STEVENS
INSTITUTE OF TECHNOLOGY
1870

PH.D. DISSERTATION DEFENSE

Candidate:	João Luís Lins
Degree:	Doctor of Philosophy
School/ Department:	Charles V. Schaefer, Jr. School of Engineering and Science / Department of Computer Science
Date:	Tuesday, November 18th, 2025
Time/Location:	11:00 AM EST / Zoom: http://stevens.zoom.us/my/xujia
Title:	Reinforcement Learning with Supervised Alignment for Grounded Truth
Chairperson:	Dr. Jia Xu, Department of Computer Science, Charles V. Schaefer, Jr. School of Engineering and Science
Committee Members:	Dr. Philippos Mordohai, Department of Computer Science, Charles V. Schaefer, Jr. School of Engineering and Science
	Dr. Jonggi Hong, Department of Computer Science, Charles V. Schaefer, Jr. School of Engineering and Science
	Dr. Abdul Rafae Khan, Department of Computer Science, School of Information Technology, Monash University

ABSTRACT

Truth is the foundation of trustworthy, ethical, and competitive AI. Without a factual basis, large language models (LLMs) may appear fluent while thinking erroneously, mimicking reasoning yet inherently misunderstanding causality, thereby growing a gray box where misconduct flourishes. Recent advances like RLHF and RLAIF strive to align LLMs with human intent but falter when labels are absent, biased, or unverifiable.

My thesis first introduces Reinforcement Learning with Supervised Alignment (RLA), which incorporates fact-based supervision into reward modeling to formally verify assumptions. This method outperforms RLAIF by up to 131% and boosts LLaMA3 performance by 50% in-domain and 16% out-of-domain. While slightly behind SFT in-domain, RLA generalizes far better, surpassing SFT by up to 50× on cross- and out-of-domain tasks, excelling at corroborating the truth.

The second part of my thesis focuses on Never Alone with Me (NAM), a socially intelligent chatbot system where I served as team lead. NAM integrates long-term memory, an animated avatar interface, and multimodal capabilities to foster deep, sustained engagement with users. These innovations tripled average conversation

duration, one of the key evaluation metrics. Together with my team members and advisor, we earned 2nd place in the 2023 Alexa Socialbot Grand Challenge 5, hosted by Amazon. To our knowledge, this marks the first time in history that Stevens Institute of Technology has earned a top award in distinguished AI competitions.