ISABEL **MACGINNITIE**

imacginnitie@hmc.edu | www.isabel.fyi

Education

Harvey Mudd College, Claremont, CA   Aug '19—May '23
B.S. in Computer Science and Mathematics (GPA: 3.73), Minor in Photography

Research Experience

Building a Model Organism of Encoded Chain of Thought   Nov '25—Present
Mentored by: Matt MacDermott (Imperial College London)   *Athena AI Alignment Mentorship Program*
- Creating a proof of concept for models to obscure their chain of thought while completing reasoning tasks, using outside-the-box semantic encoding: emojispeak, concept-word substitution, rhyming, and metaphors.
- Few-shot prompt the model with examples of the encoded CoT for math problems, evaluating results by checking for correct values in answers and measuring adherence to encoding in CoT using LLM-as-a-judge.
- From the most successful encoding technique, plan to use supervised fine tuning to create and evaluate a model organism with a broadly encoded chain of thought across a range of reasoning tasks.
- Presented preliminary results as an oral presentation at the 2025 Athena retreat in Czechia.

Evaluating Stated vs Revealed Preferences in LLMs Using Animal Ethics   Sept '25—Present
Mentored by: Catherine Brewer (Open Philanthropy)   *Supervised Program for Alignment Research (SPAR)*
- Investigating alignment between LLMs' stated values and revealed preferences using animal ethics scenarios: comparing what models say matters for moral consideration with their choices in welfare tradeoff dilemmas.
- Designed three-stage experiments across five species (dogs, pigs, chickens, shrimp, bees): eliciting freeform values, querying for values on specific proxies for sentience from welfare science (pain sensitivity, neuron counts, self-awareness), and establishing revealed preference rankings through forced choices.
- Comparing stated values, revealed preferences, and sentience proxy metrics across three model families (Claude Opus, GPT-4o, Llama 3.1 70B) to better understand model moral reasoning and behavior.

Building a Practical and Intuitive Image Complexity Metric   Sept '21—Dec '22
Mentored by: Prof Lucas Bang (3 semesters and summer REU)   *Harvey Mudd College*
- Approximated logical depth—an idealized metric of intuition-aligned complexity defined as the time it takes to decode the minimal representation of an image—using compression algorithms and timing methods.
- Developed an image generation system using OpenCV, working independently to design and implement software, producing a dataset of 68,000 images with organized complexity trends (680 sets of 100 images).
- Ran extensive testing to assess performance for each combination of algorithm and timing method and determine the most reproducible and accurate combination (LZ77 and CPU instruction count).
- Presented results as a poster at the 2022 Southern California CS REU Symposium.

Awards & Recognitions

NSF CSGrad4US Computer Science PhD Fellowship   Aug '25—Present
- Award of $159,000 for three years of full graduate funding over a five-year fellowship period [28 awardees].

CS Departmental Award for Computational Creativity   May '23
- Recognizes exceptional graduates who use computer science to make art. [3 awardees; inaugural award]

Graduate with High Distinction   May '23
- Highest honor of scholarly achievement. [105 awardees]

Industry Experience

Software Engineer, RetroRate, San Diego, CA   Jan '25—Oct '25
Early-stage fintech startup ($2.2M seed funding) tackling housing crisis with assumable mortgage technology
- First engineering hire on a team of three engineers: drove decision-making and developed systems for full-stack real estate platform using React/TypeScript frontend and Express.js REST APIs deployed on GCP, featuring complex search filters with multi-criteria property matching, interactive map visualization for 2.9M+ properties, and embeddable widget for real estate agents websites with 1,000+ daily views.
- Owned mortgage data processing architecture for 2.9M assumable and 152M non-assumable properties across 10 states, including designing, implementing, and maintaining a data reconciliation system that takes 7 independent sources into detailed assumable loan and property information with minute-level accuracy.
- Independently created a daily interest rate infrastructure for updating real-time rates for 155M properties using Kubernetes, including web scraping architecture, graceful database updates, and robust error handling.

Software Engineer, Color Health, Burlingame, CA   Oct '23—Jul '24
Health tech startup using AI and genetic testing for early cancer detection.
- Modernized frontend infrastructure from Angular to React for healthcare-grade genetic testing reports serving 400,000+ customers while maintaining strict FDA compliance requirements.
- Gained exposure to regulatory considerations around AI deployment in high-stakes healthcare domains.

Explore Intern, Microsoft, Redmond, WA   May—Aug '21
Azure Edge + Platform, Azure Watson team
- Developed design for and implemented debugging software from concept to fully functional prototype.

Projects

Futurekind AI for Animals Fellowship, Electric Sheep   Apr—Jul '25
- Inaugural fellow: 13-week program of coursework covering AI safety and animal ethics, culminating in the design of a research project: *Investigating Value Coherence in Large Language Models Using Animal Ethics*
- Designed choice experiments to evaluate AI moral reasoning using Mazeika et al.'s utility engineering framework for Thurstonian utility models to measure value coherency and exchange rates, then comparing results to social science derived sentience metrics to better understand models' valuation of animals.

Automated Highlight Reel Generation For American Football   Aug '22—May '23
Supervisors: Prof Calden Wloka, Dr. Pesh Pahalawatta (capstone project)   Harvey Mudd College & DIRECTV
- Collaborated with 6 other undergrad students to experimentally combine existing computer vision, signal processing, and video models to automate highlight reel generation for full-length NFL games.
- Used OpenCV to implement a prototype goal post detector with classical CV and machine learning.

Comparing Deep Learning and Classical Colorization on Film Photography   Mar—May '21
Mentored by: Prof Darryl Yong, Prof Doug Goodwin   Harvey Mudd & Scripps College
- Designed an independent research project comparing machine learning colorization with a classical technique from 1900s Russia [submitted as a joint final project for Comp Photo II and Scientific Computing]
- Trained a convolutional neural network colorizer with TensorFlow on an experimentally curated dataset.

Teaching Experience

Writing Center Consultant, Harvey Mudd College *(6 semesters & 1 summer term)*   Aug '20—May '23
- Gave personalized guidance on student papers, technical reports, and posters. [40 min appts, 6x per week]

CS Dept Teaching Assistant, Harvey Mudd College   Jan '21—May '23
Computer Vision (2 semesters), Programming Languages (1 semester), Data Structures (1 semester)
- Held office hours for problem sets with 0-1 other undergrad TAs. [2x per week, avg 10 students/session]