

A Unified Data Model for ICH Online Inventories: A Proposal

Maria Teresa Artese, Isabella Gagliardi

CNR IMATI – Milano, Via A. Corti 12, 20133 Milan, Italy
artese@mi.imati.cnr.it, gagliardi@mi.imati.cnr.it

Abstract. This paper proposes an investigation about problems arising when searching and managing Intangible Cultural Heritage data: how to find out which websites are dealing with them in the world, how to collect and evaluate web services provided to make integrated queries, how to model a common metadata structure able to collect ICH entities coming from different realities and provide them in Linked Open Data format. The study is based on the *QueryLab* prototype, an ongoing experimentation used to integrate different inventories to provide new and easy ways to query data.

Keywords: Living Heritage, ICH Collection, Common Metadata Model, Inventory Integration, ICH Web Services.

1 Introduction

The 2003 Convention for the Safeguarding of the Intangible Cultural Heritage led to a proliferation of ICH websites collecting cultural heritage objects related to dialects, music, dance, traditional customs, knowledges, and other cultural expressions. It is a sea of information spread all over the web.

Examples of online archives can be found in all Europe countries, created after the UNESCO Convention, while Asian countries such as South Korea, Japan and China have defined strategies to safeguard their traditions much earlier.

Unfortunately, there is no easy way to find them all or a keyword that can clearly identify them, and above all they apply very different interpretations of the guidelines provided by UNESCO in cataloguing "living goods". Yet everyone is concerned with collecting and preserving the same type of information: "living heritage" or even "living culture", as defined in the UNESCO site (<http://www.unesco.org/new/en/culture/resources/in-focus-articles/safeguarding-communities-living-heritage/>).

This paper tries to investigate problems that arise in finding and managing this type of data: find out which sites are dealing with Intangible Cultural Heritage entities in the world, find a way of interaction to overcome problems related to language and the different ways in which data is stored and displayed, study a common model for data that allows to properly design an inventory to contain such entities.

2 How to Find ICH Archives on the Web

To deal with the search for websites and online archives related to intangible cultural heritage, we focused on the Unesco website at first, useful to find the lists of inscribed elements and to dive into them (<https://ich.unesco.org/en/dive>), but it is really difficult to reach the online collections starting from every single entity. The project proposed by the ICH Digital Observatory (Sousa & Imaterial, 2017), instead, focuses directly on inventories and aims to produce, organize and analyze information on the e-Inventories of Intangible Cultural Heritage, to create a network for sharing knowledge and practices on ICH available on the web (<https://digitalich.memoriamedia.net/index.php>).

One of the main results of the project is the creation and publication of the Map of ICH e-inventories, which enables direct access to 203 digital platforms, collecting and showing information about each of them, such as country, inventory name and format, language, organization holder, ICH domains and type of multimedia resources contained. Inventories are shown as points on the world map and by clicking one of them it is possible to get information, together with the link to reach the specific resource, if available online.

A first analysis of the archives collected in the map shows that, although the proposed contents are of the same type, the methods of data collection, archiving and dissemination are very different, as well as the quality of the data collected, the level of detail and the communication, i.e. whether or not they offer the translation of the site and the contents from the original language to English. We are still a long way from a standard data model and data opening, activities that are increasingly developed in the world of inventories of tangible heritage, large museums and libraries, as shown by the large participatory collection represented by Europeana Collection.

3 Web Services for Cultural Heritage

In the previous work (Artese & Gagliardi, 2019) we focused on how to integrate data from different archives in order to provide query tools that would allow users to find information without knowing exactly, a priori, where to look for it and the language in which it is expressed. In the prototype *QueryLab* we experimented with web services provided by large inventories such as Europeana, Victoria&Albert Museum and CooperHewitt, to allow users to search through different archives simultaneously and have a single integrated response. We have seen that RestAPI web services are a very useful tool to get access to data directly at "home" and at the same time to overcome the obstacles of different languages: they allow both to access always updated data and to build software tools to use these data in different ways. From the analysis of the different services studied for the extension of the *QueryLab* prototype we can detail reflections and comparisons:

- **Europeana:** is a web portal containing digitalised museum collections of more than 3,000 institutions across Europe. It includes records of over 10 million cultural and scientific artefacts, brought together on a single platform. They offer both web services (<https://pro.europeana.eu/page/search>) and a graphic

console (<https://pro.europeana.eu/page/api-rest-console>) where users can test queries. To use the services an API Key is needed. The service offers many parameters to perform a query, but is widely documented and easy to use, supported by many examples.

- **Victoria&Albert Museum**, London: also known with the abbreviation V&A, is the world's largest museum of applied and decorative arts, design and sculpture. The on-line collection (<http://collections.vam.ac.uk/>) can be searched using web services (<https://www.vam.ac.uk/api/>) which are basic and described in a very simple way but at the same time they are the easiest to use and very effective both in results and response times. They do not require an API Key.
- **Cooper Hewitt, Smithsonian Design Museum**, New York: is the only museum in the United States devoted exclusively to historical and contemporary design and is the steward of one of the most diverse and comprehensive design collections in existence. The web site (<https://collection.cooperhewitt.org/>) offers web services (<https://collection.cooperhewitt.org/api/>) together with a simple but effective console useful to test queries. An API Key is needed.
- **Reunion des musées nationaux – Grand Palais**: (Rmn-GP) is a public cultural establishment, overseen by the French Ministry of Culture and Communication, and with the web site Images d'Art (<https://art.rmngp.fr/en>) offer artworks from the biggest French museums. Web services documentation is not so simple to understand at first glance (<https://api.art.rmngp.fr/>) but they offer query possibilities and some example. They ask for an API Key.
- **Digital Public Library of America**: collection of images, texts, videos, and sounds, in collaboration with partners all over the United States, to offer an open, distributed national digital library (<https://dp.la/>). The web service is available (<https://pro.dp.la/developers/api-codex>) and fully documented with examples and use cases, the API Key is needed.
- **Auckland War Memorial Museum**: is one of New Zealand's first Museums and tells the story of New Zealand, its place in the Pacific and its people. It has pre-eminent Māori and Pacific collections, significant natural history resources and major social and military history collections, as well as decorative arts and pictorial collections (<https://www.aucklandmuseum.com>). They have a specific page where web services are fully described and documented and they don't require an API Key (<https://www.aucklandmuseum.com/discover/collections-online/our-data>)
- Other museums, such as Royal Albert Memorial Museum, Royal Museum of Greenwich and Museum of Applied Arts & Sciences (Sidney), offer web services to reach their data but they are limited only to full exports or selections of categories or specific subsets of data, without the possibility to apply a real query. They are not still included in QueryLab.

The analysis of different web services for the realization of integrated queries allows the extension of the QueryLab prototype to several other inventories compared to the first 3 made. We observe, however, that the linked resources are always related to tangible entities, in our analysis we have not found inventories provided with web services that were related to the intangible heritage.

4 Common Metadata and Linked Open Data

The intangible inventories integrated by *QueryLab* are currently available in a local database and include *IntangibleSearch* and the "*Data Bank on Traditional/Folk Performing Arts in Asia and the Pacific*", collected by the Asia/Pacific Cultural Centre for UNESCO, both of which are included in the Map of ICH e-inventories. The data of these two inventories were used to test the first draft of the metadata structure proposed in the previous work, "ICH Light", intended to harvest and store the subset of minimal information needed to perform queries. Other archives have been identified through the Map of ICH e-inventories, to be collected into the common structure, and the metadata structure itself has been revised and updated accordingly. The inventories analyzed to be used as test case for minimal integration into ICH Light are:

- Sahapedia (<https://www.sahapedia.org/>), an open online resource on the arts, cultures and heritage of India.
- ICH Scotland (<http://ichscotland.org/>), inventory of Scotland's living culture,
- the German Nationwide Inventory of Intangible Cultural Heritage (<https://www.unesco.de/en/kultur/immaterielles-kulturerbe/german-inventory.html>)
- Living Traditions in Switzerland, recently updated list of Swiss living traditions (<https://www.lebendige-traditionen.ch/tradition/en/home.html>).

Each of these archives, and many others analyzed, includes practices and knowledges that, because of their social characteristics, should be integrated more than others. They are rituals and traditions of the past to be safeguarded and defended, but they create also a common thread among all the archives visited. They bring together the peoples of the most disparate places on earth in an incredible way, creating a connection, a common base, a unique and shareable tradition. It is precisely these "data" that highlight how parochialism has no sense of existing because people have much more in common than what divides them: from cooking to the cult of the dead, from wedding traditions to carnival rituals, tales and legends, folk songs, herbalist knowledge, and so on.

Hence the idea to evolve *QueryLab* towards a Europeana style portal, dedicated to intangible data, where the minimal data of the different inventories can flow together to be integrated, queried and shared, always keeping the link to the original resource for details, but also providing services for data entry and management, for those entities that do not have a proprietary archiving system. To achieve this, it is necessary to remodel the common data structure for intangible heritage so that it is extensible and customizable, defining a format that also opens towards the Linked Open Data environment. Starting from the analysis of similar experiences in the field of museum data, we referred in particular to the detailed experience in Szekely et al., 2013, which describes the activities and problems encountered in the data mapping of the Smithsonian American Art Museum in Linked Open Data format. In this work the reference model used is the Europeana Data Model (EDM) and the author introduces a new ontology (called SAAM) which extend EDM with a selection of sub-classes and sub-properties

to represent the attributes unique to the museums objects. Museum data have been successfully modeled into Linked Data Cloud and it is a significative example of how database-to-RDF process can be implemented for tangible heritage.

In the works of Wijesundara & Sugimoto, 2018 and Wijesundara et al., 2017 it is observed that “neither EDM nor CIDOC-CRM (the model by the International Council of Museums) has specific entities designed for expressing intangible cultural heritage”, so they purpose a new model, called CHDE (Cultural Heritage in Digital Environments) to provide a solution using *Instantiation*, because “an intangible cultural heritage is not an item physically collectable by memory institutions and the item-centric resource aggregation is not suitable. *Instantiation* acts as a bridge to aggregate those resources related to intangible cultural heritage”. According to this model a traditional dance performance is an *Instantiation* of the corresponding intangible cultural heritage entity, performed in a specific place and time. The CHDE model collects all related digital resources into one instance, which can be transformed into a digital archive record based on the One-to-One Principle of Metadata (Miller, 2010), useful to distinguish digital copies and their physical source.

This model fits perfectly with the ICH Light data structure that has been implemented so far. In fact, ICH Light provides a contextual and historical description of the ICH entity and a set of different instances related to the documentation collected during the various representations, which took place at different times by different people. These instances are the physical objects collected in the digital archive created and represent the evolution of the ICH entity over time. If represented on a timeline they show the changes and the adaptations that the living good experiences over time.

5 Conclusions

From the study carried out so far, it can be concluded that the purpose of obtaining a specific metadata model for intangible cultural heritage is a complex challenge that cannot ignore the standards developed so far, such as EDM and CIDOC-CRM, but must be differentiated by what are the specific peculiarities of this type of data. While Europeana's EDM format is well suited for tangible heritage and has also been used by important American museums and libraries such as the Smithsonian and DPLA, it is noted that neither EDM nor CIDOC-CRM are suitable to describe intangible entities and the top-down model they propose is not suitable to express the meaning of the relationship between digital objects and the entities they refer to, whether tangible or intangible. Starting from the formalism presented in Wijesundara & Sugimoto, 2018 we are working on the *QueryLab* project on two fronts:

- updating the metadata of ICH Light to have a flexible and customizable structure to suits the needs of the different inventories involved so far
- detailing the transformation from the ICH Light metadata to RDF format, using the recommended ontologies, to formulate all the properties needed to produce Linked Open Data.

Future work will be focused on improving data query methods and developing a back-end environment for organisations and communities to manage data, to increase the opportunity to preserve and share intangible cultural heritage online.

References

- Artese, M. T., & Gagliardi, I. (2019). Sharing ICH Archives: Integration of Online Inventories and Definition of Common Metadata. *Digital Presentation and Preservation of Cultural and Scientific Heritage*. Burgas (Bulgaria): Institute of Mathematics and Informatics at the Bulgarian Academy of Sciences.
- Beck, J., Büchner, M., Bartholmei, S., & Knepper, M. (2017). Performing Entity Facts. *Datenbank-Spektrum*, 17(1), 47-52.
- Carboni, N., & De Luca, L. (2016). Towards a conceptual foundation for documenting tangible and intangible elements of a cultural object. *Digital Applications in Archaeology and Cultural Heritage*, 108--116.
- Fink, E. E., Szekely, P., & Knoblock, C. A. (2014). How Linked Open Data Can Help in Locating Stolen or Looted Cultural Property. *Euro-Mediterranean Conference* (p. 228--237). Cham: Springer.
- Knoblock, C. A., Szekely, P., Ambite, J. L., Goel, A., Gupta, S., Lerman, K., . . . Mallick, P. (2012). Semi-automatically mapping structured sources into the semantic web. *Extended Semantic Web Conference* (p. 375--390). Springer.
- Miller, S. J. (2010). The one-to-one principle: challenges in current practice. *International Conference on Dublin Core and Metadata Applications*, (p. 150-164).
- Sousa, F., & Imaterial, I. R.--. (2017). Map of e-Inventories of Intangible Cultural Heritage. *Memoriamedia Review*. *Memoriamedia Review*.
- Szekely, P., Knoblock, C. A., Yang, F., Zhu, X., Fink, E. E., Allen, R., & Goodlander, G. (2013). Connecting the smithsonian american art museum to the linked data cloud. *Extended Semantic Web Conference* (p. 593--607). Springer.
- Wijesundara, C., & Sugimoto, S. (2018). Metadata model for organizing digital archives of tangible and intangible cultural heritage, and linking cultural heritage information in digital space. *LIBRES: Library and Information Science Research Electronic Journal*, 58-80.
- Wijesundara, C., Monika, W., & Sugimoto, S. (2017). A metadata model to organize cultural heritage resources in heterogeneous information environments. *International Conference on Asian Digital Libraries* (p. 81-94). Springer.
- Ziku, M. (2020). Digital Cultural Heritage and Linked Data: Semantically-informed conceptualisations and practices with a focus on intangible cultural heritage. *LIBER Quarterly*, 1-16, DOI: <http://doi.org/10.18352/lq.10315>.

Received: June 19, 2020

Reviewed: July 25, 2020

Finally Accepted: July 31, 2020