ARUNDHATI SINGH

+1 608-421-2946 | asingh278@wisc.edu | linkedin.com/in/arundhati-singh171003/ | github.com/arundhatisingh17
Research Description (Arpaci-Dusseau Lab): https://bit.ly/4sszwIW, (FabLab): https://bit.ly/3YVPq0P

Education

University of Wisconsin–Madison Madison, WI
*B.S. in Computer Science, Data Science, Mathematics Certificate (**Honors in Big Data Systems**)* *Aug. 2022 – May 2026*
[Cum. GPA: 3.703 — GRE Quantitative: 165/170]
Clubs @ UW-Madison: Women in Computing (WACM), MadCloud

Experience

Teaching Assistant: CS320, CS564, and CS574 GLC, University of Wisconsin-Madison, Jun 2024 – Present
- Supported 700+ students across semesters through code reviews, tutorials, and guidance on OOP, custom data structures, time complexity, Pandas/GeoPandas/Scikit-learn pipelines, Flask apps and Selenium & BeautifulSoup scrapers in CS320.
- Currently peer mentoring in CS 564 (Database Management) and CS 574 (Data Management for Data Science), providing targeted support on physical query execution optimizations, buffer management policies, ingestion and warehousing solutions for ETL and ELT pipeline construction, time series analysis, gradient boosting mechanisms, and relevant query languages including SQL, Elasticsearch, and MongoDB.

Software Developer @ TOPS Lab *Traffic Operations and Safety Laboratory, Madison, Wisconsin, Sep 2025 – Present*
- Extended a statewide crash-analytics platform (Java, PostgreSQL) with new API endpoints and parameterized, filtered queries to improve data-access performance for government users.
- Designed database views and server-side rendering logic to display pilot start dates, agency activity, and last-access timestamps, increasing accessibility of time-critical operational information for policy and planning stakeholders.

Research Assistant – SALIS | https://bit.ly/4sszwIW *(Advisor: Prof. Remzi Arpaci-Dusseau), Jul 2025 – Present,*
- Implemented the MRU eviction policy and resolved critical state-consistency failures within a discrete event, storage-aware simulation framework by debugging complex desynchronization issues between virtual block mappings and cache eviction priority queues.
- Engineered GPU execution tracing pipelines to capture Model FLOP Utilization (MFU), deriving a novel "Energy per FLOP" metric to quantify inference efficiency.
- Demonstrated that current LLM pricing models are structurally flawed by proving that active energy consumption, driven by the arithmetic intensity of operations, is a more accurate indicator of physical cost than flat token counts.

Research Assistant @ FabLab | https://bit.ly/3YVPq0P *(Advisor: Prof. Kevin Eliceiri), Sep 2024 – Sep 2025*
- Engineered a real-time monitoring pipeline for electron beam manufacturing equipment, processing 50K+ daily log entries to enable proactive anomaly detection.
- Optimized backend data ingestion using buffered batch transfers and asynchronous HTTP requests, successfully reducing end-to-end latency by 30%.
- Developed a low-level driver interface for legacy hardware using Python and RS-232, reverse-engineering binary protocols to verify parity bits and stream high-frequency telemetry with 24/7 reliability.
- Implemented automated error-recovery logic and handshake protocols to handle hardware disconnects, ensuring robust fault tolerance and reproducible deployments across lab setups.

Equity and Debt Research Analyst Intern *IDBI Asset Management Ltd, Mumbai*, Jun 2023 – Aug 2023
- Automated fund report generation using VBA, pulling data directly from internal databases and formatting it into structured, analysis-ready outputs used for daily performance reviews.
- Designed modular, object-oriented Python modules to analyze stock and bond data, improving data maintainability and significantly reducing access time.
- Built linear regression models on historical equity return data to extract predictive signals, categorizing companies into key industry sectors to enable sector-based analysis, enhancing model accuracy.
- Compared buy-side models to sell-side data for anomalies using quantitative risk indicators and pitched sector findings to board.

Options & Equity Analytics Pipeline Snowflake, dbt, SQL, Tableau

- Architected a centralized Snowflake data warehouse to ingest and unify high-volume raw trading data from options and equities sources, supporting risk and performance analysis for high-frequency trading workflows.
- Engineered automated ELT pipelines using dbt to transform and model unstructured inputs into production-grade tables, reducing manual processing time by approximately 60%, enabling timely refreshes for time-sensitive market insights.
- Developed advanced SQL queries and transformations to integrate disparate datasets (options, equities) into unified, analytics-ready views for downstream processing.
- Implemented statistical quality controls (z-scores, volatility bands, anomaly detection thresholds) within dbt models to automatically flag data inconsistencies.

Technical Skills

Big Data & Distributed Systems: HDFS, MapReduce, Spark (RDDs), Kafka **Programming:** Python, R, SQL, Java, C/C++, Go, **Data & Analytics:** Pandas, NumPy, scikit-learn, Excel, Tableau, Matplotlib, **ML & AI:** Regression Models, PyTorch, TensorFlow, **Databases:** MySQL, PostgreSQL, MongoDB, **Tools & DevOps:** Git, Docker, Google Cloud, AWS, **Methodologies:** Agile, CI/CD, **Certifications:** Marsh McLennan P.R.E.P (Financial Analytics & Stakeholder Communication), Bloomberg Spreadsheet Analysis (Bloomberg)