

# Linked Open Data – Eine Einführung

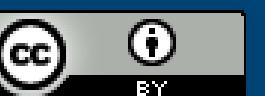
Adrian Pohl

Offene Infrastruktur, Hochschulbibliothekszentrum NRW (hbz)



Teamworkshop des Deutschen Bildungsservers, 2022-05-06

<https://slides.lobid.org/2022-05-06-lod-dipf/>



# Hello, again!

# 10 Jahre ist's her

Vorträge zur FIS Bildung-Frühjahrstagung vom 22.05.2012 zum Thema  
**„Linked Open Data“**

1. Adrian Pohl (hbz):  
Linked Open Data & Bibliotheken - Warum? Was? Wie?
2. Benjamin Zapilko (gesis):  
Anwendungsszenarien und Potenziale von Linked Open Data im wissenschaftlichen Forschungsprozess
3. Adrian Pohl (hbz):  
Zum Aufbau von Linked-Open-Data-Services im hbz

# Agenda

1. Warum Linked Open Data?
2. LOD Basics
3. Linked Open Usable Data
4. SKOS
5. Fragen und Diskussion

# 1. Warum Linked Open Data?

# Warum Linked Open Data (LOD)?

- Überführung traditioneller bibl. Praktiken in das Web
- Sichtbarkeit und Auffindbarkeit im Web erreichen
- Nachnutzbarkeit ermöglichen
- Synergieeffekte durch Verlinkung mit anderen Daten
- Verbesserung der Recherchemöglichkeiten

Pohl, Adrian / Ostrowski, Felix (2010): 'Linked Data' - und warum wir uns im hbz-Verbund damit beschäftigen." B.I.T. Online 13(3): S. 259-268. [Preprint \(PDF\)](#)

# LOD als politische Entscheidung

LOD hat zum Ziel, Austausch und Nachnutzung von Daten über Systemgrenzen hinweg zu vereinfachen. Somit richtet es sich an Organisationen,

1. deren Geschäftsmodell nicht darauf basiert,  
Nutzer\*innendaten zu horten
2. die nicht im kommerziellen Wettbewerb stehen
3. die existieren, um Information zu teilen

Siehe auch: Dorian Taylor. "Linked Data is a political agenda"

# 2. LOD Basics

# Technische und rechtliche Offenheit

Damit Daten nachgenutzt werden können, sollten sie sowohl rechtlich als auch technisch offen sein.

# Open Definition

*Wissen ist offen, wenn jedeR darauf frei zugreifen, es nutzen, verändern und teilen kann – eingeschränkt höchstens durch Maßnahmen, die Ursprung und Offenheit des Wissens bewahren.*

<https://opendefinition.org/od/2.1/de/>

# Rechtliche Offenheit

Vergabe einer **Offenen Lizenz**

(für Daten hat sich **CC0** etabliert)

# Technische Offenheit



# Linked Data: Best Practices

1. Nutze URLs als Namen für Dinge
2. Nutze HTTP-URLs, so dass Menschen sie aufrufen können
3. Wenn jemand einen URI aufruft, biete nützliche Informationen an unter Nutzung der Standards (RDF, SPARQL)
4. Nimm Links zu anderen URLs auf, so dass weitere Dinge entdeckt werden können.

Tim Berners-Lee (2006ff): Linked Data – Design Issues

# Exkurs: URIs

- *Uniform*: einheitliche Struktur ([RFC 3986](#))
- *Resource*: ein Ding, ob digital oder nicht-digital
- *Identifier*: Identifikation durch Vergabe eines Namens/einer ID

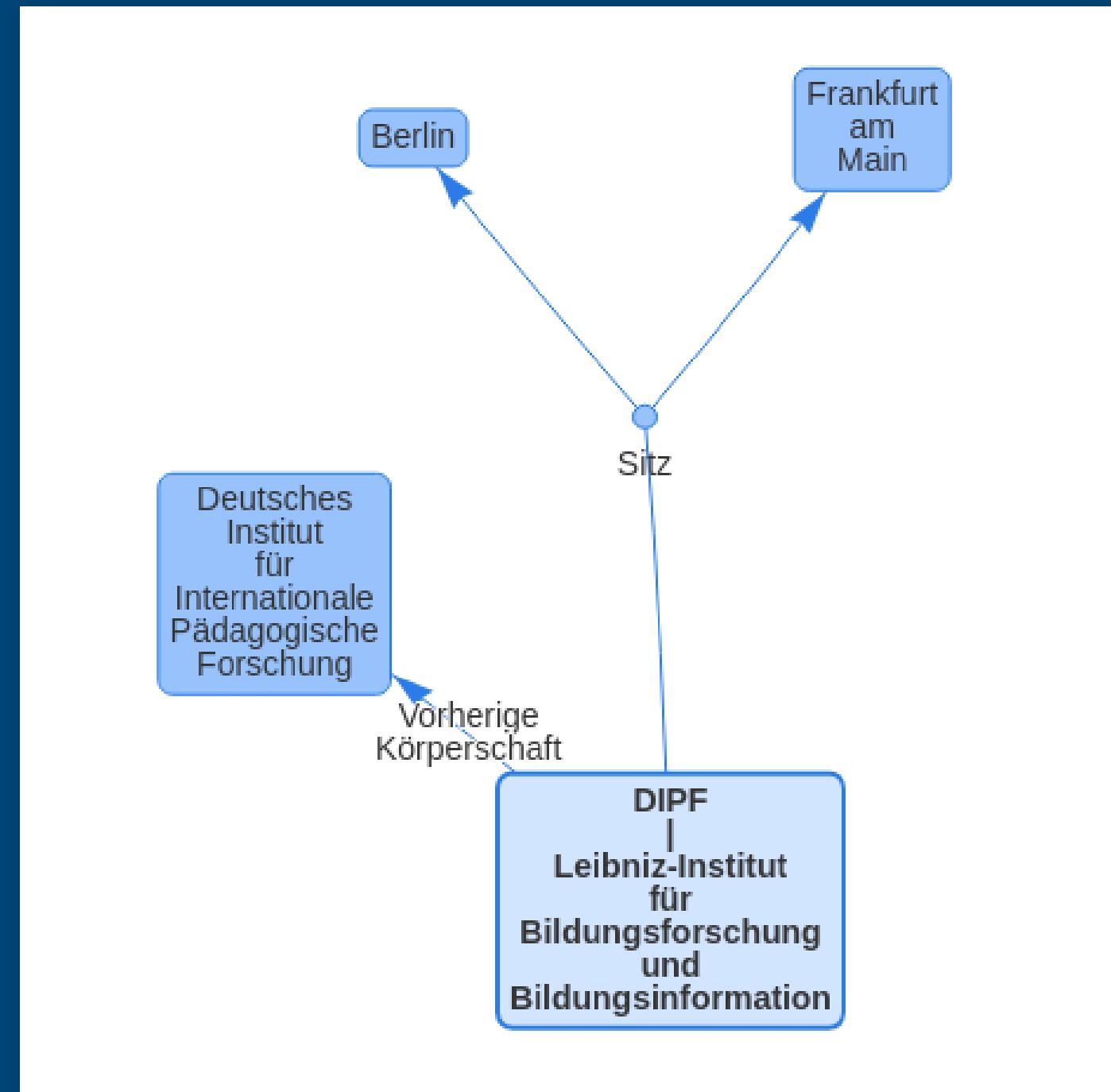
# URIs

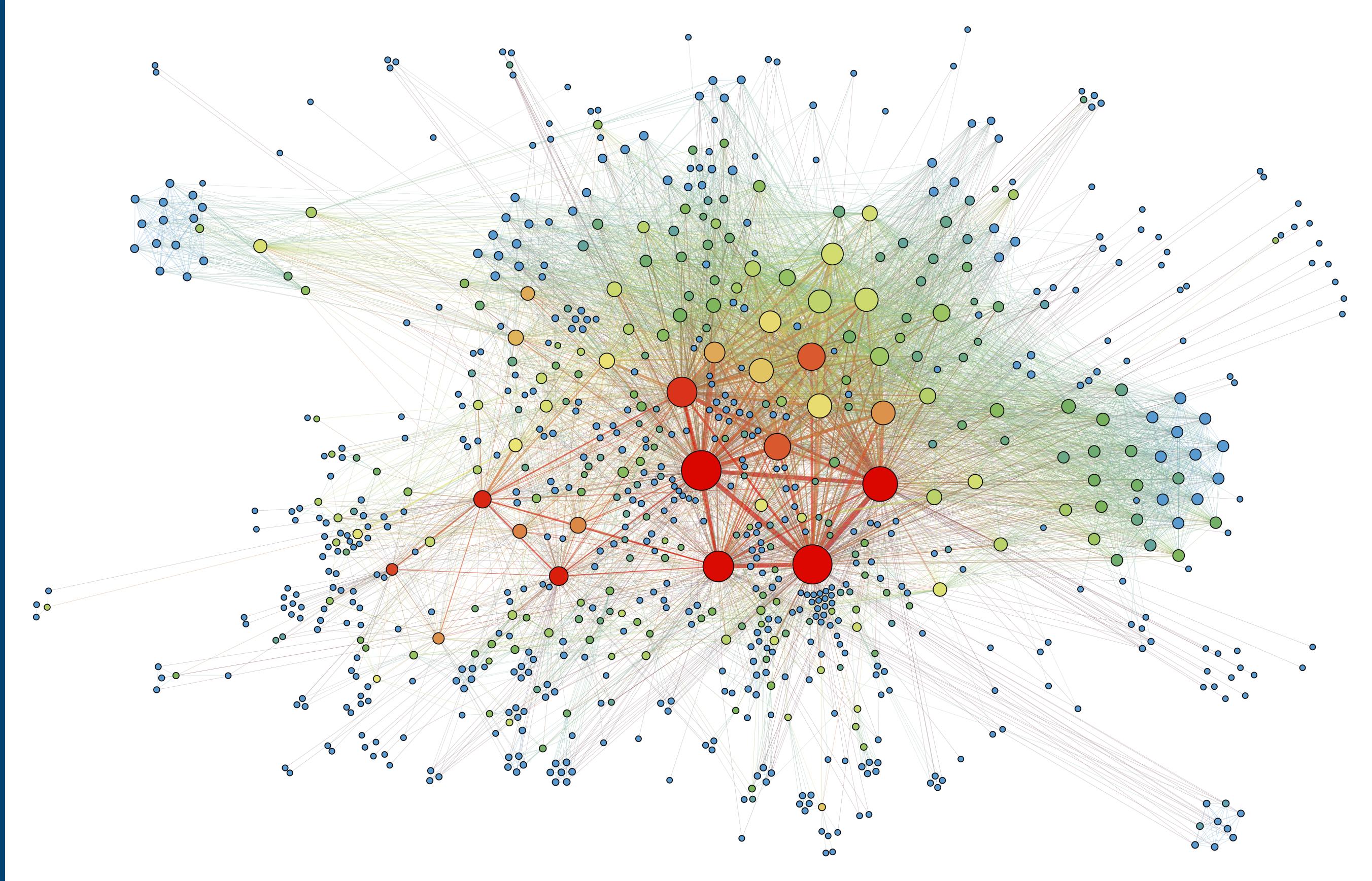
- URIs sind die Basis eines verteilten, dezentral gepflegten Linked-Data-Netzwerks
- Namensräume machen URIs eindeutig
- z.B. gibt es viele Webseiten, die "about" oder "faq" heißen und durch die Angabe der Domain disambiguiert werden
- URIs können als bloßer Identifikator genutzt werden, ohne im Browser aufgerufen werden zu können

# RDF – Resource Description Framework

- graph-basiertes Datenmodell
- zum Austausch von Daten über das Web, entworfen und gepflegt vom W3C
- abstraktes Modell mit verschiedenen Serialisierungsformaten (RDF/XML, Turtle, JSON-LD...)

# Was ist ein Graph?





# RDF: ein gerichteter, beschrifteter Graph

- Englisch: directed labeled graph
- d.h. die Kanten haben eine Richtung und ein Etikett



# Tripel: Das Graph-Atom

Alle Daten in RDF existieren in Form von Tripeln mit Subjekt,  
Prädikat, Objekt:



# Abbildung von Daten in RDF

- **Subjekt:** Ressource über die eine Aussage getroffen wird
- **Prädikat:** Beziehung zwischen Subjekt und Objekt
- **Objekt:** Ressource, zu dem die Ressource in einer Verbindung steht oder ein Literal

# URIs und Literale

- **Subjekt** und **Prädikat** sind immer ein URI
- **Objekt** ist ein URI *oder* ein Literal (String, Date, Integer)

Subjekt	Prädikat	Objekt
Ich	heiße	Adrian.
< <a href="https://orcid.org/0000-0001-9083-7442">https://orcid.org/0000-0001-9083-7442</a> >	< <a href="https://schema.org/name">https://schema.org/name</a> >	"Adrian"@de .
Ich	interessiere mich für	Linked Library Data.
< <a href="https://orcid.org/0000-0001-9083-7442">https://orcid.org/0000-0001-9083-7442</a> >	< <a href="http://xmlns.com/foaf/0.1/interest">http://xmlns.com/foaf/0.1/interest</a> >	< <a href="http://www.wikidata.org/entity/Q12345">http://www.wikidata.org/entity/Q12345</a> > .

Die verschiedenen RDF-Serialisierungen  
sind semantisch äquivalent.

Deshalb wird es ein "abstraktes" Modell genannt.

# N-Triples

```
<https://orcid.org/0000-0001-9083-7442>
<http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
<https://schema.org/namePerson> .
<https://orcid.org/0000-0001-9083-7442>
<https://schema.org/namename> "Adrian"@de .
<https://orcid.org/0000-0001-9083-7442>
<http://xmlns.com/foaf/0.1/interest>
<http://www.wikidata.org/entity/Q466> .
<https://orcid.org/0000-0001-9083-7442>
<http://xmlns.com/foaf/0.1/interest>
<http://www.wikidata.org/entity/Q28134421> .
<https://orcid.org/0000-0001-9083-7442>
<http://xmlns.com/foaf/0.1/interest>
<http://www.wikidata.org/entity/Q341> .
<https://orcid.org/0000-0001-9083-7442>
<http://xmlns.com/foaf/0.1/interest>
```

# Turtle

```
@prefix schema: <https://schema.org/name> .  
@prefix foaf: <http://xmlns.com/foaf/0.1/> .  
@prefix wd: <http://www.wikidata.org/entity/> .  
<https://orcid.org/0000-0001-9083-7442> a schema:Person ;  
    schema:name "Adrian"@de ;  
    foaf:interest wd:Q466, wd:Q28134421, wd:Q341, wd:Q380962 ,  
    wd:Q56298524 .
```

# RDF/XML

```
<?xml version="1.0" encoding="utf-8"?>
<rdf:RDF xmlns:foaf="http://xmlns.com/foaf/0.1/"
           xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
           xmlns: schema="https://schema.org/name">
    <schema:Person rdf:about="https://orcid.org/0000-0001-9083-
7442">
        <foaf:interest
            rdf:resource="http://www.wikidata.org/entity/Q28134421"/>
        <foaf:interest
            rdf:resource="http://www.wikidata.org/entity/Q341"/>
        <foaf:interest
            rdf:resource="http://www.wikidata.org/entity/Q380962"/>
        <foaf:interest
            rdf:resource="http://www.wikidata.org/entity/Q466"/>
        <foaf:interest
            rdf:resource="http://www.wikidata.org/entity/Q5000504"/>
```

Source: <https://www.wikidata.org/entity/Q5000504>

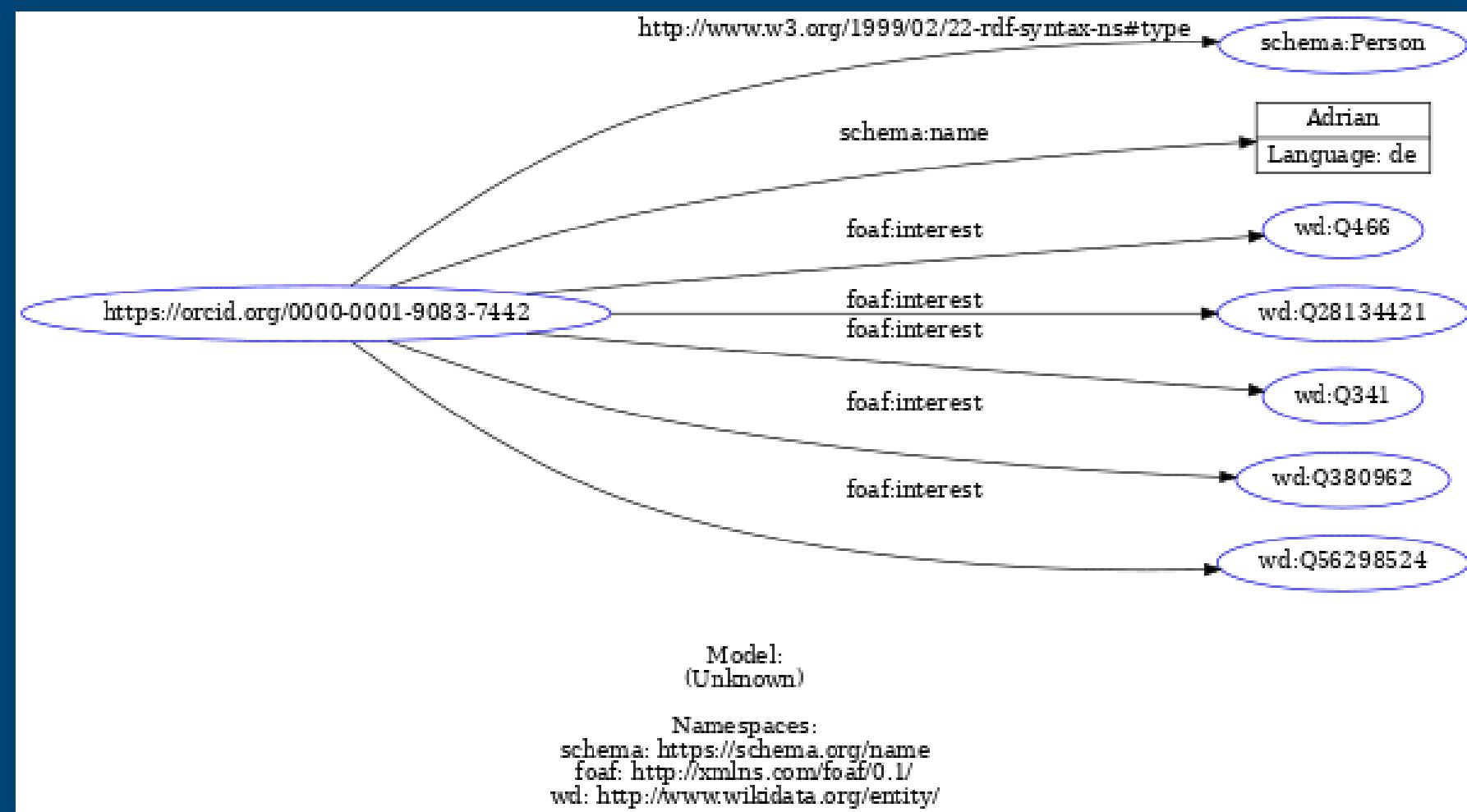
# RDFa – RDF in attributes

```
<div xmlns: schema="https://schema.org/"  
      xmlns: foaf="http://xmlns.com/foaf/0.1/"  
      about="https://orcid.org/0000-0001-9083-7442"  
      typeOf="schema:Person">  
    <span property="schema:name">Adrian Pohl</span> is  
    interested in  
    <span rel="foaf:interest"  
          resource="http://www.wikidata.org/entity/Q466">the World Wide  
          Web</span>, <span  
          rel="foaf:interest"  
          resource="http://www.wikidata.org/entity/Q28134421">Library  
          Linked Data</span>, <span  
          rel="foaf:interest"  
          resource="http://www.wikidata.org/entity/Q341">free  
          software</span>, <span  
          rel="foaf:interest"  
          resource="https://creativecommons.org/licenses/by-sa/3.0/deed.en">Creative Commons Attribution-ShareAlike 3.0 Unported License</span>
```

# JSON-LD

```
{  
  "@context": {  
    "Person": "https://schema.org/Person",  
    "name": "https://schema.org/name",  
    "interest": "http://xmlns.com/foaf/0.1/interest"  
  },  
  "@id": "https://orcid.org/0000-0001-9083-7442",  
  "@type": "Person",  
  "name": "Adrian",  
  "interest": [  
    "http://www.wikidata.org/entity/Q466",  
    "http://www.wikidata.org/entity/Q28134421",  
    "http://www.wikidata.org/entity/Q341",  
    "http://www.wikidata.org/entity/Q380962",  
    "http://www.wikidata.org/entity/Q56298524"  
  ]  
}
```

# RDF-Visualisierung



# Übung: Turtle schreiben

Erstellen Sie eine kurze Beschreibung von sich orientiert am Beispiel von **Slide 25**.

Zur Erstellung und Validierung lässt sich gut dieser Turtle-Web-Editor nutzen: <https://felixlohmeier.github.io/turtle-web-editor/>

# RDFS und OWL

- RDF Schema und Web Ontology Language
- Beschreibungssprachen zur Definition und Dokumentation von Properties und Klassen und deren Beziehungen
- Beispiele: schema.org, Dublin Core, SKOS, Bibframe, FoaF

# Beispiel: DC Terms

- URI: <http://purl.org/dc/terms/>
- Turtle-Version:  
[https://www.dublincore.org/2020/01/20/dublin\\_core\\_terms.ttl](https://www.dublincore.org/2020/01/20/dublin_core_terms.ttl)

```
@prefix dcterms: <http://purl.org/dc/terms/> .  
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .  
  
dcterms:BibliographicResource  
    dcterms:issued "2008-01-  
14"^^<http://www.w3.org/2001/XMLSchema#date> ;  
    a rdfs:Class ;  
    rdfs:comment "A book, article, or other documentary  
resource."@en ;  
    rdfs:isDefinedBy <http://purl.org/dc/terms/> ;  
    rdfs:label "Bibliographic Resource"@en .  
  
dcterms:title  
    dcterms:issued "2008-01-  
14"^^<http://www.w3.org/2001/XMLSchema#date> ;
```

```
@prefix dct: <http://purl.org/dc/terms/> .  
  
<https://slides.lobid.org/2022-05-06-lod-dipf/>  
    rdf:type dct:BibliographicResource ;  
    dct:title "Linked Open Data - Eine Einführung"@de ;  
    dct:creator <https://lobid.org/team/ap#!> .
```

# SPARQL Protocol and RDF Query Language

- Sprache/Protokoll zur Abfrage und Modifikation von RDF-Graphen
- Grundidee: Beschreibung von Subgraphen, die in dem abzufragenden Graphen enthalten sind inklusive ? Variablen
- Server antwortet mit passenden Knoten/URIs, für die das Graphmuster passt

# Gut zu SPARQLn: Wikidata

- ein Knowledge Graph
- ein Wikimedia-Projekt
- für Menschen und Maschinen
- multilingual
- kollaborativ gepflegt & offen für alle
- CC0
- basiert auf der freien Software **Wikibase**

# Wikidata-Datenmodell

# Beispiel Wikidata-Query I

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

SELECT ?item ?itemLabel
WHERE
{
    ?item ?p ?o;
        rdfs:label ?itemLabel .
} LIMIT 10
```

# Beispiel Wikidata-Query II

```
SELECT ?item ?itemLabel
WHERE
{
    ?item wdt:P31 wd:Q5; # is human
          wdt:P512 wd:Q6785216 ; # holds an MLIS
          rdfs:label ?itemLabel .
    FILTER ( langmatches(lang(?itemLabel), "en")
              && ! langmatches(lang(?itemLabel), "en-ca")
              && ! langmatches(lang(?itemLabel), "en-gb"))
}
```

Direktlink: <https://w.wiki/4hFu>

# Beispiel Wikidata-Query III

```
SELECT ?item ?itemLabel
WHERE
{
    ?item wdt:P31 wd:Q5; # is human
          wdt:P512 wd:Q6785216 . # holds an MLIS
    SERVICE wikibase:label { bd:serviceParam wikibase:language "
[AUTO_LANGUAGE],de". }
          # Wikibase-spezifischer, nicht SPARQL-konformer Label-
Dienst
}
```

Direktlink: <https://w.wiki/4hFv>

# Beispiel Wikidata-Query IV

```
#defaultView:ImageGrid

SELECT ?educationalInstitution ?image
WHERE
{
    ?item wdt:P31 wd:Q5; # is human
          wdt:P512 wd:Q6785216 ; # holds an MLIS
          wdt:P69 ?educationalInstitution .
    ?educationalInstitution wdt:P18 ?image .
}
```

Direktlink: <https://w.wiki/4hG9>

# Beispiel Wikidata-Query V

```
#defaultView:Map

SELECT ?educationalInstitution ?geo
WHERE
{
    ?item wdt:P31 wd:Q5; # is human
          wdt:P512 wd:Q6785216 ; # holds an MLIS
          wdt:P69 ?educationalInstitution .
    ?educationalInstitution wdt:P625 ?geo .
}
```

Direktlink: <https://w.wiki/4hGA>

# Übung: SPARQL-Abfrage erweitern

-> <https://w.wiki/4ixR>

# Wikidata kann für Vieles nützlich sein

- Generierung einer Liste offizieller Ländercodes
- Aufbau und Pflege einer Raumsystematik
- Sammlung bibliographischer Informationen und von Personen- und Institutionsprofilen (siehe [Scholia](#))
- Citizen Science Projekt zur Erschließung der digitalisierten Artikel einer Illustrierten ([Die Datenlaube](#))
- Erweiterung der Katalogsuche, z.B. Vorschläge von Personen bei bestimmten Suchbegriffen (siehe die [SWIB20-Präsentation von Bernd Uttenweiler](#))

# 3. Linked Open Usable Data

# Ein Imageproblem?

*Ultimately, RDF and the Semantic Web are of no interest to Web developers. They also have a really negative public perception problem. We should stop talking about them. Let's shift the focus to be on Linked Data, explaining the problems that Web developers face today, and concrete, demonstrable solutions to those problems.*

– Manu Sporny, damals Vorsitzender der RDFa Working Group beim W3C, der JSON-LD Community Group & Mitglied weiterer Semantic-Web-Gruppen, beim Schreiben an der JSON-LD-Spezifikation **Sporny (2012)**

# Linked Open Usable Data

A screenshot of a Twitter post from user @azaroth42. The post features a profile picture of a man in a suit, the name 'Rob Sanderson' in bold black text, and the handle '@azaroth42'. To the right is a blue button with the German word 'Folge ich' (Follow me) and a downward arrow. The tweet's text is in large, bold, black font: "'Open' is not necessarily simple nor useful, just available. We need fans to get LOUD: Linked Open Useful Data :)" Below the tweet are two smaller pieces of text: a globe icon followed by 'Tweet übersetzen' and the timestamp '19:39 - 23. Aug. 2016'.

"Open" is not necessarily simple nor useful,  
just available. We need fans to get LOUD:  
Linked Open Useful Data :)

Quelle: [Rob Sanderson auf Twitter](#), siehe auch Rob Sanderson's Europeanatech 2018 Keynote ([Slides](#), [Aufzeichnung](#))

# *Using data?*

- Daten werden mit existierender Software bearbeitet (ausgewertet, ergänzt, integriert etc.)
- Entwicklung neuer Software zur Interaktion mit Daten
- LOUD: Orientierung auf Bedürfnisse und Konventionen rund um Software (Entwicklung, Standards, etc.)
- Nützliche Daten: Zielgruppe kennen & eigene Angebote auf sie ausrichten
- Hauptzielgruppe: Entwickler:innen oder Nutzer:innen von Software

# LOUD = Web-APIs mit JSON-LD

- Entwickler:innen sind Kernzielgruppe von Open-Data-Publikation
- APIs sind das User Interface von Entwickler:innen

*When it comes to APIs, developers are your users. The same principles of user-centred-design apply to the development and publication of APIs (simplicity, obviousness, fit-for-purpose etc).*

Quelle: <https://apiguide.readthedocs.io/en/latest/principles/empathy.html>

# JSON-LD

- "designed to be usable directly as JSON, with no knowledge of RDF" – Es ist richtiges JSON!
- "also designed to be usable as RDF" –  
<https://www.w3.org/TR/json-ld/>

# JSON-LD = JSON + @context

```
{  
  "@context": {  
    "Person": "https://schema.org/Person",  
    "name": "https://schema.org/name",  
    "interest": "http://xmlns.com/foaf/0.1/interest"  
  },  
  "@id": "https://orcid.org/0000-0001-9083-7442",  
  "@type": "Person",  
  "name": "Adrian",  
  "interest": [  
    "http://www.wikidata.org/entity/Q466",  
    "http://www.wikidata.org/entity/Q28134421",  
    "http://www.wikidata.org/entity/Q341",  
    "http://www.wikidata.org/entity/Q380962",  
    "http://www.wikidata.org/entity/Q56298524"  
  ]  
}
```

# Übung: JSON-LD-Kontext schreiben

Ergänze einen Kontext im JSON-LD Playground:

<http://tinyurl.com/ya3ydscw>

# Warum nicht einfach JSON?

*Usable* heisst nicht nur dass die Daten strukturell gut nutzbar sind, sondern auch einige Vorteile von Linked Data mit sich bringen:

1. Implizite Dokumentation: Properties haben URIs und ihre Bedeutung lässt sich nachschlagen
2. Interoperabilität: wer faul ist, nutzt nach und muss nicht selbst spezifizieren
3. Things not strings: Präferenz von URIs über Strings führt zu Erstellung, Nutzung und Pflege von SKOS-Schemas für kontrollierte Wertelisten

# LOUD-Beispiele

- lobid
- IIIF
- Linked Art
- OERSI und AMB

# Low Hanging Fruit I

- JSON-LD lässt sich auch in HTML-Seiten einbetten mit  
`<script type="application/ld+json">`
- für viele Content-Management-Systeme gibt es Plugins, z.B.  
für WordPress **Schema** oder **Structured Content**

# 4. SKOS



- SKOS: Simple Knowledge Organization System
- Datenmodell (RDFS/OWL) für die maschinenlesbare und webbasierte Publikation kontrollierter Vokabulare als Linked Open Data 2009 als Empfehlung des World Wide Web Consortiums (W3C) veröffentlicht
- Unterstützt Mehrsprachigkeit, Abbildung von Beziehungen & einiges mehr

# Kontrollierte Vokabulare als LOD

- **auflösbare IDs:** ein HTTP URI (Uniform Resource Identifier) für ein Vokabular und jeden Wert
- **Verlinkung:** Werte sind innerhalb eines Vokabulars und zwischen Vokabularen verlinkt
- **Erweiterbarkeit:** weitere Aussagen oder Funktionen auf Basis anderer Webstandards können ergänzt werden

# Vorteile

- Webintegration
- Maschinenlesbarkeit
- Interoperabilität
- Nachnutzbarkeit
- Entkoppelung von Datenmodellierung und Softwareentwicklung

# SKOS-Namespace und -Prefix

@prefix skos:  
<<http://www.w3.org/2004/02/skos/core#>>.

# Aufbau eines SKOS-Vokabulars

# Zwei grundlegende Typen von Einträgen

- ConceptScheme:
  - das Vokabular selbst, wird genau **1 Mal** pro Vokabular definiert
  - generelle Informationen wie Lizenz, Titel
- Concept:
  - ein Wert des Vokabulars
  - Informationen wie Label, Beziehungen

# Bsp.: Concept Scheme

```
@prefix colour: <https://example.org/colour-with-hierarchy/> .  
@prefix dct: <http://purl.org/dc/terms/> .  
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .  
  
colour: a skos:ConceptScheme ;  
    dct:title "Colour Vocabulary"@en, "Farbvokabular"@de ;  
    dct:creator "Hans Dampf"@de ;  
    dct:created "2021-11-02" ;  
    dct:license  
<https://creativecommons.org/publicdomain/zero/1.0/> .
```

# Wichtige SKOS-Properties

- Verbindung zwischen Vokabular und seinen Werten:  
hasTopConcept, topConceptOf, inScheme
- Vorzugsbezeichnung: prefLabel
- Alternativbezeichnung: altLabel
- hierarchische Informationen: narrower/broader

# Bsp.: SKOS Concept I

```
colour:violet a skos:Concept ;  
    skos:prefLabel "Violett"@de, "violet"@en ;  
    skos:altLabel "Lila"@de, "purple"@en ;  
    skos:topConceptOf colour: .
```

```
colour: skos:hasTopConcept colour:violet .
```

# Bsp.: SKOS Concept II

```
colour:green a skos:Concept ;  
    skos:prefLabel "Green"@en, "Grün"@de ;  
    skos:narrower colour:grasgreen ;  
    skos:topConceptOf colour: .
```

```
colour:grassgreen a skos:Concept ;  
    skos:prefLabel "grass-green"@en, "Grasgrün"@de ;  
    skos:inScheme colour: .
```

# Sonstige Properties I

- Notation: `notation`
- Dokumentation des Vokabulars und seiner Entwicklung:  
`changeNote`, `definition`, `editorialNote`, `example`,  
`note`, `historyNote`, `scopeNote`
- Externe Relationen: `broadMatch`, `narrowMatch`,  
`exactMatch`, `closeMatch`, `mappingRelation`

# Bsp.: SKOS Concept III

```
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .  
  
<https://psyndex.de/vocab/psycharchives-types/conferenceObject>  
  a skos:Concept ;  
  skos:altLabel "Konferenzveröffentlichung"@de, "conference  
output"@en ;  
  skos:exactMatch <http://purl.org/coar/resource_type/c_c94f>  
;  
  skos:example "presentation slides, conference programs and  
abstract collections, posters"@en ;  
  skos:inScheme <https://psyndex.de/vocab/psycharchives-  
types/> ;  
  skos:narrowMatch  
<http://purl.org/spar/fabio/ConferencePoster>,  
<http://purl.org/spar/fabio/Presentation> ;  
  .
```

Quelle: <http://zpidvokabulare.surge.sh/psyndex.de/vocab/psycharchives-types/conferenceObject.de.html>

# Sonstige Properties II

- Collection, OrderedCollection, member, memberList
- hiddenLabel
- broaderTransitive, narrowerTransitive, related, semanticRelation
- Siehe auch diese **erläuterte Liste der Elemente**

# SkoHub Vocabs

- Software zum Publizieren von SKOS-Vokabularen als statische Webseiten
- leicht aufzusetzen und zu deployen
- git-basierte Redaktionsprozesse
- Integration mit GitHub und GitLab via Webhooks
- “serverless” Hosting möglich

# Beispielvokabulare

- Kompetenzzentrum Interoperable Daten (KIM):
  - Hochschulfächersystematik
  - Resource Types
  - Schulfächer
- Kerndatensatz Forschung: Interdisziplinäre Forschungsfeldklassifikation
- Wir lernen online: <https://vocabs.openeduhub.de/>
- ZPID (Test), leibniz-psychology.org
- MPIeR Categories of Matters regulated by Police Ordinances

# Low Hanging Fruit II

- existierende vom Bildungsserver gepflegte kontrollierte Vokabulare nach SKOS überführen und mit SkoHub Vocabs veröffentlichen
- kann auch über skohub.io, ein Service des hbz, publiziert werden

# 5. Fragen und Diskussion

# Ende