

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.

Defense scheduled for Fall 2026.

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

AWARDS

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

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) June 2021 – August 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 (Ducel et al. 2023a, Ducel et al. 2023b)

TEACHING

All hours were taught at the IDMC, Université de Lorraine (and indicated as "heures équivalent TD") :

Year	Title	Audience	Nb. of students	Language	CM	TD
2024-2025	De l'écrit à l'information	L1	40	FR	-	4
2025-2026	Advanced algorithmics	L2	17	FR	-	14
2025-2026	IA et ingénierie de la langue	L3	5	FR	-	15
2025-2026	Project management tools	M1	45	EN	3	5
2024-2025	Python for beginners	M1	45/22	EN	21	16
2024-2025	Introduction to Grid 5000	M1	45	EN	3	2
2023-2024	Advanced Python programming	M1	22	FR	-	18
2024-2025	Ethics for NLP	M1-M2	22	EN	-	10
2025-2026	Software projects	M2	22	EN	-	10
2025-2026	Ethics for NLP	M2	22	EN	3	-
Total					30	93

Advanced algorithmics - TD 14h 2025-2026
L2 MIASH, IDMC, Université de Lorraine

- Chained lists, binary trees, n-ary trees. Course in French.

IA et Ingénierie de la Langue - TD 15h 2025-2026
L3 MIASH (TAL), IDMC, Université de Lorraine

- Introduction to NLTK, Spacy, n-grams, tokenisation, rule-based systems (such as ELIZA).
- Creation of materials. Course in French.

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 English.

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. Course in French.

Tutor in Computer Science for Language Science Students 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, ...

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. Led to publications at LREC and TALN (Ducel et al. 2026a, Ducel et al. 2026b).
- Group 2: Evaluating stereotypical biases in NLP benchmarks. Ibrahim Al Kotob, Binesh Arakkal Remesh, Lucie Digoin-Caparros, Shayan Ahmed Shariff. Led to a publication at TALN (Ducel et al. 2026c) and a submission (currently under review).

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-Caparros, Shayan Ahmed Shariff.

Co-supervision of a bachelor's intern (3 months) 2024-2025
LORIA

- Co-supervision of Vidit 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.

RESPONSIBILITIES

International responsibilities

Student member of the Ethics committee of the ACL (international NLP association) 2024 – Now

- Contributed to organizing a tutorial at ACL 2025
- Contributed to the creation of the committee's website: <https://ethics.aclweb.org>.
- Collaborations with the other student members : Guido Ivetta, Minzhi Li, Esther Gan and Punya Syon Pandey.

Student volunteer at ACL 2023 2023
Toronto, Canada

- Poster session, registration desk.

Student volunteer at IWCS 2023 2023
IDMC, Nancy, France

- In charge of coffee breaks.

National responsibilities

Member of the Administration Council of the ATALA (French NLP association)

2025 – Today

- In charge of Workshops (LLM@Hôpital, Langue des Signes Française).

PUBLICATIONS

Type	#	Comments
International, peer-reviewed journals	1	1 Q1
National, peer-reviewed journals	1	1 Q4
International, peer-reviewed conferences	6	1 Core A*, 1 Core A, 4 Core C
National, peer-reviewed conferences	9	incl. 7 TALN, 1 LIFT
Invited talks	3	
Tutorials	1	
Workshops	4	
Total	25	

Submitted (3)

Benchmarks are a source of bias in LLMs: MMLU, CommonSenseQA and MGSM under the Microscope. Fanny Ducel, Shayan Ahmed Shariff, Ibrahim Al Kotob, Binesh Arakkal Remesh, Lucie Digoïn-Caparros, Aurélie Névéol, Karën Fort. EMNLP. Core A*. 2026.

Participation: Conception et réalisation d'expériences et d'annotations manuelles sur les *benchmarks* des modèles de langue. Analyses des résultats, rédaction de l'article.

Prompting LLMs with gender-inclusive language: a win-win strategy. Fanny Ducel, Sayaka Sato, Lucie Escasain, Aurélie Névéol, Karën Fort, Pascal Gygas. PLOS ONE. Revue SJR "Q1". 2026.

Participation: Conception et réalisation d'une expérience de psycholinguistique sur l'impact de l'écriture inclusive dans les invites des modèles de langue. Analyses des résultats, rédaction de l'article.

The Umbrella Coup: Alignment, LLMs, and the Misdirection of Ethics Discourse in NLP. Fanny Ducel, Clémentine Bleuze, Aurélie Névéol, Emily M. Bender, Karën Fort. Computational Linguistics. Revue SJR "Q1". 2026.

Participation: Discussions sur la notion d'alignement, rédaction de sections de l'article.

International, peer-reviewed journals (1)

"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. Revue SJR "Q1". 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.

National, peer-reviewed journals (1)

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. Revue nationale du TAL, SJR "Q4". September 2024. 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 (6)

Code-Switching as a Bias Indicator in LLMs: "The consequences are not the same para nosotros" Fanny Ducel, Aurélie Névéol, Vidit Khazanchi, Loïc Leclere, Arthur Pedrini, Léa Bouchet, Benjamin Caissial, Karën Fort. LREC 2026 - International Conference on Language Resources and Evaluation (LREC). Core C, Qualis B2. May 2026, Palma, Spain. URL to come.

Participation: Design and implementation of an experiment on the impact of code-switching in LLMs prompts, results analyses, writing of the article. Oral presentation of the paper.

COCOA: Creation and Exploratory Investigation of a CORpus of Claims frOm NLP Articles. Clémentine Bleuze, Fanny Ducel, Maxime Amblard, Karën Fort. LREC 2026 - International Conference on Language Resources and Evaluation (LREC). Core C, Qualis B2. May 2026, Palma, Spain. URL to come.

Participation: Adaptation and augmentation of the TALN 2025 paper.

“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. *CORE A*. 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. Oral (online) presentation of the paper.

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). *CORE C, QUALIS B2*. 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)*. *CORE A**. July 2023, Toronto, Canada.. <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). *CORE C, QUALIS B2*. June 2022, Marseille, France. <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 (9)

Les benchmarks sont une source de biais des LLM : MMLU, CommonSenseQA et MGSM au microscope. Fanny DuceL, Lucie Digoin-Caparras, Ibrahim Al Kotob, Shayan Ahmed Shariff, Binesh Arakkal Remesh, Aurélie Névéol, Karën Fort. Conférence sur le Traitement Automatique des Langues Naturelles (TALN), June 2026, Nantes, France. URL to come.

Participation: Adaptation and translation of a submitted article. Oral plenary presentation.

L’écriture inclusive dans les invites des modèles de langue : une stratégie gagnant-gagnant. Fanny DuceL, Sayaka Sato, Lucie Escasain, Aurélie Névéol, Karën Fort, Pascal Gyax. Conférence sur le Traitement Automatique des Langues Naturelles (TALN), June 2026, Nantes, France. URL to come.

Participation: Adaptation and translation of a submitted article. Oral plenary presentation.

Le code-switching comme indicateur de biais dans les LLM : "The consequences are not the same para nosotros". Fanny DuceL, Aurélie Névéol, Vidit Khazanchi, Loïc Leclere, Arthur Pedrini, Léa Bouchet, Benjamin Caissial, Karën Fort. Conférence sur le Traitement Automatique des Langues Naturelles (TALN), June 2026, Nantes, France. URL à venir.

Participation: Adaptation and translation of a submitted article. Oral plenary presentation.

« 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.

« **Vers la création d'une super-intelligence** » : **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.

Évaluation automatique des biais de genre dans des modèles de langue auto-régressifs. **Fanny Ducel**, Aurélie Névéol, 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.

Invited talks (3)

« **Les femmes ne font pas de crise cardiaque, les hommes ne peuvent pas être infirmiers** » : **des modèles de langue sexistes ?**. Semaine de l'Intelligence Artificielle / Fête de la Science – *Université Champagne-Ardenne, Reims, France*. October 2025.

Comment définir, comprendre et évaluer les biais stéréotypés des modèles de langue ?. *GT Intelligence Artificielle Soutenable, Intelligible et Vérifiable* du Plateau de Saclay – *CEA, Palaiseau, France*. March 2025.

Les enjeux éthiques de l'IA par le prisme du TAL et des biais stéréotypés, avec Karën Fort. Matinée thématique transversale « L'IA en terrain miné » – *ATILF, Nancy, France*. June 2024.

Tutorials (1)

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 (4)

Introducing MascuLead: the First Gender Bias Leaderboard. **Fanny Ducel**, Jeffrey André, Aurélie Névéol, 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.

Desiderata for Actionable Bias Research **Fanny Ducel**, Aurélie Névéol 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 Névéol 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 Névéol, 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

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>.

SCIENCE POPULARISATION

PhD Students Seminar

March 2025

LORIA, Nancy, France

- How to define, understand and evaluate stereotypical biases in language models?

Journée Filles-Maths et Sciences (Association « Elles bougent ! »)

February 2025

Lycée Loritz, Nancy, France

- Discussions about science and research with middle-school girls

"CHICHE" Talks

December 2023 – January 2024

Lycée Varoquaux, Tomblaine, France / Lycée Antoine de Saint-Exupéry, Fameck, France

- Presentations about Computer Science Research to groups of high-schoolers

PEER-REVIEWING

- Reviewer for EMNLP 2026 (3 articles)
- Ethics reviewer for ACL 2026 (2 articles)
- Reviewer for FaCCT 2026 (3 articles)
- 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)

COMPUTER SCIENCE SKILLS

- **Languages:** Python, bash, L^AT_EX, SQL, HTML, CSS
- **NLP tools:** TXM, Praat, Unitex, GREW, Antconc, GATE
- **Other:** Git, Neo4j, Gephi

LANGUAGES

French (native speaker), English (C1-C2), German (B1)