

# Projektpräsentation im Fach Semantic Web

Hochschule für Technik, Wirtschaft und Kultur Leipzig  
Masterstudiengang: Medieninformatik  
Referent: Mario Reyes Napoles  
Dozent: Prof. Dr. Thomas Riechert  
Fach: SemanticWeb

# Gliederung

1. Fragestellung
2. Verwendete Datenquellen
3. Status Quo des Projekts
4. Dcat- Vokabular
5. Next Steps

# Semantische Aufbereitung von Datenquellen

Fragestellung:

Wie können vergangene Projektdatenquellen semantisch aufbereitet werden, dass diese zukünftig für vereinfachte Recherche allen zukünftigen Kursteilnehmern zur Verfügung stehen

# Datenquellen

- Projektarbeiten aus dem letzten Semester
  - 23 Projektarbeiten
  - 64 Datensätze
- Datensätze liegen in unterschiedlichen Formaten vor
- Datensätze wurden unterschiedliche verarbeitet und verfügbar gemacht
  - Extraktion
  - RDF-Triple
  - Triplestore
  - Verlinkung

# Status Quo

- Analyse der Daten ✓
- Bewertung der Quellen ✓
- Semantische Beschreibung ✓
- Semantische Anreicherung der Datenquellen ✓
  - Dcat Vokabular
  - Openrefine



**Ziel: Die so aufbereiteten Datenquellen sollen durch einen eigens dafür eingerichteten Triplestore durchsuchbar gemacht werden.**

# OpenRefine

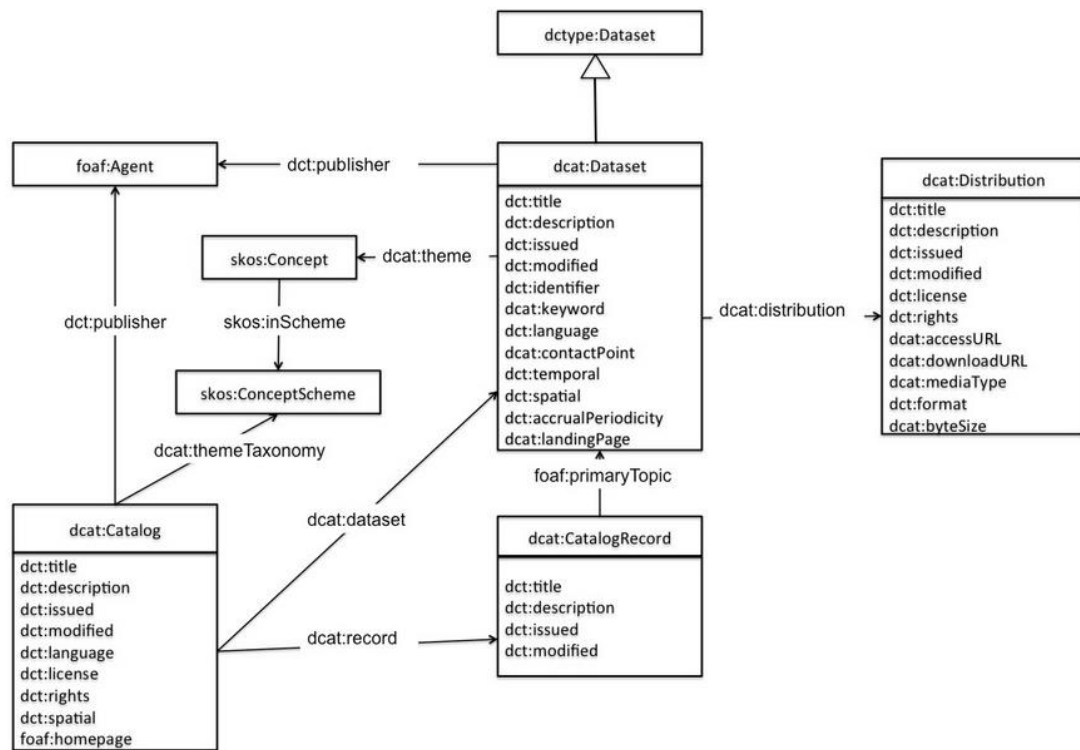
Show as: **rows** records      Show: 5 10 25 50 records      « first ‹ previous **1 - 1** next › last »

	1.Projekt	Agent	Data Catalog: Ti	Data Catalog: D	Data Catalog: Language	Data Catalog: H	DataCatalog:Da	Dataset: Title	Dataset: Descri	Dataset: Keywo	Dataset: Data
1.	Familienfreundlichste Stadtteile der Stadt Leipzig	Alexander Mueller	Datenbasis des Leipzig Data Projekts	Die Leipziger Initiative für Offene Daten ist angetreten, um die Bemühungen zur Etablierung Offener Daten als wesentlichen Teil einer sich entfaltenden Weblandschaft in der Leipziger Region voranzubringen.	<a href="http://id.loc.gov/vocabulary/iso639-1/de.html">http://id.loc.gov/vocabulary/iso639-1/de.html</a>	<a href="http://leipzig-data.de/">http://leipzig-data.de/</a>	Adressen in der Stadt Leipzig, Schulen der Stadt Leipzig, Stadtbezirke und Ortsteile	Adressen in der Stadt Leipzig	Es sind keine Kommentare, Beschreibungen oder Anmerkungen in der Wissensbasis enthalten.	Adressen, Stadt, Leipzig	<a href="http://ic1/de.ht">http://ic1/de.ht</a>
								Dataset: Title	Dataset: Description	Dataset: Keyword	Dataset: Data
								Schulen der Stadt Leipzig,	ohne berufliche Schulen und Förderschulen. 2014-07-20: Abgleich mit den Stadtseiten.	Schulen, Stadt, Leipzig	<a href="http://ic1/de.ht">http://ic1/de.ht</a>
								Dataset: Title	Dataset: Description	Dataset: Keyword	Dataset: Data
								Stadtbezirke und Ortsteile	Es sind keine Kommentare, Beschreibungen oder Anmerkungen in der Wissensbasis enthalten.	Stadtbezirke, Ortsteile, Stadt, Leipzig	<a href="http://ic1/de.ht">http://ic1/de.ht</a>

# Dcat- Vokabular

- dcat ist ein Standardvokabular in RDF-Schema für Open-Data-Kataloge
  - Die grundlegenden Begriffe des dcat Vokabulars sind *dataset*, *catalog record* und *catalog*:
  - **dataset** ist eine Sammlung von Informationen in einem maschinenlesbaren Format
  - **catalog record** besteht aus Metadaten für eine Datenmenge . Es beschreibt also den Datensatz .
  - **catalog** ist eine Sammlung von catalog records.

# Dcat





# Beispiel

RDF Skeleton

RDF Preview

This is a sample Turtle representation of (up-to) the first 10 rows

```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix rdfs: <http://www.w3.org/ns/rdfs#> .
@prefix rdfa: <http://www.w3.org/ns/rdfa#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix dcat: <http://www.w3.org/ns/dcat#> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix dc: <http://purl.org/dc/elements/1.1/> .

<http://127.0.0.1:3333/+Familienfreundlichste+Stadtteile+der+Stadt+Leipzig> a dcat:Catalog ;
    dc:title "Datenbasis des Leipzig Data Projekts" ;
    rdfs:label "Datenbasis des Leipzig Data Projekts" ;
    dc:publisher "Alexander Mueller" ;
    dc:language "http://id.loc.gov/vocabulary/iso639-1/de.html" ;
    foaf:homepage "http://leipzig-data.de/" ;
    dcat:dataset "0"@schulen der stadt leipzig, adressen in der stadt leipzig, stadtbezirke und ortsteile .

<http://127.0.0.1:3333/Adressen+in+der+Stadt+Leipzig> a dcat:Dataset ;
    dc:title "Adressen in der Stadt Leipzig" ;
    dc:description ""Es sind keine Kommentare, Beschreibungen oder
Anmerkungen in der Wissensbasis enthalten."" ;
    dcat:keyword "Adressen, Stadt, Leipzig" ;
    dc:language <http://id.loc.gov/vocabulary/iso639-1/de.html> .

<http://127.0.0.1:3333/Dataset%3A+Title> a dcat:Dataset ;
    dc:title "Dataset: Title" ;
    dc:description "Dataset: Description" ;
    dcat:dataset "Dataset: Keyword" .
```

# Next Steps

- Verlinkung der einzelnen Datenquellen
- Daten im Triplestore zur Verfügung stellen
  - [Apache Jena Fuseki 1.1.2](#)
- Dokumentation

Vielen Dank für eure Aufmerksamkeit