hbz

Wissen. Information. Innovation.

# An Introduction to Linked Open Data

Felix.Ostrowski@googlemail.com (@literarymachine)

Adrian Pohl@hbz-nrw.de (@acka47)

SWIB 2013 Pre-Conference Workshop

Monday, November 25th 2013

Hamburg

# Schedule

- Organize in teams
- Introduction: Data – Graphs – Triples
- Groupwork
- URIs and Namespaces
- Groupwork
- Open Data Principles
- Groupwork
- Identification vs. Description
- Groupwork
- Triple Stores & SPARQL
- Groupwork
- RDF Schema
- Groupwork
- Summary, Questions & Discussion

hbz

# Linked Open Data

- It's about **data** ...
- ... more precisely: about **open** data ...
- ... even more precisely: about **linked** open data!

hbz

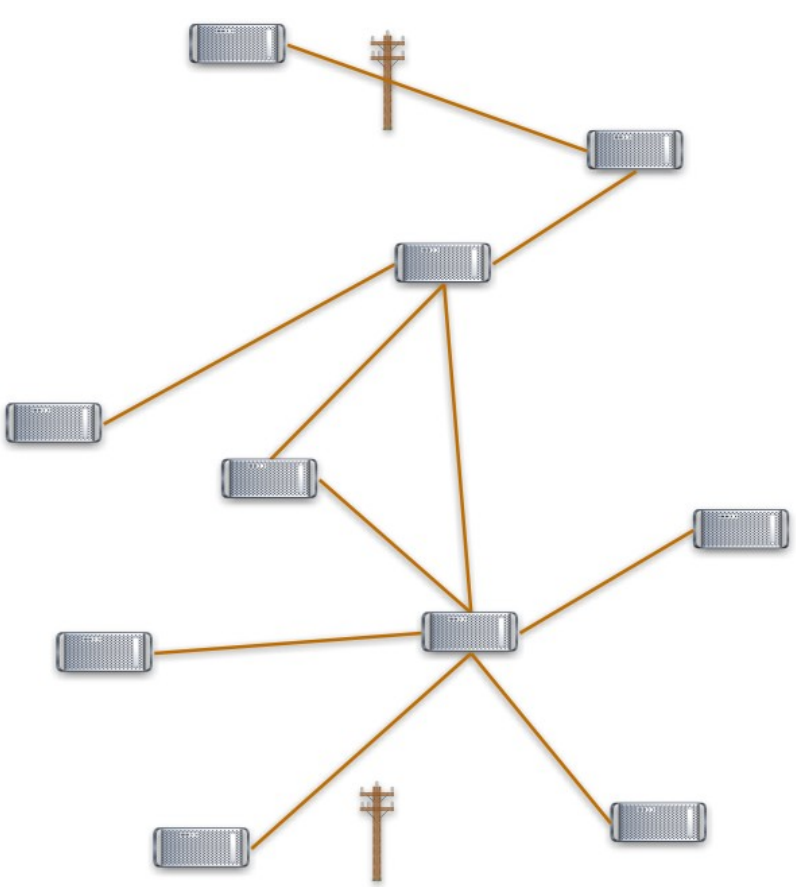
# **Data**, how we know it

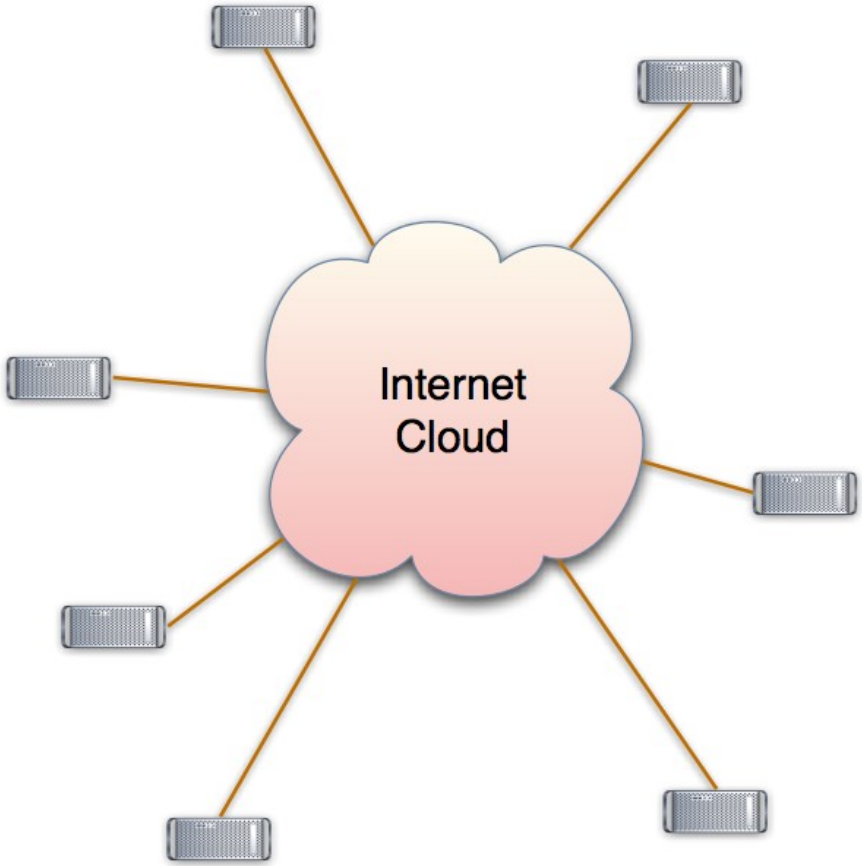

```
LDR             ------M2.01200024------h
FMT             MH
001             |a HT016905880
002a            |a 20110726
003             |a 20110729
026             |a HBZHT016905880
030             a|1uc||||||17
036a            |a NL
037b            |a eng
050             a|||||||||||||
051             m|||f|||
070             |a 294/61
070b            |a 361
080             |a 60
100             |a Allemang, Dean |9 136636187
104a            |a Hendler, James A. |9 115664564
331             |a Semantic web for the working ontologist
335             |a effective modeling in RDFS and OWL
359             |a Dean Allemang ; Jim Hendler
403             |a 2. ed.
410             |a Amsterdam [u.a.]
412             |a Elsevier MK
425a            |a 2011
433             |a XIII, 354 S. : graph. Darst.
540a            |a 978-0-12-385965-5
```

(To be honest, we might actually be the only ones knowing such data. And there aren't too many things that one can describe in this way.)

hbz

# Along came the Internet






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

hbz

# **Data**, how others know it

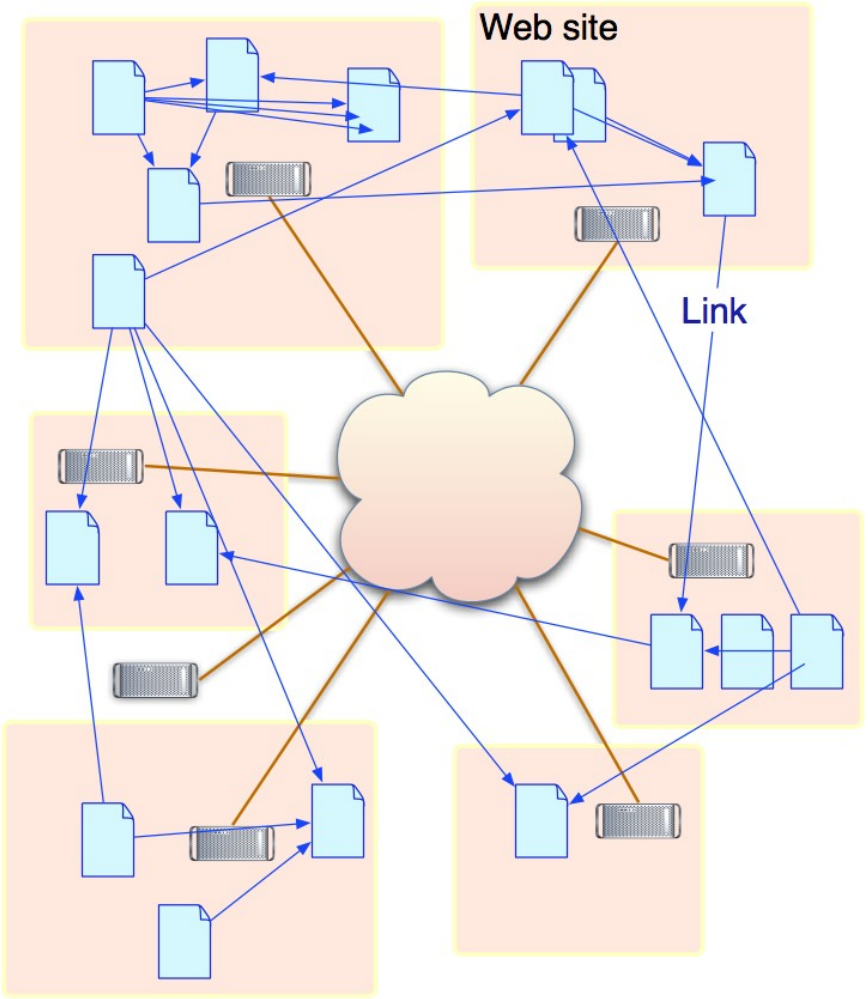
```
+-----------+-----------+----------+----------+
| id        | firstname | lastname | birthday |
+-----------+-----------+----------+----------+
| 136636187 | Dean      | Allemang | NULL     |
+-----------+-----------+----------+----------+

+-------------+-----------------------------------------+-----------+
| id          | title                                   | author    |
+-------------+-----------------------------------------+-----------+
| HT016905880 | Semantic web for the working ontologist | 136636187 |
+-------------+-----------------------------------------+-----------+
```

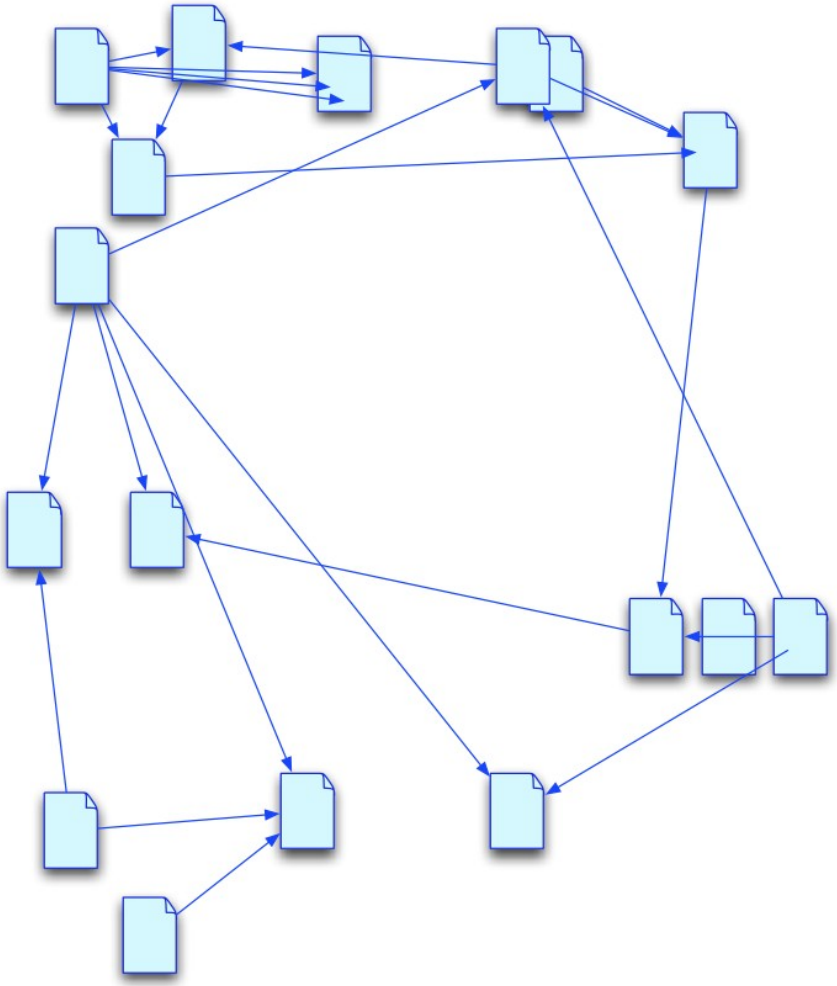
```
<book id="HT016905880">
  <title>Semantic web … </title>
  <author id="136636187">
    <firstname>Dean</firstname>
    <lastname>Allemang</lastname>
  </author>
</book>
```

(Of course, "others" does not mean "everybody". But at least you can describe many things this way. Maybe even everything.)

hbz

# The World Wide Web

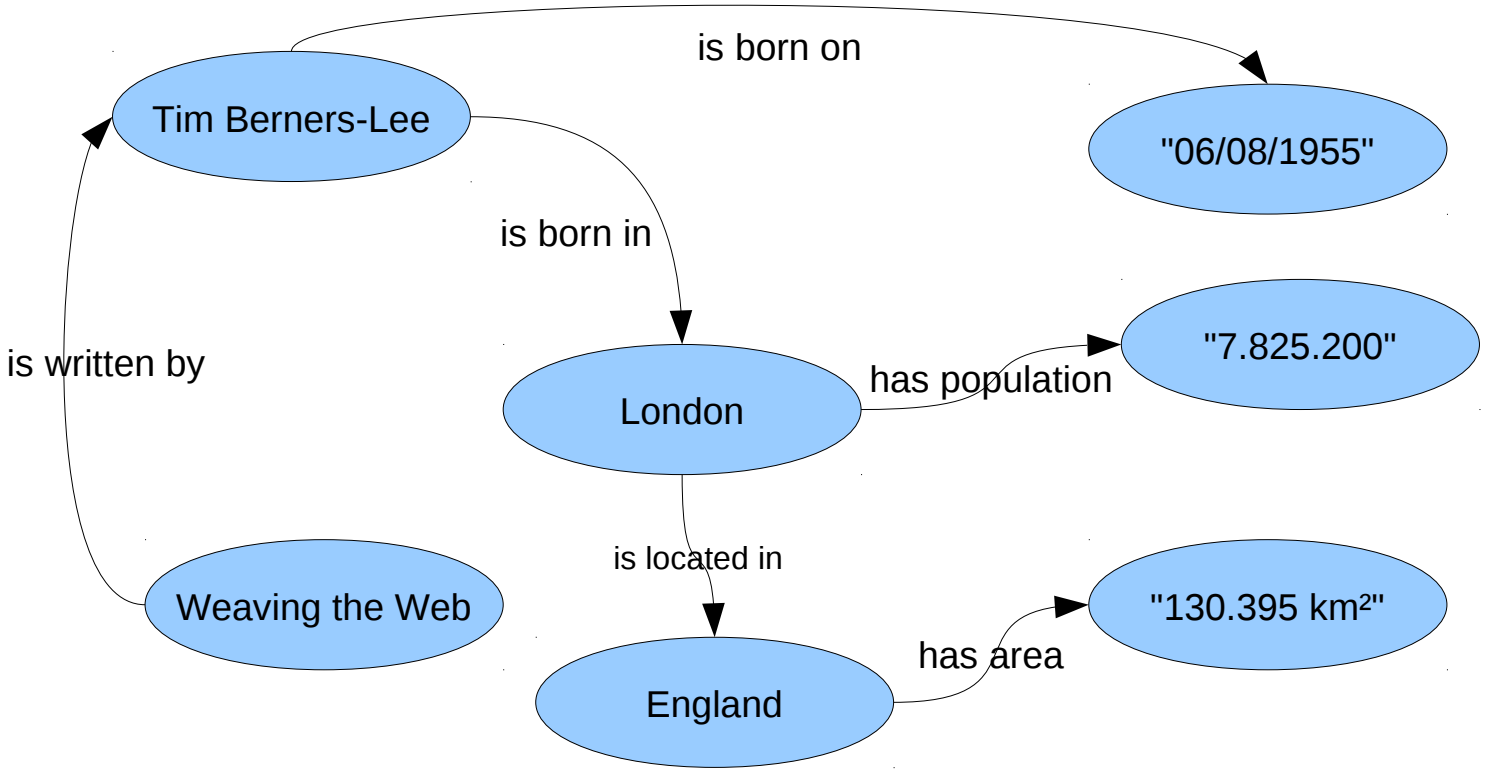


http://www.w3.org/DesignIssues/Abstractions.html

hbz

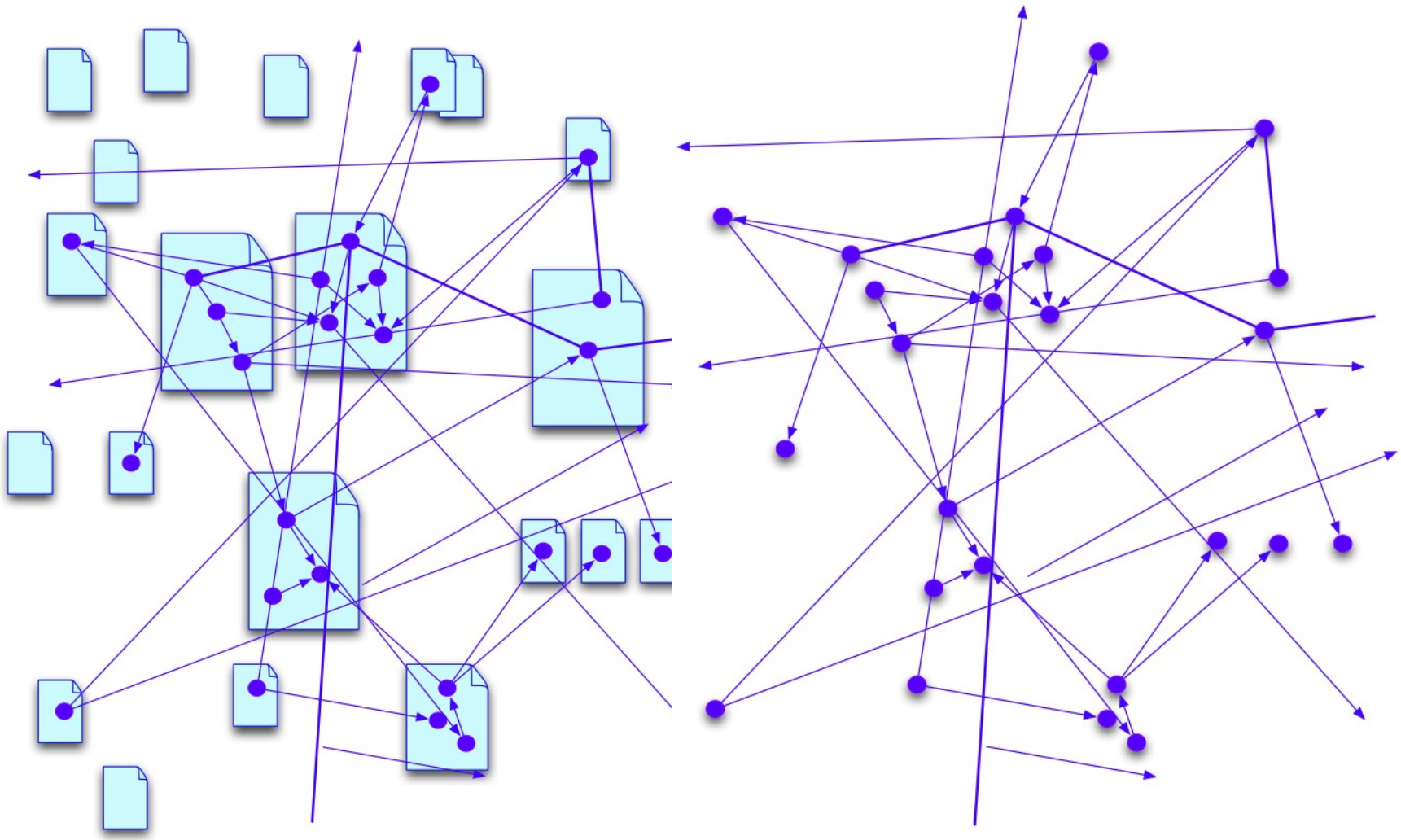
# **Data**, how the web likes it




(No wonder, it actually looks like a web. Or, if you will, a directed labelled **graph**.)

hbz

# The Giant Global Graph



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

hbz

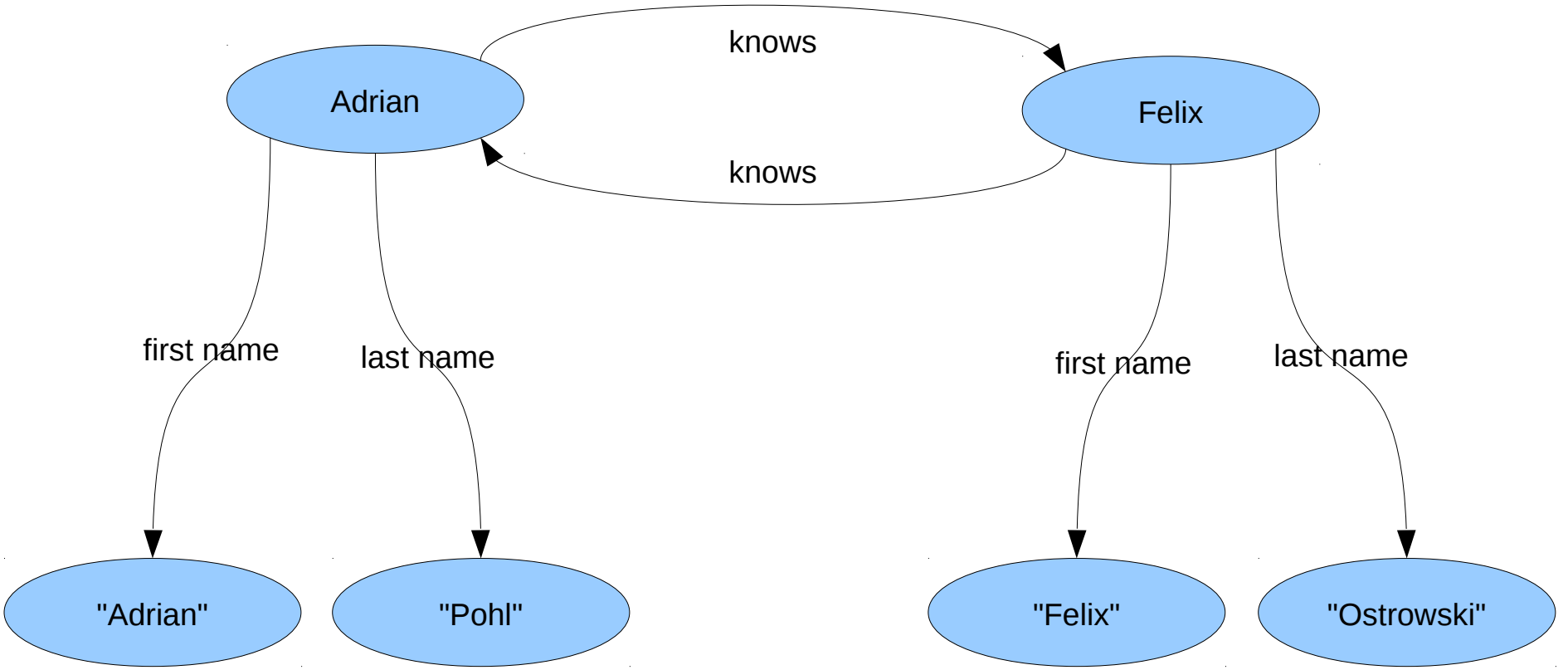
Your turn!

Draw a graph of your social network.

(For now, stick with the people on your table)

hbz

# A simple social **graph**




hbz

Obviosly a computer will have trouble interpreting such a diagram. The **graph data model** is an **abstract** one, but we can concrete it for the computer.

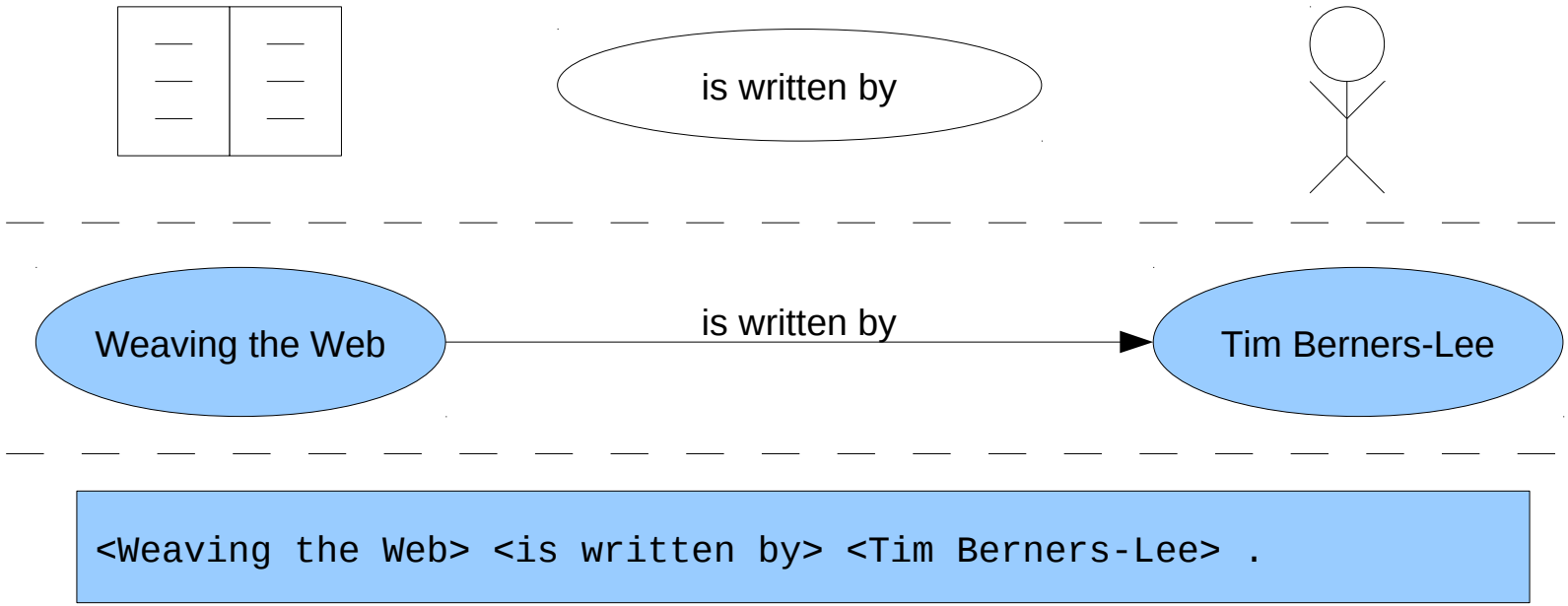
# **Graphs**, (almost) how computers like them

```
<Weaving the Web> <is written by> <Tim Berners-Lee> .
<Tim Berners-Lee> <has first name> "Tim" .
<Tim Berners-Lee> <has last name> "Berners-Lee" .
<Tim Berners-Lee> <is born on> "06/08/1955" .
<Tim Berners-Lee> <is born in> <London> .
<London> <is located in> <England> .
<London> <has population> "7825200" .
<London> <hat Fläche> "130395 km²" .
```

(This notation is called Turtle and it is one of several writing styles for a data model called **RDF**. RDF stands for "**Resource Description Framework**"; this is the de-facto standard for publishing Linked Data.
A big advantage of the Turtle notation: humans can actually read it!)

hbz

# Basic element: the **triple**




```
<Weaving the Web> <is written by> <Tim Berners-Lee> .
```

(A triple is the smallest possible graph. It's components are called **subject**, **predicate** and **object**.)

Your turn!

Open the etherpad for your group. In this etherpad, express the graph you have drawn in RDF.

# Simple social graph in **RDF**

```
<Adrian> <first name> "Adrian" .
<Adrian> <last name> "Pohl" .
<Adrian> <knows> <Felix> .
<Felix> <first name> "Felix" .
<Felix> <last name> "Ostrowski" .
<Felix> <knows> <Adrian> .
```

hbz

**What does** …

➔… <Tim Berners-Lee>,

➔… <London> and

➔… <England>

**stand for, and what does**

➔<has first name>,

➔<is located in> and

➔<has population>

**mean**?

# We need **unambigous reference**!

Authority files are a good start, but again we'll be the only ones understanding those. On the web, people use **URI**s!

(URI stands for **Uniform Resource Identifier**)

hbz

# URI

=

scheme ":" hier-part [ "?" query ] [ "#" fragment ]

# (???)

hbz

http://de.wikipedia.org/wiki/Uniform_Resource_Identifier

ftp://ftp.is.co.za/rfc/rfc3986.txt

file:///home/fo/doc/swib13/slides.odp

urn:isbn:978-1608454303

hbz

# **Graphs**, how computers really like them

```
<urn:isbn:978-0062515872> <http://purl.org/dc/terms/creator> <http://d-nb.info/gnd/121649091> .
<http://d-nb.info/gnd/121649091> <http://xmlns.com/foaf/0.1/givenName> "Tim" .
<http://d-nb.info/gnd/121649091> <http://xmlns.com/foaf/0.1/familyName> "Berners-Lee" .
<http://d-nb.info/gnd/121649091> <http://xmlns.com/foaf/0.1/birthday> "06/08/1955" .
```

(A pleasant side-effect when using HTTP-URIs – which is what Linked Data is based upon, is that they can be **dereferenced**. When following such a **link**, one should get a **description** of the resource. More on that later.)

hbz

# Graphs, (sort of) readable for humans and machines

```
@prefix dc:   <http://purl.org/dc/terms/> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix gnd:  <http://d-nb.info/gnd/> .

<urn:isbn:978-0062515872> dc:creator gnd:121649091 .
gnd:121649091 foaf:givenName "Tim" .
gnd:121649091 foaf:familyName "Berners-Lee" .
gnd:121649091 foaf:birthday "06/08/1955" .
```

(You can abbreviate URIs using **prefixes**. This also makes it easier to identify the vocabularies you use.)

hbz

# But isn't some data we had missing!?

```
<http://d-nb.info/gnd/121649091> <is born in> <London> .
<London> <is located in> <England> .
<London> <has population> "7825200" .
<London> <has area> "130395km²" .
```

(There may not be a URI for everything you want to refer to, neither for entities nor for vocabularies.)

hbz

# Don't repeat others, **link**!

- Reuse properties from existing vocabularies
- Link to things by simple URI reference
- Think Data-Library (as in Software-Library)

hbz

```
@prefix :     <#> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix dc:   <http://purl.org/dc/terms/> .

:ostrowski foaf:givenName "Felix" .
:ostrowski foaf:familyName "Ostrowski" .
:ostrowski foaf:birthday "28.05.1981" .

<> dc:creator :ostrowski .
```

(When something you want to describe does not have a URI yet, you can use Ids that are relative to the describing document. Since two documents can't be at the same place at the same time, these Ids only have to be unique within that document. "<>" stands for the document itself. You can check here if you are creating valid turtle.)

hbz

Your turn!

Reformulate your RDF using the FOAF vocabulary. Also, use DC Terms to assert that you are the authors of the describing document. You can also add further metadata about the document if you want.

hbz

# Simple social graph using FOAF

```
@prefix :     <#> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix dc:   <http://purl.org/dc/terms/> .

:adrian foaf:givenName "Adrian" .
:adrian foaf:familyName "Pohl" .
:adrian foaf:knows :felix .
:felix foaf:givenName "Felix" .
:felix foaf:familyName "Ostrowski" .
:felix foaf:knows :adrian .

<> dc:creator <Felix> .
<> dc:creator <Adrian> .
<> dc:created "25.11.2013" .
```

hbz

# Break

# **Open** Data

Your turn!

Agree on a
Creative Commons License within
your group and link your document to
that license.

(The predicate
<http://creativecommons.org/ns#license>
is well suited for this link, but searching the Web
will reveal alternatives.)

hbz

# Open licencing

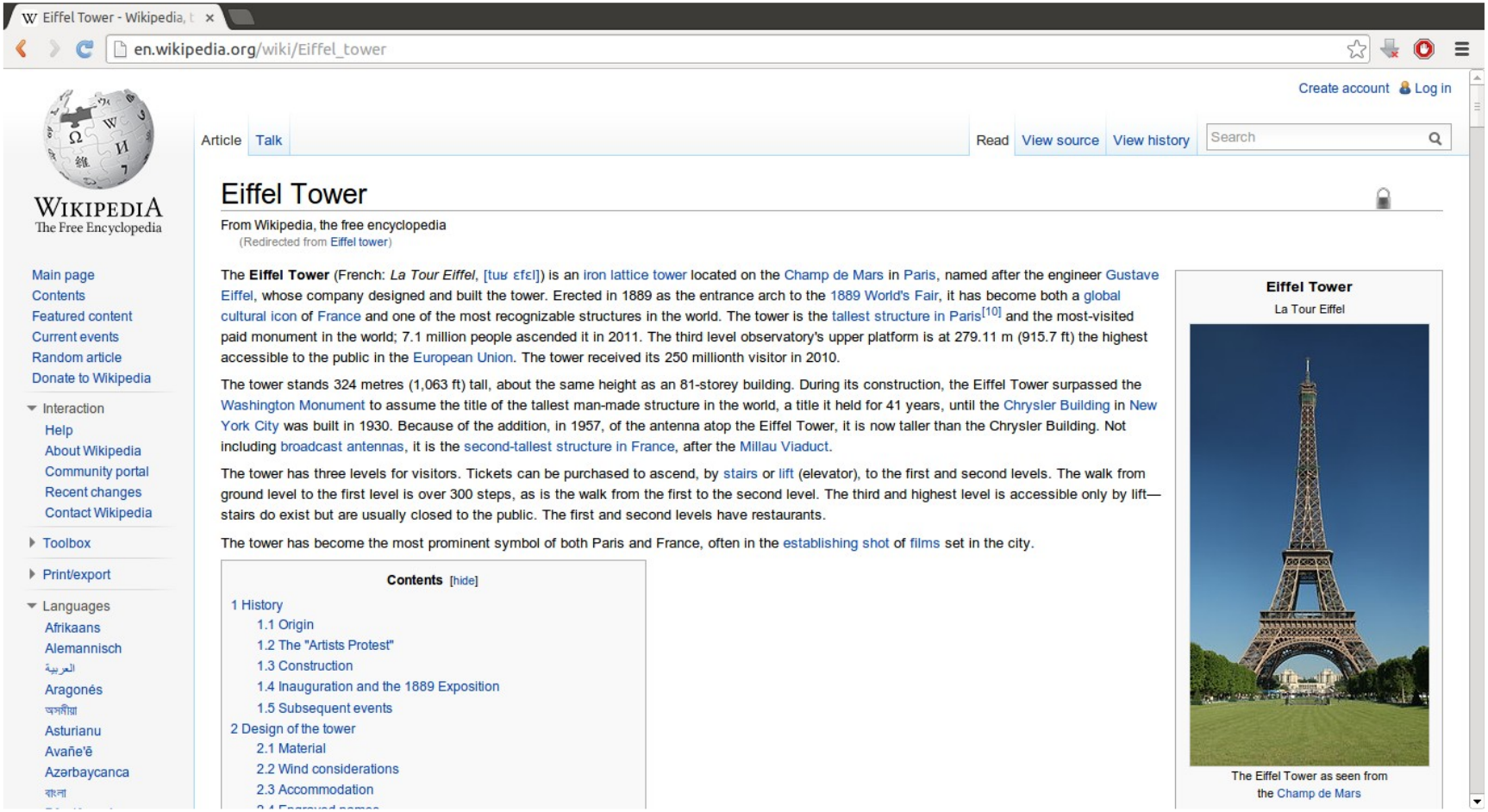
```
@prefix :     <#> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix dc:   <http://purl.org/dc/terms/> .

:adrian foaf:givenName "Adrian" .
:adrian foaf:familyName "Pohl" .
:adrian foaf:knows :felix .
:felix foaf:givenName "Felix" .
:felix foaf:familyName "Ostrowski" .
:felix foaf:knows :adrian .

<> dc:creator :felix .
<> dc:creator :adrian .
<> dc:created "25.11.2013" .
<> <http://creativecommons.org/ns#license>
      <http://creativecommons.org/publicdomain/zero/1.0/> .
```

hbz

# **Linked Data** in Action

hbz

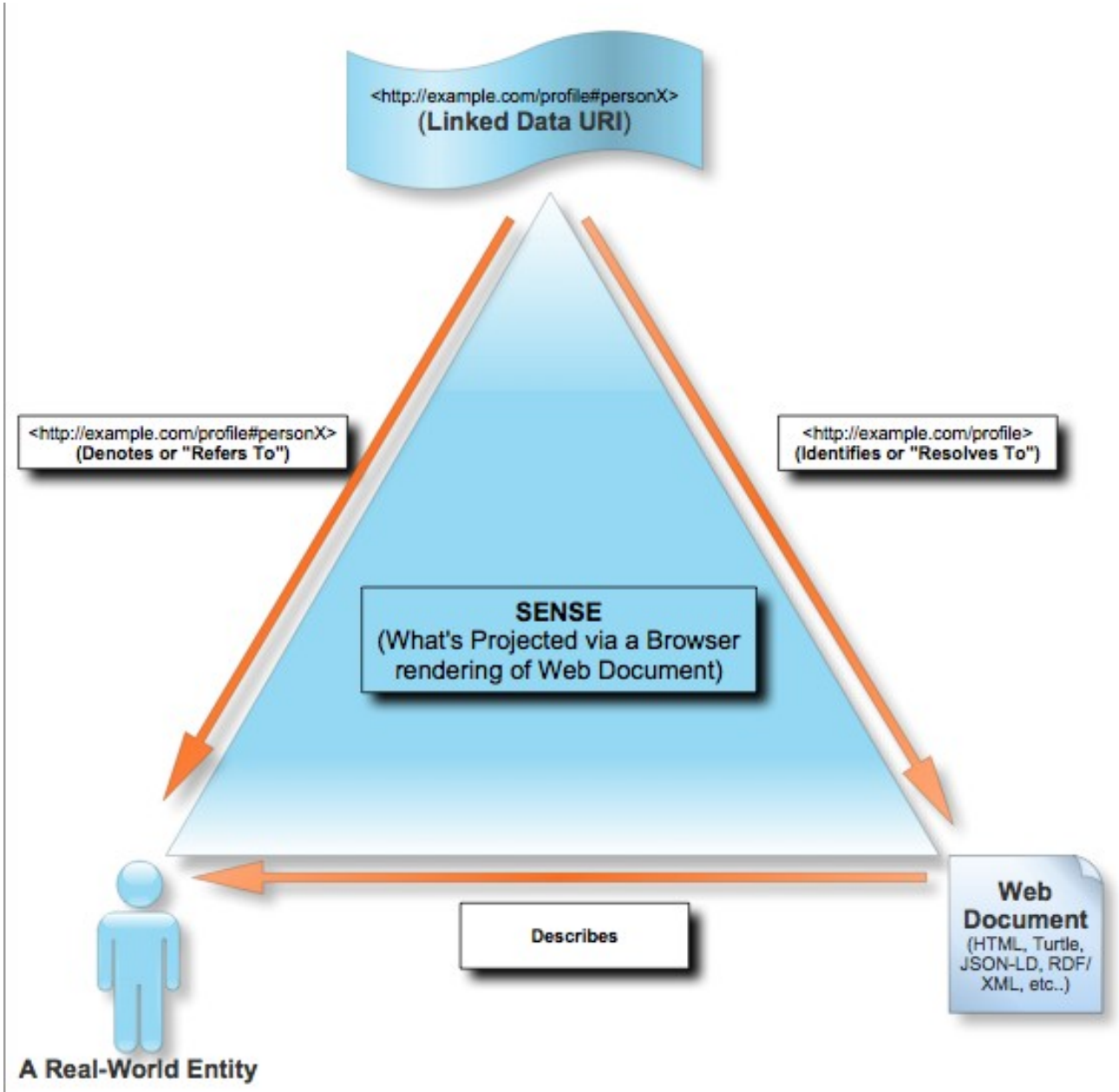
# The Treachery of Documents




Ceci n'est pas la Tour Eiffel.

hbz

**Identification** and **description** of a resource ought to be distinguished!
But in the Linked-Data-Paradigm, both are linked.




https://dl.dropboxusercontent.com/u/11096946/Screenshots/linked-data-hash-uri-semiotic-triangle-2.png

hbz

The description of a resource can be made available in various **formats**. Which format will be delivered can be decided by **Content-Negotiation**.

Your turn!

In your description, link yourself to people from other groups that you know. This doesn't have to be reciprocal.

Also, link (approximately) to the place you live or work. Use DBpedia for this.

hbz

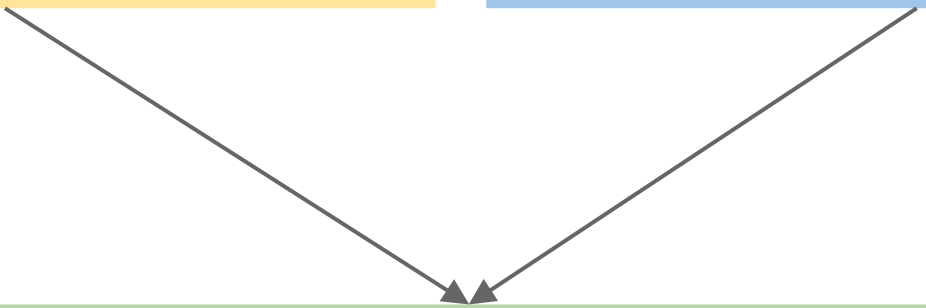
# Break

Scattered machine-readable descriptions are useful, but we can do better than that! RDF is a **distributed** data model that makes it easy to **combine** several descriptions. Furthermore, special **databases** exist that allow to **query** RDF data.

hbz

```
@prefix foaf:
   <xmlns.com/foaf/0.1/> .
@prefix ex1: <http://ex1.org/> .
@prefix ex2: <http://ex2.org/> .

ex1:adrian foaf:givenName "Adrian" .
ex1:adrian foaf:knows ex2:felix .
```

```
@prefix foaf:
   <xmlns.com/foaf/0.1/> .
@prefix there: <http://ex1.org/> .
@prefix here: <http://ex2.org/> .

here:felix foaf:givenName "Felix" .
here:felix foaf:knows there:adrian .
```



```
<http://ex1.org/adrian> <xmlns.com/foaf/0.1/givenName> "Adrian" .
<http://ex1.org/adrian> <xmlns.com/foaf/0.1/knows> <http://ex2.org/felix> .
<http://ex2.org/felix> <xmlns.com/foaf/0.1/givenName> "Felix" .
<http://ex2.org/felix> <xmlns.com/foaf/0.1/knows> <http://ex1.org/knud> .
```

hbz

Felix Ostrowski (FOAF Pro ×

graphthinking.com/foafpress/test/literarymachine.html

DEBUGINFO

a foaf:Person

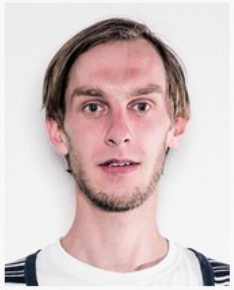
# Felix Ostrowski

## Graph thinker

Blog

### Profiles

Facebook Twitter Slideshare Github LinkedIn Delicious



### Activity

Is there something missing in the link? (30. November 2011)

Dead Simple: RDF and SPARQL using PHP (17. Januar 2011)

### About me

**INTERESTS**

Semantisches Web, Informationsarchitektur, Virtuelle Bibliothek, Linked Open Data

### Contacts

**WORK**

**Address**

- Reichenberger Straße 1, Berlin, 10999, Deutschland

**Phone:** +49 30 12092041

### Network


Knud Möller


Adrian Pohl

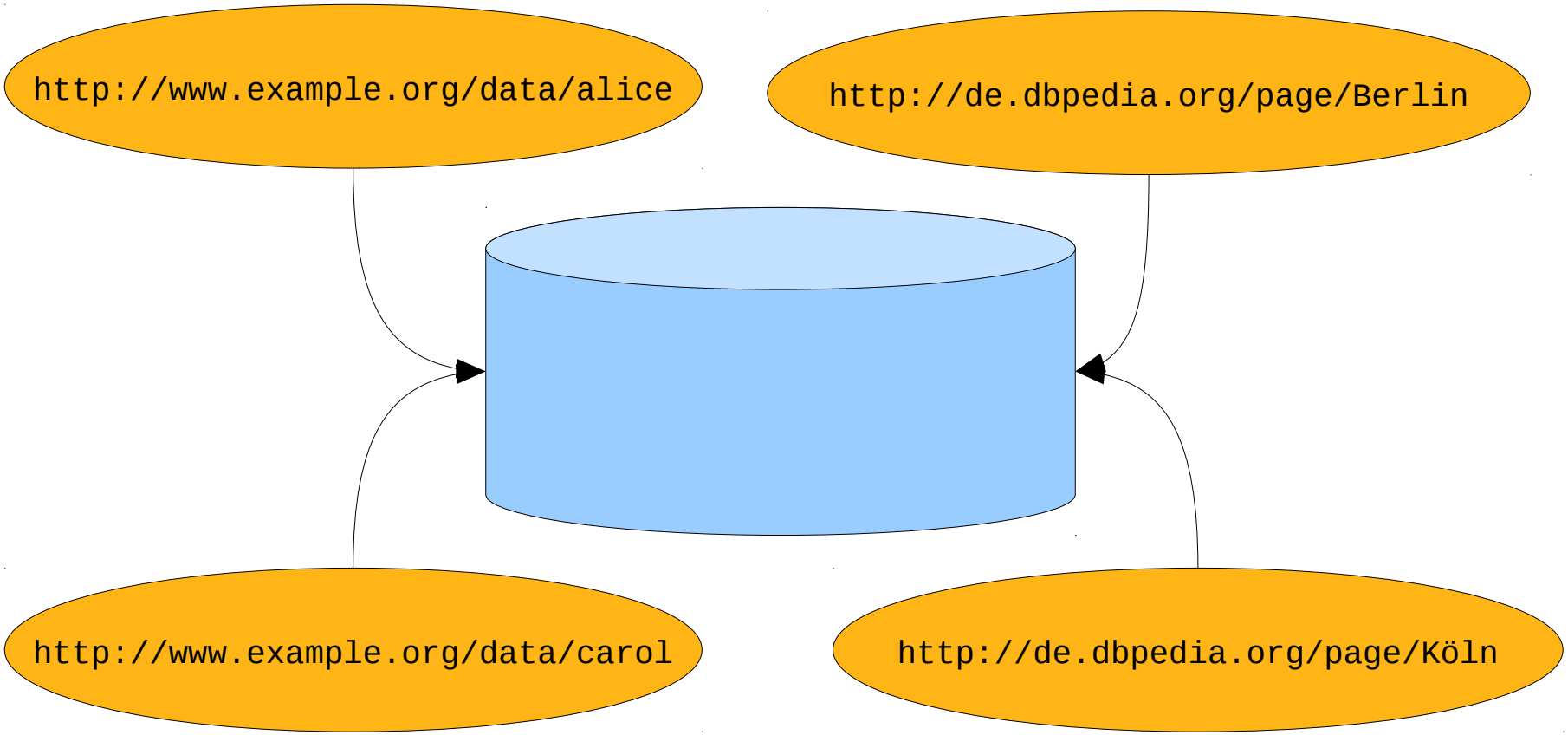

Pascal Christoph

### Interests

**SEMANTISCHES WEB**

hbz

# Triple Stores




hbz

**SPARQL** facilitates queries on the data in a triple store. The foundations for this are simply graph **patterns**. These look almost like triples, the difference being that the contain **variables**.

hbz

```
@prefix   ex: <http://example.org/people#> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .

ex:alice foaf:name "Alice" .
```

```
PREFIX   ex: <http://example.org/people#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>

SELECT * WHERE {
  ex:alice foaf:name ?name .
}
```

| name |
| --- |
| "Alice" |

hbz

```
@prefix   ex: <http://example.org/people#> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .

ex:alice foaf:name "Alice" ;
         foaf:knows ex:bob .
ex:bob   foaf:name "Bob" ;
         foaf:knows ex:carol .
ex:carol foaf:name "Carol" ;
         foaf:knows ex:alice .
```

```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>

SELECT ?name1 ?name2 WHERE {
    ?person1 foaf:knows ?person2 .
    ?person1 foaf:name ?name1 .
    ?person2 foaf:name ?name2 .
}
```

| name1 | name2 |
|---|---|
| "Alice" | "Bob" |
| "Bob" | "Carol" |
| "Carol" | "Alice" |

hbz

```
@prefix      ex: <http://example.org/people#> .
@prefix    foaf: <http://xmlns.com/foaf/0.1/> .
@prefix dbpedia: <http://de.dbpedia.org/resource/> .

ex:alice foaf:name "Alice" ;
         foaf:knows ex:bob ;
         foaf:based_near dbpedia:Berlin .
ex:bob   foaf:name "Bob" ;
         foaf:knows ex:carol ;
         foaf:based_near dbpedia:Dresden .
```

```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

SELECT ?name ?ortname WHERE {
    ?person1 foaf:knows ?person2 .
    ?person2 foaf:name ?name .
    ?person2 foaf:based_near ?ort .
    ?ort rdfs:label ?ortname .
}
```

| name | ortname |
|---|---|
| "Bob" | "Dresden"@de |

hbz

Your turn!

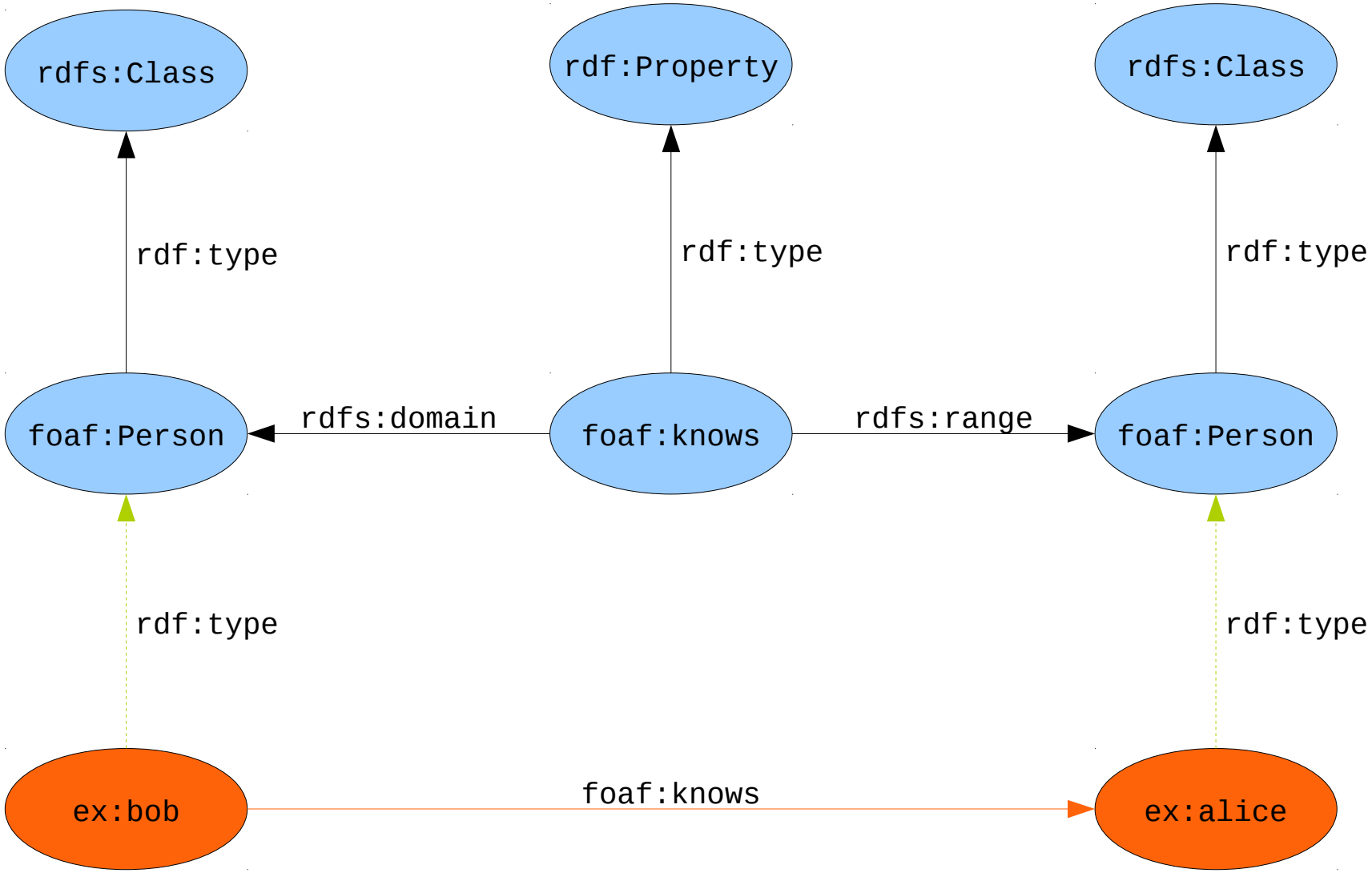
Use SPARQL to analyse your connections. For example you might want to determine who you know directly or indirectly or who comes from the same city as you.

# Break

# Let's put some **Semantic** in the **Web**

The **classes** and **properties** being used can be using **description languages** for **vocabularies**. The relatively simple RDF Schema (**RDFS**) is wide spread, but more complex issues can be expressed in the Web Ontology Language (**OWL**).

hbz


rdfs:Class
rdf:Property
rdfs:Class
rdf:type
rdf:type
rdf:type
foaf:Person
rdfs:domain
foaf:knows
rdfs:range
foaf:Person
rdf:type
rdf:type
ex:bob
foaf:knows
ex:alice

```
# Explicit triples
ex:bob foaf:knows ex:alice .
```

```
# RDF Schema
foaf:knows  rdf:type    rdfs:Property ;
            rdfs:range  foaf:Person ;
            rdfs:domain foaf:Person .
foaf:Person rdf:type    rdfs:Class .
```

```
# Implicit triple, that follow from the schema
ex:bob   rdf:type foaf:Person .
ex:alice rdf:type foaf:Person .
```

hbz

```
# Explicit triples
ex:bob ex:colleague ex:alice .
```

```
# RDF Schema as a "bridge" across vocabularies
ex:colleague rdfs:subPropertyOf foaf:knows ;
             rdfs:domain        ex:Employee ;
             rdfs:range         ex:Employee .
ex:Employee  rdf:type           rdfs:Class ;
             rdfs:subClassOf    foaf:Person .
```

```
# Implicit triple, that follow from the schema
ex:bob   foaf:knows ex:alice .
ex:bob   rdf:type foaf:Person .
ex:alice rdf:type foaf:Person .
ex:bob   rdf:type foaf:Employee .
ex:alice rdf:type foaf:Employee .
```

hbz

Your turn!

# Create an RDF Schema so that from these assertions

```
@prefix team: <http://example.org/soccer/vocab#> .
@prefix   ex: <http://example.org/soccer/resource#> .

ex:team1 team:player ex:bob .
ex:team2 team:player ex:alice .
ex:game1 team:home ex:team1 .
ex:game1 team:away ex:team2 .
```

# the following triples can be inferred.

```
@prefix team: <http://example.org/soccer/vocab#> .
@prefix   ex: <http://example.org/soccer/resource#> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .

ex:team1 rdf:type foaf:Group .
ex:team2 rdf:type foaf:Group .
ex:team1 foaf:member ex:bob .
ex:team2 foaf:member ex:alice .
ex:bob   rdf:type foaf:Person .
ex:alice rdf:type foaf:Person .
ex:game1 rdf:type team:Game .
ex:game2 rdf:type team:Game .
```

hbz

```
@prefix rdf:  <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix team: <http://example.org/soccer/vocab#> .

team:player rdf:type rdfs:Property ;
            rdfs:subPropertyOf foaf:member ;
            rdfs:domain foaf:Person ;
            rdfs:range foaf:Group .

team:home   rdf:type rdfs:Property ;
            rdfs:domain team:Game .

team:away   rdf:type rdfs:Property ;
            rdfs:domain team:Game .

team:Game   rdf:type rdfs:Class .
```

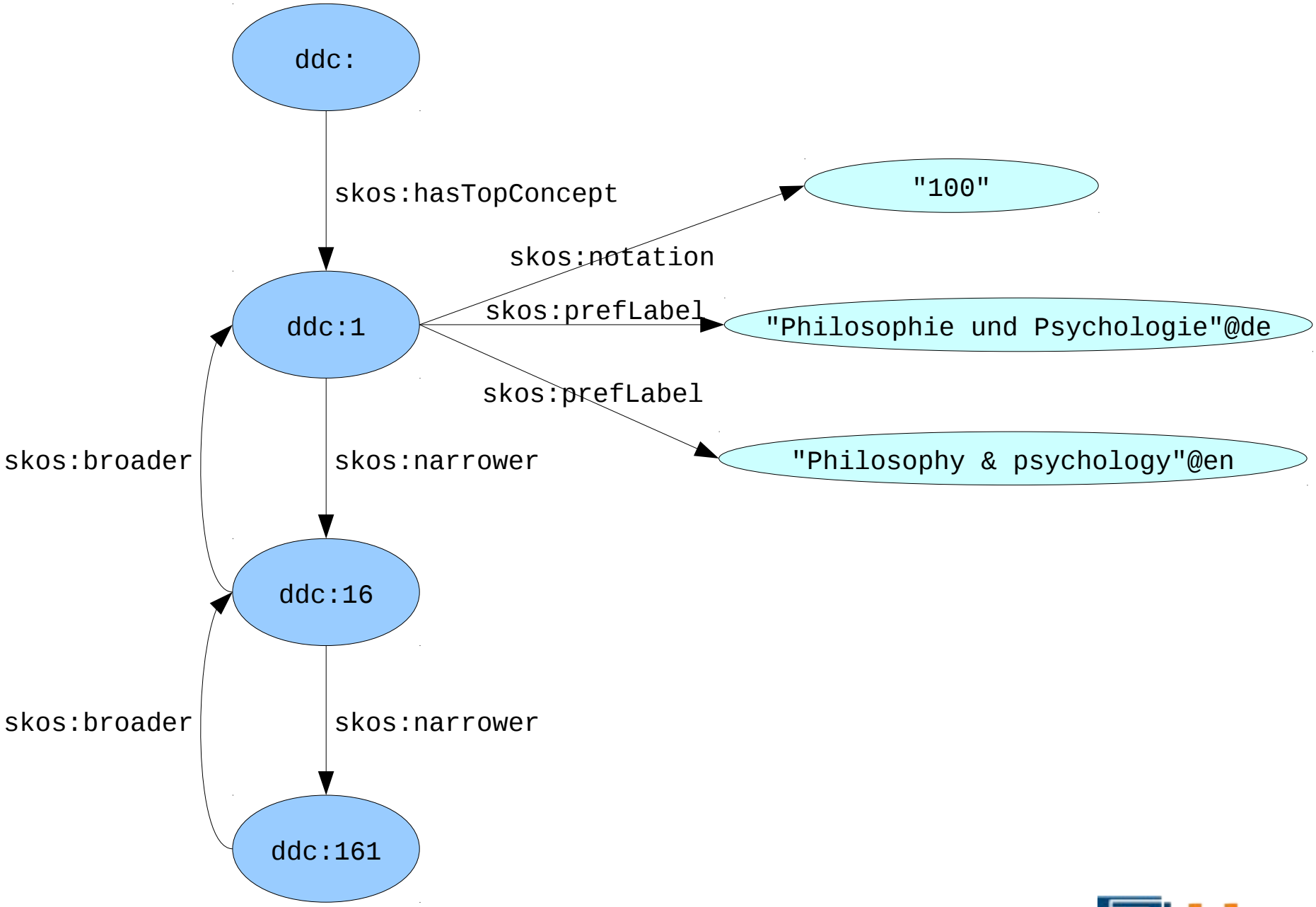
hbz

The expressiveness and the possibilities of inference of RDFS and OWL are not always needed.

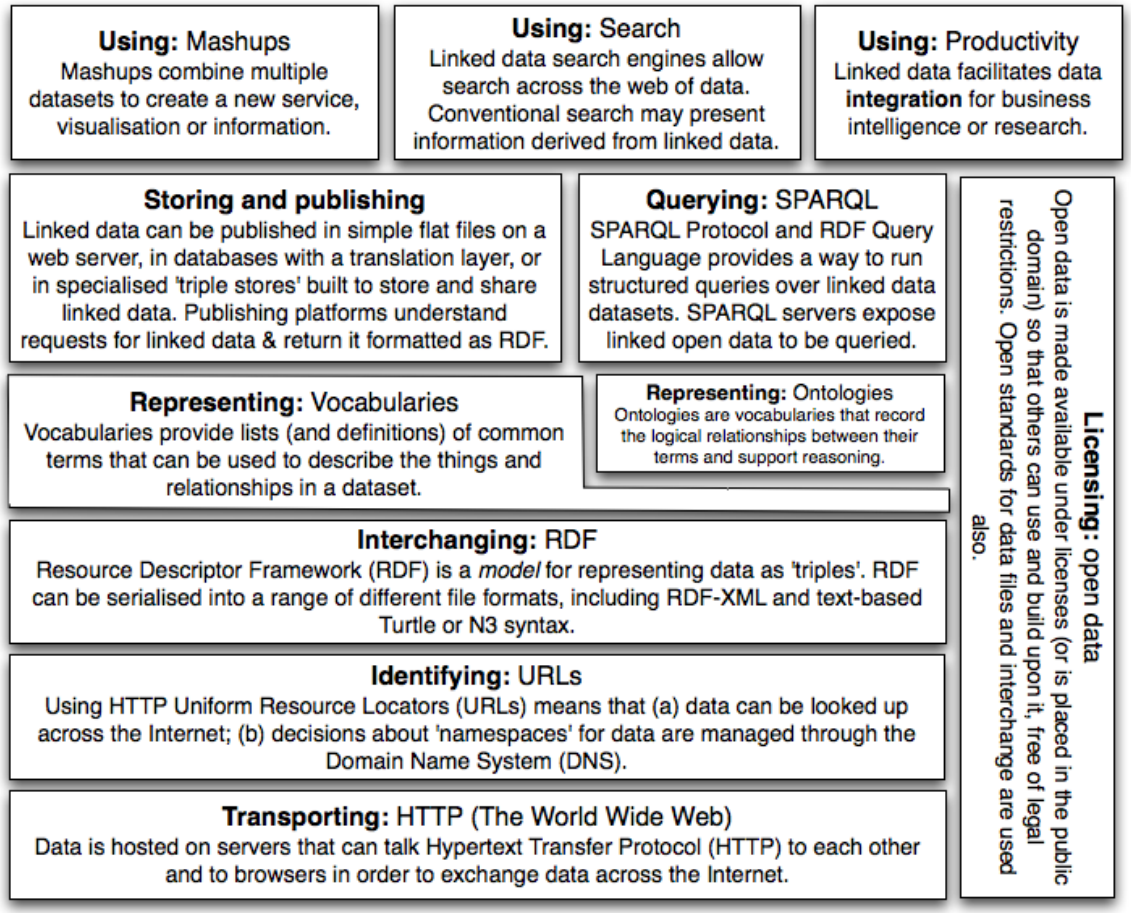
For controlled vocabularies, the **Simple Knowledge Organization System** (SKOS) is a simpler alternative that is also based on RDF.

The Dewey Decimal Classification and the Library of Congress Subject Headings have already found their way into the Linked-Data-world.

hbz


ddc:
skos:hasTopConcept
"100"
skos:notation
ddc:1
skos:prefLabel
"Philosophie und Psychologie"@de
skos:prefLabel
"Philosophy & psychology"@en
skos:broader
skos:narrower
ddc:16
skos:broader
skos:narrower
ddc:161

# Elements of Linked (Open) Data




**Elements of the Linked Open Data Stack (revision 3)** - 5th May 2011. CC BY-SA-NC
Draft sketch by Tim Davies (@timdavies / tim@practicalparticipation.co.uk) for IKM Working Paper on Linked Open Data for Development. Comments welcome. Search 'linked open data stack' on http://www.opendataimpacts.net for latest version.

Idea based on Semantic Web Stack at http://en.wikipedia.org/wiki/Semantic_Web_Stack

hbz

Thank you!

# License