

Job advertisement No. NL4XAI-ESR6

The EU funded NL4XAI Innovative Training Network is looking to employ a

Research Associate / PhD Candidate in Deep Learning and Natural Language Generation

The researcher will work under the supervision of [Claire Gardent](#) at CNRS/LORIA/Lorraine University, [Nancy](#) (France) and be co-supervised by [Albert Gatt](#) (University of Malta); he or she will be expected to enrol for a PhD at Lorraine University (Nancy, France). Both Claire Gardent and Albert Gatt are leading experts on NLG. The researcher will be part of the Lorraine computer science research unit (LORIA) at Nancy, and work alongside other students and researchers who work on models for NLG. S/he will also benefit from the wider training and research network provided by the [European NL4XAI Innovative Training Network](#).

During the course of the project, the researcher will carry out two 3 months-secondments to the University of Malta (with Albert Gatt) and one 3-months secondment to [Orange](#) in Lanion, France (with Lina Rohas-Barahona).

Claire Gardent has just been awarded an AI chair which focuses on multilingual and multisource NLG and will fund an additional 3 PhD students and an engineer over a period of 4 years (2020-2024). She also participates in the ANR Quantum Project on Question Generation (2019 – 2023) and heads the CNRS Research Network on Computational, Formal and Field Linguistics (2019 - 2023).

This is a great opportunity to join a leading NLG research group and work with top researchers to develop innovative techniques for explainable AI !

Your qualities

We are looking for candidates with a strong background in computer science, natural language processing (NLP) and/or deep learning. The candidate should have strong programming skill, should be able to think creatively and be interested in NLP.

- Successfully completed university degree (Master or comparable) in a relevant field like Computer Science or Natural Language Processing
- Proficiency in spoken and written English.
- Excellent proficiency in modern programming languages (especially Python) and deep learning libraries (Pytorch). Proven team software development skills.
- Research and development interest in Natural Language Processing
- Because of the Marie Skłodowska-Curie eligibility criteria, candidates must be in the first four years (full-time equivalent research experience) of their research careers and have not been awarded a



UNIVERSITÉ
DE LORRAINE



Campus scientifique

BP 239 – 54506 Vandœuvre-lès-Nancy Cedex

Tél. : 03 83 59 20 00 – Fax : 03 83 27 83 19 www.loria.fr

doctoral degree. They also must not have resided or carried out their main activity (work, studies, etc) in France for more than 12 months in the 3 years immediately prior to the recruitment date.

Your tasks

While neural approaches to text production show promising results, the end-to-end encoder-decoder process they mostly assume fails to provide any insight on how generation works and more specifically, on how faithful the generated text is to the input. In fact, it has repeatedly been observed that neural generators frequently output semantically inadequate texts ie text that either fails to express all relevant content or contain information that is not included in the input.

The broad goal of this PhD thesis will be to provide explainable models of text generation which permit identifying relevant mismatches between input and output. Two main directions will be explored.

- Explaining by breaking up the end-to-end decoder process in several substeps as was done for instance in some approaches to visual question answering (VQA) where VQA was decomposed into predicting the name of the objects contained in the image, generating a caption and answering the question based on both the input image and these intermediate predictions; and evaluation metrics were used to evaluate how accurate these intermediate predictions are and how well they correlate with success. The challenge here will be to identify relevant substeps and evaluation criteria for adapting the method to text generation. A natural option would be to decompose NLG into some or all of the traditional NLG modules thereby also allowing for a more fine grained evaluation of how neural NLG systems can handle the various choices that need to be made to produce a well-formed text.

- Explaining by exploiting advanced neural models and interpretation, metrics such as for instance, Hudson and Manning's MAC network, a novel neural network architecture, designed to facilitate explicit and expressive reasoning; or Samek et al's methods for computing the sensitivity of the prediction with respect to changes in the input and for decomposing the decision in terms of the input variables. A second goal of the thesis will be to investigate how these methods can be applied to analyse and explain the behaviour of NLG systems and whether and if so, how, they complement the more linguistic approach developed in the first part of the thesis.

The thesis will consider both data-to-text and text-to-text generation drawing on existing datasets such as WebNLG or the SR dataset for data-to-text generation and one or some of the many datasets available for text summarisation. While for these two data-to-text benchmarks, the text should match the input, for summarization, it should only express the key information contained in the input. Thus these two tasks will raise different questions both on how to analyse semantic adequacy and on how to explain the behaviour of a generation system. One goal of the thesis should be to come up with a unifying view for these two tasks which can explain these differences.

The PhD topic will be carried out in the context of the European NL4XAI ITN project. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 860621.

What we offer

At LORIA/Lorraine University, you will have the opportunity to further your scientific development in a dynamic and excellent research environment. We provide a scientifically and intellectually inspiring environment embedded in a leading university and one of the largest computer science research lab in France.

[LORIA](#) (Lorraine Research Laboratory in Computer Science and its Applications) is a research unit common to [CNRS](#), the [University of Lorraine](#) and INRIA which gathers around 500 people and 27 teams and is structured into 5 main departments targeting both fundamental and applied research in computer science. The Natural Language and Knowledge Processing department includes 6 teams among which Synalp, the hiring team for this PhD topic which specialises in Statistical and Symbolic Natural Language Processing with a strong focus on neural approaches to Natural Language Generation. SYNALP is well anchored in the national and international research community and Claire Gardent has been regularly involved in the activities of the top international association of the field, the ACL (Association for Computational Linguistics) as member, vice-president and president of the EACL (European Chapter of the ACL) board and more recently, as the chair of SIGGEN, the ACL Special Interest Group in Natural Language Generation. She has supervised 16 PhD students, 5 post-docs and 8 engineers and has been the principal investigator for 10 projects (4 national, 6 european). She regularly serves in all main NLP conferences as chair, area chair or reviewer; and is or has been a member of the editorial board for 5 of the main NLP journals (JoS, TAACL, CL, TAL, JoLLI 3).

Furthermore, we offer

- Funding for necessary equipment, conference and research visit travel
- A modern workplace with a collegial, attractive and versatile working environment
- The integration in the NL4XAI European Innovative Training Network providing a wide range of internal and external further education and training measures
- Discount for employees in the University canteens as well as the possibility to use the various offers of Lorraine University
- A gross salary of 2851 euros/month.

How to apply

Fill in the form available below and upload the requested documentation :

<https://nl4xai.eu/vacancies/>

If you have questions regarding the position, contact Claire Gardent by mail at claire.gardent@loria.fr.

We look forward to receiving your application!

10 December 2019



UNIVERSITÉ
DE LORRAINE



Campus scientifique

BP 239 – 54506 Vandœuvre-lès-Nancy Cedex

Tél. : 03 83 59 20 00 – Fax : 03 83 27 83 19 www.loria.fr