

Post-Doctoral Position



Sentiment Analysis and touristic event extraction

The L3i laboratory, within the Tourinflux project (<u>http://www.tourinflux.com/</u>) is seeking a postdoctoral researcher in computer science, on the topic of sentiment analysis and touristic event extraction.

Length: 10 months Expected recruitment: 1st of June 2015 or ASAP – the position is to be filled ASAP, and will last until the end of February 2016. Net salary: 2100 € monthly Location: L3i laboratory, La Rochelle, France Fields: Computer Science / Natural Language Processing / Semantic Web Keywords: E-tourism, Multilingual NLP, Event extraction, opinion mining, sentiment analysis, NLP, Normalization (schema.org, TourInFrance), Semantic Web, Data analysis, spatiotemporal events.

Job description:

The work will be conducted in the informatics, image and interaction laboratory (L3i), within the Tourinflux project, funded by a public investment program for the future (PIA). The L3i is a 120-person laboratory created in 1993. Hosted in the historical and sunny city of La Rochelle (http://www.holidays-la-rochelle.co.uk/), it is ranked A by the French research evaluation agency (AERES).

In addition to the L3i, the Tourinflux project involves 2 companies and the professional association for the digital economy, in collaboration with several actors of the French tourism industry. The project aims at providing actors of the tourist industry with a set of tools allowing them to handle both their internal data, and the information available on the Web, so as to better diagnose and influence on the perception of territories. The tools currently available are insufficient, and a lot of the data gathering, analysis and processing is currently done by hand, or via the use of various tools that are only partially satisfactory. Tourinflux aims to provide an extensive dashboard, allowing all institutions, whatever their size, to visualize and interpret the information available about their territory. This should allow them to improve their decision process and subsequently, their effectiveness.

Specifically, this postdoctoral position will focus on the automated extraction of touristic events and/or opinion mining over touristic objects. Both subjects will be considered on equal grounds and the ability of the candidate will be the main recruitment criteria. The techniques to be used will need to be applicable to any language.

1. (Semi)-automated extraction of touristic events:

Tourism information is both heterogeneous (free text, Web page, pictures, ...) and semistructured. Extracting touristic information is a major challenge at a time when the masses of unstructured information are constantly evolving, be it on the Web or within organizations. The task of the recruited postdoctoral researcher will be to extract touristic events and insert information into a given representation model. Two key challenges were identified:

- Adapting NLP techniques to the specific domain of tourism information, and extract the key features of the field
- Understanding and feeding the given models of representations for touristic data.

Recent work on the temporal analysis of natural language, so as to grasp the opening hours of touristic objects remains to be integrated. It will be a plus if the recruited postdoctoral researcher is able to integrate this work within the event extraction process.

2. Sentiment analysis over touristic objects

The postdoctoral researcher will be in charge of extracting sentiment from online data, e.g, evaluation platforms (such as tripadvisor), blogs, microblogs, ... The extracted data is to be adapated to our own opinionated mark-up language (see, eg, SentiML). This task will require three steps:

- Given a touristic object, fetch online information about it;
- Extract sentiment about the object;
- Fill our annotation schema.

Specific requirements:

Candidates must have PhD in computer science, with abilities in knowledge representation and data mining. Research experience is also required in at least 2 of the following domains:

- Natural Language Processing, Text mining
- Information Retrieval
- Big Data and Data Warehouses (eg, Hadoop)
- Modeling, Ontologies and inference engines
- Annotation and Evaluation methodologies

General requirements:

- One or more of the following programming languages: Python, C/C++, java...
- Team-work skills (knowledge of Agile methodologies would be a plus)
- Proven ability for scientific writing

To apply:

All candidates are required to send a resume, an expression of interest, and the names and contact information of at least 2 references (including email addresses) to:

mickael.coustaty@univ-lr.fr antoine.doucet@univ-lr.fr

Please note that applications must be sent at the latest on Tuesday 14 April at 23:59 CET