Fabrice Harel-Canada

SOFTWARE ENGINEER · COMPUTER SCIENCE RESEARCHER

☎ (+1) 954-805-5735 | ✉ fabricehc@cs.ucla.edu | ⌂ fabrice.harel-canada.com | ⌥ fabriceyhc

Biography

Recent UCLA Computer Science PhD (May 2025) specializing in pioneering robust and holistic evaluation frameworks for AI systems. My doctoral research, including award-winning work on human-centric LLM quality assessment, developed novel methods in AI testing, data augmentation (e.g., sibylvariant transformations), content provenance (watermarking), and applied NLP (e.g., drug detection). Passionate about advancing trustworthy and beneficial AI, I am currently applying these insights through my new venture, Kairno, aimed at enhancing research critique.

Education

University of California, Los Angeles (UCLA)

Advisor: Miryung Kim

PhD | COMPUTER SCIENCE
Jan 2020 - May 2025

- 3.92 / 4.00 GPA
- Majors: Artificial Intelligence + Software Engineering | Minor: Data Science Computing
- Thesis: *From Metrics to Meaning: Advancing Evaluation Frameworks for Robust and Human-Centric AI*

University of California, Los Angeles (UCLA)

Advisor: Miryung Kim

MS | COMPUTER SCIENCE
Sept 2018 - Dec 2019

- Major: Artificial Intelligence
- Thesis: *Is Neuron Coverage a Meaningful Measure for Testing Deep Neural Networks?*

University of Florida (UF)

BS | BUSINESS ADMINISTRATION
Aug 2008 - May 2012

- Majors: Management + Information Systems and Operations Management
- Minors: Entrepreneurship + Finance
- Thesis: *The Viability of the Student-Run Housing Cooperative: A Case Study of the Cooperative Living Organization*

Publications

[7] **Fabrice Y. Harel-Canada***, Boran Erol*, Connor Choi, Jason Liu, Gary Jiarui Song, Nanyun Peng, and Amit Sahai. "Sandcastles in the Storm: Revisiting the (Im)possibility of Strong Watermarking". In: *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (ACL 2025)*. To appear. (arXiv:2505.06827 [cs.CR]). Vienna, Austria: Association for Computational Linguistics, July 2025. URL: https://arxiv.org/abs/2505.06827.

[6] **Harel-Canada, Fabrice Y.**, Anabel Salimian, Brandon Moghanian, Sarah Clingan, Allan Nguyen, Tucker Avra, Michelle Poimboeuf, Ruby Romero, Arthur Funnell, Panayiotis Petousis, Michael Shin, Nanyun Peng, Chelsea L. Shover, and David Goodman-Meza. "Enhancing Substance Use Detection in Clinical Notes with Large Language Models". Version 1. In: *Research Square (Preprint)* (May 2025). PREPRINT (Version 1). DOI: 10.21203/rs.3.rs-6615981/v1. URL: https://doi.org/10.21203/rs.3.rs-6615981/v1.

[5] **Harel-Canada, Fabrice Y**, Hanyu Zhou, Sreya Muppalla, Zeynep Senahan Yildiz, Miryung Kim, Amit Sahai, and Nanyun Peng. "Measuring Psychological Depth in Language Models". In: *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*. Ed. by Yaser Al-Onaizan, Mohit Bansal, and Yun-Nung Chen. **Outstanding Paper Award (Top 0.25% of submissions)**. Miami, Florida, USA: Association for Computational Linguistics, Nov. 2024, pp. 17162–17196. DOI: 10.18653/v1/2024.emnlp-main.953. URL: https://aclanthology.org/2024.emnlp-main.953/.

[4] Hong Jin* Kang, **Harel-Canada, Fabrice***, Muhammad Ali Gulzar, Nanyun Peng, and Miryung Kim. "Human-in-the-Loop Synthetic Text Data Inspection with Provenance Tracking". In: *Findings of the Association for Computational Linguistics: NAACL 2024*. Ed. by Kevin Duh, Helena Gomez, and Steven Bethard. Mexico City, Mexico: Association for Computational Linguistics, June 2024, pp. 3118–3129. DOI: 10.18653/v1/2024.findings-naacl.197. URL: https://aclanthology.org/2024.findings-naacl.197/.

[3] **Harel-Canada, Fabrice**, Muhammad Ali Gulzar, Nanyun Peng, and Miryung Kim. "Sibylvariant Transformations for Robust Text Classification". In: *Findings of the Association for Computational Linguistics: ACL 2022*. Ed. by Smaranda Muresan, Preslav Nakov, and Aline Villavicencio. Dublin, Ireland: Association for Computational Linguistics, May 2022, pp. 1771–1788. DOI: `10.18653/v1/2022.findings-acl.140`. URL: `https://aclanthology.org/2022.findings-acl.140/`.

[2] Guangxuan Xu, Ruibo Liu, **Harel-Canada, Fabrice**, Nischal Reddy Chandra, and Nanyun Peng. "EnDex: Evaluation of Dialogue Engagingness at Scale". In: *Findings of the Association for Computational Linguistics: EMNLP 2022*. Ed. by Yoav Goldberg, Zornitsa Kozareva, and Yue Zhang. Abu Dhabi, United Arab Emirates: Association for Computational Linguistics, Dec. 2022, pp. 4884–4893. DOI: `10.18653/v1/2022.findings-emnlp.359`. URL: `https://aclanthology.org/2022.findings-emnlp.359/`.

[1] **Harel-Canada, Fabrice**, Lingxiao Wang, Muhammad Ali Gulzar, Quanquan Gu, and Miryung Kim. "Is neuron coverage a meaningful measure for testing deep neural networks?" In: *Proceedings of the 28th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*. 2020, pp. 851–862.

A more comprehensive list is available at: Semantic Scholar Profile

Research Experience

University of California, Los Angeles *Supervisor: David Goodman-Meza*

GRADUATE STUDENT RESEARCHER — *Dec 2023 - present*

- Led the development of a context-aware drug misuse detection system, achieving F1-scores of 0.99+ for multiple drug classes, improving clinical identification accuracy.
- Collaborated on designing a predictive model for the time and location of probable overdoses in Los Angeles County.

University of California, Los Angeles *Supervisor: Miryung Kim*

GRADUATE STUDENT RESEARCHER — *June 2019 - Dec 2023*

- Conducted a theoretical and empirical analysis demonstrating the limitations of robust LLM watermarking, constructing a modular, quality-preserving attack that no current watermarking technique can withstand.
- Developed a theoretically grounded framework to assess the psychological depth of LLMs, revealing that modern LLMs match or exceed the abilities of advanced human short story writers.
- Created automated evaluation metrics for natural language generation systems, including measures of linguistic diversity – e.g. semantic, syntactic, morphological, phonological – and as well as general engagement.
- Designed and open-sourced a new class of data augmentations called sibylvariant transforms for enhancing NLP model robustness.
- Assessed the reliability of structural testing metrics, such as neuron coverage, in evaluating deep learning models.
- Developed diversity-enhancing extensions for adversarial attack algorithms to improve robustness testing.

University of California, Los Angeles *Supervisor: Judson Caskey*

RESEARCH ASSISTANT — *July 2020 - Oct 2020*

- Conducted LDA topic modeling and phrase-level sentiment analysis using BERT transformers on millions of Amazon and Glassdoor product / employer reviews

Professional Experience

Kairno

FOUNDER AND CEO — *May 2025 - present*

- Founding Kairno to revolutionize research communication by developing an AI-powered platform for deep, automated critique of scientific works, leveraging insights from doctoral research on LLM reliability and quality assessment.

University of California, Los Angeles *Supervisor: David Smallberg*

TEACHING ASSISTANT, CS 31 : INTRO TO COMPUTER SCIENCE — *Sep 2023 - Dec 2023*

- Facilitated a robust understanding of C++ fundamentals for introductory computer science students, providing guidance on programming concepts, debugging techniques, and hands-on problem-solving support.

University of California, Los Angeles *Supervisor: Michael Burns*

TEACHING ASSISTANT, CS 130 : SOFTWARE ENGINEERING — *Mar 2023 - June 2023*

- Collaborated with a team of Google instructors to facilitate cutting-edge learning experiences and contributing to the development of students' technical skills.

University of California, Los Angeles
Supervisor: Miryung Kim

TEACHING ASSISTANT, CS 130 : SOFTWARE ENGINEERING
Sept 2020 - Dec 2020

- Facilitated hands-on learning in software engineering by guiding 50+ undergraduates through coding exercises and debugging, enhancing their grasp of complex algorithms and object-oriented programming.
- Mentored student teams through capstone projects, enhancing their understanding of full-stack development and agile methodologies.

University of California, Los Angeles
Supervisor: Carlo Zaniolo

READER, CS 143 : DATABASE SYSTEMS
Jan 2019 - Mar 2019

- Developed test cases for a Spark project in Scala, actively addressed student questions, and assisted students with project work

Verox Tech
Supervisor: Michael Simhai

PROJECT MANAGER + DEVELOPER TEAM LEAD
July 2015 – Sept 2018

- Designed and successfully launched a smarter insurance app by building, motivating, and guiding a team of 23
- Contributed 100s of check-ins, primarily consisting of C# Web APIs and database tables, triggers, sprocs, and T-SQL scripts
- Designed and implemented a 3-system legacy data migration
- Responsible for producing, coordinating and reviewing nearly every aspect of the SDLC:
    - code reviews and code optimization
    - requirements and spec generation
    - database design + SSIS + SSRS
    - technical documentation
    - UI design standards + product design
    - build and deployment scripts

People's Trust Insurance
Supervisor: Scott Thompson

BUSINESS + DATA ANALYST
Dec 2012 – Sept 2018

- Go-to data expert for executives to produce hundreds of reports and analyses for pivotal decision making at the executive level
- Developed numerous front-end / back-end application features that were successfully deployed into production to automate tedious and time-intensive workflows
- Performed continuous requirements gathering & demos with business stakeholders to support 60+ sprints of development for both applications and business intelligence teams, including the transition to a new policy and claim management system

Cooperative Living Organization

PRESIDENT
Dec 2010 – May 2012

- Appointed and oversaw the performance of 17 managers who were responsible for executing the policies and procedures I developed
- Increased house occupancy from 60 residents to full capacity (80) and increased net income 57% while expanding services & activities for residents

Skills

Programming	python, sql, java, c#, linq, html, css, js, scala, problog, prolog, LaTeX
Frameworks	huggingface, pytorch, tensorflow, keras, spark, numpy, pandas, scipy, sklearn, opencv, matplotlib, seaborn
Tools	ssms, ssis, ssrs, tfs, azure devops, visual studio, excel, aws, gcp, exchange server, moqups