

**Elaine Ménard**

**School of Information Studies, McGill University, Montréal, Québec**

**Sabine Mas**

**EBSI, Université de Montréal, Montréal, Québec**

**Inge Alberts**

**Cogniva Information, Gatineau, Québec**

# **Faceted Classification for Museum Artefacts: A Methodology to Support Web Site Development of Large Cultural Organizations**

**Abstract:** This research project aimed to provide a visual representation of the Artefacts Canada digital collection. The taxonomy creation involved the implementation of an innovative four-step methodology. The resulting user-friendly bilingual taxonomy that will provide worldwide visitors with the mean to better access Canadian virtual museum collections is presented.

**Résumé :** Ce projet de recherche a pour objectif de représenter visuellement la collection numérique d'Artefacts Canada. La création d'une taxinomie repose sur la mise en œuvre d'une méthodologie novatrice en quatre étapes. Sera présentée la taxinomie bilingue conviviale ainsi créée qui permettra aux visiteurs de partout un meilleur accès aux collections muséales numériques du Canada.

## **1. Problem Statement**

Nowadays, thousands of museums give access to their collection all over the world through their Web sites. The educational potential of these resources is immeasurable. Unfortunately, these Web sites are often too complex and sophisticated to be useful and usable for the majority of visitors. Too frequently, the needs of the users of these museum web sites are overlooked in the organization of the museum content. In order to minimize the aforementioned frustrations and difficulties, faceted classification presents great potential in providing the users with simple, flexible and scalable information access points. However, there are few methodological tools available for museums who wish to adopt such approach in the development of their Web sites.

## **2. Research Objectives**

The research project aimed to provide a visual representation of the Artefacts Canada digital collection, as well as a means for the majority of users to browse this content. *Artefacts Canada Humanities* is a database containing approximately 3.5 million records describing the different collections of Canadian museums. This database is at the core of the Canadian Heritage Information Network (CHIN)'s work to make Canada's heritage accessible. CHIN's mission is to connect Canadians and worldwide audiences to Canada's heritage and to promote the development, the presentation and preservation of Canada's digital heritage content for current and future generations. In order to provide a new visual representation for this digital collection, a faceted taxonomy model was created. The taxonomy creation involved the implementation of an innovative methodology.

### 3. Methodology

A four-step methodology was adopted for the development of the faceted taxonomy model. First, a Best Practice Review consisting in an extensive analysis of existing terminology standards in museum communities and public Web interfaces of large cultural organizations was performed. The second step of the methodology entailed a Domain Analysis, consisting in extracting and comparing relevant concepts from terminological authoritative sources. Thirdly, we proceeded to Term clustering & Entity Listing which involved the breaking up of the taxonomy domains into potential facets. An Incremental User Testing was also realized in order to validate and refine the taxonomy components (facets, values, and relationships).

The best practices review consisted in an extensive analysis of existing terminology standards in Canadian museum community and public web interfaces for the collection databases of large cultural organizations. The objectives were the following: 1) to acquire knowledge of the industry's terminology standards; 2) to assess how controlled vocabularies can be adapted in the development of the new taxonomy; and 3) to examine Web interfaces for the collection of large cultural organizations to gather important usability criteria prior to the taxonomy's development.

Since a well defined set of facets can only provide a conceptual representation of one single domain, domain analysis was conducted prior to facet selection. In the case of the museum's collection, the domain corresponds to the perspectives adopted in describing a cultural artefact. For this project, the objective of domain analysis was threefold: 1) to limit the scope and the coverage of the future taxonomy while insuring the taxonomy coverage is accurate and sufficient; 2) to explore and detect potential facets; and 3) to provide recommendations on first-level facet selection. Domain analysis consisted in a mapping of selected vocabulary source categories/fields with Artefacts Canada database fields and an analysis of Artefacts Canada database's frequency reports. The analysis of the Artefacts Canada frequency reports aimed at selecting candidate facets by identifying important concepts represented in the various fields of the actual database.

The third phase in the process of the taxonomy model development involved the choices of top-level facets and their sub-facets. Two parallel approaches were combined: a bottom-up approach and a top-down approach. Starting from a list of potential facets resulting from domain analysis, the relevant values were extracted from the Artefacts Canada frequency reports and these values were clustered into a limited number of classes. These classes were the basis from which facets, sub-facets, and candidate descriptors were selected. For the taxonomy development, frequencies of the terms in the different Artefacts Canada fields were used as a form of weighting to assess a list of candidate terms. During the frequency report analysis, we realized that a saturation point was reached beyond the first 500 single terms whose frequency of use was the highest. Beyond that point, a great majority of terms were becoming repetitive. A more conceptual analysis of Canadian museum collections was also carried out, inspired from the best practice review.

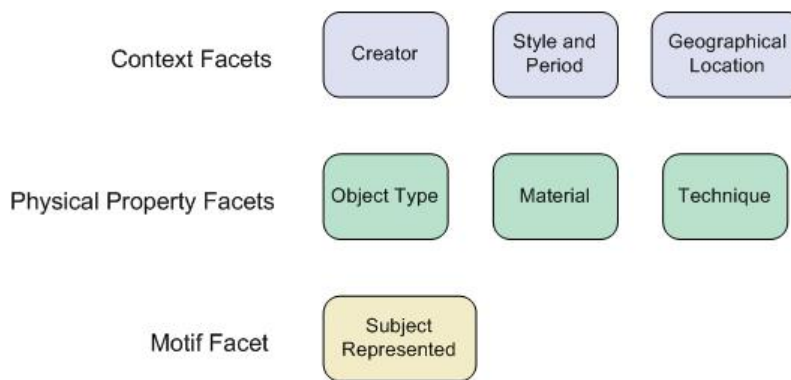
During the taxonomy development, a two-phased user testing strategy was conducted to insure the final product was clear, comprehensive, and consistent. In total, eighteen different users contributed to the development of the taxonomy. The card sorting technique was employed during the first evaluation. This data collection method is mainly used in fields such as Psychology, Cognitive Science, and Web Usability. The card sorting technique consists in a classification exercise where participants are asked to

organize a set of “cards” presenting different categories. The card sorting technique allows the confirmation of a given information structure and the understandability of the proposed labels.

The second phase of user testing entailed an evaluation of the performance of the taxonomy usage, involving a usability test under experimental conditions. First, a simulation of the retrieval process was performed. The objective of this testing was to ask a representative sample of users to complete typical retrieval tasks using the proposed taxonomy in order to measure the degree of effectiveness, efficiency, and user satisfaction. The performance testing was expected to identify usability inconveniences of the new taxonomy that may not be revealed by less formal testing. The experiment also aimed to evaluate the quality of the vocabulary, the structure of the taxonomy, and the selection of specific facets.

#### 4. Results

As depicted in Figure 1, three main perspectives have oriented the choice of the top-level facets: (1) Context Facets: Facets relating to the set of facts or circumstances surrounding the creation of the artefact; (2) Physical Property Facets: Facets that constitute properties characterizing the apparent artefact’s nature, form or function; and (3) Motif Facet: Facet representing the unifying idea being communicated or rendered by the artefact.



**Figure 1. Overview of Artefacts Canada Taxonomy Model**

These three perspectives constitute the rationale behind the designation of the seven top-level facets aiming to describe an artefact from its essential angles. The *Creator* facet refers to the entity that creates or produces the object or the work. The *Geographical Location* facet contains the name of the general regions, continents, countries and provinces where the object was created (for work of art) or use (for functional object). Within this facet, cultural information is also presented, using subdivisions such as “historical region” or “former nation”. The *Material* facet refers to what the artefact is comprised of, made with. Most of the terms regrouped in this facet represent material based on their composition or origin. The *Object Type* facet refers to all entities which can be perceived by the senses, and most specifically by the view. The *Style and Period* facet provides the names of distinct historical periods, broad cultural region styles and periods, art and architecture movements and groups & schools that are represented in the work being classified. The *Subject Represented* facet refers to the iconographic content to describe the subject or image of the object whether the optical counterpart exists in reality or in imagination. The *Technique* facet represents the processes, methods and means used to produce an object and contains techniques used to produce objects according to their materials as well as more specific techniques used for creating art objects.

## **5. Contribution**

The resulting bilingual and expandable vocabulary structure will further be used to describe the Artefacts Canada database records. Three main perspectives have oriented the choice of the top-level facets: the context facets, i.e. facets relating to the set of facts or circumstances surrounding the creation of the artefact; the physical property facets, i.e. the facets that constitute properties characterizing the apparent artefact's nature, form or function; and the motif facet, i.e. the facet representing the unifying idea being communicated or rendered by the artefact. These perspectives constitute the rationale behind the designation of the seven top-level facets aiming to describe an artefact from its essential angles. This structure offers a high-level and flexible representation of the information accessible in the database. The new taxonomy also helps representing the complexity of the content and to group objects into similar concepts to classify all records of the Artefacts Canada database. The user-friendly bilingual taxonomy will provide worldwide visitors with the mean to better access Canadian virtual museum collections. For practitioners, the methodology developed within this project is a direct contribution to support Web site development of large cultural organizations.

## **6. References**

- Marty, P. F., and M. B. Twidale. 2004. Lost in gallery space: A conceptual framework for analyzing the usability flaws of museum Web sites. *First Monday*, Vol. 9 No 9, [http://firstmonday.org/issues/issue9\\_9/marty/index.html](http://firstmonday.org/issues/issue9_9/marty/index.html) (accessed 22 October 2008).
- Mills, J. 2004. Faceted Classification and Logical Division in Information Retrieval. *Library Trends* 52, no. 3: 541-570.