

RAJEEV JAIN

✉ rajeeja@gmail.com 🌐 rajeeja 🔗 rajeeja 🌐 rajeeja.github.io

Summary

Research software engineer building high-performance tools for climate science, cancer research, and exascale computing. Lead developer of UXarray (205+ GitHub stars), 2× R&D 100 Award winner, with 16+ years optimizing scientific workflows at Argonne National Laboratory.

Experience

Principal Specialist, Research Software Engineering

2009 – Present

Argonne National Laboratory · Joint Appointment, University of Chicago

Chicago, IL

- **UXarray (2022–Present):** Lead developer for Python library enabling analysis of multi-petabyte unstructured climate datasets. Implemented conservative zonal averaging using Gauss-Legendre quadrature, Grid I/O readers for ESMF, MPAS, SCRIP, and HEALPix formats, and comprehensive testing infrastructure. Established CI/CD pipeline with GitHub Actions and regular PyPI releases. Presented tutorials at SC24, AMS 2024, and ESDS Annual Event. **205+ GitHub stars, adopted by NCAR, DOE labs, and universities worldwide.**
- **FLASH-X (2019–2024):** Implemented asynchronous HDF5 I/O achieving **40-70% reduction** in checkpoint write times on Summit supercomputer. Integrated SZ3 and ZFP compression algorithms reducing storage requirements by **50%+** with minimal accuracy loss. Built verification workflow with nightly baseline testing. Published at SC24 10th International Workshop on Data Analysis and Reduction. **R&D 100 Award 2022.**
- **CANDLE/IMPROVE (2020–2025):** Built hyperparameter optimization infrastructure running **10,000+ training experiments** across Summit, Theta, and Cori supercomputers. Developed GitHub Actions workflows for cross-study validation and Docker-based reproducibility. Maintained benchmarking framework used by **15+ researchers**. Results published in Briefings in Bioinformatics (2025). **R&D 100 Award 2023.**
- **Urban ECP & MeshKit (2009–2019):** Led bi-weekly coordination across 4 national labs for urban microclimate simulation project. Coupled WRF weather models with Nek5000 CFD simulations achieving 11% improvement in building energy prediction accuracy. Developed parallel reactor core mesh generator creating billion-element hex meshes. Published in Engineering with Computers. **Best Paper Award, IMR 2010.**

Education

M.S. in Computer Science

2020

University of Chicago

Chicago, IL

M.S. in Structural Engineering

2009

Arizona State University · Graduate Fellowship Recipient

Tempe, AZ

B.Tech in Mechanical Engineering

2006

Indian Institute of Technology (IIT), Indian School of Mines

Dhanbad, India

Technical Skills

Languages: Python, C++, Fortran, R, Bash

ML/Data Science: PyTorch, TensorFlow, NumPy, Pandas, Xarray, Scikit-learn, Parsl, Swift/T

HPC/Systems: MPI, OpenMP, HDF5, NetCDF, Docker, Singularity, Git, GitHub Actions, Jenkins

Domains: Climate modeling, cancer pharmacogenomics, computational physics, mesh generation

Selected Publications & Presentations

Partin, A., Vasanthakumari, P., Narykov, O., Wilke, A., Jain, R., et al. (2025). "Benchmarking community drug response prediction models: datasets, models, tools, and metrics." *Briefings in Bioinformatics* (forthcoming).

Jain, R., Tang, H., Dhruv, A., Byna, S. (2024). "Enabling Data Reduction for Flash-X Simulations." *10th International Workshop on Data Analysis and Reduction for Big Scientific Data*, SC24.

Jain, R., Wozniak, J., Partin, A., et al. (2024). "Cross-HPO: Optimizing Neural Networks for Cancer Drug Response Using Hyperparameter Tuning on Multiple Datasets." *Tenth Computational Approaches for Cancer Workshop*, SC24.

Chmielowiec, P., Chen, H., Jacob, R., Jain, R., et al. (2024). "UXarray: Extending Xarray with Support for Unstructured Grids." *104th AMS Annual Meeting*.

Tautges, T. J., Jain, R. (2011). "Creating Geometry and Mesh Models for Nuclear Reactor Core Geometries Using a Lattice Hierarchy-Based Approach." *Engineering with Computers*.