**Contexte et atouts du poste**

**Post Doctoral Position. Natural Language Processing: automatic speech recognition system using Deep Neural Networks without out-of-vocabulary words**

**Mission confiée**

- Missions:
  We assume that in an audio document to transcribe we have missing proper names, i.e. proper names that are pronounced in the audio document but that are not in the lexicon of the automatic speech recognition system; these proper names cannot be recognized (out-of-vocabulary proper names, OOV PNs). The purpose of this work is to design a methodology how to find and model a list of relevant OOV PNs that correspond to an audio document.

Assuming that we have an approximate transcription of the audio document and huge text corpus extracted from internet, several methodologies could be studied:

From the approximate OOV pronunciation in the transcription, generate the possible writings of the word (phome to character conversion) and search this word in the text corpus.

A deep neural network can be designed to predict OOV proper names and their pronunciations with the training objective to maximize the retrieval of relevant OOV proper names.

The proposed approaches will be validated using the ASR developed in our team.

**Principales activités**

**Keywords:** deep neural networks, automatic speech recognition, lexicon, out-of-vocabulary words.

**Bibliography**


**Skills and profile:** PhD in computer science, background in statistics, natural language processing, experience with deep learning tools (keras, kaldi, etc.) and computer program skills (Perl, Python).

**Additional information:**
Supervision and contact: Irina Illina, LORIA/INRIA (illina@loria.fr), Dominique Fohr INRIA/LORIA (dominique.fohr@loria.fr) https://members.loria.fr/IIllina/, https://members.loria.fr/DFohr/

Additional links: Ecole Doctorale IAEM Lorraine

Duration: 12 months.
Starting date: between Oct. 1st 2018 and Jan. 1st 2019
Salary: about 2,100 euros net, medical insurance included

The candidates must have defended their PhD later than Sept. 1st 2016 and before the end of 2018.

The candidates are required to provide the following documents in a single pdf or ZIP file:

- CV including a description of your research activities (2 pages max) and a short description of what you consider to be your best contributions and why (1 page max and 3 contributions max); the contributions could be theoretical or practical. Web links to the contributions should be provided.
- Include also a brief description of your scientific and career projects, and your scientific positioning regarding the proposed subject.
- The report(s) from your PhD external reviewer(s), if applicable.
- If you haven't defended yet, the list of expected members of your PhD committee (if known) and the expected date of defence.
- In addition, at least one recommendation letter from the PhD advisor should be sent directly by their author(s) to the prospective postdoc advisor.

The Inria Sophia Antipolis - Méditerranée center counts 37 research teams and 9 support departments. The center's staff (about 600 people including 400 Inria employees) is composed of scientists of different nationalities (250 foreigners of 50 nationalities), engineers, technicians and administrators. 1/3 of the staff are civil servants, the others are contractual. The majority of the research teams at the center are located in Sophia Antipolis and Nice in the Alpes-Maritimes. Six teams are based in Montpellier and a team is hosted by the computer science department of the University of Bologna in Italy. The Center is a member of the University and Institution Community (ComUE) “Université Côte d’Azur (UCA)”.

Compétences

Avantages sociaux
- Subsidised catering service
- Partially-reimbursed public transport
- French courses

Rémunération
Salary: 2653€ gross/month