# Jayjeet Chakraborty

jayjeetc.github.io | jayjeetc@ucsc.edu | github.com/JayjeetAtGithub

## Objective

To explore the field of Big data processing, Storage systems, and Distributed systems, perform research and solve problems, and spend a career building technologies for tackling critical Systems challenges.

## Education

- **University of California, Santa Cruz** — Santa Cruz, CA
  *PhD, Computer Science And Engineering* — *Expected Grad: June, 2026*

- **National Institute Of Technology, Durgapur** — Durgapur, WB, India
  *B.Tech, Computer Science And Engineering, CGPA: 7.65/10* — *Graduated: June, 2021*

- **Hem Sheela Model School, Durgapur** — Durgapur, WB, India
  *Senior Secondary Education, CBSE, Percentage: 92.6%* — *Graduated: April, 2017*

## Experience

- **IRIS-HEP Winter Fellow** — Remote
  *Princeton University and CROSS, UC Santa Cruz* — *Jan 2021 - July 2021*
  - Completed working on SkyhookDM, an Arrow-Native storage system based on Ceph. Integrated Coffea, a distributed scientific data processing framework from CERN with Skyhook to enable compute offloading in HEP data analytics. Also, working on a Pull request to contribute Skyhook to the Apache Arrow project. [ArXiv]

- **IRIS-HEP Summer Fellow** — Remote
  *Princeton University and CROSS, UC Santa Cruz* — *June 2020 - October 2020*
  - Built scalable and reproducible Popper workflows for running experiments on large datasets stored in a SkyhookDM cluster. Also, performed experiments on a SkyhookDM deployment and studied the performance gains due to push-down. [report]
  - Redesigned SkyhookDM to store and process data in Arrow IPC and Parquet format and also extended the Arrow framework with a Skyhook Dataset API to be able to natively connect to a Ceph/RADOS cluster with SkyhookDM plugins and push-down compute tasks.

- **Winter Research Intern** — Varanasi, India
  *Indian Institute Of Technology, BHU* — *December 2019 - January 2020*
  - Studied and analyzed several terrain rendering techniques and implemented the ROAM(Real-time Optimally Adapting meshes) and Incremental Delaunay Triangulation algorithms for Level Of Detail based rendering of large terrain datasets. [report]
  - Developed a visualization tool in C++ using OpenGL to render terrains from Li-DAR datasets at 60 fps and benchmarked both the algorithm implementations on GPU.

- **Google Summer Of Code Student** — Remote
  *Centre for Research in Open Source Software, UC Santa Cruz* — *May 2019 - August 2019*
  - Extended the Popper workflow engine by adding support for additional container runtimes like Singularity, added more sub-commands, implemented concurrent execution capabilities, and other CI/CD features. Also, wrote unit tests to achieve an 87% test coverage and added documentation for the newly added features. [report]
  - Contributed plugins to facilitate the execution of Popper workflows on Virtual machines, Kubernetes clusters, and Slurm based HPC clusters.

- **Software Engineering Intern** — Mumbai, India
  *LogN Software* — *May 2018 - July 2018*

- ○ Worked on developing the backend APIs for a client-facing Ionic application using Django Rest Framework and MySQL. Also, built and integrated parts of UI of the application. Also, built the payment gateway of the application using the Stripe Payments SDK.

## PUBLICATIONS

- Jayjeet Chakraborty, Carlos Maltzahn, Ivo Jimenez. Enabling Seamless Execution of Computational and Data Science Workflows on HPC and Cloud with the Popper Container-Native Automation Engine. Paper at CANOPIE-HPC Workshop 2020, 12 November, 2020. [paper]

- Jayjeet Chakraborty, Ivo Jimenez, Carlos Maltzahn, Arshul Mansoori, Quincy Wofford. Popper 2.0: A Container-native Workflow Execution Engine For Testing Complex Applications and Validating Scientific Claims. Poster at 2020 Exascale Computing Project Annual Meeting, Houston, TX, February 3-7, 2020. [poster]

- Partha Kumbhakar, Abhirup Roy Karmakar, Gour P. Das, Jayjeet Chakraborty, Chandra Sekhar Tiwary and Pathik Kumbhakar. Reversible temperature-dependent photoluminescence in a semiconductor quantum dot for development of smartphone-based optical thermometer. Nanoscale, 2021. [paper]

- Biswas, S., Chakraborty, J., Agarwal, A. and Kumbhakar, P., Gold nanostructures for the sensing of pH using a smartphone. RSC Advances Journal, 2019. [paper]

## RESEARCH

- **Student Researcher**                                                                                     Remote
  *Centre for Research in Open Source Software, UC Santa Cruz*
  - ○ Extended the Popper ecosystem by building Popper workflows for automating the MLPerf benchmark suite and for end-to-end benchmarking of bare-metal machines. Also, developed a Python library to compute the confidence intervals for measuring the variability in CPU, Memory, Disk, and Network performance of CloudLab machines.
  - ○ Currently, working on extending the Coffea and Awkward Array libraries to be able to read NanoEvents data from SkyhookDM using the Arrow Dataset API.

- **Research Assistant**                                                                                     Durgapur, India
  *Nanoscience Lab, NIT Durgapur*
  - ○ Cleaned and Pre-processed experimental data collected by scholars in the lab and built Deep learning models using them for making predictions of several parameters like pH and temperature.
  - ○ Developed algorithms using ML techniques to facilitate performing scientific experiments through smartphones. Built 2 Android apps that take pictures of samples and use these algorithms on the images to make predictions of different scientific parameters. One of the works is currently under review at the RSC Advances Journal.

## PRESENTATIONS

- Presented SkyhookDM and it's recent developments at the SNIA Storage Developers Conference 2021, September 28-29, 2021. [slides]

- Presented the IRIS-HEP Winter Fellowship project on building Arrow-Native SkyhookDM and it's integration with Coffea at the IRIS-HEP Topical Meeting, 30 June, 2021. [slides]

- Talked about "Reproducible and Automated Storage systems experimentation with Popper" at the Second K8s-HEP Meetup, December 1-2, 2020. [slides]

- Presented the paper entitled "Enabling Seamless Execution of Computational and Data Science Workflows on HPC and Cloud with the Popper Container-Native Automation Engine." at the CANOPIE-HPC Workshop, November 12, 2020. [slides]

- Presented the IRIS-HEP Summer Fellowship project on making SkyhookDM and Ceph Experiments Reproducible at the CROSS Research Symposium, October 6-9, 2020 and at IRIS-HEP Topical Meeting, October 5, 2020. [slides]

## SKILLS

**Languages**: C, C++, Python, Go, Java, JavaScript, Bash script, MATLAB.
**Tools**: Git, Docker, Jupyter, Kubernetes, Ansible, Azure, GCP, AWS, MongoDB, MySQL, Prometheus, Grafana.
**Frameworks**: Flask, React, Boost, Qt, NumPy, Pandas, Matplotlib, Seaborn, Android Framework.
**Operating Systems**: Ubuntu, Centos, MacOS.

## TEACHING/MENTORING EXPERIENCE

- Served as a mentor in Google Summer Of Code 2021 and OSRE (Open Source Research Experience) 2021 as a part of the organization "Centre for Research in Open Source Software" at UC Santa Cruz.

- Took workshops on Open Source Software Development topics like Linux, Git, GitHub, and Python programming during college fests every year.

- Mentored 2 freshman year students from IIT Roorkee during Kharagpur Winter of Code (KWoC) 2018 on Back-End Web Development projects. Also, guiding new contributors in the getpopper.io community.

## EXTRACURRICULAR ACTIVITIES

- Served as the Open Source Head of GNU/Linux User's Group, NIT Durgapur to help spread knowledge and awareness about Open Source software development between students.

- Writing Articles and Blogs on Computer Systems topics and Open Source Development on my Medium.