



Open Access Netzwerk

als Linked (Open) Data

- Linked Data Definition
- Standards für Linked Data
- Maximalziele im OA-Netzwerk
- Lösungsansätze
- D2R Server
- OA-Netzwerk als Linked Open Data: ein Beispiel
- Zukunft von OA-Netzwerk: mögliche Prozessierungsketten

Linked Data Regeln

Tim Berners-Lee

Date: 2006-07-27, last change: \$Date: 2009/06/18 18:24:33 \$

Status: personal view only. Editing status: imperfect but published.

[Up to Design Issues](#)

Linked Data

The Semantic Web isn't just about putting data on the web. It is about making links, so that a person or machine can explore the web of data. With linked data, when you have some of it, you can find other, related, data.

Like the web of hypertext, the web of data is constructed with documents on the web. However, unlike the web of hypertext, where links are relationships anchors in hypertext documents written in HTML, for data they links between arbitrary things described by RDF. The URIs identify any kind of object or concept. But for HTML or RDF, the same expectations apply to make the web grow:

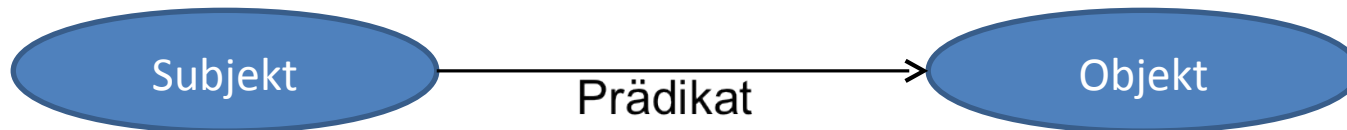
1. Use URIs as names for things
2. Use HTTP URIs so that people can look up those names.
3. When someone looks up a URI, provide useful information, using the standards (RDF*, SPARQL)
4. Include links to other URIs, so that they can discover more things.



<http://www.w3.org/DesignIssues/LinkedData.html>

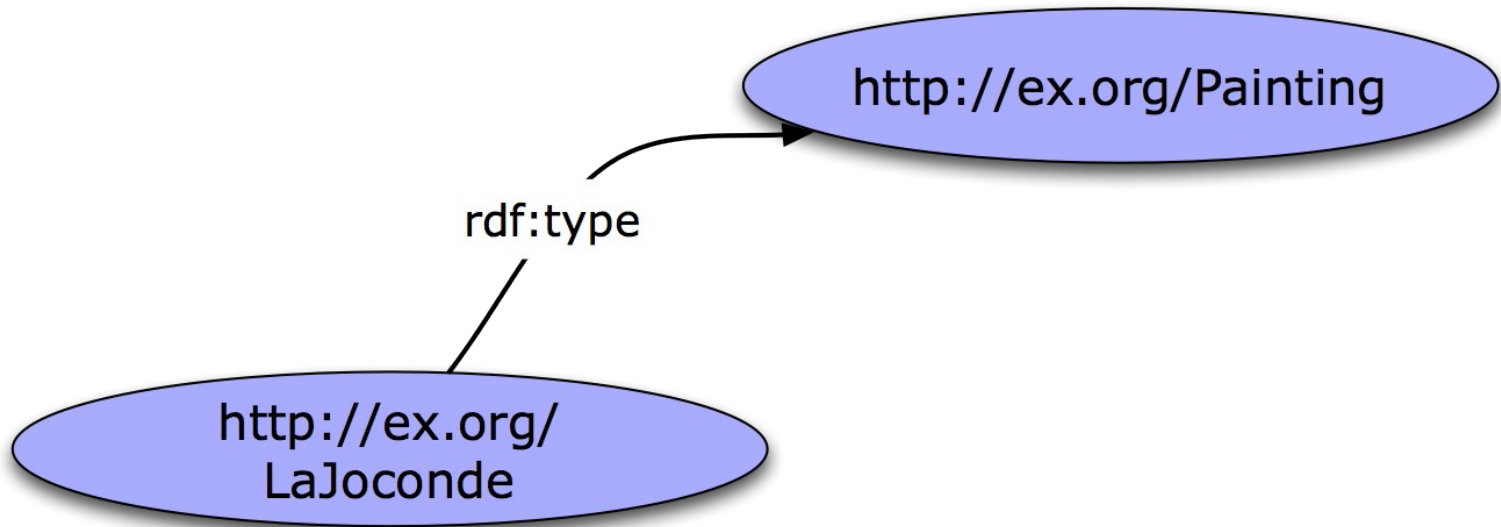
RDF (S)

- W3C-Standard für die Modellierung von Informationen
- WWW Entitäten ('Resources') werden durch URIs identifiziert
- RDF macht Aussagen über Entitäten in Form von 'Tripeln':

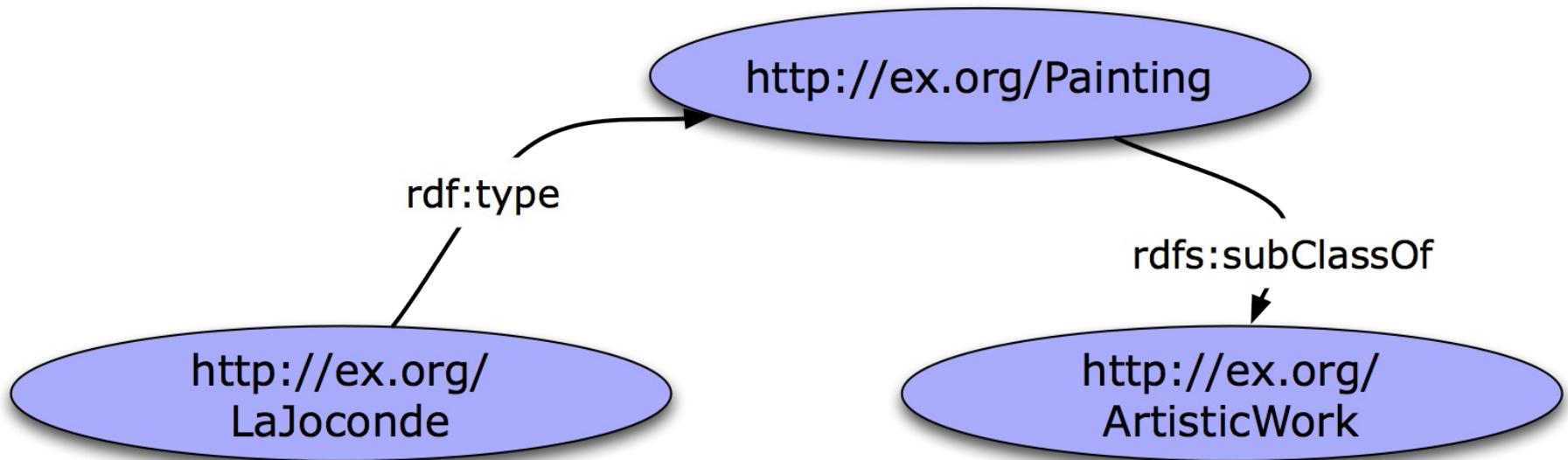


- Resource Description Framework Schema (RDFS)
- W3C-Standard, der RDF-Tripeln elementare Semantik hinzufügt
- Basis für (einfache) Ontologien: domänenspezifische Vokabularien mit (Sub-)Klassen, (Sub-)Eigenschaften und Schlussregeln

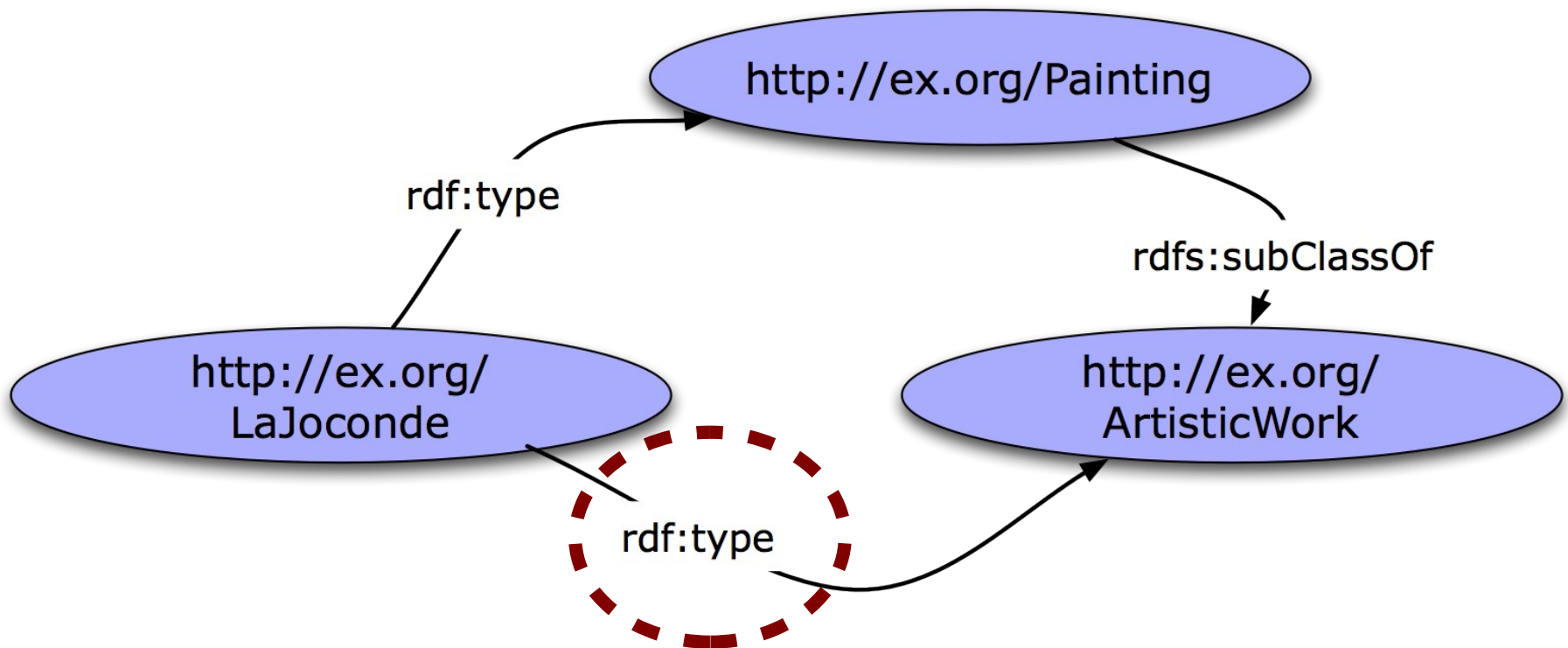
Triple Sets und Reasoning (1)



Triple Sets und Reasoning (2)



Triple Sets und Reasoning (3)



→ **Potential für Digitale Heuristiken!**

Inferenzen: Beispiele aus der Wirtschaft

- Gibt es eine direkte Zugverbindung zwischen Berlin Hbf. und dem Hauptbahnhof Sants in Barcelona?
 - Nein, weil jedes Land seine eigene Schienenführungssysteme
- Kann man eine Solaranlage in Marokko bauen, die Strom in das Schwedische Stromnetz einspeist?
 - Nein, weil auch hier einheitliche Standards fehlen
- In der Wirtschaft arbeitet man schon seit Jahren an die Standardisierung und Vernetzung der vorhandenen heterogenen Netzwerke

- Und warum nicht in der Wissenschaft?

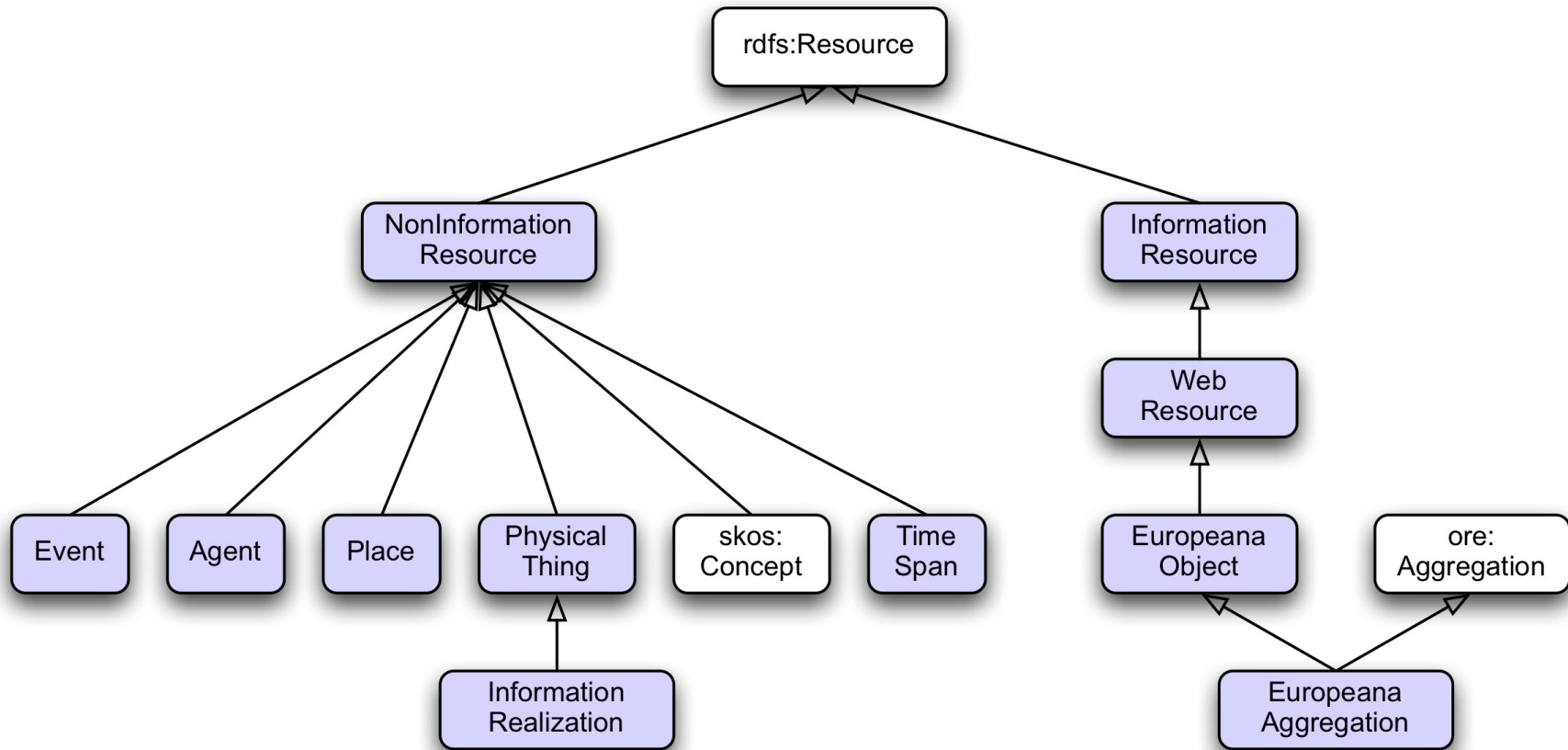
OAI-ORE

- Herbert Van de Sompel, Carl Lagoze
- Standardspezifikation seit Oktober 2008
- Erweiterung des OAI-PMH Standard
- Versionen und Formate (Volltext in PDF und HTML, Metadaten in RDF etc.)
- Teile (Kapitel, Bilder, Dateien)
- Verknüpfungen zu anderen Objekten (Zitation, Versionierung)
- Aggregationen mittels URI's identifizieren
- Dokumenten-Binnenstruktur in Resource Maps abbilden
- Austausch und Nachnutzung von digitalen Objekten

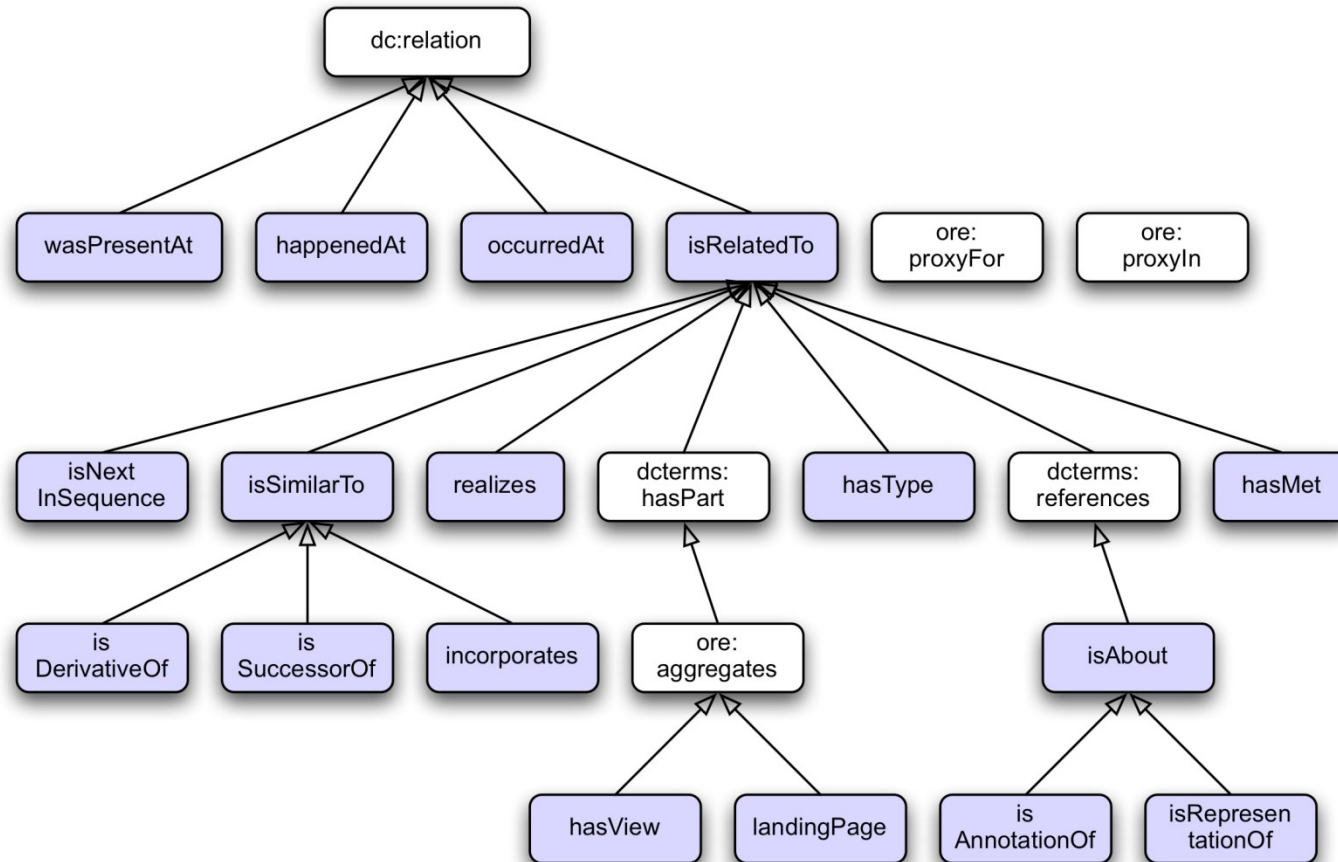
Europeana Datenmodell

- Globalziel: Granularität und Ausdrucksstärke der Originaldaten erhalten – und doch Interoperabilität herstellen
- Basiert auf OAI-ORE, DCTerms, SKOS und CIDOC-CRM
- Erlaubt Spezialisierung durch Communities (AP)
- RDF-graphenbasiert
- Unterscheidung zwischen Objekt und Beschreibungsdaten
- Mehrere Repräsentationen eines Objektes zulassen, die potentiell einander widersprechende Aussagen enthalten
- Zusammengesetzte Objekte modellierbar machen
- Standard-Metadatenformat mit Spezialisierungsoption
- Standard-Vokabular mit Spezialisierungsoption
- EDM basiert wo immer möglich auf existierenden Standards

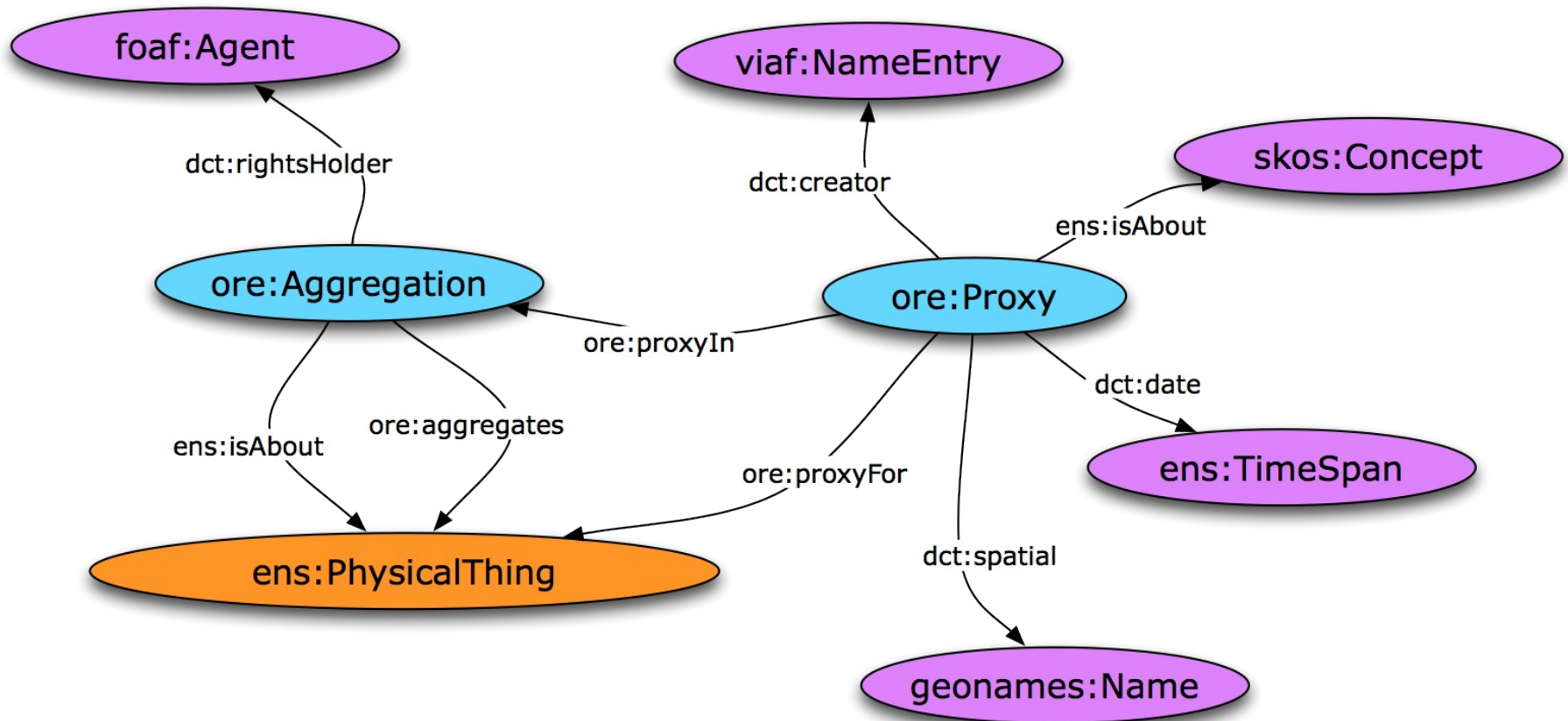
Klassen – Taxonomie EDM Version 5



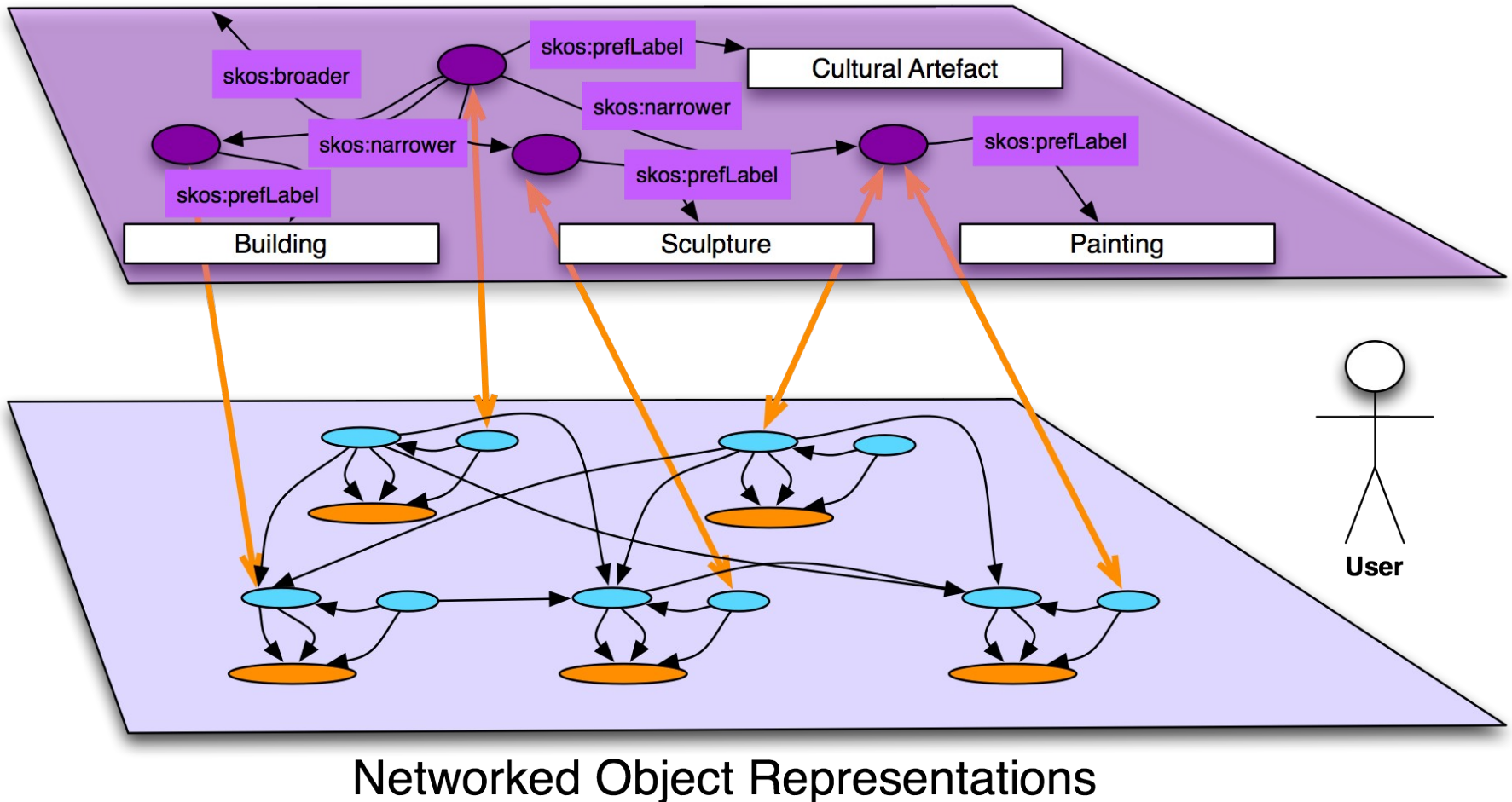
Taxonomie der EDM-Eigenschaften Version 5



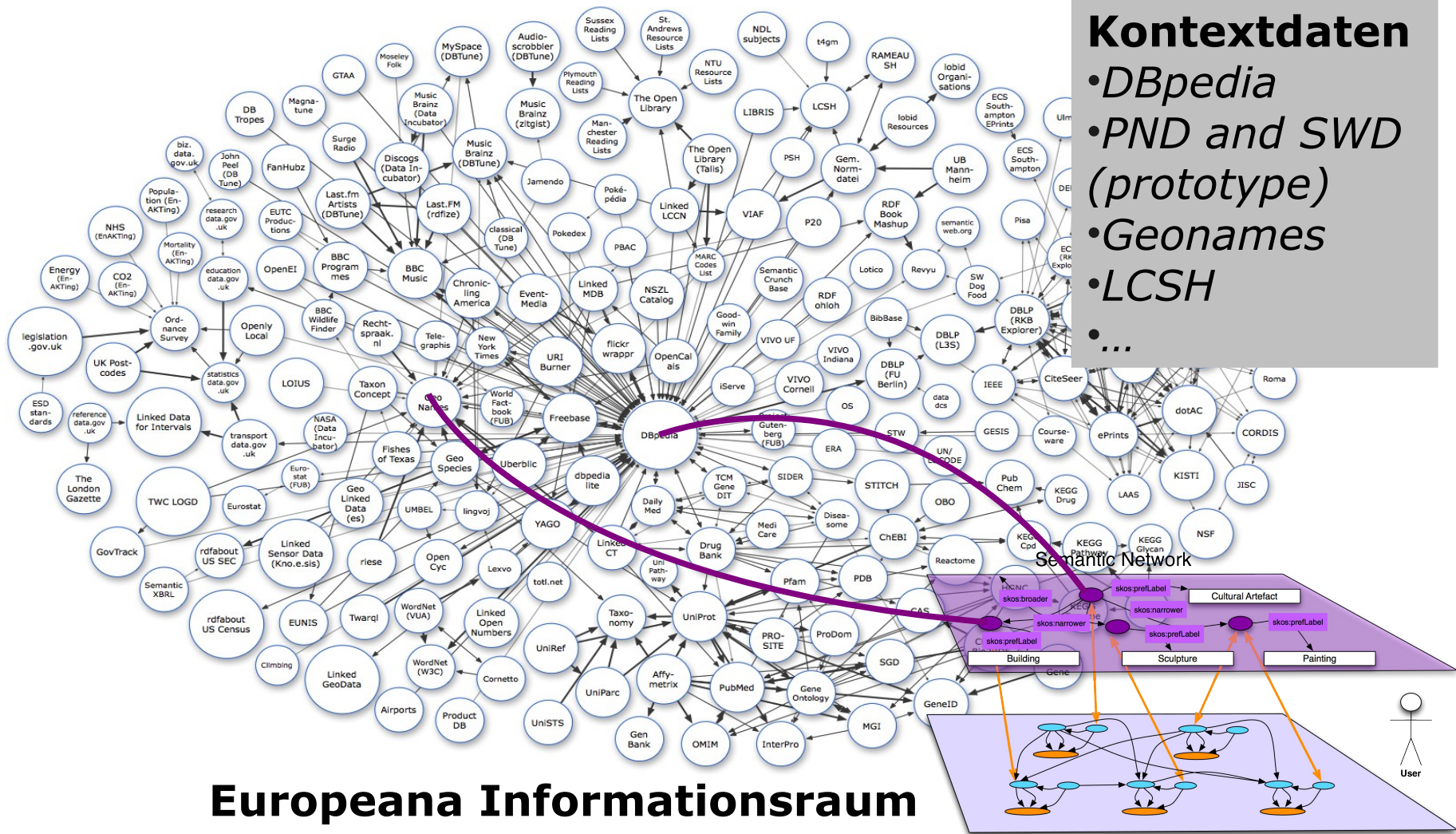
EDM-Aggregation und Kontext



EDM in der Gesamtarchitektur von Europeana Semantic Network



EDM und Linked Open Data



- ### Kontextdaten
- DBpedia
 - PND and SWD (prototype)
 - Geonames
 - LCSH
 - ...

Europeana Informationsraum

„Repositorien-Inhalte als Linked (Open) Data“

Prof. Dr. Stefan Gradmann, HU-Berlin

Julia Iwanowa, OA-Netzwerk

Networked Object Representations

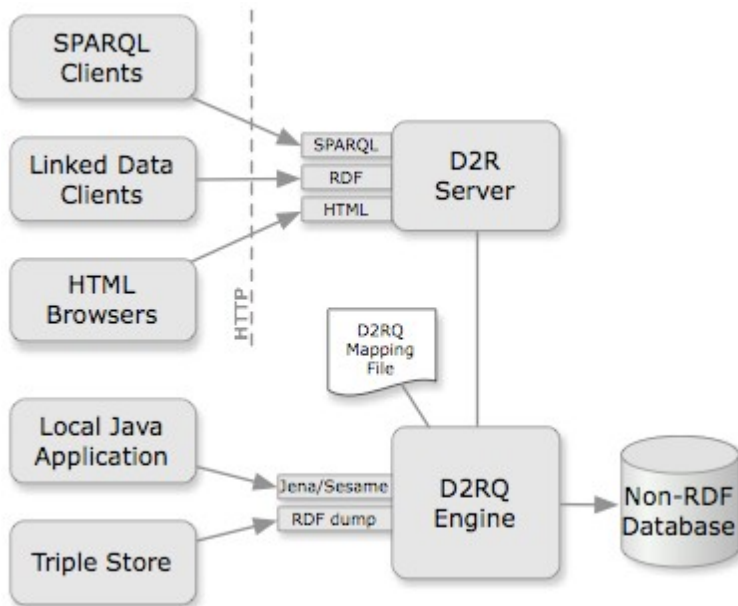
Maximalziele im OAN

- Semantische Resource Maps erstellen
- Metadatensätze nach RDF/RDFS mappen
- Europeana Data Model integrieren
- RDF Triple Store
- Automatisiertes Update
- SPARQL Schnittstelle für Client-Zugriff
- HTML und RDF Browsing der Daten
- Content als Linked Open Data freigeben

Lösungsansätze

- dezentral auf Repositorienseite
- zentral im OA-Netzwerk
- EDM Ontologie integrieren
- Jena oder Sesame Framework
- Webserver

D2RQ-Plattform



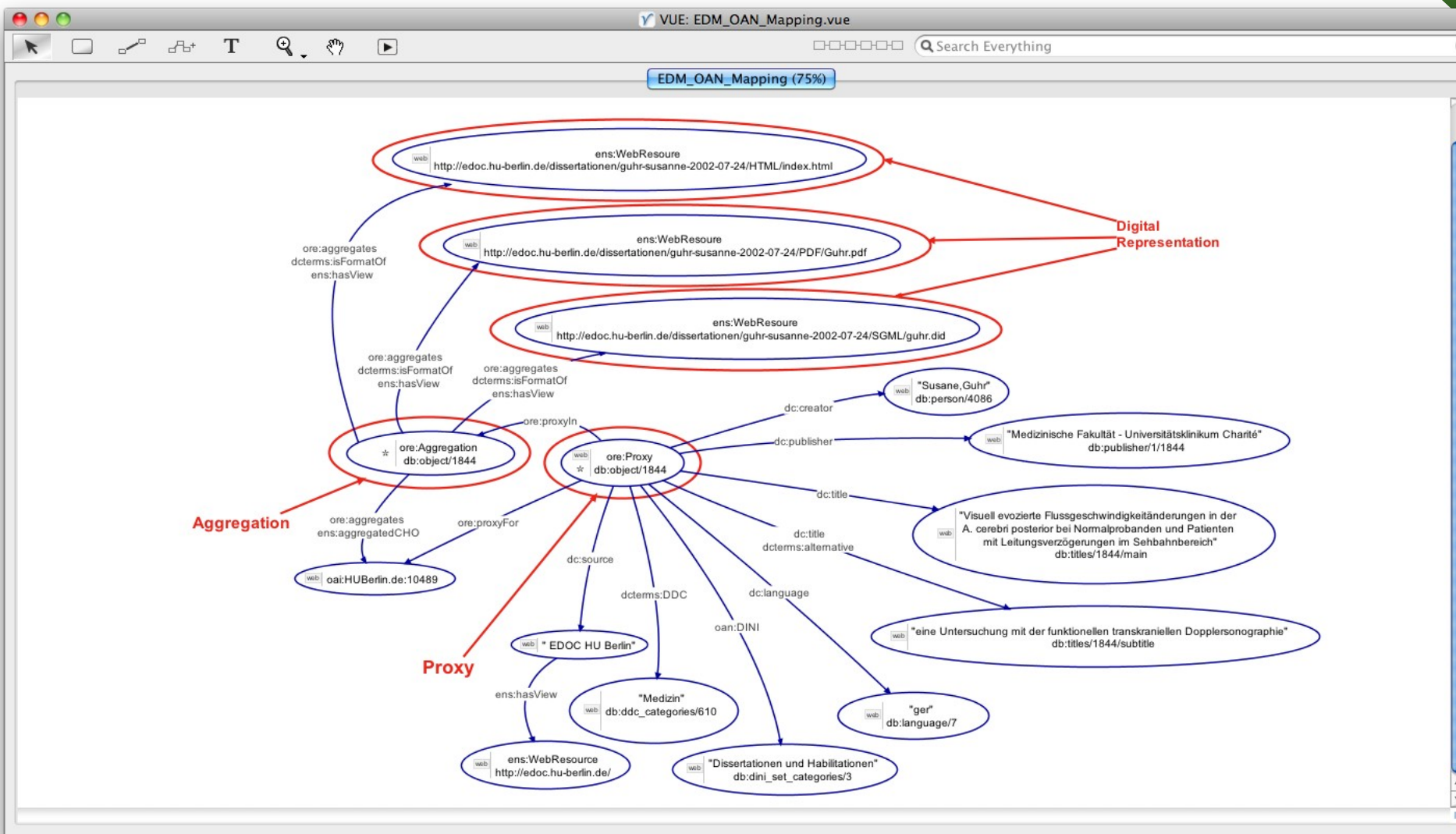
- D2RQ Mapping Language
 - Deklarative Sprache
 - Beziehungen zu beschreiben
- D2RQ Engine
 - Plug-in für Jena und Sesame API
- D2R Server
 - Linked Data Ansicht
 - HTML Ansicht zur Browsing
 - SPARQL Protocol Schnittstelle

<http://www4.wiwiss.fu-berlin.de/bizer/d2rq/spec/>

D2R-Server

- Relationale DB im Semantic Web publizieren
- Datennavigation mit HTML oder RDF Browser
- Client-Zugriff über eine SPARQL-Schittstelle
- D2RQ Plattform (D2RQ API) ermöglicht
 - Abfrage von non-RDF DB mit SPARQL query language
 - Informationszugriff auf non-RDF DB mit Jena/Sesame API
 - DB als Linked Data im Web
 - SPARQL queries über SPARQL Protocol

<http://www4.wiwiss.fu-berlin.de/bizer/d2rq/spec/>



Subjekt

Description of <http://oanet.cms.hu-berlin.de/d2r/resource/object/1844>
Resource URI: <http://oanet.cms.hu-berlin.de/d2r/resource/object/1844>

[Home](#) | [All object](#)

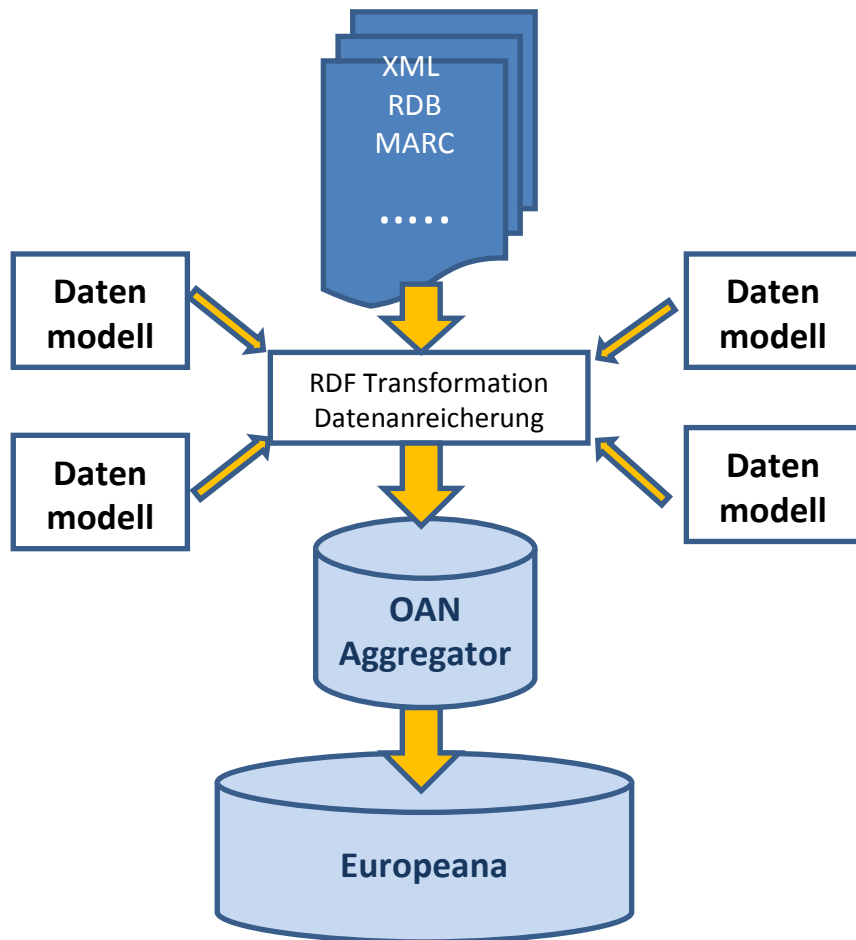
Property	Value
is dcterms:DDC of	< http://oanet.cms.hu-berlin.de/d2r/resource/ddc_categories/610 >
is dc:DINI of	< http://oanet.cms.hu-berlin.de/d2r/resource/dini_set_categories/3 >
is ore:aggregates of	< http://oanet.cms.hu-berlin.de/d2r/resource/identifier/1/1844 >
is ore:aggregates of	< http://oanet.cms.hu-berlin.de/d2r/resource/identifier/2/1844 >
is ore:aggregates of	< http://oanet.cms.hu-berlin.de/d2r/resource/identifier/3/1844 >
dc:creator	< http://oanet.cms.hu-berlin.de/d2r/resource/person/4086 >
dc:language	< http://oanet.cms.hu-berlin.de/d2r/resource/language/7 >
ore:proxyFor	< http://edoc.hu-berlin.de/OAI-2.0?verb=GetRecord&metadataPrefix=oai_dc&identifier=oai:HUBerlin.de:10489 >
is dc:publisher of	< http://oanet.cms.hu-berlin.de/d2r/resource/publisher/1/1844 >
dc:source	< http://oanet.cms.hu-berlin.de/d2r/resource/repositories/1 >
is dcterms:title of	< http://oanet.cms.hu-berlin.de/d2r/resource/titles/1844/main >
is dcterms:title of	< http://oanet.cms.hu-berlin.de/d2r/resource/titles/1844/subtitle >
rdf:type	ore:Proxy

Generated by [D2R Server](#)

OAN-Zukunftsansatz

- D2R – ein hilfreiches LD Publikationswerkzeug
- Es gibt eine Reihe von LD Publikations- und RDF Konvertierungstools
 - MARC/MODS RDFizer
 - Marc2rdf-Modeler
 - OAI-PMH RDFizer
 - RDF Crosswalk (Dublin Core)
 - DC.Metadata Gen
 - Simile Project RDFizers
 - D2R server

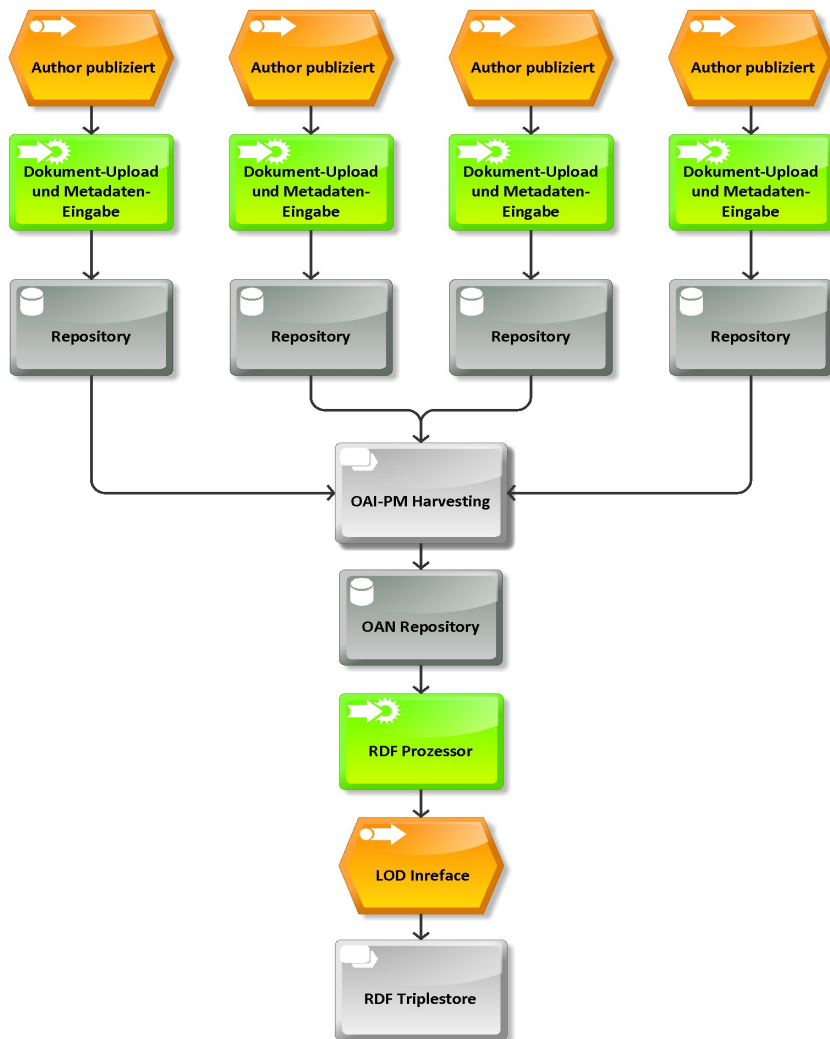
OAN-Zukunftsansatz



Aggregator für Linked Data

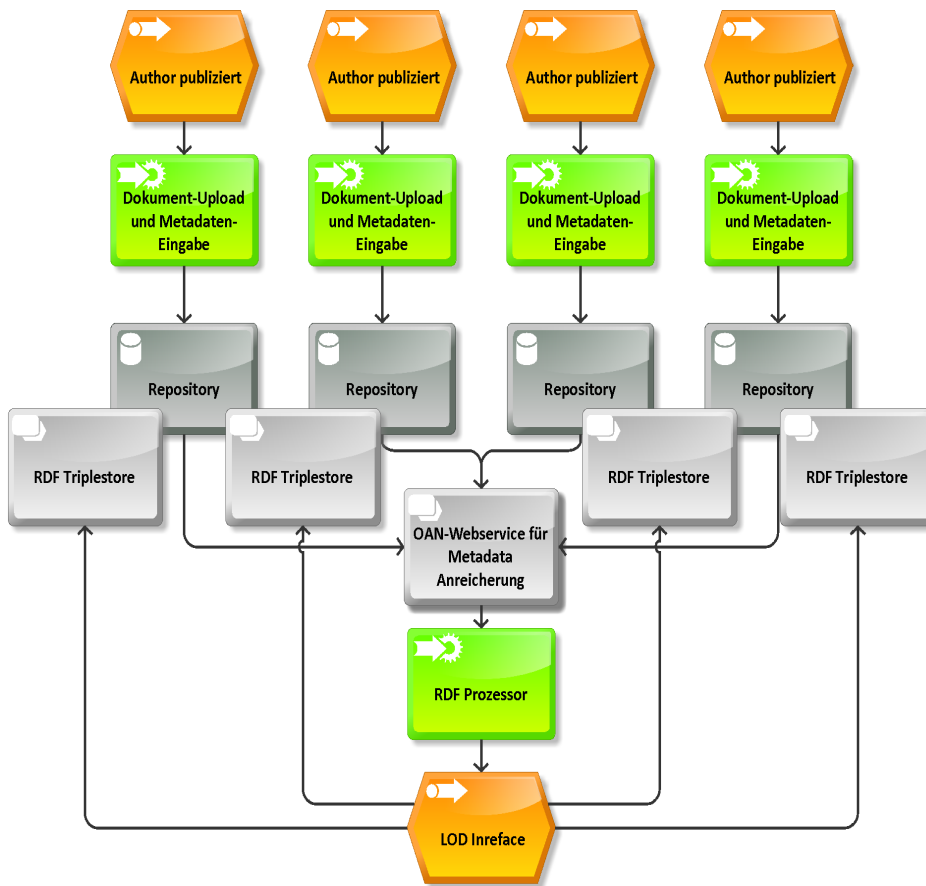
- Deutsche Digitale Bibliothek (DDB)
- Europeana

Szenario I



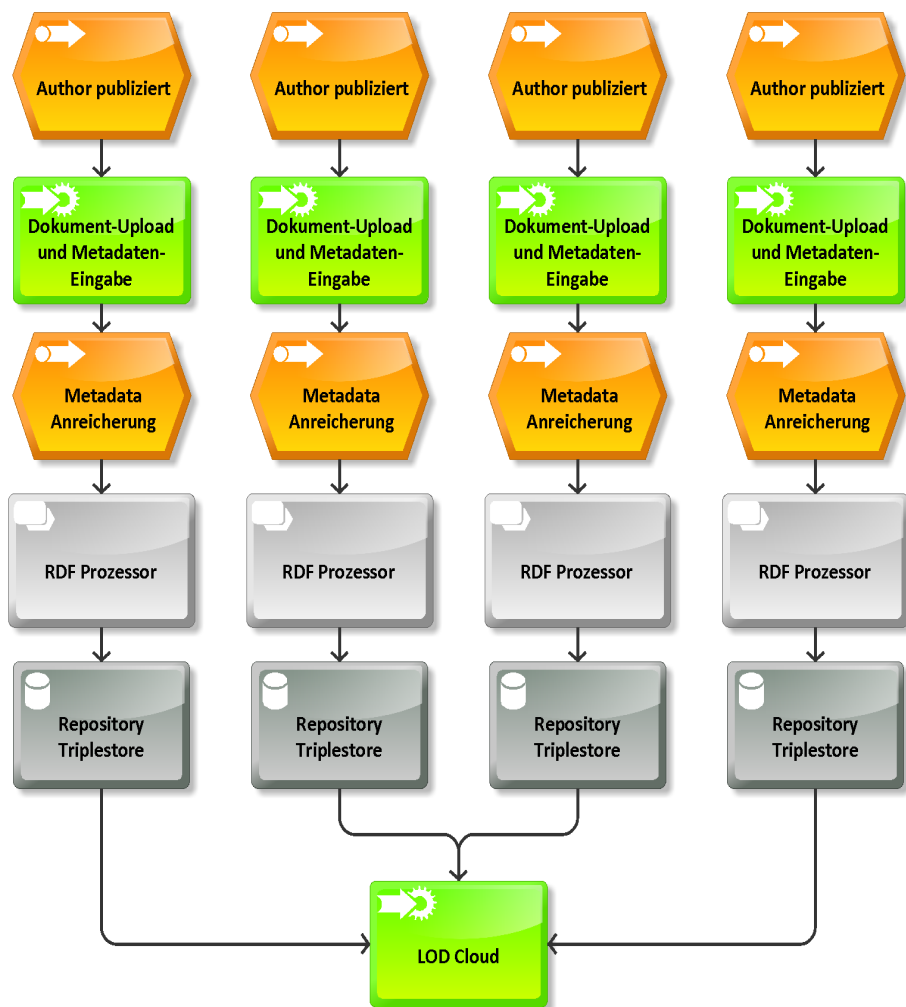
- Publikationsprozess ändert sich nicht
- Vernetzung über OAI
 - Es können nicht alle vorhandenen Daten geharvestet werden!!
 - Verlinkungen bleiben auf der Strecke
 - Vorteil: es findet eine einheitliche Datenanreicherung statt

Szenario II



- Auslesen der Daten aus den DB's
- Zentralisierte Anreicherung und Kontextualisierung der Daten
- Transformation in RDF
- Visualisieren der Daten

Szenario III



- Dezentrale Anreicherung
- Nachteile:
 - Heterogenität
 - Keine einheitlichen Standards etablierbar

Repositorien als Teil eines Workflows für die Transformation, Kontextualisierung, Anreicherung und Veröffentlichung beliebiger Inhalte als Linked (Open) Data im Semantic Web

