

Thomas Dopierre

3rd year PhD NLP student



Contact

(+33) 7 87 13 52 81

thomas.dopierre@pm.me

75 rue Fondary
75015 Paris, France

Languages

French

English

TOEIC 975/990

Programming

♥ Python: Pytorch, Transformers (Hugging Face), Tensorflow, Flask
R, C++, SQL, JavaScript (VueJS)

Technologies

OpenSource, Linux, AWS (SageMaker, Lambda, EB, ECR, DynamoDB), Web Crawling, git, Latex

Hobbies

Tabletop role-playing & board games, logic puzzles, guitar

Bicycle touring, Running, Hiking

summary

I am a 3rd year PhD student at the **Hubert Curien** laboratory, in France. My research is focused on processing and modeling user-generated short texts to improve Natural Language Understanding in human-agent interactions. I am particularly studying few-shot and semi-supervised learning, as labeled data are often costly to obtain.

experience

- 2017–Now **Meetic–Match Group** Paris, France
NLP Research Scientist
 I work on Lara, the first dating chatbot, who talks to thousands of singles everyday. She recommends profiles, provides dating tips, answers questions regarding customer care, etc. My task is to ensure her proper understanding of the human language. I also endorse the role of an adviser, as I interact with the full data science team on a daily basis.
- Jun–Aug 2016 **Data Lab–CNP Assurances** Paris, France
Actuary Intern
 Dependency & loss of income risk modeling.

education

- June 2019 **Conversational Intelligence Summer School (CISS2)** UMASS Lowell, MA
 Group project on a Question-Answering Conversational Agent-based game, combining both NLP and Computer Vision. **Group Project: 1st prize**
- 2018–Now **PhD** Laboratoire Hubert Curien, UMR CNRS 5516
 My research is focused on intent detection, a special case of text classification in the conversational AI domain. I am particularly studying few-shot learning methods, as conversational agents evolve with time, and manually exploring chat logs as well as annotating rare labels is both time and cost expensive for humans. Supervisors: Christophe Gravier and Julien Subercaze
- 2016–2017 **MSc (final year)** ENS Cachan, Paris, France
Mathematics, Statistics, Machine Learning
 Many team projects with various ML algorithms. Among other things, we studied Kernel Methods, Probabilistic Graph Models, Convex Optimization, Monte-Carlo Markov Chains. 1st prize for ML project. **Rank:1/18**
- 2012–2016 **Bachelor & MSc (1st year)** Paris-Dauphine University, Paris, France
 Mathematics, Programming, Economics & Finance. **Rank:2/63**

publications

- ACL'21 **ProtAugment: Unsupervised diverse short-texts paraphrasing for intent detection meta-learning**
 Thomas Dopierre, Christophe Gravier, Wilfried Logerais
- EACL'21 **A Neural Few-Shot Text Classification Reality Check**
 Thomas Dopierre, Christophe Gravier, Wilfried Logerais
- COLING'20 **Few-shot Pseudo-Labeling for Intent Detection**
 Thomas Dopierre, Christophe Gravier, Julien Subercaze, Wilfried Logerais