Methodological Considerations for Developing the *Art & Architecture Thesaurus* in Chinese and its Applications

Sophy Shu-Jiun CHEN
Institute of History and Philology, Academia Sinica, Taiwan

NKOS Consolidated Workshop 2020
2020. 09. 10
AAT in Chinese Language
Developed by the Academia Sinica Center for Digital Cultures (ASCDC) of Taiwan
http://aat.teldap.tw/

ITWG
The International Terminology Working Group

Pre-ITWG Meeting
1. 2008.07
2. 2009.07

ITWG Meeting
1. 2010.08
2. 2011.11
3. 2012.01
4. 2013.01
5. 2014.09
6. 2016.08
7. 2017.08
8. 2019.02
9. 2020.02
AAT as a Knowledge Tree

- AAT is a multilingual and hierarchical thesaurus structured in 7 facets and 21 hierarchies
- Current totals: 57,500 records (2016)
- Includes generic terms to describe art, architecture, design, artifacts, archival materials, archaeology, visual surrogates and other material culture
Record types in AAT: Concept

ID: 300015638
Page Link: http://vocab.getty.edu/page/aat/300015638

Terms:
- nature mortes (preferred: C. U. French: P. D. U. N.)
- naturaleza muerta (C. U. Spanish: P. D. U. N.)
- naturalezas muertes (C. U. Spanish: A. D. U. F. N.)
- stillleven (C. U. Dutch: P. D. U. N.)

Additional Notes:
Chinese (traditional): 材料 ...

Indication of term's historical usage:
- Used in use from ca. 1650
- Used from the early 18th century
- Used in use from the early 18th century in Spanish to refer to still lifes without figures
- From the Spanish term
- Early 18th-century French term

Core information in AAT concept:
(1) Preferred name
(2) Scope note
(3) Equivalence Relationships
(4) Hierarchical Relationships
(5) Associative Relationships
(6) Sources of concept term
Chinese-Traditional Version of AAT: The Purpose

- **Western Art Resource**
  - **Getty’s AAT**
  - Make Western Art Accessible to Western Users

- **Chinese Art Resource**
  - **AAT-Taiwan**
  - Mapping
  - Contributing AAT-Chinese
  - Make Chinese Art Accessible to Western Users

- **AAT-Taiwan**
  - Translating
  - Contributing Getty’s AAT
  - Make Western Art Accessible to Chinese Users

- **Digital Archives Thesaurus (DAT)**
  - Structuring DAT [Based on AAT]
  - Keywords from Taiwan Digital Archives
  - Make Chinese Art Accessible to Chinese Users

**AAT-TAIWAN website:** http://aat.teldap.tw/
Working Model of AAT-Taiwan

From English to Chinese
- Term Collection
- Semantic Analysis
- Equivalence Mapping
- Research on KOS Construction
- Concept Construction
- Translation
- Proofreading
- Disambiguation
- Expert Group

From Chinese to English
- Term Collection
- Semantic Analysis
- Equivalence Mapping
- Research on KOS Construction
- Concept Construction
- Translation
- Proofreading
- Disambiguation
- Expert Group
From English to Chinese

Make Western Art Accessible to Chinese Users
From Chinese to English

Make Chinese Art Accessible to Western Users
Conceptual Mapping of the NPM Terms to AAT

- The study takes a total of 2100 Chinese controlled vocabulary terms from the National Palace Museum in Taiwan, and a total of 35,314 English concepts of AAT, and performing semantic mapping between Chinese and English terms. The study asked research questions including:

  A. To what degree of semantic interoperability can be achieved between the two controlled vocabularies?

  B. What patterns of interoperability are involved in the field of Chinese art in the bilingual conceptual structure?

Example of mapping sheet between NPM and AAT terms
# Results of Mapping Types Between NPM and AAT

<table>
<thead>
<tr>
<th>Code</th>
<th>Mapping Type</th>
<th>Quantity (percentage)</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>BM</td>
<td>Hierarchical (narrower to broader)</td>
<td>1327 (63.19%)</td>
<td>剔犀 (ti xi) BM lacquering (A12067)</td>
</tr>
<tr>
<td>=EQ</td>
<td>Exact simple equivalence</td>
<td>638 (30.38%)</td>
<td>仰韶文化 (yang shao wen hua) =EQ Yangshao (A02002)</td>
</tr>
<tr>
<td>EQ</td>
<td>Cumulative compound equivalence</td>
<td>52 (2.48%)</td>
<td>塔(ta) EQ pagodas</td>
</tr>
<tr>
<td>~EQ</td>
<td>Inexact simple equivalence</td>
<td>37 (1.76%)</td>
<td>后妃 (hou fei) ~EQ empresses (P04003)</td>
</tr>
<tr>
<td>EQ+</td>
<td>Intersecting compound equivalence</td>
<td>33 (1.57%)</td>
<td>春景 (chun jing) EQ springs + scenes (P03001)</td>
</tr>
<tr>
<td>NM</td>
<td>Hierarchical (broader to narrower)</td>
<td>9 (0.43%)</td>
<td>玉石器 (yu shi qi) NM jades (objects) (A04004)</td>
</tr>
<tr>
<td>RM</td>
<td>Associative mapping</td>
<td>4 (0.19%)</td>
<td>臺閣 (tai ge) RM pavilions (P14002)</td>
</tr>
<tr>
<td>總計</td>
<td></td>
<td>2100 (100%)</td>
<td></td>
</tr>
</tbody>
</table>
Interoperability of Conceptual Structures for NPM in AAT

<table>
<thead>
<tr>
<th>Vocabularies</th>
<th>Structures</th>
<th>Specific category in AAT</th>
<th>No specific category in AAT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Completely covered in AAT</strong></td>
<td>Pattern 1 A02 考古學文化 (13 values) Archaeological Culture A17 時代 (66 values) Chinese Periods</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Incomplete covered in AAT</strong></td>
<td>Pattern 2 A03 窯系 (19 values) Chinese Ceramic Style</td>
<td>Pattern 3 A08 釉色 (69 values) Chinese Ceramic Glazes P02 中國繪畫技法 (34 values) Techniques of Chinese Painting</td>
<td></td>
</tr>
<tr>
<td><strong>None covered in AAT</strong></td>
<td></td>
<td></td>
<td>Pattern 4 P17 傳説動物 (7 values) fabulous Animals P18 中國節令 (8 values) Chinese Festivals</td>
</tr>
</tbody>
</table>

**Conceptual structures for the Chinese art concepts covered in AAT - Four Patterns**

1. The values listed in a unit of the NPM are completely covered by the corresponding AAT category.
2. The values listed in a unit in NPM are not completely covered by that AAT category.
3. The values in a unit in NPM are across in various places in AAT.
4. No specific category and equivalent concept in AAT for the Chinese art concepts.
The *Concept* Issues

- **Chinese-Specific Culture**
  - Create a Whole Set of Terms (Chinese Scripts)
  - **Re-Group Terms (Chinese Bronze Vessels)**
  - Term Disambiguation (Chinese Festivals)
    - Lantern Festival

- **Western-Specific Culture**
  - The Loanword Principle
  - Term Disambiguation (Translated Terms)
    - Principle 1. Add Qualifiers to Chinese Terms (The same terms exist when translated to Chinese). Chinese qualifiers are added according to the priority rules below:
      - Upper contexts
      - Combination of multiple parent nodes
      - Extracted from scope note
    - Principle 2. Remove Qualifiers from Chinese Terms (Different terms exist when translated to Chinese.)
      - Chinese qualifiers are removed according to the elimination of homograph
AAT has these 19 ritual vessels...

Example: Chinese ritual bronzes (re-grouping)

Guide Term Suggestion: We could classify them into 3 groups based on their historical usages.
Need to add 19 NEW Chinese Ritual Bronzes!

**<Wine Vessels>**

- 勺 (shao)  
- 觛 (gong)  
- 壶 (hu)
- 釂 (fou)
- 鬲 (he)
- 鼀 (zhì)
- 鬲 (ho)

**<Musical Instruments>**

- 鐗 (zhōng)  
- 鈴 (líng)  
- 鐛 (duō)
- 饒 (náo)
- 鋸 (zhēng)
- 鍲 (bō)
- 鏇 (chūn yú)

**<Food Vessels>**

- 匕 (shào)  
- 览 (gōng)  
- 壺 (hú)
- 釂 (fòu)
- 鬲 (lí)
- 鬲 (hú)
- 鬲 (li)

**<Food Vessels>**

- 釂 (shào)  
- 壺 (gōng)  
- 壺 (hú)
- 釂 (fòu)
- 鬲 (lí)
- 鬲 (hú)
- 鬲 (li)
After completion... AAT will have 38 terms with classification

- Possible Hierarchical Position for Chinese Bronze Vessels - Musical Instruments

- Wine Vessels
  - gu (ritual vessels)
  - he (vessel)
  - jia
  - jiao
  - jue
  - lei
  - pou
  - you (lidded vessel)
  - zun
  - fang yi
  - fou
  - gong
  - he
  - ho
  - hu
  - shao
  - zhi

- Food Vessels
  - ding (ritual vessels)
  - fang ding
  - dou
  - gui
  - xu
  - xu
  - yan (ritual vessels)
  - bi
  - dun
  - fu
  - li

- Water Vessels
  - jian
  - pan (ritual vessels)
  - yi (ritual vessels)
  - yu (basin-shaped vessels)
What Are **Structural Issues** in an AAT Context?

- The way we classify, organize and arrange terms
- The AAT can be used to guide users in enhancing their knowledge of specific topics and concepts, and can even affect their views on and perception of these concepts.
- In terms of Chinese and English mapping, the study has noticed some definite patterns in the structural issues which arise.

Facet and hierarchical structures of AAT
The Structural Issue

There is a conflicting structural arrangement between Western and Chinese cultures.....

To begin with, the case of ceramic glazes
Western v.s. Chinese

In the context of Western cultures, people tend to classify ceramic glazes:

- **by composition**
  alkaline glaze, blue-and-white, Bristol glaze, celadon, lead glaze
- **by form**
  clear glaze
- **by technique**
  blister glaze, crystalline glaze, matte glaze, overglaze

From a Chinese perspective, curators and research studies tend to classify ceramic glazes:

- **by color**
  under different colors with varying compositions, forms and techniques

![Tea dust glaze](image1.png)  
![Red glaze](image2.png)

Six Conjoined Vases with Tea Dust Glaze  
Ch’ien-lung Reign (1736-1795)  
Qing (1644-1911)

Monk’s Cap Ewer with Ruby Red Glaze, Hsüan-te Reign (1426-1435)  
Ming (1368-1644)

清 乾隆 茶葉末六聯瓶  
明宣德 宝石紅僧帽壺
The Problem Includes

• Some terms are shared by Western & Chinese cultures. Where should these terms be located within the hierarchical structure of the AAT?
• Some terms are arranged in different FACETS by Western & Chinese cultures
Ceramic Glaze by
**color vs. composition/origin vs. form vs. technique**

**NPM - A08 Chinese Ceramic Glaze by Color**

- 單色釉 (single color)
  - 青釉 (EQ)
  - 透明釉 (EQ)
  - 紅釉 (EQ)
  - 脈下彩録 (underslaze enamels)
    - 青花 (EQ)
    - 墨彩 (EQ)
    - 鐵録 (BM)
    - 脈裡紅 (BM)
- 彩瓷 (painted china)
  - 粉彩 (EQ)
  - 綠彩 (EQ)
  - 黃彩 (EQ)
  - 五彩 (EQ)

**NPM - A03 Chinese Kiln System**

- 定窯系 (EQ)
- 汝窯系 (EQ)
- 榮府窯系 (EQ)
- 邢窯系 (EQ)
AAT Applications: Linked Open Data (LOD) Supporting Digital Humanities
AAT as LOD (Linked Open Data) and its application in ASCDC

- Online portals for LOD Datasets, launched since 2018
  - ASCDC LOD Datasets (http://data.ascdc.tw/)
  - ASCDC LODLAB (http://lodlab.ascdc.tw/)

- Current results
  - Datasets: 13
  - Data records: 605,324
  - Triples: 5,270,200

- Topic types of ASCDC’s LOD datasets
  - Anthropology (Indigenous people in Taiwan)
  - Arts (Art Textbooks of Elementary and Public Schools in the Japanese Colonial Period, Archives of Chen Cheng-po’s Paintings and Document, Linked Taiwan Artists)
  - Audio-video files (Educational and Cultural Programs from Chinese Television System)
  - Biology (Taiwan fish species database, Taiwan’s Fish Otholits, Aquatic Animals in Taiwan)
  - History (Chinese Rare Books, Database for Names and Biographies, Database of Qing Official Titles, Wooden Slip Character Dictionary)
  - Religion (Tibetan Collection Audio Archive)
# Overview on the reused vocabularies in ASCDC LOD datasets

<table>
<thead>
<tr>
<th>External resources</th>
<th>Local resources by ASCDC</th>
</tr>
</thead>
<tbody>
<tr>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
</tr>
<tr>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
</tr>
<tr>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
</tr>
<tr>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
</tr>
<tr>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
</tr>
<tr>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
</tr>
<tr>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
</tr>
<tr>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
</tr>
<tr>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
</tr>
<tr>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Biology</th>
<th>Anthropology</th>
<th>Religion</th>
<th>Arts</th>
<th>History</th>
<th>Audio-Video</th>
</tr>
</thead>
<tbody>
<tr>
<td>Darwin Core</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
</tr>
<tr>
<td>DBpedia</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
</tr>
<tr>
<td>DC/ DCterms</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
</tr>
<tr>
<td>EDM</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
</tr>
<tr>
<td>GeoNames</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
</tr>
<tr>
<td>AAT</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
</tr>
<tr>
<td>TGN</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
</tr>
<tr>
<td>ULAN</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
</tr>
<tr>
<td>LOC</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
</tr>
<tr>
<td>Linked Open Copac and Archives Hub (LOCAH)</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
</tr>
<tr>
<td>VIAF</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
</tr>
<tr>
<td>ASCDC</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
</tr>
<tr>
<td>Linked Taiwan Artists (LTA)</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
</tr>
<tr>
<td>Database for Names and Biographies (DNB)</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
</tr>
</tbody>
</table>
Cross-dataset queries as strategy for data enrichment

HOWEVER,

- To link external resources is only one step for data enrichment, but to further reuse the data content of the linked external resources can further enlarge knowledge.
AAT as LOD Resource for Cross-dataset Queries

Dataset: The Archival fonds of Chen Cheng-po’s paintings and documents

Query: Which Western artists appear in Chen's scrapbooks? What were their nationalities and major activities?

ULAN: Auguste Renoir

AAT: Painters

AAT: Printmakers

gvp:agentTypeNonPreferred

gvp:nationalityPreferred

Auguste Renoir

雷諾瓦

http://data.ascdc.tw/Agent/archive/ith/ccp/006

http://data.ascdc.tw/archive/ith/ccp/CCP_09_01033_AP1_091

AAT: French
How can Linked Open Data enhance research? **Data Enrichment**

**Example of Cross-dataset Query for Knowledge Expansion**

**Dataset:** The Archival fonds of Chen Cheng-po’s paintings and documents

**Query:** Which Western artists appear in Chen's scrapbooks? What were their **nationalities** and major activities?

- French: 26 (57.8%)
  - Edgar Degas, *French*
  - Paul Cézanne, *French*
  - Edouard Manet, *French*
  - Camille Pissarro, *French*
- Italian: 7
  - Michelangelo, *Italian*
- German: 4
  - Albrecht Dürer, *German*
  - Peter Paul Rubens, *Flemish*
- Flemish: 3
  - Vincent van Gogh, *Dutch*
- Dutch: 1
- Spanish: 1
- Belgian: 1
- American: 1
Thanks for your attentions