Amelia Francesca Hardy

(707) 363-7008 | ahardy@stanford.edu|website|Linkedin

EDUCATION

Stanford University

Undergraduate (September 2019): Computer Science (GPA: 3.9/4.0)
Master's (December 2024): Computer Science (GPA: 4.0/4.0)

Selected Coursework:

Undergraduate Level: From Languages to Information (CS 124), Databases (CS145), Automata and Complexity Theory (CS154), Logic and Automated Reasoning (CS 157), The Modern Algorithmic Toolbox (CS 168), Matrix Methods (CME 103), Scientific Computing (CME 108)

Graduate Level: Natural Language Understanding (CS 224U), Natural Language Processing with Deep Learning (CS 224N), Mining Massive Datasets (CS 246), Apache Spark and Hadoop Lab (CS 246H), Logic and Artificial Intelligence (CS 247), Computational Genomics (CS 262), Decision Making Under Uncertainty (CS 238), Research and Policy on Postsecondary Education (EDUC 117B), Advanced Topics in Sequential Decision Making (CS 239), Natural Language Processing for Computational Social Science (CS 224C)

PhD Level: Project in Mining Massive Datasets (CS 341), Computational Models of the Neocortex (CS 379C)

WORK EXPERIENCE

Research Assistant - *Stanford Intelligent Systems Laboratory (PI: Mykel Kochenderfer)* 2023-present

- Team lead for project with Airbus, developing LLM-based tools for aviation safety. Developing HCI study of the implications of such tools, with targeted venue ACL 2025.
- Collaborated with Stanford NLP group on research applying techniques from robotics to language modeling, to elicit toxic behaviors which can ultimately be avoided through fine-tuning on these negative elicitations (under submission, NAACL 2025)
- Research project on best practices for AI benchmarking (accepted as Spotlight, NeurIPS 2024)
- Research project on using diffusion models to generate trajectory data which can be used to more efficiently populate DESPOT trees used for online POMDP planning
- Implementing (currently in progress) open-source version of Forward Search planning algorithm, to be included in POMDPs.jl

Contributor - *Stanford AI Index* *2024*

- Contributed to the Stanford AI Index report on the Responsible AI chapter

Research Assistant - *Stanford Center for Research on Foundational Models (PI: Percy Liang)* 2022-2023

- Lead team of PhD and Postdoctoral students studying dialogue benchmarking for LLMs
- Structured experiments, research questions, and metrics for dialogue benchmarking project
- Managed organizational aspects of team, including scheduling and running meetings, assigning tasks, and communicating with team members to ensure completion of research objectives

Research Assistant- *Stanford Alexa Prize Team (2nd place, 2019) (PI: Chris Manning)* 2020-2022

- Selected from competitive applicant pool to build a social chatbot for Amazon's Alexa Prize Competition
- Implemented and deployed ML models to improve bot performance
- Performed statistical analysis on user data to aid in targeting improvements
- Used insights and data collected during the competition as basis of research project on strategies to increase user initiative
- Produced open-source version of the bot, following conclusion of the competition (https://github.com/stanfordnlp/chirpycardinal)
- Served as advisor for 2020 Alexa Prize team, consulting on prioritization and structure of ML models for social dialogue

Machine Learning Engineer - *Woebot Labs (Advised by Dan Jurafsky)* 2019-2020

- Built deep contextual NLP models for crisis, mood, and activity detection (BERT, Snorkel, TensorFlow, Pytorch)
- Managed data labeling team to obtain high-quality, domain-specific datasets (Mturk)
- Full stack machine learning development on AWS and GCP

Researcher - *Stanford Intelligent Systems Laboratory (PI: Mykel Kochenderfer)* 2018-2019

- Survey and implementation of algorithms formally verifying properties of deep neural networks used in applications such as the FAA Airborne Collision Avoidance System and the Airborne Collision Avoidance System-Xu (Julia)

Researcher – *Ré Informatics Lab (PI: Chris Ré)* 2016

- Developed machine vision algorithm for DARPA's Memex project, segmenting and matching tattoos on images from a dataset of dark web advertisements (Pytorch, OpenCV)
- Implemented pipeline for segmentation and matching of these images, allowing individuals to be searched by identifying tattoos and for tattoos to be clustered as potential brands. (Python, MATLAB)

Software Development Intern (Backend) - *Google* (*Nest)* 2016

- Introduced developer-facing distributed tracing interfaces for the optimization of micro-services architecture (Zipkin)
- Worked closely with systems stack engineers to write Scala functions facilitating testing of micro-service architecture (Scala)

Software Development Intern (Backend) – *Quantescent* 2015 – 2016

- Designed financial database schema (SQL)
- Debugged and developed financial indicator software (C)
- Researched and implemented real-time, performant financial indicator algorithms (C++)

PUBLICATIONS

BetterBench: Assessing AI Benchmarks, Uncovering Issues, and Establishing Best Practices
Amelia Hardy*, Anka Reuel*, Chandler Smith, Max Lamparth, Mykel Kochenderfer (NeurIPS 2024 spotlight)

ASTPrompter: Weakly Supervised Automated Language Model Red-Teaming to Identify Likely Toxic Prompts
Amelia Hardy*, Houjun Liu*, Bernard Lange, Mykel Kochenderfer (Under submission, NAACL 2025)

More than Marketing? On the Information Value of AI Benchmarks for Practitioners
Amelia Hardy*, Anka Reuel*, Kiana Jafari Meimandi, Lisa Soder, Allie Griffith, Dylan Asmar, Sanmi Koyejo, Michael Bernstein, Mykel Koochenderfer (Under submission, IUI 2025)

Inferring Traffic Models in Terminal Airspace from Flight Tracks and Procedures
Soyeon Jung, **Amelia Hardy,** Mykel Kochenderfer (Under submission, JAIS)

Effective Social Chatbot Strategies for Increasing User Initiative
Amelia Hardy, Ashwin Paranjape, Christopher D Manning (SIGDIAL 2021)

Evaluating human-language model interaction
Mina Lee, Megha Srivastava, **Amelia Hardy**, John Thickstun, Esin Durmus, Ashwin Paranjape, Ines Gerard-Ursin, Xiang Lisa Li, Faisal Ladhak, Frieda Rong, Rose E Wang, Minae Kwon, Joon Sung Park, Hancheng Cao, Tony Lee, Rishi Bommasani, Michael Bernstein, Percy Liang (TMLR 2023)

Neural generation meets real people: Towards emotionally engaging mixed-initiative conversations
Ashwin Paranjape, Abigail See, Kathleen Kenealy, Haojun Li, **Amelia Hardy**, Peng Qi, Kaushik Ram Sadagopan, Nguyet Minh Phu, Dilara Soylu, Christopher D Manning (Alexa Prize Proceedings 2020)

Neural generation meets real people: Building a social, informative open-domain dialogue agent
Ethan A Chi, Ashwin Paranjape, Abigail See, Caleb Chiam, Trenton Chang, Kathleen Kenealy, Swee Kiat Lim, **Amelia Hardy**, Chetanya Rastogi, Haojun Li, Alexander Iyabor, Yutong He, Hari Sowrirajan, Peng Qi, Kaushik Ram Sadagopan, Nguyet Minh Phu, Dilara Soylu, Jillian Tang, Avanika Narayan, Giovanni Campagna, Christopher D Manning (SIGDIAL 2022)

INVITED TALKS

Developing an LLM-Based Cockpit Assistant, NASA Human Systems Integration August 2024

ASTPrompter: Weakly Supervised Automated Language Model Red-Teaming to Identify Likely Toxic Prompts,
ZEW - Leibniz Center For European Economic Research workshop on Designing Procedures for Red Teaming Generative AI Models September 2024

TEACHING

Head Course Assistant *CS 238 (Decision Making Under Uncertainty)*
Course Assistant *CS 157 (Introduction to Logic)*, *CS 124 (From Languages to Information), CS 227B (General Game Playing), CS 120 (Introduction to AI Safety)*
Course Development *CS 124 (From Languages to Information)*
Course Facilitator X*CS 221 (Artificial Intelligence: Principles and Techniques), XCS 224U (Natural Language Understanding), XCS 224N (Natural Language Processing with Deep Learning), XCS 224W (Machine Learning with Graphs)*

SERVICE

Reviewer for Journal of Aerospace Information Systems 2022, 2023
Reviewer for AAAI Conference on Artificial Intelligence 2024
Organizer for ICLR Workshop on Human-AI Co-Evolution 2025

LEADERSHIP

Nest Women in Engineering Leadership Award 2016

- Elected by women engineers in intern group to receive award for leadership

PROGRAMMING SKILLS

Scientific: Python, Numpy, Scipy, Julia, JuMP, R, MATLAB, LaTeX
Technical: C, C++(11), Java, Scala, Postgres-SQL, Apache Spark, Hadoop, SQL, GitHub, Git, Unix
Machine Learning: Pytorch, TensorFlow, AWS, GCP, Mechanical Turk