David Xia

davidx3@illinois.edu | Personal Website | LinkedIn - dx3 | GitHub - davidxia3

Education

- **University of Illinois Urbana-Champaign** *Aug 2023 - May 2027*
  Bachelor of Science in Computer Science Champaign, Illinois
  Bachelor of Science in Liberal Arts & Sciences Major in Mathematics, Data Optimization Concentration
  Bachelor of Science in Liberal Arts & Sciences Major in Statistics

  ◦ GPA: 3.98/4.00

Papers

SUB=Submitted, ACC=Accepted

2026 **[SUB] Revitalizing Local Democracy: A Human-Centered Audit of LLMs in City Council Journalism**
David Xia, Chris Maury
Submitted to CHI 2026 HEAL Workshop

2025 **[SUB] RSK linear operators and the Vershik-Kerov-Logan-Shepp curve** [arXiv]
Duy Phan, **David Xia**
Submitted to Electronic Journal of Combinatorics

2024 **[ACC] EveGuard: Defeating Vibration-based Side-Channel Eavesdropping with Audio Adversarial Perturbations** [arXiv]
Jung-Woo Chang, Ke Sun, **David Xia**, Xinyu Zhang, Farinaz Koushanfar
IEEE Symposium on Security and Privacy, 2025

Research Experience

- **Carnegie Mellon University Human-Computer Interaction Institute**, Research Intern *May 2025 - present*

  ◦ Designed and evaluated an automated journalism pipeline using SOTA LLMs to replicate end-to-end editorial workflows, from transcript segmentation to headline generation and topic prioritization.
  ◦ Orchestrated large-scale crowd-sourced study to establish human-evaluated ground truth validation.
  ◦ Demonstrated that LLM-based headline quality and topic prioritization can meet and exceed professional-standards.
  ◦ Presented research poster at the CMU HCII REU Symposium.
  ◦ Paper in submission to CHI 2026 Human-centered Evaluation and Auditing of Language Models (HEAL) workshop.

- **Illinois Mathematics Lab (formerly Illinois Geometry Lab)**, Student Scholar *Jan 2024 - present*

  ◦ Researched the data science perspective of sports matches at the Illinois Mathematics Lab.
  ◦ Built web scrapers with Selenium to scrape online data and analyzed the predictability of sports matches by comparing prediction accuracies of multiple sources including polls and the betting market.
  ◦ Evaluated efficiency of probabilistic predictions based on the betting market, Elo, and other mathematical models.
  ◦ Presented findings at Rose-Hulman Undergraduate Math Conference 2024, UIUC Undergraduate Research Symposium 2024, and Joint Mathematics Meeting 2025.
  ◦ Currently working on paper.

- **Illinois Combinatorics Lab for Undergraduate Experience (ICLUE)**, Research Mentee *Aug 2024 - May 2025*

- Studying algebraic combinatorics, focusing on the Robinson-Schnested-Knuth correspondence.
- Used the Vershik-Kerov-Logan-Shepp curve to apply probability bounds on the RSK linear operator.
- Paper in submission to the Electronic Journal of Combinatorics.

• **National University of Singapore SEEDER Group**, Research Intern *May 2024 - Aug 2024*

- Developed machine learning pipeline to detect, classify, and analyze diseases found in plant leaves.
- Trained U-Net segmentation models and classification models based on convolutional neural networks to isolate and classify images using the TensorFlow framework, achieving 90% segmented classification rate.
- Generated synthetic images using custom trained cGAN model to increase data variety and volume.
- Stacked deep learning classification models with traditional machine learning models to achieve an increased detection accuracy of over 95%.

• **University of California San Diego**, Research Intern *Jun 2020 - Aug 2023*

- Developed EveGuard, a software-based defense framework to protect voice privacy from vibrometry-based side channels using adversarial audio by using a perturbation generator model (PGM) to effectively suppress sensor-based eavesdropping while preserving high audio quality.
- Implemented Eve-GAN, a novel domain translation task for inferring eavesdropped signals, enabling end-to-end training of PGM, utilizing few-shot learning techniques to reduce data collection overhead.
- Achieved a protection rate of over 97% against audio classifiers, hindering eavesdropped reconstruction.
- Paper published in the IEEE Symposium on Security and Privacy 2025.

Presentations

2025 **Carnegie Mellon University Human-Computer Interaction Institute REU Symposium** LLMs for Extracting Agenda Items from City Council Meetings and Generating Newsworthy Headlines

2025 **Algebra, Geometry and Combinatorics Day** RSK linear operators and the Vershik-Kerov-Logan-Shepp curve (joint work presented by Duy Phan)

2025 **Joint Mathematics Meeting Pi Mu Epsilon Poster Session** Predicatbility in Major Sports Leagues

2024 **Illinois Mathematics Lab Research Symposium** Predictability in College Sports

2024 **University of Illinois Undergraduate Research Symposium** Predictability in College Sports

2024 **Rose-Hulman Undergraduate Math Conference** Predictability in College Sports

Service

SP26 UIUC Algorithms and Models of Computation (CS 374) Course Assistant

Grants and Awards

2025 CMU HCII REU: NSF 2349558

2025 ICLUE RTG: NSF 1937241

Languages

- English: Native
- Mandarin Chinese: Fluent, HSK6

davidx3@illinois.edu | Personal Website | LinkedIn - dx3 | GitHub - davidxia3