

Everything you need to know
about ...

D2.5

Prof. Dr. Stefan Gradmann
Humboldt-Universität zu Berlin / School of Library and Information Science
stefan.gradmann@ibi.hu-berlin.de





Europeana: Two Timelines

- **Prototype** timeline:
 - Was to be operational by November 2008
 - Entirely built with existing human resources in The Hague (TEL-/EDLOffice) and Pisa (CNR ISTI)
 - Context of D2.2
- **Europeana 1.0** Timeline
 - Operational in 2010 (Q3)
 - Built using existing resources plus (hopefully) resources made available within Europeana 1.0 and EuropeanaConnect projects
 - D2.5 is the starting point of technical planning along this timeline



Europeana: D2.5 (and D2.2)

(Makx Dekkers, Stefan Gradmann, Carlo Meghini)



- **D2.2 Initial Semantic and Technical Interoperability Requirements**
 - Was to prepare the prototype timeline with a first set of functional and data specifications
 - All of which have been transferred to either ESE (mostly) or to D2.5 (some) => disregard D2.2
- **D2.5 Europeana Outline Functional Specification For development of an operational European Digital Library**
 - Version 1.2 made publicly available in August 2008
 - Currently working on V 1.5
 - Planning and specification document: do not expect guidelines!
 - Specification on mid term level



D2.5

■ ToC

- User Groups and Requirements
- Europeana System Functional View (including basic concepts such as Actors, **Logical Data Model**, Metadata Requirements and **Functional Components Overview**)
- Individual Functional Areas

Functionality

■ Prototype

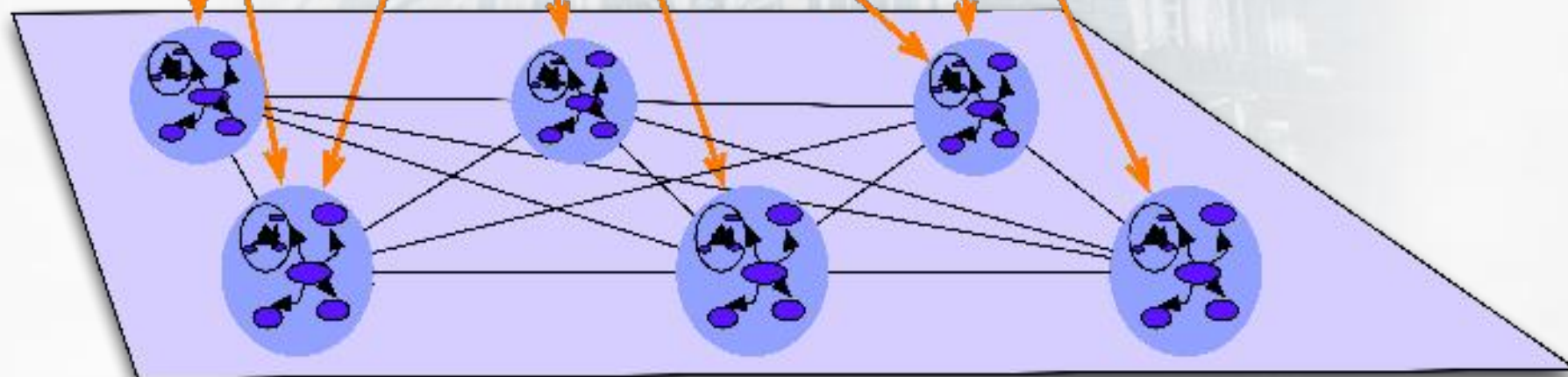
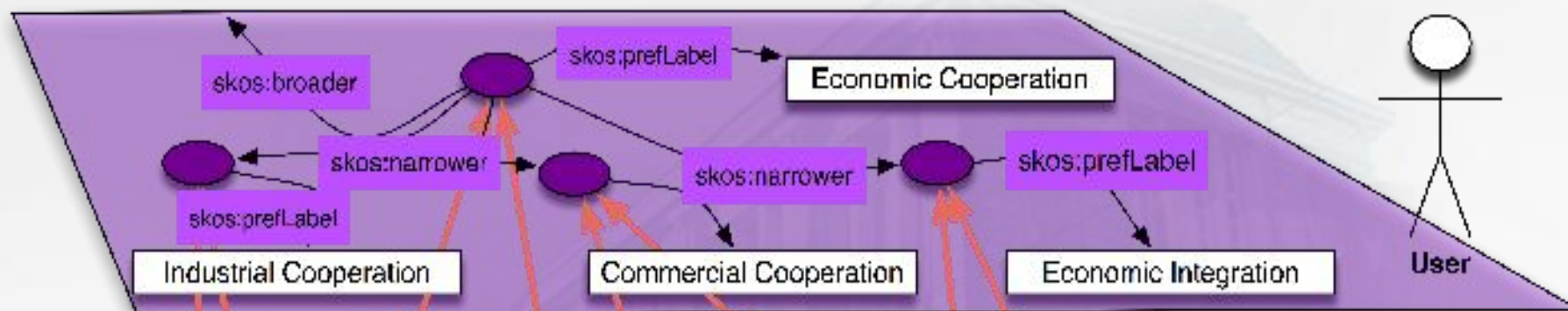
- Catalogue based **portal** including thumbnails
- With functional enhancements (e. g. faceted browsing, timeline, user tag)
- Essentially mediating access to remote original objects

■ Europeana 1.0

- API (+ portal)
- Interaction with rich contextualised surrogates may substitute access to original objects to some extent
- D2.5: "Europeana can be thought of as a **network of inter-operating object surrogates** enabling **semantics based object discovery and use.**"

Functionality: 1.0

Semantic Network



Networked Surrogates

Data

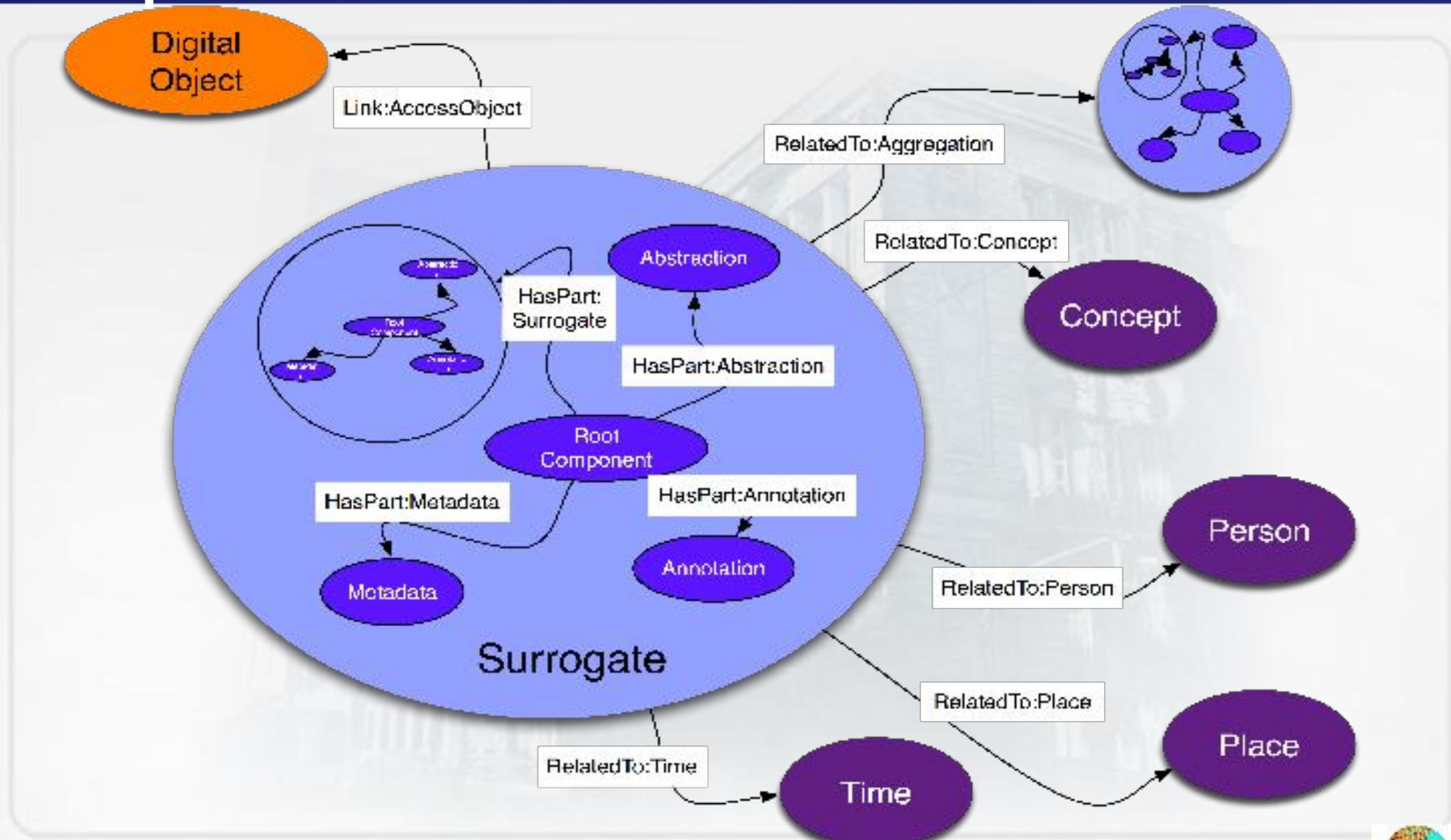
■ Prototype

- Metadata (DCQ with only a few extensions)
- All metadata attributes refer to original object on suppliers' sites, hence the extensions
 - <europeana:isShownBy> as a subproperty of <dc:relation>
 - <europeana:isShownAt> as a subproperty of <dc:relation>
 - <europeana:userTag> for content supplied by registered users

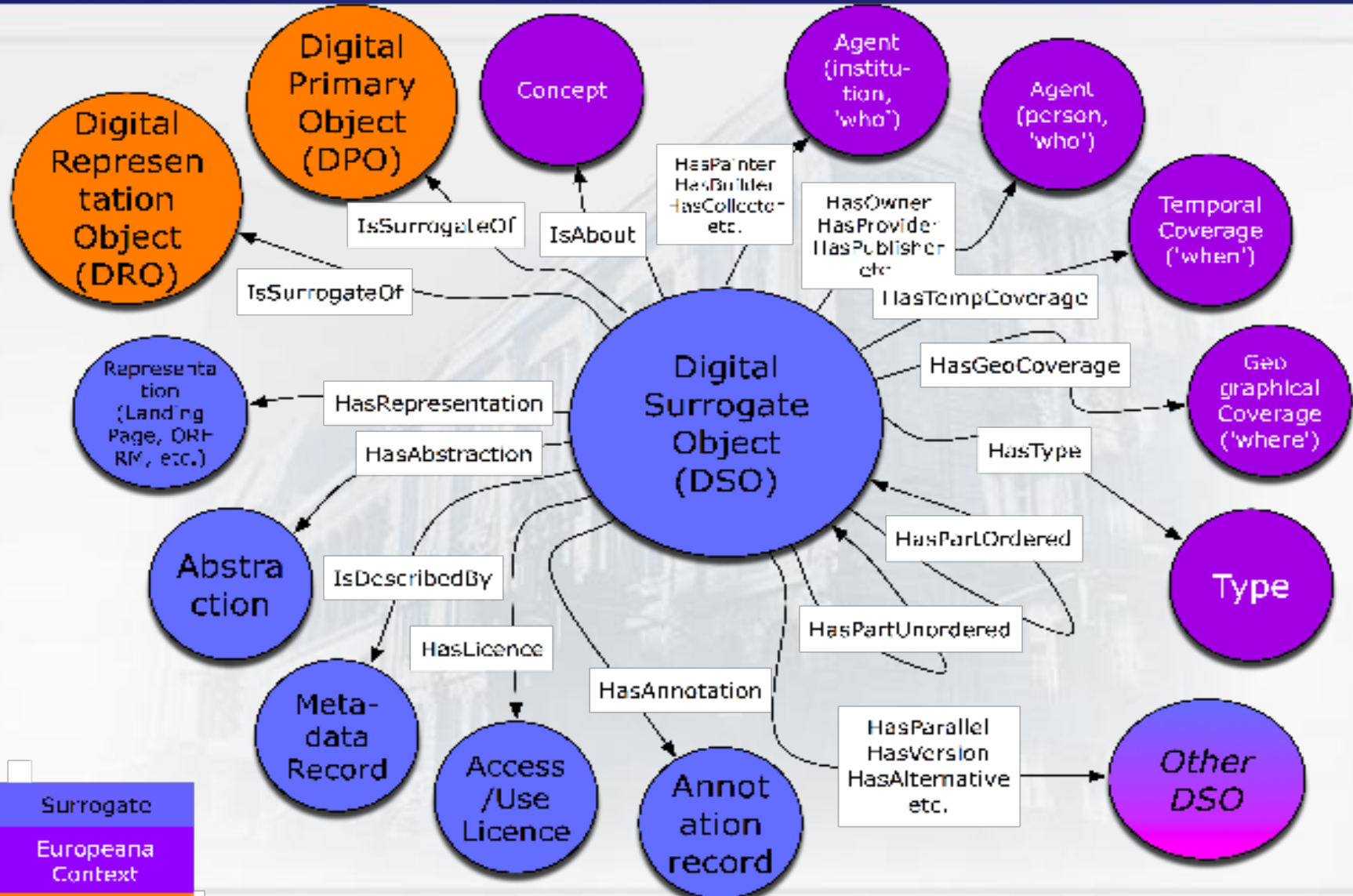
■ Europeana 1.0

- Object data and metadata as well as abstractions and other elements aggregated into rich 'surrogates' according to high level functional / technical specification D2.5
- surrogates are linked to each other and to Semantic Web nodes
- Details of underlying object model to be worked out in Europeana 1.0 taking into account CIDOC CRM, DCMI Abstract, OAI ORE

Data: Europeana 1.0 – as specified in D2.5 V1.2



From Object Model to Data Model: DSO Refined

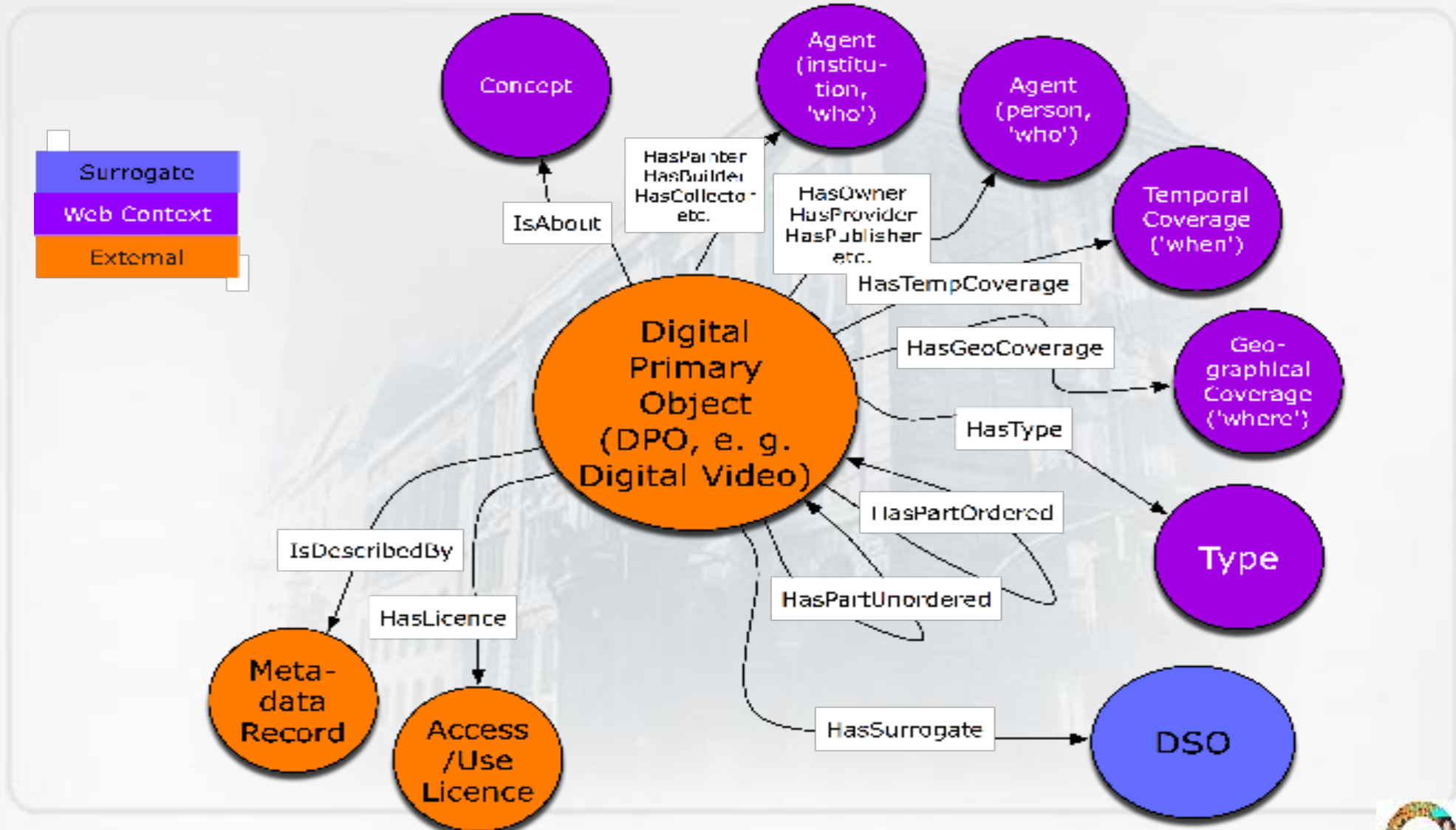


Legend for the diagram's color coding:

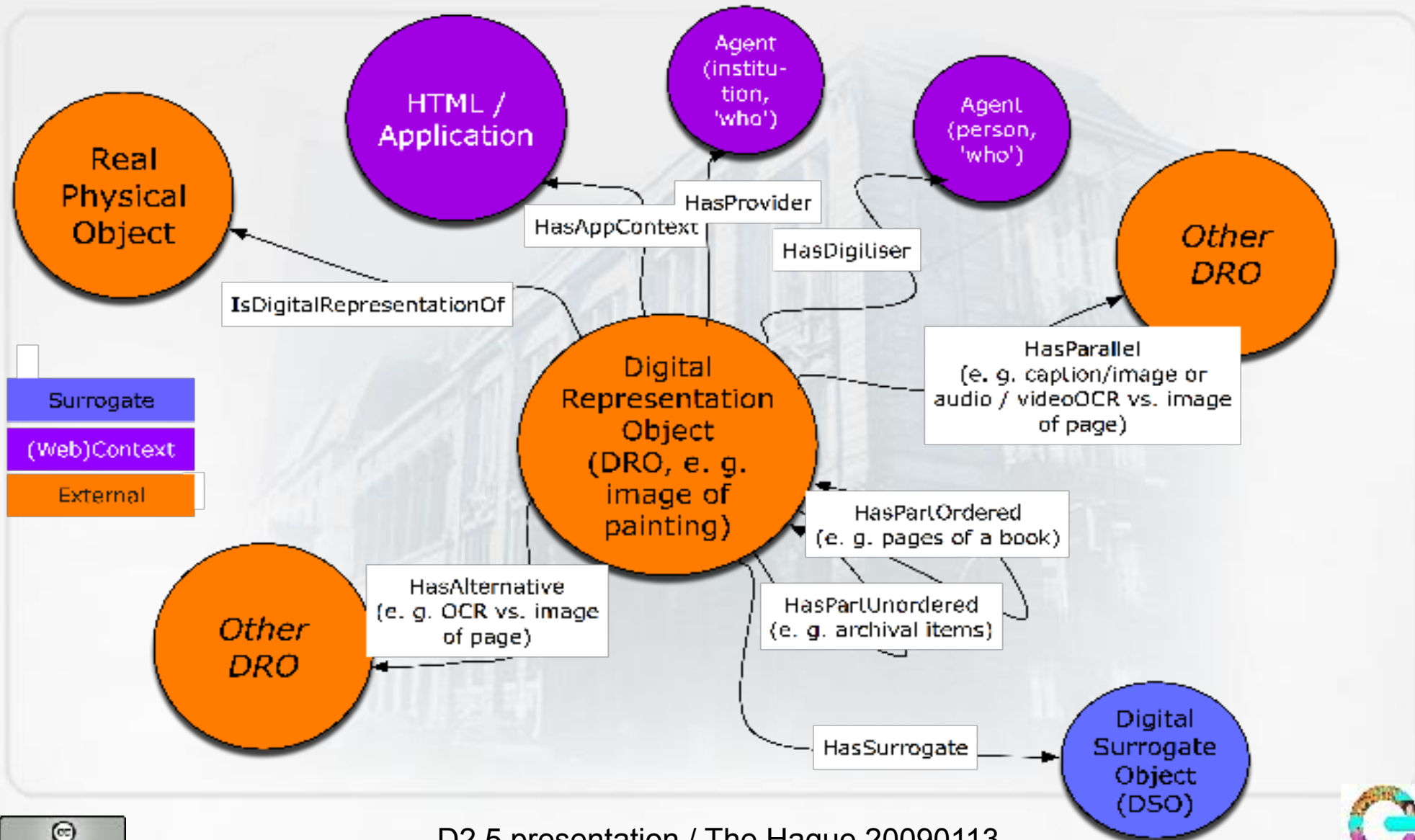
- Surrogate (Orange)
- Europeana Context (Purple)
- External (Blue)



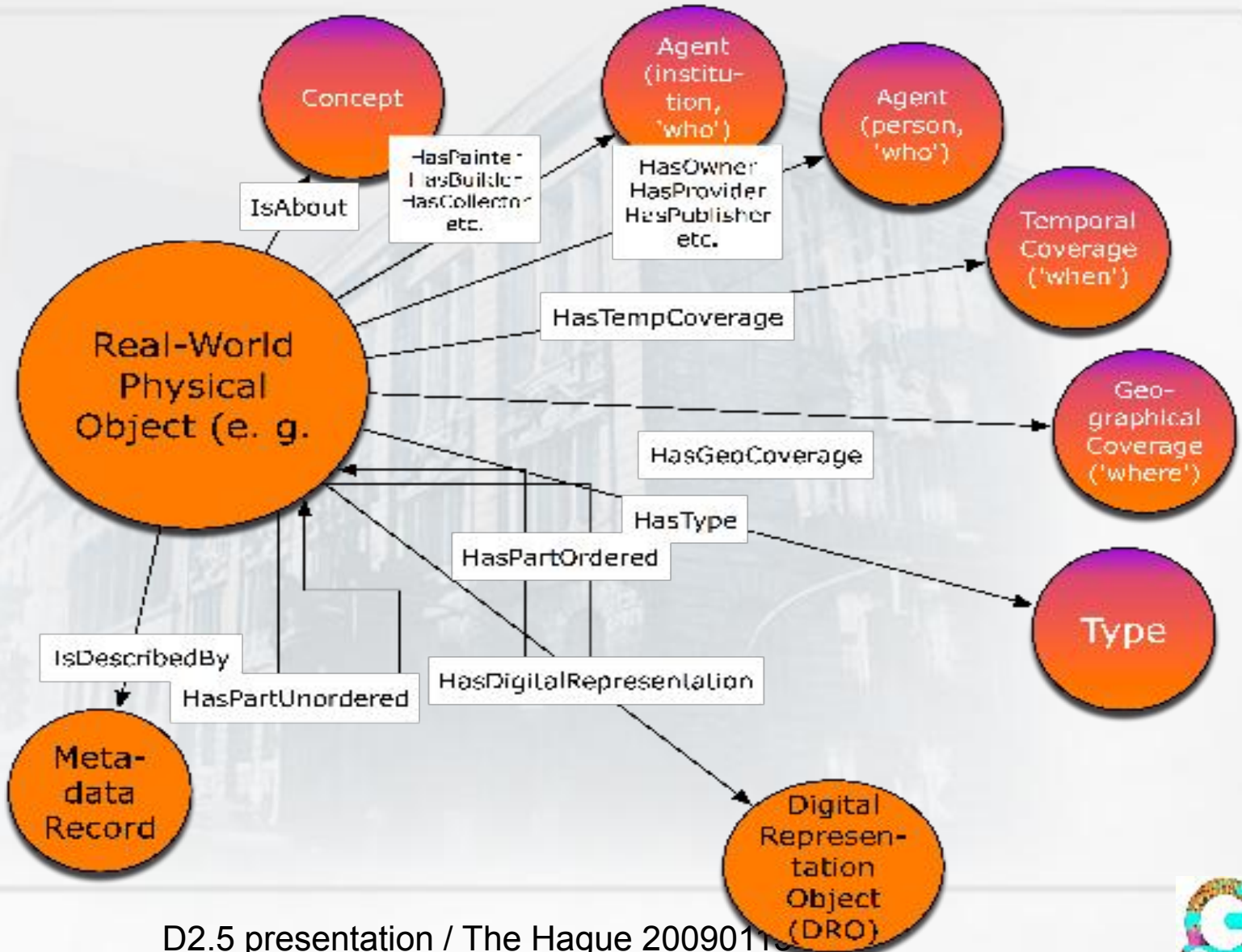
From Object Model to Data Model: DPO



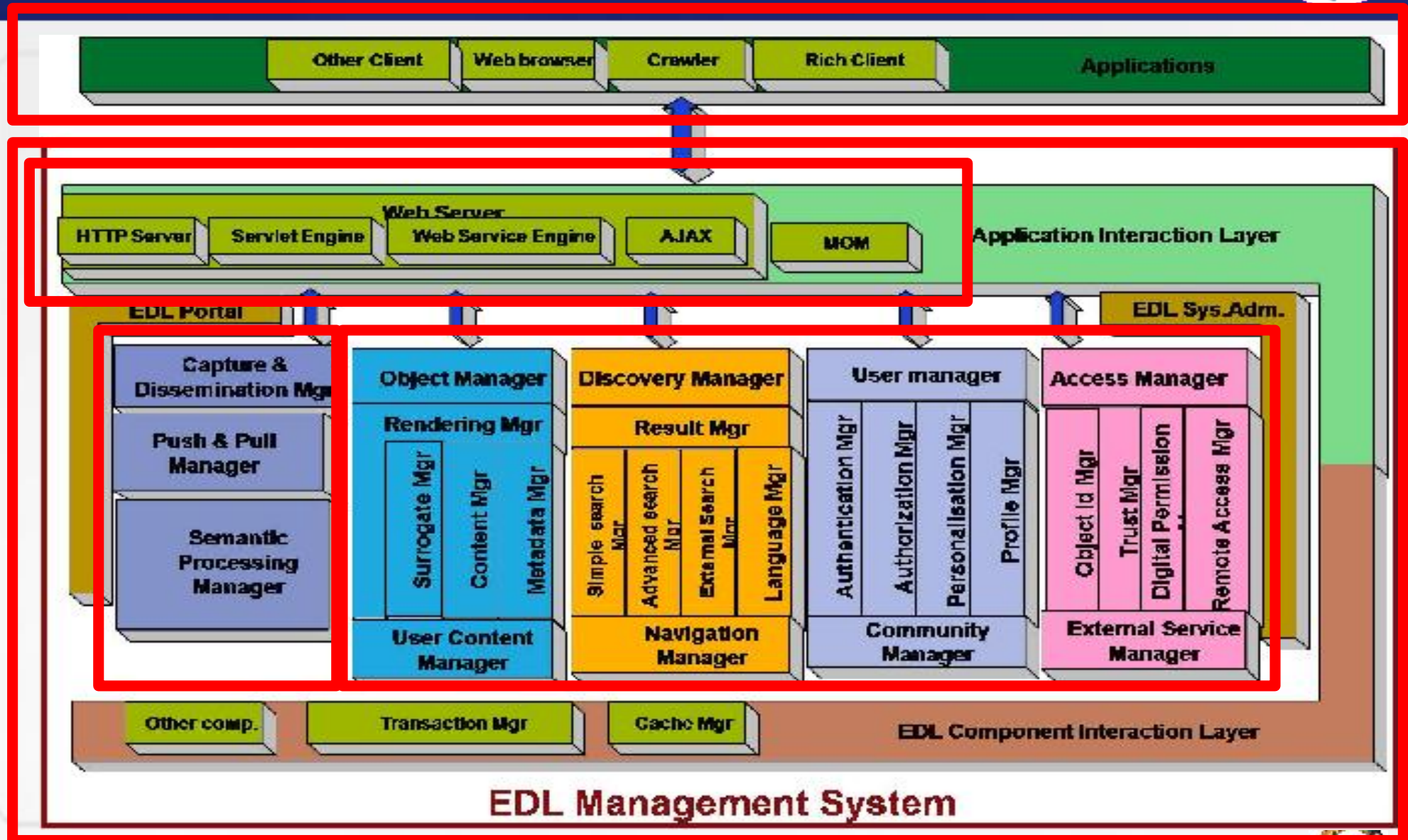
From Object Model to Data Model: DRO



From Object Model to Data Model: RPO



Component Architecture: 1.0



D2.5 Functional Component Architecture



- **1.0: standards based component architecture**
 - **OAI-PMH** for harvesting (+OAI-ORE),
 - **SRU, SOAP, JSR** and **OpenSearch** for external search and access
 - **SKOS, OWL/RDF(S)** for semantics
 - **SPARQL** for advanced search
 - **CIDOC/CRM** and **DCMI Abstract Model** for Metadata modelling
 - **SAML, LDAP** and **OpenID** for security and authentication
 - Lots of others: WSDL, WAI, UNICODE UTF-8, XSLT, NACO, UN-API "See-also", HTTP/HTTPS ...



Data Flows

■ **Prototype: One Way**

- Metadata is harvested (*not much decentral preprocessing!*)
- Transformed to Europeana ESE
- Integrated in Prototype Portal

■ **Europeana 1.0: Potentially Interactive**

- Metadata is harvested (*how much decentral preprocessing?*)
- Objects are accessed, aggregation elements extracted and combined with metadata into surrogates (*how much decentral preprocessing?*)
- Surrogate links to semantic nodes are created
- Surrogates are integrated in Europeana workspace
- Surrogates and semantic links could be sent back to suppliers



Please Keep in Mind ...

- The central Europeana data store will *only contain object surrogates and index files*. Original digital objects will stay with the content provider sites.
- Europeana will use the OAI-PMH harvesting approach. Regarding attributes, *the relevant specification is ESE*.
- Content aggregators and providers are strongly encouraged to provide more elaborate metadata so that we can *migrate* these to later versions of the ESE datamodel based on the Surrogate Object Model.
- Try to distinguish RPO and DRO layers if ever possible ...!

