# Gautier Dagan

PhD Candidate in Natural Language Processing

⌖ gautier.tech    gautierdag    in gautier-dagan    scholar

📍 Edinburgh, UK

✉ gautier.dagan@ed.ac.uk

📞 +44 742 383 6797

## EDUCATION

**The University of Edinburgh** — 2021 – (anticipated) 2025 — *Edinburgh, UK*
*Integrated PhD in Natural Language Processing*
*Supervisors: Prof. Alex Lascarides and Prof. Frank Keller*

**The University of Amsterdam** — 2017 – 2019 — *Amsterdam, NL*
*Master of Science in Artificial Intelligence; Cum Laude*

**Northwestern University** — 2013 – 2017 — *Evanston, USA*
*BA in Computer Science, Economics and International Studies (Triple Major)*

## EXPERIENCE

**Technical Member of Staff** — 01/2026 - present
*Cohere* — *London, UK*
- Member of the next-gen agents team, enabling better planning, memory and learning in LLM agents.

**Research Scientist Intern** — 07/2023 - 11/2023
*FAIR, AI at Meta* — *Paris, FR*
- Conducted research on tokenizers in the Code Generation team under the supervision of Baptiste Rozière and Gabriel Synnaeve. Published at ICML 2024.

**Senior NLP Engineer (Part-Time)** — 01/2021 - 03/2023
*Research Grid* — *London, UK*
- Led development of patent for clinical trial research automation, deployed ML APIs, managed interns, and assisted in hiring process.

**Machine Learning Engineer** — 09/2019 - 09/2020
*Mansa (acquired by Updraft)* — *London, UK*
- Implemented an OCR solution for bank statements of 10 major French banks to reduce on-boarding attrition and augment Open Banking coverage.

## SELECTED PUBLICATIONS

**Gautier Dagan**, Frank Keller, Alex Lascarides $How^2$: How to learn from procedural How-to questions preprint |paper |code

**Gautier Dagan**, Frank Keller, Alex Lascarides *Plancraft: an evaluation dataset for planning with LLM agents* COLM 2025 |paper |code

**Gautier Dagan**, Frank Keller, Alex Lascarides *Dynamic planning with a llm* NeurIPS 2024 Language Gamification Workshop (Spotlight) |paper |code

**Gautier Dagan**, Olga Loginova, Anil Batra *CAST: Cross-modal Alignment Similarity Test for Vision Language Models* COLING 2025 |paper |code

**Gautier Dagan**, Gabriel Synnaeve, Baptiste Rozière *Getting the most out of your tokenizer for pre-training and domain adaptation* ICML 2024 |paper |code

**Gautier Dagan**, Frank Keller, Alex Lascarides *Learning the Effects of Physical Actions in a Multi-modal Environment* EACL Findings 2023 |paper |code

## SKILLS

**Technologies**: Python, Git, Pytorch, AWS, Docker, Serverless, transformers, lightning

**Certifications**: AWS Machine Learning Speciality (08/2020), AWS Solutions Architect Associates (06/2020)

**Languages**: native English and French, Spanish (A2)

**Open-Source**: bpeasy, kblaunch, batche

**Competitions**: IGLU: NLP Task 2022 (4th Place), Kaggle Competition Expert (**top 1%**), RoboCup IranOpen 2018

**Reviewer**: EMNLP 22, EACL 23, ACL 23, COLING 24, NeurIPS 24, ARR 24, JAIR 24, ARR 25, COLM 25