

MILAD REZAEI HAJIDEHI

12-2B Harvard Peabody Terrace Residence, Cambridge, Massachusetts, USA

☎ +1-857-242-9597 ✉ milad@g.harvard.edu [in linkedin.com/in/miladrzh](https://www.linkedin.com/in/miladrzh)

Education

Harvard University

2025 – 01/2029 (Anticipated)

Ph.D. in Computer Science (Fully Funded), GPA: 4/4, Supervisor: Prof. Stratos Idreos

Cambridge, MA, USA

University of British Columbia

2024

M.Sc. in Computer Science (Fully Funded), GPA: 92/100, Supervisor: Prof. Margo Seltzer

Vancouver, Canada

- Thesis: CUTTANA: Scalable Graph Partitioning for Faster Distributed Graph Databases and Analytics

Amirkabir University of Technology (Tehran Polytechnic)

2020

B.Sc. in Computer Engineering, GPA: 3.92/4.0 (18.20/20), Highest Honor (by GPA)

Tehran, Iran

Publications

SWESPOKE: Declare What You Want, Get a Bespoke Software Engineering Benchmark

Qitong Wang, Milad Rezaei Hajidehi, Divesh Srivastava, Stratos Idreos. **ICLR 2027** (Under review)

Agentic Context Cracking: Token-Efficient Data Reasoning Agents via Adaptive Structuring of Unstructured Data

Milad Rezaei Hajidehi, Qitong Wang, Stratos Idreos. **CIDR 2027** (Accepted, To appear). [arXiv:2608.31082](https://arxiv.org/abs/2608.31082)

CUTTANA: Scalable Graph Partitioning for Faster Distributed Graph Databases and Analytics

Milad Rezaei Hajidehi, Sraavan Sridhar, Margo Seltzer. **VLDB 2025**

ExtMem: Enabling Application-Aware Virtual Memory Management for Data-Intensive Applications

Sepehr Jalalian, Shaurya Patel, Milad Rezaei Hajidehi, Margo Seltzer, Alexandra Fedorova. **USENIX ATC 2024**

CS-TGN: Community Search via Temporal Graph Neural Networks

Farnoosh Hashemi*, Ali Behrouz*, Milad Rezaei Hajidehi*. **WWW 2023** Companion (Oral)

*Equal Contribution

Research Experience

Data/AI Systems Laboratory (DASlab) - Harvard University

2025 – Present

AI Systems Researcher

Cambridge, MA, USA

- **Research Focus:** Research on building a reliable data ecosystem for AI agents. spanning learning (post-training and RL), planning (abstractions), memory (persistent knowledge and reusable structures), and execution (GPU kernels and scheduling). Particular interest in highly accurate deep research and data reasoning across many documents (web/enterprise).
- First work, context cracking (CIDR 2027): efficient technique for agents to speculatively structure documents as they read and prefetch facts for later questions, cutting cost in half at equal accuracy with minimal overhead.
- Co-design and instruct Harvard's graduate AI Systems course. Design and evaluate the semester project: an LLM from scratch in CUDA. Lead the GPU and LLM-operator lectures.

Systopia Lab - UBC

3 Years

Systems Researcher

Vancouver, Canada

- **Research Focus:** Built systems for large-scale data-intensive applications: reducing network overhead in distributed settings (VLDB 2025), customizing the OS memory manager from user space in out-of-core settings (USENIX ATC 2024), and finding communities in large graphs (WWW 2023).

Industry Experience

Divar - Cafebazaar Inc.

2.5 Years

Software Engineer (Infra) - Owner of the Payments Microservice

Tehran, Iran

- Built a **1TB/hour** log pipeline for **30M users**. Migrated payments to a microservice with **zero downtime**.

Honors & Awards

Best MSc Thesis in the Department | UBC CS; nominated for the Governor General of Canada's Gold Medal

2024

3rd Place | ICPC Pacific Northwest Regional (Division 1)

2020

Top 500 at Meta Hacker Cup | among 34,000 participants

2021

Bronze Medal | ICPC Asia West Regional

2018

Top 0.1% | Iranian University Entrance Exam

2015

Neekeyfar Award for Graduate Study | \$6000, Additional to Harvard PhD Fellowship

2025

Selected Projects

LLM From Scratch | *Implementation of Llama 3 in CUDA from scratch.*

GPU Superoptimizing Compiler | *DSL and Monte Carlo tree search that compiles DSL programs directly into PTX.*

Cuttana Graph Partitioner | *Billion-edge streaming graph partitioner in C++ & OpenMP (VLDB 2025).* [GitHub](#)

Hardware-Conscious Columnar Storage Engine | *Parallel execution, shared scans, and indexes, written in C.*

Distributed KV Store with Raft | *Go, Raft Protocol*

Technical Skills

Languages: C/C++, Python, Go, SQL, Bash, JavaScript

GPU & HPC: CUDA/PTX, OpenMP, Nsight, vLLM, Slurm/FASRC

Frameworks & Technologies: PyTorch, Spark, Airflow, HDFS, Kubernetes, Docker, AWS, Azure, Graph Databases

Selected Relevant Courses

Ph.D. (Harvard University)

Big Data Systems (Harvard CS265) (4/4), Advanced Algorithms (Harvard CS221) (4/4), Program Synthesis (MIT 6.5130) (4/4), Data Systems Design (Harvard CS165) (4/4), Modern Storage Systems (Harvard CS264) (4/4)

M.Sc. (University of British Columbia)

Multimodal Learning: Vision, Language, Audio (99/100), Machine Learning (97/100), Distributed Systems Abstraction (97/100), Programming Languages (90/100)

Undergraduate Math

Foundation of Combinatorics (19/20), Calculus I (17.45/20), Calculus II (20/20), Differential Equations (17.03/20), Engineering Probability & Statistics (19.5/20), Engineering Mathematics (17.81/20), Engineering Economics (20/20)

Undergraduate Computer Science

Data Structures and Algorithms (20/20), Fundamentals of Programming (20/20), Advanced Programming (19.5/20), Principles of Database Design (20/20), Design of Algorithms (20/20), Artificial Intelligence (19.32/20), Computer Networks I (18.6/20), Internet Engineering (19/20), Design Automation (FPGA) (20/20)

Miscellaneous

Reviewer, VLDB Journal **2025, 2026**

10× TA/TF: Data Structures (5×), Databases (2×), Distributed Systems, Algorithms, AI Systems Design **2017 – Present**

Varsity roster, Harvard Table Tennis Team; competed in NCTTA **2025, 2026**

Languages

Persian: Native **English:** TOEFL iBT 109/120 (R:27, L:28, S:25, W:29)

References

Professor Margo Seltzer
Professor Stratos Idreos

mseltzer@cs.ubc.ca
stratos@seas.harvard.edu