Andrew Silva

andrew-silva.com
andrew.silva@gatech.edu
https://www.linkedin.com/in/andrew-silva-6a29026a/
https://scholar.google.com/citations?user=6wuXl_cAAAAJ&hl=en

RESEARCH OBJECTIVES

My research interests on advancing **personalized, interactive learning** with an emphasis on **explainability** to enable effective two-way interaction with digital (or embodied) agents that feel personal and unique.

In 2021 I was named an **Apple Scholar** and a **DAAD AInet Fellow**.

EDUCATION

Georgia Institute of Technology, Atlanta, GA - *Ph. D in Computer Science*

August 2019 - August 2023

I completed my Computer Science Ph.D. in the School of Interactive Computing, working in the CORE Robotics lab under Dr. Matthew Gombolay.

Georgia Institute of Technology, Atlanta, GA - *Bachelor's / Master's*

August 2010 - May 2014 / January 2015 - May 2017

I completed my Bachelor's of Science in Computational Media with a certificate in Business and Entrepreneurship, and my Master's in Computer Science at Georgia Tech with a specialization in Computational Perception and Robotics.

RESEARCH EXPERIENCE

Graduate Research Assistant - *Georgia Institute of Technology*

August 2019 - August 2023

- Serve as VP of Academics within the graduate robotics organization, RoboGrads
- Developed a new approach to adaptive personalization for explanations and verified in a user-study set in a self-driving car simulator domain.
- Developed new approach for personalizing LLMs to match speaker styles (EACL 2023)
- Introduced new influence-function view for explaining neural networks (AISTATS 2022)
- Studied humans and their reactions to different types of explainable AI (IJHCI 2022)
- Introduced new approach to multi-task learning with language (RAL-ICRA 2022)
- Measured and reported societal biases in pre-trained LLMs (NAACL 2021)
- Introduced new mechanism to inject domain expertise into neural networks (AAAI 2021)
- Developed approach to convert a neural network into a decision-tree (AISTATS 2020)
- Enabled unsupervised role discovery in multi-agent systems (AAMAS 2019)

- Developed computer vision system for robots to estimate human workload (THRI 2018)
- Applied deep networks to creating video embeddings for zero-shot learning

Research Areas: reinforcement learning, interpretability, human-robot interaction, interactivity, computer vision, user studies, supervised learning, bias and fairness, AI ethics, robotics

Research Intern - *Google - Brain Robotics*

May 2022 - September 2022

- Worked on robust robot policies operating at high frequency (hosted by Pannag Sanketi and Avi Singh)
- Implemented standardized evaluation protocol to measure performance of diverse policies along standardized axes
- Developed multi-style robot policies using decision transformers for agile robots to adapt to fit diverse expert datasets

Research Areas: task learning, agile robotics, multi-style learning, decision transformers

Research Intern - *Apple, Inc.*

June 2020 - September 2020 || May 2021 - September 2021 || October 2022 - May 2023

- Worked in Interactive Intelligence (hosted by Nick Apostoloff and Barry Theobald)
- Developed (1) multi-modal punctuation prediction model for unstructured ASR, (2) a new sub-group approach to private personalized federated learning, and (3) a new approach for interactive personalization with large language models (LLMs)
- Drastically reduced overfitting of LLMs-as-reward-models and greatly increasing multi-task & personalization efficiency through new soft-prompted approach.
- Project (1) resulted in a publication at ICASSP 2021

Research Areas: natural language, multimodal learning, federated learning, personalization

Research Scientist I - *Georgia Institute of Technology*

May 2019 - August 2019

- Robot Autonomy and Interactive Learning lab (PI: Dr. Sonia Chernova)
- Cognitive Optimization and Relational Robotics lab (PI: Dr. Matthew Gombolay)
- Published ROS packages for real time face and object detection with deep networks
- Conducted three user studies to evaluate the performance of experimental algorithms

Research Areas: artificial intelligence, interactive learning, robotics, interpretability

TEACHING EXPERIENCE

Natural Language Processing - *Graduate Teaching Assistant*

Fall 2021 || **Enrollment**: 90/Semester

Head TA, hosted office hours, created and graded homework assignments, responded to students in an online forum, prepared and delivered three lectures

Interactive Robot Learning - *Graduate Teaching Assistant*

Spring 2021 || **Enrollment**: 75/Semester

Head TA, hosted office hours, created and graded homework assignments, responded to students in an online forum, prepared and delivered one lecture

Introduction to Computer Vision - *Graduate Teaching Assistant*

August 2016 - September 2017 || **Enrollment**: 240/Semester

Hosted office hours, graded assignments, and responded to students in an online forum

INDUSTRY EXPERIENCE

Applications Developer - *AT&T*

June 2014 - December 2015

- Author on Patent US10189479B2 for monitoring vehicle driver status
- Pitched and developed various web and Android projects
- Oversaw the Atlanta internship program, helping to manage 40 Summer interns

Skills: Android development, web development, leadership, management, product-pitching

PUBLICATIONS

Journal Papers

- **Andrew Silva**, Mariah Schrum, Erin Hedlund-Botti, Nakul Gopalan, and Matthew Gombolay. “Explainable Artificial Intelligence: Evaluating the Objective and Subjective Impacts of xAI on Human-Agent Interaction.” in International Journal of Human-Computer Interaction (IJHCI) (2022).
- **Andrew Silva**, Nina Moorman, William Silva, Zulfiquar Zaidi, Nakul Gopalan, and Matthew Gombolay. “LanCon-Learn: Learning with Language to Enable Generalization in Multi-Task Manipulation Domains.” in IEEE Robotics and Automation Letters (2022).
- Banerjee, Siddhartha, **Andrew Silva**, and Sonia Chernova. "Robot Classification of Human Interruptibility and a Study of Its Effects." ACM Transactions on Human-Robot Interaction (THRI) 7.2 (2018): 14.

Conference Papers

- **Silva, Andrew***, Pradyumna Tambwekar*, and Matthew Gombolay. “FedPerC: Federated Learning for Language Generation with Personal and Context Preference Embeddings.” In Findings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (EACL) 2023.
- **Silva, Andrew**, Rohit Chopra, and Matthew Gombolay. "Cross-Loss Influence Functions to Explain Deep Network Representations." International Conference on Artificial Intelligence and Statistics. PMLR, 2022.
- **Silva, Andrew,** Katherine Metcalf, Nicholas Apostoloff, and Barry-John Theobald. "Fedembed: Personalized private federated learning." arXiv preprint arXiv:2202.09472 (2022).

- **Silva, Andrew**, and Matthew Gombolay. "Encoding Human Experts' Domain Knowledge to Warm Start Reinforcement Learning." In AAAI Conference on Artificial Intelligence (AAAI) 2021.
- **Silva, Andrew**, Barry-John Theobald, and Nicholas Apostoloff. "Multimodal Punctuation Prediction with Contextual Dropout." ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2021.
- **Silva, Andrew***, Tambwekar, Pradyumna*, and Matthew Gombolay. "Towards a Comprehensive Understanding and Accurate Evaluation of Societal Biases in Pre-Trained Transformers." Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2021.
- **Silva, Andrew,** Matthew Gombolay, Taylor Killian, Ivan Jimenez, and Sung-Hyun Son. "Optimization methods for interpretable differentiable decision trees applied to reinforcement learning." In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, pp. 1855-1865. 2020.
- Paleja, Rohan, **Andrew Silva**, Letian Chen, and Matthew Gombolay. "Personalized Apprenticeship Learning from Heterogeneous Decision-Makers." Advances in Neural Information Processing Systems (NeurIPS), 2020
- **Silva, Andrew**, and Sonia Chernova. "Unsupervised Role Discovery Using Temporal Observations of Agents." Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems. International Foundation for Autonomous Agents and Multiagent Systems (AAMAS), 2019.
- Seraj, Esmaeil, **Andrew Silva**, and Matthew Gombolay. "Safe Coordination of Human-Robot Firefighting Teams." arXiv preprint arXiv:1903.06847 (2020)
- Hahn, Meera, **Andrew Silva**, and James M. Rehg. "Action2Vec: A Crossmodal Embedding Approach to Action Learning." arXiv preprint arXiv:1901.00484 (2019).
- Banerjee, Siddhartha*, **Silva, Andrew***, Feigh, Karen, & Chernova, Sonia. “Effects of interruptibility-aware robot behavior.” arXiv preprint arXiv:1804.06383 (2018).

Workshop Papers

- **Silva, Andrew***, Banerjee, Siddhartha*, Chernova, Sonia (2018) “Excuse Me, Could You Please Assemble These Blocks For Me?” Presented at the What Could Go Wrong? Workshop at Human-Robot Interaction (HRI 2018)
- Pradyumna Tambwekar, **Andrew Silva,** Matthew Gombolay (2023) “The Design and Preliminary Results of a User Study Measuring Diverse Explainability Preference” Presented at the Lifelong Learning and Personalization in Long-Term Human-Robot Interaction (LEAP-HRI) Workshop at Human-Robot Interaction (HRI 2023)

Reviewer

Association for the Advancement of Artificial Intelligence Conference (AAAI)

International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)

International Conference on Artificial Intelligence and Statistics (AISTATS)

Empirical Methods in Natural Language Processing (EMNLP)

ACM/IEEE International Conference on Human-Robot Interaction (HRI)

International Conference on Machine Learning (ICML)

Neural Information Processing Systems (NeurIPS)

Robotics: Science and Systems (RSS)

International Conference on Learning Representations (ICLR)

Association for Computational Linguistics (ACL)