



N° d'ordre NNT : 2023LYO20092

# THÈSE de DOCTORAT DE L'UNIVERSITÉ LUMIÈRE LYON 2

**École Doctorale : ED 512**  
**Informatique et Mathématiques**

Discipline : Informatique

Soutenue publiquement le 11 décembre 2023 par :

**Diego VINASCO-ALVAREZ**

---

## **Model-driven Integration of Heterogeneous n-Dimensional Urban Data.**

---

Devant le jury composé de :

Christophe NICOLLE, Professeur, Université de Bourgogne, Président

Marlène VILLANOVA-OLIVER, Maîtresse de conférences HDR, Université Grenoble Alpes,

Rapporteure

François PINET, Directeur de recherche HDR, INRAE, Rapporteur

Gilles GESQUIÈRE, Professeur, Université Lumière Lyon 2, Directeur de thèse

Sylvie SERVIGNE, Maîtresse de conférences, INSA de Lyon, Co-Directrice de thèse

John SAMUEL, Enseignant chercheur, CPE Lyon, Co-Directeur de thèse

# Contrat de diffusion

Ce document est diffusé sous le contrat *Creative Commons* « [Paternité – pas de modification](#) » : vous êtes libre de le reproduire, de le distribuer et de le communiquer au public à condition d'en mentionner le nom de l'auteur et de ne pas le modifier, le transformer ni l'adapter.

**Sujet :**

Intégration de données urbaines hétérogènes n-dimensionnelles basée modèle

**Mots-clefs :** *Données 3D, données spatio-temporelles, données urbaines, intégration des données, transformation des données, ontologies, graphes de connaissances, modèles conceptuels, basée-modèles, standards de données*

**Encadrant :** [Gilles Gesquière](#)

**Co-encadrants :** [John Samuel](#), [Sylvie Servigne](#)

**Résumé :**

L'urbanisation et l'anthropisation sont des processus de changement dynamiques et multiformes qui ont de fortes répercussions sur nos sociétés. Pour répondre au besoin de comprendre ces processus, les approches fondées sur les données---telles que les jumeaux numériques urbains et les applications de villes intelligentes---sont devenues des solutions puissantes pour modéliser, visualiser et naviguer dans le paysage et le cycle de vie urbains complexes. Souvent, ces approches reposent sur l'intégration de ces données urbaines, qui consistent en une variété d'informations provenant de différents acteurs et organisations. Dans ce contexte, l'intégration des données urbaines est un processus qui combine des données urbaines hétérogènes provenant de ces domaines d'information afin de créer des vues de données plus complètes du paysage urbain et de son évolution pour l'utilisateur. Pour fournir ces vues, le processus d'intégration doit prendre en compte à la fois la nature hétérogène de ces données et leurs caractéristiques n-dimensionnelles (nD) (i.e., 2D, 3D, temps, sémantique).

En général, l'intégration des données nécessite de relever une grande variété de défis. Dans le domaine des données urbaines nD, cette thèse identifie deux problèmes primordiaux d'intégration de données à résoudre. Premièrement, divers types d'hétérogénéité peuvent exister entre les différentes sources de données urbaines nD. Dans ce contexte, l'hétérogénéité peut prendre de nombreuses formes, telles que des différences sémantiques ou structurelles entre les normes de données. Deuxièmement, les processus de transformation ou de conversion des données sont souvent utilisés pour l'intégration des données urbaines. Toutefois, ces processus peuvent entraîner une perte de données sémantiques.

Ces problèmes donnent lieu aux questions de recherche suivantes qui orientent le travail présenté dans cette thèse : En premier, comment

les approches d'intégration de données urbaines de la nD peuvent-elles garantir que les normes de données urbaines peuvent être facilement réutilisées, même si ces normes évoluent ? Ensuite, comment limiter la perte de données lors de la transformation de données entre des formats de données nD hétérogènes ?

Afin de relever ces défis, cette thèse propose une approche sémantique robuste pour l'intégration de données basée sur des normes et des modèles où les modèles conceptuels sous-jacents aux différentes sources de données sont préservés. Au lieu d'utiliser la conversion directe des données, nous proposons une méthodologie axée sur les modèles. Les approches telles que l'Model-Driven Architecture (MDA) permettent la transformation des modèles conceptuels de données, tout en limitant la perte des informations sémantiques et de la structure du modèle d'origine. Cette approche se prête également à l'adaptabilité et à la reproductibilité, car les transformations elles-mêmes peuvent être matérialisées, partagées et modifiées si nécessaire. Les approches fondées sur des modèles peuvent également garantir que, pour une application donnée, toutes les informations pouvant être représentées par une norme donnée seront disponibles pour l'application, même si les normes et les modèles de données qu'elles définissent évoluent implicitement. Cependant, la construction d'un modèle sémantique robuste reste un défi, car le choix d'un modèle sémantique expressif nécessite une bonne compréhension des différents types d'informations qui peuvent être représentés par les diverses sources et les normes auxquelles elles appartiennent. En outre, la mise en correspondance automatisée des informations du modèle conceptuel présente un certain nombre de limites. En prenant en compte les standards hétérogènes et en constante évolution des données urbaines nD, notre objectif est de pouvoir assurer l'interopérabilité des données et la possibilité d'intégrer ces données avec d'autres données ouvertes connexes.

En réponse à ces défis, cette thèse propose également des transformations vers une représentation computationnellement compatible et lisible par machine des modèles de données urbaines nD. Cette représentation permet la formalisation de ces modèles en tant qu'ontologies calculatoire (ou graphes de connaissances). En outre, la formalité de ces modèles et leur application dans les processus de validation des données sont également explorées dans cette thèse. En particulier, comment ces processus peuvent-ils être utilisés pour garantir la haute qualité des données transformées. En outre, la nature lisible par machine de cette représentation permet à ces modèles d'être directement utilisés dans des applications de données urbaines pour fournir un contexte supplémentaire au paysage urbain.

L'objectif final de cette thèse est de s'assurer que les données urbaines intégrées par cette approche peuvent être facilement réutilisées par des applications basées sur les données urbaines telles que les jumeaux numériques urbains afin de fournir aux utilisateurs des vues plus complètes du paysage urbain en évolution. À cette fin, plusieurs contributions scientifiques et techniques ont été produites au cours de cette thèse. Ces contributions comprennent la formalisation de modèles de données urbaines nD, des extensions et des règles issues de normes évolutives pour naviguer dans des scénarios simultanés d'évolution urbaine. En outre, des flux de transformation de données reproductibles pilotés par des modèles sont développés et appliqués pour l'intégration de données urbaines nD ouvertes du monde réel. Ces modèles de données et leurs ensembles de données conformes sont mis en œuvre dans des applications de données urbaines pour naviguer entre différents scénarios d'évolution, à la fois réels et hypothétiques. Enfin et notamment, plusieurs contributions au groupe de travail sur les normes CityGML de l'Open Geospatial Consortium sont apportées dans le cadre de cette thèse, y compris la participation à la création d'un encodage de la dernière version de la norme.