

Gautier Dagan

PhD Candidate in Natural Language Processing

 [gautier.tech](https://github.com/gautier.tech)


 [gautierdag](https://www.linkedin.com/company/gautierdag)

 [gautier-dagan](https://twitter.com/gautier-dagan)

 [scholar](https://orcid.org/0009-0001-9030-4000)

 Edinburgh, UK

 gautier.dagan@ed.ac.uk

 +44 742 383 6797

EDUCATION

The University of Edinburgh

Integrated PhD in Natural Language Processing

Supervisors: Prof. Alex Lascarides and Prof. Frank Keller

2021 – (anticipated) 2025

Edinburgh, UK

The University of Amsterdam

Master of Science in Artificial Intelligence; Cum Laude

2017 – 2019

Amsterdam, NL

Northwestern University

BA in Computer Science, Economics and International Studies (Triple Major)

2013 – 2017

Evanston, USA

EXPERIENCE

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 and deployment of 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.

Co-founder

02/2016 – 01/2018

Zcruit

Evanston, USA

- Co-founded and successfully exited a predictive analytics start-up in college football recruiting.

SELECTED PUBLICATIONS

Gautier Dagan, Frank Keller, Alex Lascarides *Plancraft: an evaluation dataset for planning with LLM agents* Preprint | [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](#)

Gautier Dagan, Dieuwke Hupkes, and Elia Bruni *Co-evolution of Language and Agents in Referential Games* EACL 2021 | [paper](#) | [code](#)

Yann Dubois, **Gautier Dagan**, Dieuwke Hupkes, and Elia Bruni *Location Attention for Extrapolation to Longer Sequences* ACL 2020 | [paper](#)

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