Leonardo F. R. Ribeiro

leoribeiro.github.io

LinkedIn · Google Scholar · leonardofribeiro@gmail.com

Education

November 2018 - June 2022

Technische Universität Darmstadt, Darmstadt, Germany – *PhD in Computer Science*

Topic: Graph-based Neural Methods for Natural Language Generation. Advisor: Prof. Dr. Iryna Gurevych. Summa Cum Laude.

November 2015 - August 2018

Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil – *MSc in Computer and Systems Engineering*

Thesis: Learning Network Node Representations from Structural Identity. Advisor: Prof. Dr. Daniel Ratton Figueiredo. **Recipient of the 2018 Brazilian Artificial and Computational Intelligence Thesis Award**

September 2012 - June 2013

Universidad del País Vasco, San Sebastian, Spain – *International student*

Exchange Programme. Took computer science classes and did an internship as an exchange student.

January 2009 - August 2014

Centro Federal de Educação Tecnológica de Minas Gerais, Belo Horizonte, Brazil – *BSc in Computer Engineering*

Ranked 2nd in the class. Advisors: Prof. Dr. Rodrigo Gaiba de Oliveira, Prof. Dr. Bruno R. Silva.

February 2007 - November 2008

Colégio Técnico Padre de Man, Ipatinga, Brazil – *IT technician*

Ranked 1st in the class. Technical Course in Information Technology focused on software development.

February 2006 - June 2007

Acesita, Timóteo, Brazil – *Industrial Apprentice*

Industrial Apprentice and Trainee on Electrical and Electronic courses.

Work Experience

10/03/2022 - Now

Alexa IA, Amazon, Seattle, United States – *Applied Scientist II*

- Research on large language models (LLMs) focusing on question answering, dialog modeling, controllable text generation and summarization.

- Research focuses on multimodal neural methods for natural language generation and text summarization employing images and structured data.

11/15/2018 - 09/30/2022

UKP Lab, Technische Universität Darmstadt, Darmstadt, Germany *– Doctoral Researcher*

- Research focuses on neural graph methods for natural language generation (e.g., text generation and summarization using graph-based data). In particular, the research projects contemplate graph-based inputs and propose different approaches to better represent graph structures for text generation.
- Supervised bachelor and master theses.
- Candidate evaluation for PhD- and PostDoc-level positions.
- Responsible for a research group focused on Question Answering.

06/01/2021 - 10/01/2021

Alexa IA, Amazon, Seattle, United States (virtually) *– Applied Scientist Intern*

- Worked on improving factuality for neural abstractive summarization. We proposed a deep learning graph-based approach to classify the factual consistency of generated summaries. Advised by Markus Dreyer, Mengwen Liu and Mohit Bansal.

01/19/2015 - 11/09/2018

Labor Court of Rio de Janeiro (TRT1), Rio de Janeiro, Brazil *– IT Specialist*

- Design and development of applications (Web and mobile) for the Brazilian Federal Justice.
- Worked on the development of the Electronic Judicial Process. The system holds 6 million legal proceedings and about 900 million users registered. Technologies used: Java, Javascript, Angular, Oracle, HTML, CSS.

08/19/2013 - 01/16/ 2015

Labor Court of Minas Gerais (TRT3), Belo Horizonte, Brazil *– IT Specialist*

- Lead technical decisions and the team during the agile processes such as prioritization and estimation.
- Worked in different team sizes, tech stacks (backend, frontend, and infrastructure), and in different business contexts.

02/01/ 2013 - 30/06/2013

SGSmap Ingenieros Consultores, San Sebastian, Spain *– Software Developer Intern*

- Development of quality control applications for the production of grapes in vineyards and beverages in wineries.

03/04/2009 - 08/16/2013

Centro Federal de Educação Tecnológica de Minas Gerais, Timóteo, Brazil *– Software Developer*

- Responsible for the design, development, and maintenance of Information Technology solutions, including diverse academic applications for the institution. Technologies used: PHP, Java, Python, MySQL, Linux, Shell Script.
- Lead technical decisions and teams during academic projects of software development using agile methodologies.

01/10/2011 - 12/09/2011

Colégio Lúcia Casa Santa, Timóteo, Brazil – *Technical Instructor*

- Taught Information technology courses at the high school level.
- Courses: programming, algorithms, web development, database management.

07/04/2007 - 03/03/2009

TecWise Sistemas de Automação, Timóteo, Brazil – *Industrial Automation Technician*

- Specification, development and diagnostic on industrial automation systems (sensors, industrial networks, controllers, HMIs, SCADA systems, etc.) using many brands available on market (Rockwell, Siemens, Schneider, etc.).

02/13/2006 - 06/30/2007

Acesita SA (Aperam South America), Timóteo, Brazil – *Industrial Apprentice*

- Industrial Apprentice and Trainee on Electrical and Electronic courses. Courses: electrotechnics, electronics, industrial automation, technical drawing, metrology, computing, metallurgy.
- Developed projects using programmable logic controllers (PLC) – ABB, Siemens, Rockwell models – for automation of industrial platforms.

Publications

Leonardo F. R. Ribeiro, Mohit Bansal, Markus Dreyer. **Generating Summaries with Controllable Readability Levels.** *In Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing, Singapore. Association for Computational Linguistics.*

Georgia Chalvatzaki, Ali Younes, Daljeet Nandha, An Thai Le, Leonardo F. R. Ribeiro, Iryna Gurevych. **Learning to reason over scene graphs: a case study of finetuning GPT-2 into a robot language model for grounded task planning.** *2023 Frontiers in Robotics and AI.*

Tim Baumgärtner, Leonardo F. R. Ribeiro, Nils Reimers, Iryna Gurevych. **Incorporating Relevance Feedback for Information-Seeking Retrieval using Few-Shot Document Re-Ranking.** *In Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.*

Tim Baumgärtner, Kexin Wang, Rachneet Sachdeva, Gregor Geigle, Max Eichler, Clifton Poth, Hannah Sterz, Haritz Puerto, Leonardo F. R. Ribeiro, Jonas Pfeiffer, Nils Reimers, Gözde Şahin, and Iryna Gurevych. 2022. **UKP-SQUARE: An Online Platform for Question Answering Research.** *In Proceedings of the 60th Annual Meeting of the Association for Computational*

Linguistics: System Demonstrations, pages 9–22, Dublin, Ireland. Association for Computational Linguistics. Leonardo F. R. Ribeiro, Mengwen Liu, Iryna Gurevych, Markus Dreyer, and Mohit Bansal. 2022.

Sebastian Gehrmann, Abhik Bhattacharjee, Abinaya Mahendiran, Alex Wang, Alexandros Papangelis, Aman Madaan, Angelina McMillan-Major, Anna Shvets, Ashish Upadhyay, Bingsheng Yao, Bryan Wilie, Chandra Bhagavatula, Chaobin You, Craig Thomson, Cristina Garbacea, Dakuo Wang, Daniel Deutsch, Deyi Xiong, Di Jin, Dimitra Gkatzia, Dragomir Radev, Elizabeth Clark, Esin Durmus, Faisal Ladhak, Filip Ginter, Genta Indra Winata, Hendrik Strobelt, Hiroaki Hayashi, Jekaterina Novikova, Jenna Kanerva, Jenny Chim, Jiawei Zhou, Jordan Clive, Joshua Maynez, João Sedoc, Juraj Juraska, Kaustubh Dhole, Khyathi Raghavi Chandu, Leonardo FR Ribeiro, Lewis Tunstall, Li Zhang, Mahima Pushkarna, Mathias Creutz, Michael White, Mihir Sanjay Kale, Moussa Kamal Eddine, Nico Daheim, Nishant Subramani, Ondrej Dusek, Paul Pu Liang, Pawan Sasanka Ammanamanchi, Qi Zhu, Ratish Puduppully, Reno Kriz, Rifat Shahriyar, Ronald Cardenas, Saad Mahamood, Salomey Osei. **GEMv2: Multilingual NLG Benchmarking in a Single Line of Code.** *In Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing: System Demonstrations, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.*

Leonardo F. R. Ribeiro, Mengwen Liu, Iryna Gurevych, Markus Dreyer, Mohit Bansal. 2022. **FactGraph: Evaluating Factuality in Summarization with Semantic Graph Representations**. *In Proceedings of the 2022 Annual Conference of the North American Chapter of the Association for Computational Linguistics.*

Rachneet Sachdeva, Haritz Puerto Tim Baumgärtner, Sewin Tariverdian, Hao Zhang, Kexin Wang, Hossain Shaikh Saadi, Leonardo FR Ribeiro, Iryna Gurevych. **UKP-SQuARE v2 Explainability and Adversarial Attacks for Trustworthy QA.** *The 2nd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 12th International Joint Conference on Natural Language Processing: System Demonstrations. Association for Computational Linguistics.*

Tim Baumgärtner, Kexin Wang, Rachneet Sachdeva, Max Eichler, Gregor Geigle, Clifton Poth, Hannah Sterz, Haritz Puerto, Leonardo FR Ribeiro, Jonas Pfeiffer, Nils Reimers, Gözde Gül Şahin, Iryna Gurevych. 2022. **UKP-SQUARE: An Online Platform for Question Answering Research**. *In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics, Demo Track.*

Leonardo F. R. Ribeiro, Yue Zhang and Iryna Gurevych. 2021. **Structural Adapters in Pretrained Language Models for AMR-to-text Generation**. *In Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP).*

Leonardo F. R. Ribeiro, Jonas Pfeiffer, Yue Zhang and Iryna Gurevych. 2021. **Smelting Gold and Silver for Improved Multilingual AMR-to-Text Generation**. *In Proceedings of the 2021 Conference*

on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP).

Leonardo F. R. Ribeiro, Martin Schmitt, Hinrich Schütze and Iryna Gurevych. 2021. **Investigating Pretrained Language Models for Graph-to-Text Generation**. *In Proceedings of the 3rd Workshop on NLP for ConvAI. 2021 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP).*

Mohsen Mesgar, Leonardo F. R. Ribeiro, Iryna Gurevych. 2021. **A Neural Graph-based Local Coherence Model**. *In Proceedings of the Findings of the Association for Computational Linguistics: EMNLP 2021. Association for Computational Linguistics.*

Martin Schmitt, Leonardo F. R. Ribeiro, Philipp Dufter, Iryna Gurevych, Hinrich Schütze. 2021. **Modeling Graph Structure via Relative Position for Text Generation from Knowledge Graphs**. *In Proceedings of the TextGraphs-15 Workshop: Graph-based Methods for Natural Language Processing. 2021 Annual Conference of the North American Chapter of the Association for Computational Linguistics.*

Leonardo F. R. Ribeiro, Yue Zhang, Claire Gardent and Iryna Gurevych. 2020. **Modeling Global and Local Node Contexts for Text Generation from Knowledge Graphs**. *Transactions of the Association for Computational Linguistics (TACL), 2020.*

Anne Lauscher, Olga Majewska, Leonardo F. R. Ribeiro, Iryna Gurevych, Nikolai Rozanov, Goran Glavas. 2020. **Common Sense or World Knowledge? Investigating Adapter-Based Knowledge Injection into Pretrained Transformers**. *In Proceedings of Deep Learning Inside Out (DeeLIO): The First Workshop on Knowledge Extraction and Integration for Deep Learning Architectures, Association for Computational Linguistics.*

Leonardo F. R. Ribeiro, Claire Gardent and Iryna Gurevych. 2019. **Enhancing AMR-to-Text Generation with Dual Graph Representations**. *In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP).*

Leonardo F. R. Ribeiro, Claire Gardent and Iryna Gurevych. 2019. **Learning Dual Graph Representations for AMR-to-Text Generation**. *The First International Workshop on Deep Learning on Graphs: Methods and Applications (DLG'19). 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.*

Tobias Falke, Leonardo F. R. Ribeiro, Prasetya Utama, Ido Dagan and Iryna Gurevych. 2019. **Ranking Generated Summaries by Correctness: An Interesting but Challenging Application for Natural Language Inference**. *In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (ACL), 2019.*

Leonardo F. R. Ribeiro, Daniel R. Figueiredo. 2018. **Performance of Monte Carlo Tree Search Algorithms when Playing the Game Ataxx**. *In Proceedings of the 15th National Meeting on Artificial and Computational Intelligence. Brazil.*

Leonardo F. R. Ribeiro, Pedro H. P. Saverese, Daniel R. Figueiredo. 2017. **struc2vec: Learning Node Representations from Structural Identity**. *In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.*

Leonardo F. R. Ribeiro, Daniel R. Figueiredo. 2017. **Ranking lawyers using a social network induced by legal cases**. *Journal of the Brazilian Computer Society.*

Invited Talks

- **Improving Graph-to-Text Generation with Neural Graph Encoders and Pretrained Language Models.** May 2022, Deep Learning on Graphs for Natural Language Processing Workshop (DLG4NLP@ICLR'22).
- **Injecting structure into pretrained language models.** January 2022, The IST and Unbabel Seminar, Lisbon, Portugal.
- **Structural Adapters in Pretrained Language Models for AMR-to-text Generation.** November 2021, UKP-CIS Research Workshop, Virtual.
- **Graph-to-Text Generation with Pretrained Language Models.** October 2021, Huawei Intelligent Collaboration Workshop 2021, Leuven, Belgium.
- **Graph-to-Text Generation with Graph Neural Encoders and Pretrained Language Models.** November 2020, School of Electrical and Computer Engineering (FEEC), University of Campinas, Brazil.
- **Ranking Generated Summaries by Correctness.** June 2019, Natural Language Processing (NLP) Lab, Bar-Ilan University, Ramat Gan, Israel.

Awards and Scholarships

- **2019.** Honorable Mention on the DATAH Master Dissertation Artificial Intelligence Contest in Brazil.
- **2018.** Best Master's Thesis Award. 11th Thesis and Dissertation on Artificial and Computational Intelligence Contest in Brazil.
- **2012-2013.** Science Without Borders scholarship from the Brazilian government to study one year at the Universidad del País Vasco, Spain.
- **2016.** Honorable Mention for paper in Brazilian Workshop on Social Network Analysis and Mining (BraSNAM).
- **2014.** Graduating Student Award - Computer Engineering — 2nd Place.
- **2008.** Graduating class for Technical Degree of Informatics — 1st Place.

Advising

MSc student supervision:

- **2020.** Tobias Andreas Semmelmann. Thesis: Injecting Graph Structure into Pretrained Language Models to Improve AMR-to-text Generation.
- **2019.** Alina Dallmann. Thesis: Employing Graph Neural Networks for Text Production from Knowledge Graphs.

BSc student supervision:

- **2021.** Lennart Mischnaewski. Thesis: Reasoning over Knowledge Graphs for Generative Question Answering.

Editorial Activities

Conferences:

- **2024. Tutorial Chair** - The 62nd Annual Meeting of the Association for Computational Linguistics (ACL 2024)
- **2023. Area Chair (Summarization, Machine Learning for NLP [ARR Action Editor])** - The 18th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2024)
- **2022. Area Chair** - The 29th International Conference on Computational Linguistics (COLING 2022).

Reviewing Activities

Conferences:

[Each listed year is a reviewing load of 3-5 papers.]

- **2024.** Thirty-Seventh AAAI Conference on Artificial Intelligence (2024).
- **2023.** The 61th Annual Meeting of the Association for Computational Linguistics, Toronto, Canada.
- **2023.** 2023 Annual Conference of the North American Chapter of the Association for Computational Linguistics, Virtual.
- **2023.** Thirty-Seventh AAAI Conference on Artificial Intelligence (2023).
- **2023.** Brazilian Conference on Intelligent Systems.
- **2022.** The 2022 Conference on Empirical Methods in Natural Language Processing.
- **2022.** 36th AAAI Conference on Artificial Intelligence, Vancouver, Canada.
- **2021-2023.** ACL (Association for Computational Linguistics) Rolling Review.
- **2022.** Brazilian Conference on Intelligent Systems.
- **2021.** Brazilian Conference on Intelligent Systems, Virtual.
- **2021.** The 2021 Conference on Empirical Methods in Natural Language Processing, Punta Cana, Dominican Republic.
- **2021.** The 59th Annual Meeting of the Association for Computational Linguistics, Virtual.
- **2021.** 2021 Annual Conference of the North American Chapter of the Association for Computational Linguistics, Virtual.

- **2021.** 35th AAAI Conference on Artificial Intelligence, Virtual.
- **2020.** The 2020 Conference on Empirical Methods in Natural Language Processing, Virtual.
- **2020.** Brazilian Conference on Intelligent Systems, Rio Grande, Brazil.

Journals:

[Each listed year is a reviewing load of 1-3 papers.]

- **2021.** Journal of Machine Learning Research.
- **2021.** Language Resources and Evaluation (LREV).

Workshops:

[Each listed year is a reviewing load of 1-3 papers.]

- **2023.** Generation, Evaluation & Metrics (GEM) Workshop.
- **2023.** The 4rd Workshop on Evaluation & Comparison of NLP Systems.
- **2023.** Workshop on Graph Learning Benchmarks (GLB).
- **2022.** Generation, Evaluation & Metrics (GEM) Workshop.
- **2022.** The 3rd Workshop on Evaluation & Comparison of NLP Systems.
- **2022.** Third Workshop on Knowledge Extraction and Integration for Deep Learning Architectures.
- **2022.** Workshop on Graph Learning Benchmarks (GLB).
- **2022.** Workshop on Graph-Based Natural Language Processing (TextGraphs).
- **2021.** 15th Workshop on Graph-Based Natural Language Processing, 2021 Annual Conference of the North American Chapter of the Association for Computational Linguistics, Virtual.
- **2021.** Evaluation and Comparison of NLP Systems (Eval4NLP), The 2021 Conference on Empirical Methods in Natural Language Processing, Punta Cana, Dominican Republic.
- **2020.** Evaluation and Comparison of NLP Systems (Eval4NLP), The 2020 Conference on Empirical Methods in Natural Language Processing, Virtual.
- **2020.** 14th Workshop on Graph-Based Natural Language Processing, The 28th International Conference on Computational Linguistics, Barcelona.
- **2020.** Third Workshop on Multilingual Surface Realization, The 28th International Conference on Computational Linguistics, Barcelona.
- **2019.** Second Workshop on Multilingual Surface Realization, The Conference on Empirical Methods in Natural Language Processing, Hong Kong.

Skills

Languages: Portuguese (native), English (fluent), Spanish (advanced), German (basic).