

Dorsaf Sallami

✉ dorsaf.sallami@mila.quebec🌐 [My personal website](#)🐙 [GitHub](#)🔖 [Google Scholar](#)

Research Interests

My research focuses on responsible AI for information integrity, with an emphasis on automated fact-checking and fake news detection. I study how language models and human-AI collaboration can support verification, and how to evaluate these systems beyond accuracy, including fairness, explainability, contextual ambiguity, and societal impact.

- Member, [Centre for the Study of Democratic Citizenship](#) 🌐.

Academic Positions

Postdoctoral Researcher, Mila – Québec AI Institute

Jan 2026 – Present

- Advisors: [Reihaneh Rabbany](#) 🌐 (McGill University) and [Jean-François Godbout](#) 🌐 (Université de Montréal).
- Project lead for [Veracity](#) 🌐, an open-source AI fact-checking initiative:
 - Cited in [Canada's National Artificial Intelligence Strategy](#) 🌐 as an example of impactful Canadian AI research.
 - Added to the [Senate of Canada's resources for senators and staff](#) 🌐 in May 2026.

Education

Ph.D. Université de Montréal, Computer Science

Sept 2022 – Dec 2025

- **Supervisors:** Professor [Esma Aïmeur](#) 🌐 and Professor [Gilles Brassard](#) 🌐
- **Areas:** NLP, Responsible AI, and Fake News Detection

M.Sc. Université de Montréal, Computer Science

Sept 2020 – Aug 2022

- **Supervisor:** Professor [Esma Aïmeur](#) 🌐
- **Areas:** Fake News Detection, Multimodal Data, Explainable AI

M.Sc. Higher Institute of Applied Sciences and Technology, Tunisia

Sept 2017 – Aug 2020

- Software Engineering (Master's equivalent)
- Advanced systems design, software engineering, and applied computing

B.Sc. Higher Institute of Applied Sciences and Technology, Tunisia

Sept 2015 – Aug 2017

- Mathematics, Physics, Computer Science (Bachelor's equivalent)
- Preparatory cycle for national engineering program

Professional Development

Stanford University, Ethics, Technology, and Public Policy for Practitioners

Fall 2026

- Live cohort-based professional program on technology ethics, public policy, and responsible innovation.

Selected Publications

Journal Articles

1. [SALLAMI, Dorsaf](#), and [AÏMEUR, Esma](#). **Explain Further: Multi-Level Explanations for Fake News Detection Using Large Language Models**. International Journal of Information Technology, 2025, p. 1-8.
2. [SALLAMI, Dorsaf](#), and [AÏMEUR, Esma](#). **Exploring beyond detection: a review on fake news prevention and mitigation techniques**. Journal of Computational Social Science, 2025, vol. 8, no 1, p. 1-38.
3. [SALLAMI, Dorsaf](#), and [AÏMEUR, Esma](#). **Fairframe: a fairness framework for bias detection and mitigation in news**. AI and Ethics Journal, 2024, p. 1-17.

4. ABDALLAH, Wejden, KANZARI, Dallel, [SALLAMI, Dorsaf](#), et al. **A deep reinforcement learning based decision-making approach for avoiding crowd situation within the case of Covid'19 pandemic**. Computational Intelligence, 2022, vol. 38, no 2, p. 416–437.

Conference Papers

5. [SALLAMI, Dorsaf](#), and AÏMEUR, Esma. **Aletheia: Detect, Discuss, and Stay Informed on Fake News**. Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence, IJCAI-25, p.11109–11113.
6. AÏMEUR, Esma, BRASSARD, Gilles, and [SALLAMI, Dorsaf](#). **Too Focused on Accuracy to Notice the Fallout: Towards Socially Responsible Fake News Detection**. Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society. Vol. 8. No. 1. 2025.
7. [SALLAMI, Dorsaf](#), and AÏMEUR, Esma. **Does Gender Matter? Examining and Mitigating Gender Bias in Fake News Detection**. International Symposium on Foundations and Practice of Security. Springer, Cham, 2025. p. 249-266.
8. [SALLAMI, Dorsaf](#), BEN SALEM, Rim, and AÏMEUR, Esma. **Trust-based Recommender System for Fake News Mitigation**. In: Adjunct Proceedings of the 31st ACM Conference on User Modeling, Adaptation and Personalization. 2023. p. 104–109.
9. AMRI, Sabrine, [SALLAMI, Dorsaf](#), et AÏMEUR, Esma. **Exmulf: an explainable multimodal content-based fake news detection system**. In: International Symposium on Foundations and Practice of Security. Cham: Springer International Publishing, 2021. p. 177–187.

Workshop Papers

10. [SALLAMI, Dorsaf](#), GUEDDICHE, Ahmed, and AÏMEUR, Esma. **From Hype to Reality: Transformer-Based Models for Fake News Detection Performance and Robustness Revealed**. In 3rd Workshop on Adverse Impacts and Collateral Effects of Artificial Intelligence Technologies. IJCAI 2023.

Honors and Scholarships

- **2025–2026:** [Rector's Honour List](#) , Université de Montréal (top 10% of doctoral graduates in the program).
- **2020–2025:** First nationwide recipient of the Tunisian Government Scholarship for studies in Canada, awarded by the Tunisian Ministry of Higher Education (2020).
- **Fall & Winter 2022, 2023, 2024, 2025:** DIRO Excellence-Writing-Support Scholarship, Université de Montréal

Teaching and Supervision

Courses and Tutoring

Teaching Assistant, Université de Montréal

- **IFT2015/IFT6002 – Data Structures** *Winter 2023, 2024, 2025*
Led laboratory sessions, assessed assignments, and provided individualized support during office hours.
- **IFT6261 – Knowledge Processing** *Winter 2024*
Assessed assignments and supported students during office hours.

Tutor, Tutoring 2.0 Program, Université de Montréal

Fall 2022

- Provided academic support to undergraduate students in computer science courses.

Invited Lectures

- **Guest Lecturer**, Knowledge Processing (IFT 6261), Université de Montréal: “Information Integrity in the Age of Generative AI” — October 2026 | Montréal, Canada
- **Guest Lecturer**, Knowledge Processing (IFT 6261), Université de Montréal: “Multimodal Content-Based Fake News Detection System” — February 2022 | Montreal, Canada

Doctoral Research Mentorship

Sukanya Krishna, Ph.D. Student

- Research project: *Ambiguity and Underspecification in AI-Assisted Fact-Checking*

Harvard University

Jan. 2026 – Present

Undergraduate Research Supervision and Mentorship

Tong Wu, Undergraduate Research Intern

- Research project: *Multi-Turn Fact-Checking with Large Language Models*

McGill University

Summer 2026

Andrew Lin, Undergraduate Research Intern

- Research project: *Claim Augmentation for Fact-Checking*

McGill University

Summer 2026

Hugh Kelly, Undergraduate Research Intern

- Research project: *Source Preferences Affect Trust in Large Language Models*

McGill University

Winter 2026

Arpad Botond Rigo, Undergraduate Research Intern

- Research project: *Simulation of Fake News Propagation*

Université de Montréal

Summer 2023

Academic Service and Leadership

Editorial Service

- **Topic Editor** — “*AiOfAi: Adverse Impacts and Collateral Effects of Artificial Intelligence Technologies*” [🔗](#), *Frontiers in Artificial Intelligence*, 2026

Workshop and Event Organization

- **Workshop Organiser** — *Disinformation and Misinformation under the lens of Generative AI Safety* [🔗](#), AAAI 2027
- **Workshop Organiser** — *5th Workshop on Adverse Impacts and Collateral Effects of Artificial Intelligence Technologies*, [🔗](#), The ACM Web Conference 2027
- **Social Event Organiser** — *Who Decides What Counts as True?* [🔗](#), ACM FacCT 2026, June 2026, Montréal; sponsored by Mastercard.
- **Workshop Organiser** — *4th Workshop on Adverse Impacts and Collateral Effects of Artificial Intelligence Technologies* [🔗](#), The ACM Web Conference 2026 — *Proceedings* [🔗](#)
- **Publicity Chair** — *3rd Workshop on Adverse Impacts and Collateral Effects of Artificial Intelligence Technologies* [🔗](#), 32nd International Joint Conference on Artificial Intelligence (IJCAI), 2023 — *Proceedings* [🔗](#)

Research Coordination and Expert Contributions

- **Designated Coordinator** — *IVADO Thematic Semester* [🔗](#): Information Integrity in the Age of Generative AI, 2027
- **Contributing Expert** — *Global Trust Challenge* [🔗](#), 2026

Reviewing

- **Reviewer for** — International Joint Conference on Artificial Intelligence (IJCAI); AAAI/ACM Conference on AI, Ethics, and Society (AIES); Conference on Neural Information Processing Systems (NeurIPS); AAAI; and AI and Ethics Journal

Media Coverage and Public Engagement

- **Co-Author, Opinion Article** — “*Construire la confiance, au-delà de l’urgence d’adopter l’intelligence artificielle*” [🔗](#), *Le Devoir*, July 2026
- **Media Interview**, Knowable Magazine article, “*AI shows promise in the fight against fake news*” [🔗](#), interviewed by journalist Katya Zimmer — June 2026

- **Research Media Coverage**, UdeMNouvelles article “[Fake news: do the tools meant to protect us actually mask their flaws?](#)” [🔗](#) featuring my research on socially responsible fake news detection; also covered by [Tech Xplore](#) [🔗](#) and shared by [Crickit.ai](#) [🔗](#) — March 2026
- **Content Contributor** — CYBER project (HEC Montréal, IMC2, Centre de formation en langues des affaires), “[Comment éviter la propagation de fausses informations en ligne ?](#)” [🔗](#) — 2025

Talks and Presentations

Invited Talks and Seminars

- **Invited Speaker**, LATECE Seminar, Université du Québec à Montréal (UQAM): “[Peut-on confier la vérification des faits à l’IA ? Promesses, limites et perspectives](#)” [🔗](#) — October 2026 | Montréal, Canada | Presented in French
- **Invited Speaker**, Community of Inquiry in AI Ethics, Curtin University: “Why Accuracy Isn’t Enough When AI Decides What Counts as True” — July 2026 | Online
- **Invited Speaker**, First Canada–Denmark Municipal Knowledge Exchange Webinar: “RAG4MuniESG: ESG Assessment for Municipal Sustainability through Retrieval-Augmented Generation, a Quebec Case Study” — June 2025 | Online

Conference and Public Presentations

- **Poster Presenter**, 8th AAAI/ACM Conference on AI, Ethics, and Society (AIES): “Too Focused on Accuracy to Notice the Fallout: Towards Socially Responsible Fake News Detection” — October 2025 | Madrid, Spain
- **Presenter**, Forum INCYBER Canada (delegation of the Centre for Cybersecurity Belgium) — October 2025 | Montreal, Canada
- **Poster Presenter**, 34th International Joint Conference on Artificial Intelligence (IJCAI): “Aletheia: Detect, Discuss, and Stay Informed on Fake News” — August 2025 | Montreal, Canada
- **Poster Presenter**, Journée de la Recherche du DIRO: “CoALFake: Collaborative Active Learning with Human-LLM Co-Annotation for Cross-Domain Fake News Detection” — May 2025 | Montreal, Canada
- **Presenter**, 17th International Symposium on Foundations & Practice of Security (FPS): “Does Gender Matter? Examining and Mitigating Gender Bias in Fake News Detection” — December 2024 | Montreal, Canada
- **Presenter**, Innovate4Cities Conference: “What Do Policy, Planning, and Reporting Documents Reveal About Urban Sustainability Initiatives? Patterns and Insights from AI-Assisted Textual Analysis” — September 2024 | Montreal, Canada
- **Presenter**, 15th International Sustainability Transitions Conference (IST): “Nature and Biodiversity in Urban Sustainability Transition Planning and Reporting Documents: Patterns and Insights from AI-Assisted Textual Analysis” — June 2024 | Oslo, Norway (Online)
- **Presenter**, Cyberconference: “Beyond the Facts: Fake News, Generative Artificial Intelligence, and Cybersecurity” — April 2024 | Montreal, Canada
- **Presenter**, Inaugural Event of the Multidisciplinary Institute for Cybersecurity and Cyber Resilience (IMC²): “Explainable System for Countering Fake News and Doxxing” — October 2023 | Montreal, Canada
- **Presenter**, 3rd Workshop on Adverse Impacts and Collateral Effects of Artificial Intelligence Technologies: “From Hype to Reality: Transformer-Based Models for Fake News Detection Performance and Robustness Revealed” — August 2023 | Macao, China (Online)
- **Poster Presenter**, 31st ACM Conference on User Modeling, Adaptation and Personalization (UMAP) — October 2022 | Limassol, Cyprus
- **Presenter**, Annual Canadian Celebration of Women in Computing Conference (CanCWIC): “Multimodal Content-Based Fake News Detection System” — October 2022 | Toronto, Canada

Research and Industry Experience

Circle Medical, AI Research Intern

- Designed and implemented an evaluation framework for 10/CPT code prediction and HPI/Plan summary generation
- Developed quantitative metrics and procedures to measure model performance systematically.
- Benchmarked and compared several LLMs (OpenAI, Anthropic).

Montreal
June 2025 – Sep 2025

AlphaFixe Capital (Mitacs), AI Research Intern (Part-time)

- Research intern in Artificial Intelligence applied to sustainable finance
- Python, NLP, Data Analysis, Data collection, LLM

Montreal
Jan 2024 – May 2025

Videotron, Data Analyst Intern

- Data Cleaning, Data Integration, Data Visualization
- Python, NetSuite, Databases

Montreal
May 2023 – Aug 2023

Société Générale, BI - Data Analyst Intern (Part-time)

- Data Cleaning, Data Integration, Data Visualization
- Python, JS, GitHub, JSON, Databases, REST/POST Data visualisation, Power BI

Montreal
Sep 2021 – April 2022

ITPEAC, End of Studies Internship

- An approach to optimizing decision-making in a crisis based on deep reinforcement learning. Case Study: COVID-19

Tunisia
Feb 2020 – July 2020

Extra-Curricular Activities

- **2021–2025:** Lab Leader, [AI Lab for Cybersecurity](#), Montreal
- **2019–2020:** Training Manager, *ARSII Association*, Tunisia
- **2018–2020:** Treasurer, *ISSATSo Data Science Club*, Tunisia
- **Personal Interest:** When I'm not debugging code or mentoring students, you'll probably find me hiking up a trail 🏔️

Technical Skills

Programming Languages: Python, Java **Web Development:** Angular.js, Flask **Machine Learning:** PyTorch, NLTK, Scikit-learn, Keras, OpenCV **Documentation:** $\LaTeX$, Markdown

Languages

Arabic (Native), French (Professional), English (Professional)