

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

Advisor: [Thiago A. S. Pardo](#)Co-Advisor: [Fabrício Benevenuto](#)

August 2019 - December 2023

MSc, Computer Science and Computational Mathematics

University of São Paulo

Thesis: Semantic clustering of aspects for opinion mining [pdf](#)Advisor: [Thiago A. S. Pardo](#)

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. 3rd Workshop on Computational Approaches to Discourse. In *29th International Conference on Computational Linguistics (COLING 2022)*. [see](#)
2. 6th Workshop on Online Abuse and Harms. In *The 2022 Conference of the North American Chapter of the Association for Computational Linguistics - Human Language Technologies (NAACL-HLT 2022)*. [to appear](#)
3. The NAACL Student Research Workshop. In *The 2022 Conference of the North American Chapter of the Association for Computational Linguistics - Human Language Technologies (NAACL-HLT 2022)*. [see](#)
4. 22th Workshop on Computational Models of Natural Argument . In *9th International Conference on Computational Models of Argument (COMMA 2022)*. [to appear](#)
5. 1st Workshop on Argumentation & Machine Learning. In *9th International Conference on Computational Models of Argument (COMMA 2022)*. [see](#)
6. 5th International Workshop Challenges and Applications of Automated Extraction of Socio-political Events from Text. In *2022 Empirical Methods in Natural Language Processing (EMNLP 2022)*. [see](#)

7. 4th International Workshop Challenges and Applications of Automated Extraction of Socio-political Events from Text. In *59th Annual Meeting of the Association for Computational Linguistics and 11th International Joint Conference on Natural Language Processing (ACL-IJCNLP 2021)*. [see](#)

Organizing Committee

1. Committee Member
5th International Workshop Challenges and Applications of Automated Extraction of Socio-political Events from Text. In *The 2022 Empirical Methods in Natural Language Processing (EMNLP 2022)*. [see](#)
2. Accessibility Chair
16th International Conference on Web and Social Media (ICWSM'22). [see](#)
3. Diversity, Equity, and Inclusion Chair
15th International AAAI Conference on Web and Social Media (ICWSM'21). [see](#)

Publications

Journals

Vargas, F.A., Benevenuto, F., Pardo, T.A.S (2022). A multilingual discourse-tagging guideline using rhetorical structure theory. *Dialogue & Discourse Journal*, pp. 1-63. [submitted](#)

Vargas, F.A., Carvalho, I., Góes, F.R., Pardo, T.A.S and Benevenuto, F. (2022). An expert annotated corpus and a contextual-aware offensive lexicon for Brazilian hate speech detection. *Natural Language Engineering Journal*, pp. 1-22. [accepted/to appear](#)

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 2022)*, pp. 1-10. [accepted/to appear](#)

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 2022)*, pp. 1-10. [accepted/to appear](#)

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 2022)*, pp. 68-78. Fortaleza, Brazil. [pdf](#)

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

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 2021)*, pp. 1442-1451. Held Online. [pdf](#)

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 2020)*, pp. 511-520. Rio Grande, Brazil. [pdf](#)

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

Vargas, F.A. and Pardo, T.A.S. (2017). Hierarchical clustering of aspects for opinion mining: a corpus study. *9th Escola Brasileira de Linguística Computacional*, pp. 69-84. São Leopoldo, Brazil. [pdf](#)

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. [pdf](#)

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. [pdf](#)

Vargas, F.A., Pardo, T.A.S. (2020) Studying dishonest intentions in texts. *st International Workshop on Deceptive AI*. In Communications in Computer and Information Science book series (CCIS, volume 1296). pp. 166-178. Held Online. [pdf](#)

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. [pdf](#)

Cambráia, C. N., **Vargas, F. A.**, Tavares, P. C., Santos, T. N. (2014) A arquitetura da repressão: as narrativas nos inquéritos policiais/militares. In *Anais do Encontro do grupo de estudos e trabalho em história e linguagem: política das narrativas*, III, pages 83-103, Belo Horizonte. Fafich/UFGM. [pdf](#)

Book published/organized

Bittencourt, A.; Costa, F. A.; **Vargas, F. A.** (2014). Por que amo Llansol? 1st ed. Belo Horizonte: Publicações Viva Voz, FALE/UFGM. 34p. ISBN 978-85-7758-227-3. [pdf](#)

Technical reports

Vargas, F.A. and Pardo, T.A.S. (2017). Empirical study on clustering and hierarchical organization of aspects for opinion mining. *Book of Technical Reports of the Institute of Mathematical and Computer Sciences*, University of São Paulo, Brazil. no.418. 48p. [pdf](#)

Previous Positions

2019/2023: **Graduate Research Assistant - Dept. of Computer Science.** University of São Paulo. [pdf](#)

2021/2021: **Graduate Teaching Assistant in Neural Networks and Deep Learning.** University of São Paulo. to appear

2020/2020: **Graduate Teaching Assistant - Compilers.** University of São Paulo. [pdf](#)

2016/2016: **Graduate Teaching Assistant**

University of São Paulo. [pdf](#). 2015/2017: **Graduate Research Assistant - Dept. of Computer Science.**

University of São Paulo. [pdf](#) Graduate Research Assistant - Dept. of Computer Science.
 2014/2015: **System Analyst**.
 Unisys Corporation.
 2011/2014: **Undergraduate Research Assistant - Dept. of Linguistics**.
 Federal University of Minas Gerais. [pdf](#)
 2009/2010: **System Analyst**.
 Grupo Barcelos.

Conference participation

2021: Recent Advances in Natural Language Processing (RANLP). [pdf](#)
 2021: 15th International AAAI Conference on Web and Social Media (ICWSM) . [pdf](#)
 2020: 58th annual meeting of the Association for Computational Linguistics (ACL). [pdf](#)
 2020: Conference on Empirical Methods in Natural Language Processing (EMNLP). [pdf](#)
 2020: 9th Brazilian Conference on Intelligent Systems (BRACIS). [pdf](#)
 2020: 1st Workshop Automated Extraction of Socio-political Events from News.In 12th International Conference on Language Resources and Evaluation (LREC). [pdf](#)
 2017: 9th Brazilian School on Computational Linguistics (EBRALC). [pdf](#)
 2013: 29th Congresso Latino Americano de Sociologia: Crisis e Emergencias Sociales en America Latina.[pdf](#)

Awards received

2013: **Best Undergraduate Student Award**, Federal University of Minas Gerais. [pdf](#)
 2013: **Best Extension Project Student Award**, Federal University of Minas Gerais. [pdf](#)
 2012: **Best Undergraduate Student Award**, Federal University of Minas Gerais. [pdf](#)

Hackathon, Datathon and Summer Schools

2nd Advanced Language Processing School, Université Grenoble Alpes, France, 2022.[pdf](#)
 1st Datathon and Hackathon “Antisemitism on Social Media”, Indiana University Bloomington, [NLP Lab](#), United States, 2020.[pdf](#)
 10th Lisbon Machine Learning School, Portugal, 2020.[pdf](#)
 4th Brazilian Advanced School in Big Data Analysis, University of São Paulo, 2020. [pdf](#)
 22th Computer Science Summer School, University of Sao Paulo, 2019.[pdf](#)
 4th Computer Science Summer School, Federal University of Minas Gerais, 2015. [pdf](#)
 10th Linguistics Summer School, Federal University of Minas Gerais, 2011.[pdf](#)

Participation in examination committees

2021: Computer science undergraduate student. University of São Paulo. Lucas Silva Loureiro. *Sistema de Recomendação de Livros*. [pdf](#)
 2021: Graduate student in computer science. University of São Paulo. Vivian Vieira. *Aprendendo sobre veracidade de informações: Desenvolvimento e avaliação de curso EAD que aplica método científico na análise de notícias*. [pdf](#)
 2021: Graduate student in computer science. University of São Paulo. Laudeir José Lopes de Oliveira. *Explorando o pensamento computacional com aprendizado de máquina: Elaboração de um material didático para uma oficina introdutória*. [pdf](#)
 2020: Computer science undergraduate student. University of São Paulo. Luana Belisario. *Análise discursiva aplicada à mineração de opinião: classificação de subjetividade para a língua portuguesa*. [pdf](#)
 2019: 27th USP International Symposium on Scientific and Technological Initiation. *Member of the examination committee of the undergraduate projects*. [pdf](#)

Talks as an invited speaker

- **Natural Language Processing and Data Science**
 University Center Methodist Izabela Hendrix, Belo Horizonte, Brazil, 2018. [pdf](#)
 - **Opinion Mining**
 University of Sao Paulo, São Carlos, Brazil, 2016. [pdf](#).

Technical-Scientific Productions **OpCluster-PT**: Fine-grained opinion extraction and clustering tool. ([code](#)) ([web version](#))
OPCovidBR: Aspect-based sentiment analysis dataset of Covid-19 Tweets. ([dataset](#))
HateBR: Large-scale abusive language dataset of Brazilian Instagram comments. ([dataset](#))
Deceiver: Multilingual discourse-annotated corpus of fake and truth news. ([dataset](#))
MOL: Multilingual Offensive Lexicon ([lexicon](#))

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

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