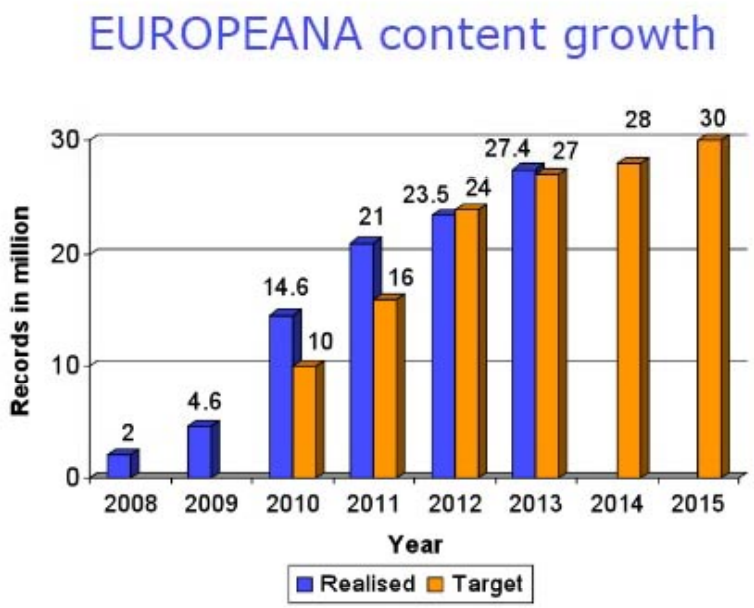
Europeana Creative.
EDM Endpoint. Custom Views.

Dr. Nikola Ikonomov
with Boyan Simeonov, Jana Parvanova and Vladimir Alexiev

Content is everything in the digital age



Europeana - a search portal which links the user to the actual digitized object on the institutions webpage.
Is this enough? How can this huge potential of data be further unleashed?

Europeana – the next step

- Recently opened up its dataset of over 25 million cultural objects to the public.
- the material can be used for any purpose – creative, educational commercial….
- the free availability of this data offers a new boost to the digital economy. The sector of cultural and creative industries (particularly from the learning, tourism and games sectors) represents 3.3% of EU GDP and is worth over €150 billion in exports.
- new and exciting opportunities for app developers, designers and other digital innovators have been created.

Obstacles and challenges

- Organizational: lack of sustainable environment for experimentation. .
- Technical : Europeana API doesn't retrieve content.
- Legal : Licensing framework doesn't cover content but only metadata.
- Cultural: no real-world case studies.

The Europeana Creative project

In September 2012, Europeana encouraged the development of innovative applications by:

- publishing the metadata for 25 million cultural heritage objects under the terms of the Creative Commons Public Domain Dedication 0 (CC0) license;
- providing free and open access to the metadata through API and Linked Open Data.
- *Europeana Creative* will not only use this metadata, but also many of the digital objects themselves, which are available for re-use together with the necessary licenses.

Project objectives 1

- Establish the Europeana Open Labs Network for experimentation beginning with four Open Lab affiliates and a virtual space.
- Create the Europeana Content Re-use Framework which will extend the Europeana Licensing Framework from metadata to content.
- Implement the infrastructure and essential services needed to support re-use of cultural resources, at the same time boosting creativity and long-term business development.
- Create five pilot applications in the thematic areas of History Education, Natural History Education, Tourism, Social Networks and Design.

Project objectives 2

- Identify, incubate, and spin-off five viable projects into the Creative Industry sector.
- Identify business models allowing key stakeholders within the Europeana ecosystem to develop products and services based on the Content Re-use Framework.
- Evaluate all work streams, while the delivery of pilots and services supports best practice learning.
- Project results will be promoted via an extensive stakeholder engagement campaign in the Creative Industries sector and the Europeana Network.

The role of Ontotext

- Creation of Content Retrieval Services as part of the Content Re-use Framework - a sophisticated platform for the retrieval of Europeana metadata and related digital content objects.
- Deployment of OWLIM platform for metadata storage and search. Europeana Creative will make use of this triple store as a central data integration repository.
- Geographic mapping to be used for the creation and management of cultural routes as required by the Tourism pilot.
- Linking to external web resources (e.g. Freebase, DBpedia, Wikipedia, VIAF, Getty, Geonames).
- Creation of services for geo-referencing of data that are required by the Social Networks Pilot. Geo-referenced cultural objects can also be used for location-based services.

ESE schema

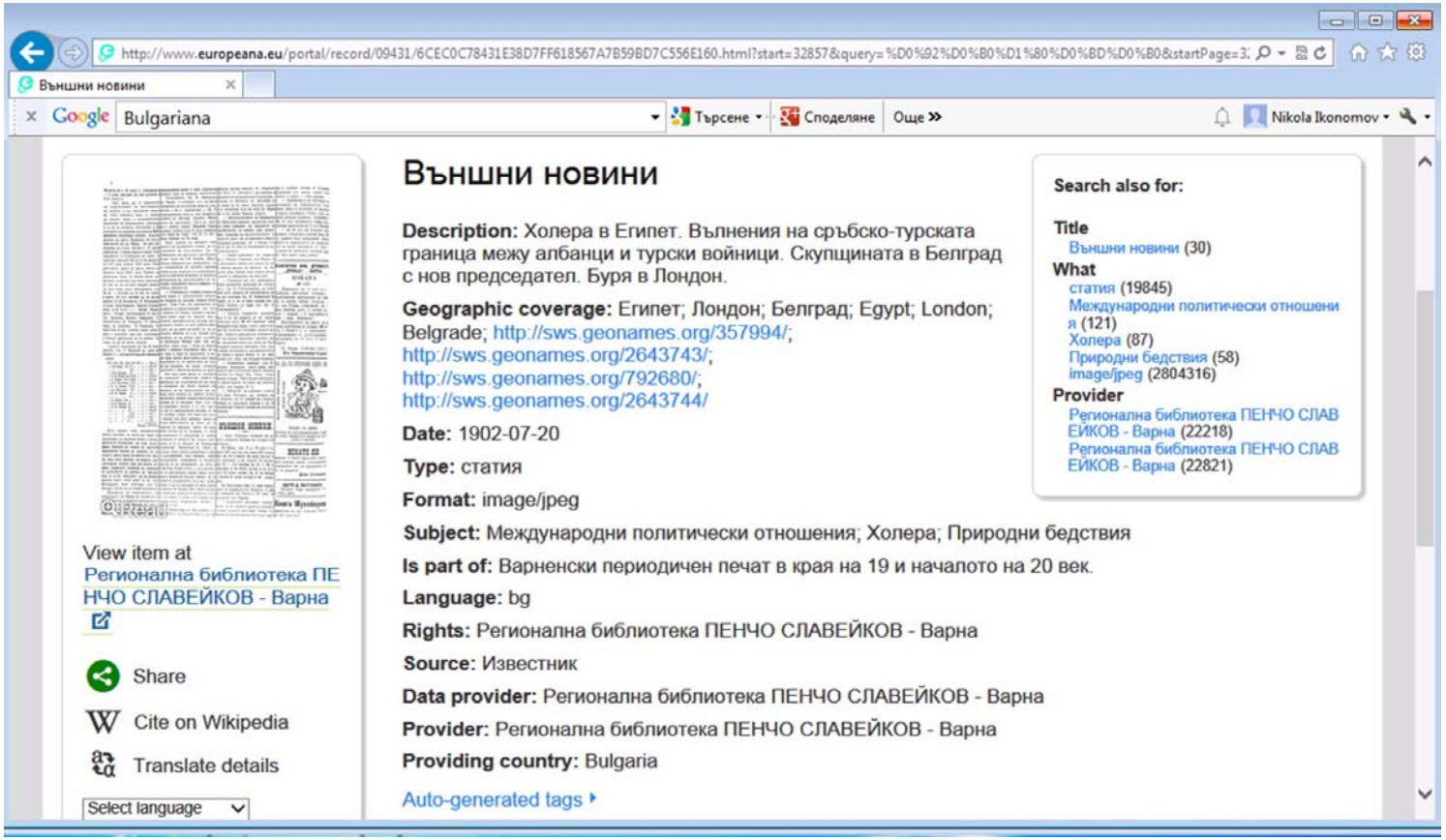
Source	Element	Element Refinement(s)
DC	title	alternative
DC	creator	
DC	subject	
DC	description	tableOfContents
DC	publisher	
DC	contributor	
DC	date	created; issued
DC	type*	
DC	format	extent; medium
DC	identifier	
DC	source	
DC	language*	
DC	relation	isVersionOf; hasVersion; isReplacedBy; replaces; isRequiredBy; requires; isPartOf; hasPart; isReferencedBy; references; isFormatOf; hasFormat; conformsTo
Europeana		isShownBy; isShownAt
DC	coverage	spatial; temporal
DC	rights	
DC terms	provenance	
Europeana	userTag	
Europeana	unstored	
Europeana	object	
Europeana	language*	
Europeana	provider	
Europeana	type*	
Europeana	uri	
Europeana	year	
Europeana	hasObject	
Europeana	country	

- Each of the different heritage sectors represented in Europeana uses different data standards, and ESE reduces these to the lowest common denominator.

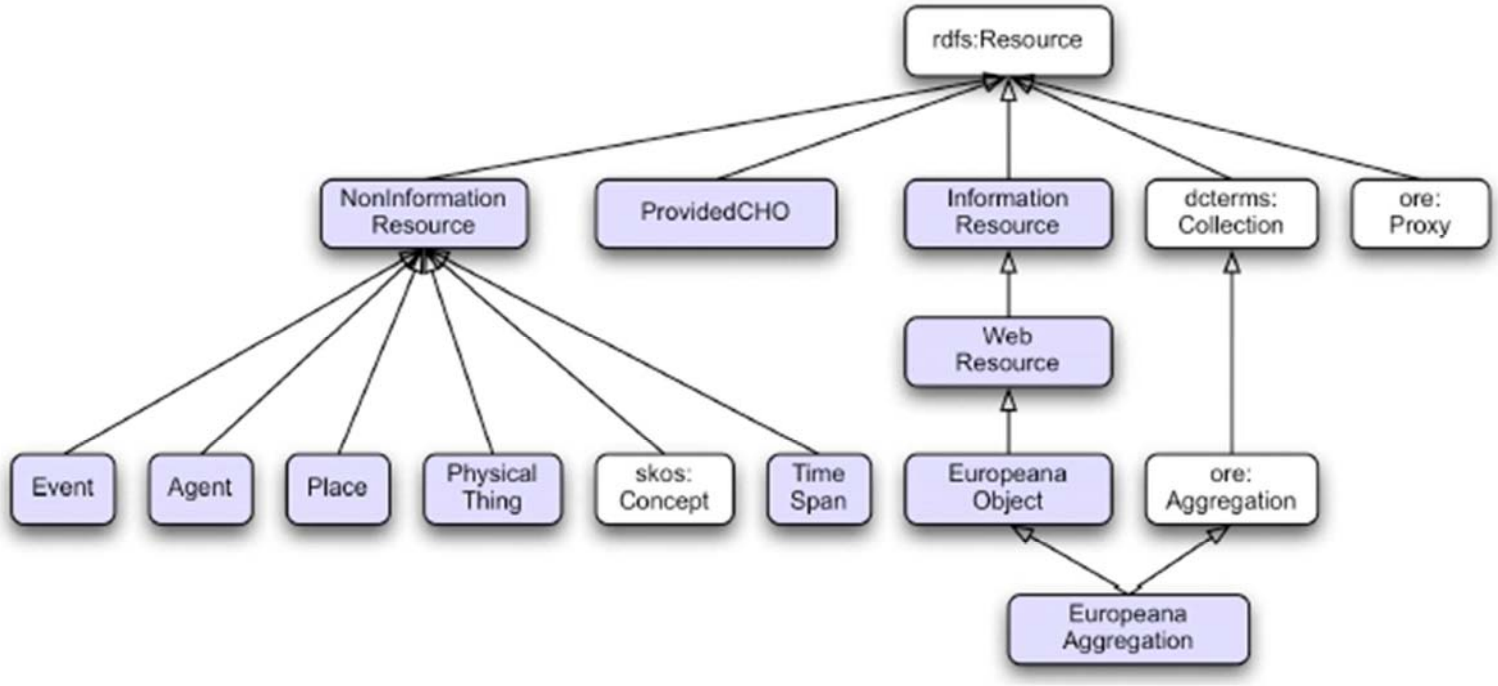
- 15 original Dublin Core (DC) metadata elements, a subset of the DC terms

- a set of twelve elements which were created to meet Europeana's specific needs.

How does one CH object looks like in Europeana?



EDM - the new data structure for Europeana



Classes introduced by EDM are in blue, classes in white are re-used from other schemas.

What makes EDM different?

- not built on any particular community standard.
- allows for richer and truly semantic representations of CHO.
- Integrates ontologies like SKOS, Dublin Core, and FOAF.
- uses RDFS as meta-model and URIs to identify information about CHO.
- uses OAI-ORE as ontology framework thus making Europeana compatible with the Semantic Web paradigm and enables it to become part of the emerging LOD community.
- has 3 core classes of resources that result from the package of data provided to Europeana: *ProvidedCHO, WebResource* and *Aggregation*. Each of them can be subject to corresponding metadata, leading to three core "metadata sections".

Ontotext's EDM Endpoint

- Endpoint - a specific location for accessing a web service using a specific protocol and data format. EDM (SPARQL) endpoint - a URI to which queries can be sent, and which returns answers to the queries as a response.
- In December 2012 Ontotext launched the Europeana's EDM Endpoint, which allows CH data from data.europeana.eu to be accessed from applications and to be explored through the UI of the Forest framework. One billion explicit statements have been loaded in the OWLIM repository and with inference they stack up to about 4 billion distinct retrievable statements.
- Ever since Ontotext launched the Endpoint, the positive reactions from the community and the number of visitors constantly increased.

OWLIM

- The backbone of the EDM Endpoint is OWLIM - a family of semantic repositories, or RDF database management systems, with the following characteristics:
  - native RDF engines, implemented in Java;
  - delivering full performance through both Sesame and Jena;
  - best scalability, loading and query evaluation performance.
- OWLIM comes in 3 editions: OWLIM-Lite, OWLIM-SE (Standard Edition) and OWLIM-Enterprise. All editions share the same inference mechanisms and semantics.

Forest framework

- Forest - web interface that brings together tools for execution of SPARQL queries, visual exploration of RDF triples, full text search and integration of popular third-party data visualization components.
- Forest is an optimal platform for exploring the data in an OWLIM repository. It is the front end for linked data web services offered by Ontotext.
- Forest is used in various client applications providing a flexible architecture for development of custom functionalities like user-friendly RDF views.

Visualization of a CH object

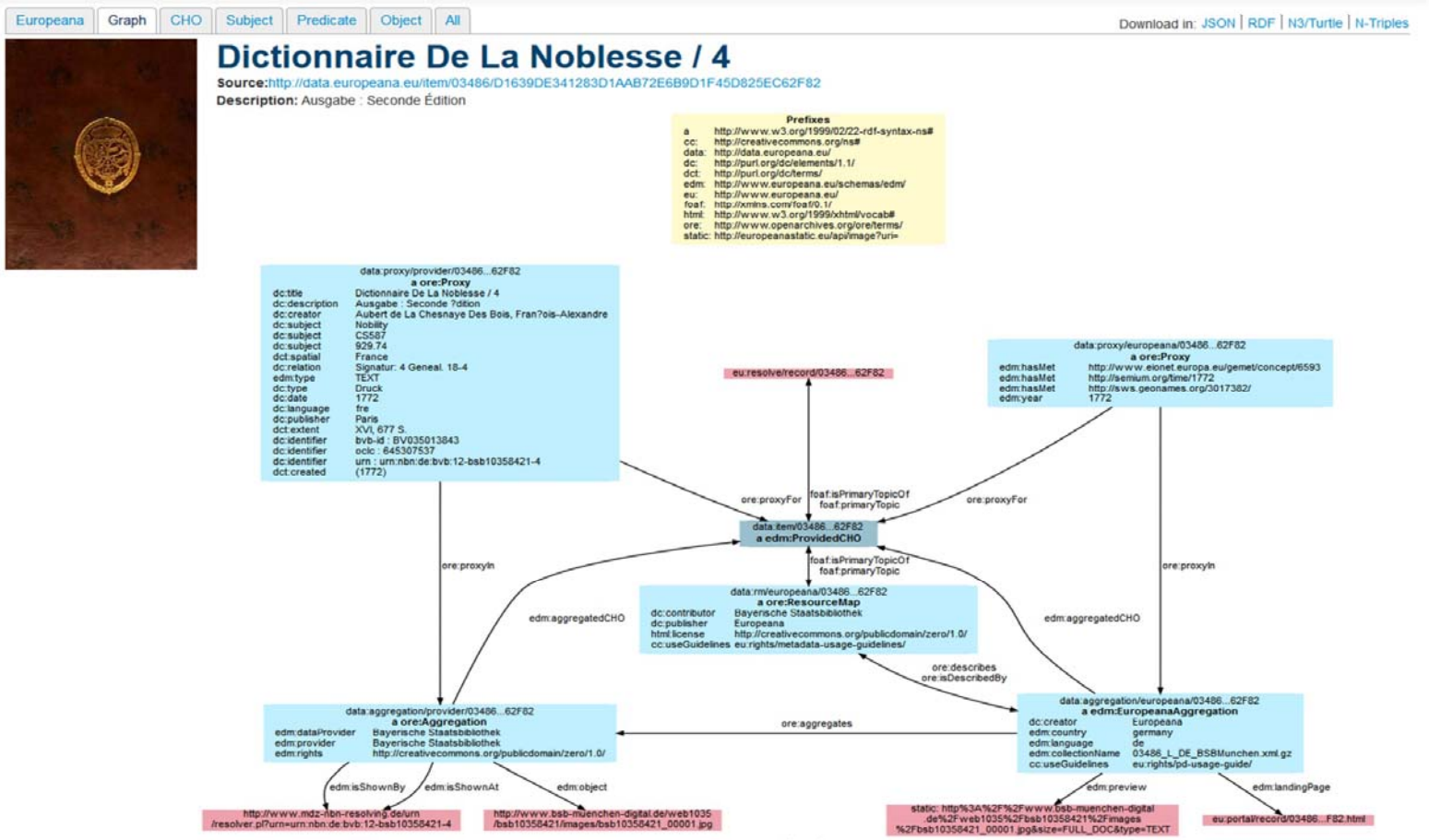


Fig. 3

ontotext

DIPP 2013, Veliko Tarnovo, September 2013

What is typical for custom views?

- All triples of all nodes comprising the CHO are shown. Download can be done in JSON, RDF, Turtle, NTriples.
- A view from the Europeana portal is shown, including links to the original (source).
- A CHO Graph and connections between the nodes are shown. URLs are shortened. Data is compressed into the node.

Thank you for your attention!

Questions?

nikola.ikonomov@ontotext.com