

Manuskripte für Europeana und neue Tools für Geisteswissenschaftler: Das DM2E-Projekt

Evelyn Dröge

Institut für Bibliotheks- und Informationswissenschaft, Humboldt-Universität zu Berlin
Vortrag auf der Herbsttagung 2013 der Fachgruppe Dokumentation
im Deutschen Museumsbund

Mit Input von Kai Eckert, Steffen Henricke, Julia Iwanowa,
Sam Leon, Christian Morbidoni, Vivien Petras, Violeta
Trkulja und anderen aus dem DM2E-Projekt.



Digitised Manuscripts to Europeana

11 Projektpartner aus ganz Europa.
Koordination: IBI, Humboldt-Universität.

Projektziele:

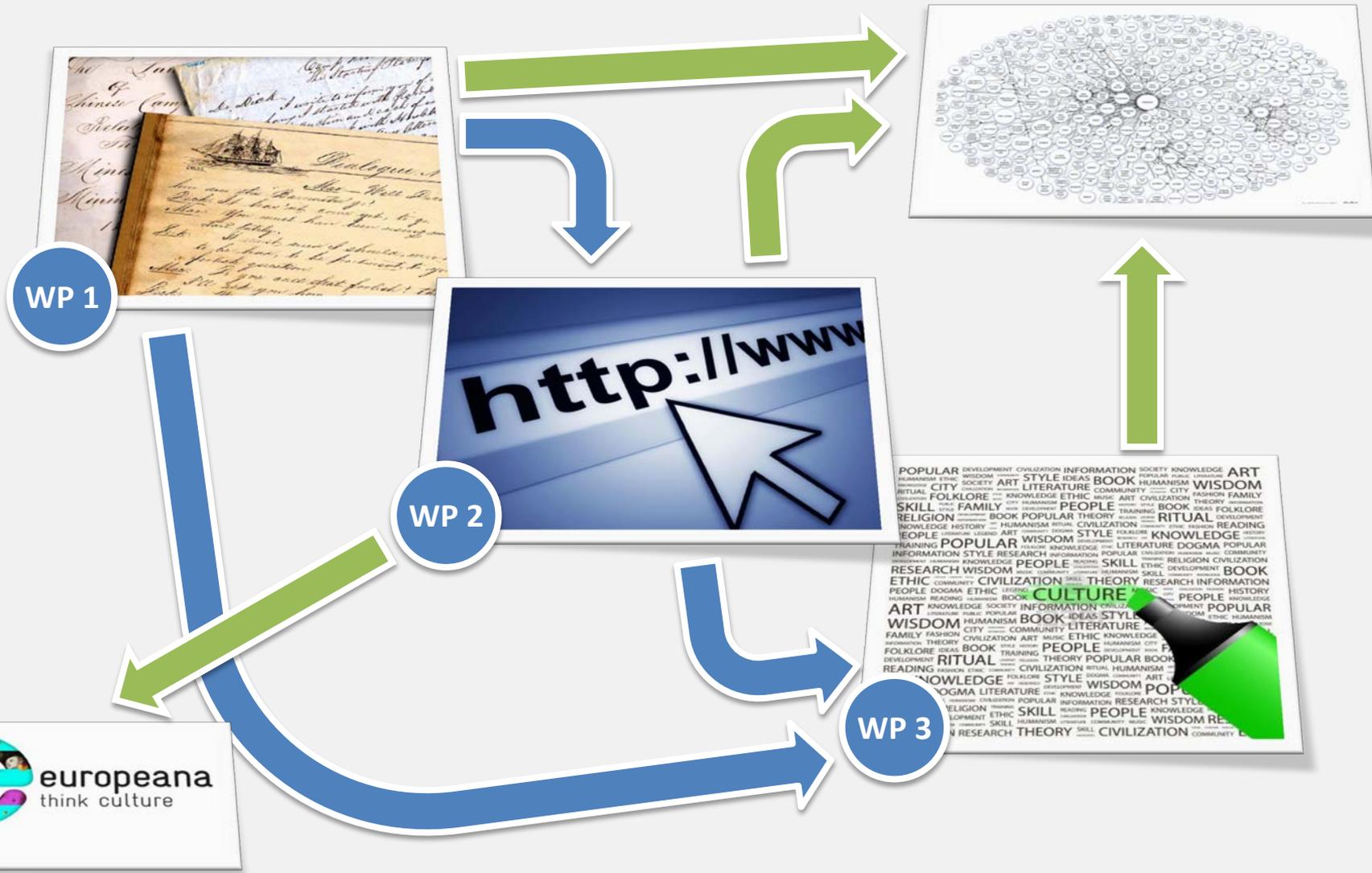
- Lieferung von Manuskripten (z.B. Wittgensteins Nachlass) nach Europeana
- Datenlieferung nach Europeana zu erleichtern
- Einbindung der Digital Humanities Community (OpenGLAM)
- Erstellung von innovativen Tools für Geisteswissenschaftler

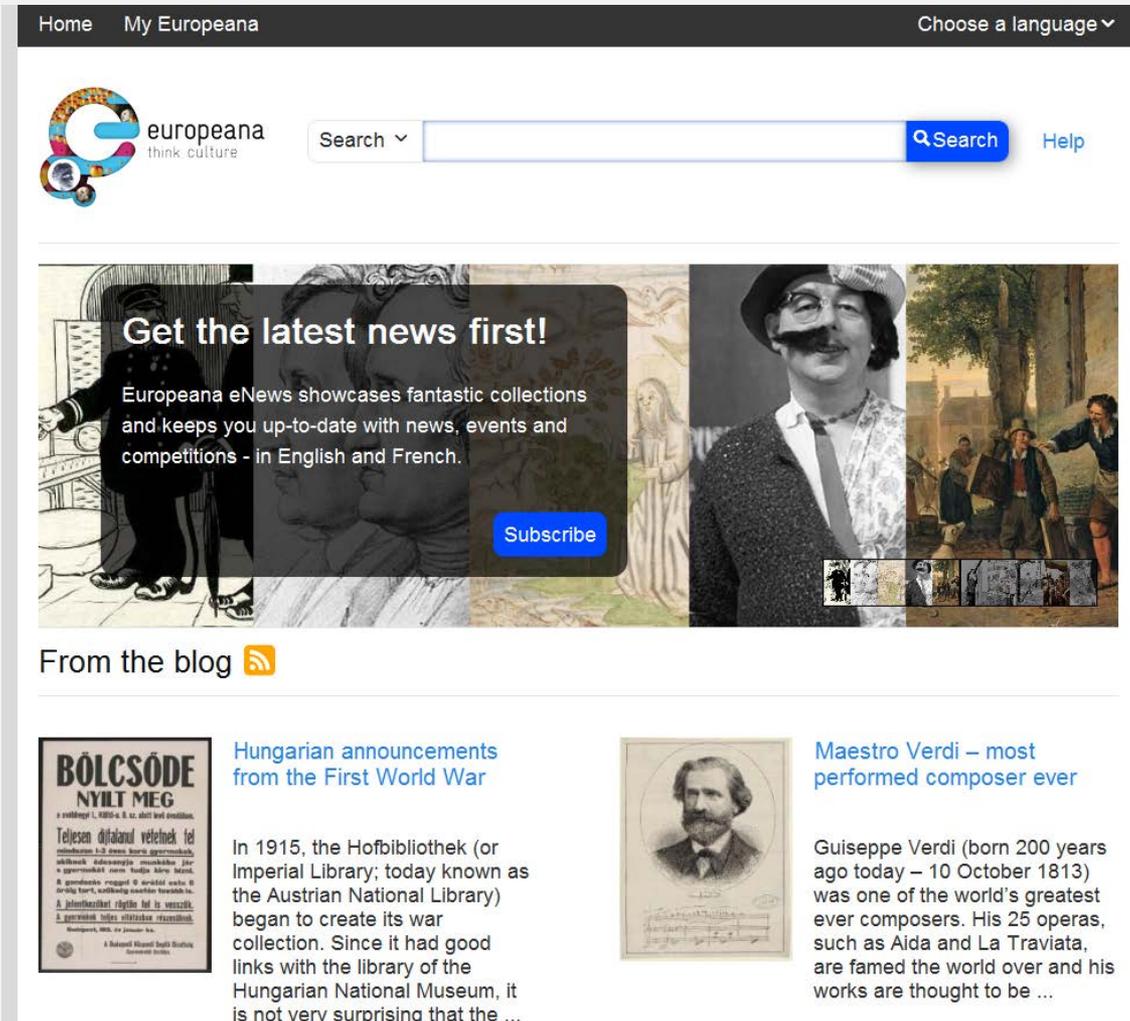
Bisher erarbeitet:

- OmNom: Tool für die Lieferung von Daten nach Europeana
- Pundit: Tool für die Erstellung semantischer Annotationen
- DM2E-Modell: Spezialisierung des EDM für Manuskripte
- Scholarly Domain Model: Repräsentation von geisteswissenschaftlicher Arbeitsweise



Projektaufbau



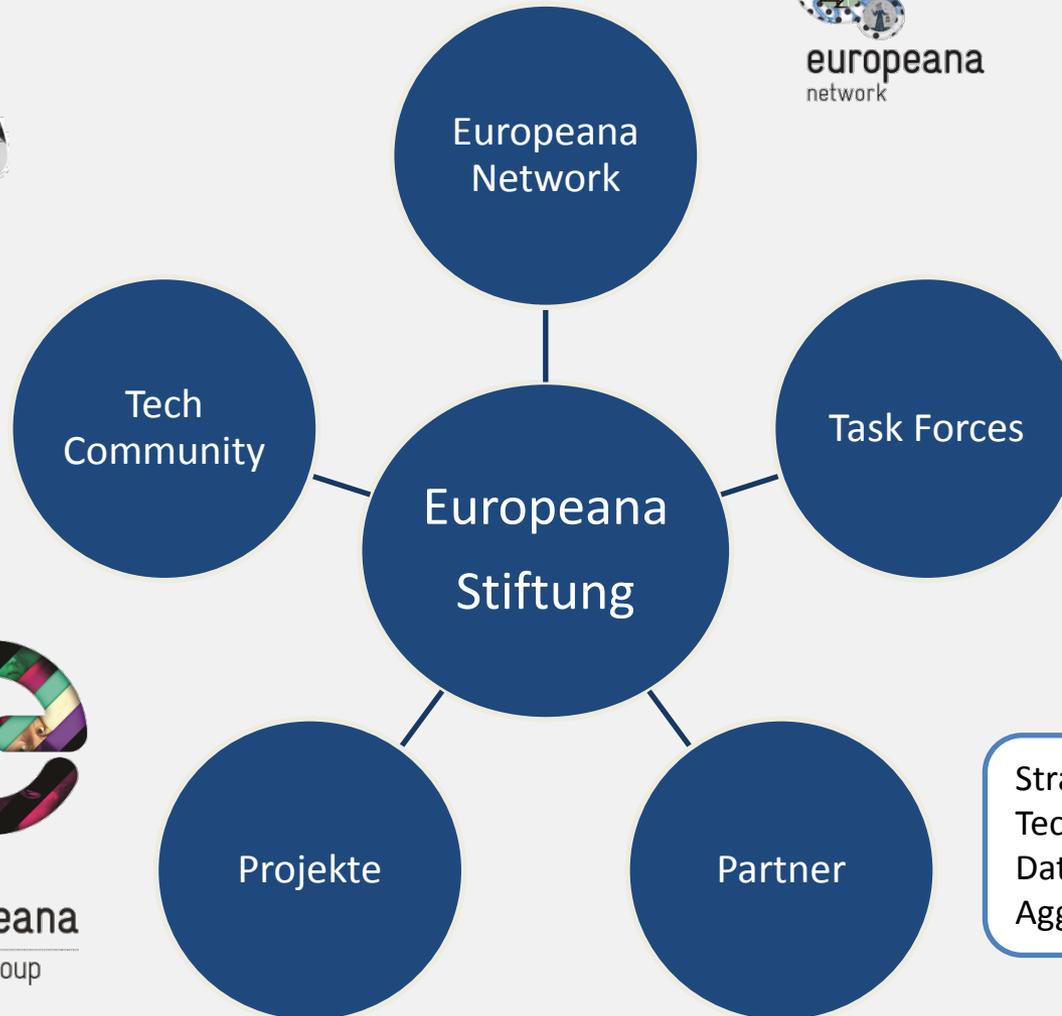


The screenshot shows the Europeana website interface. At the top, there are navigation links for 'Home' and 'My Europeana', and a 'Choose a language' dropdown menu. Below this is the Europeana logo with the tagline 'think culture'. A search bar with a search icon and a 'Search' button is present, along with a 'Help' link. The main content area features a large banner with the text 'Get the latest news first!' and a 'Subscribe' button. Below the banner, there is a section titled 'From the blog' with an RSS icon. Two featured articles are visible: 'Hungarian announcements from the First World War' and 'Maestro Verdi – most performed composer ever'. The first article includes a historical document snippet and text about the Hofbibliothek. The second article includes a portrait of Giuseppe Verdi and text about his birth and works.

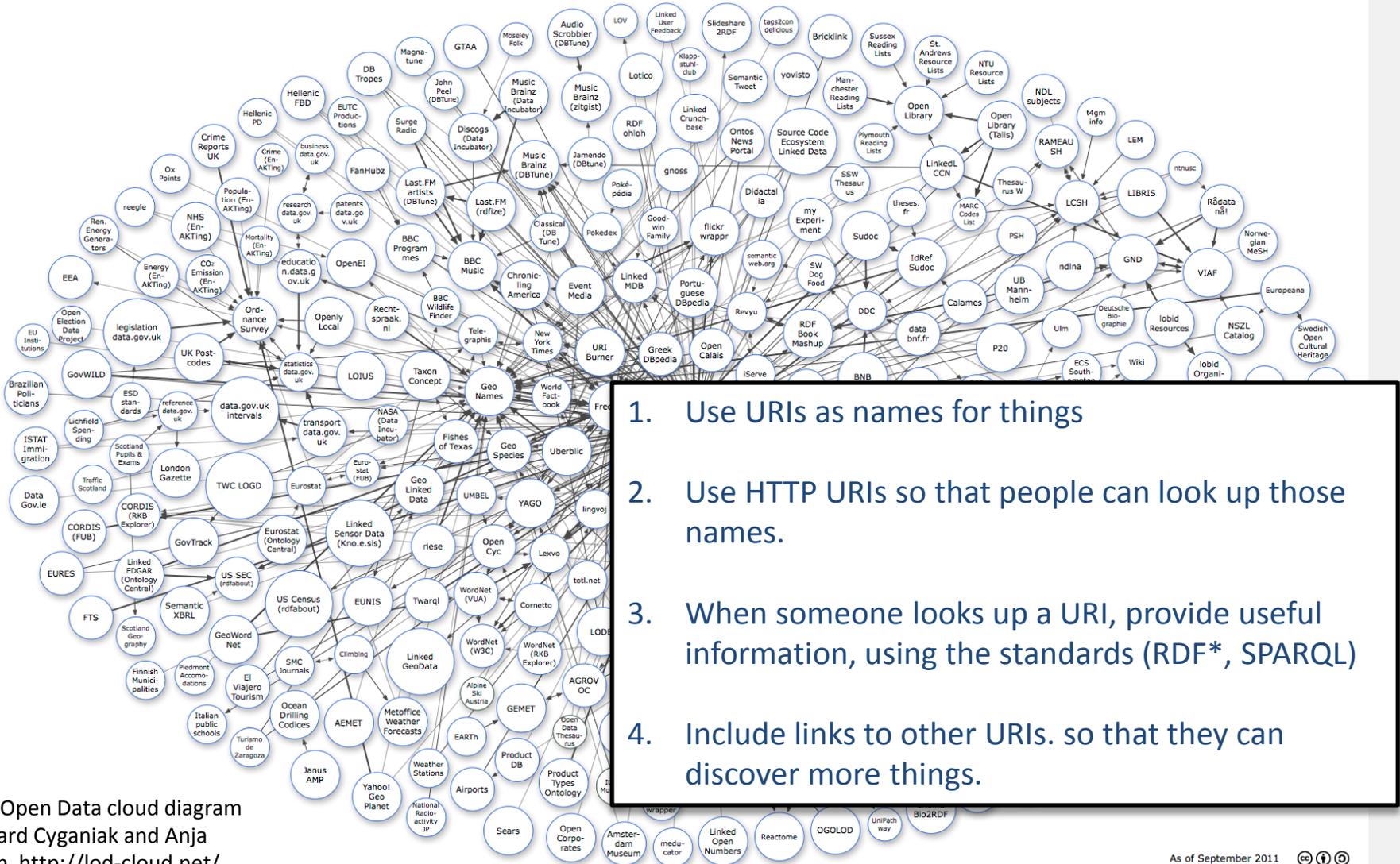
“A digital library that is a single, direct and multi-lingual access point to the European cultural heritage.”

Europäisches Parlament
27. September 2007

Das Europeana Ökosystem



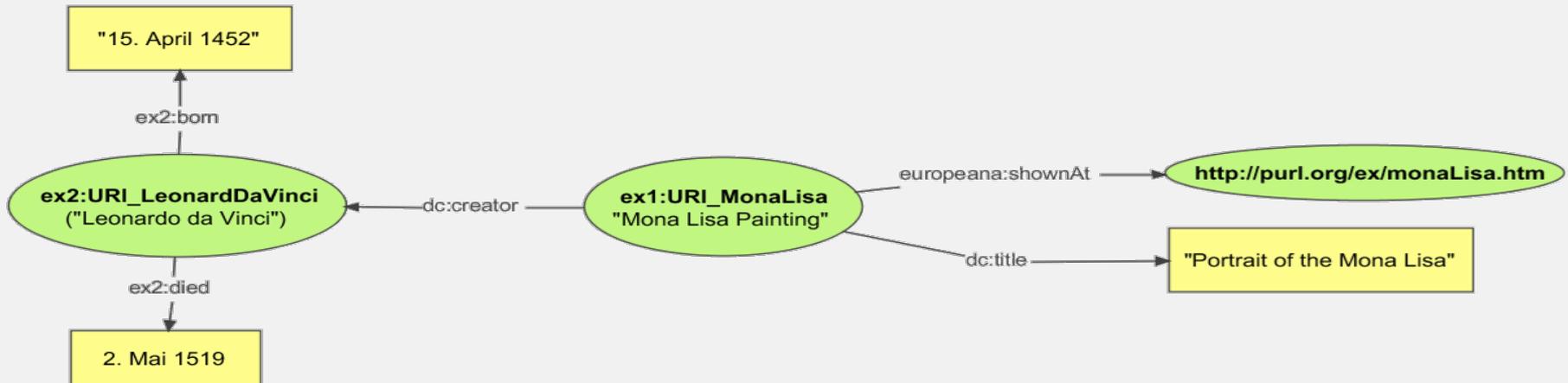
Strategische oder
Technologiepartner,
Datenlieferanten &
Aggregatoren



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.

Linking Open Data cloud diagram by Richard Cyganiak and Anja Jentzsch. <http://lod-cloud.net/>

- Resource Description Framework (RDF)
 - Modellierung von Informationen als Statements in Form von Tripeln (Subjekt – Prädikat – Objekt)
 - Ressourcen über URIs eindeutig referenzierbar
 - Verlinkungen auf Statements anderer



Das EDM und seine Spezialisierungen

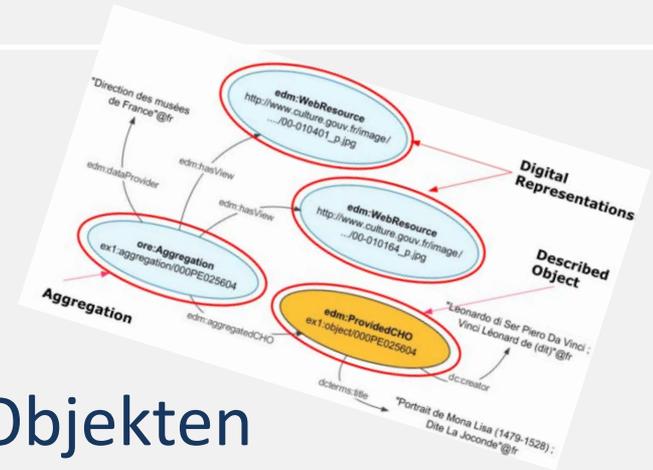
Das EDM (Europeana Data Model)...

... vereint diverse Standards und Vokabulare

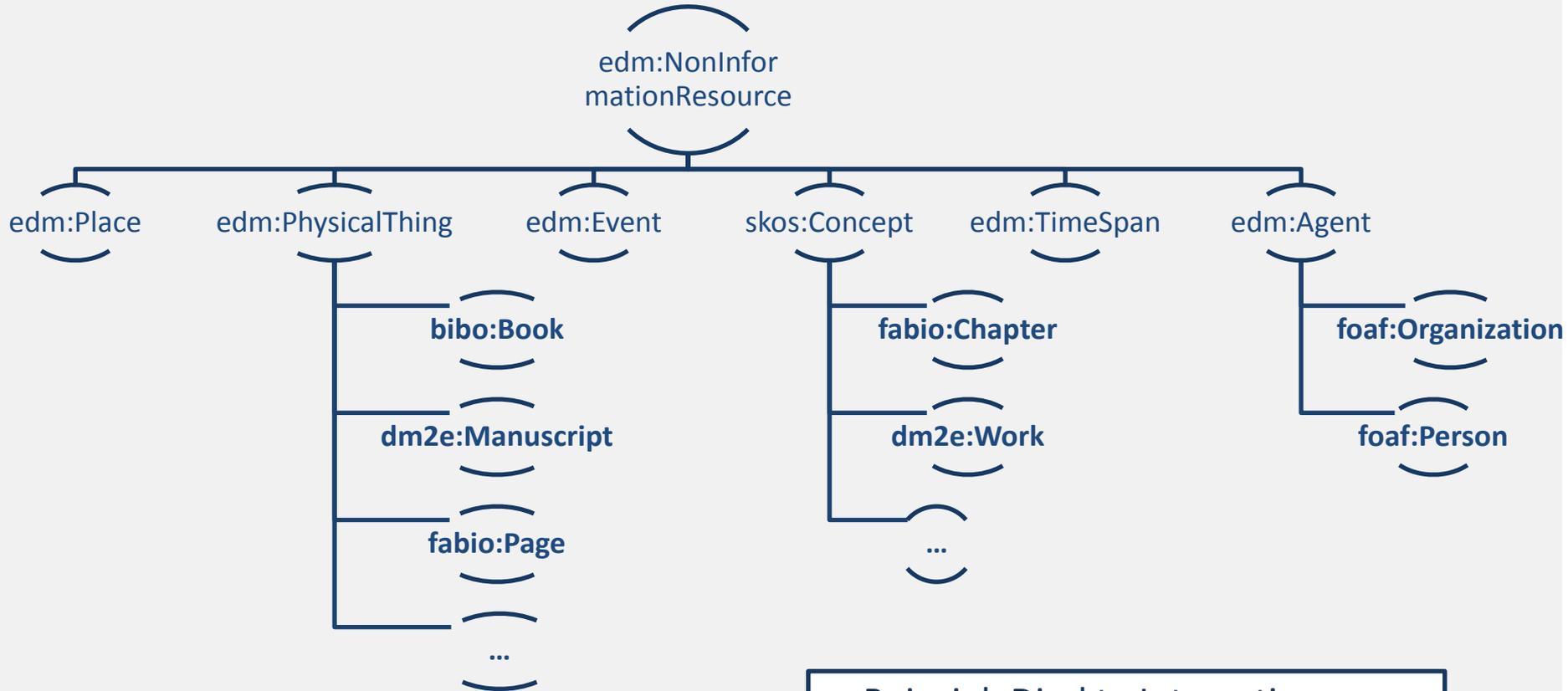
... ermöglicht die Repräsentation von Objekten aus Büchereien, Archiven und Museen

... ist so generisch wie möglich gehalten

... kann für unterschiedliche Domänen spezialisiert werden



DM2E-Modell: Klassenspezialisierungen



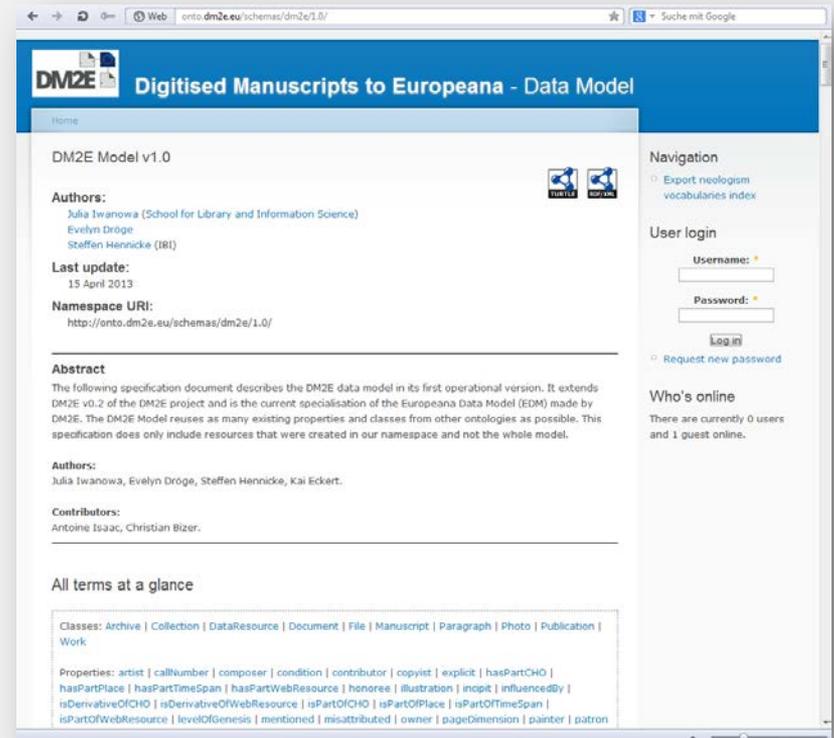
Beispiel: Direkte Integration von Klassen in den EDM-Aufbau

Das DM2E-Modell

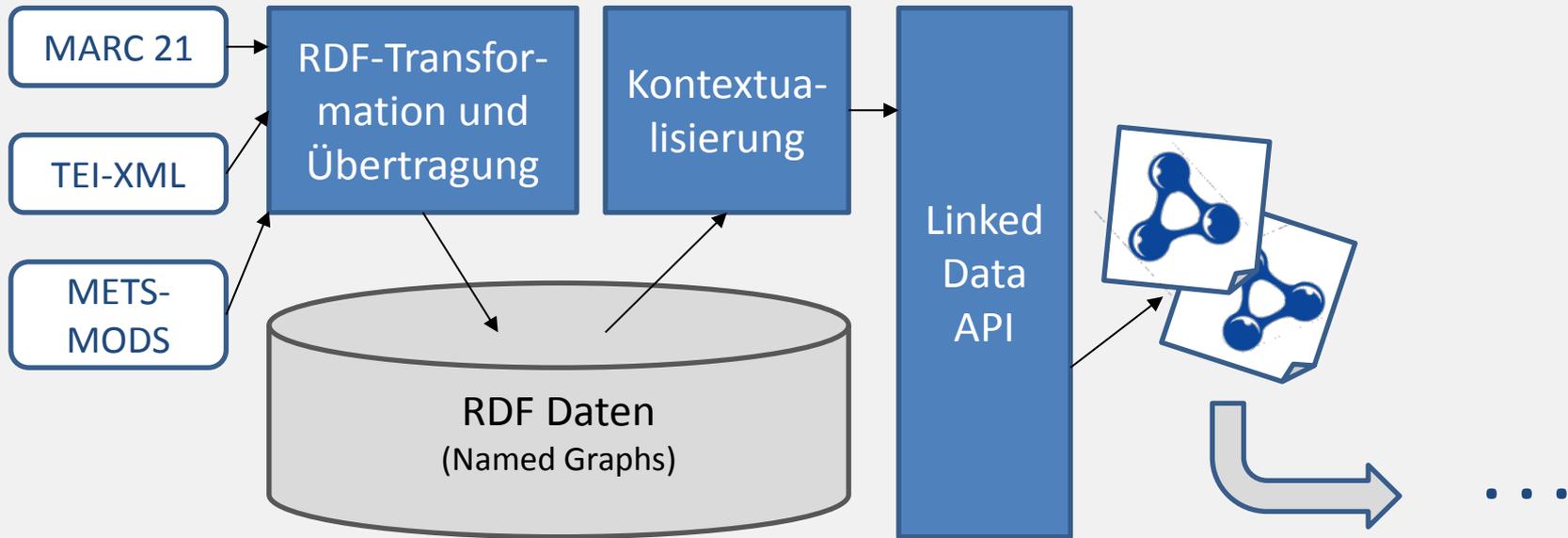
- Namespaces:
 - dm2e: `<http://onto.dm2e.eu/schemas/dm2e/1.0/>` .
 - dm2edata: `<http://data.dm2e.eu/data/>` .
- Nachnutzung externer Ressourcen:
 - Bibliographic Ontology (Bibo), FRBR-aligned Bibliographic Ontology (FaBiO), Publishing Roles Ontology, VIVO Ontology, VoID
- Wie werden sie nachgenutzt?
 - Neue Property: dm2e:scopeNote
 - Originaldefinitionen der nachgenutzten Terme
- Modellierung analog zu EDM (OAI-ORE)
 - Umfangreiche Verwendung von Properties statt Klassen
Beispiel: 52 neue Properties für `edm:ProvidedCHO`
 - Nach Möglichkeit direkte Nachnutzung externer Ressourcen

DM2E-Modell: Eine Spezialisierung des EDM für Manuskripte.

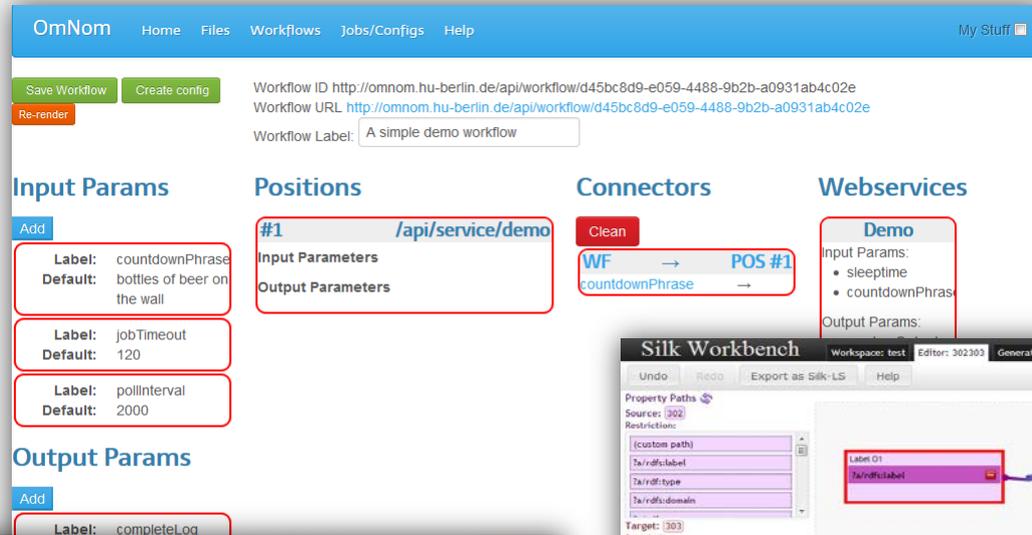
- Stabile Version: 1.0
 - Online: onto.dm2e.eu/dm2e
- Modelldokumentation:
<http://dm2e.eu/document/>
 - Modellbeschreibung (pdf)
 - OWL-Datei
- Weiterentwicklungen in Modellversion 1.1



Aufbau der Mapping-Infrastruktur



Der Weg nach Europeana über OmNom



OmNom Home Files Workflows Jobs/Configs Help My Stuff

Save Workflow Create config Re-render

Workflow ID <http://omnom.hu-berlin.de/api/workflow/d45bc8d9-e059-4488-9b2b-a0931ab4c02e>
Workflow URL <http://omnom.hu-berlin.de/api/workflow/d45bc8d9-e059-4488-9b2b-a0931ab4c02e>
Workflow Label: A simple demo workflow

Input Params

Add

- Label: countdownPhrase
Default: bottles of beer on the wall
- Label: jobTimeout
Default: 120
- Label: pollInterval
Default: 2000

Output Params

Add

- Label: completeLog

Positions

#1 /api/service/demo

Input Parameters
Output Parameters

Connectors

Clean

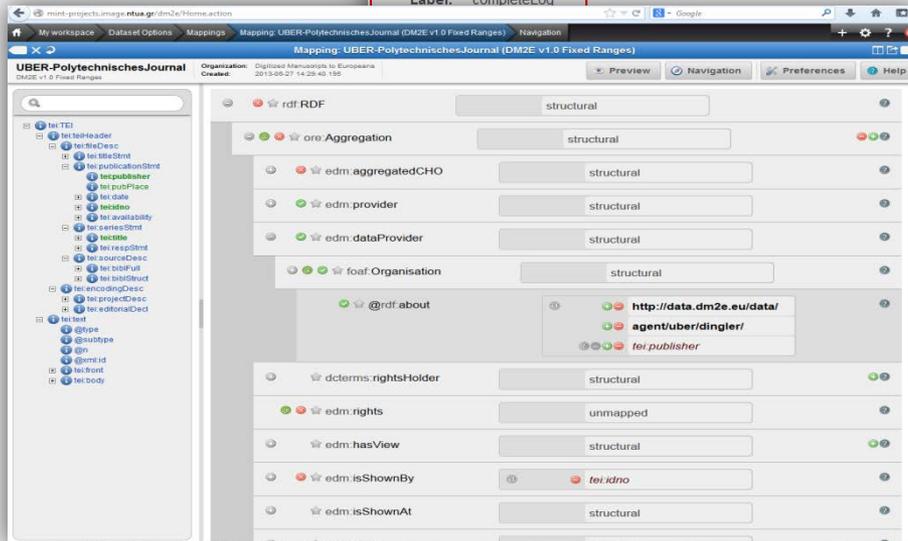
WF → POS #1
countdownPhrase →

Webservices

Demo

Input Params:
• sleeptime
• countdownPhras

Output Params:



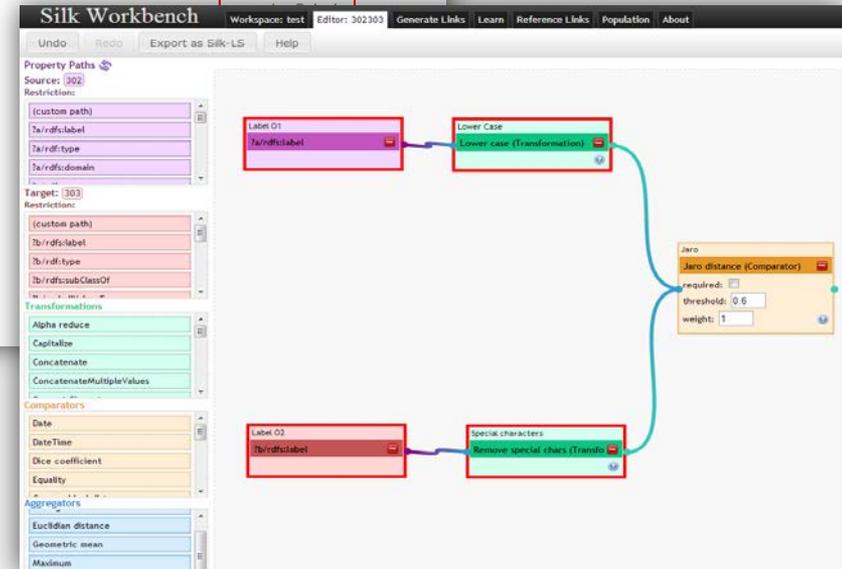
my-workspace Datasets Options Mappings Mapping: UBER-PolytechnischesJournal (DM2E v1.0 Fixed Ranges) Navigation

UBER-PolytechnischesJournal Organization: Digitalisat-Museum für Europeana Created: 2013-05-27 14:29:40 155

rdf RDF structural

oro Aggregation structural

- edm aggregatedCHO structural
- edm provider structural
- edm dataProvider structural
- foaf Organisation structural
- http://data.dm2e.eu/data/ agent/uber/dingler/ tei publisher
- dcterms:rightsHolder structural
- edm:rights unmapped
- edm:hasView structural
- edm:isShownBy tei:tdno structural
- edm:isShownAt structural



Silk Workbench Workspace: test Editor: 302303 Generate Links Learn Reference Links Population About

Undo Redo Export as Silk-LS Help

Property Paths

Source: [302] Restriction: [(custom path) %r/rdf:label]

Target: [303] Restriction: [(custom path) %b/rdf:label]

Transformations

- Alpha reduce
- Capitalize
- Concatenate
- ConcatenateMultipleValues

Comparators

- Date
- Date Time
- Dice coefficient
- Equality

Aggregators

- Euclidian distance
- Geometric mean
- Maximum

Workflow diagram showing transformations: Label 01 (%r/rdf:label) → Lower Case (Lower case (Transformation)) → Jaro distance (Comparator) → Label 02 (%b/rdf:label) → Special characters (Remove special chars (Transformation))

Das Ergebnis: Linked Data

Ms-114_17r

URI of this Resource Map: http://lelystad.informatik.uni-mannheim.de:8080/data/resourcemap/oib/wittgenstein-repo/Ms-114_17r

Ms-114_17r

URI: http://lelystad.informatik.uni-mannheim.de:8080/data/resource/oib/wittgenstein-repo/Ms-114_17r

Property	Value
is http://www.europeana.eu/schemas/edm/aggregatedCHO of	<ul style="list-style-type: none"><http://lelystad.informatik.uni-mannheim.de:8080/data/aggregation/oib/wittgenstein-repo/Ms-114_17r>
Description	<ul style="list-style-type: none">Page pagename_Ms-114_17r pageref_Ms-114_37 from Wittgenstein Nachlass MS 114: Ms-114.xml - A machine-readable transcription [Work in progress]
Identifier	<ul style="list-style-type: none">Ms-114_17r
is http://www.europeana.eu/schemas/edm/isNextInSequence of	<ul style="list-style-type: none"><http://lelystad.informatik.uni-mannheim.de:8080/data/resource/oib/wittgenstein-repo/Ms-114_16v>
http://www.europeana.eu/schemas/edm/isNextInSequence	<ul style="list-style-type: none"><http://lelystad.informatik.uni-mannheim.de:8080/data/resource/oib/wittgenstein-repo/Ms-114_17v>
Is Part Of	<ul style="list-style-type: none"><http://lelystad.informatik.uni-mannheim.de:8080/data/resource/oib/wittgenstein-repo/Ms-114>
Language	<ul style="list-style-type: none">ger
Same As	<ul style="list-style-type: none"><http://dm2e.eu/resource/oib/wittgenstein-repo/Ms-114_17r>
http://www.europeana.eu/schemas/edm/type	<ul style="list-style-type: none">TEXT
Type	<ul style="list-style-type: none">TEXT
Type	<ul style="list-style-type: none">http://www.europeana.eu/schemas/edm/ProvidedCHO

Ms-114_17r

URI: http://lelystad.informatik.uni-mannheim.de:8080/data/aggregation/oib/wittgenstein-repo/Ms-114_17r

Property	Value
http://www.europeana.eu/schemas/edm/aggregatedCHO	<ul style="list-style-type: none"><http://lelystad.informatik.uni-mannheim.de:8080/data/resource/oib/wittgenstein-repo/Ms-114_17r>
http://www.europeana.eu/schemas/edm/dataProvider	<ul style="list-style-type: none">Wittgenstein Archives at the University of Bergen (WAB)
http://www.europeana.eu/schemas/edm/hasView	<ul style="list-style-type: none"><http://wab.uib.no/cost-a32_fax/114/Ms-114_17r.jpg>
http://www.europeana.eu/schemas/edm/isShownAt	<ul style="list-style-type: none"><http://wab.uib.no/cost-a32_fax/114/Ms-114_17r.jpg>
http://www.europeana.eu/schemas/edm/provider	<ul style="list-style-type: none">Wittgenstein Archives at the University of Bergen (WAB)
Rights	<ul style="list-style-type: none">Copyright holders: The Master and Fellows of Trinity College, Cambridge; Oxford University Press, Oxford; University of Bergen, Bergen. Released under the Creative Commons General Public License Attribution, Non-Commercial, Share-Alike version 3 (CCPL BY-NC-SA).
http://www.europeana.eu/schemas/edm/rights	<ul style="list-style-type: none"><http://creativecommons.org/licenses/by-nc/3.0/>
Same As	<ul style="list-style-type: none"><http://dm2e.eu/aggregation/oib/wittgenstein-repo/Ms-114_17r>
Type	<ul style="list-style-type: none">Aggregation



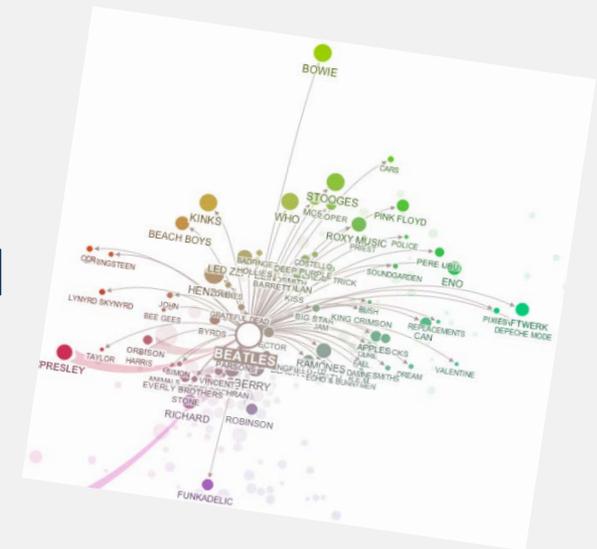
This page shows information obtained from the SPARQL endpoint at <http://lelystad.informatik.uni-mannheim.de:8080/openrdf-sesame/repositories/dm2etest>.

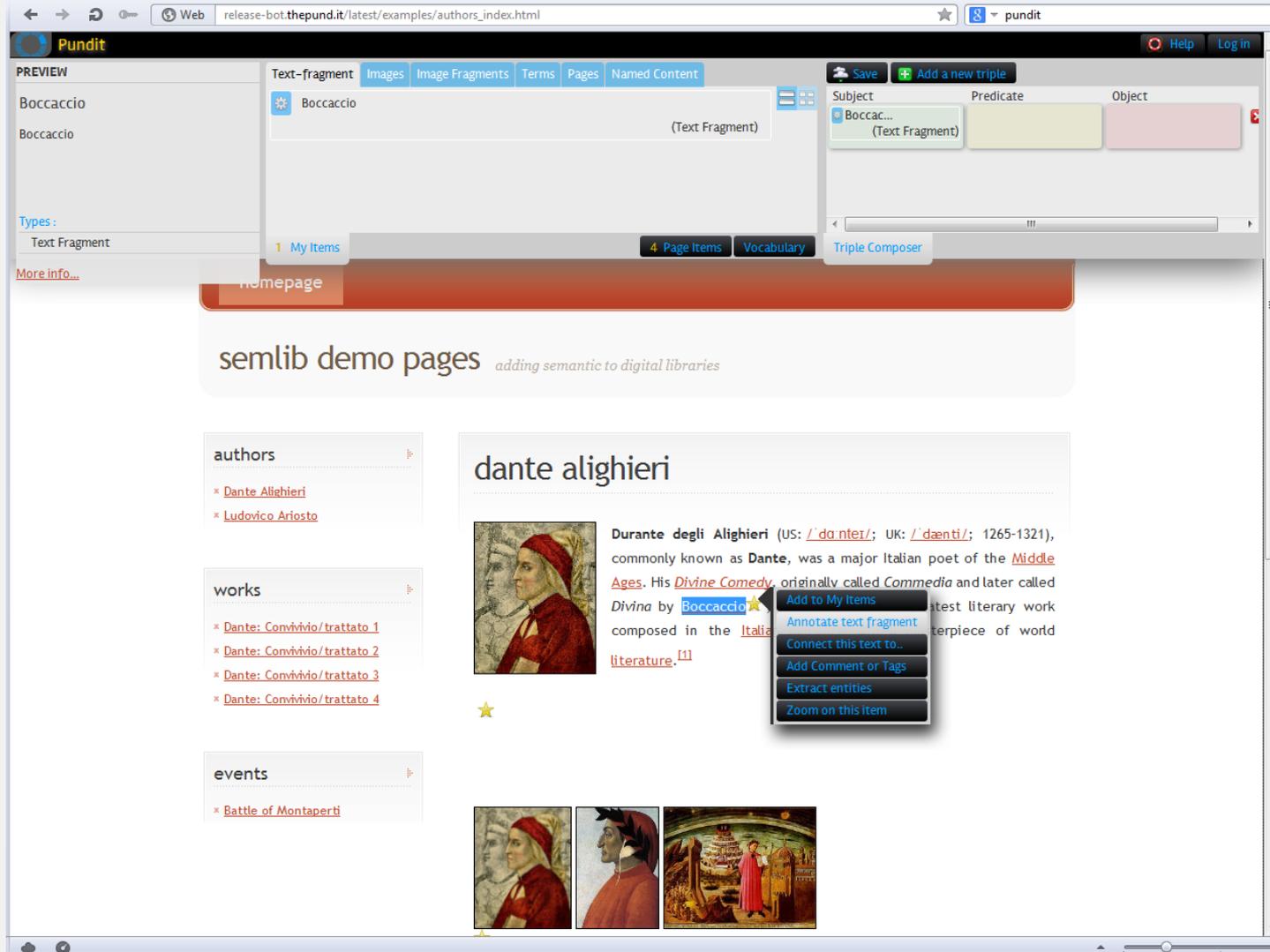
 As Turtle |  As RDF/XML |  Browse in Disco |  Browse in Graphite Browser

Der nächste Schritt: Pundit

Pundit: Ein web-basiertes Annotationssystem
<http://www.thepund.it>

- Annotationen von Websites oder Bildern
 - Einteilung in nutzerspezifische Fragmente
 - Annotationen in Form von RDF-Tripeln
- Ergänzungen über andere Vokabulare
- Erstellung von Links zu anderen Dokumenten
- Bereitstellung der RDF-Daten über API
- Visualisierung der RDF-Daten über Edgemaps





The screenshot shows the Pundit web interface. At the top, a browser window displays the URL `release-bot.thepund.it/latest/examples/authors_index.html`. The Pundit interface includes a navigation menu with options like 'Text-fragment', 'Images', and 'Image Fragments'. A 'Triple Composer' is visible, showing a triple with 'Boccaccio (Text Fragment)' as the subject. The main content area features a 'semilib demo pages' header and a profile for 'dante alighieri'. A contextual menu is open over the name 'Boccaccio' in the text, offering actions such as 'Add to My Items', 'Annotate text fragment', and 'Connect this text to...'. The profile text describes Dante Alighieri and his work, mentioning 'Boccaccio' as the author of a text fragment.



- DM2E läuft bis Januar 2015
- Was geschieht im Moment?
 - Mapping der Providerdaten
 - Fertigstellung von DM2E-Modell Version 1.1
 - Nutzertests: OmNom und Pundit
 - Datenkontextualisierung



Vielen Dank für Ihre Aufmerksamkeit!

Evelyn Dröge

Institut für Bibliotheks- und
Informationswissenschaft
Humboldt-Universität zu Berlin
www.ibi.hu-berlin.de

Digitised Manuscripts to Europeana
www.dm2e.eu

evelyn.droege@ibi.hu-berlin.de



Literatur

- Berners-Lee, T. (2006). Linked Data - Design Issues. W3C Website. <http://www.w3.org/DesignIssues/LinkedData>. [13.10.2013]
- Heath, T., & Bizer, C. (2011). Linked Data: Evolving the Web into a Global Data Space. Synthesis Lectures on the Semantic Web: Theory and Technology (Bd. 1). Morgan & Claypool.

Bilder

- Oxygen480-actions-mail-mark-task.svg (Folie 2): <http://commons.wikimedia.org/wiki/File:Oxygen480-actions-mail-mark-task.svg>. Urheber: Nuno Pinheiro et al.
- Projektaufbau (Folie 3): Kai Eckert.
- Europeana Screenshot und Logos (Folien 4, 5 & 8): europeana.eu und Projektseiten.
- Linking Open Data cloud diagram (Folie 6): Richard Cyganiak and Anja Jentzsch. <http://lod-cloud.net/>.
- RDF-Graph (Folie 7): Sam Leon (erstellt mit VUE).
- DM2E-Tools, Logos & Screenshots (Folien 13, 14, 16): Konstantin Baierer, Kai Eckert, Sam Leon, Christian Morbidoni.
- Edgemaps Screenshot (Folie 15): <http://mariandoerk.de/edgemaps/>. Urheber: Marian Dörk.
- Footstep (Folie 17): <http://openclipart.org/detail/left-footprint-by-anonymous>.
- IBI (Folie 18): http://commons.wikimedia.org/wiki/File:Berlin,_Mitte,_Dorotheenstrasse,_Handelskammer_Berlin_02.jpg. Urheber: [Beek100](#).