**Thesis offer**
**Construction 4.0 platform using Process Mining to digitize, describe, visualize and improve the buildings' construction process.**


**Scientific supervision**:  Exposito Ernesto (LIUPPA),  Christian Laborderie (SIAME)
**Contract type:** contract over a period of 36 months as a PhD student
**Monthly remuneration**: according to the current national legislation in France
**Start of the contract**: September 2023


### Résumé du projet

Conception d'une plateforme Construction 4.0 utilisant le Process Mining pour numériser, décrire, visualiser et améliorer le processus de construction des bâtiments. La plateforme propose une approche générique qui peut être mise en œuvre sur n'importe quel site de construction et qui commence par l'identification des phases de construction, l'installation de capteurs sur le site de construction, la collecte de données, le stockage des données, la transformation des données (construction de journaux d'événements), l'analyse (application du Process Mining) et la visualisation des données sur la construction, leurs contextes et leurs processus (méthode appliquée) ; afin de découvrir et de décrire le processus de construction des bâtiments (en générant un modèle théorique à partir du processus réel de construction des bâtiments), et de permettre aux parties prenantes d'effectuer un diagnostic mettant en œuvre la vérification de la conformité entre la réalité (données du processus de construction des bâtiments) et le modèle de processus précédemment découvert, et enfin de permettre d'améliorer ces processus de construction en tenant compte des différents éléments qui peuvent être analysés en utilisant l'exploration de processus (en utilisant les connaissances acquises lors de la phase de découverte et de conformité pour améliorer le processus).

### Project Summary

Design of a Construction 4.0 Platform using Process Mining to digitalize, describe, visualize and improve the buildings' construction process. The platform proposes a generic approach to be implemented in any construction site, and that starts from the identification of the construction phases, the installation of captors/sensors in the construction site, the collection of data, storage of data, transformation of data (construction of events logs), analysis (application of process mining) and visualization of data about construction, their contexts, and processes (method applied); in order to discover and describe the buildings' construction process (generating a theoretical model from the real building construction process), and

allow stakeholders to perform diagnosis implementing conformance checking between reality (buildings' construction process data) and the process model previously discovered, and finally allowing to enhance these construction processes considering the different elements that can be analyzed using process mining (using insights from the discovery and conformance phase to improve the process).

I.    **Project description**

**Context**

Building activities are as old as humans themselves, however, the construction of permanent shelters started once humans abandoned nomadic life; since then, buildings have become key parts of human lives, so much that approximately 90 percent of the time is spent indoors. The construction of these infrastructures has increased over time mostly at the same rate as the population does, however, the digitalization of it has not increased at the same rhythm; in fact, the building process generates a huge amount and diversity of data and events that are normally processed manually or even isolated using, for example, Building Information Modelling (BIM), technical documents, safety reports, data related to resources, equipment, construction materials, costs, etc. Unfortunately, on a large scale, the analysis of all these different data can be a complex task, which can lead to late identification of issues and bottlenecks, decrease in efficiency, compliance issues, and of course to slowly improve the building process.

For this reason, it is proposed the design of a Construction 4.0 Platform using Process Mining to digitalize, describe, visualize and improve the buildings' construction process. The platform proposes a generic approach to be implemented in any construction site, and that starts from the identification of the construction phases, the installation of captors/sensors in the construction site, the collection of data, storage of data, the transformation of data (construction of events logs), analysis (application of process mining) and visualization of data about construction, their contexts, and processes (method applied); in order to discover and describe the buildings' construction process (generating a theoretical model from the real buildings' construction process), and allow stakeholders to perform diagnosis implementing conformance checking between reality (buildings' construction process data) and the process model previously discover, and finally allowing to enhance these construction processes considering the different elements that can be analyzed using process mining (using insights from the discovery and conformance phase to improve the process).

The project is being carried out in close collaboration between the Université de Pau et des Pays de l'Adour (UPPA), NOBATEK/INEF4 which is a technological research center and construction companies. The computer science (LIUPPA) and engineering science (SIAME) laboratories are involved in supervising the thesis.

**Objectives**

This interdisciplinary project will merge Computer Science and Civil Engineering domains, to incorporate top down (semantic web) and bottom up (process mining) approaches to understand and guide construction life cycles, based on the following contributions :

1. **Construction 4.0 Framework**
- **Construction Phases identification:** this step represents one of the most important steps of the project, this will allow us to identify the different construction phases of a building with the idea of creating a generic model/framework, describing the nature of each of the phases identified as well as the needs, actors and other related processes.
- **Data collection and treatment regulations**: this phase will take into account the regulations of data collection and treatment that need to be taken into account for the construction 4.0 context.
- **Input/Sensors/captors identification**: this phase will be based on the identification of the ideal input/sensor/captor sources that would allow the digitalization of all of the construction phases identified. For this, it is necessary to describe the nature of the measurements obtained and the technical requirements for their deployment.

2. **Construction 4.0 Platform: Data Warehousing, BI/BA, and Process mining**
- **Data sources:** once the construction phases have been identified the ideal input data sources, including potential IoT sensors will be identified or deployed within the construction site, with the necessary connections and configurations, to extract the data of interest.
- **Collection and storage of data:** once the data sources are identified, this phase will manage a data warehouse, assuring the collection of the information as well as the storage in a common format of the data sources to allow at least the identification of the construction phase, the sensor and the event that occur (who, what, when and where).
- **Transformation of data:** Once the building records are collected and stored in a common format, it is possible to apply process mining to these data, however, it is necessary to apply some prior modifications to deliver them as needed. The process mining task starts with an Event Log and inside this structure, a process is described as follows, a process consists of cases, a case consists of events such that each event relates to precisely one case and each sequence of activities executed for a case is a trace. Each line in the event log presents one event. Events within a case are ordered. Events can have attributes (e.g., activity, time, cost, resource, etc)[2]. The event log structure can be summarized as follows:
    A. Case ID: indicates at which case or instance belongs to an event or activity. In this case, will be the construction phase identified.
    B. Activity: Action captured by the event.
    C. Timestamp: Indicate the time when the event took place.
- **Process discovery:** once the data have been transformed to the necessary structure 'event log', it will be possible to apply a process mining algorithm to "discover" the process. This technique takes an event log and produces a model without using any a-priori information. In the end, the discovery model will be "representative" of the behavior seen in the event log [2]. For this task, we are thinking of using an online existing solution that already implements process mining, open-source or otherwise. After this, it will be possible to visualize globally all the interactions made in the construction process, it will be visually described all the sequence of interactions from

the different phases of the construction process, discovering the most frequent one highlighting, for example, the average time spent between activities, and also the less frequent ones that may represent early problems in the buildings' construction process such as bottlenecks and decrease on efficiency.

- **Process conformance:** For this second type of analysis available, it is necessary that both the process model and an event log are available. Conformance checking relates events in the event log to activities in the process model (previous step) and compares both. The goal is to find commonalities and discrepancies between the modeled behavior and the observed behavior. Conformance checking is relevant for business alignment and auditing. For example, the event log can be replayed on top of the process model to find undesirable deviations suggesting fraud or inefficiencies.[2] In Fig. 1 It gets explained that the global conformance measures quantify the overall conformance of the model and log. Local diagnostics are given by highlighting the nodes in the model where model and log disagree.
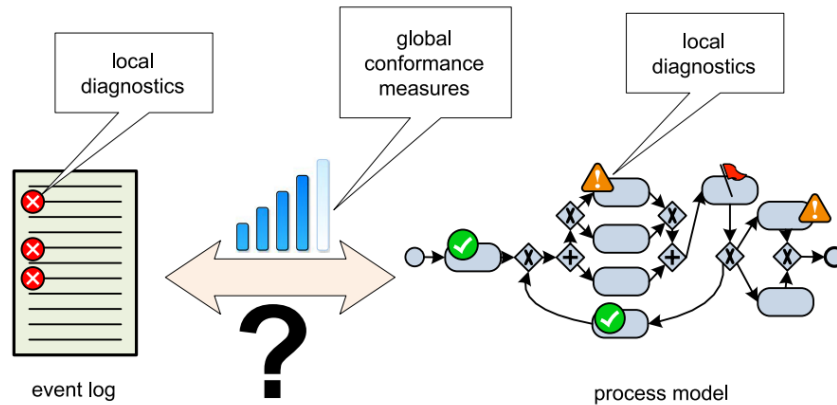


Figure 1: Conformance checking: comparing observed behavior with modeled behavior[2]

- **Continuous Process enhancement:** the main idea of digitalizing and discovering the buildings' construction process, is the possibility to improve it, and this is possible also applying process mining, in fact, if the process model discovered and then checked by the conformance checking step does not reflect the reality, it can be applied either some measures enhancement to the model. The first type of enhancement is repair (modifying the model to better reflect reality), for example, if two activities are modeled sequentially but in reality, can happen in any order, then the model may be corrected to reflect this. Another type of enhancement is an extension (adding a new perspective to the process model by cross-correlating it with the log). An example is the extension of a process model with performance data, with information about resources, decision rules, quality metrics, etc [2].
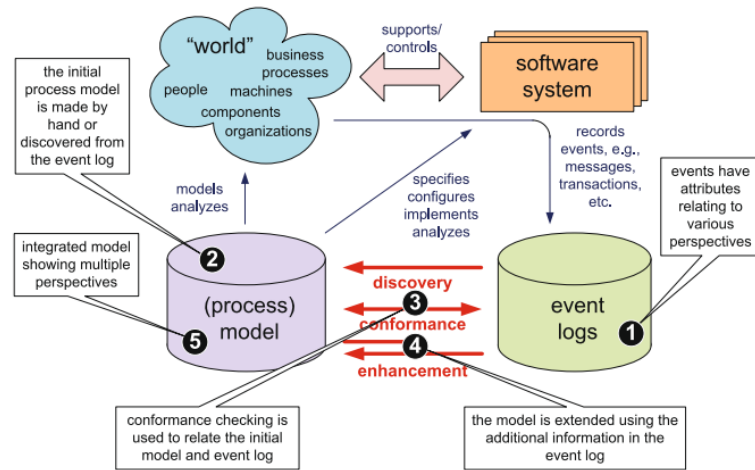
Figure 2: The organizational, case, and time perspectives can be added to the original control-flow model using attributes from the event log[2]

**State of the art**

We are living in a time of huge digital transformations that are having an impact on all types of organizations and construction is not the exception, in fact, "construction 4.0" have been a field of study that has increased since it was first mentioned in Germany on 2016 as exposed by a literary review carried out in 2020 by the Department of Civil and Environmental Engineering, Universidad del Bío-Bío, Chile[3], this review also explains that the definition of Construction 4.0 derives from the foundation of Industry 4.0 but focuses on and relates to the construction sector. Therefore, Construction 4.0 is all technological changes related to the implementation of new work methods which are related to processes, materials, and markets. Hereafter Fig. 3 describes a diagram to represent the concept of "Construction 4.0".
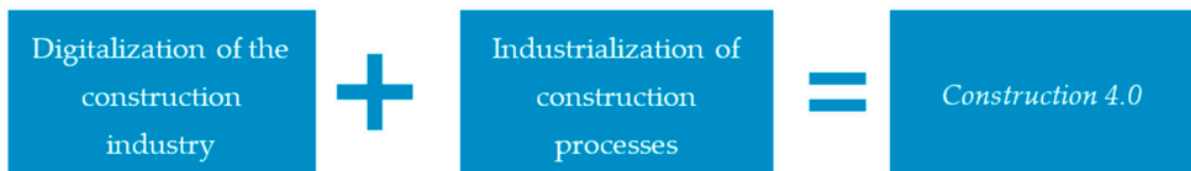


Figure 3: Diagram of "Construction 4.0" pillars

A study carried out in 2019 in the Dept. of Civil and Environmental Engineering, Nazarbayev University, Astana, Kazakhstan [4] proposed a framework (Fig. 4) that incorporates current technological advancement related to the construction industry, legislative requirements, barriers, enterprise transformation requirements and so on. They stated that the field of Construction 4.0 would make a great impact in the construction industry through an improved value chain of construction projects, productivity improvement, and safe and sustainable construction.
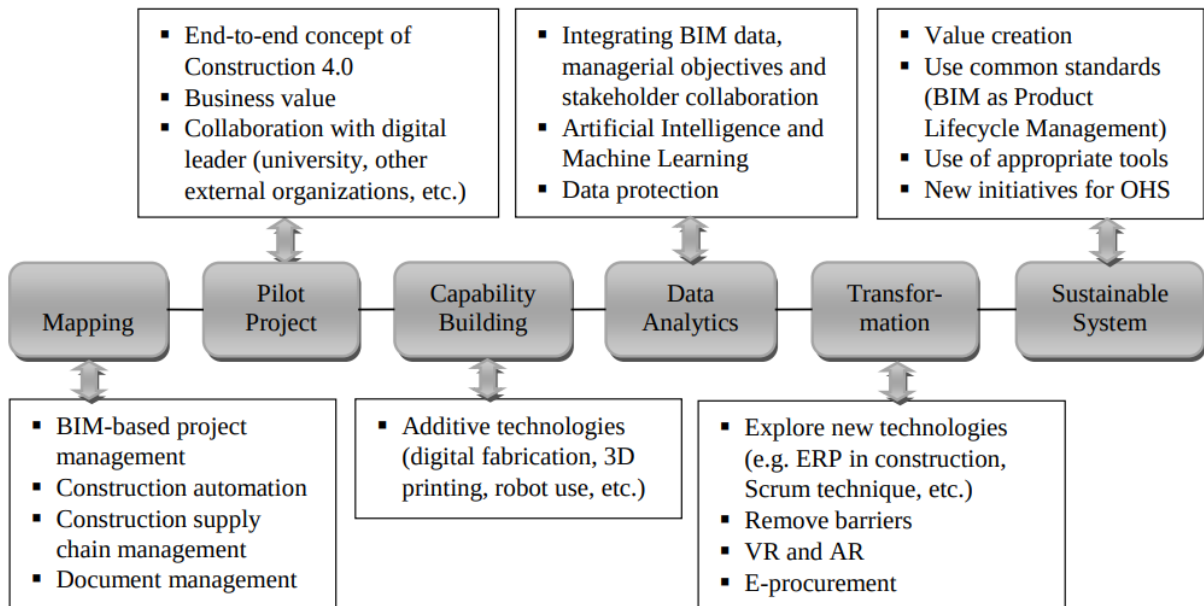
Figure 4: Framework to implement the concept of Construction 4.0

Another study performed in 2020 at the International Center for Numerical Methods in Engineering (CIMNE), Barcelona, Spain [5] proposes a methodological-technological Framework for Construction 4.0, which the main objective is attending to the following three scenarios (1) Physical Domain/Automation, (2) Simulations and Modelling and (3) Digitalisation and Virtualisation. Scenario 1 refers to the digital end-to-end engineering integration, considering technologies for the automation of the physical manufacturing environment, whereas the second one includes modeling and simulation tools for the design, construction, and operation of buildings and infrastructures. Finally, the third scenario refers to digitalization and virtualization tools for industrial processes and products, to link domains 1 and 2 [5]
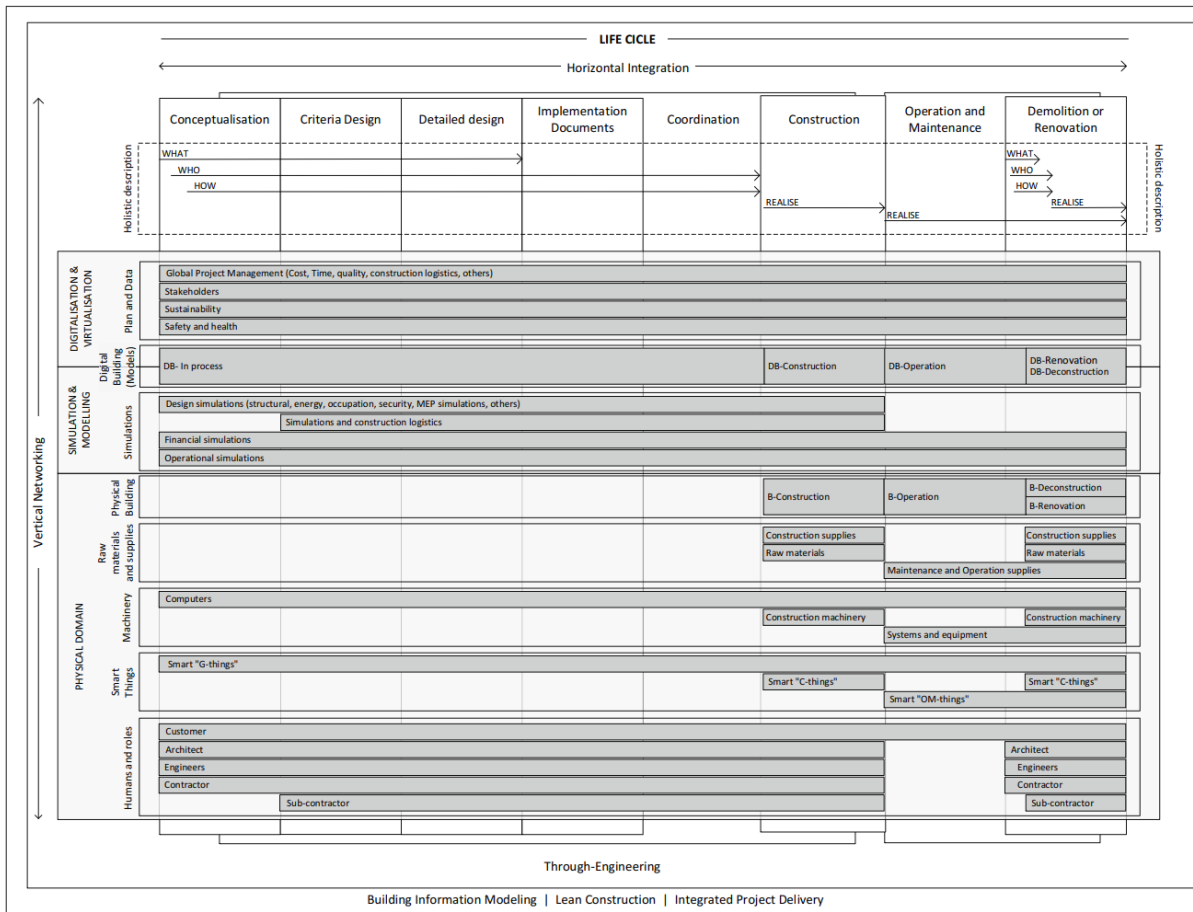
**LIFE CICLE**

Figure 5:  Construction 4.0 methodological framework [5]

## II.    **Methodology and research activities planning**

The project will be carried out following an agile approach called Rational Unified Process (RUP) which is an agile Software Engineering Process, and has four phases: inception, elaboration, construction, and transition. Each phase has one or more iterations and is completed with a milestone. At that point, the project progress is assessed, and major decisions are made based on that progress. The focus of the iterations in each phase is to produce technical deliverables that will fulfill the phase objectives [6]. The phases of the project are described in table 1 and the expected deliverables associated to these activities are listed in Table 2.

**Table 1: PhD Plan**

| Phase 0: State of the art | |
|---|---|
| M1 - M6 | Develop a detailed state of the art, regarding the construction phases, describing the nature of it, as well as the physical and logical resources, such as actors, materials, and other processes that need digitalization. |

| Phase 1: Inception | |
|---|---|
| M6 - M12 | Positioning of work, based on related works and existing potential solutions, well adapted to the identified constraints |
| | Solution proposal specification and design of use case scenarios |

| Phase 2: Elaboration: Construction 4.0 Prototype | |
|---|---|
| M12 - M18 | **Construction 4.0 Framework:** Work collaboratively between the civil engineering and computer sciences research groups to map the digitalization of the construction phases with the help of the ideal sensor/captor previously identified, by developing a generic framework that can be implemented in any construction site. |
| | **Deployment of captors/sensors basic sensors:** as part of the prototype it will be deployed some sensors in the construction site with the necessary connections and configurations, to extract the data of interest. |
| | **Data warehouse:** deployment of a data warehouse prototype platform that allows the storage of the data coming from the different sensors/captors deployed. |
| | **Data treatment:** manual cleaning and transformation of the sensors/captors data into a new structure called event logs (case id, activity, and timestamp) |
| | **Data Analysis:** as part of the prototype it will be tested different process mining platforms open source or otherwise to compare the results of the analysis performed such as process discovery, process conformance, and process enhancement in order to set the basis for the final phase of the project. |

| Phase 3: Construction: Construction 4.0 Platform | |
|---|---|
| | **Deployment of captors/sensors:** in the construction site, once the construction phases have been identified the ideal sensors /captors will be deployed, depending on the required |

| | | |
|---|---|---|
| M18 - M24 | nature of the measure, with the necessary connections and configurations, to extract the data of interest. |
| | **Data warehousing:** deployment of a data warehouse platform that allows the storage of the data coming from the different sensors/captors and assures a common format for the buildings' construction phases events. |
| | **Data treatment:** automatic cleaning and transformation of the sensors/captors data into a new structure called event logs (case id, activity, and timestamp) |
| | **Data Analysis:** configurations and connections of the process mining platform chosen (platform with the best results in the previous phase) on the Construction 4.0 platform. Integration of automatic forwarding of events logs of the construction process, as well as the necessary configurations to deploy all the analysis available such as process discovery, process conformance, and process enhancement. |
| **Phase 4: Deployment and Evaluation : Construction 4.0 Platform** | | |
| M24 - M36 | **Deployment** of the solution in a real case environment<br>**Collection and Evaluation** of results<br>**Manuscript** elaboration<br>**Publications and dissemination** |

**Table 2: List of deliverables**

| | Deliverables | Format | Target audience | Date |
|---|---|---|---|---|
| | **D1: Inception** | | | |
| **1** | **State of the art** | Report | Internal | Y1 |
| | State of the art (Survey articles) | 1 journal<br>1 Conference | International | Y1 |
| | **D2: Elaboration** | | | |
| 2 | Construction 4.0 Framework | Architectural design | Internal | Y2 |
| | Construction 4.0 Platform Prototype | Cyber-Physical Platform specification | Internal | Y2 |
| | Positioning Article | 1 journal<br>1 Conference | International | Y2 |

| | D3: Construction, transition. evaluation | | | |
|---|---|---|---|---|
| 3 | Construction 4.0 Platform Implementation, Deployment and evaluation | 1 o 2 journals 1 Conference | International | Y3 |
| | PhD manuscript | Dissertation | International | Y3 |

The thesis is proposed in the frame of the DATABUILDER project which is a research collaborative project between SIAME Laboratory, LIUPPA and NOBATEK/INEF4. The project is also supported by two large construction companies ETCHART Construction and EIFFAGE Construction, deeply involved in the project by providing concrete use cases and real construction project data.

**Qualifications**

Candidates should hold a Master degree in Computer Science, Information technology, Data Science or similar engineering field.

We are looking for a candidate who has knowledge in the field of Ontology and semantic web technologies. The candidate must master the English language with proficiency, and basic knowledge of the French language. Finally, the candidate should show a positive attitude with a flexible approach to work. He should be able to communicate across multiple teams and should be curious and autonomous. Travels in France or abroad for meetings or conferences can be expected.

**Working environment of the PhD student**

The PhD student will be located in the LIUPPA laboratory and will share his time between the LIUPPA and SIAME and will work in close relation with NOBATEK/INEF4.

**How to apply**

Application including motivation letter, CV, recommendations and master degree official transcript should be sent to Ernesto EXPOSITO: ernesto.exposito@univ-pau.fr and Christian LA BORDERIE: christian.laborderie@univ-pau.fr.

**References**

[1] "Process Mining and Execution Management Software," *Celonis*. [Online]. Available: https://www.celonis.com/. [Accessed: Jul. 15, 2021]

[2] W. van der Aalst, *Process Mining: Discovery, Conformance and Enhancement of Business Processes*. Berlin Heidelberg: Springer-Verlag, 2011 [Online]. Available: https://www.springer.com/gp/book/9783642193453. [Accessed: Jan. 25, 2021]

[3] E. Forcael, I. Ferrari, A. Opazo-Vega, and J. A. Pulido-Arcas, "Construction 4.0: A Literature Review," *Sustainability*, vol. 12, no. 22, p. 9755, Jan. 2020, doi: 10.3390/su12229755.

[4] M. Hossan and A. Nadeem, *TOWARDS DIGITIZING THE CONSTRUCTION INDUSTRY: STATE OF THE ART OF CONSTRUCTION 4.0*, vol. 6. 2019.

[5] F. Muñoz-La Rivera, J. Mora-Serrano, I. Valero, and E. Oñate, "Methodological-Technological Framework for Construction 4.0," *Arch. Comput. Methods Eng.*, vol. 28, no. 2, pp. 689–711, Mar. 2021, doi: 10.1007/s11831-020-09455-9.

[6] A. Anwar, "Rational Unified Process," *Int. J. Softw. Eng.*, p. 17, 2014.