

# Yuchen Liang

yuchenl3@andrew.cmu.edu | [github.com/yliang412](https://github.com/yliang412) | [linkedin.com/in/yliang412](https://www.linkedin.com/in/yliang412)

## Education

---

### Carnegie Mellon University

2020/08 – 2024/05

Bachelor of Science in Computer Science • Concentration in Computer Systems

Pittsburgh, PA

*Relevant Coursework:*

Operating Systems Design and Implementation • Advanced Database Systems • Distributed Systems  
• Parallel Computer Architecture • Algorithm Design and Analysis • Program Analysis • Machine Learning

## Research Experience

---

### CMU Database Group, Adaptive Query Optimization [github.com/cmu-db/optd](https://github.com/cmu-db/optd)

2023/11 – present

Advisor: Dr. Andy Pavlo

Pittsburgh, PA

- Building optd, a Cascades-style cost based query optimizer framework that uses runtime statistics to adaptively guide plan search during re-optimization.

### CMU Database Group, PostgreSQL Extension Manager [github.com/cmu-db/pgextmgrext](https://github.com/cmu-db/pgextmgrext)

2023/02 – 2023/05

Advisor: Dr. Andy Pavlo

Pittsburgh, PA

- Explored database extension ecosystems with a focus on PostgreSQL hooks. Formulated extension management framework for compatibility checking and dependency management. Designed better APIs for easy, safe, and extensible PostgreSQL extension creation and management.

## Publication

---

Abdelkareem Bedri, **Yuchen Liang**, Sudershan Boovaraghavan, Geoff Kaufman, and Mayank Goel. 2022. FitNibble: A Field Study to Evaluate the Utility and Usability of Automatic Diet Monitoring in Food Journaling Using an Eyeglasses-based Wearable. In *27th International Conference on Intelligent User Interfaces (IUI '22)*. Association for Computing Machinery, New York, NY, USA, 79–92. <https://doi.org/10.1145/3490099.3511154>

## Work Experience

---

### Neon [github.com/neondatabase/neon](https://github.com/neondatabase/neon)

2024/06 – 2024/12

Software Engineering Intern

Remote / Pittsburgh, PA

- Neon is a managed PostgreSQL platform built on top of a key-value storage engine that supports point-in-time recovery, autoscaling, and copy-on-write database branching. It is an open-source alternative to Amazon Aurora.
- **Ephemeral Endpoint Support in Storage.** Implement a leasing mechanism to temporarily block the background task from garbage-collecting data needed by the lightweight, short-lived endpoints. Directly enabled safe usage of time-travel query, branch restoration, and other customer-facing features.
- **Direct IO for the Storage Engine.** Eliminate all kernel page cache usage during IO to ensure predictable latencies and avoid extra memory copy. Under high memory pressure, we achieved a 20% decrease in CPU utilization while maintaining similar latency for read operations.

### Amazon Web Services, OpenSearch Serverless

2023/05 – 2023/08

Software Development Engineer Intern

Boston, MA

- OpenSearch Serverless is a managed search and log analytics service that handles petabyte-scale data operations without the need to manually provision, manage, and scale computation resources.
- Implemented workflows for dynamically rebalancing replicas among indexing workers to ensure worker health and efficient ingestion. Demonstrated improvement in balance metrics after adding the workflow to the pipeline.

### Argo AI

2022/05 – 2022/08

Software Engineering Intern (Simulation)

Pittsburgh, PA

- Integrated new features into Argo's autonomous vehicle simulation pipelines to support the latest on-vehicle features.
- Investigate solutions to utilize public traffic datasets for automatic scenario generation.

## Teaching Experience

---

**Database Systems, Head Teaching Assistant** [15445.courses.cs.cmu.edu](https://15445.courses.cs.cmu.edu)

2022/08 - 2024/05

Carnegie Mellon University

Pittsburgh, PA

- Developed the SQL homework to help students learn advanced SQL features while getting familiar with two full-featured DBMS, SQLite and DuckDB.
- Lead the development of the BusTub course projects.

## Selected Projects

---

**Bustub** [github.com/cmu-db/bustub](https://github.com/cmu-db/bustub)

2022/08 - 2024/05

- Developed query execution course projects and improved testing infrastructure in the Bustub RDBMS.
- Designed and supported order by clause, top-k optimization, left outer joins, and composite-key hash joins in the query optimization and execution layer.
- Redesigned the disk-based extendible hash index project to support multi-level directories and thread-safe access.

**TicTacTOS**

2023/10 - 2023/12

- Implemented a UNIX-inspired preemptive kernel on x86 architecture from scratch. The kernel enables private process virtual address space through paging, supports multi-processing and multithreading through a scheduler, uses a zero-fill-on-demand allocation strategy, and provides a system call interface to the user.
- Developed a Pthreads-like thread library that allows users to run multithreaded C programs. Implemented a thread management interface and synchronization primitives like mutex, condition variable, semaphore, and reader-writer lock.

**Concurrent Adaptive Radix Tree** [yliang412.github.io/cart](https://yliang412.github.io/cart)

2024/04 - 2024/05

- Implemented a concurrent Adaptive Radix Tree, an efficient in-memory index data structure, using the optimistic lock coupling synchronization approach. Achieved 2x speedup for writer-only low contention scenarios and mixed read/write high contention scenarios.

## Awards

---

**Dean's List, School of Computer Science**

Spring 2022, Spring 2023