Maria Leonor Pacheco

maria.pacheco@colorado.edu
https://mlpacheco.github.io
https://www.linkedin.com/in/mlpacheco/

Education

2022 **Ph.D. Computer Science,** Purdue University.
Thesis: *Neural-Symbolic Modeling for Natural Language Discourse*

2013 **B.Sc. Computer Science and Engineering,** Universidad Simon Bolivar.
Honors thesis: *Information Extraction from Twitter During Mass Emergency Situations.*

Employment History

Jul 2022 – **Visiting Professor,** University of Colorado Boulder. Remote, USA.

Jul 2022 – **Postdoctoral Researcher,** Microsoft Research. NY, USA.

Jan 2017 – May 2022 **Graduate Research Assistant,** Natural Language Processing Group, Dept. of Computer Science, Purdue University. IN, USA.

May 2021 – Aug 2021 **Research Intern,** Microsoft Research. WA, USA.

May 2019 – Jul 2019 **Research Intern,** Conversational Search Group at Bing, Microsoft. WA, USA.

Jun 2018 – Aug 2018 **Research Intern,** Twitter Cortex (ML group) at Twitter. CA, USA.

Aug 2015 – Dec 2016 **Graduate Teaching Assistant,** Dept. of Computer Science, Purdue University.

Nov 2013 – Jul 2015 **Data Scientist,** Predictvia. Caracas, Venezuela.

Sep 2011 – Jul 2012 **Visiting Student Researcher,** Knowledge Systems Lab, Information and Management Systems Eng. Dept., Nagaoka University of Technology. Niigata, Japan.

Publications (* Indicates co-first authorship. The fields I publish in are conference-driven.)

Journal Articles, Conference and Workshop Proceedings

1. **Pacheco**, **M. L.**, Roy, S., & Goldwasser, D. (2022). Hands-on interactive neuro-symbolic NLP with DRaiL. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing: System Demonstrations (to appear).*

2. **Pacheco**, **M. L.**, Islam, T., Mahajan, M., Shor, A., Yin, M., Ungar, L., & Goldwasser, D. (2022). A holistic framework for analyzing the COVID-19 vaccine debate. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2022* (pp. 5821–5839). doi:10.18653/v1/2022.naacl-main.427

3. Mehta, N., **Pacheco**, **M. L.**, & Goldwasser, D. (2022). Tackling fake news detection by continually improving social context representations using graph neural networks. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics, ACL 2022* (pp. 1363–1380). doi:10.18653/v1/2022.acl-long.97

4. **Pacheco**, **M. L.**, von Hippel, M., Weintraub, B., Goldwasser, D., & Nita-Rotaru, C. (2022). Automated attack synthesis by extracting finite state machines from protocol specification documents. In *43rd IEEE Symposium on Security and Privacy, SP 2022.* doi:10.1109/SP46214.2022.9833673

5. Roy, S., **Pacheco**, **M. L.**, & Goldwasser, D. (2021). Identifying morality frames in political tweets using relational learning. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, EMNLP 2021* (pp. 9939–9958). doi:10.18653/v1/2021.emnlp-main.783

6. Lee, I.-T., **Pacheco**, **M. L.**, & Goldwasser, D. (2021). Modeling human mental states with an entity-based narrative graph. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2021* (pp. 4916–4926). doi:10.18653/v1/2021.naacl-main.391

7. Widmoser, M. *., **Pacheco, M. L. ***, Honorio, J., & Goldwasser, D. (2021). Randomized deep structured prediction for discourse-level processing. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main volume, EACL 2021* (pp. 1174–1184). doi:10.18653/v1/2021.eacl-main.100

8. **Pacheco**, **M. L.**, & Goldwasser, D. (2021). Modeling content and context with deep relational learning. *Transactions of the Association for Computational Linguistics, TACL, 9*, 100–119. Presented at NAACL-HLT 2021. doi:10.1162/tacl_a_00357

9. Lee, I.-T., **Pacheco**, **M. L.**, & Goldwasser, D. (2020). Weakly-supervised modeling of contextualized event embedding for discourse relations. In *Findings of the Association for Computational Linguistics: EMNLP 2020* (pp. 4962–4972). doi:10.18653/v1/2020.findings-emnlp.446

10. Jain, A., **Pacheco**, **M. L.**, Lancette, S., Goindani, M., & Goldwasser, D. (2020). Identifying collaborative conversations using latent discourse behaviors. In *Proceedings of the 21th annual meeting of the Special Interest Group on Discourse and Dialogue, SIGDIAL 2020* (pp. 74–78). Retrieved from https://www.aclweb.org/anthology/2020.sigdial-1.10

11. Jero, S., **Pacheco**, **M. L.**, Goldwasser, D., & Nita-Rotaru, C. (2019). Leveraging textual specifications for grammar-based fuzzing of network protocols. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 33, pp. 9478–9483). doi:10.1609/aaai.v33i01.33019478

12. Lee, I.-T., Goindani, M., Li, C., Jin, D., Johnson, K. M., Zhang, X., … Goldwasser, D. (2017). PurdueNLP at SemEval-2017 task 1: Predicting semantic textual similarity with paraphrase and event embeddings. In *Proceedings of the 11th international workshop on Semantic Evaluation (SemEval-2017)* (pp. 198–202). doi:10.18653/v1/S17-2029

13. Zhang, X. *., **Pacheco, M. L. ***, Li, C., & Goldwasser, D. (2016). Introducing DRAIL – a step towards declarative deep relational learning. In *Proceedings of the workshop on Structured Prediction for NLP* (pp. 54–62). doi:10.18653/v1/W16-5906

14. **Pacheco**, **M. L.**, Lee, I.-T., Zhang, X., Zehady, A. K., Daga, P., Jin, D., … Goldwasser, D. (2016). Adapting event embedding for implicit discourse relation recognition. In *Proceedings of the CoNLL-16 shared task* (pp. 136–142). doi:10.18653/v1/K16-2019

Refereed Workshops without Published Proceedings

1. **Pacheco**, **M. L.**, Islam, T., Ungar, L., Yin, M., & Goldwasser, D. (2022). An interactive framework for identifying latent themes in large text collections. *NeurIPS 2022 Workshop on Interactive Learning for Natural Language Processing*.

2. **Pacheco**, **M. L.**, & Goldwasser, D. (2020). Neural-symbolic modeling for natural language discourse. *ICML 2020 Workshop on Bridge Between Perception and Reasoning*. Retrieved from https://logicalreasoninggnn.github.io/papers/16.pdf

3. Widmoser, M. *., **Pacheco, M. L. ***, Honorio, J., & Goldwasser, D. (2020). Randomized deep structured prediction for argumentation mining. *EMNLP 2020 workshop on Structured Prediction for NLP*.

4. **Pacheco**, **M. L.**, Dalal, I., & Goldwasser, D. (2018). Leveraging representation and inference through deep relational learning. *NeurIPS 2018 Workshop on Relational Representation Learning*. Retrieved from https://r2learning.github.io/assets/papers/CameraReadySubmission%2042.pdf

Invited Talks, Seminars and Symposia

- **Neuro-Symbolic Discourse Processing.** Invited talk at Universidad de Puerto Rico Mayaguez.
- **Neural-Symbolic Modeling for NLP.** Keynote at LatinX in NLP Workshop, NAACL 2022.
- **Automated Attack Synthesis by Extracting Finite State Machines from Protocol Specification Documents.** Microsoft Research Redmond (Mar 2022), MIT CSAIL (Apr 2022), Trail of Bits (Apr 2022), co-presented with Max von Hippel.
- **Neural-Symbolic Modeling for Natural Language Discourse.** Microsoft Research Redmond (Nov 29, 2021), Microsoft Research NYC (Feb 18, 2022).
- **Reasoning About Entities and Events with Narrative Graphs.** University of North Carolina at Chapel Hill (Oct. 2, 2021)
- **Modeling Content and Context with Deep Relational Learning.** Boston Computation Club (June 13, 2021), Orlando Machine Learning and Data Science Meetup (March 20, 2021), Microsoft Research Redmond (Jun. 17, 2021), 1st annual retreat, Purdue Center for Programming Principles and Software Systems (Aug. 31, 2020)
- **DRaiL: Towards Declarative Deep Relational Learning.** Midwest Speech and Language Days / Computational Linguistics Colloquium, Toyota Technological Institute at Chicago (May 4-5, 2017)
- **Improving Protocol Vulnerability Discovery via Semantic Interpretation of Textual Specifications.** The 18th Annual CERIAS Security Symposium, Purdue University (Apr. 18-19, 2017)
- **Event Embeddings for Implicit Discourse Relations.** Universidad Simon Bolivar, Caracas, Venezuela (May 17, 2016)

Tutorials

- **NS4NLP: Neuro-Symbolic Methods for NLP *(extended version)*** Tutorial at the 29th International Conference in Computational Linguistics, COLING-22 (Oct. 17, 2021) taught with Sean Welleck, Dan Goldwasser, Vivek Srikumar, Yejin Choi and Dan Roth.
- **NS4NLP: Neural-Symbolic Modeling for NLP.** Tutorial at the 30th International Joint Conference on Artificial Intelligence, IJCAI-21 (Aug. 20, 2021) taught with Dan Goldwasser.

Teaching Experience

Purdue University

- **Instructor.** Problem Solving and Object Oriented Programming (2 semesters). Taught lectures of 50-100+ students, developed assignments and exams, and managed a team of 20+ graduate and undergraduate TAs.
- **Teaching Assistant.** Problem Solving and Object Oriented Programming (2 semesters). Guided lab sections of 10-20 students, held office hours, developed and graded assignments.

Awards and Scholarships

- **2021 Microsoft Research Dissertation Grant.** $25,000 towards dissertation expenses.
- **2017 Teaching Academy Graduate Teaching Award.** Selected by the CS Dept. at Purdue.
- **2011-2012 JASSO Student Exchange Support Program.** Scholarship for short-term study in Japan.

Academic and Institutional Service

Senior PC Member

- AAAI 2023

Academic and Institutional Service (continued)

Reviewer	ACL Rolling Review (2021-present), AAAI (2021, 2023), EMNLP (2021, 2022), NAACL (2021), ACL (2021), ACM Computing Surveys (2020-2021), LatinX in AI / Queer in AI Workshops (2022)
Organizing	**Presentation Chair**, LatinX in NLP Workshop, co-located with NAACL 2022
	Co-Chair, NAACL 2022 Queer in AI Workshop
	Co-Chair, NAACL 2022 Student Research Workshop
	Social Media Co-Chair, SIGDIAL 2021
Diversity & Inclusion	**Leadership Team**, Purdue Women in Science Programs (2021 - 2022)
	Board Member (Diversity), Purdue CS Graduate Student Board (2021 - 2022)
	Diversity Chair, Purdue CS Graduate Student Board (2020 - 2021)
	Committee Member, Purdue CS Dept. DEI Committee (2020 - 2022)
	Global Ambassador, Purdue Graduate School (2018 - 2022)

Skills

Languages	Spanish (Native), English (Fluent), Japanese (N3), German (A2).
Coding	Python, Java, C, C++, R, Matlab, Git, SQL, LaTeX, …