

AN UPDATE FROM THE RESEED PROJECT

L. Jeanson^{1,2,*}, M. Cotte³, J.L. Kerouanton¹, F. Laroche²

¹ Centre François Viète, Université de Nantes, 2, rue de la Houssinière 44300 Nantes, France
(loic.jeanson, jean-louis.kerouanton)@univ-nantes.fr

² Laboratoire des Sciences du Numérique de Nantes (LS2N), Ecole Centrale de Nantes
Université de Nantes, 1, rue de la Noë, 44300 Nantes, France - (loic.jeanson, florent.laroche) @ls2n.fr

³ MCC Heritage, 23 rue du Général Chapelle, 07300 Tournon sur Rhone, France - cotte.michel@orange.fr

Commission II, WG II/8

KEY WORDS: Linked Data, 3D scanning, Data integration, Cultural Heritage

ABSTRACT:

Following a first presentation of the works performed by our project group at ICOMOS General Assembly in 2017, it was time for us to bring an update, oriented on our technological and methodological choices and implementations. This paper presents the evolution of our data management system.

1. INTRODUCTION

The ReSeed project, which name is a acronym in French, standing for Semantic reverse engineering of digital cultural heritage objects, is public funded by the ANR (French National Agency for Research). Its members decided to join together to explore and maybe develop digital tools and approaches regarding the heritage understanding and analysis of cultural heritage objects. The group formed mostly over past common projects based on reverse engineering (Ali et al., 2015), Knowledge engineering, and cultural heritage lifecycle analysis (Laroche et al., 2008). From an administrative perspective, the Ecole Centrale de Nantes, UTT (Technology University of Troyes), UTC (Technology University of Compiègne) as well as University of Nantes joined forces in order to combine their experiences in data and knowledge engineering, mechanical engineering and epistemology, and history of sciences and technology. They have been strengthened by institutions: *l'Inventaire Général* (The general Inventory) from the French Cultural Ministry, and the *Musée des Arts et Métiers* (Arts and Crafts Museum) ; companies have also joined : MCC Heritage, a cultural heritage sites and objects counselling office, and DeltaCAD, a 3D point cloud management software development company.

The first intuition was the need of digital modelling in order to help with :

- 3D animations for the in order to explain complex systems
- Collaboration / multi-user activities and hypothesis
- Data integration from multiple sources

Three case studies have been chosen by the project members in order to approach a wide range of issues.

- A physical mock-up (at 1/10) of a colonial metallic house (ca. 1890) from the firm Moreau frères.

*Corresponding author

- A series of seven meridian telescopes (ca. 1880- ca. 1890) from Gautier, instrument maker for the *Observatoire de Paris* at that time, spread in observatories across France
- *L'Observatoire du Pic du Midi*, a high-mountain (2800m) scientific station in the Pyrénées, that will be presented with more details in the next parts of this paper.

The case studies vary in size and granularity, but also in the angles of the research questions. This is not the place to detail them for the three objects. As we did in the early communication at ICOMOS (Cotte et al., December 2017), the *Observatoire du Pic du Midi* will be our focus. The perspective is one of a counselling office for a potential heritage site candidate. On a heritage analysis point of view, the objective is - very classically - to assess the site's relevance and to determine where reside the values, to be able to compare it with national and international sites. On a technical point of view the objectives are multiples and evolving with time as this paper will show. In a first section, a short recap of the early phases is needed in order to understand the evolution we have since performed, that will be described in a second section. Finally, the lately emerging problems and intended roadmap will constitute our last section.

2. EARLY PHASES

As we started with no digital data, we started in parallel classical heritage analysis with the conception of a digital restitution tool. It has been described with more details in (Cotte et al., December 2017) but for more than mere a reference, the conceptual and technical approach of the digital restitution tool will be described in this section. To put the site in context, it all started as a temporary scientific station, with traces as early as 1706 when Plantade observed the sun crown during an eclipse. The station's first building was erected in 1874, in order to store safely the astronomical instruments, followed one year later by the first grading of the peak. The high-altitude localization of the *Pic du Midi* did not discourage regular stays, even in winters, to happen as soon from 1875. In 1880, the first building *Bâtiment Nansouty* was finished, allowing for the first wintering in the winter 1880-1881. The various phases of



Figure 1. A view from the *Observatoire du Pic du Midi* in the late 70's taken from a helicopter

later construction and activities that made the site as complex as it is today will not be presented in this paper, but for an estimate, the site is nowadays roughly a complex structure, build on and around the mountain, of large dimensions: ca. 250m of length and 30m wide. It on some buildings have up to seven floors. The peak nowadays also integrates a telecommunication station built between 1960 and 1962. The historiographic study and heritage value analysis produced a bibliography with reference data of interest on the site's history. Various research activities have been performed on this scientific platform, in various scientific domains: astronomy and astrophysics, significantly, but also botanic, aerology and meteorology. As one can expect, such research activities have been run by several scientists. Another possible expectation on this site is the constant need for heavy maintenance of the buildings, in order to last through a series of freezes and thaws, combined with the constant need for updating the scientific instrumentation and platforms in order to follow the progress of both sciences and techniques/technologies.

With this short first elements started our digital works between data management in a database and the making of a 3D model.

2.1 The data model

Data collected from the source was structured in a relational database, with a fix, predetermined data model, based, with more or less rigor to the CIDOC CRM. It models, our previously stated interests over the sites, classified under following denominations :

CIDOC-CRM		Label (en)	Label (fr)
#	Label		
E53	Place	Place	<i>Lieu</i>
E39	Actor	Person	<i>Personne</i>
E71	MMT ¹	Building	<i>Bâtiment</i>
E40	Legal Body	Organization	<i>Organisation</i>
E71	MMT	Instrument	<i>Instrument</i>
E71	MMT	Facility	<i>Installation</i>
E11	Modification	Works	<i>Travaux</i>
E28	CO ²	Field	<i>Domaine</i>
E55	Type	Type	<i>Type</i>
E7	Activity	Activity	<i>Activité</i>
E84	IC ³	Source	<i>Source</i>
E82	AA ⁴	Appellation	<i>Appellation</i>

Table 1. Types of entities in the data model used in the early phase

It is visible that our mapping is, on several levels, very approximative, sometimes basing ourselves on vague categories, or mapping several different types of entities to the same CIDOC-Entity. We knew this was a first approach in order to get our hands-on data and its manipulation.

In a similar manner, we worked with 22 relationships mapped again more or less rigorously to CIDOC-CRM Properties in order to create our relationships. The database enable for data entry with the profile of n-quads (as for in n-triples format), dedicated to the source of the data. This way, we were able to link every data entry in our relationship database to the source document containing the assertion. No automatization was used in the building of our database. All the data extraction and compiling was done by ourselves and concerned at this stage of a dozen of reference documents (with a strong influence of the book from Emmanuel Davoust on the history of the scientific station (Davoust, 2000))

2.2 The 3D models

Concurrently from a video recording of the site made from a helicopter, we extracted pictures in to build a very rough photogrammetric model.

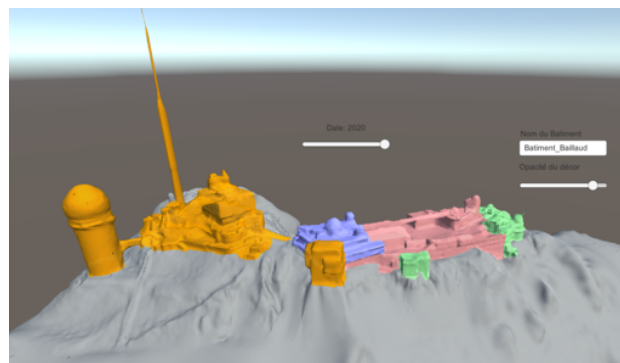


Figure 2. The global photogrammetric model in the Unity application

It was then segmented after the various named buildings emerging from the mountain. Then manual restitution, with a equivalent roughness level was performed in order to be able to show the evolution of the site through the time. The remodeling of past states was performed manually with reference dated pictures for model. The very low quality photogrammetry model and our remodeling seemed to us on comparable level of precision : a very low. We had no intention for esthetics or accuracy, we wanted a first sketch, supporting heritage experts explanations of the sites.

2.3 Models combining

With those two modellings, one focused on (very rough) geometry and another one on conceptually representing the life of the site, we thought is we had enough for a first implementation. We produced a *Unity* application containing the various models and a temporal slider, as can be seen in 2. We restricted the access on the database to very simple information : The name of the building (if it changed with time), and the shape of

¹Man Made Thing

²Conceptual Object

³Information Carrier

⁴Actor appellation

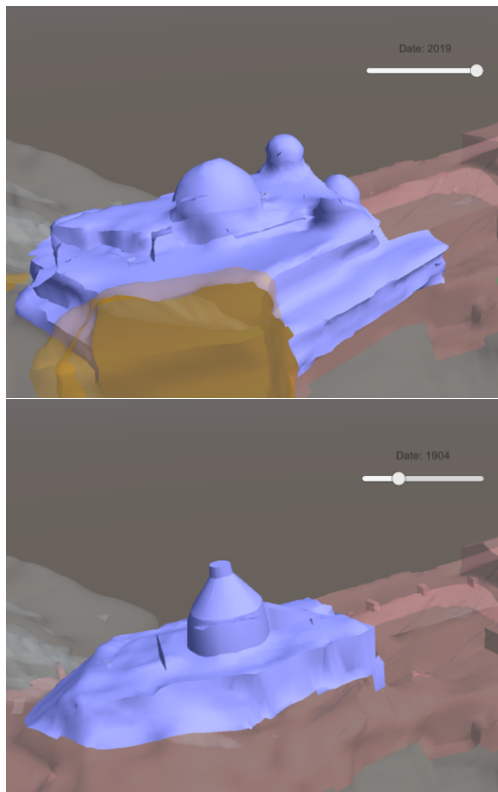


Figure 3. Top : A close up of the *Batiment Baillaud*, from the low quality photogrammetry model — Bottom : A close up of the remodeling of the state in 1904 of the same building

the building (if *Works* were to happen in the year specified by the slider, then the relevant temporal model would appear the virtual environment instead of the previous one).

This worked great, but we were not able to go any further on this path for two main reasons:

- The connection of a SQL database with the *Unity* game engine was achieved with an *Unity* extension, that was no longer maintained and got depreciated with the next *Unity* update: we were completely dependent on others to continue to work with this pipeline.
- Our very rough conceptual and geometrical approach was not satisfying anymore. We needed to precise more specific details, we needed more detailed classification possibilities. Also, we wanted to try to use directly Semantic Web technologies, and not only its formalism in a relational database.

So we needed for several technological and methodological changes that will be presented in the next section.

3. CURRENT STATE

This section reflects works done between April 2018 and April 2019. It impacts significantly our 3D data and probably even more our database modelling of the site. This section will try to reflect those changes, starting with our 3D scanning of the *Observatoire du Pic du Midi*, and then two sections dedicated to our data management, one focused on the tools we used, another on the stronger integration we have been trying to achieve.

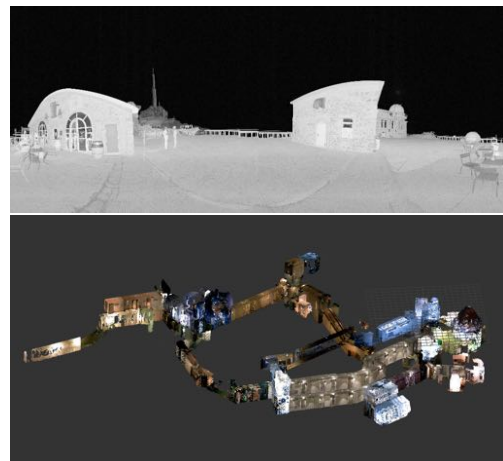


Figure 4. Top : Example of exterior scan of the terrace and buildings — Bottom : A sample of the interior scan of the buildings

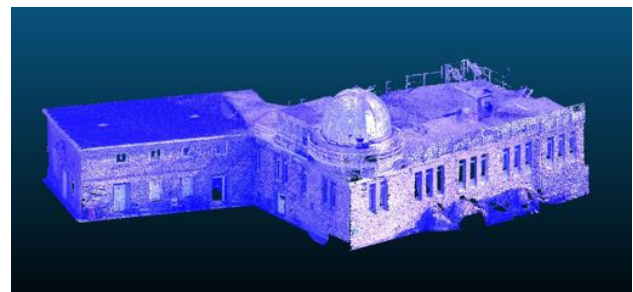


Figure 5. A segmentation of the *Vaussenat* building

3.1 3D model approach change

The photogrammetry model produced from the video we first used was insufficient, so we decided to go on a field trip and 3D scan the site. Because the team had not much experience with the digitization of such a big site, we took 3 different systems: a Faro Focus 3D, a Matterport Pro2 and a Leica BLK360. We underwent scanning the exterior and the interior of the scientific station.

The main challenge was to deal with the various file formats and the global aligning and merging of the stations. We ended up with .e57 and .pts files, so only point clouds, with sufficient density to identify the buildings clearly (for example 4). The merges stations were then segmented in order to get the geometry of individual buildings (as can be seen in 5).

Although *Unity* is able to deal with point clouds, the amount of data we are now dealing with required first to develop another approach. Since the rough shape of the buildings was enough for in our first approach, we kept the same goal in terms of level of detail : no intention for aesthetics or heavy resolution. This time, we only wanted the precision to be largely improved. In this sense, we developed a Python written algorithm for automatic reconstruction and simplification of the global shape of the buildings, based on the methodology proposed by H el ene Macher's thesis (Macher, 2017). The reconstruction includes floors, ceilings and walls in a first series of steps, and then is able to go further up to the identification of openings (doors and windows), but for this further steps, the settings is more dependent on the scan quality and on the design of the openings

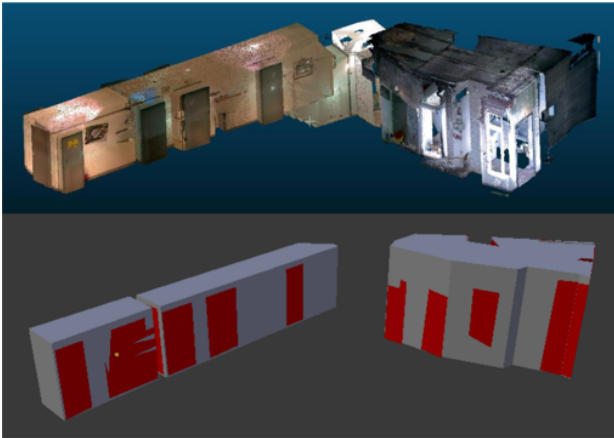


Figure 6. An example of building reconstruction

so needed more human assistance.

Although, with a tweaks (such as (Leon, 2018)), it is possible to import point clouds within *Unity*, we decided to explore another tool for the representation of our 3D data. This will be developed in the next section, but before exposing it, we need to present the database system change that was performed.

3.2 Database change

This change reflects more a will on our side to explore linked data technology in order to integration ontology formatted data from reference sources (such as BnF (French National Library) and DBPedia, for now). Since our data was already more or less formatted with standardized vocabulary the transition was not so painful as it could be if starting from more personalized data structure.

We then used Jena, an open-source library from the Apache foundation to manipulate our graph data. This system was thought from the beginning to be multi-user, so in order to avoiding any duplication of information, we added a step into our process: hashing our triple. With this we still are working with our n-quads system, but this time, the 4 components added to the triples is not the source of the data, but a unique ID. This system has several advantages:

- Possibility to identify easily and uniquely each triplet
- No duplication (if they have the same content, they will have the same hash)
- Stronger possibility of enriching the metadata: we now can directly link a triple to more metadata (source, date added, etc.)

ID	Subject	Predicate	Object
rsd:42b	rsd:PicMidi	a	CRM:E71-MMO
rsd:b1f	rsd:42b	dc:created	2019-03-20
rsd:8bc	rsd:42b	dc:references	rsd:142
rsd:f56	rsd:42b	dc:creator	rdata:admin

Table 2. Example of a triple with a few metadata

We also developed an interface to help us search and add more triple along the way : we have a graph data manipulation system, that helps us structure our triples, helping with auto-completion and linked open vocabularies propositions. By searching through the already existing ontologies, we fell upon the 3Dontology that enabled us to integrate even more our data.

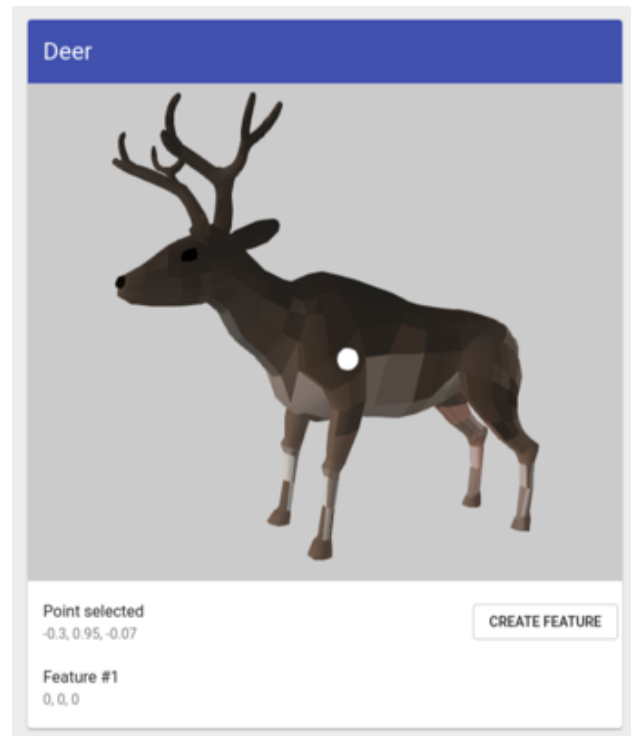


Figure 7. Prompting and point selection of a X3D data from our database

3.3 3D data in our database

A big gap remaining in our data integration is the decoupling of the geometrical shapes and the more text-based data. The 3Dontology was a complete rdf implementation of the X3D file format : it was the a perfect vocabulary in order to improve our heterogeneous integration. Adding a triple containing a simple geometry (we made our first tests with a X3D light shape of a Deer that was openly accessible as an example by the X3D community), we were able to :

- prompt the 3D data and color from our database interface
- basically manipulate the data (rotation, translation, zoom)
- select specific points on the geometry to which we could link a comment or metadata (point selection + metadata specification led to the creation of a new triple which subject was the specific point)

A trivial example of this operation can be seen in 7. X3D file format and the 3Dontology seemed to us being of high interest because of their ability to handle :

- faceted geometries
- point clouds
- several solid bodies and their mechanical links

Unfortunately, since the end of 2018, the 3Dontology website is unreachable and we have not been able to identify trends for analogue approaches that could perpetuate such approach.

4. NEXT STEPS

To sum things up, we now have a database management system dealing with graph data, that prompts and let us manipulate indiscriminately 3D data, their metadata or more conceptual modeling of our heritage site. The discontinuity of our 3D vocabulary is a heavy problem that we have to tackle as an urgent priority. The first tests that we performed convinced us that such an integrated approach makes sense. We now need to rely on community of practitioners and ontologists that share our views in order to have a stable data management system in the long run.

Taking a step back, it appeared to us that our efforts until now led to a proper data management concept, but has not really been of much help for building of our analysis. Our incremental approach made us refine our geometries and classification, but nothing huge appeared in our heritage analysis. Maybe such systems can be of better help in the processes of comparisons. With the case study of the Series of Meridian Telescope, we intend to test this hypothesis.

Another next step that we want to take on the *Observatoire du Pic du Midi* is altering the interface of our database, in order to simultaneously be able to prompt regular text triples, 3D data and automatic produced time lines, in order for our system to become a richer way to explore and analyze data.

5. CONCLUSION

This short paper contains a summary of our technical evolution in the Reseed project. The *Observatoire du Pic du Midi* is a wonderful case study, by its overall dimensions, its complex granularity, it's quite long history and its structural complexity. We will be trying to update the CIPA community at a next step of our incremental process and hope to be able to make a demo of our systems in the next years.

ACKNOWLEDGEMENTS

Many thanks to all the Reseed team members, especially Martin Boutroux, Alexandre Trendel and Gregoire Dupré Latour.

REFERENCES

- Ali, M. I. O., Laroche, F., Remy, S., Bernard, A., 2015. Structuration de la connaissance dans le cadre de la rétro-conception de systèmes mécaniques. *Colloque de l'AIP PRIMECA -Avril 2015 -La Plagne*.
- Cotte, M., Quantin, M., Jeanson, L., Bourgeois, N., Laroche, F., December 2017. Analysis of site integrity by 3D models and integrated database, Case Study: the Pic-du-Midi high-mountain observatory (France). *Symposium of ICOMOS General Assembly*.
- Davoust, E., 2000. *L'observatoire du Pic du Midi: cent ans de vie et de science en haute montagne*. CNRS.
- Laroche, F., Bernard, A., Cotte, M., 2008. Advanced industrial archaeology: a new reverse-engineering process for contextualising and digitising ancient technical objects. *Virtual and Physical Prototyping*, 3.2, 105–122.

Leon, D., 2018. Pointcloudexporter. [https://github.com/leon196/PointCloudExporte\(1 June 2019\)](https://github.com/leon196/PointCloudExporte(1 June 2019)), under GNU General Public License.

Macher, H., 2017. From point cloud to building information model (BIM) : 3D semi-automatic reconstruction of existing buildings. Theses, Université de Strasbourg.

2019