The Future of 3D Point Clouds: a new perspective

Discrete spatial datasets known as point clouds often lay the groundwork for decision-making applications. But can they become the next big thing?

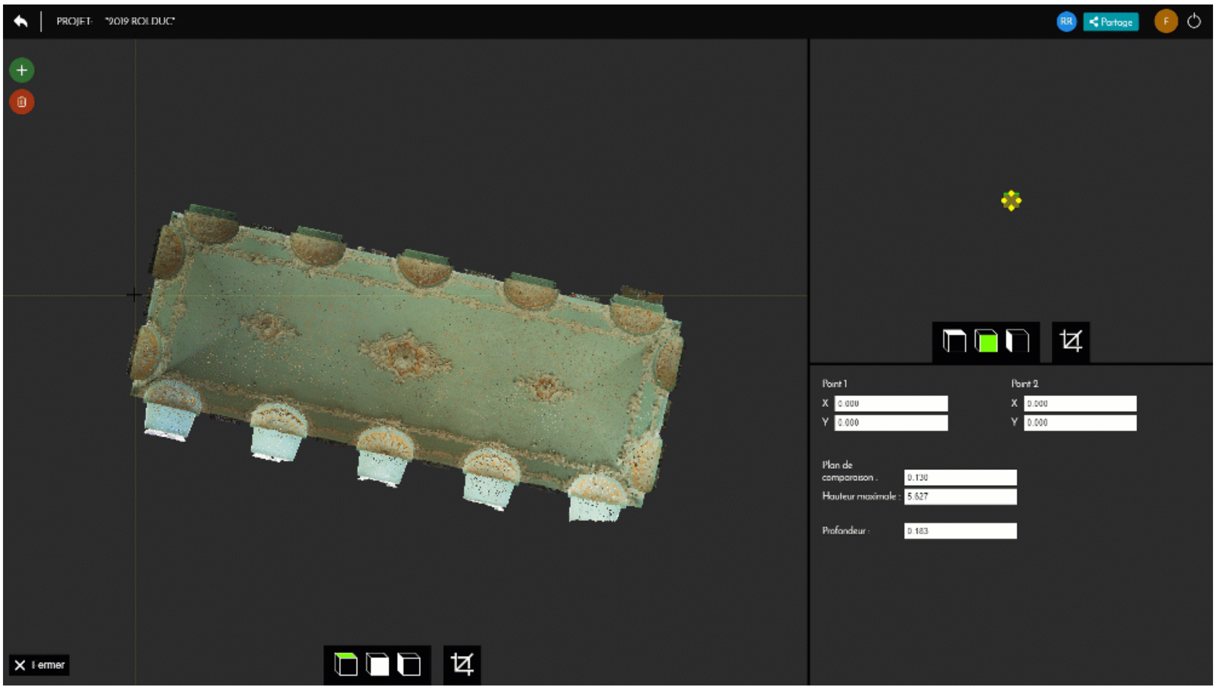
Different renderings of a point cloud. From left to right, raw point cloud, shaded, colored, voxelized, semantized

I am a big point cloud enthusiast. I first discovered their existence 10 years ago, and since then, I have been tweaking my practices through the evolution of Reality Capture to always get sharper datasets. But I still remember my first surveys with terrestrial laser scanners, and quickly getting these amazing (and still amazing) 3D point clouds.

In the process of 3D scanning an abandonned wool washing facility. © Photo [R. Robroek]

Published in Towards Data Science

But then… the dream is confronted to reality. How does one effectively consider these entities? At that time, the processing—read manual overloaded repetitive digitization—was composed of several heavily manual steps such as filtering, registration, cleaning, segmenting, classifying, meshing, digitizing … It evolved for some parts (mainly registration, filtering and meshing) but the main bottleneck that I had back then is still unresolved : why do we bother changing the nature of the data (E.g. point cloud to vector) per application ?



Manual digitization process to create a dwg file within [Flyvast](https://www.flyvast.com) online point cloud software.

Is there not a more efficient workflow ? Let me bring you on a research journey to materialize thoughts in solutions.

The Genese

Back in 2015, after 2 years as a 3D Laser scanning engineer, I decided to dedicate myself to teaching & research to try and solve this issue. I jumped into Academia and started investigating the current state of developments, looking for bricks that eventually need some mortar. Well, at that time, I quickly realized that no working attempt addressed the root of the problem. And that my endeavor would need a bit more than some hours.

The Observation

"when we open our eyes on a familiar scene, we form an immediate impression of recognizable objects, organized coherently in a spatial framework".

In 1980, Treisman defines in simple terms the complex mechanism behind our human sight-perception. For non-impaired human-being, it is often the primary source of information which our cognitive decision system can use to act on. This is extendable using our brain which quickly adapts to new surroundings and only uses the most important material captured though our eyes. In fact, the brain receives just three "images" every second, which are sorted and combined with prior knowledge to create the reality that we experience.

Does this image even make sense? I am sure you will find a meaning to these.

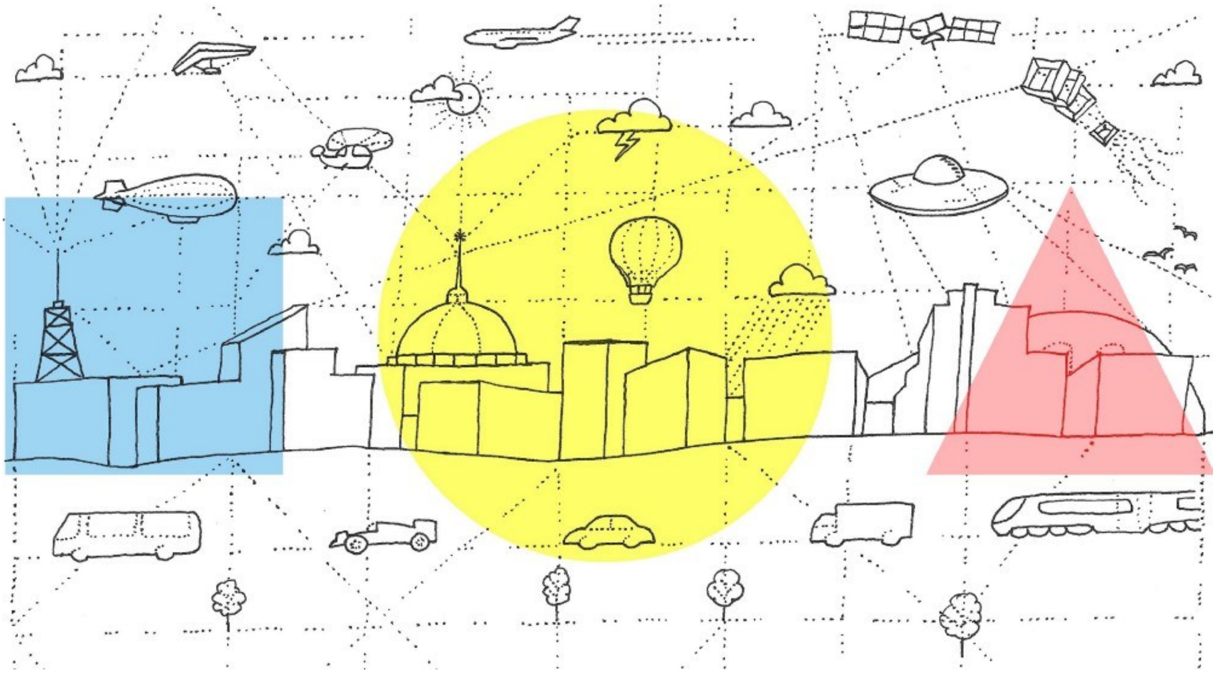
This mechanism is exceptionally fast and efficient allowing to brake when we see a red light, or simply to read this article and understand the spatial organization of words. Even more impressive, our vision can be adapted for an "orientation attention"—energy saving mode where the brain does not develop a full understanding of the surroundings—or a "discover attention"—which runs slower as the brain collects data from our memory to obtain a full understanding of the scene.

With today's computational power and high level of dematerialization, virtually replicating such a process is not only very attractive but seems feasible. While the operation is genuinely hard to mimic, studying how we interact with our environment permits to better grasp the boundaries and usable mechanisms.

Published in Towards Data Science

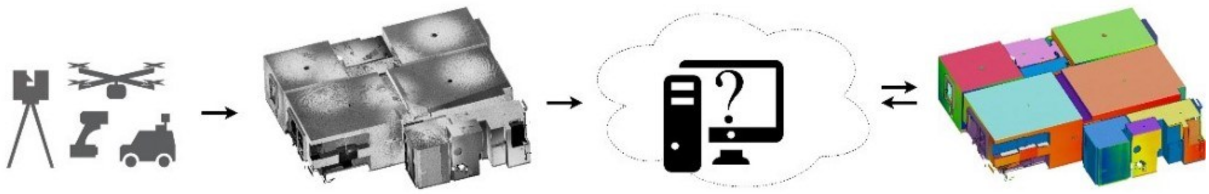
The comparison

It first translates into the use of sensors that can capture key inputs usable by a computer.



Each vector in this image is guided by sensors (artificial or natural) that gather key insights for their usage.

We then aim at a procedure based on gathered data and accessible information repositories to produce a “semantic representation”: a depiction of a scene integrating concepts and their meaning. In such a scenario, a spatial sensor plays the role of our eyes to obtain a digital spatial asset further refined into a semantic representation using available knowledge.



The sensor plays the role of our eyes, the spatial framework becomes a semantic representation, and the scene is tagged familiar using available knowledge

This availability is often a first complication. Our online cognitive perception uses our memory and is structured to access needed evidence in a very short time. Mirroring this stage using a computer is extremely complex and aiming for a solution as generalist as possible is an important challenge.

The second bottleneck when trying to virtualize a cognitive decision system is the creation of a semantic representation as in the figure below. Gathering and attaching domain knowledge to underlying spatial data is linked to colossal integration and mining complications regarding data types, sources or representations.

3D point cloud representation vs 3D semantic representation

The Data

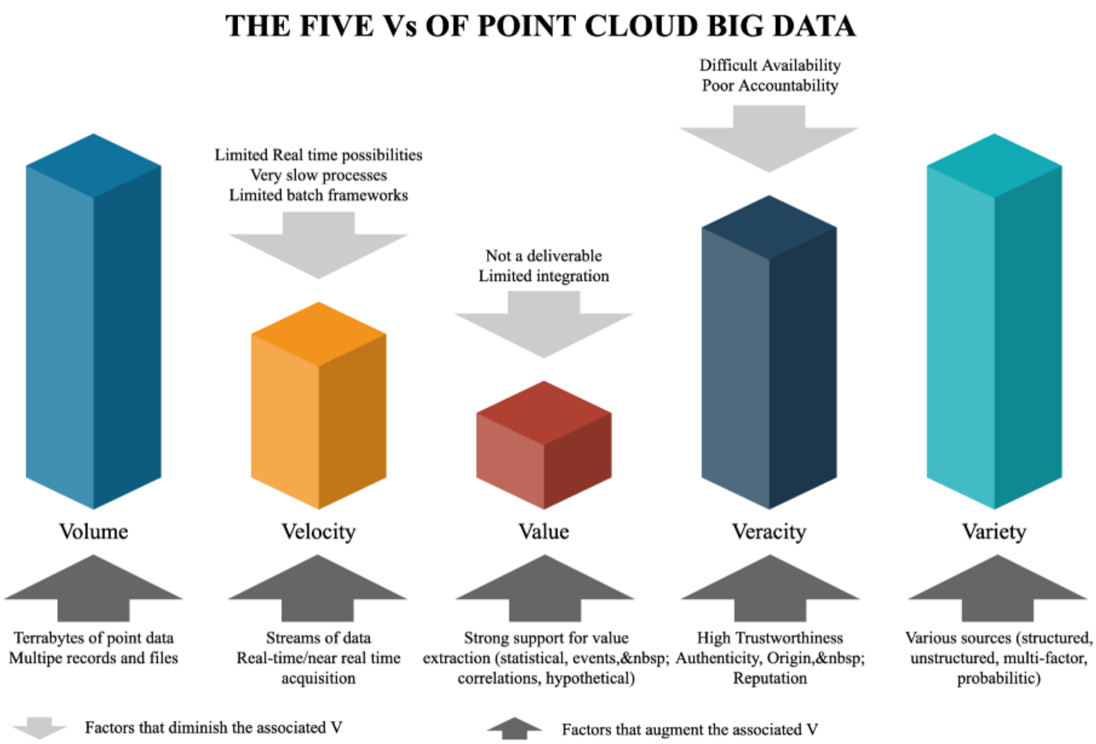
3D Point Clouds

The main challenge revolves around the specificity of the data collected by the sensor(s). Single raster images or video streams are great when depth cues are not necessary, but emulating our 3D visual cognition demands a richer data basis. Reality Capture devices permit to obtain such an exhaustive 3D spatial information primarily as a point cloud: a {X, Y, Z} (+ attributes) spatial ensemble which digitally represents the recorded environment w.r.t the sensor strengths and limitations. The landscape of these instruments and acquisition methodologies is mature enough to allow digital replicas of the real world ranging from the object scale to the country scale as illustrated below.

Real-time Multi-scale point cloud of different datasets captured and combined

Point Cloud Big Data

The acquisition of these so-called point clouds has become easier, faster and is even accessible from very low-cost solutions. All these hardware evolution were unfortunately not followed by their software counterpart, which are heavily impacted by the 5 V's of Big Data problematics as illustrated below.

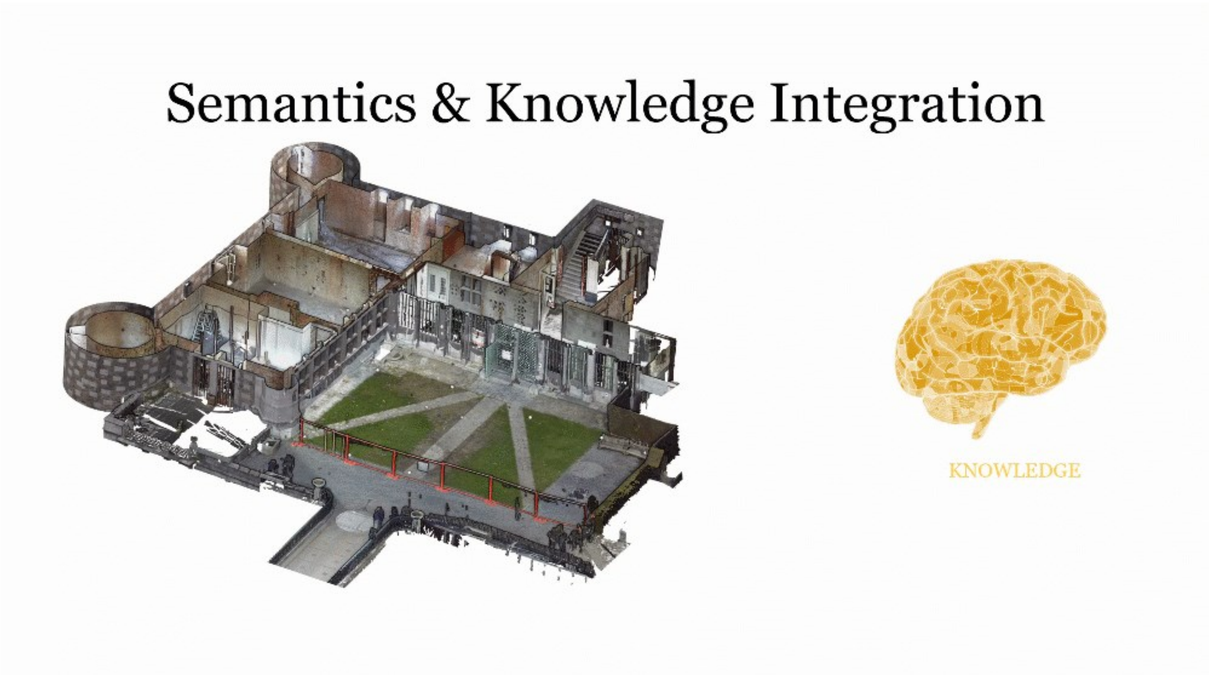


The Five Vs of Big Data in the context of point clouds.

Connecting numerous sensors/approaches creates heterogeneous point cloud datasets (Variety) and participate in the constitution of massive data repositories (Volume). In turn, it reduces the processing efficiency (Velocity) and creates new needs to turn huge amounts of point data into trustworthy (Veracity) and actionable information (Value).

The Deliverables

Point cloud acquisition and processing workflows are usually application-dependent following a classic progression from data gathering to deliverable creation. While the collection step may be specific to the sensor at hands, point-cloud-as-a-deliverable upsurges, becoming one de-facto choice for many industries. This task-oriented scenario mainly considers these as a spatial reference—which is used by experts to create other deliverables—thus being a project's closest link to reality. It brings accurate real-world information which could allow decision-making based on digital-reality instead of interpreted or not up-to-date information.



Today, the "brain" is an expert behind a desk that will process the point cloud to extract deliverables. What we want is to integrate this knowledge directly within the data to give a semantic meaning to spatial entities

Moreover, the procedures to convert point clouds in application-specific deliverables are very costly in time/manual intervention. It is getting ever more complicated for the human expertise to handle adequately the large and complex volumes of information, often contradictory disseminated among different actors/supports of one project. Thus, it is key for a sustainable system that big point cloud data translates into more efficient processes opening a new generation of services that help decision-making and information extraction.

We need to find ways for massive automation and structuration to avoid task-specific manual processing and non-sustainable collaboration.

The Collaboration

As humans, we thrive on massive collaboration. Our greatest achievements are often building on a efficient exchange of information, services and more. Point clouds are often very large depending on how much data is collected—usually in the realms of Gigabytes, if not Terabytes—and are usually destined to be archived as a reusable support to create new type of data and products. This can lead to a dead-end with exponential storage needs, incompatibility between outputs, loss of information and complicated collaboration.

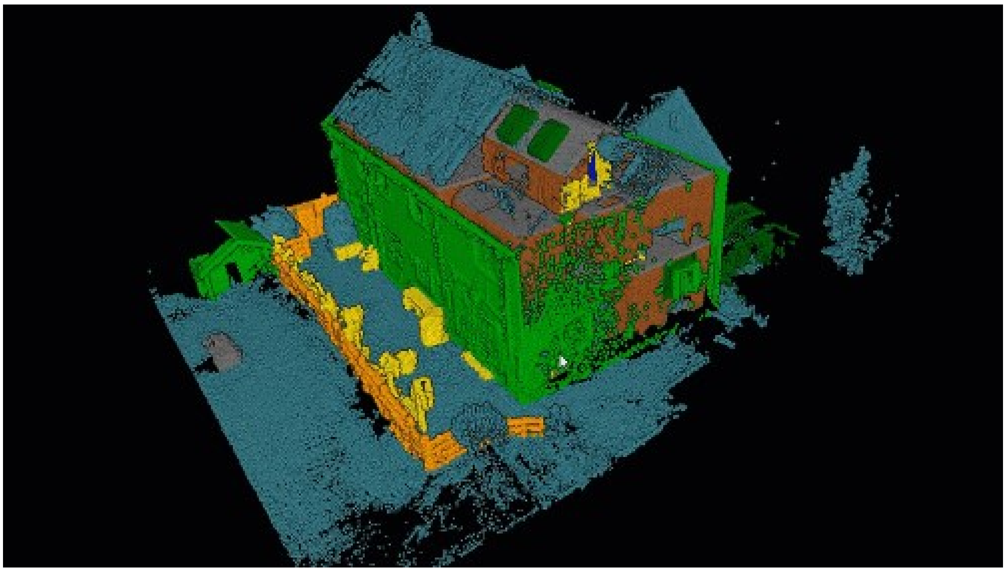
These practices also show limited to no attempt to generalize a framework which could in turn play as a common ground for further interoperability and generalization. This lack is counterproductive and could lead in term to a chaotic data repartition among actors and worsen the dependency to several outsourced service each aiming an application independently. This emphasize a strong need to study interoperable scenarios in which one point cloud could be used by many users from different domains, each having a different need.

This will in turn introduce new constraints at the acquisition level to define the needed exhaustivity of the 3D representation for use with reasoning engines. Of course, this serialize additional challenges for interconnecting processes and insuring a compatibility with the different sources, volumes and other data-driven parameters.

The automation

In this continuum, the reflexion to go from a human-centered process to an autonomous workflow orient research to develop automation and AI to speed-up inference processes. This is crucial to the development of point clouds in 3D capture workflows, where objects need to be identified.

Robotics research has made a leap forward providing autonomous 3D recording systems, where we obtain a 3D point cloud of environments with no human intervention. Of course, following this idea to develop autonomous surveying means demand that the data can be used for decision-making. The collected point cloud without context does not permit to take a valid decision, and the knowledge of experts is needed to extract the necessary information and to creates a viable data support for decision-making. Automating this process for fully autonomous cognitive decision systems is very tempting but poses many challenges mainly link to Knowledge Extraction, Knowledge Integration and Knowledge Representation from point cloud. Therefore, point cloud structuration must be specifically designed to allow the computer to use it as a base for information extraction, using reasoning and agent-based systems.



Result of my AI-powered automatic object's recognition in an unsupervised fashion. Each color represent a different class recognized automatically

The identification

And this is were I want to get to: we need intelligence within our virtual datasets ! This, in order to avoid brain work and manual processes, but also for the sake of interoperability. There are many applications that will use point clouds differently, but extracting deliverables per application doesn't seem to be the most efficient (and eco-friendly in terms of storage footprint). However, if experts knowledge is formalized and integrated within point clouds, you can only guess how centralized and efficient an infrastructure becomes !

Published in Towards Data Science

The outcome

So yes, point clouds are huge; yes we need specific “tricks” to store them and process them, but so were videos back some decades ago ! What does this imply in your specific industry? That you will soon be able to work with a “brain representation” of the 3D captured environment for you to query as you deem. But of course, the big landscape in 3D sensors make the recognition process a wide research exploration field! Exciting!

Conclusion

These thoughts are based on the award-winning thesis “The Smart Point Cloud”, that holds the more technical details about a working solution to these problematic. But of course, as it is often the case with long research works, you come out of it with more questions that you had in the beginning.

What an exciting decade we are experiencing. We are in a fascinating era where the usage of machine learning gives tremendous possibilities to solving challenges that were considered Science Fiction 10 years ago. The 5 key-points takeaways are the following:

- 3D Point Clouds is close to unrefined oil coming from sensors
- Semantic injection should aim at providing a large domain connectivity
- The underlying data structures and algorithms are tailored to this end
- Interoperabily, modularity and efficiency is key for collaboration
- Efficient automated object detection should build on these goals

Published in Towards Data Science

References

1. Poux, F. The Smart Point Cloud: Structuring 3D intelligent point data, Liège, 2019.
2. Poux, F.; Valembois, Q.; Mattes, C.; Kobbelt, L.; Billen, R. Initial User-Centered Design of a Virtual Reality Heritage System: Applications for Digital Tourism. *Remote Sens.* **2020**, *12*, 2583, doi:10.3390/rs12162583.
3. Poux, F.; Neuville, R.; Nys, G.-A.; Billen, R. 3D Point Cloud Semantic Modelling: Integrated Framework for Indoor Spaces and Furniture. *Remote Sens.* **2018**, *10*, 1412, doi:10.3390/rs10091412.
4. Billen, R.; Jonlet, B.; Luczfalvy Jancsó, A.; Neuville, R.; Nys, G.-A.; Poux, F.; Van Ruymbeke, M.; Piavaux, M.; Hallot, P. La transition numérique dans le domaine du patrimoine bâti: un retour d'expériences. *Bull. la Comm. R. des Monum. sites Fouill. 30* **2018**, 119–148.
5. Poux, F.; Billen, R. Voxel-based 3D Point Cloud Semantic Segmentation: Unsupervised geometric and relationship featuring vs deep learning methods. *ISPRS Int. J. Geo-Information* **2019**, *8*, doi:10.3390/ijgi8050213.
6. Kharroubi, A.; Hajji, R.; Billen, R.; Poux, F. Classification And Integration Of Massive 3d Points Clouds In A Virtual Reality (VR) Environment. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *42*, 165–171, doi:10.5194/isprs-archives-XLII-2-W17-165-2019.
7. Bassier, M.; Vergauwen, M.; Poux, F. Point Cloud vs. Mesh Features for Building Interior Classification. *Remote Sens.* **2020**, *12*, 2224, doi:10.3390/rs12142224.
8. Poux, F.; Ponciano, J. J. Self-Learning Ontology For Instance Segmentation Of 3d Indoor Point Cloud. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*; ISPRS, Ed.; Copernicus Publications: Nice, 2020; Vol. XLIII, pp. 309–316.
9. Poux, F.; Mattes, C.; Kobbelt, L. Unsupervised segmentation of indoor 3D point cloud: application to object-based classification. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*; 2020; Vol. XLIV–4, pp. 111–118.
10. Poux, F.; Billen, R.; Kaspryzk, J.-P.; Lefebvre, P.-H.; Hallot, P. A

Built Heritage Information System Based on Point Cloud Data: HIS-PC. *ISPRS Int. J. Geo-Information* **2020**, *9*, 588, doi:10.3390/ijgi9100588.

11. Poux, F.; Billen, R. A Smart Point Cloud Infrastructure for intelligent environments. In *Laser scanning: an emerging technology in structural engineering*; Lindenbergh, R., Belen, R., Eds.; ISPRS Book Series; Taylor & Francis Group/CRC Press: London, United States, 2019; pp. 127–149 ISBN in generation.

12. Tabkha, A.; Hajji, R.; Billen, R.; Poux, F. Semantic Enrichment Of Point Cloud By Automatic Extraction And Enhancement Of 360° Panoramas. *ISPRS - Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *XLII-2/W17*, 355–362, doi:10.5194/isprs-archives-XLII-2-W17-355-2019.

13. Poux, F.; Neuville, R.; Hallot, P.; Van Wersch, L.; Jancsó, A. L.; Billen, R. Digital investigations of an archaeological smart point cloud: A real time web-based platform to manage the visualisation of semantical queries. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch.* **2017**, *XLII-5/W1*, 581–588, doi:10.5194/isprs-Archives-XLII-5-W1-581-2017.

14. Poux, F.; Hallot, P.; Jonlet, B.; Carre, C.; Billen, R. Segmentation semi-automatique pour le traitement de données 3D denses: application au patrimoine architectural. *XYZ* **2014**, *141*, 69–75.

15. Novel, C.; Keriven, R.; Poux, F.; Graindorge, P. Comparing Aerial Photogrammetry and 3D Laser Scanning Methods for Creating 3D Models of Complex Objects. In *Capturing Reality Forum*; Bentley Systems: Salzburg, 2015; p. 15.