

# Derrick Quinn

## Curriculum Vitae

6340 NW Kelshar Dr  
Topeka, KS, 66618  
☎ +1 (785) 221 1636  
✉ dq55@cornell.edu  
in derrick-quinn

### Overview

**Area of interest:** Algorithm-hardware co-design of heterogeneous systems for data-intensive applications.

**Vision:** My research treats system-level interactions as upstream considerations that inform design, rather than downstream effects that occur once a design has been completed, which is the conventional approach. Following this approach, my vision is to develop generalized architectures for the building blocks of AI systems, promoting interoperability, adaptability, and scalability across diverse computing environments.

### Education

- 2024–2028 **Ph.D. Electrical and Computer Engineering**, *Cornell University*, Ithaca, NY  
Advisor: Mohammad Alian
- 2018–2024 **B.S. Computer Science**, *University of Kansas*, Lawrence, KS
- 2018–2024 **B.S. Mathematics**, *University of Kansas*, Lawrence, KS

### Expertise

- Classical Computer Architecture
- Design and Programming of Heterogeneous Architectures (e.g., Intel SPR)
- Modern Deep Learning Accelerator Design
- Hardware Evaluation Tools (e.g., Gem5, others)
- Systems for State-of-the-Art Deep Learning Algorithms (e.g., Transformers, Recommendation Systems)

### Professional Experience

- 2024–Current **Research Assistant**, *Department of Electrical and Computer Engineering, Cornell University*, Ithaca, NY
- 2022–2024 **Research Assistant**, *Department of Electrical Engineering and Computer Science, University of Kansas*, Lawrence, KS
- 2021–2022 **Algebra Tutor and Class Assistant**, *Kansas Algebra Program; Department of Mathematics, University of Kansas*, Lawrence, KS
- 2020 **Engineering Intern**, *Dell Technologies*, Overland Park, KS

### Publications

- [3] **Accepted: DReX: Accurate and Scalable Dense Retrieval Acceleration via Algorithmic-Hardware Codesign**, *ISCA: International Symposium on Computer Architecture (2025)*, Tokyo, JP  
Derrick Quinn\*, Emine Ezgi Yücel\*, Martin Prammer, Zhenxing Fan, Kevin Skadron, Jignesh Patel, José Martínez, Mohammad Alian
- [2] **Accelerated Retrieval Augmented Generation**, *ASPLOS: International Conference on Architectural Support for Programming Languages and Operating Systems (2025)*, Rotterdam, NL  
Derrick Quinn, Mohammad Nouri, Neel Patel, Alireza Salemi, Sukhan Lee, Hamed Zamani, Mohammad Alian
- [1] **XFM: Accelerated Software-Defined Far Memory**, *MICRO: International Symposium on Microarchitecture (2023)*, Toronto, Ontario, CA  
Neel Patel, Amin Mamandipoor, Derrick Quinn, Mohammad Alian