



OFFRE D'ALLOCATION DE THESE

ÉCOLE DOCTORALE SCIENCES EXACTES ET LEURS APPLICATIONS - ED 211

Avenue de l'université BP 1155 64 013 PAU Cedex – France

SUJET DE THESE

TITRE : Localisation automatique des points d'intérêt par la mise en correspondance du texte narratif et de la trace GPS.
Automatic Location Of Points Of Interest Through Matching Narrative Text And GPS Track.

RESUME :

Le laboratoire d'informatique de l'Université Pau et Pays de l'Adour, France (LIUPPA) et, le Laboratoire COGIT de l'Institut National de l'Information Géographique et Forestière font appel à candidature pour un poste de doctorat en Traitement Automatique de la Langue et Analyse Spatiale.

The Computer Science Lab. of the Pau and Pays de l'Adour University, France (LIUPPA) and the COGIT Lab. of the French National Institute of Geographic and Forest Information (IGN) invite applications for a PhD position in natural language processing and spatial analyses.

Mots clés: DSL, interopérabilité, SOA, agents, données réparties et hétérogènes

CONDITIONS D'EXERCICE

Laboratoire : Laboratoire d'Informatique de l'Université de Pau et des Pays de l'Adour (LIUPPA)
Computer Science Lab. of the Pau and Pays de l'Adour University, France (LIUPPA)

Site web : <http://liuppa.univ-pau.fr/>

Directeur de thèse : Mauro Gaio (LIUPPA)
Co-Directeur de thèse : Sébastien Mustière (COGIT-IGN)
En collaboration avec Sophie Mercier (LMAP)

Lieu Avenue de l'Université, 64000 Pau

Date début : 1 Sept 2017

Durée : 36 mois

Employeur : Université de Pau et des Pays de l'Adour (UPPA)

Salaires mensuel brut : 1685 € (*pour contrat doctoral sur crédits UPPA et collectivités locales*)

SAVOIR-FAIRE DU LABORATOIRE

La recherche menée au LIUPPA est de nature appliquée avec des domaines de prédilection tels que :

- Génie logiciel
- Agents et composants logiciels
- Sécurité informatique
- Systèmes d'information

- Réseaux et protocoles
- Traitement des documents électroniques
- Image et la visualisation scientifique.

MISSION - ACTIVITES PRINCIPALES

La thèse débutera en septembre 2017 et se terminera en septembre 2020. Le candidat poursuivra son travail de recherche au sein du projet ANR Choucas impliquant trois centres de recherche, le Laboratoire d'Informatique de l'Université de Pau et des pays de l'Adour (LIUPPA)(lieu de déroulement de la thèse), le laboratoire COGIT de l'Institut National de l'Information Géographique et Forestière (IGN) au Grenoble Informatics Lab. (LIG), ainsi que la Gendarmerie Nationale en particulier avec le Peloton de Haute Montagne (PGHM).

Le projet Choucas « Intégration de données hétérogènes et raisonnement spatial pour aider à la localisation des victimes en montagne » est un projet pluridisciplinaire qui regroupe trois thèmes de recherche : la structuration et l'interrogation de données hétérogènes, le raisonnement spatial flou et la géovisualisation de données géographiques imparfaites.

The position starts in September 2017 and ends in September 2020. The candidate will pursue a PhD project in the ANR project Choucas involving 3 Research Centers, the Computer Science Lab. of the Pau and Pays de l'Adour University, France (LIUPPA) (location of the position), the COGIT Lab. of the French National Institute of Geographic and Forest Information (IGN) and the Grenoble Informatics Lab. (LIG) in collaboration with the National Gendarmerie with its special High Mountain unit (PGHM).

The project Choucas « Integration of heterogeneous data and spatial reasoning for helping to locate victims in the mountains », is a pluridisciplinary project that bridges three research themes: structuring and interrogation of heterogeneous data, fuzzy spatial reasoning and geovisualization of imperfect geographical data.

Le travail de thèse s'intéressera en particulier à la partie textuelle non structurée des données et établir un pont entre la théorie spatio-linguistique, les approches d'analyse et de traitement des données spatiales, le traitement probabiliste piloté par les données et l'analyse sémantique. La nouveauté de l'approche consiste à mettre la sémantique spatiale au centre de la modélisation, mettant l'accent sur l'adoption d'une approche de construction de la grammaire en utilisant par exemple des grammaires d'arbres adjoints.

Le travail de thèse s'intéressera également à la correspondance d'informations extraites du texte avec la géométrie d'une ou de plusieurs traces GPS associées. L'originalité de cet objectif réside dans la définition d'un processus d'appariement de données très hétérogènes : le texte et les géométries. Cela nécessitera d'abord le traitement géométrique et la modélisation des traces GPS pour extraire leurs caractéristiques spatiales, ce qui peut être considéré comme un problème de raisonnement spatial qualitatif. Cela peut alors s'appuyer sur des techniques d'optimisation, comme celle utilisant des modèles de Markov cachés pour faire correspondre les traces GPS et les données routières pour la navigation automobile.

Le traitement probabiliste piloté par les données sera réalisé en collaboration avec le laboratoire de mathématiques et d'applications. De Pau, France (LMAP).

The PhD project is in particular concerned with the unstructured textual part of the data and the bridges between spatio-linguistic theory and spatial data processing approaches and data-driven

probabilistic processing and semantic parsing. The novelty of its approach consists in putting spatial-semantics as the modeling focal point, putting an emphasis on adopting a constructional approach to grammar by using tree-based grammar frameworks.

The PhD project is also concerned with the matching of information extracted from the text to the geometry of one or several associated GPS tracks. The originality of this goal lies in the definition of a matching process of very heterogeneous data: text and geometries. This will first require the geometric processing and modeling of GPS tracks to extract their spatial characteristics, which may be thought of as a qualitative spatial reasoning issue. This may then rely on optimization techniques, like the one using Hidden Markov Models to match GPS tracks and road data for car navigation.

The data-driven probabilistic processing will be conducted in cooperation with the Mathematics and Its Applications Lab. of Pau, France (LMAP).

COMPETENCES REQUISES

EXIGENCES ET QUALIFICATIONS POUR LE POSTE

- Master en Informatique ou Géomatique ou équivalent
- Solides compétences en programmation
- Connaissances en traitement du langage naturel et linguistique computationnelle
- Connaissances en analyse et traitement des données spatiales
- Connaissances en traitement probabiliste piloté par les données serait appréciée
- Le candidat doit maîtriser l'anglais ou le français et avoir une bonne connaissance de l'autre langue

JOB REQUIREMENTS AND QUALIFICATIONS

- *Master's degree in Computer Science or Geomatics or equivalent*
- *Solid programming skills*
- *Background in natural language processing and computational linguistic*
- *Background in spatial data analyses and processing*
- *Background in data-driven probabilistic processing would be appreciate*
- *The candidate should be fluent in English or in French and have a good knowledge of the other language*

CRITERES D'EVALUATION DE LA CANDIDATURE

Traitement du dossier : Jury de sélection

Les candidats seront sélectionnés d'abord sur dossier. Un entretien sera organisé après la première phase de sélection du dossier de candidature.

- Adéquation entre le diplôme de Master (ou équivalents) et le sujet de thèse
- Notes et classements en Master, et régularité dans le cursus universitaire
- Maîtrise de l'anglais
- Capacité du candidat à présenter ses travaux
- Expériences professionnelles de type stage(s) en laboratoire ou autre ; éventuels travaux de recherche déjà réalisés (rapports, publications).

CONSTITUTION DU DOSSIER DE CANDIDATURE, DATE LIMITE DE DEPOT

Envoyer par email un dossier de candidature comprenant :

Les demandes devront inclure :

- Une Lettre de motivation décrivant l'intérêt pour le poste
- Les noms et les coordonnées de deux référents
- Curriculum Vitae (CV) avec des publications (le cas échéant)
- Copie du diplôme de Master avec les listes détaillées des notes.

Applications should include:

- Cover letter outlining interest in the position
- Name and address of two referees
- Curriculum Vitae (CV) with publications (if applicable)

- Copy of MA degree with the detailed lists of marks or grades.

DATE LIMITE DE DEPOT DU DOSSIER : 31 mai 2017

CONTACT

NOM : Mauro Gaio

MAIL : mauro.gaio@univ-pau.fr