

# Antoine BOURGOIS

@ antoine.bourgois@protonmail.com |  GitHub |  Google Scholar |  LinkedIn |  Paris, France

## EDUCATION

---

**École normale supérieure (ENS-PSL)** Paris, France  
*Ph.D. in Language Sciences - Natural Language Processing* 2025 – 2028

**École des Hautes Études en Sciences Sociales (EHESS)** Aubervilliers, France  
*M.A. in History of Science, Technology and Knowledge* 2022 – 2024

**Coursework:** Philosophy of Science, History of Science, Science and Technology in Society, Environmental Humanities, Quantitative Sociology, Social Network Analysis, Digital Sociology, Quantitative Approaches and Mathematical Modeling in the Social Sciences and Humanities

**M1 Thesis:** The University of Technology of Compiègne: a limit case of a humanities and social sciences teaching project in an engineering school. Sociohistory of a curricular construction (1973-2023)

**Université de Technologie de Compiègne (UTC)** Compiègne, France  
*Engineering Degree (M.Sc. equivalent), Mechanical Engineering* 2019 – 2022

**Coursework:** Materials Science, Computer-Aided Design, Programming, Signal Processing, Statistics, History of Science, Philosophy, Ethics, Sociology of Organizations

**English:** TOEIC 950/990

**Institut Universitaire de Technologie de Lille (IUT)** Lille, France  
*University Diploma of Technology (DUT), Mechanical and Production Engineering* 2017 – 2019

## EXPERIENCE

---

**LATTICE (CNRS & ENS-PSL & Sorbonne Nouvelle)** Paris, France  
*Doctoral Researcher (Contractual Ph.D. Student)* 2025 – 2028

- Thesis: Computational study of characters in French-language novels.

*Study Engineer* June 2024 – October 2025

- Developed a computational pipeline for literary works: entity detection, coreference resolution, neural model training, performance evaluation, and large-scale inference over a corpus of novels.
- Created Propp-fr, a Python library.

*Natural Language Processing Research Intern* February 2024 – June 2024

- Reviewed the state of the art in named entity recognition and coreference resolution.
- Annotated corpora.

**RESORTECS, Textile Recycling / Circular Economy** Brussels, Belgium  
*Test Engineer* Spring 2022, 6 months, Internship

- Optimized the assembly and disassembly processes of thermoplastic sewing threads.
- Evaluated the thermomechanical impact on the breaking strength of seams.

**HYCCO, Hydrogen Fuel Cells** Albi, France  
*R&D Engineer* Autumn 2021, 6 months, Internship

- Characterized the electrochemical corrosion resistance of composite materials.

**VALLOUREC, Metallurgy and Industrial Equipment** Aulnoye-Aymeries, France  
*Assistant R&D Engineer* 2018 – 2019, 12 months, Apprenticeship

- Qualified industrial heat treatment equipment.

## PUBLICATIONS

---

**2026 | Closing the Gap at CRAC 2026: Two-Stage Adaptation for LLM-Based Multilingual Coreference Resolution** 

**Antoine Bourgois**, Olga Seminck, Thierry Poibeau. *Proceedings of the 2nd Joint Workshop on Computational Approaches to Discourse, Context and Document-Level Inferences and Computational Models of Reference, Anaphora and Coreference (CODI-CRAC 2026)*, San Diego, California, USA.

**2026 | Toward an Ontological Representation of Fictional Characters** 

**Antoine Bourgois**, Jean Barré, Olga Seminck, Thierry Poibeau. *Journal of Computational Humanities Research*, 2026, 1–15.

**2025 | Modeling the Construction of a Literary Archetype: The Case of the Detective Figure in French Literature** 

Jean Barré, Olga Seminck, **Antoine Bourgois**, Thierry Poibeau. *Anthology of Computers and the Humanities*, Vol. 3, 2025, pp. 983–999.

**2025 | The Elephant in the Coreference Room: Resolving Coreference in Full-Length French Fiction Works** 

**Antoine Bourgois**, Thierry Poibeau. *Proceedings of the Eighth Workshop on Computational Models of Reference, Anaphora and Coreference (CRAC 2025)*, EMNLP 2025, Suzhou, China.

**2025 | GLaRef@CRAC2025: Should We Transform Coreference Resolution into a Text Generation Task?**

Olga Seminck, **Antoine Bourgois**, Yoann Dupont, Mathieu Dehouck, Marine Delaborde. *Proceedings of the Eighth Workshop on Computational Models of Reference, Anaphora and Coreference (CRAC 2025)*, EMNLP 2025, Suzhou, China.

## TALKS

---

**February 2026 | Propp (Pattern Recognition and Ontologies for Prose Processing): A Computational Analysis Pipeline for Works of Fiction**

*AI & Social Sciences Seminar, Computational Social Science group, Institut Polytechnique de Paris (CSS @ IP-Paris)*

**May 2025 | Propp-fr: An Automatic Processing Pipeline for French Literary Works**

*Mate-SHS, CNRS Humanities and Social Sciences engineers' network*

**March 2025 | BookNLP-fr: A Computational Analysis Pipeline for French Fiction**

*Seminar "Artificial Interpretations", CSLF and MSH Mondes*

## SOFTWARE

---

**Propp: Pattern Recognition and Ontologies for Prose Processing** 

A Natural Language Processing framework for narrative analysis. Includes the trained French coreference pipeline, available as the open-source `propp_fr` Python library.