Fanny Jourdan

I'm researcher at IRT Saint Exupery (Technological Research Institut) for the DEEL team. I am interested in Explainability, Interpretability and Logic for Natural Language Processing algorithms.

🌐 fanny-jourdan ✉ fanny.jourdan@irit.fr
 fanny-jourdan in Fanny Jourdan 🐦 @Fannyjrd_

EDUCATION

PhD on Computer Science at **Paul Sabatier University**	2021 - 2024
Fairness and Explainability on Natural Language Processing	Toulouse, France
Supervisors: Nicholas Asher and Laurent Risser	

Master's Degree on Economics at **Paris Dauphine University**	2020 - 2021
Digital economics, management, innovation, regulation theory	Paris, France

Master's Degree on Mathematics at **École polytechnique**	2018 - 2020
Data Sciences, Statistical theory, Natural Language Processing,	Paris, France
Optimisation, Causal Inference	

Bachelor's Degree on Mathematics at **Paris-Sud University**	2015 - 2018
Fundamental mathematics: Algebra, Topology, Calculus	Paris, France
Holomorphic function, Probability	

WORK EXPERIENCE

Researcher at IRT Saint Exupery	March 2024 - Today
Explainability (XAI) and logic for Natural Language Processing algorithms.	Toulouse, France

Research intern at Criteo AI Lab	May 2020 - September 2020
Study of the causal effect of displays (advertising banners sent by Criteo) on the sales of associated products.	Paris, France

Data scientist intern at EDF R&D	April 2019 - August 2019
Analysis and disaggregation of individual electrical consumption time series.	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	2021 - 2023
Mathematics for Undergraduate students on Biology Bachelor	Toulouse, France
Teaching assistant at Paul Sabatier University	2021 - 2022
Logic for Undergraduate students on Computer Science Bachelor	Toulouse, France
Oral examiner at Parc de Vilgenis High School	2018 - 2020
"Colles" on mathematics for Undergraduate students in "Higher school preparatory classes"	Massy, France

TALKS AND POSTERS PRESENTATIONS

IA Pau Festival - Keynote Speaker	December 1, 2023
Understanding and reducing discriminative biases in Natural Language Processing algorithms.	Pau, France
NLTP group, Utrecht university - Invited speaker	October 16, 2023
COCKATIEL: COntinuous Concept ranKed ATtribution with Interpretable ELements for explaining neural net classifiers on NLP tasks	Utrecht, Netherlands
TrustNLP Workshop ACL - Oral	July 14, 2023
Are fairness metric scores enough to assess discrimination biases in machine learning?	Toronto, Canada
TrustNLP Workshop ACL - Poster session (2 posters)	July 14, 2023
- Are fairness metric scores enough to assess discrimination biases in machine learning?	Toronto, Canada
- COCKATIEL: COntinuous Concept ranKed ATtribution with Interpretable ELements for explaining neural net classifiers on NLP tasks	
ACL - Findings spotlight	July 9, 2023
COCKATIEL: COntinuous Concept ranKed ATtribution with Interpretable ELements for explaining neural net classifiers on NLP tasks	Toronto, Canada
European Workshop on Algorithmic Fairness - 2 Spotlights	June 7-9, 2023
- Are fairness metric scores enough to assess discrimination biases in machine learning?	Zurich, Switzerland
- How Optimal Transport Can Tackle Gender Biases in Multi-Class Neural Network Classifiers for Job Recommendation	

MobiliT.AI - Poster session	May 30-31, 2023
COCKATIEL: COntinuous Concept ranKed ATtribution with	Toulouse, France
Interpretable ELements for explaining neural net classifiers on NLP tasks	

3IA Doctoral workshop - Talk	November 22-23, 2021
Bias in Natural Language Processing	Toulouse, France

"Journée du GDR TAL 2021" - Talk	October 5, 2021
Bias in Natural Language Processing	Toulouse, France

BROADER PUBLIC SCIENTIFIC COMMUNICATION

- Finalist of MT180 contest (french version of 3MT) - see the video, *March 15, 2024*.
- Interview for "Mardi quiz" at Quai des Savoirs Toulouse, *March 5, 2024*.
- 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	2021 - Today
The commission initiates and launches actions to combat prejudice about	Toulouse, France
AI and to fight against gender stereotypes.	

Cofounder and treasurer at **AlumnIA association**	2020 - 2023
Creation and management of the budget of the association of alumni	Paris, France
of the Data Science master at École polytechnique Paris.	

Volunteer at Femmes & Mathématiques association	2018 - 2019
Organization of "Filles & maths : une équation lumineuse" events to help	Paris, France
high school girls discover the trades related to mathematics and not	
undermine their ambitions.	

OTHER SKILLS

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

Programming Languages:
Python (working with PyTorch, TensorFlow, Transformers), R, SQL, LaTeX.