

Sujet de thèse : Formalisation semi-automatique d'un référentiel de connaissances support de jumeau numérique pour la rénovation de bâtiments : application au projet ISOBIM

Contexte : le projet ANR ISOBIM vise à proposer un processus collaboratif pour la rénovation par ISOLation extérieure basé sur les paradigmes Lean et BIM. Dans ce cadre, un des objectifs est de pouvoir assurer la gestion des données du projet dans un référentiel formel capable de couvrir les phases de conception, calepinage, planification et suivi de réalisation. En outre, ce référentiel devra supporter le jumeau numérique de bâtiment permettant, en particulier, la simulation 4D dans la phase d'exécution d'un chantier.

La modélisation des informations du bâtiment (ou BIM pour Building Information Modelling) peut être définie comme « a modeling technology and associated set of processes to produce, communicate and analyse building models » (Sacks, Eastman, Lee, & Teicholz, 2018). Par sa nature pluridisciplinaire, le domaine de la construction regroupe des métiers différents et donc des données de nature différentes qu'il est nécessaire de regrouper au sein d'une même modélisation (modèles de données classiques, modèles de connaissance formels, comme les ontologies, les taxonomies, les hiérarchies de concepts).

Objectifs et problématique :

La modélisation des informations de conception est largement traitée dans le BIM. En effet, au-delà des standards couramment utilisés dans la conception mécanique (STEP(Pratt, 2005), IGES (Association & others, 1996), DXF (Autodesk, 2010), ...), des standards ont été définis pour la modélisation des données de construction tels que IFC (ISO, 2013) ou CIS/2 (Reed, 2003). Dans le cadre du projet ISOBIM, il est envisagé d'**intégrer 3 points de vue métiers (conception, planification et suivi de la réalisation) et leurs contraintes associées pour supporter le processus de rénovation**. L'ensemble des activités d'ISOBIM requiert en effet un système d'information adéquat implémentant différentes fonctionnalités, parmi lesquelles des algorithmes de clusterisations efficaces, un modèle de données capable de stocker des données multimodales : géométriques (panneaux issus des possibilités de calepinage), de planning (ressources, stocks, gammes, ...) et de suivi de chantier (statut d'une tâche, délai estimé, personne en charge de la tâche ...). Ce système d'information pourra être support au **jumeau numérique travaillant sur le processus de rénovation**.

Nous nous intéresserons donc aux questions de recherche suivantes :

- 1) **Quelle structure de données serait la plus adaptée pour notre un jumeau numérique ?**
- 2) **Comment obtenir cette structure en réduisant le temps/la difficulté de modélisation nécessaire pour son obtention ?**

Pour répondre à ces deux questions, ces travaux chercheront à explorer l'usage des outils de fouille de données afin de proposer les apports suivants :

- La proposition d'une **formalisation de la connaissance sous forme ontologique pour structurer les concepts du jumeau numérique du processus de rénovation** ;
- La proposition d'une **méthodologie semi-automatique permettant d'obtenir cette formalisation, basée sur des algorithmes de clustering**. Cette méthodologie reposera sur de la fouille de données, avec deux méthodes déjà envisagées. Dans un premier temps, la méthode Formal Concept Analysis (FCA) (Ganter, Stumme, & Wille, 2004) serait utilisée pour la modélisation de la connaissance à partir d'un ensemble d'exemples ou de données structurées. Dans un second temps, son extension Relational Concept Analysis (RCA) (Rouane-Hacene, Huchard, Napoli, & Valtchev, 2013), (Wajnberg, Valtchev, Lezoche, Panetto, Blondin Massé, 2019) serait employée pour l'extraction des relations entre chaque domaine de connaissances.

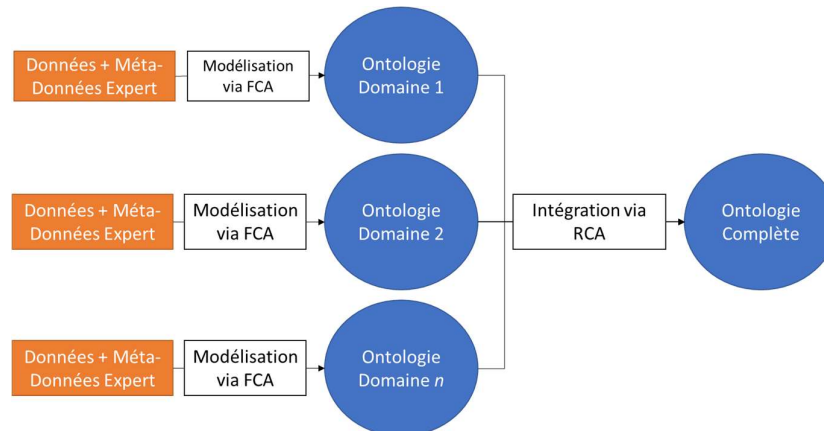


Figure 1. Modélisation et intégration semi-automatique pour la construction de l'ontologie

Profil recherché :

- Etudiant titulaire d'un master en informatique ou en génie industriel,
- Connaissances en modélisation de données et de connaissances.
- Capacité de travail en équipe

Lieu de la thèse : La thèse se déroulera au CRAN-Epinal -

Rémunération : ~1768.55 brut (1400€ net)

Démarrage envisagé : septembre 2021

Contacts :

Hind BRIL EL HAOUZI

William DERIGENT

Mario LEZOCHÉ

hind.el-haouzi@univ-lorraine.fr (Directrice de thèse)

william.derigent@univ-lorraine.fr

mario.lezoche@univ-lorraine.fr

Références :

- Association, U. S. P. D., & others. (1996). Initial Graphics Exchange Specification IGES 5.3. *ANSI*. Retrieved July, 12, 2008.
- Autodesk. (2010). *DXF Reference*. Consulté à l'adresse <https://www.autodesk.com/techpubs/autocad/acadr14/dxf/>
- Ganter, B., Stumme, G., & Wille, R. (2004). Formal Concept Analysis: Theory and applications-I. UCS Special Issue. *Journal of Universal Computer Science*, 10(8), 926.
- ISO. (2013). Industry Foundation Classes (IFC) for data sharing in the construction and facility management industries. Consulté à l'adresse <https://www.iso.org/standard/51622.html>
- Pratt, M. J. (2005). ISO 10303, the STEP standard for product data exchange, and its PLM capabilities. *International Journal of Product Lifecycle Management*, 1(1), 86. <https://doi.org/10.1504/IJPLM.2005.007347>
- Reed, K. (2003). The role of the CIMSteel integration standards in automating the erection and surveying of structural steelwork. *NIST SPECIAL PUBLICATION SP*, 15-20.
- Rouane-Hacene, M., Huchard, M., Napoli, A., & Valtchev, P. (2013). Relational concept analysis: Mining concept lattices from multi-relational data. *Annals of Mathematics and Artificial Intelligence*, 67(1), 81-108. <https://doi.org/10.1007/s10472-012-9329-3>
- Sacks, R., Eastman, C., Lee, G., & Teicholz, P. (2018). *BIM handbook: a guide to building information modeling for owners, designers, engineers, contractors, and facility managers*. John Wiley & Sons.
- Wajnberg, M., Valtchev, P., Lezoche, M., Panetto, H., Blondin Masse, A. (2019). Mining process factor causality links with multi-relational associations. *10th International Conference on Knowledge Capture, K-CAP'19, Nov 2019, Marina Del Rey, CA, United States*. pp.263-266, (10.1145/3360901.3364446). (hal-02377662v2)



Thesis subject: Semi-automatic formalisation of a knowledge repository supporting the digital twin for building renovation: application to the IsoBIM project

Context:

the ANR ISOBIM project aims to propose a collaborative process for renovation by external ISOLation based on Lean and BIM paradigms. In this context, one of the objectives is to be able to ensure the management of project data in a formal repository capable of covering the design, scheduling, planning and implementation monitoring phases. In addition, this repository will have to support the digital building twin allowing, in particular, 4D simulation in the execution phase of a building site.

Building Information Modelling (BIM) can be defined as "a modeling technology and associated set of processes to produce, communicate and analyse building models" (Sacks, Eastman, Lee, & Teicholz, 2018). Because of its multidisciplinary nature, the construction field brings together different trades and therefore different types of data that need to be grouped together within the same modelling (classic data models, formal knowledge models, such as ontologies, taxonomies, concept hierarchies).

Objectives and issues:

The modelling of information related to the design phase is largely dealt in the BIM. Indeed, beyond the standards commonly used in mechanical design (STEP (Pratt, 2005), IGES (Association & others, 1996), DXF (Autodesk, 2010), ...), standards have been defined for the modelling of construction data such as IFC (ISO, 2013) or CIS/2 (Reed, 2003). Within the framework of the IsoBIM project, it is envisaged to integrate 3 business points of view (design, planning and construction monitoring) and their associated constraints to support the renovation process. All of IsoBIM's activities require an adequate information system implementing various functionalities, including efficient clustering algorithms, a data model capable of storing multimodal data: geometric (panels from the layout possibilities), planning (resources, stocks, ranges, etc.) and site monitoring (task status, estimated time, person in charge of the task, etc.). This information system could be used to support the digital twin of the renovation process.

We are therefore interested in the following research questions:

- 1) Which data structure would be the most suitable for our digital twin?
- 2) How can this structure be obtained automatically to reduce the modelling time/difficulty?

To answer these two questions, this work will seek to explore the use of data mining tools to propose the following contributions:

- A formalization of knowledge in ontological form to structure the concepts of the digital twin of the renovation process.
- A semi-automatic methodology to obtain this formalisation, based on clustering algorithms.

This last point will be based on data mining, with two methods already envisaged. Firstly, the Formal Concept Analysis (FCA) method (Ganter, Stumme, & Wille, 2004) would be used for knowledge modelling based on a set of examples or structured data. In a second step, its extension Relational Concept Analysis (RCA) (Rouane-Hacene, Huchard, Napoli, & Valtchev, 2013) (Wajnberg, Valtchev,

Lezoche, Panetto, Blondin Massé, 2019) would be used for the extraction of relationships between each knowledge domain.

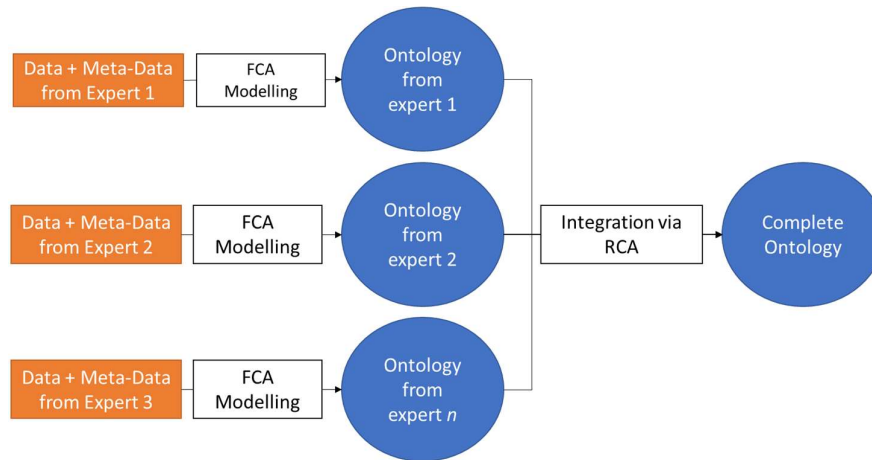


Figure 1: Modelling and semi-automatic integration for the global ontology construction

Candidate profile:

- Student with a master's degree in computer science or industrial engineering,
- Knowledge in data and knowledge modelling.
- Ability to work in a team

Location of the thesis: The thesis will take place at the CRAN-Epinal -

Salary: ~ 1768.55 brut (1400€ net)

Planned start-up: September 2021

Contacts :

Hind BRIL EL HAOUZI

William DERIGENT

Mario LEZOCHÉ

hind.el-haouzi@univ-lorraine.fr (Phd Supervisor)

william.derigent@univ-lorraine.fr

mario.lezoche@univ-lorraine.fr

Bibliography :

Association, U. S. P. D., & others. (1996). Initial Graphics Exchange Specification IGES 5.3. *ANSI*. Retrieved July, 12, 2008.

Autodesk. (2010). *DXF Reference*. Consulté à l'adresse <https://www.autodesk.com/techpubs/autocad/acadr14/dxf/>

Ganter, B., Stumme, G., & Wille, R. (2004). Formal Concept Analysis: Theory and applications-J. UCS Special Issue. *Journal of Universal Computer Science*, 10(8), 926.

ISO. (2013). Industry Foundation Classes (IFC) for data sharing in the construction and facility management industries. Consulté à l'adresse <https://www.iso.org/standard/51622.html>

Pratt, M. J. (2005). ISO 10303, the STEP standard for product data exchange, and its PLM capabilities. *International Journal of Product Lifecycle Management*, 1(1), 86. <https://doi.org/10.1504/IJPLM.2005.007347>

Reed, K. (2003). The role of the CIMSteel integration standards in automating the erection and surveying of structural steelwork. *NIST SPECIAL PUBLICATION SP*, 15-20.

Rouane-Hacene, M., Huchard, M., Napoli, A., & Valtchev, P. (2013). Relational concept analysis: Mining concept lattices from multi-relational data. *Annals of Mathematics and Artificial Intelligence*, 67(1), 81-108. <https://doi.org/10.1007/s10472-012-9329-3>

Sacks, R., Eastman, C., Lee, G., & Teicholz, P. (2018). *BIM handbook: a guide to building information modeling for owners, designers, engineers, contractors, and facility managers*. John Wiley & Sons.

Wajnberg, M., Valtchev, P., Lezoche, M., Panetto, H., Blondin Masse, A. (2019). Mining process factor causality links with multi-relational associations. *10th International Conference on Knowledge Capture, K-CAP'19, Nov 2019, Marina Del Rey, CA, United States*. pp.263-266, {10.1145/3360901.3364446}. {hal-02377662v2}