# A Knowledge Management Architecture for Digital Cultural Heritage

MAURO DRAGONI, SARA TONELLI, and GIOVANNI MORETTI, Fondazione Bruno Kessler

15

The increasing demand of technological facilities for galleries, museums, and archives has led to the need for designing practical and effective solutions for managing the digital life cycle of cultural heritage collections. These facilities have to support users in addressing several challenges directly related to the creation, management, preservation, and visualization of digital collections. Such challenges include, for example, the support for a collaborative management of the produced information, their curation from a multilingual perspective to break the language barriers and make collections available to different stakeholders, and the development of services for exposing structured version of data both to users and machines. Platforms satisfying all of these requirements have to support curators activities and, at the same time, provide facilities for engaging the virtual consumers of the produced data. In this article, we propose a description of an abstract architecture for managing digital collections built on a set of components, services, and APIs able to address the challenges mentioned previously. An instantiation of this architecture is discussed, and we present a use case concerning the management of a digital archive of verbo-visual art. Lessons learned from this experience are reported to outline future activities.

CCS Concepts: • **Information systems** → **Data management systems**; **Digital libraries and archives**; • **Computing methodologies** → *Ontology engineering*; • **Applied computing** → *Enterprise ontologies, taxonomies and vocabularies*; *Arts and humanities*;

Additional Key Words and Pharses: Knowledge management, digital heritage, curation, visualization, services

**ACM Reference Format:**
Mauro Dragoni, Sara Tonelli, and Giovanni Moretti. 2017. A Knowledge Management Architecture for Digital Cultural Heritage. *ACM J. Comput. Cult. Herit.* 10, 3, Article 15 (July 2017), 18 pages.
http://dx.doi.org/10.1145/3012289

## 1 INTRODUCTION

In the past years, cultural heritage institutions have been faced with the challenge of integrating the digital component in their daily activities. Galleries, libraries, archives, and museums typically own rich and structured datasets developed over many years and organized by domain, which in principle could be connected with the databases of other institutions and then made available online to a larger audience. In this scenario, however, several challenges are still open. For example, most information management systems used by Museums are not freely available. Besides, a lack of technical skills especially in small museums makes basic data manipulation and conversion hard to achieve. Relevant standardization efforts such as that of Europeana[1] contribute to raise museums' awareness of the importance of knowledge sharing among cultural heritage institutions, but they

[1] http://www.europeana.eu/.

Authors' addresses: M. Dragoni; email: dragoni@fbk.eu; S. Tonelli; email: satonelli@fbk.eu; G. Moretti; email: moretti@fbk.eu.

http://dx.doi.org/10.1145/3012289

are mainly focused on data repositories, leaving open the problem of which tools should be used to support knowledge management and address domain experts' requirements.

In this article, the problem of designing and developing technological platforms for managing cultural heritage collections is addressed. Our goal is to design an architecture based on a knowledge management layer that domain experts can use to model the complex elements and connections inside a cultural heritage archive. This layer should foster collaboration and support experts with many services for data enrichment and quality check based on Semantic Web technologies. At the same time, the complex information stored in the system should be made accessible online through a data visualization layer, offering the possibility to browse the archive content in a user-friendly way. In particular, the main contributions of the article are the following:

— The presentation of an abstract architecture for managing cultural heritage collections in digital format, to be used as a reference for building real-world infrastructures. In the discussed design schema, the components of the technological platform used for addressing the main challenges and requirements are presented.
— An instantiation of the architecture mentioned previously, consisting in a two-layered platform: a knowledge management tool for supporting curators in managing the life cycle of digital collections and the connected application for presenting such collections online to the final users.
— The implementation of such an instantiation in a project on verbo-visual art, as a case study to collect feedback from museums' personnel and plan future system extensions.

The rest of the article is structured as follows. In Section 2, we review the state of the art on the technologies used for curation of digital cultural heritage, including tools and metadata schemata. Section 3 presents the general idea of the architecture for managing the full life cycle of digital collections by discussing the requirements behind the conceptual idea, the main components, and the connected services. In Sections 4 and 5, we describe how the main components of the architecture have been instantiated in two tools: one for supporting curators in managing the knowledge contained in digital collections and one for allowing users to browse their content online. The two tools have been used in a project called *VVV* (verbo-visual-virtual), described in Section 6. Finally, Section 7 concludes the article with a summary of the findings.

This article extends a previous work concerning the use of knowledge management solutions within the cultural heritage field presented in Dragoni et al. (2016) and other works related to the MoKi tool. In Dragoni et al. (2016), we presented a conceptual model supporting linking activities with the aim of enriching a cultural heritage collection. Here, such a conceptual model has been integrated in a complete architecture including a data management tool, a visualization layer, and several services available through APIs. Finally, with respect to previous works related to the MoKi tool, for the first time we present (i) the full platform consisting of a suite of functionalities and services for managing the full cycle of the knowledge information in the cultural heritage domain, and (ii) the integration of the knowledge management platform with a full-feature user-friendly visualization solution.

## 2 RELATED WORK

Given the growing interest in enriching cultural heritage data using Semantic Web languages and techniques, several works has been proposed to address this issue. Hyvönen (2012) presents an overview on why, when, and how Linked Open Data (LOD) and Semantic Web technologies can be employed in practice in publishing cultural heritage collections and other content on the Web.

An example is provided by Szekely et al. (2013), where the authors define a specific ontology that extends the EDM ontology. This extension is motivated by the lack of properties needed to represent specific data of the Smithsonian museum. The data is then linked to knowledge bases such as DBpedia, the Getty vocabulary, "Ulan" (Union List of Artist Names), and the list of artists of the Rijksmuseum in Amsterdam. De Boer et al. [2012, 2013] present the Amsterdam Museum Linked Open Data project, where the data of the museum have been automatically converted into RDF data. The authors use the EDM ontology for defining the semantics of the data coupled with vocabularies such as the Dublin Core vocabulary.

On the one hand, we share with these works the idea of starting from raw cultural heritage data available in different formats, such as tables, texts, and CSV, and then transforming these data into semantic data. We all rely on the EDM ontology as a starting point, and we extend it depending on the purpose of our data transformation task. Moreover, a data interlinking step with DBpedia is addressed both in these works and in our approach. On the other hand, several differences arise. More precisely, we not only translate the raw data into RDF data and interlink it with DBpedia instances but also address also an enrichment step by leveraging structured information from Web resources, such as information about the artists' date and place of birth.

Other works deal with the definition of specific ontologies and annotation tools. Benjamins et al. (2004) present an ontology of Humanities, which they integrate in a semiautomatic tool for the annotation of cultural heritage data, to ease knowledge acquisition.

Other related works have been proposed to enable multilingual access to cultural heritage information. Dannells et al. (2013) address this task by adopting Semantic Web languages. They process museum data extracted from two distinct sources and make these data accessible in natural language, thanks to a grammar-based system designed to generate coherent texts from Semantic Web ontologies in 15 languages. Here the task was performed by combining natural language processing techniques and Semantic Web technologies.

Finally, the LOD cloud contains some examples of semantic data from museums. Among others, the British Museum[2] has published its data collection using the CIDOC-CRM ontology to allow data manipulation and reuse, and it provides a SPARQL access point to query the available database. Although in the use case presented in Section 6 we also aim at linking our cultural heritage data with other information sources in the LOD cloud, our setting is different from the British Museum one in that our small collection faces several issues (e.g., data sparseness and quality) that do not concern well-documented collections.

On the knowledge management side, available tools are mainly thought for knowledge engineers. The most popular open source tools available are *Protégé* (Gennari et al. 2003) and the *NeOn Toolkit* (Espinoza et al. 2008), and they are both designed to support collaboration between experts in knowledge management. Therefore, interfaces and other functionalities do not account for requirements of experts in the specific domains to be modeled. This aspect has a significant impact, limiting the possibility to make knowledge management accessible to a wider audience.

In addition to the ontology management tools described earlier, a lighter way for modeling knowledge is the use of tools for managing vocabularies with the possibility of directly exposing them. One of such tools is *VocBench* (Stellato et al. 2015). Even if this tool includes a set of facilities for domain experts, it does not offer connection to external machine translation services or the alignment with other vocabularies.

Finally, concerning the online exposure of digital collections, the literature reports experiences in developing systems addressing specific domains, such as intangible cultural heritage (Dagnino et al. 2014) or architecture (De Luca et al. 2011; Liu et al. 2006). Such architectures are usually focused on one specific goal, such as learning, semantic inference, or cultural transmission, and do not address the end-to-end process covered by our architecture, which starts with the definition of the conceptual model and ends with data navigation, trying to account for heterogenous curator and user requirements.

## 3 ARCHITECTURE

In this section, we present the general architecture of the knowledge management platform for digital collections. We start with a presentation of the main requirements that such a platform should satisfy for supporting a complete digital curation process (Section 3.1). Then we describe which are the main components of the platform (Section 3.2) and which external services can be plugged to them (a preliminary list is presented in Section 3.3).

[2] http://collection.britishmuseum.org.

## 3.1 Requirements

We present five important requirements concerning the creation and maintenance of digital collections, which represent the road map for the design of the proposed architecture.

*Collaboration support.* Crafting cultural heritage collections in a digital scenario has become a teamwork activity, as it requires a wide range of skills. For this reason, collaborative aspects in digital collection modeling have been investigated in the past, focusing particularly on the collaborative creation of domain-specific ontologies (e.g., see Palma et al. (2011) and Dimitrova et al. (2008)). The requirements and features that have emerged from these studies highlighted the complexity of the collaborative process, since it typically involves domain experts and ontology engineers who need to make shared decisions also when they work in (geographically) distributed teams of contributors. Supporting collaboration requires fostering user awareness on the evolution of the modeling artifacts, supporting the communication of modeling choices and decisions taken among the involved actors.

*Multilingual management.* A requirement strictly related to collaborative aspects of knowledge modeling is the need to overcome language barriers. Thus far, research and development activities have been concentrated on monolingual environments, and in most cases the default language has been English. Although English admittedly plays a primary role in international communications, the varied scenario of cultural heritage includes both international institutions and small local museums, whose personnel and visitors do not necessarily understand English.

*Data exposure.* The growth of complex and distributed platforms made the role of Web services very relevant to give access to the created knowledge. One of the main advantages of using Web services is the possibility of exposing all information in several standard formats that may also evolve over time. Therefore, an effective exposure service is necessary when there is an intensive usage of digital collections in very dynamic contexts, such as with the goal of creating virtual exhibitions or disseminating culture globally.

*Data linking.* The growing amount of data being exposed on the Web in the cultural heritage domain, like in the Europeana effort, makes it necessary to exploit, whenever possible, the information already available online. This can be performed by setting links between different knowledge sources describing artifacts, enriching "local" knowledge bases with information coming from external ones. The main advantage is to avoid the redefinition of already existing information. For example, given an artifact record, we do not need to insert information about the author's life if we can set a link between the object in our digital collection and DBpedia.[3] One advantage is that updates of the linked resource are instantly available also in the local digital collection, avoiding the publication of deprecated information. A survey offering different approaches concerning the linking problem may be found in Euzenat and Shvaiko (2007) and Bellahsene et al. (2011).

*User engagement.* To increase the number of users who contribute to maintaining the collected information, a knowledge management platform should make the user experience as engaging as possible. In fact, an effective interaction with users of digital collections is a key aspect for the success of a platform. For these reasons, an architecture should include modern and engaging technologies for making digital contents very attractive. An example is the recent adoption of personalized storytelling applications for museums that allow the creation of visiting paths based on user preferences and experiences, and can also be used to design virtual exhibitions.

## 3.2 Overview

Here we provide a general description of the platform for managing the entire life cycle of a cultural heritage collection. The aim of this schema is to provide a reference architecture by taking into account the requirements described in the previous section.

[3] http://wiki.dbpedia.org/.

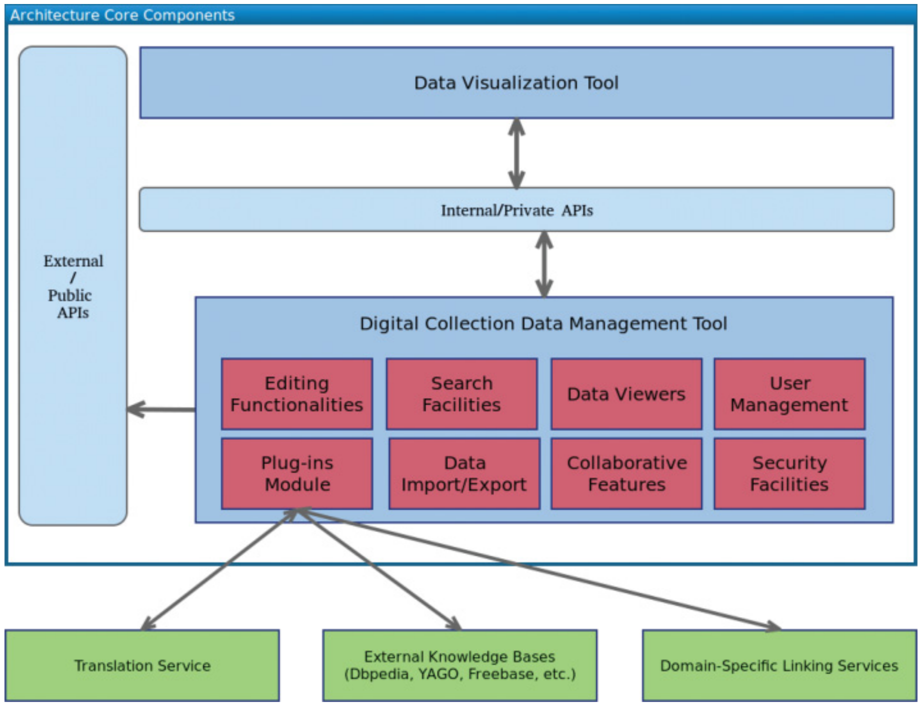



Fig. 1. Abstract structure of the platform architecture.

Figure 1 shows the diagram of the architecture, comprising two main blocks: (i) the core components (described in this section, blue boxes in Figure 1) and (ii) a list of potential external services (described in Section 3.3, green boxes in Figure 1) that can be used for extending the capabilities of the core set.

In the set of core components, we include the following modules.

*Digital collection data management tool.* The data management tool is the main component for managing the entire life cycle of the artifacts included in the digital collection. This tool contains all facilities adopted by curators for creating, updating, and preserving the entire collection. In this module, eight facilities are foreseen: the *editing functionalities* include the set of interfaces, customized for the type of artifact that has to be managed, used by curators for the data entry task. These functionalities are accompanied by a set of *search facilities* and *data viewers*. The former allow lookup operations on the collection, whereas the latter show the content of the created collection.

The multiuser aspect of digital collection curation is supported by a set of *collaborative features*, enabling both the collaborative editing of different types of data and the notification mechanisms for keeping all involved actors updated about the status of the collection. Roles and permissions are assigned to the curators involved in the digital collection management through the *user management* and *security* facilities. The content of the data management tool can be exported to third-party software or enriched with additional data that can be imported from other collections. All of these operations can be performed through the *data import/export* facilities.

Finally, the interaction with external services is enabled through the *plug-ins module*, which connects with external services providing information to enrich the content of the digital collection. Through the set of plug-ins, it is easier to add new functionalities to the tool or to enhance the full architecture with updated versions of services, libraries, and so on. Examples of such services are provided in Section 3.3.

*Data visualization tool.* The data visualization tool refers to the software used for publishing the curated digital collections on the Web or through other types of media. Such a tool should contain all facilities for showing the data from different perspectives (e.g., hyperlinks between documents, graphical summarization of the artifacts, browsing facilities) by also taking into account the need for making the user's experience engaging. Software for the transposition of museum collections to the Web can range from museum Web sites to virtual (or augmented) reality facilities.

The data management and data visualization tool are logically separated because they address the requirements of different types of users. In the data management tool, the implemented functionalities are designed from the curator's perspective, whereas data visualization is oriented toward improving and maintaining the engagement level of final users. In addition, the same data management tool can underlie different visualizations, depending on the scenario in which the architecture is deployed.

*Internal/private APIs.* Both the data management tool and the visualization layer are part of the core architecture. However, for maintaining the high modularity of the platform and to easily plug different instantiations of the visualization layer, we foresee the implementation of an internal API layer, which supports the communication between the presentation tools and the digital collection stored in the data management tool. These APIs are marked as "private" because the set of methods is available only for the components of the core architecture.

Another advantage of using an API layer for retrieving information from the data management tool is the possibility of managing the security aspects of each artifact. For instance, the curators can decide (i) which artifacts can be published online and (ii), for each artifact, which data fields can be exposed and to which components. This way, through the use of the APIs, it is possible to have a real-time synchronization between the actions carried out by the curators on the digital collections and the content that is actually exposed.

*External/public APIs.* In addition to the internal communication between the core components of the platform, a public exposure of the digital collection would have several advantages. For instance, the (full or partial) exposure of artifact information enables the connection between the collection metadata and a wider knowledge context, such as the LOD cloud. Furthermore, the exposure of artifact data can be part of a sustainable model of museums, such as by fostering the development of mobile apps based on these data.

### 3.3 Services

As introduced in the previous section, the architecture foresees the interaction with third-party services providing specific functionalities that are not directly integrated in the general platform. Such services, represented by the green boxes placed at the bottom of Figure 1, are connected with the data management tool through the plug-ins module using well-known protocols like REST or SOAP.

Here we present three services meant to support the implementation of the requirements described in Section 3.1. This list of services is not limited to the ones described in the following, but it represents a starting set of features that are often required in the cultural heritage community.

*Translation service.* Support for multilinguality is mandatory for breaking language barriers and reaching as many people as possible. In the past decade, research in statistical machine translation has led to the implementation of several public services and models that can be used for translating natural language content into different languages. Example of services available through APIs are Google Translate[4] and Microsoft Bing,[5] whereas an example of an open source toolkit for machine translation is Moses (Koehn et al. 2007).

Translation services can be very useful to translate on the fly the description and the metadata of an artifact, and save this information in the artifact record. This can be performed by including a connector activating the

[4] https://cloud.google.com/translate/docs.
[5] https://www.microsoft.com/en-us/translator/translatorapi.aspx.

machine translation service on demand. The architecture can include a range of translation services, which can be reached by a curator through a plug-in based mechanism, so that comparisons of the output translation are also possible. For instance, this would allowswitching from a general-purpose translation service to a domain-specific one (Bosca et al. 2014).

*Linking to general-purpose knowledge bases.* Curators can exploit links to the LOD cloud to enrich digital records with information freely available over the Web. In particular, the metadata associated with each artifact can be extended or completed semiautomatically, reducing manual effort. Examples of these metadata are the authors' name, places where artifacts have been created or where they are hosted, and temporal information about the epoch in which the artifact was created or artists were born.

Such metadata can be the anchor to set links between each artifact and external knowledge bases containing general-purpose information. Examples of these knowledge bases are the ones shown in the related block in Figure 1. Although we mentioned only the most widely used ones, such as DBpedia,[6] YAGO,[7] and Freebase,[8] others are available in the LOD cloud and could be integrated in the architecture. One of the advantages of setting these links is the possibility to connect each artifact with knowledge that final users can exploit for their personal goals (i.e., for learning purposes), even if the curator who originally set the link had not foreseen this specific use.

*Linking to domain-specific knowledge bases.* The knowledge bases mentioned previously enable one to enrich artifact descriptions with general information; however, more specific information can be obtained through the connection with domain-specific knowledge bases. In the cultural heritage field, several repositories have been created and made available to this purpose. An example is the Getty vocabularies,[9] containing structured terminology related to art, architecture, archival materials, and conservation. For instance, these enable the automated linking between the technique used for creating an artifact and the entity in the Getty Art & Architecture Thesaurus describing this technique.

## 4 MOKI-CH: A KNOWLEDGE MANAGEMENT TOOL FOR CULTURAL HERITAGE

The "Digital Collection Data Management Tool" showed in Figure 1 can be implemented in several ways. Here we present MoKi-CH, a collaborative knowledge management tool for cultural heritage collections, as one of the possible instantiations of such system.

In Section 4.1, we introduce the general architecture of the tool by explaining the building blocks and the rationale behind its development. In Section 4.2, we describe in more detail the facilities that have been implemented in MoKi-CH for supporting curators in managing digital collections.

### 4.1 Overview of the MoKi Tool

MoKi[10] is a collaborative MediaWiki-based (Wikimedia Foundation 2012) tool for modeling ontological knowledge expressed through the well-known ontology language called *OWL*.[11]MoKi is grounded on three main pillars, which we briefly illustrate with the help of Figure 2:

— Each basic entity of the ontology (i.e., concepts, object and datatype properties, and individuals) is associated with a Wiki page. For instance, the concept Mountain in Figure 2 is associated with a Wiki page that contains its description.

[6] http://wiki.dbpedia.org/.
[7] https://www.mpi-inf.mpg.de/departments/databases-and-information-systems/research/yago-naga/yago/.
[8] https://www.freebase.com/.
[9] http://www.getty.edu/research/tools/vocabularies/index.html.
[10] http://moki.fbk.eu.
[11] https://www.w3.org/TR/owl-ref/.

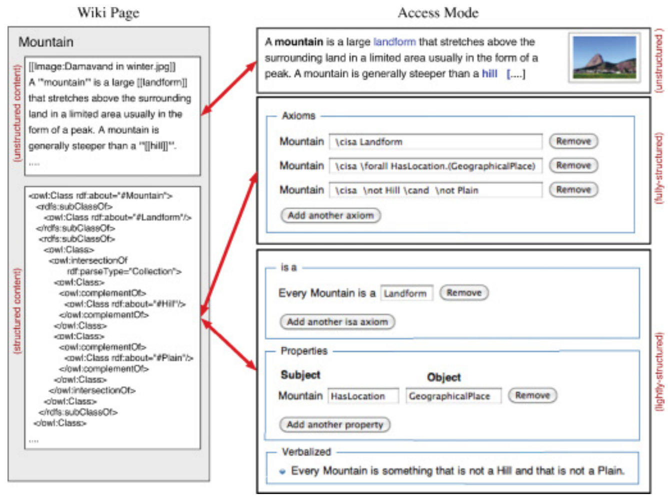



Fig. 2. A page and the access modes in MoKi.

— Each Wiki page describes an entity by means of both unstructured (e.g., free text, images) and structured (e.g., OWL axioms) content.
— Three different system versions and interfaces (with different degrees of complexity) are available in a so-called multimode access to address the needs of users with different skills and competencies and permit their collaboration.

The three different types of users foreseen in the multimode access are (i) "knowledge engineers," the people in charge of managing the formalization of all information inserted into the knowledge base; (ii) "domain experts," the people most involved in the modeling task with the goal of providing all of the knowledge in a raw shape; and (iii) "language experts," in the case of multilingual management, people who deal with the multilingual layer of the created knowledge base by providing and checking translated data, and have access only to a subset of the functionalities available to domain experts. A comprehensive description of the core version of MoKi and the three technological pillars of the tool are presented in Dragoni et al. (2014) and Ghidini et al. (2012).

MoKi has been designed as a general-purpose tool for collaborative knowledge management. However, due to its high flexibility supporting different kinds of experts in modeling their domains, MoKi interfaces can be easily adapted to different user needs. In Section 3, we introduced a set of five requirements that a platform should account for to support users in managing cultural heritage collections. Based on these requirements, we created a customized version of MoKi, called *MoKi-CH*, including a set of domain-specific functionalities that we describe in the next section.

With MoKi-CH, the modeling workflow is as follows. First, an underlying conceptual model must be defined to represent and connect all information of the cultural heritage collection. The version of MoKi-CH presented in this section exploits the conceptual model described in Dragoni et al. (2016). Then the model is instantiated into a Wiki-based tool like MoKi-CH by associating a page to each entity modeled within the digital collection. In particular, curators manage the life cycle of pages related to the five concepts that in the conceptual model were defined as the basic ones in the cultural heritage domain: "Author," "Artwork," "Technique," "Location," and "Title." Depending on the kind of entity that has to be modeled, the related page is also equipped with a customized set of facilities to ease the work of curators.

### 4.2 User Facilities for Managing Cultural Collections

The customization effort based on MoKi-CH was carried out based on the requirements described in Section 3.1, and particularly with the following goals: (i) grant the collaboration between curators to facilitate their work from different locations; (ii) support the exposure of information outside the platform in a structured way, allowing both users and machines to consume the produced data; (iii) provide mechanisms for supporting multilinguality to make the platform able to expose information in different languages; and (iv) integrate linking components for connecting all possible information managed by curators to the LOD cloud. Details on the customization are provided in the following.

*Collaborative editing facilities.* Given the complexity of managing digital collections, it is necessary to give different users access to the data to enable multiple consistency checks and collaborative work. Therefore, we expanded the standard Wiki features already implemented in MoKi with the possibility of assigning specific tasks to different experts for maintaining the digital collection. This way, experts are able to manage the entire life cycle of the built knowledge base by monitoring and possibly approving the suggested changes.

These facilities are grouped into two sets of features. The first group is aimed at monitoring the activities performed on each entity page. Every time a change is committed, approval requests are created and saved into the page of the changed artifact. Each approval request contains the identification of the expert in charge of approving the change, the date in which the change has been carried out, and a natural language description of the change. Connected to this, a mechanism was implemented to manage the approvals and maintain the history of all approval requests of each artifact.

The second set of features, instead, contains the facilities for managing the discussions associated with the page of each artifact. A user interface for creating the discussions has been implemented together with a notification procedure alerting users when new topics or replies related to the discussions they are following are posted.

To summarize, the set of the collaborative editing functionalities that has been implemented includes:

- *Discussions*: These are used to discuss challenging issues related to the modeling of the digital collection. Through this facility, it is possible to discuss about a single artifact or about (a part of) the collection. Comments in the discussion pages are organized in threads, with details on the user and date/time associated with each comment.
- *Watchlist*: The watchlist is used to monitor artifacts. Any change performed on monitored artifacts is presented via notifications (with messages and email alerts) to the domain expert(s).
- *Notifications*: Notifications inform experts about artifact changes that are relevant to them. Email or message notifications are automatically sent in case of changes to pages in the experts' watchlist. Experts can also send specific notifications, soliciting a confirmation or revision of some artifact from selected experts.
- *History and revision*: These are used to track changes and comments added to a specific artifact.

*Multilinguality management features.* Multilinguality concerns two main aspects: (i) the knowledge base *content* and (ii) the *interface* of the knowledge management tools. In the first case, tools are equipped with facilities for translating (automatically through external services or manually) information provided by domain experts to improve as much as possible the dissemination of information. In the second case, tools' interfaces are provided in several languages to enable the collaboration between users speaking different languages.

MoKi-CH is equipped with a set of features for enabling both manual and automatic translation of labels and descriptions associated with each artifact. Manual translations can be provided by experts through a custom interface, whereas automatic suggestions are available through a connector that invokes external translation services. Specifically, the component sends the request to APIs exposed by third-party translation services (e.g., Microsoft Bing[12]), and after the retrieval of the result, the representation of the artifact is updated with the

[12] http://www.bing.com/translator.

returned information. Further details about the translation services currently supported by MoKi-CH can be found in Dragoni et al. (2013, 2011).

MoKi-CH has been also equipped with two further components to manage the language of the ontology and that of the tool interface. If a new language has to be added to the ontology, MoKi-CH invokes, for each artifact described in the collection, the translation of its label and description in the new language. This is taken into account also when exporting the dump of the knowledge base in a Semantic Web format (OWL, SKOS, etc.) by giving the possibility to choose the export language among the available ones. Finally, for facilitating the work of language experts, we have implemented the possibility of side-by-side comparison of two lists of translations. This way, the language expert in charge of revising the translations is relieved from navigating through each artifact pages by speeding up the revision process.

*LOD service.* One of the main challenges for promoting cultural heritage is to find a standard and easy way for distributing content over the Web. The Semantic Web community defined the concept of Linked Data as a method for distributing structured data, enabling an easy interlinking between different knowledge sources and fostering the reuse of existing knowledge. LOD services are built on standard Semantic Web formats such as RDF[13] and Turtle,[14] and can be accessed through Web services like REST. To permit the exposure of a structured version of the digital collection to third-party services, MoKi has been equipped with a service that exposes artifacts information by using LOD formats. Such a service offers the possibility to remotely perform operations on the digital collection, such as the retrieval of the collection or its expansion. The service provides a RESTful interface for receiving the requests, and the results are exposed in the SKOS language.[15]

*Entity linking component.* Linking the created collection to external knowledge bases enables the semiautomated enrichment of the local knowledge with information extracted from external resources. This is possible thanks to the ontology linking component of MoKi-CH, which offers to experts a set of suggestions to match candidates (artworks, artists, etc.) with external entities from sources selected by the user. Technical details about how such a component works can be found in Dragoni (2015).

From the artifact, technique, or author page, it is possible to invoke an external service that receives as input a bean containing all information related to the concept that has to be mapped, and that returns a list of candidate entities that can be mapped. The expert only has to select which entity, if any, is a good mapping to the local one.

Figure 3 shows an example list of candidate mappings, issued when a new request is sent. For each technique, it is possible to manage the translations of the label and to define a link with a definition contained in the Getty Thesaurus. When a curator wants to align a concept with the Getty Thesaurus, a request is sent to the API service of the Getty portal, returning a set of candidate definitions that might be suitable for creating the new link (see Figure 3). The curator is able to click on the link for opening the Getty portal to obtain more information about the chosen definition and then may click on the *Create Mapping* button for adding the new link to the technique page.

*Customization of the user interface.* The user interface has been customized to facilitate the work of curators. The most important change was the introduction of three entity forms necessary to support knowledge management in the arts domain: “Artwork,” “Author,” and “Technique.” These forms are used for managing the life cycle of artworks, the techniques used for creating them, and the related author pages contained in the digital collection. They contain a general set of facilities for curators, which can be enriched case by case based on the curators’ requirements. Each form also contains a check box, which gives curators the possibility to publish the content of specific fields out of MoKi-CH (e.g., to make them visible through the data visualization tool). This

[13] https://www.w3.org/TR/2004/REC-rdf-concepts-20040210/.
[14] https://www.w3.org/TR/turtle/.
[15] http://www.w3.org/2004/02/skos/.

Fig. 3. Result of link request to the Getty API service.

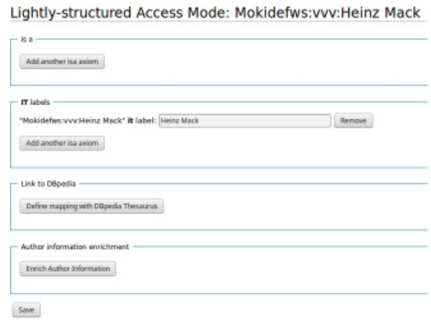



Fig. 4. Mask for modeling “Author” information.

feature is crucial to properly manage work that is typically in progress, like the curation of cultural heritage data. In fact, it may be necessary to check the consistency of records, or complete them, thus keeping them offline during the revision process.

Figure 4 shows the form related to the management of the author pages. In this form, the two most important facilities are the ones related to the creation of a mapping between the current author entity and DBpedia, and the possibility of analyzing unstructured information from external textual resources. In the first case, the procedure, and the result, is the same as shown in Figure 3 concerning the creation of a new link with the Getty Thesaurus. As for the import of external information, the system is able to extract structured information from external sources and to present them to the curator. Then the curator can choose which information should be added to the author page.

Another type of form concerns the modeling of artwork entities. The characteristic of the artwork form is to put in relation the different types of entities that are modeled within the digital collection. Indeed, for each artwork, it is necessary to define which author created the artwork and which technique was used. To maintain the information of the digital collection consistently, the form is equipped with facilities assisting curators in choosing the right entity to link. While typing the author’s name, MoKi-CH performs a back-end search in real time for showing appropriate entities for the author field. This way, curators are able to check the correctness of the inserted information in real time.

## 5 DATA VISUALIZATION LAYER

The knowledge management layer is designed to interact with the visualization layer so that the information stored in the knowledge base can also be used to navigate the digital archive built on it. The visualization layer can be directly embedded in a Web site to give public access to the archive information. For this reason, whereas the GUI of MoKi-CH is quite basic since it is meant for experts, the viewer was designed to ease navigation: it should be as user-friendly as possible while retaining the complexity of the underlying information. A seamless communication between the knowledge management and the visualization layer is of key importance to the overall architecture and is made possible through internal APIs (see Section 3.2).

Although the layer implemented in our system architecture is just one of the possible viewers that can be built on top of the knowledge management system, it was designed with different users in mind, trying to account for general requirements of a wide public. Specifically, data exploration was foreseen according to the following three different paradigms.

*Collection-centered exploration.* This exploration starts from the overview of the whole collection (Figure 5(a)), where each artwork is a record represented by an image and a small set of metadata (i.e. title, date, artist, unique

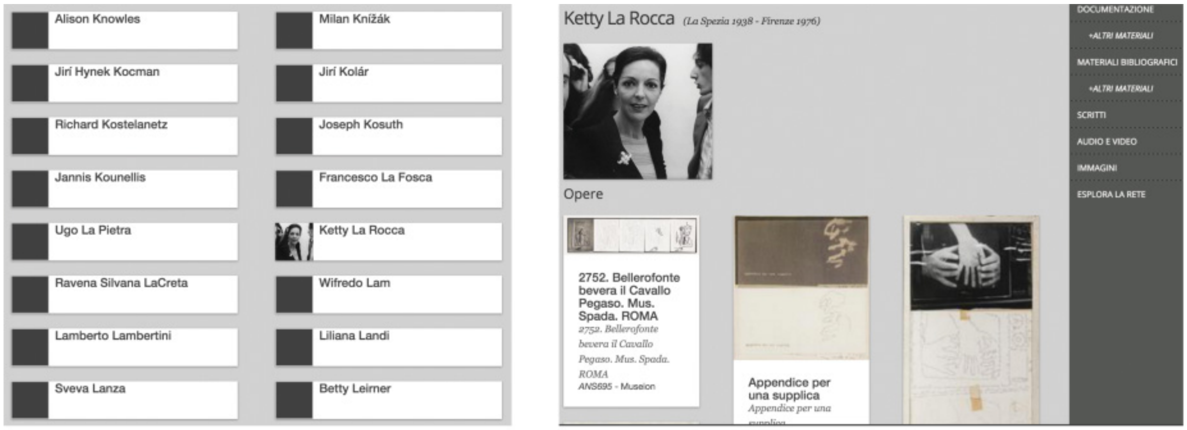

(a) Collection-centred view (b) Image detail

Fig. 5. Two views for collection-centered exploration. The general view (Fig. 5(a)) and the image detail (Fig. 5(b)).

identifier, and the museum hosting the artwork). The records are sorted by ID, so the order in which they are displayed does not necessarily match chronological or other criteria. The proposed navigation pattern complies with the well-known paradigm "Overview first, zoom and filter, details on demand" (Heer and Shneiderman 2012), in that users can click on an image and open a new view (Figure 5(b)), where the material, the technique, and the dimension are displayed, together with a description of the content in plain text and a low-resolution image. This can be zoomed in to see all details thanks to a pyramid representation of the artwork: the image is conveyed by a multiple mosaic of images, each at a different zoom level, so that each tile is stored in a separate file. Pyramid representation is very important when dealing with images of artworks since museums consider these important assets and do not generally want to release them in high resolution. If other material related to the artwork of interest is present in the knowledge base, such as bibliographic material, it is shown on the right-hand bar and can be further explored. A search functionality is also available to look for information in the stored metadata. Although this view is meant for users who do not have a specific interest and just want to serendipitously browse through the collection, it has some drawbacks. For instance, it does not provide any form of aggregated records, and scrolling the page to view very large collections may be tedious.

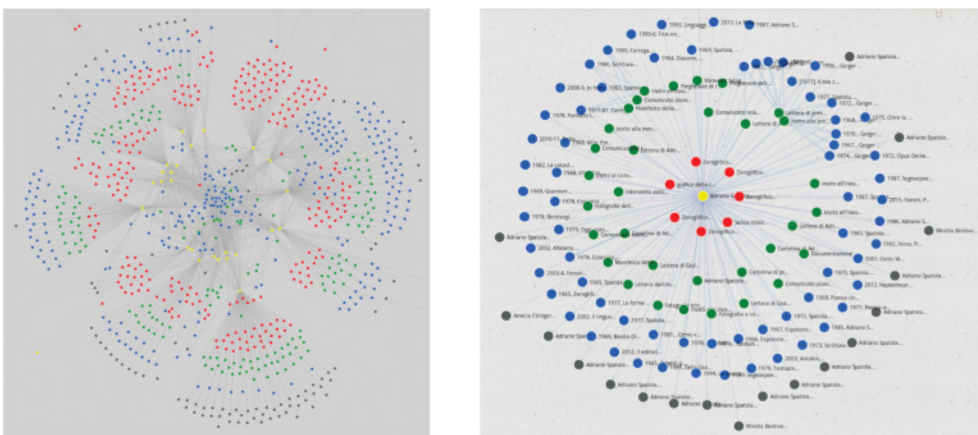
*Artist-centered exploration.* The *Artists* item first shows the list of artists in the repository, then by selecting one of them, it is possible to see all information in the database related to the given person. The displayed information ranges from basic biographical information (e.g., date and place of birth/death) to the list of works and additional audio/visual material related to the chosen artist. Relations are an important piece of information of a knowledge management system for the cultural heritage domain, especially when such systems contain heterogeneous connections between different types of entities, such as artists, artworks, and bibliographic material. In addition, an archive can store (explicitly or implicitly) information on events such as exhibitions and collaborations among artists. All of this can be stored in MoKi-CH through relations of different types, and the *Viewer* must allow users to easily access and visualize this complex data organization. Nevertheless, different types of users may be interested in viewing the collection at different granularities, with more or fewer details, so the Viewer should be flexible enough to show only the amount of information needed. Therefore, the artist-centered exploration can start from the list of works and related material (see Figure 7(b)), or from the artist's ego network (see the following).

*Network-based exploration.* Networks are particularly suitable for the visualization of museum knowledge bases, as they can provide a view at a glance of very complex information, making it possible to visually compare




(a) Artist-centred view (b) Artist detail

Fig. 6. Two views for artist-centered exploration.



(a) Network-based view of the collection (b) Network-based view of the artist

Fig. 7. Two views for network-based exploration.

the coverage of different collections or the production of different artists. In addition, outliers (e.g., artists who do not collaborate with others) or highly connected items can be easily spotted in a network-based visualization. Furthermore, they can accommodate in one layer different types of information, such as by coloring nodes according to different categories or edges representing different relations. As well, current libraries for network visualization allow users to zoom in and look for information at node level, making it very easy to change search granularity. For all of these reasons, a network visualization is present both for each author and to represent the whole collection (*Explore* tab) in our Viewer. In the artist-centered view, an ego network is displayed (Figure 7(b)), in which the artist is the central node and the material related to him or her is arranged in different circles and colors. The default view is the simplest one, with only direct edges connecting artists and artworks; however, complexity can be increased through a sliding bar, by adding indirect material and links. The first view is aimed at maximizing readability, whereas the second is primarily targeting domain experts who want to have as much information as possible. In Figure 7(b), for example, the artist is the yellow circle, whereas the red ones are the artworks. Green nodes are documents written by the artist (i.e., with a *direct* connection to him, displayed as the inner circle) and written about the artist (see the outer blue circle, representing *indirect* connections). In addition, the blue nodes, representing bibliographic material, make a distinction between articles written by or about the artist (i.e., inner and outer circle, respectively).

Networks are displayed with the Force layout implementation of the d3.js library (Bostock et al. 2011). This implementation allows the straightforward creation of networks composed by nodes and links, guaranteeing speed, customizability, and interactivity. In particular, the Force layout relies on an efficient algorithm for the graph disposition using the Verlet integration (Verlet 1967), the Barnes-Hut approximation (Barnes and Hut 1986) to speed up the objects' grouping process, and Tim Dwyer's method for graph layout constraints (Dwyer 2009). Data used by d3.js in the interactive visualization of networks are provided by querying the knowledge management system and performing data manipulation to get all possible relations between the queried entities.

The same library and the same options to explore the knowledge base content with incremental complexity are used in the collection-centered view (Figure 7(a)). This displays all elements stored in MoKi-CH as nodes in a network, whose color is consistent with the author-centered view. In this case, all artists belonging to the collection of interest are present, and the global network highlights information that would not be visible starting from the artworks' or the artists' view (i.e., which artists have been part of the same exhibitions, have worked collaboratively, or are most represented in the collection). The search functionality also works in this context, allowing users to zoom in and look for specific nodes. Note that the displayed information is retrieved from the knowledge management system only if the related content was flagged as "complete" by curators.

## 6 THE EXPERIENCE WITH THE VVV PROJECT

### 6.1 The VVV Project

The presented architecture was first designed and implemented in the framework of the VVV project[16] (Marchetti et al. 2013), which started in 2013 as a joint effort between two museums in Trentino-Alto Adige (Italy) and Fondazione Bruno Kessler in Trento, with the goal to create a unified virtual collection of "Archivio di Nuova Scrittura" (ANS) (Ferrari 2012). The collection, albeit international, is mainly centered around the artworks of Italian artists active between 1950 and 1990, and finds its origin in the collecting activity of Paolo Della Grazia, an entrepreneur with a passion for interdisciplinary forms between art and poetry. Della Grazia's focus on verbo-visual poetry was influenced by the collaboration with Ugo Carrega, one of the most representative artists of visual poetry in Italy in the 1960s. The ANS fulcrum is represented by works linked to concrete poetry, visual poetry, Fluxus, and conceptual art.

Toward the end of the 1990s, Della Grazia decided to donate to a public institution the archive, which had been steadily growing and needed an appropriate site. The collection was split into two parts, one assigned to MUSEION,[17] the Museum of Contemporary Art in Bolzano, and the other to MART,[18] the Museum of Modern and Contemporary Art in Rovereto, meaning that a collection originally conceived as a single archive was now divided in two and hosted by two different institutions.

Given the possibility offered by current digital technologies to access art collections online, the VVV project was launched to create a unified virtual collection of ANS works, where all information about the collection would be semantically enriched and made available through a Web-based platform. Researchers in digital humanities and knowledge management at Fondazione Bruno Kessler have therefore worked in the past 2 years to provide a semantically consistent representation of the data about the archive stored in the knowledge management systems of the two different museums, possibly add information automatically retrieved from the Web, and implement a navigation platform to display and search works in a virtual exhibition. In addition, two art

[16] https://dh.fbk.eu/projects/vvv-verbo-visuale-virtuale-la-piattaforma-di-ricerca-interattiva-dellarte-verbo-visuale.
[17] http://www.museion.it.
[18] http://www.mart.trento.it.

curators were in charge of retrieving from the archive all information available for a selected number of artists, converting them in digital format, and manually inserting all possible direct and indirect links with artifacts and additional material (this work could not cover all archive content due to time constraints). Overall, the digital ANS archive now contains 445 artists and 2,617 artworks (1,336 from MART and 1,281 from MUSEION), and it is available at http://www.verbovisualevirtuale.org, a platform based on the data visualization layer presented in Section 5. The curators performed additional research on 7 artists, adding 220 related documents, 347 bibliographic records, 108 exhibition descriptions, and 100 multimedia entries (e.g., audio or video recordings). The general network displayed in the *Explore* view now contains 3,900 links, connecting the materials of the archive through different relation types.

*6.1.1 Project Challenges.* The VVV project brought about several challenges, which were taken into account and partly motivated the implementation of MoKi-CH:

(1) MART and MUSEION had already started to create records about the ANS artworks in the respective data management systems. However, this was done incrementally, at different times, and by different people, which had an impact on the quality of the available information. Some records were found to be inconsistent even for important elements such as the artwork title or author. Here, experts had to be supported in the management of such items by providing facilities to interlink information about collection items with as many external data sources as possible to improve the qualitative description of each item. In VVV, this was particularly relevant to guarantee that the two portions of the archive were curated with the same quality.
(2) In addition to quality, the quantity of available information was an important issue. Some of ANS artists are not very well known, and it was not easy to find information about them on the Web. When it was available, it had to be manually extracted for enriching the related content in the knowledge base. Addressing this challenge implies the design and implementation of semiautomated techniques to support domain experts in the enrichment of artworks information, like the ones provided by MoKi-CH.
(3) The exposure of information in semantic formats requires artworks to be classified through a classification schema. For this reason, the use of ontologies in the cultural heritage domain has gained a lot of attention in the past years. However, well-known classification schemas (e.g., the Europeana Data Model[19]) may be too generic to capture the peculiarities of minor collections like ANS. In this case, a conceptual modeling activity was needed to tailor existing models to the specific needs of a collection and was supported by the creation of a specific ontology presented in Dragoni et al. (2016).
(4) Finally, managing multilinguality was an issue in the project because MART records were all in Italian, whereas MUSEION had recorded part of the information in German, given that in South Tyrol both Italian and German are official languages. This made it crucial to include functionalities for automated translation in the knowledge management architecture.

Given that the project duration was only 2 years, data collection and manipulation had to start right at the beginning of the project. Therefore, the two art curators responsible for the task first adopted the FileMaker[20] software, which they commonly use for digital archive curation. This implied quite a lot of data manipulation, including (i) exporting the database content related to ANS and stored in the two commercial record management systems adopted by MART and MUSEION, (ii) harmonizing the two databases, and (iii) converting them into the format required by FileMaker so that they could be imported, manually expanded, and corrected by the project curators. Once the correction and enrichment work was complete, the FileMaker content had to be exported in another database that contains all unified data and is used to display information in the data visualization layer.

[19]http://www.europeana.eu/.
[20]http://www.filemaker.com.

This workflow was clearly too complex and could not account well in case of updates to the single databases (the export and conversion operations had to be repeated at specific intervals). Therefore, we decided to investigate the possibility of having an integrated framework, simplifying the import/export pipeline and making communication between the components seamless. This led to the implementation of the architecture introduced in Section 3.

### 6.2 Evaluation

The proposed architecture was evaluated from two different points of view. The first was a quantitative evaluation, aimed at assessing the effectiveness of the data linking and enrichment modules in the knowledge management layer. The second was a qualitative evaluation, based on a focus group organized with the two museums' personnel.

In the quantitative evaluation, which was presented in Dragoni et al. (2016), we took into account a subset of 287 artists present in MART records and performed automated linking to DBpedia using the entity linking component of MoKi-CH (see Section 4.2). Only 47% of the artists were linked, as many minor authors in the collection do not have a Wikipedia page (and a DBpedia entry). However, the quality of the linking reached 0.812 accuracy. We then focused on the 148 artists with no Wikipedia page and manually looked for possible sources of information available online. This was done to identify which additional knowledge bases could be connected to the MoKi-CH linking component. For 93 artists, we found an online description of the author biography. Finally, we automatically aligned the concepts related to techniques with the Getty thesaurus using the mapping service implemented in our platform, and we obtained an accuracy of 0.976.

A second qualitative evaluation was performed with the two art curators involved in the project and the museums' personnel in charge of the archive. After about a year and a half from the beginning of the project (i.e., when the data manipulation workflow based on FileMaker was already consolidated), we proposed the alternative knowledge management workflow based on MoKi-CH. We first carried out a system demonstration and organized a focus group in which the demo participants could comment on the platform functionalities. In general, the feedback was positive, particularly with regard to the time that could be spared by completing records through semantic linking. In addition, the possibility to manually correct translations or to choose the best piece of information among a set of options was deemed very useful, as art experts would never trust a completely automatic enrichment process. As well, direct communication between the knowledge management layer and the visualization layer, made possible through internal APIs, was seen as crucial to reduce the import/export effort based on FileMaker dumps. The possibility to flag complete records to allow their visualization online had not been foreseen and was added to the architecture as an outcome of the focus group. Another suggestion, which has not yet been implemented, is the possibility to perform image-centered linking (i.e. to retrieve information starting from the archive images and not only from artist or artwork names). Although interesting, this new task would be computationally more demanding and would require additional research in the field of image processing, as most established linking techniques are focused on text.

## 7 CONCLUSIONS

In this article, we presented an architecture based on a knowledge management platform to effectively support curators in handling the life cycle of cultural heritage collections while engaging users with different backgrounds and expectations in the online exploration of digital archives. The platform is composed of (i) a set of modules implementing different sets of facilities to support domain experts in editing and curating cultural heritage data, (ii) services in charge of connecting the platform with external data sources, and (iii) components aiming to make the modeled data available online. After describing the general model of the proposed architecture, we detailed the two main components: a knowledge management system, MoKi-CH, which is a specific implementation of the MoKi system tailored to the cultural heritage domain, and a data visualization layer, which retrieves information from the underlying knowledge management tool and displays the archive content in different engaging views. We explored the adoption of these tools within the VVV project and showed the advantages in terms of effort and data quality with respect to traditional data management workflows. In the future, we plan to extend the platform to address more complex scenarios—for example, in distributed contexts, where issues like scalability must be properly addressed.

## REFERENCES


Josh Barnes and Piet Hut. 1986. A hierarchical O (N log N) force-calculation algorithm. *Nature* 324, 446–449.

Zohra Bellahsene, Angela Bonifati, and Erhard Rahm (Eds.). 2011. *Schema Matching and Mapping*. Springer.

V. R. Benjamins, J. Contreras, M. Blázquez, J. M. Dodero, A. Garcia, E. Navas, F. Hernández, and C. Wert. 2004. Cultural heritage and the Semantic Web. In *The Semantic Web: Research and Applications*. Lecture Notes in Computer Science, Vol. 3053. Springer, 433–444.

Alessio Bosca, Vassilina Nikoulina, and Marc Dymetman. 2014. A lightweight terminology verification service for external machine translation engines. In *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics (EACL'14)*. 49–52. http://aclweb.org/anthology/E/E14/E14-2013.pdf.

Michael Bostock, Vadim Ogievetsky, and Jeffrey Heer. 2011. D$^3$ data-driven documents. *IEEE Transactions on Visualization and Computer Graphics* 17, 12, 2301–2309.

Francesca Maria Dagnino, Leontios J. Hadjileontiadis, Michela Ott, and Francesca Pozzi. 2014. An integrated platform supporting intangible cultural heritage learning and transmission: Definition of requirements and evaluation criteria. *Journal of Computing and Information Technology* 22, 4, 277–292. http://cit.srce.unizg.hr/index.php/CIT/article/view/2461.

Dana Dannells, Aarne Ranta, Ramona Enache, Mariana Damova, and Maria Mateva. 2013. Multilingual access to cultural heritage content on the Semantic Web. In *Proceedings of the 7th Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities*. 107–115.

Viktor de Boer, 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*. Lecture Notes in Computer Science, Vol. 7295. Springer, 733–747.

Victor de Boer, Jan Wielemaker, Judith van Gent, Marijke Oosterbroek, Michiel Hildebrand, Antoine Isaac, Jacco van Ossenbruggen, and Guus Schreiber. 2013. Amsterdam Museum Linked Open Data. *Semantic Web* 4, 3, 237–243. DOI: http://dx.doi.org/10.3233/SW-2012-0074

Livio De Luca, Chawee Busayarat, Chiara Stefani, Philippe Véron, and Michel Florenzano. 2011. A semantic-based platform for the digital analysis of architectural heritage. *Computers and Graphics* 35, 2, 227–241. DOI: http://dx.doi.org/10.1016/j.cag.2010.11.009

Vania Dimitrova, Ronald Denaux, Glen Hart, Catherine Dolbear, Ian Holt, and Anthony G. Cohn. 2008. Involving domain experts in authoring OWL ontologies. In *The Semanti Web—ISWC 2008*. Lecture Notes in Computer Science, Vol. 5318. Springer, 1–16.

Mauro Dragoni. 2015. Multilingual ontology mapping in practice: A support system for domain experts. In *The Semantic Web—ISWC 2015*. Lecture Notes in COmputer Science, Vol. 9367. Springer, 169–185. DOI: http://dx.doi.org/10.1007/978-3-319-25010-6_10

Mauro Dragoni, Alessio Bosca, Matteo Casu, and Andi Rexha. 2014. Modeling, managing, exposing, and linking ontologies with a Wiki-based tool. In *Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC'14)*. 1668–1675. http://www.lrec-conf.org/proceedings/lrec2014/summaries/769.html.

Mauro Dragoni, Elena Cabrio, Sara Tonelli, and Serena Villata. 2016. Enriching a small artwork collection through semantic linking. In *The Semantic Web: Latest Advances and New Domains*. Lecture Notes in Computer Science, Vol. 9678. Springer, 724–740. DOI: http://dx.doi.org/10.1007/978-3-319-34129-3_44.

Mauro Dragoni, Chiara Di Francescomarino, Chiara Ghidini, Julia Clemente, and Salvador Sánchez Alonso. 2013. Guiding the evolution of a multilingual ontology in a concrete setting. In *The Semantic Web: Semantics and Big Data*. Lecture Notes in Computer Science, Vol. 7882. Springer, 608–622. DOI: http://dx.doi.org/10.1007/978-3-642-38288-8_41

Mauro Dragoni, Chiara Ghidini, Giannis Stoitsis, Miguel Angel Sicilia, and Salvador Sanchez-Alonso. 2011. Recommendations for revising existing ontologies and schemas. Organic.Lingua EU Project, Deliverable D3.1.1.

Tim Dwyer. 2009. Scalable, versatile and simple constrained graph layout. *Computer Graphics Forum* 28, 991–998.

Mauricio Espinoza, Asunción Gómez-Pérez, and Eduardo Mena. 2008. Enriching an ontology with multilingual information. In *Proceedings of the 5th European Semantic Web Conference (ESWC'08)*. 333–347. *http://dl.acm.org/citation.cfm?id=1789394.1789428*

Jérôme Euzenat and Pavel Shvaiko. 2007. *Ontology Matching*. Springer.

Daniela Ferrari. 2012. *Archivio di Nuova Scrittura Paolo della Grazia. Storia di una Collezione/Geschichte einer Sammlung*. Silvana Editoriale, Milan, Italy.

J. H. Gennari, M. A. Musen, R. W. Fergerson, W. E. Grosso, M. Crubézy, H. Eriksson, N. F. Noy, and S. W. Tu. 2003. The evolution of Protege maiuscolo: An environment for knowledge-based systems development. *International Journal of Human-Computer Studies* 58, 1, 89–123.

Chiara Ghidini, Marco Rospocher, and Luciano Serafini. 2012. Modeling in a Wiki with MoKi: Reference architecture, implementation, and usages. *International Journal of Advanced Life Sciences* 4, 111–124. https://dkm.fbk.eu/rospocher/files/iaria_ijals2012.pdf.

Jeffrey Heer and Ben Shneiderman. 2012. Interactive dynamics for visual analysis. A taxonomy of tools that support the fluent and flexible use of visualizations. *ACM Queue* 10, 2, 1–26.

Eero Hyvönen. 2012. *Publishing and Using Cultural Heritage Linked Data on the Semantic Web*. Morgan & Claypool. http://dx.doi.org/10.2200/S00452ED1V01Y201210WBE003

Philipp Koehn, Hieu Hoang, Alexandra Birch, Chris Callison-Burch, Marcello Federico, Nicola Bertoldi, Brooke Cowan, et al. 2007. Moses: Open source toolkit for statistical machine translation. In *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics (ACL'07)*. 177–180. *http://dl.acm.org/citation.cfm?id=1557769.1557821*

Yong Liu, Congfu Xu, Zhigeng Pan, and Yunhe Pan. 2006. Semantic modeling for ancient architecture of digital heritage. *Computers and Graphics* 30, 5, 800–814. DOI: http://dx.doi.org/10.1016/j.cag.2006.07.008

Alessandro Marchetti, Sara Tonelli, and Rachele Sprugnoli. 2013. The verbo-visual virtual platform for digitizing and navigating cultural heritage collections. In *Proceedings of the 2nd Annual Conference on Collaborative Research Practices and Shared Infrastructures for Humanities Computing (AIUCD'13)*.

Raúl Palma, Oscar Corcho, Asunción Gómez-Pérez, and Peter Haase. 2011. A holistic approach to collaborative ontology development based on change management. *Web Semantics: Science, Services and Agents on the World Wide Web* 9, 3, 299–314. http://www.websemanticsjournal.org/index.php/ps/article/view/200.

Armando Stellato, Sachit Rajbhandari, Andrea Turbati, Manuel Fiorelli, Caterina Caracciolo, Tiziano Lorenzetti, Johannes Keizer, and Maria Teresa Pazienza. 2015. VocBench: A Web application for collaborative development of multilingual thesauri. In *The Semantic Web: Latest Advances and New Domains*. Lecture Notes in Computer Science, Vol. 9800. Springer, 38–53. http://dx.doi.org/10.1007/978-3-319-18818-8_3

Pedro A. Szekely, 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*. Lecture Notes in Computer Science, Vol. 7882. Springer, 593–607.

Loup Verlet. 1967. Computer experiments on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. *Physical Review* 159, 1, 98.

Wikimedia Foundation 2012. MediaWiki. Retrieved May 25, 2017, from http://www.mediawiki.org.

Received April 2016; revised August 2016; accepted October 2016