

Summary

I am a 7th-year Ph.D. candidate in Computer Science at The Ohio State University, with 4 years of professional experience in developing and testing web and server applications, and over 9 years of experience in academic research. As a Graduate Research Assistant under Dr. Rajiv Ramnath, I contribute to the NSF-EAGER Bridging the Last Mile and ICICLE AI Institute projects, where I design AI-driven systems to help researchers execute scientific workflows more efficiently. My work includes developing innovative tools such as the **HPC Application Resource Predictor (HARP)** and **Smart Scheduler**, which enable intelligent resource prediction and workflow optimization across the **edge-to-center computing continuum**. In addition, my current research explores **integrating Large Language Models (LLMs) and agentic AI systems** into cyberinfrastructure to create **self-adaptive orchestration layers** capable of autonomous workflow coordination and domain-aware decision-making. I have advised and mentored **over 20 undergraduate and master's students** and guided **more than 200 undergraduates** through teaching, summer capstone courses, and outreach programs, including **K-12 AI education initiatives** as part of **ICICLE's NextGen activities**. With multiple award-winning publications and a strong focus on interdisciplinary innovation, I combine **technical expertise, mentorship, and research leadership** to advance **AI-integrated and agentic cyberinfrastructure systems**.

Education History

Doctor of Philosophy (Computer Science and Engineering) The Ohio State University (Computer Science and Engineering) Columbus, OH	Aug, 2019 - Present GPA: 3.93/4.0
Master of Science (Computer Science and Engineering) The University of Minnesota Duluth, MN	Aug, 2016 - Jul, 2018 GPA: 3.95/4.0
Bachelor's in Technology (Computer Science and Engineering) Vallurupalli Nageswara Rao Vignana Jyothi Institute of Engineering & Technology (VNRVJIET) Affiliated to Jawaharlal Nehru Technological University, Hyderabad (JNTU-Hyd) Hyderabad, Telangana, India	Sept, 2009 - May, 2013 GPA: 3.90/4.0

Work History

Graduate Research Assistant The Ohio State University Columbus, OH	Aug, 2019 – Present
SGX3 Summer Fellow Texas Advanced Computing Center (TACC), University of Texas at Austin Austin, TX	June, 2025 – Aug, 2025
SGX3 Summer Fellow Texas Advanced Computing Center (TACC), University of Texas at Austin Austin, TX	June, 2024 – Aug, 2024
SGX3 Summer Fellow Texas Advanced Computing Center (TACC), University of Texas at Austin Austin, TX	June, 2023 – Aug, 2023
Teaching Research Fellow (Summer India) Vallurupalli Nageswara Rao Vignana Jyothi Institute of Engineering & Technology (VNRVJIET) Hyderabad, Telangana, India	May, 2022 – July, 2022
Research Intern Oak Ridge National Laboratory (remote) Columbus, Ohio	June, 2021 – Aug, 2021
Web Application Developer Spherexx.com Tulsa, OK	Aug, 2018 – July, 2019
IT Intern CHART Industries, Tulsa, OK	May, 2017 – July, 2017

Current Academic Research

Integrating Artificial Intelligence into Research Workflows: A Comprehensive Approach for Resource Provisioning and AI-Driven Cyberinfrastructure Development Aug'19-present

Relevant Coursework: Artificial Intelligence (statistical models to build baselines for task-specific classifiers and regressors), Operating Systems (batch processing, hyper-parameters settings for performance optimizations), Advanced Computer Architecture, High-Performance deep neural network processing, and Research project capstone (application of research methods and best practices in research; teamwork, written and oral communication)

Developed an AI-driven framework for optimizing resource allocation and scheduling in HPC and deep learning workflows, improving efficiency and potentially reducing energy and time costs.

- **AI-Powered Cyberinfrastructure Tool for Researchers:** Led a project to develop an AI-driven tool that streamlines pre-processing, tool exploration, and software optimization by leveraging past experiences and resources to support researchers in their workflows.
- **Smart Scheduler for Deep Learning Applications:** Designed a framework for optimized resource allocation in deep learning environments, integrating SLURM to estimate job runtimes and enhance scheduling efficiency, tailored for TensorFlow and PyTorch frameworks.
- **HARP Framework for HPC Resource Prediction:** Created an AI-driven system to predict HPC resource needs, supporting scientific workflows through historical data analysis, improving adherence to FAIR principles for software generalization across cyberinfrastructures.
- **ML-Driven HPC Resource Allocation:** Developed machine learning models to predict optimal resource allocation for applications like genome sequencing, reducing the time and computational resources needed to adjust system configurations in high-performance computing environments.
- **DNN Resource Estimators:** Developing DNN architecture-aware resource estimators that leverage XLA features—specifically drawn from HLO graphs—to incorporate hardware-aware characteristics influencing resource requirements, and scaling these estimators across **edge and center environments** to account for diverse resource constraints such as **latency, memory, and power consumption**.
- **Rules Engine for Model Adaptation:** Implemented a rules-based decision framework that integrates predictive insights from HARP and Smart Scheduler to autonomously trigger **model retraining, fine-tuning, and estimator updates** for edge and center deployments, enabling adaptive optimization based on **dynamic workload, latency, and resource feedback** across the computing continuum.
- **Agentic System Wrappers with LLM Interfaces (Currently Exploring):** Extending existing frameworks into **agentic, LLM-powered systems** that serve as **interactive user interfaces** for scientists—enabling natural language-driven workflow creation, adaptive scheduling, and self-updating orchestration across the computing continuum.

AI-Driven Model Adaptation Framework for Edge Deployments in Computational Ecology

Aug'24-present

Developed a toolkit to enable field researchers in **computational ecology** to deploy and evaluate AI models directly on **edge devices**, supporting tasks such as **intrusion detection** and **blank filtering** in camera-trap imagery. The project addresses a key challenge: ecologists often lack access to tools that test, compare, and fine-tune models under **real-world hardware constraints** such as **latency, memory, and power**. By bridging this gap, the toolkit empowers non-ML experts to select the most suitable model for their use case while minimising dependency on AI engineers. Although initially built for ecology, the framework's architecture is generalizable to similar domains such as **precision agriculture** and **environmental monitoring**.

My Role: Conducted **requirement gathering** with ecologists, trained **domain-centric ML models**, and **tested models** with the Field Planner, and validated **edge deployments** using TACC's infrastructure.

Core Components:

- **ML Field Planner (TAPIS UI):** Guided interface for selecting a **model, dataset, and target device**, then launching controlled tests to benchmark models under edge constraints.
- **CKN Dashboard for Camera Traps:** A knowledge-network visualization system that enables researchers to **compare performance and efficiency metrics**—including **runtime, memory utilization, latency, precision, and recall**—across experiments executed via the ML Field Planner..
- **ML Edge Server:** Plug-and-play edge software stack (Dockerized components) that assembles workflow stages—e.g., **delete blanks only or delete blanks + alert authorities** (via a near-edge laptop) when a **target species** is detected.
- **CT-Controller (Edge Deployment Manager):** Simple control UI to **deploy, manage, and verify** models/workflows on edge devices; supports quick iteration and on-site validation.

Relevant Coursework: Natural Language Processing (regular expression parsing, word vectors, part-of-speech tagging, dependency parsing, statistical analysis of data distribution to obtain Hearst Patterns, word similarity analysis using vector and WordNet representations)

WordNet is a free and open-source lexical database widely used in NLP applications such as machine translation. However, since its last update in 2006, many new words have emerged, leading to inaccuracies in NLP systems that depend on it. This project explored **automated methods for updating WordNet** without manual intervention, contributing to **SemEval 2016 – Task 14** and **SemEval 2018 – Task 9**.

- Applied **pattern recognition algorithms** (Hearst Patterns and regular expressions) along with **state-of-the-art vector space models** (word2vec) to identify optimal insertion points for new word senses in WordNet.
- Published a **peer-reviewed paper** on this research in **SemEval@NAACL-HLT 2018**, demonstrating improved performance in automated lexical hierarchy expansion.

Professional Experience

Web Application Developer

Aug'18–Aug'19

- Developed a **full-stack voice bot application** to assist in generating new leads, booking appointments, and providing property information within a **lease management system**.
- Built the application using **C#, .NET MVC Framework**, and integrated a **Python-based machine learning classifier** (NumPy, Pandas, Scikit-learn) for intelligent response automation.

Teaching Assistant

Aug'16–May'18

- Assisted the lead professor in **conducting classes, labs, examinations, and discussions**, while evaluating student progress in projects and exercises.
- **Tutored students** in computer programming and web technologies including **C++, HTML, CSS, JavaScript**, and **Natural Language Processing**.

IT Intern

May'17–Jul'17

- Analyzed and resolved software bugs in the in-house application **CACHE**, developed using **ASP.NET**, which automated quotation generation for the company's products.
- Improved **server-side performance and workflow automation**, reducing manual intervention in product configuration and quotation processes.

Software Engineer

Jun'13–Jul'16

During my tenure at CA Technologies, I worked across multiple product development cycles in both **development and quality engineering roles**.

- Developed **shippable product deployments** using **Java** and **InstallAnywhere** across several enterprise products.
- Collaborated with QA teams to design and implement **automated test suites using Selenium**, expediting product build verification.
- Conducted **unit testing** with **TestNG**, maintained code integrity using **CA Harvest**, and utilized **LoadRunner** for performance testing and optimization.
- Ensured **cross-platform compatibility** across **Oracle, DB2, and MySQL databases** and **Windows, UNIX, and AIX operating systems**.
- Gained comprehensive experience in the **Agile–Scrum product lifecycle**, including backlog planning, sprint execution, performance testing, customer demonstrations, and team training.

Papers Published:

- Hassan, A. Z., Vallabhajosyula, M. S., & Pedersen, T. (2018). Umduluth-cs8761 at semeval-2018 task 9: Hypernym discovery using hearst patterns, co-occurrence frequencies and word embeddings. *arXiv preprint arXiv:1805.10271*.
- Vallabhajosyula, Manikya Swathi, and Rajiv Ramnath. "Towards Practical, Generalizable Machine-Learning Training Pipelines to build Regression Models for Predicting Application Resource Needs on HPC Systems." Practice and Experience in Advanced Research Computing. 2022. 1-5.
- S. Vallabhajosyula and R. Ramnath, "Establishing a Generalizable Framework for Generating Cost-Aware Training Data and Building Unique Context-Aware Walltime Prediction Regression Models," 2022 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom), Melbourne, Australia, 2022, pp. 497-506, doi: 10.1109/ISPA-BDCloud-SocialCom-SustainCom57177.2022.00070.

- [BEST PAPER] Vallabhajosyula, Manikya Swathi, and Rajiv Ramnath. "Insights from the HARP Framework: Using an AI-Driven Approach for Efficient Resource Allocation in HPC Scientific Workflows." *Practice and Experience in Advanced Research Computing*. 2023. 341-344.
- Manikya Swathi Vallabhajosyula, Carlos Guzman, Rajiv Ramnath and Joe Stubbs "Demonstrating HARP (HPC Application Resource Predictor) Framework for Predicting Walltime for Single Node Applications"
- Vallabhajosyula, Manikya Swathi, Sandeep Satish Budhya, and Rajiv Ramnath. "Reference Implementation of Smart Scheduler: A CI-Aware, AI-Driven Scheduling Framework for HPC Workloads." *Practice and Experience in Advanced Research Computing 2024: Human Powered Computing*. 2024. 1-4.
- [BEST PAPER] Joe Stubbs, Sowbaranika Balasubramaniam, Samuel Khuvis, Sachith Withana, Swathi Vallabhajosyula, Richard Cardone, Christian Garcia, Nathan Freeman, Carlos Guzman, Beth Plale, Rajiv Ramnath and Tanya Berger-Wolf. "ML Field Planner: Analyzing and Optimizing ML Pipelines For Field Research". *Practice and Experience in Advanced Research Computing 2024: Human Powered Computing*. 2025
- Salil Reddy, Ronald Davis, Swathi Vallabhajosyula, and Rajiv Ramnath. "Building a Lab-scale Cyberinfrastructure for Fun and Profit". *Practice and Experience in Advanced Research Computing 2024: Human Powered Computing*. 2025.
- Nick Cliffel, Manikya Swathi Vallabhajosyula, Jingxin Wang, Zhao Zhang and Rajiv Ramnath, "Unified Component API: Supporting Standardized Middleware Integration and Flexible Interoperability in Cyberinfrastructure" *Science Gateways 2025 (SG25)*, Green Bay, WI.
- [Extended Abstract – WHPC Workshop] Manikya Swathi Vallabhajosyula, Rajiv Ramnath, "Predicting Resources for AI Workloads in HPC: Methods, Challenges, and Opportunities", *Super Computing (SC'25)*, St. Louis, MI.

MS Thesis/Project Reports:

1. (MS Thesis) "Hypernym Discovery over WordNet and English Corpora - using Hearst Patterns and Word Embeddings", S. Vallabhajosyula, T. Pederson <https://hdl.handle.net/11299/200144>

Mentoring Bachelors/Masters Students towards their Project Reports:

2. "Information Extraction from Gene Sequencing Scientific Documents", Rishabh Chana, Advisor: Rajiv Ramnath, co-advisors: Bryan Carstens, Jian Chen, May 2021. **
3. "Generating Knowledge graphs on scientific execution workspace", Akhilesh Gulati, Advisor: Rajiv Ramnath, co-advisors: Bryan Carstens, Jian Chen, December 2020. **
4. "HARP - HPC Application Runtime Predictor", Prasanna, Saishree Miriyala, Advisor: Rajiv Ramnath, co-advisors: Bryan Carstens, Jian Chen, December 2022. **
5. "Reference Implementation of Smart Scheduler: Configuring a pre-SLURM Database for Enabling Smart Scheduler", Akanksha Jain, Advisor: Rajiv Ramnath May, 2024. **
6. "Reference Implementation of Smart Scheduler: Establishing a Smart Scheduler Framework Backend – Intelligence Plane", Sandeep Satish Budhya, Advisor: Rajiv Ramnath May 2024. **
7. "AI-Enabled Smart Scheduler for Optimizing Resource Utilization in HPC", Bhargavi Dwivedi (MS-ECE), Advisor: Dr. Rajiv Ramnath**.
8. A Project on Exploring and Building Profilers for HPC Environments (Bachelor's), Maaz Baig, Advisor: Dr Rajiv Ramnath. **
9. "Parallelizing Training and Inferencing for Machine Learning Models": A Study on HLO Graphs for Network Optimization, Shashwat Rao, Advisor: Dr. Rajiv Ramnath. **
10. "GPU Resource Prediction Dataset for DNN Workloads", Rahul Vaidya, Advisor: Dr. Rajiv Ramnath. **
11. A Project on Exploring Different Regression Models for Runtime Prediction on Black-Box and White-Box Resource Profiles for DNN Training Time Estimation, Nachiappan Ramasamy, Advisor: Dr. Rajiv Ramnath. **

(** These are the MS projects I have worked on and provided guidance for. Please contact the Computer Science and Engineering (CSE) department at The Ohio State University to access these documents.)

POSTERS Presented:

1. Vallabhajosyula, Manikya Swathi (2018): Identifying Hypernym for a New Sense in WordNet. figshare. Poster. <https://doi.org/10.6084/m9.figshare.22565089.v1>
2. Vallabhajosyula, Manikya Swathi; Ramnath, Rajiv (2020): EAGER: Bridging the Last Mile. figshare. Poster. <https://doi.org/10.6084/m9.figshare.11777808.v3>
3. Vallabhajosyula, Manikya Swathi; Ramnath, Rajiv (2021): Modeling A Framework To Estimate Resource Requirements For Scientific Workflows. figshare. Poster. <https://doi.org/10.6084/m9.figshare.22363183.v1>
4. Chana, Rishabh; Manikya Swathi Vallabhajosyula; Ramnath, Rajiv (2021): Using synthesized data to train machine learning models used in genome engineering pipeline. figshare. Poster. <https://doi.org/10.6084/m9.figshare.22565113.v1>

5. Manikya Swathi Vallabhajosyula; Ramnath, Rajiv (2022): Building an AI-powered Assistant for Computational Scientists. figshare. Poster. <https://doi.org/10.6084/m9.figshare.11777796.v1>
6. Vallabhajosyula, Manikya Swathi; Ramnath, Rajiv (2023): Modeling A Framework To Estimate Resource Requirements For Scientific Workflows. figshare. Poster. <https://doi.org/10.6084/m9.figshare.22363183.v1>
7. Manikya Swathi Vallabhajosyula, and Rajiv Ramnath. "Towards Characterizing DNNs to Estimate Training Time using HARP (HPC Application Resource (runtime) Predictor." Practice and Experience in Advanced Research Computing. 2023. 483-485
8. Manikya Swathi Vallabhajosyula, Rajiv Ramnath, and Joe Stubbs. "Custom Cost, Loss, And Reward Functions to Train Regression Models for Estimating Execution Resources using HARP." Science Gateways 2023 (SG23), Pittsburgh, PA.
9. Orchestrating a DNN training job using an iScheduler Framework: a use case, Swathi Vallabhajosyula, Sandeep Satish Budhya, Maaz Baig, Akanksha Jain, Rajiv Ramnath, Practice and Experience in Advanced Research Computing. 2024
10. Orchestrating End-to-End AI-Model Development using TAPIS and Smart Scheduler, Swathi Vallabhajosyula; Nathan Freeman; Christian Garcia; Joe Stubbs; Rajiv Ramnath. Science Gateways 2023 (SG23), Bozeman, MO.
11. Manikya Swathi Vallabhajosyula, Rahul Vaidhya, Nachiappan Ramasamy and Rajiv Ramnath. *Hybrid black-box and white-box approaches for efficient resource prediction for AI workloads in high-performance computing* [Project poster]. ISC High Performance Conference (ISC-HPC 2025), Hamburg, Germany.
12. Manikya Swathi Vallabhajosyula, Karen Tomko, and Rajiv Ramnath. *Intelligence Plane: A framework for machine learning application life-cycle management* [Research poster]. ISC High Performance Conference (ISC-HPC 2025), Hamburg, Germany.
13. Manikya Swathi Vallabhajosyula. *Motivated by challenges: Harnessing AI to revolutionize resource and workflow management in high-performance computing* [Women in HPC poster]. ISC High Performance Conference (ISC-HPC 2025), Hamburg, Germany.
14. AI-Driven Resource Optimization for High-Performance Computing: A Comprehensive Framework. International-HPC-Summer School, Lisbon, Portugal (IHPCSS-2025)
15. Beyond Automation: Integrating Agentic Capabilities into MLOps with ICICLE Infrastructure, Swathi Vallabhajosyula, Gautam Molakalmuru, Nick Cliffel, Rajiv Ramnath. Practice and Experience in Advanced Research Computing. 2025
16. Gautam Gururaj Molakalmuru, Manikya Swathi Vallabhajosyula, Samuel Khuvis, Joe F Stubbs and Rajiv Ramnath, "Operational Considerations for Real-Time ML Pipelines on Edge Devices", Science Gateways 2025 (SG25), Green Bay, WI.
17. Manikya Swathi Vallabhajosyula, Gautam Gururaj Molakalmuru, Neelesh Karthikeyan, Agulugaha Isuru Gamage, Samuel Khuvis, Nathan Freeman, Joe Stubbs, Beth Plale and Rajiv Ramnath, "Beyond Accuracy: The ML Field Planner's Framework for AI Model Selection in Conservation", Science Gateways 2025 (SG25), Green Bay, WI.
18. [Portal] Manikya Swathi Vallabhajosyula, Gautam Gururaj Molakalmuru, Agulugaha Isuru Gamage, Neelesh Karthikeyan, Nick Cliffel, Samuel Khuvis, Christian Garcia, Nathan Freeman, Joe Stubbs, Beth Plale and Rajiv Ramnath, "ML Field Planner with TAPIS: Configuring and Analyzing AI Models for Animal Ecology", Science Gateways 2025 (SG25), Green Bay, WI.
19. [Doctoral Showcase] Manikya Swathi Vallabhajosyula, Rajiv Ramnath, "AI-Driven Resource Optimization for High-Performance Computing: A Comprehensive Framework", Super Computing (SC'25), St. Louis, MI

Software Proficiency

Programming / Testing:	C, C++, C#, Core Java, J2EE, Python, Perl, Shell, Batch, TestNG, Selenium
Web Development:	HTML5, CSS, JavaScript, ASP.NET, ReactJS, VueJS, Node.js, REST APIs, Docker / Singularity Images, Kafka
Servers and Frameworks:	Tomcat, JBOSS, TAPIS API Framework, Product Automation Framework (CA Internal), Flask, FastAPI, Selenium WebDriver
Operating Systems:	Windows Server, UNIX, AIX, Linux (Ubuntu / CentOS)
Profilers & HPC Tools:	NVIDIA-SMI, Nsight Systems, SLURM, TensorBoard, TAU
Tools and IDEs:	InstallAnywhere, HP ALM, HP LoadRunner & Performance Center (PC), Rally, VersionOne, Eclipse, Maven, Visual Studio Code, Microsoft Office, Harvest, Embedded Entitlement Manager (CA), Rational Rose, REST Client, Selenium IDE with Firebug
Databases:	MySQL, Oracle 11g/12c, DB2 LUW, DB2 z/OS, MS-SQL, MongoDB, PostgreSQL
Application Lifecycle Management:	Agile Scrum, Scrumban, Waterfall, CI/CD (GitHub Actions, Jenkins), MLOps (MLflow, Weights & Biases)
AI / ML Frameworks and Libraries:	TensorFlow, PyTorch, Keras, Distributed Training with Horovod, Hugging Face Transformers, Scikit-learn, Pandas, NumPy, OpenCV, NetworkX

Awards, Certifications & Workshops

Titles	: ICICLE NextGen (Student Group) Manager (June 2023–present) – Collaborated with more than 280 students from various universities, hosted tutorials for broader impact and project integrations and participated in outreach programs showcasing AI's power on the cloud and edge to K-12 and undergraduate students as part of HACK/IO series
Professional Awards	: - Above & Beyond September 2014 @ CA Technologies (Broadcom) - Above & Beyond March 2016 @ CA Technologies (Broadcom)
Academic Awards	: - PEARC25 Best Paper: Systems & Applications Software Track and Phil Andrews Award for Best Paper Overall; - PEARC23 Best Paper: Systems & Systems Software Track (Overall & Student Categories) and Phil Andrews Award for Best Paper Overall, - Honorable mention for Best Research at The OSU College of Engineering 2024; - Honorable mention for Best Research CSE Annual Research Poster Exhibit. - Undergraduate Third-best Academic Project (CSE),
Soft Skill Certifications	: Leaders at All Levels @ CA Technologies , JAWS
Travel Awards	: - PEARC 2024 & 2025 Student Volunteer Travel Awards – Provided support to attend PEARC as a dedicated student volunteer, facilitating active involvement in workshops and panel organization. - Science Gateways 2023, 2024 & 2025 Travel Grants – Enabled participation at the Science Gateways conferences, supporting presentation and engagement with scientific gateway communities focused on research computing and education. - Supercomputing (SC) 2025 SCinet Travel Grant – Awarded to participate in SCinet, the high-performance networking infrastructure team for the annual Supercomputing Conference, providing technical experience in building critical systems for the SC event. - International High Performance Computing Summer School (IHPCSS) Grant (Lisbon, Portugal) . - Summit of AI Institute Leadership (SAIL) Graduate Student Researcher Keynote Speaker and GSR Travel Grant, 2025
Scholarships	: Fully Funded Master's Program , University of Minnesota – Awarded full tuition and stipend support for both years of M.S. program.