

Specialising the EDM for Digitised Manuscripts

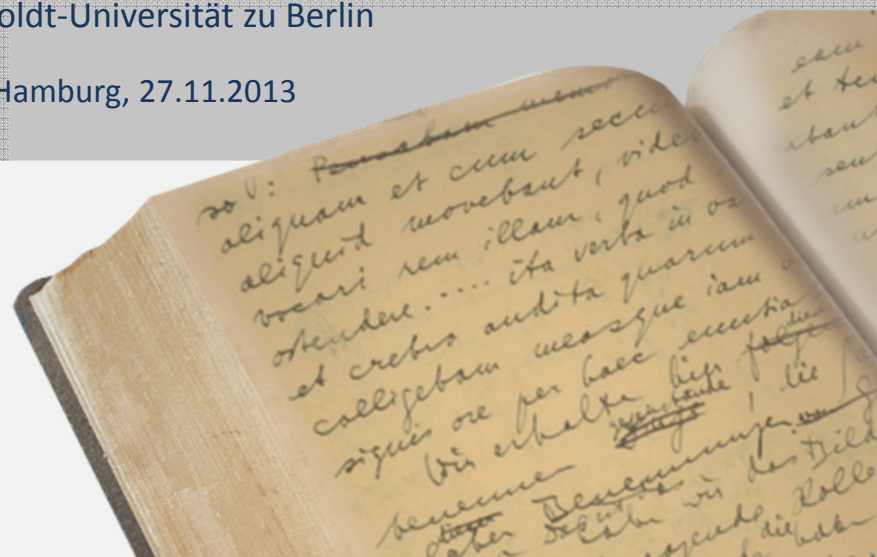
Kai Eckert¹, Steffen Hennicke², Evelyn Dröge², Julia Iwanowa², Violeta Trkulja²

¹Universität Mannheim, ²Humboldt-Universität zu Berlin

Semantic Web in Libraries - Hamburg, 27.11.2013



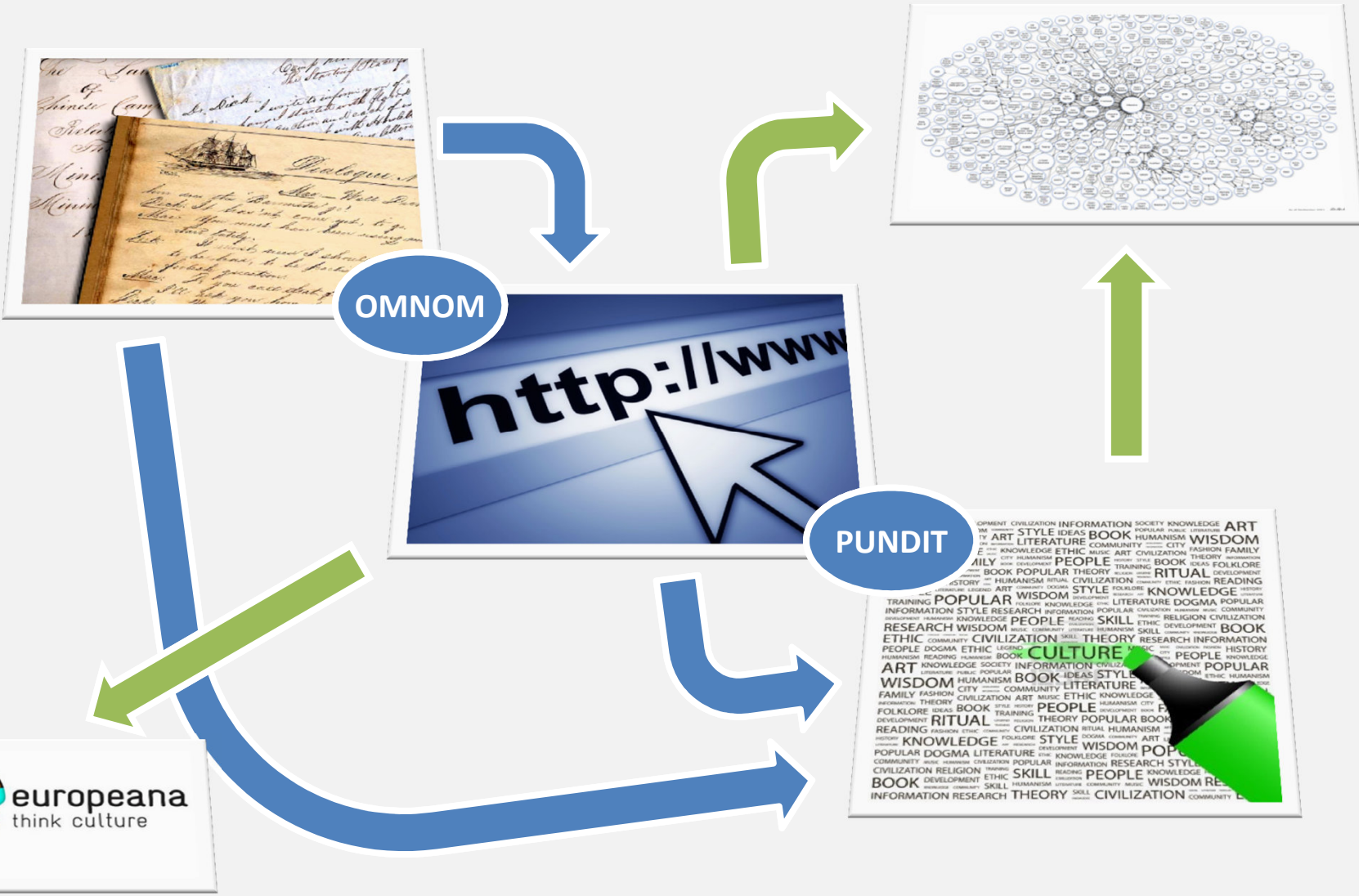
co-funded by the European Union



Digitised Manuscripts to Europeana

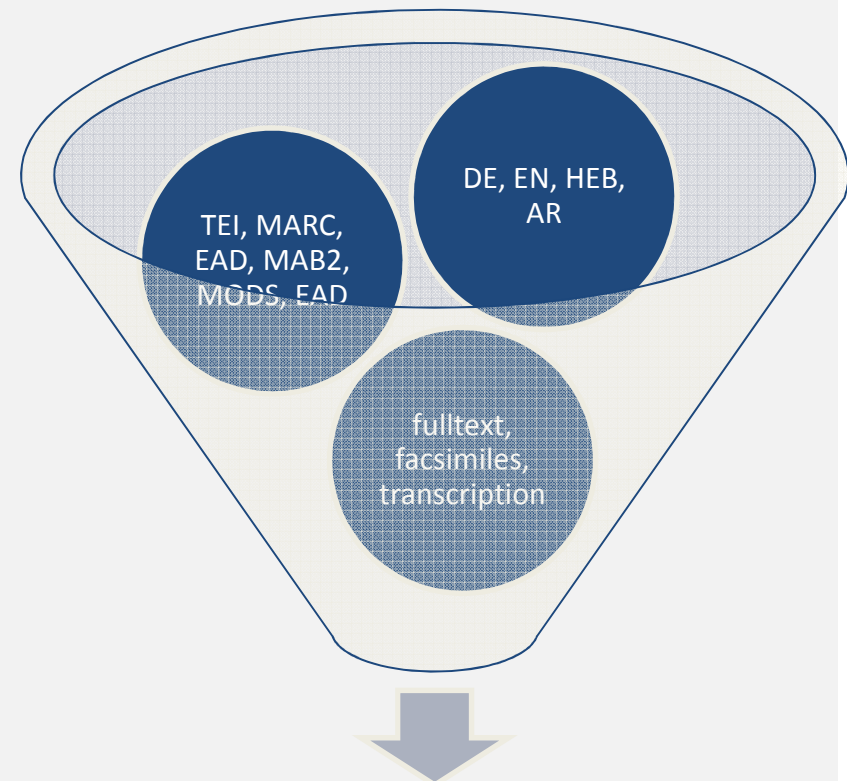
- EU-funded Europeana satellite project
- Duration: Three years (2012 – 2015)
- Partners from Germany, Austria, Norway, Greece, UK and Italy
- DM2E works on:
 - a tool-chain for **data migration** to Europeana and the LOD Web (**OMNOM**),
 - a digital **research environment** for the Digital Humanities (**PUNDIT**),
 - an **open community** of cultural heritage professionals (**OPENGLAM**)

DM2E: Infrastructure



DM2E: Provided Content

- Metadata about manuscripts:
 - Described by: TEI, MAB2, MARC, EAD, METS/MODS Database content
 - In different languages
 - 118.000+ items
 - 20.006.930+ pages



DM2E Model



DM2E: Data Model

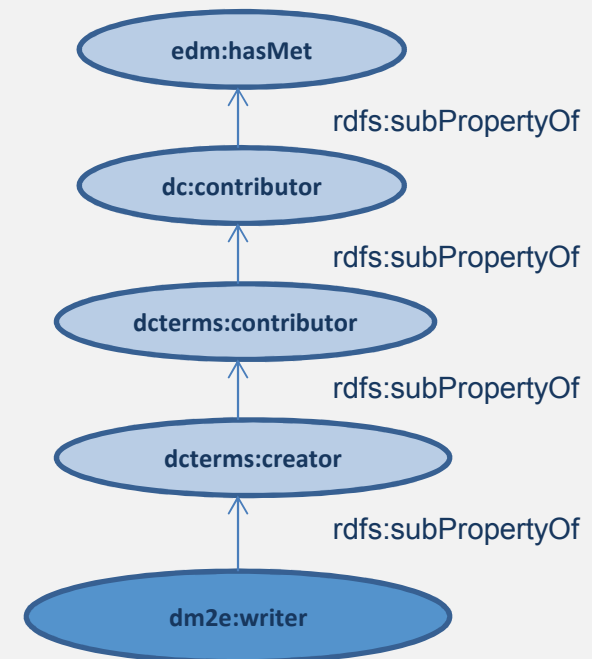
- Semantically and structurally **heterogeneous** data
 - e.g. EAD, METS, TEI, MARCXML and MAB2, relational databases, proprietary schemas

- The **Europeana Data Model (EDM)** is made for this scenario!
 - provides a generic semantic interoperability layer
 - enables the definition of “**applications profiles**” which may address the needs of specific communities

- The **DM2E Data Model (DM2E)**
 - is an “application profile” of the EDM for the domain of handwritten manuscripts
 - retains rich descriptions by specialising the EDM

DM2E: Specialisation approach

- RDF(S) allows the specialisation of EDM classes and properties
 - use of `rdfs:subClassOf`
 - use of `rdfs:subPropertyOf`
- An “application profile” typically also includes
 - additional ontological restrictions
 - documentation



DM2E: Specialisation Guidelines

- Empirical analysis of provided source metadata
- Iterative mappings to the EDM
- Close cooperation with data providers
 - agree on shared conceptualisations
- Create rich and connected representations
 - retain original semantics as much as possible
 - use existing URIs of resources
 - assign a class to the resources (rdf:type)

DM2E: Interoperability approach

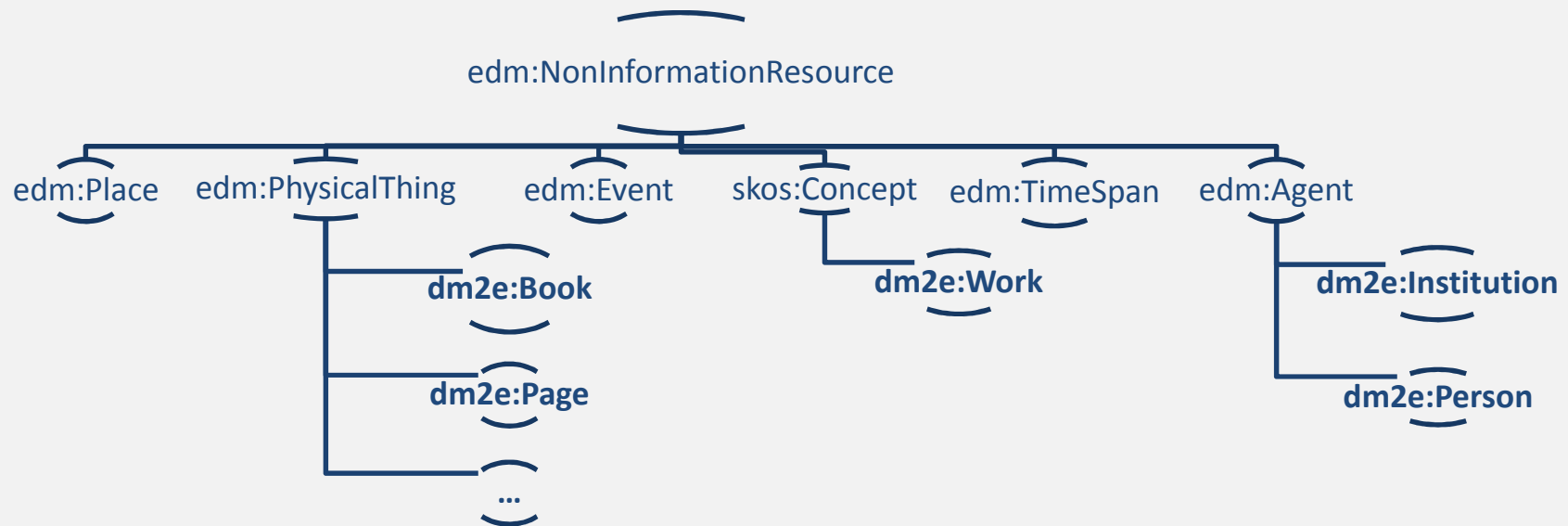
- Create new classes or properties in the **DM2E-Namespace** only if there is no other suitable option available
 - reuse existing namespaces (ontologies)
 - mind existing semantics (scope notes, domains, ranges)

- Types, roles and relations between **agents**
 - Friend-of-a-Friend (**FOAF**) [**FOAF**] (**types of agents**)
 - Publishing Roles Ontology (**PRO**) [**SPAR**] (**roles of agents in the publication process**)
 - **VIVO** [**VIVO**] (**types of agents**)

- Detailed semantics on **bibliographic entities**
 - FRBR-aligned Bibliographic Ontology (**FaBiO**) [**SPAR**]
 - Citation Typing Ontology (**CiTO**) [**SPAR**]
 - Bibliographic Ontology (**BIBO**) [**BIBO**]

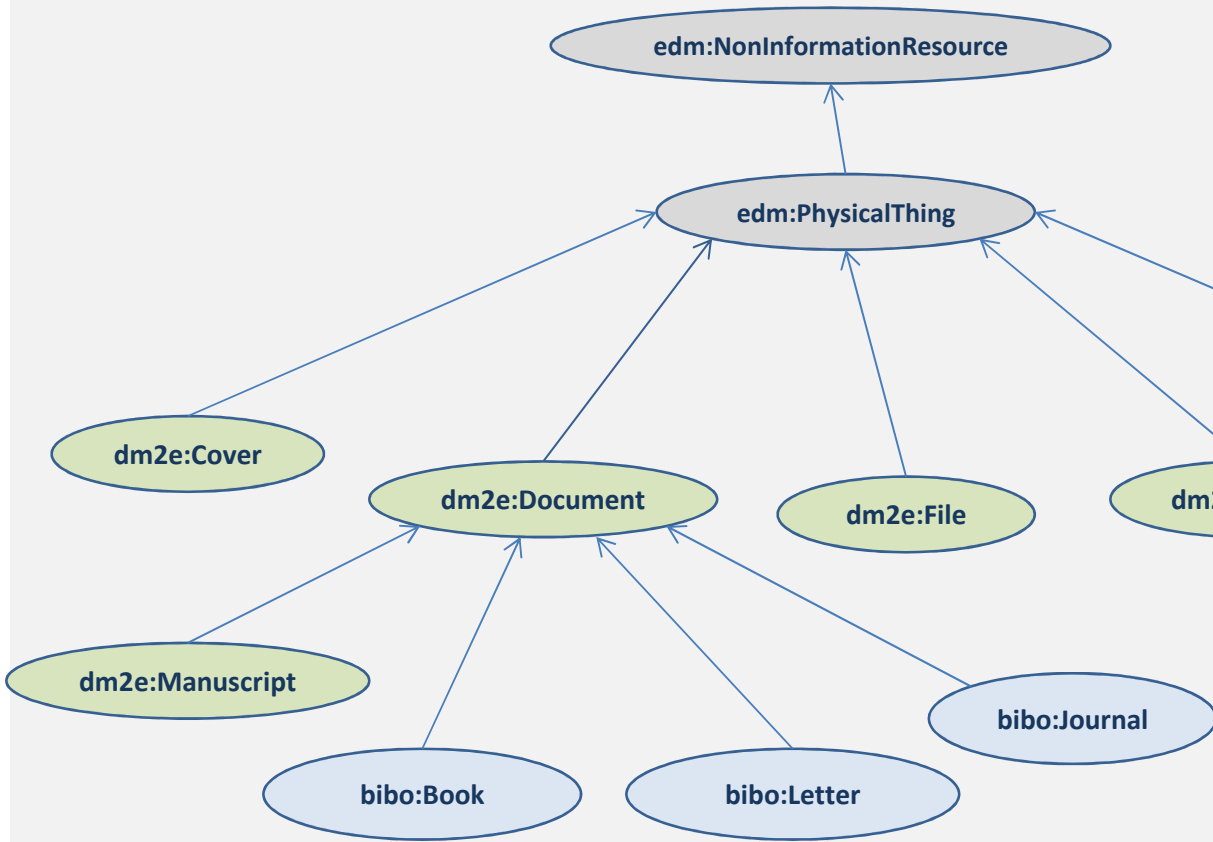
DM2E Model: Class-Specialisation

- 23 new or reused classes, mainly for
 - physical and conceptual parts of a handwritten manuscripts
 - as found in our source metadata
 - different types of Agents



edm:PhysicalThing

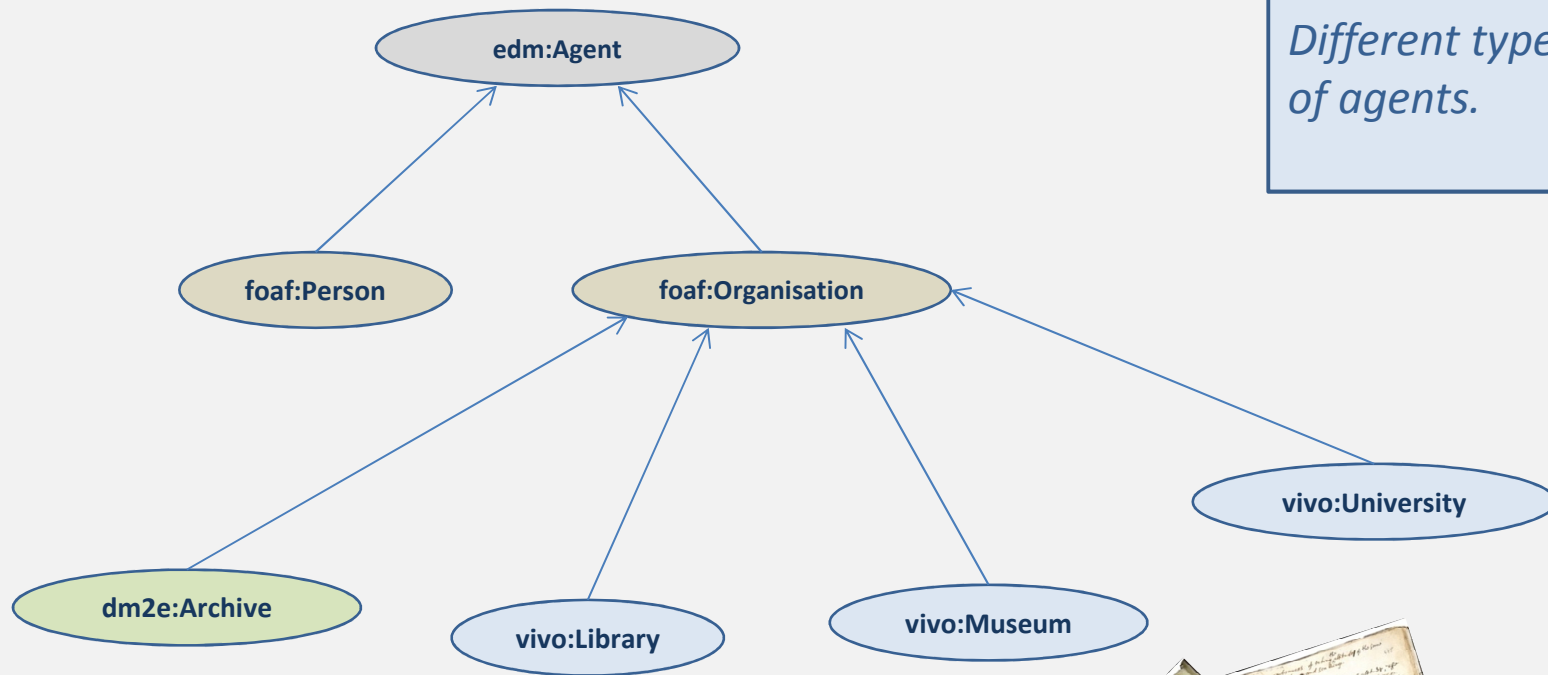
Physical and tangible aspects of handwritten manuscripts.



- is-a
- <http://www.europeana.eu/schemas/edm/>
- <http://onto.dm2e.eu/schemas/dm2e/1.0/>
- <http://purl.org/ontology/bibo/>



Contextual Resources: Agent



Different types of agents.

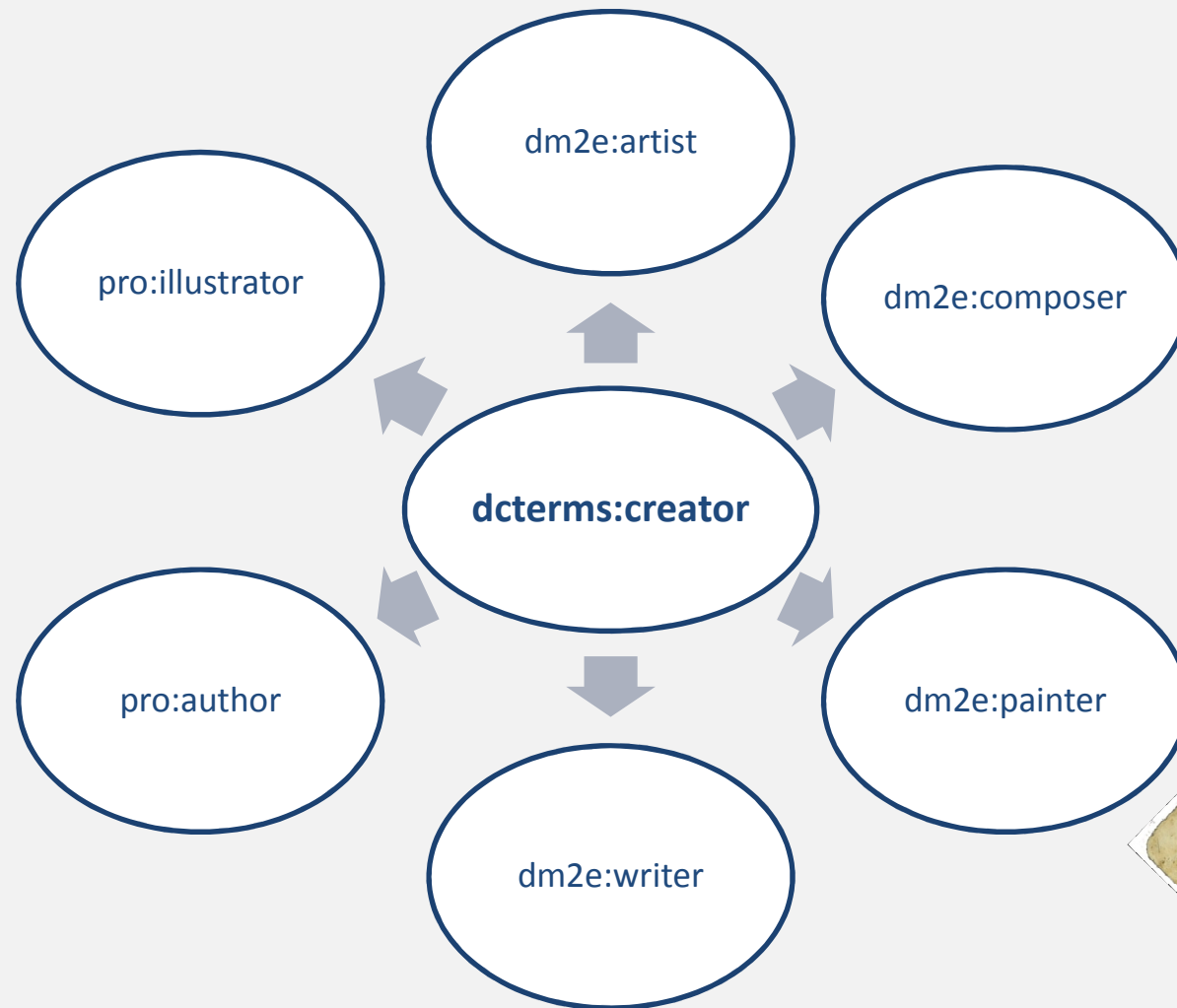


- is-a
- <http://www.europeana.eu/schemas/edm/>
- <http://xmlns.com/foaf/0.1/>
- <http://onto.dm2e.eu/schemas/dm2e/1.0/>
- <http://vivoweb.org/ontology/core#>

DM2E Model: Properties-Specialisation

- **Property-centric** modelling
 - more than 50 new properties
- Documentation for the DM2E Data Model contains only EDM properties which are utilized
 - to keep the documentation clear
 - e.g. dcterms:replaces, dc:source, or dc:conformsTo are not used
- **Domain and Range Restrictions**
 - some OWL-Restrictions on properties in order to encourage the use of specific resources of a specific type, e.g.
 - CHO hasPart CHO
 - WebResource hasPart WebResource
- Some EDM-Properties are **mandatory** in DM2E
 - dc:type: at least one of the physical (e.g. dm2e:Page) or logical (e.g. dm2e:Paragraph) aspects
 - dc:subject: ideally an URI from a controlled vocabulary

DM2E Model: Property Extensions



Example: Adding new properties as subproperties for `dcterms:creator`



Outlook: Uncertain Statements

Part of the next model version: How to deal with uncertain timespans and presumably creators?

- Problem: Confidence declarations for RDF-statements need Named Graphs or Reification
- Solution:

Agents

„The creator of the CHO is presumably Goethe.“

```
res1 dc:creator presumableAgent1.

presumableAgent1 a PresumableAgent;
  isPresumably goethe;
  confidence 0.8.
```

Confidence is optional

Timespans

„The timespan was somewhere in the 1920ies and lasted 2 years.“

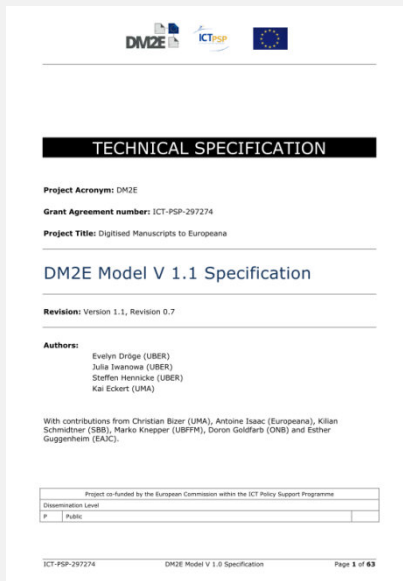
```
timeSpan1 a edm:TimeSpan.
  uncertainBegin 1920;
  uncertainEnd 1929;
  duration 2.
```

Duration is optional

Documentation: PDF and OWL

The PDF and the OWL representations can be accessed via the project's website:

dm2e.eu/document/#DM2EModelSpecification

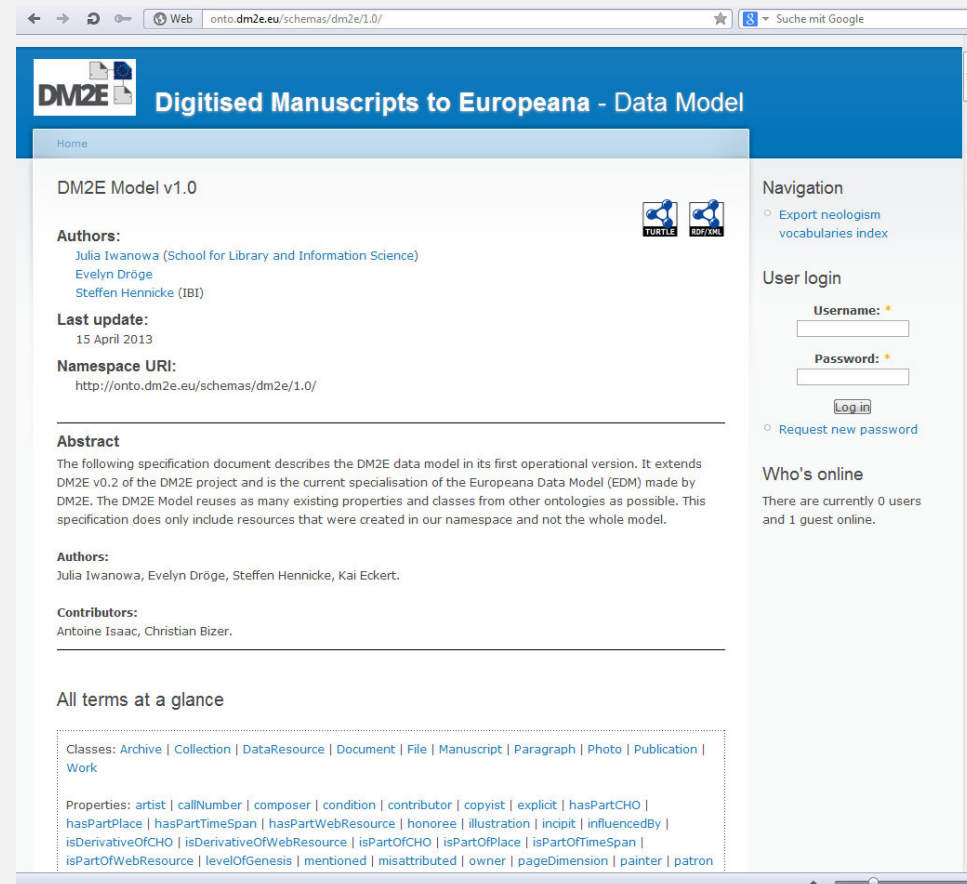


```
<?xml version="1.0"?>
<rdf:RDF
  xmlns:rdf="http://protege.stanford.edu/plugins/owl/protege#"
  xmlns:via="http://via.web.org/ontology/core#"
  xmlns:pr="http://pr1.org/par/par/"
  xmlns:sc="http://www.w3.org/2003/11/owl#"
  xmlns:dc="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:libo="http://pr1.org/ontology/libo/"
  xmlns:foaf="http://xmlns.com/foaf/0.1/"
  xmlns:docname="http://pr1.org/doc/parma/"
  xmlns:scgs="http://www.w3.org/2001/XMLSchema#"
  xmlns:DOICCE-Lite="http://www.lea-cmr.it/ontology/DOICCE-Lite.owl#"
  xmlns:orm="http://www.cidoc-crm.org/crm/crm_v5.0.2_english_label.rdf#"
  xmlns:wgsn4_pos="http://www.w3.org/2003/01/scn/wgsn4_pos#"
  xmlns:dm2e="http://onto.dm2e.eu/schemas/dm2e/1.0/"
  xmlns:sg="http://www.w3.org/2004/02/sga/core#"
  xmlns:xsp="http://www.owl-ontologies.com/2005/08/07/xsp.owl#"
  xmlns:edm="http://www.europeana.eu/schemas/edm/"
  xmlns:kocho="http://pr1.org/net7/kocho/vocab#"
  xmlns:do="http://pr1.org/doc/elements/1.1/"
  xmlns:voiid="http://dfe.org/na/voiid#"
  xmlns:rda0r2="http://sdvocab.info/elements/2/"
  xmlns=""http://onto.dm2e.eu/schemas/dm2e/1.0/DM2E_v1.0.owl#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:ore="http://www.openarchives.org/ore/terms/"
  xmlns:sgul="http://pr1.stanford.edu/ontology/builtins/3.4/sgul.owl#"
  xmlns:fabio="http://pr1.org/par/fabio/"
  xmlns:svt="http://www.w3.org/2003/11/par#"
  xmlns:scs="http://www.w3.org/2001/04/scs-schema#"
  xmlns:svle="http://pr1.stanford.edu/ontology/3.3/svle.owl#"
  xml:base="http://onto.dm2e.eu/schemas/dm2e/1.0/DM2E_v1.0.owl">
  <owl:Ontology rdf:about="">
    <rdfa:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
      This is the DM2E Model V1.0. It is the first operational version, it extends DM2E v0.2 of the DM2E
      project and is the current specialization of the Europeana Data Model (EDM) made by DM2E. The DM2E
      Model reuses as many existing properties and classes from other ontologies as possible. This
      specification does only include resources that were created in our namespace and not the whole model.
    </rdfa:comment>
    Authors:
      Julia Iwanowa, Evelyn Dröge, Steffen Hennicke, Kai Eckert.
    Contributors:
      Antoine Isaac, Christian Bizer. </rdfa:comment>
    <owl:versionInfo rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
      >DM2E v1.0</owl:versionInfo>
    </owl:Ontology>
    <owl:Class rdf:about="http://onto.dm2e.eu/schemas/dm2e/1.0/DataResource">
```

Documentation: Online

- Human & machine readable
- Version 1.0

onto.dm2e.eu



DM2E Digitised Manuscripts to Europeana - Data Model

Home

DM2E Model v1.0

Authors:
 Julia Iwanowa (School for Library and Information Science)
 Evelyn Dröge
 Steffen Hennicke (IBI)

Last update:
 15 April 2013

Namespace URI:
<http://onto.dm2e.eu/schemas/dm2e/1.0/>

Abstract
 The following specification document describes the DM2E data model in its first operational version. It extends DM2E v0.2 of the DM2E project and is the current specialisation of the Europeana Data Model (EDM) made by DM2E. The DM2E Model reuses as many existing properties and classes from other ontologies as possible. This specification does only include resources that were created in our namespace and not the whole model.

Authors:
 Julia Iwanowa, Evelyn Dröge, Steffen Hennicke, Kai Eckert.

Contributors:
 Antoine Isaac, Christian Bizer.

All terms at a glance

Classes: [Archive](#) | [Collection](#) | [DataResource](#) | [Document](#) | [File](#) | [Manuscript](#) | [Paragraph](#) | [Photo](#) | [Publication](#) | [Work](#)

Properties: [artist](#) | [callNumber](#) | [composer](#) | [condition](#) | [contributor](#) | [copyist](#) | [explicit](#) | [hasPartCHO](#) | [hasPartPlace](#) | [hasPartTimeSpan](#) | [hasPartWebResource](#) | [honoree](#) | [illustration](#) | [incipit](#) | [influencedBy](#) | [isDerivativeOfCHO](#) | [isDerivativeOfWebResource](#) | [isPartOfCHO](#) | [isPartOfPlace](#) | [isPartOfTimeSpan](#) | [isPartOfWebResource](#) | [levelOfGenesis](#) | [mentioned](#) | [misattributed](#) | [owner](#) | [pageDimension](#) | [painter](#) | [patron](#)

Navigation
 Export neologism
 vocabularies index

User login
 Username: *

 Password: *

[Request new password](#)

Who's online
 There are currently 0 users and 1 guest online.

Summary

- The DM2E Data Model is an application profile of the EDM for the domain of Manuscripts
- DM2E v1.0: Latest and first operational version
- DM2E v1.1: Next version under development
- Work is on-going and feedback welcome!

Thank you for your attention!

Questions and Feedback:

Steffen Hennicke,

Julia Iwanowa,

Evelyn Droege.

`vorname.nachname@ibi.hu-berlin.de`