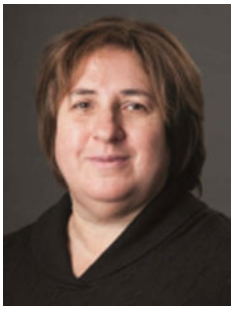
Laying the Ground for DOLMEN: Offering a Simple Standardization Starts with Understanding What Museums Do

Alexandre Fortier* and Elaine Ménard**

McGill University, School of Information Studies, 3661 Peel Street Montreal, Quebec H3A 1X1 Canada,
*<alexandre.fortier@mail.mcgill.ca>, **<elaine.menard@mcgill.ca>

Alexandre Fortier is a postdoctoral fellow under the supervision of Elaine Ménard within the context of DOLMEN, a project investigating the use of linked open data to catalogue museum artefacts. His contributions to DOLMEN will be to help model the description of museum artefacts and to study users' perception of the developed model. During his postdoctoral training, he looks forward to merging his keen interests for information organization and information behaviour. Fortier received his doctorate from the University of Western Ontario in 2016.

Elaine Ménard is Associate Professor at the School of Information Studies, McGill University, Montreal, Canada. Her teaching expertise includes cataloguing, indexing, classification and information retrieval. Her main research interests deal with cross-language information retrieval, image indexing and metadata. Ménard has published in a number of scholarly journals, including the *Journal of Information Ethics*, *Knowledge Organization*, *The Indexer, Library Hi Tech*, and *Documentation et bibliothèques*. She has presented papers at conferences such as ACFAS, CAIS-ACSI, ISKO, NASKO, ISKO France, ISKO UK, LIDA, and ASIS&T.



Fortier, Alexandre and Elaine Ménard. 2017. "Laying the Ground for DOLMEN: Offering a Simple Standardization Starts with Understanding What Museums Do." *Knowledge Organization* 44, no. 7: 485-493. 28 references.

Abstract: For most museums, online access to their collections is still a challenge. In museum databases, descriptions include descriptive metadata, along with other information that is often irrelevant to the public. Information that would help users to navigate from an object to one sharing similar characteristics is often absent. The conceptual model developed by the International Committee for Documentation, CIDOC-CRM, which provides a formal structure for linking museum objects, is still not widely adopted by institutions, due to its complexity. This project aims to provide a simpler model that could be more easily adopted. For this phase of the project, a sample of 266 Canadian museums with humanities collections (archaeology, ethnology, history, fine and decorative arts) was identified. It is composed of every museum that, during the fall of 2016, was offering to the public at least a part of its collection online. From each museum, a minimum of ten objects was selected, ensuring that the variety of the collections was represented, and extracted the metadata used in the object descriptions. This inventory, which aimed to provide a comprehensive picture of what museums already offer in terms of metadata associated to their online collections, exposed a lack of standardization and interoperability.

Received: 19 June 2017; Revised: 19 September 2017; Accepted 29 September 2017

Keywords: museums, museum objects, linked open data, information

1.0 Introduction

Through their collections, museums allow people to explore arts, cultures, history and sciences. Access to museums, however, remains strewn with barriers—physical or monetary, among others—for many. A virtual access to museum collections, even though it does not fully replicate the real experience, can palliate some of the barriers that exist between collections and those who are unable to access them. Museums, according to the Code of Ethics of the International Council of Museums (2017): "have a particular responsibility for making collections and all relevant information available as freely as possible, having regard to restraints arising for reasons of confidentiality and security." This responsibility, however, relates to a general access to collections, and museums have no obligation to afford a virtual access.

For museums, presenting virtual collections constitutes considerable challenges arising from many factors, such as the discrepancy of descriptions, lack of financial resources and database incompatibility. To complicate things further, museums often work in silos and use dif-

ferent description schemes. This lack of interoperability results in the near impossibility to exchange data among museums, which makes it harder for an institution to benefit from the work that has already been done somewhere else. Object descriptions could include multiformat information, such as text documents, images, audio tracks, videos, collection items and learning objects, but museums still struggle to make their object descriptions available to users in different forms and languages. In addition, even though museum objects possess many interesting features that could be exploited through rich descriptions, museums often have to settle for limited descriptions that only include some textual data, usually chosen for internal purposes and not necessarily for dissemination to a large public.

This paper presents results from the first phase of DOLMEN (Linked Open Data: Museums and Digital Environment), an ongoing academic research project that proposes to develop a linked open data model that will allow Canadian museums, among others, to disseminate the rich and sophisticated content emanating from their various databases. Specifically, three objectives have been established for the project:

1. To understand the characteristics necessary for the description of museum objects of any kind;
2. To define a model for the description of museum objects using linked open data; and,
3. To strengthen data exchange networks among various cultural and heritage institutions.

DOLMEN proposes to examine the fundamental elements for the description of museum objects and model them by using linked open data. This project is a stepping stone toward implementing the semantic web, as envisioned by Berners-Lee, Hendler and Lassila (2001) more than a decade ago, with the aim of making cultural heritage collections more accessible to future generations. The first phase of the study aims to better understand the elements used for the description of museum objects of all kinds within humanities collections (archaeology, ethnology and history, fine and decorative arts). More specifically, this paper attempts to answer the following research question: in general, what are the metadata elements used by Canadian museums to describe their online collections?

Recently, the use of linked open data has started to attract the attention of cultural heritage institutions (libraries, archives and museums) to provide access to their resources. For museums, linked open data offer a feasible solution to overcome the obvious lack of compatibility between different databases by establishing links between them. Linked data is not a defined standard per se, but a model for the publication of structured data on the Internet (van Hooland and Verborgh 2014). The model consists of machine-interpretable triples in which a resource (the subject) has a specific relationship (the predicate) with another resource (the object). Elements in a triple use uniform resource locators for names, allowing a browser to locate a resource, no matter where it is stored. Open data, in turn, can be freely used, reused and redistributed. Linked open data effectively takes away the constraints of existing web approaches in which web users are forced to follow a pre-defined path chosen by the hosting organization to access the descriptions of museum objects. Thus, search engines can explore a collection of web resources to provide sophisticated and complete digital content. In addition, linked open data will make it easier to offer descriptions of museum objects in more than one language, as objects can easily be linked to multilingual controlled vocabularies and authority files. The desire to transmit and share digital content requires museums to envision a collaborative work logic, both among themselves and with other data providers. Unfortunately, Canadian museums, among others, often feel helpless in the face of fast-paced technological evolution.

For this first phase of the study, an inventory of metadata used to describe online collections of Canadian museums was carried out. This examination aimed to provide a comprehensive picture of what Canadian museums already offer in terms of metadata associated with their online collections and open content that may be linked. The results of this first phase of the study provide one of the foundations of the DOLMEN linked open data model.

2.0 Background

Museum object collections can be described in many ways: by nature of objects, type of materials used, discipline, time periods, etc. Composed of several steps, the documentation of a museum collection usually begins with an extensive inventory process, that is to say, a precise and detailed object examination leading to the creation of records in which all the objects of the collection are listed and described. In a cultural institution such as a museum, the goal of inventory is to ensure administrative conservation and preservation of the identity of the objects acquired by the institution or held in custody. This process also serves to establish beyond any doubt that the object belongs to the institution or that the institution only holds the object in temporary custody. In general, a museum inventory is followed with the completion of a detailed description of the objects, similar to the cataloging process performed for a library catalogue. Given their mission of transmitting knowledge to contemporary and

future generations, museums must, to the extent of the means available to them, document collections by adding descriptive and historical information. Thus, in the museum context, "documentation" of collections is used to describe the fundamental process of cataloging (CHIN 2015).

Recording some elements relating to museum objects (such as the identification, designation, description, dimensions, materials, history, manufacturing process and function) belongs to the documentation process. The typical documentation process usually takes three standards into consideration: 1) standards in data structure, that is to say the categories of data that will be contained in a record; 2) standards in data content, that is to say the rules for recording the information (i.e., order, syntax and format for entering data values in a record.); and finally, 3) standards in data value, that is to say the vocabularies (e.g., authority files, thesauri and classification systems) used in the process of cataloging and retrieval of museum collections.

2.1 Standards in data structure and data content

The challenges of describing museum collections have been discussed for decades. The main factor that could explain the difficulties encountered by museums is the nonexistence of one universally agreed upon standard that would capture all useful information. Moreover, since museum objects vary enormously, providing established documentation standards adapted to the needs of every collection is complex. A lack of controlled vocabularies and authority lists suitable for all collections is yet another problem faced by museums, even if various international documentation standards exist and have proven their usefulness to fit the needs and requirements of many museums. Among these standards, it is worth mentioning the Information Categories from the International Council of Museums' International Committee for Documentation (CIDOC) published in 1995. These guidelines (CIDOC 1995) "can be adopted by an individual museum, national documentation organization, or system developer, as the basis for a working museum documentation system." This data structure standard comprises twenty-two information groups, each containing one or more information categories. These guidelines also suggest the use of established terminologies for elements such as object and personal names, materials and techniques, among other information units, to facilitate searching for information across collections.

Another important example of a data structure standard is the Art Information Task Force's (AITF) Categories for the Description of Works of Art (CDWA), funded by J. Paul Getty Trust. CDWA (Getty Research Institute 2014) is "a set of guidelines for the description of art, architecture, and other cultural works." This framework (Getty Research Institute 2014) "also provides a framework to which existing art information systems may be mapped, upon which new systems may be developed, or upon which data may be linked in an open environment." This standard includes 540 categories and subcategories of information elements. Among these categories, some are considered essential to describe a work while others are being considered optional. CDWA also suggests the use of authority information about persons, places, concepts and subjects that may be essential for the retrieval of the work. CDWA recommends however that this information is recorded in separate authority files rather than in records about the work itself: "The advantage of storing ancillary information in an authority file is that this information needs be recorded only once, and it may then be linked to all appropriate work records."

The Museum Documentation Association (MDA)'s SPECTRUM is another example of a data structure standard. It is free to download for non-commercial usage. It encompasses all those areas of museum activity that produce information. This standard is the result of contributions from documentation practitioners in museums throughout the United Kingdom. The current version of SPECTRUM (4.0) encompasses twenty-one collection management procedures, as well as SPECTRUM Advice fact sheets on primary procedures and related topics. Among the procedures, cataloging, for example, supposes "the compilation and maintenance of key information, formally identifying and describing objects" (Collections Trust 2017). For SPECTRUM, cataloging may also "include information concerning the provenance of objects and also collections management documentation e.g. details of acquisition, conservation, exhibition and loan history, and location history." This current version will be replaced by SPECTRUM 5.0 in 2017.

Since the 1980s, the Visual Resources Association (VRA) has published the VRA Core Categories (Core 4), a standard meant to describe images. Based on the Dublin Core metadata model, Core 4 encompasses a list of elements for art and architectural images. Core 4 is built around three record types: work, image and collection. A work is a unique event or object of cultural production (a building, a vase, a painting, a performance). An image is the visual representation of the object or event, in part or in whole (a digital image of an artwork, a photograph of a building). In "Core 4, a Work and an Image each have their own record. These records are related with the Relation attribute. The third record type, Collection, allows for collection-level cataloging of groups of materials

such as groups of works or groups of images" (Visual Resources Association 2017).

Finally, sponsored by the Visual Resources Association Foundation, *Cataloging Cultural Objects* (*CCO*) was published in 2006 as a data standard for the cultural heritage community: "The primary focus of CCO is art and architecture, including but not limited to paintings, sculpture, prints, manuscripts, photographs, built works, installations, and other visual media. CCO also covers many other types of cultural works, including archeological sites, artifacts, and functional objects from the realm of material culture." *CCO* is intended to provide rules for "describing, documenting, and cataloging cultural works and their visual surrogates." In order words, CCO not only prescribes the choice of terms, but it recommends the order, syntax and form in which data should be entered (Visual Resources Association 2006).

In Canada, the Data Dictionaries of the Canadian Heritage Information Network (CHIN 2013) provide a basis for the data structure to be used in the description of collections: "They can be used by a wide range of museums to help them to identify their institution's information needs and standardize their documentation. Each data field in the CHIN Data Dictionaries is described by a field label, a mnemonic, and a name. Fields include a definition, entry rules, related fields, a data type, examples, a discipline, authority lists, a source, and other information."

Several versions of the CHIN Humanities Data Dictionary exist. The complete Canadian Heritage Information Network's Humanities Data Dictionary includes 654 fields for collections in history, ethnology, archeological specimens and fine and decorative arts, that is to say, all fields that museums can use internally for collection management. Another version, the Artefacts Canada Humanities Data Dictionary, is a subset of 143 fields that consist of those fields that can be used for public display. Finally, a lighter version exists (a subset of forty-nine fields) that are the required and recommended fields for Artefacts Canada database contribution (CHIN 2013). The CHIN Data Dictionaries can be mapped to other similar standards for cultural heritage information (e.g., the VRA Core, SPECTRUM) or to other general standards such as Dublin Core. A correspondence can also be established with the CIDOC Conceptual Reference Model (CIDOC-CRM).

Another standard, the Info-Muse Network documentation system, is mainly used by museums in Quebec and, to some extent, elsewhere in Canada. It is developed in collaboration with museums and experts from the different scientific validation committees for the tools. Info-Muse recommends (Société des musées Québécois 2006) "a means of dividing up and organizing the data, based on the type of collection to be documented and the type of data to be recorded in each field block, and then offers a means of breaking the data down into the smallest meaningful units, known as fields."

The preceding examples constitute the most popular standards in data structure used on the Canadian and international scenes. However, these standards are often customized by museums in order to meet specific local needs. For example, additional information fields might be added to complete the description of collections, thereby giving birth to an in-house schema that perfectly fulfills local needs but that is also difficult to export and reuse in other institutions.

These standards not only recommend the elements necessary to describe museum objects, but also how these elements need to be recorded to ensure consistency. For example, a standard format for data is recommended for fields that contain names and dates. Consistency is particularly crucial, since such information is mostly used in queries employed to retrieve a specific record. Similar to cataloging rules used by libraries, such as the second edition of the *Anglo-American Cataloging Rules* or the new Resource Description and Access, museums use cataloging rules to determine how data are entered in fields: "Cataloging rules dictate the order, syntax, and format the museum uses to record data—word order, punctuation, how to record vague or unknown data, diacritics, rules for recording titles, names of people, places, and organizations, capitalization, date formats, and many other directives that make for consistent documentation" (CHIN 2017). For example, the Data Dictionaries of the Canadian Heritage Information Network (CHIN), previously mentioned, follow international conventions for data format and prescribe the entry rules by which to enter the data for easy retrieval. The format rules from the *Cataloging Cultural Objects* (*CCO*) standard have been incorporated into the CHIN Data Dictionaries (CHIN 2013).

2.2 Standards in data value

Similar to the use of controlled vocabularies in library catalogues (e.g., *Library of Congress Subject Headings*, Library of Congress Authorities), museums use authority lists to control the terms and their variants when documenting their collection. The main advantage of an authority list is the possibility to disambiguate similar or identical terms (e.g., different artists or works with the same name), or to collocate terms that belong together (e.g., an artist's name or work title in another language). In the museum context, an authority list may be used for artist names during data entry to warrant that the name is spelled consistently, or to ensure that a certain version (e.g., when they vary from one language to another) is consis-

tently used. The authority list also makes sure that "preferred" terms or names and all variants are linked so that the term can be cross-referenced. Some authorities include rich supplemental information (e.g., an artist name authority with information on the artist's dates, technique and biography). Finally, authority lists of terminology have been created for some of the key fields and adopted as standards.

Traditionally, museums make use of many different authority lists, depending on the unit of information. For example, in the context of Canadian museums, the most familiar authority lists used include:

- *Getty Art & Architecture Thesaurus* (*AAT*) commonly used for controlling terminology in a wide variety of fields, including object names, materials, techniques, cultures, time periods, and more. Most *AAT* terms are available in English only, with variants in other languages often included, although rarely in French (Getty Research Institute 2017).
- *Getty Cultural Objects Name Authority* (CONA), an authority for titles of architectural works, paintings, sculpture, etc. The CONA interface and most fields within records are available in English only. When the object comes from a culture speaking a language other than English, variations in the local language are included (Getty Research Institute 2017).
- Getty Union List of Artist Names ® (ULAN), an authority that includes proper names and biographical information about artists. The ULAN interface and most fields within records are available in English only (Getty Research Institute 2017).
- Artists in Canada, an authority available on the CHIN Professional Exchange website, includes proper names and biographical information about Canadian artists. *Artists in Canada* is available in English and French (CHIN 2016a).
- Parks Canada Classification System (CHIN 2016b), "a bilingual museum classification system and vocabulary standard used in Canada for humanities collections. It helps museums catalog collections to identify, name and classify objects using definitions and illustrations. This classification system is based on an object's original function (the purpose for which the object was created)." It is available in English and French.
- *Nomenclature 4.0 for Museum Cataloging* is a function-based classification system and vocabulary for man-made objects in museum collections. It is available in English only (AASLH 2017).

These controlled vocabularies are only examples of those used by museums. It should also be noted that many institutions develop their own controlled vocabularies.

2.3 Linked open data projects in museums

Once the description of objects is completed, museums can offer, via a search engine, access to these descriptions in whole or in part. So far, search engines available on many museum websites are mostly designed to perform simple searches in the database associated with the museum. The search results are therefore limited to the contents of this unique database, often to a limited number of descriptive metadata, such as the title or name of the object, and the name of its creator. Creating links between different databases for descriptions of museum objects would offer a range of new possibilities, almost without limits. This endeavor would be facilitated by the use of linked open data.

During the last few years, projects have started to digitize cultural heritage materials through the use of semantic web technologies (Clough et al. 2008; Dekkers et al. 2009). Some related open data projects specifically targeting the museum objects have recently been launched (Oard et al. 2014). For example, the Europeana digital library created in 2008 by the European Commission, brings together various digital resources (books, audiovisual material, photographs, archives documents, etc.) of national libraries from twenty-seven countries. Europeana Labs published its own model of open data with mapping guidelines for institutions wanting to map their data to this model (2015). The Amsterdam Museum Linked Open Data set comprises more than seventy thousand object descriptions. The institution's thesaurus and person authority files used in the object metadata are included in the linked data set. The data is mapped to the Europeana Data Model, utilizing Dublin Core, SKOS, RDA-group2 elements and the OAI-ORE model to represent the museum data. Vocabulary concepts are mapped to GeoNames and DBpedia. The two main contributions of this dataset are the inclusion of internal vocabularies and the fact that the complexity of the original dataset is retained (de Boer et al. 2012). In 2010, Japan launched the Linked Open Data for Academia (LODAC) project, bringing together fifteen of the museums in Japan and providing them with the appropriate data model to enable them to publish the RDF data and to connect to the data hub (Kamura et al. 2011). LODAC allows for an integrated multilingual access to diverse digital archives of Japanese prints. In 2014, the Smithsonian American Art Museum began mapping its museum records as linked open data as well as a growing body of related data published by organizations worldwide (Szekely et al. 2013).

Some pilot projects have also emerged in Canada. The National Network of Documentary Heritage (RPCPD) used the RDF/XML format with a sample of digital re-

sources (documents of the First World War) in order to share the metadata provided by the five partner institutions. Thus, data about several types of documents (e.g., sheet music for songs about war, photographs of battalions, war diaries), and also additional resources (e.g., sound recordings, films, portraits of Canadians Celebrities who participated in the First World War) were left "linked" to the museum objects (RPCPD 2015). However, these projects are not yet widespread, which illustrates the pressing need to carry on extensive research in this direction. Most Canadian museums demonstrate some concerns—if not apprehensions—toward the possibility of embarking in the linked open data endeavor, feeling not prepared due to a lack of means and expertise. Nevertheless, the use of linked open data offers many advantages such as "improved data visibility, data linked with external resources, easy resource annotation process and reuse of data" (Hallo et al. 2016). The DOLMEN project intends to help surmount this setback in implementing linked open data in the context of Canadian museums.

3.0 Methods

For this phase of the research project, a population of 3,133 Canadian museums with humanities collections (archaeology, ethnology and history, fine and decorative arts) was identified. The sample is composed of every museum that, during the fall of 2016, was offering to the public at least a part of its collection online. This sample comprises museums of different sizes from the ten Canadian provinces and three territories. These museums hold diversified collections, in order to offer the maximum variety of museum objects that will possibly be described by the model that will be developed afterwards.

The information (name of museum, civic address, URL, presence or absence of online collections) for museums was compiled in an Microsoft® Excel spreadsheet. Among these, only 266 museums offered online collections. From each one of these online collections, ten objects or more (depending on the heterogeneity of the collection) were selected, ensuring that the variety of the collections was represented, and the metadata used in the object descriptions were extracted. This criterion-based sampling made it possible to study a wide variety of cases rich in information (Patton 2014) and, thus, to gain a comprehensive understanding of the description practices used among Canadian museums. For this phase of the project, a degree of saturation in terms of descriptive elements was sought. In addition to metadata associated with museum objects, other elements, such as the presence of images, hyperlinks to other objects in the same collection, and to objects from other collections of the museum or to other museums, were recorded. Metadata were collected between October 1 to November 15, 2016. Figure 1 presents two examples of metadata collected from the McCord Museum online collections.

Chest of drawers	**Shoes**
Owen McGarvey	Expo 67 hostess uniform,
1870–1880, 19th century	British Pavilion
Wood, walnut, mahogany;	Roger Nelson
marble; metal; ceramic;	1967, 20th century
glass; Assembled	Leather
240 x 131 x 61 cm	Gift of British Pavilion Ex-
Gift of Mrs. Audrey Smith	po 1967
M987.66.2.A-B	M967.98.3.1-2

Figure 1. Examples of metadata extracted from museum objects.

Once metadata were collected, a content analysis approach (Weber 1985) was used to overcome two issues that arose: the variations in the labels used by different museums and the absence of labels in many collections. This approach has the advantage of providing insight into large amounts of data in order to develop research hypotheses. Content analysis has been chosen for this study because it is the most appropriate method to investigate a large set of heterogeneous data (Landry 2002). The collected metadata were analyzed on the basis of their content and grouped together using IBM® SPSS®, a dedicated computer program used for statistical analysis. This comparative analysis of the metadata allowed for the identification of any form of internal and external standardization. This analysis took place between November 15, 2016 and January 31, 2017. The results of the content analysis are presented in the next section.

4.0 Results

The results of the metadata analysis extracted from the online collections of objects revealed that 66.2% (n=176) of the museums in the sample describe their objects using structured metadata, while the remaining museums only offer a brief, textual description. A very large portion of museums, however, do not go further than a basic descriptive description, close to one based on the International Standard Bibliographic Description: title, creator, date and material description. A high number of unique terms related to the level of specificity of a particular field was also observed. Many terms represent, for instance, the title of a work (e.g., *A summer shower)*, the year or the creation date (e.g., 1934), or the particular size of an object (e.g., nine inches). Very few museums in the sample offer metadata that could link objects to one another through characteristics such as a specific origin (geographic or cultural), period, movement, technique or

object category. On the basis of 176 museums using structured metadata, twenty-five metadata types were identified. Table 1 offers a synthesis of the findings focusing on the representativeness of a metadata type in the resources analyzed.

Metadata type	Representativeness
Title or Object name	93.8% (n=165)
Date	74.4% (n=131)
Identification number	67.6% (n=119)
Description	65.9% (n=116)
Creator (person or corporate body)	50.0% (n=88)
Dimensions	45.5% (n=80)
Materials or Medium	44.3% (n=78)
Subject	38.1% (n=67)
Category of objet	34.7% (n=61)
Acquisition mode	34.1% (n=60)
Geographic origin	15.3% (n=27)
Collection	13.6% (n=24)
Copyright of image	10.2% (n=18)
Condition details	9.7% (n=17)
Image credit line	8.5% (n=15)
Cultural origin	6.8% (n=12)
Period	5.7% (n=10)
Brand	5.7% (n=10)
Technique	5.1% (n=9)
Dates linked to creator	4.5% (n=8)
Citizenship of creator	3.4% (n=6)
Biography of creator	2.3% (n=4)
History of object	1.7% (n=3)
Place of use	1.1% (n =2)
Role of creator	1.1% (n =2)

Table 1. Metadata type and frequency.

The content analysis also revealed that four types of objects are present in the collections: works of art, artisanal functional artifacts, industrial functional objects and documents. These categories are not mutually exclusive, and a collection might have more than one type of objects. The results also indicated that the frequency of use of metadata varies by object type, and that some metadata are associated with the four types while others are unique to a particular type.

Following the first phase of the analysis, commonalities among different metadata types were identified, and then examined to see whether it was possible to group them together. The frequency analysis allowed us to identify the most important concepts (and the associated metadata types) represented in the description of museum objects. Consequently, it was possible to group them into ten categories covering the main aspects of museum objects: metadata relating to the description and the composition of the object; metadata relating to the place of creation and use of the object; and metadata relating to the artistic, cultural and academic context in which the object was created. This first grouping of metadata types that appears essential to describe museum objects ("fields with a high or medium representativeness") and their definition is presented in Table 2.

Metadata	Definition
Maker/Creator/Artist/ Manufacturer/Brand	Refers to the entity that creates, manufactures or produces the object or the work of art.
Material or medium	Refers to what the work is comprised of, made with. This represents materials based on their composition or origin.
Subject	Contains identification, description (or interpretation) of what is depicted by a work or image. It may include a concept, place, an activity, an event, a person, etc.
Object name	Contains the common name of the object or work of art.
Geographical origin	Contains the name of the general regions, continents, countries and further subdivisions where the object or the work of art is created.
Cultural origin	Indicates the historical, social, economic, religious or other cultural origins of the object or the work of art.
Style and period	Provides the names of distinct historical periods, broad cultural region styles and periods, art and architecture movements and groups and schools that are represented in the object or work of art.
Technique	Represents the processes, methods and means used to produce an object or a work of art.
Copyright	Contains the name of the entity that holds the copyright of the digital image of the object or the work of art and the copyright date.
Credit	Contains the credit line or acknowledgment to be used with the digital image of the object or the work of art.

Table 2. Main metadata and definitions.

In the sample of museums examined for this study, the description of objects varied considerably from one museum website to another. The examination revealed significant disparities from one museum to another, particularly in the ways in which information is organized and

presented, and in existing searching and browsing functionalities. As a consequence, users of museum websites can quickly be overwhelmed by information excessively oriented to museum professionals (terminology or categorization of the information that is often too difficult to be understood by a non-professional) or by websites offering an interface that requires an extended learning process. Results exposed a clear lack of standardization among museums, not only the types of metadata used but also in their value. Inconsistencies were also observed in the data value of many metadata types within descriptions created by the same institution (e.g., "armchairs," "arm chairs," "bergères," "elbow chairs" and "fauteuils" to describe the same category of object in English). This adds to the complexity of exchanging data among museums, therefore multiplying the colossal task of having to produce individual descriptions for the multitude of artifacts housed in their collections.

5.0 Discussion and conclusion

The preceding inventory aimed to provide a comprehensive picture of what Canadian museums already offer in terms of metadata associated with their online collections. Museum databases show extremely heterogeneous data structures, which necessitate advanced mapping and standards integration for them to benefit from the interoperability enabled by the technologies of the semantic web. Over the years, museums around the world have built databases with metadata describing billions of objects, their history, the people who created them and the entities that represent them (de Boer et al. 2012). However, these data are stored in proprietary databases and are not universally usable and exchangeable.

Recently, museums have turned to the web with the intention of finding a solution to make their data more accessible. In the specific case of museums, structuring metadata with RDF involves the selection of one or more vocabularies to model the domain. This process is complex because many museums have various data standards that often include unique fields that reflect their own describing needs. Museum data also show many inconsistencies often because different people have kept them updated over time. The CIDOC-CRM, literally a conceptual reference model, aims to promote a common understanding of cultural heritage information by providing a common and flexible semantic framework for how information about museum objects can be related (ICOM 2017). However, this model is still not widely used by museums, possibly because of its complexity. The model, which is often regarded as one hailing from documentalists, proposes a paradigm shift that may make museum conservators reluctant to adopt it (see, for example, Szabados, Briatte and Letricot 2012).

Cultural heritage organizations such as museums stand to gain a great deal by engaging in the thoughtful management of linked open data. DOLMEN is a project that seeks to explore how the traditional description of museum objects can be simplified and possibly improved by the use of linked open data. To reach this objective, the expectations and frustrations of museum website users are studied in a parallel, ongoing phase. The first phase of the study (examination of metadata standards and an investigation of museum website users' expectations) provides a foundation to develop a linked open data model that will allow Canadian museums, among others, to disseminate the rich and sophisticated content emanating from their various databases. This will ensure the continued exchange and use of museum data in the information society, enabling the data to be utilized not only internally for museum website management but also externally by various users for education, research, learning and creative activities. This is crucial, as very few museums are ready or equipped to tackle this challenge. This could also facilitate the public's understanding of museums and their fundamental mission and could revolutionize how museums will provide information in the future. The DOLMEN project will help museums make cultural heritage collections more accessible to future generations, which is, for most museums, a fundamental part of their mission.

References

AASLH (American Association for State and Local History). 2017. "Chenhall No. 4." http://community.aaslh.org/nomenclature

Berners-Lee, Tim, James Hendler, and Ora Lassila. 2001. "The Semantic Web." *Scientific American* 284, no. 5:34-43.

CHIN (Canadian Heritage Information Network). 2013. "CHIN Data Dictionaries." https://app.pch.gc.ca/application/ddrcip-chindd/description-about.app?lang=en

CHIN (Canadian Heritage Information Network). 2015. "Introduction to Documentation of Heritage Collections." http://canada.pch.gc.ca/eng/1443534877754

CHIN (Canadian Heritage Information Network). 2016a. "Artists in Canada." https://app.pch.gc.ca/application/aac-aic/?lang=en

CHIN (Canadian Heritage Information Network). 2016b. "Parks Canada Descriptive and Visual Dictionary of Objects." https://app.pch.gc.ca/application/dvp-pvd/appli/descr-eng.php#a_propos-about

Clough, Paul, Jennifer Marlow, and Neil Ireson. 2008. "Enabling Semantic Access to Cultural Heritage: A Case Study of Tate Online." In *Proceedings of the ECDL 2008 Workshop on Information Access to Cultural Heritage, Aarhus, Denmark.European Conference on Research and Advanced Technology on Digital Libraries*, ed. M. Larson, K. Fernie, J. Oomen and J. Cigarran. Amsterdam: University of Amsterdam, Information and Language Processing Group, 978-90.

Collections Trust. 2017. "Spectrum." http://collections trust.org.uk/spectrum/

CIDOC (Committee for Documentation of the International Council of Museum). 1995. International Guidelines for Museum Object Information: The CIDOC Information Categories. http://icom.museum/fileadmin/user_upload/pdf/Guidelines/CIDOC guidelines1995.pdf

de Boer, Victor, Jan Wielemaker, Judith van Gent, Michiel Hildebrand, Antoine Isaac, Jacco van Ossenbruggen and Guus Schreiber. 2012. "Supporting Linked Data Production for Cultural Heritage Institutes: The Amsterdam Museum Case Study." In *The Semantic Web: Research and Applications; 9th Extended Semantic Web Conference, ESWC 2012, Heraklion, Crete, Greece, May 27-31, 201; Proceedings*, ed. Elena Simperl, Philipp Cimiano, Axel Polleres, Oscar Corcho, and Valentina Presutti. Lecture Notes in Computer Science 7295. Berlin: Springer, 733-47.

Dekkers, Max, Stefan Gradmann, and Carlo Meghini. 2009. Europeana Outline Functional Specification for Development of an Operational European Digital Library. In *Europeana Thematic Network Deliverable 2.5*.

Europeana Labs 2015. "Europeana Linked Open Data."

Getty Research Institute. 2014. "Categories for the Description of Works of Art." http://www.getty.edu/research/publications/electronic_publications/cdwa/index.html

Getty Research Institute. 2017. "Getty Vocabularies." http://www.getty.edu/research/tools/vocabularies/

Hallo, María, Sergio Luján-Mora, Alejandro Maté, and Juan Trujillo. 2016. "Current state of Linked Data in digital libraries." *Journal of Information Science* 42:117-27.

ICOM (International Council of Museums). 2017. "Code of Ethics." http://icom.museum/professional-standards/code-of-ethics/

Kamura, Tetsuro, HideakiTakeda, Ikki Ohmukai, Fumihiro Kato, Toru Takahashi, and Hiroshi Ueda. 2011. "Study Support and Integration of Cultural Information Resources with Linked Data." In *Proceedings of the 2011 Second International Conference on Culture and Computing, 20-22 Oct. 2011*. Washington, D.C.: IEEE Computer Society, 177-78.

Landry, Réjean 2002. "L'analyse de contenu." In *Recherche sociale: de la problématique à la collecte des données, 3e éd.*, ed. Benoît Gauthier. Sainte-Foy: Presses de l'Université du Québec, 329-56.

Oard, Douglas W., Amalalia S. Levi, Ricardo L. Punzalan, and Robert Warren. 2014. "Bridging Communities of Practice: Emerging Technologies for Content-Centered Linking." Paper presented at MW2014: Museums and the Web 2014 The annual conference of Museums and the Web, April 2-5, 2014, Baltimore, MD, USA. https://terpconnect.umd.edu/~oard/pdf/mw2014.pdf

Patton, Michael Quinn. 2015. *Qualitative Research & Evaluation Methods: Integrating Theory and Practice*, 4th ed. Thousand Oaks, Calif.: Sage.

RPCPD (Réseau pancanadien du patrimoine documentaire). 2015. "Démonstration de faisabilité de la Visualisation des Données ouvertes liées (LOD)."

Szabados, Anne-Violaine, Katell Briatte and Rosemonde Letricot. 2012. "Utiliser l'ontologie CIDOC CRM pour l'information relative au patrimoine culturel." In *THATCamp Paris 2012: Non-actes de la non-conférence des humanités numériques*. Paris: Éditions de la Maison des sciences de l'homme.

Société des Musées Québécois. 2006. "Documenting Your Collections Info-Muse Network Documentation Guide." http://www.musees.qc.ca/fr/professionnel/guidesel/doccoll/en/index.htm

Sproull, Nathalie. L. 1995. *Handbook of Research Methods: A Guide for Practitioners and Students in the Social Sciences,* 2nd ed. Metuchen, N.J.: Scarecrow.

Szekely, Pedro, Craig A. Knoblock, Fengyu Yang, Xuming Zhu, Eleanor E. Fink, Rachel Allen, and Georgina Goodlander. 2013. "Connecting the Smithsonian American Art Museum to the Linked Data Cloud." In *The Semantic Web: Semantics and Big Data; 10th International Conference, ESWC 2013, Montpellier, France, May 26-30, 2013; Proceedings,* ed. Philipp Cimiano, Oscar Corcho, Valentina Presutti, Laura Hollink, and Sebastian Rudolph. Lecture Notes in Computer Science 7882. Berlin: Springer, 593-607.

Visual Resources Association. 2006. "VRA Core." http://core.'vraweb.org/index.html.

Visual Resources Association. 2017. "The CCO Commons: Cataloging Cultural Objects." http://cco.vrafoundation.org/index.php.

Weber, Robert Philip. 1985. *Basic Content Analysis*. Sage University Papers Series. Quantitative Applications in the Social Sciences 07-049. Beverly Hills, Calif.: Sage.