

João Augusto Leite

 jaleite.com

 j.leite@sheffield.ac.uk

 [joao-augusto-leite](https://www.linkedin.com/in/joao-augusto-leite)

 [joaoaleite](https://github.com/joaoaleite)

ABOUT ME

PhD candidate and Research Associate at the University of Sheffield specializing in large language models in the context of factuality, including fact-checking, credibility analysis, and mitigating persuasion and hallucination. I combine rigorous academic research with over five years of experience developing machine learning pipelines. My hands-on expertise covers the full LLM lifecycle, including training (pre-training, fine-tuning, alignment), and inferencing (advanced prompting, reasoning, retrieval-augmented generation).

EXPERIENCE

Research Associate

The University of Sheffield

Jun 2024 – Present

Sheffield, UK

- Developed novel methods to detect Persuasion Techniques and Coordinated Inauthentic Behaviour in social media.

Data Scientist

PicPay

Oct 2021 – Oct 2022

Brazil (Remote)

- Enhanced search functionality by leading the end-to-end development of a vector-based retrieval system with query augmentation and a personalised recommender system. I conducted the evaluation of these features through rigorous A/B testing, leading to their successful deployment in production for millions of daily users.
- Built and maintained dashboards to track search KPIs and model performance.

Data Scientist

Birdie.ai

Nov 2020 – Oct 2021

São Carlos, Brazil

- Developed machine learning pipelines to extract features from product reviews and derive actionable insights for businesses. I led the development of several ML models for tasks such as named entity recognition, aspect-based sentiment analysis, aspect clustering, ranking, disambiguation, and ontology building.
- Designed and led data labelling tasks with a dedicated team to improve our machine learning pipelines.

Data Scientist Intern

Birdie.ai

Aug 2019 – Nov 2020

São Carlos, Brazil

- Developed web crawlers to retrieve product reviews from various retailers (e.g., Amazon, Ebay).

EDUCATION

PhD, Computer Science

The University of Sheffield

May 2026 (expected)

Sheffield, UK

- *Award:* Fully funded by an EPSRC Doctoral Training Partnership (DTP) Scholarship.
- *Research:* Automatic content verification methods assisted by large language models.

MSc, Computer Science

Universidade Federal de São Carlos

2024

São Carlos, Brazil

- *Research:* Noisy self-supervised methods for hate speech detection.

BSc, Computer Science

Universidade Federal de São Carlos

2021

São Carlos, Brazil

- Relevant coursework: Machine Learning, Natural Language Processing, Computer Vision, Parallel Processing, Algorithms, Data Structures.

TECHNICAL SKILLS

Programming: Python, C, Bash, SQL

ML/DL Frameworks: PyTorch, Transformers, TRL, LangChain, Scikit-learn, SpaCy, W&B

Developer Tools: Git, DVC, Docker, Slurm, AWS (S3, EC2)

Languages: English (Full Professional Proficiency), Portuguese (Native)

AWARDS

Google and ZenDesk Scholarship

Covered fees to attend the Lisbon Machine Learning School (LxMLS).

2024

EPSRC Doctoral Training Partnership (DTP) Scholarship

Covered full international fees and provided a stipend to support my Ph.D.

2022

Best Undergraduate Paper

Awarded by the Computer Science department at Universidade Federal de São Carlos.

2020

SoBigData TransNational Access (TNA)

Covered fees for a technical visit at the University of Sheffield to conduct a short-term research project.

2019

SELECTED PUBLICATIONS

Asterisk (*) denotes equal contribution.

- **A survey on automatic credibility assessment of textual credibility signals in the era of large language models.**
I Srba, O Razuvayevskaya, **JA Leite**, R Moro, IB Schlicht, S Tonelli, et al. *Transactions on Intelligent Systems and Technology*, 2025.
- **Weakly supervised veracity classification with LLM-predicted credibility signals.**
JA Leite, O Razuvayevskaya, K Bontcheva, C Scarton. *EPJ Data Science*, 2025.
- **A Cross-Domain Study of the Use of Persuasion Techniques in Online Disinformation.**
JA Leite, O Razuvayevskaya, C Scarton, K Bontcheva. *WWW (The Web Conf)*, 2025.
- **Comparison between parameter-efficient techniques and full fine-tuning: A case study on multilingual news article classification.**
O Razuvayevskaya*, B Wu*, **JA Leite***, F Heppell*, I Srba, C Scarton, et al. *Plos one*, 2024.
- **EUvsDisinfo: A Dataset for Multilingual Detection of Pro-Kremlin Disinformation in News Articles.**
JA Leite, O Razuvayevskaya, K Bontcheva, C Scarton. *CIKM*, 2024.
- **Noisy Self-Training with Data Augmentations for Offensive and Hate Speech Detection Tasks.**
JA Leite, C Scarton, DF Silva. *RANLP*, 2023.
- **Toxic language detection in social media for Brazilian Portuguese: New dataset and multilingual analysis.**
JA Leite, DF Silva, K Bontcheva, C Scarton. *AACL-IJCNLP*, 2020.

A full list of publications is available on my [Google Scholar profile](#).

PROFESSIONAL ACTIVITIES

Invited Presentations

- "Weakly Supervised Veracity Classification with Credibility Signals" - Universidade de São Paulo and Universidade Federal de São Carlos (2024)
- "Toxic language detection in social media for Brazilian Portuguese" - Universidade Federal de São Carlos (2022)
- "Building data pipelines enriched by AI models" - The Developer's Conference, Data Track International (2021)

Teaching

- **Graduate Teaching Assistant**, The University of Sheffield (Jul 2023 – Dec 2023). Marked undergraduate work, conducted demonstrations, and designed lab materials for the COM3110 Text Processing module.

Reviewer

- ACL (2025, 2024), CIKM (2024), ICWSM (2024), AACL-IJCNLP (2022), KDD (2022)