



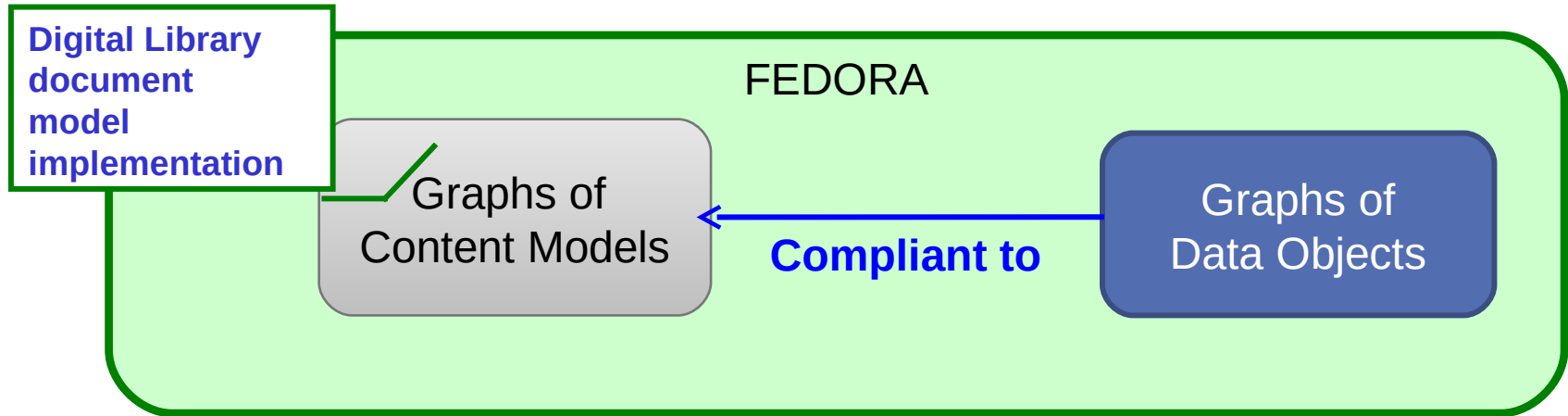
Customized
OAI-ORE and OAI-PMH exports
of compound objects for Fedora
repositories

Alessia Bardi, Sandro La Bruzzo, Paolo Manghi
name.surname@isti.cnr.it
<http://nemis.isti.cnr.it/groups/infrascience>

Scenario and motivations

- Digital Library Systems (DLSs) need to export content
 - Export of **compound objects**: packages of information objects with an identity
 - **OAI protocols** (typical solution)
- DL management systems (DLMSs) need to support export protocols
- DLMS issues:
 - Absence of OAI-exports: some DLMS do not provide support for OAI protocols (Relational Databases)
 - Pre-defined OAI exports of compound objects: shape of objects as exported by OAI protocols is pre-defined by the DLMS (Fedora)

Fedora: Flexible Extensible Digital Object Repository Architecture



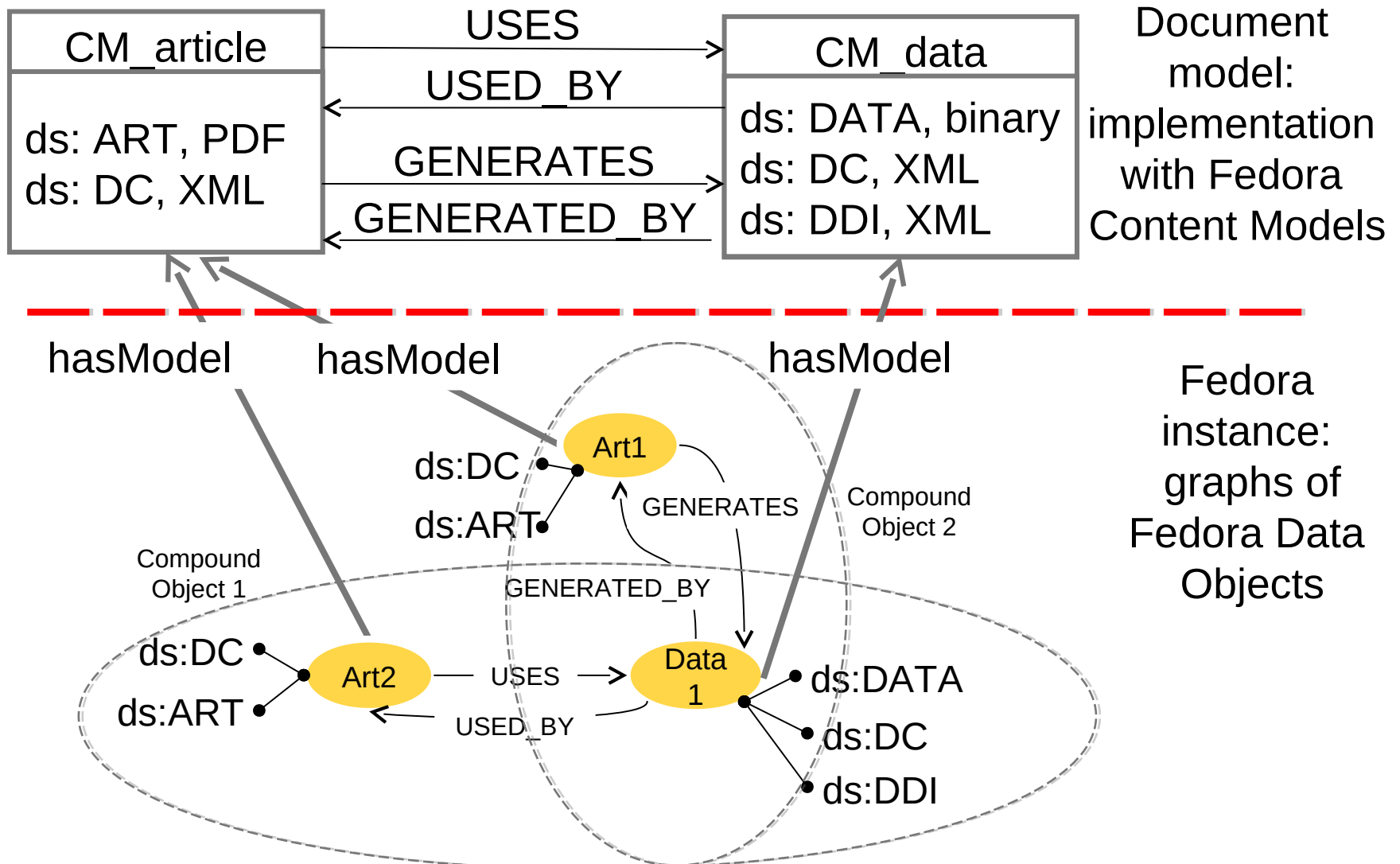
A Fedora instance manages graphs of Data Objects

Data Objects conform to Content Models

Content Models define:

- Type and name of mandatory datastreams (files)
- Allowed relationships to other Data Objects

Fedora objects and compound objects



Existing OAI-PMH Solutions

	Basic OAI-PMH Provider	OAI Provider
PMH-Set	Each Fedora Data Object is annotated with relationships to the Sets it belongs to	
PMH-Item	Fedora Data Object	Fedora Data Object
PMH Metadata Records	Datastream	Datastream
PMH Metadata Format	Dublin Core	Any format in existing datastreams
Compound Object Boundaries	Fedora Data Object	Fedora Data Object

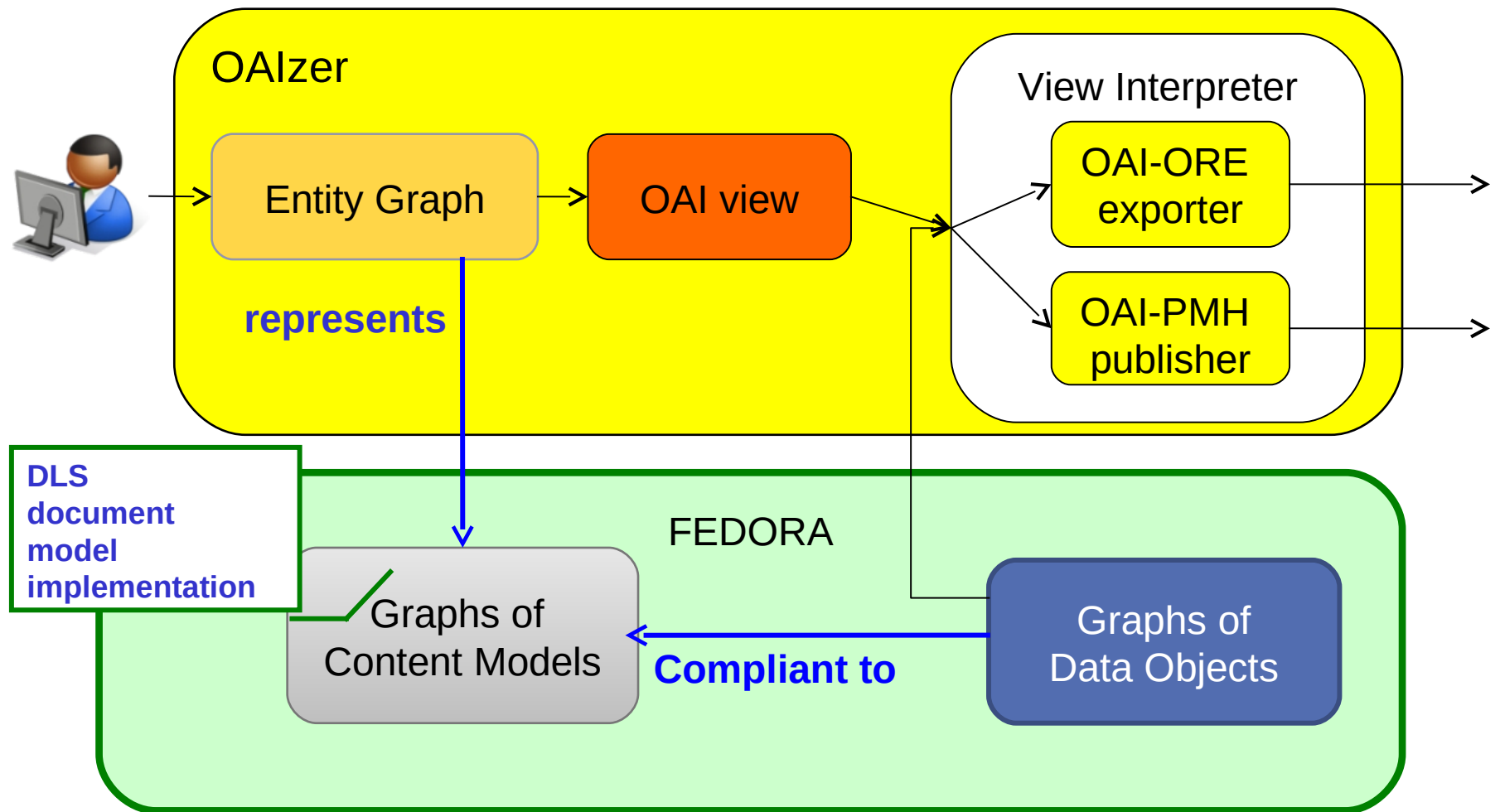
Existing OAI-ORE Solutions

	OREProvider	Fedora2ORE
ORE Aggregation	Defined by annotation	Sub graph visited starting from a given Fedora Data Object
ORE Aggregated Resources	Datastreams with a given name (e.g., DC)	Fedora Data Objects in the sub graph
ORE Proxy	not supported	
Relationships between Aggregated Resources	NO	
Compound Object Boundaries	Fedora Data Objects annotated with the same tag	Navigation depth

exporting compound objects: the OAIzer solution

- OAIzer is a generic software for the customization of compound objects in Digital Libraries
 - Software layer over the Digital Library Management System
 - ORE exports include relationships and their semantics
 - PMH metadata formats can be generated on request
- DB-OAIzer:
 - works on RDBMS;
 - prototype in the OpenAIRE project, to be integrated with OpenAIRE+ (Open Access Infrastructure for Research in Europe)
- Fedora-OAIzer:
 - for Digital Library Systems based on Fedora frameworks with Content Models

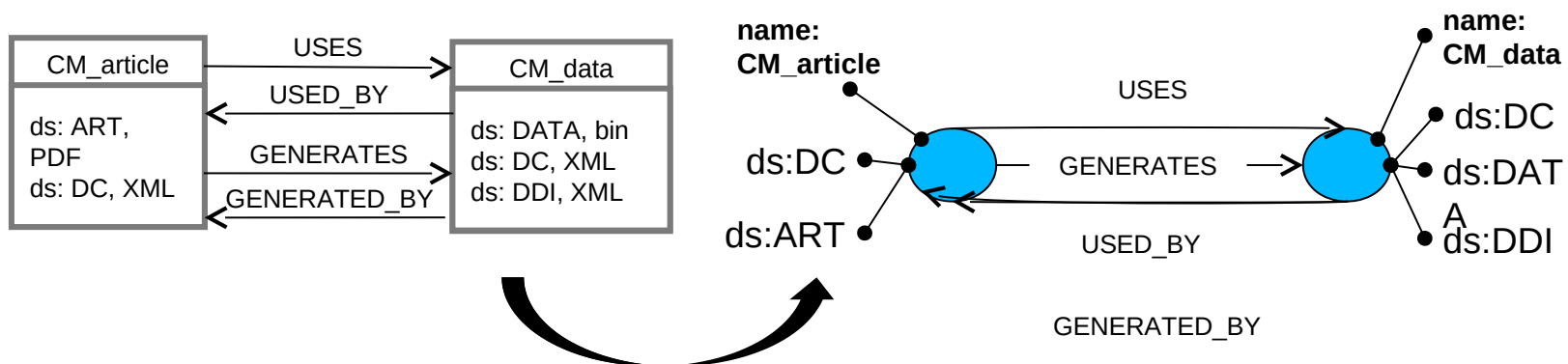
Fedora-OAIzer: High-level architecture



PHASE 1: ENTITY GRAPH

Entity Graph: Graph representation of the document model

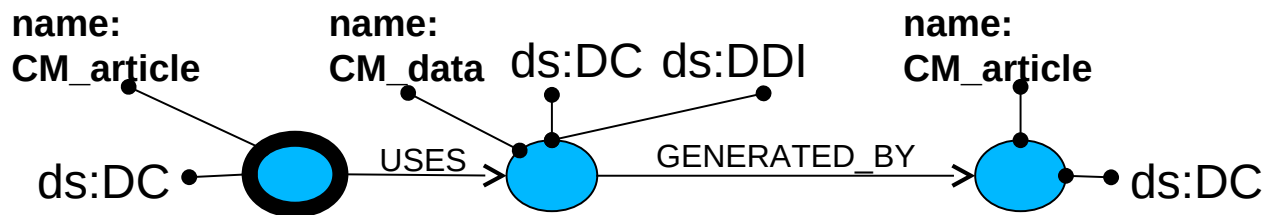
- Created exploiting information about Content Models
- Graph nodes represent Content Models
- Properties of a node represent the datastreams defined by the Content Model
- Edges between nodes represent the relationships defined by the Content Models



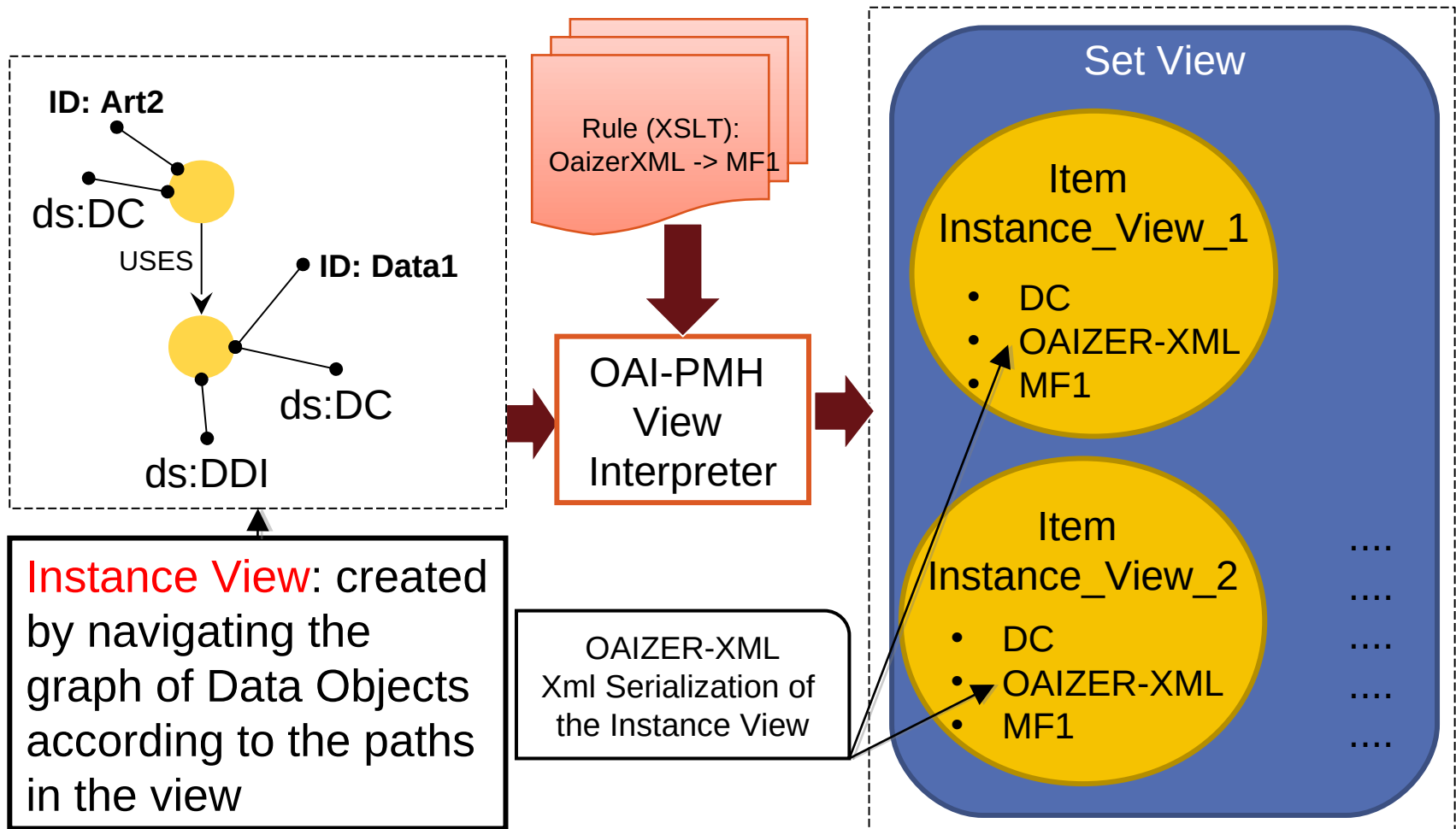
Phase 2: definition of the oai view

OAI View: Tree representation of the structure of the compound objects to export

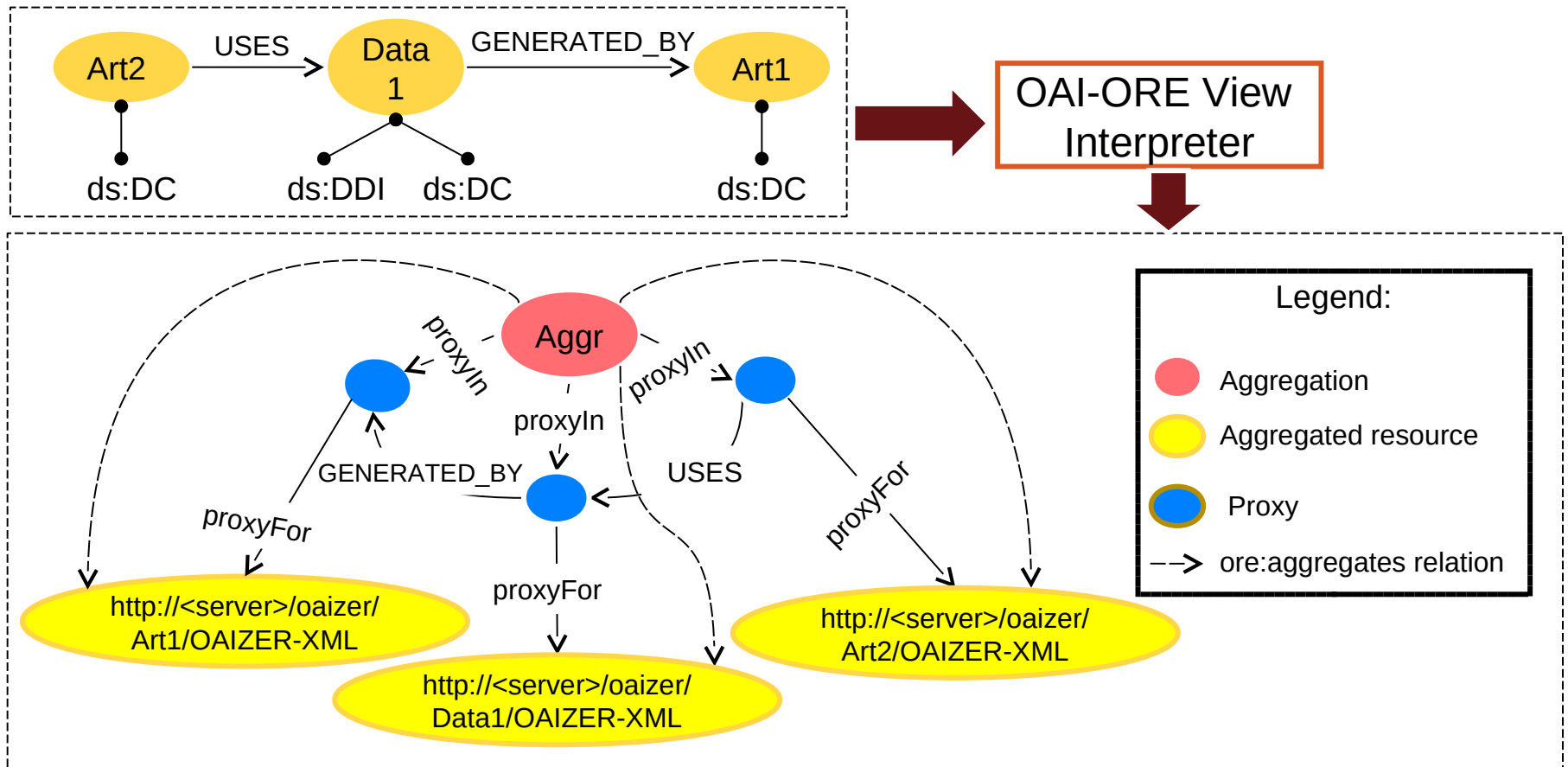
- Sub graph of the entity graph with entities and relationships to include in the compound object
 - Ex. The DATA datastream of CM_data is excluded
- The root node is the entry point of the OAI view



PHASE 3.1: INTERPRETATION OF THE OAI VIEW for oai-pmh



Phase 3.2: interpretation of the OAI view for OAI-ORE



ConclusionS

D

adapt the exports to your document model!

- Customized, domain-dependent exports of compound objects
- Future work:
 - Better graphical support for the definition of views in case of entity graphs with cycles
 - Implementations for different back-ends
 - Integration of DB-OAIzer with

