

# Fanny DUCCEL

Nancy/Orsay, France | fanny.ducel@universite-paris-saclay.fr | fannyducel.github.io

## EDUCATION

---

**PhD Candidate in Computer Science** 2023 – Now

*LISN, Université Paris-Saclay, Orsay, France*

Thesis topic: Evaluation of stereotypical biases in autoregressive language models.

Under the supervision of Aurélie Névéol and Karën Fort.

**Master in Linguistics and Computer Science, with highest honours** 2021 – 2023

*Sorbonne Université, Paris, France*

- M2 Thesis on the Evaluation of stereotypical biases in auto-regressive language models.
- M1 Thesis on the Analysis of claims in NLP articles by unsupervised learning.

**Bachelor in Language Sciences, Computer Science focus, with highest honours** 2018 – 2021

*Sorbonne Université, Paris, France*

## RESEARCH EXPERIENCE

---

**Research Internship (M2)** March 2023 – September 2023

*LORIA, Nancy, France* – supervision: Karën Fort and Aurélie Névéol

- Literature review on stereotypical biases in autoregressive language models with a critical approach
- Experiments with manipulation of masked and autoregressive language models for bias assessment, reproduction and adaptation of bias mitigation and evaluation experiments
- Participation in the international research project MultiCrowsPairs, along with 20 other researchers and students from 8 different countries

**Research Assistant (part-time)** September 2022

*CERES, Sorbonne Université, Paris, France* – supervision: Gaël Lejeune

- Participation in a project to map researchers according to their publications : wrote Python scripts to test new vectorization and clustering techniques, created graphs in Gephi, reorganized the Github repository with proper documentation

**Research Internship (M1)** July 2022 – September 2022

*LORIA, Nancy, France* – supervision: Karën Fort and Maxime Amblard

- Literature review on the linguistic concepts of *hedging* and *claims*
- Creation and pre-processing of ACL articles corpus
- Extraction of *claims* in the corpus and classification by using *clustering* methods in Python
- Evaluation of unsupervised learning methods

**Research Internship (L3)** Juin 2021 – Août 2021

*LORIA, Nancy, France* – supervision: Karën Fort, Gaël Lejeune and Yves Lepage

- Manual annotation of NLP articles (TALN, LREC, ACL)
- Automatic classification of these articles depending on whether the studied language is mentioned
- Work subsequently accepted at LREC and TALN

## TEACHING

---

**Project Management Tools** - CM 2h, TD 5h 2025-2026

*M1 TAL, IDMC, Université de Lorraine*

- Introduction to Grid 5000 : first connection, key notions (site/cluster/job), essential commands (file transfer, resource reservation).
- Introduction to  $\LaTeX$ : essential commands, first document.
- Creation of materials.
- Course in English.

**Software projects** - TD 10h 2025-2026

*M2 TAL, IDMC, Université de Lorraine*

- Advising and evaluating groups.
- Course in English.

**Ethics in NLP** - CM 2h 2025-2026

*M2 TAL, IDMC, Université de Lorraine*

- Presentation of my research projects.

- Course in English.
- Python for beginners** - TD 16h 2025-2026  
*M1 SC, IDMC, Université de Lorraine*
- Introducing labs, helping out students to complete exercises and grading their assignments. Course in French.
- De l'Écrit à l'Information** - TD 4h 2024-2025  
*L1 MIASH, IDMC, Université de Lorraine*
- Introduction to linguistics. Course in French.
- Ethics in NLP** - TD 10h 2024-2025  
*M1/M2 TAL, IDMC, Université de Lorraine*
- Introduction to ethics concepts, supervising group projects, evaluation and grading of oral presentations.
  - Course in English.
- Python for beginners** - CM 14h (21h HETD) 2024-2025  
*M1 TAL/SC, IDMC, Université de Lorraine*
- Lectures to introduce Python notions: data structures, functions, libraries, introduction to object-oriented.
  - Creation of some of the materials.
  - Course in English.
- Python for beginners** - TD 16h 2024-2025  
*M1 TAL/SC, IDMC, Université de Lorraine*
- Introducing labs, helping out students to complete exercises and grading their assignments.
  - Creation of some of the materials.
  - Course in English.
- Introduction to Grid 5000** - CM 2h, TD 2h 2024-2025  
*M1 TAL, IDMC, Université de Lorraine*
- Introduction to Grid 5000 : first connection, key notions (site/cluster/job), essential commands (file transfer, resource reservation).
  - Creation of materials.
  - Course in English.
- Advanced Python Programming** - TD 18h 2023-2024  
*M1 TAL/SC, IDMC, Université de Lorraine*
- Introducing labs, helping out students to complete exercises and grading their assignments. In French.
- Tutor in Computer Science for Language Science Studentse** 2021-2023  
*Sorbonne Université*
- Shifts of 3 hours per week with students of the Bachelor in Language Sciences needing help to understand and complete their exercises and projects in Python programming, Web development (HTML and CSS), language engineering (TXM, Antconc, GATE), first order logic, ... In French.

## SUPERVISION

---

- Co-supervision of NLP master's students internships (2 months)** 2024-2025  
*Université de Lorraine*
- Group 1: Racist bias detection in language models through code-switching. Loïc Leclere, Arthur Pedrini.
  - Group 2: Evaluating stereotypical biases in NLP benchmarks. Ibrahim Al Kotob, Binesh Arakkal Remesh, Lucie Digoin-Caparro, Shayan Ahmed Shariff.
- Co-supervision of NLP master students' supervised projects** 2024-2025  
*Université de Lorraine*
- Co-supervision of two groups of 4 M1 students in NLP. Participation in meetings, scientific contribution and input.
  - Group 1: Racist bias detection in language models through code-switching. Léa Bouchet, Benjamin Caissial, Loïc Leclere, Arthur Pedrini.
  - Group 2: Evaluating stereotypical biases in NLP benchmarks. Ibrahim Al Kotob, Binesh Arakkal Remesh, Lucie Digoin-Caparro, Shayan Ahmed Shariff.
- Co-supervision of a bachelor's intern (3 months)** 2024-2025  
*LORIA*
- Co-supervision of Vedit Khazanchi for his work on racist bias evaluation through English-Hindi code-switching.
- Co-supervision of a bachelor's intern (2 months)** 2024-2025  
*LORIA*
- Co-supervision of Jeffrey André in his work on reproducing, extending and promotion of my gender bias framework.
- Co-supervision of a master's intern (6 months)** 2023-2024  
*LORIA*
- Co-supervision of Clémentine Bleuze who worked on "Analysing claims in NLP research". Discussions and weekly meetings with the intern, scientific contribution to the project..

## PUBLICATIONS

---

### Journals

**“You’ll be a nurse, my son!” Automatically Assessing Gender Biases in Autoregressive Language Models in French and Italian** Fanny Ducel, Aurélie Névéol, Karën Fort. Language Resources and Evaluation. October 2024. <https://link.springer.com/article/10.1007/s10579-024-09780-6>

Participation: Design and implementation of the experiment which includes an automatic gender detection for French and Italian, stereotypical bias metrics, writing the article.

**La recherche sur les biais dans les modèles de langue est biaisée : état de l’art en abyme.** Fanny Ducel, Aurélie Névéol, Karën Fort. Revue TAL, September 2024. [https://www.atala.org/content/tal\\_64\\_3\\_5](https://www.atala.org/content/tal_64_3_5)

Participation: Reading and manually annotating a hundred articles on stereotypical biases in Language Models, writing a literature review paper with meta-data analysis.

### International peer-reviewed conferences

**“Women do not have heart attacks !” Gender Biases in Automatically Generated Clinical Cases in French.** Fanny Ducel, Nicolas Hiebel, Olivier Ferret, Karën Fort, Aurélie Névéol. In Findings of the Association for Computational Linguistics: NAACL 2025, April 2025, Albuquerque, United States. <https://aclanthology.org/2025.findings-naacl.398/>

Participation: Design and implementation of an experiment to adapt French gender detection system, analysis of results for biomedical documents, writing of the article.

**Your Stereotypical Mileage May Vary: Practical Challenges of Evaluating Biases in Multiple Languages and Cultural Contexts.** Karën Fort, Laura Alonso Alemany, Luciana Benotti, Julien Bezançon, Claudia Borg, Marthese Borg, Yongjian Chen, Fanny Ducel, et al. Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024), May 2024, Turin, Italy. <https://aclanthology.org/2024.lrec-main.1545/>

Participation: Meetings and discussions with 22 researchers from 8 different countries. Corrections in the English and French corpora, writing of a section of the article, designing the poster.

**The Elephant in the Room: Analyzing the Presence of Big Tech in Natural Language Processing Research.** Mohamed Abdalla, Jan Philip Wahle, Terry Ruas, Aurélie Névéol, Fanny Ducel, Saif Mohammad, Karën Fort. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (ACL)*, July 2023, Toronto, Canada. - *CORE A\**. <https://aclanthology.org/2023.acl-long.734/>

Participation: Frequent meetings and discussions with 6 other scientists from 3 different countries involved in the project. Extraction, analysis and formatting of raw data by Python scripts, manual annotations, proofreading and correction of the article.

**Do we Name the Languages we Study? The #BenderRule in LREC and ACL articles.** Fanny Ducel, Karën Fort, Gaël Lejeune, Yves Lepage. LREC 2022 - International Conference on Language Resources and Evaluation (LREC), June 2022, Marseille, France. - *CORE C, QUALIS B2*. <https://aclanthology.org/2022.lrec-1.60.pdf>

Participation: Oral plenary presentation at LREC, manual annotation of NLP articles, constitution and pre-processing of a corpus of articles from ACL Anthology, implementation of automatic classifiers in Python, analysis and formatting of results, writing of the article with an ethical approach calling for diversity, realization of the slides for the presentation.

### National peer-reviewed conferences

**« Les femmes ne font pas de crise cardiaque ! » Étude des biais de genre dans les cas cliniques synthétiques en français.** Fanny Ducel, Nicolas Hiebel, Olivier Ferret, Karën Fort, Aurélie Névéol. Conférence sur le Traitement Automatique des Langues Naturelles (TALN), July 2025, Marseille, France. <https://aclanthology.org/2025.jeptalnrecital-trad.1/>

Participation: Translation of the NAACL 2025 paper’s abstract.

**« De nos jours, ce sont les résultats qui comptent » : création et étude diachronique d’un corpus de revendications issues d’articles de TAL.** Clémentine Bleuze, Fanny Ducel, Maxime Amblard, Karën Fort. Conférence sur le Traitement Automatique des Langues Naturelles (TALN), July 2025, Marseille, France. <https://aclanthology.org/2025.jeptalnrecital-taln.1/>

Participation: Took part in the writing process and manual annotation campaigns.

**La linguistique appliquée pour une IA plus éthique.** Fanny Ducel, Aurélie Névéol, Karën Fort. Naturel et Artificiel en Linguistique Appliquée : une époque de paradoxes (NÉALA), July 2025, Nancy, France. <https://neala2025.sciencesconf.org/586094/document>

Participation: Writing of a long abstract (position paper style), oral presentation.

**Évaluation automatique des biais de genre dans des modèles de langue auto-régressifs.** Fanny Ducel, Aurélie Névél, Karën Fort. Conférence sur le Traitement Automatique des Langues Naturelles (TALN), July 2024, Toulouse, France. <https://aclanthology.org/2024.jeptalnrecital-taln.5/>

Participation: Design and implementation of an experiment which includes an automatic gender detection for French and Italian, stereotypical bias metrics, writing and presenting the article at the conference.

**Langues « par défaut » ? Analyse contrastive et diachronique des langues non citées dans les articles de TALN et d'ACL.** Fanny Ducel, Karën Fort, Gaël Lejeune, Yves Lepage. Conférence sur le Traitement Automatique des Langues Naturelles (TALN), June 2022, Avignon, France. <https://aclanthology.org/2022.jeptalnrecital-taln.14.pdf>

Participation: This is a second part of the work presented at LREC, so the tasks are very similar, but on different data. Production and presentation of the poster at TALN and same tasks as for LREC.

## Tutorials

**Navigating Ethical Challenges in NLP: Hands-on strategies for students and researchers.** Luciana Benotti, Fanny Ducel, Karën Fort, Guido Ivetta, Min-Yen Kan and Minzhi Li. ACL, August 2025, Vienna, Austria. <https://aclanthology.org/2025.acl-tutorials.5/>

Participation: Writing sections in the tutorial proposal, updating the tutorial material and resources and presenting the tutorial.

## Workshops

**Introducing MascuLead: the First Gender Bias Leaderboard.** Fanny Ducel, Jeffrey André, Aurélie Névél, Karën Fort. Ethic and Alignment of (Large) Language Models (@ CORIA-TALN), July 2025, Marseille, France. <https://inria.hal.science/hal-05282981>.

Participation: Part of the demonstrator's website design process and leaderboards creations, writing and presentation of the article.

**« Un corpus qui fait tomber la barrière de la langue » : annotation de revendications dans des articles de TAL.** Clémentine Bleuze, Fanny Ducel, Karën Fort, Maxime Amblard. Journées de lancement LIFT 2, November 2024, Orléans, France. <https://inria.hal.science/hal-04880335>

Participation: Proofreading the paper, manual annotations.

**Desiderata for Actionable Bias Research** Fanny Ducel, Aurélie Névél and Karën Fort. New Perspectives on Bias and Discrimination in Language Technology, November 2024, Amsterdam, the Netherlands. <https://inria.hal.science/hal-04755691>

Participation: Writing a position paper about current issues in bias research and proposing solutions.

**What ChatGPT tells us about ourselves.** Aaron Boussidan, Fanny Ducel, Aurélie Névél and Karën Fort. Journée Éthique et TAL 2024, April 2024, Nancy, France. <https://inria.hal.science/hal-04521121>.

Participation: Designing an experiment on *ChatGPT*, writing a long abstract, creating the poster.

**Bias Identification in Language Models is Biased.** Fanny Ducel, Aurélie Névél, Karën Fort. Workshop on Algorithmic Injustice, June 2023, Amsterdam, the Netherlands. <https://inria.hal.science/hal-04171198>

Participation: Literature review on biases in language models, annotation of 60+ articles, meta-data extraction, writing of an extended abstract of a literature review for a workshop.

## Master's Theses

**Évaluation des biais stéréotypés dans les modèles de langue auto-régressifs : état de l'art et exploration de techniques extrinsèques** Fanny Ducel. [https://fannyducel.github.io/\\_pages/MemoireM2Fanny.pdf](https://fannyducel.github.io/_pages/MemoireM2Fanny.pdf)

**Analyse des *claims* dans les articles de Traitement Automatique des Langues à l'aide d'une méthode par apprentissage non supervisé.** Fanny Ducel. 2022. <https://tinyurl.com/2p976fxf>.

## INVITED TALKS

---

**« Les femmes ne font pas de crise cardiaque, les hommes ne peuvent pas être infirmiers » : des modèles de langue sexistes ?** Octobre 2025

AI Week / Fête de la Science – Université Champagne-Ardenne, Reims, France

<b>Comment définir, comprendre et évaluer les biais stéréotypés des modèles de langue ?</b> <i>GT Intelligence Artificielle Soutenable, Intelligible et Vérifiable du Plateau de Saclay – CEA, Palaiseau, France</i>	Mars 2025
<b>Les enjeux éthiques de l'IA par le prisme du TAL et des biais stéréotypés, avec Karèn Fort</b> Matinée thématique transversale « L'IA en terrain miné » – <i>ATILF, Nancy, France</i>	Juin 2024

## PEER-REVIEWING

---

- Reviewer for LREC 2026 (4 articles)
- Reviewer for the Workshop CHOMPS@AAACL-IJCNLP 2025 (1 article)
- Reviewer for Transactions on Asian and Low-Resource Language Information Processing (1 article)
- Reviewer (1 article) and ethics reviewer (4 articles) for EMNLP 2025
- Reviewer for TALN 2025 (2 articles)
- Reviewer (4 articles) and ethics reviewer (8 articles) pour ACL 2025
- Reviewer for TALN 2024 (1 article)
- Member of the program committee and reviewer for the ATALA Day Ethics and NLP 2024 (2 long abstracts)
- Primary and secondary reviewer for LREC-COLING 2024 (3 articles)
- Secondary reviewer for Undone Computer Science 2024 (1 article)
- Secondary reviewer for GdR-LIFT 2023 (2 long abstracts)
- Secondary reviewer for ACL 2023 (1 article)
- Secondary reviewer for TALN 2023 (1 article)

## SCIENCE POPULARISATION

---

<b>PhD Students Seminar</b> <i>LORIA, Nancy, France</i>	March 2025
<ul style="list-style-type: none"> <li>• <i>How to define, understand and evaluate stereotypical biases in language models?</i></li> </ul>	
<b>Journée Filles-Maths et Sciences (Association « Elles bougent ! »)</b> <i>Lycée Loritz, Nancy, France</i>	February 2025
<ul style="list-style-type: none"> <li>• Discussions about science and research with middle-school girls</li> </ul>	
<b>"CHICHE" Talks</b> <i>Lycée Varoquaux, Tomblaine, France / Lycée Antoine de Saint-Exupéry, Fameck, France</i>	December 2023 – January 2024
<ul style="list-style-type: none"> <li>• Presentations about Computer Science Research to groups of high-schoolers</li> </ul>	

## ORGANIZATION AND RESPONSIBILITIES

---

<b>Member of the Administration Council of the ATALA</b>	2025 – Today
<ul style="list-style-type: none"> <li>• In charge of Workshops.</li> </ul>	
<b>Student volunteer for the ACL Ethics Committee</b>	2024 – Today
<ul style="list-style-type: none"> <li>• Organization of the ACL 2025 tutorial, creation of webpages for the committee's website.</li> </ul>	
<b>Student volunteer at ACL 2023</b> <i>Toronto, Canada</i>	2023
<ul style="list-style-type: none"> <li>• Poster session, registration desk.</li> </ul>	
<b>Student volunteer at IWCS 2023</b> <i>IDMC, Nancy, France</i>	2023
<ul style="list-style-type: none"> <li>• In charge of coffee breaks.</li> </ul>	

## AWARDS

---

Best Master's Thesis Award from the French Applied Linguistics Organization (AFLA), 2024 edition.

## SKILLS

---

- **Languages:** Python, bash, L<sup>A</sup>T<sub>E</sub>X, SQL, HTML, CSS
- **NLP Tools:** TXM, Praat, Unitex, GREW, Antconc, GATE
- **Other tools:** Git, Neo4j, Gephi