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

Sophy Shu-Jiun Chen

Institute of History and Philology, Academia Sinica, Taiwan, <sophy@sinica.edu.tw>

Sophy Shu-Jiun Chen is the Assistant Research Fellow at the Institute of History and Philology, Academia Sinica, and also the Executive Secretary of Academia Sinica Center for Digital Cultures(ASCDC). She received her M.A. degree in Information Studies from the University of Sheffield, UK in 1997, and Ph.D. degree in Library and Information Science from the National Taiwan University in 2012. Dr. Chen is also an Adjunct Assistant Professor of the Department and Graduate Institute of Library and Information Science, National Taiwan University. Her research interests include digital libraries, metadata, ontologies, Linked Data, knowledge organization, and digital humanities. She initiated the Research Project of Chinese-language AAT (Art & Architecture Thesaurus) (https://aat.teldap.tw/) with the Getty Research Institute, USA since 2008, and the Linked Open Data Lab in Aacdemia Sinica (https://lodlab.ascdc.tw/).

Abstract: A multilingual thesaurus' development needs the appropriate methodological considerations not only for linguistics, but also cultural heterogeneity, as demonstrated in this report on the multilingual project of the Art & Architecture Thesaurus (AAT) in the Chinese language, which has been a collaboration between the Academia Sinica Center for Digital Culture and the Getty Research Institute for more than a decade. After a brief overview of the project, the paper will introduce a holistic methodology for considering how to enable Western art to be accessible to Chinese users and Chinese art accessible to Western users. The conceptual and structural issues will be discussed, especially the challenges of developing terminology in two different cultures. For instance, some terms shared by Western and Chinese cultures could be understood differently in each culture, which raises questions regarding their locations within the hierarchical structure of the AAT. Finally, the report will provide cases to demonstrate how the Chinese-Language AAT language supports online exhibitions, digital humanities and linking of digital art history content to the web of data.

1.0 Introduction

The Art & Architecture Thesaurus (AAT) (Getty Research Institute 2017) is a multilingual thesaurus that was developed by the Getty Research Institute (GRI) in the 1970s, published in print in 1990, and became an online database from 1997. In addition to the English language, AAT also includes Dutch, Spanish, German, Chinese, and some French and Italian. AAT collects concepts and terms for describing items of art, architecture, and material culture, which are used for communication of visual arts, as well as in cataloging, information retrieval, data analysis, and re-

search platforms. It was released as Linked Open Data in 2014. The chronological scope of the Thesaurus spans from ancient times to the present. The geographical scope was first based in the West, and then continuously expanded to cover cultures from all over the world. By the end of 2020, it included about 71,000 records and 400,400 terms (Molholt and Petersen 1993; Harpring 2010; Franzon 2014; J. Paul Getty Trust 2020a, 2020b).

To develop a more comprehensive AAT multilingual work, GRI established the International Terminology Working Group (ITWG) in 2010. The initial members included Taiwan's Academia Sinica Center for Digital Cul-
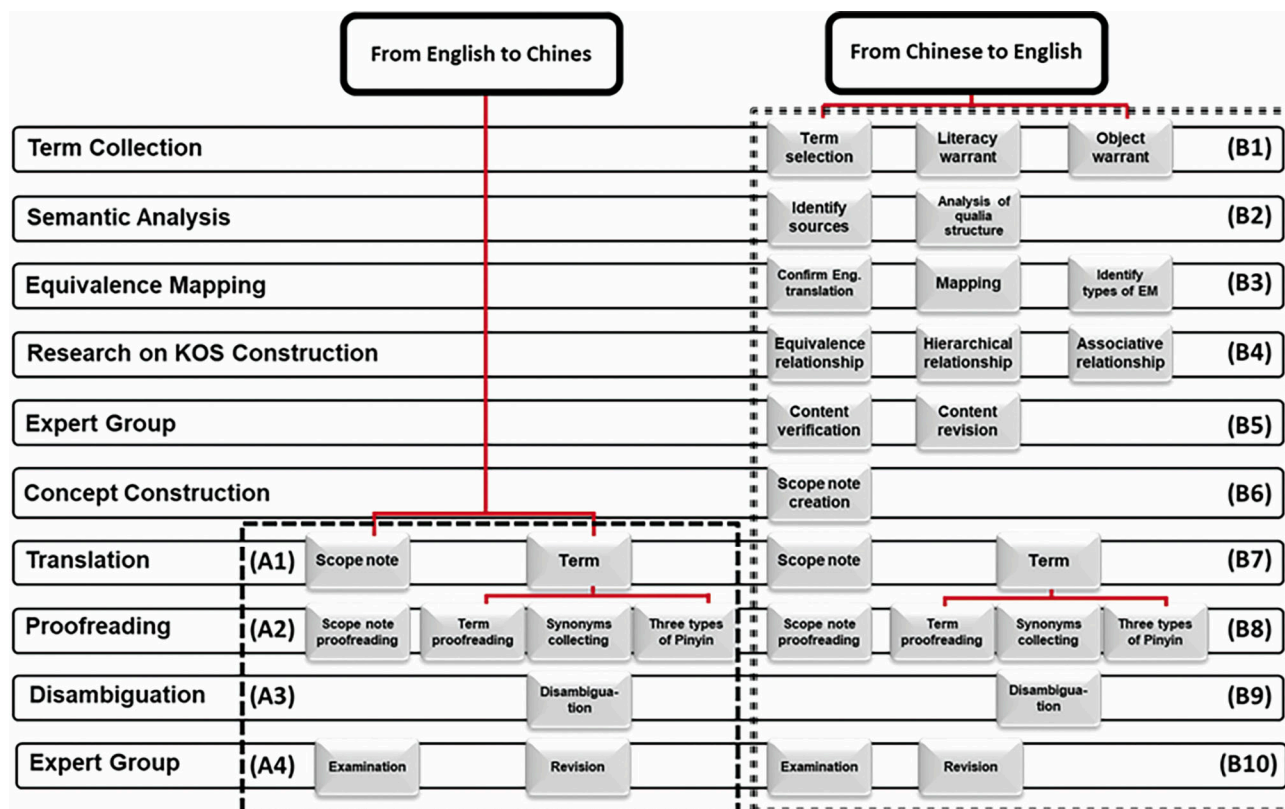
268

Knowl. Org. 48(2021)No.4
S. S.-J. Chen. Methodological Considerations for Developing the *Art & Architecture Thesaurus* in the Chinese Language ...

tures (ASCDC), the Berlin State Museums' Institute for Museum Research, the Netherlands' Institute for Art History, and Chile's Centro de Documentación de Bienes Patrimoniales, respectively representing the Chinese, German, Dutch, and Spanish language translation projects. The Chinese AAT project, known as "AAT-Taiwan" (Academia Sinica Center for Digital Cultures 2021) collaborated with GRI, not only on translating AAT vocabulary from English to Chinese, but also on supplementing and contributing to AAT terminologies in both Chinese and English (Chen, Cheng and Chen 2011; Baca 2013; Chen and Chen 2015; Chen, Zeng and Chen 2016). From its inaugural meeting in 2010, ITWG has convened nine times to discuss and share the diverse members' experiences with editorial, translation, and technical issues, problem-solving strategies, as well as application of the AAT and other Getty Vocabularies in digital humanities research (Harpring, Cobb and Garcia, 2013; Baca and Gill 2015; Andree and Weda 2016).

Since the publishing of Getty AAT as Linked Open Data in 2014, AAT has already proved itself as a reusable vocabulary and structured resource for many research platforms and databases to normalize, manage and enrich data content, and to create semantic relations between different data types (Baca 2013; Tudhope and Binding 2018). For instance, in museum, conservation and archaeological field research, AAT can be applied as an indexing tool to organize museum collections, conservation content online (Chapman 2012; Dijkshoorn, Jongma, Aroyo, Van Ossenbruggen, Schreiber, Ter Weele and Wielemaker 2018; Velios 2016) and to register archaeological findings in a linked data knowledge base. (López, Lerones, Llamas, Gómez-García-Bermejo and Zalama 2018). In collaboration with the Getty Conservation Institute (GCI), more and more art conservation-related concepts are being coined in the AAT and used to structure the Art & Archaeology Technical Abstracts (AATA) online database (Brown 1987; Velios 2016). However, the methodological construction of multilingual AAT terms has not been fully discussed. This report will focus on the Chinese language AAT projects (AAT-Taiwan) in the past ten years, discussing 1) methodology, 2) conceptual issues, and 3) structural issues, as well as the applications of AAT.

2.0 Methodology

A multilingual thesaurus is a type of knowledge organization system (KOS), which involves methods of interoperability between two or more languages. According to a study by Zeng and Chan (2004), there are eight kinds of methods in establishing interoperability among KOS, including: derivation/modeling, translation/adaptation, satellite and leaf node linking, direct mapping, co-occurrence mapping, switching, linking through a temporary union list, and link-

ing through a thesaurus server protocol. Among them, the translation method is usually adopted for development and interoperability between KOS of different languages. The mapping method is especially used for semantic interoperability of heterogeneous environments. For instance, many large-scale interoperability projects between different KOS, such as MACS (Multilingual Access to Subjects), AGROVOC, and HILT (High Level Thesaurus) are based on the mapping method (Tudhope, Koch and Heery 2006; Zeng and Chan 2004; Doerr 2004; MuCulloch, Shiri and Nicholson 2005).

AAT-Taiwan aims to integrate Chinese language and art concepts into AAT through translation and direct mapping. To carry out this task in a more systematic and consistent way, a practical methodology for editorial work has been developed, including A) from English to Chinese (see left in Figure 1), and B) from Chinese to English (right in Figure 1). The whole procedure includes ten tasks, in the order of (1) term collection, (2) semantic analysis, (3) equivalence mapping, (4) research on KOS construction, (5) expert group consulting, (6) concept construction, (7) translation, (8) proofreading, (9) disambiguation, and (10) expert group review and approval.

2.1 From English to Chinese

The main goal of converting AAT into the Chinese language is to transform the structured vocabulary of English concepts into Traditional Chinese and contribute to Getty's AAT database. This would expand the multilinguality of the original vocabulary and enable Chinese language users to utilize Western art resources that were formerly based in English. Among the tasks listed above (Figure 1, left), the first six tasks have already been completed by the English AAT. The workflow steps that this project has designated are the last four (A7-A10), as described below:

A7) Translation
Chinese translation is carried out for the English "preferred terms" and "scope note" for AAT concepts. Based on Hulme's principle of literary warrant, at least three authoritative reference sources are found for each Chinese "preferred term" (Petersen 1990).

A8) Proofreading
Detailed proofreading is carried out for each translated Chinese scope note, as well as for the Chinese preferred and non-preferred terms. At this stage, the search for Chinese synonyms is expanded to collect all potential Chinese terms corresponding to the concept. Then, based on the number of sources found for each term, as well as the principles of literary warrant and cultural warrant (Molholt 2001; Barité 2018), the Chinese preferred term is deter-

Knowl. Org. 48(2021)No.4

269

S. S.-J. Chen. Methodological Considerations for Developing the *Art & Architecture Thesaurus* in the Chinese Language ...

Figure 1. The methodology for AAT translation.

mined and adjusted as necessary to correspond to a cultural translation of the AAT English term; the remaining terms are regarded as the Chinese non-preferred terms. Finally, three systems of Pinyin romanization are generated for the preferred Chinese terms, including Hanyu Pinyin without tones, Hanyu Pinyin with tones, and Wade-Giles Pinyin, to make English language users pronounce and type a Chinese term properly.

A9) Disambiguation
Homographs may occur and require disambiguation, including instances where English concepts are not differentiated within the Chinese language (for example, the English terms "threads" and "lines" are both translated into Chinese as the preferred term "xian" (線)). If there are two or more Chinese term homographs, regardless of whether they are Chinese preferred terms or non-preferred terms, they will be distinguished by adding a qualifier.

A10) Expert Review
Some of the original vocabulary concepts and scope notes are very domain-specific, making it difficult for a more literal translation to accurately reflect the correct meaning or usage of the concept. Therefore, appropriate domain experts are invited to review the translated content in order to improve the accuracy and professionalism of the trans-

lation results. For example, the lists of terms under the AAT term "ritual vessels" (AAT300265801), "ceremonial containers" (AAT300197576), and the "Objects Facet" (AAT300264092) were found to be incomplete, after review by experts in the history of the Chinese Bronze Age.

2.2 From Chinese to English

As the concepts and structures of the original AAT are based on Western material cultural perspectives (Soergel 1995), this has caused the problem of not reflecting the overall human material civilization of both the East and the West. In order to make up for the shortcomings of specialized and localized cultural vocabulary, the project cooperates with domain experts to jointly supplement and add localized concepts related to Chinese and Taiwanese material cultures to AAT, such as: "Chinese Festivals," "Chinese scripts," and "festivals of Taiwanese indigenous people." The main tasks include:

B1) Term Collection
A set of new Chinese terms relevant to a specific cultural concept (such as "Chinese scripts") are collected in order to be added as a group to AAT. Authoritative literature sources and primary source objects in both Chinese and English are collected to ensure that each term has literary

270

Knowl. Org. 48(2021)No.4
S. S.-J. Chen. Methodological Considerations for Developing the *Art & Architecture Thesaurus* in the Chinese Language ...

warrant, cultural warrant and institutional warrant (Barité 2020). For example, the project collected seven terms related to Chinese festivals, originating from the controlled vocabulary and indexing terms of the National Palace Museum in Taipei, and expanded upon them.

B2) Semantic Analysis

After confirming the list of newly added localized vocabulary and their scope, the semantics of each concept are analyzed. The purpose is mainly to find a definition to describe each new term based on domain literature and verify the nature of the Chinese concept itself. The definition of each concept must have three or more authoritative sources, following the AAT's guidelines. AAT-Taiwan uses Pustejovsky's (1995) theory of qualia structure to structurally analyze concepts, utilizing the four basic roles for lexical items (formal (what an object is), constitutive (what it is composed of), agentive (its origin) and telic (its purpose) (Romeo, Mendes and Bel 2013)) to further explore the people, events, places, times, objects, etc. implied in each concept, thereby confirming the characteristics that each new concept needs to include when defining and describing it. This information can be used to write follow-up scope notes and to establish semantic relations (equivalence, hierarchical, associative relationships).

B3) Equivalence Mapping

Based on the seven types of mapping proposed by the ISO 25964-2 standard (ISO 2013), the project maps Chinese concepts to the existing AAT vocabulary, with the resulting relationships consisting of exact simple equivalence, inexact simple equivalence, cumulative compound equivalence, intersecting compound equivalence, hierarchical (narrower to broader), hierarchical (broader to narrower), and associative mapping, in order to confirm whether the new concept has an equivalent term in the existing English AAT that can be reused, or whether it is necessary to create a new entry, including Chinese and English versions.

B4) Construction of a Set of Concepts

In the existing AAT, each concept not only provides its scope note, but also indicates the equivalence, hierarchical, and associative relationships which demonstrate the knowledge structure of the concept. At this step in the process, these three types of semantic relations are established for the new Chinese concepts, enabling integration into the existing AAT.

B5) Expert Review

For those concepts which are very domain-specific (such as Chinese scripts, festivals of Taiwanese indigenous people, etc.), an expert review process is conducted to improve the accuracy and completeness of the Chinese and English scope notes, as well as the preferred and non-preferred terms.

B6) Record Construction

According to the steps mentioned above, for each entry, a full concept-based record, expressed in both Chinese and English, is formally established and contributed to the AAT.

B7-B10) Local Concept Translation, Proofreading, Disambiguation, and Expert Review

After establishing an entry for a Chinese local concept, the task of full English translation follows. The work includes four parts (B7-B10): translation, proofreading, disambiguation, and expert review. These are equivalent to steps A7-A10. The only difference is that the processed object is not an English concept, but the Chinese concept constructed from steps B1 to B6.

3.0 The conceptual issues

Studies indicate that the more two knowledge organization systems differ in language and culture, the greater heterogeneity they will have in their conceptual structures (Liang and Sini 2006). In comparison with science-based multilingual knowledge organization systems such as those in medical or agricultural fields, concepts in art and cultural fields seem to have a closer and more dependent relationship to their respective cultures. Chen and Chen (2015) conducted lexical-semantic mapping to compare the Chinese controlled vocabulary developed by the National Palace Museum (NPM) with that of the AAT. The study found that only about one-third of NPM had "exact equivalence" with AAT vocabulary, and three-fifths (3/5) of NPM had a "hierarchical (narrower to broader)" (BM) relationship with AAT vocabulary. A BM relationship indicates that when two semantically closest terms in NPM and AAT were paired together, the NPM term had greater specificity versus the broader AAT term. A portion of these cases occurred because the NPM term is unique to Chinese culture and therefore had no existing English counterpart in AAT. Examples include specific Chinese festivals, Chinese painting techniques, Chinese scripts, etc. Similarly, AAT concepts that arise from different cultures often face challenges during the process of matching or translating into Chinese. Some concepts even exist in both Chinese and Western cultures but possess differences in semantic scope due to cultural differences. In general, three types of conceptual issues were encountered when constructing an English-Chinese AAT, as discussed in the following sub-sections.

Knowl. Org. 48(2021)No.4
S. S.-J. Chen. Methodological Considerations for Developing the *Art & Architecture Thesaurus* in the Chinese Language ...

271

3.1 Lack of equivalent terms in Chinese or English

In the process of producing a Chinese-language AAT, it was found that many terms exist only in Western culture with no current equivalent Chinese term. Furthermore, most of these cases are specific terminologies in specialized fields. Adding on top of that the existing differences between English and Chinese writing systems, this has become a very difficult problem in Chinese translation. According to the AAT editorial guidelines (J. Paul Getty Trust 2018), when the target language lacks an equivalent, it is recommended to use one of three options: (1) loan terms; for example, AAT directly borrows the French term "lits à la romaine" (beds of Roman type) to use as the English term [ID: 300142594]; (2) literal translations and coined terms, which consist of literally translating the term from the source to the target language, but without authoritative warrant for the usage of the term as a loan term; for example, the AAT English term "berry spoons" [ID: 300043155] is directly translated into French as "petites louches à fruits rafraîchis; or (3) retain the source language descriptor as the target-language descriptor without warrant; for example, directly using the English term "slipper chairs" [ID: 300038225] as the Dutch term.

To elucidate the issue of appropriately converting AAT concepts from English to Chinese, the project has formulated five categories based on an extension of the principles in the AAT editorial guidelines. Under the category A of Loan Terms are: (A1) loanwords, which directly copy the source term including non-Chinese characters, (A2) orthographic transcription, which phonetically imitates the source term (A3) transcription-translation hybrid, which combines phonetic transcription and semantic translation. Under the category B of Literal Translations are: (B1) literal translation, which directly translates word-to-word, and

(B2) communicative translation, which provides a semantic description (Table 1).

Orthographic transcription (A2) is often the initial technique used to deal with loan terms. Because the result primarily reproduces the sounds of the source language with little regard for the meaning of the Chinese characters used to transcribe it, Chinese users are therefore unable to understand the transcription's meaning unless they had previously known the term. Therefore, if only such homophonic translations are available, it may not be possible to introduce foreign terms in a truly satisfactory way. Based on the overall principles of loan terms and the best practical work of Chinese translation, the project has prioritized literal translations (B1) and communicative translations (B2), because they can better express the meaning of the foreign term fully. Literal translation emphasizes closeness to the source language; communicative translation emphasizes closeness to the target language.

3.1 The difficulties of making Western art resources accessible to Chinese users

The terms defined in the AAT so far are generally derived from Western cultural contexts and concepts. Inevitably, many terms refer to similar concepts with subtle differences. For cultural natives, the distinctions between some terms (such as "bar" vs "saloon" vs "beer hall") may seem so commonsensical that they become overlooked. But it is vital to point out in the scope notes the most salient distinctions for non-Western audiences, for whom such differences do not exist in their native culture. The project recommends that existing AAT scope notes be supplemented with more detailed cultural background information, explicitly pointing out distinctions between similar terms without only relying on a qualifier label to achieve clear disambiguation. For instance, it is found that several terms under the category of

AAT Editorial Guidelines	Extension by the Chinese-language AAT	Examples
A. Loan terms	A1. loanwords	(Eng) MP3 (Chi) MP3
	A2. orthographic transcription	(Eng) krewes (Chi) ke lu (克魯)
	A3. transcription-translation hybrid	(Eng) cork (pigment) (Chi) ke ke lan (科克藍)
B. Literal translations and coined terms	B1. literal translation	(Eng) vertical wick lamps (Chi) zhi li deng xin deng (直立燈芯燈)
	B2. communicative translation	(Eng) retablos[1] (Chi) ji tan zhuang shi hua (祭壇裝飾畫)

Table 1. An extension of AAT Editorial Guidelines for lack of equivalent terms.

272

S. S.-J. Chen. Methodological Considerations for Developing the *Art & Architecture Thesaurus* in the Chinese Language ...

Knowl. Org. 48(2021)No.4

"social and sitting spaces" [ID: 300004388] in the Objects Facet [ID: 300264092], such as "living rooms" [ID: 300165461], "drawing rooms" [ID: 300004417], "parlors" [ID: 300004441], and "keeping rooms" [ID: 300004503] have similar contents in their scope notes. For users with an East Asian cultural background, their distinctions are difficult to understand and distinguish. To solve this problem, the project provided supplementary clarifications to the scope notes. The clarifications generally fell under three types: (1) Historical linguistic development or etymology of words; (2) Cultural background relating to a particular time and/or place; and (3) Modifying the definition to emphasize unique or special features of a term that are more instructive for foreign audiences, as demonstrated in the following examples (Table 2).

3.2 Cultural partiality in cross-cultural terms

Many vocabularies for cultural concepts are shared between Eastern, Western, or other cultures, but the details of their definitions and connotations may differ. AAT currently contains many such types of vocabularies (such as "lantern festivals," "hermits," etc.) where the scope notes were originally written from a particular cultural perspective that make it difficult for users of a different cultural background to recognize it as the same term used within their own culture. To improve AAT concepts that possess multiple culturally-specific meanings, possible approaches include: (1) supplementing the scope notes to include a more comprehensive definition, or (2) adding a new entry for the concept as it is understood in other cultures, distinguishing their differences in meaning or usage.

Taking "hermits" and "yin shi" (隱士) as an example, these two terms are generally accepted as equivalents in their respective languages, as well as in AAT. But "hermits" in Chinese culture historically refer to people who choose "to bar one's gates and earn one's own living without reliance on the emolument of office, to display a lack of regard for the social status which could be attained only by entering officialdom, and to devote one's life to self-cultivation, scholarship or artistic pursuits made one a recluse" (Mote 1960, 203). In other words, they were knowledgeable and qualified to be officials, but refused to seek positions in the bureaucracy against their will. They can be regarded as intellectual elites with moral integrity. However, the concept of "hermits" in the AAT is defined in the scope note as follows: "Those who live a life of seclusion and solitude, particularly those who do so for religious motives, such as the early Christian hermits." We can see that the AAT definition of "hermits" is very broad but emphasizes those who seclude themselves for religious motives, whereas the historical Chinese definition specifically refers to intellectuals who seclude themselves to avoid officialdom, a secular motive. In fact, other Chinese words exist for people who seclude for religious reasons. One possible method of dealing with this situation would be to continue treating the terms as equivalent but modifying the AAT scope notes to be more inclusive of other cultural definitions.

Another example is "lantern festivals" [ID: 300263480], a term under "festivals" [ID: 300073472] in the "Activities Facet" [ID: 300264090]. This term refers to festivals involving lanterns and parades held in Western Africa and Caribbean countries on religious holy days, which self-evidently reflects the cultural perspectives of those regions. The con-

	Original AAT definition	Suggested definition
drawing rooms (ID: 300004417)	Formal reception rooms, as in a home or hotel, or rooms to which one may retire for privacy or rest.	Formal reception rooms, as in a home or hotel, or rooms to which one may retire for privacy or rest. *Popular in the 18th and 19th century residences of wealthier families, often decorated with the finest furnishings as a symbol of social status. Originally from "withdrawing room," a private room to withdraw to* (Chambers 2020).
parlors (ID: 300004441)	Rooms intended for conversation and the reception of guests in residences, hotels, monasteries, or churches, particularly common in the 19th century. The contemporary living room is similar to the parlor in that often another room is provided for daily family relaxation.	Rooms intended for conversation and the reception of guests in residences, hotels, monasteries, or churches, particularly common in the *18th and* 19th century, *often decorated with the finest furnishings as a symbol of social status. Mostly replaced in residential architecture since the 20th century by the living room. The term has developed multiple historical meanings, from the 14th century "apartment in a monastery for conversations with outside persons," to the late 14th century "room off a main hall offering some privacy," by early 15th century to "public room of a private house," by 16th century to "the ordinary sitting room of a family," to a "show room for a business" (as in ice cream parlor) by 1884* (Harper 2021).

Table 2. Suggestions for modification of scope notes to make Western art resources accessible to Chinese users.

Knowl. Org. 48(2021)No.4

S. S.-J. Chen. Methodological Considerations for Developing the *Art & Architecture Thesaurus* in the Chinese Language ...

273

cept of "lantern festival" celebrations also exist in East Asian cultures, but since the AAT term limits the cultural geographic scope, that term cannot be reused or linked with such data. When East Asian users search for "lantern festivals," they may be confused at the term being defined by a different cultural perspective. To remedy this, our project proposed a "Chinese festivals" category at the same hierarchical level, with a sub-entry "Lantern Festival (Chinese festival)" [ID: 300400803]. This new "lantern festival" entry's scope notes describe the Chinese cultural definition, origin, and related activities, filling the hole in AAT's cultural knowledge (Figure 2). Furthermore, to differentiate the Chinese and African "lantern festivals" and avoid misunderstandings, the terms were appended with the "2100 distinguished from" associative relationship, in accordance with AAT Editorial Guidelines.

4.0 Structure issues

AAT is essentially a hierarchical thesaurus that began and was designed from Western cultural backgrounds and conceptual categories (Soergel 1995) in order to collect concept terms from material cultures, regardless of region. Because it encompasses such a broad scope, there are cases where the hierarchical structure is unable to include some cultural vocabulary representing Eastern cultural concepts, or where Chinese and Western conceptual frameworks are different. Overall, we have encountered three main structure issues:

4.1 Creating new groups of terms

The main purpose of the AAT's structure is to incorporate in a hierarchical manner concept vocabulary related to material

Figure 2. Localizing similar cultural vocabulary concepts in AAT-TW; adding an associative relationship between "Lantern Festival (Chinese Festival)" and "lantern festivals (African festivals)"

274

Knowl. Org. 48(2021)No.4

S. S.-J. Chen. Methodological Considerations for Developing the *Art & Architecture Thesaurus* in the Chinese Language ...

culture from the entire world, regardless of region. Although its approach can fully reflect the cultural vocabulary of a people or environment, it can also lead to the vocabulary hierarchy being too large and complicated, or to vocabulary related to certain cultures being so specific and narrow that it can only be understood by a few users with special cultural backgrounds, which complicates the difficulty and accuracy of work for the various international AAT translation projects. Our project focused on Chinese concepts that were important but absent in the AAT, using the above-mentioned methodology to supplement a more comprehensive vocabulary, and appropriately integrate them into the AAT's conceptual structure. Examples of the concepts involve Chinese scripts, Chinese festivals, festivals of Taiwanese indigenous people, etc. But how granular should the thesaurus be when cataloguing terms from a particular culture? Taking "Chinese scripts" as an example, terms such as "clerical script," "standard script," and "seal script" in the field of Chinese scripts are important calligraphic styles. Many museums across the world hold collections of important Chinese calligraphic works, including the Metropolitan Museum of Art, Freer Gallery of Art, Princeton University Art Museum, and British Museum. However, by 2010, the AAT had not provided any specific vocabulary for Chinese scripts. Conversely, under "scripts (writing)," there were already completed categories for Arabic scripts, hieroglyphics, Latin and Greek alpha-

bet scripts, etc. From this, we can envision that the addition of the "Chinese scripts" vocabulary group holds considerable significance for the integrity of AAT's classification of cultural scripts. Our project cooperated with scholars of calligraphy to develop the controlled vocabulary and semantic structure for "Chinese scripts", including oracle bone script, bronze script, seal script, clerical script, running script, cursive script, etc. The terms were organized in a three-tier hierarchy to represent the knowledge structure, for example, "Chinese scripts" → "seal script" → "oracle bone script" (Figure 3).

In the context of the AAT, another consideration when determining the degree of granularity to which a culture's terms are represented in the thesaurus is "organizational warrant" (Barité 2018), which refers to whether a term has been used by organizations in their collections. In international AAT translation projects (such as AAT-Taiwan), each team constructs a local AAT website that collects and localizes the contents of the AAT vocabulary, linking to the original English AAT. When a team introduces specialized vocabulary specific to a particular culture, a flexible approach can be adopted to investigate the practical needs of international users outside that culture, upon which the appropriate structural granularity of the new vocabulary can be evaluated. Take, for example, the Chinese traditional brushwork term "ts'un" [ID: 300178918]. It is located in the AAT under "brushwork" [ID: 300054094] in the "Activities Facet" [ID:
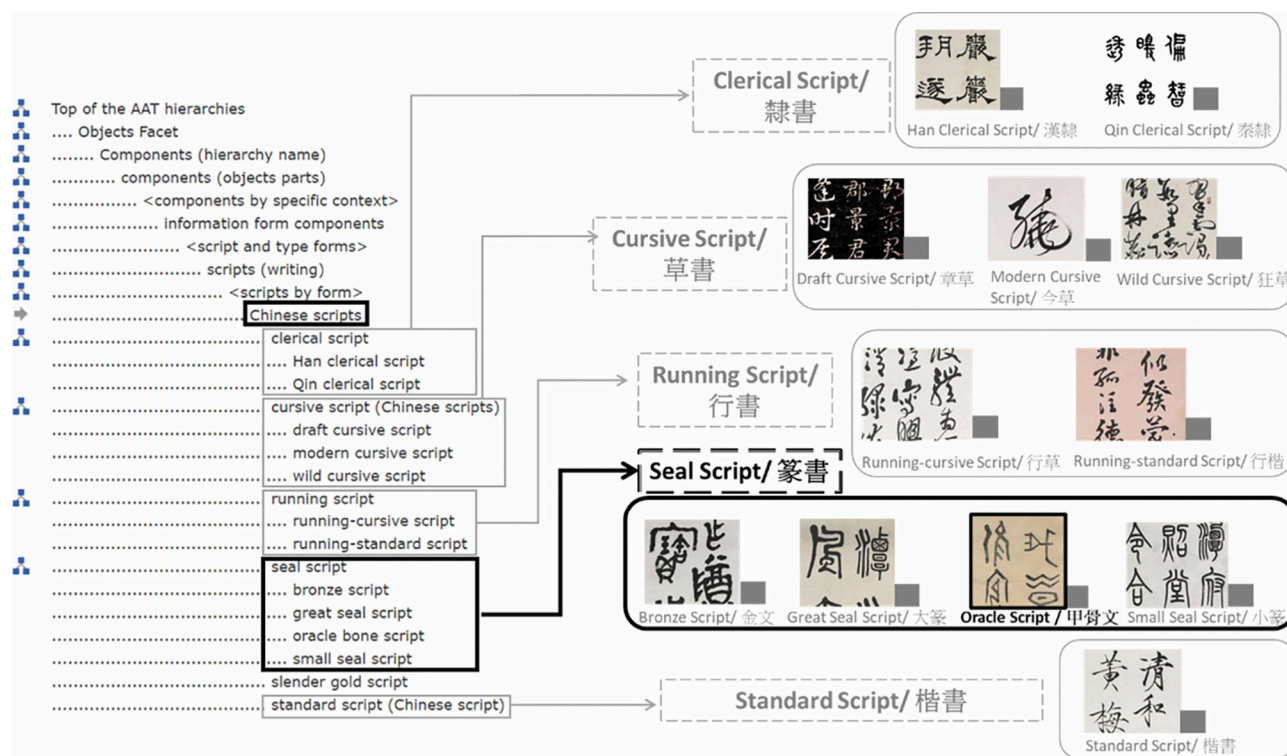

Figure 3. Concepts of Chinese scripts as three-tier hierarchy in current AAT knowledge structure (e.g.: "Chinese scripts" → "seal script" → "oracle bone script").

Knowl. Org. 48(2021)No.4

275

S. S.-J. Chen. Methodological Considerations for Developing the *Art & Architecture Thesaurus* in the Chinese Language …

300264090] and defined in its scope notes as the "type of brush strokes used in Chinese painting to indicate wrinkles, markings, and modeling on mountains, rocks, trees, and stones." In Eastern cultures, many terms exist to describe traditional ink painting techniques. Under the NPM-CV, "ts'un" encompasses ten techniques, such as "hemp fiber stroke," "raindrop stroke," "axe stroke," and "folded-belt stroke." These "ts'un" concept vocabulary are very important in cultural environments and art institutions such as East Asian museums and art galleries, but for users of Western cultural backgrounds, their importance is relatively insignificant. In order to gauge the usage of the term "ts'un," our project examined the webpages of large American and European art institutions to see what terms they used to describe their East Asian art collections. We found that for Western museums with larger collections of Chinese paintings, including the Metropolitan Museum of Art and Cleveland Museum of Art's online collections, "ts'un" was being used in the metadata of Eastern art objects (Table 3). Therefore, we proposed the addition of "ts'un" with ten sub-entries under "brushwork" in the "Activities Facet."

4.2 Re-arranging groups of terms

In the AAT's current structure, all vocabulary included under a "guide term" are organized according to A-Z alphabetical order, which in certain situations can fail to reflect the detailed relationships within a group of terms. Under this standardized order, it is easy to lose the conceptual structure originally present in certain groupings of cultural vocabulary. Taking as an example the Chinese "ritual vessels" [ID: 300265801] under "ceremonial containers" [ID: 300197 576], in the "Objects Facet" [ID: 300264092], all 19 of its sub-terms are listed in A-Z alphabetical order. But in research on traditional Chinese Shang dynasty bronzeware, the preferred categorization of bronzeware is according to their practical functions, such as division into wine vessels, water vessels, food vessels, bronze musical instruments, etc. These relationships in the original Eastern vocabulary classification are thus lost under the current AAT ordering. To fix this, our project proposed to AAT to adjust the vocabu-

lary structure, creating the divisions of "wine vessels," "water vessels," "food vessels," faithfully restoring the relationships present in the cultural classification. (Figure 4).

4.3 Creating associative relationships between terms

Associative relationships are semantic ties or connections between concepts in the AAT that are not hierarchical or equivalent relationships. These related concepts are also called Related Terms (RT) in the terminology of the thesaurus (J. Paul Getty Trust 2018). Although the AAT had formulated over 30 types of associative relationships in its guidelines, through the course of constructing the Chinese-language AAT, our project found deficiencies in relationships (Chen, Kuo and Yang 2013). Take, for example, "Qingming Festival" [ID: 300400806] under "Chinese festivals." The Qingming Festival has connections to the activities of "spring rolls", "tomb-sweeping", and "wearing willows", but the AAT guidelines originally had no associative relationship that could connect terms under the "Activities Facet," thus precluding these terms from being connected. To supplement the existing associative relationships, our project thus proposed the addition of new activity-activity relationship types: "2203 activity -associated with- activity," "2205 activity -causes- activity," and "2206 activity -is required- activity," which were published in the AAT in October 2015.

5.0 The applications of the Chinese AAT

The AAT is a multilingual, hierarchically structured vocabulary. In addition to providing knowledge retrieval functions, it also assists digital humanities research, digital curation, and interoperable information exchange.

Currently, the AAT-Taiwan has 35,450 entries, including 34,419 Translated and 2,990 added, arranged in the seven facets of AAT: Associated Concepts, Physical Attributes, Styles and Periods, Agents, Activities, Materials, and Objects (Figure 5). [2]

The project's own applications of the AAT can be seen in the following major areas.

Type of brushstroke (ts'un)	National Palace Museum (Taipei)	Metropolitan Museum of Art	Cleveland Museum of Art
axe-cut	v	v	
folded-belt	v	v	v
hemp fiber	v	v	
ox-hair	v		
Rain drop	v	v	

Table 3. Inclusion of ts'un terms in institutional websites.

276

Knowl. Org. 48(2021)No.4

S. S.-J. Chen. Methodological Considerations for Developing the *Art & Architecture Thesaurus* in the Chinese Language ...
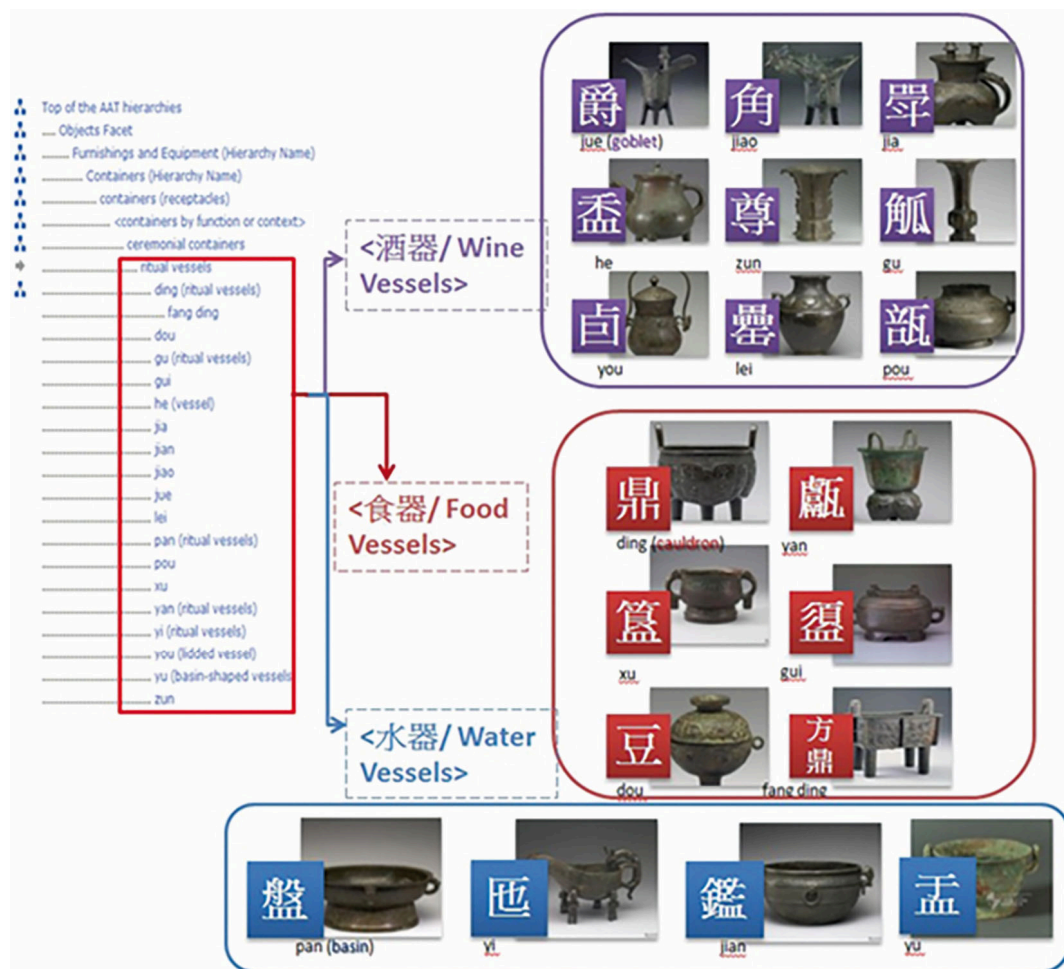
Figure 4. Revised AAT vocabulary structure proposed for Chinese "ritual vessels," divided into "wine vessels," "food vessels," and "water vessels" based on function.
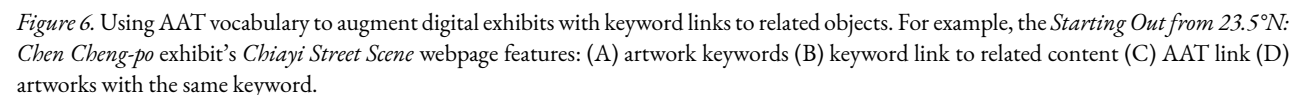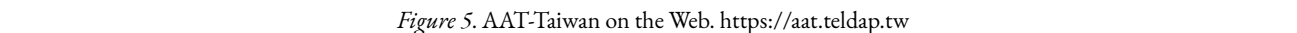
5.1 Usage in digital curation

Thanks to its rich terminology, the AAT can be utilized in the digital exhibition of artworks as a controlled vocabulary and to standardize data retrieval. It can also be provided to audiences as a keyword glossary, allowing them to link to different sections of an exhibit, expanding contents and functionality. For example, the Starting Out from 23.5°N: Chen Cheng-po (https://chenchengpo.asdc.sinica.edu.tw/index) online exhibition uses about 70 oil paintings, watercolors, sketches, and other artworks to showcase different stages in the artist's career. Each artwork uses AAT vocabulary, following the three-level theory of iconology developed by Erwin Panofsky (1892-1968). The first level corresponds to the "natural subject matter" illustrated in each image, which then constitute that image's basic keywords. For instance, Chen Cheng-po's 1934 painting Chiayi Street Scene contains the AAT terms "townscapes" [ID: 300236227], "streetscapes" [ID: 300249 570], "utility poles" [ID: 300006446], "signboards" [ID: 300 207704], "parasols" [ID: 300046218], "staffage" [ID: 30026

4353], etc. These AAT terms on the webpage are directly linked to the corresponding term on the AAT-Taiwan webpage, facilitating audiences' understanding of art-specific vocabulary and meanings, thus adding knowledge value. Furthermore, a list of other Chen Cheng-po artworks containing the same AAT keywords is also integrated and displayable on the webpage, extending access to other high-resolution images in the collection and encouraging viewers to explore the exhibit (Figure 6).

5.2 Adding value to Linked Open Data

Since 2014, AAT has become a major player in the semantic web, facilitating exchange, retrieval, and reuse in the Linked Open Data (LOD) format, and enabling digital humanities institutions worldwide to create data linkages and add value to their work. The AAT dataset is currently the most widely reused LOD resource in the digital humanities research field. The Academia Sinica Center for Digital Cultures (ASCDC) established a Linked Open Data Lab (LODLab)

Knowl. Org. 48(2021)No.4
S. S.-J. Chen. Methodological Considerations for Developing the *Art & Architecture Thesaurus* in the Chinese Language ...

277

Figure 5. AAT-Taiwan on the Web. https://aat.teldap.tw

Figure 6. Using AAT vocabulary to augment digital exhibits with keyword links to related objects. For example, the *Starting Out from 23.5°N: Chen Cheng-po* exhibit's *Chiayi Street Scene* webpage features: (A) artwork keywords (B) keyword link to related content (C) AAT link (D) artworks with the same keyword.

278

Knowl. Org. 48(2021)No.4

S. S.-J. Chen. Methodological Considerations for Developing the *Art & Architecture Thesaurus* in the Chinese Language ...

Types of Reused Linked Open Vocabularies	ASCDC's LOD Datasets (by Subject)					
	Biology	Anthropology	Religion	Art	History	Audio-video
Darwin core	●					
DBpedia	●	●	●	●	●	●
DC/ DCterms	●	●	●	●		●
EDM	●	●	●	●		●
GeoNames	●					
AAT	●	●	●	●	●	●
TGN				●	●	
ULAN				●		
LOC					●	
Linked Open Copac and Archives Hub (LOCAH)				●		
Temporal Gazetteer (TGAZ, Harvard)				●	●	
VIAF	●	●	●	●	●	●
ASCDC	●	●	●	●	●	●
Linked Taiwan Artists (LTA)				●		
Database for Names and Biographies (DNB)				●	●	

Table 4. Mapping of the Reused Linked Open Vocabularies to the ASCDC's LOD Datasets. (The bold line highlights LOD datasets enriched by reusing AAT term resources.)

in 2018 to convert datasets from various cooperating projects to LOD. It currently comprises 13 LOD datasets in the 6 areas of biology, anthropology, religion, art, history, and audio-video. In the process of data conversion, consideration was given to the specific needs and characteristics of each cooperating project's content, so that the resultant data could correspond and link to AAT LOD (Table 4). In this way, LOD's cross-institutional, knowledge-sharing characteristics added value by supplementing the original data and stimulating possibilities for new research and interpretation. Taking as an example the semantic data conversion of the Paintings and Documents of Chen Cheng-po archive, each artwork entry in the archive contained data on materials, techniques, etc. that could correspond and link to AAT terms, allowing users to search AAT term definitions to understand the meaning of specialized vocabulary, in effect expanding the data content.

5.3 Supporting digital humanities research

The AAT itself is an open dataset structured by semantic relationships. Its structured data content allows external resources to link to AAT vocabulary, thus assisting digital humanities research and inspiring new knowledge connections

and possibilities for deeper research. For example, a previous study applied AAT to the converted linked data research on "Paintings and Documents of Chen Cheng-po" from the Archives of the Institute of Taiwan History in Academia Sinica (Chen 2019), and found that although the Archives had the clippings of over 90 Western artists' artworks that Chen Cheng-po had collected over his lifetime, there was nothing in the Archives attributing the artists' nationalities. Since Chen Cheng-po's artworks were mainly painted in Western style, the collected artists' nationalities would faithfully show the influence of artists from different countries and regional styles, as well as illustrate developments in Western painting. Because the AAT's Styles and Periods Facet contains terms for nationalities, during the process of converting the Archive's "Artwork Clippings of Western Artists" file into semantic data, the names of artists were extracted from the metadata and matched with names in the Getty Center's Union List of Artists Names (ULAN) and AAT, thereby systematically identifying nationality data to enable statistical analysis and data visualization. It was discovered from the data that, for artists whose nationality could be found, most were French (26 artists), followed by Italian (7), Dutch (4), German (3), etc. These results reflect how the creation and development of Chen Cheng-po's Western style was influenced by French
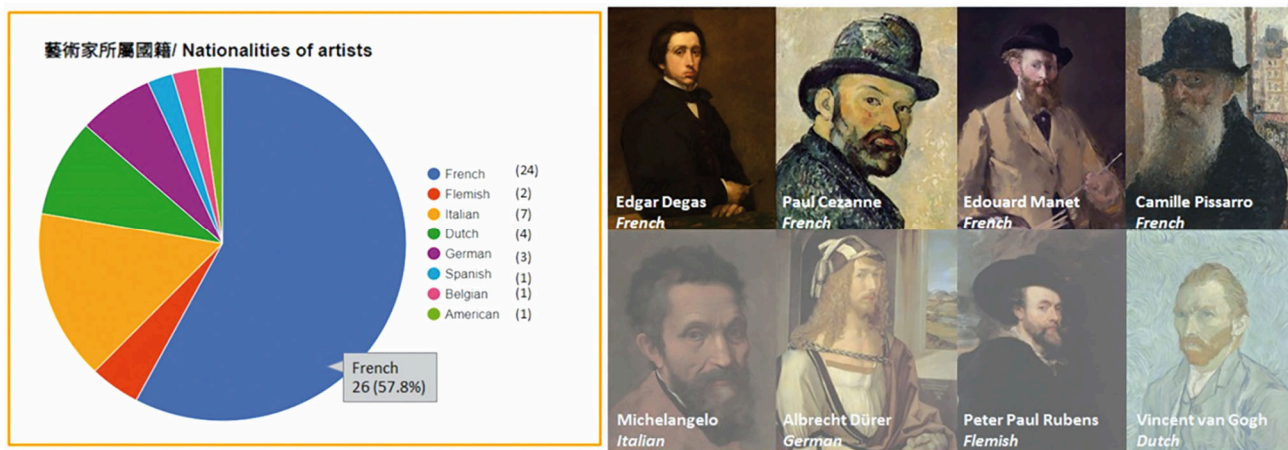
Knowl. Org. 48(2021)No.4

279

S. S.-J. Chen. Methodological Considerations for Developing the *Art & Architecture Thesaurus* in the Chinese Language ...

Figure 7. Using AAT vocabulary for nationalities in order to assist querying and statistical analysis of Western authors in art clippings from the "Paintings and Documents of Chen Cheng-po" archive file, thereby identifying Chen Cheng-po's artistic influences.

art, which was indeed a reflection of France being a focus of international art movements during Chen Cheng-po's time, attracting attention and emulation from artists all over the world. Thus, we see how borrowing the controlled vocabulary of AAT can assist the process of digital humanities research (Figure 7).

6.0 Conclusion

In view of the fact that global digital collections are developed by different organizations with different focuses and purposes, the structures and levels of detail in the metadata and controlled vocabularies they use are also different. This has resulted in digital collections using different terms to express the same concept, preventing users from finding what they need from a single search; some collections even lack controlled vocabulary to index their contents. This project is based on the actual needs of international cooperative research in developing a methodology suitable for an English-Chinese multilingual thesaurus in the field of art. The results of this research will serve as best practices to implement bilingual thesauri for the exchange of Chinese and English terminology in the art domain. The work done through mapping and analysis of the Chinese and English lexical semantics will be further linked to other languages, including French, Spanish, German, Italian, Dutch and other globally important languages. In addition, this research adopts the perspective of English users who use Chinese art resources, and Chinese users who use Western art resources when carrying out the translation of controlled vocabularies between Chinese and English. Once fully implemented, it is expected to resolve the problems faced by major Chinese art collection institutions worldwide in describing their collections, which have arisen due to the lack of controlled vocabulary in this field; in addition, it will provide principles for reference in similar studies.

Notes

1. The word "retablos" is in fact of Spanish origin and borrowed by AAT as preferred term to describe the small religious panel paintings produced in Latin American countries, seen as term "retablos" in AAT, online resource: http://vocab.getty.edu/page/aat/300262837

2. The translation project of AAT into Traditional Chinese is a long-term engagement and cannot finish all terms within a few years. Due to the limited budget of the AAT-Taiwan editorial group, ASCDC can only execute the translation of some hundred terms per year. This gradual process leads to the discrepancy between the numbers of entire entries and translated terms.

References

Academia Sinica Center for Digital Cultures. 2021. Art & Architecture Thesaurus Taiwan. Retrieved from https://aat.teldap.tw/

Andree, Karen and Reem Weda. 2016. "The Dutch Art & Architecture Thesaurus Put into Practice: the Example of Anet, Antwerp." In *Proceedings of the 2016 International Conference on Dublin Core and Metadata Applications*: 16.

Baca, Murtha. 2013. "Multilingual AAT and Digital Art History Work at the Academia Sinica." *Visual Resources* 29(4): 273-5.

Baca, Murtha and Melissa Gill. 2015. "Encoding Multilingual Knowledge Systems in the Digital Age: The Getty Vocabularies." *Knowledge Organization* 42(4): 232-43.

Barité, Mario. 2018. "Literary Warrant." *Knowledge Organization* 45(6): 517-36.

280

Knowl. Org. 48(2021)No.4
S. S.-J. Chen. Methodological Considerations for Developing the *Art & Architecture Thesaurus* in the Chinese Language ...

Barité, Mario. 2020. "Towards a General Conception of Warrants: First Notes." *Knowledge Organization* 46(8): 647-55.

Brown, Jessica S. 1987. "Computerising Art and Archaeology Technical Abstracts: Present Possibilities and Future Potential." *In ICOM Committee for Conservation: 8th Triennial Meeting*, Sydney, Australia, 6-11 September, 1987. Preprints Vol. 1: 235-7.

Chambers, Deborah. 2020. *Cultural Ideals of Home: The Social Dynamics of Domestic Space*. New York: Routledge.

Chapman, Emma 2012. *Speaking the Same Language: Using Controlled Vocabularies to Search Museum Collections Databases.* (Master's dissertation, Victoria University of Wellington, Wellington).

Chen, Shu-Jiun. 2019. "Semantic Enrichment of Linked Archival Materials." *Knowledge Organization* 46(7): 530-47. doi:10.5771/0943-7444-2019-7-530.

Chen, Shu-jiun, Ching-ju Cheng and Hsueh-hua Chen. 2011. "Methodologies for Multilingual Information Integration in the Domain of Chinese Art." *IFLA Journal* 37(4): 296-304.

Chen, Shu-jiun and Hsueh-hua Chen. 2015. "Lexical-Semantic Mapping Between Chinese and English Controlled Vocabularies in the Domain of Chinese Art." *Journal of Library and Information Studies* 13(2):161-208. Retrieved from https://ediscilinas.usp.br/pluginfile. php/5500894/mod_resource/content/1/v13-2-s3.pdf

Chen, Shu-Jiun, Chiao-Chen Kuo and Ya-Pei Yang. 2013. "Making Chinese Art Accessible to Western Users: A Brief Report from AAT-Taiwan Project." In *International Terminology Working Group (ITWG) Meeting, 7-10 January 2013, Getty Research Institute, Los Angeles, California, USA*. Making Chinese Art Accessible to Western Users- A Brief Report from AAT Taiwan Project - [PPTX Powerpoint] (fdocuments.net)

Chen, Shu-jiun, Marcia Lei Zeng and Hsueh-hua Chen. 2016. "Alignment of Conceptual Structures in Controlled Vocabularies in the Domain of Chinese Art: A Discussion of Issues and Patterns." *International Journal on Digital Libraries* 17(1): 23-38.

Dijkshoorn, Chris, Lizzy Jongma, Lora Aroyo, Jacco Van Ossenbruggen, Guus Schreiber, Wesley Ter Weele and Jan Wielemaker. 2018. "The Rijksmuseum Collection As Linked Data." *Semantic Web* 9(2): 221-30.

Doerr, Martin. 2004. *Semantic Interoperability: Theoretical Considerations*. TR 345, Institute of Computer Science, FORTH, 2004. Retrieved from: ftp://139.91.151.170/ tech-reports/2004/2004.TR345_Semantic_Interopera bility_Theoretical_Considerations.pdf

Franzon, Eric. 2014. "First of Four Getty Vocabularies Made Available as Linked Open Data." Retrieved from https://www.dataversity.net/first-four-getty-vocabular ies-made-available-linked-open-data/

Getty Research Institute. 2017. *Art & Architecture Thesaurus*. Retrieved from https://www.getty.edu/research/ tools/vocabularies/aat/

Harper, Douglas. 2021. *Online Etymology Dictionary, s. v. "parlor*. Retrieved from https://www.etymonline.com/ word/parlor

Harpring, Patricia. 2010. "Development of the Getty Vocabularies: AAT, TGN, ULAN, and CONA." *Art Documentation: Journal of the Art Libraries Society of North America* 29(1): 67-72.

Harpring, Patricia, Joan Cobb and Gregg Garcia. 2013. *ITWG 2013 Editorial and Technical Issues*. Retrieved from: http://citeseerx.ist.psu.edu/viewdoc/download?doi =10.1.1.392.7477&rep=rep1&type=pdf

International Organization for Standardization. 2013. *ISO 25964-2:2013: Information and documentation -- Thesauri and interoperability with other vocabularies -- Part 2: Interoperability with other vocabularies.*

J. Paul Getty Trust. 2018. *Editorial Guidelines: Art & Architecture Thesaurus (AAT)*. Retrieved from: https://www. getty.edu/research/tools/vocabularies/guidelines/index. html#aat

J. Paul Getty Trust. 2020a. *AAT: Frequently Asked Questions*. Retrieved from https://www.getty.edu/research/ tools/vocabularies/aat/aat_faq.html

J. Paul Getty Trust. 2020b. *About the AAT*. Retrieved from https://www.getty.edu/research/tools/vocabularies/aat/ about.html

Liang, Anita C. and Margherita Sini. 2006. "Mapping AGROVOC and the Chinese Agricultural Thesaurus: Definitions, Tools, Procedures." *New Review of Hypermedia and Multimedia* 12(1): 51-62. doi: 10.1080/ 13614560600774396

López, Facundo José, P. Martín Lerones, J. Llamas, J. Gómez-García-Bermejo and E. Zalama. 2018. "Linking HBIM Graphical and Semantic Information Through the Getty AAT: Practical application to the Castle of Torrelobatón." In *IOP Conference Series: Materials Science and Engineering* 364(1): 012100. https://iopscience.iop.org/article/ 10.1088/1757-899X/364/1/012100/pdf

Molholt, Pat. 2001. "The Art and Architecture Thesaurus: Controlling Relationships Through Rules and Structure." In *Relationships in the Organization of Knowledge* edited by Carol Bean and Rebecca Green. Dordrecht: Springer, 153-70.

Molholt, Pat and Toni Petersen. 1993. "The Role of the 'Art and Architecture Thesaurus' in Communicating About Visual Art." *Knowledge Organization*, 20(1):30-4.

Mote, Frederick. W. 1960. "Confucian Eremitism in the Yüan Period." In *The Confucian Persuasion*, edited by Arthur F. Wright. Stanford, CA: Stanford University Press, 202-40.

Knowl. Org. 48(2021)No.4

S. S.-J. Chen. Methodological Considerations for Developing the *Art & Architecture Thesaurus* in the Chinese Language ...

281

McCulloch, Emma, Ali Shiri and Dennis Nicholson. 2005. "Challenges and Issues in Terminology Mapping: A Digital Library Perspective." *The Electronic Library* 23(6): 671-7. doi: 10.1108/02640470510635755

Petersen, Toni. 1990. "Developing a New Thesaurus for Art and Architecture." *Library Trends* 38(4): 644-58.

Pustejovsky, James. 1995. *The Generative Lexicon*. Cambridge, MA: MIT Press.

Romeo, Lauren, Sara Mendes and Núria Bel. 2013. "Using Qualia Information to Identify Lexical Semantic Classes in an Unsupervised Clustering Task." *arXiv preprint arXiv:1303.2449*. https://arxiv.org/abs/1303.2449

Soergel, Dagobert. 1995. "The Art and Architecture Thesaurus (AAT): A Critical Appraisal." *Visual Resources* 10(4): 369-400.

Tudhope, Douglas and Ceri Binding. 2018. "A Study of Multilingual Semantic Data Integration." In *DCMI'18 Proceedings of the International Conference on Dublin Core and Metadata Applications, Porto, Portugal, September10-13 2018.* DCMI, 27-8. https://dcpapers.dublincore.org/pubs/article/view/3978/2180

Tudhope, Douglas, Traugott Koch and Rachel Heery. 2006. *Terminology Services and Technology. JISC State of the Art Review*. http://www.ukoln.ac.uk/terminology/TSreview-jisc-final-Sept.html

Velios, Anthanasios. 2016. "Online Event-Based Conservation Documentation: A Case Study from The IIC Website." *Studies in Conservation* 61(1): 13-25.

Zeng, Marcia Lei and Lois Mai Chan. 2004. "Trends and Issues in Establishing Interoperability Among Knowledge Organization Systems." *Journal of the American Society for Information Science and Technology* 55(5): 377-95.