

Géoréférencement d'images anciennes

à l'aide des indications de localisation fournies par leurs métadonnées

Mots-clés: Web de données, référentiels de données géographiques vectorielles, analyse spatiale, extraction d'informations topographiques à partir de textes, programmation.




Contexte


De plus en plus d'institutions patrimoniales et culturelles publient les catalogues des collections dont elles ont la garde sur le Web afin de faciliter la découverte de ces collections. Ces catalogues se composent, pour chaque ressource patrimoniale ou culturelle, de métadonnées permettant d'en découvrir le contenu même lorsque aucune copie numérique de la ressource n'est disponible: titre, auteur, éditeur, date, mots-clés, etc. Désenclaver ces jeux de métadonnées suppose de les interconnecter avec des référentiels communs, comme des thésaurus sur les types d'œuvres architecturales, des index d'auteurs ou des jeux de données géographiques. Ceci permet de désambiguïser les chaînes de caractères utilisées pour désigner les auteurs, les types d'œuvres ou les lieux en leur substituant les identifiants de ressources décrites par les référentiels et faisant autorité. En outre, ceci permet de naviguer aisément au travers des collections et de leurs référentiels associés en suivant les liens ainsi explicités.

Le projet [ALEGORIA](#) vise à valoriser des fonds iconographiques institutionnels décrivant le territoire français à différentes époques allant de l'entre-deux-guerres à nos jours. Ces clichés s'accompagnent de métadonnées, et lorsqu'une photographie représente une entité géographique, sa légende peut contenir des indications sur la localisation de cette entité ou sur le lieu de prise de vue. On retrouve ce type d'informations dans les métadonnées de la photographie présentée en figure 1.

OpenmuséeNiépce Thèmes ▾ Les projets Créer mon compte | Me connecter | Ma sélection | Q

[Retour](#)



© musée Nicéphore Niépce

Légende : Quai de la Monnaie, Pont St Laurent et hôpital de Chalons-sur-Saône, vue prise du quai Ste Marie

Date : entre 1840 et 1855
Auteur : Fortuné-Joseph PETIOT-GROFFIER
Technique : tirage sur papier albuminé
Mots-clés : [vue d'architecture](#) - [Chalons-sur-Saône](#) - [pont](#) - [quai](#) - [hôpital](#) - [Saône](#) - [Fortuné-Joseph PETIOT-GROFFIER](#)
Copyright : [musée Nicéphore Niépce](#), [Ville de Chalons-sur-Saône](#)

Thèmes :

- Chalons-sur-Saône

Choisir cette image


- Ajouter à ma sélection
-  Télécharger

Figure 1: Une photographie conservée au musée Nicéphore Niépce et ses métadonnées (<http://www.open-museeniepce.com/recherche-photos/photo,2611>)

Principales hypothèses et objectifs du stage

Une des réalisations prévues par le projet [ALEGORIA](#) est un moteur d'indexation et de recherche multimodal et à grande échelle, couplant recherche par contenu et par métadonnées dans les fonds d'images numérisées et documentées. Ce moteur s'appuiera, entre autres, sur les indications de localisation fournies par les métadonnées des images.

L'objectif de ce stage est de faciliter l'exploitation par le moteur d'indexation des indications de localisation disponibles dans les métadonnées. Une première étape consistera à extraire des métadonnées des jeux d'images du projet ALEGORIA les indications de localisation, représentées sous la forme d'entités spatiales nommées [M15]. Il conviendra en outre de déterminer si ces entités spatiales nommées désignent le contenu de l'image ou le lieu de sa prise de vue. Enfin, une dernière étape consistera à résoudre ces entités spatiales nommées, c'est-à-dire à associer à chaque mention d'entité spatiale nommée l'identifiant de l'entité géographique du monde réel à laquelle elle fait référence [PAB17]. L'utilisation d'un jeu de données géographiques de référence comportant des indications de localisation directes (coordonnées ou géométrie), permettra de localiser précisément les entités spatiales mentionnées dans les métadonnées.

Verrous scientifiques

Une première difficulté réside dans le manque d'informations de contexte dans les métadonnées, dans la mesure où ce sont des textes relativement courts. Les approches d'extraction et de résolution des entités spatiales choisies devront s'adapter à cette contrainte.

Le traitement des entités spatiales relatives constitue une seconde difficulté majeure. En effet, il s'agit d'entités spatiales identifiées et localisées par référence à une autre entité spatiale nommée, comme "la rive droite de la Saône" par exemple. L'approche d'extraction proposée devra être en mesure de détecter ces entités spatiales relatives. En outre, dans la mesure où elles ne figurent pas dans les référentiels géographiques, il conviendra de proposer une approche pour dériver leur localisation à partir des données géographiques disponibles.

Renseignements pratiques

- **Formation** : Master 2 ou troisième année d'école d'ingénieur en informatique ou en géomatique avec une forte composante informatique.
- **Durée et période de stage** : 5 mois, au cours du printemps et de l'été 2018.
- **Lieu de stage** : Equipe [LaSTIG/Strudel](#), [Institut national de l'information géographique et forestière](#) (IGN), Saint-Mandé (métro 1, station Saint Mandé).
- **Indemnités de stage**: Stage gratifié selon la législation française.
- **Modalités de candidature** : Envoyer par email un fichier PDF avec votre curriculum vitae, une lettre de motivation **ciblée sur le sujet**, vos relevés de notes des deux dernières années d'études.
- **Encadrement de stage** :
 - Nathalie Abadie (IGN/COGIT): nathalie-f.abadie@ign.fr
 - Carmen Brando (EHESS): carmen.brandot@ehess.fr

Bibliographie

- [M15] **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.
- [PAB17] **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.