

Fanny Ducel

<https://fannyducel.github.io>
fanny.ducel@lisn.upsaclay.fr

Education

PhD in Computer Science

2023 –

Université Paris-Saclay

Gif-sur-Yvette, France

- Evaluation of stereotypical biases in autoregressive language models.
- Under the supervision of Aurélie Névéol and Karĕn Fort.

Master in Linguistics and Computer Science, with highest honours

2021 – 2023

Sorbonne Université

Paris, France

- M2: Valedictorian. Annual grade : 17.9/20. Thesis on the Evaluation of stereotypical biases in auto-regressive language models.
- M1: Valedictorian. Annual grade : 17.35/20. 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 - 35h/week)

March 2023 - September 2023

LORLA, supervised by Karĕn Fort and Aurélie Névéol

Nancy, France

- Achieved 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 CDD - 17h/week)

September 2022

CERES, Sorbonne Université, supervised by Gaĕl Lejeune

Paris, France

- 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 - 35h/week)

July 2022 - September 2022

LORLA, supervised by Karĕn Fort and Maxime Amblard

Nancy, France

- Establishing a state of the art 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 - 35h/week)

June 2021 - August 2021

LORLA, supervised by Karĕn Fort, Gaĕl Lejeune and Yves Lepage

Nancy, France

- 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

Lecturer for a Python Programming class (Master's level) - 13h September 2024 - November 2024

Université de Lorraine

Nancy, France

- * Introduction to Python for beginners in NLP/Cognitive Science masters: data structures, functions, libraries, introduction to object-oriented. In English.

Tutor for Python Programming labs (Master's level) - 16h September 2024 - November 2024

Université de Lorraine

Nancy, France

- * Introducing labs, helping out students to complete them and correcting their assignments. Beginner group. In English.

Introduction to Grid5000 (4h)

September 2024

Université de Lorraine

Nancy, France

- * First connection, key notions (site/cluster/job), essential commands (file transfer, resource reservation)

Tutor for Python Programming labs (Master's level) - 18h September 2023 - December 2023

Université de Lorraine

Nancy, France

- * Introducing labs, helping out students to complete them and correcting their assignments. Advanced group.

Tutor in Computer Science for Language Science Students

September 2021 - March 2023

Sorbonne Université

Paris, France

- * 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, ...

Other Contributions

Peer-reviewing

- * Secondary 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)

Participation in conferences organization

- Student volunteer at ACL 2023 (registration desk, poster session)
- Student volunteer at IWCS 2023 (in charge of coffee breaks)

Science popularisation

December 2023 / January 2024

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

- "CHICHE" talks

- 1h-1h30 oral presentations to three groups of high-schoolers / four groups of high-schoolers

Invited talks

June 2024

ATILF, Nancy, France

- Jointly with Karën Fort, on ethics and stereotypical biases in NLP, during a half-day workshop on AI.

Invited talks

June 2024

ATILF, Nancy, France

- Jointly with Karën Fort, on ethics and stereotypical biases in NLP, during a half-day workshop on AI.

Co-supervision of a Master's intern

March 2024 - August 2024

LORIA, Nancy, France

- Discussions and weekly meetings with the intern, scientific contribution to the project.

Member of the editorial board of the "Éthique et TAL" blog

- Blog post authoring and reviewer for blog posts.

Publications

International peer-reviewed conferences

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), Mai 2024, Turin, Italie. <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 Névél, **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://arxiv.org/pdf/2305.02797.pdf>

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

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), Juillet 2024, Toulouse, France. <https://hal.science/hal-04621134v1>

Participation : Designing and implementing the experience 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.

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, to be published (accepted article).

Participation : Designing and implementing the experience 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

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

Workshops

Desiderata for Actionable Bias Research Fanny Ducel, Aurélie Névéol and Karën Fort. New Perspectives on Bias and Discrimination in Language Technology, November 2024, Amsterdam, Netherlands. To be published (accepted).

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, Netherlands.

Participation : Established a state of the art on biases in language models, annotated above 60 articles, extracted meta-data, wrote the extended abstract of a literature review for a workshop.

Master 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. <https://tinyurl.com/2p976fx>

Specialized Skills

Languages: Python, SQL, SPARQL, Voice XML, UML, HTML, CSS, Java, C++, \LaTeX , Markdown

NLP tools: TXM, Antconc, Unitex, Praat, GREW-match, GATE, HTK, MBROLA, espeak, audeep

Other tools: Neo4j, GitHub, Apache Solr, GANT, Agile (Scrum), Gephi, Twitter API

Python libraries: Scikit-learn, SpaCy, NLTK, CoreNLP, Gensim, Pandas, LangChain, Pytorch

Language Skills

French (Native), **English** (C1), **German** (B1), **Spanish** (A0-A1), **Hindi** (A0)

Other Interests

Research interests: NLP, Ethics, Biases, Linguistic diversity, Sociolinguistics, Sociology, Psychology, Philosophy.