu.se