

**DESCRIPTION SUJET DE THESE  
POUR UN CO-FINANCEMENT  
RÉGION GRAND EST  
« CAMPAGNE 2021 »**

TITRE DE LA THÈSE :

Méthodes formelles d'extraction et de réutilisation des connaissances provenant de sources hétérogènes pour l'interopérabilité sémantique des architectures distribuées

Université : Université de Lorraine

Laboratoire : Centre de Recherche en Automatique de Nancy (CRAN)

Département : Ingénierie des Systèmes Eco-Techniques (ISET)

Directeurs proposés : Mario LEZOCHÉ (Mdc HDR), Hervé PANETTO (PU)

Mots clés : Formalisation de la connaissance, Fouille de données multi relationnelles, Analyse de concepts relationnels, Systèmes cyber physique, Interopérabilité sémantique des plateformes

## Contexte et objectifs

Le tourisme est un enjeu stratégique pour la France. Première destination touristique mondiale depuis plusieurs années avec 89 millions de visiteurs par an, le secteur pèse plus de 7% du PIB de l'Hexagone. Dans la région du Grand Est il y a eu l'émergence d'une nouvelle politique régionale en matière touristique. Elle apparaît comme un enjeu majeur tant en termes d'attractivité touristique et économique, que de participation à la définition d'une véritable image du Grand Est. Définie à l'issue d'un travail de convergence, la stratégie régionale s'accompagne de l'élaboration d'un Schéma régional de développement touristique (SRT). Voté en mars 2018, il répond aux exigences d'attractivité, d'excellence et de proximité afin de promouvoir une filière qui représente à elle seule 4% de l'emploi régional.

Lorsque l'on étudie chacune de ses grandes composantes (transport, hébergement, restauration et loisirs), de nouveaux acteurs émergent régulièrement sur le marché. Ce rythme d'innovation soutenu a pour conséquence de modifier les rapports de forces en présence. Comme on a bien vu le tourisme est toujours un pilier stratégique mis en avant par les pouvoirs publics, pilotés par l'État au niveau national et régional, tout en s'appuyant sur les collectivités locales à des échelons inférieurs comme les offices de tourisme de France (OTF), les comités départementaux du tourisme (CDT) et les comités régionaux du tourisme (CRT). Cette partie institutionnelle du tourisme a souffert et souffre encore du grand boom des nouvelles technologies. Capter l'intérêt de nouveaux voyageurs à l'heure du numérique et suivre les dernières tendances toujours plus rapides nécessitent une approche novatrice intégrant les nouvelles habitudes digitales des consommateurs.

La transformation du marché du tourisme a commencé dès le début des années 2000 avec la démocratisation d'internet ; Booking et Expedia apparaissent en 1996 et Trip Advisor en 2000 par exemple. L'apparition des OTA (Online Travel Agencies) et des comparateurs en ligne a eu un impact direct sur la chaîne de valeur existante qui comptait un grand nombre d'intermédiaires. De nombreuses plateformes se sont ainsi développées dans le but de centraliser le trafic internet, capter et monétiser la donnée client afin de se positionner comme un point de passage obligatoire pour les acteurs du tourisme. Au travers d'un commissionnement sur les ventes allant parfois jusque + 30%, ces acteurs sont devenus une force et une faiblesse du modèle actuel entraînant une majoration des prix de ventes pour pouvoir prendre en compte les commissions. Le voyageur souhaitant pouvoir gérer simplement et de bout en bout la réservation de ses prochaines vacances via un unique point de contact. En tant qu'acteur fortement impliqué dans le développement économique des massifs, et dans notre région visant le développement du Massif des Vosges, le Syndicat National des Moniteurs du Ski de France (SNMSF) qui regroupe 17 000 moniteurs au sein des Écoles du Ski Français (ESF) sur l'ensemble du territoire national, participe au développement du Tourisme depuis 1945.

En 2016, le SNMSF a lancé une place de marché « Mon Séjour en Montagne » qui repose sur des valeurs fortes d'équité pour les acteurs touristiques et de transparence et fluidité pour le client. Cette plateforme collaborative porte la vision du SNMSF d'un tourisme durable, équitable et solidaire.

Le projet ambitionne de rendre la plateforme un moyen plus équitable pour poser les jalons d'un système global, intelligent et connecté (Smart Montagne) permettant la collecte et le traitement des données en temps réel afin d'extraire de la nouvelle connaissance à des fins touristiques. Ces données et informations seront ensuite agrégées et transformées par des systèmes d'inférence automatique pour formaliser les connaissances implicites existantes.

## Descriptif du sujet de thèse et des impacts socio-économiques

Pour apporter à ce projet, déjà en marche, des contributions scientifiques précises et concrètes on utilisera une approche de l'ingénierie des systèmes interopérables (Ramos, 2011) et (Morel, 2003) qui consiste à s'appuyer sur différents types et niveaux d'abstraction

ou de modèles. Ces modèles doivent exprimer et formaliser non seulement l'aspect "structurel" des composants du système, mais aussi leur comportement (Maier, 1998), qui peut être limité par les exigences spécifiques du domaine du système (règles commerciales). Un autre type de contrainte peut être induit par le ou les protocoles d'interopérabilité qui peuvent imposer des règles strictes pour doter les systèmes interopérables de propriétés telles que l'autonomie, la confidentialité et la transparence (Zdravković et al, 2016).

L'objectif de ce projet de recherche est double : d'une part, modéliser des données provenant de sources hétérogènes et, d'autre part, étudier les problèmes posés par l'ingénierie dirigée par les modèles dans les systèmes coopératifs. Impliquant une coopération concernant les "systèmes d'acteurs" désireux d'interopérer. Les systèmes collaboratifs sont désormais organisés en réseaux, soit sous forme de systèmes complexes (Camarinha-Matos, 2014). Les systèmes complexes envisagés seront composés de réseaux de Systèmes Cyber Physique, des capteurs intelligents, qui récupéreront des données en insérant le contexte et formeront ainsi des réseaux d'information (Cardin, 2016).

### Verrous scientifiques adressés par la thèse

Face à ce défi les verrous scientifiques concernent :

1. Le manque de formalisation (en autres mathématique) de l'agglomération des informations dans les modèles de systèmes et des systèmes d'informations qui en émergent, ainsi que la définition de la sémantique des concepts et des relations qu'ils mettent en œuvre, pour assurer leur compréhension commune, et faciliter leur interopération en minimisant les pertes sémantiques ;
2. L'adaptation (voire l'extension) des outils de nature algébrique et/ou géométrique (théorie des treillis, théorie des catégories, algèbre homologique) dans le contexte de l'analyse des concepts formels, pour le traitement de données hétérogènes en constante évolution. Ce qui constitue une approche récente qui n'a pas encore été pleinement développée (même du point de vue mathématique) pour ce type de données.

### Résultat scientifiques attendus et impact socio-économique :

Le sujet de thèse proposé se positionne clairement dans la continuité de la thèse effectuée par Mickael Wajnberg de 2017 à 2020 au CRAN ayant posé les fondements d'une méthodologie d'extraction de connaissance polyvalente et portant plus spécifiquement sur la gestion des règles d'association dans des contextes relationnels. La thèse proposée entre dans le cadre du Domaine d'Expertise Stratégique « Industrie 4.0 ». Les précédents travaux ont démontré l'intérêt d'une approche holistique de l'ensemble des ressources informationnelles et ont permis le développement d'une méthodologie focalisée sur l'optimisation de la gestion de connaissance. La présente proposition vise donc à poursuivre les travaux engagés par l'élaboration d'une méthode formelle d'extraction et de réutilisation des connaissances provenant de sources hétérogènes pour l'interopérabilité sémantique des architectures distribuées. Cette méthode sera intégrée comme brique méthodologique au processus de gestion des ressources d'information aptes à l'aide à la décision.

Le cas d'application sera constitué par les systèmes d'informations et de remontés des données de au moins une des stations de ski des Vosges.

### Retombées socio-économiques attendues

Les attendus pour les stations de ski des Vosges concernent la validation de l'applicabilité concrète de la méthodologie et de sa valeur ajoutée par rapport aux approches alternatives présentes chez eux. Les retombées estimées sont autant opérationnelles (développement de



l'approche générique de formalisation de la connaissance, développement de la méthode d'interopérabilité sémantique entre les sources hétérogènes de données, développement algorithmique, développement de modèles de connaissance, développement de la plate-forme web et Data Hub destinés aux touristes et aux entreprises du domaine du tourisme etc.) qu'organisationnelles (meilleure gestion des ressources, meilleure utilisation des bases de connaissances évolutive remplie de données provenant de systèmes cyber physiques et de d'autres plateformes open source) que business (mise en application des résultats, déploiement sur les autres stations & automatisation / généralisation). Du fait de cette démonstration, cette méthodologie devra ensuite être industrialisée et déployée sur un grand nombre de systèmes de production.

CENTRE DE  
RECHERCHE EN  
AUTOMATIQUE DE  
NANCY



UNIVERSITÉ  
DE LORRAINE

CRAN UMR 7039

CAMPUS SCIENCES - BP 70239 - 54506 Vandœuvre-lès-Nancy - Cedex

Tél. : + 33 (0)383684419 - Fax : + 33 (0)383684437

E-mail : [cran-secretariat@univ-lorraine.fr](mailto:cran-secretariat@univ-lorraine.fr) - Web : [www.cran.univ-lorraine.fr](http://www.cran.univ-lorraine.fr)