

JEAN-BAPTISTE LAMARE

☎ 412-626-8707 | ✉ jblamare@gmail.com

🌐 linkedin.com/in/jblamare | 🐙 github.com/jblamare | 🎓 Google Scholar | 🌐 Personal Website

EDUCATION

Carnegie Mellon University, School of Computer Science	Pittsburgh, PA
<i>M.S. Language Technologies - Focus: Natural Language Processing, Computer Vision, Audio Processing</i>	Aug. 2017 – Aug. 2019
Ecole Polytechnique, Department of Computer Science	Palaiseau, France
<i>B.S. & M.S. Data Science - Focus: Machine Learning, NLP, Databases, Mathematics, Computer Science</i>	Aug. 2012 – Aug. 2017

EXPERIENCE

Senior AI Scientist II	Jan. 2024 – Present
Senior AI Scientist I	Nov. 2021 – Jan. 2024
<i>Covera Health</i>	New York, NY
<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%.Developed NER models for report structuring, entity extraction, and entity linking with common ontologies (SNOMED, 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 to evaluate open LLMs on clinical tasks (LLaMA, Mixtral, Qwen, Gemma, Phi...).Fine-tuned 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.Contributed to demos involving LLMs enhanced with RAG and Agents for new applications.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).	
Senior Research Scientist	Apr. 2021 – Nov. 2021
Research Scientist	Aug. 2019 – Apr. 2021
<i>Enlitic</i>	San Francisco, CA
<ul style="list-style-type: none">Developed self-supervised, supervised, and unsupervised NLP models for classification, NER, and entity linking, using transformers and RNNs to detect >70 abnormalities in clinical reports, extract and link entities.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.	
Graduate Research Assistant	Aug. 2017 – Aug. 2019
<i>Carnegie Mellon University</i>	Pittsburgh, PA
<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.	

PUBLICATIONS

RSNA 2022 <i>Automated DL Pipeline for Pathology Characterization in MRI MSK Reports</i>	November 2022
RSNA 2022 <i>Automated Structure Detection for Radiology Report Analysis</i>	November 2022
EMNLP 2020 - Clinical NLP WS <i>On the diminishing return of labeling clinical reports</i>	November 2020
AVSS 2018 - T4S WS <i>Accident forecasting in CCTV traffic camera videos</i>	November 2018
ICMR 2018 <i>Multimodal filtering of social media for temporal monitoring and event analysis</i>	June 2018

TECHNICAL SKILLS

Programming Languages: Python, Java, C++, SQL (Postgres), JavaScript, HTML/CSS
Libraries: PyTorch, Tensorflow, Transformers, Adapters, LangChain, vLLM, MLFlow, ONNX, pandas, NumPy, spaCy, Stanza
Developer Tools: Git, Docker, Kubernetes, AWS, VS Code, PyCharm
Languages: English (fluent), French (native), German (intermediate), Spanish (beginner), Italian (beginner)