Chung Peng (Anderson) Lee

Email: cl6486@princeton.edu
URL: https://anderson-lee-git.github.io/

EDUCATION

2025 - now **Princeton University**
Ph.D. in Computer Science   *Adviser: Aleksandra Korolova*

- **Relevant Courses** – Efficient Systems for Foundation Models, Deep Learning Theory

2021 - 2024 **University of Washington**   GPA: 3.98 / 4.00
B.S. in Computer Science   *Adviser: Jamie Morgenstern*

- Magna Cum Laude, Departmental Honor
- **Relevant Courses** – Machine Learning, Advanced Machine Learning, Deep Learning, Data Structures & Parallelism, Algorithms, Data Management (SQL), Explainable AI, Probabilistic Graphical Models, Natural Language Processing, Operating System

PUBLICATIONS

CONFERENCE PAPERS

2025 **Chung Peng Lee**, Rachel Hong, Harry H. Jiang, Aster Plotnik, William Agnew, and Jamie Morgenstern (2025), "How do data owners say no? A case study on data consent mechanisms in web-scraped vision-language AI training datasets," The 40th Annual AAAI Conference on Artificial Intelligence – AI for Societal Impact Track (AISI 2026). (**Oral**) [Paper]

Yiwei Yang*, **Chung Peng Lee***, Shangbin Feng, Dora Zhao, Bingbing Wen, Anthony Z. Liu, Yulia Tsvetkov, and Bill Howe (2025), "Escaping the SpuriVerse: Can Large Vision-Language Models Generalize Beyond Seen Spurious Correlations?" The Thirty-Ninth Annual Conference on Neural Information Processing Systems Datasets and Benchmarks Track (NeurIPS D&B 2025). [Paper]
* denotes equal contribution.

WORKSHOP PAPERS

2024 **Chung Peng Lee**, Rachel Hong, and Jamie Morgenstern (2024), "Fairness through partial awareness: Evaluation of the addition of demographic information for bias mitigation methods," The Forty-First International Conference on Machine Learning – Next Generation of AI Safety Workshop (ICML NGAIS 2024). [Paper]

WORK EXPERIENCE

06/2024 - 09/2024 ***Stripe*** | *Software Engineer (Machine Learning) Intern*   Seattle, WA
Airflow, Flyte, Chronon, Hadoop, (Py)Spark, Databricks, Bazel, AWS S3

- Integrated Airbnb's Chronon feature store to training and serving ML models to allow cross-function feature sharing, reduce vanilla feature computation pipeline maintenance, and ensure offline-online consistency
- Reduced 2-days feature latency to within-1-day with <0.1% cost in loss saving (metrics) by revisiting charge related features and re-implementation on Chronon
- Increased projected 0.6M USD annual loss savings by incorporating new features through feature sharing

| 06/2023 – 08/2023 | **SeekOut** | *Software Engineer Intern* | Bellevue, WA |

Azure, Azure OpenAI, CosmosDB, NoSQL, Rest API, Kusto, C#, TypeScript, React

- Delivered an end-to-end external-facing admin tool in production to allow hiring managers in client's organization to curate skills content pertained to their internal job board with skills recommendation

TEACHING

2024	CSE 446/546: Machine Learning (Teaching Assistant, Winter 2024)
2023	CSE 446/546: Machine Learning (Teaching Assistant, Fall 2023)
	CSE 446/546: Machine Learning (Teaching Assistant, Spring 2023)
	CSE 123: Introductory Programming III (Teaching Assistant, Winter 2023)
2022	CSE 121: Introductory Programming I (Teaching Assistant, Fall 2022)

SERVICE

2025	*Program Committee*, AAAI – AI for Societal Impact Track (AISI 2026)
	Reviewer, NeurIPS Regulatable ML Workshop (NeurIPS RegML 2025)
2024	*Reviewer*, ICML Next Generation of AI Safety Workshop (ICML NGAIS 2024)