

Jean-Baptiste LAMARE

Senior AI Scientist I | Deep Learning and NLP

[in linkedin.com/in/jlamare](https://www.linkedin.com/in/jlamare) github.com/jblamare [Google Scholar](https://scholar.google.com/citations?user=...) [Personal Website](https://www.jblamare.com) [i French citizen, H-1B visa](#)

EDUCATION

CARNEGIE MELLON UNIVERSITY (CMU), SCHOOL OF COMPUTER SCIENCE - PITTSBURGH, PA	2017 - 2019
M.S. Language Technologies - Focus : NLP, Computer Vision, Audio Processing	
ECOLE POLYTECHNIQUE, DEPARTMENT OF COMPUTER SCIENCE - PALAISEAU, FRANCE	2012 - 2017
B.S. & M.S. Data Science - Focus : Machine Learning, NLP, Databases, Mathematics, Computer Science	

PROFESSIONAL EXPERIENCE

COVERA HEALTH New York, NY	SENIOR AI SCIENTIST II (January 2024 - Present) SENIOR AI SCIENTIST I (November 2021 - January 2024) <ul style="list-style-type: none">Managed several end-to-end NLP projects to extract pathologies from clinical reports, from project definition and guideline creation to model training and deployment with F1 scores over 90%.Contributed to developing NER models for report structuring, pathology/location/severity entity extraction, and entity linking with common ontologies such as SNOMED and ICD10. These models are a core piece of a company-wide process to ingest and understand reports.Implemented distillation techniques to reduce our deployed models' size and latency by up to 45x.Developed prompt engineering methods and leaderboards to evaluate open LLMs on clinical tasks (LLaMA, Mixtral, Qwen...) and fine-tuned our LLMs for radiology discussion and QA. These models solved new zero-shot tasks with F1 scores over 80%.Created an internal image-text dataset and trained multimodal LLaVA and BLIP models for QA.Designed company-wide demos showcasing our pipelines, LLMs, and multimodal models.Developed and maintained shared repositories and python packages for deep learning model training (Keras-like), deployment, and NLP pipeline creation (spaCy-like).	2021 - PRESENT
ENLITIC San Francisco, CA	SENIOR RESEARCH SCIENTIST (April 2021 - November 2021) RESEARCH SCIENTIST (August 2019 - April 2021) <ul style="list-style-type: none">Developed self-supervised, supervised, and unsupervised NLP models for multiclass/multilabel classification, NER, and entity linking, using various architectures (transformer, BERT, RNN...) to detect over 70 abnormalities in clinical reports as well as extract and match entities to our ontology.Trained supervised multilabel computer vision and multimodal models at the intersection of CV and NLP to detect over 20 abnormalities and body parts in radiological images, specifically X-rays and CTs.Designed end-to-end labeling pipelines including interface design, case selection, and quality assurance, that would be used by over 15 experts to label more than 100,000 studies.Led a successful blind test demonstrating model outputs on unseen datasets for a prospective client.	2019 - 2021
CMU Pittsburgh, PA	GRADUATE RESEARCH ASSISTANT <ul style="list-style-type: none">Developed a Deep Learning useful-tweets extraction system for public safety events to identify tweets containing information that could help the authorities respond to such crises.Designed and implemented a web interface allowing users to visualize extracted tweets in real-time and to provide feedback in an active learning paradigm to fine-tune the model.	2017 - 2019

PUBLICATIONS

RSNA 2022 Scientific Pres.	AUTOMATED DL PIPELINE FOR PATHOLOGY CHARACTERIZATION IN MRI MSK REPORTS <i>J. Lamare, I. Daga, K. Kannan, D. Torres, J. Pawar, A. Das, B. Odry</i>	NOVEMBER 2022
RSNA 2022 Scientific Pres.	AUTOMATED STRUCTURE DETECTION FOR RADIOLOGY REPORT ANALYSIS <i>M. Sandora, J. Lamare, A. Sehanobish, I. Daga, K. Kannan, N. Abraham, A. Das, B. Odry</i>	NOVEMBER 2022
EMNLP 2020 Clinical NLP WS	ON THE DIMINISHING RETURN OF LABELING CLINICAL REPORTS <i>J. Lamare, T. Olatunji, L. Yao</i>	NOVEMBER 2020
AVSS 2018 T4S WS	ACCIDENT FORECASTING IN CCTV TRAFFIC CAMERA VIDEOS <i>A. Shah, J. Lamare, T. Anh, A. Hauptmann</i>	NOVEMBER 2018
ICMR 2018	MULTIMODAL FILTERING OF SOCIAL MEDIA FOR TEMPORAL MONITORING AND EVENT ANALYSIS <i>P. Huang, J. Liang, J. Lamare, A. Hauptman</i>	JUNE 2018

SKILLS

Programming Languages	Python, Java, C++, SQL, HTML, Javascript, CSS
Libraries and Tools	PyTorch, Tensorflow, Transformers, ONNX, spaCy, Stanza, Git, AWS, Docker, Kubernetes
Languages	English (fluent), French (native), German (intermediate), Spanish (beginner)

MISCELLANEOUS

Hobbies Competitive tennis and skiing, guitar, piano, chess, baking, traveling.