

# **SILKNOW. Designing a Thesaurus about Historical Silk for Small and Medium-Sized Textile Museums.**

A. León

*Garin 1820 S.A., Moncada, Spain*

M. Gaitán

*Art History Department, Universitat de València, Spain*

I. Insa

*PhD candidate, Art History, Universitat de València, Spain*

J. Sebastián

*Art History Department, Universitat de València, Spain*

E. Alba

*Art History Department, Universitat de València, Spain*

**ABSTRACT:** The cultural heritage domain in general and silk textiles in particular, are characterized by large, rich and heterogenous data sets. Silk heritage vocabulary comes from multiple sources that have been mixed up across time and space. This has led to the use of different terminology in specialized organizations in order to describe their artifacts. This makes data interoperability between independent catalogues very difficult. Moreover, the interaction level of existing resources is low, most complex queries are not possible and results are poorly shown. In this regard, a recent EU-funded research project titled SILKNOW is building a multilingual thesaurus related to silk textiles. It is being carried out by experts in textile terminology and art historians, and computationally implemented by experts in text mining, multi-/cross-linguality and semantic extraction from text. This paper presents the rationale behind the realization of this thesaurus.

## 1 BACKGROUND

The cultural heritage domain is characterized by large, rich and heterogeneous data sets where different organizations use different terminology to describe their objects (Amin et al., 2010; Owens and Atherton Cochrane, 2004). GLAMs and Cultural Heritage institutions aim to conserve and disseminate their collections; in order to do so, prior knowledge is extremely important. The registration of a cultural asset in an inventory or its inclusion in a catalog assumes its recognition as an element that requires conservation and protection. Controlled vocabularies stand out as essentials to provide access to museum collections not only to inside users (curatorial staff, conservators, education department), but also to external users who wish to know more about a subject without knowing the specific term of its search (Baca, 2004). Nowadays, Cultural Heritage institutions strive to obtain controlled vocabularies based on their own collections (Schreiber et al., 2008), such as the The Textile Museum Thesaurus from the Textile Museum in Washington, D.C. or the Museon Arlaten (Arles, département des Bouches-du-Rhône). While specific thesauri are useful for each institution, the result is a multitude of vocabularies in different languages that are difficult to standardize, especially when there is a strong need not only to make heritage accessible to the general public but also to facilitate exchanges across collections and institutions (Isaac et al., 2007). In this regard, some efforts have been made such as the Getty AAT (<https://www.getty.edu/research/tools/vocabularies/aat/>) or the UNESCO thesaurus. Concerning silk heritage terminology, it comes from various sources that have changed across time and space. Moreover, it changes according specialties (weavers or historians), nationalities (Europe or North

America), or disciplines (ethnographic specialists vs art historians), etc. (Anderson, 2006). For example, local variations of a term are rarely taken into account (e.g. *espolín* has different meanings in some regions of Spain). To meet these challenges, SILKNOW is building a multilingual thesaurus dedicated to the specific vocabulary of historic silk textiles, which will also include local term variants. The thesaurus will help heritage institutions to provide access to and preserve silk heritage in the digital environment. Participating and collaborating institutions will radically improve their cataloging practices and digital data retention. In addition, the thesaurus will serve as an example of the benefits of shared cataloging frameworks and data interoperability.

## 2 THE SILKNOW THESAURUS

### 2.1 *Objectives*

The SILKNOW thesaurus aims to improve silk knowledge by building a web portal that will allow advanced search and representation of information related to this heritage. It intends to use as many scientific references as possible, in order to generate quality research that eventually will allow investigators to expand their studies. This thesaurus is targeted to researchers, students and cultural heritage professionals. For example, a researcher may use the thesaurus to connect terms that she or he could have found in historical documentation, providing updated, standard naming for these terms. In addition, when planning exhibitions, a curator often needs to either write the exhibition catalogue or ask for a loan to another museum. In these cases, artworks coming from other countries can be identified in their vernacular languages. Having a thesaurus will help curators to standardize terminology. Art history and conservation students will be familiarized with a variety of textile terms. It will also help to disseminate scientifically accurate terms that come from a standardized vocabulary whose final purpose is to improve the understanding, conservation and dissemination of silk heritage.

### 2.2 *Methodology*

To compile the thesaurus, inductive and deductive methods were undertaken (Nielsen 2004). Around 80% of inductive methodology was used. This means that terms were included in the thesaurus as soon as they were found in the literature. The other 20% was deductive, based on museum records and previous knowledge from the researchers. In order to be as accurate as possible, the compilers consulted a number of sources, mainly written sources that can be divided into: specialized textile dictionaries that cover the domain of silk heritage; general dictionaries that were useful when defining Preferred Terms (PT) and synonyms; theses related to historical fabrics; books about silk and textiles that had glossaries with specific terminology; historical sources that helped the compilers to define historical terms and their evolution over time and their current use; and, finally, other thesauri that, albeit quite generic, in some cases they were useful to specify certain terms. These terms were added manually by the researchers and adapted into SKOS format by automatic, computer-based extraction.

Next, terms and concepts were controlled and described by adding scope notes, qualifiers and synonyms. Regarding concepts, a PT was chosen to refer a unique concept, these terms are based in specialized literature. Whenever a term had more than one meaning (polysemy), all its meanings were included. In some cases, polysemy was not easy to undertake, since some concepts may refer to a weaving technique and to a fabric at the same time, such as *satin* or *twill*. Adding qualifiers to those terms was the best solution for these problems. Since there are no specialized

silk thesauri, when entering terms the degree specificity was as the highest possible. Terms were included based on an estimation about those that will most likely appear in museums' catalogues, taking into account that not all museums catalog the same way, and that the thesaurus may not be used only by curators, but also by conservators. In order to make clearer what those concepts meant, scope notes were added following specialized literature. Special care was taken to follow the same syntactic style. As far as possible, compilers tried to add the etymology of the words and describe their uses in time and space. Finally, these definitions were reviewed by international experts.

The next logical step was to categorize those terms. The SILKNOW thesaurus is based on the Getty AAT structure. Following it, we have established three different relationships.

- 1) Hierarchical: when the relationship between terms is broader and narrower, SILKNOW used the 7 facets and subfacets of the AAT. When possible, parents were also placed according to the AAT structure. As silk heritage terminology is extensive and not easy to classify, compilers had to add new guide terms and subfacets in order to place some terms with maximum accuracy.
- 2) Equivalence: this relationship was appropriate when different names refer to the same concept as they are synonyms or quasi-synonyms.
- 3) Associative relationships: used when some terms are conceptually very related, but not hierarchically.



Figure 1. The SILKNOW Thesaurus.

### 2.3 *A multilingual thesaurus*

The SILKNOW thesaurus is symmetrical, which means that all terms needed to be translated. First steps included the direct translation of terms, synonyms and associated concepts. As it was done with the master spreadsheet, translators (who also are textile specialists) used specialized sources, which in some cases provided translations in other languages (such as the Castany Saladrigas dictionary, 1949). Some issues arose when translating some terms; for example, some terms only existed in one language, or needed more than one translation.

### 2.4 *Extension*

The current version of the SILKNOW thesaurus has 545 terms as per May 2019. The implementation of the multilingual web-based thesaurus is due in 2020. When the thesaurus is completed it will be shared with the AAT in order to include new terms to their vocabulary. A first coverage calculation was made with 2 different Spanish databases in order to find out if the terms included in the thesaurus could be found in museum records. The result was a 60.8% of matches between the thesaurus terms and those two databases.

### 3 CONCLUSIONS

Conservation of cultural heritage includes a set of direct and indirect actions in order to ensure the physical persistence of objects; therefore, prior knowledge is extremely important. The registration of a cultural asset in an inventory or its inclusion in a catalog assumes its recognition as an element that requires conservation and protection. Nowadays, museums around the world are generating tools that allow the development of a systematic and coherent cataloging of museum collections and the improvement of a standardized model, in order to avoid the lack of common criteria when dealing with these kinds of records.

The SILKNOW thesaurus is designed to generate synergies across European textile museums. It will allow not only to connect various textiles located in diverse collections but also to trace similar artistic trends. Moreover, it will help to standardize a wide and fragile heritage, hence enhancing its conservation. Understanding that conservation requires proper management is essential. SILKNOW, as an open access tool, seeks to have a strong impact not only for museum collections and researchers, but also among the general public. If society becomes more engaged with its heritage, this will lead to the development of a true awareness and defense of heritage, turning it into an agent of innovation (Bakhski et al., 2008). Society must become guardian of its own heritage, facilitating the transmission between the past and the future (Andreu Pintado, 2014).

This is the first step to generate a common framework of knowledge and dissemination of historical textiles in general, and silk heritage in particular.

#### 3.1. *References*

- Amin, A., Hildebrand, M., van Ossenbruggen, J., Hardman, L., 2010. Designing a thesaurus-based comparison search interface for linked cultural heritage sources. Proc. 15th Int. Conf. Intell. User Interfaces IUI'10, 249–258.
- Anderson, C.G., 2006. Talking about Textiles: The Making of the Textile Museum Thesaurus. Textile Society of America Symposium Proceedings 302.
- Andreu Pintado, J., 2014. Arqueología en directo: canales de comunicación y transferencia de resultados en la investigación sobre patrimonio arqueológico: la ciudad romana de los Bañales un castillo en Zaragoza. Visibilidad Divulg. Investig. Desde Las Humanidades Digit. Exp. Proy., 17-41.
- Baca, M., 2004. Fear of Authority? Authority Control and Thesaurus Building for Art and Material Culture Information. Cat. Classif. Q. 38, 143–151.
- Bakhski, H., McVittie, E., Simmie, J., 2008. Arts and Humanities Research and Innovation. Arts and Humanities Research Council.
- Isaac, A., Zinn, C., Matthezing, H., Van de Meij, H., Schlobach, S., Wang, S., 2007. The value of usage scenarios for thesaurus alignment in cultural heritage context, in: Cultural Heritage on the Semantic Web Workshop at the 6th International Semantic Web Conference (ISWC 2007).
- Owens, L.A., Atherton Cochrane, P., 2004. Thesaurus Evaluation. Cat. Classif. Q. 37, 87–102.
- Schreiber, G., Amin, A., Aroyo, L., van Assem, M., de Boer, V., Hardman, L., Hildebrand, M., Omelayenko, B., van Ossenbruggen, J., Tordai, A., Wielemaker, J., Wielinga, B., 2008. Semantic annotation and search of cultural-heritage collections: The MultimediaN E-Culture demonstrator. Semantic Web Chall. 20062007 6, 243–249.