



Fanny Jourdan

[fanny-jourdan](#) | [Fanny Jourdan](#) | [fanny-jourdan](#) | fanny.jourdan@irit.fr | [@Fannyjrd](#)

SUMMARY

I'm PhD Candidate in computer science and mathematics at [IRIT](#) (Institut de Recherche en Informatique de Toulouse) and [ANITI](#) (Artificial and Natural Intelligence Toulouse Institute). I am interested in Explainability, Interpretability and Fairness of Natural Language Processing algorithms. I am looking for a research position (academic or private sector) in these topics for October 2024.

EDUCATION

PhD on Computer Science at Paul Sabatier University <i>Fairness and Explainability on Natural Language Processing</i> <i>Supervisors: Nicholas Asher and Laurent Risser</i>	2021 - 2024 Toulouse, France
Master's Degree on Economics at Paris Dauphine University <i>Digital economics, management, innovation, regulation theory</i>	2020 - 2021 Paris, France
Master's Degree on Mathematics at École polytechnique <i>Data Sciences, Statistical theory, Natural Language Processing, Optimisation, Causal Inference</i>	2018 - 2020 Paris, France
Bachelor's Degree on Mathematics at Paris-Sud University <i>Fundamental mathematics: Algebra, Topology, Calculus</i> <i>Holomorphic function, Probability</i>	2015 - 2018 Paris, France

WORK EXPERIENCE

Research intern at Criteo AI Lab <i>Study of the causal effect of displays (advertising banners sent by Criteo) on the sales of associated products.</i>	May 2020 - September 2020 Paris, France
Data scientist intern at EDF R&D <i>Analysis and disaggregation of individual electrical consumption time series.</i>	April 2019 - August 2019 Paris, France

PUBLICATIONS

1. **Fanny Jourdan**, Louis Bethune, Agustin Picard, Laurent Risser, Nicholas Asher, "TaCo: Targeted Concept Removal in Output Embeddings for NLP via Information Theory and Explainability" preprint.
2. **Fanny Jourdan**, Agustin Picard, Thomas Fel, Laurent Risser, Jean-Michel Loubes, Nicholas Asher, "COCKATIEL: COntinuous Concept ranKed ATtribution with Interpretable ELEments for explaining neural net classifiers on NLP tasks" to appear in Proceedings of *the Findings of the Association for Computational Linguistics (ACL 2023)*, Toronto, Canada, 2023, ACL.
3. **Fanny Jourdan**, Laurent Risser, Jean-Michel Loubes, Nicholas Asher, "Are fairness metric scores enough to assess discrimination biases in machine learning?" to appear in Proceedings of *Third Workshop on Trustworthy Natural Language Processing (TrustNLP ACL2023)*, Toronto, Canada, 2023, ACL.
4. **Fanny Jourdan**, Titon Tshiongo Kaninku, Nicholas Asher, Jean-Michel Loubes, Laurent Risser, "How Optimal Transport Can Tackle Gender Biases in Multi-Class Neural Network Classifiers for Job Recommendations" in *Algorithms*, 16.3, p. 174

TEACHING

Teaching assistant at Paul Sabatier University <i>Mathematics for Undergraduate students on Biology Bachelor</i>	2021 - 2023 Toulouse, France
Teaching assistant at Paul Sabatier University <i>Logic for Undergraduate students on Computer Science Bachelor</i>	2021 - 2022 Toulouse, France
Oral examiner at Parc de Vilgenis High School <i>"Colles" on mathematics for Undergraduate students in "Higher school preparatory classes"</i>	2018 - 2020 Massy, France

TALKS AND POSTERS PRESENTATIONS

IA Pau Festival - Keynote Speaker <i>Understanding and reducing discriminative biases in Natural Language Processing algorithms.</i>	December 1, 2023 Pau, France
NLTP group, Utrecht university - Invited speaker <i>COCKATIEL: COntinuous Concept ranKed ATtribution with Interpretable ELEments for explaining neural net classifiers on NLP tasks</i>	October 16, 2023 Utrecht, Netherlands
TrustNLP Workshop ACL - Oral <i>Are fairness metric scores enough to assess discrimination biases in machine learning?</i>	July 14, 2023 Toronto, Canada
TrustNLP Workshop ACL - Poster session (2 posters) - <i>Are fairness metric scores enough to assess discrimination biases in machine learning?</i> - <i>COCKATIEL: COntinuous Concept ranKed ATtribution with Interpretable ELEments for explaining neural net classifiers on NLP tasks</i>	July 14, 2023 Toronto, Canada

ACL - Findings spotlight
COCKATIEL: Continuous Concept ranKed ATtribution with Interpretable ELeMents for explaining neural net classifiers on NLP tasks

July 9, 2023
Toronto, Canada

European Workshop on Algorithmic Fairness - 2 Spotlights
- *Are fairness metric scores enough to assess discrimination biases in machine learning?*
- *How Optimal Transport Can Tackle Gender Biases in Multi-Class Neural Network Classifiers for Job Recommendation*

June 7-9, 2023
Zurich, Switzerland

MobiliT.AI - Poster session
COCKATIEL: Continuous Concept ranKed ATtribution with Interpretable ELeMents for explaining neural net classifiers on NLP tasks

May 30-31, 2023
Toulouse, France

3IA Doctoral workshop - Talk
Bias in Natural Language Processing

November 22-23, 2021
Toulouse, France

"Journée du GDR TAL 2021" - Talk
Bias in Natural Language Processing

October 5, 2021
Toulouse, France

BROADER PUBLIC SCIENTIFIC COMMUNICATION

- Interview for [ActuIA](#), December 11, 2023.
- Interview for [Toulouse University Press Service](#), February 9, 2023.
- Participated as a guest in the online roundtable discussion: [Reducing Bias and Inequality: Can AI Play a Role?](#), March 22, 2022.
- Participated as a guest in the roundtable discussion [Women in AI](#), March 8, 2022.

EXTRACURRICULAR EXPERIENCES

PhD student representative at [ANITI Diversity Commission](#)
The commission initiates and launches actions to combat prejudice about AI and to fight against gender stereotypes.

2021 - Today
Toulouse, France

Cofounder and treasurer at **AlumnIA association**
Creation and management of the budget of the association of alumni of the Data Science master at École polytechnique Paris.

2020 - 2023
Paris, France

Volunteer at [Femmes & Mathématiques association](#)
Organization of "Filles & maths : une équation lumineuse" events to help high school girls discover the trades related to mathematics and not undermine their ambitions.

2018 - 2019
Paris, France

OTHER SKILLS

Languages:

French - *Native*
English - *Professional working*
Spanish - *Elementary*

Programming Languages:

Python (working with PyTorch, TensorFlow, Transformers), R, SQL, \LaTeX .