

Nathan Godey

Post-doctoral Associate – Cornell

Astoria, New York

+1 347-992-8697

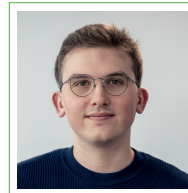
✉ godeynathan@gmail.com

🌐 nathangodey.github.io

🔄 NathanGodey

in [nathan-godey-114298175](#)

🔑 YlbsNP8AAAAJ



I work on the representations of language models, characterizing their geometry, finding where standard architectural choices silently constrain training or inference, and proposing alternatives.

Academic Experience

- Sep 25 – Sep 27 **Post-doctoral Associate**, Cornell University, New York City
Supervised by Yoav Artzi, as part of the Empire AI initiative.
- Feb 25 – May 25 **Post-doctoral Researcher**, Inria, Paris
Project: Training several bilingual large language models (LLMs) from scratch on national clusters. Development of a custom codebase to train models from 1B to 24B parameters.

Education

- Dec 21 – Dec 24 **PhD in Natural Language Processing**, Inria & Sorbonne, Paris
Thesis: *"Improving Representations for Language Modeling"*.
Supervision: Éric de la Clergerie and Benoît Sagot.
- Sep 20 – Sep 21 **Master, Machine Learning**, ENS Paris-Saclay & CNAM, Paris area
GPA: 4.0 (17.3/20)
Courses: Deep Learning, NLP & Speech, Computer Vision, Optimal Transport, Neuroscience, Metaheuristics, Graphs in ML.
- Sep 17 – Sep 20 **Engineering Degree, Maths & CS**, École des Ponts, Paris area
GPA: 3.65
Courses: Machine Learning, Algorithmics, Optimization, Probability, Statistics.

Publications

- ACL 2025 *Gaperon: A Peppered English-French Generative Language Model Suite*
Nathan Godey, Wissam Antoun, Rian Touchent, Rachel Bawden, Éric de la Clergerie, Benoît Sagot, Djamé Seddah
- ACL 2025 *Biomed-Enriched: A Biomedical Dataset Enriched with LLMs for Pretraining and Extracting Rare and Hidden Content*
Rian Touchent, **Nathan Godey**, Éric de la Clergerie
- ICLR 2025 *Q-Filters: Leveraging Query-Key Geometry for Efficient Key-Value Cache Compression*
SLLM **Nathan Godey**, Alessio Devoto, Yu Zhao, Simone Scardapane, Pasquale Minervini, Éric de la Clergerie, Benoît Sagot
- COLM 2024 *Why do small language models underperform? Studying LM Saturation via the Softmax Bottleneck*
Nathan Godey, Éric de la Clergerie, Benoît Sagot
- COLING 2024 *On the Scaling Laws of Geographical Representation in Language Models*
Nathan Godey, Éric de la Clergerie, Benoît Sagot
- ICLR 2024 *Headless Language Models: Learning without Predicting with Contrastive Weight Tying*
Nathan Godey, Éric de la Clergerie, Benoît Sagot
- EACL 2024 *Anisotropy Is Inherent to Self-Attention in Transformers*
Nathan Godey, Éric de la Clergerie, Benoît Sagot
- ACL 2023 *Is Anisotropy Inherent to Transformers?*
SRW **Nathan Godey**, Éric de la Clergerie, Benoît Sagot

EMNLP 2022 *MANTa: Efficient Gradient-Based Tokenization for Robust End-to-End Language Modeling*
Nathan Godey, Roman Castagné, Éric de la Clergerie, Benoît Sagot

Awards

2025 **Best PhD Dissertation Award** - French National NLP Association (ATALA)

Academic Visits

Apr 24 – Jul 24 **Visiting PhD Student**, *University of Edinburgh*, UK
3-month collaboration with Edoardo Ponti as part of my PhD research on gradient-based KV cache compression.

Teaching

Dec 23 – Mar 24 **Advanced NLP**, *Lecturer*, Master SCIA / EPITA, Paris
Wrote and gave a state-of-the-art NLP lecture for AI Masters' final year students.
[Link to the lecture](#)

Dec 22 – May 23 **Python for Mathematics**, *Teaching Assistant*, Sorbonne Université, Paris
Taught numerical techniques in Python in various applied maths domains.

Dec 21 – May 22 **Linear Algebra**, *Teaching Assistant*, Sorbonne Université, Paris
Supervised Linear Algebra exercises in blackboard sessions.

Industry Experience

Jan 20 – Jul 20 **AI Engineer intern**, *Stellia AI*, Paris
Chatbot and search engine development for an artificial teaching assistant.

Jul 19 – Dec 19 **Software Engineer intern**, *Arcane*, Paris
Back-end and front-end development.

Talks

- Seminar, Naver Labs Europe (11/2024)
- *Topics in NLP* course, EPFL NLP Group (11/2024)
- PhD seminar, Inria (02/2024)
- DeepVoice seminar, Sorbonne Université (06/2022)

Blog Posts

Dec 2023 Chinchilla Scaling Laws Calculator
Mar 2023 How word frequency affects language models

Service

Reviewer EMNLP'22-23, ACL'23-25, EACL'24, NAACL'24, ICLR'24, ICLR'26 (Gold), COLM'26.
Area Chair ACL 2025.

Key Skills

Programming Python, C++, SQL, JavaScript
Frameworks PyTorch, Lightning, Transformers, Slurm
Hardware Multi-GPU, Multi-node, Tensor Parallelism, NVIDIA/AMD
Tools Git, Linux, LaTeX