Dances with the AAT: Moving Vocabularies into the Semantic Web

Sophy S. J. Chen, Lu-Yen Lu
Academia Sinica Digital Center

International Terminology Working Group (ITWG) Meeting
Los Angeles, USA. 11-18 Aug. 2017
Working Status

– Updated status of AAT-TW Translation
– The Issues

LOD Projects

– TaiUC LOD Project with Getty Vocabularies
– Semantic Annotation of the Photo Collection with Getty Vocabularies
– Digital Humanities Project with Getty Vocabularies: Buddhist Art History
Works on AAT-Taiwan in 2016/ 2017  
(Sep. 2016- Aug. 2017)  

Contribution of 5,000 translated Concepts

• Verification and Contribution of 5,000 records to AAT (Apr.-Aug. 2017). Now 20,359 records to Getty (57% from the whole 35,485 records).

Example: AAT 3000009966 fleur-de-lis (motif)

Note: A decorative motif in the form of a stylized iris or lily, usually with three curled petals, but occasionally with five, one of which forms the central axis to the symmetrical design. The design is based on an iris, which was equated with the lily and had the same symbolic references as the lily; it probably developed from the honeysuckle design motif.

Terms:
- fleur-de-lis (motif) (preferred, C,U,English-P,D,U,SN)
- fleurs-de-lis (motifs) (C,U,LC,English,AD,U,PN)
- fleurs-de-lys (motifs) (C,U,English,UF,U,U)
- fleur de lis (motif) (C,U,English,UF,U,U)
- (motif) (Dutch)

Franse lelie (C,U,Dutch,UB)
Franse lelies (C,U,Dutch,UB)
heraldische lelie (C,U,English)
heraldische lelies (C,U,English)
oflores de lis (C,U,Spanish,UB)
flor de lis (C,U,Spanish-UB)

詞彙：
- 鹿尾花紋樣 (紋飾) (P., U.繁體中文-偏好, D,U,U)
- 百合花飾 (U.繁體中文-非偏好, UF,U,U)
- 鹿尾 (U.繁體中文-非偏好, UF,U,U)
- 燕子花 (U.繁體中文-非偏好, UF,U,U)
- 藍蝴蝶 (U.繁體中文-非偏好, UF,U,U)
- 鹿尾花形紋章 (U.繁體中文-非偏好, UF,U,U)
- yuān wei hua wen yang (P., U.韻氏拼音-偏好, UF,U,U)
- yuān wèi huǎ wén yāng (P., U.漢語拼音-偏好, UF,U,U)
- yuan wei hua wen yang (P., U.漢語拼音(無聲調)-偏好, UF,U,U)
Works on AAT-Taiwan in 2016/2017
(Sep. 2016-Aug. 2017)
The Facet Distributions of AAT-Taiwan Contribution

- Details of the AAT-TW contribution

### Facets

<table>
<thead>
<tr>
<th>Facets</th>
<th>1-7th contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concepts</td>
<td>19,189</td>
</tr>
<tr>
<td>Guide Terms</td>
<td>1,159</td>
</tr>
<tr>
<td>Hierarchy Names</td>
<td>11</td>
</tr>
<tr>
<td>Total</td>
<td>20,359</td>
</tr>
</tbody>
</table>

### Record Type

<table>
<thead>
<tr>
<th>Record Type</th>
<th>Records</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concepts</td>
<td>19,189</td>
<td>94.2%</td>
</tr>
<tr>
<td>Guide Terms</td>
<td>1,159</td>
<td>5.7%</td>
</tr>
<tr>
<td>Hierarchy Names</td>
<td>11</td>
<td>0.1%</td>
</tr>
<tr>
<td>Total</td>
<td>20,359</td>
<td>100%</td>
</tr>
</tbody>
</table>
Works on AAT-Taiwan in 2016/ 2017
(Sep. 2016- Aug. 2017)

3,000 selected AAT-TW concepts for developing the Faceted Thesaurus of Chinese Cultural Heritages (FTCCH)

- Cooperation with Dr. Claire Huang, University of Science and Technology Beijing, China (Jan.-Mar. 2017)

<table>
<thead>
<tr>
<th>Facets</th>
<th>Hierarchies</th>
<th>Total</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical Attributes</td>
<td>Design Elements</td>
<td>44</td>
<td>1.5%</td>
</tr>
<tr>
<td>Styles and Periods</td>
<td>Styles and Periods</td>
<td>145</td>
<td>5.1%</td>
</tr>
<tr>
<td>Activities</td>
<td>Processes and Techniques</td>
<td>151</td>
<td>5.4%</td>
</tr>
<tr>
<td>Materials</td>
<td>Materials</td>
<td>234</td>
<td>8.5%</td>
</tr>
<tr>
<td>Objects</td>
<td>Objects Groupings and Systems</td>
<td>76</td>
<td>2.6%</td>
</tr>
<tr>
<td></td>
<td>Objects Genres</td>
<td>137</td>
<td>4.8%</td>
</tr>
<tr>
<td></td>
<td>Components</td>
<td>428</td>
<td>15.0%</td>
</tr>
<tr>
<td></td>
<td>Built Environment</td>
<td>363</td>
<td>12.8%</td>
</tr>
<tr>
<td></td>
<td>Furnishing and Equipment</td>
<td>854</td>
<td>30.0%</td>
</tr>
<tr>
<td></td>
<td>Visual and Verbal Communication</td>
<td>407</td>
<td>14.3%</td>
</tr>
</tbody>
</table>
What we’ve supported:
1. Subject ID
2. Record Type
3. Preferred Name
4. Scope Note
5. Non-preferred Name
6. Qualifier
7. Facet Code
Problem by working

- 20,359 records contributed to Getty since 2012
- List of 7,075 updated AAT records in “Download Center for AAT”
- Updated in what aspect? Refined, new added, revised
- Solution: indicating the type of renewing in the scope note of the updated AAT records.

![Excel Table Example](image)

Remarks example:
- Refined
- New added
- Revised

In order to determine changes to records between last year’s release and this current release, please consult the Revision History and the Changes spreadsheet (you may be prompted to reenter your password).
Problem by working

- Case 1: Original scope notes changed after contribution
  
  Example: AAT 300020110 Corinthian (architectural style)

**Note:** Refers to the third of both the three Greek architectural orders and the later five traditional classical orders of architecture that, with Doric, Ionic, Tuscan, and Composite, was used by the Romans and through the Renaissance and beyond. It seems to have been influenced by Egyptian architecture, though it probably developed in Greece in interior architecture, and was used in exterior architecture by the third century BCE. In Greek architecture it is characterized by a form that is lighter and more ornate than Doric or Ionic, a bell-shaped capital with acanthus stalks emerging to support graceful volutes, and a column that is seated on a base and usually fluted. It was the most common of Greek styles in Roman architecture, and in Roman and later architecture the style was often modified. For the assemblage forming associated columns, use "Corinthian order."

**Additional Notes:**

Chinese (traditional) ..... 指稱三種希臘建築柱式的第三式，亦為之後五種傳統古典建築柱式的第三式，此柱式與其他古典柱式包括多利克柱式(Doric)、愛奧尼亞柱式(Ionic)、托斯坎/托斯卡納柱式(Tuscan)，以及複合柱式(Composite)，同為羅馬人及稍晚文藝復興與之後的時期所沿用。雖然此柱式似乎受到埃及式建築的影響，但其應該發源於希臘建築的內部結構中，並在西元前3世紀前後，開始運用在建築外部結構上。在希臘的建築當中，此柱式的特徵為較多利克柱式與托斯坎/托斯卡納柱式更為輕盈且裝飾更為繁複的形式，其柱頭呈鐘狀，並以其若葉形狀的葉梗支撐優雅的卷曲柱飾，其柱子則固定在底座上，柱身刻有凹槽。這是羅馬式建築所採用的許多希臘風格中最為普遍的一種，且此風格在羅馬式建築與之後的建築風格當中經常被修改。此柱式與「科林斯柱式」(Corinthian order)有所不同，因為柱式柱式乃是嚴格指稱一種特定的系統，或是按照既定的規則與比例加以安排各部份位置及功能的建築集合體。
Problem by working

- Case 2: Scope note newly added after contribution

Example: **AAT 300017934 Araucanian (culture or style)**

Note: Culture and style of the South American Indians that are now concentrated in the fertile valleys and basins of south-central Chile, from the Biobío River in the north to the Toltén River in the south. Although the pre-Columbian Araucanians did not themselves recognize political or cultural unity above the village level, the Spanish distinguished three Araucanian populations geographically: the Picunche living in the north between the Choapa and Biobío rivers, the Mapuche inhabiting the middle valleys, and the Huilliche dwelling in the south between the Toltén River and Chiloé Island. At the time of the Spanish arrival in Chile, most of central Chile was settled by scattered populations of Mapuche farmers who grew corn (maize), beans, squash, potatoes, and other vegetables.

Terms:
- **Araucanian (culture or style)** (preferred, C,U,English-P,D,U,A)
- **Araucanos** (C,U,LC,English,UF,U,U)
- **阿剌卡尼亞** (U,Chinese (traditional)-P,D,U,U)
- **Araucano** (C,U,Dutch-P,D,U,U)
  (Spanish-P,D,U,U)
LOD Projects & Future Plans

* TaiUC LOD Project with Getty Vocabularies
* Semantic Annotation of the Photo Collection with Getty Vocabularies
* Digital Humanities Project with Getty Vocabularies: Buddhist Art Images
AAT-Taiwan application program 1

TaiUC LOD Project with AAT

Fishing in the Data Ocean: the TaiUC LOD Project

The Data Model: EDM (Europeana Data Model)
Red Bigeye Fish (*Pristigenys niphonia*) enriched using contextual entities
Enhancing the accessibility to data consumer

- Textual Data released under CC0; image released after request of the original rights holder
- Download: for entire LOD datasets in datahub (in N-triples); for single digitized item in TaiUC SPARQL endpoint (in JSON, RDF, turtles, N-triples)

Source: http://data.ascdc.tw/uc/bio/as/00426782

Download in: JSON | RDF | Turtle | N-Triples

https://datahub.io/zh_TW/dataset/asdc
AAT-Taiwan application program 2

**Semantic Annotation of the Photo Collection with AAT**

- Cyber Island ([http://cyberisland.teldap.tw/](http://cyberisland.teldap.tw/)), an UGC online platform for professional and hobbyist photographers to storage, display photos.
- We have graduate students to give semantic annotation of over 1,000 selected photos in the system. They are asked to tag the image based on visual experience and follow 5 types of annotation (agent, object, places, time, activity) which are based on linked data. We’ve collected over 4,000 tags.
- Then, we linked these thousand tags to 4 authority files. (AAT, DBpedia, TGN and GeoNames).

<table>
<thead>
<tr>
<th>type</th>
<th>properties</th>
<th>linked database</th>
<th>examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agent/ Person</td>
<td>dbpedia-owl:mainCharacter</td>
<td>AAT, DBpedia</td>
<td>AAT: 300025607 farmers</td>
</tr>
<tr>
<td>Objects</td>
<td>schema:temporalCoverage</td>
<td>AAT, DBpedia</td>
<td>AAT: 300132399 flowers</td>
</tr>
<tr>
<td>Places</td>
<td>bci:hasObject</td>
<td>TGN, GeoNames, AAT, DBpedia</td>
<td>TGN: 7488234 Nangang Dist. AAT: 300008632 farmland</td>
</tr>
<tr>
<td>Time/ Season</td>
<td>schema:contentLocation</td>
<td>AAT, DBpedia</td>
<td>AAT: 300343621 dawn</td>
</tr>
<tr>
<td>Activity/ Event</td>
<td>premis:hasEvent</td>
<td>AAT, DBpedia</td>
<td>AAY: 300053027 cleaning</td>
</tr>
</tbody>
</table>
Work flow

ASCDC’s content management system for photo images

plug in

ASCDC annotation system

input

Reusing of controlled vocabularies

DBpedia
GeoNames
TGN
AAT-TW

input

tagging, mapping to external authority resources

花/ AAT: 300132399 flowers

Tagging by types

Agent/Person

Time/Season

Objects: A_ 花 (flowers)

Place

Activity/Event

Getty Vocabularies: LOD

flowers (plant components)

Source: http://vocab.getty.edu/aat/300132399

Subject (123)  Predicate  Object  All

Statements in web

URI: AAT: 300132399
Image tagging combining with IIIF

Compatible to IIIF standard; uploading in IIIF image viewer
Provide Knowledge Graph of the Image Tags combining with IIIF

More info about the tag

Visualize Tags in AAT's structure

More related Photos
AAT-Taiwan application program 3
Digital Humanities Project with AAT: Buddhist Art Images

Buddhist steles
Buddhist caves
Buddhist stupas
3-layered metadata design

1. Item/Object
   - ID
   - Description
2. Iconographic Information
   - Depicted person
   - Depicted object
   - Depicted place
   - Gestures
   - Image source
   - Image theme
3. Hierarchical Relations
   - Parallel
   - Part to whole
   - Narrow-broader
   - Broader-narrow

Metadata/Item

Metadata/Iconography

Metadata/Hierarchical Relations

Work of Buddhist Art
Hierarchy under Buddhist Iconography as of 22 September 2016

Top of the IA hierarchies

Legend, Religion, Mythology
Buddhist iconography

+ Buddhist allegory, symbolism, theme
  Four Encounters (Buddhist theme)
  Four Noble Truths (Buddhist theme)
  Six Paths of Rebirth (Buddhist theme)

Buddhist characters
Amit$01abha (Mahayana deity)
Amit$01ayus (Amit$01ayus Buddha)
$01Ananda (Buddhist character)

+ Avalokiteshvara (Buddhist bodhisattva)
  Amoghapasa (Buddhist character)
Bishamoni (Buddhist character)

+ Buddha (primary character of Buddhism)
  Buddha Preaching
  Reclining Buddha
  Seated Buddha
  Standing Buddha

Concepts suggested for IA:
1. Jātaka/ 本生
2. Life of Buddha/ 佛傳
3. Nidāna/ 嫁緣
4. Avadana/ 譬喻

Concepts suggested for IA:
1. Bodhisattva/ 菩薩
2. Deva/ 天
3. Monk/ 僧人
Central Cave in Xiaonanhai Caves, China

Historical map of the Northern and Southern Dynasties, 5th-8th c.
Thanks for your attention