Supporting and Enhancing Research on Cultural Heritage in France: the PATRIMA Project

Dominique Laurent

ETIS Laboratory - ENSEA / UCP / CNRS, Cergy-Pontoise, France
`http://www-etis.ensea.fr`

Abstract. In this paper, we first overview the French project on heritage called PATRIMA, launched in 2011 as one of the *Projets d'investissement pour l'avenir*, a French funding program meant to last for the next ten years. The overall purpose of the PATRIMA project is to promote and fund research on various aspects of heritage presentation and preservation. Such research being interdisciplinary, research groups in history, physics, chemistry, biology and computer science are involved in this project. The PATRIMA consortium involves research groups from universities and from the main museums or cultural heritage institutions in Paris and surroundings. More specifically, the main members of the consortium are the two universities of Cergy-Pontoise and Versailles Saint-Quentin and the following famous museums or cultural institutions: Musée du Louvre, Château de Versailles, Bibliothèque nationale de France, Musée du Quai Branly, Musée Rodin.

In the second part of the paper, we focus on two projects funded by PATRIMA named EDOP and Parcours and dealing with data integration. The goal of the EDOP project is to provide users with a data space for the integration of heterogeneous information about heritage; Linked Open Data are considered for an effective access to the corresponding data sources. On the other hand, the Parcours project aims at building an ontology on the terminology about the techniques dealing with restoration and/or conservation. Such an ontology is meant to provide a common terminology to researchers using different databases and different vocabularies.

Keywords: Cultural heritage; restoration; preservation; data integration; ontology.

1 Introduction

Five years ago the French government launched a national call for *investissements d'avenir* (or investments for the future in English) for a global amount of 35 billions Euros, among which 7.9 are devoted to research and 11 to higher education. Up to now 224 projects are running in this framework, and the PATRIMA project is one of them. This project started in 2011, initiated by the universities of Cergy Pontoise and Versailles Saint-Quentin, both located next to Paris and close to each other.

The overall purpose of the PATRIM A project is to promote and fund research on various aspects of heritage presentation and preservation. Such research being interdisciplinary, research groups in literature, history, physics, chemistry and computer science are involved in this project. The PATRIM A consortium involves research groups from universities and from the main museums or cultural heritage institutions in Paris and surroundings. More specifically, additionally to the two universities of Cergy Pontoise and Versailles Saint-Quentin, the main members from the cultural heritage side of the consortium are *Château de Versailles, Musée du Louvre, Musée Rodin, Musée du Quai Branly, Bibliothèque nationale de France*.

The remainder of this paper is structured as follows, in Section 2 we introduce the PATRIM A project in more details to give the reader a rough idea of the main facets of research issues addressed in this project. Then, in Section 3, the focus is put on two ongoing projects funded by PATRIMA in the domain of data integration: the *Espace de Données pour les Objects du Patrimoine Culturel* (EDOP) project (Data Space for Objects on Cultural Heritage in English) and the *PAtrimoine culturel et Restauration-Conservation : Ontologie pour l'Usage d'un Référentiel commun aux différentes Sources de données* (Parcours) project (Cultural Heritage and Restoration-Conservation: an Ontology for referring to different data sources in English). In Section 4 we propose concluding remarks about the project.

2   The PATRIMA Project

The PATRIMA project has two distinct components, called respectively *laboratoire d'excellence* or labex for short and *équipement d'excellence* or equipex for short, each corresponding to different calls in the framework of *investissements d'avenir*. These two components of the project were accepted for a duration of 8 years, with a budget of 7 million Euros for the labex and 8 million Euros for equipex. They are both under a foundation called *Fondation des Sciences du Patrimoine*[1] and each component is meant to fund different kinds of actions as listed below:

— The labex funds mainly post-docs, Ph.D students, master internships, event organization and participation, publications, and more generally, scientific collaborations between the different laboratories in the project.
— On the other hand, the equipex is devoted to the funding of some of the necessary technical equipments for the project achievements.

The main objectives of the PATRIMA project are to investigate the following three issues related to cultural heritage:

1. improve the knowledge on cultural heritage,
2. improve the techniques for preservation and restoration, and
3. improve cultural outreach either in the museums or through médias such as the internet.
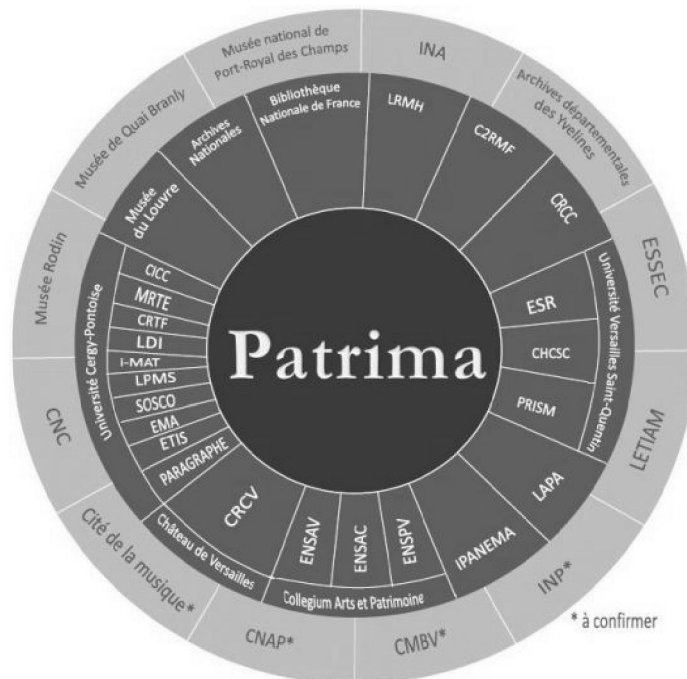
---

[1]   www.sciences-patrimoine.org

Fig. 1. The partners of PATRIMA

To reach these goals, institutions from the academic and heritage sides are involved and researchers from different domains, ranging from literature and history to chemistry or computer science, participate in research under the project. Figure 1 displays all institutions and laboratories involved in the project, where the inner dark part of the disc shows the founding partners and laboratories and the outer bright part shows partners that joined the project after its submission. Without going into the details of every participating institution or laboratory, the following items are meant to give an idea on the most renown members of the project. The main academic institutions or laboratories are the following:

— The University of Cergy Pontoise with 10 laboratories in literature (CRTF), linguistics (LDI), history (CICC), chemistry (SOSCO, I-MAT), physics (LPMS), didactics (EMA), humanities and information communication

— (PARAGRAPHE), computer science (ETIS), environment and geography (MRTE);

— The University of Versailles Saint-Quentin with 3 laboratories in literature (ESR), history (CHCSC) and computer science (PRISM);

— LETIAM (University Paris 11), a laboratory in chemistry, specializing in techniques for preservation;

— LAPA (CNRS, CEA[2]), a laboratory specializing in metal protection;
— ESSEC, one of the top-three business schools in France.

On the other hand, the following museums, institutions and laboratories from the cultural heritage side are also involved in the project:

— Château de Versailles, Musée du Louvre, Musée Rodin, Musée du Quai Branly,
— Bibliothèque nationale de France, Archives nationales, INA (Institut National de l'Audiovisuel),
— IPANEMA, LRMH, C2RMF, CRCC four laboratories in the domain of preservation and restoration.

As earlier mentioned, the PATRIMA project currently funds about 30 Ph.D theses among which we cite the following as examples:

— Historical and scientific studies of Islamic ceramics from the 13th to the 18th centuries,
— Pigment characterization using non invasive techniques,
— A robot in the museum.

As examples of research projects funded by PATRIMA, we mention

— the VESPERA project whose goal is the digitalization of the maps of the château de Versailles,
— the ALBATRE project which aims at studying chemical properties of alabaster and the way it has been used in arts,
— a project on seals in the Middle Ages.

In the next section we give details on two other ongoing projects funded by PATRIMA: one called EDOP is about information systems and is the subject of a Ph.D thesis, and the other one is a collaborative research project whose purpose is to build an ontology on terms used for preservation and restoration.

3 Two Ongoing PATRIMA Projects

3.1 The EDOP Project

The project named *Espace de Données pour les Objects du Patrimoine Culturel* (or Data Space for Objects on Cultural Heritage in English) and abbreviated into EDOP, aims at building a data space for integrating all kinds of data associated to the cultural heritage objects. This is a three year project (2012-2015) whose funding is mainly devoted to a Ph.D. This project involves:

— the two laboratories in computer science ETIS[3] and PRISM[4] respectively from the University of Cergy Pontoise and the University of Versailles Saint-Quentin,

---

[2] CNRS is the acronym for the French Council for Scientific Research and CEA is the acronym for the French Agency for Atomic Energy and Alternative Energies.

— Bibliothèque nationale de France[5], Château de Versailles[6] and Musée Rodin[7].

The difficulty in the context of the project is that, due to the size and the heterogeneity of the data to be integrated, traditional approaches to data integration cannot be successfully applied. This is so because, in the context of the project, it is almost impossible to design a data model for all objects of interest, and even if such a model could be found, the links between the actual data models and the common model would be very difficult to set.

To overcome these difficulties, the purpose of the project is to build a data space using Linked Open Data techniques ([1]). Such an approach is indeed more flexible than traditional approaches to data integration because no global data model is needed. The integration is achieved through semantic links between the knowledge bases to be integrated, and these links are expressed in terms of the knowledge bases themselves (and not in terms of a common model). More precisely, assuming that RDF[8] (Resource Description Framework) semantic descriptions of the contents of each data source are available, the semantic links between data sources are also expressed using the RDF framework, and it is hoped that these links can be (partially) *automatically* generated. The data to be integrated in a first stage are the following:

— At Bibliothèque nationale de France: all data about authors, books, documents stored in the French national library. This huge data set has been the subject of two integration projects called Data.bnf.fr (http://data.bnf.fr) and Gallica (http://gallica.bnf.fr).
— At Château de Versailles: textual data describing people, plants and bibliographic items related to the castle and its history.
— At Musée Rodin: all data related to the object owned by the museum.
— The data sources DBpedia ([2]) and Yago ([3]).

Figure 2 depicts the architecture of the targeted integration system. The links mentioned in this figure are those links that are built based on RDF semantic descriptions of the data.

The overall approach works as follows: given a user query Q, based on the local knowledge base, Q is first translated into queries to each data component of the data space. These queries are processed by the components and their answers are merged to produce the global answer to Q. This global answer is then integrated to the local knowledge base (with the goal of possibly enriching the links between the data source semantics) and presented to the user.

---

3   http://www-etis.ensea.fr
4   http://www.prism.uvsq.fr
5   http://www.bnf.fr
6   http://chateauversailles-recherche.fr
7   http://www.musee-rodin.fr
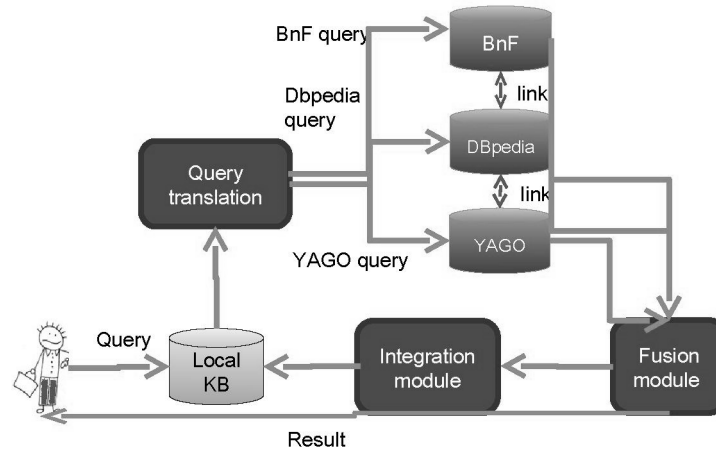8   http://www.w3.org/2001/sw/wiki/RDF

Fig. 2. The integration architecture of the EDOP project

3.2 The Parcours Project

The project named *PAtrimoine culturel et Restauration-Conservation : Ontologie pour l'Usage d'un Référentiel commun aux différentes Sources de données* (Cultural Heritage and Restoration-Conservation: an Ontology for referring to different data sources in English) and abbreviated into Parcours, aims at providing users from different laboratories and institutions with a "unified global access" to data related to techniques and practices in preservation and restoration of heritage objects. As the EDOP project, this project is about data integration, but now the main problem is that, additionally to have this unified access, users wish to keep with their own data stored in their own database.

To reach this goal, it has been decided that the first step would be the definition and construction of an ontology unifying the terms used by all practitioners. The participating institutions and laboratories are ETIS from the academic side and the following laboratories from the heritage side:

— LRMH[9] [*Laboratoire de Recherche des Monuments Historiques*, or Laboratory for Research on Historical Monuments in English) working on issues related to the preservation and restoration of materials such as stone, wood, glass, concrete, metals or textiles.
— C2RMF[10] (*Centre de Recherche et de Restauration des Musées de France*, or Center for Research and Restoration of French Museums in English) working on research, restoration and archives for almost all kinds of art pieces.
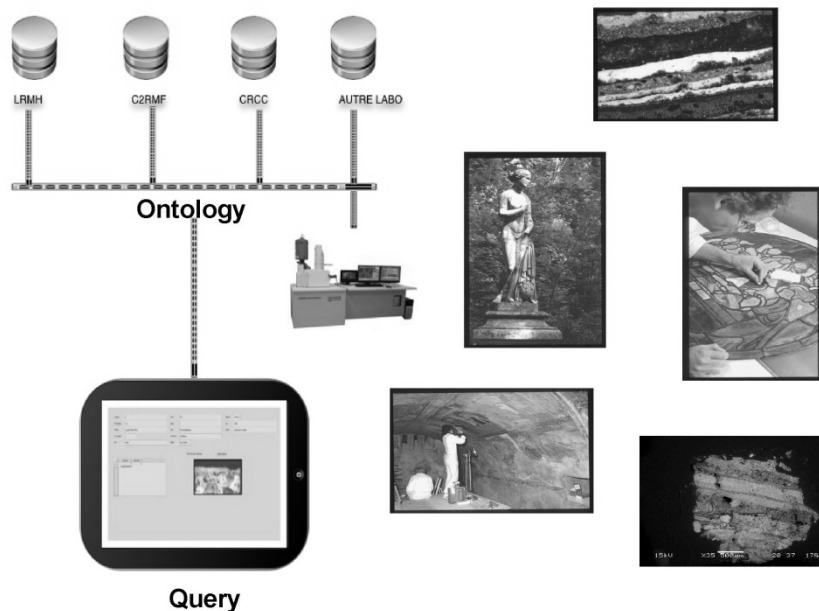

Fig. 3. The architecture of the integration system for the Parcours project

— CRCC[11] (*Centre de Recherche sur la Conservation des Collections*, Center for Research and Preservation of Collections in English) specializing on the preservation and restoration of photographs, films, digital supports, and objects of natural history.

Three main steps have been identified in this project: (*i*) build an ontology federating the terminologies used by researchers and practitioners in the domains of preservation and restoration, (*ii*) integrate the data from each institution based on the ontology (while preserving each local data source and their current proper use), and (*iii*) provide a tool box for the analysis and mining of these integrated data. Figure 3 shows the expected global architecture of the targeted system, along with on its right hand side, pictures of restoration activities.

4 Conclusion

In this paper, we have presented the overall purpose of a project on heritage called PATRIMA, funded in the general framework of the French *investissements d'avenir programme*. This project allows academics and people from cultural institutions and museums to collaborate in various scientific domains; examples of such collaborations in the domain of data integration have been discussed.

---

[11] http://www.crcc.cnrs.fr

The fact that the PATRIMA project involves different kinds of institutions and laboratories, addressing different research domains is of course a very exciting and fruitful feature, at least as important as the amount of money that has been and will be put to fund research! A last point worth mentioning is that extending the project to partners from countries other than France is an important issue that we are considering. In this respect, the H2020 framework from the European Union offers opportunities that will be investigated in the next future.

References

1. Sôren Auer, Christian Bizer, Georgi Kobilarov, Jens Lehmann, Richard Cyganiak, and Zachary G. Ives. Dbpedia: A nucleus for a web of open data. In *ISWC/ASWC*, volume 4825 of *Lecture Notes in Computer Science*, pages 722-735. Springer, 2007.
2. Christian Bizer, Jens Lehmann, Georgi Kobilarov, Sôren Auer, Christian Becker, Richard Cyganiak, and Sebastian Hellmann. Dbpedia - a crystallization point for the web of data. *J. Web Sem.,* 7(3):154-165, 2009.
3. Fabian M. Suchanek, Gjergji Kasneci, and Gerhard Weikum. Yago: a core of semantic knowledge. In *WWW*, pages 697-706, 2007.