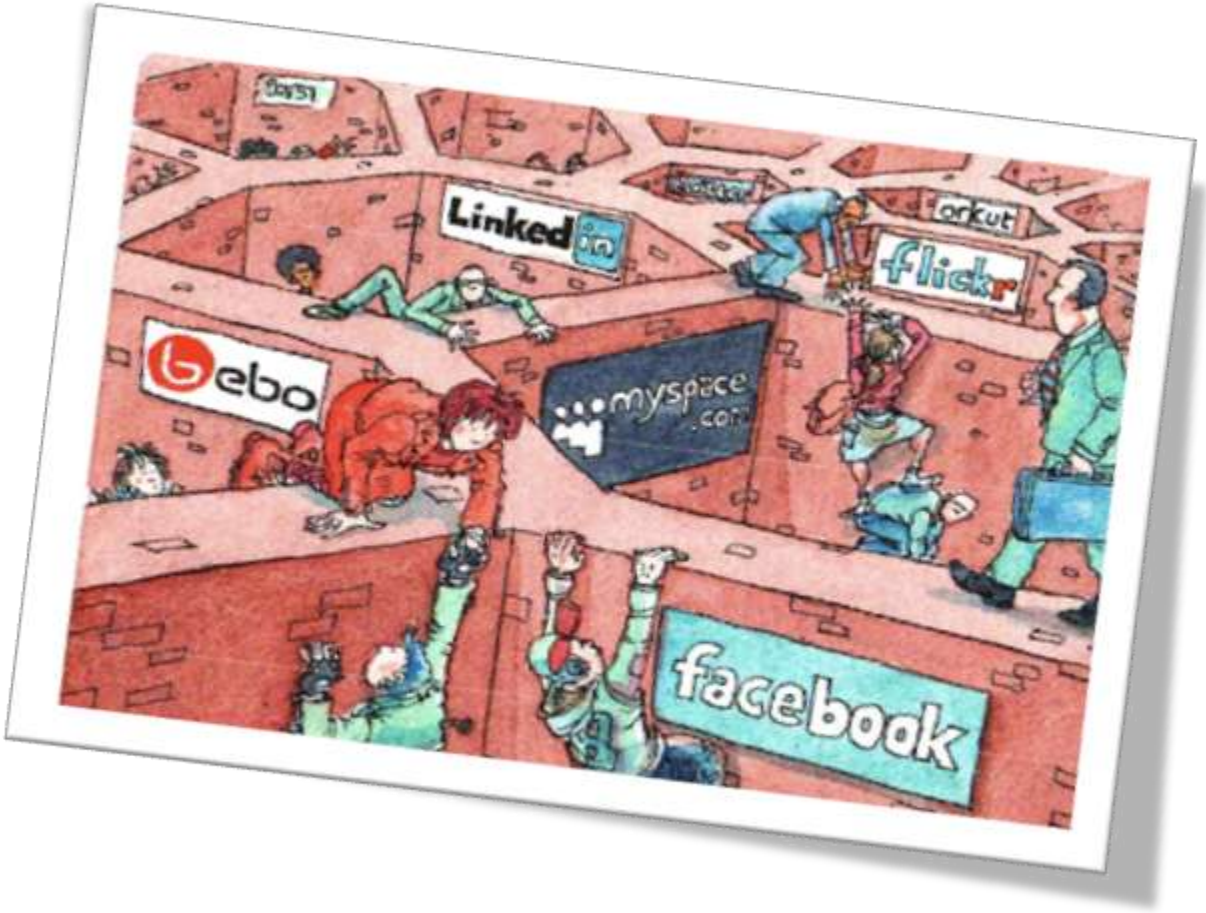
Freie Universität Berlin

Linked Data, Microformats, RDFa, Microdata

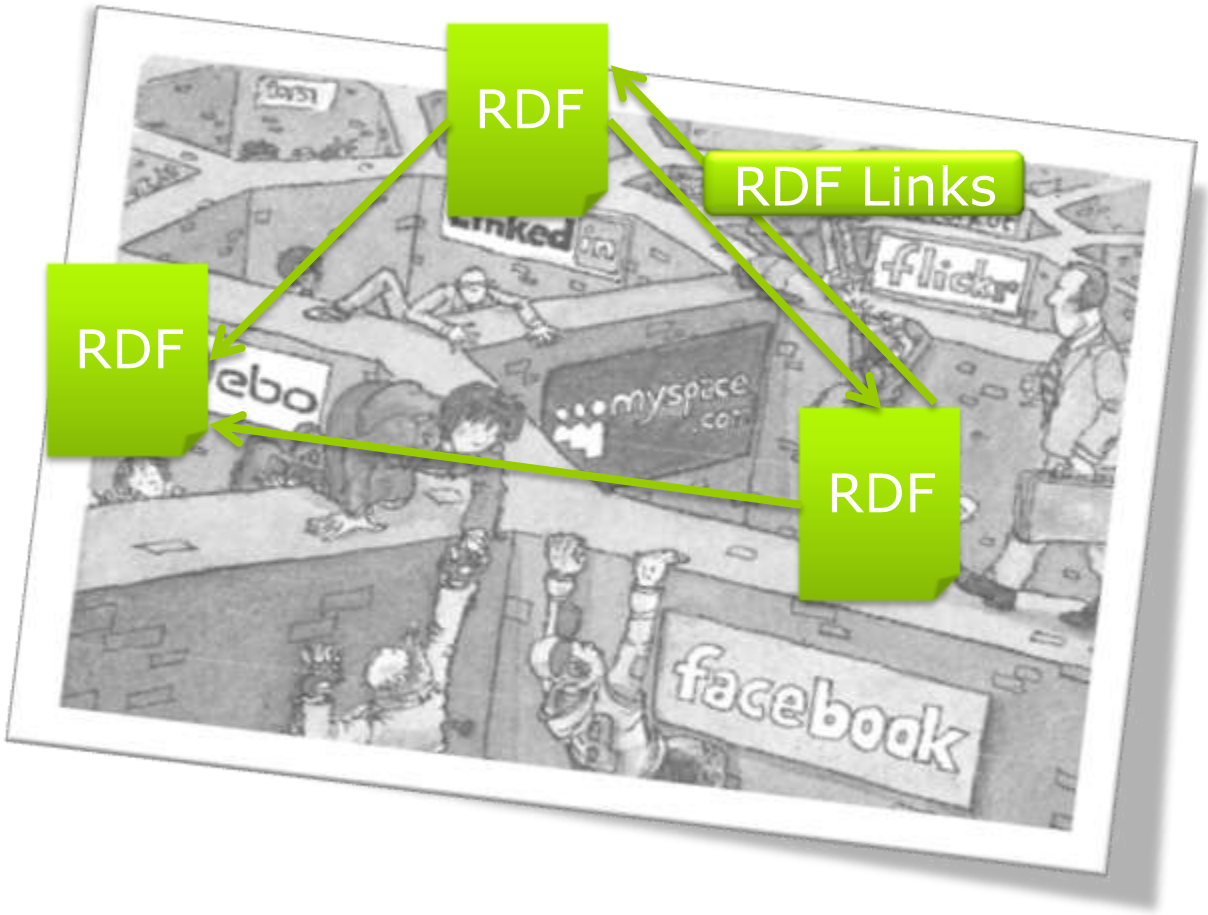
Markus Luczak-Rösch
Freie Universität Berlin
Institut für Informatik
Netzbasierte Informationssysteme
markus.luczak-roesch@fu-berlin.de

Probleme?

→ Data Silos



Web of Linked Data



Linked Data Prinzipien

Freie Universität Berlin

1. URIs als Namen für alle "Dinge"

   http://dbpedia.org/resource/Berlin

2. http:// URIs damit man im Web auf diese Namen zugreifen kann

   Content Negotiation

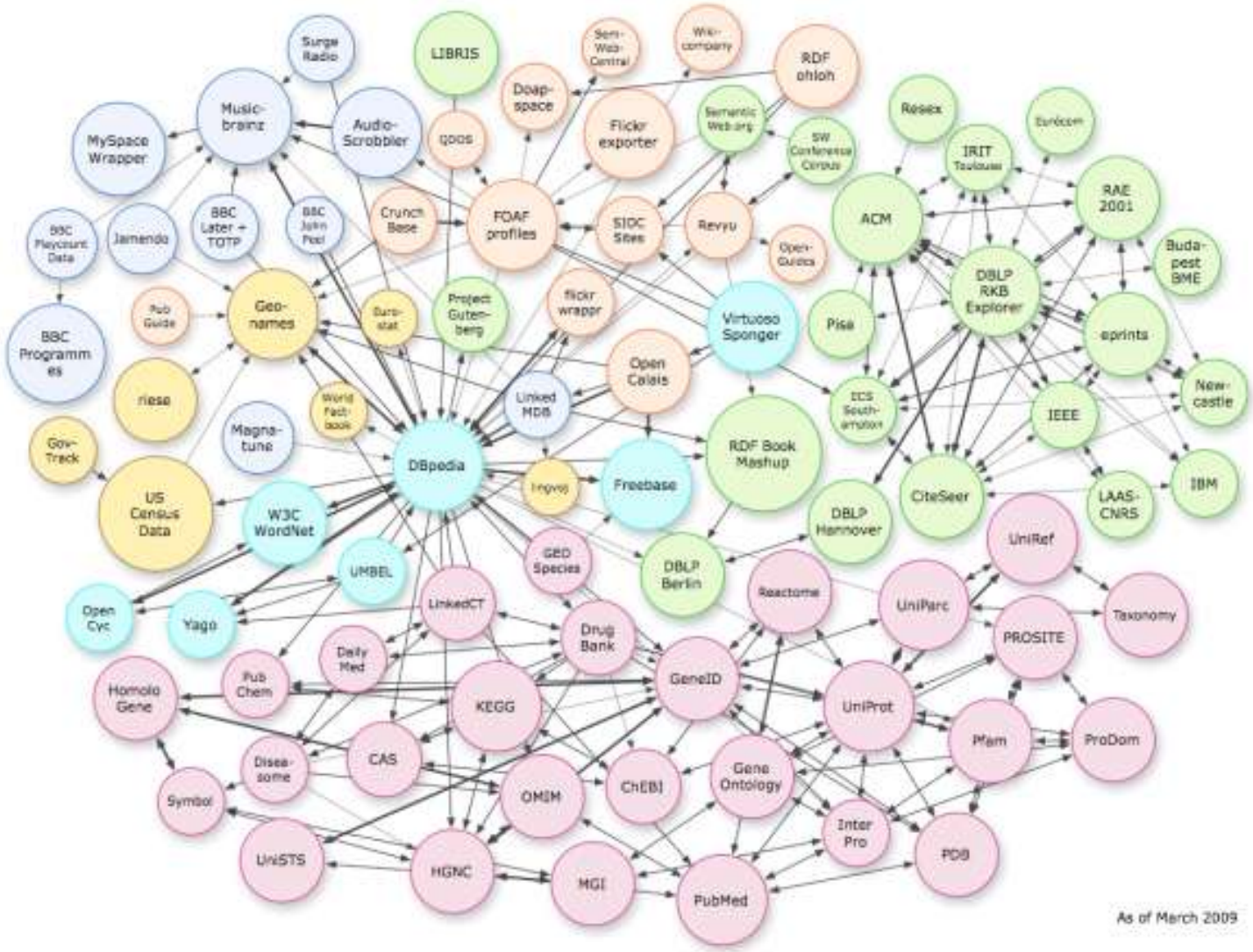
3. Wenn eine URI aufgerufen wird sollen sinnvolle Informationen entsprechend der Standards (RDF, SPARQL) geliefert werden

   http://dbpedia.org/page/Berlin
   http://dbpedia.org/data/Berlin

4. Links zu anderen URIs, damit Nutzer mehr "Dinge" finden können

   yago-res:Berlin S
   owl:sameAs P
   dbpedia:Berlin O

Linking Open Data Cloud



Linked Data erzeugen

- Domäne festlegen
- Identifier
- Vokabular
- RDF-Links

Domäne festlegen, Daten verstehen

Freie Universität Berlin

- Personen
- Geo
- Ämter
- …
- Omas Kuchen

Domäne festlegen, Daten verstehen

- Klaus Wowereit
  - ist: regierender Bürgermeister von Berlin
- Berlin
  - lat…
  - long…
  - …
- Thing
  - Mayor
  - City
  - …

Identifier

- Klaus Wowereit
- Mayor

http:// URIs

http://…

http://…

Identifier

- Klaus Wowereit
- Mayor

eigener Namensraum

http://dbpedia.org/…

http://dbpedia.org/…

Identifier

- Klaus Wowereit
- Mayor

implementierungsunabhängig

http://dbpedia.org/…

http://dbpedia.org/…

http://www4.wiwiss.fu-berlin.de:2020/demos/dbpedia/cgi-bin/resources.php?id=Klaus_Wowereit

Identifier

- Klaus Wowereit
- Mayor

Slash oder Hash?

http://dbpedia.org/daten.rdf
#Klaus_Wowereit

http://dbpedia.org/
Klaus_Wowereit

http://dbpedia.org/konzepte.owl
#Mayor

http://dbpedia.org/Mayor

Identifier

- Klaus Wowereit
- Mayor

Faktenwissen oder Konzeptwissen?

http://dbpedia.org/resource/
Klaus_Wowereit

http://dbpedia.org/ontology/
Mayor

- Klaus Wowereit

http://dbpedia.org/resource/Klaus_Wowereit ← Ressource

http://dbpedia.org/page/Klaus_Wowereit ← HTML-Seite

http://dbpedia.org/data/Klaus_Wowereit ← Daten

Freie Universität Berlin

- Klaus Wowereit

http://dbpedia.org/resource/Klaus_Wowereit ← Ressource

http://dbpedia.org/page/Klaus_Wowereit ← HTML-Seite

http://dbpedia.org/data/Klaus_Wowereit ← Daten

http://dbpedia.org/data/Klaus_Wowereit.rdf

http://dbpedia.org/data/Klaus_Wowereit.ntriples

Vokabular

- Wiederverwenden
  - Geo
  - FOAF
  - GoodRelations
  - SIOC
  - DOAP
  - …

http://xmlns.com/foaf/0.1/Person

- Entwickeln
  - Thing
    - Person
      - OfficeHolder
      - …
    - …

http://dbpedia.org/ontology/Person

http://dbpedia.org/ontology/OfficeHolder

Vokabular

- Wiederverwenden nicht wiedererfinden!
- Mischen!

  - Geo
  - FOAF
  - Dublin Core
  - DBpedia Ontology
  - …

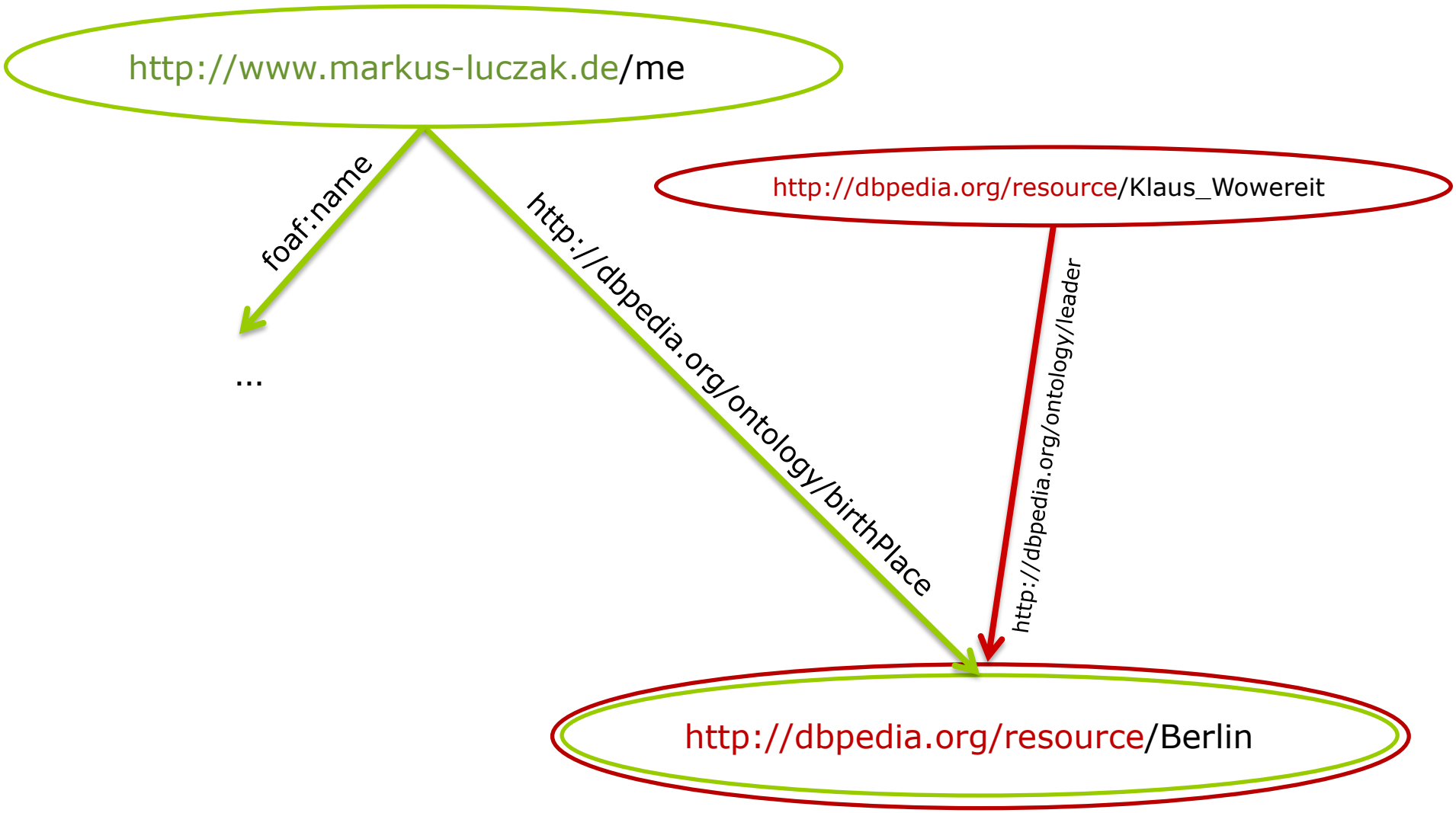
http://xmlns.com/foaf/0.1/Person

http://www.w3.org/2003/01/geo/wgs84_pos#lat

http://dbpedia.org/ontology/leader

http://dbpedia.org/ontology/City

RDF-Links

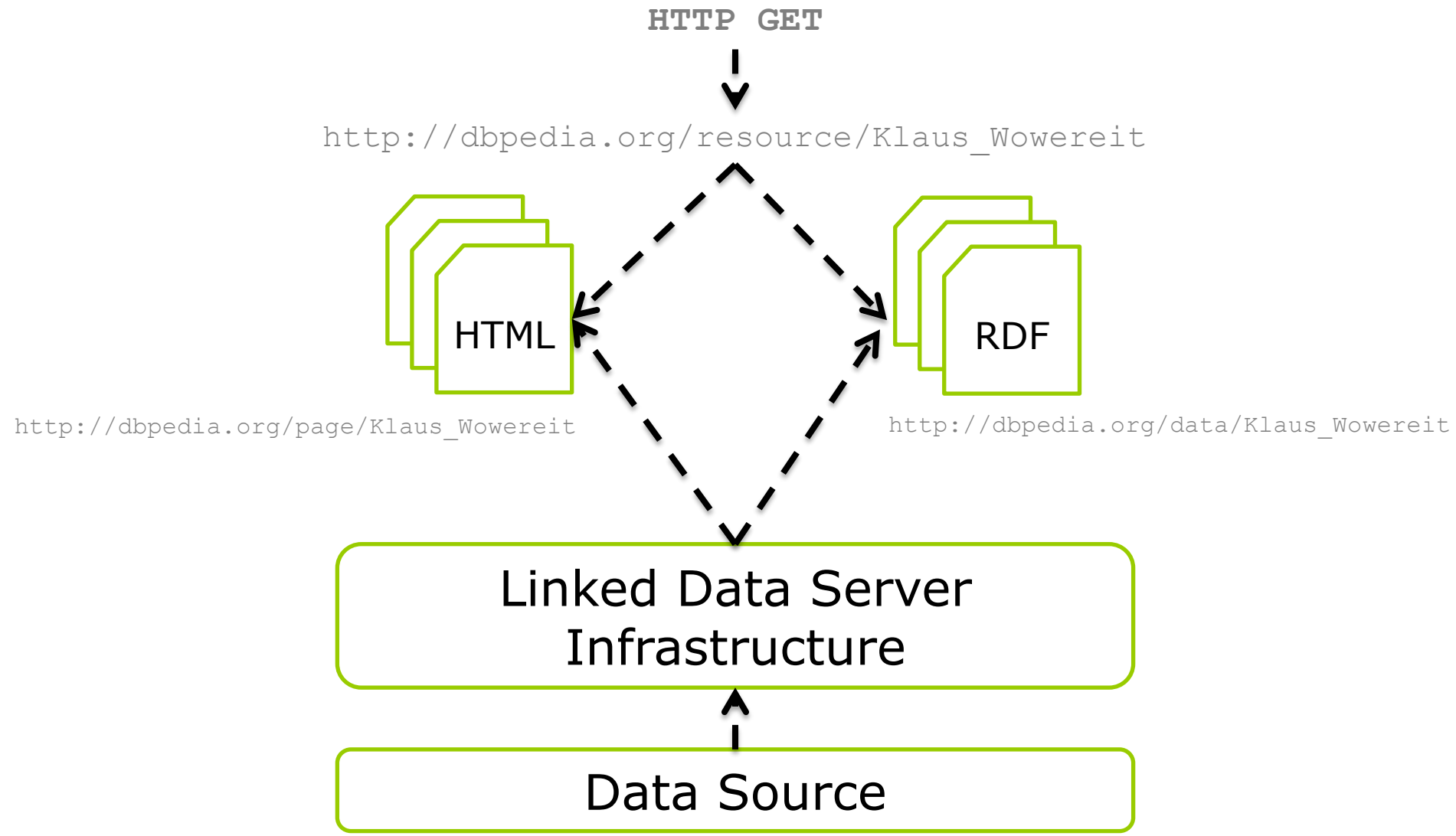

http://www.markus-luczak.de/me
foaf:name
...
http://dbpedia.org/ontology/birthPlace
http://dbpedia.org/resource/Klaus_Wowereit
http://dbpedia.org/ontology/leader
http://dbpedia.org/resource/Berlin

RDF-Links

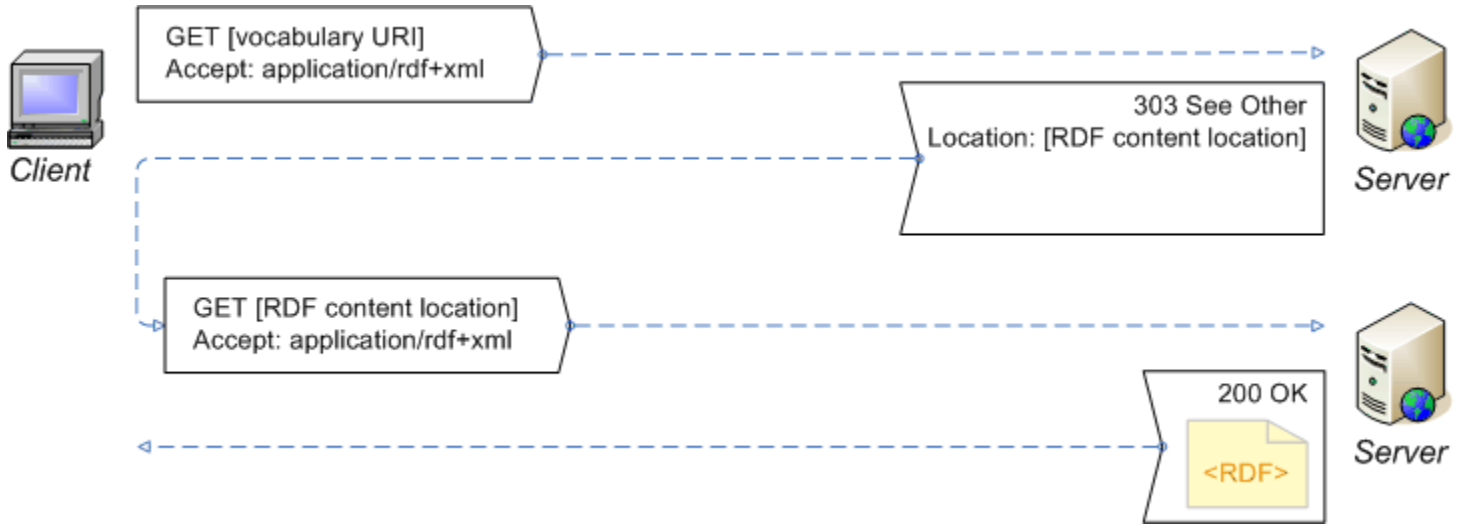
- owl:sameAs

- …
- foaf:homepage
- foaf:topic
- foaf:based_near
- foaf:maker/foaf:made
- foaf:depiction
- rdfs:seeAlso
- …

Linked Data Infrastruktur



Linked Data Infrastruktur

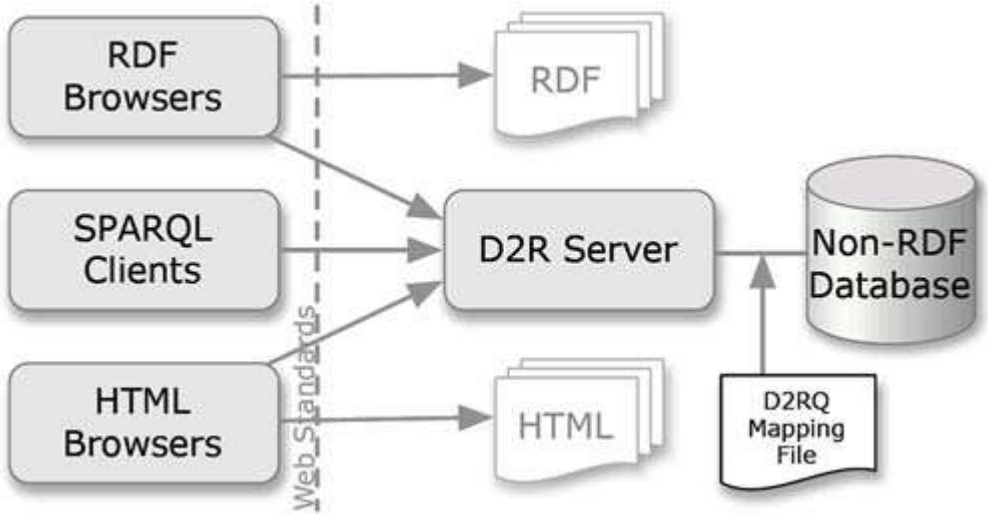


Linked Data anbieten/serven

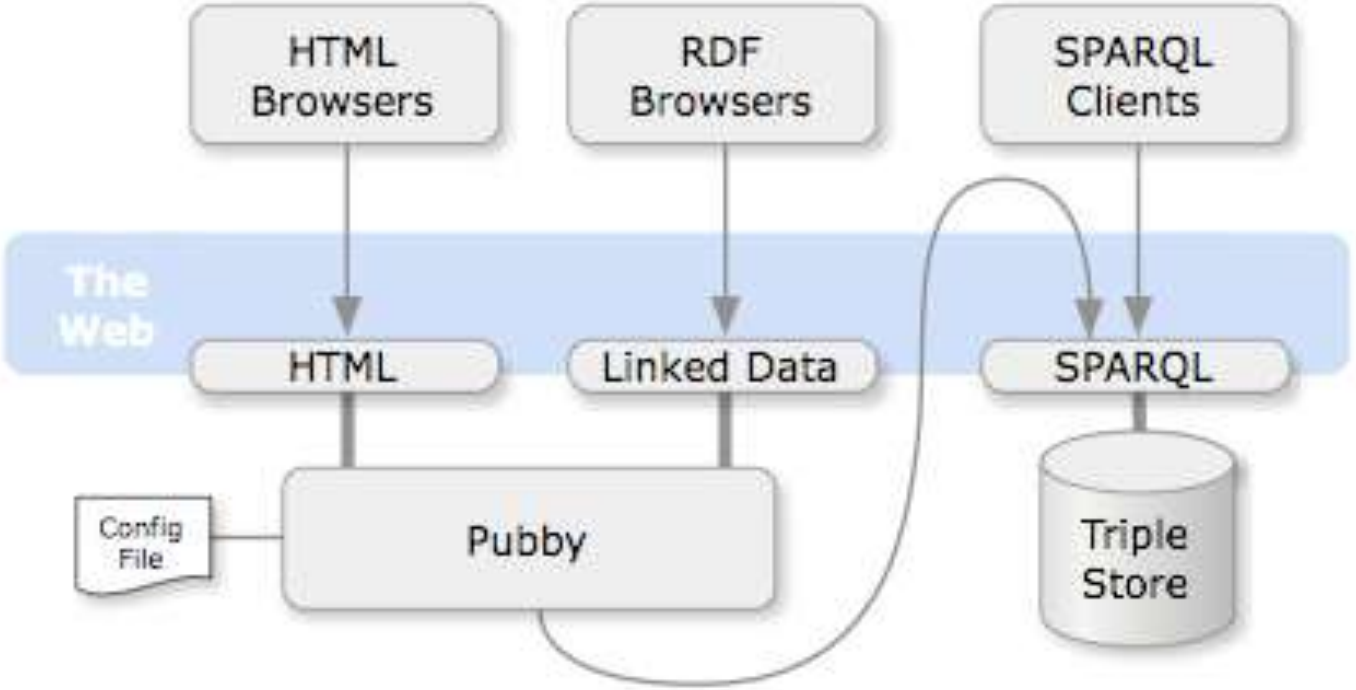
- statische RDF-Dateien
- relationale Datenbanken
- Linked Data Server
- API-Wrapper
- RDFa

RDB2Linked Data (Bsp.: D2R Server)

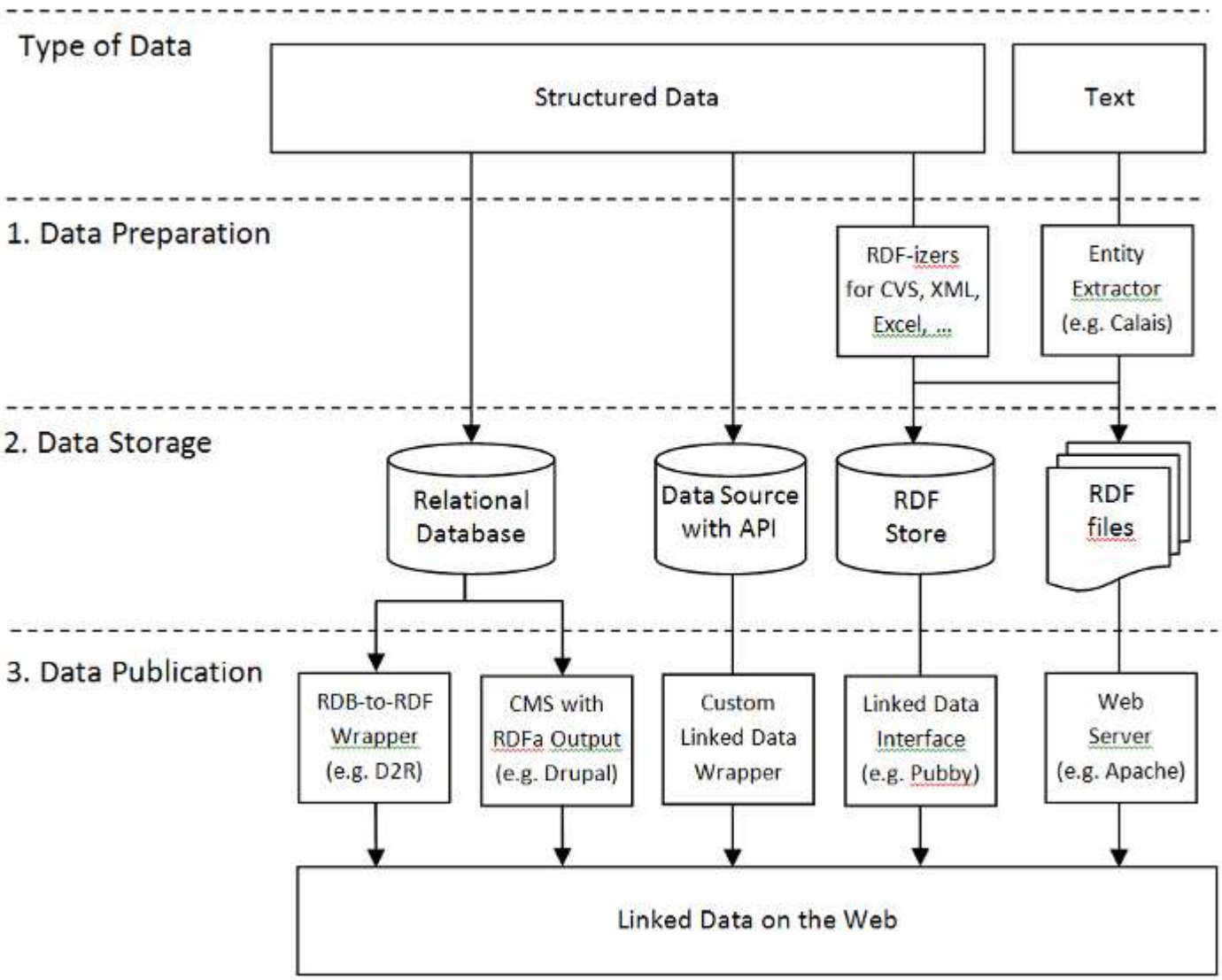
- Mappen eines relationalen Schemas auf RDF (-Klassen und -Beziehungen)
  - deklarative Mapping-Sprache (D2RQ)
  - SPARQL-Endpoint
  - Linked Data Server



Linked Data Server (Bsp.: Pubby)



Linked Data Infrastruktur



Linked Data konsumieren

- URIs finden
- Zusätzliche Daten finden
- SPARQL-Endpoints finden

URIs finden

- Gibt es (bereits) einen Identifier für das, was mich interessiert?
  - Lookup-Services für Datensets
    - http:// lookup.dbpedia.org
    - http://rkbexplorer.com
    - …
  - Web of Data Search Engines
    - http://sindice.com
    - http://ws.nju.edu.cn/falcons/objectsearch/index.jsp

zusätzliche Daten finden

- Wo finde ich weitere Daten für die URI, die ich zur Hand habe?
  - Links verfolgen
    - rdfs:seeAlso
    - owl:sameAs
  - Co-Referenz-Dienst verwenden
    - http://sameas.org
  - Web of Data Search Engines

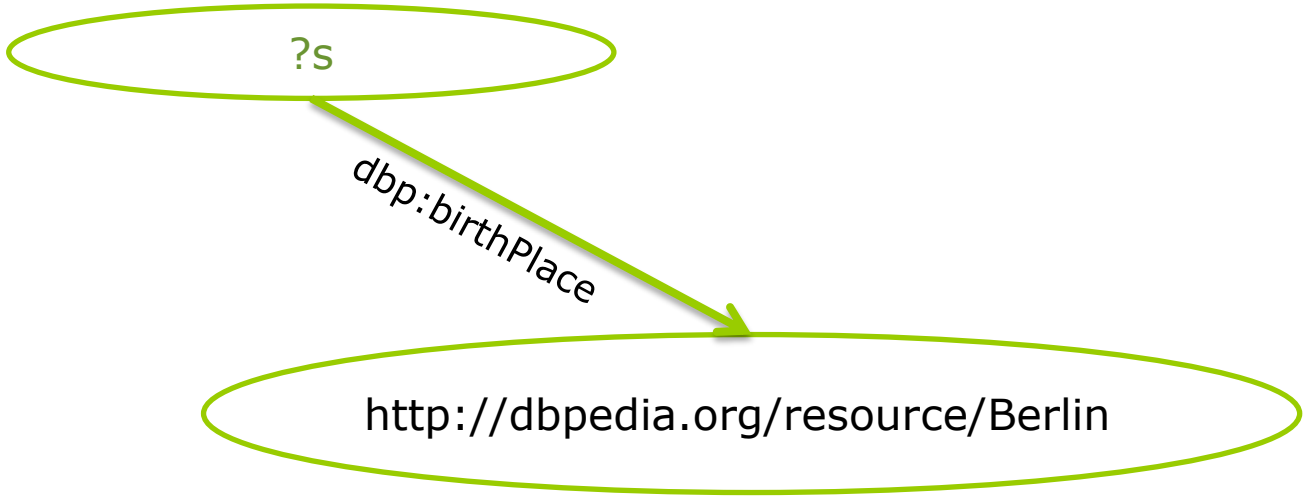
SPARQL-Endpoints finden

- http://www.w3.org/wiki/SparqlEndpoints (depr.)
- http://thedatahub.org/group/lodcloud

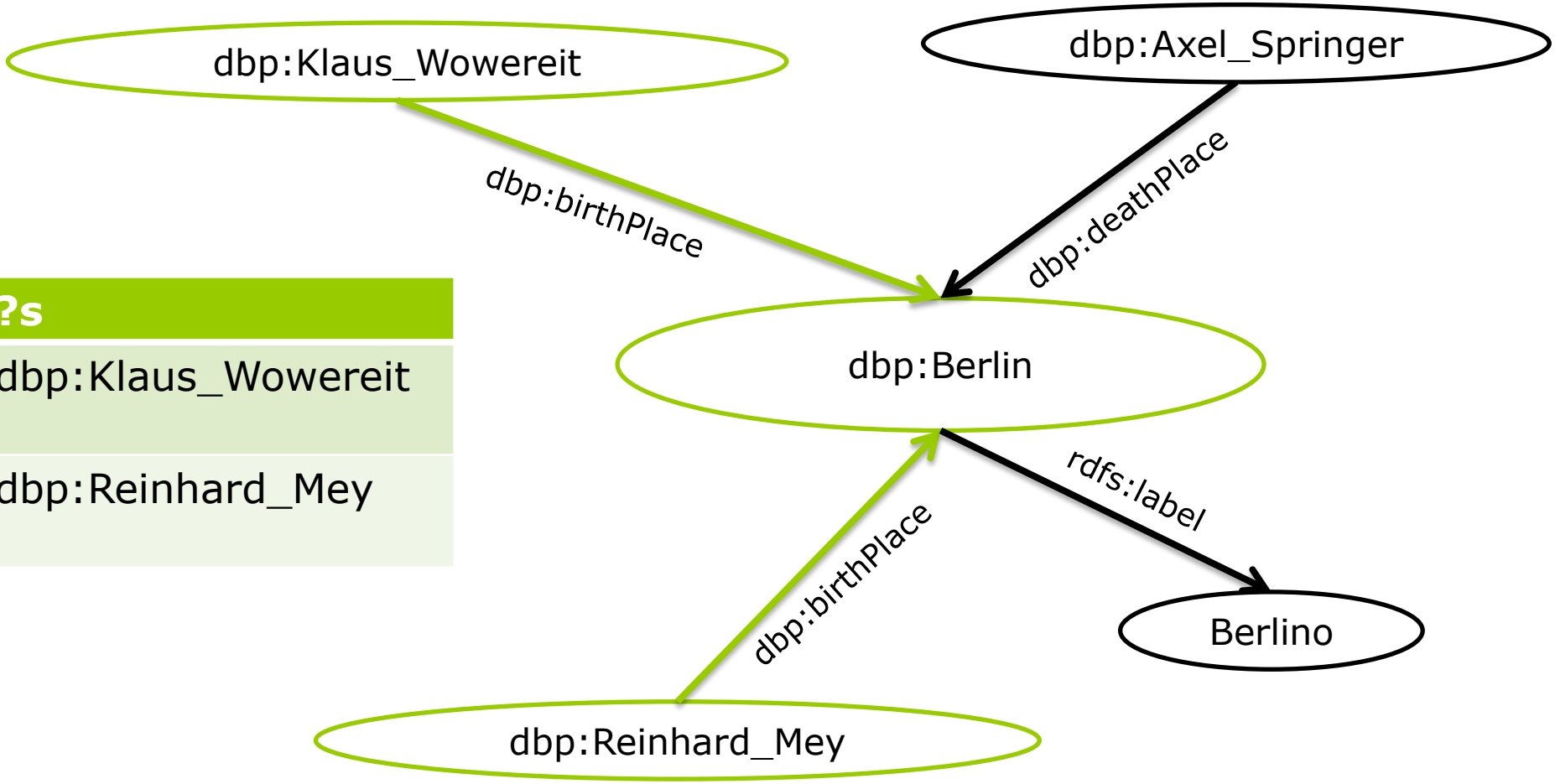
SPARQL-Anfragen über Linked Data

- SPARQL: Anfragesprache für RDF-Daten

- Grundsatz: Pattern-Matching
  - beschreibe Graphpattern
  - frage RDF-Graph mit diesem Pattern an
  - Subgraphen, die Pattern matchen kommen in die Ergebnismenge



SPARQL-Anfragen über Linked Data



?s
dbp:Klaus_Wowereit
dbp:Reinhard_Mey

SPARQL-Anfragen über Linked Data

Freie Universität Berlin

- Anfrage via RESTful Service mit Parameter *query*

```
GET /sparql?query=PREFIX+rdf… HTTP/1.1
Host: dbpedia.org
```

- Antwort liefert ein SPARQL-Result-Format (XML oder JSON)

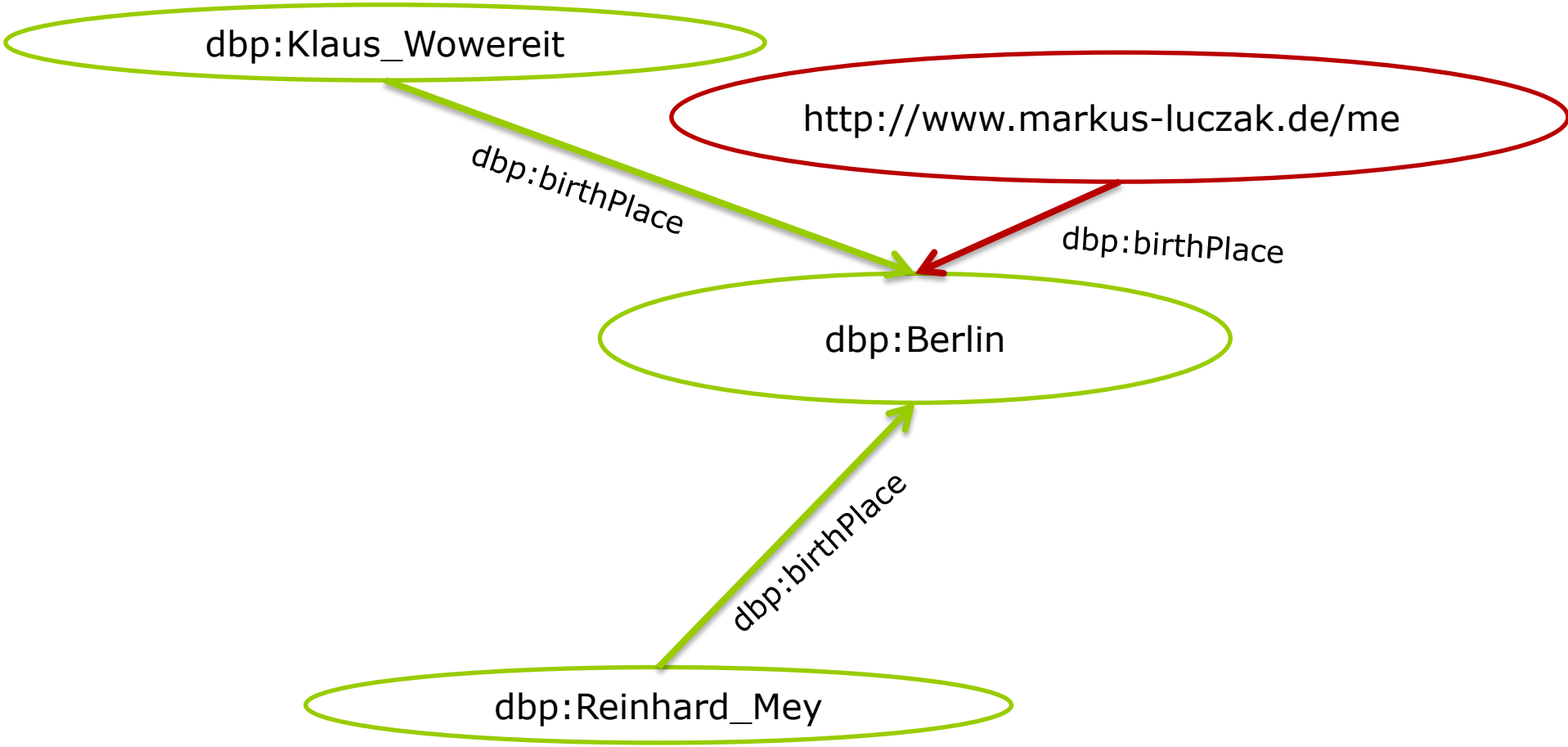
```
<?xml version="1.0"?>
<sparql xmlns="http://www.w3.org/2005/sparql-results#">
  <head>
    <variable name="x"/>
    <variable name="hpage"/>
  </head>

  <results>
    <result>
      <binding name="x"> ... </binding>
      <binding name="hpage"> ... </binding>
    </result>

    <result>
      <binding name="x"> ... </binding>
      <binding name="hpage"> ... </binding>
    </result>
    ...
  </results>

</sparql>
```

http://www.w3.org/TR/rdf-sparql-XMLres/


```
{
  "head": { "vars": [ "book" , "title" ]
  } ,
  "results": {
    "bindings": [
      {
        "book": { "type": "uri" , "value": "http://example.org/book/book6" } ,
        "title": { "type": "literal" , "value": "Harry Potter and the Half-Blood Prince" }
      } ,
      {
        "book": { "type": "uri" , "value": "http://example.org/book/book5" } ,
        "title": { "type": "literal" , "value": "Harry Potter and the Order of the Phoenix" }
      } ,
      {
        "book": { "type": "uri" , "value": "http://example.org/book/book4" } ,
        "title": { "type": "literal" , "value": "Harry Potter and the Goblet of Fire" }
      } ,
      {
        "book": { "type": "uri" , "value": "http://example.org/book/book3" } ,
        "title": { "type": "literal" , "value": "Harry Potter and the Prisoner Of Azkaban" }
      } ,
      {
        "book": { "type": "uri" , "value": "http://example.org/book/book2" } ,
        "title": { "type": "literal" , "value": "Harry Potter and the Chamber of Secrets" }
      } ,
      {
        "book": { "type": "uri" , "value": "http://example.org/book/book1" } ,
        "title": { "type": "literal" , "value": "Harry Potter and the Philosopher's Stone" }
      }
    ]
  }
}
```


http://www.w3.org/TR/rdf-sparql-json-res/

SPARQL-Anfragen über Linked Data

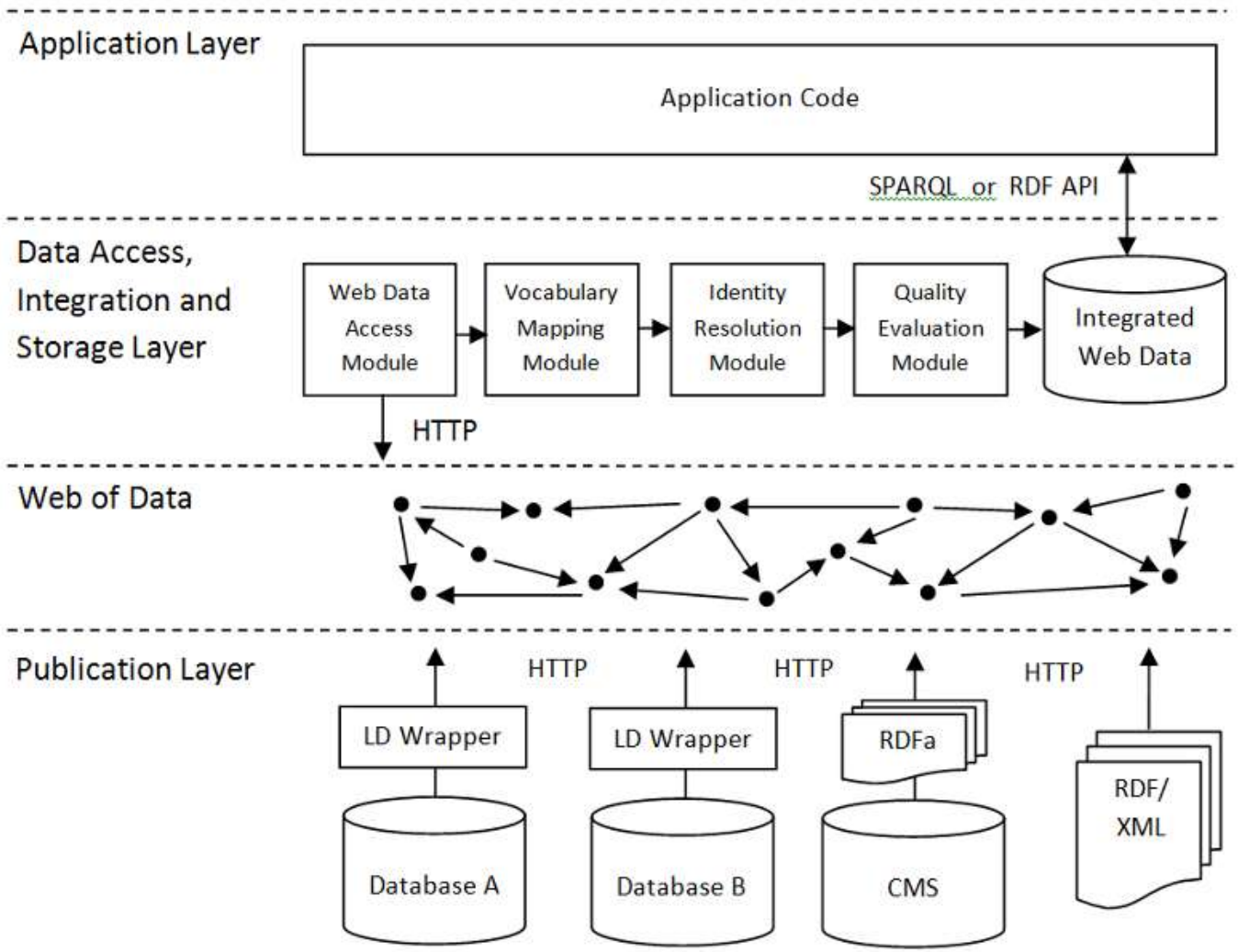


SPARQL-Anfragen über Linked Data

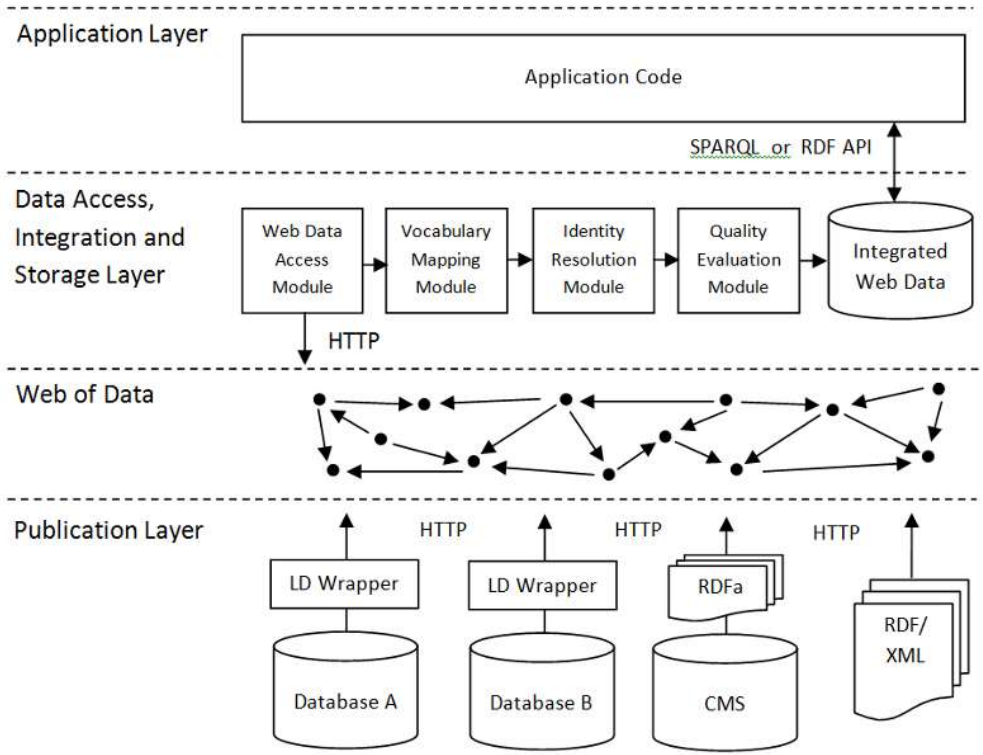
- verteilte Datenhaltung erzeugt Herausforderungen für Anfragen

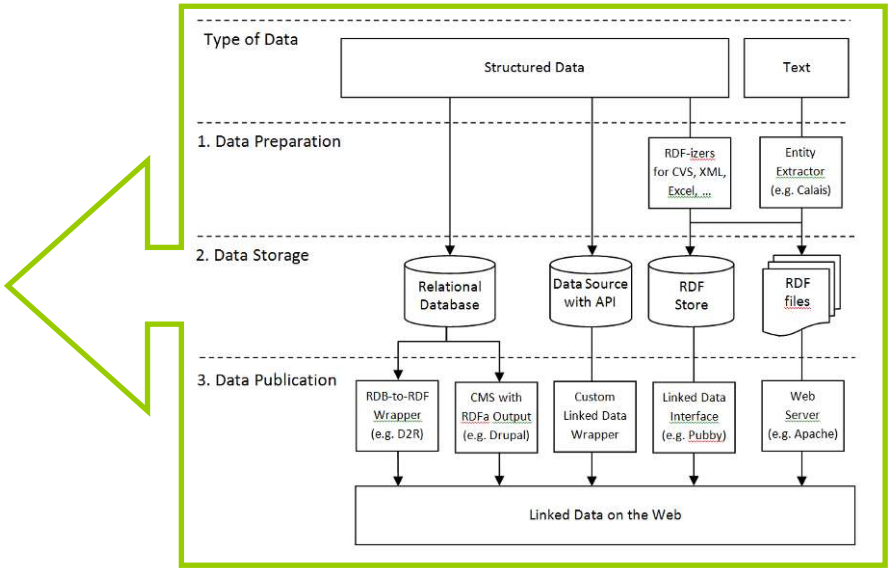
- Anfrageansätze
  - follow-up Anfragen ← anwendungsspezifisch, proprietät
  - zentrale Datensammlung anfragen ← eher trivial
  - föderierte Anfrage ← interessanter
    - Idee: Mediator anfragen, der Subanfragen an relevante Quellen verteilt und Ergebnis integriert
  - Link-Traversierung zur Anfragezeit ← sehr interessant
    - Idee: verfolge Links in gefundenen Ergebnissen, um dynamisch den Datenbestand zu vergrößern und beziehe diese zusätzlichen Daten dann in die Anfrageauswertung mit ein

Linked Data Anwendungen

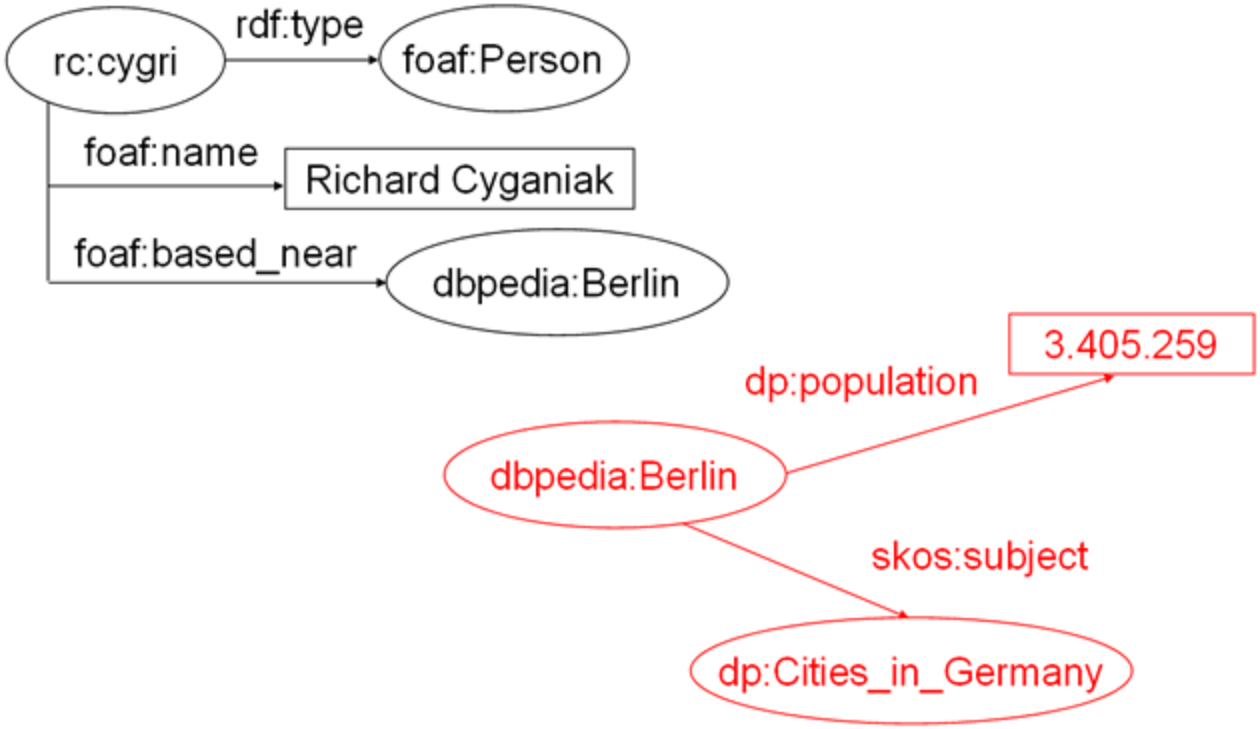


Linked Data Anwendungen

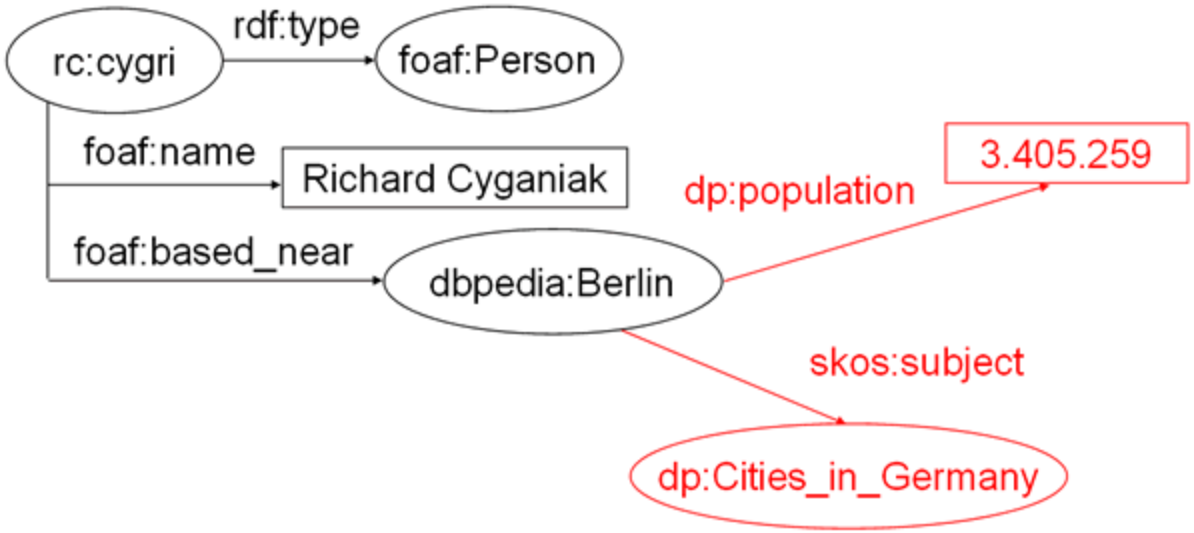




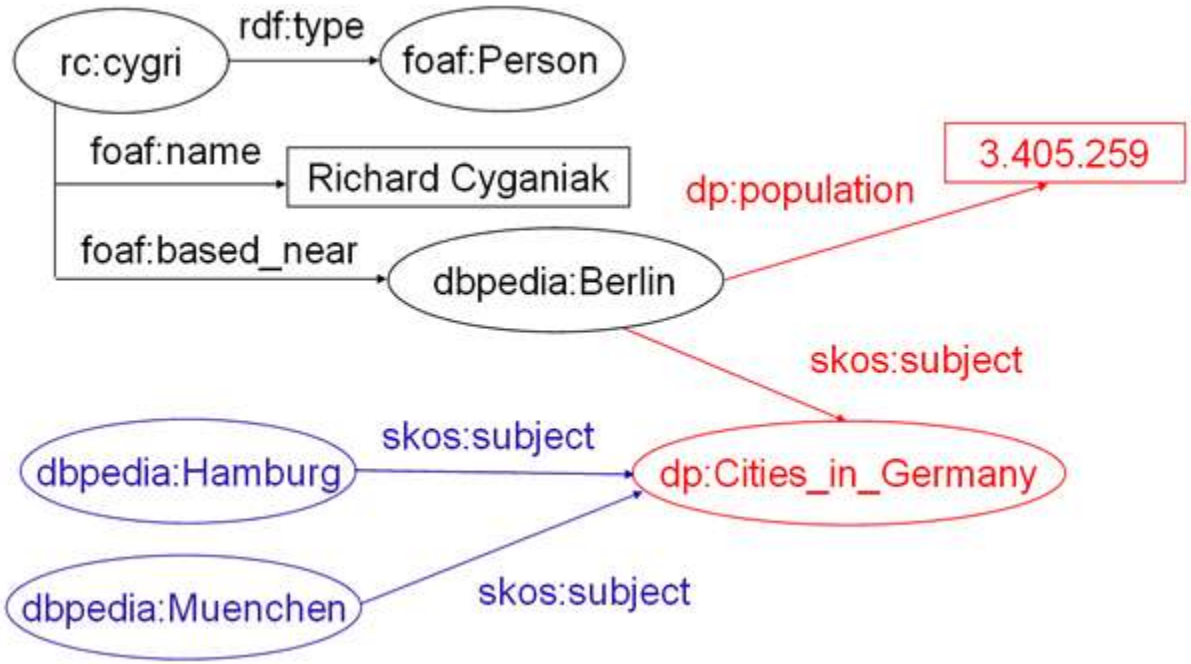
Linked Data Integration



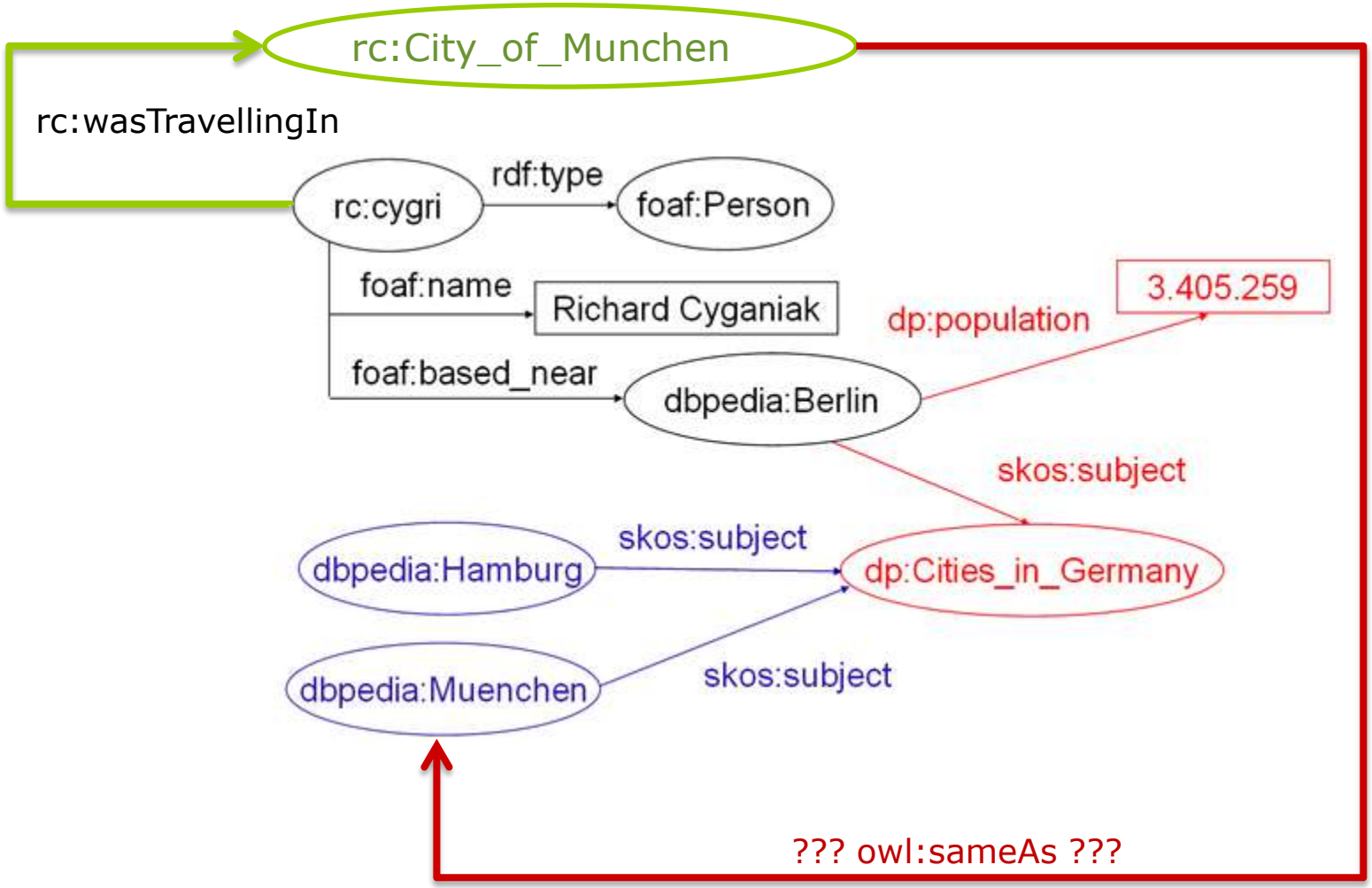
Linked Data Integration



Linked Data Integration



Linked Data Integration



Instanzmapping

- owl:sameAs
- owl:equivalentClass
- owl:equivalentProperty

Schemamapping

Shopping-Ergebnisse für **galaxy tab**

Samsung **Galaxy Tab** 10.1 16 GB - Android 3.0 (Honeycomb ...
★★★★½ 10 Erfahrungsberichte - 354 € - 59 Anbieter

Samsung **Galaxy Tab** WiFi 16 GB - Android 2.2 1 GHz
★★★★½ 18 Erfahrungsberichte - 279 € - 100 Anbieter

Samsung **Galaxy Tab** 16 GB - Android 2.2 1 GHz
★★★★☆ 33 Erfahrungsberichte - 222 € - 94 Anbieter

oliver kahn birthday

About 222,000 results (0.43 seconds)

June 15, 1969
Oliver Kahn, Date of birth
Show details

Oliver Kahn (Eine Vuvuzela zum Geburtstag) "Happy **Birthday** Olli ...

www.youtube.com/watch?v=kEdvFUoPJdo
15 Jun 2010 - 2 min - Uploaded by Hallodeutschland10
Oliver Kahn ist heute(15.06.2010) . 41 geworden und die

Microformats

- strukturierte Daten **auf einfachste Weise** in HTML einbetten
  - zwei einfache Einheiten
    - Entities
    - Properties
  - drei wichtige HTML-Attribute
    - class
    - rel
    - rev
  - vordefinierte Standardvokabulare (Erweiterung durch poshformats; microformats $\subset$ poshformats, POSH = Plain Old Semantic HTML)

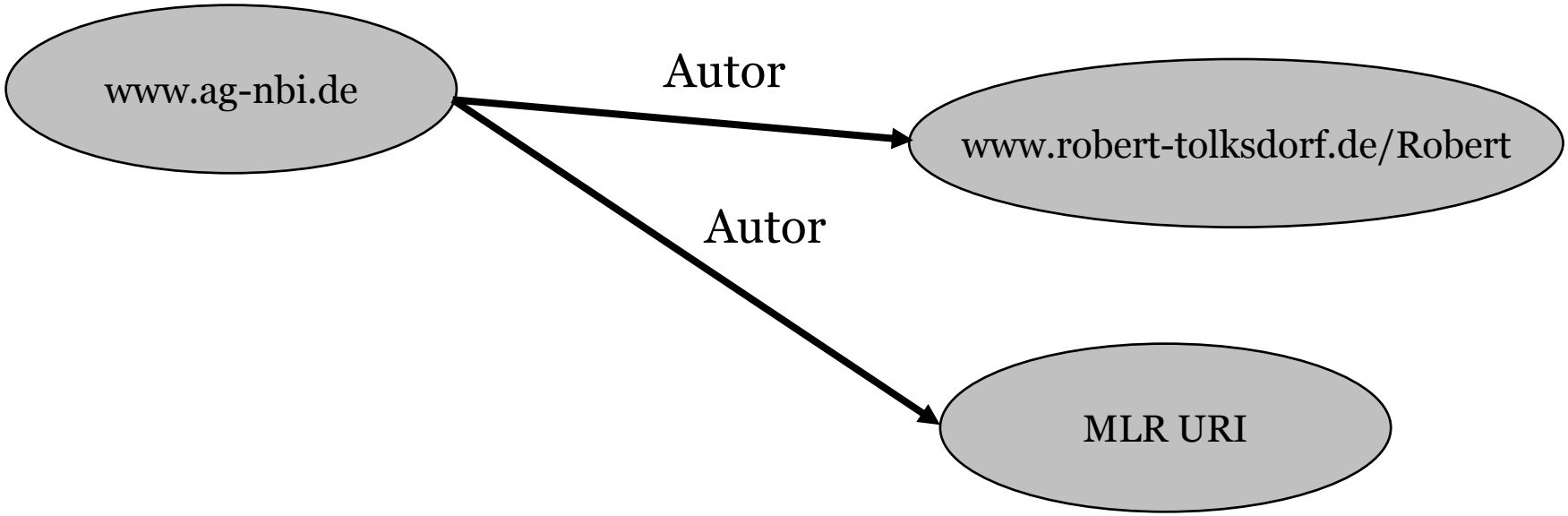
Beispiel

- Robert Tolksdorf und Markus Luczak-Rösch haben die Arbeitsgruppen-Web-Site erstellt.

```
<div>
        <a href=„http://www.robert-tolksdorf.de/Robert">
                Robert Tolksdorf
        </a>
        und
         <a href="http://www.markus-luczak.de/#me">
                Markus Luczak-Rösch
        </a>
        haben die Arbeitsgruppen-Web-Site erstellt.
</div>
```

```
<div>
        <a href="http://www.robert-tolksdorf.de/Robert"
                rel="author">
                Robert Tolksdorf
        </a>
        und
         <a href="http://www.markus-luczak.de/"
                rel="author">
                Markus Luczak-Rösch
        </a>
        haben die Arbeitsgruppen-Web-Site erstellt.
</div>
```

Microformats



Microformats

```
<div>
        <a href="http://www.markus-luczak.de/" >
                Markus Luczak-Rösch
        </a>
        arbeitet für die
        <a href="http://www.fu-berlin.de/" >
                FU Berlin
        </a>.
</div>
```

```
<div><span class="vcard">
        <a class="fn url"
                href="http://www.markus-luczak.de/" >
                Markus Luczak-Rösch
        </a></span>
        arbeitet für die
        <span class="vcard">
                <a class="fn org url"
                        href="http://www.fu-berlin.de/" >
                        FU Berlin
                </a>
        </span>.
</div>
```

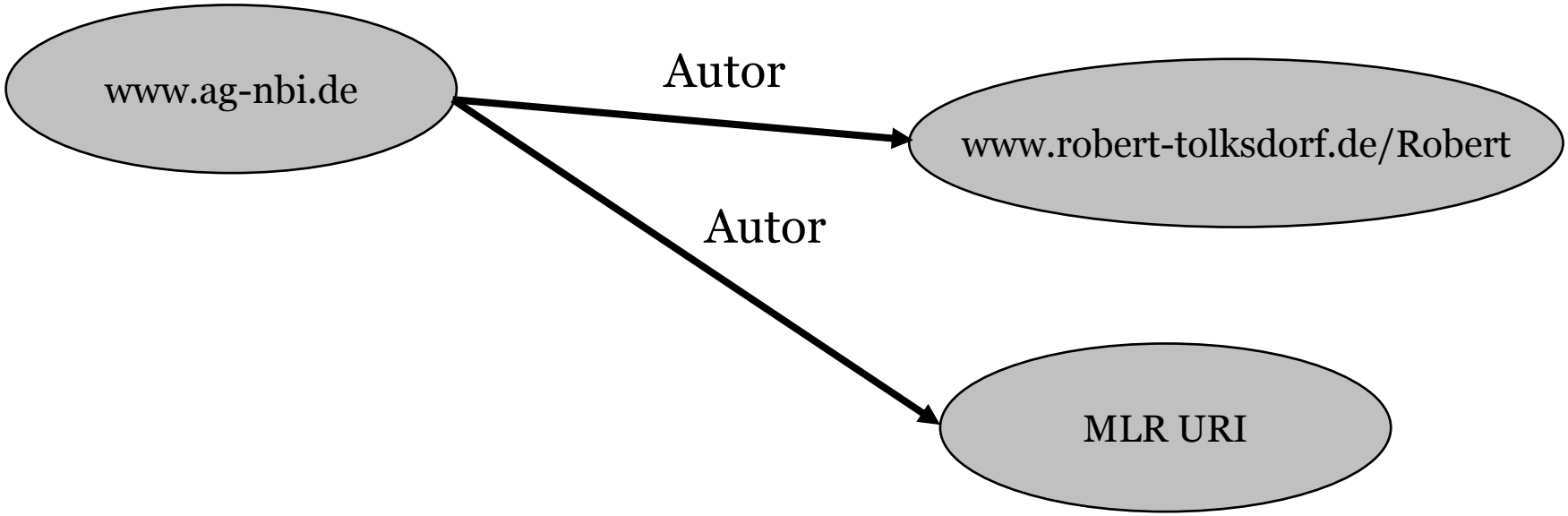
```
<div class="vcard">
        <span class="fn" >
                Markus Luczak-Rösch
        </span>
        arbeitet für die
        <span class="org" >
                FU Berlin
        </span>.
</div>
```

```
<div class="vcard">
        <a class="fn url"
                href="http://www.markus-luczak.de/" >
                Markus Luczak-Rösch
        </a>
        arbeitet für die
        <span class="org vcard" >
                <a class= "url fn org"
                        href= "http://www.fu-berlin.de" >
                FU Berlin</a>
        </span>.
</div>
```

- **RDF Daten** in HTML einbetten
  - vollständiges RDF Datenmodell
  - Mischung von Schemata möglich
  - wichtige HTML-Attribute
    - about
    - property
    - rel
    - rev

```
<div about="" rel="dcterms:author">
        <a href=„http://www.robert-tolksdorf.de/Robert">
                Robert Tolksdorf
        </a>
        und
         <a href="http://www.markus-luczak.de/#me">
                Markus Luczak-Rösch
        </a>
        haben die Arbeitsgruppen-Web-Site erstellt.
</div>
```

<> dcterms:author < http://www.robert-tolksdorf.de/Robert >,
< http://www.markus-luczak.de/#me > .



Microdata

- einfache Art, um strukturierte Daten („maschinenlesbare Labels“) in HTML einzubetten
  - Standard ab HTML 5
    - itemprop
    - itemref
    - content
    - itemscope
    - itemtype
  - verschachtelte Name-Wert-Paare
    - Gruppe von Name-Wert-Paaren → Item
    - atomares Name-Wert-Paar → Property
  - Zugriff via Microdata DOM API

Microdata

```
<div>
        <a href=„http://www.robert-tolksdorf.de/Robert">
                Robert Tolksdorf
        </a>
        und
         <a href="http://www.markus-luczak.de/#me">
                Markus Luczak-Rösch
        </a>
        haben die Arbeitsgruppen-Web-Site erstellt.
</div>
```

Microdata

```
<div><span itemscope>
        <a itemprop="url"
                href="http://www.robert-tolksdorf.de/Robert">
                <span itemprop="name">
                        Robert Tolksdorf </span>
        </a> </span>
        und <span itemscope>
        <a itemprop="url"
                href="http://www.markus-luczak.de/#me">
                 <span itemprop="name">
                Markus Luczak-Rösch </span>
        </a> </span>
        haben die Arbeitsgruppen-Web-Site erstellt.
</div>
```

Microdata

```
<div><span itemscope
        itemtype="http://example.org/people/professor">
        <span itemprop="name">
                Robert Tolksdorf
        </span> </span>
        und
        <span itemscope
                itemtype="http://example.org/people/lecturer">
                <span itemprop="name">
                Markus Luczak-Rösch</span> </span>
        haben die Arbeitsgruppen-Web-Site erstellt.
</div>
```

Microdata

```
<div><span itemscope
        itemtype="http://example.org/people/professor„
        itemid= "http://www.robert-tolksdorf.de/Robert">
        <a itemprop="name"
                href="http://www.robert-tolksdorf.de/Robert">
                Robert Tolksdorf
        </a> </span>
        und
        …
</div>
```

Microdata

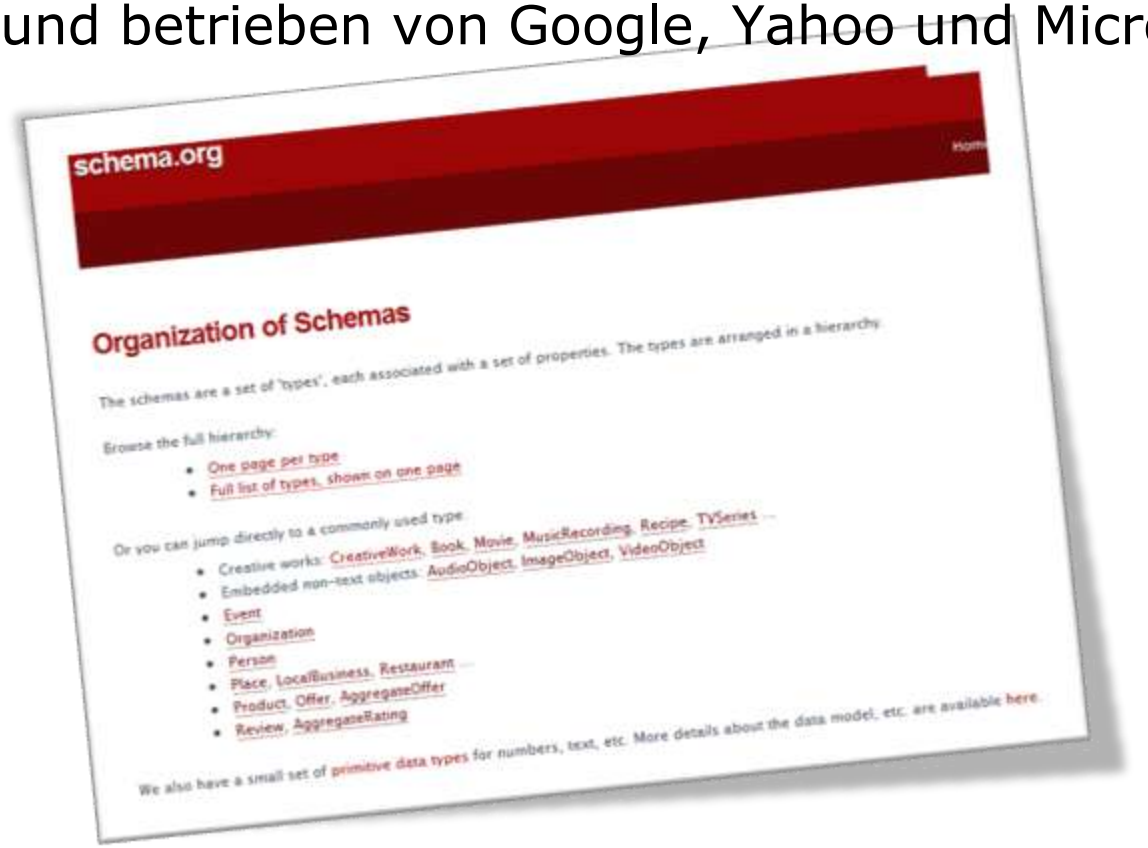
```
<div><span itemscope
        itemtype="http://example.org/people/professor„
        itemid= "http://www.robert-tolksdorf.de/Robert">
        <span itemprop="name">
                Robert Tolksdorf
        </span>
         <meta itemprop="knows"
                content="Markus Luczak-Rösch" />
        </span>
        und
        …
</div>
```

Microdata

```
<div><span itemscope
        itemtype="http://example.org/people/professor„
        itemid= "http://www.robert-tolksdorf.de/Robert">
        <span itemprop="name">
                itemref="meta-knows"
                Robert Tolksdorf
        </span>
         <meta id="meta-knows" itemprop="knows"
                content="Markus Luczak-Rösch" />
        </span>
        und
        …
</div>
```

Freie Universität Berlin

- Microdata Vokabular für die (meisten) prominenten Anwendungsfälle im Web
  - erstellt und betrieben von Google, Yahoo und Microsoft



Microdata, RDFa, Microformats

Freie Universität Berlin

Feature	RDFa 1.1	Microdata 1.0	Microformats 1.0
Relative Complexity	High	Medium	Low
Data Model	Graph	Tree	Tree
Item optionally identified by IRI	Yes	Yes	No
Item type optionally specified by IRI	Yes	Yes	No
Item properties specified by IRI	Yes	Yes	No
Multiple objects per page	Yes	Yes	Yes
Overlapping objects	Yes	Yes	No
Plain Text properties	Yes	Yes	Yes
IRI properties	Yes	Yes*	No
Typed Literal properties	Yes	No	No
XML Literal properties	Yes	No	No
Language tagging	Yes	Yes	Inconsistent

http://manu.sporny.org/2011/uber-comparison-rdfa-md-uf/

Microdata, RDFa, Microformats

Freie Universität Berlin

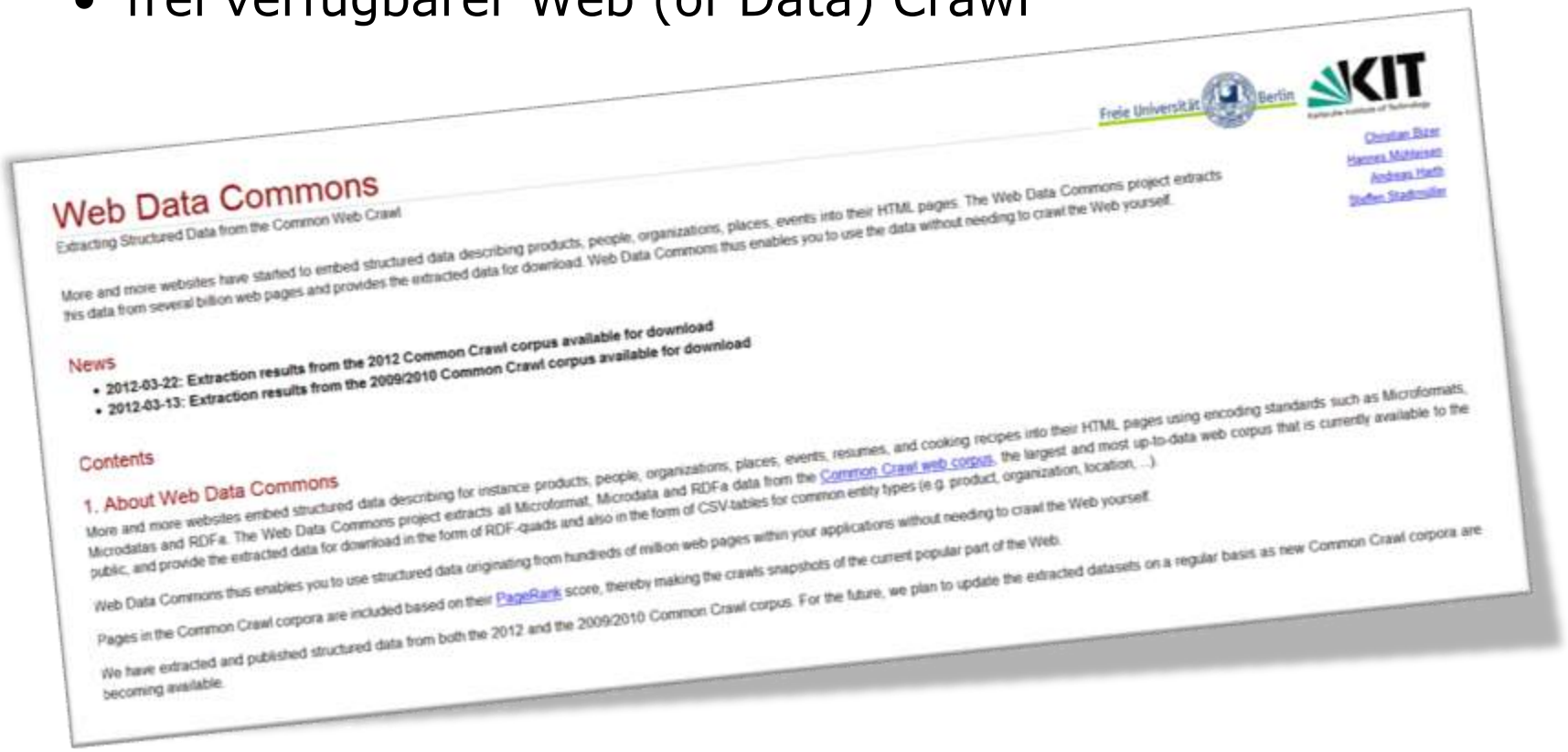
Feature	RDFa 1.1	Microdata 1.0	Microformats 1.0
Override text and IRI content	Yes	No	Text only
Clear mapping to RDF	Yes	Problematic	No
Target Languages	8 (XHTML1, HTML4, HTML5, XHTML5, XML, SVG, ePub, OpenDocument)	2 (HTML5, XHTML5)	4 (XHTML1, HTML4, HTML5, XHTML5)
New Attributes	8 about, datatype, profile, prefix, property, resource, typeof, vocab	5 itemid, itemprop, itemref, itemscope, itemtype	0
Re-used Attributes	5 content, href, rel, rev, src	5 content, src, href, data, datetime	4 class, title, rel, href
Multiple IRI types per object	Yes	RDF only	No
Multiple statements per element	Yes	No	Yes
"Locally scoped" vocabulary terms	Yes, via vocab	Yes, via itemscope	No
Item Chaining	Yes	Basic	No

Microdata, RDFa, Microformats

Freie Universität Berlin

Feature	RDFa 1.1	Microdata 1.0	Microformats 1.0
Transclusion	No	Yes	Yes, via include pattern
Compact IRIs	Yes	No	No
Prefix rebinding	Yes	No	No
Vocabulary Mashups	Yes	No	No
HTML5 time element support	Not yet	Yes	No
Different attributes for different property types	Yes property for text, rel/rev for URLs, resource/content for overrides	No	Yes class for text and rel for URLs
Transform to JSON	Yes (RDFa API)	Yes (Parser and Microdata DOM API)	No
DOM API	Yes	Yes	No
Unified Parser	Yes	Yes	No

- Extraktion strukturierter Daten aus dem Common Crawl Korpus
  - frei verfügbarer Web (of Data) Crawl



Web Data Commons

Extracting Structured Data from the Common Web Crawl

More and more websites have started to embed structured data describing products, people, organizations, places, events into their HTML pages. The Web Data Commons project extracts this data from several billion web pages and provides the extracted data for download. Web Data Commons thus enables you to use the data without needing to crawl the Web yourself.

Freie Universität Berlin KIT

Christian Bizer
Hannes Mühleisen
Andreas Harth
Steffen Stadtmüller

News

- 2012-03-22: Extraction results from the 2012 Common Crawl corpus available for download
- 2012-03-13: Extraction results from the 2009/2010 Common Crawl corpus available for download

Contents

1. About Web Data Commons

More and more websites embed structured data describing for instance products, people, organizations, places, events, resumes, and cooking recipes into their HTML pages using encoding standards such as Microformats, Microdatas and RDFa. The Web Data Commons project extracts all Microformat, Microdata and RDFa data from the Common Crawl web corpus, the largest and most up-to-data web corpus that is currently available to the public, and provide the extracted data for download in the form of RDF-quads and also in the form of CSV-tables for common entity types (e.g. product, organization, location, ...).

Web Data Commons thus enables you to use structured data originating from hundreds of million web pages within your applications without needing to crawl the Web yourself.

Pages in the Common Crawl corpora are included based on their PageRank score, thereby making the crawls snapshots of the current popular part of the Web.

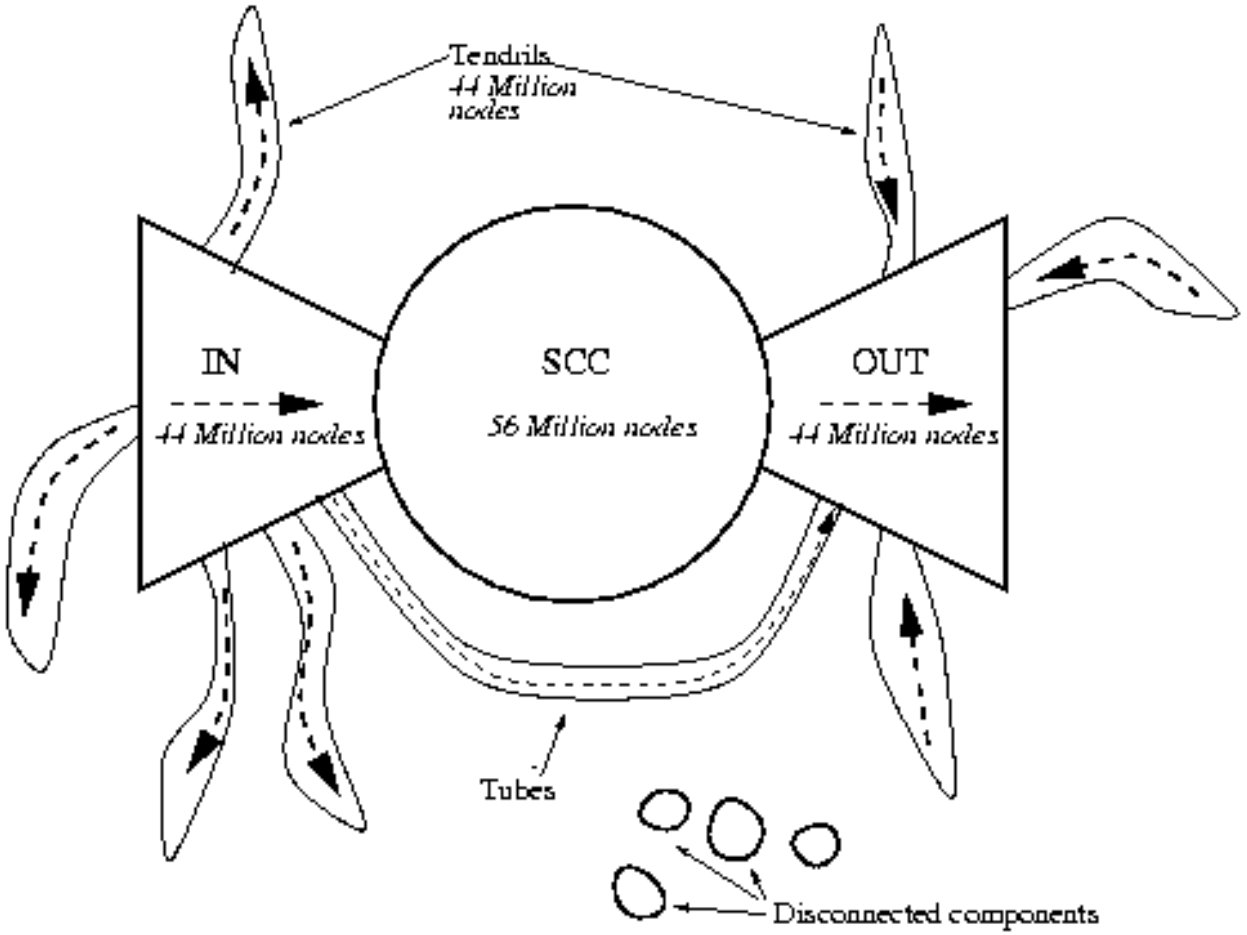
We have extracted and published structured data from both the 2012 and the 2009/2010 Common Crawl corpus. For the future, we plan to update the extracted datasets on a regular basis as new Common Crawl corpora are becoming available.

"…the web is not the ball of highly-connected spaghetti we believed it to be; rather, the connectivity is strongly limited by **a high-level global structure**."

Graph structure in the Web. Andrei Broder, Ravi Kumar, Farzin Maghoul, Prabhakar Raghavan, Sridhar Rajagopalan, Raymie Stata, Andrew Tomkins, and Janet Wiener. Comput. Netw. (June 2000). http://www9.org/w9cdrom/160/160.html

http://mrg.bz/DnycUn

A well-known topology of the Web



Graph structure in the Web. Andrei Broder, Ravi Kumar, Farzin Maghoul, Prabhakar Raghavan, Sridhar Rajagopalan, Raymie Stata, Andrew Tomkins, and Janet Wiener. Comput. Netw. (June 2000). http://www9.org/w9cdrom/160/160.html

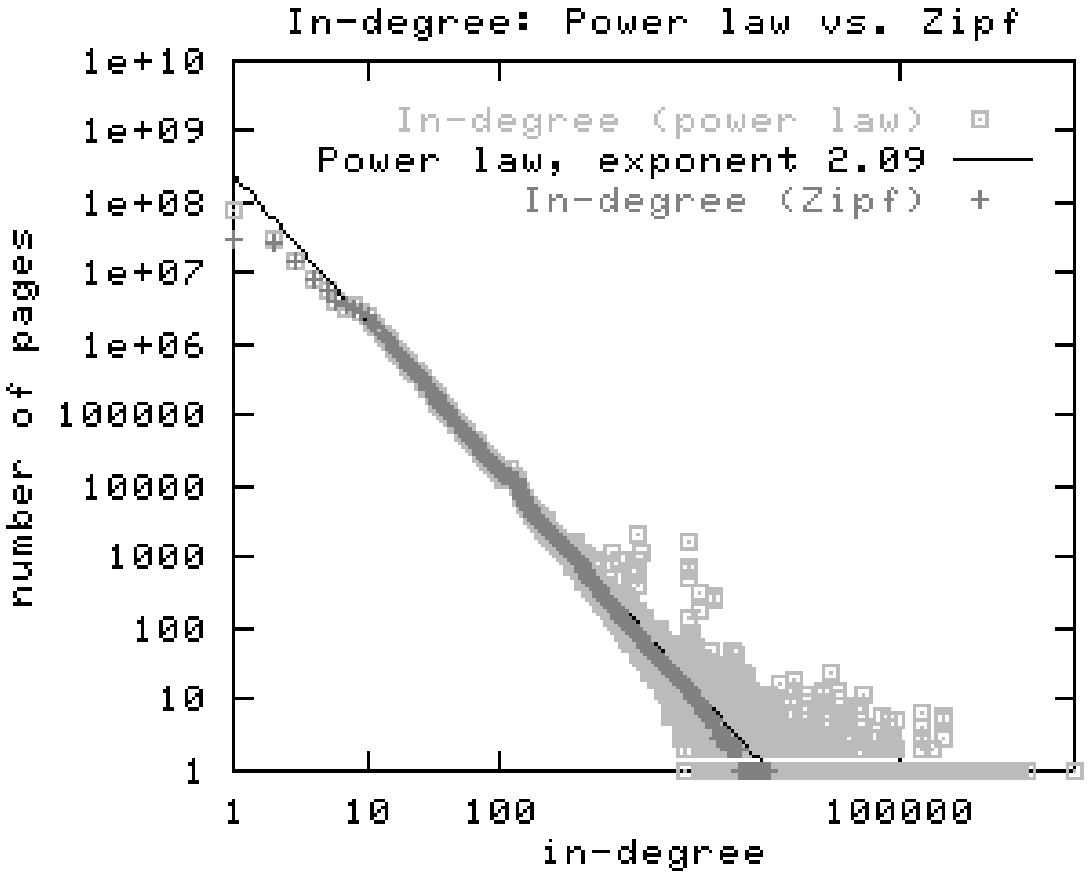
http://whatever.com/abc

http://example.org/index.html

Node A

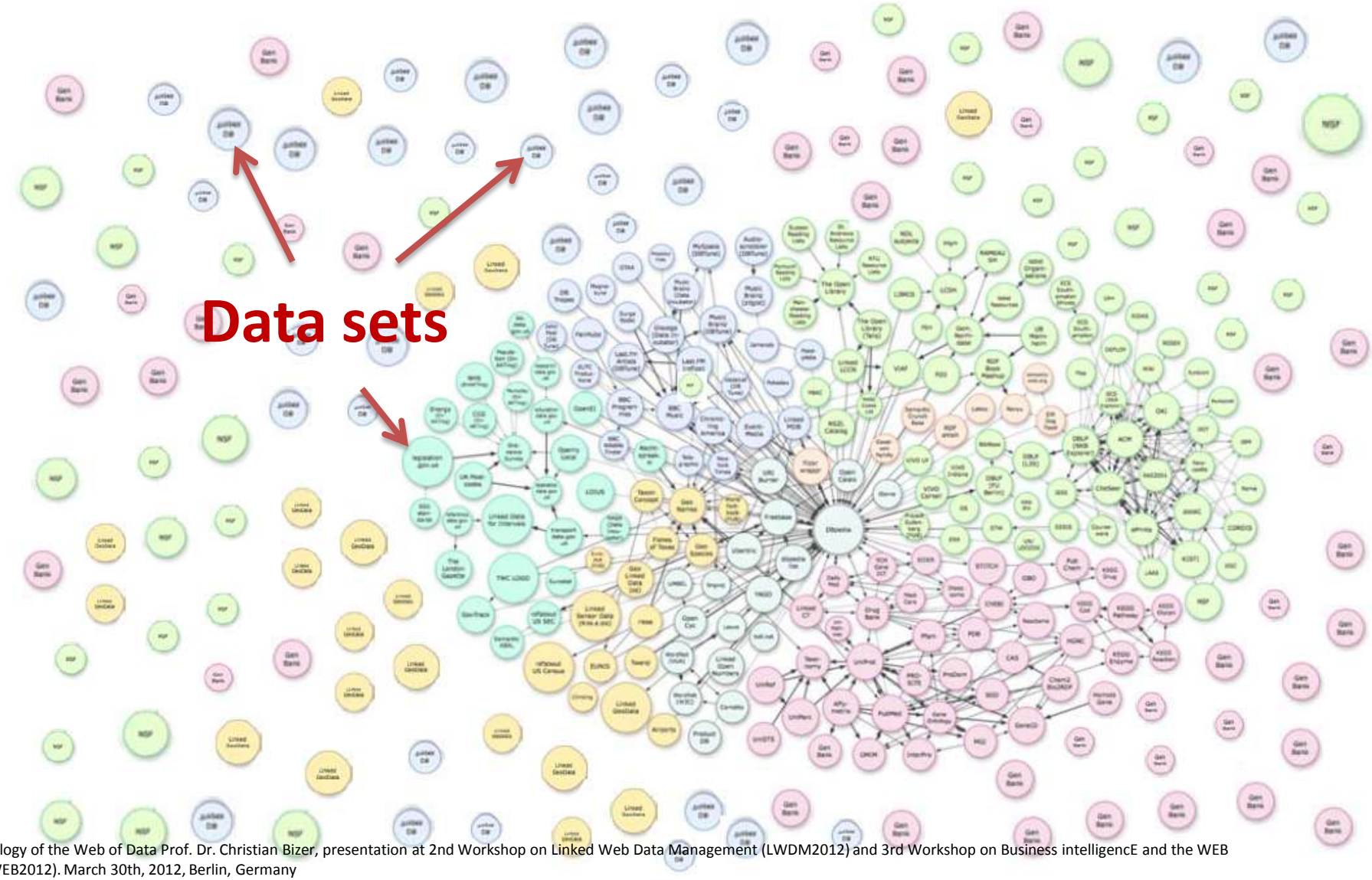
Node B

The Web is scale free

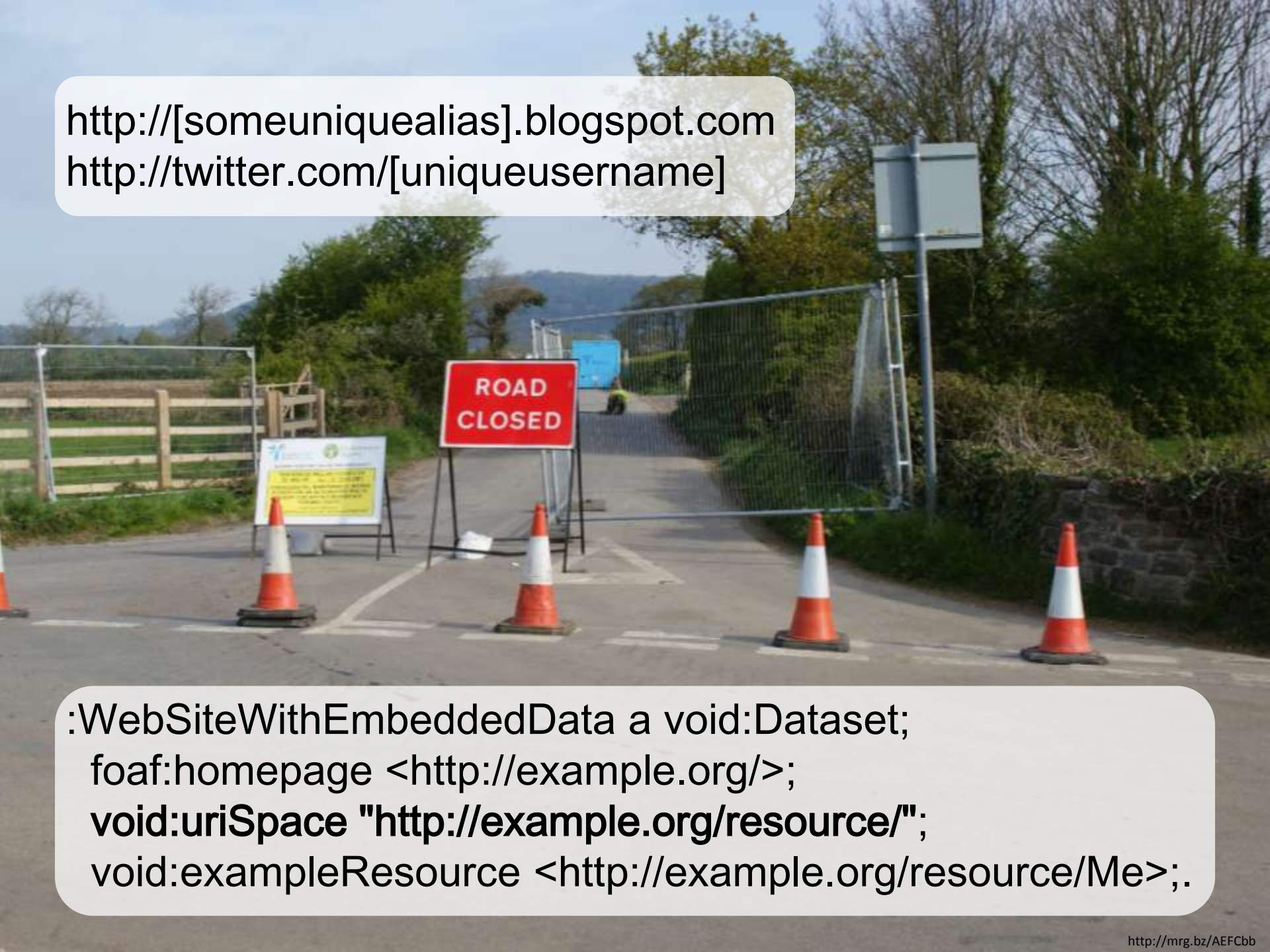


Graph structure in the Web. Andrei Broder, Ravi Kumar, Farzin Maghoul, Prabhakar Raghavan, Sridhar Rajagopalan, Raymie Stata, Andrew Tomkins, and Janet Wiener. Comput. Netw. (June 2000). http://www9.org/w9cdrom/160/160.html

Topology of the Web of Data II



Topology of the Web of Data Prof. Dr. Christian Bizer, presentation at 2nd Workshop on Linked Web Data Management (LWDM2012) and 3rd Workshop on Business intelligencE and the WEB (BEWEB2012). March 30th, 2012, Berlin, Germany



http://[someuniquealias].blogspot.com
http://twitter.com/[uniqueusername]

```
:WebSiteWithEmbeddedData a void:Dataset;
  foaf:homepage <http://example.org/>;
  void:uriSpace "http://example.org/resource/";
  void:exampleResource <http://example.org/resource/Me>;.
```

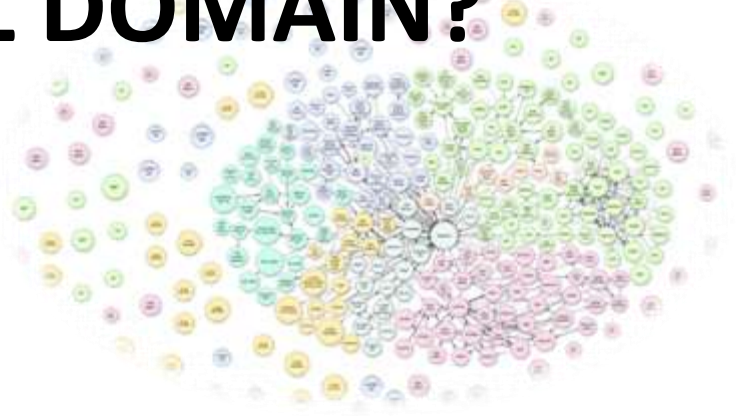
http://mrg.bz/AEFCbb

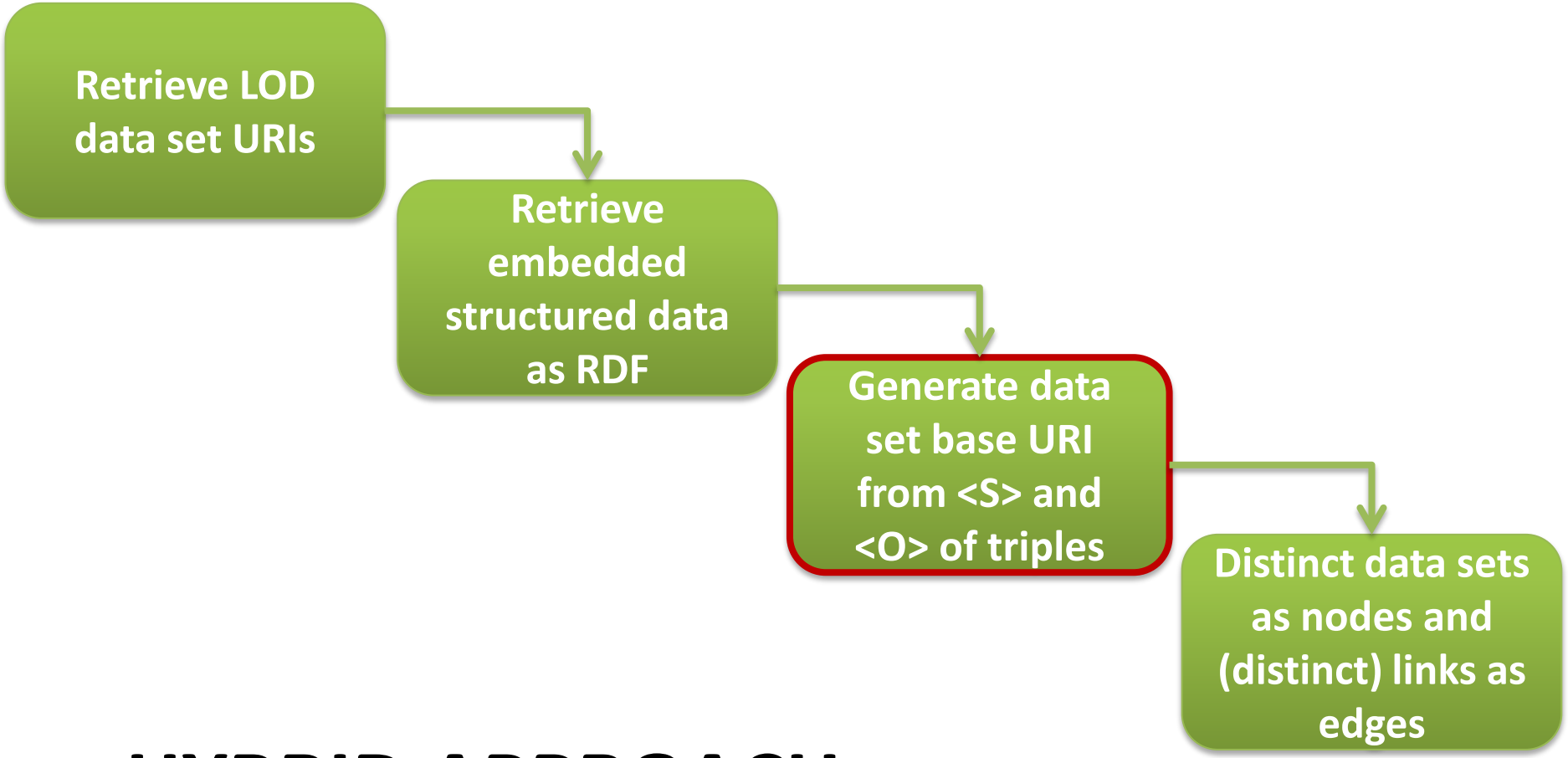
http://www4.wiwiss.fu-berlin.de/stitch/
http://www4.wiwiss.fu-berlin.de/dblp/

---

fu-berlin.de

DATA SET = PAY LEVEL DOMAIN?





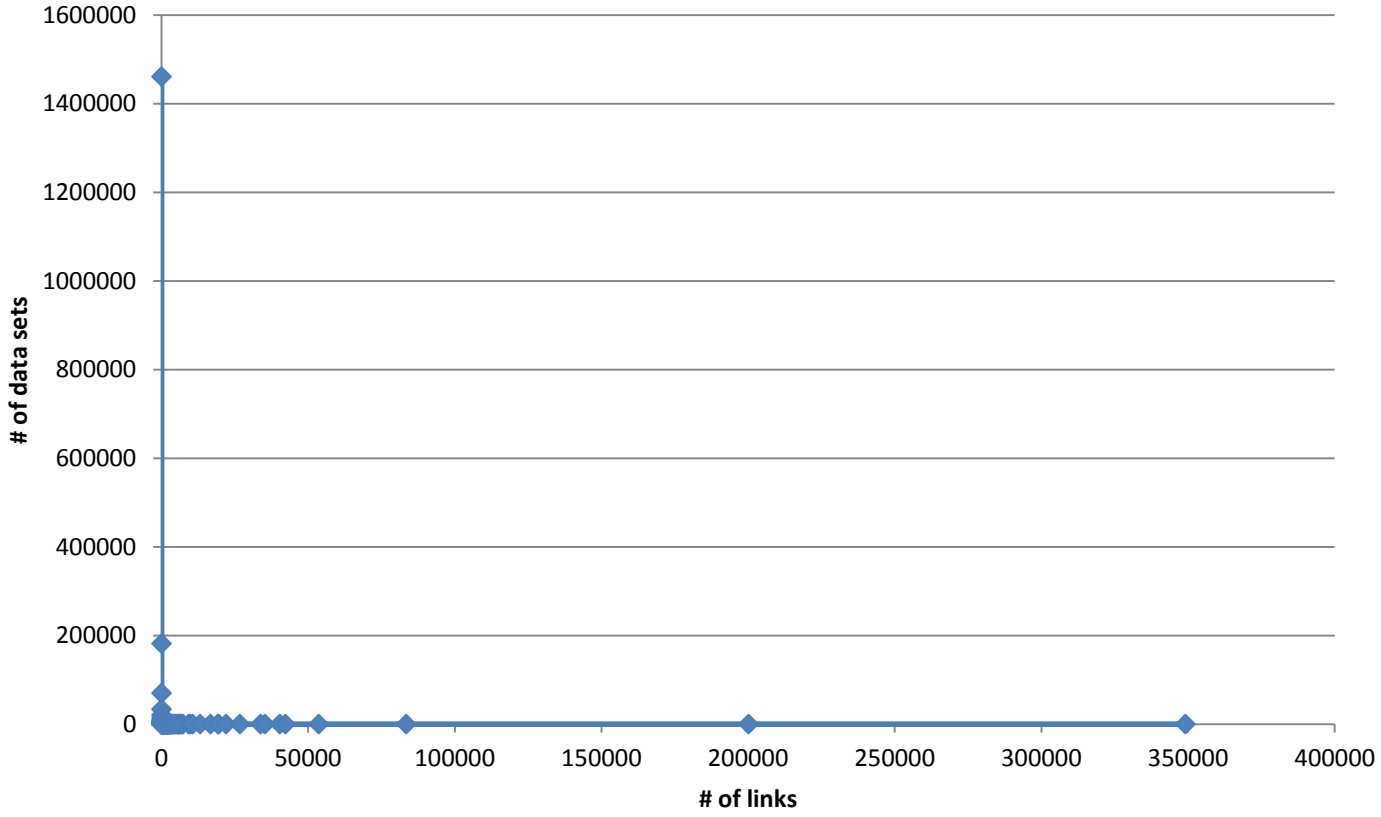
HYBRID APPROACH: LOD DATASET OR SUBDOMAIN

	WDC 2012 Crawl
# of LOD data sets	328
# of distinct links incl. Links served by a third party	2.680.692
# of distinct links served by source or target data set	2.668.851
# of nodes	1.835.909
#nodes with degree ≥ 1	**1.835.698** (99, 99%)
# of edges	2.254.269

Node id (data set base URI, "http://" stripped to save space)	**in-degree**	**out-degree**	**degree**
www.youtube.com/	13327	335757	349084
www.flickr.com/	8669	191465	200134
dbpedia.org/ *	34	83397	83431
friendfeed.com/	1628	51964	53592
www.w3.org/	42067	164	42231
twitter.com/	32935	7371	40306
www.facebook.com/	35240	44	35284
soundcloud.com/	491	33269	33760
xmlns.com/	26712	1	26713
purl.org/	21872	123	21995
rdfs.org/	19517	0	19517
profile.typepad.com/	19192	0	19192
www.last.fm/	2452	14253	16705
www.identi.ca/	3	13157	13160
www.linkedin.com/	10627	1	10628
technorati.com/	9786	0	9786
www.europages.it/	1	9250	9251
www.lastfm.de/	134	7147	7281
www.meadiciona.com/	42	7122	7164
www.lastfm.fr/	54	6623	6677
www.twitter.com/	6459	0	6459
www.threadless.com/	19	6153	6172
www.europages.fr/	2	6053	6055
www.lastfm.com.br/	48	5803	5851
www.europages.com/	4	5605	5609
www.myspace.com/	5507	3	5510
www.lastfm.se/	13	5465	5478
www.xing.com/	5470	1	5471
blip.fm/	208	4643	4851
http://rdf.data-vocabulary.org/	4823	0	4823
...			
www.bbc.co.uk/ programmes/ *	10	1105	1115

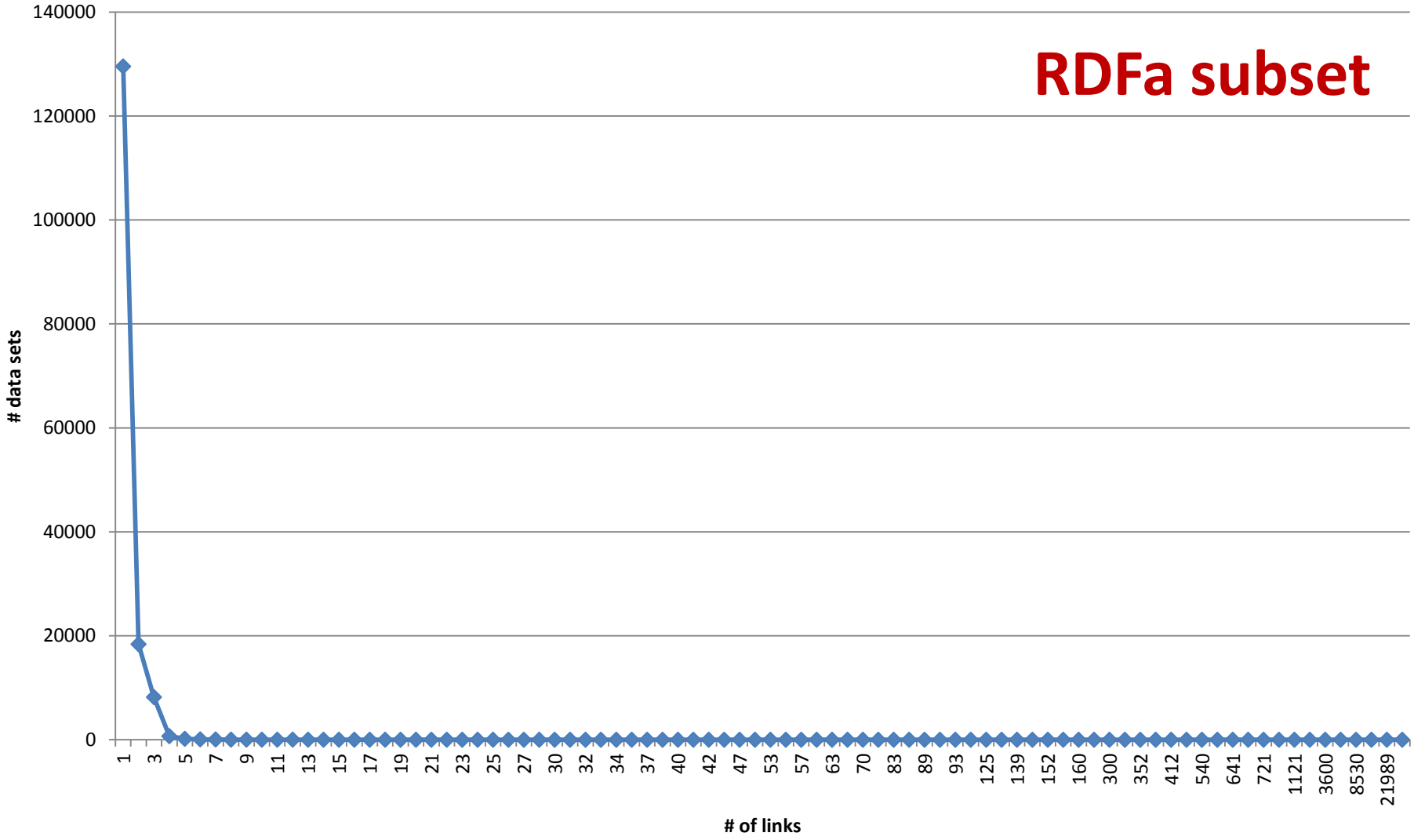
	WDC 2012 Crawl RDFa Subset
# of LOD data sets	328
# of distinct links incl. Links served by a third party	225.081
# of distinct links served by source or target data set	216.313
# of nodes	157.638
# of nodes with degree ≥ 1	**157.422** (99, 86%)
# of edges	189.653

Node id (data set base URI, "http://" stripped to save space)	**in-degree**	**out-degree**	**degree**	**modularity class**
dbpedia.org/ *	34	83397	83431	0
xmlns.com/	26712	1	26713	7
purl.org/	21866	123	21989	4
rdfs.org/	19516	0	19516	7
www.w3.org/	8393	137	8530	7
rdf.data-vocabulary.org/	4823	0	4823	38
www.n49.ca/	0	3600	3600	752
d1.scribdassets.com/	1381	1	1382	12
www.biologeek.com/	1	1120	1121	53
www.bbc.co.uk/ programmes/ *	2	1105	1107	60
d.scribd.com/	703	18	721	12
www.safecreative.org/	692	10	702	8
thedatahub.org/	0	641	641	60
www.dns.com.tw/	1	614	615	277
www.wind-watch.org/	0	540	540	277
www.nachi.org/	1	501	502	54
www.kaltura.com/	408	4	412	111
www.memebox.com/	0	410	410	277
www.ilcomuneinforma.it/	0	352	352	882
www.viewchange.org/	0	347	347	277
www.spineuniverse.com/	0	300	300	201
creativecommons.org/	151	27	178	35
chopan.tumblr.com/	159	1	160	4
alioth.debian.org/	1	158	159	47
www.rakuten.de/	149	3	152	4
files.rakuten.de/	150	0	150	4
elmcity.cloudapp.net/	0	139	139	54
www.youtube.com/	122	11	133	277
blog-pl.piechotka.com.pl/	0	125	125	35
www.blogger.com/	106	2	108	4

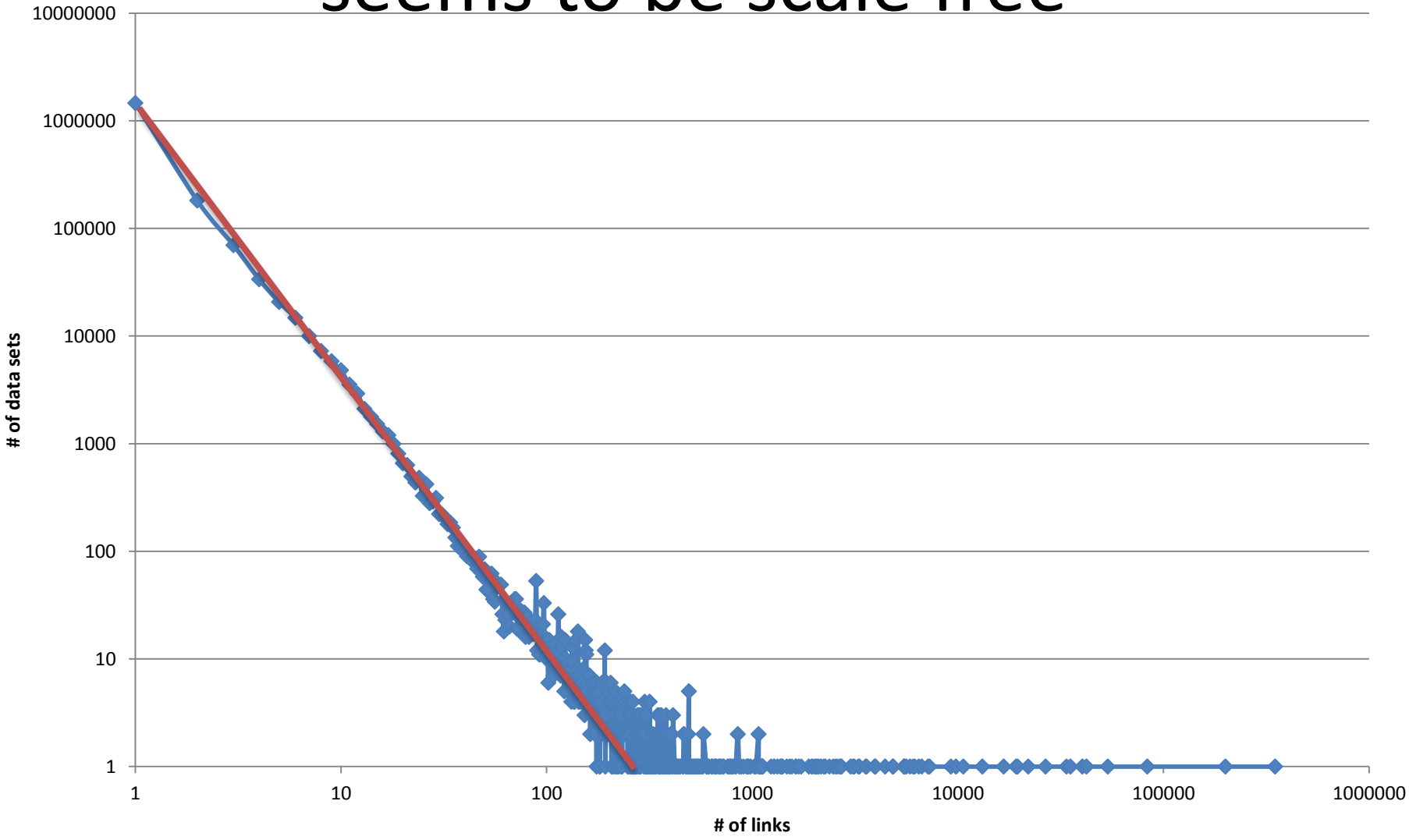


THE WEB OF DATA, A SCALE FREE NETWORK?

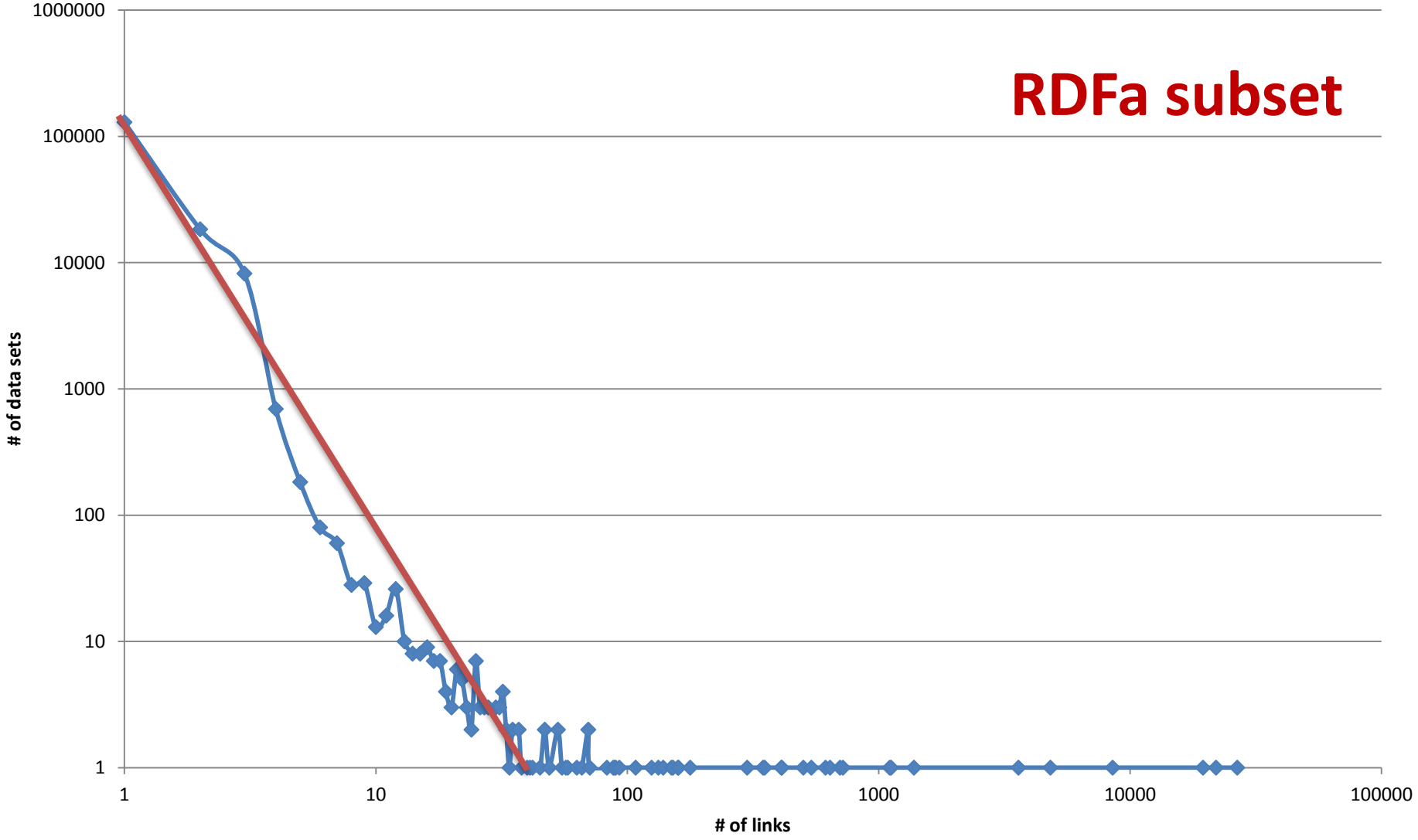
The Web of Data, a scale free network?

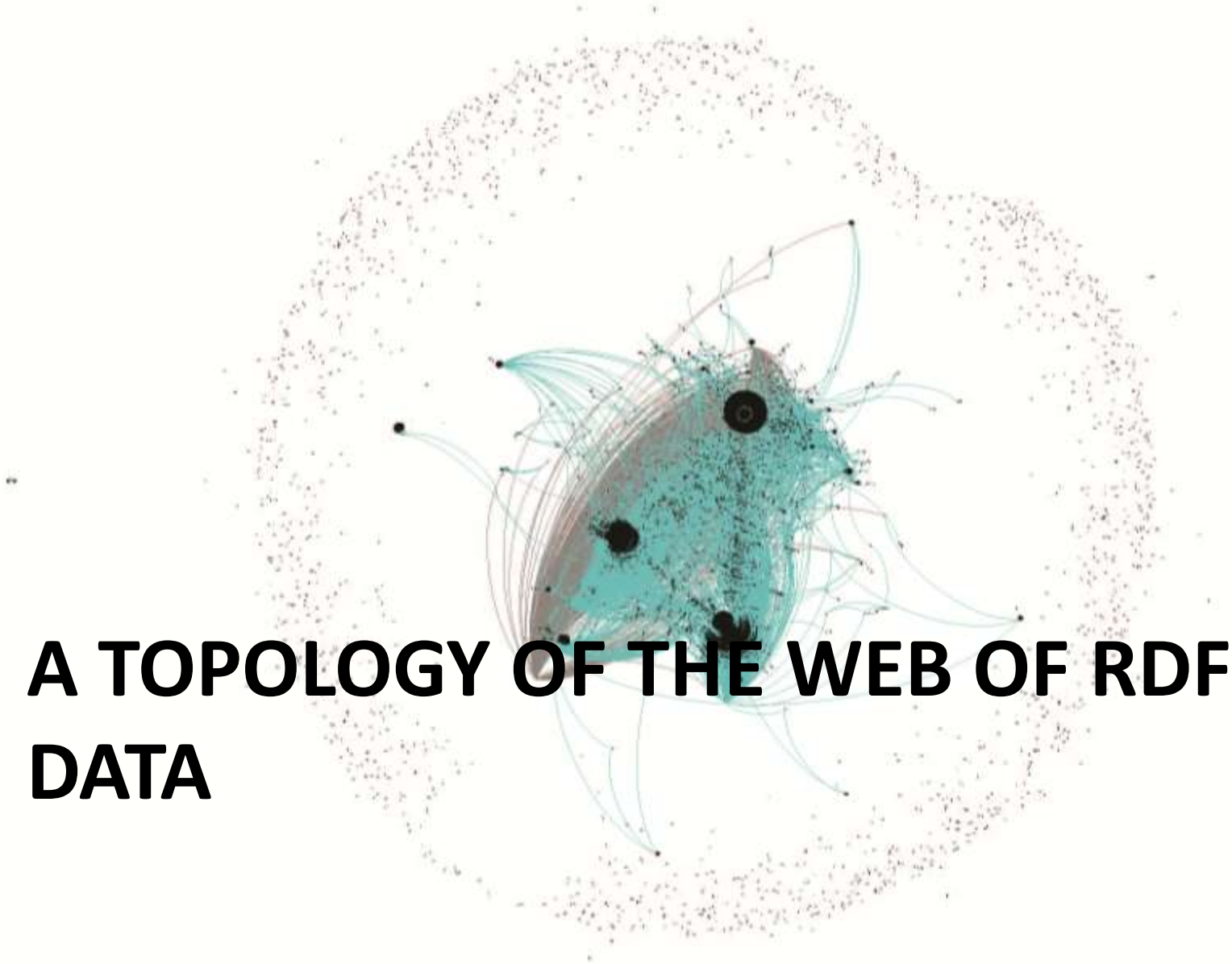


The “entire embedded“ Web of Data seems to be scale free



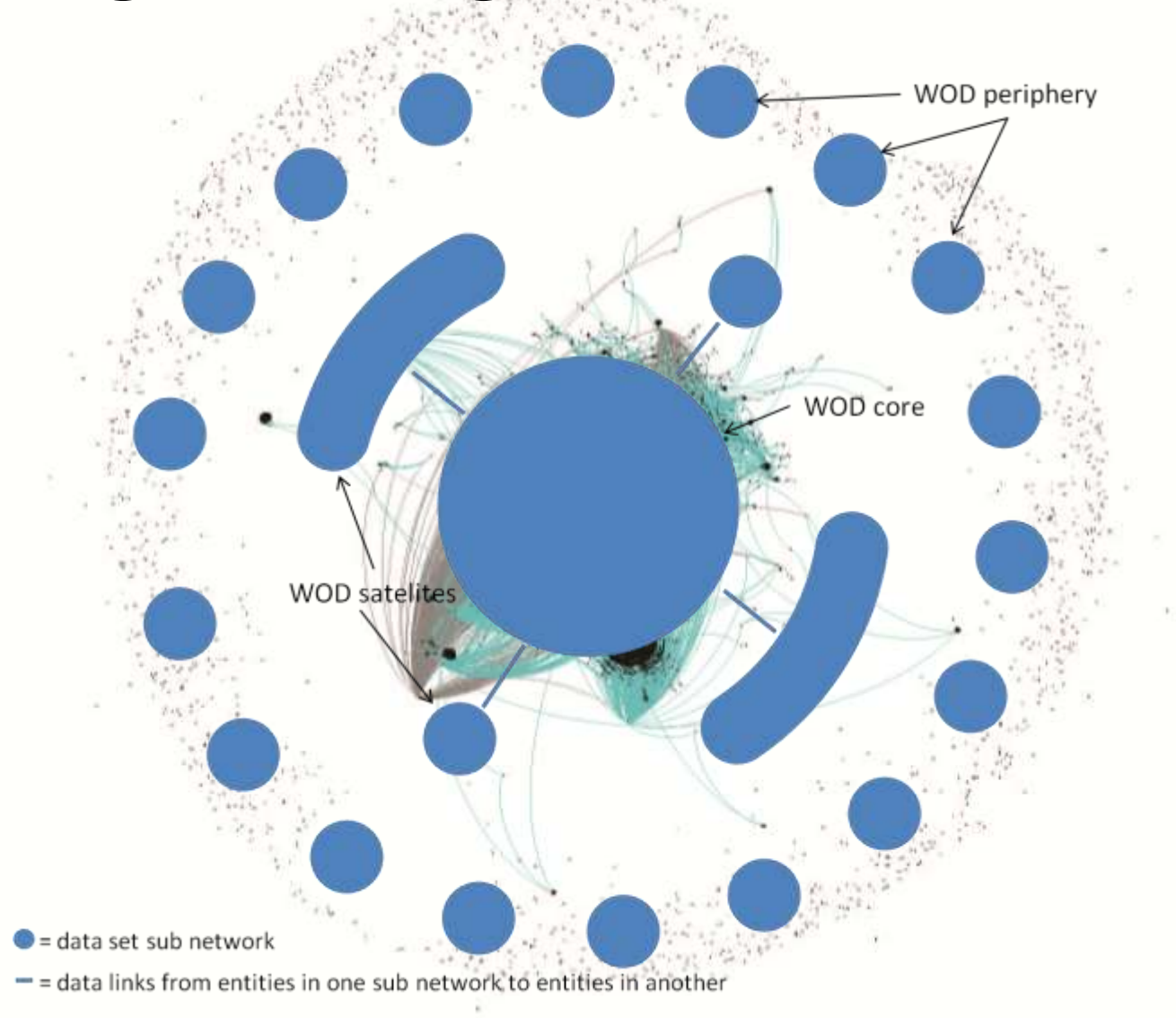
The “RDFa subset“ of the Web of Data is rather not scale free



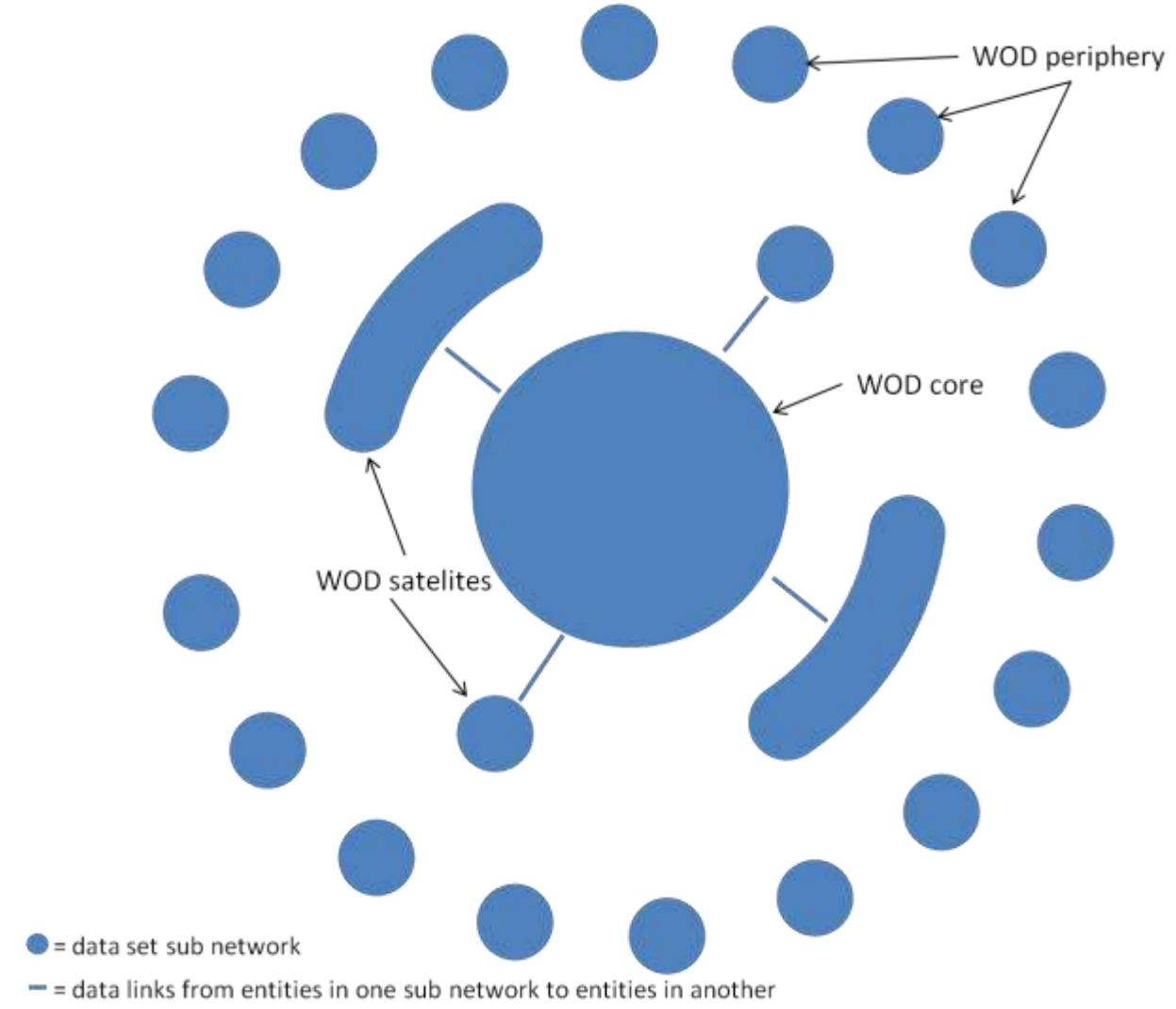


A TOPOLOGY OF THE WEB OF RDF DATA

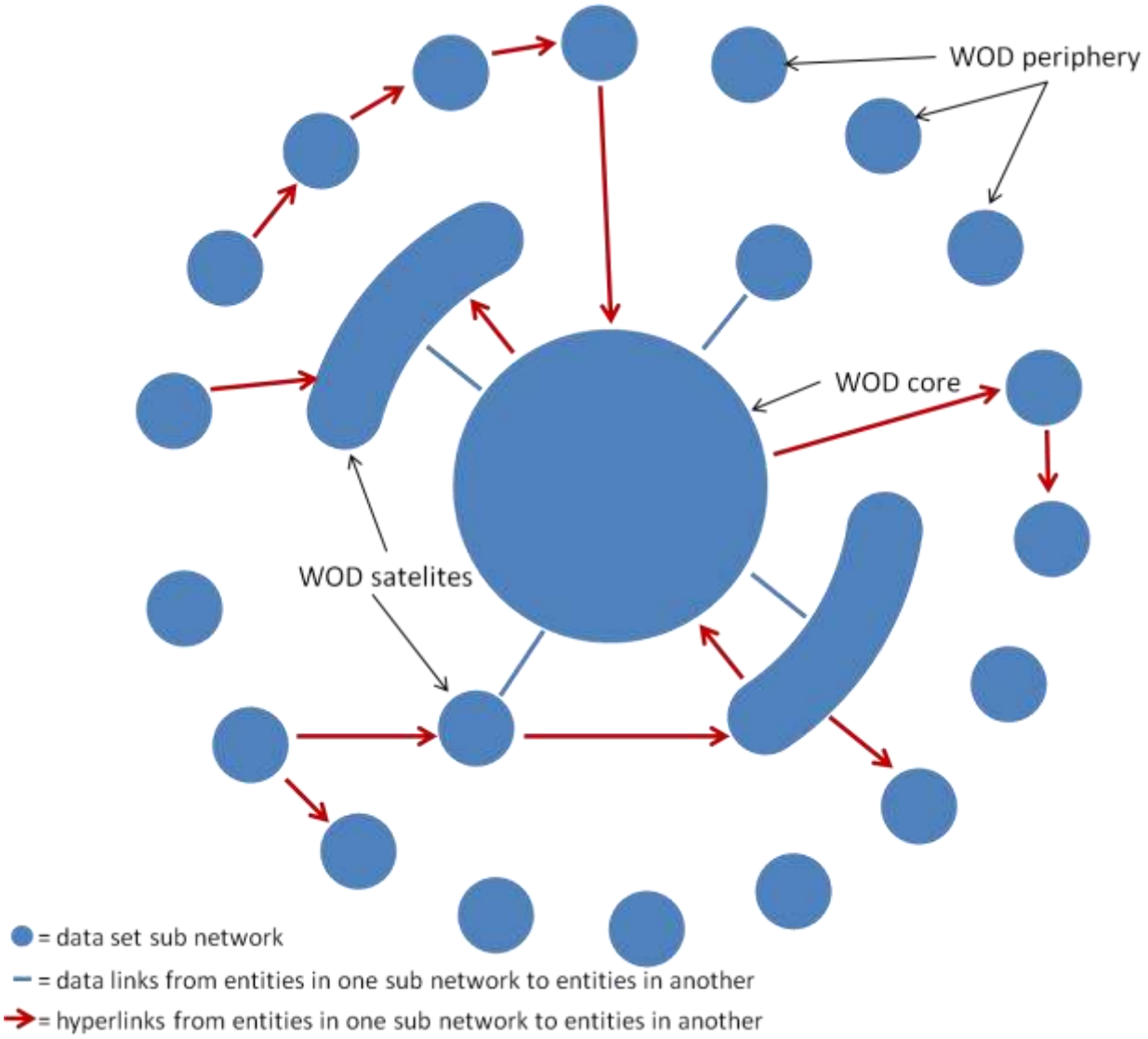
A high-level global structure?



A high-level global structure?



There is only **one Web**



Zusammenfassung

- Web of Data = Linked Data + Microformats + RDFa + Microdata

- Einfachheit (Microformats, Microdata) vs. Flexibilität (Linked Data, RDFa)

Merken!

- Linked Data Prinzipien
- URIs
- Content Negotiation
- SPARQL-Basics
- vergleich der Datenmodelle (Graph vs. Baum)

Nächste Woche

- Tutorium: Ihre Fragen
- Vorlesungstermin: Projektauftrakt