

## Francielle Vargas

francielleavargas@usp.br +55 31 999685891

[\[page\]](#) [\[dblp\]](#) [\[linkedin\]](#) [\[google scholar\]](#) [\[github\]](#)

### Education

*PhD, Computer Science and Computational Mathematics*

University of São Paulo

August 2019 - current

*MSc, Computer Science and Computational Mathematics*

University of São Paulo

Thesis: Semantic clustering of aspects for opinion mining. [see](#)

August 2015 - November 2017

*BA, Linguistics*

Federal University of Minas Gerais

August 2010 - June 2014

*BSc, Information Systems*

Pontifical Catholic University of Minas Gerais

August 2006 - December 2009

### Research Groups

[LOCUS - Social Computing Lab](#), 2021/2023

Computer Science Department, Federal University of Minas Gerais, Brazil.

[NILC - Interinstitutional Center for Computational Linguistics](#), 2015/2023

Institute of Mathematical and Computer Sciences, University of São Paulo, Brazil.

### Organizing and Program Committee

#### Program Committee

1. NLP Applications Track. *EMNLP 2022*. [see](#)
2. Natural Language Generation Track. *EMNLP 2022*. [see](#)
3. Discourse and Pragmatics Track. *EMNLP 2022*. [see](#)
4. Resources and Evaluation Track. *EMNLP 2022*. [see](#)
5. Computational Social Science and Cultural Analytics Track. *EMNLP 2022*. [see](#)
6. Sentiment Analysis, Stylistic Analysis, and Argument Mining Track. *EMNLP 2022*. [see](#)
7. 5th Southern California Natural Language Processing Symposium. [see](#)
8. 3rd Workshop on Computational Approaches to Discourse. *COLING 2022*. [see](#)
9. 9th Workshop on Argument Mining. *COLING 2022*. [see](#)
10. 6th Workshop on Online Abuse and Harms. *NAACL-HLT 2022*. [see](#)
11. NAACL Student Research Workshop. *NAACL-HLT 2022*. [see](#)
12. 22th Workshop on Computational Models of Natural Argument. *COMMA 2022*. [see](#)
13. 1st Workshop on Argumentation & Machine Learning. *COMMA 2022*. [see](#)
14. 5th International Workshop Challenges and Applications of Automated Extraction of Socio-political Events from Text. *EMNLP 2022*. [see](#)
15. 4th International Workshop Challenges and Applications of Automated Extraction of Socio-political Events from Text. *ACL-IJCNLP 2021*. [see](#)

## Organizing Committee

1. Data(set) Co-chair. *ICWSM 2023*. [see](#)
  2. Steering Committee. 5th International Workshop Challenges and Applications of Automated Extraction of Socio-political Events from Text. (*EMNLP 2022*). [see](#)
  3. Accessibility Chair. *ICWSM 2022*. [see](#)
  4. Diversity, Equity, and Inclusion Chair. *ICWSM 2021*. [see](#)
- 

## Selected Publications

### Journals

**Vargas, F.A.**, Pardo, T.A.S., Benevenuto, F., Schmeisser, W., Zaczynska, Z., Rabinovich, Z. (2022). Discourse-annotation guideline using rhetorical structure theory: multilingual examples. *Natural Language Engineering Journal*, pp. 1-63. [submitted](#)

**Vargas, F.A.**, Carvalho, I., Góes, F.R., Pardo, T.A.S and Benevenuto, F. (2022). Contextual-aware and expert resources for Brazilian Portuguese hate speech detection.. *Language Resources and Evaluation*, pp. 1-25. [submitted](#)

### Conferences

**Vargas, F.A.**, D'Alessandro, J., Rabinovich, Z., Benevenuto, F., Pardo, T.A.S (2022). Rhetorical structure approach online deception detection: A survey. *13th Conference on Language Resources and Evaluation (LREC)*, pp. 5906-5915. Marseille, France. [see](#)

**Vargas, F.A.**, Carvalho, I., Góes, F.R., Pardo, T.A.S and Benevenuto, F. (2022). HateBR: A expert annotated corpus of Brazilian Instagram comments for offensive language and hate speech detection. *13th Conference on Language Resources and Evaluation (LREC)*, pp. 7174-7183. Marseille, France. [see](#)

Machado, M., Pardo, T.A.S., Ruiz, E. E. S., Di Felippo, A., **Vargas, F.A.** (2022). Implicit aspect clues in Portuguese texts: analysis and categorization. *15th International Conference on Computational Processing of Portuguese (PROPOR)*, pp. 68-78. Fortaleza, Brazil. [see](#)

**Vargas, F.A.** Benevenuto, F. , Pardo, T.A.S (2021). Towards discourse-aware models for multilingual fake news detection. *International Conference Recent Advances in Natural Language Processing - Deep Learning for Natural Language Processing Methods and Applications - Student Research Workshop (RANLP)*, pp. 210-218. Held Online. [see](#)

**Vargas, F.A.**, Góes, F.R., Carvalho, I., Benevenuto, F. and Pardo, T.A.S. (2021). Contextual-lexicon approach for abusive language detection. *International Conference Recent Advances in Natural Language Processing - Deep Learning for Natural Language Processing Methods and Applications (RANLP)*, pp. 1442-1451. Held Online. [see](#)

**Vargas, F.A.**, Santos, R. R. S and Rocha, P. R. (2020) Identifying fine-grained opinion and classifying polarity on coronavirus pandemic. *9th Brazilian Conference on Intelligent Systems (BRACIS)*, pp. 511-520. Rio Grande, Brazil. [see](#)

**Vargas, F.A.** and Pardo, T.A.S. (2018). Aspect clustering methods for sentiment analysis. *13th International Conference on Computational Processing of Portuguese (PROPOR)*, pp. 365-374. Canela, Brazil. [see](#)

### Workshops

Nurse, J. R., **Vargas, F. A.**, Hassan, N, Wang, D., Andriotis, P., Ghenai, A., Jaidka, K., Mustafaraj, E., Joseph, K., Welles, B. F. (2021) Towards a diverse, inclusive, accessible and equitable AAAI international conference on web and social media (ICWSM). *Workshop Proceedings of the 15th International AAAI Conference on Web and Social Media (ICWSM)*, pages 1-14. Held Online. [see](#)

**Vargas, F.A.** and Pardo, T.A.S (2020). Linguistic rules for fine-grained opinion extraction. *Workshop Proceedings of the 14th International AAAI Conference on Web and Social Media (ICWSM)*, pp. 1-6. Held Online. [see](#)

Pardo, T.A.S., Baptista, J., Duran, M.S., Nunes, M.G.V., Nóbrega, F.A.A., Aluísio, S.M., Di Felippo, A., Seno, E.R.M., Silva, R.R., Anchiêta, R.T., Brum, H.B., Dias, M.S., Martins, R.S.O., Maziero, E.G., Souza, J.W.C., **Vargas, F.A.** (2017). The coreference annotation of the CSTNews corpus. In the *2nd Workshop on Evaluation of Human Language Technologies for Iberian Languages*. Workshop Proceedings of the 33th Conference of the Spanish Society for Natural Language Processing (SEPLN). 102-112. Murcia, Spain. [see](#)

---

## Previous Positions

### Academic Positions

2019/2022: **Graduate Research Assistant - Dept. of Computer Science.**  
University of São Paulo. [see](#)  
2021/2021: **Graduate Teaching Assistant in Neural Networks and Deep Learning.**  
University of São Paulo. [see](#)  
2020/2020: **Graduate Teaching Assistant - Computing Theory and Compilers.**  
University of São Paulo. [see](#)  
2016/2016: **Graduate Teaching Assistant in Computer Seminars.**  
University of São Paulo. [see](#)  
2015/2017: **Graduate Research Assistant - Dept. of Computer Science.**  
University of São Paulo. [see](#)  
2011/2014: **Undergraduate Research Assistant - Dept. of Linguistics.**  
Federal University of Minas Gerais. [see](#)

### Industry Positions

2021/2023 **Research Fellow.** Sinch.  
2021/2022 **Data Scientist.** Webex-Cisco via Telus International.  
2014/2015: **System Analyst.** Unisys Corporation.  
2009/2010: **System Analyst.** Grupo Barcelos.

---

## Conference participation

2022: 13th International Conference on Language Resources and Evaluation (LREC). [see](#)  
2021: 16th Recent Advances in Natural Language Processing (RANLP). [see](#)  
2021: 15th International AAAI Conference on Web and Social Media (ICWSM) . [see](#)  
2020: The 2022 Empirical Methods in Natural Language Processing (EMNLP). [see](#)  
2020: 9th Brazilian Conference on Intelligent Systems (BRACIS). [see](#)  
2020: 12th International Conference on Language Resources and Evaluation (LREC). [see](#)  
2017: 9th Brazilian School on Computational Linguistics (EBRALC). [see](#)  
2013: 29th Congresso Latino Americano de Sociologia: Crisis e Emergencias Sociales en America Latina (ALAS). [see](#)

---

## Mentoring

2020: **Computer Science Undergraduate Student**, Lucas Sobral Fontes Cardoso: *Title: Web resources for an aspect identification and clustering tool in fine-grained opinion mining.*  
**University of São Paulo.** [certificate](#)

---

<b>Awards received</b>	2013: <b>Best Undergraduate Student Award</b> , Federal University of Minas Gerais. <a href="#">see</a> 2013: <b>Best Undergraduate Project Award</b> , Federal University of Minas Gerais. <a href="#">see</a> 2012: <b>Best Undergraduate Student Award</b> , Federal University of Minas Gerais. <a href="#">see</a>
<b>Short Courses</b>	2019: Introduction on the Stars - Astrophysics. <a href="#">certificate</a>
<b>Hackathon, Datathon and Summer Schools</b>	2nd Advanced Language Processing School, Université Grenoble Alpes, France, 2022. <a href="#">certificate</a> 1st Datathon and Hackathon “Antisemitism on Social Media”, Indiana University Bloomington, <a href="#">NLP Lab</a> , United States, 2020. <a href="#">certificate</a> 10th Lisbon Machine Learning School, Portugal, 2020. <a href="#">certificate</a> 4th Brazilian Advanced School in Big Data Analysis, University of São Paulo, 2020. <a href="#">certificate</a> 22th Computer Science Summer School, University of Sao Paulo, 2019. <a href="#">certificate</a> 4th Computer Science Summer School, Federal University of Minas Gerais, 2015. <a href="#">certificate</a> 10th Linguistics Summer School, Federal University of Minas Gerais, 2011. <a href="#">certificate</a>
<b>Participation in examination committees</b>	2021: <b>Computer Science Graduate Course</b> . University of São Paulo. Vivian Vieira. <i>Aprendendo sobre veracidade de informações: Desenvolvimento e avaliação de curso EAD que aplica método científico na análise de notícias</i> . <a href="#">certificate</a> 2021: <b>Computer Science Graduate Course</b> . University of São Paulo. Laudeir José Lopes de Oliveira. <i>Explorando o pensamento computacional com aprendizado de máquina: Elaboração de um material didático para uma oficina introdutória</i> . <a href="#">certificate</a> 2021: <b>Computer Science Undergraduate Course</b> . University of São Paulo. Lucas Silva Loureiro. <i>Sistema de Recomendação de Livros</i> . <a href="#">certificate</a> 2020: <b>Computer Science Undergraduate Course</b> . University of São Paulo. Luana Balador Belisario. <i>Análise discursiva aplicada à mineração de opinião: classificação de subjetividade para a língua portuguesa</i> . <a href="#">see</a> 2020: <b>Computer Science Undergraduate Course</b> . University of São Paulo. Paulo André Pimenta Aragão. <i>Revisão Sistemática da Literatura sobre Software Educativos para Ensino de Engenharia de Software</i> . <a href="#">see</a> 2020: <b>Computer Science Undergraduate Course</b> . University of São Paulo. Tarcisio Ladeia de Oliveira. <i>Mapeamento de Ambientes Externos com o Uso de Drones</i> . <a href="#">see</a> 2020: <b>Computer Science Undergraduate Course</b> . University of São Paulo. Rafael Farias Pinho. <i>Planejamento de rotas para VANTs com desvio de obstáculos e incertezas</i> . <a href="#">see</a> 2019: <b>27th USP International Symposium on Scientific and Technological Initiation</b> . <i>Member of the examination committee of the undergraduate projects</i> . <a href="#">certificate</a>
<b>Talks as an invited speaker</b>	<b>Natural Language Processing and Data Science</b> , University Center Methodist Izabela Hendrix, 2018. <a href="#">see</a> <b>Opinion Mining</b> , University of Sao Paulo, 2016. <a href="#">see</a>
<b>Technical-Scientific Resources</b>	<b>OpCluster</b> : Fine-grained opinion extraction and clustering tool. ( <a href="#">tool</a> ) ( <a href="#">web version</a> ) <b>OPCovidBR</b> : Aspect-based sentiment analysis dataset of Covid-19 Tweets. ( <a href="#">dataset</a> ) <b>AspectPRO</b> : Aspect-based annotated corpus of web consumer reviews. ( <a href="#">dataset</a> ) <b>HateBR</b> : Large-scale abusive language dataset of Brazilian Instagram comments. ( <a href="#">dataset</a> ) <b>Deceiver</b> : Multilingual discourse-annotated corpus of fake and truth news. ( <a href="#">dataset</a> ) <b>PRO</b> : Taxonomies for aspect-based sentiment analysis in Portuguese ( <a href="#">lexicon</a> ) <b>MOL</b> : Multilingual offensive lexicon ( <a href="#">lexicon</a> )

**Recommendation Letters**    **Prof. Dr. Fabricio Benevenuto.** *Computer Science Department, Federal University of Minas Gerais.* [letter](#)  
**Dra. Norah Abokhodair.** *Microsoft Research.* [letter](#)  
**Prof. Dra. Roseli Romero.** *Institute of Mathematics and Computer Sciences, University of São Paulo.* [letter](#)  
**Prof. Dr. Carlos Ramisch.** *Laboratoire d'Informatique et Systèmes – LIS – UMR CNRS 7020 – Aix Marseille Université.* [letter](#)

---

**Languages**                    Python, SQL, RDF(s), OWL, XML, JSON, Unix/Bash.

---