



Louis Martin

PH.D. STUDENT

(+33) 6 45 82 43 31 | louismartin@gmail.com | louismartin.eu | [louismartin](https://www.linkedin.com/in/louismartin)
[louismartincs](https://github.com/louismartincs) | [Louis Martin](https://www.instagram.com/LouisMartin)

**PhD Student Specialized in Artificial Intelligence and Natural Language Processing
with Strong Research and Technical Skills.**

Experience

Facebook AI Research & Inria Paris

Paris, France

PH.D. STUDENT

May. 2018 - Present

- Created a demo based on my research in Text Simplification that was **presented to Mark Zuckerberg**. It was also **showcased by Facebook's CTO** at the company-wide *All Hands* conference, in front of 15k employees.
- My research involved **challenging large-scale engineering skills**: I trained CamemBERT Language Model in parallel on 256 GPUs on 138GB of training text, I scaled my text mining pipeline using neural semantic embeddings to up to 1 billion sentences by optimizing disk IOs, CPU preprocessing and GPU nearest neighbor search, all of this on up to 1000 GPUs with preemption-aware logic.
- My latest text simplification method achieves **state-of-the-art results** on all simplification benchmarks while being unsupervised.
- Integrated my code and models into Facebook's codebase and infra. My code is **currently being ported in production**.
- Major success of the release of our **open-source CamemBERT model** [↗](#) (>20k downloads/month, multiple industry applications).
- Presented my research in person to the **French Minister for People with disabilities** (my research is core to the Cap'FALC [↗](#) project, involving the French Government, UNAPEI association, Inria and Facebook AI Research).
- My work was featured in 9 news articles/podcasts, I gave 3 interviews (VentureBeat [↗](#), Le Monde [↗](#), France Culture [↗](#), ActuaIA [↗](#), ...).

Facebook AI Research

Paris, France

NLP RESEARCH INTERN

May. 2017 - Nov. 2017

- Integrated a neural conversational model in Facebook Messenger (PHP and Python) as an internal chatbot demo.
- Trained a neural model for question generation on 100k question/answer pairs as an additional module to the chatbot.

GoPro

Paris, France

COMPUTER VISION R&D INTERN

Feb. 2016 - Jul. 2016

- Developed and trained a compact CNN architecture (3MB) with state-of-the-art performance for face recognition.
- I implemented **on-device neural inference** for Android (Java and Tensorflow).
- My work led to the submission and acceptance of a **US patent (sole inventor)** on compact neural models.

BNP Paribas

Paris, France

ANALYST - DEVELOPER INTERN (ASSET & LIABILITY MANAGEMENT)

Sep. 2015 - Jan. 2016

- Developed tools in C++ and PL/SQL for the simulation of liquidity, interest rates and cross-currency risks.

Education

Facebook AI Research, Inria Paris, Sorbonne Université

Paris, France

PH.D. IN COMPUTER SCIENCE

May. 2018 - Aug. 2021

- Subject: *Deep Learning for Text Simplification using Unsupervised Multilingual Approaches*. Ph.D. at Facebook AI Research and Inria Paris (ALMAnaCH team) supervised by Antoine Bordes (Facebook), Benoît Sagot (Inria) and Éric de la Clergerie (Inria).

École Normale Supérieure Paris-Saclay

Cachan, France

MASTER OF SCIENCE - MVA (MATHEMATICS, VISION AND MACHINE LEARNING)

Oct. 2016 - Apr. 2017

- Computer Vision (C. Schmid, J. Ponce), Deep Learning (I. Kokkinos), Reinforcement Learning (A. Lazaric), Sparsity and Compressed Sensing (G. Peyre), Kernel Methods (J. Mairal, J.-P. Vert).

University of New South Wales

Sydney, Australia

POSTGRADUATE EXCHANGE - MACHINE LEARNING

March. 2015 - Jul. 2015

- Artificial Intelligence, Machine Learning, Bayesian Inference.

École Centrale Paris (now CentraleSupélec)

Châtenay-Malabry, France

ENGINEERING DEGREE (MS EQUIVALENT) - APPLIED MATHEMATICS

Sept. 2013 - May. 2017

- Machine Learning, Natural Language Processing, Distributed Optimization, Stochastic Processes, Non-linear PDEs.

Teaching

- 2019 - 2020 **African Master's in Machine Intelligence** - *Teaching Assistant* - Deep NLP (Antoine Bordes), class on site in Rwanda.
- 2020 **MVA Master** - *Teaching Assistant* - Deep Learning (Vincent Lepetit). Lecture hall class and NLP final class project.
- 2019 - 2020 **OpenClassrooms** - *Data Science Mentor*. Weekly mentoring for people in reconversion in Data Science.

Honors & Awards

- 2020 **Outstanding Reviewer (EMNLP 2020)**
- 2015 **3rd position out of 3242 in Kaggle Bike Sharing Demand Competition** ↗: Entered the competition one week before the end, and finished 3rd using neural networks for the prediction of shared bikes demand.
- 2014 **Merit Scholarship**, Granted by École Centrale Paris, for top 1% of students with very good academic results with a significant positive impact on society (thanks to volunteering).

Miscellaneous

- 2019 **[Side Project] Creation of Greenhouse Gas Dashboard**. Built an internal dashboard at Facebook to monitor and raise awareness on realtime CO2 emissions of our GPU cluster, with very positive internal reception.
- 2013 - 2014 **[Volunteering] President of Mada ECP Association**. Led a group of 12 students. We developed, gathered funds (50k€), and built onsite a refuge to help children of the streets of Madagascar. Rewarded with merit scholarship.
- 2012 **[Music] Classical Percussions at the French Conservatory**. Played percussions and drums, including a performance at the famous Paris concert hall *La Cigale* as the final stage of the *Fallenfest* competition.

Skills and tools

- **Programming:** Python (master - NumPy, PyTorch, pandas...), C/C++, Java, R, Matlab, JavaScript, PHP, HTML5, SQL.
- **Open-Source** (🔗 github.com/louismartin): EASSE (maintainer), fairseq (contributor), HuggingFace Transformers (contributor), submitit (early contributor), nevergrad (early contributor).
- **Tools:** Vim (master), Git (master), Linux, web servers, AWS, OVH, tmux, LaTeX, Slurm, VS Code, macOS.
- **Languages Spoken:** French (Native), English (Fluent), German (Written and Spoken), Spanish (Written and Spoken).

Publications & Patents

- [1] **Louis Martin**, Angela Fan, Éric de la Clergerie, Antoine Bordes (Facebook), Benoît Sagot. *MUSS: Multilingual Unsupervised Sentence Simplification by Mining Paraphrases*. Submitted to NAACL 2021
- [2] Fernando Alva-Manchego*, **Louis Martin***, Antoine Bordes, Carolina Scarton, Benoît Sagot, Lucia Specia. *ASSET: A Dataset for Tuning and Evaluation of Sentence Simplification Models with Multiple Rewriting Transformations*. **ACL 2020**
- [3] **Louis Martin**, Benoît Sagot, Eric de la Clergerie, Antoine Bordes. *Controllable Sentence Simplification*. **LREC 2020**
- [4] **Louis Martin***, Benjamin Muller*, Pedro Javier Ortiz Suárez*, Yoann Dupont, Laurent Romary, Éric Villemonte de la Clergerie, Djamé Seddah, Benoît Sagot. *Les modèles de langue contextuels Camembert pour le français: impact de la taille et de l'hétérogénéité des données d'entraînement*. **TALN 2020** (nominated for best paper award)
- [5] **Louis Martin***, Benjamin Muller*, Pedro Javier Ortiz Suárez*, Yoann Dupont, Laurent Romary, Éric Villemonte de la Clergerie, Djamé Seddah, Benoît Sagot. *CamemBERT: a Tasty French Language Model*. **ACL 2020**
- [6] Fernando Alva-Manchego*, **Louis Martin***, Carolina Scarton, Lucia Specia. *EASSE: Easier Automatic Sentence Simplification Evaluation*. **EMNLP 2019** - System Demonstrations
- [7] Ganesh Jawahar, Benjamin Muller, Amal Fethi, **Louis Martin**, Éric Villemonte de La Clergerie, Benoît Sagot, Djamé Seddah. *ELMoLex: Connecting ELMo and lexicon features for dependency parsing*. **CoNLL 2018** - Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies
- [8] **Louis Martin**, Samuel Humeau, Pierre-Emmanuel Mazaré, Antoine Bordes, Éric Villemonte de La Clergerie, Benoît Sagot. *Reference-less Quality Estimation of Text Simplification Systems*. **INLG 2018** - 1st Workshop on Automatic Text Adaptation (ATA)
- [9] **Louis Martin**. *Systems and Methods for Compact Convolutional Neural Networks*. **US Patent** 10,185,891 (2019)

*Equal contribution.