

De l'étude des nomenclatures territoriales à la modélisation des dynamiques des territoires administratifs en France

Christine Plumejeaud-Perreau, Pascal Cristofoli, Claude Motte

▶ **To cite this version:**

Christine Plumejeaud-Perreau, Pascal Cristofoli, Claude Motte. De l'étude des nomenclatures territoriales à la modélisation des dynamiques des territoires administratifs en France. Revue Internationale de Géomatique, 2015, Modéliser les dynamiques spatiales, 25 (3), pp.355-392. 10.3166/RIG.25.355-392 . hal-01228634

HAL Id: hal-01228634
https://hal.science/hal-01228634

Submitted on 15 Dec 2023

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.



Distributed under a Creative Commons Attribution - ShareAlike 4.0 International License

De l’étude des nomenclatures territoriales à la modélisation des dynamiques des territoires administratifs en France.

Christine Plumejeaud[1], Pascal Cristofoli[2], Claude Motte[2]

1. *Laboratoire LIENSs, UMR CNRS 7266*
   2 rue Olympe de Gouges, 17000 La Rochelle, France
   christine.plumejeaud-perreau@univ-lr.fr

2. *LaDéHiS - Laboratoire de démographie et d'histoire sociale*
   Centre de recherches historiques (UMR 8558 EHESS-CNRS)
   Ecole des hautes études en sciences sociales
   190 avenue de France, 75013 Paris
   motte@ehess.fr, cristofo@ehess.fr

RESUME. Ce chapitre est le fruit d'une collaboration interdisciplinaire entre informaticiens, géomaticiens et historiens pour définir une nomenclature territoriale et modéliser ses évolutions dans un système d'information qui soit opérationnel. Les nomenclatures territoriales sont des structures hiérarchiques ou semi-hiérarchiques composées d'entités territoriales formant des imbrications de maillages en constante évolution. Le problème du suivi des unités au cours du temps a été largement abordé dans de précédents travaux, il est connu comme le "split tract problem" (Howenstine, 1993). Il est ici abordé sous un jour nouveau, en traitant également de l'évolution des relations d'appartenance entre unités, tout en discutant de la représentativité de ces entités pour les sociétés et populations qui les définissent et les habitent. Les enjeux attachés à ce problème sont décrits, ainsi que le vocabulaire et les concepts afférents, notamment le concept d'identité. La discussion est assortie d’exemples issus du Code Officiel Géographique français ou la Nomenclature des Unités Territoriales Statistiques européennes (NUTS). Enfin, un cas d'étude vient illustrer la modélisation proposée pour la constitution d’une géonomenclature historique des entités territoriales françaises.

ABSTRACT. This chapter is the result of an interdisciplinary collaboration between computer scientists, GIS specialists and historians to define a territorial nomenclature and model its evolution in an information system that is operational. Territorial classifications are hierarchical structures or semi-hierarchical made territorial entities forming nested meshes evolving. The problem of tracking units over time has been widely discussed in previous works, it is known as the "split tract problem" (Howenstine, 1993). It is discussed here in a new light, while also dealing with the development of relations between units belonging, while discussing the representativeness of these entities for governments and people that define them and live in. Issues attached to this problem are described, as well as the vocabulary and concepts thereon, including the concept of identity. The discussion is accompanied by

examples from the French Official Geographical Code (COG) and the European Nomenclature of Territorial Units for Statistics (NUTS). Finally, a case study illustrates the proposed model for the formation of a historical geonomenclature of French territorial entities

MOTS-CLES : Nomenclature territoriale – unité territoriale – maillage – changement – démographie – identité – modélisation – généalogie – métadonnées – imperfection des sources.

KEYWORDS: Territorial nomenclature – territorial unit – mesh – zoning – split tract problem – demography – identity – modelling – genealogy – metadata –imperfect data sources

DOI:10.3199/JESA.45.1-n © Lavoisier 2012 AR_DOI

1. Introduction

Ce chapitre traite des transformations dans le temps des nomenclatures territoriales, qui sont à la fois des outils pour la gouvernance de l'espace, mais aussi un reflet d'une organisation humaine de l'espace, anisotrope par nature, et en constante évolution. En effet, ces nomenclatures définissent différents *niveaux*, ou *partitionnements*, de l'espace géographique étudié, qu'on appellera *maillage* ou *zonage*, et ces partitions sont complètes ou bien incomplètes (dans ce second cas, la partition n'a pas obligation à couvrir complètement l'espace). Les *mailles* ou *zones* ou *entités territoriales* qui composent chaque subdivision sont à la fois des parcelles de territoire et des territoires à part entière qui entretiennent entre eux des relations hiérarchiques (c'est-à-dire de gouvernance) mais également de compétition. L'évolution de ces relations se traduit et peut être analysée à travers les restructurations territoriales.

Ces *entités* ne seraient que fictives au sens que lui confère Barry Smith, (Smith, 1999), (ce sont des "*fiat objects*"), puisque leurs délimitations, et leurs attributs (noms, centre, statut) sont le produit d'une intention humaine incarnant un pouvoir. Elles sont porteuses d'éléments statistiques associés à des documents signifiants pour l'histoire économique, sociale et politique des territoires sous-jacents, comme par exemple les registres de naissance, mariage et décès attachés aux communes. Au fil du temps, elles ne cessent d'être redéfinies. Il devient alors difficile de rendre compte de l'histoire, de retracer les évolutions d'un territoire. Ce problème aussi connu sous le nom de *split-tract problem* (Howenstine, 1993) a fait déjà l'objet de nombreuses études considérant essentiellement un niveau de découpage évoluant au cours du temps. On pourra se référer par exemple aux travaux de (Franck, 2001), concernant la vie des entités dites "socio-économiques" ou ceux de (Gregory, 2002) qui établissent la nécessité d'établir des dictionnaires de changement des unités territoriales. Il s'agit ici de réfléchir à la façon dont nous pouvons également rendre compte des imbrications complexes et des évolutions de ces découpages structurés en nomenclatures qui co-existent et évoluent elles aussi dans le temps, dans une approche interdisciplinaire combinant les regards et analyses croisées d'historiens, de géographes, et de statisticiens. Notre propos est illustré avec des exemples

extraits d'une part de la Nomenclature des Unités Territoriales Statistiques (NUTS) pour l'espace européen, et d'autre part de l'histoire de la structuration administrative de l'espace du territoire français.

Dans un premier temps, l'intérêt d'une reconstitution des évolutions des nomenclatures territoriales est exposé en détail. Une définition complète de l'objet d'étude est proposée afin de mieux saisir les enjeux et difficultés de cette problématique. Dans un second temps l'état de l'art décrit et compare les approches qui traitent de ce sujet, notamment à travers les concepts d'identité, de processus et d'événements territoriaux. Dans un troisième temps, nous discutons des avantages, des inconvénients et de l'expérimentation de plusieurs modèles informatiques basés sur des concepts précédemment exposés à partir d'un cas d'étude qui porte sur la reconstitution de l'histoire des communes. Enfin, nous concluons par une synthèse et une mise en perspective de ces diverses expériences, pour proposer des recommandations en ce qui concerne le développement de modèles informatiques plus proches de la complexité des phénomènes qui sont ici modélisés.

2. Enjeux liés à l'étude de l'évolution de nomenclatures territoriales

L'objectif est d'expliquer l'intérêt de comprendre les dynamiques territoriales observées sous le prisme des évolutions des nomenclatures. Pour cela nous définissons ce que le terme de nomenclature sous-tend, quels sont ses éléments structurants (les maillages, les zonages, les entités territoriales, les relations de gouvernance, etc.). Cette définition amène plusieurs éléments de compréhension de la dynamique des entités territoriales, et permet d'aborder les enjeux liés à cette question dans un second temps.

2.1. Les nomenclatures territoriales comme objet d'étude

D. Pumain et T. Saint-Julien (1997) définissent l'espace géographique comme l'ensemble de lieux et de relations entre les lieux, elles-mêmes définies par les interactions entre des acteurs sociaux localisés. Il est produit par l'organisation des sociétés et de la nature et, en retour, il est l'agent du maintien et du développement des sociétés sur leur territoire :

> *« Comme pour l'espace de la physique, il y a deux manières d'envisager l'espace géographique pour le concevoir et le représenter :*
>
> *- soit comme un simple contenant, repère immuable dans lequel on situe les objets et analyse leurs relations : c'est l'espace-support ou espace absolu. Il est isotrope et homogène, c'est-à-dire qu'il a les mêmes propriétés dans toutes les directions ;*
>
> *- soit comme un ensemble de relations dont les propriétés, variables dans le temps et dans l'espace, sont définies par la nature et la forme des interactions entre les objets et les unités spatiales : c'est l'espace relatif ou espace-produit. »* (Pumain & Saint-Julien, 1997)

Les découpages territoriaux de l'espace géographique sont souvent considérés comme l'espace support nécessaire à la collecte des informations statistiques et/ou documentaires sur un territoire à un moment donné. Cependant, force est de constater que ces découpages ne sont pas immuables, ce qui invite à mieux définir la nature même de ces objets et des déterminants de leur production.

2.1.1. Les découpages territoriaux : maillages, zonages, et subdivisions

Les découpages territoriaux reflètent une organisation humaine de l'espace en constante évolution, mais aussi *anisotrope* par nature. L'espace est marqué par des similarités, des pôles d'attractivité et des zones de discontinuités, culturels, linguistiques, économiques… (Grasland, 2002). Ce problème est développé par (Brunet, François et Grasland, 1997). Cette anisotropie peut aussi être expliquée par les sciences sociales du comportement, qui décrivent les processus de compétition ou de communication entre les humains, processus qui s'inscrivent à la fois dans le temps et dans l'espace. Ainsi, dans une tentative de définition d'un nouveau paradigme pour l'organisation spatiale des sociétés, (Villeneuve, 1972) rapproche la géographie du comportement de l'analyse spatiale à travers le concept de proximité optimale, qui s'explique comme la réponse au principe général de survie et aux contraintes que l'espace invariant définit. Les décisions individuelles ou collectives seraient un compromis entre les tensions centrifuges (les processus de compétition, au niveau biotique) et les tensions centripètes (les processus de communication, au niveau culturel) qui engendrent le principe fondamental d'organisation des formes spatiales. Lorsque l'étude des formes spatiales porte sur les *zonages*, il apparaît qu'en réalité leur forme est étroitement liée au peuplement, et qu'ils matérialisent les frontières et les zones de discontinuités (linguistique, culturelle, économique) que l'occupation humaine produit. Cette observation s'applique aussi dans le cas de l'étude des *maillages*, (Grasland, 1998), qui sont une forme particulière de zonage, définissant une partition complète et totale de l'espace. Les zonages (ou *subdivisions* territoriales) ne sont donc en aucun cas une représentation de l'espace isotrope ou homogène. Nous envisageons donc ici un découpage territorial comme un *espace-produit*.

2.1.2. Niveaux, nomenclatures et relations de gouvernance

Cet espace-produit est lui-même modélisé et analysé selon plusieurs *niveaux* de granularité : l'échelle spatiale peut être comprise comme une échelle de représentation, ou bien comme un niveau d'observation et d'analyse. C'est cette seconde acception qui nous intéresse ici car elle sous-tend l'usage d'une structuration de l'espace en niveaux hiérarchiques. (Claval, 1968) a montré que combiner les échelles favorise notre compréhension de l'espace terrestre ou zonal. Toute valeur en un point donné est fonction de processus agissant à différents niveaux, locaux, régionaux et internationaux (par exemple, la localisation d'une entreprise industrielle). La compréhension d'une partie de l'espace terrestre passe obligatoirement par l'examen de niveaux différents et imbriqués. C'est l'objet de la théorie formelle des systèmes hiérarchiques qui postule qu'un espace géographique

forme un système dont les composantes sont reliées par des relations dissymétriques donc non égalitaires (Takahara et al., 1980).

Cette théorie est largement développée, et, par exemple, dans le domaine des systèmes d'information géographiques, elle amène à produire une formalisation des partitions de l'espace, très utile pour modéliser les hiérarchies spatiales (Rigaux et Scholl, 1995). Ce formalisme s'applique lorsque les niveaux s'imbriquent pour former une hiérarchie stricte, c'est-à-dire lorsque toute unité (exceptée la racine) possède une et une seule unité supérieure. Ce formalisme offre une formalisation des opérations d'agrégation sur les attributs attachés aux zonages. En effet, lorsque les attributs sont des effectifs dénombrables, comme par exemple le décompte de la population ou de logements, ils ont la propriété de pouvoir s'additionner pour calculer l'effectif de la maille supérieure. Ainsi, la population d'un département est, en théorie[1], la somme des populations des communes qui lui sont rattachées. Il apparaît que la NUTS (Nomenclature des Unités Territoriales Statistiques), établie par l'Union Européenne, depuis 2003 (Parlement européen, 2003), utilisée par les instituts statistiques nationaux ou européens (Eurostat), est effectivement une hiérarchie spatiale stricte, pour laquelle ce type de formalisme est très utile, notamment pour reconstituer par agrégation les valeurs d'attributs à des niveaux supérieurs à celui de leur mesure effective.

Ces relations d'agrégation matérialisent des relations de gouvernance : il s'agit de savoir, par exemple, à qui une entité territoriale reverse ses impôts, et/ou si elle connaît plusieurs tutelles simultanément. Grâce au suivi de ces changements d'appartenance, il est possible d'analyser des causes et des effets de conflits, politiques et/ou économiques.

Les *nomenclatures* sont une matérialisation de ce système hiérarchique, et se présentent sous la forme d'un ensemble de découpages territoriaux[2] organisés de façon hiérarchique. Chaque nomenclature est non seulement un outil d'analyse scientifique mais aussi un outil d'action de gouvernance (Guibert, 1971), que celle-ci soit du domaine du religieux, de l'administratif ou de l'économique.

2.1.3. Diversités des nomenclatures, observation d'appartenances multiples

Les niveaux ne sont pas nécessairement strictement emboîtés, ni ne forment une partition complète et totale de l'espace. Lorsque, par exemple, un historien s'intéresse à la nomenclature française des unités fiscales à l'époque Moderne, selon les dictionnaires de Saugrain (1709 et 1720), il ne liste pas moins de 23 subdivisions de l'espace de niveau 2 (bailliages, comtés, diocèses civils, duchés, élections, etc.) qui cohabitent ensemble pour constituer le Royaume de France (niveau 1), et 38

---

[1] En réalité, pour des niveaux de recensement différents, le nombre d'individus dénombrés sur une même aire peut diverger. Par exemple, les populations sans domicile fixe peuvent être localisées au niveau supérieur (le département ou l'état), sans avoir été décomptées dans les ménages des communes.

[2] Ils sont aussi appelées divisions ou subdivision (Motte & Pelissier, 2003) ou bien encore maillage (Zadora-Rio, 2008).

subdivisions de l'espace de niveau 3, (dont les paroisses, précurseurs des communes actuelles (Chareille, Rodier et Zadora-Rio, 2004). Il observe que, d'une part, ces subdivisions ne forment pas des partitions complètes et couvrantes de l'espace (l'ensemble des duchés ne couvre pas la France) et que, d'autre part, les entités entretiennent des relations d'appartenance multiples avec les unités de niveau supérieur. Le *patchwork* formé à cette époque par les découpages territoriaux a été remplacé en 1790, dès les premiers mois de la révolution française, et c'est cette nouvelle organisation strictement hiérarchisée des unités administratives qui a été codifiée par l'INSEE, un siècle et demi plus tard en 1943 dans le Code Officiel Géographique (COG) (Lang, 2010). Stable durant plus de deux siècles, cette hiérarchie est remise en cause par les récentes recompositions territoriales locales qui ont eu lieu en France : La décennie 1990 a institué les EPCI (Établissement Public de Coopération Intercommunale) à fiscalité propre : les communautés de communes (pour l'espace rural), les communautés d'agglomération (villes de plus de 15000 habitants et bassins de 50 000), et les communautés urbaines (pour les plus grandes villes). Les lois sur l'aménagement du territoire ont ensuite créé des « territoires de projet », développés parallèlement aux EPCI, et constituant un second niveau, moins formel, d'intercommunalité, les Pays et les Agglomérations.

Ces dernières recompositions procèdent de justifications multiples. La principale a trait au morcellement du maillage communal français, dont l'inadaptation aux évolutions des structures socio-spatiales contemporaines est de plus en plus problématique. Il semble en effet nécessaire de rationaliser la gestion territoriale, en construisant des territoires plus vastes que la commune, d'une taille suffisante pour atteindre une masse critique en matière de population et donc de potentialités. Le choix et le succès de l'intercommunalité tient, en France, aux successifs échecs des tentatives de réformes du maillage communal.

Globalement, même si ce n'est pas explicite dans les lois, une forme de hiérarchie et d'emboitement spatial se dessine : un premier niveau, constitué par les EPCI ; un second, constitué par le Pays, fédérant plusieurs EPCI. A terme, l'idée du législateur est de mailler l'ensemble du territoire français en EPCI et en Pays. La réussite numérique de ces nouvelles structures permet d'approcher cet objectif. Mais, d'une part, l'avenir des Pays demeure incertain dans le contexte du débat actuel sur les réorganisations territoriales, et, d'autre part, la création de ces maillages intercommunaux est progressive, puisque les différentes structures ne sont pas obligatoires et reposent pour partie sur l'initiative des acteurs locaux. En conséquence, il existe un certain nombre de cas – parfois transitoires – ne respectant pas un emboitement hiérarchique strict (Plumejeaud et al., 2011-b).

2.2 Enjeux

En résumé, les différents maillages territoriaux, qu'ils soient administratifs, fiscaux, religieux, s'inscrivent dans des nomenclatures, et sont l'illustration de relation de gouvernance et d'une réalité sociale, économique, et démographique en

perpétuelle évolution. A l'intérieur de ces nomenclatures, les entités qui composent les mailles ou zones de ces subdivisions entretiennent entre elles des relations multiples d'appartenance, de gouvernance. Retracer les transformations de ces nomenclatures, c'est donc à la fois retranscrire les recompositions territoriales "horizontales" au sein d'un même niveau de subdivision au cours du temps, mais également retranscrire les modifications des relations "verticales" entre entités territoriales. Examiner ces transformations spatiales dans une perspective historique permet d'envisager l'étude des dynamiques spatiales des sociétés, y compris dans ses aspects les plus pragmatiques liés aux compétitions de pouvoir ou à la collecte d'informations statistiques ou historiques.

L'établissement d'un référentiel géographico-administratif faciliterait l'indexation d'archives et autres documents (photographie) ainsi que leur contenu, en les liant à une entité à telle période. Par exemple, lorsqu'une source délivre des informations sur la collecte ou le montant d'une imposition dans le cadre d'un « bailliage », il est utile de rattacher ces informations à l'entité « bailliage » relevant de la division fiscale et non au « bailliage » de la division judiciaire, utile aussi de connaître le niveau de cette entité dans la nomenclature des subdivisions financières de l'époque moderne ce qui sera possible s'il existe une codification claire et unique de cette entité.

Il existe un certain nombre d'homonymie au niveau des entités territoriales de base (par exemple trois communes sont dénommées "Fontaine" en France en 2012), ainsi que dans l'intitulé des subdivisions auxquelles appartiennent ces entités : ainsi « paroisse » fait référence aussi bien à la paroisse religieuse de la période moderne qu'à celle de la période contemporaine (telle que définie par la Commission sociale des évêques de France, 1998) mais aussi à la paroisse administrative de la division financière de l'Ancien Régime. Etablir une codification univoque des entités territoriale, être en mesure de reconnaître à quelle emprise spatiale, à quel nom, quel code, et quel statut il est fait référence est l'ambition du projet Géonomenclature initié par (Motte et Pélissier, 2003). Ce projet constitue une extension (temporelle, en amont de 1943, et dans les différentes divisions du territoire, fiscale, religieuse, judiciaire, militaire, géographique) du COG de l'INSEE qui diffuse également un historique des recompositions des communes sur son site : « *le* ***fichier historique des communes*** *relate les changements de nom, les créations, les disparitions et les changements de département intervenus pour la France métropolitaine et les départements d'outre-mer (DOM) depuis 1943 de façon exhaustive. Les autres modifications (échanges de parcelles, transferts de chef-lieu, changements d'appartenance à un échelon supracommunal ou créations de cantons infracommunaux) sont répertoriées depuis 1994. Par ailleurs, les échanges de parcelles sans incidence démographique pour lesquels la superficie des communes est inchangée ne sont pas enregistrés de manière exhaustive.* »[3]

[3] Extrait le 8 mai 2012 de :
http://www.insee.fr/fr/methodes/default.asp?page=nomenclatures/cog/doc_fhistor.htm

Un référentiel commun est également nécessaire pour reconstituer des séries statistiques harmonisées, afin de restituer l'évolution de variables socio-économiques (telles que la population, le PIB, le chômage) sur un zonage de référence choisi. En effet, l'information statistique territoriale se présente le plus souvent sous la forme d'un ensemble de triplets (code_unité, code_indicateur, valeur) où le code_unité est celui valide à l'époque de la publication des données, mais plus forcément à l'époque de leur utilisation, puisque la codification des unités statistiques change au cours du temps. La Nomenclature des Unités Territoriales Statistiques européenne (NUTS) en fournit de très nombreux exemples. On observe ainsi, sur la figure 1, au Danemark entre 1999 et 2006, un remaniement complet visant à introduire un niveau régional qui n'existait pas. Les unités sont toutes recodées, et sont parfois recomposées (DK005 et DK006 fusionnent pour donner DK022). En l'absence d'un simple dictionnaire des unités, il devient difficile de connaître à quel territoire se rapporte tel ou tel chiffre. Si l'on dispose d'un dictionnaire du changement, il reste possible de reconstituer des séries statistiques longues, au moins au moyen de règles simples de proportionnalité. Ainsi, en 1999, les valeurs de DK022 pour un indicateur de compte[4] sont les valeurs additionnées de DK006 et DK005 en 1999.

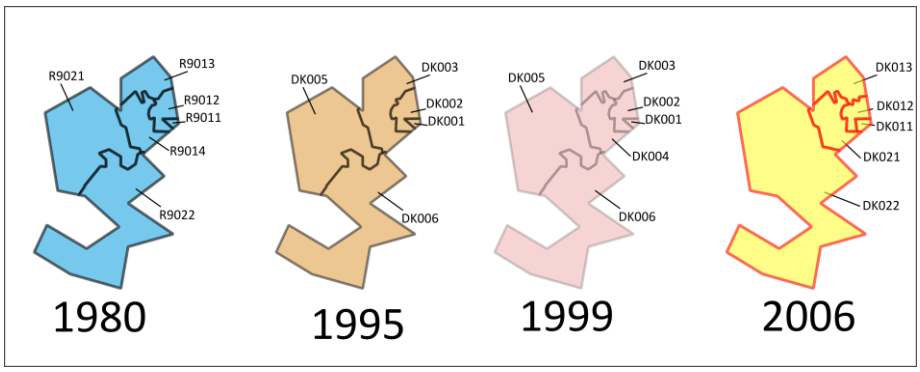


Figure 1. Histoire administrative du Danemark, Niveau 3 de la NUTS.

Par ailleurs, si l'on s'inscrit dans une logique d'analyse multi-niveau telle que (Mathian et Piron, 2001) l'ont proposé et telle que développée dans les travaux d'Hypercarte (Plumejeaud et al., 2011-a), il est nécessaire de disposer de l'histoire des appartenances des différentes entités spatiales. Il s'agit ici de mesurer un écart statistique d'une unité territoriale à un contexte, et d'observer l'évolution de cet écart au cours du temps pour une variable donnée (le PIB par exemple). Si le contexte est celui que constitue l'unité d'appartenance, la connaissance des changements

[4] Un indicateur de compte (ou stock) est additif : il résulte du comptage d'individus statistiques sur une entité ; ce n'est jamais un ratio.

d'appartenance au cours du temps permettra de comprendre l'évolution des écarts statistiques des unités étudiées suivant les différents contextes dans lesquels leur territoire a pu s'inscrire.

3. Problèmes conceptuels et méthodologiques

Une définition claire des objets de l'étude s'inspirant de travaux de recherche antérieurs est un préalable nécessaire à toute tentative d'opérationnalisation d'une solution informatique.

3.1 Conceptualiser une nomenclature

Nous entendons par nomenclature, l'ensemble des découpages territoriaux existant à un instant donné pour servir à l'exercice d'un pouvoir : religieux, administratif, fiscal, etc. Lorsque chaque niveau de la nomenclature s'imbrique parfaitement sans recouvrement ni omission dans un autre niveau, cette structure peut être représentée en informatique par un *arbre* (un graphe acyclique orienté), dont les entités territoriales constituent les sommets, et les relations hiérarchiques directes les arcs. Plus généralement, comme on l'a vu, il peut y avoir recouvrement, omission, et appartenance multiple, nous utilisons alors le terme de « *hiérarchie agrégative* » pour parler d'une nomenclature. Les maillages ou zonages en constituent les niveaux, et les entités territoriales les éléments de ces niveaux.

Un travail fondamental mené dans le cadre de la gestion de données multidimensionnelles complexes pour des entrepôts de données a donné lieu à des définitions plus précises de ces hiérarchies agrégatives (Pedersen, Jensen et Dyreson, 2001), (Banerjee and Davis, 2009). D'après ces travaux, une hiérarchie peut posséder une des trois qualités fondamentales suivantes, toutes ensembles ou l'une sans l'autre, ou même aucune : elle peut être *stricte*, et/ou *onto* et/ou *couvrante* :

- « *stricte* » signifie que tout élément d'un niveau n'appartient qu'à un seul autre élément de niveau supérieur,
- « *onto* » correspond au fait que l'arbre représentant cette hiérarchie est équilibré et que tout élément d'un niveau non élémentaire possède un sous-élément de niveau directement inférieur,
- « *couvrante* » signifie que tout élément d'un niveau appartient à au moins un élément de niveau supérieur.

Ainsi, une hiérarchie est dite *stricte* lorsque chaque unité possède une et une seule entité supérieure, *onto* lorsque toute unité de niveau non élémentaire englobe au moins une unité, et *couvrante* lorsque toute unité possède au moins une unité de niveau supérieur. Lorsque l'organisation arborescente des unités territoriales par rapport à la dimension spatiale est *stricte* et *onto*, l'agrégation ascendante depuis les feuilles vers la racine de l'arbre des données (Rigaux et Scholl, 1995) est permise, et elle offre des possibilités de vérification avec les invariants suivants :

- (i) la géométrie d'une unité non élémentaire est l'union des géométries des unités qui la composent ;

- (ii) la valeur d'une variable (quantitative absolue) d'une unité non-élémentaire est constituée par l'agrégation des variables associées aux unités qui la composent.

3.2 Concevoir la dynamique des entités territoriales

Concevoir la dynamique des entités territoriales, c'est nécessairement conceptualiser l'entité territoriale, et le changement. Il s'agit de cerner quelles sont les cibles de notre modèle, que nous qualifions de *descriptif*.

3.2.1. Modéliser une entité territoriale

Qu'est-ce qu'une entité territoriale ? A priori, c'est la *maille*, la *zone*, la *parcelle* territoriale qui constitue l'atome essentiel d'un découpage territorial, observé à son niveau. Cette chose est notre objet d'étude. Pouvons-nous en parler comme d'un *objet géographique* ? D'après (Smith, 1999), tout objet géographique possède une emprise spatiale. D'après (Couclelis, 2010), les objets géographiques correspondent également à une intention (*a purpose*), ils ont une fonction et sont composés d'autres objets, et ces objets sont identifiables. D'après (Hornsby et Egenhofer, 1998), ils ont une identité. Les objets sont des abstractions de la réalité perçus à travers des observations.

Dans notre cas aussi, l'entité territoriale est un objet géographique, elle possède une emprise spatiale, c'est un artefact à travers lequel s'exerce un pouvoir (elle correspond donc bien à une intention) ; elle remplit une fonction (selon les cas, administrative, militaire, religieuse, économique, etc.) ; elle est formée de sous-entités qui la compose, ou tout simplement des éléments du territoire que ses limites englobent. Mais possède-t-elle une identité, comme l'affirme K. Hornsby ?

Avant de répondre à cette question, il reste à souligner que d'après certains philosophes, comme Barry Smith, les entités territoriales ne sont pas des objets tangibles, et que même si elles sont perçues à travers des observations, ce n'est pas comparable à la perception d'un objet de notre univers familier, telle que la chaise. En fait, il existe une distinction entre objets authentiques (*bona fide*) et objet fictifs (*fiat objects*) au sens où le développe Barry Smith (1999) :

> ***Bona fide*** *: les entités authentiques, comme les tables, des chaises, des personnes, etc., qui existent en vertu de discontinuités physiques intrinsèques dans la constitution matérielle de la réalité.*
>
> ***Fiat object*** *: les objets qui surgissent grâce au fait que l'on a tracé des frontières spatiales qui, par exemple, séparent deux parcelles de territoire. Les frontières nationales ainsi que les frontières d'un comté fournissent des exemples de limites créées par fiat dans ce sens, au moins dans ces cas où, comme le Colorado, le Wyoming ou l'Utah, elles ne correspondent à aucune discontinuité qualitative de la réalité sous-jacente.*

Les entités territoriales, comprises au sens de catégories (plus ou moins emboitées ou hiérarchisées) formant une nomenclature, ne seraient donc que des objets fictifs ? Il est bien entendu que l'entité « *département* » n'a pas de chair, de texture, d'odeur, ne se voit pas, ne s'entend pas, ne se goûte pas. En revanche, les paysages du département du Gers, la texture de son sol, les parfums de son terroir, la forme de ses habitations et les habitudes de vie de ses habitants sont bien différents de ceux de la Meuse ou de la Haute-Savoie.

3.2.2. Identités et territoires

Quels sont les éléments qui permettraient d'affirmer qu'une entité territoriale possède une *identité* ? Il s'agit d'abord de définir *l'identité.* Voici la définition au sens informatique qu'en donne K. Hornsby (Hornsby et Egenhofer, 1998) :

> *Object **identity** is that trait that distinguishes an object from all others. Identity provides a way to represent the individuality or uniqueness of an object, independent of its attributes and values. The concept of a unique object identity has been recognized as a tool to help track changes to objects, particularly in proving existence or non-existence of an object as well as being able to track similarities or differences in objects during scenarios of change.*

Autrement dit, l'identité est un élément qui permet de différencier deux objets de même nature. C'est un moyen pour considérer son unicité, indépendamment de la valeur de ses attributs, et ce concept permet de suivre les changements qui affectent un objet dans le temps. En cela, les entités territoriales possèdent une identité et le « *département du Gers* » possède bien une *identité* qui la différencie du « *département de la Meuse* ». Ainsi, le contenant n'a pas d'identité mais le contenu en a une.

En définitive la question du point de vue est ici centrale, car on peut distinguer plusieurs natures d'identité :

- une *identité administrative* traduite dans les statistiques de population, de ressortissants à telle juridiction, à telle recette fiscale, à tel diocèse, de l'emploi, du revenu des terres etc.

- une *identité de géographie physique et humaine* reflétée par ses caractéristiques : structure du sol, réseau hydrographique et viaire, structure du peuplement (groupé/dispersé), style des habitations, etc.

- une *identité d'appartenance* révélée par exemple par la résistance aux changements, le refus de fusion et l'opiniâtreté à obtenir un rétablissement, ou tout simplement par l'attachement au nom qui désigne l'entité et l'importance du chef-lieu.

Toutes les entités territoriales connaissent au moins la première nature, et à des degrés divers la seconde nature. L'identité d'appartenance est en revanche plus prégnante dans le cas des divisions administratives. Encore faut-il noter entre elles une *graduation d'intensité*, la plus forte étant incarnée dans la commune. Voici ce qu'en dit le duc de Praslin en 1831 (Praslin, 1831) :

> *« La commune est le véritable élément virtuel de la société tandis que le canton, l'arrondissement et le département sont des fictions de l'administration. »*

Les archives (textes de lois, arrêtés ministériels, discours aux assemblées, dossiers de demande d'érection ou proposition de fusion) de la fin du 18e siècle jusqu'à aujourd'hui concernant les modifications du nom des communes, les mouvements de fusion/scission et le choix de leur nom et du chef-lieu de l'entité résultante sont très éclairants sur cet aspect. Il y a en France un attachement quasi viscéral, « charnel » au maillage mis en place pendant les années 1789-1801. Le volume de ces archives, l'imagination de l'argumentation, la force et même la violence des propos illustrent sans contestation la réalité de ce sentiment d'appartenance à un territoire : la disparition du nom de la nomenclature administrative et la suppression du statut sont comparées à une « mort civile », ce qu'énonce clairement le Baron Mounier (Mounier, 1835), quelques années plus tard :

> *« La commune correspond dans l'ordre politique à la famille dans l'ordre social. ... Transférer une commune d'un canton à un autre est une mesure d'ordre qui ne touche ni aux droits ni aux biens de la commune. Il ne peut en résulter que plus ou moins de facilités dans les relations entre les justiciables et leurs juges. ... Mais il en est tout autrement lorsqu'il s'agit de la réunion de communes, c'est, si l'on ose s'exprimer ainsi, une sentence de mort civile qui doit être prononcée. »*

Cet attachement à l'entité communale est une particularité que les gouvernements successifs ont échoué à remettre en cause. Une des raisons historiques est sans doute à trouver dans le processus de création des communes. Une commune hérite de manière quasi automatique de l'espace de la sociabilité paroissiale tissée petit à petit au cours des siècles précédents. Cet espace est consacré en circonscription administrative[5], puis laïcisé par le transfert de la tenue des registres d'État civil du curé au maire en septembre 1792, et enfin concrètement défini comme territoire lors de l'établissement du cadastre durant la première moitié du 19e siècle. Si le maillage du territoire national en circonscriptions supérieures a été l'œuvre révolutionnaire des administrateurs de l'époque (janvier-février 1790), en revanche la mise en place du tissu communal s'est faite en continuité de ce qui existait auparavant (novembre-décembre 1789).

Lorsque les limites entre deux communes sont contestées, « *le droit principal de l'une ou l'autre sera toujours l'état non contesté où étaient les choses en 1790 et antérieurement.* » (Circulaire ministérielle du 13 mars 1806 relative à la délimitation des territoires des communes). Les articles 5 et 7 de la loi municipale du 18 juillet 1831 (conservation de la propriété des biens de la commune et jouissance exclusive des biens dont les fruits étaient auparavant perçus en nature) étaient manifestement une mesure destinée à rassurer les populations qui craignaient qu'à l'issue d'une

[5] « Il y aura une municipalité en chaque ville, bourg, paroisse ou communauté de campagne », décret de l'assemblée nationale 12 novembre 1789 » et « les dénominations de ville, bourg et village sont supprimées et celle de commune leur est substituée », décret du 16 octobre 1793.

fusion « *leurs intérêts matériels soient systématiquement négligés par l'autorité municipale et que toutes les ressources soient dépensées au profit du chef-lieu* » ainsi que cela est rappelé dans la circulaire ministérielle du 29 août 1849. Quant au choix du nom, la fusion de communes se traduit souvent par l'adjonction des noms (exemple de la commune de Saint-Rémy-en-Bouzemont-Saint-Genest-et-Isson, issue de la fusion de trois communes en 1836), ainsi aucun ne disparaît de la nomenclature. Mais lequel porter en premier ? Celui qui emporte le chef-lieu ? En principe, oui, puisque la circulaire relative à la prescription du recensement de 1806 imposait que « *les arrondissements* [compris dans le sens de circonscriptions] *doivent être désignés par le nom de leur chef-lieu*. » [6] D'où l'importance de la localisation du chef-lieu ce qui donne lieu à des rivalités et des compétitions de longue durée avant d'établir un consensus.

Le département, bien que de création révolutionnaire « ex-nihilo », s'est lui aussi peu à peu attaché à son nom ou plutôt à l'image identitaire que celui-ci renvoie : la connotation négative des adjectifs tels que haut, bas, inférieur ou du nord a conduit certains départements à demander une modification de leur nom. Ainsi par exemple les Côtes-du-Nord sont devenue en 1990 Côtes-d'Armor, la Loire-Inférieure en 1957 Loire-Atlantique. Demandes pas toujours satisfaites, comme le montre dernièrement la demande, rejetée par la Commission nationale de Révision des noms, du département de l'Ariège de voir adjoindre à son toponyme un terme faisant référence à son identité pyrénéenne comme cela est le cas des trois autres départements pyrénéens (Pyrénées-Atlantiques, Hautes-Pyrénées et Pyrénées-Orientales).

Résistance aux modifications, attachement au nom et au choix du chef-lieu sont les signes d'une identité d'appartenance qu'il est possible aussi de percevoir dans l'existence d'un gentilé[7]. Commune, département et même Région, de création encore plus récente mais dont le nom fait référence à une ancienne province (Bretagne, Aquitaine, Provence, Bourgogne, etc.) ancrant ainsi son territoire dans une entité historique, se sont donné un gentilé. Mais il n'en existe pas pour les ressortissants d'un tribunal, ou les fidèles d'un diocèse, ou les imposés d'une recette fiscale !

Enfin, les territoires connaissent une identité d'appartenance d'autant plus forte qu'ils sont incarnés dans une personnalité. Le maire, le préfet de département et le préfet de région sont les représentants la puissance de l'État auprès des populations locales, mais ils font aussi remonter vers le pouvoir central les spécificités des territoires qui leur sont confiés.

En conclusion, il est possible de se représenter l'évolution au cours du temps de cette troisième nature d'identité, l'identité d'appartenance, pour différents types d'entités territoriales, en distinguant des entités qui jamais ne seront véritablement investies par celle-ci, tandis que des entités purement fictives au départ (au sens que donne Barry Smith) peuvent évoluer, avec une identité d'appartenance qui se

[6] Cette règle a été cependant et est de plus en plus souvent enfreinte.
[7] Le gentilé est le nom donné aux habitants d'un lieu.

manifeste plus intensément au cours du temps, comme les départements et les régions, figure 2.

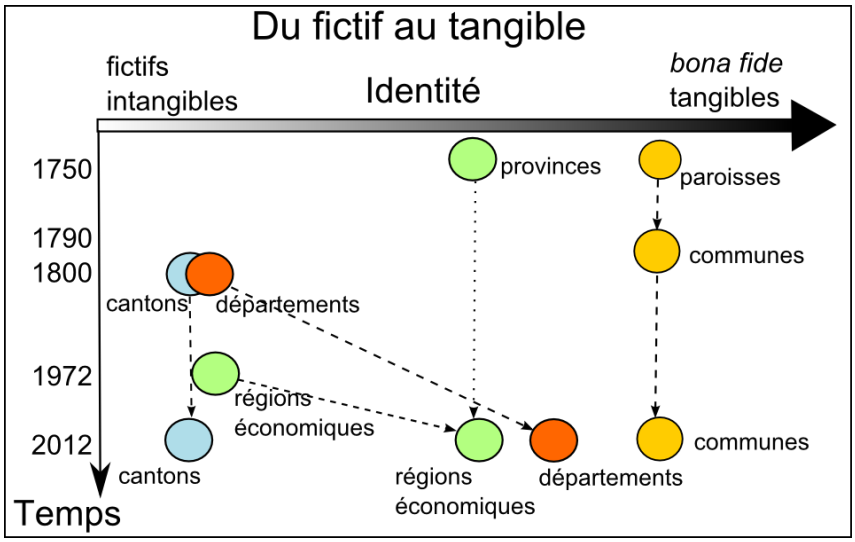


Figure 2. Du fictif au tangible : incarnation des entités territoriales dans le temps.

Il est facile de mettre en oeuvre des découpages cantonaux ou des modifications de la carte judiciaire car ils ne relèvent que d'une simple décision administrative. En revanche, modifier les communes, les départements et dans une moindre mesure les régions requiert en supplément l'assentiment des habitants car ces derniers s'identifient à ces territoires. En bref, plus ce sentiment d'appartenance est fort, plus toute mutation est longue à mettre en œuvre... La refonte de la carte administrative en France fait l'objet depuis pratiquement sa mise en place en 1790 de projets qui, de manière récurrente, réapparaissent sans grand succès. La dernière tentative en date, qui propose la réunion de communes en « communautés de communes », a certes été plébiscitée puisque sur les 36 683 communes françaises, seules 1 380 restent en dehors de toute intercommunalité (DATAR, 2010). Ce succès est relatif si l'on considère que l'intégration dans une communauté ne remet pas en cause le statut communal, ni la figure tutélaire du maire. Les communes persistent donc, même si l'espoir à long terme du législateur est qu'elles disparaissent sans émoi, peu à peu dépossédées de leurs domaines de décisions et de leurs ressources propres.

Peut-on alors considérer que les entités territoriales perdurent dans le temps (en opposition aux *occurants*) ? Voici ce que dit (Smith, 1999) :

> *Les continuants (qui comprennent ce qu'Aristote appelait « substances » ou parfois « choses », « corps » ou « grandeurs étendues ») demeurent identiques à eux-mêmes à travers le temps. Leur existence est continue d'un instant à l'autre et d'un jour au suivant. Voici quelques exemples de*

> *continuants : vous et moi, mon rocher préféré, la planète Terre, le lac de Constance, les couches supérieures de l'atmosphère terrestre, le système solaire, Pudding Island, ce que les Anglais appellent « Ulster », ce que les Irlandais appellent « Ulster ». Les continuants occupent l'espace et se divisent en parties étendues. Un continuant est identique à lui-même du début à la fin de son existence. Donc, un continuant n'a pas de parties temporelles.*

Cette définition inclue bien l'Ulster, une entité territoriale, comme exemple d'objet *perdurant* (*continuant*) dans le temps. Elle s'applique sans conteste aux objets géographiques mais semble difficile à appliquer aux entités territoriales administratives qui perdurent dans le temps alors même qu'elles connaissent de nombreux changements. Encore une fois ici, la commune française en est l'exemple type.

3.2.3. Définition de la commune retenue dans GéoNomenclature

L'examen des archives parlementaires et des textes de lois ou instructions gouvernementales au cours des premières années de la mise en place du nouveau système institutionnel en France (1790-1801) aident à définir les éléments qui fondent la commune en tant qu'*entité identitaire et administrative de base*. Il s'agit de la prise en compte et de l'application :

- de trois notions héritées de la période de l'Ancien Régime (*entité identitaire*) :

  - Notion de territoire fiscal : ceux dont les habitants sont cotisés sur les rôles d'imposition du chef-lieu sont réunis en une seule municipalité ;

  - Notion religieuse : le clocher est la référence du chef-lieu ;

  - Notion de communauté : le respect des convenances locales ;

- de trois principes généraux (*entité administrative) :*

  - Une architecture administrative hiérarchisée à 4 niveaux : départements, districts (puis arrondissements), cantons et communes ;

  - Une volonté politique d'attribuer au même échelon de base les fonctions de représentation nationale (élections) et d'administration du local (gestion des biens communaux, maintien de la sécurité publique) ce qui entraîne un équilibre difficile à établir puis à maintenir entre une administration générale (forte et centralisée) et une administration du et par le local ;

  - Une laïcisation de l'espace administratif par le transfert de la responsabilité de la tenue de l'état civil des mains du curé (Baptêmes, Mariages, Sépultures) à celle du maire (Naissances, Mariages, Décès).

Dans Géonomenclature, cinq attributs définissent une commune :

- Un nom
- Un territoire
- Une population
- Un chef-lieu
- Une appartenance à un échelon supérieur (le canton)

Ces cinq éléments, *à des ajustements spécifiques près et par l'énoncé de postulats pour leur observation*, peuvent être repris pour tout autre objet géographico-administratif y compris ceux de l'Ancien Régime. Au cours du temps passé et celui à venir, ces éléments ont connu et connaîtront de nombreux changements. La structure de la base et le code Géonomenclature historique sont tous deux conçus pour permettre la prise en compte de ces changements et des mises à jour nécessaires après le recueil de nouvelles informations.

3.2.4. Modéliser le changement

Dans (Cheylan et al., 1999), le temps est défini ainsi :

> *Un milieu indéfini où paraissent se dérouler irréversiblement les existences dans leur changement, les événements et les phénomènes dans leur succession.*

Donc formaliser le temps c'est aussi définir les *changements* et les *événements* qui se produisent dans l'espace géographique étudié. (Cheylan et al., 1999) précise que le *changement* ne se produit qu'à la condition que, pour une proposition P et des instants distincts t et t', P soit vraie à t, mais fausse à t'. Grossièrement, un changement peut être simplement défini comme une différence significative d'état pour l'objet d'étude (la propriété P de la définition) qui peut être tangible, une forêt par exemple, ou bien intangible (un bassin d'emploi) entre deux moments d'observation distincts. Pour le changement, qui est un terme générique, deux notions se distinguent : celle *d'évolution* ou bien celle de *mutation*. L'évolution, selon [Sanders, 1999], caractérise dans son sens premier un changement graduel dans le temps, marqué par une suite d'états différents observés. Mutation est utilisée plutôt pour qualifier un changement brusque, où la granularité temporelle de l'observation ne permet pas l'étude des états intermédiaires. Il est à noter également que le terme « mutation » possède un sens un peu différent dans le contexte de la programmation orientée-objet : il signifie alors changement de classe, (Chignoli, Crescenzo et Lahire, 1997). La mutation signifie alors qu'un objet typé suivant une classe change de type au cours de son existence. Là aussi la durée de ce changement n'est pas explicitée, ni la description de la phase transitoire, si elle existe. Les zonages territoriaux sont transformés par décret, à un instant donné. Aucun changement graduel n'est donc à modéliser ici : ce sont des processus de durée nulle, des mutations, qui provoquent une discontinuité dans la forme de l'espace. Si en réalité, cette mutation est effectivement précédée de changements graduels qui contribuent à la décision administrative, ceux-ci ne sont pas enregistrés dans les sources analysées, car ils sont d'ordre politique, sociétal, économique et leur complexité ne peut être rendue par celles-ci.

(Langran, 1992) fait remarquer qu'une description précise de la nature des changements, de leur localisation et de leur datation dans un espace d'étude géographique est au cœur d'un SIG temporel. Elle suggère ainsi que les *états*, les *événements*, et les *indices* sont les trois principales entités d'un SIG temporel. Les indices (ou preuves) permettent de découvrir les changements et d'en donner une mesure. Les états décrivent la distribution spatiale des phénomènes géographiques.

Les événements sont les causes de transformations de l'espace géographique (comme les inondations, ou les feux de forêt). Ainsi, la modélisation des événements est nécessaire mais également complémentaire de la modélisation des états de l'espace géographique.

Dans sa définition classique, (Cheylan *et al.*, 1999), un *événement* est forcément bref à l'échelle temporelle considérée. En UML, la notion d'événement intervient dans les diagrammes d'Etat-Transition, et un événement peut être la cause d'une transition qui sera toujours de ***durée nulle***. D'après le glossaire établi par Jensen (Jensen *et al.*, 1994), un événement est un *occurant*, quelque chose qui se produit instantanément à un instant t :

> *An event is an instantaneous fact, i.e. something occurring at an instant. An event is said to occur at a chronon t if it occurs at any instant during t.*

Cependant, les travaux de (Galton, 2004), mettent en évidence qu'un événement est également un objet, dont l'emprise spatio-temporelle est celle des objets impliqués dans l'événement. Il définit un événement comme un épisode borné dans le temps qui produit des effets remarquables sur un objet, ou un groupe d'objets localisés, ou un endroit du monde. Un événement est instantané (marqué par un instant) ou dure dans le temps (marqué par un intervalle temporel), mais ceci dépend de l'échelle d'observation. Par exemple, à l'échelle du siècle, une inondation est vue comme un événement ponctuel, alors que durant les jours de l'inondation, l'inondation est un événement qui dure. De plus, Galton observe que les événements peuvent aussi être composés de sous-événements, de granularité plus fine. Il cite comme exemple la Seconde Guerre mondiale qui peut être vue aussi comme une succession de batailles. Il introduit l'objet temporel pour généraliser la notion d'événement (ou *processus*, qu'il considère comme des synonymes), et ainsi l'histoire du monde se formule alors sous la forme d'une succession d'événements, ou *chronique*. Formaliser le temps avec des événements nécessite une algèbre appropriée pour raisonner : (Allen, 1984, Allen 1994) proposent également une algèbre pour traiter des événements. Ils généralisent les relations proposées pour les intervalles temporels avec le postulat suivant : les événements ont une durée ***non nulle***, et donc ils peuvent être modélisés sous la forme d'intervalles temporels. Pour (Smith, 1999), les événements sont les "*occurants*" :

> *Les occurrents (qu'Aristote appelle « accidents » ou parfois « événements », « processus » ou « états ») surviennent ou se déroulent dans le temps. Voici quelques exemples d'occurrents : des sifflements, des rougeurs d'embarras, des courses, mon actuelle migraine, votre connaissance du français, le processus de paix en Irlande du Nord. Tandis que les continuants peuvent exister par eux-mêmes, les occurrents ont besoin, pour exister, du support des continuants.*

La modélisation des processus de changements en tant que cause des transformations de l'espace géographique est activement défendue (Claramunt et Thériault, 1995), (Wachowicz, 1999), (Worboys, 2005), (Galton, 2004). Les deux principaux avantages de cette approche sont d'une part la construction d'un lignage (ou généalogie, ou graphe historique) des unités qui constituent le support (Spery,

Claramunt et Libourel, 2001) et d'autre part la modélisation du temps comme une dimension à part entière, avec une topologie (un événement suit ou précède un autre), qui autorise une vision à plusieurs niveaux de granularité de cette dimension. C'est pourquoi cette dernière tendance correspondrait à une vraie intégration du temps dans un espace 4-D (3D pour l'espace et 1D pour un temps linéaire orienté), et est qualifiée de *space-time geography* en anglais.

Décrire les événements permet de donner du sens au changement, et de progressivement passer d'un modèle purement descriptif des dynamiques territoriales à un modèle explicatif. En effet, l'observation de co-localisations spatiales et/ou temporelles d'événements et le tri par nature de ces événements devient possible, offrant ainsi de meilleures capacités d'analyse.

2.3 Typologie des événements et processus territoriaux

Il existe plusieurs typologies des processus de redécoupages administratifs, mais il n'existe pas de consensus entre les informaticiens (Hornsby et Egenhofer, 1998), les géomaticiens (Claramunt et Thériault, 1995), (Spery, Claramunt et Libourel, 2001), (Plumejeaud et al., 2011-b), les statisticiens (Lang, 2010) et les historiens (Motte, Séguy et Théré, 2003). Nous comparons ici ces typologies : contexte d'utilisation, limites, points communs et particularités.

Outre le fait que ces typologies ont en commun de modéliser des changements brusques sans transition, elles distinguent les événements qui n'affectent qu'une entité territoriale (comme le changement de nom) de ceux qui affectent plusieurs entités territoriales dans un échange de parcelles. Ainsi, Claramunt et Thériault distinguent trois catégories fondamentales de changement :

- I : l'évolution d'une entité indépendante ;

- II : le changement impliquant des relations fonctionnelles entre plusieurs entités interdépendantes;

- III : les restructurations territoriales impliquant des unités géographiques.

Toutefois, les travaux de Plumejeaud (Plumejeaud et al., 2011-b) ou l'INSEE proposent d'une part de considérer certains de ces changements attributaires comme des conséquences d'autres événements impliquant eux plusieurs entités. Par exemple, concernant le changement de nom d'une commune, l'INSEE codifie différemment cette opération selon qu'elle intervient sur une commune de façon isolée (code 10) ou que ce changement engage plusieurs communes (code 14, transfert de chef-lieu).

Le transfert de chef-lieu (changement de catégorie II), qui représente un transfert de pouvoir d'une localité à une autre, n'est pas modélisé au niveau européen (NUTS), mais l'est en France, par (Motte, Séguy et Théré, 2003) au niveau communal, et par l'INSEE, qui numérote entre 50 et 54 les transferts de chef-lieux aux niveaux communal, cantonal, départemental, régional. De même le changement d'appartenance n'est pas modélisé au niveau européen, mais en France, toujours par (Motte, Séguy, et Théré 2003) et par l'INSEE, à tous les niveaux (codes 40 à 43).

L'ensemble des travaux considérés modélise les changements de la catégorie III, de la manière suivante (Figure 3).

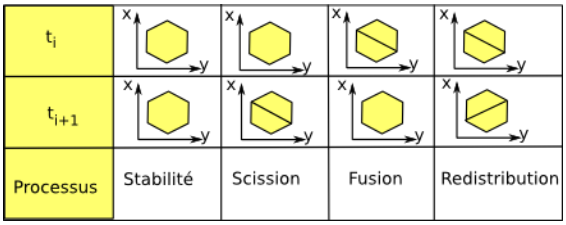


Figure 3. Processus de restructuration territoriale, (Claramunt et Thériault, 1995).

Des différences importantes existent à la fois en terme de vocabulaire, mais également de concepts. Il y a ceux qui considèrent un monde ouvert, comme (Spéry, Claramunt et Libourel, 2001), qui modélise le *Domaine Public*, espace où l'on peut prendre (*Extraction*) ou rétrocéder (*Intégration*) du terrain. Et sinon, ceux qui considèrent qu'aucune entité ne peut être créée ex-nihilo (Motte, Séguy et Théré, 2003), dans un monde fermé. Les différences conceptuelles portent également sur le fait qu'une entité puisse avoir une histoire, et une identité propre : c'est le parti-pris des travaux de (Hornsby et Egenhofer, 1998) qui modélise un objet inexistant sans histoire, ou existant avec histoire ou existant sans histoire[8]. Avec son langage de description, Hornsby introduit un ensemble de primitives permettant de désactiver une entité à un moment et de la réactiver (le terme "*réincarnation*" est employé) plus tard. De même, la représentation de compositions d'entités, comme des pays intégrant une fédération qui s'active, qui la quittent lorsque cette fédération se désintègre (comme l'URSS en 1991), permet de modéliser des entités qui sont actives, puis absorbées, puis réactivées par libération du composé. Ce type d'approche est celle qu'adoptent (Motte, Séguy et Théré, 2003) qui indiquent qu'il n'y a jamais dilution totale de l'identité d'une commune lors d'une réunification (ou fusion, ou réunion) et les termes *Absorbante, Absorbée, Absorption, Rétablissement* (pour *réincarnation*) sont employés. De façon implicite, c'est également l'approche de l'INSEE, car l'organisme utilise un code spécial pour notifier le rétablissement d'une commune (code 21), et distingue lors d'une réunification la commune absorbante de celle qui est absorbée (celle-ci perd généralement le chef-lieu, et n'apparaît plus qu'en second voir plus du tout dans le toponyme de l'entité réunie). Cependant, par rapport à la codification des événements de réunion que propose Claude Motte, l'INSEE distingue la *fusion simple* de la *fusion-association*. Cette dernière est une opération au cours de laquelle sont créées des communes *associées*, statut spécial qui a été institué par la loi du 16 Juillet 1971, dite "Loi Marcellin" et abrogé par la loi n° 2010-1563 du 16 décembre 2010 qui réforme le statut des

[8] Par rapport à la compréhension intuitive de ces termes, "histoire" est à entendre comme "vie", activité de l'entité, et "existence" comme enregistrement dans une base de données ou un registre officiel.

collectivités territoriales. Ce statut particulier permettait à la commune associée (qui est supprimée) certains droits et privilèges sur son territoire créant une sorte de sous-entité sur la commune ayant une part d'autonomie.

Cette vision identitaire n'est absolument pas partagée au niveau européen avec la NUTS, qui utilise un vocabulaire déniant à l'unité statistique territoriale toute identité : elle est fusionnée avec une autre, et chacune perd son identité au profit de la nouvelle entité créée.

En dehors des fusions ou scissions simples n'impliquant que deux unités, les différentes approches modélisent aussi les échanges de parcelles qui conduisent à des évolutions de frontières, mais qui n'ont pas forcément d'incidences démographiques. (Motte, Séguy et Théré, 2003) n'ont considéré ainsi que les échangent de parcelles ayant une incidence démographique, contrairement à l'INSEE qui distingue les deux cas. Le terme qui peut être utilisé serait *Rectification*, ou *Redistribution*, lorsque ces échanges n'impliquent aucune disparition ou apparition d'entité. Sinon, il n'y a pas de consensus sur le terme à utiliser : (Motte, Séguy et Théré, 2003) et l'INSEE décrivent la *création* d'une commune à partir d'entités qui cèdent des parcelles, ou la *suppression* d'une entité à partir d'entité qui absorbent des parcelles, tandis que (Spery, Claramunt et Libourel, 2001) et (Plumejeaud et al., 2011-b) décrivent l'événement comme une *réallocation*, de façon identique pour tous les acteurs impliqués dans l'événement.

4. Difficultés liées à l'opérationnalisation de solutions informatiques à partir d'un cas pratique

Le projet GeoNomenclature (Motte et Pélissier, 2003) se propose de recenser et restituer l'évolution des nomenclatures administratives de la France métropolitaine de l'époque moderne à nos jours dans une base de données spatio-temporelle. Il prévoit la création d'un code géo-historique associé permettant à différents profils d'utilisateurs d'indexer facilement des événements ou des documents dans le temps et dans l'espace. Ce travail doit tenir compte des incertitudes et de l'absence de sources inévitables sur une si longue période. L'examen des textes officiels et l'analyse des différentes sources disponibles a permis de distinguer les différentes dimensions sur lesquelles se fonder pour définir et suivre les évolutions des entités territoriales (nom – territoire – chef lieu – population – appartenance à l'échelon administratif supérieur), comme développé dans le paragraphe 3.2.3. La volonté d'organiser ces informations pour les restituer clairement a été l'occasion d'expérimenter plusieurs modèles de structuration des informations et de les confronter, [Rat Patron, 2011], [Plumejeaud, 2012][9].

[9] Ce projet a bénéficié de l'acquis d'un travail précédent : le site Cassini et la base de données associée, fondés sur une identification précise et minutieuse des communes ayant existé depuis 1793, de leurs ressorts administratifs et de leur population.

4.1 Quelques questions

4.1.1 L'identité est-elle nécessaire ?

Est-il possible de construire un système d'information géo-historique viable capable de rendre compte du changement sans avoir recours à la notion d'identité ? Un tel système, qui renierait toute identité à une entité territoriale, peut sembler séduisant du point de vue logique et garantir une certaine objectivité dans le traitement des informations. Il suppose que chaque changement d'une caractéristique d'une entité induit la création d'une entité nouvelle et la conservation de l'entité ancienne dans le système (versionement).

L'analyse de la mise en œuvre concrète de ce principe sur l'histoire administrative de la France révèle à l'usage des défauts majeurs.

D'une part, ce n'est pas un système de pensée partagé naturellement par un utilisateur et il est souvent difficile de se « situer » dans les données, surtout lorsqu'on y accède avec une connaissance limitée. Le versionement des informations est nécessaire pour prendre en compte l'histoire des évolutions, mais la complexité de certaines situations et événements rendent leur restitution quasiment inextricable.

D'autre part, l'absence d'identité engendre une complexité *informatique* difficile à manœuvrer. Dès lors que le toponyme, la géométrie, le statut, l'appartenance (ou lien hiérarchique) à une autre entité et le chef-lieu d'une entité change, il faut dupliquer entièrement l'entité, et ses attributs. Ceci est coûteux en terme d'espace (en particulier lorsqu'une géométrie, de type multi-polygone, est composée de 200 points en moyenne, soit 3 Ko). De même, l'interrogation des données pose de nombreux problèmes. Les recensements associés à chaque version d'entité deviennent difficiles à récupérer pour un ensemble d'entités ayant évolué au cours du temps, puisqu'il faut suivre les liens de filiation entre les différentes versions d'entités pour récupérer les données. Une requête de ce genre suppose des parcours de graphe de tailles différentes pour chacune des entités. La complexité des modifications en cascade qu'un tel paradigme peut entrainer pose également problème. Ainsi, par exemple, en cas de changement de nom en date *d2*, un canton *can1* est dupliqué et devient le canton *can2* par exemple. Ce canton gouvernait les communes *c1, c2, c3* et *c4*. Ces communes n'ont pas changé de canton (réellement) mais par un artifice, elles ont maintenant un autre canton supérieur (*can2*). La conséquence logique serait donc la duplication de *c1, c2, c3 et c4* et donc de leurs attributs, en rouge sur la Figure 4.

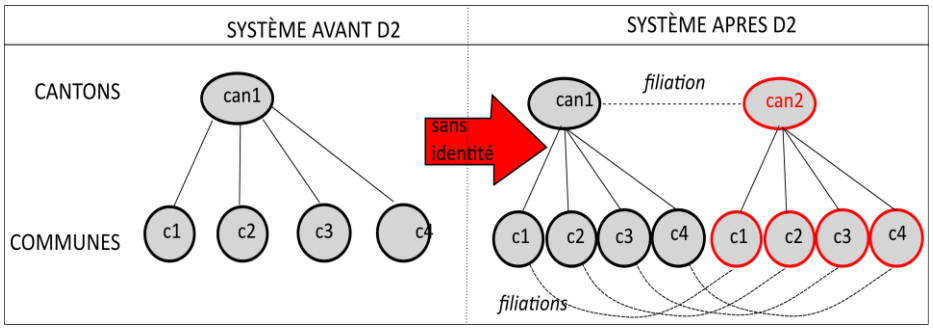


Figure 4. Modéliser des changements territoriaux sans identité.

Les informations pertinentes sont alors noyées au milieu de répétitions indigestes. Par ailleurs, sur un tel volume de données, la probabilité d'introduire des erreurs et des incohérences augmente fortement, leur détection est aléatoire et leur correction une gageure, même pour un expert. Lorsque les attributs sont versionnés, et qu'on accorde une identité aux entité territoriales indépendante des valeurs des attributs, comme sur la Figure 5, le système d'information se contente de dupliquer l'attribut modifié, en lui donnant sa période de validité (après *d2*), rattaché à l'entité qu'il décrit, ce qui est plus économique, plus lisible et maintenable.

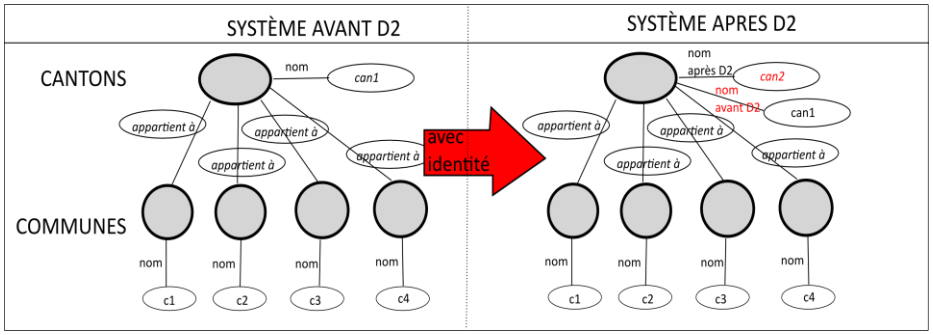


Figure 5. Modéliser des changements territoriaux avec identité.

En conclusion, les tests de structuration réalisés sans tenir compte de l'identité ont révélé le caractère coûteux et finalement impraticable d'une solution privée du concept d'identité, du moins aux niveaux les plus bas de la hiérarchie administrative (communes et cantons), ce qui représente au moins 50 000 unités.

4.1.2 Comment procéder à l'identification ?

Si le choix d'une solution privilégiant l'identité des entités géographiques s'avère pertinent pour l'utilisateur, l'expert ou le géomaticien, il est nécessaire de

préciser la façon de repérer ces entités au sein de la masse de données rassemblées. Partant des cinq éléments opérationnels concourant à la définition d'une entité territoriale, plusieurs moyens peuvent être mobilisés pour procéder à l'identification effective des entités à partir des sources dont nous disposons, sachant que ces dernières sont bien souvent incomplètes, imprécises et incertaines.

Dans le cadre de la Nomenclature des Unités Territoriales Statistiques, créée en 1980 en Europe pour définir le maillage des niveaux départementaux, régionaux et étatiques, et qui a connu au moins cinq versions depuis lors, la thèse de (Plumejeaud et al., 2011-b) a proposé un algorithme d'appariement des versions par comparaison des attributs de chaque entité composant le maillage : nom, code, empreinte spatiale, et appartenance. Cette proposition dérive d'une approche par fusion multi-critères mettant en jeu la théorie des croyances (Olteanu et al., 2009), plus à même de prendre en compte l'imperfection de l'information, mais qui avait été évaluée sur l'appariement de ponctuels ou de linéaires, et dont la mise en œuvre pouvait s'avérer coûteuse en calcul. Enfin, dans le domaine de la représentation de connaissance, les travaux fondés sur des ontologies spécifiques pour la prise en compte des événements territoriaux (Kauppinen, Väätäinen et Hyvönen, 2008), (Gantner, 2011), et qui se basent eux aussi sur un paradigme identitaire, ont réussi à produire des requêtes capables de retracer la vie d'une entité, lorsqu'elles sont formulée en SPARQL. Ce type de système ne nécessite pas de construire une base de données où les versions d'unités sont appariées, puisque l'identification se fait à la volée, sur des critères logiques. En revanche, (Gantner, 2011) déplore les performances très moyennes des raisonneurs sur ontologies. Par ailleurs, l'imperfection des données, comme les lacunes sur les attributs spatiaux (ignorance des limites de communes dans le passé par exemple) ne sont pas pris en compte non plus par ce type d'approche.

En ce qui concerne le projet Cassini, l'identification des communes a été réalisée par les historiens-démographes. Ce travail d'expertise s'est fondé sur l'observation des critères « objectifs » (toponymes, population,… cf. ci-dessous) et sur le croisement et la hiérarchisation des sources (actes officiels, listes de communes, recensements). Il s'est aussi appuyé sur des micro-analyses du contexte local, des analyses partielles dérivées de la mise en série des sources, ainsi que par l'adoption d'hypothèses contrôlées par un retour aux documents originaux ou bien par la consultation de sources complémentaires et/ou locales. Ce travail scientifique progressif, de certification par des experts en fait, a permis d'établir *in fine* une liste raisonnée des communes ayant existé entre 1793 et 2006, ainsi que de leurs relations. Il s'agit d'une base intelligible pour n'importe quel utilisateur du site web associé. Ce travail est un résultat en soi, invisible aux yeux du modélisateur mais pourtant essentiel (Dagiral et Peerbaye, 2012), qui éclaire les dimensions multiples de l'histoire administrative depuis deux siècles. Les réflexions et les enseignements de cette phase d'identification préfigurent l'analyse de l'évolution administrative et démographique de la France en même temps qu'ils en nourrissent les hypothèses (Motte, Séguy et Théré, 2001).

4.1.3 Difficultés soulevées par l'imperfection des informations

La reconstitution de l'histoire administrative de la France s'est fondée sur un ensemble de sources de grande qualité, existant à l'échelle du pays entier sur plus de deux siècles (c'est une des particularités du cas Français)[10]. Elles ont cependant posé de nombreuses difficultés pour aboutir à une restitution cohérente des données relatives à chacun des cinq attributs d'une entité territoriale. C'est particulièrement le cas au niveau des communes[11] :

- Dans bien des cas, le *nom* est source de confusion et ne peut suffire à distinguer une entité. Chaque source nous confronte aux problèmes d'homonymies, d'évolutions toponymiques (qu'elles soient issues de changements de nom officiels, de variantes ponctuelles ou d'usage locaux), ainsi qu'aux erreurs de typographie des contemporains, des administrateurs ou bien des personnes chargées des retranscriptions[12]. Un travail d'appariement est nécessaire pour assurer la cohérence des informations et la mise en série des données, et cette tâche est considérablement facilitée par le recours aux identités. Le toponyme d'une commune est bien sûr la variable principale, mais les autres attributs participent à cette phase de désambigüisation.

- Le *chef-lieu* est le siège du corps municipal. Il donne généralement son nom à la commune et est à l'origine le lieu où est situé le clocher. Cet attribut peut servir efficacement de base pour l'identification (les actes officiels sont censés le préciser clairement). Le positionnement du chef-lieu retenu sur le territoire est de plus un élément déterminant pour l'INSEE[13], et est particulièrement utile en ce qui concerne la localisation des communes sur les cartes anciennes (les clochers sont représentés sur la carte de Cassini et la carte de l'état-major). De nombreuses exceptions subsistent néanmoins : les paroisses urbaines ne donnent pas naissance à des communes et à l'inverse, en campagne, certains hameaux importants, souvent fortement imposés, mais sans clocher sont pourtant érigés en commune. De même,

[10] Bulletin des lois, recensements, nomenclature administrative, dossiers communaux aux archives, etc.

[11] Les difficultés sont plus accentuées au début de la période d'observation (1793-mi-19e siècle). Elles sont liées à la méconnaissance des premiers territoires communaux (les paroisses d'avant 1789), mais aussi jusqu'à l'établissement du cadastre (indument appelé napoléonien). Le principe du cadastre, établi dès 1791, est certes repris sous l'Empire napoléonien mais il ne sera réellement mis en œuvre que par les circulaires ministérielles de 1821 et de 1828 et l'opération ne sera achevée pour l'ensemble des territoires qu'après 1850. Ces documents, malheureusement détruits ou perdus pour certaines communes, ont fait l'objet de corrections ou de modifications non datées et parfois en incohérence avec les dates de mutation officiellement décidées.

[12] Ces « erreurs » peuvent être de simples fautes d'orthographe, ou résulter de choix plus ou moins systématiques suivant les sources (traitement différencié des particules selon les sources, emploi d'un nom d'usage plutôt que du toponyme officiel, etc.). Elles peuvent être aussi le résultat de pratiques sociales concurrentes, notamment dans un contexte où la langue française est imposée par un pouvoir fortement centralisateur.

[13] C'est le choix opéré par l'INSEE pour notifier l'évolution du Code officiel géographique par exemple.

les querelles de chef-lieu sont nombreuses lors des remembrements communaux et ont une influence non négligeable sur la toponymie des communes.

- L'examen des sources concernant la *population* peut contribuer à la levée de certaines ambigüités. A partir de 1790, le pouvoir a fait réaliser à intervalles réguliers des recensements de population sur l'ensemble du territoire national [Lang, 2006][14]. Bien souvent, l'appariement des listes de communes issues des recensements aux listes officielles des communes permet d'affiner la connaissance de l'histoire de l'évolution de chaque commune : la cohérence en termes de population permet de conforter les hypothèses d'appariement toponymiques, alors que la détection de l'absence d'un toponyme ou d'une variation de population d'amplitude « anormale » (inattendue ?) est l'occasion d'un retour aux sources officielles pour retrouver la trace d'une évolution administrative éventuelle. Pour l'Ancien Régime, les dénombrements de population existants ne pourront jouer un tel rôle car leur fréquence, leur emprise territoriale, et les objets mesurés (nombre de feu, individus,…) varient considérablement.

- Le *canton* est le niveau immédiatement supérieur à la commune dans la structure administrative générale et a – entre autre fonctions – une compétence en matière électorale. Les élections à la représentation nationale sont fondées sur les communes[15] et un canton est généralement une circonscription électorale regroupant plusieurs communes. Toutefois, par soucis d'équité, le nombre de représentants élus est aussi fonction du nombre d'électeurs (pour le suffrage censitaire, comme pour le suffrage universel qui lui succède) : certaines communes importantes ont du être divisées en plusieurs cantons, ou bien une partie de leur territoire a été intégrée dans un canton rural limitrophe. Ces exceptions brisent la rationalité d'une structure à stricte emboitement hiérarchisé. Il a donc fallu prévoir une procédure spéciale pour intégrer ces multi-appartenances. Ce cas de figure s'avère très utile pour penser l'intégration des circonscriptions d'Ancien Régime qui connaissent de nombreuses situations similaires. Pour les autres niveaux administratifs, l'emboîtement hiérarchisé fonctionne correctement (arrondissements, départements, régions).

- L'identité d'une commune passe par une définition de son emprise *territoriale*. Cette dernière permet une meilleure compréhension des rapports à l'espace, au voisinage et aux phénomènes de recomposition qui jalonnent l'évolution d'une entité[16]. Nous disposons pour la période actuelle d'une cartographie précise des territoires communaux à l'échelle de la France (limites administratives de la *BD Carto*). Il est en revanche quasi-impossible de déterminer précisément les limites des premières communes issues des paroisses d'Ancien Régime[17]. Le projet

[14] L'ensemble de ces recensements nationaux (36 depuis 1793) concernant les quelques 43 000 communes identifiées a été intégré dans la base de données.

[15] Les variantes du début de période (assemblées de municipalités, assemblées primaires) s'articulent elles aussi autour des communes.

[16] Des schémas utilisant des formes fictives d'entités territoriales peuvent toutefois être très utiles et figurer avec clarté des situations et des événements complexes (Motte, Séguy et Théré, 2003).

[17] Cf. note 11.

Géonomenclature a pris le parti de reconstituer les limites communales de l'ensemble des communes françaises depuis la révolution. Les contours communaux ont été élaborés, sans prétention d'exactitude, mais en respectant certaines règles générales[18] de sorte que chacune des communes identifiées dispose sur l'ensemble de la période d'une emprise territoriale fictive mais réaliste. Ce travail ouvre la voie à l'établissement d'une véritable cartographie statistique historique au niveau communal.

Au-delà des aspects particuliers à chacune des dimensions définissant les entités, l'exploitation des sources a dû faire face plus généralement à trois types de difficultés : l'imprécision des dates et de la toponymie, l'absence ponctuelle de sources pour des parties d'échantillon et des lacunes dans les sources existantes consultées[19].

Les textes officiels notifiant les changements intervenus (nom, territoire, chef-lieu) ne sont pas tous conservés si bien que l'existence ou la disparition d'une commune est alors constatée seulement par sa présence ou son absence sur les documents des recensements quinquennaux. Si les documents existent, nombre d'entre eux sont imprécis s'agissant des dates de création, de suppression, de modification de dénomination et/ou de territoire. Parfois encore, le croisement des sources révèle des incohérences manifestes entre les dates officielles des mutations et leur mise en application : une population communale est par exemple recensée alors que la date de suppression décrète le changement avant le recensement, un changement de nom officiel n'est pas pris en compte dans les textes administratifs postérieurs[20], etc. Du point de vue des toponymes, le principe est que l'entité territoriale doit porter le nom de son chef-lieu. Cette règle souffre en réalité de nombreuses exceptions auxquelles s'ajoutent les imprécisions de la désignation du chef-lieu communal après un remembrement : l'omission de l'information concernant le chef lieu ou la rédaction fluctuante des actes ne permet pas toujours de déterminer avec certitude quelle est l'entité qui survit (et partant son nom officiel) et celle qui disparaît[21]. L'ensemble de ces difficultés induit de fait des lacunes dans les informations nécessaires à l'identification des entités territoriales. Ces informations doivent alors être estimées à partir d'hypothèses issues d'un croisement opportun avec d'autres sources ou déduites d'une mise en cohérence avec les informations déjà présentes dans la base.

---

[18] Les limites fictives des anciennes communes (aujourd'hui disparues) ont été créées en combinant plusieurs facteurs : partant de la localisation des chefs lieux, un premier découpage respectant la proportionnalité des populations des différents morceaux d'une commune est réalisé. Les limites obtenues sont ensuite réaménagées manuellement en suivant les tracés des chemins ou autres cours d'eau observables sur les cartes anciennes disponibles.

[19] Ces lacunes sont relativement peu fréquentes pour la période contemporaine, mais deviennent la règle pour les circonscriptions d'Ancien Régime.

[20] Pour certaines communes, un rectificatif ou une date d'effet paraît plus d'un siècle après, cf., Marcilly-les-Vitteaux et Marcilly-et-Dracy, (1861-2001) en Côte d'Or.

[21] La terminologie est variée : réunion, union, suppression, rattachement, absorption, création, érection, scission.

4.1.4 Comment rendre compte des relations d'appartenance ?

Par rapport à la modélisation des appartenances multi-niveaux, (Mathian et Piron, 2001), proposent un formalisme mathématique adéquat pour la représentation de l'appartenance d'une unité territoriale à des classes, sous forme de graphes de partition ou bien de proximité, et facilitant l'usage de méthodes statistiques multidimensionnelles. Ce formalisme ne tient toutefois pas compte du temps.

La modélisation de hiérarchies sur la dimension spatiale a également été abordée dans le cadre des systèmes d'information décisionnels spatiaux dits « SOLAP » (Bédard et al., 1997). La question de l'évolution des maillages ou des supports géographiques est traitée (Tchounikine et al., 2005) à l'aide d'une grille régulière fine de l'espace qui sert de PPCD_spatial pour organiser les transferts de données entre les différentes couches temporelles. Des fonctions de mappage relativement simples sont utilisées pour réaffecter les valeurs d'indicateurs statistiques entre les différentes couches temporelles (suivant la proportion de la surface intersectée entre une couche temporelle et le PPCD_spatial). Ce type d'approche règle partiellement le problème de l'implémentation puisque c'est ici un unique support matriciel prédéfini à l'avance qui sert de PPCD_spatial. Mais le choix de l'échelle de cette matrice peut s'avérer plus délicat que prévu, car la finesse des maillages varie spatialement et temporellement, et on doit procéder à une approximation de la forme irrégulière d'un polygone. Ce traitement du changement spatial ne tient pas non plus compte des changements d'organisation dans la hiérarchie, et, enfin, aucun lien sémantique n'est établi entre les unités géographiques : on ne sait pas si ce sont les mêmes ou bien des nouvelles. Or, dans l'étude de l'évolution d'un phénomène démographique par exemple, produire un lien continu entre des séries de données statistiques et les territoires auxquels elles sont liées est particulièrement important. (C. Plumejeaud et al., 2011-b) proposent de versionner ces liens d'appartenance, datés en fonction de leur période de validité, et cette approche a été retenue pour GéoNomenclature.

4.1.5 Comment maintenir la cohérence du modèle ?

L'élaboration d'un modèle résulte d'une confrontation des sources de données existantes aux objectifs assignés à l'outil final. Plusieurs modèles peuvent être produits à partir d'une même base d'observations et il convient de juger de leur cohérence en fonction de l'usage auquel ils sont destinés.

Le projet Géonomenclature vise à créer un référentiel historique à destination du public. En l'absence d'une source décrivant par le menu la « biographie » et les relations de chaque entité territoriale, le modèle associé à ce projet a pour principale fonction de mettre en cohérence et en série des états et listes successifs afin de reconstituer l'évolution temporelle et spatiale des entités administratives françaises depuis la Révolution, puis durant l'Ancien Régime. Le maintien de la cohérence de ce modèle se joue à plusieurs niveaux et à différents moments de son élaboration.

Une composante importante de ce travail consiste à assurer la compréhension du modèle et des informations contenues dans la base par les utilisateurs potentiels. A ce titre, le choix de travailler avec des entités identifiées facilite l'accès aux

informations de la base et sa prise en main par les tiers. Ce choix a l'inconvénient d'introduire une phase intermédiaire entre la récolte des données et la restitution aux utilisateurs : un travail d'expertise et de mise en cohérence (essentiellement fondé sur la base de l'identification des entités territoriales) est nécessaire. C'est seulement à ce prix qu'une base de référence efficace de l'histoire administrative a pu être constituée.

Il est nécessaire d'assurer la cohérence du modèle dans son ensemble, mais aussi celle de ses parties. Il doit ainsi être capable de restituer une source particulière dans son état originel. Chaque source de donnée obéit à des logiques propres (organisation interne, limites, contexte,…). Le pari est de réussir à intégrer ces informations et respecter ces logiques tout en assurant une cohérence à l'ensemble de l'édifice. Le cas du dénombrement de l'an VIII est exemplaire à ce titre : il a valeur d'acte officiel de dénomination pour les communes. Pourtant, il s'agit d'un recensement qui contient à l'évidence des erreurs (communes non citées, communes comptés avec une autre) et qui présente un nombre important de dénominations de communes manifestement erronées ou du moins rapidement abandonnées. Une commune actuelle peut très bien ne disposer que de ce document comme acte officiel de dénomination et n'avoir jamais porté le nom de 1801. Il faut bien pourtant trouver le moyen d'intégrer ce recensement dans la série chronologique des populations.

L'ajout et/ou la mise à jour d'informations constitue un risque important pour la cohérence du modèle. Sa portée générale dépend notamment d'un choix raisonné de nouvelles sources de données, qui même si elles relèvent de la problématique, devront se confronter aux différentes exigences induites dans le modèle : systématisme, exhaustivité, granularité, emprise spatiale, …

Même en envisageant des protocoles précis pour l'ensemble des cas de figures d'évolutions et en automatisant la plus grande partie des modifications, l'introduction de nouvelles informations est une opération bien souvent délicate tant il est difficile pour l'utilisateur de se « situer » dans les données et de mesurer les impacts des modifications. Ceci sans compter les contradictions inévitables qui invalideront des données, des hypothèses ou des expertises préalablement validées et qui nécessiteront un nouvel examen et éventuellement une nouvelle « mise en ordre ».

Pour toutes ces dimensions, il est crucial de gérer la qualité des informations présentes dans la base de données. Il est nécessaire de conserver la *provenance des informations* enregistrées, en lien avec l'expert qui valide telle ou telle hypothèse. Ces *métadonnées* décrivent cette provenance, et exigent la citation des sources à la manière des historiens (qu'elles soient bibliographiques ou autres). Elles permettent de disposer à tout moment d'un état d'avancement de la saisie (sources intégrées) et données consolidées (appariements assurés). Il s'agit également d'être en mesure d'historiciser les différents points de vue pris sur les données. A ce titre le système d'information se doit d'être *bi-temporel*, c'est-à-dire de conserver à la fois la date de validité des entités, de leurs attributs et des relations qu'elles entretiennent, mais également la date d'enregistrement dans le système de ces informations.

4.2 Un cas d'étude : le modèle de GéoNomenclature

Le modèle de Géonomenclature repose sur le paradigme que les entités ont une vie (et donc un identifiant système unique au cours du temps), avec des attributs versionés au cours du temps. Ainsi dans le cas de Charmeceau (cf. Figure 6), on établit une filiation avec la commune de Charmesseaux existant de 1822 à 1832. C'est dans le même système un seul et même objet, dont l'attribut nom change en 1822. Cet objet interrompt définitivement sa période de validité lorsque que Charmesseaux est absorbée dans Trancautlt en 1972, à son tour absorbée dans Val d'Orvin en 1972 avant un rétablissement de Trancault en 1999. Trancault a donc deux périodes de validité : 1832-1972 et 1999 à aujourd'hui.

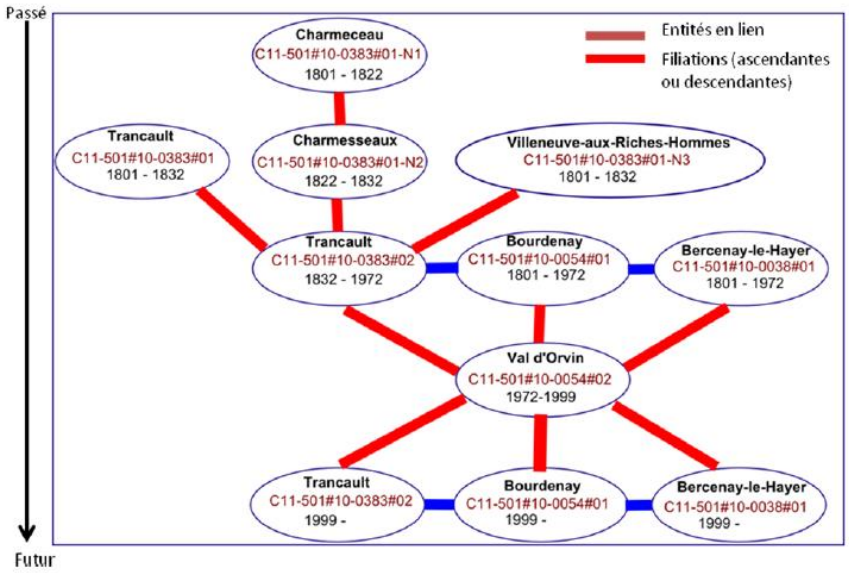


Figure 6. Base Géonomenclature, cas de Charmeceau

La Figure 7 présente une vue générale du modèle. Les entités (*EntiteTerritoriale*) sont associées à leurs attributs datés (*Limite*, *ChefLieu*, *Toponyme*, *Recensement*, *Appartenance*) par leur temps de validité. L'évolution des liens entre les éléments du modèle est retranscrite par des classes d'association datées comme *Gouvernance* pour l'association entre le chef-lieu et l'entité qui est chef-lieu, ou *ZonageComposition* pour l'entité et le(s) zonage(s) dont elle fait partie. Les événements territoriaux (*Evenement*) et leurs conséquences (*Motif*) pour les unités territoriales impliquées sont enregistrés à l'aide d'un type énuméré qui respecte la nomenclature INSEE (*MutationMotif*). Le *CodeHistorique* est unique permet de l'archivage de documents ou d'artefacts associés à chaque version d'entité.

Ce modèle répond à plusieurs préoccupations :

- l'enregistrement de nomenclatures autres que celles de l'INSEE, non forcément strictes et onto, grâce à la réification de l'association *Appartenance*, et la description nécessaire de la composition des zonages (*ZonageComposition*). Dans cet esprit, les niveaux (et leur existence) sont matérialisés par les classes *Subdivision* et *Division*, qui modélisent l'imbrication des niveaux constituant la hiérarchie dans une nomenclature donnée. Le niveau existe tant que le législateur n'en décide pas autrement, indépendamment de sa composition.

- la gestion de données imparfaites, qui se traduit par l'introduction de métadonnées, ainsi que d'un modèle de datation des éléments original,

- la compatibilité avec d'autres bases (celle de l'INSEE et celle de Cassini[22]), via le rappel des identifiants dans ces bases au niveau de l'entité territoriale (*EntiteTerritoriale.numOrdre* et *EntiteTerritoriale.codeINSEE1999*)

- une optimisation des performances et de l'espace mémoire utilisé.

Le schéma intègre une connaissance sur les contributeurs, les sources de données utilisées, et une typologie des lacunes, désignées et référencées par l'attribut « *metadata* ». Ces relations permettent de retracer la provenance de l'information et de la qualifier. Le modèle s'inspire directement de la norme ISO 19115 pour les métadonnées géographiques. À chaque élément du modèle est associé une métadonnée « *ProvenanceInformation* ». Celle-ci indique toujours la date d'enregistrement dans le système des informations, autorisant ainsi un archivage de la connaissance (*dateMiseAJour*), signale si l'information est lacunaire (*lacune* vaut vrai), est une hypothèse (*hypothèse* vaut vrai) et laisse la place au commentaire de l'expert (*commentaire*).

Afin de tenir compte de façon plus rigoureuse et systématique de l'incertitude pouvant exister sur la datation des éléments, nous proposons un type complexe de date, *DateHistorique* qui définit :

- si la date est connue comme un instant ou par un intervalle (attribut *estInstant*),

- si la date précède, est exactement valide à, ou succède à l'instant défini, ou bien appartient à la période définie (grâce à l'opérateur de type énuméré *OperateurDate*).

[22] http://cassini.ehess.fr/cassini/fr/html/1_navigation.php

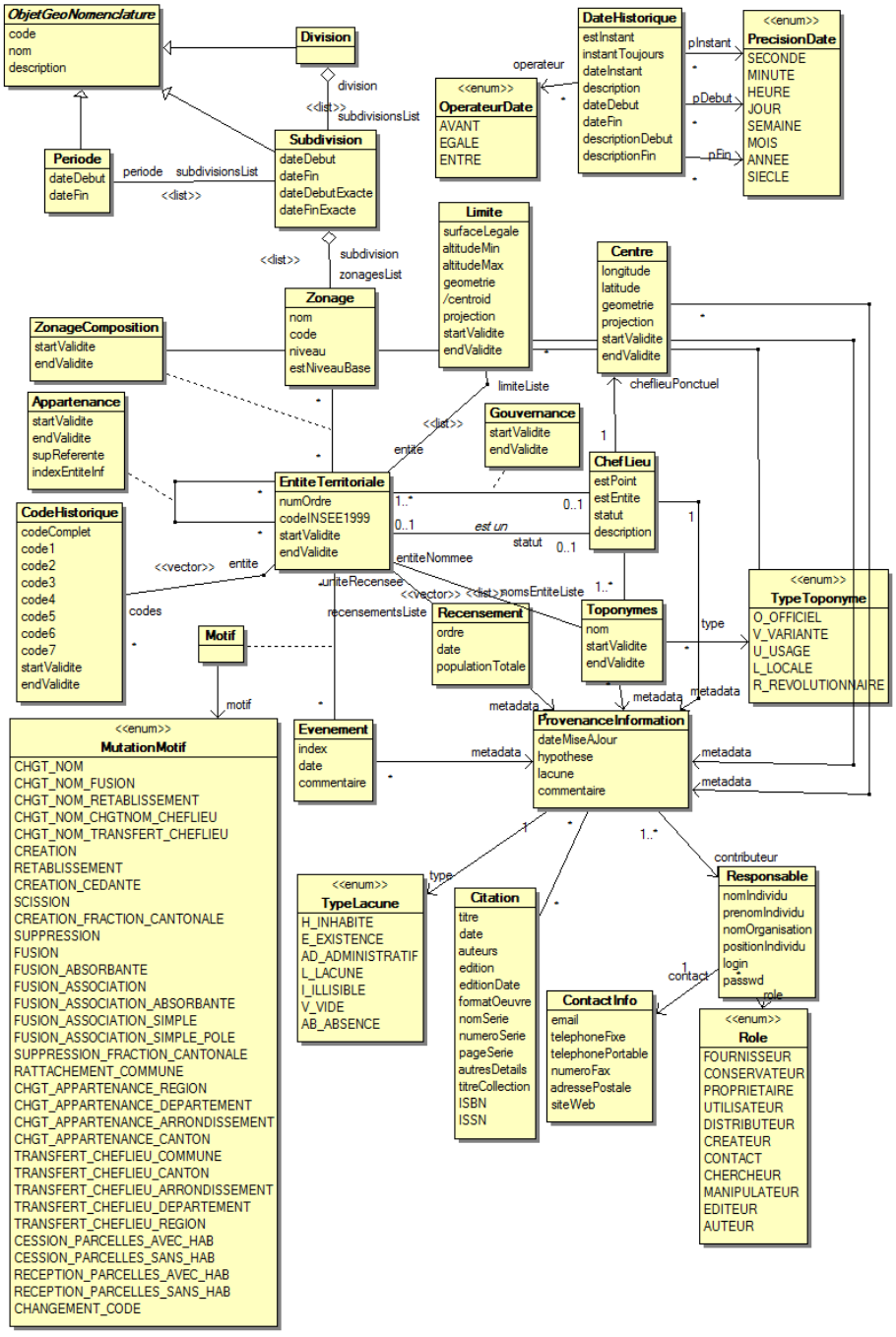


Figure 7. Modèle des données historiques dans Géonomenclature.

L'instant ou l'intervalle sont décrits par des dates de type *Date* du langage SQL, avec une précision variable (jour, mois, année, siècle), définie avec le type énuméré *PrecisionDate*, et éventuellement un champ textuel *description** qui définit un jour (par exemple, 'Sacre de Napoleon premier') ou une période ('L'Empire'). Ce nouveau type de date, identifié uniquement par un entier, est référencé par tous les champs de datation du modèle : l'index étant posé sur cet identifiant défini comme clé primaire de l'objet *DateHistorique*, les performances de recherche sont assurées. De plus, une gain d'espace mémoire important est ainsi réalisé dans la mesure où beaucoup de champs référencent les mêmes dates. *DateHistorique* est initialisé avec deux dates particulières : la date « toujours valide », id = 1, qui signifie que les éléments associés restent encore valides aujourd'hui, et la date « inconnue », id=2, spécifiant expressément que cette date n'est pas connue (c'est une lacune). Afin de faciliter les opérations de tri, de recherche et d'affichage des données associées à ces dates, nous avons également défini des procédures stockées en PL-SQL (leur pré-complication au sein du SGBD garantie leur performance).

5. Conclusion et perspectives

Ce chapitre, qui a pour thème l'évolution des nomenclatures, qu'elles soient administratives ou fiduciaires, religieuses, scientifiques, met d'abord en évidence les enjeux qu'impliquent la compréhension et la connaissance de leurs évolutions. La pluralité des maillages du territoire et leurs changements reflètent des niveaux d'organisation différents de l'espace, qui coopèrent, interagissent ou sont en compétition pour la maîtrise d'un territoire et des sociétés qui l'habitent. Modéliser ces niveaux, les entités territoriales qui les composent et l'évolution des relations entre ces niveaux pose de nombreux problèmes conceptuels et méthodologiques, voire épistémologiques.

Nous avons pour cela recensé le vocabulaire en usage pour décrire l'objet d'étude, la nomenclature et ses composants, les niveaux ou zonages, et les entités qui forment les mailles des zonages. Également, ce chapitre discute de la modélisation des entités et de différents points du vue concernant le fait que cette entité ait ou non une identité intrinsèque, indépendante des attributs qui la décrivent. La façon de décrire le changement à travers la notion d'événement, est influencée par le positionnement épistémologique vis à vis du paradigme identitaire.

Il apparaît que dès lors que l'identité d'une entité est fortement ancrée dans le monde social, il devient indispensable de prendre en compte cette notion dans la structuration des informations. Il s'agit en effet d'une « donnée » à part entière, dont le périmètre ne se limite pas à la simple combinaison « d'attributs » mais bien plus à une réalité sociale dotée d'une histoire et dont les attributs ne constituent que les symptômes visibles.

Qu'elle soit réalisée par le biais d'une expertise humaine et/ou des procédures automatisées, la phase de *construction* des « identités informatiques » correspond à la mise en ordre des informations réunies afin de les rendre intelligibles et opérationnelles. Partant des données, cette phase intermédiaire de création de classes

d'équivalence en adéquation avec une réalité sociale constitue de fait une proto-analyse des concepts et phénomènes observés et de leur complexité. Le résultat de ce travail garantit un accès facilité aux données. Techniquement, les modalités d'interrogation des données sont beaucoup plus simples à gérer et moins coûteuses. C'est aussi le moyen de ne pas laisser l'utilisateur final réaliser lui-même l'ordonnancement indispensable à la compréhension des données. C'est enfin le moyen de préparer la phase d'analyse ultérieure, en mettant à profit cette première étude formelle et les premières hypothèses qu'elle suggère. Il est ainsi possible d'envisager des analyses systématisées concernant les entités identifiées, leur pertinence, l'évolution de leurs attributs, leurs inter-relations et les relations avec les niveaux supérieurs. Mais tout dépend bien entendu des usages et fonctions du système d'information spatio-temporel ainsi créé. Disposer d'un code géographique historique intelligible reposant pour partie sur des objets identifiés sera par exemple un atout pour permettre l'appropriation du système par les utilisateurs. C'est tout aussi important si l'on désire construire des séries statistiques temporelles. Cette opinion, nous la défendons à travers Géonomenclature, un cas d'étude des évolutions de nomenclatures territoriales de tout type en France, pour lequel un modèle répondant aux exigences d'une information imparfaite a été conçu.

Ce type de modélisation, lorsqu'elle se trouve renseignée avec l'histoire des communes de France par exemple, ouvre également des perspectives pour l'analyse de corrélations entre changement topographique, démographique et administratif. En effet, si l'on suppute facilement que les opérations de fusion, ou scission n'ont pas eu lieu au hasard sur le territoire, qu'elles sont le fruit d'évolutions structurelles et démographiques, de luttes d'influence, il serait appréciable d'être en mesure de mesurer l'aspect systémique de ces mécanismes. La base de données topographique vectorielle constituée par le projet GéoPeuple (Ruas et al., 2012) qui autorise une analyse fine des évolutions des structures du territoire (réseau viaire, ferré, bâti, industrie) sera à cet égard fort utile.

Remerciements

Cette réflexion n'aurait pas pu aboutir sans le concours du GdR MAGIS que nous remercions très chaleureusement, ni le soutien bienveillant des différentes équipes de recherche qui ont accueilli les auteurs durant cette longue période de rédaction de trois années, comme Steamer, le COGIT à l'IGN et le LaDéHiS-CRH à l'EHESS.

Bibliographie

Allen, J.F., Towards a general theory of action and time. *Artificial Intelligence*, vol. 23, pp. 123–154, 1984.

Allen, J.F & Ferguson, S., "Actions and events in interval temporal logic", Technical report TR521, Computer Science Department, University of Rochester, 1994.

Article L 2113-1 du code général des collectivités territoriales d'après la loi n° 2010-1563 *http://www.legifrance.gouv.fr/WAspad/UnArticleDeCode?code=CGCTERRL.rcv&art=L 2113-1*

Article L 2113-2 du code général des collectivités territoriales d'après la loi n° 2010-1563 *http://www.legifrance.gouv.fr/WAspad/UnArticleDeCode?code=CGCTERRL.rcv&art=L 2113-2*

Banerjee, S., Davis K.C. Modeling Data Warehouse Schema Evolution over Extended Hierarchy Semantics. *In:* S. Spaccapietra, E. Zimányi and Il-Yeol Song. eds. *Journal on Data Semantics XIII*, 2009, LNCS 5530, pp. 72–96.

Bédard, Y., Lam, S., Proulx M.J., Caron, P.Y. and Letourneau, F., « Data Warehousing for Spatial Data: Reseach Issues », In : *Proceedings of the International Symposium Geomatics in the Era ofRradarsat* (GER'97), 1997 Ottawa pp. 25-30

Brunet R., R Ferras & H Théry. Les mots de la géographie - dictionnaire critique. Reclus-Documentation Française, Paris & Montpellier, 1992.

Brunet R., François J-C. & Grasland C. Entretien avec Roger Brunet : La discontinuité en géographie ; origines et problèmes de recherche. L'Espace Géographique, vol. 4, pages 297–308, 1997.

Chareille P., Rodier X. & Zadora-Rio E. Analyse des transformations du maillage paroissial et communal en Touraine à l'aide d'un SIG. *Histoire & mesure*, vol. 19, no. 3/4, pages 317–344, 2004.

Claramunt C. & Thériault M. Managing Time in GIS : An Event-Oriented Approach. In *Proceedings of the International Workshop on Temporal Databases*, pages 23–42, London, UK, 1995. Springer-Verlag.

Clifford J., Dyreson C.E., Isakowitz T., Jensen C.S., and Snodgrass R.T.. "On the Semantics of Now in Databases." ACM Transactions on Database Systems 22(2), 1997, pp. 215—254

Cheylan J-P & Lardon S. Towards a Conceptual Data Model for the Analysis of Spatio-Temporal Processes : The Example of the Search for Optimal Grazing Strategies. In *COSIT*, pages 158–176, 1993.

Chignoli, R., Crescenzo P. & Lahire P. Lien entre classes dans les langages à objets. Rapport de recherche 97-22, Laboratoire I2S, Université de Nice-Sophia-Antipolis, 1997.

Claval P., Régions, nations, grands espaces. Genin, Paris, 1968.

Clifford J., Dyreson C.E., Isakowitz T., Jensen C.S., and Snodgrass R.T.. "On the Semantics of Now in Databases." ACM Transactions on Database Systems 22(2), 1997, pp. 215—254

Couclelis, H., Ontologies of geographic information, *International Journal of Geographical Information Science*, 24:12, pp. 1785-1809, 2011

Commission sociale des évêques de France, « Église et société face à l'aménagement du territoire », Bayard Éditions, 1998, pp. 157-158.

Dagiral E., Peerbaye A. « Les mains dans les bases de données », *Revue d'anthropologie des connaissances* 1/2012 (Vol. 6, n° 1), p. 191-216. www.cairn.info/revue-anthropologie-des-connaissances-2012-1-page-191.htm. DOI : 10.3917/rac.015.0229.

DATAR, « Aménager le changement », Territoires 2040 n°1, La Documentation française, 2010, 99p.

Frank, A.U., Cheylan J-P., Raper J., Life and motion of socio-economic units. *European Science Foundation GISDATA.* Taylor and Francis, London, May 2001.

Galton A. Fields and Objects in Space, Time, and Space-time. *Spatial cognition and computation*, vol. 4, no. 1, pages 39–68, 2004.

Gantner, F. A Spatiotemporal Ontology for the Administrative Units of Switzerland. PhD thesis, University of Zurich, Switzerland, April 2011.

Grasland, C. Les maillages territoriaux : niveau d'observation ou niveaux d'organisation. In Actes des entretiens J. Cartier, volume 76-77-78 de *Les découpages du territoire*, pages 115–132. INSEE-METHODES, 1998.

Gregory, I., Time-variant GIS Databases of Changing Historical Administrative Boundaries: A European Comparison. *Transactions in GIS*, vol. 6, n°2, pp 161-178 (2002)

Gaudillère B. Atlas historique des circonscriptions électorales françaises, Ecole Pratique des Hautes Etudes, Hautes Etudes médiévales et modernes, 74, Genève, Droz, 1995.

Guibert B., Laganier J., Volle M. Essai sur les nomenclatures industrielles. Economie et Statistique, n°20, 1971

Hornsby, K., Egenhofer M.J. Identity-Based Change Operations for Composite Objects. In T. Poiker & N. Christman, editeurs, *Eighth International Symposium on Spatial Data Handling*, pages 202–213, Vancouver, Canada, July 1998.

Howenstine E. Measuring Demographic Change : The Split Tract Problem. *The Professional Geographer*, vol. 45, no. 4, pages 425–430, 1993.

Jensen C.S., Dyreson C., Böhlen M. & al. The Consensus Glossary of Temporal Database Concepts. *Temporal Databases : Research and Practice*, vol. 1399, pp. 367–405, 1998.

Journal Officiel, Loi 71-588 du Juillet 1971 sur les fusions et regroupements de communes. http://www.legifrance.gouv.fr/jopdf/common/jo_pdf.jsp?numJO=0&dateJO=19710718&pageDebut=07091&pageFin=&pageCourante=07091

Kauppinen, T., Väätäinen, J., and Hyvönen, E. Creating and using geospatial ontology time series in a semantic cultural heritage portal. In Proceedings of the 5th European semantic web conference on The semantic web: research and applications (Berlin, Heidelberg, 2008), ESWC'08, Springer-Verlag, pp. 110–123.

Langran, G.E. Time in geographic information systems. Taylor and Francis, Seattle, WA, USA, 1992.

Lang G., *Le code officiel géographique (COG). Avant, pendant et autour*. Direction de la coordination statistique et des relations internationales, INSEE, 2010, accessible sur http://projetbabel.org/gl/cog_index.htm

Lang G., *La loi n°51-711 du 7 juin 1951 modifiée sur l'obligation, la coordination et le secret en matières statistiques*, Direction de la coordination statistique et des relations internationales, série Documents de travail, n° D2006-01, INSEE, juin 2006.

Mathian H., Piron M. Echelles géographiques et méthodes statistiques multidimensionnelles. In Léna Sanders, editeur, *Modèles en analyse spatiale*. Hermès, 2001.

Mounier, baron de. *Rapport de la commission spéciale chargée d'examiner le projet de loi relatif aux attributions municipales, séance du 19 mars 1835*, Archives parlementaires, cité dans Dalloz, Jurisprudence générale, Répertoire méthodique et alphabétique de législation, de doctrine et de jurisprudence, Paris, 1848, art. Communes, p. 236 et sv.

Motte C., Séguy I., Théré C. *Communes d'hier, communes d'aujourd'hui. Les communes de la France métropolitaine, 1801-2001. Dictionnaire d'histoire administrative*. Institut National d'Études Démographiques, 2003.

Motte C., Pélissier J.-P. *Géonomenclature historique des lieux habités*, Direction des archives de France, 2003.

Motte C., Vouloir M-C., « Frontières administratives et identités communales. Le cas de la France, XVIIIe-XXe siècles », *The Historical Review*, Athènes, vol. V, 2008

Motte C., Évolution des divisions administratives de la France in *Évolution des Divisions administratives en Europe, XIXe-XXIe siècles*. Palgrave Macmillan, collectif à paraître 2013.

Olteanu-Raimond A.-M., Mustière S., Ruas A., 2009, Fusion des connaissances pour apparier des données géographiques, *Revue Internationale de Géomatique* vol.19 n°3, 2009, pp.321-349

Parlement français. Loi n° 2010-1563 du 16 décembre 2010 de réforme des collectivités territoriales *http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000023239624&dateTexte=&categorieLien=id*

Parlement européen. Établissement d'une nomenclature commune des unités territoriales statistiques (NUTS). *Journal officiel*, vol. L 154, no. 0041, 2003.

Pedersen, T.B., Jensen, C.S., Dyreson C.E., A Foundation for Capturing and Querying Complex Multidimensional Data. Information Systems, 2001, vol. 26, pp. 383–423.

Peuquet D. It's About Time : A Conceptual Framework for the Representation of Temporal Dynamics in Geographic Information Systems. *Annals of the Association of American Geographers*, vol. 83, no. 3, pages 441–461, 1994.

Plumejeaud C., Mathian H., Gensel J., Grasland C., Spatio-temporal analysis of territorial changes from a multi-scale perspective, *International Journal of Geographical Information Systems*. vol. 25, n° 11, October 2011, pp. 1597-1612

Plumejeaud C., Gensel J., Villanova-Oliver M., Ben Rebah M., Vergnaud G., Modélisation des hiérarchies territoriales multiples évolutives, *Revue Internationale de Géomatique*, Vol. 21/2, pp. 183-201, ed. Hermès, 2011

Plumejeaud, C. L2.3-4 Intégration des données démographiques et administratives. Rapport technique du projet ANR GéoPeuple, 2012, en ligne sur http://geopeuple.ign.fr/documents/rapports/L2.3-4 Intégration des données.pdf

Praslin, duc de. *Rapport présenté à la chambre des pairs sur le projet de loi relatif à l'organisation municipale, séance du 1er mars 1831*, Archives parlementaires, cité dans Dalloz, Jurisprudence générale, Répertoire méthodique et alphabétique de législation, de doctrine et de jurisprudence, Paris, 1848, art. Communes, p. 230-236.

Pumain D. & Saint-Julien T. L'analyse spatiale - 1. localisations dans l'espace. Collection Cursus. Armand Colin, Paris, 1997.

Pumain D. Les enjeux des zonages. http ://www.sfds.asso.fr/ressource.php ?fct=ddoc&i=679, Avril 2010.

Spery L., Claramunt C. & Libourel T. A Spatio-Temporel Model for the Manipulation of Lineage Metadata. *GeoInformatica*, vol. 5, no. 1, pages 51–70, 2001.

Raffestin, C. Pour une géographie du pouvoir. Litec, Paris, 1980.

Rat Patron, A. *Définition d'un prototype "Géonomenclature" et schéma conceptuel de données*. Rapport remis au Laboratoire de Démographie et d'Histoire sociale, 2011, par la Société DIGITECH.

Rigaux P., Scholl M. Multi-Scale Partitions : Application to Spa- tial and Statistical Databases. In SSD '95 : Proceedings of the 4th International Symposium on Advances in Spatial Databases, numéro 3-540-60159-7, pages 170–183, London, UK, 1995. Springer-Verlag.

Ruas, A., Plumejeaud C., Grosso E., Costes B., Vouloir M-C., Motte C., LeBras H., Cord M., Thome N., Guyomard J., 2012, "GéoPeuple : analyse de l'évolution des territories et de leur démographie sur 200 ans", dans SAGEO'2012, Liège. A paraître.

Smith, B. (1999) Les objets sociaux, Philosophiques, Volume 26, numéro 2, automne 1999, p. 315-347

Spery, L., Claramunt, C., Libourel, T., A Spatio-Temporal Model for the Manipulation of Lineage Metadata. *GeoInformatica*, vol. 5, no. 1, pages 51–70, 2001.

Takahara, Y, D. Macko & Mihajlo D. Mesarovic. Théorie des systèmes hiérarchiques à niveaux multiples. Economica, 1980.

Terrier, C. Les découpages territoriaux : problèmes épistémologiques et méthodologiques. In Violette Rey & Thérèse Saint-Julien, editeurs, Terri- toires d'Europe, la différence en partage. ENS Editions, Lyon, 2005.

Villeneuve P.Y. Un paradigme pour l'étude de l'organisation spatiale des sociétés. Cahiers de géographie du Québec, vol. 16, no. 38, pages 199–211, 1972.

Wachowicz M., Object-oriented design for temporal GIS. Taylor & Francis, Bristol, PA, USA, 1999.

Worboys M.F. Event-oriented approaches to geographic phenomena. *International Journal of Geographical Information Science*, vol. 19, pages 1–28, 2005.

Yuan M. Use of a three-domain representation to enhance GIS support for complex spatio-temporal queries. *Transactions in GIS*, vol. 3, pages 137–159, 1999.

Zadora-Rio Elisabeth dir. - *Des paroisses de Touraine aux communes d'Indre-et-Loire : la formation des territoires*, 34e supplément à la Revue archéologique du centre de la France, Tours 2008