

Francielle Vargas, MSc and PhD in Artificial Intelligence

francielleavargas@usp.br
<https://francielleavargas.github.io/>

Main Skills Natural Language Processing
 Computational Social Science
 Machine Learning
 Responsible AI / Explainability and Interpretability

Education *PhD, Computer Science and Computational Mathematics*
 University of São Paulo
 Thesis: Socially Responsible and Explainable Fact-Checking and Hate Speech Detection.
 Artificial Intelligence and Natural Language Processing
 2019 - 2024

MSc, Computer Science and Computational Mathematics
 University of São Paulo
 Thesis: Semantic Clustering of Aspects for Opinion Mining. [see](#)
 Artificial Intelligence and Natural Language Processing
 2015 - 2017

BA, Linguistics
 Federal University of Minas Gerais

BS, Information Systems
 Pontifical Catholic University of Minas Gerais

Experience 2024-2024: *Google PhD Fellowship*, Federal University of Minas Gerais, Brazil
 2024-2024: *Visiting Researcher*, University of Southern California, USA
 2023-2023: *Invited Researcher*, Leibniz Institute for the Social Sciences, Germany
 2021-2022: *Data Scientist*, Cisco Webex, USA
 2015-2023: *Graduate Researcher*, University of São Paulo, Brazil
 2014-2015: *System Analyst*, Unisys, Brazil
 2010-2014: *Research Assistant*, Federal University of Minas Gerais, Brazil
 2008-2010: *System Analyst*, Grupo Barcelos, Brazil

Courses 2024: *2nd Mexican NLP Summer School*, NAACL 2024, Mexican. [see](#)
 2023: *2nd Summer School on Deep Learning in NLP*, RANLP 2024, Bulgaria. [see](#)
 2022: *2nd Advanced NLP School*, Université Grenoble Alpes, France. [see](#)
 2021: *4th Advanced School in Big Data Analysis*, ICMC-USP, Brazil. [see](#)
 2020: *10th Lisbon Machine Learning School*, INESC, Portugal. [see](#)
 2019: *Introduction on the Stars - Astrophysics*, IFSC-USP. [see](#)
 2016: *The R language with an emphasis on probability*, ICMC-USP. [see](#)
 2015: *4th Summer School in Computer Science*, DCC-UFMG, Brazil. [see](#)
 2014: *Software Engineering Specialization*, PUC Minas. [see](#)

Context-Aware and Expert Data Resources for Brazilian Portuguese Hate Speech Detection

Francielle Vargas, Isabelle Carvalho, Thiago Pardo, Fabrício Benevenuto
Natural Language Engineering Journal. [see](#)

Discourse Annotation Guideline for Low-Resource Languages

Vargas, F., Schmeisser-Nieto, W., Rabinovich, Z., Benevenuto, F., Pardo, T.A.S.
Natural Language Engineering Journal. [see](#)

HausaHate: An Expert Annotated Corpus for Hausa Hate Speech Detection

Vargas, F., Guimarães, S., Alves, D., Muhammad, H. S., Abdulmumin, I., Ahmad, I. S.,
Mohamed, D., Pardo, T.A.S., Benevenuto, F.
Proceedings of the 8th Workshop on Online Abuse and Harms (WOAH @ NAACL 2024).
[see](#)

Extended Multimodal Hate Speech Event Detection During Russia-Ukraine Crisis

Thapa, S, Rauniyar, K., Jafri, F. A., Veeramani, H., Jain, R., Jain, S., Vargas, F., Hürriyetoğlu,
A., Naseem, U.
Workshop Proceedings of the 14th Recent Advances in Natural Language Processing (CASE
@ RANLP 2024). [see](#)

2023

Predicting Sentence-Level Factuality of News and Bias of Media Outlets

Vargas, F., Jaidka, K., Pardo, T.A.S., Benevenuto, F.
Recent Advances in Natural Language Processing (RANLP). [see](#)

Socially Responsible Hate Speech Detection: Can Classifiers Reflect Social Stereotypes?

Vargas, F., Carvalho, I., Hürriyetoğlu, A., Pardo, T.A.S., Benevenuto, F.
Recent Advances in Natural Language Processing (RANLP). [see](#)

NoHateBrazil: A Brazilian Portuguese Text Offensiveness Analysis System

Vargas, F., Carvalho, I., Schmeisser-Nieto, W., Benevenuto, F., Pardo, T.A.S.
Recent Advances in Natural Language Processing (RANLP). [see](#)

Multimodal Hate Speech Detection

Thapa, S., Jafri, F. A., Hürriyetoğlu, A., Vargas, F., Lee, R. K., Naseem, U.
Workshop Proceedings of the 17th Conference of the European Chapter of the Association
for Computational Linguistics (CASE @ EACL 2023). [see](#)

2022

HateBR: A Large Expert Annotated Corpus of Brazilian Instagram Comments for Abusive Language Detection

Vargas, F., Carvalho, I., Góes, F.R., Pardo, T.A.S., Benevenuto, F.
13th Conference on Language Resources and Evaluation (LREC). [see](#)

Rhetorical Structure Approach for Online Deception Detection: A Survey

Vargas, F., D'Alessandro, J., Rabinovich, Z., Benevenuto, F., Pardo, T.A.S.
13th Conference on Language Resources and Evaluation (LREC). pp.5906-5915.

Studying Dishonest Intentions in Brazilian Portuguese Texts

Vargas, F., Pardo, T.A.S.
Deceptive AI. Springer: Communications in Computer and Information Science. [see](#)

Extended Multilingual Protest News Detection

Hurriyetoglu, A., Mutlu, O., San, F. D., Uca, O., Gurel, A. S., Radford, B., Dai, Y., Hettiarachchi, H., Stoehr, N., Nomoto, T., Slavcheva, M., Vargas, F., Javid, A., Beyhan, F., Yoruk, E.

Workshop Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (CASE @ EMNLP 2022). [see](#)

2021

Contextual-Lexicon Approach for Abusive Language Detection

Vargas, F., Góes, F.R., Carvalho, I., Pardo, T.A.S., Benevenuto, F.

Recent Advances in Natural Language Processing (RANLP). [see](#)

Towards Discourse-Aware Models for Multilingual Fake News Detection

Vargas, F., Benevenuto, F., Pardo, T.A.S.

Recent Advances in Natural Language Processing (RANLP). [see](#)

Implicit Opinion Aspect Clues in Portuguese Texts: Analysis and Categorization

Machado, M. T., Pardo, T.A.S., Ruiz, E. E. S., Di Felippo, A., Vargas, F.

15th International Conference on the Computational Processing of Portuguese (PROPOR).

[see](#)

2020 and before

Linguistic Rules for Fine-Grained Opinion Extraction

Vargas, F., Pardo, T.A.S.

Workshop Proceedings of the 14th International AAAI Conference on Web and Social Media (SocialSens @ ICWSM 2020). [see](#)

Identifying Fine-Grained Opinion and Classifying Polarity on Coronavirus Pandemic

Vargas, F., Santos, R.S.S., F., Rocha, P.R.

9th Brazilian Conference on Intelligent Systems (BRACIS 2020). [see](#)

Aspect Clustering Methods for Sentiment Analysis

Vargas, F., Pardo, T.A.S.

13th International Conference on the Computational Processing of Portuguese (PROPOR 2018). [see](#)

Data Resources [HateBR](#): A Benchmark Expert Hate Speech Dataset of Brazilian Instagram comments
[FactNews](#): A Benchmark Dataset for Sentence-Level Factuality and Media Bias Prediction
[HausaHate](#): A Benchmark Expert Dataset for Hausa Hate Speech Detection
[OPCovidBR](#): A Dataset of Covid-19 Tweets For Aspect-Based Sentiment Analysis
[AspectPRO](#): A Dataset of Consumer Reviews for Aspect-Based Sentiment Analysis
[PRO](#): A Set of Taxonomies for Aspect-Based Sentiment Analysis
[MOL](#): A Context-Aware Multilingual Offensive Lexicon

Automated Methods [B+M](#): An Optimized Bag-of-Words Model by Contextual Lexicon Attention for Explainable Hate Speech Detection
[SSA](#): A Post-hoc Explanation Countermeasure to Assess Social Bias in HS Classifiers
[FACTual](#): Sentence-Level News Source Reliability Estimation

Computer Programs [OPCluster-PT](#): A System for Extraction and Clustering of Fine-Grained Opinions
[NoHateBrazil](#): A Brazilian Portuguese Text Offensiveness Analysis System

Invited Speaker 2024: [Invited Researcher] [MOLA - Language & Morality Lab](#)
University of Southern California (USC). Los Angeles, California
Title: *Fighting Misinformation and Polarization: Socially Responsible and Explainable Language Technologies for Fact-Checking and Hate Speech Detection*. [see](#)

2024: [Keynote Speaker] [Conference cum Conclave on Emerging trends in Journalistic and Media Practices](#)
DG Vaishnav College (DDGDV). Chennai, India
Title: *Predicting Sentence-Level News Source Reliability for Fact-Checking*. [see](#)

2023: [Invited Speaker] [Conference on Harmful Online Communication](#)
Leibniz Institute for the Social Sciences (GESIS), Cologne, Germany
Title: *Countering Harmful Online Communication in Brazil: Predicting Fine-Grained Factuality of News and Offensive Context of Social Media Comment*. [see](#)

Organizing and Program Committee

Organizing Committee

1. Data(set) Co-chair. *ICWSM 2023*. [see](#)
2. Accessibility Chair. *ICWSM 2022*. [see](#)
3. Diversity, Equity, and Inclusion Chair. *ICWSM 2021*. [see](#)
4. Steering Committee. 6th CASE@*RANLP 2023*. [see](#)
5. Steering Committee. 5th CASE@*EMNLP 2022*. [see](#)

Program Committee

- *EMNLP (2022) (2023) (2024)*
- *NAACL (2022) (2024)*
- *ACL (2023)*
- *LREC (2024)*
- *COLING (2025)*

- *ICWSM* (2025)
- *CIKM* (2024)(2023)
- *HICSS*(2023)
- *WOAH* (NAACL-HLT 2022) (ACL 2023) (NAACL 2024)
- *FEVER* (EACL 2023) (EMLP 2024)
- *CODI* (COLING 2022) (ACL 2023) (EACL 2024)
- *ARGMINING-CMNA* (COMMA 2022) (COLING 2022) (ICLP 2023)
- *SoCal NLP* (2022)

Conference Participation EMNLP 2024. to appear
 NAACL 2024. [see](#)
 RANLP 2023. [see](#)
 ACL 2023. [see](#)
 LREC 2022. [see](#)
 RANLP 2021. [see](#)
 ICWSM 2021. [see](#)
 ACL 2020. [see](#)
 EMNLP 2020. [see](#)
 LREC 2020. [see](#)
 BRACIS 2020. [see](#)

Adviser **CS Undergraduate:** Lucas Sobral Fontes Cardoso, University of São Paulo, 2020.
Thesis: *Web resources for an aspect identification and clustering tool in fine-grained opinion mining.* [see](#)

Recommendation Letters Fabricio Benevenuto. *DCC-UFMG.* [letter](#)
 Norah Abokhodair. *Microsoft Research.* [letter](#)
 Roseli Romero. *ICMC-USP.* [letter](#)
 Carlos Ramisch. *LIS-CNRS-Aix Marseille Université.* [letter](#)

Coding Python, NLTK, spaCy, Gensim, Keras, Tensorflow, PyTorch, SQL, RDF(s), OWL, Unix.

Language **Portuguese:** Native.
English: Advance.

Awards & Honors Google Latin America Research Awards (LARA 2024).
 NAACL Diversity & Inclusion Award (NAACL 2024).
 Outstanding Academic Project and Honorable Mention (UFMG 2013). [see](#)
 Outstanding Academic Project and Honorable Mention (UFMG 2012). [see see](#)