

Léo Jacqmin

PHD CANDIDATE

Lannion, France

☎ +33 (0)7 48 60 21 96 | ✉ leo.jacqmin@orange.com | 🏠 jacqle.github.io | 🌐 jacqle | in leo-jacqmin

Education

Aix-Marseille University

PHD IN COMPUTER SCIENCE

Marseille, France

October 2021 - September 2024 (est.)

- Supervised by Lina Maria Rojas Barahona (Orange) and Benoit Favre (Aix-Marseille Université)

IDMC - University of Lorraine

MSC IN NATURAL LANGUAGE PROCESSING

Nancy, France

September 2019 - August 2021

- Highest honors, GPA: 16.7/20, rank: 2/22

University of Liège

BA IN TRANSLATION AND INTERPRETING

Liège, Belgium

September 2015 - June 2018

- Exchange semester at the University of Innsbruck

Publications

L. Jacqmin, G. Marzinotto, J. Gromada, E. Szczekocka, R. Kołodyński, G. Damnati, “SpanAlign: Efficient Sequence Tagging Annotation Projection into Translated Data applied to Cross-Lingual Opinion Mining”, in Proceedings of the Seventh Workshop on Noisy User-generated Text (W-NUT 2021), 2021.

E. Calò*, **L. Jacqmin***, T. Rosemblatt*, M. Amblard, M. Couceiro, A. Kulkarni, “GECKo+: a Grammatical and Discourse Error Correction Tool”, in Actes de la 28e Conférence sur le Traitement Automatique des Langues Naturelles. Volume 3: Démonstrations, 6 2021.

Experience

Orange

PHD RESEARCHER

Lannion, France

October 2021 - September 2024

- Thesis' title : Dialogue state tracking in long dialogue contexts and in alignment with knowledge bases

Cognate AI

NLP CONSULTANT

Remote

June 2021 - August 2021

- Built a pedagogical tool for the automatic correction of written productions of learners of English
- Combined a deep learning model with rule-based approaches to provide error correction and personalized feedback

Orange Labs

RESEARCH INTERN

Lannion, France

March 2021 - August 2021

- Researched multitask approaches to multilingual aspect-based sentiment analysis
- Leveraged heterogeneous data with various annotation schemes to train and evaluate Transformer models for sequence tagging
- Proposed and implemented methods to increase robustness via domain adaptation and data augmentation using label projection

ATILF - CNRS

RESEARCH INTERN

Nancy, France

June 2020 - August 2020

- Collected and cleaned over 30 GB of raw textual data for training word embeddings
- Automated a training and evaluation pipeline for sets of word embeddings (Word2vec, fasttext, GloVe, ELMo and BERT)
- Researched data distribution for definition modeling with SentencePiece using data stored in RDF format

Seed-Up

NLP CONSULTANT

Remote

March 2020 - June 2020

- Helped build a rule-based text simplification system for French to make the web more accessible
- Managed freelance linguists for the production of syntactic simplification rules

Projects

GECKo+

- Implemented a writing assistant tool using two state-of-the-art models for grammar error correction and sentence ordering
- Built a web app that corrects spelling, grammatical, and discourse mistakes given a user input text using Flask, CSS and HTML
- Deployed this web app using Docker and Azure: *gecko-app.azurewebsites.net*

Languages

French Native Proficiency
English Full Professional Proficiency (TOEFL 108/120)
German Professional Working Proficiency

Skills

Programming languages Python, SQL, HTML, CSS
Development tools Bash (sed, awk, grep, ...), Git, Docker, Azure, Unix (Ubuntu)
Libraries scikit-learn, PyTorch, spaCy, Hugging Face Transformers, pandas, NumPy, Matplotlib, Gensim