

Einführung in Linked Open Data

Felix Ostrowski, graphthinking GmbH
Adrian Pohl, hbz

16. DINI Jahrestagung, 27. und 28. Oktober 2015
Deutsche Nationalbibliothek Frankfurt/Main

Sie sind dran!

Füllen Sie ihr Namensschild aus und übertragen Sie es anschließend in das Etherpad Ihrer Gruppe.

<http://etherpad.lobid.org/p/dini-ws-X>

```
@prefix z: <http://etherpad.lobid.org/p/dini-ws-z/export/txt#> .
```

```
@prefix schema: <http://schema.org/> .
```

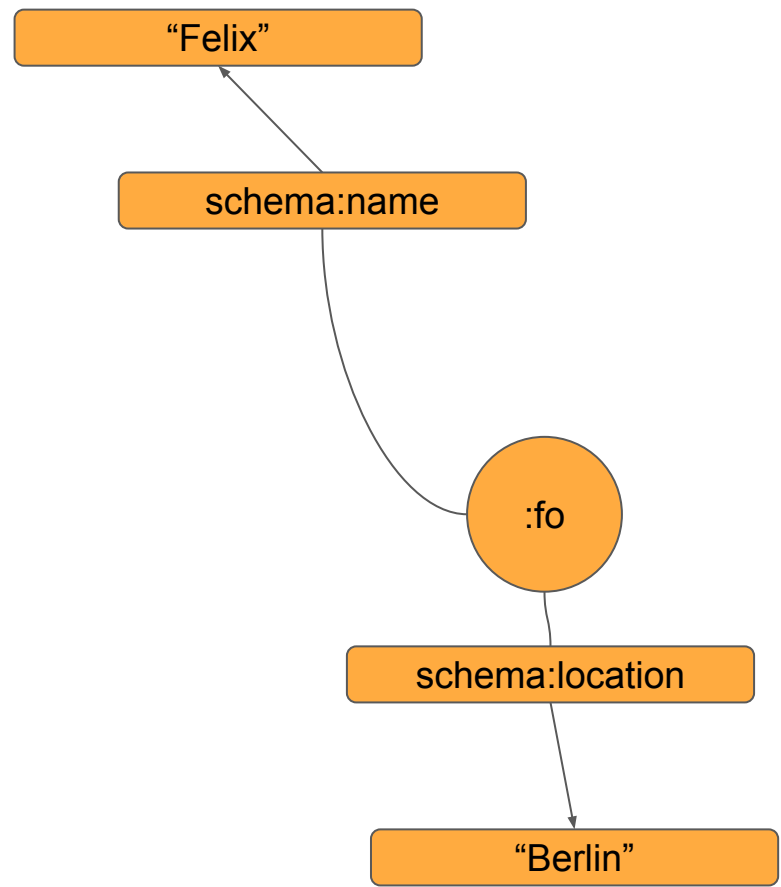
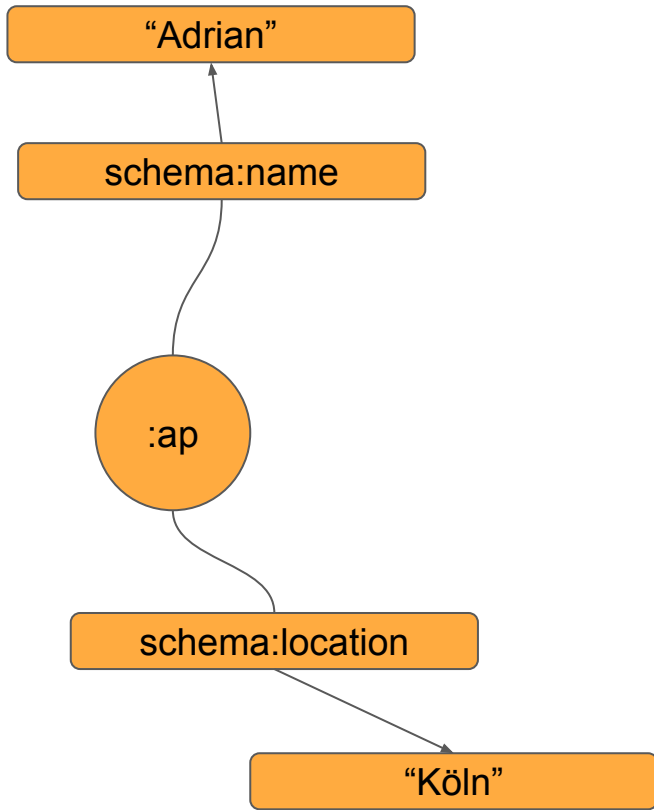
```
z:fo schema:name "Felix" .
```

```
z:fo schema:location "Berlin" .
```

```
1 @prefix a: <http://etherpad.lobid.org/p/dini-ws-a/export/txt#> .
2 @prefix b: <http://etherpad.lobid.org/p/dini-ws-b/export/txt#> .
3 @prefix c: <http://etherpad.lobid.org/p/dini-ws-c/export/txt#> .
4 @prefix d: <http://etherpad.lobid.org/p/dini-ws-d/export/txt#> .
5 @prefix e: <http://etherpad.lobid.org/p/dini-ws-e/export/txt#> .
6 @prefix f: <http://etherpad.lobid.org/p/dini-ws-f/export/txt#> .
7 @prefix z: <http://etherpad.lobid.org/p/dini-ws-z/export/txt#> .
8
9 @prefix schema: <http://schema.org/> .
10
11 z:fo schema:name "Felix" .
12 z:fo schema:location "Berlin" .
13
14 z:ap schema:name "Adrian" .
15 z:ap schema:location "Köln" .
16
```

Resource Description Framework (RDF)

Zur **Beschreibung** beliebiger Dinge wird das **Resource Description Framework** (RDF) verwendet. Dieses basiert auf dem Konzept von **Tripeln**, welche aus **Subjekt, Prädikat und Objekt** bestehen. Es ist ein **abstraktes Modell**, für das verschiedene **Schreibweisen** existieren - wir verwenden heute [Turtle](#).



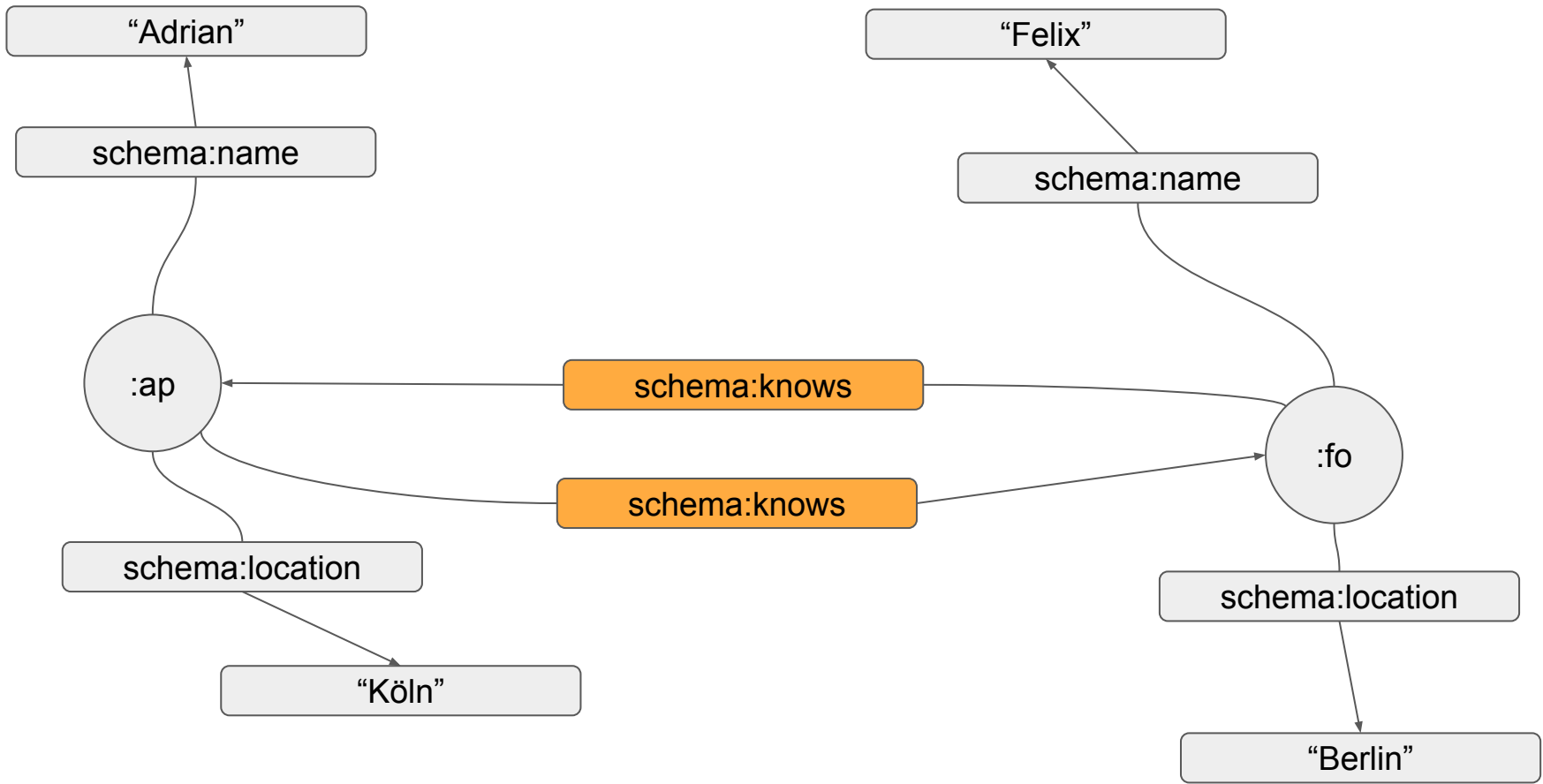
Sie sind dran!

Stellen Sie sich den anderen Mitgliedern Ihrer Gruppe vor. Reichen Sie dazu den Faden weiter und halten Sie anschließend die Bekanntschaft im Etherpad fest.

```
1 @prefix a: <http://etherpad.lobid.org/p/dini-ws-a/export/txt#> .
2 @prefix b: <http://etherpad.lobid.org/p/dini-ws-b/export/txt#> .
3 @prefix c: <http://etherpad.lobid.org/p/dini-ws-c/export/txt#> .
4 @prefix d: <http://etherpad.lobid.org/p/dini-ws-d/export/txt#> .
5 @prefix e: <http://etherpad.lobid.org/p/dini-ws-e/export/txt#> .
6 @prefix f: <http://etherpad.lobid.org/p/dini-ws-f/export/txt#> .
7 @prefix z: <http://etherpad.lobid.org/p/dini-ws-z/export/txt#> .
8
9 @prefix schema: <http://schema.org/> .
10
11 z:fo schema:name "Felix" .
12 z:fo schema:location "Berlin" .
13 z:fo schema:knows z:ap .
14
15 z:ap schema:name "Adrian" .
16 z:ap schema:location "Köln" .
17 z:ap schema:knows z:fo
18 |
```


Linked Data

Dinge werden in RDF anhand von **Uniform Resource Identifiers** (URIs) **benannt**. Durch die **Verknüpfungen** zwischen Dingen entstehen **Graphen**. Soziale **Netzwerke** wie Facebook oder LinkedIn sind wohlbekannte Vertreter dieses Ansatzes.



Sie sind dran!

Erweitern Sie Ihren Bekanntenkreis, indem Sie den Faden über Ihre Gruppe hinaus weitergeben. Dokumentieren Sie auch diese neuen Verbindungen im Etherpad.

Rich text editor toolbar: Bold (B), Italic (I), Underline (U), Strikethrough (ABC), Bulleted list, Numbered list, Indent, Outdent, Undo, Redo, Color picker

Navigation icons: back, forward, star, settings, code, user profile

```
1 @prefix a: <http://etherpad.lobid.org/p/dini-ws-a/export/txt#> .
2 @prefix b: <http://etherpad.lobid.org/p/dini-ws-b/export/txt#> .
3 @prefix c: <http://etherpad.lobid.org/p/dini-ws-c/export/txt#> .
4 @prefix d: <http://etherpad.lobid.org/p/dini-ws-d/export/txt#> .
5 @prefix e: <http://etherpad.lobid.org/p/dini-ws-e/export/txt#> .
6 @prefix f: <http://etherpad.lobid.org/p/dini-ws-f/export/txt#> .
7 @prefix z: <http://etherpad.lobid.org/p/dini-ws-z/export/txt#> .
8
9 @prefix schema: <http://schema.org/> .
10
11 z:fo schema:name "Felix" .
12 z:fo schema:location "Berlin" .
13 z:fo schema:knows z:ap .
14 z:fo schema:knows a:xy .
15
16 z:ap schema:name "Adrian" .
17 z:ap schema:location "Köln" .
18 z:ap schema:knows z:fo
19 z:ap schema:knows b:xy| .
20
```

The simple power of the Link

Schon ein einziger **Link** kann einen **Graphen** stark **erweitern**, da viele weitere Verknüpfungen zum Netzwerk hinzukommen können.

Sie sind dran!

Ersetzen Sie den Namen ihres Wohnortes durch einen Verweis auf DBpedia.

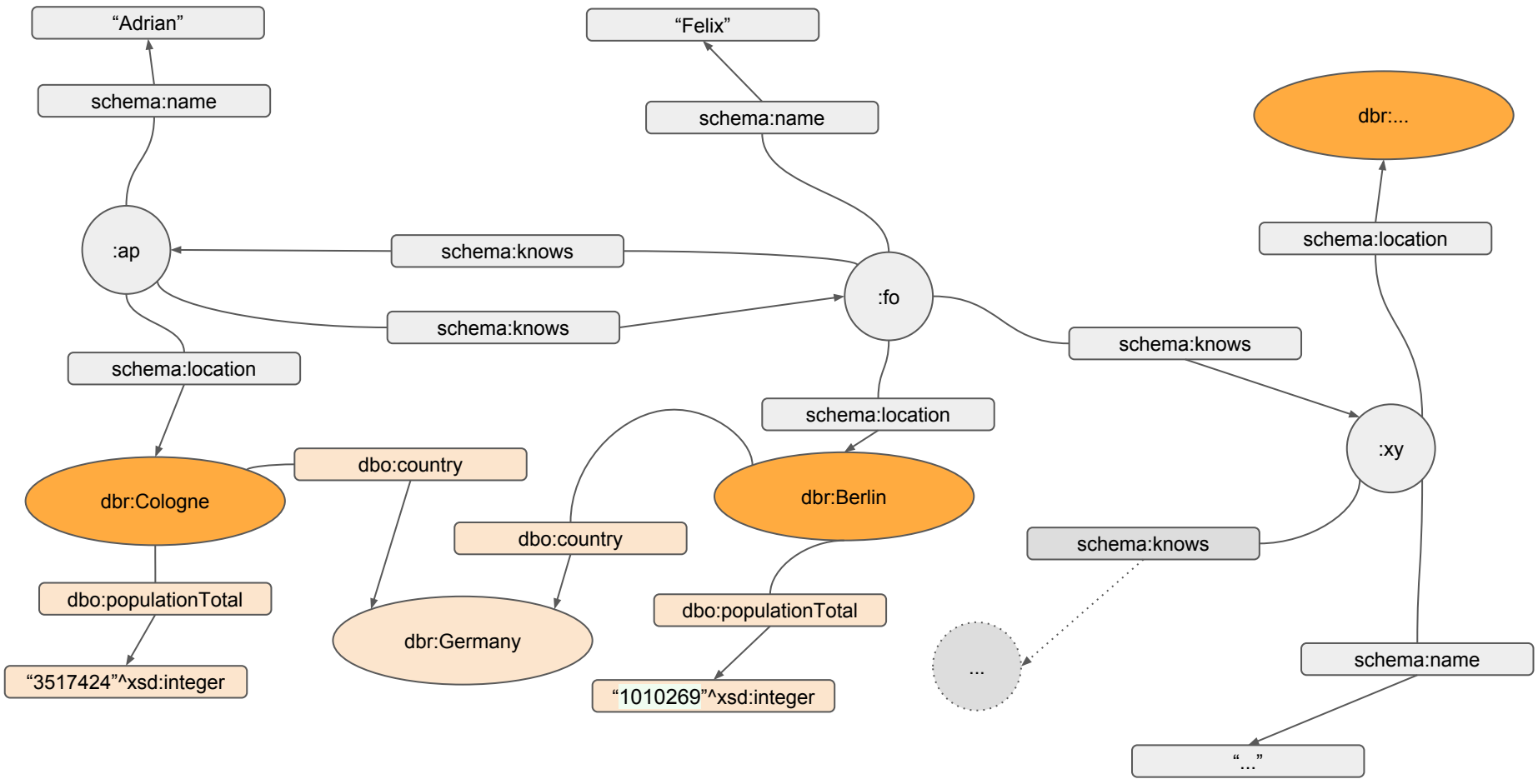
B *I* U ~~S~~ ↶ ↷

↶ ⌚ ☆ ⚙️ </> 👤 1

```
1 @prefix a: <http://etherpad.lobid.org/p/dini-ws-a/export/txt#> .
2 @prefix b: <http://etherpad.lobid.org/p/dini-ws-b/export/txt#> .
3 @prefix c: <http://etherpad.lobid.org/p/dini-ws-c/export/txt#> .
4 @prefix d: <http://etherpad.lobid.org/p/dini-ws-d/export/txt#> .
5 @prefix e: <http://etherpad.lobid.org/p/dini-ws-e/export/txt#> .
6 @prefix f: <http://etherpad.lobid.org/p/dini-ws-f/export/txt#> .
7 @prefix z: <http://etherpad.lobid.org/p/dini-ws-z/export/txt#> .
8
9 @prefix schema: <http://schema.org/> .
10 @prefix dbr: <http://dbpedia.org/resource/> .
11
12 z:fo schema:name "Felix" .
13 z:fo schema:location dbr:Berlin .
14 z:fo schema:knows z:ap .
15 z:fo schema:knows a:xy .
16
17 z:ap schema:name "Adrian" .
18 z:ap schema:location dbr:Cologne .
19 z:ap schema:knows z:fo
20 z:ap schema:knows b:xy .
21 |
```


Der Giant Global Graph

Durch die Verwendung von **HTTP-URIs** baut Linked Data auf einer Technologie auf, dessen globale **Skalierbarkeit** belegt ist. In Anlehnung an das World Wide Web (WWW) spricht man deshalb auch vom **Giant Global Graph** (GGG). Wie im WWW gilt auch im GGG: **Anyone can say anything about anything!**



Pause

Sie sind dran!

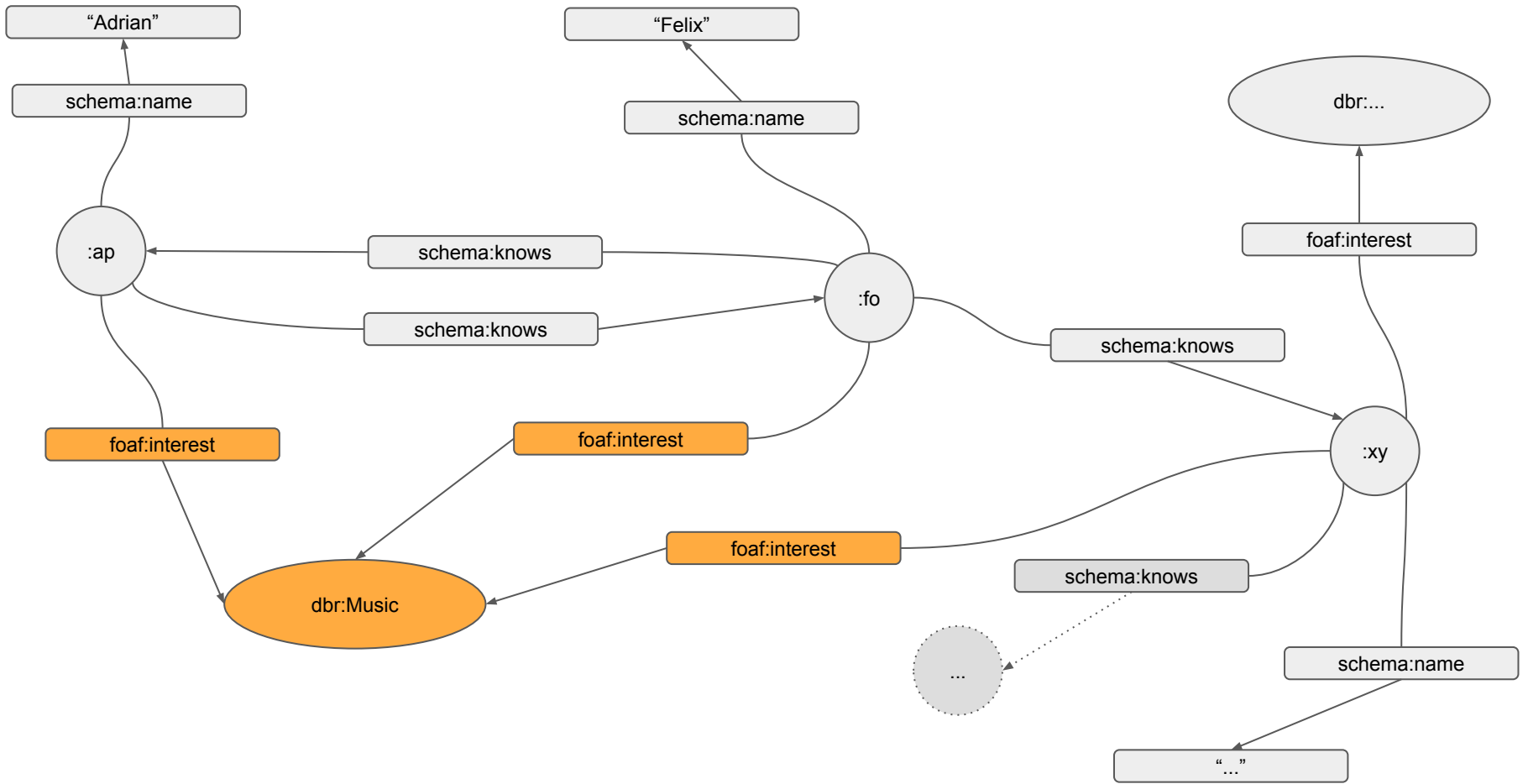
Erweitern Sie ihr Profil um Interessen, die Sie hegen.
Verweisen Sie dazu wiederum auf die
entsprechenden DBpedia-Einträge.

B *I* U

```
1 @prefix a: <http://etherpad.lobid.org/p/dini-ws-a/export/txt#> .
2 @prefix b: <http://etherpad.lobid.org/p/dini-ws-b/export/txt#> .
3 @prefix c: <http://etherpad.lobid.org/p/dini-ws-c/export/txt#> .
4 @prefix d: <http://etherpad.lobid.org/p/dini-ws-d/export/txt#> .
5 @prefix e: <http://etherpad.lobid.org/p/dini-ws-e/export/txt#> .
6 @prefix f: <http://etherpad.lobid.org/p/dini-ws-f/export/txt#> .
7 @prefix z: <http://etherpad.lobid.org/p/dini-ws-z/export/txt#> .
8
9 @prefix schema: <http://schema.org/> .
10 @prefix dbr: <http://dbpedia.org/resource/> .
11 @prefix foaf: <http://xmlns.com/foaf/0.1/> .
12
13 z:fo schema:name "Felix" .
14 z:fo schema:location dbr:Berlin .
15 z:fo schema:knows z:ap .
16 z:fo schema:knows a:xy .
17 z:fo foaf:interest dbr:Music .
18 z:fo foaf:interest <http://dbpedia.org/resource/Python_(programming_language)> .
19
20 z:ap schema:name "Adrian" .
21 z:ap schema:location dbr:Cologne .
22 z:ap schema:knows z:fo .
23 z:ap schema:knows b:xy .
24 z:ap foaf:interest dbr:Music .
25 z:ap foaf:interest dbr:Resource_Description_Framework .
26
```

Zusätzliche Vokabulare

- FOAF
- DBpedia-Ontologie
- ... und viele mehr (siehe [Linked Open Vocabularies \(LOV\)](#))
- prefix.cc -> Namensräume finden



Bibliothekarisches relevante Vokabulare

- [Dublin Core Metadata Terms](#)
- [Simple Knowledge Organisation System \(SKOS\)](#)
 - Bsp.: [Standard Thesaurus Wirtschaft \(STW\)](#), [Umweltthesaurus UMTHESES](#), [Agrovoc](#), viele kleine kontrollierte Wertelisten (z.B. [RDA value vocabularies](#))
 - SKOS-Schemas in BARTOC: [http://bartoc.org/en/search/advanced?f\[0\]=field_format%3A24](http://bartoc.org/en/search/advanced?f[0]=field_format%3A24)
- [GND Ontology](#)
- [Learning Resource Metadata Initiative](#)
- [RDA Elements sets](#)
- [Bibframe](#)
- ...

Sie sind dran!

Erlauben Sie die Nachnutzung Ihrer Daten.
Dokumentieren Sie dazu Ihre Urheberschaft und
lizenzieren Sie die Daten mit CC0.

B *I* U

```
1 @prefix a: <http://etherpad.lobid.org/p/dini-ws-a/export/txt#> .
2 @prefix b: <http://etherpad.lobid.org/p/dini-ws-b/export/txt#> .
3 @prefix c: <http://etherpad.lobid.org/p/dini-ws-c/export/txt#> .
4 @prefix d: <http://etherpad.lobid.org/p/dini-ws-d/export/txt#> .
5 @prefix e: <http://etherpad.lobid.org/p/dini-ws-e/export/txt#> .
6 @prefix f: <http://etherpad.lobid.org/p/dini-ws-f/export/txt#> .
7 @prefix z: <http://etherpad.lobid.org/p/dini-ws-z/export/txt#> .
8
9 @prefix schema: <http://schema.org/> .
10 @prefix dbr: <http://dbpedia.org/resource/> .
11 @prefix foaf: <http://xmlns.com/foaf/0.1/> .
12 @prefix dcterms: <http://purl.org/dc/terms/> .
13
14 <> dcterms:creator z:fo .
15 <> dcterms:creator z:ap .
16 <> schema:license <https://creativecommons.org/publicdomain/zero/1.0/> .
17
18 z:fo schema:name "Felix" .
19 z:fo schema:location dbr:Berlin .
20 z:fo schema:knows z:ap .
21 z:fo schema:knows a:xy .
22 z:fo foaf:interest dbr:Music .
23 z:fo foaf:interest <http://dbpedia.org/resource/Python\_\(programming\_language\)> .
24
25 z:ap schema:name "Adrian" .
26 z:ap schema:location dbr:Cologne .
27 z:ap schema:knows z:fo .
28 z:ap schema:knows b:xy .
29 z:ap foaf:interest dbr:Music .
30 z:ap foaf:interest dbr:Resource_Description_Framework .
31
```

“open” / “offen”?

*“Open data and content can be **freely used, modified, and shared by anyone for any purpose**”*

[The Open Definition](#)

Offenheit ist eine Frage...

- ...des **Zugangs**: keine Passwort-, Mengen- oder sonstige Beschränkungen
- ...der **Lizenz**: einzige mögliche Einschränkungen sind Attribution und Share-Alike
 - CC0, CC-BY, CC-BY-SA
 - keine non-commercial (NC) Lizenzen
- ...des **Formats**: keine proprietären Formate ohne frei zugängliche Spezifikation

Daten lizenzieren

Gesamtlieferungen unter der Lizenz [CC0 1.0](#)

Folgende Datenabzüge sind laut [Geschäftsmodell](#) unter der [Lizenz CC0 1.0](#) kostenfrei verfügbar.

Die Bestandsdaten der [ZDB](#) sind zum größten Teil ebenfalls frei nachnutzbar. Eine entsprechende [Kennzeichnung](#) erfolgt im Datensatz selbst.

Gesamtabzug (Stand: Oktober 2015)	Bereitstellung (Aktualisierung)	Formate	Anzahl der Datensätze/Datengröße (gezippt)
Titeldaten der DNB*	Februar, Juli, Oktober	MARC 21	ca. 13,8 Mio. / ca. 3,1 GB
		MARC21-xml	
Titeldaten der DNB* **	Februar, Juni, Oktober	RDFxml u. RDF/turtle	ca. 11,9 Mio. / ca. 1,2 GB
Gemeinsame Normdatei (GND)	Februar, Juni, Oktober	MARC 21	ca. 11,1 Mio. / ca. 1,2 GB
		MARC21-xml	
		RDFxml u. RDF/turtle	ca. 0,8 GB
Titeldaten der ZDB	März u. September	MARC 21	ca. 1,8 Mio. / ca. 430 MB
		MARC21-xml	
		RDFxml u. RDF/turtle	ca. 185 MB
Bestandsdaten der ZDB	März u. September	MARC 21	ca. 14,7 Mio. / ca. 1,1 GB
		MARC21-xml	

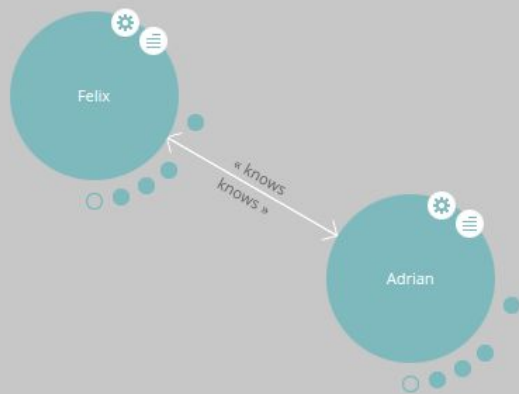
Demo

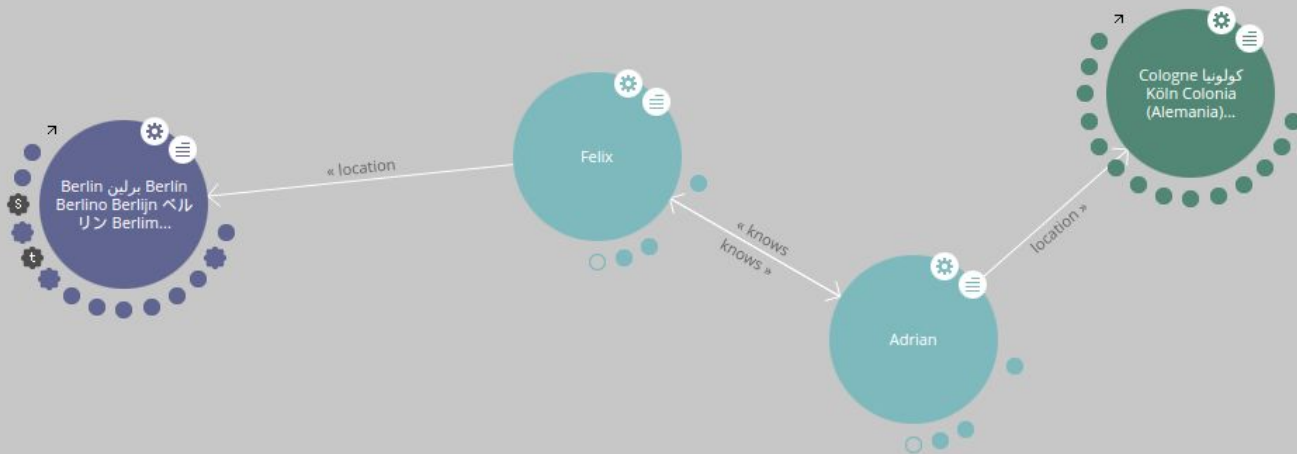
Browsen, Einsammeln und Abfragen

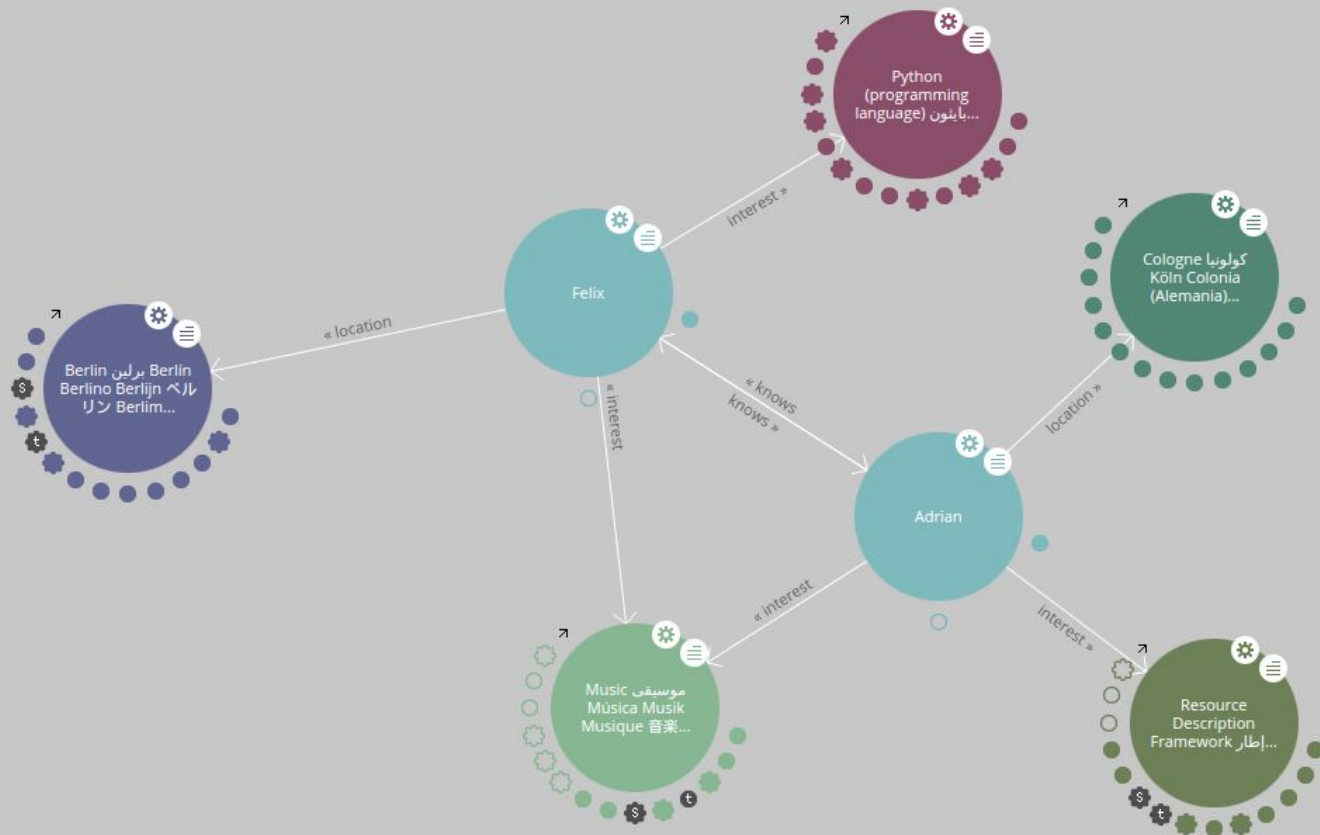
lodlive

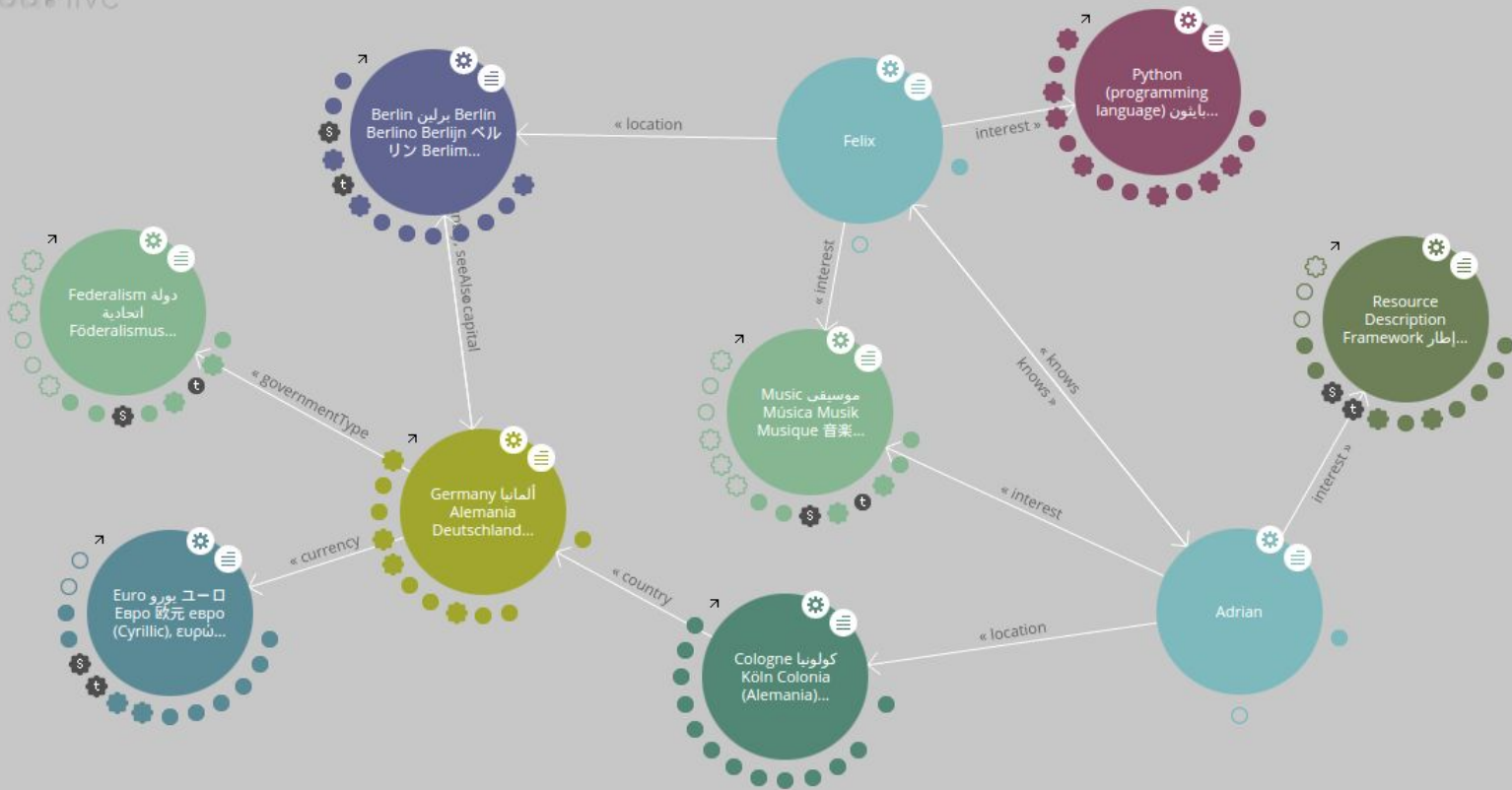


lodlive









Triplestores & SPARQL

Verstreute maschinenlesbare Beschreibungen sind nützlich, aber da geht noch mehr! RDF ist ein **verteiltes Datenmodell**, das es einfach macht, mehrere Beschreibungen zu **kombinieren**. Mit **Triplestores** existieren spezielle **Datenbanken**, um anhand von **SPARQL Abfragen** über die aggregierten Daten ausführen zu können.

Namen der Teilnehmer

```
PREFIX schema: <http://schema.org/>
```

```
SELECT * WHERE {  
  ?person schema:name ?name .  
}
```

?person	?name
http://etherpad.lobid.org/p/dini-ws-z/export/txt#fo	Felix
http://etherpad.lobid.org/p/dini-ws-z/export/txt#ap	Adrian

Completed in 1736 ms. 2 results

Bekanntschaften

```
PREFIX schema: <http://schema.org/>
```

```
SELECT * WHERE {  
  ?wer schema:knows ?wen .  
}
```

?wer	?wen
http://etherpad.lobid.org/p/dini-ws-z/export/txt#fo	http://etherpad.lobid.org/p/dini-ws-z/export/txt#ap
http://etherpad.lobid.org/p/dini-ws-z/export/txt#fo	http://etherpad.lobid.org/p/dini-ws-a/export/txt#xy
http://etherpad.lobid.org/p/dini-ws-z/export/txt#ap	http://etherpad.lobid.org/p/dini-ws-z/export/txt#fo
http://etherpad.lobid.org/p/dini-ws-z/export/txt#ap	http://etherpad.lobid.org/p/dini-ws-b/export/txt#xy

Completed in 277 ms. 4 results

Namen von Bekanntschaften

```
PREFIX schema: <http://schema.org/>
```

```
SELECT ?namewer ?namewen WHERE {  
  ?wer schema:knows ?wen .  
  ?wer schema:name ?namewer .  
  ?wen schema:name ?namewen .  
}
```

?namewer	?namewen
Adrian	Felix
Felix	Adrian

Completed in 2403 ms. 2 results

Wohnorte und Länder

```
PREFIX schema: <http://schema.org/>
PREFIX dbo: <http://dbpedia.org/ontology/>

SELECT * WHERE {
  ?person schema:location ?place .
  ?place dbo:country ?country .
}
```

?country	?person	?place
http://dbpedia.org/resource/Germany	http://etherpad.lobid.org/p/dini-ws-z/export/txt#ap	http://dbpedia.org/resource/Cologne
http://dbpedia.org/resource/Germany	http://etherpad.lobid.org/p/dini-ws-z/export/txt#fo	http://dbpedia.org/resource/Berlin

Completed in 591 ms. 2 results

Gleiche Interessen

```
PREFIX schema: <http://schema.org/>
PREFIX dbo: <http://dbpedia.org/ontology/>
PREFIX z: <http://etherpad.lobid.org/p/dini-ws-z/export/txt#>
```

```
SELECT * WHERE {
  z:fo foaf:interest ?interest .
  ?person foaf:interest ?interest .
  FILTER (?person != z:fo)
}
```

?interest	?person
http://dbpedia.org/resource/Music	http://etherpad.lobid.org/p/dini-ws-z/export/txt#ap

Completed in 165 ms.

1 result

Großstädter

```
PREFIX schema: <http://schema.org/>
PREFIX dbo: <http://dbpedia.org/ontology/>

SELECT * WHERE {
  ?person schema:location ?place .
  ?place dbo:populationTotal ?population .
  FILTER (?population > 100000) .
}
```

?person	?place	?population
http://etherpad.lobid.org/p/dini-ws-z/export/txt#fo	http://dbpedia.org/resource/Berlin	3517424
http://etherpad.lobid.org/p/dini-ws-z/export/txt#ap	http://dbpedia.org/resource/Cologne	1010269

Completed in 635 ms. 2 results

Sie sind dran!

Erkunden Sie Ihr soziales Netzwerk, nutzen Sie dafür den SPARQL-Endpoint und die interaktive LodLive-Visualisierung.

Nutzung: Mashups

Mashups nutzen verschiedene Datenquellen zur Generierung eines Services, einer Visualisierung o. ä.

Nutzung: Suche

Linked-Data-Suchmaschinen ermöglichen eine Suche über das Web of Data. Konventionelle Suchmaschinen können Informationen aus LOD integrieren (schema.org).

Nutzung: Datenintegration

Linked Data erleichtert die Integration unterschiedlicher Daten zu Analyse- und Recherchezwecken.

Speicherung & Publikation

Linked Data kann u. a. in Dateien auf einem Webserver, als Markup von Webseiten (RDFa) oder als RDF-Graph in Triple Stores publiziert werden.

Abfrage: SPARQL

SPARQL Protocol and RDF Query Language (SPARQL) ermöglicht strukturierte Abfragen über RDF-Graphen in einem Triple Store.

Ontologien/Vokabulare: RDFS & OWL

Vokabulare und Ontologien liegen Beschreibungen von Dingen mit RDF zugrunde. Zu diesem Zweck definieren sie Klassen und Eigenschaften sowie deren Relationen. RDF-Schema (RDFS) und die Web Ontology Language (OWL) sind Sprachen zur Beschreibung von Ontologien/Vokabularen.

Datenmodell: RDF

Das Resource Description Framework (RDF) ist ein Datenmodell, mit dem Aussagen in Tripeln repräsentiert werden können.

Identifikation: URI

Die Nutzung von HTTP Uniform Resource Identifiers (URIs) bringt mit sich, dass
a) Daten über das Internet abgerufen werden können und
b) Namensräume über das Domain Name System (DNS) verwaltet werden.

Transport: HTTP

Die Daten werden auf Webservern exponiert, die das Hypertext Transfer Protocol (HTTP) benutzen.

Lizenzierung: Open Data
Die Daten werden unter Bedingungen publiziert, die eine Nutzung, Weiterverwendung und Verbreitung der Daten ermöglichen.

Ende

Fragen gerne jetzt oder später an

felix.ostrowski@gmail.com

pohl@hbz-nrw.de



Diese Folien stehen unter einer Creative-Commons-Namensnennung-Lizenz:
[CC-BY 4.0 International](https://creativecommons.org/licenses/by/4.0/)

Referenzen

DINI-KIM-Glossar: <http://dini-ag-kim.github.io/glossar/glossar.html>

Europeana (2012): Linked Open Data - What is it? <https://vimeo.com/36752317>

Kreutzer, Till (2011): Open Data – Freigabe von Daten aus Bibliothekskatalogen. Ein Leitfaden. Hg. v. Hochschulbibliothekszentrum des Landes Nordrhein-Westfalen. URL: <https://www.hbz-nrw.de/dokumentencenter/veroeffentlichungen/open-data-leitfaden.pdf>

Pohl, Adrian / Danowski, Patrick (2013): Linked Open Data in der Bibliothekswelt: Grundlagen und Überblick. In: dies. (Hg.): (Open) Linked Data in Bibliotheken. URL: <http://dx.doi.org/10.1515/9783110278736.1>