

Extraction de règles de conception de cartes d'un corpus de la cartographie

Mots clés

TAL, informatique, gestion des connaissances, cartographie

Contexte

Le laboratoire COGIT de l'Institut national de l'information géographique et forestière (IGN) étudie les problématiques liées à l'utilisation de données topographiques pour la fabrication de produits (cartes géographiques, lot de données) ou de services répondant à des besoins particuliers, spécifiés par les utilisateurs de ces produits.

Ce stage a pour objectif d'extraire les règles de conception de cartes telles qu'elles sont décrites dans un corpus textuel de la cartographie, disponible dans une version électronique. Ce corpus a été formé à partir de sources différentes : un manuel de cartographie utilisé dans une école d'ingénieurs topographes qui détaille particulièrement la fabrication de cartes topographiques et des notes de cours de différents enseignements en cartographie de l'université, ciblé sur les cartes thématiques.

Nous nous intéressons particulièrement aux règles de conception de la carte, l'ordre des étapes, les concepts (en relation avec l'ontologie de la cartographie OntoCarto) sur lesquels s'appuient ces modes opératoires et leurs articulations, les principes de représentation (en relation avec la base de règles OntoCartoRules). Pour l'extraction des règles, sera mis en œuvre l'outil SEMEX, une plateforme d'exploration sémantique et d'aide à l'acquisition de règles métiers candidates développée à Paris 13.

Ce stage est co-encadré par les laboratoires COGIT de l'IGN et LIPN (Laboratoire d'informatique de Paris Nord) à Villetaneuse.

Sujet

L'objectif du stage est d'identifier les règles de conception d'une carte (topographique ou thématique) et de réécrire ces règles dans un formalisme plus rigide compatible avec la base de connaissances déjà construite.

Pour ce travail, il faudra préciser la notion de règle dans ce corpus ; reconnaître les variations de vocabulaire autour des noms des concepts du domaine ; repérer les indices de désambiguïsation propres à ce corpus ; définir les marqueurs spécifiques au corpus qui annoncent la définition de règles et le caractère plus ou moins impérieux de cette règle. En effet certaines règles sont essentielles à la compréhension de la carte par le lecteur, d'autres correspondent plus à des habitudes ou des préférences du concepteur. L'objectif final est d'identifier les structures des phrases correspondant, pour ce corpus, à des règles et si possible leur impériosité et leur portée, en particulier déterminer si la règle s'applique à toutes les cartes ou plus spécifiquement à une carte topographique ou à une carte thématique.

L'ontologie OntoCarto intègre déjà les concepts correspondant aux travaux sur les variables visuelles réalisés par Bertin sur lesquels sont fondées les règles de la sémiologie cartographique.

Bibliographie

Bertin, J. (1967). *Sémiologie graphique : les diagrammes, les réseaux, les cartes*.

Dominguès, C., S. Christophe, et L. Jolivet (2009). "Connaissances opérationnelles pour la conception automatique de légendes de cartes". 20èmes Journées Francophones d'Ingénierie des Connaissances (IC'2009), Hammamet, Tunisie.

Dominguès, C., O. Corby, et F. Soualah-Alila. "Raisonnement sur une ontologie cartographique pour concevoir des légendes de cartes". 12^e Conférence Internationale Francophone sur l'Extraction et la Gestion de Connaissance (EGC'2012), 31 janvier - 3 février, Bordeaux, France (à paraître).

manuel d'utilisation de SemEx :

http://www-lipn.univ-paris13.fr/~guisse/ontorule/SemEx/SemEx_Manual.pdf

Compétences particulières et formation requise

Ce stage s'adresse aux étudiants de master 2 ou de 3^{ème} année d'école d'ingénieurs avec une spécialisation en informatique ou en traitement automatique du langage naturel.

Lieu du stage

IGN/laboratoire COGIT
73 avenue de Paris
94165 Saint-Mandé Cedex
métro : Saint-Mandé - ligne 1

Durée et rémunération

durée : 5 à 6 mois
début : mars/avril 2012
rémunération : 30% du SMIC

Prolongements éventuels

Le COGIT propose chaque année des sujets de thèse ainsi que des stages de post-doctorant.

Responsable du stage

Catherine Dominguès
IGN/SR/COGIT, 73 avenue de Paris, 94165 SAINT-MANDE Cedex
tél : 01 43 98 85 44 mël : catherine.domingues@ign.fr

Le stage est co-encadré par : François Lévy
LIPN, Institut Galilée, Avenue J.B. Clément, 93430 VILLETANEUSE
tél : 01 49 40 35 78 mël : Francois.Levy@lipn.univ-paris13.fr

Pour candidater

Le dossier de candidature sera envoyé par mail. Il devra se composer d'un curriculum vitae et d'une lettre de motivation, accompagnés des relevés de notes des années de M1 et M2 (ou deux dernières années d'école d'ingénieurs) et d'une description des enseignements suivis (un lien vers le site internet de la formation est le bienvenu).