Interaction with Linked Digital Memories

Ilaria Torre

Department of Computer Science, Bioengineering, Robotics and Systems Engineering
University of Genoa, Italy
ilaria.torre@unige.it

Abstract. This paper presents an ongoing project on Web of Things and Linked Data. Physical objects are provided with a digital memory that enables them to store data about themselves and to link other resources in the Linked Data cloud. This possibility radically transforms the interaction with real objects and the user experience with the Web. In this scenario, personalization seems an opportunity and a requirement. The physical objects used in the project are museum artefacts. The paper describes architecture and main modules.

Keywords: digital object memories, Linked Data, personalization, museum

1 Introduction

The availability of technologies for the identification and connection of physical objects and the possibility to access them via HTTP are changing the modality to explore and experience the Web. The Web is going to become an extension of reality, making boundaries between physical and virtual objects more and more faint. As packaging is often mixed with products, web data associated to objects are going to fade into objects.

According to the paradigm of the Web of Things, real world objects can be connected in a network and accessed by using the web standards and protocols. The Semantic Web of Things (SWoT) uses web standards and protocols to expose physical objects as web resources and exploits the technologies of the semantic web to ensure semantic interoperability between systems that access and interact with heterogeneous physical objects.

Furthermore, the Linked Data (LD) principles provide a publishing paradigm in which structured data can be easily exposed, shared and linked by using HTTP URIs for identifying data and the RDF graph-based model to represent them. As a consequence, by using a URI that identifies the digital representation of physical objects, virtual and physical objects can be connected in a network of Linked Data [8]. Lower level identification systems enable the association of the URI to the physical object. The main technologies that are used for identification are RFID, QR-Code, Semacode and visual recognition.

Building a digital representation of a physical object makes objects digital memories which can be passive or active and can be used for different purposes, ranging from storing temporary data obtained by sensors to storing and representing complex information. For example, memories can store information about the provenance of the object, about its history, its use ant its position. Moreover they can also store the flow of messages generated by users while interacting with the object or while interacting with other users through the object. Barthel et al. [2] refer to them as Digital Object Memories.

The availability of information about an object while interacting with that object has many applications, such as job support, emergency handling, learning by doing, entertainment and edutainment. The linking with other pieces of information related to that object in the cloud of Linked Open Data (LOD) extends the possible applications and radically transforms the interaction with real objects and the user experience with the Web. In the domain of cultural heritage there are several examples of applications that experiment the Internet of Things technologies [1], [2] and examples of initiatives and institutions that follow the principles of LD to expose and connect artefacts with related data [5], [12]. However there are still few applications that combine them [6], [10].

This contribution presents an ongoing project which combines these approaches with the aim of enhancing the interaction with museum artworks by exploring the world "behind" the artwork. The key features of the project are: (i) search and exploration of the object memories and of the LOD cloud (ii) personalization and adaptation of content and user interface to the usage context and to the media used for interacting with the artwork.

The paper is structured as follows: the next section presents a background concerning LD and Object Memories and describes an architecture for representing them. Section 3 describes the main components of the prototype under development and presents an example scenario with two use cases. Finally, Section 4 sketches the next steps and concludes the paper.

2 LD and Object Memory: Background and Architecture

The so-called five-star rating scheme identifies increasing levels of structuring and linking of data towards a Web of Linked Data [3], with the requirement, for all levels, of an open licence to be Open Data.

There are several European projects and groups that work on LD. Organizations in the field of CH are particularly interested in the possibility to share and link knowledge about CH artefacts and repositories. Examples of cultural heritage libraries that make available their data according to the principles of LD are the British National Library[1] and the German National Library[2]. Both provide free access to their RDF/XML data, under a Creative Commons licence. Other active institutions are the museums. For example, the Amsterdam Museum has its entire collection represented as LD [5]. Several tools and services have been developed to support organizations to publish their data as RDF. Indeed, only a fraction of open data is currently repre-

[1] http://www.bl.uk/bibliographic/datafree.html
[2] http://www.dnb.de/EN/Service/DigitaleDienste/LinkedData/linkeddata_node.html

sented as RDF[3]. Examples of services aimed to facilitate the conversion of tabular data to semantically enriched data are Any23[4], Tabels[5], RDF Refine[6]. Moreover, D2R[7] server is a tool for accessing relational databases as RDF-graphs. They are particularly useful in scenarios when huge legacy data of an institution have to be extracted and transformed into RDF.

Linking data to other datasets or to more general vocabularies such as DBPedia is even a more costly task. The task of interlinking automation is faced in European and international projects. E.g., the 7FT LOD2 project[8] (2010-2014) dedicates a WP to the issue of reuse, interlinking and knowledge fusion, addressing the automatic and semi-automatic link creation with minimal human interaction. Example of frameworks and tools for supporting the search and interlink of data are Linked Media Framework[9] and the DBpedia Lookup Service[10].

Considering physical objects, additional requirements have to be satisfied in order to manage them as LD. A layered architecture makes easier the separation of functions by making layers autonomous.

In our project, we defined a three-layered architecture consisting of a *physical layer*, a *layer of the object memory* and a *LD layer*. The physical layer includes: the physical object (named entity in **Fig. 1**), the specific modality of interaction with the object, such as pointing, scanning, touching [11] or using a mediation device (e.g., a smartphone), and the modality of identification of the object, such as RFID, QR-Code, Semacode and techniques for visual object recognition. The layer of the object memory manages the description of the object and the way to access it, according to the Object Memory Model (OMM) [2]. The LD layer describes the physical object according to the LD principles.

OMM is a proposal of reference model for representing structure and content of object memories submitted to W3C [2]. In OMM, a digital object memory is described as a repository of digital data that is linked with a physical artefact, and may be populated with static or dynamic data from entities that interact virtually or physically with the artefact. This repository may exist at the artefact itself, outside, or both. The model includes an OMM header, a list of blocks and a set of metadata for describing blocks. The header contains the unique identifier for the object memory. In our architecture it is an HTTP URI. The model provides a list of blocks and a set of metadata that can be extended. Each block concerns a specific information fragment (a topic such as the object provenance, content and the level of protection required). A set of metadata is used to annotate each block, mostly based on Dublin Core: the

[3] For example, on PublicData portal, an aggregator of dataset descriptions from numerous other European data portals, only 459 out of more than 17.000 datasets are available as RDF [8].
[4] http://any23.apache.org/
[5] http://idi.fundacionctic.org/tabels/
[6] http://refine.deri.ie/
[7] http://d2rq.org/d2r-server
[8] http://lod2.eu/
[9] http://semanticweb.org/wiki/Linked_Media_Framework
[10] http://wiki.dbpedia.org/lookup/

block ID, Namespace, Format (and encoding), Title and Description of the block, Subject (free text tags, ontology concepts) and Link (used to split and distribute blocks in case of space constraints). The OMM layer is useful to manage aspects such as memory distribution, changing identifiers or physical relationships.

In our architecture, over this layer we have the LD layer. It is obtained from a subset of blocks of the OMM layer, automatically extracted and published as RDF (by LD wrappers). This layer is optional. It has to be implemented if the owner of the object wishes it to be accessed by RDF browsers and crawlers and if the owner wishes the object to be linked by other data in the LOD cloud. In order that wrappers publish data as LD, mappings between OMM metadata and RDF vocabularies have to be defined, together with links pointing to external and internal data sources.

Specific applications implementing the architecture can define how and where to store the object memory: locally in the object, or on a server, which can be easily accessed via HTTP. For example, Tales of Things[11], a platform using OMM, stores object memories on a server. Users can add stories about their objects and connect to other people who share similar experiences. Users have to print blank tags, stick them to the physical objects they want to have a digital memory and populate the memory by using an App which provides tools for accessing the object memory via HTTP.

Linking the object memory to other data in the LOD cloud presents opportunities for personalization, similarly to personalizing the access to open corpora, but with the advantage of managing structured data, which makes possible using filters in SPARQL queries and reasoning on data. This enables combining information from different sources and providing users with personalized mashup services, as in [12], or performing recommendations based on similarity measures between data as in [13].

3 Components of the System

To validate the approach, we are working on developing the modules in charge of identifying objects and accessing them. They include:

1) A <u>configuration tool</u> for the identification of the physical object and the association to its digital memory,
2) An <u>authoring system</u> that enables the writing, structuring, annotating and interlinking of the artefacts,
3) A system for the <u>personalized access</u> to the object memory and to the linked data, and for the adaptation of the interaction.
   Fig. 1 displays a diagram showing data flow and connections among architecture layers and components for the personalized access to the object memory. It will we described below, when presenting an example scenario.

1. Configuration tool. The layered architecture in Sec. 2 does not specify any particular identification system to be used. In our prototype we use the method of QR-code. To this end we have developed a tool that generates the QR-code and prints labels that can be attached to artworks. The QR-code contains the URI of the object memory.

[11] http://talesofthings.com/

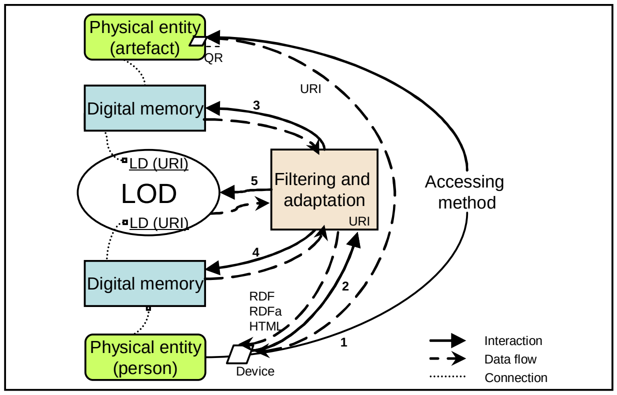

Fig. 1. Interaction diagram showing data flow and connections among architecture layers and components for personalized access to the object memory.

2. Authoring system for the knowledge definition.

The authoring system consists in a set of back-end modules made available to the administrator to populate the object memory associated to each artwork.

These modules provide support on three main tasks.

i) Querying the LOD cloud, in order to find vocabularies and reference ontologies that can be used to describe the physical object and to (inter)link the object with other resources. This task can be done manually by the administrator, using search tools on ontologies, such as DBPedia and Yago, or in a semi-automatic way, by using DBpedia Lookup APIs[12] and the Linked Media Framework[13]. Once they are collected, they are validated by the administrator who will use them in the next step.

ii) Editing the object memory upon the physical layer, by using the metadata specified in the OMM architecture, and defining mappings with vocabularies and links identified in the previous step. Related resources discovered on LOD will be used by LD wrappers in two ways: to connect to a similar resource (sameAs) and to connect to a resource that is used as a description (hasTopic).

iii) Adding optional materials about the object (e.g. movies, technical data, etc.). Additional materials can be provided in different versions, thus enabling the personalization of the returned data. Notice that also these materials are objects. Therefore they should be identified with an HTTP URI too.

3. Personalization module.

When interacting with physical objects, the interaction modality and the mediation device used to interact with the object has usually constraints which make very critical the task of personalization. A further reason that makes personalization important

[12] The Web service is based on a Lucene index providing a weighted label lookup, which combines string similarity with a relevance ranking (similar to PageRank) in order to find the most likely matches for a given term [4].

[13] The Linked Media Framework provides tools for querying LD and for integrating the returned results within the application.

is that when the user interacts with a physical object, s(he) has often time constraints and noise around him/her. In our project, this module manages two main tasks.

i) Personalized access to the object memory. The OMM model has been designed to support a basic level of personalization, used to filter content before downloading it. It is useful to adapt the content to the context domain, to the user profile an in case of limited bandwidth or connectivity. The set of metadata for each block provides information about data creator, data format, data content and blocks' location.

ii) Adaptation of the interaction. Physical objects can be of different type and can have different capabilities. Objects with processing capability can make recursive queries contacting other physical objects in the local network or in the LOD cloud.

Objects equipped with processing capabilities can also perform another type of adaptation, which concerns the adaptive composition of content to be displayed to the user. For this task, the project will exploit the features of the MM-Learning framework [9] that we developed to perform adaptive composition of materials and activities in order to suite the device constraints.

Example scenario. The idea at the basis of object memories is that objects can store all relevant information about them, accumulated along time. This concept is already widely accepted and applied to goods and products. However, it can be extended to all kinds of objects and to people too. Museum artefacts can have a long history. Several properties describe an object and each one may have an history. Typical properties are the artwork title, its author, date of creation and place, artistic current, the artistic phase of the author's life. Other information concern provenance, e.g., the origin, whether it was commissioned, purchased, moved from a place to another, stolen, retrieved. Other information concern whether it was repaired, which techniques have been used for restauration, its current quoted value, price fluctuation, etc. Other information can be about experts' reviews or visitors' comments.

A layered architecture, with information fragments organized in topics has two main advantages. First, it enables all these pieces of information to be managed in a flexible way. For example, it enables a LD wrapper to publish and make discoverable just a subset of the information stored in the object memory. Moreover it makes easier the filtering of content, based on the context of use and on the user profile. For example in the context of a museum, topics concerning restauration and quoted value can be filtered out, even though they could be re-filtered in if the visitor is an art merchant.

Considering **Fig. 1**, we sketch two use cases.

1) Context: *museum*; application used to access the object memory: *normal browser + QR reader*. Flow: (n1) the user/device interacts with the physical object (QR-code) and acquires the URI of the object. (n2) Http Request to the URI:

- If LD layer is *not set*, the URI is an active web page (public or private) which accesses the object memory, filters content and returns an HTML document.
- If the LD layer is *set*, different configuration options are available, according to LD principles: a 303 redirection with content negotiation is used to return the HTML page describing the artwork requested by the brwoser, as above; or the URI uses RDFa embedded in a HTML document (in this case the wrapper has to

use the mapping files to retrieve the RDF vocabularies and producing the RDF triples and links).

In both cases, the URI is a dynamic page which accesses the object memory (n3) and filters content based on the context of use (in this scenario, it filters out restauration and quoted vaule, that do not match the museum context).

2) Context: *museum*; application used to access the object memory: *specific App* which works as a RDF browser; LD layer: *set* . Flow: n1 and n2 are the same as in the first use case. Given that LD layer is set, the returned URI will be a RDFa document or a 303 URI. However, in this use case, the 303 redirect does not return a HTML document but RDF, since RDF browsers accept RDF+HTML. Moreover, in this case, before accessing the object memory (n3) and filtering content based on the context of use, the personalization module will access the user profile (n4) provided by the App. **Fig. 1** displays the case when the user profile is stored in an object memory representig the user. Given the user profile, the personalization module will decide whether to filter out other content or re-filter in content previously rejected. A further step is the exploration of the LOD cloud (n5) in order to follow RDF links and discover new links, as discussed in Sec. 2 and 3.

4 Next Steps and Conclusion

The project started on Jan 2013. Currently we have defined the architecture and the main modules that will perform the service. The modules for the identification of the physical object and for the association to its digital memory have already been developed. Modules 2 and 3 are under development.

An issue that has still to be faced in the workflow concerns the digital memory update, given changes in the physical object. Another issue that the current project does not address specifically but which is an interesting extension is enabling the object memory to be written by people and even to be updated, which requires defining roles and validation criteria.

The infrastructure will be useful to experiment new models of edutainment, based on using mobile devices, making education a pervasive process, which can be built and adapted to the context of interaction and to the specific user. Considering this underlying objective, another aspect of the project that requires to be detailed and improved is the provision of learning activities related to the visit of the museum or related to the artefact.

The experimental evaluation will have to test the validity of the approach and the performance of the prototype. Validating the approach can mean several features to be analyzed/compared and in particular: the user's satisfaction, her/his retention of what seen/experienced, the correctness of personalization. Concerning the performance, the evaluation will have to test in particular the time to return and visualize results. Emaldi et al [6] found that the time employed by the smartphone to visualize the information mainly depends on the number of triples to be processed and not on the number of triples retrieved. The evaluation should test the benefit of filtering content before download, based on the OMM features.

Acknowledgments. This research has been funded by the University of Genoa, within PRA 2012 projects, Prot. N. 2142.

References

1. Angelaccio, M., Basili, A., Buttarazzi A., Liguori ,W.: Smart and Mobile Access to Cultural Heritage Resources: A Case Study on Ancient Italian Renaissance Villas. In Proc. of the 2012 IEEE 21st International Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE '12), Washington, DC, USA, 310-314 (2012)
2. Barthel, R., Kröner, A., Haupert, J.: Mobile interactions with digital object memories, Pervasive and Mobile Computing 9 281–294 (2013)
3. Berners-Lee, T.: Linked Data - Design Issues. http://www.w3.org/DesignIssues/LinkedData.html (2006)
4. Bizer, C., Lehmann, J., Kobilarov, G., Auer, S., Becker, B., Cyganiak, R., Hellmann, S.: DBpedia - A crystallization point for the Web of Data. Web Semantics: Science, Services and Agents on the World Wide Web 7(3), 154-165 (2009)
5. de Boer, V., Wielemaker, J., van Gent, J., Oosterbroek, M., Hildebrand, M, Isaac, A.:, van Ossenbruggen, J., Schreiber, G. (2013). Amsterdam Museum Linked Open Data, in Semantic Web Journal, In Press, DOI 10.3233/SW-2012-0074
6. Emaldi, M., Lázaro, J., Laiseca, X, López-de-Ipiña, D.: LinkedQR: Improving Tourism Experience through Linked Data and QR Codes. Proc. of Ubiquitous Computing and Ambient Intelligence (UCAmI 2012), LNCS 7656, 371-378, Springer, December (2012)
7. Ermilov, I., Auer, S., Stadler, C.: Crowd-Sourcing the Large-Scale Semantic Mapping of Tabular Data. Proceeding of the ACM Web Science http://svn.aksw.org/papers/2013/WebSci_CSV2RDF/public.pdf (2013)
8. Heath, T., Bizer, C.: Linked Data: Evolving the Web into a Global Data Space. Synthesis Lectures on the Semantic Web, 1:1, 1-136. Morgan & Claypool (2011)
9. Mercurio, M., Torre, I., Torsani, S.: Design of Adaptive Micro-Content in Second Language Acquisition. Journal of e-Learning and Knowledge Society, 7(3), 109-119 (2011)
10. Parra, G., Klerkx, J., Duval, E.: More! Mobile interaction with linked data. In Proc. of the Int. Workshop on Data-Centric Interactions on the Web, within the 13th IFIP TC13 Conference on Human-Computer-Interaction (INTERACT 2011) vol. 817, 37-47 (2011)
11. Rukzio, E., Broll, G., Leichtenstern, K., Schmidt, A.: Mobile interaction with the real world: an evaluation and comparison of physical mobile interaction techniques, Proc. of the Conf. on Ambient Intelligence, AmI-07, Darmstadt, Germany, 7–10 Nov. (2007)
12. Solomou, G., D., Kalou, A., K., Koutsomitropoulos, D., A., Papatheodorou, T., S.: A Mashup Personalization Service Based on Semantic Web Rules and Linked Data. In Proc. of the Conference on Signal Image Technology & Internet-Based Systems (SITIS '11). IEEE Computer Society, Washington, DC, USA, 89-96 (2011)
13. Thalhammer, A.: Leveraging linked data analysis for semantic recommender systems. In Proceedings of the 9th international conference on The Semantic Web: research and applications (ESWC'12) Springer-Verlag, Berlin, Heidelberg, 823-827 (2012)