

Fanny Ducl

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

Formation

Doctorat en Informatique

2023 –

Université Paris-Saclay

Gif-sur-Yvette, France

- Évaluation des biais stéréotypés dans les modèles de langue autorégressifs.
- Dirigée par Aurélie Névéol et Karĕn Fort.

Master Langue et Informatique

2021 – 2023

Sorbonne Université

Paris, France

- M2 : Major de promotion. Moyenne annuelle de 17,9/20. Mémoire sur l'Évaluation des biais stéréotypés dans les modèles de langues auto-régressifs : état de l'art et exploration de techniques extrinsèques
- M1 : Major de promotion. Moyenne annuelle de 17,35/20. Mémoire sur l'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 - 35h/semaine)

Mars 2023 - Septembre 2023

LORLA, encadrée par Karĕn Fort et Aurélie Névéol

Nancy, France

- 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-es et étudiant-es de 8 pays différents

Assistante de Recherche (CDD à temps partiel - 17h/semaine)

Septembre 2022

CERES, Sorbonne Université, supervisée par Gaël Lejeune

Paris, France

- 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 - 35h/semaine)

Juillet 2022 - Septembre 2022

LORLA, encadrée par Karĕn Fort et Maxime Amblard

Nancy, France

- 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 - 35h/semaine)

Juin 2021 - Août 2021

LORLA, encadrée par Karĕn Fort, Gaël Lejeune et Yves Lepage

Nancy, France

- 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

Cours magistral de Python débutants (14h)

Septembre 2024 - Novembre 2024

Université de Lorraine

Nancy, France

- Introduction à Python pour débutants de M1 TAL/SC : structures de données, fonctions, bibliothèques, introduction à l'orienté objet. Cours en anglais.

Chargée de TD de programmation Python débutants (16h)

Septembre 2024 - Novembre 2024

Université de Lorraine

Nancy, France

- Explications, aide et correction des TD d'étudiant-es en Master 1 de Traitement Automatique des Langues ou Master 1 de Sciences Cognitives. Les notions de programmation abordées sont génériques. Cours en anglais.

Introduction à Grid5000 (4h)

Septembre 2024

Université de Lorraine

Nancy, France

- Introduction à Grid5000 : première connexion, notions clés (site/cluster/job), commandes essentielles (transfert de fichiers, réservation de ressources)

Chargée de TD de programmation Python avancée (18h)

Septembre 2023 - Décembre 2023

Université de Lorraine

Nancy, France

- Explications, aide et correction des TD d'étudiant-es en Master 1 de Traitement Automatique des Langues ou Master 1 de Sciences Cognitives. Les notions de programmation abordées sont génériques.

Tutrice en informatique pour la Licence Sciences du Langage

Septembre 2021 - Mars 2023

Sorbonne Université

Paris, France

- Permanence de 3 heures par semaine auprès d'étudiant-es de la Licence 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, ...

Autres contributions dans la communauté scientifique

Participation à des comités de lecture

- 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)

Participation à l'organisation de conférences

- Étudiante volontaire à ACL 2023 (gestion de session de posters, gestion du bureau d'enregistrement)
- Étudiante volontaire à IWCS 2023 (gestion des pauses cafés)

Médiation

Décembre 2023 / Janvier 2024

Lycée Varoquaux, Tomblaine, France / Lycée Fameck, France

- Interventions CHICHE (sensibilisation aux sciences du numérique, à la recherche en informatique)
- Présentations de 1h-1h30 à trois groupes de seconde générale / à quatre groupes de seconde générale

Présentations invitées

- *Les enjeux éthiques de l'IA par le prisme du TAL et des biais stéréotypés*, avec Karën Fort, dans le cadre d'une matinée thématique transversale de l'ATILF *L'IA en terrain miné*.

Prix et distinctions

- Prix du meilleur mémoire de M2, édition 2024, de l'Association Française de Linguistique Appliquée.

Co-encadrement d'une stagiaire de M2

Mars 2024 / Août 2024

LORIA, Nancy, France

- 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

Membre du Comité de Rédaction du blog Éthique et TAL

- Autrice d'articles sur le blog, participation aux relectures et à l'acceptation d'articles.

Participation au Comité d'Éthique d'ACL

Septembre 2024 -

- Participation en tant qu'étudiante volontaire dans le comité d'éthique, aide à la préparation et à la rédaction d'une proposition de tutoriel, aide à la création d'un site web dédié au comité d'éthique.

Publications

Conférences internationales avec comité de lecture

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 : 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

classifieurs 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

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://hal.science/hal-04621134v1>

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.

Revue scientifique

"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*, à paraître (article accepté).

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.

Workshops

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. À paraître.

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.

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 méta-données 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 Ducl. 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 Ducl. 2022. <https://tinyurl.com/2p976fxf>.

Compétences informatiques

Langages : Python, SQL, SPARQL, Voice XML, UML, HTML, CSS, Java, C++, L^AT_EX, Markdown

Outils de TAL : TXM, Antconc, Unitex, Praat, GREW-match, GATE, HTK, MBROLA, espeak, audeep

Autres outils : Neo4j, GitHub, Apache Solr, GANT, Agile (Scrum), Gephi, Twitter API

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