Requirements for Visualization and Indexing of Museum Content

Regine Stein
Philipps-Universität Marburg – Bildarchiv Foto Marburg (UNIMAR)
1. Current state in Europeana
   - Visualization
   - Indexing

2. What is needed: a bit of theory

3. Current approach in Athena Plus

4. Museum-specific requirements
1. Current state in Europeana
   – Visualization
   – Indexing

2. What is needed: a bit of theory

3. Current approach in Athena Plus

4. Museum-specific requirements
Villa Dürckheim (Äußeres - Nordfassade von Norden)

Alternative Title:
Palais Dürckheim

Creator:
Henry van de Velde, van de Velde, Henry, Henry Clemens van de Velde (Henry van de Velde, van de Velde, Henry; Henry Clemens van de Velde) To: 1957-10-25+02:00 Date of birth: 1957-10-25+02:00 Date of death: 1957-10-25+02:00, Veide, Henry van de (Veide, Henry van de)

Contributor:
Dürckheim-Montmartin, Friedrich von (Graf)

Date:
1928; Part of: 2 quarter of the 20th century; From: 01-01-1928 — To: 31-12-1928; 1936; Part of: Mid 20th century; From: 01-01-1936 — To: 31-12-1936; 1945; Part of: Mid 20th century; From: 01-01-1945 — To: 31-12-1945; 2000/2007

Date of creation:
1912/1913

Type:
Wohn- und Bürohaus; Architektur (Bauwerk); Architecture; Art Nouveau; arkitektur; Baukunst (Objektgattung); Art Nouveau

Format:
verputzt

Identifier:
obj/32057917-fmd471220

Rights:
Foto: Gaasch, Uwe

Provenance:
Weimar, Cranachstraße 47

http://www.europeana.eu/portal/record/2026116/Partage_Plus_ProvidedCHO_Bildarchiv_Foto_Marburg_obj_32057917_fmd471220.html
**Villa Dürckheim (Äußeres - Nordfassade von Norden)**

**Alternative Title:**
Palais Dürckheim

**Creator:**
Henry van de Velde, van de Velde, Henry, Henry Clemens van de Velde (Henry van de Velde; van de Velde, Henry; Henry Clemens van de Velde)

Date of birth: 1957-10-25+02:00 Date of death: 1957-10-25+02:00, Velde, Henry van de (Velde, Henry van de)

**Contributor:**
Dürckheim-Montmartin, Friedrich von (Graf)

**Date:**
1928; Part of: 2 quarter of the 20th century, From: 01-01-1928 — To: 31-12-1928; 1936; Part of: Mid 20th century, From: 01-01-1935 — To: 31-12-1935; 1936; Part of: Mid 20th century, From: 01-01-1936 — To: 31-12-1936; 1945; Part of: Mid 20th century, From: 01-01-1945 — To: 31-12-1945; 2000/2007

**Type:**
Wohn- und Bürohaus; Architektur (Bauwerk); Architecture; Art Nouveau; arkitektur; Baukunst (Objektgattung); Art Nouveau

**Format:**
verputzt

**Identifier:**
obj/32057917-fmd471220

**Rights:**
Foto: Gaasch, Uwe

**Provenance:**
Weimar, Cranachstraße 47

http://www.europeana.eu/portal/record/2026116/Partage_Plus_ProvidedCHO_Bildarchiv_Foto_Marburg_obj_32057917_fmd471220.html
Issues

• Visualization
  – Display of ALL information – no distinction between indexing and display
  – No distinction between information coming from the content provider and automatically enriched information
  – No possibility to display uncertainty
  – No possibility to transmit type of contribution, roles, extent

• Indexing
  – Only “administrative” facets available
  – No browsing aids for the user in terms of content
1. Current state in Europeana
   - Visualization
   - Indexing

2. What is needed: a bit of theory

3. Current approach in Athena Plus

4. Museum-specific requirements
"Give end-users consistent, reliable access to information" (CDWA)

Users want to retrieve

- **all** documents in an information system that match the search request, without any information loss (recall)
- **only** documents in an information system that match the search request, without any unwanted results (precision)

Recall and precision are measures for the quality of information retrieval.

The quality of "retrieval of documents relies heavily on the quality of their representation. If the documents are represented poorly or inadequately, the quality of the searching will likewise be poor". Mai 2000, 270

Precision is getting more and more important in the context of large-scale aggregation of metadata.
Factors which affect recall and precision

- Recall is decreased by
  - synonyms occurring in natural language
    => use controlled vocabularies employing synonym control
  - low exhaustivity – too few indexing terms assigned
    => consider rules or procedures for more in-depth indexing
  - inconsistent indexing – dissimilar indexing of similar objects
    => consider supervised training of indexers or machine aided feedback

- Precision is counteracted by
  - homonyms occurring in natural language
    => use controlled vocabularies employing homonym control
  - low indexing specificity – specific indexing terms are missing
    => follow the basic rule of indexing the most specific term
  - broadly related descriptors as indexing terms
    => apply the most specific term, do not add broader terms from a hierarchy chain
A more recent challenge is to harmonise collection-specific rules with the requirements arising from large-scale aggregation of metadata.

One general indexing principle is true for all kinds of documents:

"The single most important principle of subject indexing [...] is that a topic should be indexed under the most specific term that entirely covers it. Thus, an article discussing the cultivation of oranges should be indexed under *oranges* rather than under *citrus fruits* or *fruit.*" Stock, 350

This main principle applies to all kinds of "documents" and to values of all types of descriptive metadata, such as

- the LIDO element *Object/Work Type*
- and ALL LIDO elements referring to concepts.
Indexing the specific term presupposes broader terms to be automatically added as search entries (not as indexing terms!) to assure retrieval of the specifically indexed item when using the broader term for search.
1. Current state in Europeana
   - Visualization
   - Indexing

2. What is needed: a bit of theory

3. Current approach in Athena Plus

4. Museum-specific requirements
Core Elements for the Delivery of Metadata to Europeana

*Mandatory elements*

- Metadata language
- Europeana type
- Object / Work Type
- Object Title
- Record Information:
  - Record ID - Record Type - Record Source - Record Info link
- Resource Information / Visual Surrogates
  - Link DCHO (Master) - Link DCHO (Thumbnail) - Rights Resource
1. Current state in Europeana
   - Visualization
   - Indexing

2. What is needed: a bit of theory

3. Current approach in Athena Plus

4. Museum-specific requirements
### Overview of the LIDO Metadata Elements

<table>
<thead>
<tr>
<th>Mandatory</th>
<th>Mandatory if applicable</th>
<th>Recommended</th>
</tr>
</thead>
</table>
| **Mandatory elements**

*These elements are mandatory in each metadata record and the information is used for browsing the objects in a museum-specific portal.*

- **Metadata language**
- **Object / Work Type**
- **Classification**
- **Object Title**
- **Location**
- **Event Information**
- **Record Information**
- **Resource Set**

| **Mandatory if applicable elements**

*These elements are mandatory if applicable for the described object and the information is used for browsing the objects in a museum-specific portal.*

- **Repository Name**
- **Event Information ->**
  - **Actor**
  - **Cultural Context**
  - **Date**
  - **Period**
  - **Place**
  - **Material / Technique**
- **Subject Information ->**
  - **Theme-Concept**
  - **Actor**
  - **Date**
  - **Place**

| **Recommended elements**

*These elements are recommended if applicable and the information is used for displaying the object information in a museum-specific portal.*

- **Inventory Number**
- **Object Description**
- **Object Measurement**
- **Event Information ->**
  - **Event Name**
  - **Actor->Role**
  - **Actor->Attribution**
  - **Actor->Extent**
  - **Method**
  - **Material/Technique->Extent**
- **Subject Information ->**
  - **Extent Subject**
- **Related Works**
- **Rights Work**
- **Resource Set ->**
  - **Date Taken**
  - **Creator of Resource**
  - **Resource Description**
More specified vocabulary links!

Use Getty vocabularies wherever possible – most appropriate, most elaborate and most sustainable source!

- AAT for
  - Object / Work Type and CH type classification
  - Cultural context
  - Periods
  - Materials and Techniques

- ULAN (second best: VIAF) for
  - Repository Name
  - Event Actor and Subject Actor

- TGN for
  - Repository Location
  - Event Place and Subject Place
Requirements

Mandatory vocabulary links!!

- Cultural Heritage Type: AAT Facet Codes
- Thematic Context: UDC Summary
- Repository Location: TGN nations for the “part of” relation
- Event Type: LIDO Event types
- Event Place: TGN nations for the “part of” relation
### Cultural Heritage Type: AAT Facet Codes

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>V.RD</td>
<td>Settlements and Landscapes</td>
</tr>
<tr>
<td>V.RG</td>
<td>Built Complexes and Districts</td>
</tr>
<tr>
<td>V.RK</td>
<td>Single Built Works</td>
</tr>
<tr>
<td>V.RM</td>
<td>Open Spaces and Site Elements</td>
</tr>
</tbody>
</table>

**Built Environment**

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>V.TC</td>
<td>Furnishings</td>
</tr>
<tr>
<td>V.TE</td>
<td>Costume</td>
</tr>
<tr>
<td>V.TH</td>
<td>Tools and Equipment</td>
</tr>
<tr>
<td>V.TK</td>
<td>Weapons and Ammunition</td>
</tr>
<tr>
<td>V.TN</td>
<td>Measuring Devices</td>
</tr>
<tr>
<td>V.TQ</td>
<td>Containers</td>
</tr>
<tr>
<td>V.TT</td>
<td>Sound Devices</td>
</tr>
<tr>
<td>V.TV</td>
<td>Recreational Artifacts</td>
</tr>
<tr>
<td>V.TX</td>
<td>Transportation Vehicles</td>
</tr>
</tbody>
</table>

**Visual and Verbal Communication**

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>V.VC</td>
<td>Visual Works</td>
</tr>
<tr>
<td>V.VK</td>
<td>Exchange Media</td>
</tr>
<tr>
<td>V.VW</td>
<td>Information Forms</td>
</tr>
</tbody>
</table>
Requirements

Mandatory vocabulary links!!

• Thematic Context: UDC Summary
  -> http://udcdata.info/064676

<table>
<thead>
<tr>
<th>UDC Summary Linked Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>URI</td>
</tr>
<tr>
<td>Notation</td>
</tr>
<tr>
<td>Caption</td>
</tr>
<tr>
<td>Application Note</td>
</tr>
<tr>
<td>Broader class</td>
</tr>
<tr>
<td>Narrower classes</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
Further functionalies

- Provide appropriate grouping and order of information
- Display actor roles, attribution and uncertainty of information
- Allow for display values so that providers can take control how their items are displayed, particularly use the following display elements as they include information about uncertainty:
  - objectMeasurementsSet -> displayObjectMeasurements
  - eventActor -> displayActorInRole
  - eventDate -> displayDate
  - eventPlace -> displayPlace
  - eventMaterialsTech -> displayMaterialsTech
Further functionalities

- Allow provider to set order (-> @lido:sortorder) so that providers can e.g. control order of appearance of artists – very important for art museums.
- Display events according to an ordering on event type vocabulary
- Allow for multiple digital resources with related information: description, creator (photographer), date taken, rights
- Practical Mapping Exercise -