Sara McAllister

PhD Candidate, Carnegie Mellon University

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

Research Interests

Sara McAllister is a PhD candidate at Carnegie Mellon University, advised by Nathan Beckmann and Greg Ganger, expected to graduate in Summer 2025. 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". 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. - Presennt*

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 Electrical Engineering and Computer Science (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 a team	*Harvey Mudd*
2019	**Outstanding Undergraduate Researcher Award**, Honorable Mention	*CRA*
2019	**Best Malware**, Most creative malware during capture the flag (CTF) competition	*Yelp*

Refereed Journal Publications

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*

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*

Refereed Conference Publications

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: 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%*

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 and Posters

Efficient and Sustainable Data Retrieval — *NDSEG Fellows, 31 July 2023*
Overcoming Write Limitations to achieve Sustainable Flash Caching — *University of Toronto, 20 Mar 2023*
FairyWREN: A Superb Cache Co-optimized for Write-Limited Flash — *PDL Retreat, 7 Nov 2022*
Caching on Flash: Kangaroo and Beyond — *Meta, 11 Mar 2022*
Kangaroo: Caching Billions of Objects on Flash — *Microsoft Research, 22 Nov 2021*
Kangaroo: Caching Billions of Objects on Flash — *SOSP, 27 Oct 2021*
Building a Stronger, More Just Academic Community Through Mandatory Anti-bias Learning — *UPitt Diversity Forum, 28 Jul 2021*
Kangaroo: Caching Billions of Objects on Flash — *Cache@Scale, 4 Mar 2021*

Leadership and Service

CS-JEDI and Other DEI Initiatives — *Carnegie Mellon University, July 2020 - Mar. 2023*
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

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

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

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

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

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

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

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

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*

Invited Speaker

Computer Systems (CMU 18-213/613) — *Fall 2022*
Data Center Computing (CMU 18-847C) — *Spring 2022*

CS-JEDI: Intro to Justice, Equity, Diversity, and Inclusion in Computer Science (CMU 15-996) *Spring 2022*
Computer Systems (CMU 18-213) *Fall 2021*
Storage Systems (CMU 18-746) *Fall 2021*

Mentoring

Yucong (Sherry) Wang (CMU ECE undergraduate student) *Fall 2022 - Present*
Akshath Karanam (CMU ECE masters student) *Fall 2022*
Priyal Suneja (Univesity of Washington CS PhD student) *Fall 2021 - Summer 2022*
Julian Tutuncu-Macias (CMU CS undergraduate student) *Fall 2019 - Spring 2021*
Sheng Xu (CMU CS masters student) *Spring 2020*
Karina Mejia (Ontario High School) *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