

Zhaoheng (Billy) Li

Thomas M. Siebel Center, Rm 2119
201 North Goodwin Avenue
Urbana, IL 61801

billyzhaohengli.github.io
zl20@illinois.edu
+1 217-721-0260

Research	Systems for AI and ML; Interactive Data Analytics; Computational Notebooks; Vector Search & Databases	
Education	Ph.D., Computer Science University of Illinois at Urbana Champaign • Dissertation Proposal: Data Infrastructure for Emerging AI and Science Applications • Advisor: Yongjoo Park	Aug 2021 - May 2026 (expected)
	B.S., Statistics and Computer Science University of Illinois at Urbana Champaign • <i>Magna Cum Laude and with Highest Distinction in Statistics and Computer Science</i> • <i>High Distinction in Mathematics</i> • Thesis: REFORM: Fast and Adaptive Solution for Subteam Replacement • Advisor: Hanghang Tong	Aug 2017 - May 2021
Professional Experience	Research Assistant University of Illinois at Urbana Champaign CreateLab	Aug 2021 - Present
	Research Intern ByteDance Inc. Systems Infrastructure Lab	May 2025 - Aug 2025
	Research Intern ByteDance Inc. Systems Infrastructure Lab	May 2024 - Oct 2024
	Software Engineer Intern Google LLC Google Cloud, Google BigQuery	May 2023 - Aug 2023
	Software Engineer Intern Google LLC Google Cloud, Systems & Storage Infrastructure	May 2022 - Aug 2022
	Research Assistant University of Illinois at Urbana Champaign IDEA Lab	Aug 2020 - May 2021
	Software Engineer Intern Google LLC Google Ads, Google Local Services	May 2020 - Aug 2020
	Software Engineer Intern Google Canada Corporation	May 2019 - Aug 2019

Google Ads, Google Ads Experiments

Awards	Best Demo Award , SIGMOD 2025 <i>With Supawit Chockchowwat, Hanxi Fang, and Yongjoo Park</i>	June 2025
	Travel Grant , PyCon 2025	March 2025
	Travel Award , ICDE 2023	March 2023
	Finalist , C1Games Terminal Game AI competition	September 2019
	Best Data Visualization Award , Illini Datathon	February 2019
	Top in China , Cambridge CIE AS Sociology exam	October 2016
	National Finalist , High School Mathematical Contest in Modeling (HIMCM)	February 2015

Publications

1. **Zhaoheng Li**, Wei Ding, Silu Huang, Zikang Wang, Yuanjin Lin, Ke Wu, Yongjoo Park, Jianjun Chen
[Cloud-Native Vector Search: A Comprehensive Performance Analysis](#)
ArXiv Preprint
2. Supawit Chockchowwat, Sumay Thakurdesai, **Zhaoheng Li**, Matthew S. Krafczyk, Yongjoo Park
[Chipmink: Efficient Delta Identification for Massive Object Graph](#)
VLDB 2026 (research): 52nd International Conference on Very Large Databases, Boston, USA, 2026
3. Shengya Huang, **Zhaoheng Li**, Derek Werner, Yongjoo Park
[MojoFrame: Dataframe Library in Mojo Language](#)
ICDE 2026 (research): 42nd IEEE International Conference on Data Engineering, Montreal, Canada, 2026
4. Raunak Shah, **Zhaoheng Li**, Yongjoo Park
[QStore: Quantization-Aware Compressed Model Storage](#)
VLDB 2026 (research): 52nd International Conference on Very Large Databases, Boston, USA, 2026
5. **Zhaoheng Li**, Silu Huang, Wei Ding, Yongjoo Park, Jianjun Chen
[SIEVE: Effective Filtered Vector Search with Collection of Indexes](#)
VLDB 2025 (research): 51st International Conference on Very Large Databases, London, United Kingdom, 2025
6. **Zhaoheng Li**, Supawit Chockchowwat, Ribhav Sahu, Arret Sheth, Yongjoo Park
[Kishu: Time-traveling for Computational Notebooks](#)
VLDB 2025 (research): 51st International Conference on Very Large Databases, London, United Kingdom, 2025
7. **Zhaoheng Li**, Supawit Chockchowwat, Hanxi Fang, Yongjoo Park
[Demo of Kishu: Time-traveling for Computational Notebooks](#)
(SIGMOD 2025 Best Demo Award)
SIGMOD 2025 (demo track): ACM Special Interest Group on Management of Data, Berlin, Germany, 2025

8. Mingyuan Wu, Meitang Li, Jingcheng Yang, Jize Jiang, Kaizhuo Yan, **Zhaoheng Li**, Hanchao Yu, Minjia Zhang, Klara Nahrstedt
A-ha Moment Revisited: Are VLMs Truly Capable of Self Verification in Inference-time Scaling?
NeurIPS 2025 (short paper): Multimodal Algorithmic Reasoning workshop, Thirty-Ninth Annual Conference on Neural Information Processing Systems, San Diego, USA, 2025
9. Mingyuan Wu, Jize Jiang, Haozhen Zheng, Meitang Li, **Zhaoheng Li**, Beitong Tian, Bo Chen, Yongjoo Park, Minjia Zhang, Chengxiang Zhai, Klara Nahrstedt
Cache-of-Thought: Master-Apprentice Framework for Cost-Effective Vision Language Model Reasoning
EMNLP 2025 (main conference): 30th Conference on Empirical Methods in Natural Language Processing, Suzhou, China, 2025
10. **Zhaoheng Li**, Pranav Gor, Rahul Prabhu, Hui Yu, Yuzhou Mao, Yongjoo Park
ElasticNotebook: Enabling Live Migration for Computational Notebooks
VLDB 2024 (research): 50th International Conference on Very Large Databases, Guangzhou, China 2024
11. **Zhaoheng Li**, Supawit Chockchowwat, Hanxi Fang, Ribhav Sahu, Sumay Thakurdesai, Kantanat Pridaphatrakun, Yongjoo Park
Demo of ElasticNotebook: Migrating Live Computational Notebook States
SIGMOD 2024 (demo track): ACM Special Interest Group on Management of Data, Santiago, Chile, 2024
12. Supawit Chockchowwat, **Zhaoheng Li**, Yongjoo Park
Transactional Python for Durable Machine Learning: Vision, Challenges, and Feasibility
SIGMOD DEEM 2023 (short paper): 7th Workshop on Data Management for End-to-End Machine Learning, Seattle, USA, 2023
13. **Zhaoheng Li**, Xinyu Pi, Yongjoo Park
S/C: Speeding up Data Materialization with Bounded Memory
ICDE 2023 (research): 39th International Conference on Data Engineering, Anaheim, USA, 2023
14. **Zhaoheng Li**, Xinyu Pi, Mingyuan Wu, Hanghang Tong
REFORM: Fast and Adaptive Solution for Subteam Replacement
IEEE BigData 2021 (research): IEEE International Conference on Big Data, 2021
15. **Zhaoheng Li**, Xinyu Pi, Mingyuan Wu
Subteam Replacement: Problem Definition and Fast Solution
SIGMOD SRC 2021 (extended abstract): ACM Special Interest Group on Management of Data Student Research Competition, Xi'an, China, 2021

Teaching	ML and Data Systems (CS 598) Teaching Assistant, University of Illinois at Urbana-Champaign	Spring 2024
	Game Development (CS 498) Teaching Assistant, University of Illinois at Urbana-Champaign	Spring 2022
	Introduction to Computer Science (CS 125) Course Assistant, University of Illinois at Urbana-Champaign	Spring 2018

Invited Talks	JupyterCon 2025 , San Diego, USA	November 2025
	VLDB 2025 , London, United Kingdom	September 2025
	SIGMOD 2025 , Berlin, Germany	June 2025
	PyCon 2025 , Pittsburgh, USA	May 2025
	UIUC NCSA , Urbana, USA	April 2025
	VLDB 2024 , Guangzhou, China	August 2024
	Bytedance Inc. , San Jose, USA	August 2024
	SIGMOD 2024 , Santiago, Chile	June 2024
	ICDE 2023 , Anaheim, USA	April 2023
	IEEE BigData 2021 (Virtual)	December 2021
	SIGMOD SRC 2021 , Xi'An, China (Virtual)	June 2021