

THOMAS DOPIERRE

NLP Researcher (PhD) | Senior Software Engineer

[in linkedin.com/in/thomasdopierre](https://www.linkedin.com/in/thomasdopierre) github.com/tdopierre
☎ + (33) 7 87 13 52 81 ☎ + (44) 7 462 651 441 @ thomas.dopierre@pm.me
📍 London, United Kingdom 🏠 thomasdopierre.com

I'm a **Natural Language Processing (NLP)** researcher and **senior software engineer** with a PhD focused on user-generated short text modeling for conversational AI. I currently work at Bloomberg in London, where I design and deploy large-scale NLP systems, including query understanding and financial research agents. My expertise lies in few-shot and semi-supervised learning, with experience spanning from academic research to production-grade ML systems.

EXPERIENCE

-
- | | |
|--|--|
| Now 2022 | AI Senior Software Engineer, BLOOMBERG LP, London, UK <ul style="list-style-type: none">> Implemented automatic training, evaluation, and deployment pipelines of some of our text classification models, substantially reducing the required human intervention.> Created and deployed the first version of Bloomberg's Research Agent, capable of answering financial research-related queries in natural language.> Designed and implemented the core of our Query Understanding system, a heavily asynchronous application with dynamic dependencies.> Added caching layers on expensive functions, reducing the runtime latency of our query understanding system by 25%. <div style="display: flex; gap: 5px;">Python Docker NLP Transformers</div> |
| 2022 2017 | Research Data Scientist, MEETIC – MATCH GROUP, Paris, France <p>Key ML projects include :</p> <ul style="list-style-type: none">> Creation of the Natural Language Understanding part Lara, the first dating chatbot, who talks to thousands of singles everyday.> Building Computer Vision models to help users select their best profile pictures.> Development of a tool to analyze raw user queries, using state-of-the-art NLP models. I handled both the back-end and front-end parts of the tool.> Creation of a transformer-based text classification model to identify users with toxic behaviors.> Application of modern NLP methods to profile recommendation systems, to improve matching between singles looking for their other half. <div style="display: flex; gap: 5px;">Python AWS NLP PyTorch Transformers Computer Vision Spark</div> |
| June 2016 August 2016 | Actuary Intern, CNP ASSURANCES, Paris, France <p>Dependency & loss of income risk modeling.</p> <div style="display: flex; gap: 5px;">SQL SAS Excel</div> |

EDUCATION

-
- | | |
|----------------------------|---|
| 2021 2018 | PhD, LABORATOIRE HUBERT CURIEN, UMR CNRS 5516, Saint-Étienne, France <p>Research focused on intent detection, a special case of text classification in the conversational AI domain, with a large emphasis on few-shot learning methods, as conversational agents evolve with time, and labeled data are costly. Supervisors : Christophe Gravier and Julien Subercaze</p> <div style="display: flex; gap: 5px;">NLP Few-Shot Learning Semi-Supervised Learning Intent Classification Paraphrase Generation</div> |
| June 2019 | Conversational Intelligence Summer School (CISS2), UMASS LOWELL, Massachusetts <p>One-week summer school covering state-of-the-art Conversational NLP methods.</p> <div style="display: flex; gap: 5px;">NLP Conversational AI Visual Question Answering</div> |
| 2017 2016 | MSc - Final Year, ÉCOLE NORMALE SUPÉRIEURE, Paris, France <p>Kernel Methods, Probabilistic Graph Models, Convex Optimization, Monte-Carlo Markov Chains, Statistical Learning. Multiple team projects, including clustering, computer vision, and signal processing.</p> <div style="display: flex; gap: 5px;">Machine Learning Mathematics Deep Learning Python R MATLAB</div> |

2016 | Bachelor & MSc (1st year), UNIVERSITÉ PARIS-DAUPHINE, Paris, France
2012 | Mathematics, Programming, Economics & Finance.
Mathematics Finance C++ Java

💡 INTERESTS

Board games Strategic games (Splendor, 7Wonders, Risk, Azul, Chess), Tabletop role-playing games
Sports & Travel Running, Hiking, Bicycle touring
Music Practiced guitar for 10+ years

📖 PUBLICATIONS

2022 Kaustubh Dhole, Varun Gangal, ... **Thomas Dopierre**, ... (Multiple Authors)
[🔗 NL-Augmenter : A Framework for Task-Sensitive Natural Language Augmentation](#)
2021 **Thomas Dopierre**, Christophe Gravier, Wilfried Logerais
[🔗 PROT AUGMENT : Unsupervised diverse short-texts paraphrasing for intent detection meta-learning, ACL 2021](#)
2021 **Thomas Dopierre**, Christophe Gravier, Wilfried Logerais
[🔗 A Neural Few-Shot Text Classification Reality Check, EAACL 2021](#)
2020 **Thomas Dopierre**, Christophe Gravier, Julien Subercaze, Wilfried Logerais
[🔗 Few-shot Pseudo-Labeling for Intent Detection, COLING 2020](#)

📋 SKILLS

Programming Python (🔥 PyTorch, 😊 Transformers, Tensorflow, Flask, FastAPI), Docker, C++, JavaScript (VueJS)
Software IntelliJ's Pycharm, \LaTeX , Prodigy
AWS SageMaker, S3, DynamoDB, Lambda, Beanstalk, ECR, EC2
OS Mac OS X, Windows, Linux
Languages **French** (native) – **English** (fluent, TOEIC : 975/990)

★ COMPETITIONS & AWARDS

2022 **Meetic Hackathon** – 1st prize
Group project in which I implemented a multi-modal (i.e. text / image / speech) zero-shot classification model to derive hobbies from user data.
2019 **Summer School** – Group project : 1st prize
Developed a Visual Question-Answering game where users had to find the correct picture by asking questions about it. We implemented models in Pytorch, and the game's front-end on Telegram.
2019 **Meetic Hackathon** – 1st prize
Group project where we implemented a Computer Vision system to improve the quality of matches.
2017 **Data Challenge** – **École Normale Supérieure** – 1st prize
We built a Neural Network to predict the age of patients using their brain's activity (EEG) during their sleep.

🎓 TEACHING AND MENTORING

2016 | College Tutor, UNIVERSITÉ PARIS-DAUPHINE, Paris, France
2014 | Helped 1st and 2nd year college students in various mathematical fields.
Probabilities Statistics Linear Algebra