PROMISE EKPO

Cornell Tech, New York City, NY | poe6@cornell.edu :| Personal Website | Google Scholar

EDUCATION

Cornell University (Cornell Tech, NYC)

Ph.D. in Computer Science | Advisor: Prof. Angelique Taylor | GPA: 3.9/4.0 *Expected May 2027*

Princeton University

M.Sc. in Computer Science | Advisor: Prof. Jaime Fernández Fisac | GPA: 3.7/4.0. *June 2022*

University Of Benin

B.Eng. in Computer Engineering | GPA: 4.79/5.0 *January 2020*

PUBLICATIONS & THESES

[1] T. Tanjim, **Promise Ekpo**, H. Cao, J. St. George, K. Ching, H. R. Lee, A. Taylor. *"Human-Robot Teaming Field Deployments: A Comparison Between Verbal and Non-verbal Communication."* IEEE RO-MAN Workshop, 2025. *(***Accepted)** *Conducted statistical analysis of field data with 115 healthcare workers)*

[2] **Promise Ekpo**. *"Investigating Persuasiveness in Large Language Models."* M.A. Thesis, Princeton University, 2023 [PDF].

[3] *Yashika Batra,* **Promise Ekpo**, *Giuliano Pioldi, Purnjay Marjur, Arman Ibrayeva, Angelique Taylor, "RFM-HRI: A Multimodal Dataset of Medical Robot Failure, User Reaction, and Recovery Preferences for Item Retrieval Tasks." [PDF]* **(Accepted, THRI Journal, 2026)**

[4] *Integrated artificial intelligence in healthcare and the patient's experience of care. Oluwatosin Ogundare, Tolu Owadokun, Temitope Ogundare, Promise Ekpo, Ha Linh Nguyen & Stephen Bello,* [PDF]. *(***Accepted, Scientific Reports (Nature))**

[5] **Promise Ekpo**, S. Agarwal, F. Grimm, Jiachang Liu, L. Molu, A. Taylor. *"AdaFair-MARL: Enforcing Adaptive Fairness Constraints in Multi-Agent Reinforcement Learning.* **(Submitted, IEEE Conference on Decision and Control (CDC), 2025)**

[6] **Promise Ekpo**, B. La, T. Wiener, S. Agarwal, A. Agrawal, G. Gonzalez-Pumariega, L. Molu, A. Taylor. *"Skill-Aligned Fairness in Multi-Agent Learning for Collaboration in Healthcare."* 2025. [PDF]. *(***Submitted, NeurIPS 2026 Evaluations and Datasets Track, 2026)**

[7] **Promise Ekpo**, A. Taylor, L. Molu. *"A Generalized Nash Equilibrium-Seeking Scheme for Trauma Resuscitation."* [PDF] . *(***Accepted, IFAC World Congress, 2026)**

[8] Giuliano Pioldi, Yashika Batra, Yuanchen Bai, Purnjay Maruur, Arman Ibrayeva, Promise Ekpo, Angelique Taylor. "REPAIR-Bench: A Benchmark for Robot Error Perception And Interaction Recovery." [PDF] **(Submitted to IROS Conference, 2026)**

RESEARCH EXPERIENCE

Graduate Researcher, AirLAB | Cornell Tech

Advisor: Prof. Angelique Taylor *Sept 2023 – Present*

REPAIR-Bench: A Benchmark for Robot Error Perception And Interaction Recovery (IROS 2025, Submitted)

- Built REPAIR-Bench on 214 interaction trials from 41 participants, spanning four induced failure types with synchronized multimodal signals including facial action units (OpenFace), head pose, speech transcripts, and post-interaction affect and recovery reports.
- Formalized 3 novel evaluation tasks capturing the full failure lifecycle: longitudinal failure detection across interdependent sessions, visual failure-type classification, and user-centered recovery preference prediction
- Demonstrated that hierarchical recurrent modeling improves failure detection over single-session baselines (strict F1: 0.80 vs. 0.68), with failure localization median absolute error of 2.97s.

- Fine-tuned Mistral-7B with QLoRA for user-centered recovery prediction, achieving Hit@5=0.76 and F1@5=0.32, enabling adaptive recovery strategies grounded in user behavior.

RFM-HRI Dataset: Multimodal Robot Failure Response Dataset (THRI 2025, Submitted)

- Built a multimodal dataset capturing 214 human robot interactions from 41 laypersons and healthcare workers responding to crash cart robot failures in an item retrieval task.
- Developed an annotation framework with facial action units (OpenFace) and speech transcription (Whisper) to quantify user reactions and preferred recovery modes in embodied AI.
- Demonstrated that user responses to robot failures vary significantly by context, from confusion in the first trials to frustration in the last set of trials for participants.
- Discovered that participants' preferred recovery strategies combine verbal correction with transparency cues.
- Tracked user emotions using the Self-Assessment Manikin (SAM) to record valence and perceived control.

Failure-Predictive Clarification System (In preparation)

- Designed a hybrid safety framework combining failure prediction with risk weighted ambiguity detection to prevent robot failure chains before they occur through clarification dialogue, improving accuracy in controlled settings.
- Prioritized safety clarifications over knowledge preferences, reducing unnecessary interruptions while maintaining task success in household robotics tasks,expansion to cross domain applications (Kitchen, emergency room)

Fair-GNE: Generalized Nash Equilibrium-Seeking Fairness *(L4DC 2025 submission, Under review (preprint available))*

- **Achieved improvement in workload balance over fixed-penalty baselines (JFI: 0.89 vs 0.33, $p<0.01$, Cohen's $d=8.99$) while maintaining 86% task success, demonstrating that adaptive constraint enforcement outperforms static penalty tuning.**
- Designed an automatic penalty parameter learning using the Lagrangian dual ascent RL method to eliminate manual hyperparameter tuning in fairness in MARL with heterogeneous teams.
- Developed Fair-GNE framework with theoretical convergence guarantees through KKT conditions, achieving 95-100% KKT satisfaction and 88-100% constraint satisfaction across fairness thresholds

FairSkillMARL: Skill-Aligned Fairness Framework and MARLHospital SImulator *(AAMAS 2025 submission, Under review (preprint available))*

- **Achieved 40-60% higher skill-task alignment than workload-only baselines with 70-85% reduction in workload range disparity, maintaining success rates of 83% vs 80-86%**
- Developed composite fairness metric $L3=\alpha L1+(1-\alpha)L2$ balancing workload equality and skill-task alignment, demonstrating statistical significance ($p=0.095$, $r=0.47$)
- Built a MARLHospital simulation environment, the MARL benchmark supporting heterogeneous agents with skill heterogeneity, energy levels, and customizable team compositions.

Distributed GNE for Trauma Resuscitation *(IFAC 2026 submission)*

- **Developed distributed GNE algorithm achieving state consensus at $||x_i||\approx 0.35$ within $t<10$ time units for n=7 healthcare workers, with heterogeneous dual variables ($\lambda_i \in [0,15.5]$) reflecting network topology**
- Formalized clinical worker attributes (skill proficiency, alacrity, workload fairness, communication efficiency) as mathematical quantities in distributed game-theoretic framework.

Human-Robot Teaming Field Deployments *(IEEE RO-MAN 2025)*

- **Participated in field deployment of robotic crash cart with 115 healthcare workers at Weill Cornell Medicine, conducting statistical analysis of workload and attitudes comparing verbal vs non-verbal robot communication modalities**
- Analyzed between-subjects experimental data using MANOVA with post-hoc tests (Bonferroni, Games-Howell), demonstrating significant reduction in mental demand ($p=.042$) and effort ($p=.021$) with speech-based guidance

- Conducted quantitative evaluation of NASA-TLX workload measures (45 responses) and ACIR-Q workplace attitudes (23 responses), providing insights for multimodal robot communication design.

Large-Scale Distributed Training Infrastructure

- Designed SLURM workflows managing 200+ concurrent multiple GPU jobs on Unicorn/G2 clusters.
- Conducted statistical analysis with experimental protocols with bootstrap resampling, Welch's t-tests with Bonferroni correction, and matched random seeds across 500K-700K timesteps

Graduate Researcher | Princeton University

M.A. Thesis: Investigating Persuasiveness in Large Language Models | Advisor: Prof. Jaime Fernández Fisac , AI Safety & Alignment , Sept 2021 – June 2022

- **Quantified persuasiveness abilities of GPT-3/4 using a game-theoretic framework, demonstrating 12-50% belief shift toward false statements (prior/0.58: → posterior/0.79 for GPT-3), addressing misalignment and safety failures** in pre-trained LLMs.
- Discovered systematic social bias such as ChatGPT generated misleading arguments for USA 10/10 times but refused for Nigeria 9/10 times on identical false claims
- Trained LLMs with PPO on persuasiveness reward signal over 5000 iterations, observing emergent manipulation strategies (authority establishment initiating trust, emotional appeals, reward hacking).
- Designed novel reward functions for persuasiveness and truthfulness for evaluating persuasiveness as an extension of the Google BIG-BENCH benchmark, **contributing to AI safety frameworks for responsible deployment**
  Technical Stack: Hugging Face Transformers, PyTorch, OpenAI API (GPT-3/4), Python

Principal Investigator: Prof. Barbara Engelhardt August 2021 – August 2022

Beehive Research Laboratory, Dept. of Computer Science, Princeton University, USA

Main Project: Multi-Group Reinforcement Learning for Personalized Healthcare

This method involved using multi-group Gaussian process regression models in a fitted Q-iteration (FQI) reinforcement learning framework for personalization, deriving optimal policies for different subgroups of patients according to their comorbidities, and addressing fairness concerns in medical treatment recommendations across diverse patient populations.

- Developed a novel algorithm built on the intersection of reinforcement learning and Gaussian processes for electrolyte repletion, applying probabilistic modeling approaches to handle uncertainty in clinical decision-making and contributing to responsible AI in healthcare contexts

INDUSTRY EXPERIENCE

AI Safety and Verification Intern | Siemens

Robust Intelligence Team | PI: Christof Budnick | Manager: Georgi Markov

June 2023 – August 2023

- Developed explainable AI framework using GNNExplainer on graph attention networks for medical diagnosis, providing detailed explanations for disease inference predictions on real-world datasets.
- Engineered multimodal safety monitor (Project Titanium) combining vision-based AI and real-time localization systems (RTLS) to ensure worker safety in robotic warehouses
- Designed data fusion algorithms for seamless integration of vision and RTLS systems, optimizing I/O compatibility for real-time safety monitoring

Data Engineer | Infinion Technologies (Microsoft Gold Partner)

July 2020 – March 2021

- Led database migration from multiple servers to Azure Data Warehouse for Union Bank, implementing ETL processing with SQL Server Integration Services (SSIS)

IT Support Intern | Shell Nigerian Exploration and Production Company

July 2018 – Nov 2018

- Conducted sentiment analysis for Shell Eco-Marathon monitoring using Python, collaborating with External Relations team on reporting

TECHNICAL EXPERTISE

- **Optimization & Game Theory:** Expert in Lagrangian methods, dual ascent, KKT conditions, generalized Nash equilibria, variational inequalities. Experience in multi-objective RL for fairness and equilibrium-seeking algorithms.
- **AI Safety & Explainability:** GNNExplainer, graph attention networks, LLM safety evaluation, persuasiveness quantification, multimodal safety monitoring,
- **Distributed Computing:** SLURM workflows for 200+ concurrent multi-GPU jobs.
- **Deep Learning:** PyTorch, TensorFlow, JAX, EPyMARL, OpenAI Gym, Stable-Baselines3, HuggingFace Transformers, Graph Neural Networks.
- **RL & Post-Training:** RLHF, DPO, PPO, reward modeling, automatic KL penalty tuning, multi-objective optimization, preference learning, VLMS
- **Statistical Analysis & Data Engineering:** MANOVA, Welch's t-tests, Bonferroni correction, Cohen's d, Jensen-Shannon divergence, Azure Data Warehouse, ETL (SSIS), SQL
- **LLM Post-Training & RLHF Experience:** Implemented SFT and DPO pipelines for instruction-following language models using Hugging Face Transformers and Google Vertex AI
- **Vision-Language Models:** Gemini, Qwen2.5-VL, Florence-2 (object detection, image captioning)

HONORS & AWARDS

- NCWIT Aspirations in Computing (AiC) Program Grand Prize Winner **($8,500)** | 2026
- **NeurIPS Women in Machine Learning Grant ($1100) + Netflix & RTC Grant ($3,200) | 2025**
- **CMD-IT/ACM Richard Tapia Conference 2025, 1st Place Best Poster($1500)(Graduate) | 2025**
- **Cadence Technology Awards – Women in Tech Category** (8% acceptance) ($**5,000**)| 2023
- **Gordon Wu Fellowship, Princeton University** (0.03% acceptance) | 2021
- Princeton Data Driven Social Science Fellowship ($2,000) | 2023
- NeurIPS WiML Travel Grant & Black in AI Travel Grant | 2022
- EducationUSA Scholarship, Opportunity Funds Program, U.S. Embassy Lagos | 2020
- Dean's Honor Award, Best Graduating Student, Computer Engineering, University of Benin | 2020

PROFESSIONAL SERVICE & LEADERSHIP

- Conference Reviewer: NeurIPS, ICLR, AAAI, L4DC, AAMAS, IFAC, IEEE RO-MAN (2024-2025)
- Research Mentor: Supervising 5 undergraduate/MS researchers on MARL fairness projects (2023-Present)
- Collaborations: Microsoft Research (Lekan Molu), Weill Cornell Medical College (trauma resuscitation deployment)

CERTIFICATES

- Microsoft Certified : Azure Data Engineer Associate 2021
- IBM Data Science Professional Certificate 2019