



**HAL**  
open science

## Leveraging heterogeneous Cultural Heritage data to promote tourism

Marie-Noelle Bessagnet, Patrick Etcheverry, Annig Le Parc-Lacayrelle, Christophe Marquesuzaà, Landy Rajaonarivo, Philippe Roose, André Sales, Christian Sallaberry

### ► To cite this version:

Marie-Noelle Bessagnet, Patrick Etcheverry, Annig Le Parc-Lacayrelle, Christophe Marquesuzaà, Landy Rajaonarivo, et al.. Leveraging heterogeneous Cultural Heritage data to promote tourism. Open Source Geospatial Research and Education Symposiums (OGRES), Oct 2018, Lugano, Switzerland. hal-01976405

**HAL Id: hal-01976405**

**<https://hal.science/hal-01976405>**

Submitted on 10 Jan 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Leveraging heterogeneous Cultural Heritage data to promote tourism

Marie-Noelle Bessagnet, Patrick Etcheverry, , Annig Lacayrelle, Christophe Marquesuzaà,  
Landy Rajaonarivo, Philippe Roose, André Sales Fonteles, Christian Sallaberry

Univ. Pau & Pays Adour /E2S UPPA

Laboratoire d'Informatique de L'Université de Pau et des Pays de l'Adour

2 allée du Parc Montaury, 64600 Anglet – France

{FirstName.LastName@univ-pau.fr}

## Abstract

The objective of our works is to propose an open source software able to manage data source heterogeneity and to generate other ones in order to propose context oriented and on the fly itineraries according to several criterions (time, means of transport, personal interest, current close events, etc.). In this first step, we propose a software architecture and an implementation in order to manage data source heterogeneity.

## 1. Introduction

This paper presents part of the European FEDER Project TCVPYR, which aims to promote tourism in the French Pyrenees region by leveraging elements of its cultural heritage. The TCVPYR is a multidisciplinary project involving scientists from various domains, including: computers scientists, geographers, historians and anthropologists. In order to achieve its goal, some of the TCVPYR researchers are currently collecting georeferenced cultural heritage data in different areas of the Pyrenees. This data, together with data from the local governments and the Open Data, is intended to feed a mobile application that promotes the tourism in the region. The mobile application will allow tourists, but also scientists, to access cultural heritage data in the form of points of interest (POI). Moreover, these POI are to be presented according to the user profile and the environmental context, including spatiotemporal aspects such as his current location. For example, the application may suggest to a tourist an itinerary with several POI taking into account his interests, means of locomotion, available time and mobile features (battery level, internet connexion...). All data collected in this project as well as the final application will be published as open data and open source.

The first major challenge in this project consists in managing the heterogeneity of the several different sources of data we plan to use. For example, the French government currently employs at least two different software each one with different data model (format) to storage its data. Similarly, a plethora of data models are used in the countless sources of Open Data. In order to handle this challenge, we propose a unified model that allows us to gather all this data we plan to use and to import it into a single database suitable for further exploitation.

The second major challenge is related to leverage the cultural heritage data stored in our unified database. In particular, we plan to elaborate a set of algorithms and heuristics to propose contextualized itineraries (i.e., sequences of POI to be visited) on the fly to the users of the mobile application. This recommendation must take into account the user profile, his location, his device features and also some other parameters such as available time for visiting the POI and means of transport. A recommendation must also be enriched with information extracted from open data. For example, if an exhibition is happening close to the user location, the user itinerary may be enriched with a link to a mobile ticketing service. Another important approach we plan to adopt in order to leverage the cultural heritage data is to publish it as open data. In this direction, we are currently developing a set of tools and API that will allow anyone to download our data in open popular formats (e.g., JSON, XML or CSV). Moreover, we are investigating and experimenting on the creation robots to publish our data directly on well-known and visited open data providers such as Wikipedia and OpenStreetMap.

## 2. Related works

The contribution presented in this paper deals with patrimonial data valorization through tourism vector. As presented in (Benhamou and Thesmar, 2011) which highlight the close links between culture, tourism and economy, we are also positioning the cultural heritage at the service of the tourism and the attractiveness of the territory. This issue is also related to the "globalism" phenomenon addressed in papers such as (Baumann, 2002) and (Angelaccio, 2017). The Social Smart Tourism project presented in (Angelaccio, 2017) aims at developing services to improve the quality of life by highlighting places and resources insufficiently known. Our project shares some common goals without directly focusing on the social aggregation aspect.

[Hui et al, 2015] propose an algorithm called PERSTOUR for recommending personalized tours using POI popularity and user interest preferences, which are automatically derived from real-life travel sequences based on geotagged photos. The algorithm uses trip constraints such as time limits and the need to start and end at specific POIs. Such approach is interesting as in our proposition, available time is also an important information in order to propose an itinerary. [Ramalho et al., 2015] propose TripBuilder, an unsupervised framework for planning personalized sightseeing tours in cities using collected and categorized POIs from Wikipedia and albums of geo-referenced photos from Flickr. Such traces give information about tourists' behaviors. Spatio-temporal information obtained from photo metadata are compared with city POIs in order to obtain a measure of personal interest for the user given her preferences and visiting time-budget is maximized. [Colomo et al., 2017] present POST-VIA 360, a platform to support the whole life-cycle of tourism loyalty after the first visit. The objective of the platform is to collect and analyse data and to propose (using positioning and bio-inspired recommender systems) relevant recommendations.

As we can see on previous works [Fonteles, 2016] or [Rajaonarivo, 2017], many works and prototypes are now available in order to increase the value of tourism oriented data (including POI, traces) and propose context and user oriented itineraries. Unfortunately, few of them use public data and/or are freely available.

### 3. Proposal

This project is part of the European ERDF Project TCVPYR. This project is a multidisciplinary one involving geographers, historian, anthropologists and computer scientists. The objective consists in collecting data on the topic of thermal cities and increasing their value for the public. Both data and applications will be open-source. The final application will permit tourists to express some requirements and to get itinerary propositions according to the time they have, their means of transport, their personal interest and knowledge skills, etc.

In order to collect patrimonial information, researchers draw up an inventory of georeferenced cultural heritage data using two different systems: RenabLP2 and Gertrude. These two systems allow geographers to structure, to organize and to describe all the patrimonial collected data. RenabLP2 is used in the Occitanie region and Gertrude is both used in the new Aquitanian and in the Occitanie regions.

The idea is to gather and to merge all the collected data with these two systems in a same database (named TCVPYR in Figure 1) that could be used for touristic purposes (Figure 4) but also that could be shared through open data platforms (Figure 3).

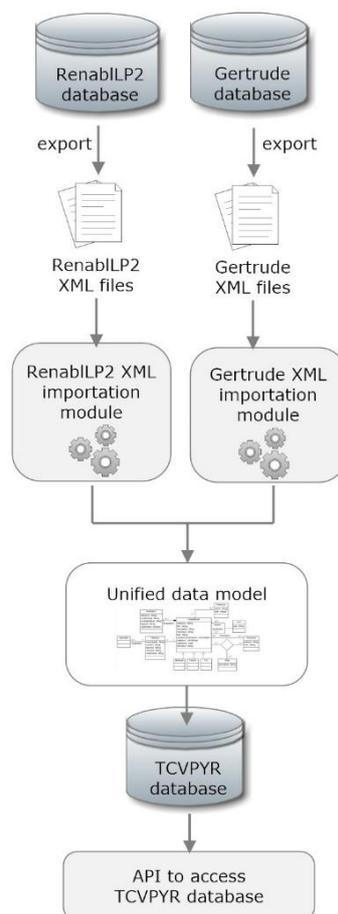


Figure 1: Patrimonial data merge process

Even if RenabLP2 and Gertrude systems can export data in XML format, resulting exports are very different in terms of structure, data organization or information naming. In order to

emphasize these heterogeneous data within applications dedicated to tourism, we designed a unified and homogeneous data model (Figure 2).

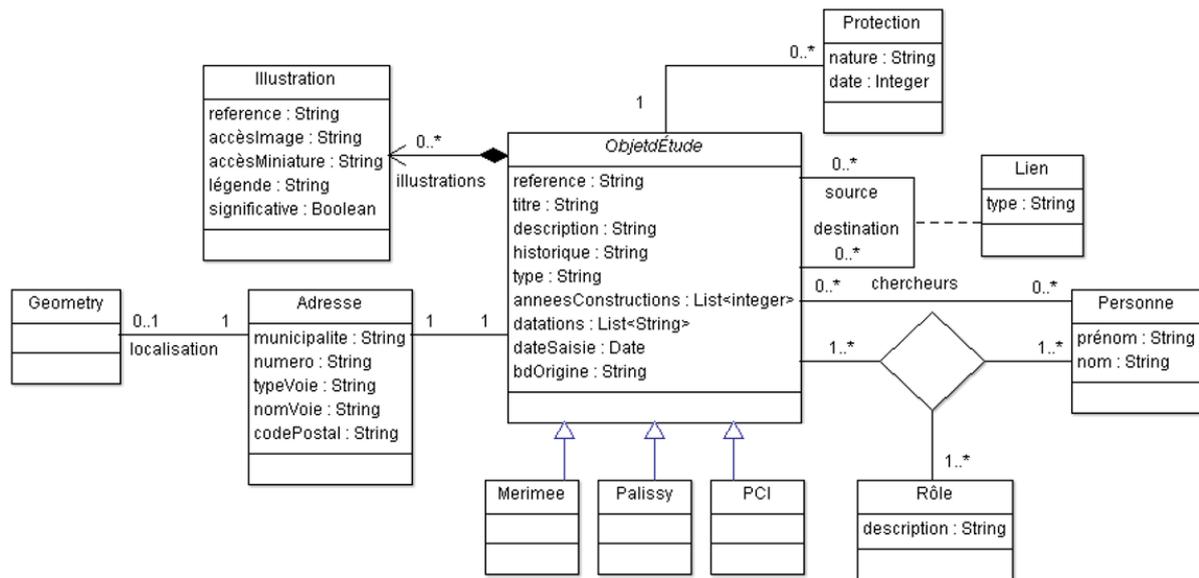


Figure 2: Unified patrimonial data model for tourist use

This model can describe all the study objects collected from the field: buildings, furnitures and all elements related to intangible cultural heritage. Each object ("ObjetdÉtude" class) has a description ("description" property), an history ("historique" property) and can be geolocated with specific geographic coordinates ("Adresse" and "Geometry" classes). Some object can also be represented by one or more images ("Illustration" class). The model allows also geographers to reference, for each object, persons ("Personne" class) related with it: the architect of a building, the sculptor of a statue and so on, according to this/her role ("Rôle" class). Of course, several objects can be interrelated ("Lien" association class): a building can be composed of several buildings, a furniture can be located in a building, etc.

This model is not intended to describe all the complexity of the patrimonial domain but only the set of information that can be later exploited and highlighted within a tourist context.

#### 4. Prototype

In order to leverage all the cultural heritage data merged in our unified database, we plan to develop two applications: an application dedicated to share all the collected data through open data platforms (Figure 3) and another one dedicated to tourists (Figure 4).

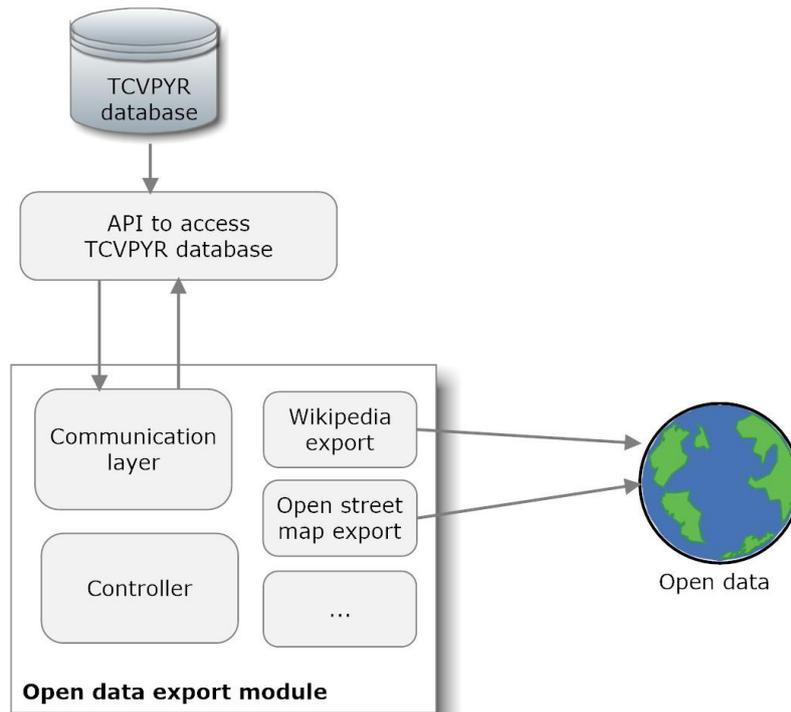


Figure 3: Patrimonial data dissemination through open data platforms.

The application dedicated to share the collected patrimonial data exploits the API developed to access the TCVPYR database. The communication layer of this application connects with the TCVPYR API in order to get the data that must be shared on open data platforms. We plan to develop several export modules, each one aiming at distributing patrimonial data on a specific open data platform such as Wikipedia or Open Street map.

We developed a first prototype which is able to publish patrimonial data of specific patrimonial objects on Wikipedia. This component, developed in Java, queries the TCVPYR unified database and uses the JWiki API (<https://github.com/fastily/jwiki>) to publish each point of interest in the form of a Wikipedia article. Figure 4 shows an example of result for a patrimonial object (Champsaurai Castle) stored in the TCVPYR database.



Figure 4: An example of a patrimonial object exported to Wikipedia.

Concerning the dissemination of patrimonial data through the tourist vector, we plan to develop a classic client/server application allowing a tourist to access patrimonial data during a trip. The idea consists in providing customized data according to the user profile but also the context in which he moves (Figure 5).

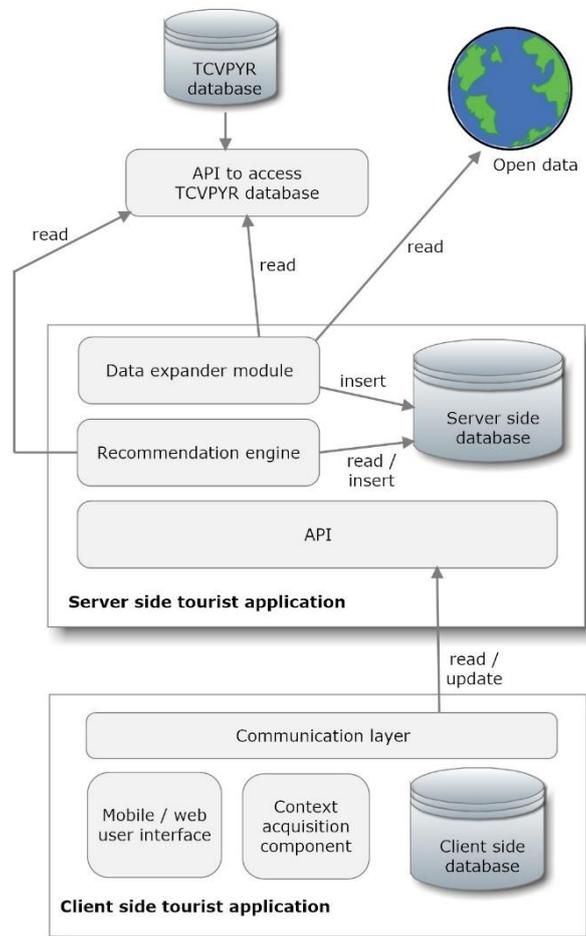


Figure 5: Patrimonial data dissemination through a tourist application.

The client side application takes the form of a mobile application that will exploit the mobile components to get contextual information: location, battery level, internet connexion type, etc. The user will also be able to define personal preferences such as the time available to achieve his trip, his/her means of travel, preferences in terms of things to see during his/her trip...

According to all this gathered information, the server side application will search for relevant patrimonial information matching with the tourist context and profile. To carry out this task the server side application relies on a recommendation engine that will exploit all the collected data concerning the user in order to search for relevant information in the TCVPYR database but also on open data platforms.

We are currently working towards implementing the technical architecture of this application and we have developed a first prototype allowing the user to access, without any filter, all the patrimonial data stored in the TCVPYR database (Figure 6). The next step will consist in designing the recommendation engine to take into account the user context and preferences in order to enhance information retrieval.

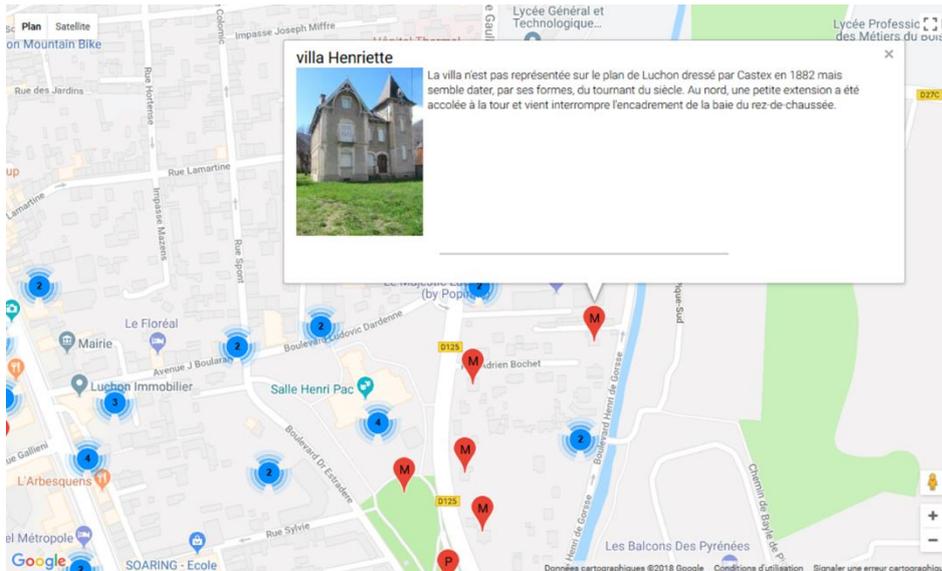


Figure 6: An example of a first prototype for the tourist application.

## 5. Conclusion and future work

This paper proposes a unified model, and architecture and a software solution in order to get a basic solution for further data integration. We are also proposed a tool in order to extract tourism oriented data and export them to a Wikipedia format.

The next step will now be the on the fly itinerary proposition according to the complex context (user profile, means of locomotion, time, interests, events, etc.

## 6. Acknowledgement

These work are supported by the European Regional Development Fund (ERDF). Project TCV-PYR (2017-2020), in collaboration with Occitanie-Méditerranée and Nouvelle-Aquitaine Regions.

## 7. Bibliography

Rajaonarivo, 2017 - Landy Rajaonarivo, Matthieu Courgeon, Eric Maisel, Pierre De Loor -  
 Inline Co-Evolution between Users and Information Presentation for Data Exploration  
 - IUI '17 Proceedings of the 22nd International Conference on Intelligent User Interfaces  
 Pages 215-219 - Limassol, Cyprus — March 13 - 16, 2017 - doi>10.1145/3025171.3025226

Fonteles, 2016 - André Sales Fonteles, Sylvain Bouveret & Jérôme Gensel (2016) Trajectory  
 recommendation for task accomplishment in crowdsourcing – a model to favour different  
 actors, Journal of Location Based Services, 10:2, 125-  
 141, DOI: 10.1080/17489725.2016.1184770

Kwan Hui Lim, Jeffrey Chan, Christopher Leckie and Shanika Karunasekera - Personalized  
 Tour Recommendation Based on User Interests and Points of Interest Visit Durations -

Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2015)

Igo Ramalho Brilhante, Jose Antonio Macedo, Franco Maria Nardini, Raffaele Perego, Chiara Renso - On planning sightseeing tours with TripBuilder - Information Processing & Management Volume 51, Issue 2, March 2015, Pages 1-15

Ricardo Colomo, Palacios Francisco José García-Peñalvo, Vladimir Stantchev, Sanjay Misra - Towards a social and context-aware mobile recommendation system for tourism - Pervasive and Mobile Computing - Volume 38, Part 2, July 2017, Pages 505-515

Benhamou F., Thesmar, D. (2011). Valoriser le patrimoine culturel de la France. Conseil d'analyse économique. ISBN: 978-2-11-008595-5, Paris.

Bauman Zygmunt, Society under siege, Ed. Wiley, 264 pages, ISBN: 978-0-745-62984-1, Nov. 2002.

Angelaccio Michele. (2017). Social Smarttourism: A Sustainable Development Methodology for Smart Urban Lands. Athens Journal of Tourism. 4. 111. 10.30958/ajt.4.2.2.