National Documentation Centre

Search Culture.gr

Open access in Cultural Heritage digital content

Haris Georgiadis, Ph.D.
Senior Software Engineer | EKT
hgeorgiadis@ekt.gr
The beginning..

- **42 institutions** documented & digitalized items of Cultural Heritage & created on-line digital repositories
  - Co-funding projects by Greece and EU structural funds
- **The role of EKT with respect to this content was to**
  - collect it
  - validate and certify it
  - enriched it semantically
  - publish it in searchculture.gr
  - preserve it
The ecosystem architecture

- **Harvester** (Java) - http://harvester.ekt.gr
- **Validator** (Java) - http://validator.ekt.gr
- **Aggregator** (Java) - http://searchculture.gr
- **Semantics** (Java) - http://semantics.gr

- Persistence/Service Layer
  - Metadata Storage
  - Digital File Storage
  - Batch Harvesting Service
  - OAI-PMH (REST API)

- Web Layer
  - Management GUI
  - REST API

- Persistence/Service Layer
  - Validation Process Storage
  - Validation Result Parser
  - Validator Back-End Connector

- Web Layer
  - Management GUI

- Persistence/Service Layer
  - EDM Storage and Indexing Service
  - Batch Metadata Ingestion Service
  - Thumbnails Storage
  - OAI-PMH Service

- Web Layer
  - Management WEB GUI
  - OAI-PMH Service
  - Public Portal

- Persistence/Service Layer
  - Organizations and Collections Storage
  - Organizations and Collections Management WEB GUI

- Web Layer
  - Management WEB UI
  - Public portal
  - REST API

- Persistence/Service Layer
  - OAI-PMH Providers/Repositories
  - OAI-PMH (REST API)

- Web Layer
  - Management GUI

- Persistence/Service Layer
  - Validation Process Storage
  - Validation Result Parser
  - Validator Back-End Connector

- Web Layer
  - Management GUI

**harvester, aggregator (searchculture.gr), validator (back-end & front-end), system for enrichment (semantics.gr)**
Open access in Cultural Heritage digital content

Search Culture.gr

Explore 157959 items of a rich variety of types* from 42 institutions...

*The classification of the content against a vocabulary of types was held by EKT
Open access in Cultural Heritage digital content

150,000+ items | 53 collections | 42 institutions
Features

- **Aggregates digital content** from registered digital repositories
- **Public portal**
  - Modern & effective search engine
    - facet queries
    - thumbnails
  - Dissemination as Open Linked Data (Open Linked Data)
  - Multilingual (English / Greek)
  - Responsive design
  - Rights statements per item, mostly CC, defining copyright, accessibility and reusability status
- Presents the participating institutions & their collections
- Supports hierarchical navigation
- **User friendly administration web GUI**
  for easy management of providers and ingestion workflows
Challenges for aggregators

Improve searchability

- Multiple ways for discovering content
- Multilingual search / navigation
- Advanced search / filers on results
- Homogenization and classification of content

Increase the size of the aggregated content

- Automatic validation
- User friendly administration environment
- Automatic ingestion workflows
- Automatic enrichment/classification
SearchCulture.gr deals with this problem using **two strategies:**

- EKT has set specific requirements against which it validates and certifies the aggregated content using **Validator**
- EKT using **Semantics.gr** enriches and homogenizes the aggregated content with respect to **dc:type** against a hierarchical bilingual vocabulary of types.

Metadata records are enriched with a separated field ‘**EKT Type**’ that holds references to terms of the vocabulary.
1. Key-based search

2. Narrower search in specific fields

3. Multiple key-based criteria

4. Choose collection

5. Choose EKT type

6. Choose license for digital content
Search Results - facet filtering
ΕΚΤ τύπος: Ζωγραφική (EL)
Painting (EN)

Τίτλος: Κεφάλι νίκου

Σκηνοθέτης: Σαράουχης Πάνης

Περιγραφή: Ηνωμένη αριθμού Π.6179Π.6179
Εθνική Πινακοθήκη - Μουσείο Αλέξανδρου Σούτζου
Λάθος σε ξίφος

Τύπος: Ζωγραφική

Περιγραφή: Προσωπογραφία Μορφή

Προμηθευτής: Εθνική Πινακοθήκη - Μουσείο Αλέξανδρου Σούτζου
National Gallery – Alexandros Soutzos Museum (EN)

Πηγή: Εθνική Πινακοθήκη - Μουσείο Αλέξανδρου Σούτζου
National Gallery – Alexandros Soutzos Museum (EN)

Σχεδίαση: εικόνα

Δημιουργικότητα: 2006-12-19T00:00:00+02:00

Ημερομηνία: 1941

Ευρωπαϊκός τύπος: εικόνα

Μορφή: 35,5 x 24,5 εκ.

Γλώσσα: Αδιάλειπτη λεξικοποίηση πληροφοριών
Institutions and collections

42 institutions contributed with 56 collections / repositories in total

Ephorate of Antiquities of Chania – Archaeological Museum of Chania
http://www.chaniamuseum.culture.gr/
The Archaeological Museum of Chania is subject to the Ephorate of Antiquities of Chania. It was founded in the early 20th century. Its purpose is the preservation, protection and promotion of cultural heritage of the area and includes exhibits from the Neolithic era (4th millennium BC) up to the Roman period (3rd century AD).

Repository
Repository of Archaeological Museum of Chania – Digitized items database
http://chaniamuseum.culture.gr
The Portal of the Archaeological Museum of Chania provides access to digital collections and particularly to 614 items of the permanent exhibition distributed in 18 individual categories according to their type (Vessels - utensils, sculptures - reliefs, mosaics, inscriptions, coins, jewelry - gems, Others, figurines, miniature, clay tablets - seals, weapons - tools). In addition, eight of the above categories are presenting their objects in 3D application. Each object has been scientifically documented and the information concerning it are contained (data source, description, date, bibliographic reference). There are also photos or 3D presentations that allow an interactive rotation of the object. Finally, in the portal are educational programs, an entertainment platform, cultural routes and information about the opening hours and the activities of the Museum.

Baron Michael Tossizza Foundation
http://metsovomuseum.gr/
THE BARON MICHAEL TOSSELLA FOUNDATION was founded in 1947 with the aim of maintaining Metsovo's cultural heritage and sustaining its economic growth. Today, the Foundation works to achieve the same goals as in the past, adjusted to Metsovo's modern needs and conditions. The foundation remains focused on the Tossizza Cheese Factory, the Ski Resort, the Metsovo Folk Art Museum, Aghios Nikolaos Monastery and the Student House in Athens for students from Europe, as well as on many recent initiatives, such as the Children’s Library, the
EKT enriches and homogenizes the aggregated content against a hierarchical bilingual Vocabulary of Types that has created. Metadata records are enriched with a separated field ‘EKT Type’ that holds references to terms of the vocabulary. Therefore, the content of SearchCulture.gr is been classified to a compact and balanced set of types which gives it semantic cohesion and improves its searchability, presentation and multilingualism.

### 2D Graphics (77914)
- Banknote (7)
- Chart (1)
- Clipping (1)
- Collage (1)
- Drawing/Sketch (8008)
  - Architectural drawing (1126)
  - Technical drawing (575)
- Greetings card (40)
- Identifying marking (732)
- Images and ornament for layout features (2896)
- Master (2344)
- Mosaic (906)
- Mural painting (720)
- Painting (10862)

### 3D Artefacts and Realia (44370)
- Album (19)
- Archaeological object (23861)
- Archaeological site (1483)
- Architectural element (2272)
- Arm (475)
- Artifact (1)
- Bone (238)
- Building (34)
- Ceramic ware (3765)
- Coin (2128)
- Costume (mode of fashion) (6596)
- Disk (51)
- Doll (364)
- Fragment (2044)

### Text (34693)
- Announcement (481)
- Article (2580)
- Biography (9)
- Book (2516)
- Book page (1284)
- Catalogue (110)

### Sound Recordings (882)
- Sounds (15)

### Projected and Video Material (804)
The problems before enrichment

- Some types are in Greek others in English
- Some types are in capitals
- Some types are synonyms with each other
- Some providers give very general types some very detailed

- Low searchability
- Narrow options for navigation and discoverability
- Poor presentation
• A platform for establishing vocabularies and thesauri
  – public portal
  – dissemination as Open Linked Data
• Tool for semantic enrichment
  – supports automatic classification of metadata items against vocabulary terms
  – uses information retrieval and self-improving techniques
  – can be used by repositories and aggregators for controlled homogenization & enrichment of their content
<table>
<thead>
<tr>
<th>Category</th>
<th>Semantic class</th>
<th>Creator</th>
<th>Default language</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concepts</td>
<td>skos:Concept</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>National Documentation Center</td>
<td></td>
</tr>
</tbody>
</table>

### Item types by EKT

<table>
<thead>
<tr>
<th>Content</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Διαδικτυακά γραφικά</td>
<td>2D Graphics</td>
<td></td>
<td>RDF</td>
</tr>
<tr>
<td>Άλλο</td>
<td>Other</td>
<td></td>
<td>RDF</td>
</tr>
<tr>
<td>Αντίγραφο</td>
<td>Replica</td>
<td></td>
<td>RDF</td>
</tr>
<tr>
<td>Απόκομμα</td>
<td>Clipping</td>
<td></td>
<td>RDF</td>
</tr>
<tr>
<td>Αυτοκόλλητη ητικέτα</td>
<td>Sticker</td>
<td></td>
<td>RDF</td>
</tr>
<tr>
<td>Αφιερωματικές κάρτες Αγίων</td>
<td>Santino</td>
<td></td>
<td>RDF</td>
</tr>
<tr>
<td>Αφίσα</td>
<td>Poster</td>
<td></td>
<td>RDF</td>
</tr>
<tr>
<td>Γραμματόσημο</td>
<td>Stamp</td>
<td></td>
<td>RDF</td>
</tr>
<tr>
<td>Διάγραμμα</td>
<td>Chart</td>
<td></td>
<td>RDF</td>
</tr>
<tr>
<td>Διακόσμηση εγγράφου</td>
<td>Images and ornament for layout features</td>
<td></td>
<td>RDF</td>
</tr>
<tr>
<td>Εικόνα</td>
<td>Picture</td>
<td></td>
<td>RDF</td>
</tr>
<tr>
<td>Εκτύπωση</td>
<td>Print</td>
<td></td>
<td>RDF</td>
</tr>
<tr>
<td>Έργο θρησκευτικής τέχνης</td>
<td>Religious visual works</td>
<td></td>
<td>RDF</td>
</tr>
<tr>
<td>Φορητή Εικόνα</td>
<td>Icon</td>
<td></td>
<td>RDF</td>
</tr>
<tr>
<td>Ευχετήρια κάρτα</td>
<td>Greetings card</td>
<td></td>
<td>RDF</td>
</tr>
</tbody>
</table>
• **SearchCulture.gr collections are registered** in Semantics.gr

• Per collection, for metadata field **dc:type**, the tool allows the user to map **all the distinct values** to terms of the EKT Types Vocabulary
  
  • automatic OAI-PMH **harvesting** of distinct values
  
  • **automatic mapping** attempt based on string matching
  
  • a **secondary field** may be used for more precise mapping
  
  • The tool improves the automatic suggestion effectiveness based on previous experience
  
  • After the mapping is done, **SearchCulture.gr** enriches the collection according to the mapping rules.
Real example (1)

Homogenization based on **dc:type** and **dc:subject** (secondary field for filtering)
Real example (2)

Homogenization based on **dc:type** and **dc:title** (secondary field for filtering)
Improved searchability & search multilingualism
New ways for navigation → improved **discoverability** of content
Thank you!

Learn more

helios-eie.ekt.gr/EIE/handle/10442/15324