



HAL
open science

Immersion de divisions territoriales et description de leur évolution dans le Web sémantique

Camille Bernard, Marlène Villanova-Oliver, Jérôme Gensel, Philippe Genoud,
Hy Dao

► **To cite this version:**

Camille Bernard, Marlène Villanova-Oliver, Jérôme Gensel, Philippe Genoud, Hy Dao. Immersion de divisions territoriales et description de leur évolution dans le Web sémantique. Journées francophones d'Ingénierie des Connaissances (IC 2020), Association française pour l'intelligence artificielle, Jun 2020, Angers, France. hal-02885484

HAL Id: hal-02885484

<https://hal.science/hal-02885484>

Submitted on 30 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Immersion de divisions territoriales et description de leur évolution dans le Web sémantique

Camille Bernard¹, Marlène Villanova-Oliver¹, Jérôme Gensel¹, Philippe Genoud¹, Hy Dao²

¹ UNIV. GRENOBLE ALPES, CNRS, GRENOBLE INP, LIG
camille.bernard@univ-grenoble-alpes.fr,
marlene.villanova-oliver@univ-grenoble-alpes.fr,
jerome.gensel@univ-grenoble-alpes.fr, philippe.genoud@univ-grenoble-alpes.fr

² UNIV. DE GENÈVE, DÉPARTEMENT DE GÉOGRAPHIE ET ENVIRONNEMENT
hy.dao@unige.ch

Résumé :

Partout dans le monde, les découpages géographiques, dont les territoires font l'objet, sont soumis à des modifications de noms, d'affiliations, de frontières, etc. Ces changements sont un obstacle à la comparabilité des données statistiques (socio-économiques, sanitaires, environnementales...) mesurées sur ces territoires sur de longues périodes. Afin d'apporter une solution conceptuelle et opérationnelle à cette problématique, nous proposons un framework, nommé *Theseus*, qui s'appuie sur les technologies du Web sémantique pour décrire l'évolution des découpages géographiques dans le Web des données ouvertes et liées (*Linked Open Data* (LOD) en anglais). Notre approche couvre la modélisation des zones géographiques et de leurs changements au cours du temps, mais aussi la détection et la description automatique des changements, ainsi que l'exploitation de ces descriptions dans le LOD Cloud au moyen de requêtes SPARQL.

Mots-clés : Web Sémantique, Ontologies, Évolution, Données géographiques.

1 Introduction

Depuis la mise en place de directives officielles visant à l'ouverture des données créées par les institutions publiques un peu partout dans le monde, le volume du Web des données ouvertes et liées augmente continuellement. En particulier, le secteur public, les organismes gouvernementaux et notamment les instituts officiels de statistique et de cartographie déversent de plus en plus de contenus. Ces institutions publient des statistiques géo-codées à travers des découpages géographiques permettant aux responsables politiques d'accéder à des analyses fines du territoire dont ils ont la charge. Les découpages géographiques sur lesquels reposent ces données statistiques sont construits pour les besoins de la statistique, mais dérivent généralement de structures électorales ou administratives. On parle dans ce contexte de *Nomenclatures Statistiques Territoriales* (acronyme TSN en anglais). Les TSN codifient les unités géographiques qui, sur plusieurs niveaux d'imbrication (par exemple, en France les niveaux régional, départemental, communal, etc.), composent un territoire observé.

Une TSN fournit ainsi un catalogue d'Unités Territoriales (UT, également appelées domaines statistiques), souvent présentés selon une hiérarchie qui se reflète dans les codes numériques ou alphabétiques qui leur sont attribués. Une TSN est utilisée pour compiler des données statistiques sur un ou plusieurs territoire(s) étudié(s), à un ou plusieurs niveau(x) de division géographique. Les TSN sont utilisées par les instituts statistiques (locaux, régionaux, nationaux, internationaux) ou d'autres organisations pour établir et publier des statistiques comparables et intégrées dans l'espace.

Partout dans le monde, les divisions territoriales sont soumises à des modifications de noms, d'affiliations, de frontières, etc. Par exemple en 2016, en France, les régions administratives ont fusionné en grandes régions. Ces changements territoriaux sont très fréquents, d'autant plus que le maillage géographique est fin. Or, ces changements sont un obstacle à la comparabilité de données statistiques (socio-économiques, sanitaires, environnementales...)

ayant été agrégées spatialement par régions géographiques (des données ponctuelles agrégées par régions administratives, par exemple). Ce problème de comparabilité, bien connu sous le nom de *problème d'agrégation spatiale* (*Modifiable Areal Unit Problem* en anglais) (Openshaw, 1984) décrit le phénomène selon lequel des données collectées dans différents découpages géographiques (ou différentes versions d'un même découpage) ne sont pas comparables en raison des différences potentielles entre les zones géographiques observées.

Face à ces problèmes de changements territoriaux pouvant entraîner des ruptures dans les séries chronologiques, les instituts statistiques choisissent de fournir une estimation des données passées calculées pour les toutes dernières versions des découpages géographiques ou de s'affranchir des frontières territoriales, en créant des cartes de chaleur par exemple. Les traces des changements territoriaux (changements des frontières des régions, changements de noms de communes, etc.) sont alors comme effacées, ignorées, bien qu'en général, de tels changements aient un sens et soient capitaux pour comprendre un territoire. Il est alors crucial de conserver et d'enrichir ces informations relatives aux changements territoriaux, par devoir de mémoire, mais aussi pour fiabiliser les traitements des données relatives à ces territoires. D'une part, l'objectif est de fournir les connaissances permettant une lecture du territoire fidèle à la réalité du terrain et assurant une certaine traçabilité des décisions politiques. D'autre part, il s'agit d'éviter toute erreur d'analyse des données par exemple lorsque les zones géographiques auxquelles elles se rapportent changent alors même qu'elles conservent leur identifiant dans les jeux de données statistiques (dans les TSN, des unités de même code peuvent ne plus couvrir tout à fait le même espace territorial). Ainsi, un enjeu est de structurer la connaissance permettant de comprendre l'organisation et l'évolution des territoires, dans l'optique de fournir à différents acteurs (statisticiens, chercheurs, décideurs ou encore citoyens) des descriptions riches aidant, par suite, à une meilleure exploitation et une meilleure compréhension des données statistiques géo-codées.

Afin d'apporter une solution conceptuelle et opérationnelle à cette problématique, nous proposons un framework, nommé *Theseus*, qui s'appuie sur les technologies du Web sémantique pour automatiser la représentation de découpages géographiques et de leurs évolutions au cours du temps, sous forme de données ouvertes et liées (*Linked Open Data* en anglais (LOD)). Nous adoptons les technologies du Web sémantique afin de bénéficier de la base de connaissances distribuées que constitue le LOD Web (ou LOD Cloud) nous permettant d'enrichir les informations sur les changements territoriaux avec des métadonnées et d'autres ressources disponibles dans le Web pouvant contribuer à expliquer les changements (par exemple, des événements historiques). Ces technologies garantissent également l'interopérabilité syntaxique et sémantique entre des systèmes échangeant des TSN produites par les agences statistiques. Le framework *Theseus* se compose d'un ensemble de modules permettant la gestion du cycle de vie des TSN dans le LOD : de la modélisation des zones géographiques et de leurs changements au cours du temps, à la détection automatique des changements via un algorithme d'appariement de versions de TSN (Bernard *et al.*, 2018a), jusqu'à l'exploitation de ces descriptions dans le LOD Cloud au moyen de requêtes SPARQL. L'ensemble des modules logiciels est articulé autour de deux ontologies nommées TSN Ontology (<http://purl.org/net/tsn>) et TSN-Change Ontology (<http://purl.org/net/tsnchange>) (Bernard *et al.*, 2018b; Bernard, Camille *et al.*, 2018), conçues pour une description spatiale et temporelle non ambiguë des structures géographiques et de leurs modifications au cours du temps. Ces deux ontologies s'appuient sur des ontologies standards telles que l'ontologie OWL-Time (pour la dimension temporelle des données) et l'ontologie GeoSPARQL (pour la représentation de la dimension spatiale des données). Les graphes de connaissances générés par *Theseus* améliorent la compréhension des dynamiques territoriales, en fournissant aux décideurs politiques, aux techniciens, aux chercheurs et au grand public des descriptions sémantiques fines des changements territoriaux, exploitables pour des analyses fiables et traçables.

Cet article présente l'ensemble du travail que nous avons réalisé pour la publication de TSN évolutives dans le LOD : ensemble qui prend la forme d'un framework jamais présenté jusqu'alors. Cet objectif soulève des défis qui s'inscrivent à différentes étapes du cycle de vie établi par (Villazón-Terrazas *et al.*, 2011) pour la publication des données liées gouvernementales. Dans une première partie, nous présentons ces défis et montrons en quoi des approches existantes contribuent à les relever en partie, sans toutefois y répondre pleinement.

Notre approche, qui comble les lacunes de travaux existants, est incarnée par le framework Theseus dont nous exposons les grands principes. Nous justifions des choix de modélisation retenus pour les ontologies TSN et TSN-Change qui offrent une solution inédite pour la représentation des versions de nomenclatures et la description des changements. Nous montrons comment ces ontologies ont été utilisées pour produire des graphes de connaissances sur les évolutions de trois nomenclatures territoriales officielles. Ceci illustre l'applicabilité et la généralité de notre approche pour verser sous forme de graphes RDF des divisions territoriales évolutives et la description de leurs changements dans le LOD. Enfin, nous présentons quelques unes des requêtes que nous avons prédéfinies pour interroger ces graphes de connaissances avant de conclure et de donner les perspectives envisagées pour ce travail.

2 Problématique et défis visés

2.1 Défi 1 : représenter les divisions territoriales évolutives de façon générique et interopérable

De nombreux instituts statistiques dans le monde publient désormais leurs statistiques sous forme de LOD (par exemple, les données statistiques italiennes¹, les données statistiques écossaises², les données statistiques du ministère britannique des communautés et des collectivités locales³, les données statistiques aragonaises⁴ ou les données statistiques japonaises⁵). Quand il s'agit d'associer une composante géographique à leurs données, ces instituts créent souvent leur propre ontologie pour la description des zones. Un état de l'art sur ces différentes initiatives est présenté dans (Bernard, 2019). Nous pouvons citer par exemple, la *Territorio Ontology*⁶ de l'Institut national italien de statistique, la *Geography Ontology*⁷ du gouvernement écossais, l'ontologie de l'*Ordnance Survey* d'Irlande⁸ (Debruyne *et al.*, 2017). Ceci se traduit par une prolifération contre-productive de vocabulaires non alignés. L'initiative la plus approfondie est celle de l'*Office for National Statistics* (ONS) du Royaume-Uni qui propose des vocabulaires⁹ pour représenter les zones géographiques dans le contexte de la publication de données statistiques. Dans cette approche, on peut noter un certain niveau d'abstraction sur les termes utilisés pour décrire les UT, ce qui permet de décrire d'autres UT que celles du Royaume-Uni. Néanmoins aucun de ces vocabulaires ne permet la description de n'importe quelle structure TSN et des niveaux qui la composent. Il est nécessaire d'ajouter de nouveaux concepts à l'ontologie Geography¹⁰ si l'on veut décrire de nouveaux territoires et de nouvelles hiérarchies territoriales que ceux énumérés. Aucune des initiatives étudiées n'offre un niveau d'abstraction suffisant pour décrire sémantiquement n'importe quelle structure hiérarchique de type TSN. Du côté des recommandations du W3C, l'ontologie¹¹ RDF Data Cube (QB) est largement utilisée pour décrire des statistiques dans le LOD. Basée sur la norme ISO SDMX, elle assure une certaine interopérabilité pour l'échange et le partage de données statistiques et de métadonnées entre les organisations (Cyganiak & Reynolds, 2014). En particulier, l'extension de QB pour les composants spatio-temporels (QB4ST)¹² (Atkin-

1. <http://datiopen.istat.it/index.php?language=eng>

2. <http://statistics.gov.scot/>

3. <http://opendatacommunities.org/data>

4. <http://opendata.aragon.es/>

5. <http://data.e-stat.go.jp/lodw/en>

6. <http://datiopen.istat.it/odi/ontologia/territorio/>

7. <http://statistics.gov.scot/vocabularies/>

8. Ireland's Administrative boundaries ontology :<http://ontologies.geohive.ie/osi#>

9. <http://statistics.data.gov.uk/vocabularies>

10. <http://statistics.data.gov.uk/graph/ontology/geography>

11. <http://purl.org/linked-data/cube#>

12. <http://www.w3.org/ns/qb4st/>

son, 2017) fournit des termes canoniques pour définir la dimension spatiale et la dimension temporelle des données. Deux concepts fournissent un moyen d'exprimer des entités imbriquées (par exemple, les pays contenant des unités administratives).

Le vocabulaire QB4OLAP¹³, quant à lui, permet la représentation de données multidimensionnelles, y compris, pour la dimension spatiale, la description de la hiérarchie des niveaux géographiques (Etcheverry & Vaisman, 2012). Cependant, en ce qui concerne la dimension temporelle des données, il s'agit d'une approche discrète qui ne représente pas les processus d'évolution des éléments dans le temps, ni les changements et événements sous-jacents. Dans (Plumejeaud *et al.*, 2011), nous avons proposé un modèle de base de données relationnelle offrant un niveau d'abstraction suffisant pour décrire sémantiquement toute TSN. Ce modèle constitue un socle pertinent pour notre objectif mais doit être adapté en vue d'une immersion dans le Web sémantique. Par suite, visant le Web des Données Ouvertes et Liées, il doit être possible de référencer chaque UT avec précision. Ceci prend tout son sens dans le contexte de représentation de la connaissance qui est le nôtre, des connaissances relative aux évolutions des territoires. Ici, la notion d'UT doit être rapprochée de celle de *version d'UT*, elle-même liée à la notion de version de TSN. Cela nécessite de créer des identifiants uniques (URI dans le LOD) pour chaque ressource que constitue chaque représentation de ce qu'a été une UT au cours de sa vie (*i.e.*, chaque version d'UT). Ainsi, dans une perspective applicative, nous visons un modèle permettant à des producteurs de données de déclarer précisément pour chaque observation ou mesure quelle est exactement la version de l'UT concernée, et à quelle division géographique (hiérarchie) elle appartient (*i.e.*, version de TSN).

2.2 Défi 2 : représenter finement les changements territoriaux

En ce qui concerne la nature évolutive des zones géographiques couvertes par les observations et les TSN auxquelles ces zones évolutives appartiennent, aucune des ontologies décrites ci-dessus ne fournit un vocabulaire suffisamment générique et riche pour décrire la façon dont n'importe quelle TSN, mais aussi les niveaux territoriaux et les UT qui la composent, évoluent au fil du temps. Pourtant, représenter finement le changement territorial vise à en comprendre les raisons et à garantir le transfert fiable de données statistiques d'une version de TSN à l'autre. Il manque donc un modèle qui puisse aider à savoir comment un territoire a évolué dans le temps, un pré-requis essentiel à une interprétation et une compréhension plus justes des valeurs statistiques observées ou mesurées sur un territoire évolutif. La plupart des contributions des agences limitent leur description à des changements d'attributs isolés et ne proposent aucun vocabulaire pour décrire et regrouper les changements qui interviennent lors d'un même événement.

Parmi les initiatives déjà évoquées, nous revenons ici sur deux d'entre elles qui traitent du problème de l'évolution dans le temps :

- le Royaume-Uni (ONS) propose un vocabulaire¹⁴ pour représenter l'évolution des zones géographiques dans le cadre de la publication de données statistiques. Néanmoins, aucun des concepts de l'ontologie des changements de frontières de l'ONS ne permet de nommer les changements qui ont un impact sur plusieurs UT en même temps. Les concepts décrivant les événements de changement sont peu nombreux et se limitent aux termes *Recoding change*, *Boundary change*.
- Les vocabulaires du Bureau des statistiques du Japon¹⁵ représentent les zones géographiques et leur évolution, y compris la description des changements de zone géographique, dans le contexte de la publication de données statistiques. Cette proposition se limite à la description des changements dans les municipalités, par un simple littéral.

Ainsi, les changements intervenant dans les structures des TSN sont rarement décrits et lorsqu'ils le sont, les descriptions sont faites UT par UT, sans lien entre les changements. Il est donc difficile d'identifier chacun des composants de la TSN (territoires, niveaux et UT)

13. <http://purl.org/qb4olap/cubes>

14. <http://statistics.data.gov.uk/def/boundary-change>

15. <http://data.e-stat.go.jp/lod/sac/>, <http://data.e-stat.go.jp/lod/terms/sacs#>, <http://data.e-stat.go.jp/lod/sace/>

qui change. Il manque également un support pour décrire des changements qui se propagent de niveau en niveau (par exemple, le changement des limites d'une UT peut également avoir un impact sur les limites de ses sous-UT). Compte tenu de la nature imbriquée des éléments d'une TSN, cet aspect est un élément favorisant une compréhension plus globale d'une évolution territoriale observée selon une approche *top-down* ou *bottom-up*.

La littérature sur la représentation de l'évolution territoriale propose des solutions pour l'expression de lien de filiation (Del Mondo *et al.*, 2010; Harbelot *et al.*, 2013) entre unités géographiques et la caractérisation du changement (à travers des typologies (Claramunt & Thériault, 1995; Del Mondo *et al.*, 2010; Plumejeaud *et al.*, 2011; Del Mondo *et al.*, 2013) et des structures de représentations dédiées (Kauppinen *et al.*, 2008; Lopez-Pellicer *et al.*, 2012; Lacasta *et al.*, 2014)). Ces travaux ont inspiré notre contribution visant à offrir un cadre conceptuel pertinent pour une représentation sémantiquement riche de l'évolution. Les modèles ontologiques que nous proposons répondent ainsi aux défis 1 et 2 que nous venons de présenter. Ils sont décrits dans (Bernard, Camille *et al.*, 2018; Bernard, 2019) et brièvement rappelés dans la Section 4.

2.3 Défi 3 : automatiser la description des changements territoriaux

Dans (Lacasta *et al.*, 2014), les auteurs proposent un processus semi-automatique pour créer et alimenter le modèle représentant de l'évolution d'unités administratives correspondant à des Domaines Juridictionnels proposés dans (Lopez-Pellicer *et al.*, 2012). Le processus repose sur une approche qui permet de définir les caractéristiques des sources de données à intégrer afin de gérer les Domaines Juridictionnels de n'importe quel pays. Toutefois, pour peupler le modèle ontologique, il est nécessaire d'utiliser en entrée un dictionnaire énumérant tous les changements individuels. Cette contrainte est également observée dans le travail de (Kauppinen *et al.*, 2008). Ceci constitue un inconvénient majeur pour les systèmes d'information statistique actuels, car l'établissement manuel de la liste des changements observés dans toute une nomenclature est coûteuse en temps.

D'autres approches telles que (Plumejeaud, 2011) et (Harbelot *et al.*, 2015) consistent à développer des algorithmes et programmes afin de *détecter* les changements entre deux versions, au lieu de les retranscrire à partir d'une liste préétablie. Détecter un changement entre deux UT puis le décrire sémantiquement nécessite de traiter la question de l'identité des UT. D'une part, pour qu'un changement soit décrété, il faut définir quelles sont les variations à observer pour conclure à un changement. Ensuite, selon ses variations, une seconde décision doit être prise concernant l'impact de ce changement. En effet, un changement peut être considéré comme suffisamment important en regard de l'identité d'une UT pour que soit questionnée la continuité de l'existence-même de cette UT. Par exemple, une UT, dont le nom ou la géométrie ou les deux sont modifiés, reste-elle la même UT après ces changements ? Il n'y a pas de réponse unique à cette question, car la définition de ce qui constitue l'identité d'une zone géographique varie d'une nomenclature à une autre. Il s'agit donc d'adopter une approche configurable pour répondre aux différents besoins. Des questions en termes de performance des algorithmes sont aussi à considérer puisque, s'agissant de détecter des changements dans des nomenclatures territoriales, les calculs vont impliquer des géométries à comparer.

A partir de ces trois défis, nous avons fait le constat qu'un système paramétrable était à développer pour que des utilisateurs tels que des agences de statistique puissent facilement, et en fonction des caractéristiques des nomenclatures qu'elles gèrent, créer de nouvelles versions de leur TSN, documenter leurs changements, déverser cette connaissance dans le Web LOD et l'exploiter (référencement, requêtes). Dans ce but, nous avons conçu et développé un framework, nommé Theseus présenté dans la partie suivante.

3 Le framework Theseus

3.1 Vue d'ensemble

Theseus¹⁶ est un framework conçu et développé pour gérer l'ensemble du cycle de vie des TSN : de la modélisation des données à leur exploitation dans le LOD. Lors de la conception de ce framework, nous suivons les recommandations du W3C pour la publication de données liées (Hyland *et al.*, 2014) et l'approche de (Villazón-Terrazas *et al.*, 2011) pour la publication de données gouvernementales sous la forme de données liées. La Figure 1 illustre les cas d'utilisation auxquels Theseus permet de répondre (Bernard, 2019).

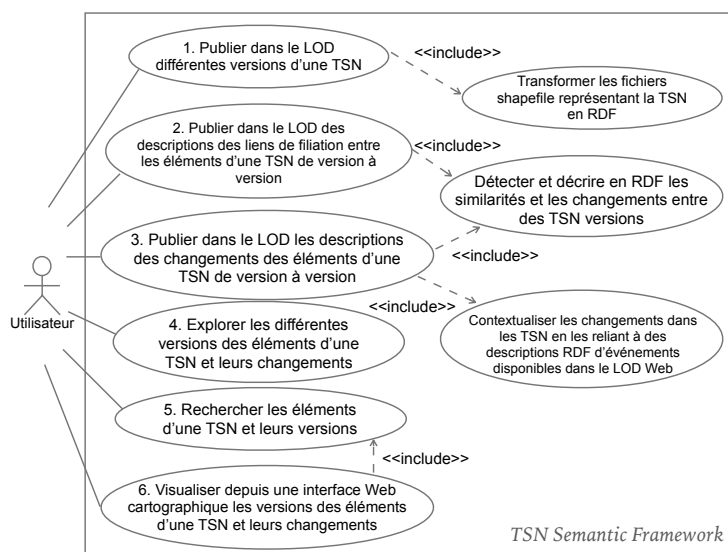


FIGURE 1 – Cas d'utilisation du framework Theseus.

Ce framework automatise la publication dans le LOD de toute TSN et de ses différentes versions, la détection des changements qui impactent leurs unités et la publication de ces changements dans le LOD. Il encapsule deux ontologies que nous avons conçues afin de décrire les TSN et leurs évolutions dans le temps : les ontologies TSN-Ontology et TSN-Change Ontology (voir Section 4). Il encapsule également une implémentation, appelée *TSN Semantic Matching Algorithm* (Bernard *et al.*, 2018a), de l'algorithme de (Plumejeaud *et al.*, 2011) qui détecte et décrit les similitudes et changements entre deux versions consécutives d'une TSN. L'algorithme original a été modifié pour produire des descriptions RDF d'une TSN et de ses changements dans le temps, sur la base des concepts définis dans l'ontologie TSN-Change.

Ce framework sémantique est, à notre connaissance, le premier à assurer la gestion de versions de structures de type TSN. De plus, bien plus que de simplement lier les éléments d'une TSN à travers les versions, Theseus fournit aux utilisateurs des descriptions sémantiques riches des changements des UT au fil du temps.

3.2 Pré-requis et portée

Le Framework Theseus prend en entrée plusieurs fichiers géospatiaux (fichiers au format shapefile ESRI¹⁷), un pour chaque version de TSN, et les transforme en graphes RDF.

16. Ce nom a été choisi en référence à la question philosophique identitaire soulevée par le bateau de Thésée reconstruit entièrement au fil des ans, ses planches ayant été cassées et remplacées les unes après les autres, soulevant alors une question relative à l'identité de ce bateau, préservée ou non alors que le bateau est été entièrement reconstruit ?

17. <https://www.esri.com/library/whitepapers/pdfs/shapefile.pdf>

Concernant la description des changements, qui opère en partie sur la géométrie des UT, une condition préalable pour des résultats optimaux est que toutes les géométries soient de même niveau de généralisation et dans le même référentiel spatial. Cela a en effet un impact sur la pertinence des résultats comme discuté dans (Bernard, 2019).

Dans cet article, nous montrons comment le framework permet de décrire les changements de nature géographique et liés à une nomenclature. Il s'agit donc de changements portant par exemple sur la géométrie, la surface, l'emplacement d'une capitale, l'organisation de la structure spatiale, les informations toponymiques, etc. Plus largement, il est possible de faire évoluer le framework pour permettre la détection de changements liés à d'autres informations, qualifiées de thématiques, qui seraient associées à une UT (telles que son nombre d'habitants, ou tout autre mesure la concernant).

Comme évoqué précédemment la question identitaire est centrale pour qualifier finement le changement. Il est possible de paramétrer au sein du framework une définition de ce qui constitue l'identité des UT de la TSN traitée. Cette définition apparaît sous la forme d'une liste d'attributs d'UT pondérés. Cette définition est établie par un expert de la nomenclature qui détermine d'abord quels attributs doivent être considérés lors de la comparaison des UT de version à version et quelle importance ont ces attributs dans la définition de l'identité des UT, via l'attribution de poids à chacun des attributs de la liste. Ainsi, il est par exemple possible de définir l'identité d'une UT comme étant constituée à 40% de son code dans la nomenclature, à 40% de sa géométrie, à 10% de son nom et à 10% de son unité englobante.

Cette souplesse répond à l'hétérogénéité avec laquelle est définie l'identité dans les nomenclatures. Dans certaines TSN, le nom d'une UT est primordial et si ce nom change, l'UT n'est plus considérée comme la même UT, bien que sa géométrie soit restée la même. Au contraire, dans une autre TSN, bien que le nom ait changé, l'UT est considérée comme "la même UT" car ses frontières sont inchangées. La liste définissant l'identité des UT dans la TSN est un paramètre en entrée d'une fonction de calcul de similarité, implémentée au sein du framework pour calculer automatiquement si l'identité d'une UT est préservée dans la version suivante de la nomenclature.

Deux UT sont considérées comme similaires si la fonction renvoie un score de correspondance supérieur à un seuil global, fixé par l'expert de la nomenclature qui estime le taux de variation admis, permettant de conclure que l'UT persiste malgré ces variations. Cela se traduit alors par un lien de filiation de type *Continuation* entre les deux UT : l'identité de l'UT est conservée dans la version suivante de la nomenclature.

3.3 Architecture globale du framework

Le Framework Theseus est composé de plusieurs modules organisés en quatre niveaux fonctionnels comme illustré par la Figure 2. Les niveaux correspondent aux étapes de la procédure établie par (Villazón-Terrazas *et al.*, 2011) pour la publication de données gouvernementales dans le LOD Cloud :

1. Spécifier et modéliser : au cœur même du Framework Theseus se trouve son modèle de données, composé des ontologies TSN et TSN-Change. Nous avons spécifié ce modèle selon la méthodologie pour la conception d'ontologies exposée dans (Bachimont *et al.*, 2002) : un corpus de TSN et d'ontologies existantes (présenté dans (Bernard, 2019)) nous a aidés à déterminer les concepts et le vocabulaire à utiliser pour être proche des désignations utilisées notamment par les agences produisant des statistiques géocodées. Le modèle ontologique TSN/TSN-Change fournit aux modules du framework une représentation formelle des TSN et de leur évolution dans le temps. Il est brièvement mis en regard de l'état de l'art et expliqué dans la Section 4 .
2. Générer : pour la génération de données RDF à partir des ontologies TSN et TSN-Change, plusieurs modules logiciels sont utilisés. Le logiciel *Geotriples* vise à transformer les fichiers géospatiaux (shapefile) TSN en triplets RDF, en utilisant les concepts définis par l'ontologie TSN. Les modules *TSN Change Detector* et *TSN Change annotator*, que nous avons développés pour implémenter l'algorithme TSN Semantic Matching, détectent les changements dans les géométries et autres attributs, et décrivent

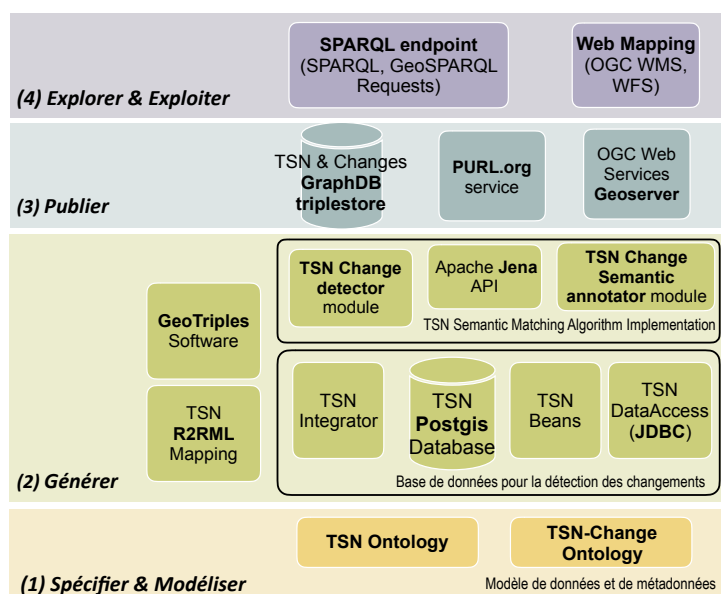


FIGURE 2 – Les modules du framework Theseus.

ces changements en RDF via les concepts de l'ontologie TSN-Change. Ce sont notamment ces modules qui ont permis de produire les connaissances relatives à trois nomenclatures territoriales comme expliqué dans la Section 5.

3. Publier : à cette étape, trois logiciels existants sont utilisés : le service purl.org (pour fournir un URI persistant aux données) ; le triplestore GraphDB (pour publier les données RDF créées) ; le logiciel GeoServer (pour publier les géométries des UT liées à leur représentation LOD).
4. Explorer et exploiter : afin de fournir aux utilisateurs des outils pour explorer et exploiter les données publiées, nous avons mis en place divers moteurs de recherche, en utilisant des mécanismes Web standard : un point d'accès SPARQL pour effectuer des requêtes SPARQL et GeoSPARQL sur les données RDF (disponible depuis l'URL <http://steamerlod.imag.fr/sparql>) ; une interface utilisateur (UI) pour visualiser et interroger les données géospatiales, via des requêtes OGC standard aux Web services WMS et WFS générés par une instantiation du logiciel GeoServer se connectant à la base de données TSN PostGIS de notre framework.

4 Représenter des nomenclatures territoriales et leurs évolutions

4.1 Etat de l'art

La représentation de la dynamique des territoires nécessite de modéliser les entités et les relations spatiales, spatio-temporelles et de filiation entre elles (Del Mondo *et al.*, 2010). La modélisation des entités spatio-temporelles, dont on souhaite capturer l'évolution, repose sur des fondamentaux (voir (Bernard, 2019), Chapitre 3) que nous rappelons brièvement car ils constituent un cadre utile à la compréhension de l'approche décrite dans cet article.

Parler d'évolution requiert de déterminer ce qui fait qu'une entité est toujours ou non la même, dans le temps, donc ce qui constitue son identité. La question de l'identité des entités spatio-temporelles est discutée dans les travaux de (Hornsby & Egenhofer, 2000; Harbelot *et al.*, 2013).

Par ailleurs, des typologies de processus spatio-temporels ((Claramunt & Thériault, 1995; Del Mondo *et al.*, 2010; Plumejeaud *et al.*, 2011; Del Mondo *et al.*, 2013)) introduisent des termes pour la description de changements observables d'entités (comme par exemple, une

fusion d'unités pour en créer une nouvelle, ou à l'inverse une *scission*, donnant lieu à de nouvelles entités à partir d'une autre). Ces travaux définissent également une sémantique pour caractériser les liens de filiation entre entités, selon qu'il s'agit de relation de *Continuation* (une entité subit un changement mais continue d'être cette même entité) ou de *Dérivation* (la modification d'une entité entraîne la perte de son identité, ce n'est alors plus la même entité suite au changement).

Du côté des ontologies de haut niveau, l'ontologie BFO (Basic Formal Ontology (Grenon & Smith, 2004)) offre un cadre théorique pour la représentation d'entités et de leurs évolutions. BFO se structure en deux sous-ontologies exploitées à ces fins : les ontologies pour *continuants* aussi appelées *SNAP* permettant de représenter des entités qui ont une existence continue et une capacité à supporter (persister de façon identique) à travers le temps même en subissant différentes sortes de changements (par exemple, un personne, la planète Terre); et les ontologies pour *occurrents* aussi appelées *SPAN* permettant de décrire des processus ou des événements (par exemple, un sourire, le passage d'une tempête de pluie sur une forêt) (Grenon & Smith, 2004).

Les concepts introduits brièvement ci-dessus se retrouvent dans plusieurs modèles de la littérature visant à représenter l'évolution dans le temps d'unités territoriales. Différentes approches s'inscrivent dans le Web sémantique comme nous le faisons. Elles traitent d'espaces géographiques particuliers comme de parcelles de couverture terrestre (Harbelot *et al.*, 2013, 2015), de régions historiques (Kauppinen & Hyvönen, 2007; Kauppinen *et al.*, 2008), ou encore de domaines juridictionnels ou régions administratives (López-Pellicer *et al.*, 2008; Lopez-Pellicer *et al.*, 2012; Lacasta *et al.*, 2014)). Dans (Bernard, 2019) un état de l'art de ces différentes approches ontologiques est proposé. Il analyse notamment les travaux dont s'est nourrie notre proposition d'un système basé sur les deux ontologies TSN et TSN-Change. Nous rappelons dans la section suivantes leurs grand principes.

4.2 Les Ontologies TSN et TSN-Change

Notre proposition s'inspire tout d'abord des systèmes de gestion de version utilisés dans le développement logiciel et de la façon dont ces systèmes gèrent les changements (Redmond *et al.*, 2008; Völkel & Groza, 2012).

Notre modèle ontologique est, au sens de (Grenon & Smith, 2004), une Trans-Ontologie SNAP-SPAN. Elle décrit les entités spatiales et leur structure, et dépeint leur vie (ou leur histoire) dans le temps, en gérant à cet effet respectivement les vues SNAP et SPAN.

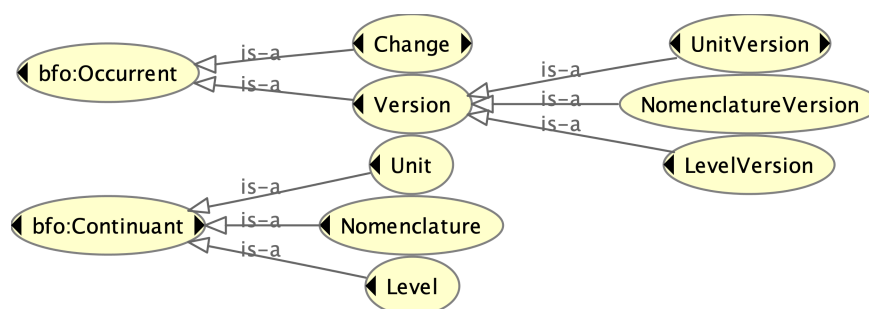


FIGURE 3 – Principaux concepts du Modèle Ontologique TSN hérités de l'ontologie BFO (Concepts Occurrents et Continuants de (Grenon & Smith, 2004)).

D'un côté, une TSN et ses composants sont des entités continues (voir Figure 3) : les concepts de l'ontologie TSN *Unit*, *Nomenclature*, *Level* héritent du concept *Continuant* de BFO. De l'autre, tous les composants versionnés d'une TSN et les changements qu'ils subissent sont des entités occurrentes (qui dépendent des objets continus précédemment cités dont elles sont des versions). Ensemble, les composants versionnés et les changements dépeignent la vie des UT au fil du temps. Ces tranches de vie reposent sur des constructions

4D conformément à l'approche perdurantiste de l'ontologie pour les *fluents* (Welty *et al.*, 2006) aussi exploités dans (Harbelot *et al.*, 2015) et dont nous reprenons les principes (voir Figure 4, *cadre 1*). Cependant, là où les ontologies pour les *fluents* utilisent le terme "tranche de temps" *Slice*, nous préférons utiliser le terme "version", pour être aussi proches que possible du vocabulaire des statisticiens qui constituent des utilisateurs cibles (Figure 4, *cadre 3* présentant notre approche).

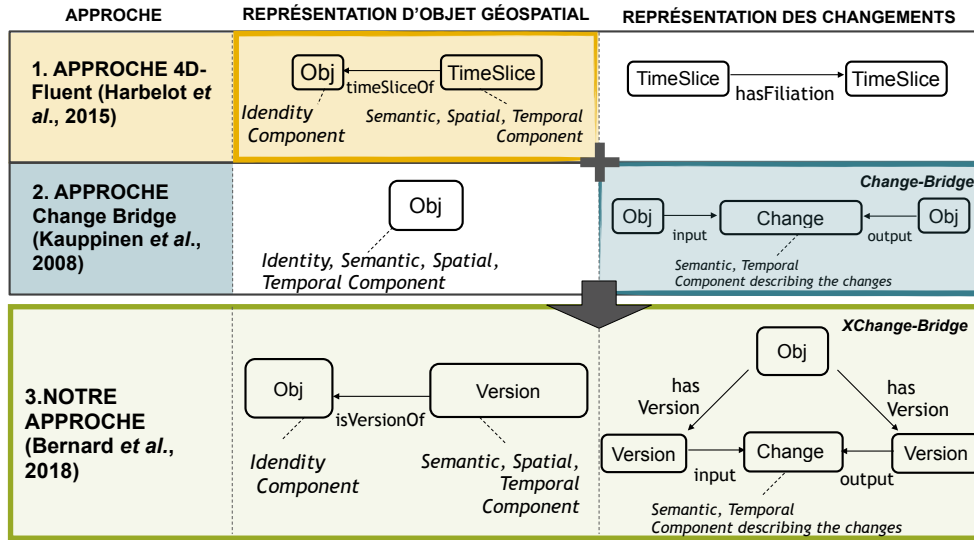


FIGURE 4 – Combinaison d'approches à l'origine du modèle ontologique TSN/TSN-Change

Nous adoptons également l'approche Change Bridges de (Kauppinen *et al.*, 2008), pour gérer l'union des versions successives (Figure 4, *cadre 2*). Au sein de ces ponts de changement (que nous appelons *XChange-Bridges*, *X* pour *eXtended*), nous décrivons les différences entre deux versions et caractérisons la nature des changements territoriaux, en utilisant une typologie des changements basée sur les typologies de (Claramunt & Thériault, 1995) et (Plumejeaud *et al.*, 2011) que nous étendons afin de décrire des changements de frontières des territoires (par exemple, l'élargissement de l'UE). La structure d'un changement territorial *XChange-Bridges* est présentée dans la figure 5. Comme (Plumejeaud *et al.*, 2011), nous considérons qu'un changement est rarement isolé et indépendant des autres changements qui se produisent simultanément au sein des autres unités à l'intérieur d'une zone donnée. Sur la base de cette observation, nous créons des graphes de changement territorial à plusieurs niveaux qui décrivent et relient des changements concomitants qui ont un impact sur plusieurs niveaux de la TSN. Ainsi, nous fournissons aux analystes une représentation détaillée d'un changement territorial dans une TSN, via les informations suivantes :

- la connaissance des versions d'UT impactées. Le prédicat *tsnchange :input* (propriété inverse *tsnchange :before*) pointe vers un *TSNComponent* (*LevelVersion* et *UnitVersion*) qui change; le prédicat *tsnchange :output* pointe vers un *TSNComponent* créé ou modifié après l'événement de changement.
- des liens pour découvrir les conséquences d'un changement. Par exemple, un changement de frontières d'une UT sera relié aux changements affectant ses sous-UT. Le prédicat *tsnchange :lowerChange* (propriété inverse *tsnchange :upperChange*) permet de relier des changements affectant l'élément courant à des changements subis par des sous-élément à l'élément courant.
- une représentation détaillée des causes de ce changement en le reliant à des ressources dans le LOD cloud (événements historiques par exemple). Le prédicat *isCausedBy* indique les raisons contextuelles du changement (à condition que ces informations soient disponibles dans le LOD, dans DBpedia par exemple). Nous recommandons l'utilisation de l'ontologie Linking Open Descriptions of Events (LODE) (Shaw *et al.*,

2009) pour la représentation des événements (historiques) causant les changements des UT.

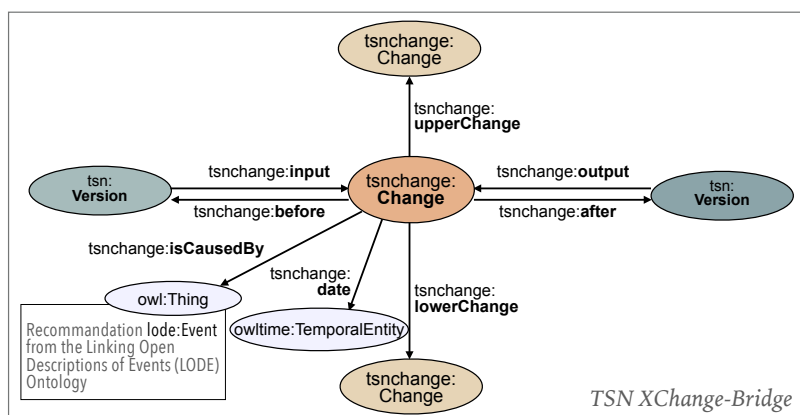


FIGURE 5 – Proposition de structure X-ChangeBridge pour la description de changement territoriaux dans des TSN.

Pour la représentation des propriétés spatiales et temporelles des éléments d'une TSN (par ex., la géométrie des UT), les ontologies TSN et TSN-Change s'appuient sur les ontologies GeoSPARQL (Perry & Herring, 2012) de l'OGC et OWL Time (Cox & Little, 2017) du W3C. Les ontologies TSN et TSN-Change constituent une approche innovante (combinant l'approche Change-Bridge et l'approche des ontologies de flux) pour la représentation de changements territoriaux dans le LOD via des graphes RDF de changements territoriaux multi-niveau.

5 Production de connaissances reversées dans le Web des Données

Certains jeux de données correspondant à des TSN sont publiés dans le Web LOD au format RDF, tels que, par exemple des versions de la nomenclature européenne NUTS. Néanmoins, ces jeux de données sont, soit des versions isolées, ne reflétant pas les différents stades de l'évolution de la nomenclature (le NUTS-RDF Geovocab¹⁸, ou encore une instanciation de l'ontologie Eionet RAMON¹⁹), soit, lorsqu'ils décrivent des changements, ces descriptions ne sont que très partielles (Linked NUTS (Correndo & Shadbolt, 2013)).

Nous avons testé notre approche en l'appliquant à la production de connaissances pour trois TSN :

- La Nomenclature européenne des Unités Territoriales Statistiques (NUTS) (versions 1999, 2003, 2006 et 2010) de l'Institut statistique européen d'Eurostat disponibles dans le Web sous forme de fichiers vectoriels (shapefile ESRI) contenant la liste des UT de chaque version avec leurs attributs (code, nom, niveau, géométrie, ...).
- La nomenclature australienne *Australian Statistical Geography Standard (ASGS)*, élaborée par le Bureau australien des statistiques, qui comprend sept divisions imbriquées du territoire australien, dans les versions 2011 et 2016.
- La nomenclature des unités administratives suisses (SAU) de l'Office fédéral de la statistique, qui décrivent les cantons, districts et communes de Suisse en 2017 et 2018.

Après avoir instancié l'ontologie TSN pour chacune des versions de nomenclatures retenues, nous avons exécuté l'algorithme TSN Semantic Matching Algorithm pour ces différents jeux de données (et donc instancié l'ontologie TSN-Change).

18. <http://nuts.geovocab.org/>

19. <http://rdfdata.eionet.europa.eu/ramon/nuts.rdf> <http://rdfdata.eionet.europa.eu/ramon/nuts2008.rdf> <http://rdfdata.eionet.europa.eu/ramon/nuts2003.rdf>

En ce qui concerne la question de l'identité, ces trois nomenclatures utilisent des définitions similaires. Nous avons néanmoins montré dans (Bernard, 2019), chapitre 10) l'impact du choix des attributs constituant l'identité d'une UT et des poids associés sur la détection et la caractérisation des changements en simulant des définitions d'identité différentes.

Du point de vue des territoires décrits, les deux jeux de données NUTS et SAU, bien qu'ils soient tous deux basés en Europe, diffèrent par la taille des UT qu'ils contiennent. Les plus grandes UT de la NUTS sont les États membres de l'UE, tandis que les plus grandes UT de SAU sont les cantons, ce qui équivaut aux plus petites régions de la NUTS (niveau 3). Les plus petites unités de SAU sont des communes, des unités beaucoup plus petites et plus nombreuses que dans la NUTS. Nous avons fait le choix du troisième jeu de données australien ASGS parce qu'il couvre une région du monde très différente, un vaste territoire dont les géométries des UT sont composées de 4 000 sommets en moyenne (contre une moyenne de 28 sommets dans la NUTS).

Ces nomenclatures nous ont permis d'évaluer la capacité de notre système Theseus à générer des descriptions de filiations et de changements pour des jeux de données de taille importante et de tester son aptitude à calculer, notamment, des distances surfaciques entre deux géométries très précises (donc composées de nombreux sommets) ou de niveaux de généralisation différents (voir (Bernard, 2019), chapitre 10 pour une discussion sur les performances de l'algorithme TSN). Notre programme a ainsi détecté et décrit les similarités et changements entre les versions des trois nomenclatures. Ces connaissances nouvelles ont été versées dans le LOD. Le triplestore Theseus compte 156 162 triplets pour les 4 versions traitées de la NUTS (<http://purl.org/steamer/nuts>), 76 504 pour les 2 versions de SAU (<http://purl.org/steamer/sau>), et 89 974 triplets pour les 2 versions de la nomenclature australienne ASGS (<http://purl.org/steamer/asgs>).

A titre d'exemple, la Figure 6 donne le nombre de changements détectés entre les deux versions 2017 et 2018 de la nomenclature SAU. Notre algorithme détecte automatiquement tous les changements recensés dans la liste officielle qui présente 33 "mutations" (numérotées de 3580 à 3622), incluant chacune des changements tels que des fusions d'unités, des modifications de code ou de noms, etc. Pour chacun de ces changements, des descriptions sémantiquement plus riches (grâce à notre typologie des changements) et plus complètes (grâce au chaînage multi-niveau des changements) sont fournies ce qui explique les 361 changements décrits. Nous avons confronté nos résultats à la liste officielle des changements recensés²⁰ et créé un catalogue confrontant nos résultats aux descriptions textuelles présentes dans la liste officielle²¹.

Matching of the versions	SAU 2017 - 2018
Number of Feature Change	336
Number of GeometryChange	65
Number of NameChange	0
Number of IdentifierChange	30
Number of SubUnitChange	19
Number of SuperUnitChange	7
Number of Structure Change	25
Number of Split	0
Number of Merge	14
Number of Redistribution	1
Number of IdentificationRestructuration	6
Total Number of Change	361

FIGURE 6 – Nombre de changements détectés entre les deux versions 2017 et 2018 de SAU

20. <https://www.bfs.admin.ch/bfs/fr/home/statistiques/catalogues-banques-donnees/publications.assetdetail.4123244.html>

21. http://purl.org/steamer/tsndoc/resources/sau_2017_2018_tsn_change_descriptions.pdf

Les graphes RDF créés constituent des catalogues d'UT et de leurs évolutions. Ils améliorent la compréhension des dynamiques territoriales, en fournissant aux statisticiens des descriptions pour comprendre les motivations et l'impact du redécoupage, et les moyens de référencer précisément les territoires pour lesquels ils produisent des données.

6 Exploitation des graphes créés et génération de nouvelles connaissances

Les graphes RDF créés par notre framework, basés sur le modèle ontologique TSN/TSN-Change, réutilisant lui-même GeoSPARQL et OWL Time, fournissent une représentation des relations spatiales, temporelles et de filiation entre les éléments d'une TSN. Ainsi, en utilisant le module SPARQL endpoint de notre framework <http://steamerlod.imag.fr/>, les utilisateurs peuvent interroger ces trois types de relations.

Nous avons, à titre d'exemples de ce qu'il est possible d'obtenir comme information, défini un ensemble de requêtes SPARQL permettant l'exploration des graphes produits pour les jeux de données NUTS, SAU et ASGS http://purl.org/steamer/tsndoc/resources/tsn_sparql_requests.pdf. Nous illustrons simplement ici par un exemple la capacité offerte par notre approche à produire de la connaissance.

Ainsi, la requête présentée par la figure 7 permet de reconstruire la ligne de vie d'une unité territoriale (ici l'unité ayant pour code ES63 dans la nomenclature NUTS) en parcourant les 4 versions successives de cette nomenclature que nous avons instanciées et pour lesquelles nous avons généré les liens de filiation et documenté les changements. La Figure 8 montre une représentation graphique automatiquement générée à partir du graphe résultat de la requête. Les nœuds rouges de la ligne supérieure représentent les versions successives de l'UT de code ES63 dans la NUTS. Le nœud étiqueté `nuts:V2003_L2_ES64`, représenté en rouge également, est un nœud en sortie du changement de type *extraction*, selon notre terminologie, qui a eu lieu entre les versions 1999 et 2003 de la NUTS. Ce changement a donné lieu à des modifications de la géométrie et du nom de l'UT ES63. Comme le montre cette vue, il s'agit du seul changement qui a affecté l'UT ES63 au cours de sa vie depuis 1999, dénotant une certaine stabilité territoriale pour cette unité (au moins jusqu'à la version NUTS 2010).

```
1 PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
2 PREFIX tsnchange: <http://purl.org/net/tsnchange#>
3 PREFIX tsn: <http://purl.org/net/tsn#>
4 CONSTRUCT { ?TU_input tsnchange:inputUnitVersion ?change ;
5             tsnchange:hasNextVersion ?TU_output .}
6 FROM <http://purl.org/steamer/nuts> WHERE{{
7     ?TU_input tsnchange:inputUnitVersion ?change .
8     ?change tsnchange:unitVersionAfter ?TU_output .
9     ?TU_input tsn:hasIdentifier "ES63".}
10 UNION {?TU_input tsnchange:hasNextVersion ?TU_output .
11         ?TU_input tsn:hasIdentifier "ES63".}}
```

FIGURE 7 – Requête SPARQL retournant la ligne de vie d'une unité territoriale, ici l'unité ES63 telle que codifiée dans la NUTS

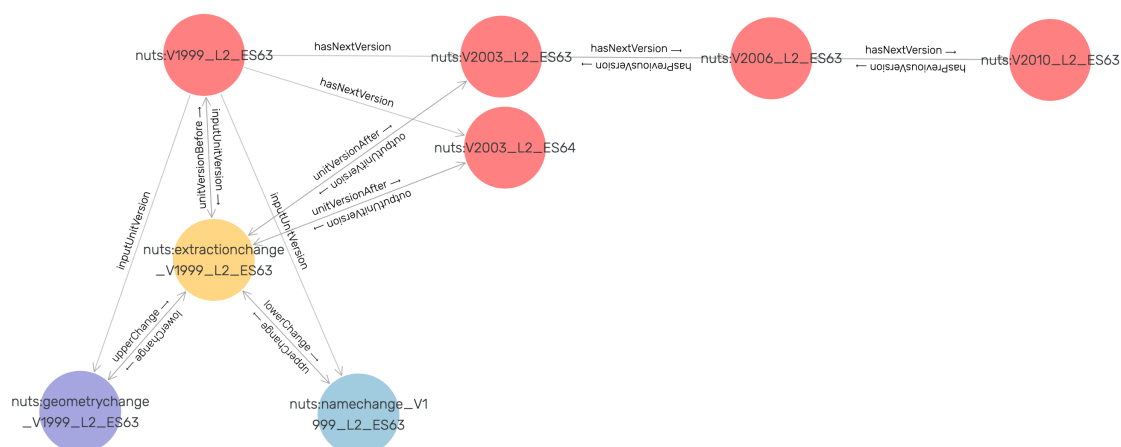


FIGURE 8 – Visualisation de la ligne de vie de l'UT ES63 (résultat de la requête précédente), visible aussi à http://purl.org/steamer/nuts_ES63_lifeline

Cette capacité à reconstituer la ligne de vie de chaque UT dans le temps par une lecture horizontale des graphes, en suivant la même entité à travers le temps, à travers les versions, est complétée par la capacité à parcourir les graphes selon une lecture verticale : nous avons créé d'autres requêtes qui fournissent une représentation de la propagation d'un changement à travers les niveaux de la nomenclature. Un exemple de requête est donné dans (Bernard, 2019), page 172), et le résultat associé visible depuis http://purl.org/steamer/nuts_V1999_ES6_change_graph.

En exploitant les données disponibles dans le LOD, il est possible d'enrichir les graphes générés, par exemple en recherchant dans d'autres graphes de connaissances tels que DBPedia ou Wikidata (le versant LOD de Wikipedia) des informations complémentaires (telles que le nom des gouvernants, le nombre d'habitants, voire (et surtout) une explication des changements...) sur les unités territoriales à partir de leur nom. En outre, on constate, dans ces graphes de données encyclopédiques, la quasi absence de représentation ontologique des changements territoriaux au cours du temps. En effet, s'il existe bien une classe DBPedia pour la représentation d'événements (<http://dbpedia.org/ontology/Event>), les événements décrits sont essentiellement des événements de type bataille militaire ou événement sportif (Hienert & Luciano, 2015). De fait, lorsque les changements subis par des régions sont décrits, ils ne le sont qu'à travers des champs de texte libre (tels que `dbo:abstract`), donc non structurés et non directement exploitables via des requêtes SPARQL. Ceci milite pour un référencement croisé des graphes de connaissances dédiées à l'évolution des unités territoriales statistiques proposés par notre approche, et des graphes de connaissances plus générales et grand public tels que ceux accessibles via DBPedia.

7 Conclusion et Perspectives

Dans cet article nous avons présenté une solution conceptuelle basée sur deux ontologies (TSN et TSN-Change), et logicielle (le framework Theseus, incluant notamment un algorithme, nommé TSN Semantic Matching Algorithm) pour représenter, dans le Web Sémantique, des connaissances relatives aux évolutions des territoires décrits par des nomenclatures territoriales statistiques. Notre approche est suffisamment générique pour couvrir tous les cas de structure hiérarchique de type TSN. En utilisant notre approche, les agences statistiques peuvent créer des unités spatiales déversables dans le LOD sous la forme de ressources référençables. Cela permet l'association précise de leurs données statistiques aux unités du territoire concernés qui sont désormais disponibles dans toutes les versions de la nomenclature. Des liens de filiation sont établis entre versions et nœuds de changements, contribuant à une meilleure compréhension des dynamiques territoriales. Le modèle intègre enfin le moyen de décrire et relier des changements qui affectent plusieurs niveaux de la TSN pour en donner une vision plus précise et plus complète, selon une approche ascendante ou descendante.

Nos descriptions des changements territoriaux sont générées automatiquement grâce à l'algorithme *TSN Semantic Matching Algorithm*, capable de s'adapter aux spécificités des TSN décrites, notamment sur la question de l'identité des unités territoriales.

Parmi les perspectives de ce travail, nous prévoyons d'étudier l'applicabilité de la démarche à d'autres types de nomenclatures utilisées notamment dans la description de zones médico-sociales, de zones trans-frontalières, de zones d'emploi, de zones urbaines. Un des points à explorer concerne la question de la modélisation de hiérarchies différentes des TSN car elles peuvent être non couvrantes (certaines unités peuvent avoir une ou plusieurs unités supérieures à un niveau territorial qui n'est pas immédiatement supérieur), non strictes (une UT peut avoir plusieurs UT englobantes), non hiérarchiques, etc.

Du côté de la restitution des connaissances, nous travaillons à la création d'un outil de visualisation des changements. Des outils logiciels spécialisés dans la gestion de version tels que *GeoGIG* ou le projet *GitHub Inc.* fournissent un moyen de visualiser les différences de géométries, UT par UT (Negretti, 2015; Boundless, 2014b,a). Néanmoins, ils ne permettent pas de regrouper des changements qui affectent plusieurs objets géographiques en même temps (par exemple, une fusion de deux UT), ou d'adjoindre une sémantique donnant le contexte et la nature de ce changement territorial. Notre objectif est de doter le framework *Theseus* d'un module de géovisualisation des changements intégrant les apports et originalités de l'approche de représentation des connaissances territoriales évolutives basée sur TSN et TSN-Change.

Références

- ATKINSON R. (2017). QB4ST : RDF Data Cube extensions for spatio-temporal components.
- BACHIMONT B., ISAAC A. & TRONCY R. (2002). Semantic commitment for designing ontologies : a proposal. In *International Conference on Knowledge Engineering and Knowledge Management*, p. 114–121 : Springer.
- BERNARD C. (2019). *Immersing evolving geographic divisions in the semantic Web*. PhD thesis, Université Grenoble Alpes.
- BERNARD C., PLUMEJEAUD-PERREAU C., VILLANOVA-OLIVER M., GENSEL J. & DAO H. (2018a). An ontology-based algorithm for managing the evolution of multi-level territorial partitions. In *Proceedings of the 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, SIGSPATIAL '18*, p. 456–459, New York, NY, USA : ACM.
- BERNARD C., VILLANOVA-OLIVER M., GENSEL J. & DAO H. (2018b). Modeling changes in territorial partitions over time : Ontologies tsn and tsn-change. In *Proceedings of the 33rd Annual ACM Symposium on Applied Computing, SAC '18*, p. 866–875 : ACM.
- BERNARD, CAMILLE, VILLANOVA-OLIVER, MARLENE, GENSEL, JÉRÔME & DAO, HY (2018). Ontologies pour représenter l'évolution des découpages territoriaux statistiques. *Rev. Int. Geomat.*, **28**(4), 409–437.
- BOUNDLESS (2014a). *GeoGit in Action : Distributed Versioning Geospatial Data* - Boundless.
- BOUNDLESS (2014b). *Introducing Versio : Version Control for Spatial Data* - Boundless.
- CLARAMUNT C. & THÉRIAULT M. (1995). Managing Time in GIS An Event-Oriented Approach. In C. J. VAN RIJSBERGEN, J. CLIFFORD & A. TUZHILIN, Eds., *Recent Advances in Temporal Databases*, p. 23–42. London : Springer London. DOI : 10.1007/978-1-4471-3033-8_2.
- CORRENDO G. & SHADBOLT N. (2013). Linked nomenclature of territorial units for statistics. *Semantic Web*, **4**(3), 251–256.
- COX S. & LITTLE C. (2017). Time Ontology in OWL - W3C Recommendation 19 October 2017.
- CYGANIAK R. & REYNOLDS D. (2014). The RDF Data Cube Vocabulary.
- DEBRUYNE C., MEEHAN A., CLINTON É., MCNERNEY L., NAUTIYAL A., LAVIN P. & O'SULLIVAN D. (2017). Ireland's authoritative geospatial linked data. In *International Semantic Web Conference*, p. 66–74 : Springer.
- DEL MONDO G., RODRÍGUEZ M., CLARAMUNT C., BRAVO L. & THIBAUD R. (2013). Modeling consistency of spatio-temporal graphs. *Data & Knowledge Engineering*, **84**, 59–80.
- DEL MONDO G., STELL J. G., CLARAMUNT C. & THIBAUD R. (2010). A graph model for spatio-temporal evolution. *J. UCS*, **16**(11), 1452–1477.
- ETCHEVERRY L. & VAISMAN A. A. (2012). QB4OLAP : a new vocabulary for OLAP cubes on the semantic web. *Proceedings of COLD*.

- GRENON P. & SMITH B. (2004). SNAP and SPAN : Towards Dynamic Spatial Ontology. *Spatial Cognition & Computation*, **4**(1), 69–104.
- HARBELOT B., ARENAS H. & CRUZ C. (2013). Continuum : A spatiotemporal data model to represent and qualify filiation relationships. In *Proceedings of the 4th ACM SIGSPATIAL International Workshop on GeoStreaming*, p. 76–85 : ACM.
- HARBELOT B., ARENAS H. & CRUZ C. (2015). LC3 : A spatio-temporal and semantic model for knowledge discovery from geospatial datasets. *Web Semantics : Science, Services and Agents on the World Wide Web*, **35**, 3–24.
- HIENERT D. & LUCIANO F. (2015). Extraction of Historical Events from Wikipedia. In E. SIMPERL, B. NORTON, D. MLADENIC, E. DELLA VALLE, I. FUNDULAKI, A. PASSANT & R. TRONCY, Eds., *The Semantic Web : ESWC 2012 Satellite Events*, volume 7540, p. 16–28. Berlin, Heidelberg : Springer Berlin Heidelberg.
- HORNSBY K. & EGENHOFER M. J. (2000). Identity-based change : a foundation for spatio-temporal knowledge representation. *International Journal of Geographical Information Science*, **14**(3), 207–224.
- HYLAND B., ATEMEZING G. A. & VILLAZÓN-TERRAZAS B. (2014). Best Practices for Publishing Linked Data.
- KAUPPINEN T. & HYVÖNEN E. (2007). Modeling and reasoning about changes in ontology time series. In *Ontologies*, p. 319–338. Springer.
- KAUPPINEN T., VÄÄTÄINEN J. & HYVÖNEN E. (2008). *Creating and using geospatial ontology time series in a semantic cultural heritage portal*. Springer.
- LACASTA J., LOPEZ-PELLICER F. J., FLORCZYK A., ZARAZAGA-SORIA F. J. & NOGUERAS-ISO J. (2014). Population of a spatio-temporal knowledge base for jurisdictional domains. *International Journal of Geographical Information Science*, **28**(9), 1964–1987.
- LÓPEZ-PELLICER F. J., FLORCZYK A. J., LACASTA J., ZARAZAGA-SORIA F. J. & MURO-MEDRANO P. R. (2008). Administrative units, an ontological perspective. In *Advances in Conceptual Modeling—Challenges and Opportunities*, p. 354–363. Springer.
- LOPEZ-PELLICER F. J., LACASTA J., FLORCZYK A., NOGUERAS-ISO J. & ZARAZAGA-SORIA F. J. (2012). An ontology for the representation of spatiotemporal jurisdictional domains in information retrieval systems. *International Journal of Geographical Information Science*, **26**(4), 579–597.
- NEGRETTI M. (2015). Operation-based revision control for geospatial data sets. p.15.
- OPENSHAW S. (1984). The modifiable areal unit problem. *Concepts and Techniques in Modern Geography*, **38**.
- PERRY M. & HERRING J. (2012). OGC GeoSPARQL - A Geographic Query Language for RDF Data. p.75.
- PLUMEJEAUD C. (2011). *Modèles et méthodes pour l'information spatio-temporelle évolutive*. PhD thesis, Université de Grenoble.
- PLUMEJEAUD C., MATHIAN H., GENSEL J. & GRASLAND C. (2011). Spatio-temporal analysis of territorial changes from a multi-scale perspective. *International Journal of Geographical Information Science*, **25**(10), 1597–1612.
- REDMOND T., SMITH M., DRUMMOND N. & TUDORACHE T. (2008). Managing Change : An Ontology Version Control System. p.10.
- SHAW R., TRONCY R. & HARDMAN L. (2009). LODE : Linking Open Descriptions of Events. *ASWC*, **9**, 153–167.
- VILLAZÓN-TERRAZAS B., VILCHES-BLÁZQUEZ L. M., CORCHO O. & GÓMEZ-PÉREZ A. (2011). Methodological Guidelines for Publishing Government Linked Data. In D. WOOD, Ed., *Linking Government Data*, p. 27–49. New York, NY : Springer New York.
- VÖLKEL M. & GROZA T. (2012). SemVersion : an RDF-based ontology versioning system. p.9.
- WELTY C., FIKES R. & MAKARIOS S. (2006). A reusable ontology for fluents in OWL. In *FOIS*, volume 150, p. 226–236.