

Fanny DUCEL

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

FORMATION

Doctorat en Informatique

2023 – Aujourd’hui

LISN, Université Paris-Saclay, Orsay, France

Intitulé de la thèse : Mesurer les biais stéréotypés dans les modèles de langue auto-régressifs.

Sous la direction d’Aurélie Névéol et Karën Fort.

Master Langue et Informatique – mention Très Bien

2021 – 2023

Sorbonne Université, Paris, France

- Mémoire master 2 : Évaluation des biais stéréotypés dans les modèles de langue auto-régressifs : état de l’art et exploration de techniques extrinsèques.
- Mémoire master 1 : Analyse des *claims* dans les articles de traitement automatique des langues à l’aide d’une méthode par apprentissage non supervisé.

Licence Sciences du Langage, spécialité Informatique – mention Très Bien

2018 – 2021

Sorbonne Université, Paris, France

EXPÉRIENCES EN RECHERCHE

Stage de recherche (M2)

Mars 2023 – Septembre 2023

LORIA, Nancy, France – encadrée par Karën Fort et Aurélie Névéol

- Établissement d’un état de l’art sur les biais stéréotypés dans les modèles de langues auto-régressifs avec une approche critique
- Expériences de manipulation de modèles de langues masqués et auto-régressifs pour l’évaluation des biais, reproduction et adaptation d’expériences de limitation et quantification de biais
- Participation au projet de recherche international MultiCrowsPairs, constitué de 20 autres chercheur·euses et étudiant·es de 8 pays différents

Assistante de Recherche (temps partiel)

Septembre 2022

CERES, Sorbonne Université, Paris, France – encadrée par Gaël Lejeune

- Participation à un projet de cartographie de chercheurs et chercheuses selon leurs publications : réalisation de scripts Python pour tester de nouvelles techniques de vectorisation et clustering, création de graphes avec Gephi, réorganisation du répertoire Github avec documentation.

Stage de recherche (M1)

Juillet 2022 – Septembre 2022

LORIA, Nancy, France – encadrée par Karën Fort et Maxime Amblard

- Réalisation d’un état de l’art sur les notions linguistiques de *hedging* et *claims*
- Constitution et pré-traitement d’un corpus d’articles ACL
- Extraction de *claims* en corpus, utilisation de méthodes de *clustering* en Python
- Évaluation de méthodes d’apprentissage non-supervisé

Stage de recherche (L3)

Juin 2021 – Août 2021

LORIA, Nancy, France – encadrée par Karën Fort, Gaël Lejeune et Yves Lepage

- Annotation manuelle d’articles de TAL (TALN, LREC, ACL)
- Classification automatique de ces articles selon la présence de mention de la langue étudiée
- Travaux acceptés par la suite à LREC et TALN

ENSEIGNEMENT

Outils pour la gestion de projet - CM 2h, TD 5h

2025-2026

M1 TAL, IDMC, Université de Lorraine

- Introduction à Grid 5000 : première connexion, notions clés (site/cluster/job), commandes essentielles (transfert de fichiers, réservation de ressources).
- Introduction à L^AT_EX : commandes essentielles, réalisation d’un premier document.
- Réalisation des supports.
- Cours en anglais.

Projets logiciels - TD 10h

2025-2026

M2 TAL, IDMC, Université de Lorraine

- Suivi et évaluation de projets de différents groupes.
- Cours en anglais.

Éthique du TAL - CM 2h 2025-2026

M2 TAL, IDMC, Université de Lorraine

- Présentation de mes travaux de recherche.
- Cours en anglais.

Python pour débutant·es - TD 16h 2025-2026

M1 SC, IDMC, Université de Lorraine

- Explications, aide et correction des TD d'étudiant·es.

De l'Écrit à l'Information - TD 4h 2024-2025

L1 MIASH, IDMC, Université de Lorraine

- Introduction à la linguistique.

Éthique du TAL - TD 10h 2024-2025

M1/M2 TAL, IDMC, Université de Lorraine

- Introduction de notions sur l'éthique, accompagnement de travaux de groupes, évaluation de présentations orales en Master 1 et Master 2 de Traitement Automatique des Langues.
- Cours en anglais.

Python pour débutant·es - CM 14h (21h HETD) 2024-2025

M1 TAL/SC, IDMC, Université de Lorraine

- Introduction à Python : structures de données, fonctions, librairies, introduction à l'orienté objet.
- Réalisation d'une partie des supports.
- Cours en anglais.

Python pour débutant·es - TD 16h 2024-2025

M1 TAL/SC, IDMC, Université de Lorraine

- Explications, aide et correction des TD d'étudiant·es.
- Réalisation d'une partie des supports.
- Cours en anglais.

Introduction à Grid 5000 - CM 2h, TD 2h 2024-2025

M1 TAL, IDMC, Université de Lorraine

- Introduction à Grid5000 : première connexion, notions clés (site/cluster/job), commandes essentielles (transfert de fichiers, réservation de ressources).
- Réalisation des supports.
- Cours en anglais.

Programmation Python avancée - TD 18h 2023-2024

M1 TAL/SC, IDMC, Université de Lorraine

- Explications, aide et correction des TD d'étudiant·es.

Tutorat en Informatique 2021-2023

Sorbonne Université

- Permanence de 3 heures par semaine auprès d'étudiant·es de la Licence Sciences du Sciences du Langage pour aider à la réalisation des TD et projets de programmation Python, développement Web (HTML et CSS), ingénierie de la langue (TXM, Antconc, GATE), logique du premier ordre, ...

ENCADREMENT

Co-encadrement de stagiaires de M1 TAL (2 mois) 2024-2025
Université de Lorraine

- Co-encadrement de stagiaires de M1 TAL, dans la continuité de leurs projets tutorés.
- Groupe 1 : Détection de biais racistes des modèles de langue par l'analyse du code-switching. Loïc Leclerc, Arthur Pedrini.
- Groupe 2 : Évaluation des biais stéréotypés présents dans les benchmarks de TAL. Ibrahim Al Kotob, Binesh Arakkal Remesh, Lucie Digoin-Caparros, Shayan Ahmed Shariff.

Co-encadrement de groupes de projets tutorés de M1 TAL
Université de Lorraine

2024-2025

- Co-encadrement de deux groupes de 4 étudiant·es en M1 TAL dans leurs projets tutorés. Suivi de groupes, participation aux réunions et contribution scientifique.
- Groupe 1 : Détection de biais racistes des modèles de langue par l'analyse du code-switching. Léa Bouchet, Benjamin Caissial, Loïc Leclerc, Arthur Pedrini.
- Groupe 2 : Évaluation des biais stéréotypés présents dans les benchmarks de TAL. Ibrahim Al Kotob, Binesh Arakkal Remesh, Lucie Digoin-Caparros, Shayan Ahmed Shariff.

Co-encadrement d'un stagiaire de L3 Informatique (3 mois)
LORIA

2024-2025

- Co-encadrement de Vudit Khazanchi dans son travail sur l'évaluation de biais racistes à l'aide de *code-switching* anglais-hindi.

Co-encadrement d'un stagiaire de L3 TAL (2 mois)
LORIA

2024-2025

- Co-encadrement de Jeffrey André dans son travail sur la reproduction, l'extension et la valorisation de mon outil d'évaluation de biais de genre en français.

Co-encadrement d'une stagiaire de M2 TAL (6 mois)
LORIA

2023-2024

- Co-encadrement de Clémentine Bleuze dans son travail sur l'Analyse de claims dans la recherche en TAL. Suivi de l'étudiante, participation aux réunions hebdomadaires et contribution scientifique à son projet.

PUBLICATIONS

Revues scientifiques

"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*. Octobre 2024. <https://link.springer.com/article/10.1007/s10579-024-09780-6>

Participation : Conception et réalisation de l'expérience avec création d'un système de détection automatique du genre en français et en italien, mesure de biais stéréotypés genrés, rédaction de l'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*, Septembre 2024. https://www.atala.org/content/tal_64_3_5

Participation : Lecture et annotation manuelle d'une centaine d'articles sur les biais stéréotypés dans les modèles de langue, rédaction d'un état de l'art avec analyses de méta-données.

Conférences internationales avec comité de lecture

"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*, Avril 2025, Albuquerque, États-Unis. <https://aclanthology.org/2025.findings-naacl.398/>

Participation : Élaboration et implémentation de l'expérience avec adaptation d'un système de détection automatique du genre en français, analyses de résultats pour le domaine médical, rédaction de l'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 Bezanson, Claudia Borg, Marthense 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 : Réunions et discussions avec 22 scientifiques de 8 pays différents. Correction des corpus anglais et français, rédaction d'une section de l'article, réalisation du 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)*, Juillet 2023, Toronto, Canada. - CORE A*. <https://aclanthology.org/2023.acl-long.734/>

Participation : Réunions et discussions fréquentes avec 6 autres scientifiques de 3 pays différents prenant part au projet. Extraction, analyse et mise en forme de données brutes par scripts Python, annotations manuelles, relecture et correction de l'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), Juin 2022, Marseille, France. - *CORE C, QUALIS B2*. <https://aclanthology.org/2022.lrec-1.60.pdf>

Participation : Présentation orale plénière à LREC, annotation manuelle d'articles de TAL, constitution et pré-traitement d'un corpus d'articles à partir de l'ACL Anthology, implémentation de classificateurs automatiques en Python, analyses et mise en forme des résultats, rédaction de l'article avec une approche éthique appelant à la diversité, réalisation des transparents pour la présentation.

Conférences nationales avec comité de lecture

« 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), Juillet 2025, Marseille, France. <https://aclanthology.org/2025.jeptalnrecital-trad.1/>.

Participation : Traduction du résumé de la publication NAACL 2025.

« 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), Juillet 2025, Marseille, France. <https://aclanthology.org/2025.jeptalnrecital-taln.1/>.

Participation : Participation à la rédaction et aux annotations manuelles.

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), Juillet 2025, Nancy, France. <https://neala2025.sciencesconf.org/586094/document>

Participation : Rédaction d'un résumé long (de type papier de positionnement), présentation orale de l'article.

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), Juillet 2024, Toulouse, France. <https://aclanthology.org/2024.jeptalnrecital-taln.5/>

Participation : Conception et réalisation de l'expérience avec création d'un système de détection automatique du genre en français et en italien, mesure de biais stéréotypés genrés, rédaction et présentation orale de l'article.

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), Juin 2022, Avignon, France. <https://aclanthology.org/2022.jeptalnrecital-taln.14.pdf>

Participation : Il s'agit d'un deuxième volet du travail présenté à LREC, les tâches sont donc très similaires, mais sur des données différentes. Réalisation et présentation du poster à TALN, et mêmes tâches que pour le travail LREC. Obtention d'une bourse ATALA pour participer à la conférence.

Tutoriels

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, Août 2025, Vienne, Autriche. <https://aclanthology.org/2025.acl-tutorials.5/>

Participation : Rédaction de sections dans la proposition de tutoriel, mise à jour des ressources et animation du tutoriel.

Ateliers

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

Participation : Aide à la conception d'un démonstrateur en ligne et de *leaderboards*, rédaction et présentation de l'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, Novembre 2024, Orléans, France. <https://inria.hal.science/hal-04880335>

Participation : Aide à la correction du résumé long, annotations manuelles.

Desiderata for Actionable Bias Research Fanny Ducel, Aurélie Névéol and Karën Fort. New Perspectives on Bias and Discrimination in Language Technology, Novembre 2024, Amsterdam, Pays-Bas. <https://inria.hal.science/hal-04755691>

Participation : Rédaction d'un papier de positionnement sur les obstacles rencontrés dans la recherche sur les biais et propositions de solutions.

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

Participation : Participation à la mise en place d'une expérience sur *ChatGPT*, rédaction du résumé long et réalisation du poster.

Bias Identification in Language Models is Biased. Fanny Ducel, Aurélie Névéol, Karën Fort. Workshop on Algorithmic Injustice, Juin 2023, Amsterdam, Pays-Bas. <https://inria.hal.science/hal-04171198>

Participation : Réalisation d'un état de l'art sur les biais dans les modèles de langues avec annotation d'une soixantaine d'articles, extraction de métadonnées et rédaction d'un résumé étendu critique de cet état de l'art.

Mémoires

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>.

SÉMINAIRES INVITÉS

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

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

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

PARTICIPATION À DES COMITÉS DE LECTURE

- Relectrice pour LREC 2026 (4 articles)
- Relectrice pour l'atelier CHOMPS@ACL-IJCNLP 2025 (1 article)
- Relectrice pour Transactions on Asian and Low-Resource Language Information Processing (1 article)
- Relectrice (1 article) et relectrice éthique (4 articles) pour EMNLP 2025
- Relectrice pour TALN 2025 (2 articles)
- Relectrice (4 articles) et relectrice éthique (8 articles) pour ACL 2025
- Relectrice pour TALN 2024 (1 article)
- Membre du comité de programme et relectrice pour la journée Éthique et TAL 2024 (2 résumés)
- Relectrice de premier et deuxième niveaux pour LREC-COLING 2024 (3 articles)
- Relectrice de deuxième niveau pour Undone Computer Science 2024 (1 article)
- Relectrice pour le GdR-LIFT 2023 (2 résumés longs)
- Relectrice de deuxième niveau pour ACL 2023 (1 article)
- Relectrice de deuxième niveau pour TALN 2023 (1 article)

MÉDIATION

Séminaire doctorant·es Mars 2025
LORIA, Nancy, France

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

Journée Filles-Maths et Sciences (Association « Elles bougent ! ») Février 2025
Lycée Loritz, Nancy, France

- Matinée de tables rondes avec des collégiennes de 4e

Intervention CHICHE

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

Décembre 2023 – Janvier 2024

- Sensibilisation aux sciences du numérique, à la recherche en informatique
- Présentations à 7 groupes de seconde générale

ORGANISATION ET RESPONSABILITÉS

Membre du Conseil d'Administration de l'ATALA

2025 – Aujourd’hui

- Co-chargée des Journées d’Étude.

Étudiante volontaire au Comité d'éthique d'ACL

2024 – Aujourd’hui

- Participation à l’organisation et la tenue d’un tutoriel à ACL 2025, participation à la création du site web du comité.

Étudiante volontaire à ACL 2023

2023

Toronto, Canada

- Gestion de session de posters, gestion du bureau d’enregistrement.

Membre du comité d'organisation de la conférence IWCS 2023

2023

IDMC, Nancy, France

- Gestion des pauses café

PRIX ET DISTINCTIONS

Prix du meilleur mémoire de Master de l’Association Française de Linguistique Appliquée (édition 2024).

COMPÉTENCES INFORMATIQUES

- **Langages** : Python, bash, L^AT_EX, SQL, HTML, CSS
- **Outils de TAL** : TXM, Praat, Unitex, GREW, Antconc, GATE
- **Autres** : Git, Neo4j, Gephi