
Visualiser l'évolution des toponymes anciens à partir de sources de données patrimoniales

Perrine Pittet¹, Damien Vurpillot¹, Béatrice Markhoff², Marion Lamé³, Benoist Pierre¹.

1. Intelligence des Patrimoines, UMR 7323 CESR, Université de Tours
{damien.vurpillot|perrine.thuringer|benoist.pierre}@univ-tours.fr

2. LIFAT, Université de Tours
markhoff@univ-tours.fr

3. UMR 7324 CITERES-LAT, Université de Tours
marion.lame@univ-tours.fr

RÉSUMÉ. L'étude de l'évolution des noms de lieux offre un outil de recherche essentiel afin de mieux appréhender l'organisation des territoires et leurs transformations au cours du temps. La normalisation des toponymes, par des cartes et des nomenclatures officielles, relève d'un effort relativement récent à l'échelle du territoire français (fin 18ème siècle). Pour les périodes plus anciennes, le croisement de sources issues de plusieurs domaines de recherche devient nécessaire afin de retracer l'évolution des toponymes avec une marge d'incertitude plus ou moins grande, amenée à fluctuer avec l'apport de nouvelles informations. Cet article introduit le développement d'un outil s'appuyant sur un modèle sémantique pour visualiser l'évolution des toponymes anciens à partir de sources de données patrimoniales.

ABSTRACT. The evolution of place names offers an essential research field to better understand the organization of territories and their transformations over time. Standardization of place names, through maps and official nomenclatures, is a relatively recent effort across french territory (end of the 18th century). For older periods, crossing sources coming from several fields of research becomes necessary in order to trace the evolution of toponyms with a margin of uncertainty more or less important and that is intended to fluctuate with the contribution of new information. This article introduces a semantic model to visualize the evolution of ancient place names from heritage data sources.

MOTS-CLÉS : données patrimoniales, évolution des toponymes anciens, CIDOC CRM, CRMgeo, GeoSPARQL, projet HeritageS

KEYWORDS: heritage data, ancient place names evolution, CIDOC CRM, CRMgeo, GeoSPARQL, HeritageS project

1. Introduction

Le projet HeritageS, financé par le programme ARD Intelligence des Patrimoines¹ a mis, parmi ses objectifs, la valorisation des données de la recherche interdisciplinaire sur les patrimoines naturel et culturel du Val de Loire. Cela se traduit par le développement d'une plateforme numérique de données hétérogènes et d'un écosystème d'applications et de services numériques à destination tant des chercheurs que des acteurs socio-économiques du territoire. Ce challenge central du projet vise à rendre interopérables des données patrimoniales produites par près de 33 laboratoires au carrefour de nombreux domaines de recherche qui rencontrent parfois des difficultés légitimes à communiquer entre eux. L'évidente hétérogénéité sémantique de ces données nécessite une approche d'intégration sémantique (Noy, 2004). Le choix s'est porté sur une intégration hybride (Wache et al., 2001) fondée sur l'ontologie de référence pour les ressources patrimoniales, CIDOC CRM (Doerr, 2003), et sur ses extensions. Le modèle de connaissance dérivé de cette ontologie doit permettre à la plateforme d'alimenter les applications et services de l'écosystème et doit ainsi s'adapter pour répondre à leurs besoins spécifiques.

Dans ce contexte, un outil de visualisation en ligne à dimension spatio-temporelle est en développement au sein de l'écosystème HeritageS. Il permettra d'appréhender l'évolution spatio-temporelle de ressources toponymiques sur un territoire en croisant les sources de données patrimoniales intégrées à la plateforme. Il s'agira de faire converger des mentions textuelles de lieux avec la représentation spatiale qui coïncide le mieux en fonction de la période concernée et des éléments de référence disponibles pour inférer le rapprochement toponymique. Fonctionnellement, l'outil devra permettre d'afficher pour chaque lieu visualisé à une date t sur une carte : (1) le centroïde du lieu symbolisé par un point accompagné du toponyme de référence à la date t ; (2) l'emprise spatiale du lieu symbolisée par un polygone autour du centroïde; (3) un encart comprenant : les liens vers une copie numérique des documents administratifs publics ou privés à portée historique qui ont permis de restituer le toponyme de référence et l'emprise spatiale à cette date ainsi que les variations orthographiques du toponyme. Cette démarche s'inscrit dans le processus plus complet de création d'une base de connaissances où sont regroupés et recoupés les données scientifiques produites issues de l'interprétation des toponymes par les experts du domaine et les référentiels utiles à ces interprétations. Cette démarche précède l'étape qui vise à lever des ambiguïtés d'interprétation qui n'avaient pu être résolues de manière satisfaisante dans la première phase d'interprétation.

2. Problématique liée aux données patrimoniales

Les données patrimoniales considérées ici sont des données numériques multimédia issues de projets de recherche interdisciplinaires. Ces données sont de

¹ *Intelligence des Patrimoines*, <https://intelligencedespatrimoires.fr>

trois types : des sources primaires numérisées et étudiées dans le cadre des projets de recherche (ex : numérisations 2D et 3D de documents et d'objets anciens), des données dérivées produites par ces projets à partir de ces sources primaires (ex : transcriptions, analyses, modèles), et leurs métadonnées associées (ex: notices). Toutes ces données se trouvent intégrées dans la plateforme HeritageS, en outre, les métadonnées sont alignées au CIDOC CRM. Néanmoins, n'alimenteront l'outil de visualisation que les sources primaires dont le contenu présente des informations toponymiques pertinentes, c'est-à-dire éventuellement rapprochées dans les métadonnées associées à des toponymes normalisés. Ces données constituent le corpus que nous cherchons à croiser au référentiel toponymique dont la plateforme HeritageS dispose par ailleurs, afin de proposer de renforcer ou de confirmer un rapprochement toponymique.

Un référentiel toponymique est défini ici comme un document scientifique créé au sein du projet HeritageS, faisant autorité et qui référence l'intitulé ainsi que l'emprise spatiale des différents toponymes issus de documents administratifs publics ou privés, faisant quant à eux autorité lors de leur création. Ce référentiel toponymique est construit à partir de deux types de sources primaires: celles en provenance du corpus lui-même et celles extérieures à celui-ci, mais présentes et exploitables au sein du projet HéritageS. Il s'agit dans tous les cas de (1) documents administratifs cartographiques : cartes, cadastres et autres représentations graphiques conventionnelles de données administratives concrètes ou abstraites référencées géographiquement, reconnues dans le cadre du projet comme faisant autorité ; (2) des documents administratifs textuels : chartes, cartulaires, bullaires, et autres documents administratifs référençant des toponymes, reconnus comme faisant autorité toujours dans le cadre du projet.

Cet ensemble de documents ainsi organisés nous permet de définir deux types de toponymes : (1) des toponymes de référence : toponymes présents dans le référentiel toponymique (ex: "Beaumont" était le toponyme de référence de Beaumont-la-Ronce selon la carte de Cassini² de 1793); (2) des toponymes alternatifs : toponymes apparaissant dans les documents primaires du corpus présentant des variations orthographiques ou des formes grammaticales (ex : pluriel, déclinaisons latines) différentes des toponymes de référence (ex: "Beaumont la Ronse" dans le poème Le poème de Ronsard³ en 1555). Nous considérons donc qu'à une date t un lieu est identifié par un toponyme de référence, défini par son emprise spatiale sur une période donnée et qu'un toponyme de référence peut être associé à un ou plusieurs toponymes alternatifs.

Cependant, l'étude des toponymes à partir de ces ressources comporte plusieurs inconvénients. D'une part, les toponymes qui apparaissent dans les sources primaires du corpus ne sont pas toujours mis en lien avec un toponyme de référence dans les métadonnées associées. Or on trouve de nombreuses variations orthographiques et formes grammaticales pour un même lieu, parfois dans un même

² Carte de Cassini : <https://www.geoportail.gouv.fr/donnees/carte-de-cassini>

³ Le Voyage de Tours ou les amoureux, Pierre de Ronsard, 1555

document, notamment pour les documents en français ancien issus de corpus du Moyen-Âge et de la Renaissance. Par exemple, nous trouvons “Beaumont” et “Beaumont de la Ronce” dans le même acte de vente en 1476⁴. Lorsqu’ils sont rapprochés à un toponyme de référence dans les métadonnées, celui-ci n’est généralement pas issu d’un référentiel toponymique contemporain de la source primaire mais d’un référentiel plus récent comme Geonames⁵.

D’autre part, les toponymes de référence ne sont pas toujours géoréférencés (Timár *et al.*, 2014) et c’est souvent le cas pour des référentiels anciens. Le cas échéant, il ne s’agit souvent pas de référentiels cartographiques contemporains de la source primaire du corpus mais des référentiels plus récents, et/ou approximatifs (ex: rattaché au centroïde de la commune), ce qui ne permet pas de connaître la véritable emprise spatiale du lieu à cette époque.

3. Approche

Pour définir le modèle qui servira à l’outil de visualisation, nous avons choisi d’étendre le modèle d’intégration sémantique de la plateforme HeritageS par le biais de l’extension CRMgeo (Hiebel *et al.*, 2017) du CIDOC CRM et de l’ontologie Geosparql (Battle & Kolas, 2011) (cf. Fig.1). L’apport du CRMgeo est le concept de volume spatio-temporel (cf. Fig.1 en bleu E92_SpaceTimeVolume) qui permet de définir un lieu par une projection spatiale et une projection temporelle. Il devient donc possible de décrire l’évolution d’un lieu comme la succession de plusieurs volumes spatio-temporels. L’ontologie Geosparql (cf. Fig.1. en vert) apporte la possibilité de spécifier la géométrie de l’emprise spatiale de lieu dans des formats compatibles avec la plupart des systèmes d’information géographiques. Son alignement avec le CIDOC CRM se fait par le biais d’une relation d’héritage entre le concept geo:SpatialObject et le concept E53_Place. Enfin, afin de modéliser l’alignement de toponymes à des documents administratifs produits à la même époque, nous avons choisi de définir des périodes toponymiques et spatiales de référence. Une période toponymique de référence correspond à la période hypothétique d’usage d’un toponyme de référence. Celle-là démarre à la date de création du premier document administratif connu dans lequel ce toponyme apparaît et se termine à la date de création du document suivant référençant un nouveau toponyme. La période spatiale de référence correspond à la période hypothétique d’existence de l’emprise spatiale d’un lieu. Celle-là démarre à la date de création du premier document cartographique administratif connu qui la définit géographiquement et se termine à la date de création du document cartographique suivant définissant une emprise spatiale différente. Nous avons donc créé deux concepts ToponymReferencePeriod et SpatialReferencePeriod qui héritent du concept E4_Period (cf. Fig.1 en rouge).

⁴ AD37, Minute notariale du 5 décembre 1476, Jaloignes, Jehan

⁵ GeoNames : <http://www.geonames.org/>

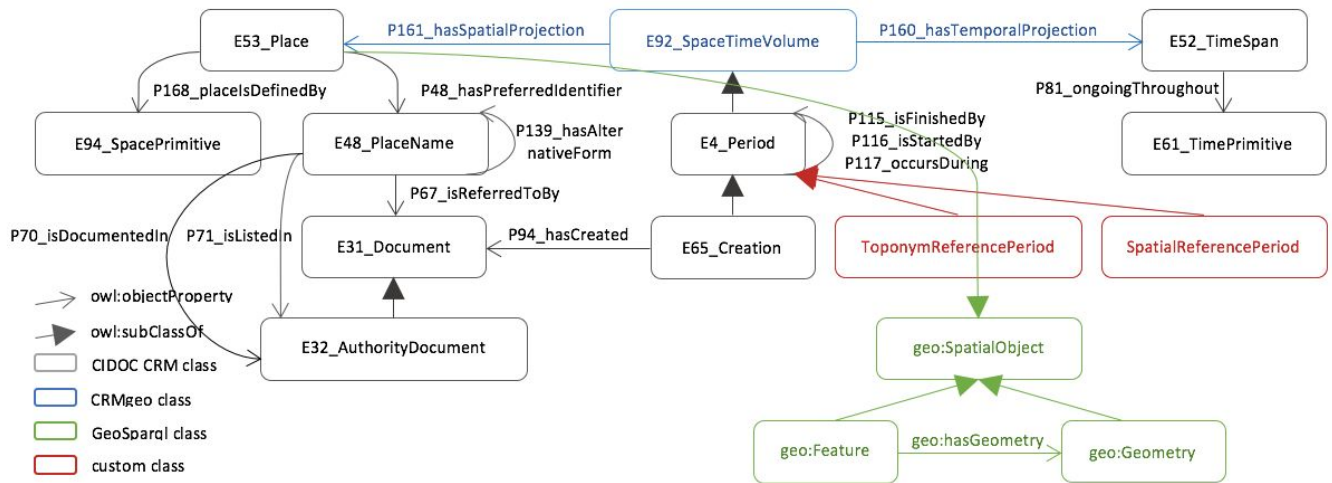


FIGURE 1. Modèle sémantique CIDOC-CRM dédié à l'évolution des toponymes

4. Modèle

Un lieu se trouve représenté par un volume spatio-temporel (E92_SpaceTimeVolume) défini spatialement (P161_hasSpatialProjection) par une projection spatiale (E53_Place) et temporellement (P160_hasTemporalProjection) par une projection temporelle (E52_TimeSpan). Un toponyme de référence (E48_PlaceName) identifie (P48_isPreferredIdentifierOf) une projection spatiale et apparaît dans une ou plusieurs sources primaires du corpus (E31_Document) elles-mêmes listées (P71_isListedIn) dans le référentiel toponymique (E32_AuthorityDocument). Une projection spatiale est décrite géographiquement (P70_isDocumentedIn) par un document administratif cartographique (E31_Document) et est définie dans le référentiel toponymique (E32_AuthorityDocument). Un document, retenu pour le corpus ou externe au corpus, fut créé (P94_hasCreated) par une activité de (E65_Creation) à une date donnée (E52_TimeSpan). Un toponyme de référence peut disposer d'un ou plusieurs toponymes alternatifs (P139_hasAlternativeForm). Un toponyme (de référence ou alternatif) peut être cité (P67_isReferredToBy) par un ou plusieurs documents présents dans le corpus (E31_Document). La projection spatiale du lieu est définie (P168_placesDefinedBy) par une primitive spatiale (E94_SpacePrimitive). La projection temporelle du lieu est définie (P81_ongoingThroughout) par une primitive temporelle (E61_TimePrimitive). Un toponyme de référence identifie (P48_hasPreferredIdentifier) la projection spatiale pendant une période de référence (ToponymReferencePeriod). Cette période de référence démarre (P116_isStartedBy) à la création (E65_Creation) du document administratif (E31_Document) dans

lequel le toponyme de référence (E48_PlaceName) apparaît et se termine (P115_isFinishedBy) à la création (E65_Creation) du premier document administratif (E31_Document) lui succédant dans le temps, et dans lequel apparaît un toponyme différent, les deux toponymes étant indexés dans (E32_Authority_Document). La projection spatiale d'un lieu (E53_Place) est limitée dans le temps sur une période spatiale de référence (SpatialReferencePeriod). Cette période spatiale démarre à la création (E65_Creation) du document administratif cartographique (E31_Document) dans lequel l'emprise spatiale du toponyme de référence est décrite, et se termine à la création (E65_Creation) du premier document administratif cartographique (E31_Document) lui succédant et dans lequel apparaît le changement de l'emprise spatiale précédemment définie, les deux emprises spatiales étant toutes deux indexées dans le référentiel toponymique (E32_Authority_Document).

Afin d'aligner un toponyme de référence et une projection spatiale à un toponyme cité dans un document source, deux règles sont définies: (R1) Un toponyme alternatif est associé à un toponyme de référence si la date de création de la source primaire où il est cité est incluse (P117_occursDuring) dans la période de référence de ce toponyme. (R2) Un toponyme alternatif est associé à une projection spatiale si la date de création de la source primaire où il est cité est incluse dans la période spatiale de référence.

Remerciements

Les auteurs remercient Johann Forte et Xavier Rodier pour leur aide et leurs conseils dans l'élaboration de cette proposition.

Références

- Battle, R., & Kolas, D. (2011). Geosparql: enabling a geospatial semantic web. *Semantic Web Journal*, 3(4), 355-370.
- Doerr, M. (2003). The CIDOC conceptual reference module: an ontological approach to semantic interoperability of metadata. *AI magazine*, 24(3), p. 75.
- Hiebel, G., Doerr, M., & Eide, Ø. (2017). CRMgeo: A spatiotemporal extension of CIDOC-CRM. *International Journal on Digital Libraries*, 18(4), 271-279.
- Noy N. F. (2004). Semantic integration: a survey of ontology-based approaches. *ACM Sigmod Record*, 33(4), p. 65-70.
- Timár, G., Mészáros, J., & Molnár, G. (2014). A simple solution for georeferencing the Cassini map series of France. In *9th International Workshop on Digital Approaches to Cartographic Heritage*.
- Wache, H. et al.(2001). Ontology-based integration of information-a survey of existing approaches. *IJCAI-01 workshop: ontologies and information sharing*. vol. 2001 p. 108-117.

Annexe

Nous illustrons l'utilisation du modèle avec l'exemple du site de Beaumont-la-Ronce. L'objectif est ici de trouver quel est le toponyme de référence associé au toponyme "Beaumont la Ronse" cité dans le poème "Le Voyage de Tours" de Pierre de Ronsard datant de 1555. Dans les ressources d'HeritageS, le référentiel toponymique ayant précédé la création de ce poème est le registre de chancellerie de Louis XII, datant de 1479 dans lequel le toponyme de référence est "Beaumont de la Ronce". Le référentiel toponymique lui succédant est la carte de Capitaine datant de 1790 dans laquelle le toponyme de référence est "Beaumont la Ronce". Nous décrivons l'ensemble de ces informations à l'aide du modèle présenté plus haut (cf. Fig. 2). La période toponymique de référence délimitée par les deux référentiels s'étend de 1479 à 1790 (cf. Fig. 2 en rouge). L'application de la règle (R1) permet d'associer le toponyme "Beaumont la Ronse" au toponyme de référence "Beaumont de la Ronce" (cf. Fig. 2 en violet).

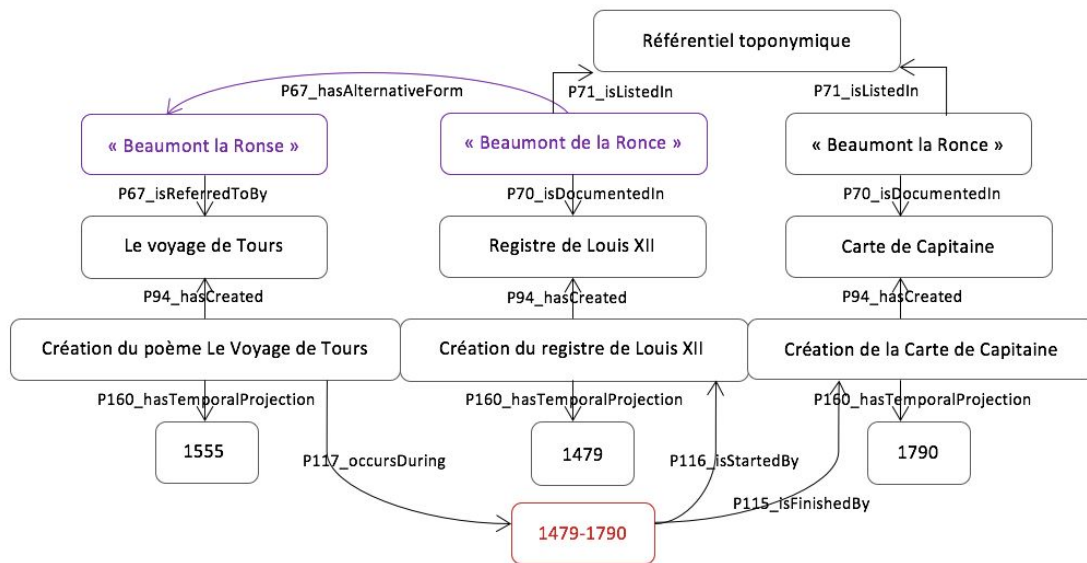


Figure 2. Modèle appliqué à l'exemple de Beaumont-la-Ronce

La période spatiale de référence pourra être de la même manière définie à partir des documents administratifs cartographiques produits à la même époque ou sur une période la plus proche possible dans le temps. La règle (R2) du modèle sera alors appliquée pour associer la projection spatiale restituée au toponyme. Le volume

spatio-temporel de Beaumont-la-Ronce en correspondant au lieu désigné par le poème de Ronsard en 1555 pourra alors être construit. Selon le même processus, d'autres volumes spatio-temporels de Beaumont-la-Ronce seront définis à partir d'autres données patrimoniales. Leur évolution pourra être visualisée dans l'outil cartographique en cours d'élaboration, sans que cela préjuge d'emplois pour le moment inconnus et absents des sources disponibles.