Sara McAllister

Doctoral Candidate
Carnegie Mellon University
Computer Science Department
5000 Forbes Avenue, Pittsburgh, PA 15213

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

Biography

I research sustainable datacenters from a computer systems perspective, particularly focused on caching and storage systems. My work includes improves efficiency and sustainability through hardware-software co-design and grounding design 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, a 2023 EECS Rising Star, and a 2025 Siebel Scholar. For my service work, I 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

2025 Siebel Scholar, for outstanding academic performance and leadership — *2024*
2023 Rising Star, in Electrical Engineering & Computer Science (EECS) — *2023*
NDSEG Exemplary Poster Presentation, in computer and computational sciences at NDSEG fellows conference — *2023*
SIGCSE Best Paper Award, for CS-JEDI paper — *2023*
CMU Graduate Student Service Award, for the development of 15-996 CS-JEDI — *2022*
SOSP Best Paper Award, for Kangaroo paper — *2021*
NDSEG Graduate Fellowship, DoD sponsored 3-year fellowship — *2021*
NSF Graduate Research Fellowship (GRFP), NSF sponsored 3-year fellowship — *2021*
Harvey Mudd Class of '94 Award, for an outstanding CS graduate in coursework, research, and service — *2019*
Harvey Mudd Computer Science Departmental Honors — *2019*
Harvey Mudd Clinic Team Award, for outstanding performance on an industry-sponsored team capstone project — *2019*
CRA Outstanding Undergraduate Researcher Award, Honorable Mention — *2019*

Employment History

Assistant Professor, University of Wisconsin-Madison – *Madison, Wisconsin* — *Starting Aug 2026*

Visiting Faculty Researcher, Google – *Sunnyvale, CA* — *Starting Sept 2025*
Managers: David Culler and Hank Levy
Researching large-scale data systems for emerging applications

Graduate Research Assistant, Carnegie Mellon University – *Pittsburgh, PA* — *Aug 2019 - Aug 2025*
Advisors: Nathan Beckmann and Greg Ganger
Researched reducing IO to create sustainable caching and storage systems at scale

Research Intern, Microsoft Research – *Redmond, WA* — *Summer 2022*
Mentor: Amar Phanishayee
Researched serving LLMs more efficiently especially under failure

Research Intern, Microsoft Research – *Redmond, WA* — *Summer 2021*
Mentor: Daniel Berger
Researched in-kernel disaggregated CXL memory solutions

Software Engineering Intern, Yelp – *San Francisco, CA* — *Summer 2019*
Database Team
Designed, implemented, and rolled out a library to manage MySQL database permissions

Clinic (Capstone) Project, Harvey Mudd College – *Claremont, CA* — *Aug. 2018 - May 2019*
Sponsored by Pure Storage
Technical lead for team of 4 designing failover mechanisms for NFS VMs running on a two-controller system

Undergraduate Research Assistant, UNC Chapel Hill – *Chapel Hill, NC* — *May 2018 - Aug. 2018*
Advisor: Don Porter
Researched theoretical and experimental analysis of write-optimized dictionaries

Software Engineering Intern, Facebook – *Menlo Park, CA* — *Summer 2017*
Developer Experience Team
Developed a Python library to restart and repair development servers

Research Assistant, Harvey Mudd College – *Claremont, CA* — *May. 2016 - Jun. 2017*
Advisor: Anna Ahn
Researched three-legged walking and led data analysis of wearable devices

Peer-Reviewed Publications

FairyWREN: A Sustainable Cache for Emerging Write-Read-Erase Flash Interfaces — *ACM TOS*
Sara McAllister, Yucong Wang, Benjamin Berg, Daniel S. Berger, George Amvrosiadis, Nathan Beckmann, Gregory R. Ganger — *2025*

A Call for Research on Storage Emissions — *Energy Informatics Review*
Sara McAllister, Fiodar Kazhamiaka, Daniel S. Berger, Rodrigo Fonseca, Kali Frost, Aaron Ogus, Maneesh Sah, Ricardo Bianchini, George Amvrosiadis, Nathan Beckmann, Gregory R. Ganger — *2025*

FairyWREN: A Sustainable Cache for Emerging Write-Read-Erase Flash Interfaces — *OSDI 2024*
Sara McAllister, Yucong Wang, Benjamin Berg, Daniel S. Berger, George Amvrosiadis, Nathan Beckmann, Gregory R. Ganger — *Acceptance Rate: 18%*

A Call for Research on Storage Emissions — *HotCarbon 2024*
Sara McAllister, Fiodar Kazhamiaka, Daniel S. Berger, Rodrigo Fonseca, Kali Frost, Aaron Ogus, Maneesh Sah, Ricardo Bianchini, George Amvrosiadis, Nathan Beckmann, Gregory R. Ganger — *Acceptance Rate: 46%*

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

Taking off the compute-colored glasses: Storage is vital to datacenter sustainability — *5 June 2025*
PDL Summer Talk Series – *Pittsburgh, PA*

Stanford University – *Stanford, CA* *22 Apr 2025*
Microsoft Research – *Redmond, WA* *21 Apr 2025*
University of Wisconsin, Madison – *Madison, WI* *1 Apr 2025*
University of Toronto – *Toronto, Canada* *27 Mar 2025*
Cornell Tech – *New York City, NY* *20 Mar 2025*
MIT – *Boston, MA* *6 Mar 2025*
University of Maryland, College Park – *College Park, MD* *27 Feb 2025*
Google – *Sunnyvale, CA* *25 Feb 2025*

Scaling the bandwidth-per-TB wall with Declarative Storage Interfaces
Salesforce – *Database Reading Group, San Francisco, CA* *4 Dec 2024*
PDL Retreat – *Bedford, PA* *15 Oct 2024*
PDL Retreat – *Bedford, PA* *6 Nov 2023*

A Call for Research on Storage Emissions
PDL Retreat – *Bedford, PA* *15 Oct 2024*
Western Digital – *San Jose, CA* *5 Sep 2024*
HotCarbon – *Santa Cruz, CA* *9 July 2024*

FairyWREN: A Sustainable Cache for Write-Read-Erase Interfaces
OSDI – *San Jose, CA* *12 July 2024*
PDL Retreat – *Bedford, PA* *7 Nov 2023*
PDL Retreat – *Pittsburgh, PA* *7 Nov 2022*

Towards Understanding the Carbon Impact in End-to-end Sensing Pipelines
HotEthics – *Co-presented with Harash Desai, San Diego, CA* *29 Apr 2024*

Overcoming Write Limitations to achieve Sustainable Flash Caching
AMD – *Santa Clara, CA* *29 Mar 2024*
Salesforce – *Database Reading Group, San Francisco, CA* *27 Mar 2024*
University of California, Berkeley – *Berkeley, CA* *25 Jan 2024*
Stanford – *Stanford, CA* *24 Jan 2024*
University of California, Santa Cruz – *Santa Cruz, CA* *11 Jan 2024*
McGill University – *Montreal, Canada* *16 Nov 2023*
Microsoft – *Pittsburgh, PA* *2 Nov 2023*
MIT – *Boston, MA* *10 Oct 2023*
NDSEG 2021 Fellows Conference – *San Antonio, TX* *31 July 2023*
University of Toronto – *Toronto, Canada* *20 Mar 2023*

CS-JEDI: DEI education by PhD students, for PhD students
McGill – *Montreal, Canada* *31 Oct 2023*

Caching on Flash: Kangaroo and Beyond
Meta – *Menlo Park, CA* *11 Mar 2022*

Kangaroo: Caching Billions of Objects on Flash
Microsoft Research – *Redmond, WA* *22 Nov 2021*
SOSP – *Remote* *27 Oct 2021*
Cache@Scale – *Industry Remote 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 – *Co-presented w/ Bailey Flanigan and Catalina Vajiac, Pittsburgh, PA* *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

Computer Systems – Sustainable Computing (CMU 15-213/513)	*Fall 2024*
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

Theo Gregersen. CMU CS PhD student	*Fall 2024 - Present*
Yiwei Chen. CMU ECE masters student	*Fall 2024 - Present*
Tim Kim. CMU CS PhD student	*Spring 2024 - Present*
Sanjith Athlur. CMU CS PhD student	*Spring 2024 - Present*
Lucy Wang. CMU ECE undergraduate student	*Spring 2024 - Present*
Sophia (Qingyang) Cao. CMU CS undergraduate student	*Fall 2023 - Present*
Sarvesh Tandon. CMU ECE masters student	*Fall 2023 - Present*
Suhas Thalanki. CMU computational data science masters capstone	*Spring 2024 - Fall 2024*
Sriya Ravi. CMU computational data science masters capstone	*Spring 2024 - Fall 2024*
Yu Liu. CMU computational data science masters capstone	*Spring 2024 - Fall 2024*
Sherry (Yucong) Wang. CMU ECE undergraduate student, After degree: Salesforce	*Fall 2022 - Spring 2024*
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*

Service

Reviewer

Transactions on Storage (TOS)	*2024*

External Reviewer

ACM SIGMETRICS	*2025*
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 – *Informal Survey, CS-JEDI course, advisor-advisee feedback form*	*2020-2023*
Parallel Data Lab (PDL) Meeting Coordinator	*2021*
PhD Orientation Committee – *CMU CS Department's Introductory Course (IC)*	*2020*

Community Outreach

TechNights volunteer – *CS program for middle school girls* *2019-2020*

AP CS Remote Talk at Eagan High School – *On CS career opportunities* *2021*

Science Bus Volunteer and Treasurer – *Teaching 4/5th graders at under-resourced schools science lessons* *2015 - 2018*

STEAM:coders Site Coordinator and Instructor – *CS program for students from disadvantaged communities* *2016*