# Sara McAllister

*PhD Candidate, Carnegie Mellon University*

sjmcalli@cs.cmu.edu | saramcallister.github.io | saramcallister

## Biography

Sara McAllister is a PhD candidate at Carnegie Mellon University, advised by Nathan Beckmann and Greg Ganger. She is interested in computer systems, particularly caching and storage systems. Her work includes a focus on improving efficiency and sustainability through hardware-software co-design and grounding design choices in mathematical modeling. Her work has appeared at OSDI and SOSP, including receiving a Best Paper Award at SOSP 2021 for her paper "Kangaroo: Caching Billions of Tiny Objects on Flash". She is a 2021 NDSEG fellow and a 2023 EECS Rising Star. Sara also strives to increase inclusion in computer science, including by creating a DEI course for CS PhD students. Due to these efforts, she was awarded CMU's Graduate Student Service Award in 2022 and a Best Paper Award at SIGCSE 2023.

## Education

**Carnegie Mellon University** — *Pittsburgh, PA*

PhD in Computer Science, Advisors: Nathan Beckmann and Greg Ganger — *Aug 2019. - Summer 2025 (Expected)*

**Carnegie Mellon University** — *Pittsburgh, PA*

Masters in Computer Science Research — *Aug 2019. - May 2022*

**Harvey Mudd College** — *Claremont, CA*

B.S. in Computer Science, Graduated with High Distinction — *Aug. 2015 - May 2019*

## Honors and Awards

| | | |
|---|---|---|
| 2023 | **Rising Star in EECS** | *EECS* |
| 2023 | **Exemplary Poster Presentation**, In computer and computational sciences at fellows conference | *NDSEG* |
| 2023 | **Best Paper Award** | *SIGCSE* |
| 2022 | **Graduate Student Service Award**, For the development of 15-996 CS-JEDI | *CMU* |
| 2021 | **Best Paper Award** | *SOSP* |
| 2021 | **Graduate Fellowship**, NDSEG | *DoD* |
| 2021 | **Graduate Research Fellowship**, GRFP | *NSF* |
| 2019 | **Class of '94 Award**, Outstanding CS graduate in a combination of course work, research, and service | *Harvey Mudd* |
| 2019 | **Departmental Honors**, Computer Science Department | *Harvey Mudd* |
| 2019 | **Clinic Team Award**, Outstanding performance on an industry-sponsored team capstone project | *Harvey Mudd* |
| 2019 | **Outstanding Undergraduate Researcher Award**, Honorable Mention | *CRA* |
| 2019 | **Best Malware**, Most creative malware during capture the flag (CTF) competition | *Yelp* |

## Publications

**DéjàVu: KV-cache Streaming for Fast, Fault-tolerant Generative LLM Serving** — *ICML 2024*

Fonteini Strati, **Sara McAllister**, Amar Phanishayee, Jakub Tarnawski, Ana Klimovic — *Acceptance Rate: 27.5%*

**Towards Understanding the Carbon Impact in End-to-end Sensing Pipelines** — *HotEthics 2024*

Harsh Desai*, **Sara McAllister***, Nathan Beckmann, Brandon Lucia (* = co-first author)

**CS-JEDI: Required DEI Education, by CS PhD Students, for CS PhD Students** — *SIGCSE 2023*

Bailey Flanigan, Ananya Joshi, **Sara McAllister**, Catalina Vajiac — *Acceptance Rate: 35%*

**Kangaroo: Theory and Practice of Caching Billions of Tiny Objects on Flash** — *ACM ToS*

**Sara McAllister**, Benjamin Berg, Julian Tutuncu-Macias, Juncheng Yang, Sathya Gunasekar, Jimmy Lu, Daniel S. Berger, Nathan Beckmann, Gregory R. Ganger — *August 2022*

**Kangaroo: Caching Billions of Tiny Objects on Flash** — *SOSP 2021*

**Sara McAllister**, Benjamin Berg, Julian Tutuncu-Macias, Juncheng Yang, Sathya Gunasekar, Jimmy Lu, Daniel S. Berger, Nathan Beckmann, Gregory R. Ganger — *Acceptance Rate: 16%*

**External-memory Dictionaries in the Affine and PDAM Models** — *ACM ToPC*

Michael A. Bender, Alex Conway, Martin Farach-Colton, William Jannen, Yizheng Jiao, Rob Johnson, Eric Knorr, **Sara McAllister**, Nirjhar Mukherjee, Prashant Pandey, Donald E. Porter, Jun Yuan, Yang Zhan — *September 2021*

**The CacheLib Caching Engine: Design and Experiences at Scale** *OSDI 2020*
Benjamin Berg, Daniel S. Berger, **Sara McAllister**, Isaac Grosof, Sathya Gunasekar, Jimmy Lu, Michael Uhlar, Jim Carrig, Nathan Beckmann, Mor Harchol-Balter, Gregory R. Ganger *Acceptance Rate: 18%*

**Small Refinements to DAM Can Have Big Consequences for Data-Structure Design** *SPAA 2019*
Michael A. Bender, Alexander Conway, Martin Farach-Colton, William Jannen, Yizheng Jiao, Rob Johnson, Eric Knorr, **Sara McAllister**, Nirjhar Mukherjee, Prashant Pandey, Donald E. Porter, Jun Yuan, Yang Zhan *Acceptance Rate: 40%*

## Talks

**Overcoming Write Limitations to achieve Sustainable Flash Caching**
AMD (Remote) – *Research and Advance Development (RAD) and Xilinx Labs* *29 Mar 2024*
Salesforce (Remote) – *Database Reading Group* *27 Mar 2024*
UC Berkeley – *Hosted by Natacha Crooks* *25 Jan 2024*
Stanford – *Hosted by Keith Winstein* *24 Jan 2024*
UC Santa Cruz – *Hosted by Andrew Quinn* *11 Jan 2024*
McGill (Remote) – *Hosted by Oana Balmau* *16 Nov 2023*
Microsoft Pittsburgh – *Hosted by Jeff Butler* *2 Nov 2023*
MIT – *Hosted by Frans Kaashoek* *10 Oct 2023*
NDSEG 2021 Fellows Conference – *Recieved best poster-presentation award* *31 July 2023*
University of Toronto – *Hosted by Bianca Schroeder* *20 Mar 2023*

**Scaling the bandwidth-per-TB wall with Declarative Storage Interfaces**
PDL Retreat – *Presented to a large group of industry attendees* *6 Nov 2023*

**FairyWREN: A Sustainable Cache for Write-Read-Erase Interfaces**
PDL Retreat – *Presented to a large group of industry attendees* *7 Nov 2023*
PDL Retreat – *Presented to a large group of industry attendees* *7 Nov 2022*

**CS-JEDI: DEI education by PhD students, for PhD students**
McGill (Remote) – *Hosted by Oana Balmau* *31 Oct 2023*

**Caching on Flash: Kangaroo and Beyond**
Meta (Remote) – *Core Data Tech Talk* *11 Mar 2022*

**Kangaroo: Caching Billions of Objects on Flash**
Microsoft Research (Remote) – *Hosted by Daniel Berger* *22 Nov 2021*
SOSP (Remote) *27 Oct 2021*
Cache@Scale (Remote) – *Industry Caching Meetup hosted by Meta* *4 Mar 2021*

**Building a Stronger, More Just Academic Community Through Mandatory Anti-bias Learning**
University of Pittsburgh Diversity Forum (Remote) – *Co-presented w/ Bailey Flanigan and Catalina Vajiac* *28 July 2021*

## Teaching

**Carnegie Mellon University**
Storage Systems (15-746/18-746) *TA, Fall 2023*
Parallel Computer Architecture and Programming (15-418/618) *TA, Spring 2022*
Diversity, Equity, and Inclusion in Computer Science and Society (15-996) *Co-Creator and TA, Spring 2021*

**Harvey Mudd College**
Programming Languages (CS131) *Grader and Tutor, Spring 2019*
Introduction to Computer Systems (CS105) *Grader and Tutor, Fall 2018*
Introduction to Computer Systems (CS105) *Grader and Tutor, Spring 2018*
Data Structures and Programming Development (CS70) *Grader and Tutor, Fall 2017*
Principles of Computer Science (CS60) *Grader and Tutor, Spring 2017*
Introduction to Biology and Computer Science (CS5 Green) *Grader and Tutor, Fall 2016*

**Guest Lecturer**

| | |
|---|---|
| Storage Systems – Overcoming Flash's Write Limitations to Achieve Sustainable Caching (CMU 15/18-746) | *Fall 2023* |
| Graduate Computer Architecture – Sustainable Computing (CMU 15-740) | *Fall 2023* |
| Computer Systems – Kangaroo: Caching Billions of Tiny Objects on Flash (CMU 18-213/613) | *Fall 2022* |
| Data Center Computing – Kangaroo Discussion (CMU 18-847C) | *Spring 2022* |
| CS-JEDI – Panel on Allyship (CMU 15-996) | *Spring 2022* |
| Computer Systems – Kangaroo: Caching Billions of Tiny Objects on Flash (CMU 18-213) | *Fall 2021* |
| Storage Systems – Kangaroo: Caching Billions of Tiny Objects on Flash (CMU 18-746) | *Fall 2021* |

## Mentoring

| | |
|---|---|
| **Lucy Wang.** CMU ECE undergraduate student | *Spring 2024 - Present* |
| **Suhas Thalanki.** CMU computational data science masters capstone | *Spring 2024 - Present* |
| **Sriya Ravi.** CMU computational data science masters capstone | *Spring 2024 - Present* |
| **Yu Liu.** CMU computational data science masters capstone | *Spring 2024 - Present* |
| **Sophia (Qingyang) Cao.** CMU CS undergraduate student | *Fall 2023 - Present* |
| **Sarvesh Tandon.** CMU ECE masters student | *Fall 2023 - Present* |
| **Sherry (Yucong) Wang.** CMU ECE undergraduate student | *Fall 2022 - Present* |
| **Akshath Karanam.** CMU ECE masters student, After degree: Salesforce | *Fall 2022* |
| **Priyal Suneja.** Univesity of Washington CS PhD student | *Fall 2021 - Summer 2022* |
| **Julian Tutuncu-Macias.** CMU CS undergraduate student, After degree: Goldman Sachs | *Fall 2019 - Spring 2021* |
| **Sheng Xu.** CMU CS masters student, After degree: Amazon Web Services | *Spring 2020* |
| **Karina Mejia.** Ontario High School | *Summer 2016* |

## Leadership and Service

**Faculty Hiring Committee** — *Carnegie Mellon University*
One of two student committee members helping solicit student perspectives on faculty candidates — *Spring 2024*

**CS-JEDI and Other DEI Initiatives** — *Carnegie Mellon University*
Developed and implemented inclusivity initiatives with 2 other PhD students including an informal climate survey, a mandatory DEI class for CS PhD students, an advisor-advisee feedback form, and being awarded CMU's Graduate Student Service Award — *July 2020 - Mar. 2023*

**PhD Admissions Committee** — *Carnegie Mellon University*
PhD student in charge of reading applications for systems area in the Computer Science Department — *Dec. 2021 - Mar. 2022*

**Parallel Data Lab (PDL) Meeting Coordinator** — *Carnegie Mellon University*
Invited and scheduled talks for PDL weekly talk series — *Fall 2021*

**Introductory Course (IC) Committee** — *Carnegie Mellon University*
Co-organizer for first virtual orientation in the Computer Science Department — *Fall 2020*

**Faculty Search - Student Committee** — *Harvey Mudd College*
Interviewed each invited faculty candidate for the Computer Science Department — *Spring 2019*

**Mentor and Proctor (Residential Assistant)** — *Harvey Mudd College*
Led residential activities and crisis response in East Dorm with 82 residents — *Fall 2016 - Spring 2019*

**Science Bus Volunteer and Treasurer** — *Harvey Mudd College*
Instructed 4th and 5th graders from under-resourced schools in hands-on science lessons and managed ∼$3000 of grant money (April 2016 - May 2017) — *Aug. 2015 - May 2018*

**STEAM:coders Site Coordinator and Instructor** — *Harvey Mudd College*
Led CS-related activities for 25 middle-school ages students from disadvantaged communities — *Summer 2016*

## Professional Experience

**Graduate Research Assistant** — *Carnegie Mellon University*
Advisors: Nathan Beckmann and Greg Ganger — *Aug. 2019 - Present*

- Researched caching systems to decrease cost and increase sustainability of providing internet services at scale
- Explored new memory and storage hardware interfaces, particularly for caching applications

### Research Intern

*Microsoft Research*

Mentor: Amar Phanishayee

*Summer 2022*

- Researched serving large generative ML models more efficiently

### Research Intern

*Microsoft Research*

Mentor: Daniel Berger

*Summer 2021*

- Researched in-kernel disaggregated memory solutions using CXL

### Software Engineering Intern

*Yelp*

Database Team

*Summer 2019*

- Designed and implemented a Python library to manage MySQL database permissions
- Planned and started gradual roll out system, fully rolled out after internship across production

### Clinic (Capstone) Project

*Harvey Mudd College*

Sponsored by Pure Storage

*Aug. 2018 - May 2019*

- Designed and implemented failover mechanisms for NFS VMs running on a two-controller system
- Technical lead, about file systems and network partitioning, on a team of 4

### Undergraduate Research Assistant

*UNC Chapel Hill*

Advisor: Don Porter

*May 2018 - Aug. 2018*

- Investigated theoretical and experimental analysis of write-optimized dictionaries

### Software Engineering Intern

*Facebook*

Devleper Experience Team

*Summer 2017*

- Developed and tested a Python library to restart and repair development servers
- Created a React and Hack PHP user interface to receive and store user inputs

### Research Assistant

*Harvey Mudd College*

Advisor: Anna Ahn

*May. 2016 - Jun. 2017*

- Led data analysis of a three-legged walking study