

Development of AAT-Taiwan and Reusing Getty's Vocabularies in the Semantic Web

Hsi-yuan Chen, Sophy S. J. Chen, Lu-yen Lu
Academia Sinica Center for Digital Cultures (ASCDIC)
Academia Sinica, Taiwan (R.O.C.)

International Terminology Working Group (ITWG) Meeting
06-07 Feb. 2020, Los Angeles, USA

AAT-Taiwan

AAT-TAIWAN and ASCDC in Academia Sinica

<http://aat.teldap.tw/>

The screenshot shows the AAT-Taiwan website interface. At the top, there are navigation links: 首頁 (Home), 瀏覽 (Browse), 使用說明 (Usage Instructions), 關於 AAT-TAIWAN (About AAT-TAIWAN), 聯絡與建議 (Contact & Suggestions), and ENGLISH. The main content area features a search bar with fields for TERM (詞彙), ID (語彙號), and SN (聯號註). Below the search bar are radio buttons for AND/OR and a dropdown menu for 指定層面 (指定層面: 全部層面). The interface is divided into several facets: 所有詞彙 (TOP OF THE AAT HIERARCHIES), 關聯概念層面 (ASSOCIATED CONCEPTS FACET), 物理特質層面 (PHYSICAL ATTRIBUTES FACET), 風格與時代層面 (STYLES AND PERIODS FACET), 代理者層面 (AGENTS FACET), 活動層面 (ACTIVITIES FACET), 材料層面 (MATERIALS FACET), and 物件層面 (OBJECTS FACET). A featured entry for '銅圓註' (銅圓註) is displayed, including a small image of a bronze vessel and a brief description. The footer contains links for 網站首頁 (Home), 網站導覽 (Site Map), 著作權聲明 (Copyright Statement), 隱私權聲明 (Privacy Statement), and 管理權登入 (Manager Login).

- In cooperation with Getty Research Institute(GRI) since 2008, Academia Sinica Center for Digital Cultures (ASCDIC) joined as a member of the multilingual **translation project of AAT terms into traditional Chinese language.**
- The current AAT-Taiwan contains **35,450 concepts in 7 facets.**
- Translated Chinese terms are contributed to Getty's AAT and its LOD datasets to enrich the data content.

Participation of ASCDC in ITWG

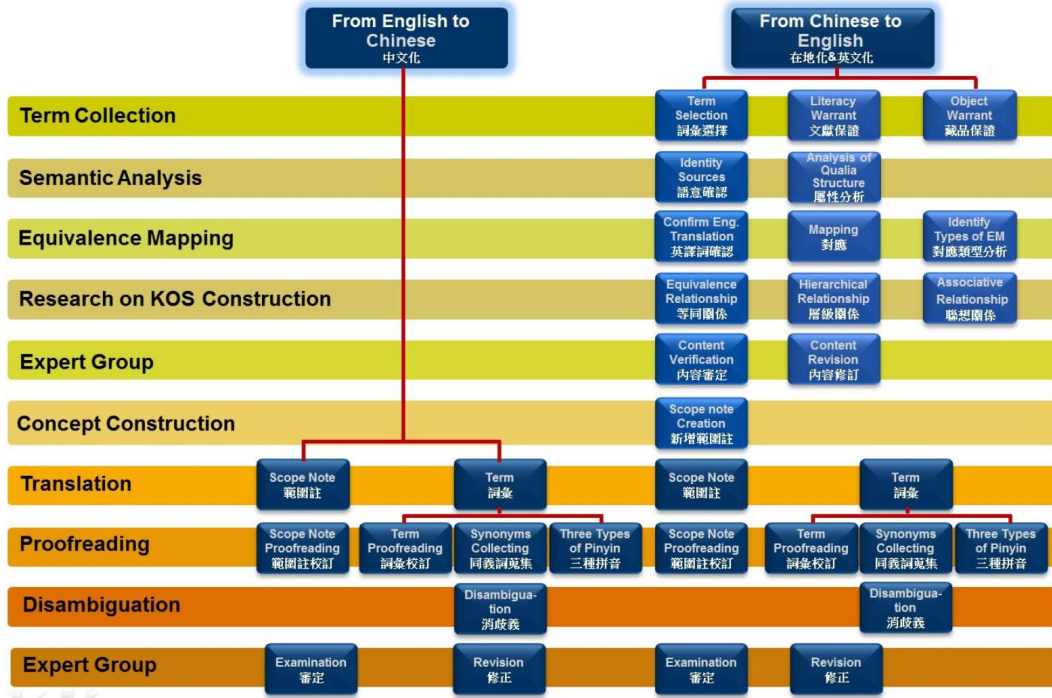
Pre-ITWG Meeting

2008.07, 2009.02, 2009.07

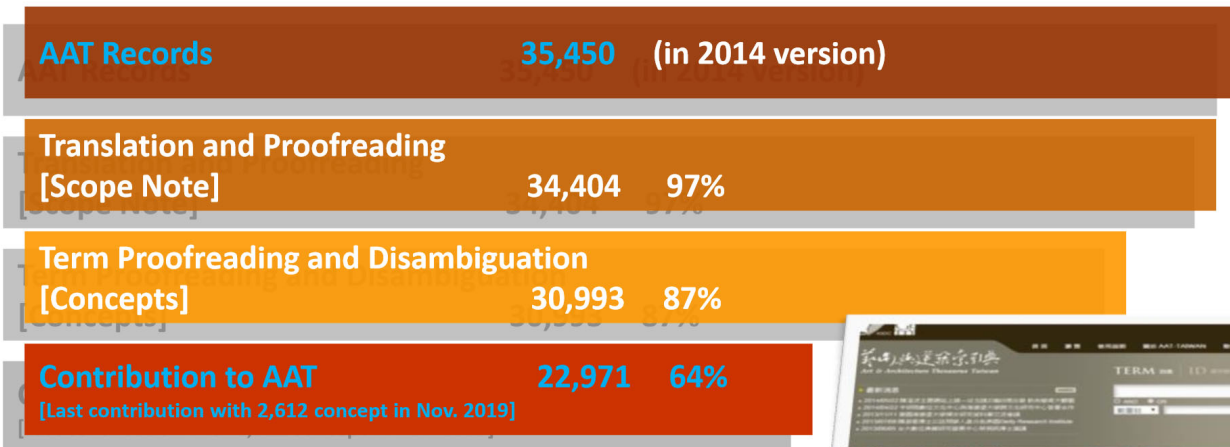
ITWG Meeting

2010.08, 2011.11, 2012.01, 2014.09, 2013.01, 2016.08, 2017.08

Research Model of the AAT-Taiwan



Current Status of AAT-Taiwan



Status: According to AAT-Taiwan database in Jan 2020

Add on local Images

Translation of scope note in Chinese

Translation of preferred and non-preferred terms

translated terms

Add related external resources from Taiwan Union Catalog

外部連結：
 Getty AAT: 300015637
 聯合目錄：肖像畫、肖像、hsiao hsiang hua、xiào xiàng huà、xiao xiang hua、portraits、portrait、portraiture、Porträt、portretten、portret、portrait、ritratti、ritratto、porträtt、retratos、retrato、retrato de cuerpo entero、retrato individual

引用：**Add related external resources from Taiwan Union Catalog**
 [肖像畫 portraits(編號300015637)]。《藝術與建築索引典》。
<http://aat.teldap.tw/AATFullDisplay/300015637> (2018/06/10瀏覽)。

Extensions for Cultural-Specific Concepts Create a Whole Set of Terms



Lantern Festival
(AAT:300400803)

Term Group of the "Chinese festivals" (9 Terms)

e.g.: Chinese New Year, Lantern Festival, Spring Sacrifice Day, Qingming Festival, Dragon Boat Festival, Double Seventh Day, Mid-Autumn Festival, Double Ninth Festival



Zun
(AAT:300265850)

Term Group of the "Chinese scripts" (16 Terms)

e.g.: Clerical script, Qin clerical script, Han clerical script, Cursive script, Wild cursive script, Running script, Running-cursive script, Seal script, Bronze script, Great seal script, Oracle bone script, Small seal script



Seal Script
(AAT:300343588)

Term Group of the "Chinese ritual vessels" (19 Terms)

e.g.: Ding, Fang ding, Dou, Gu, Gui, He, Jia, Jian, jiao, Jues, Lei, Pan, Pou, Xu, Yan, Yi, You, Yu, Zun



Malilikid
(AAT-TW:903002)

Term Group of the "Cultural festivals", "Styles and periods of Taiwanese indigenous people" (40 Terms)

e.g.: Millet harvest festival, Sea/river worship, End-of-year festival, Malilikid, Pas-taai, Mikong, Amis, Ataya, Paiwan, Bunun, Rukai, Kavalan, Seedig...

Contribution of term group of "Taiwanese indigenous people" in near future

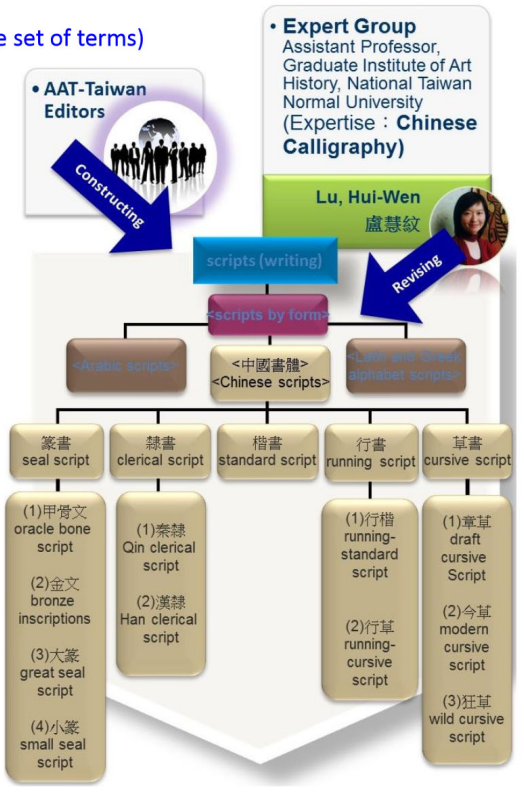
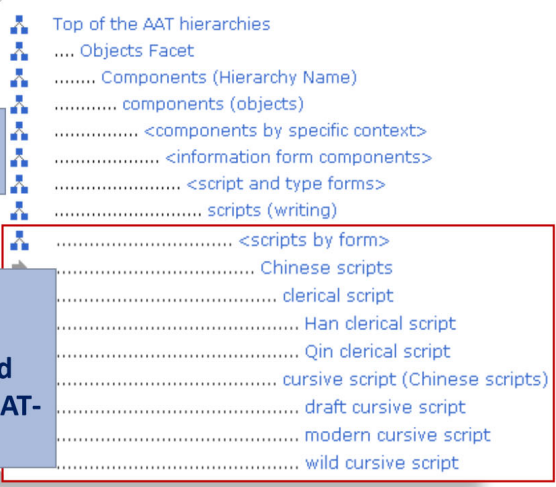
Example: Chinese scripts (create a whole set of terms)



* These terms are ma...

Original structure in AAT

Chinese scripts structure -- established and contributed by AAT-Taiwan



Semantic View (JSON, RDF, N3/Turtle, N-Triples)

ID: 300343585

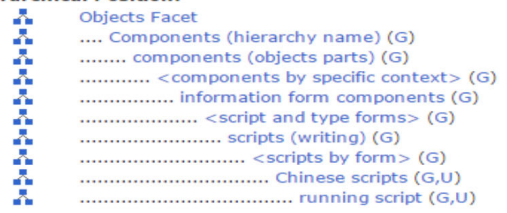
running script (Chinese scripts, <scripts by form>, ... Components (hierarchy name))

Note: Semicursive Chinese script that developed out of the Han dynasty lishu script at the same time that the standard kaishu script was evolving (1st-3rd century CE). The characters of Chinese running script are not abbreviated or connected, but strokes within the characters are often run together. The characteristic continuous brushstrokes of this script are more easily written than the more formal clerical and standard/regular script, but the running characters are more controlled and legible than Chinese cursive script. This is the most common script used in daily informal writing.

Terms:

- running script (preferred, C,U,English-P,D,U)
- semi-cursive script (C,U,English,AD,U)
- semicursive script (C,U,English,UF,U,U)
- 行書 (C,U,Chinese (traditional)-P,UF,U)
- xíng shū (C,U,Chinese (transliterated Hanyu Pinyin)-P,UF,U)
- kǎi shū (C,U,Chinese (transliterated Hanyu Pinyin),UF,U,U)
- xíng shu (C,U,Chinese (transliterated Pinyin without tones)-P,UF,U)
- hsing shu (C,U,Chinese (transliterated Wade-Giles)-P,UF,U)
- k'ai shu (C,U,Chinese (transliterated Wade-Giles),UF,U,U)
- lopend schrift (C,U,Dutch-P,D,U,U)
- kaishu (C,U,Dutch,UF,U,U)
- gyōsho (C,U,Japanese (transliterated),UF,U,U)

Hierarchical Position:

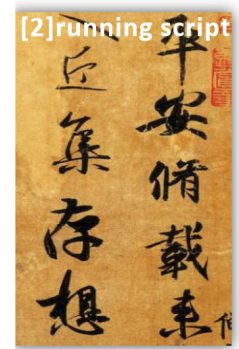


Additional Notes:

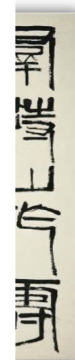
Chinese (traditional) 介於楷書、隸書與草書之間的書體，書體特色為：具有連續書寫的筆勢，較隸書、楷書生動便捷，但字形不像草書放縱難以辨識，是社會最常用的手寫書體，多用於非正式的私人書寫活動。
Dutch Semicursief Chinees schrift dat zich ontwikkelde vanuit het lishuschrift van de Han-dynastie in dezelfde tijd waarin het reguliere kaishuschrift ontstond (1ste-3de eeuw n. Chr.). De karakters van het Chinese lopende...

Record Type: concept

[2]running script



"Three Passages: P'ing-an, Ho-ju, and Feng-chü", Wang Xizhi (303-361), Chin dynasty



"Five-Character Couplet in Seal Script", Ch'i Pai-shih (1864-1957), Republic of China

Linked Open Data (LOD) Applications Supporting Digital Humanities

Application & Getty's vocabularies

LOD at ASCDC

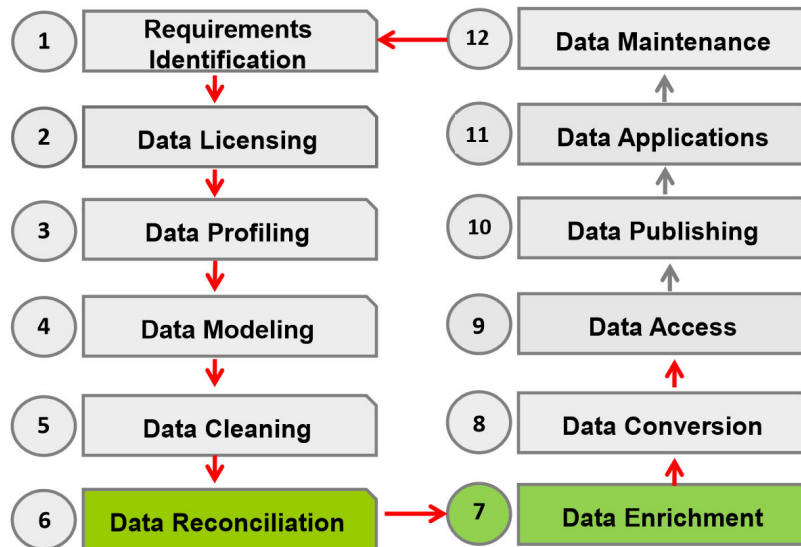
- Since 2018, ASCDC has established the **LOD Lab** to drive the research and application of LOD in order to publish digital archives of Taiwan through the semantic web and forge meaningful linkage with global data resources.
- Online portals for LOD Datasets, launched since 2018
 - **ASCDC LOD Datasets** (<http://data.ascdc.tw/>)
 - **ASCDC LODLAB** (<http://lodlab.ascdc.tw/>)
- Current results
 - Datasets: 12
 - Data records: 602,089
 - Triples: 5,252,571
- LOD dataset topics
 - **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 Documents, Linked Taiwan Artists)
 - **Audio-video files** (Educational and Cultural Programs from Chinese Television System)
 - **Biology** (Taiwan fish species database, Taiwan's Fish Otoliths, Aquatic Animals in Taiwan)
 - **History** (Chinese Rare Books, Database for Names and Biographies, Database of Qing Official Titles)
 - **Religion** (Tibetan Collection Audio Archive)

The screenshot shows the 'Download Datasets' page of the ASCDC LOD Lab. At the top, it states: 'The platform will transform open data provided by the institutes into LOD listings.' Below this, the statistics are displayed: 12 Datasets, 602,089 Data records, and 5,252,571 Triples. A grid of 12 dataset thumbnails follows, each with a title and a brief description:

- Archive of the Art Textbooks of Elementary and Public Schools in the Japanese Colonial Period** (National Taipei University of Education Department of Digital Technology Design)
- The Tibetan Collection Audio Archive Research Project** (Tibetan Studies Institute of Liberal Arts)
- The Digital Collection of Taiwan's Fish Otoliths** (National Museum of Marine Biology and Aquatics)
- The National Taiwan Museum's Foreman Aboriginal Collection Digitizing Project** (National Taiwan Museum)
- Taiwan Fauna database research: Taiwan digital fish species database** (Biodiversity Research Center, Academia Sinica)
- Linked Taiwan Artists** (National Taiwan University Center for Digital Culture)
- Linked Data for Name Authority Database at IH P** (Institute of History and Philology, Academia Sinica)
- Constructing LOD (Linked Open Data) for Chinese Rare Books** (Institute of History and Philology, Academia Sinica)

The Linked Open Data Lifecycle Framework (v2.4)

The method and procedure of this research is based on the “Linked Open Data Lifecycle Framework, LODLiFrame”, which contains the following steps:



Data Reconciliation
Data Enrichment
Application & Getty's vocabularies
□ External resources
□ Local resources by ASCDC

Overview of reused resources in ASCDC LOD datasets

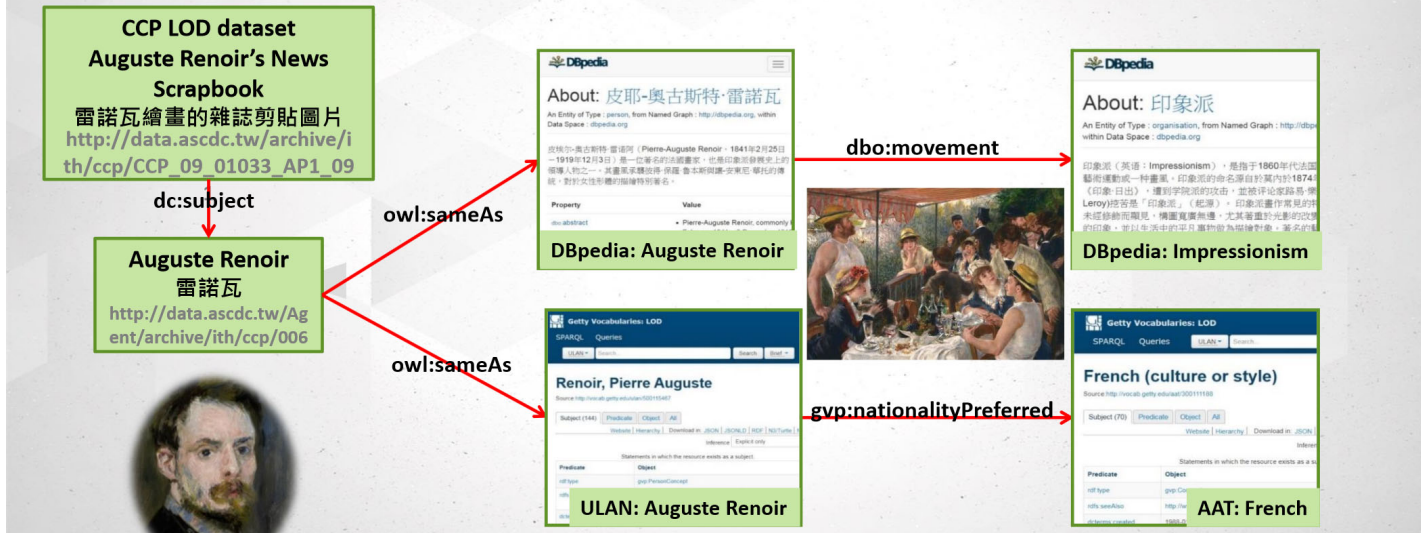
	Biology	Anthropology	Audio-Video	Religion	Arts	History
Darwin Core	✓					
DBpedia	✓	✓	✓	✓	✓	✓
DC/DCterms	✓	✓	✓	✓	✓	
EDM	✓	✓	✓	✓	✓	
GeoNames	✓	✓	✓			
AAT (type, material, technic, script type)	✓	✓	✓	✓	✓	✓
TGN (birth place, depicted location)					✓	✓
ULAN (artists)					✓	
LOC						✓
Linked Open Copac and Archives Hub (LOCAH)					✓	
Temporal Gazetteer (TGAZ, Harvard)					✓	✓
VIAF	✓	✓	✓	✓	✓	✓
ASCDC	✓	✓		✓	✓	✓
Linked Taiwan Artists (LTA)					✓	
Database for Names and Biographies (DNB)					✓	✓

How can the Linked Open Data enhance research? Data Enrichment Cross-datasets Queries as Strategy for Knowledge Expansion

Dataset: Chen Cheng-po's paintings and documents

Query 1: Which western artists appear in Chen's scrapbook? What are their nationalities?

Query 2: What are the artistic styles of these artists?



Application & Getty's vocabularies

Example of cross-datasets query for knowledge expansion

Dataset: Chen Cheng-po's paintings and documents

Example 1: Which Western artists appear in Chen's scrapbook? What are their nationalities?

LOD Datasets interface showing a SPARQL query and its results.

Query examples:

```

1 PREFIX dct: <http://purl.org/dc/terms/>
2 PREFIX dc: <http://purl.org/dc/elements/1.1/>
3 PREFIX owl: <http://www.w3.org/2002/07/owl#>
4 PREFIX foaf: <http://xmlns.com/foaf/0.1/>
5 PREFIX gvp: <http://vocab.getty.edu/ontology#>
6 SELECT ?ulan, ?person, ?nationality
7 FROM <http://data.ascdc.tw/archive/ith/ccp/>
8 FROM <http://vocab.getty.edu/ulan/>
9 FROM <http://vocab.getty.edu/aat/>
10 WHERE {
11 ?s dct:isPartOf <http://data.ascdc.tw/>

```

SPARQL query reasoning:

- 1) Locate every URI containing value as `http://data.ascdc.tw/archive/ith/ccp/CCP_09_01` in the field of `dct:isPartOf` in Item level.
- 2) Look for the URIs beginning with `http://data.ascdc.tw/agent/archive/ith/ccp/` in the field of `dc:subject` in the URIs obtained in step 1.
- 3) Look for person's name in the field of `rdfs:label` and `ULAN` URI in the field of `owl:sameAs` in the URIs obtained in step 2.
- 4) Look for ID Agent URI in the field of `foaf:foaf` of the matched `ULAN` URI queried in the step 3.
- 5) Look for the data in the field of `gvp:nationalityPreferred` of the ID Agent URI and their preferred label (`gvp:label`) obtained in the Step 4. Search results will appear in the fields of Item, `Person`, `ULAN` and

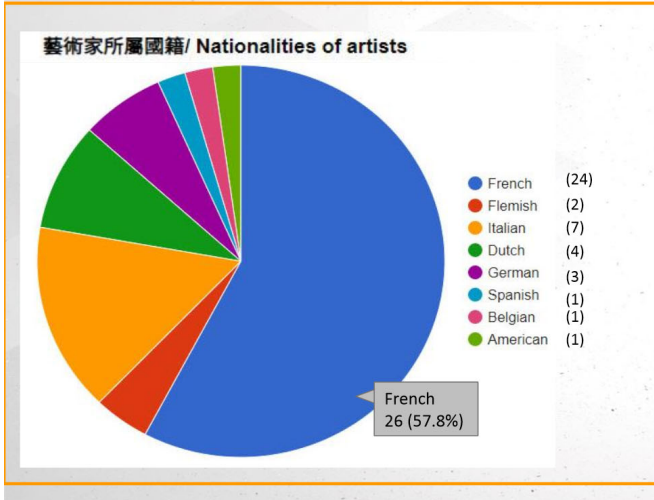
Results of cross-dataset query: name (label) and nationality of the Western artists in ULAN

ulan	person	nationality
http://vocab.getty.edu/ulan/500001924	Pissarro, Camille	French (culture or style)
http://vocab.getty.edu/ulan/500003416	Lhotte, Andre	French (culture or style)
http://vocab.getty.edu/ulan/500004535	Deraïn, André	French (culture or style)
http://vocab.getty.edu/ulan/500004793	Cézaire, Paul	French (culture or style)
http://vocab.getty.edu/ulan/500008410	Signac, Paul	French (culture or style)
http://vocab.getty.edu/ulan/500008582	Vlaeminck, Maurice de	French (culture or style)
http://vocab.getty.edu/ulan/500009979	Camoin, Charles	French (culture or style)

How can the Linked Open Data enhance research? **Data Enrichment** Example of Cross-datasets Query for Knowledge Expansion

Dataset: Chen Cheng-po's paintings and documents

Query 1: Which western artists appear in Chen's scrapbook? What are their nationalities?



IIIF as method of Establishing Image-based DH Research

<https://buddhism.ascdc.sinica.edu.tw/buddhism>

- A LOD-based knowledge-base of Asian Buddhist art images with focus on sculptural works (Stupa, Caves and Stelae) in India and China.
- Images are presented in 3-layered architectural structure as: whole structure (layer A), different sides of wall (layer B), image section on a single wall (layer C).
- Semantic data modeling based on CIDOC-CRM/Linked Art Model; AAT terms for enriching data content as materials and method of applied techniques; TGN for describing to works' locations and integrate to the historical GIS.
- Presentation of images based on IIIF standards which allow researchers:
 - Exchange of materials
 - Collaboration
 - Reuse of data
 - Reaggregate content
 - Comparison
 - Spatio-temporal information
 - Microscopic analysis
 - Representation of structural spaces
 - Semantic Annotation

Image of parent-layer (Layer A)

Image of the main work (Layer B)

Images of same layer (Layer B, other wall sides)

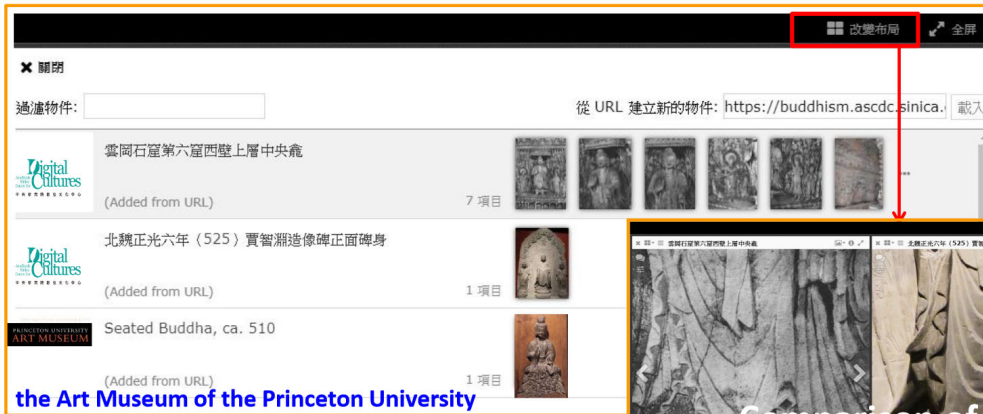
Images of child-layer (Layer C)

Metadata (Layer B)

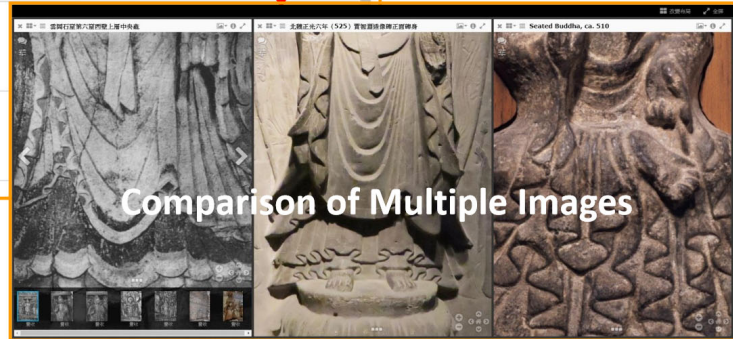
Xiaonanhai caves, eastern wall in the middle caves, China, ca. 550
小南海石窟中窟東壁, 河南省 (約西元550年)

Content and Beyond

Combining LOD-based semantic queries with IIIF-based interoperability



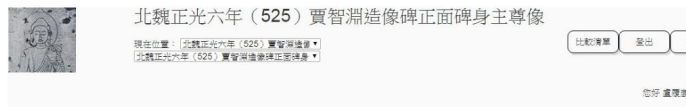
Using IIIF viewer as method to conduct stylistic analysis on the drapery study of the Buddhist arts



Query example: what are the Buddhist art works in the dynasty of the Northern Wei (ca. 386-534) and artistic style of the drapery modeling on those art works?

Content and Beyond

- Development of different **motivation subcategories for IIIF annotation of images**, including **“Tagging”**, **“Describing”** and **“Commenting”** based on W3C’s Web Annotation Data Model.
- Expand **“tagging”** semantics: currently, **4 annotation types** under subcategory of **“Tagging”** are expanded and **mapped to AAT**.



Example: Jia zhi-yuan Buddhist stele, ca. 525, Northern Wei Dynasty

Anno. Type	AAT concept
1 人物/ figures	http://vocab.getty.edu/aat/300189808
2 物件/ objects	http://vocab.getty.edu/aat/300311889
3 地點/ places	http://vocab.getty.edu/aat/300264550
4 動作/ events	http://vocab.getty.edu/aat/300069084

Historical GIS Map with TGN information

- Embedded with **multi-layered historical GIS maps from 20 historical Chinese dynasties** with a temporal coverage of **ca. 3500 years**.
- Reusing **geographical information in TGN** to mark the location of each work on map



Thanks for Listening!

Academia Sinica Center for Digital Cultures (ASDC)
Academia Sinica, Taiwan (R.O.C.)