



Un SIG collaboratif pour la recherche historique

Partie 2

Exemple d'application : conception de l'atlas historique numérique et analyses de données attributaires de l'Italie du Risorgimento.

F. BERETTA, C. BUTEZ

LABORATOIRE DE RECHERCHE HISTORIQUE RHÔNE-ALPES – LARHRA UMR 5192

INSTITUT DES SCIENCES DE L'HOMME – ISH

14, AVENUE BERTHELOT - 69363 LYON CEDEX 07

FRANCESCO.BERETTA@ISH-LYON.CNRS.FR - CHARLOTTE.BUTEZ@ISH-LYON.CNRS.FR

Dans une première partie de l'article, nous avons présenté la conception et l'architecture logicielle d'une plate-forme de travail collaborative permettant de traiter des données géo-historiques mutualisées et cumulatives, réalisée en appliquant la méthode *SyMoGIH*. Nous aimerions maintenant illustrer de quelle manière l'historien peut se servir concrètement de cette plate-forme pour stocker, visualiser et analyser ses données. Comme nous avons souhaité nous placer dans une perspective de *web* de données et d'*Open Access*, nous utiliserons des données issues du réseau de bibliothèques animé par la *Bayerische Staatsbibliothek*, qui met à disposition un point d'accès *SPARQL*. Nous ne présenterons pas ici un parcours abouti mais une recherche en cours au sein du Pôle histoire numérique du *Larhra* : il s'agit donc d'un chantier qui soulève plus de questions qu'il n'apporte de réponses et que nous espérons pouvoir mener en collaboration avec d'autres instituts de recherche intéressés par ce sujet.

Un projet d'atlas historique numérique

Parmi les utilisations possibles des SIG en histoire, nous en retiendrons deux : d'une part, le fait de disposer de « *fonds de cartes* » pertinents, c'est-à-dire adaptés à la période étudiée, afin de mettre en scène ses données en évitant l'anachronisme ; d'autre part, la possibilité d'appliquer une démarche d'analyse spatiale permettant de croiser les données spatiales avec les données attributaires. Dans le premier

cas, il s'agira, par exemple, de représenter des batailles ou d'autres épisodes significatifs sur un fonds de carte contemporain aux événements. Dans le deuxième, il s'agira d'étudier la localisation et la configuration de phénomènes politiques, sociaux ou économiques en fonction de l'évolution des territoires politiques qui en sont le théâtre.

Plus particulièrement, nous avons retenu ici le cas de l'Italie au XIX^e siècle, au temps du *Risorgimento*, qui voit un processus d'unification territoriale se réaliser, réunissant

par étapes les différents États de la péninsule issus du Congrès de Vienne (1815) pour les intégrer au sein du nouveau royaume d'Italie, proclamé en 1861. Ce processus est désormais « *visible* » dans son évolution diachronique grâce aux données produites et accessibles sur le site de ressources géo-historiques du LARHRA : <http://geo-larhra.org>.

La mise en place de cet atlas historique numérique se veut une réponse aux deux utilisations des SIG en histoire évoquées précédemment. En effet, pour le

chercheur qui souhaite visualiser et traiter les données géolocalisées qu'il a produites (batailles et autres événements, lieux de naissance ou de publications de livres, etc.) se pose la question de trouver les fonds de cartes numériques adaptés. Une première approche permettant de dépasser la simple projection des données dans *OpenStreetMap*TM ou *Google Maps*TM est celle d'avoir recours aux atlas historiques traditionnels que l'on peut géoréférencer.

Cette démarche permet de visualiser les données et de les explorer en constante relation avec les territoires représentés. Toutefois, il est difficile de produire des cartes numériques avec cette méthode, d'autant plus que les atlas représentent souvent plusieurs couches temporelles sur la même planche ; de surcroît aucune analyse spatiale n'est rendue possible par ce procédé.

Il s'agit donc de chercher des fonds vecteurs adaptés à la période traitée, géoréférencés et, dans l'idéal, libres, modifiables et gratuits. Or, si plusieurs jeux de données vecteurs concernant les territoires des États, produites par les organisations internationales ou les collectivités publiques, sont disponibles pour les données contemporaines, seuls les vecteurs produits par l'entreprise suisse *Euratlas*, disponibles sous licence commerciale, réunissent le géoréférencement des géométries avec le traitement de l'évolution des frontières des États sur deux millénaires. Ces géométries sont très utiles pour une première visualisation de données historiques, mais elles présentent quelques défauts qui les rendent inutilisables pour une analyse plus poussée, du moins dans le sens évoqué ci-dessus. D'une

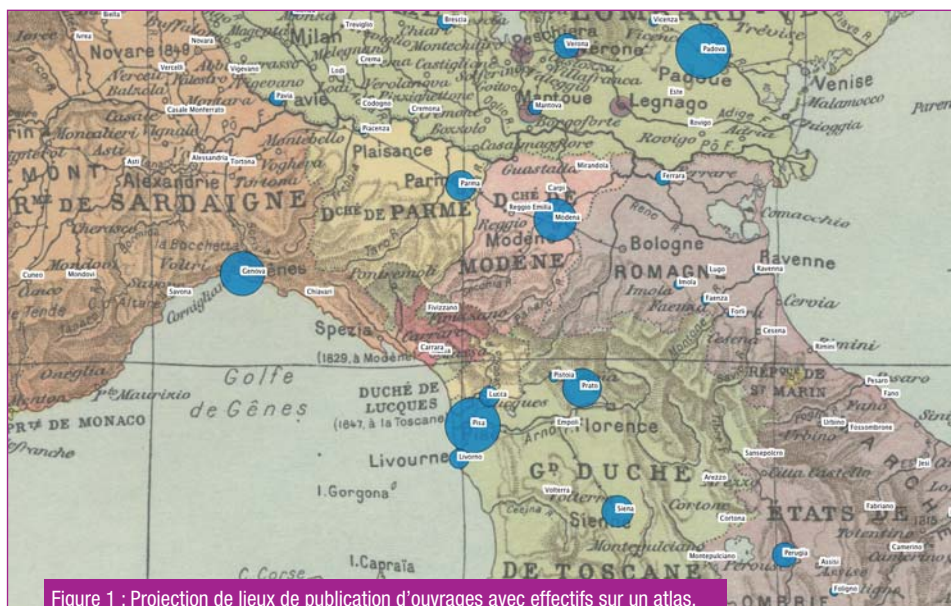


Figure 1 : Projection de lieux de publication d'ouvrages avec effectifs sur un atlas, Vidal de La Blache, Paul, Atlas historique & géographique Vidal-Lablache, p.47.

part, les coupes chronologiques s'effectuent tous les siècles, en 1800, 1920, etc., ce qui ne permet pas de prendre en compte, par exemple, l'évolution des territoires en Italie au XIX^e siècle, qui se caractérise par des modifications significatives autour des guerres d'indépendance (en 1859-1861, en 1866 et en 1870 avec la prise de Rome). D'autre part, la qualité du géoréférencement de ces données empêche leur exploitation pour une analyse spatiale car, dans plusieurs cas, les villes côtières se retrouvent dans la mer et seront donc « perdues » lors de comptage des lieux.

La partie historique de l'atlas

C'est à partir de cette demande des chercheurs, et du constat de l'insuffisance des données disponibles, qu'il nous a paru utile d'entamer la réalisation d'un atlas historique numérique des États dont il s'agit maintenant de décrire la mise en place, puis d'illustrer l'utilisation. La question de départ est donc :

comment procéder pour produire des géométries géoréférencées prenant en compte l'évolution historique des frontières des États ?

La méthode *SyMoGIH* couplée avec les technologies SIG, telle que nous l'avons présentée dans le précédent article, fournit les outils conceptuels et technologiques pour entreprendre la construction de l'atlas envisagé. Dans notre approche, le traitement historique de l'information revêt une priorité absolue, passant avant toute illustration sous forme de géométries : celles-ci ne servent qu'à représenter une réalité géo-historique décrite préalablement par le système d'information historique, c'est-à-dire par une ontologie qui explicite les rapports entre les territoires des États, les différentes formes de leurs frontières, la localisation des lieux habités dans ces territoires, etc.

Par conséquent, dans un premier temps, toute l'attention du chercheur doit porter sur l'identification des États, de leurs territoires, des rapports

politiques entre entités territoriales. Ce travail s'effectue dans la base de données collaborative, ou *base d'hébergement de projets* (BHP). On identifie ou on crée des lieux de type *territoire d'État*. On récolte des informations décrivant leur évolution, qui peuvent aller des rattachements d'un territoire à un État (par exemple un archipel) à des absorptions partielles ou totales d'un territoire par un autre, de l'exercice de la souveraineté sur telle enclave à des événements qui justifient ces modifications. Les matériaux traditionnels de la recherche sont mobilisés : bibliographie, archives, cartes anciennes, atlas historiques. Les atlas peuvent fournir un bon point de départ pour la reconstitution, mais le chercheur doit mener une réflexion critique à leur égard, car on sait que la représentation cartographique est influencée par les conceptions, politiques ou autres, des auteurs ainsi que par des choix et simplifications inévitables.

Une fois ces relations historiques entre lieux décrites dans le système d'information, se pose la question de l'évolution de leur forme effective en fonction des époques, selon la distinction essentielle à notre méthode entre lieux et formes concrètes (cf. *Géomatique expert*, n.91, p. 31-32). C'est à ce moment qu'intervient la question délicate de la notion de frontière à laquelle de nombreux travaux ont été consacrés (voir Références). La notion de frontière a évolué dans le temps pour indiquer une zone ou région sur laquelle s'exercent un ou plusieurs pouvoirs, ou un front idéal occupé par les armées en face d'un État adverse, ou une ligne de démarcation plus ou moins précise séparant les territoires. Dans certains cas,

comme en Italie, la notion (sans le mot) de frontière en tant que limite entre États s'est imposée dès le XVI^e siècle avec la création d'organismes chargés de veiller sur ces délimitations. Mais, dans d'autres cas, il s'agit souvent de régions plus ou moins définies, pouvant comprendre de vastes territoires. C'est à ce moment que se pose également la question de l'échelle, car si on se positionne sur le terrain, à grande échelle, les réalités peuvent être bien compliquées et entremêlées sans qu'on puisse ramener une frontière à une ligne.

Pour produire les formes concrètes des États nous avons donc opéré plusieurs choix : nous nous sommes placés à l'échelle d'un atlas historique, en admettant les simplifications qu'impose ce type de traitement ; nous avons veillé à ce que les localités appartenant à un État se trouvent bien intégrées dans son territoire, en renseignant ces inclusions par des informations ; nous avons choisi l'instant du passage d'une forme concrète à une autre en nous référant, en règle générale, à la date officielle du traité de paix correspondant, même si l'on sait que, souvent, les passations de pouvoir ont lieu avant ou après les traités. Une plus grande finesse dans le traitement des chronologies, de même que celle de l'évolution du territoire mesurée sur le

terrain, relèvent d'une échelle plus grande, qu'il n'est pas possible de réaliser au niveau d'un atlas. En même temps, le système d'information permet de stocker et de documenter tout détail supplémentaire, qui pourra être retrouvé plus tard. Concrètement, toutes ces informations, ces objets (lieux, formes concrètes) ainsi que leurs associations sont saisis directement via l'interface *web* du système d'information.

Par exemple, pour les États pontificaux, nous avons créé deux formes concrètes qui les « *instancient* » : respectivement l'une valable du 9 juin 1815 (acte final du Congrès de Vienne) jusqu'au 4 novembre 1860 (veille du vote par les populations des Marche et de l'Ombrie de leur rattachement au royaume de Sardaigne), et l'autre du 5 novembre 1860 (le jour du referendum) jusqu'au 19 septembre 1870 (veille de la prise de Rome).

La production des géométries

Jusqu'ici, l'ensemble des informations n'a été traité qu'au niveau du système d'information géo-historique, tout en munissant les données des indispensables références aux sources et à la bibliographie. Ce n'est qu'à ce point qu'intervient la

BHP - Interface de gestion des données

Objets | Bibliothèque de connaissances | Références | Associations | Gazetteers | Navigation | Déconnexion

Etats pontificaux (734 - 1870) - NAF12369 [Mémoriser la clé] [Ajouter une information]

Identification | Propriétés | Localisation | Classes | Associations | Formes concrètes | Documentation et Web | Commentaires | Projets | Annotations

(Retour à la liste)

Liste des formes concrètes associées

#	Forme concrète	Clé Cifs	Est associé	Début de l'association	Fin de l'association	Raison de l'association	Notes	Créateur	Date de création
102	Etats pontificaux (1860-1870)	Caf24126	Non	1860-11-05	1870-09-19	Représentation du territoire des Etats Pontificaux de 1860 à 1870		valentin	2010-07-02 11:4
103	Etats pontificaux (1815-1860)	Caf24102	Non	1815-06-09	1860-11-04			valentin	2010-11-23 11:4

20 | Page 1 of 1 | Affichage de 1 à 2 sur un total de 2

Figure 2 : Saisie des formes concrètes des Etats pontificaux dans l'application web, copie d'écran.

production des géométries : leur fonction consiste, comme nous l'avons dit, à fournir une représentation spatiale géoréférencée des données stockées dans le système d'information. En fonction des différents cas de figure, on pourrait choisir pour représenter les frontières une ligne (si la limite est connue), un polygone (pour un territoire ou une marche) ou des points avec un *buffer* (en cas de zones dont l'étendue est incertaine). En Europe, au XIX^e et XX^e siècles, et en particulier en Italie, il a paru judicieux d'utiliser des lignes, puisque la notion de frontière, à l'échelle et selon la méthode indiquée ci-dessus, se prête bien à ce type de traitement. Pour les géométries servant de base au travail, nous avons eu recours à des données contemporaines libres de droits et gratuites, choisies en fonction d'un degré de précision et d'un 'poids' informatique correspondant à l'échelle retenue.

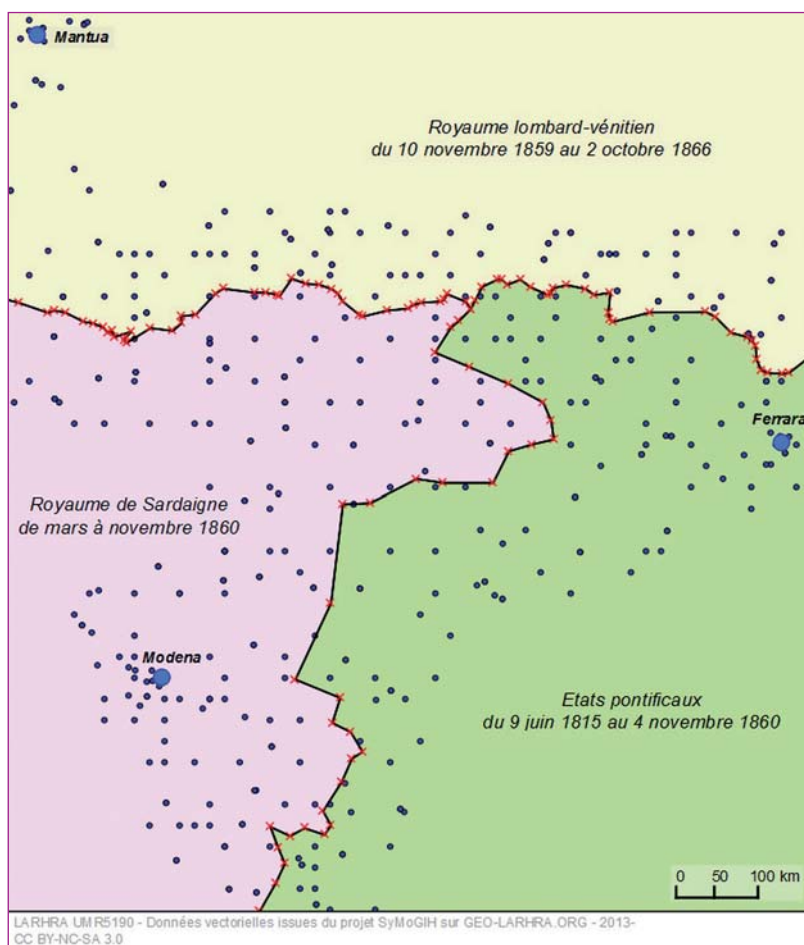


Figure 3 : Édition d'une frontière en Italie. Les lieux habités sont représentés par des points de couleur bleue.

Dans la production des géométries, nous avons adopté une démarche qui consiste à remonter dans le temps, c'est-à-dire à prendre comme point de départ les découpages contemporains tout en les modifiant au fur et à mesure que l'on recule dans le passé. Chaque modification des formes concrètes donne lieu à une mise à jour des géométries, souvent par section de lignes, et de leurs associations avec les formes concrètes correspondantes. Ce travail est effectué en se connectant directement avec le logiciel *QuantumGIS* sur la base de données *PostgreSQL/PostGIS* qui contient les géométries, en utilisant quelques fonctions *ad hoc* écrites en *pl/pgSQL* (*Procedural Language/PostgreSQL Structured Query*

Language) qui assurent la mise à jour des associations et la cohérence topologique des données.

La Figure 3 illustre les différents niveaux de cette démarche, parfois complexe. Pour représenter les territoires de cette partie de l'Italie jusqu'au 4 novembre 1860, date retenue pour la modification des formes concrètes, il faut articuler les représentations de trois territoires qui possèdent des formes n'ayant pas les mêmes bornes chronologiques. Les géométries qui représentent ces formes concrètes sont dessinées à partir des données géo-historiques précédemment stockées dans la BHP, en tenant compte en particulier de l'appartenance de telle ou telle localité au territoire en question.

Selon la méthode présentée dans la première partie de l'article (p. 33), les géométries ainsi produites sous forme de lignes sont ensuite regroupées en polygones, à l'aide de fonctions et de vues implémentées dans la base *PostgreSQL*, afin d'exprimer les relations entre territoires et formes concrètes décrites dans le système d'information géo-historique. L'atlas ainsi conçu représente les évolutions chronologiques déterminées par les moments historiques significatifs retenus par les chercheurs et documentées par des informations qui apparaissent lorsqu'on glisse la souris au-dessus des dates correspondantes. De plus, ces mêmes géométries peuvent être téléchargées par les utilisateurs du système sous forme de *shapefiles*.



Figure 4 : Principales batailles ayant impliqué les Zouaves pontificaux (données de Laurent Gruaz).

Deux exemples d'application

Nous donnerons ici un exemple pour chacun des deux usages d'un atlas historique numériques évoqués ci-dessus. D'une part, il faut disposer d'un fonds de carte chronologiquement pertinent afin d'illustrer des événements historiques. Dans le cadre de la thèse de doctorat de Laurent Gruaz, ayant pour sujet la *Prosopographie des officiers français engagés dans les Zouaves pontificaux (1860-1870)*, il était question, entre autres, de localiser les principales batailles ayant impliqué ce régiment constitué de volontaires français, mais aussi belges et néerlandais, venus défendre l'État pontifical contre les partisans de l'unification italienne.

La carte ainsi réalisée (figure 4) tire pleinement profit de la représentation géométrique des deux formes concrètes décrites

ci-dessus, correspondant aux étendues différentes du territoire des États pontificaux, qui subit une première perte en 1860 et disparaît définitivement en 1870, suite à la prise de Rome. La démarche d'illustration est ici analogue à celle des atlas historiques traditionnels mais le fonds de carte peut s'adapter aisément à la chronologie en fonction des événements représentés.

D'autre part, l'atlas historique numérique permet d'effectuer des analyses spatiales. Pour cet exemple, nous avons eu recours aux données accessibles sur *Internet* afin d'illustrer notre méthode. Il s'agit aussi d'attirer l'attention sur l'intérêt que représente, pour la recherche, la possibilité d'intégrer dans un même contexte des données attributaires et spatiales produites par les chercheurs avec celles récupérées *online* grâce à des producteurs s'inscrivant dans la mouvance *open data*.

Le *Bibliotheksverbund Bayern* est un réseau de plus de cent cinquante bibliothèques allemandes réunies autour de la *Bayerische Staatsbibliothek*, qui met à la disposition du public un point d'accès *SPARQL* permettant d'interroger le catalogue collectif. Grâce à une série de requêtes effectuées depuis un serveur *Fuseki* (projet *Jena d'Apache*), nous avons pu extraire l'année et le lieu de publication de quelque cinquante mille ouvrages publiés en langue italienne entre 1816 et 1915, conservés dans les bibliothèques du réseau bavarois. Grâce au *gazetteer* qui fait partie de la BHP (cf. le site web <http://geo-larhra.org>), nous avons pu géolocaliser en très peu de temps plus de quarante-cinq mille livres, moyennant une jointure sur le nom du lieu d'édition.

Étant donné le caractère purement illustratif de notre démarche, nous n'avons retenu que les lieux comportant au moins cinq occurrences d'ouvrages et avons renoncé à une analyse exhaustive. Cette population de livres ne reflète, en effet, que la politique d'achat et de conservation des bibliothèques bavaroises, et seulement indirectement celle de la production d'ouvrages en Italie à l'époque en question.

Nous avons regroupé les livres par lieu d'édition et par période de cinq années (lustres). Ce regroupement montre que, parmi les lieux d'édition des livres en langue italienne détenus par les bibliothèques bavaroises, sept villes totalisent plus de deux mille ouvrages édités sur un siècle. Ce groupe de tête se détache nettement des suivants, la huitième place, Palerme, ne totalisant que sept cent quarante ouvrages. Dans l'ordre décroissant, on trouve à la première place Milan,

avec plus de huit mille ouvrages, puis Florence, Rome, Turin, Venise, Naples et Bologne. On relèvera la nette progression de l'effectif des livres imprimés à Rome après sa proclamation comme capitale du royaume d'Italie, en 1871 (figure 5, ligne rouge).

La problématique suivante consiste à pouvoir distinguer les livres édités sur le territoire italien de ceux édités à l'étranger, hors de la péninsule ou des régions italophones. On va ici recourir à une première analyse spatiale très simple, qui consiste à définir grâce au logiciel QGIS, par croisement entre ces points et le polygone du territoire, l'ensemble des lieux d'édition se situant tant à l'intérieur qu'à l'extérieur de l'État italien, dans sa forme de 1914, date retenue comme l'aboutissement du processus d'unification du Risorgimento, à la veille de la Première guerre mondiale.

Il est ainsi possible de calculer la part des ouvrages en langue italienne publiés dans le reste de l'Europe. Si leur volume est peu élevé (il se monte à 6 % environ de l'ensemble des ouvrages conservés dans les bibliothèques du réseau bavarois), l'observation de la distribution par lustres des effectifs de ces différentes catégories (Figure 6) permet de relever quelques moments significatifs, autour des mouvements révolutionnaires de 1830 et de 1848. On assiste à cette époque à un fléchissement du nombre des ouvrages publiés en Italie (ligne orange) et à une augmentation concomitante des effectifs des ouvrages publiés hors du territoire du futur État italien (ligne rouge). La nature des données ne permet pas de savoir si ce phénomène relève de la politique des achats

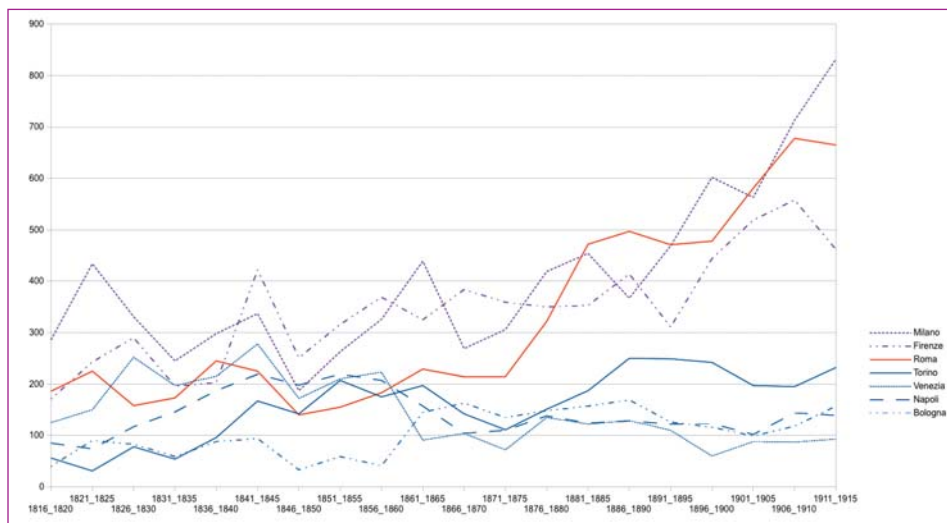


Figure 5 : Distribution par lustres des livres issus de villes avec un effectif dépassant les deux mille ouvrages.

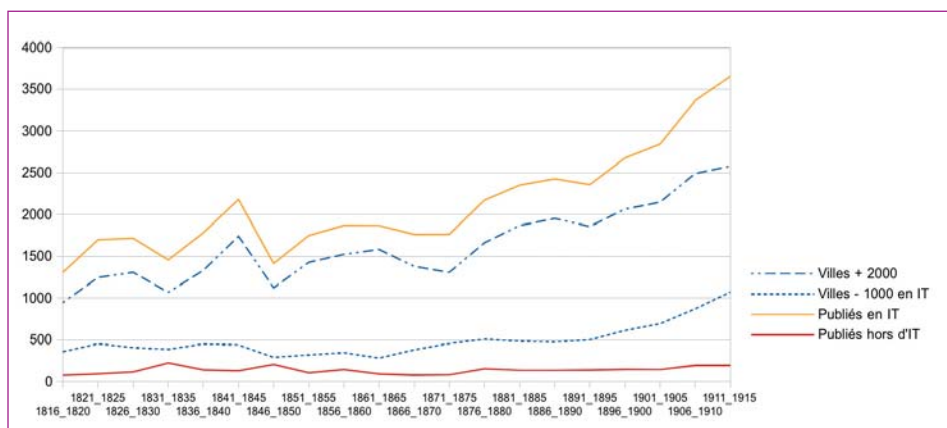


Figure 6 : Distribution des ouvrages regroupés en fonction de l'effectif et de la publication sur le territoire italien.

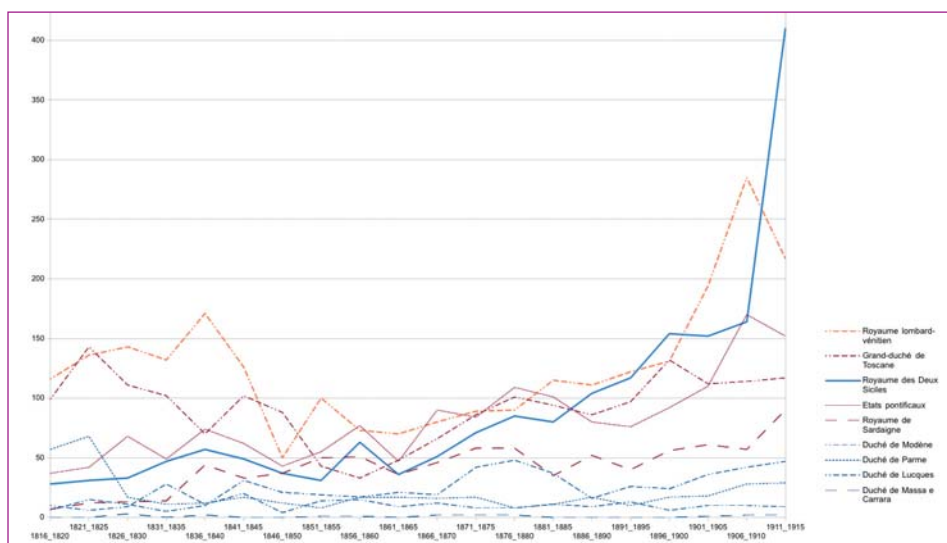
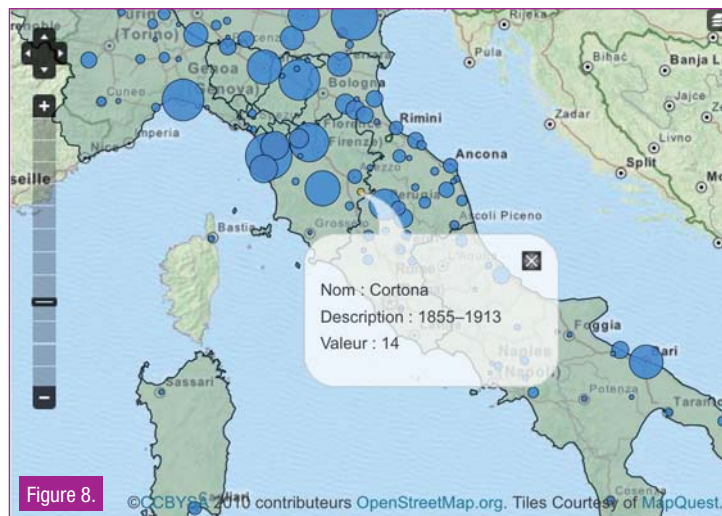


Figure 7 : Évolution des effectifs des villes périphériques regroupées par État d'appartenance.



des bibliothèques allemandes ou de la production effective des livres, mais le fait mérite d'être relevé.

Une fois défini l'ensemble des lieux de publication se situant sur le territoire italien, on peut opérer une analyse spatiale dans QGIS et croiser les lieux de publication avec la structure des États en 1816, suite à la réorganisation de l'Italie après le Congrès de Vienne (figure 7). Ce regroupement du nombre de publications par État ne prend pas en compte les sept villes principales, qui forment un ensemble à part, et résume donc les données des publications issues des centres périphériques ou secondaires. Si on observe l'évolution de ces centres durant un siècle, on remarque que l'unification de l'Italie, accomplie en 1870, n'a pas eu d'impact important sur l'effectif des livres issus de ces régions sauf pour ce qui concerne le territoire de l'ancien royaume de Naples et de Sicile (ligne bleue). La forte croissance constatée après 1925 s'explique, comme le montrent les données, par l'essor de l'éditeur Laterza à Bari. Il importe de souligner, à ce sujet, la fonction heuristique et non explicative de cette démarche : pour l'historien, il s'agit de visu-

aliser des phénomènes, de les rendre perceptibles, mais il faudra ensuite en chercher les causes dans l'historiographie et dans les sources.

Notre démarche n'avait pour but que d'illustrer les avantages et le potentiel que présente l'utilisation d'un atlas historique numérique. Pour notre projet, il s'agit maintenant d'avancer dans la production de l'atlas, d'abord pour le continent européen et pour l'époque contemporaine, puis au-delà. Mais un tel objectif ne peut pas être atteint par une seule institution de recherche, en particu-

lier lorsqu'on aborde le traitement de régions géographiques dont l'évolution territoriale est particulièrement riche comme l'Allemagne ou l'Europe orientale. Pour cette raison, nous espérons arriver à fédérer d'autres producteurs de données géo-historiques dans un réseau européen, afin de réaliser collectivement cet important projet. Nous souhaiterions également que l'on puisse placer ces données à disposition des chercheurs et du public, dans une logique d'*open access*. La possibilité de projeter directement ses propres données attributaires géolocalisées sur une carte du territoire italien après 1815 dans l'atlas disponible sur le site *geolarhra*, ainsi que le montre la figure 8, constitue un premier pas dans cette démarche.

La méthode *SyMoGIH* et la plateforme collaborative ainsi mises en place fournissent à l'historien une base solide conforme aux standards technologiques actuels, pour qu'il puisse tirer pleinement profit dans sa recherche des potentialités offertes par ces outils numériques. |

Références

Atlas historique de l'Italie sur www.Geo-Larhra.org :

<http://geo-larhra.ish-lyon.cnrs.fr/?q=atlas-historique/regroupement-de-territoires/evolution-des-territoires-en-italie>

Parmi la littérature abondante sur la notion de frontière mentionnons :

- ▶ Stopani Antonio : *La production des frontières, États et communautés en Toscane (XVIe-XVIIIe siècle)*. Rome, École française de Rome, 2008 ;
- ▶ Nordman Daniel : *Frontières de France: de l'espace au territoire : XVIe-XIXe*. Paris, Gallimard, 1998 ;
- ▶ Carl Grundy-Warr (éd.) : *World boundaries, t.3 : Eurasia*. New York, Routledge, 1994.

Données vectorielles contemporaines utilisées :

Eurostat Gisco Données Nuts, régions 2006 (échelle 1:10 000 000), accessibles sur le portail <http://epp.eurostat.ec.europa.eu>.

Les instructions d'interrogations SPARQL sont disponibles sur <http://lod.b3kat.de/doc/sparqlguide.html>.

Enfin, le serveur *Fuseki*, projet Jena d'Apache http://jena.apache.org/documentation/serving_data.