Léo Jacqmin

Lannion, France

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

Education

Aix-Marseille University   *Marseille, France*

PhD in Computer Science   *October 2021 – Present*

- Thesis' title : Dialogue state tracking in long dialogue contexts and in alignment with knowledge bases
- Advised by Lina M. Rojas Barahona (Orange) and Benoit Favre (Aix-Marseille University)

University of Lorraine   *Nancy, France*

MSc in Natural Language Processing   *September 2019 – August 2021*

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

University of Liège   *Liège, Belgium*

BA in Translation and Interpreting   *September 2015 – June 2018*

- Erasmus exchange semester at the University of Innsbruck

Experience

Orange   *Lannion, France*

Industrial PhD Candidate   *October 2021 – Present*

- Implemented and evaluated Transformer models (T5) for robust language understanding in task-oriented dialog systems
- Authored research in NLP, surveying, reviewing, and publishing two articles in top conferences
- Placed first out of six research teams in the 11th Dialog System Technology Challenge track on spoken dialog understanding

Orange   *Lannion, France*

R&D Intern   *March 2021 – August 2021*

- Researched multitask approaches to multilingual opinion mining based on pre-trained language models (BERT)
- Proposed and implemented a transfer learning approach to analyze comments in languages with no access to annotated data
- First authored one paper at EMNLP workshop

Cognate AI   *Remote*

NLP Consultant   *January 2021 – March 2021*

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

ATILF - CNRS   *Nancy, France*

Research Intern   *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

Projects

GECko+: a Grammatical and Discourse Error Correction Tool

- Implemented an English 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: *github.com/psawa/gecko-app*

Skills

Programming	Python, SQL, HTML, CSS, LaTeX
Development tools	Shell scripting, Git, Docker, Azure, Slurm, regex
Data science libraries	scikit-learn, PyTorch, Hugging Face Transformers, pandas, NumPy, spaCy, NLTK, Gensim
Languages	French (native), English (C2), German (B2), Dutch (A2), Spanish (A1)

Publications

L. Jacqmin, L. Rojas Barahona, B. Favre "Do you follow me?": a Survey of Recent Approaches in Dialogue State Tracking, In Proceedings of the 23rd Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL), 2022.

L. Jacqmin, "« Est-ce que tu me suis ? » : une revue du suivi de l'état du dialogue", in Actes de la 29e Conférence sur le Traitement Automatique des Langues Naturelles. Volume 2 : 24e REncontres jeunes Chercheurs en Informatique pour le TAL (RECITAL), 2022.

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. Rosemplatt[*], 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 (TALN). Volume 3: Démonstrations, 6 2021.