Paul Lerner

Email: lerner@isir.upmc.fr
LinkedIn: paul-lerner-466997a8
GitHub: github.com/PaulLerner
Website: https://paullerner.github.io/
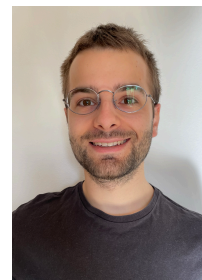
EDUCATION

Paris-Saclay University, CNRS, LISN (ex-LIMSI) Orsay, France
Ph.D. in Computer Science, Advisors: Olivier Ferret and Camille Guinaudeau 2020–Current

My Ph.D. is entitled *Knowledge-based Visual Question Answering about Named Entities.* It is centered around this task, which consists in answering questions about Named Entities grounded in a visual context. This task raises challenges for (Multimodal) Information Retrieval, Representation Learning, and Knowledge Representation. I collected and annotated a dataset to benchmark this task and set a first baseline where the text and image modalities are fused in a late fashion [1], [2]. This work was extended thanks to the internship of Salem Messoud on Multimodal Re-ranking [3]. I then worked on early fusion for Multimodal Retrieval, which relied on a pre-training task [4]. Finally, I studied Cross-Modal Retrieval as a complementary way to visually recognize named entities [5].

Paris Descartes University Paris, France
M.S. in Artificial Intelligence, with high honors (double degree with ESILV) 2018–2019

Various fields of Artificial Intelligence: machine learning and text mining but also decision theory and multi-agent systems.

ESILV Courbevoie, France
Engineer in Computer Science 2014–2019

ESILV is a generalist engineering school, one of the best "post-bac" schools in France (i.e., with integrated preparatory classes). It's focused on group work and cross-skills as it's part of a university pole along with a management and a web-design school.

Czech Technical University Prague, Czech Republic
B.S. in Computer Science (exchange with ESILV) 2016

Thanks to the ERASMUS program, I spent a semester at the Czech Technical University, where I studied C programming, Graph Theory, and learned a few words of Czech.

EXPERIENCE

Paris-Saclay University, CNRS, LIMSI Orsay, France
Research Engineer, Advisors: Hervé Bredin and Camille Guinaudeau October 2019–2020

TV series are filled with natural, multi-party dialogues challenging for many Speech and Natural Language processing tasks such as Speaker Diarization and Named Entity Recognition. Moreover, the intrinsic multimodality of videos gives rise to a number of different tasks which can be solved using one or several of the following: text, audio, and image. All being largely available through external sources given the nature of the data, which allows for a semi-automatic annotation [6]. This was done in the context of the PLUMCOT project, along with active learning and multimodal speaker identification experiments.

Télécom ParisTech Paris, France
Research Intern, Advisor: Laurence Likforman-Sulem March–September 2019

Parkinson's disease is the second most common neurodegenerative disease. Several researchers have attempted to provide a Clinical Decision Support System that would confirm or question a neurologist's diagnosis based on

a cheap and non-invasive handwriting exam, as many symptoms of Parkinson's Disease are observable through handwriting. We pursue their work by exploring deep learning techniques [7].

Sorbonne University (ex-Pierre and Marie Curie), CNRS, ISIR Paris, France

Engineer Intern, Advisors: Beatrice Biancardi and Catherine Pelachaud April–September 2018

In the context of Beatrice Biancardi's Ph.D., we aimed at designing an Embodied Conversational Agent able to manage the impression it makes on its human interlocutor. The agent uses different behavioral strategies to be perceived as warm and/or competent. I implemented a Reinforcement Learning algorithm using the user's engagement in the interaction as a reward [8]. To test our hypotheses, we designed and realized an experiment at Cité des sciences et de l'industrie, Paris, to evaluate the system [9].

TEACHING

- **Teaching Assistant** at Paris-Saclay University 2021–Current
  Introduction to Machine Learning and Natural Language Processing, Professors: François Landes and Kim Gerdes.
- **Teaching Assistant** at Polytech Paris-Saclay 2020–Current
  Introduction to imperative programming in C++ to students in preparatory classes, Professor: Frédéric Voisin.
  Programming in C++ to students in first year of engineering school, Professor: Joël Falcou.

SERVICE

- **Reviewer** for Pattern Recognition, ICMR 2023, RECITAL-RJCRI 2023, ACMMM 2022 –
- **Fixed-term workers' representative** at LIMSI's board 2019–2020

PUBLICATIONS

[1] **P. Lerner**, O. Ferret, C. Guinaudeau, H. Le Borgne, R. Besançon, J. G. Moreno, and J. Lovón Melgarejo, "ViQuAE, a Dataset for Knowledge-based Visual Question Answering about Named Entities", in *Proceedings of The 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, ser. SIGIR'22, New York, NY, USA: Association for Computing Machinery, 2022.

[2] **P. Lerner**, O. Ferret, C. Guinaudeau, H. Le Borgne, R. Besançon, J. G. Moreno, and J. Lovón Melgarejo, "Un jeu de données pour répondre à des questions visuelles à propos d'entités nommées en utilisant des bases de connaissances", in *Actes de la Conférence sur le Traitement Automatique des Langues Naturelles (TALN) 2022.*, Avignon, France: ATALA, 2022.

[3] **P. Lerner**, S. Messoud, O. Ferret, C. Guinaudeau, H. Le Borgne, R. Besançon, J. G. Moreno, and J. Lovón Melgarejo, "Un jeu de données pour répondre à des questions visuelles à propos d'entités nommées", *Traitement Automatique des Langues*, vol. 63, no. 2, 2023.

[4] **P. Lerner**, O. Ferret, and C. Guinaudeau, "Multimodal Inverse Cloze Task for Knowledge-Based Visual Question Answering", in *Advances in Information Retrieval (ECIR 2023)*, Cham: Springer Nature Switzerland, 2023, pp. 569–587, ISBN: 978-3-031-28244-7.

[5] **P. Lerner**, F. Olivier, and C. Guinaudeau, "Recherche cross-modale pour répondre à des questions visuelles", in *18e Conférence en Recherche d'Information et Applications*, H. Zargayouna, Ed., Paris, France: ATALA, 2023, pp. 74–92.

[6] **P. Lerner**, J. Bergoënd, C. Guinaudeau, H. Bredin, B. Maurice, S. Lefevre, M. Bouteiller, A. Berhe, L. Galmant, R. Yin, and C. Barras, "Bazinga! A Dataset for Multi-Party Dialogues Structuring", in *Proceedings of the Language Resources and Evaluation Conference*, Marseille, France: European Language Resources Association, Jun. 2022, pp. 3434–3441.

[7] **P. Lerner** and L. Likforman-Sulem, "Classification of Online Handwriting Time Series for Parkinson's Disease Diagnosis using Deep Learning", in *Proceedings of the 4th Junior Conference on Data Science and Engineering (JDSE)*, 2019.

[8] M. Mancini, B. Biancardi, S. Dermouche, **P. Lerner**, and C. Pelachaud, "Managing Agent's Impression Based on User's Engagement Detection", in *Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents*, ser. IVA '19, Paris, France: Association for Computing Machinery, 2019, pp. 209–211, ISBN: 9781450366724.

[9] B. Biancardi, M. Mancini, **P. Lerner**, and C. Pelachaud, "Managing an Agent's Self-Presentational Strategies During an Interaction", *Frontiers in Robotics and AI*, vol. 6, 2019, ISSN: 2296-9144.

Programming Skills

- **Deep learning:** PyTorch
- **Languages:** Python, C, C++, C#, Java
- **Python libraries:** Faiss, Transformers, spaCy, Scikit-learn, NumPy, Matplotlib, Seaborn
- **Computer clusters**: slurm
- **Databases:** SPARQL, Elasticsearch

Languages

- **French:** native
- **English:** fluent
- **Spanish:** elementary