
Des clochers d'Ancien Régime aux paroisses constitutionnelles et premières communes (1789-1804)

Construction et peuplement d'une base de connaissances sur les redécoupages administratifs durant la Révolution française

Mots clés : interdisciplinarité, Web sémantique, extraction d'information à partir de textes, données géohistoriques, évolution du territoire.

A. Contexte

Les ressources cartographiques et iconographiques anciennes sont de plus en plus intégrées dans des Systèmes d'Information Géographique (SIG) historiques pour servir de référentiels spatiaux, supports d'études sur l'évolution du territoire (Duménieu, 2015). Toutefois, ces travaux délaissent souvent les archives textuelles qui sont pourtant des sources d'information sur l'espace ancien majeures (parfois même ce sont les seules existantes). Si de nombreux travaux existent pour annoter et spatialiser manuellement le contenu de documents textuels anciens (par exemple <http://commons.pelagios.org/>), très peu s'intéressent à l'extraction et la structuration automatique d'informations géographiques.

Ce stage s'intéresse à une période importante de la Révolution française : la redéfinition partielle des paroisses religieuses au moment de la mise en place de l'Église constitutionnelle (1791). Des textes d'archive permettent de reconstituer très précisément ce remodelage à partir du tissu paroissial d'Ancien Régime, duquel émergeront plus tard les premières communes. L'extraction et le géocodage des entités géographiques citées dans ces textes, ainsi que la modélisation formelle de la réorganisation paroissiale permettra, pour la première fois, de cartographier et d'analyser quantitativement ce phénomène.

B. Principales hypothèses et objectifs

L'extraction d'information à partir de textes consiste à mettre en œuvre des techniques de traitement automatique du langage naturel (TALN) ou de fouille de textes pour en extraire des informations structurées que l'on souhaite mettre à profit pour une application donnée (recherche d'information, analyses, visualisation, etc.). L'évolution récente de ces approches d'extraction accompagne celle du Web de données qu'elles contribuent à alimenter en fournissant des outils pour la création automatique d'ontologies, le peuplement de bases de connaissances ou encore l'interrogation de bases de connaissances en langage naturel (Gangemi, 2013).

Ce stage vise à explorer les potentialités de l'extraction d'information à composante spatiale dans des textes pour construire et peupler une base de connaissances sur l'évolution du découpage administratif à partir de textes réglementaires décrivant les remembrement des

paroisses d'Ancien Régime. Les informations disponibles dans les textes sont à la fois très riches et partielles. Elles sont donc difficilement intégrables dans les modèles de données des SIG qui requièrent des données fortement structurées a priori. La convergence des possibilités offertes par l'extraction d'information et par les modèles de gestion de connaissances du Web sémantique, fondés sur l'hypothèse du monde ouvert, offre donc un cadre plus propice à l'extraction et la modélisation de ce phénomène.

Des résultats convaincants pourront aboutir à la rédaction d'un article scientifique à la suite du stage.

C. Verrous scientifiques

Il s'agira d'extraire, typer, désambiguïser et structurer les informations sur les entités spatiales nommées décrites par les textes (noms, relations de filiation, d'inclusion ...) pour les intégrer dans une base de connaissances et vérifier la cohérence des informations extraites, inférer de nouveaux faits, répondre à des requêtes sur le découpage administratif et son évolution. Ceci suppose de proposer des solutions pour :

- Adapter les approches existantes d'extraction d'information à composante spatiale (Moncla, 2015 ; Zenasni et al., 2015 ; Kergosien et al., 2014) à partir de textes à des corpus anciens caractérisés par un langage soutenu et une structuration juridique - portions de textes courtes, très structurées, utilisant un vocabulaire précis;
- Représenter, stocker et manipuler des informations à composante spatiale qualitative (références spatiales indirectes, positionnement relatif, relations de filiation, etc.) selon les standards du Web de données (Atemezing et al., 2014) ;
- Désambiguïser les entités géographiques extraites (Paris et al., 2017), c'est à dire :
 - ◆ Typer les entités géographiques extraites selon qu'il s'agit de paroisses de l'Ancien Régime, de paroisses constitutionnelles ou autre (hameau, domaine, etc.).
 - ◆ Lier les paroisses de l'Ancien Régime aux lieux nommés extraits de la carte de Cassini,
 - ◆ Lier les paroisses constitutionnelles avec les données de la base de données de l'historique des communes de France (LaDéHiS).
- Détecter et corriger d'éventuelles incohérences spatiales ou temporelles dans les informations extraites, améliorer le typage des entités géohistoriques.

D. Corpus de travail

L'Église romaine condamnant la Révolution, une Constitution Civile du Clergé fut mise en place afin de réorganiser les rapports entre les prêtres, les paroisses et la monarchie constitutionnelle française. Cela se concrétisa notamment par une redéfinition territoriale des évêchés, correspondant désormais aux nouveaux départements, et des paroisses. Le déclassement

d'une cinquantaine de sièges épiscopaux et la suppression de « plusieurs centaines de paroisses »¹ ont ainsi démultiplié les fermetures d'églises, notamment dans les villes où, suivant le nouvel ordre géométrique et égalitaire des choses, le nombre de paroisses et de succursales se trouvait généralement « disproportionné » par rapport à celui de la population résidente (Bacocchi et Julia, 2009). Les cités épiscopales et les grandes villes riches en sanctuaires, tout comme les petites villes et les bourgs de moins de six mille habitants obligés de réduire à une seule leurs paroisses, furent tout particulièrement affectés par ces mesures législatives². Les décrets de circonscription des paroisses sont disponibles dans les fonds des archives parlementaires et présentent un intérêt particulier dans la mesure où ils s'appuient, pour la remembrer, sur l'organisation paroissiale d'Ancien Régime. Or, celle-ci est mal connue: on la devine, sans réellement en connaître les contours, à partir des clochers représentés sur les cartes des Cassini. De plus, ces décrets donnent un premier aperçu sur les liens entre paroisses et communes. En effet, les paroisses d'Ancien Régime supprimées, les registres paroissiaux désormais sécularisés (on parle depuis d'état-civil) passèrent aux chefs-lieux des paroisses constitutionnelles.

Dans un premier temps, on travaillera sur les textes de quatre séances de l'Assemblée Nationale qui comprennent, entre autres, la circonscriptions des paroisses du Puy-de-Dôme (Fig. 1).

Cette zone du Puy-de-Dôme a fait l'objet d'une vectorisation complète à partir de la carte de Cassini (Fig. 2), par le groupe de recherche GeoHistoricalData³ : on dispose désormais d'une base de données géographique vectorielle de référence sur cette zone où sont décrits les chefs-lieux de paroisses de l'Ancien Régime, leurs succursales, ainsi que d'autres types de lieux nommés (hameaux, villages, oratoires). Celle-ci sera mise à profit pour désambiguïser les mentions d'entités spatiales nommées extraites des textes.

¹ L'ordre de grandeur est celui donné par Timothy Tackett, *La Révolution, l'Église, la France*, trad. de l'américain par Alain Spiess, Paris, Éditions du Cerf, 1986 [1985], p. 31. Si l'importance des suppressions est très mal connue à l'échelle nationale, on dispose cependant pour les paroisses rurales d'une étude d'ensemble (Hildesheimer 1991) qui permet de prendre la mesure des réactions locales face aux projets du Comité de division.

² Un dépouillement systématique des décrets de circonscription paroissiale entérinés par l'Assemblée nationale fait apparaître, en première approximation, que jusqu'au mois de septembre 1791, et suivant un rythme qui s'est accéléré au printemps 1791, l'ensemble des territoires ayant fait l'objet d'une procédure aboutie de circonscription constitutionnelle a perdu environ 60 % de ses paroisses d'Ancien Régime. Cette réduction porta en priorité sur les régions Nord, Ile-de-France et Centre, et toucha tout particulièrement les cités épiscopales et les localités de moins de 6 000 habitants dont les paroisses urbaines furent systématiquement réduites à une seule. Voir A. P., t. 20, p. 351-352 ; t. 22, p. 101-102, 422, 476-477, 516-517, 739-744 ; t. 23, p. 112, 172, 221-223, 651, 657-658 ; t. 24, p. 31, 86-87, 143-144, 292-293, 493-494, 559-560, 578 ; t. 25, p. 1-2, 232-233, 235, 326-327, 375-376, 412-413, 432, 553, 555-556, 575-576, 864 ; t. 26, p. 29-30, 575-576, 694-697 ; t. 27, p. 139-140, 189, 251-253, 759-763 ; t. 28, p. 595-596 ; t. 29, p. 259-260, 472-477, 641-644 ; t. 30, p. 25-26, 93, 303-304, 559-561, 627-629 ; t. 31, p. 123-129.

³ <http://geohistoricaldata.org/>

Assemblée nationale, séance du mercredi 1^{er} juin 1791, au soir, M. Lanjuinais, au nom du comité ecclésiastique, présente un projet de décret sur la circonscription des paroisses de plusieurs villes et bourgs de divers départements, Art. 25. Département du Puy-de-Dôme. District de Clermont, *Archives Parlementaires*, 1^{re} série, t. 26, 1887, p.696a-697a

Assemblée nationale, séance du mercredi 15 juin 1791, au soir, Un membre du comité ecclésiastique propose un projet de décret relatif à la circonscription des paroisses des villes de Lyon, Le Puy, Compiègne, Châtillon-sur-Indre et Chambly, et des districts de Riom, Ambert et Beaugency, *Archives Parlementaires*, 1^{re} série, t.27, Art.7. Département du Puy-de-Dôme, district de Riom, p.252a et p.253a.

Assemblée nationale, séance du mardi 16 août 1791, au soir, Projets de décrets relatifs à la circonscription de diverses paroisses, Deuxième projet, Département du Puy-de-Dôme, district de Billom, *Archives Parlementaires*, 1^{re} série, t.29, Art 1 à 4, p. 473a-b et p. 475a-b

Assemblée nationale, séance du mercredi 21 septembre 1791, Rapport relatif à la circonscription de plusieurs paroisses... 1. Décret relatif à la circonscription des paroisses du district de Besse (Puy-de-Dôme), *Archives Parlementaires*, 1^{re} série, t.31, p.123a-124a et p.129a

« Nohanent qui comprendra, outre son ancien territoire, les hameaux de Chanat, la Moutaire, Lagnat, Lasgoutas et l'Etang.
 « Olby.
 « Olloix.
 « Omme.
 « Orcet, qui comprendra le Cendre. Il y aura au Cendre un oratoire.
 « Orcines.
 « Orcival, à laquelle sera réunie celle de Saint-Bonnet, dont l'église sera conservée comme succursale.
 « Planzat.
 « Pont-du-Château, dont les deux paroisses seront réunies sous l'invocation et dans l'église de Saint-Martin.
 « Prondioes, qui comprendra, outre son ancien territoire, celui du lieu de Peyrol. Il y aura à Peyrol une succursale.
 « Queuille (1a), à laquelle est réunie la paroisse de Perpezat, en ce qui n'est pas réuni à Rochefort. Il y aura une succursale à Perpezat.

Extrait de l'article 25, "Département du Puy-de-Dôme. District de Clermont", Assemblée nationale, séance du mercredi 1^{er} juin 1791, au soir, M. Lanjuinais, au nom du comité ecclésiastique, présente un projet de décret sur la circonscription des paroisses de plusieurs villes et bourgs de divers départements, *Archives Parlementaires*, 1^{re} série, t. 26, 1887, p.696b

Fig. 1: Inventaire et extrait des sources parlementaires concernant le Puy-de-Dôme



Fig. 2: Extrait de la feuille 52 de la carte de Cassini sur le Puy-de-Dôme
<http://catalogue.bnf.fr/ark:/12148/cb40860545f>

E. Compétences et formation requises

Formation:

Master 2 ou troisième année d'école d'ingénieur en informatique ou en géomatique avec une forte composante informatique.

Compétences et connaissances:

- Un bon niveau en programmation Java est essentiel.
- En raison de la nature du corpus de documents, la maîtrise du français est nécessaire.
- Extraction d'information à partir de textes.
- Données géographiques vectorielles.
- Web de données, notamment RFDS et OWL.
- Un intérêt pour les sciences sociales est un plus.

Durée et période de stage :

5 mois, au cours du printemps et de l'été 2018.

Lieu du stage :

Equipe LaSTIG/COGIT de l'Institut National de l'Information Géographique et Forestière (IGN), à Saint-Mandé (métro 1, station Saint Mandé).

Le stage se déroulera dans l'équipe COGIT menant des recherches en géomatique sur l'intégration, la diffusion, la représentation et l'analyse de données géographiques sous forme de référentiels vectorisés et à grande échelle.

Indemnités de stage :

Stage gratifié selon la législation française.

Modalités de candidature :

Envoyer CV, lettre de motivation ciblée sur le sujet et relevés de notes des 2 dernières années d'études par email au format PDF et en un seul fichier.

Encadrement du stage :

- Nathalie Abadie [COGIT/IGN] : nathalie-f.abadie@ign.fr
- Éric Kergosien [GERiiCO/SID/Université Lille 3] : eric.kergosien@univ-lille3.fr
- Bertrand Duménieu [CRH/EHESS] : bertrand.dumenieu@ehess.fr
- Stéphane Baciocchi [CRH/EHESS] : stephane.baciocchi@ehess.fr

F. Bibliographie

Atemezing, G.A., N. Abadie, R. Troncy and B. Bucher (2014) Publishing Reference Geodata on the Web: Opportunities and Challenges for IGN France. , Terra Cognita 2014, 6th International Workshop on the Foundations, Technologies and Applications of the Geospatial Web. In Conjunction with the 13th International Semantic Web Conference.

Baciocchi S., D. Julia (2009), « Reliques et Révolution française (1789-1804) », in P. Boutry, P-A. Fabre et D. Julia (eds.), *Cultes et usages chrétiens des corps saints des Réformes aux révolutions*, volume 2, Paris, éditions de l'École des hautes études en sciences sociales, coll. « En temps & lieux », p. 483-585

Duménieu, B. (2015) Un système d'information géographique pour le suivi d'objets historiques urbains à travers l'espace et le temps. Thèse en mathématiques et applications aux sciences de l'homme de l'École des Hautes Etudes en Sciences Sociales.

Gangemi, A. (2013) A comparison of knowledge extraction tools for the semantic web. In Extended Semantic Web Conference, pp. 351-366. Springer Berlin Heidelberg.

Hildesheimer, F. (1991) "La circonscription des paroisses (1790-1793)", in *La naissance et les premiers pas des départements, actes du colloque de Compiègne, 27-28 octobre 1990*, Beauvais, Archives de l'Oise, 1991, p. 279-308

Kergosien, E., Laval, B., Roche, M., Teisseire, M. (2014) Are opinions expressed in land-use planning documents? International Journal of Geographical Information Science, vol. 28, issue 4, jan. 2014. [Rank A, IF: 1.61 in 2012].

Moncla, L. (2015) Automatic Reconstruction of Itineraries from Descriptive Texts. Thèse de doctorat de l'Université de Pau et des Pays de l'Adour et de l'Université de Saragosse.

Paris, P-H, Abadie, N., Brando C. (2017). Linking spatial named entities to the Web of Data for geographical analysis of historical texts. Journal of Map & Geography Libraries. Volume 13, 2017 - Issue 1: Semantic Historical Gazetteers: A Place for Places - Papers from the DH2016 GeoHumanities SIG Workshop.

de Viguerie J. (1988) « La circonscription des paroisses pendant la Révolution : premières recherches », in *Histoire de la paroisse. Actes de la 11e rencontre d'histoire religieuse. Fontevraud, 2-3 octobre 1987*, Angers, Presses de l'Université, 1988, p. 101-121

Zenasni, S., Kergosien, E., Roche M., and Teisseire, M. (2015) Discovering Type of Spatial Relations with a Text Mining Approach. In Proceedings of ISMIS 2015 (22nd International Symposium on Methodologies for Intelligent Systems), Springer LNAI conference proceedings, p. 442-451, Lyon, France.