

# Sara McAllister

PhD Candidate

Carnegie Mellon University

✉ sjmcalli@cs.cmu.edu | 🏠 saramcallister.github.io | 🎓 saramcallister

## Research Interests

I am interested in computer systems, particularly caching and storage systems. My work includes a focus on improving efficiency and sustainability through hardware-software co-design and grounding design choices in mathematical modeling. My work has appeared at OSDI and SOSP, including receiving a Best Paper Award at SOSP 2021. I am a 2021 NDSEG fellow and a 2023 EECS Rising Star. I also strive to increase inclusion in computer science, including by creating a DEI course for CS PhD students. Due to these efforts, I was awarded CMU's Graduate Student Service Award in 2022 and a Best Paper Award at SIGCSE 2023.

## Education

### Carnegie Mellon University

PHD IN COMPUTER SCIENCE, ADVISORS: NATHAN BECKMANN AND GREG GANGER

Pittsburgh, PA

Aug 2019. - Summer 2025 (Expected)

### Carnegie Mellon University

MASTERS IN COMPUTER SCIENCE RESEARCH

Pittsburgh, PA

Aug 2019. - May 2022

### Harvey Mudd College

B.S. IN COMPUTER SCIENCE, GRADUATED WITH HIGH DISTINCTION

Claremont, CA

Aug. 2015 - May 2019

## Honors and Awards

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

## Publications

### FairyWREN: A Sustainable Cache for Emerging Write-Read-Erase Flash Interfaces

Sara McAllister, Yucong Wang, Benjamin Berg, Daniel S. Berger, George Amvrosiadis, Nathan Beckmann, Gregory R. Ganger

OSDI 2024

Acceptance Rate: 18%

### A Call for Research on Storage Emissions

Sara McAllister, Fiodar Kazhamiaka, Daniel S. Berger, Rodrigo Fonseca, Kali Frost, Aaron Ogus, Maneesh Sah, Ricardo Bianchini, George Amvrosiadis, Nathan Beckmann, Gregory R. Ganger

HotCarbon 2024

Acceptance Rate: 46%

### DéjàVu: KV-cache Streaming for Fast, Fault-tolerant Generative LLM Serving

Fonteini Strati, Sara McAllister, Amar Phanishayee, Jakub Tarnawski, Ana Klimovic

ICML 2024

Acceptance Rate: 27.5%

### Towards Understanding the Carbon Impact in End-to-end Sensing Pipelines

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

HotEthics 2024

### CS-JEDI: Required DEI Education, by CS PhD Students, for CS PhD Students

Bailey Flanigan, Ananya Joshi, Sara McAllister, Catalina Vajiac

SIGCSE 2023

Acceptance Rate: 35%

### Kangaroo: Theory and Practice of Caching Billions of Tiny Objects on Flash

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

ACM ToS

August 2022

## Kangaroo: Caching Billions of Tiny Objects on Flash

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

🏆 SOSP 2021

Acceptance Rate: 16%

## External-memory Dictionaries in the Affine and PDAM Models

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

ACM ToPC

September 2021

## The CacheLib Caching Engine: Design and Experiences at Scale

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

OSDI 2020

Acceptance Rate: 18%

## Small Refinements to DAM Can Have Big Consequences for Data-Structure Design

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

SPAA 2019

Acceptance Rate: 40%

## Talks

---

### FairyWREN: A Sustainable Cache for Write-Read-Erase Interfaces

OSDI

12 July 2024

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

### A Call for Research on Storage Emissions

HotCarbon

9 July 2024

### Towards Understanding the Carbon Impact in End-to-end Sensing Pipelines

HotEthics – Co-presented with Harsh Desai

29 Apr 2024

### 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

### 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

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

## Service

### External Review Committee Member

USENIX Annual Technical Conference (ATC)	2024
--	------

### Faculty Hiring Committee

Carnegie Mellon University, Computer Science Department	2024
Harvey Mudd College, Computer Science Department	2019

### PhD Admissions

Carnegie Mellon University, Computer Science Department	2022
---	------

### Student Organizer

DEI initiatives in CMU’s CS Department – <i>Informal Survey, CS-JEDI course, advisor-advisee feedback form</i>	2020-2023
Parallel Data Lab (PDL) Meeting Coordinator	2021
PhD Orientation Committee – <i>CMU CS Department’s Introductory Course (IC)</i>	2020

## Outreach

<b>Science Bus Volunteer and Treasurer</b> – Harvey Mudd College	2015 - 2018
--	-------------

- Instructed 4th and 5th graders from under-resourced schools in hands-on science lessons
- Managed ~\$3000 of grant money (April 2016 - May 2017)

- Led CS-related activities for 25 middle-school ages students from disadvantaged communities

## 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