EUROMUSE: A web-based system for the management of MUSEum objects and their interoperability with EUROpeana

Varvara Kalokyri, Giannis Skevakis

Laboratory of Distributed Multimedia Information Systems & Applications (MUSIC, TUC), Technical University of Crete, Greece

Chania, November 2011
Motivation

- Expansion of the traditional model of museums to include high resolution images of their exhibits
  - Acceleration in the digitization of information.
  - Increasing capacity of digital information storage.

- Interconnection of museum objects with the outside world for exploration from any place with Internet connectivity

- Museums don’t have the ability to purchase a complex system for publishing, organizing and authoring museum objects

- Lack of semantic linkage with already published global vocabularies

- Lack of a system that supports standard-based metadata description of museum objects
Motivation

- **Europeana foundation**
  - Started by the presidents of major European nations
  - Financed by the member states and EU programs

- **Europeana internet portal**
  - Web portal exposing increasingly impressive amounts of digitized cultural heritage objects from various sources throughout Europe.

- **Main purpose**: enable people to explore the digital resources of Europe's museums, libraries, archives and audio-visual collections.

- Contribution from 1500 European institutions (Louvre, British Library)

- More than 15,000,000 items
Overview

- Motivation
- **Contribution**
- Background technologies
- Functional specification
- The ESE-CHO Application Profile
- System Architecture
- GUI Design Specification
- Implementation
- Evaluation
- Conclusion
Contribution

- **EuroMuse**
  - Allows the publishing of cultural heritage objects
  - Facilitates the authoring and metadata enrichment of cultural heritage objects
  - Establishes the interoperability between museums and Europeana
  - Provides seamless ingestion of legacy metadata.
Contribution

- Main features
  - Web-based management system
  - Shared access for multiple users with concurrency control mechanisms
  - Supports a rich metadata element set (ESE-CHO Application Profile)
  - Semantic linkage of the objects with well-established controlled vocabularies
  - Supports a variety of the most popular multimedia formats.
Contribution

- Developed in the context of Natural Europe project.
  - Connect the digital collections of *European Natural History Museums* with *Europeana*

- Used by NHMs
  - Natural History Museum of Crete
  - National Museum of Natural History
  - Jura-Museum Eichstaett
  - Arctic-Center
  - Estonian Museum of Natural History
  - Hungarian Natural History Museum

- Already includes over 2000 fully described museum objects

- Target: 16,000 museum objects
Overview

- Motivation
- Contribution
- Background technologies
  - Functional specification
  - The ESE-CHO Application Profile
  - System Architecture
  - GUI Design Specification
- Implementation
- Evaluation
- Conclusion
XML/XML Schema
- Set of rules for encoding documents in machine-readable form.
- Description of a type of XML document.

XML Beans
- Java to XML binding framework

RDF
- Standard language for representing information about resources in the World Wide Web

Europeana Semantic Elements (ESE)

Google Web Toolkit (GWT)

Simple Knowledge Organization System (SKOS)
Europeana Semantic Elements (ESE)

- Dublin-Core based application profile
- Generic terms applied to heterogeneous materials
- Metadata as simple text descriptions

<table>
<thead>
<tr>
<th>Mandatory elements</th>
<th>Recommended elements</th>
<th>Optional elements</th>
<th>Elements supplied by Europeana</th>
</tr>
</thead>
<tbody>
<tr>
<td>dc:title</td>
<td>dcterms:alternative</td>
<td>dc:format</td>
<td>europeana:country</td>
</tr>
<tr>
<td>dc:description</td>
<td>dc:creator</td>
<td>dcterms:extent</td>
<td>europeana:language</td>
</tr>
<tr>
<td>dc:language</td>
<td>dc:contributor</td>
<td>dcterms:medium</td>
<td>europeana:uri</td>
</tr>
<tr>
<td>europeana:dataProvider</td>
<td>dc:date</td>
<td>dc:identifier</td>
<td>europeana:usertag</td>
</tr>
<tr>
<td>europeana:isShownAt</td>
<td>dcterms:created</td>
<td>dc:rights</td>
<td>europeana:year</td>
</tr>
<tr>
<td>europeana:isShownBy</td>
<td>dcterms:issued</td>
<td>dcterms:provenance</td>
<td></td>
</tr>
<tr>
<td>europeana:provider</td>
<td>dcterms:temporal</td>
<td>dc:relation</td>
<td></td>
</tr>
<tr>
<td>dc:subject</td>
<td>dc:publisher</td>
<td>dcterms:conformsTo</td>
<td></td>
</tr>
<tr>
<td>dc:type</td>
<td>dc:source</td>
<td>dcterms:hasFormat/isFormatOf</td>
<td></td>
</tr>
<tr>
<td>dc:coverage</td>
<td>dcterms:isPartOf</td>
<td>dcterms:hasVersion/isVersionOf</td>
<td></td>
</tr>
<tr>
<td>dcterms: spatial</td>
<td>europeana:object</td>
<td>dcterms:isReferencedBy/references</td>
<td></td>
</tr>
<tr>
<td>europeana:rights</td>
<td></td>
<td>dc:hasPart</td>
<td></td>
</tr>
<tr>
<td>europeana:type</td>
<td></td>
<td>dcterms:isReplacedBy/replaces</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>dcterms:isRequiredBy/requires</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>dcterms:tableOfContents</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>europeana:unstored</td>
<td></td>
</tr>
</tbody>
</table>
Background technologies

Europeana Semantic Elements (ESE)

Example

```
<metadata>
  <record>
    <dc:coverage>Greece</dc:coverage>
    <dc:creator>Trichas, A.</dc:creator>
    <dcterms:issued>09/11/2010</dcterms:issued>
    <dc:description>Mediterranean monk seal</dc:description>
    <dc:format>TIFF</dc:format>
    <dc:identifier>35651</dc:identifier>
    <dc:language>English</dc:language>
    <dc:publisher>Natural History Museum, Crete</dc:publisher>
    <dc:rights>NHMC</dc:rights>
    <dc:source>Natural History Museum, Crete</dc:source>
    <dc:subject>Zoology</dc:subject>
    <dc:subject>MAMMAL</dc:subject>
    <dc:title>Monk seal Monachus monachus</dc:title>
    <dc:type>Image</dc:type>
    <dcterms:alternative>Monachus monachus</dcterms:alternative>
    <dcterms:created>12/10/2002</dcterms:created>
    <dc:spatial>NHMC, Knossou Avenue, Irakleiou</dc:spatial>
    <dc:collectionName>NHMC data_set_1</dc:collectionName>
    <dc:language>el</dc:language>
    <dc:identifier>NHMC_data_set_1</dc:identifier>
    <dc:subject>MAMMAL</dc:subject>
    <dc:subject>Greece</dc:subject>
    <dc:subject>Zoology</dc:subject>
    <dc:creator>Trichas, A.</dc:creator>
    <dc:coverage>Greece</dc:coverage>
    <dc:language>en</dc:language>
    <dc:subject>Zoology</dc:subject>
    <dc:subject>Greece</dc:subject>
    <dc:subject>MAMMAL</dc:subject>
  </record>
</metadata>
```

- "Μονάχους μονάχους" @ el
- "Monk seal Monachus monachus" @ en
- Trichas, A.
- Natural History Museum Of Crete
- 12/10/2002
- NHMC, Knossou Avenue, Irakleiou
- Greece
- Natural History Museum Of Crete
- 09/11/2010
- 35651
- NHMC_data_set_1
- IMAGE
- Monachus monachus
- Image
- Zoology
- MAMMAL
- Greece
- English
- Mediterranean monk seal (common name).

Http://www.nhmc.uoc.gr/zooology/35651_tn.jpg
Http://www.nhmc.uoc.gr/zooology/35651.html
What is GWT?
- framework for Rich Internet Applications (RIA)
- web application with desktop characteristics
- extensive use of Ajax
- enhanced interactivity and improved user experience
- Better responsiveness and information flow
- Provides a set of core Java APIs to write RIAs.
- Compiles client-side Java code to highly optimized cross-browser compatible JavaScript.
- Developed and used by Google.
- Open source and completely free.

Recommended by Europeana
Google Web Toolkit (GWT)

- Why use GWT?
  - UI development similar to Swing (desktop applications).
  - No need to write HTML / JavaScript, but can if desired.
  - Use CSS for formatting.
  - Full IDE support – easy development and debugging.
  - Send complex Java types to and from the server
    - Data gets serialized across network.
  - Documentation

- Learning Time

- Slow compilation
Simple Knowledge Organization System (SKOS)

- Model for **expressing** the basic structure and content of thesauri, classification schemes, taxonomies, or any other type of structured controlled vocabulary.
- RDF format
- Concepts can be composed and published on the Web
- Machine readable and linked with data on the Web

**Concept**

- `skos:prefLabel`
- `skos:altLabel`
- `skos:broader`
- `skos:narrower`

Background technologies
Overview

- Motivation
- Contribution
- Background technologies
- Functional specification
- The ESE-CHO Application Profile
- System Architecture
- GUI Design Specification
- Implementation
- Evaluation
- Conclusion
Functional requirements

- Easy to install/use
- Empower the collaboration among curators
- Metadata Unification
- CHO Publishing
- Thumbnail creation
- Multilingual tool
- Semantic linkage with well-established controlled vocabularies
- Schema-independent
- Flexibility and modularity
- Highly customizable
Functional requirements

- Create-Read-Update-Delete (CRUD)
  - CHO
  - CHO Metadata
  - CHO Collection
  - CHO Collection Metadata

- Support different user categories
  - Specific access rights
    - Guest (Review CHO collection/CHO metadata)
    - Curator (CRUD CHO collection/CHO metadata)
    - Administrator (CRUD user accounts, CHO collection/CHO metadata)
NHM Metadata Lifecycle

![Flowchart](image)

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Existing Metadata</th>
<th>Published CHO(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>2</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>3</td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td>4</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
Overview

- Motivation
- Contribution
- Background technologies
- Functional specification
- The ESE-CHO Application Profile
- System Architecture
- GUI Design Specification
- Implementation
- Evaluation
- Related Work
- Conclusion
ESE-CHO Application Profile

 A superset of Europeana Semantic Elements (ESE)

 The ESE-CHO Application Profile consists of the following parts:
   Basic information (e.g. title, creator)
   Life Cycle information (e.g. date created)
   Technical information (e.g. medium, extent)
   Relation information (e.g. is part of)
   Collection information (e.g. title, subject)
   Europeana information (e.g. data provider, country)
Overview

- Motivation
- Contribution
- Background technologies
- Functional specification
- The ESE-CHO Application Profile
- System Architecture
- GUI Design Specification
- Implementation
- Evaluation
- Conclusion
EuroMuse has been developed as a RIA.

RIAs require complex code on the client side to handle user interaction and other operations.
Client Side Logic

- Resides on the user's web browser
- Interaction with the users.
- Presentation of the information.
- Communication with the server.

Design Patterns:
- Model View Presenter
- Observer
- Mediator
System Architecture

Server Side Logic

- Resides on the web server hosting the system.
- Follows a **multi-layered architectural pattern**.
  - **Strict interaction** between layers
- 3 layers
  - Service
  - Business Logic
  - Data
Vocabulary Access Module

- Responsible for the **indexing** and **accessing** of vocabularies and authority files
  - taxonomic classification
  - publicly sourced authority files of person/places

- DCMI-Type Vocabulary

- MARC Relators (Library of Congress)
  - Relation between a name and a bibliographic resource

- EuroMuse currently uses the taxonomic classification of **Catalogue of Life**
  - started in 2001 by Species 2000 and ITIS
  - comprehensive catalogue of all known species of organisms on Earth
  - 99 taxonomic databases, 3000 specialists, 1.4 million species
Vocabulary Access Module

- Published the full database in **RDF** format (D2R Server)
  - compliant to SKOS
  - allows the semantic linkage of the CHOs to the taxonomic classification nodes
Concurrent Control Module

- Responsible for the concurrency of the application

- Why do we need concurrency?
  - EuroMuse allows multiple users to interact with the system at the same time.
  - Avoid issues when multiple users try to update the same data.

- The **Concurrency Control** mechanism is based on pessimistic.

- Introduced view/edit mode.
  - In view mode data is not locked and no changes can be made
  - In edit mode data must be locked

- Automatic lock expiration and lock refreshing
Overview

- Motivation
- Contribution
- Background technologies
- Functional specification
- The ESE-CHO Application Profile
- System Architecture
- **GUI Design Specification**
- Implementation
- Evaluation
- Conclusion
GUI Design Specification

- Main GUI Specification
GUI Design Specification

- GUI Specification of CHO Metadata

![GUI Design Diagram]
Overview

- Motivation
- Contribution
- Background technologies
- Functional specification
- The ESE-CHO Application Profile
- System Architecture
- GUI Design Specification
- Implementation
- Evaluation
- Conclusion
Object: Salmo trutta (TNHM fish images) [VIEW MODE]

Title
Salmo trutta

Alternative Title
sea trout
Language English
Meriforell
Language Eesti

Description
Sea trout.
Language English
Meriforell on lõhelaste sugukonda kuuluv kalaliik. Elupaik: jõgi.
Language Eesti

Subject
animalia
chordata
actinopterygii
Implementation

- http://147.27.41.103/music/mmat

System requirements
- Servlet Container (Tomcat, etc.)
- Imagemagick for image processing
- Xuggle for video processing
Overview

- Motivation
- Contribution
- Background technologies
- Functional specification
- The ESE-CHO Application Profile
- System Architecture
- GUI Design Specification
- Implementation
- **Evaluation**
- Conclusion
Evaluation

- TUC-MUSIC (2011, March)
  - think aloud evaluation

- Athens (2011, March)
  - think aloud evaluation

- Natural Europe curators summer school, Chania (2011, June)
  - hands-on evaluation

- Lisbon (2011, September)
  - hands-on evaluation

- Constant testing from curators who already use EuroMuse for their CHOs.
Overview

- Motivation
- Contribution
- Background technologies
- Functional specification
- The ESE-CHO Application Profile
- System Architecture
- GUI Design Specification
- Implementation
- Evaluation
- Conclusion
Future Work

- Europeana Data Model (EDM)
  - Semantic expansion of the existing metadata
- Describing a museum object with multiple media objects
- User access rights per collection/museum department
- Client/Server side caching for optimization

Conclusion

DEMO