

Emily J. Herron

Oak Ridge National Laboratory

1 Bethel Valley Rd, MS-6085, Oak Ridge, Tennessee 37831-6085, USA

emily.j.herron@gmail.com • (865) 341-3679 • github.com/herronej • linkedin.com/in/emily-j-herron

Education

Bredesen Center, The University of Tennessee, Knoxville, Tennessee, USA

Ph.D. in Data Science & Engineering

Aug 2018 – Dec 2023

- Thesis: Generalized Differentiable Neural Architecture Search with Scaling and Stability Improvements
- Advisor: Dr. Steven R. Young
- Focus: Machine learning, deep learning, neural architecture search, high performance computing
- Cumulative GPA: 3.95 / 4.00

Mercer University, Macon, GA, USA

B.S. in Computational Science

Aug 2014 – May 2018

- Graduated Summa Cum Laude
- Cumulative GPA: 3.94 / 4.00

Experience

Analytics and AI Methods at Scale Group, Oak Ridge National Laboratory (ORNL)

Postdoctoral Research Associate

Jan 2024 – Present

- Developed VISTA, an agentic AI framework for molten-salt thermophysical-property analysis under the DOE Genesis Mission, with a PydanticAI agent loop (planning, memory, tool-use), a FastMCP server fleet, retrieval-augmented reasoning over a curated ChromaDB corpus, and a Next.js frontend with human-in-the-loop credential elicitation for SLURM job submission.
- Designing VISTAGuard, a unified security and safety-alignment mediation framework for agentic AI in DOE scientific workflows, with capability-aware safety gates spanning prompt, tool, RAG/memory, code, HPC, instrument, and federation boundaries, interpreter-level trust-tag propagation, and a quarantined-LLM architecture for structurally-isolated classification of untrusted content and SciAgentBench, a companion red-teaming evaluation pipeline for agentic AI safety in scientific workflows.
- Designed and implemented HARMONY, a distributed evolutionary architecture-search framework for efficient hybrid Transformer–Mamba–Mixture-of-Experts language models, scaling to 16,384 GPUs on the Frontier exascale supercomputer.
- Architected SciTrust, a trustworthiness-evaluation framework for scientific foundation models across truthfulness, adversarial robustness, safety, and ethics, establishing reusable benchmarks for systematic model comparison.
- Engineered CHUNKS, a multi-agent LLM pipeline for real-time spectroscopic data analysis at ORNL’s Center for Nanophase Materials Sciences, comprising eight LLM-backed agents across Engineering, Analysis, and Knowledge MCP servers, a 100K+ publication RAG corpus, YAML material knowledge bases, and UI for live experiment integration.
- Applied supervised fine-tuning and Group Relative Policy Optimization (GRPO) to reasoning LLMs for materials-science hypothesis generation in collaboration with AllenAI, building a chain-of-thought dataset of 100K+ documents and designing reward functions and evaluation pipelines.
- Built large-scale distributed training and inference systems with PyTorch, DeepSpeed, FSDP, and vLLM, implementing hybrid parallelism (FSDP + DDP, pipeline parallelism) for 4B+ parameter models on both AMD ROCm (Frontier MI250X) and NVIDIA CUDA platforms.

Learning Systems / Computational Data Analytics Group, Oak Ridge National Laboratory

Graduate Research Assistant

Aug 2018 – Dec 2023

- Developed stability and performance improvements to CDARTS NAS algorithm

- Researched and implemented selection algorithms for ORNL’s MENNDL NAS software utilizing ORNL’s Titan and Summit supercomputers
- Collaborated on creation of NIEHS document mining pipeline: comparing and testing document embeddings and implementing NAS for in-PDF text detection

Intern

Jun 2018 – Aug 2018

- Carried out text mining project dealing with classification of NIEHS research publications
- Trained Hierarchical Attention Networks and other models including Naive Bayes and Random Forests
- Leveraged frameworks including Hierarchical Attention Networks and Tf-Idf Based Weighting

Big Data X REU, University of Chicago & Illinois Institute of Technology

Undergraduate Researcher

May 2017 – Jul 2017

- Contributed to development of automated pipeline for extracting metadata and predicting contextual relationships between files in large scientific repositories
- Developed collection of modules for extracting metadata from images
- Module tested on over 3,500 images; results published in poster, placed 3rd in ACM’s undergraduate Student Poster Competition at Super Computing (SC) 2017 Conference

Mercer Engineering Research Center, Warner Robins, GA

Intern

May 2016 – May 2018

- Applied variety of machine learning methods to classification of aircraft flight regime data
- Presented classification results and findings to group of engineers on weekly basis
- Researched and developed augmented reality-based remote collaboration and video streaming application for Microsoft Hololens

Technical Skills

- **Programming Languages:** Python, C/C++, Bash
- **ML/DL Frameworks:** PyTorch, Hugging Face Transformers, vLLM, LangChain, PydanticAI, Pydantic, Scikit-learn
- **Agentic AI & LLM Systems:** Model Context Protocol (MCP), FastMCP, multi-agent orchestration, tool-use, retrieval-augmented generation, ChromaDB, agent safety / red-teaming, reward modeling, GRPO post-training
- **HPC & Distributed Computing:** SLURM, MPI, DeepSpeed, DDP, FSDP, pipeline parallelism
- **Tools & Technologies:** Git, Docker, Weights & Biases, Jupyter, VS Code, Next.js
- **Platforms:** CUDA (NVIDIA), ROCm (AMD), Frontier (MI250X), Summit, Titan
- **Specializations:** Agentic AI for Science, Foundation Models for Science, LLM Safety & Trustworthy AI, High Performance Computing, Neural Architecture Search

Publications

*Note: * indicates first authorship. Citation counts from Google Scholar as of May 2026.*

Peer-Reviewed Publications

1. * E. Herron, S. Dash, and F. Wang, “HARMONY: Large-Scale Architecture Search for Efficient Hybrid Language Models,” *ISC High Performance 2026*, June 2026.
2. * E. Herron, V. Lama, S. Bouknight, and T. Ghosal, “From Rules to Reasoning: A Survey of Large Language Model-Based Approaches to Scientific Hypothesis and Idea Generation,” *ACM Computing Surveys*, 2026.
3. T. de Haan *et al.*, “AstroMLab 4: Benchmark-Topping Performance in Astronomy QA with a 70B-Parameter Domain-Specialized Reasoning Model,” in *2025 Workshop on Machine Learning for Astrophysics*, ICML 2025. arXiv: 2505.17592 [astro-ph.IM]. [**4 Citations**]
4. J. Yin *et al.*, “ChatHPC: Empowering HPC Users with Large Language Models,” in *Journal of Supercomputing*, vol. 81, 2025. DOI: [10.1007/s11227-024-06637-1](https://doi.org/10.1007/s11227-024-06637-1). [**21 citations**]
5. * E. Herron, J. Yin, and F. Wang, “SciTrust: Evaluating the Trustworthiness of Large Language

Models for Science,” in *AI4S: 5th Workshop on Artificial Intelligence and Machine Learning for Scientific Applications*, SC24, 2024. [6 citations]

6. M. Chai, * E. Herron, E. Cervantes, and T. Ghosal, “Exploring Scientific Hypothesis Generation with Mamba,” in *Proceedings of the 1st Workshop on NLP for Science (NLP4Science)*, ACL, 2024, pp. 197–207. (Equal contribution) [14 citations]
7. * E. J. Herron, S. R. Young, and D. Rose, “ICDARTS: Improving the Stability of Cyclic DARTS,” in *2022 21st IEEE International Conference on Machine Learning and Applications (ICMLA)*, Nassau, The Bahamas, 2022. [2 citations]
8. * E. J. Herron, S. R. Young, and T. E. Potok, “Ensembles of Networks Produced from Neural Architecture Search,” in *International Conference on High Performance Computing*, Springer, 2020, pp. 223–234. [28 citations]

Preprints & Under Review

9. * E. Herron, J. Yin, and F. Wang, “SciTrust 2.0: A Comprehensive Framework for Evaluating Trustworthiness of Large Language Models in Scientific Applications,” 2025, Available: [arXiv:2510.25908](https://arxiv.org/abs/2510.25908). [Preprint]
10. S. Dash, E. Herron, and F. Wang, “Recipes for Distributed Training of Mixture of Experts Models,” *Submitted to AI4S: 6th Workshop on Artificial Intelligence and Machine Learning for Scientific Applications*, Nov. 2025. [Preprint]
11. J. Huang, E. Herron, I. Ciucă, and T. Ghosal, “Evaluation Methods in LLM-Based Scientific Hypothesis Generation: Current Methods, Gaps, and Next-Generation Design Principles,” 2025. [Preprint]
12. S. Dash, Y. Yang, E. Herron, and M. Zhang, “Roasting SMOREs: Training Mixture of Experts Foundation Models for Science on a Large-Scale Supercomputer,” 2025. [Preprint]
13. * E. Herron, D. Rose, and S. Young, “ICDARTS: Improving the Stability and Performance of Cyclic DARTS,” 2023. Available: [arXiv:2309.00664](https://arxiv.org/abs/2309.00664). [Preprint]
14. J. Duncan, F. Fallas, C. Gropp, * E. Herron, M. Mahbub, P. Olaya, E. Ponce, T. Samuel, D. Schultz, S. Srinivasan, M. Tang, V. Zenkov, Q. Zhou, E. Begoli *et al.*, “The Sensitivity of Word Embeddings-based Author Detection Models to Semantic-preserving Adversarial Perturbations,” 2021. DOI: 10.48550/ARXIV.2102.11917. [1 citation]

Presentations

- **Functional Agents for Functional Materials**
ORPA Research Symposium 2026 (Talk)
- **Towards Secure and Trustworthy Agentic AI Systems for Scientific Discovery**
ORNL Collaboration Catalyst Future of Computing 2026 (Poster)
- **Agentic AI Systems in Scientific Workflows**
CNMS Postdoc Research Symposium 2026 (Invited Talk)
- **HARMONY: Evolutionary Design of Efficient Hybrid Transformer-Mamba-MoE Language Models**
ORPA Research Symposium 2025 (Talk)
- **AstroBench: Probing Large Language Models with the Challenges of Astrophysics**
ORNL AI4Science Workshop 2025 (Poster)
- **Functional Agents for Functional Materials**
ORNL AI4Science Workshop 2025 (Poster); ORNL AI Expo 2025 (Poster)
- **SciTrust: Evaluating the Trustworthiness of Large Language Models for Science**
SC24 AI4Science Workshop; ORPA National Postdocs Appreciation Week 2024 Poster Session; Monterey Data Conference 2024 (Poster)
- **Generalized Differentiable Neural Architecture Search with Performance and Stability Improvements for Scientific Applications**
SOS26 2024, Cocoa Beach, FL (Poster)
- **ICDARTS: Improving the Stability of Cyclic DARTS**

21st IEEE International Conference on Machine Learning and Applications (ICMLA), 2022, Nassau, The Bahamas

- **Ensembles of Neural Networks Produced from Neural Architecture Search**
Women in High Performance Computing Workshop, SuperComputing 2020 (Virtual); International Conference on High Performance Computing 2020 (Virtual)
- **Applying Image Feature Extraction to Cluttered Scientific Repositories**
3rd Place, ACM Student Research Competition, SuperComputing 2017, Denver, CO

Awards & Scholarships

- Argonne Training Program on Extreme-Scale Computing (ATPESC) 2025 - Accepted Scholar Jul 2025
- Oak Ridge Leadership Computing Facility - Director's Discretion Project, 20,000 Summit Hours 2022–2023
- University of Tennessee - Bredesen Center Data Science & Engineering Fellowship Mar 2018
- SuperComputing Conference - ACM Student Poster Competition, Semifinalist Nov 2017
- Mercer University - Outstanding Student in Computational Science 2016–2018
- Mercer University - Summa Cum Laude, President's and Dean's Lists 2014–2018

Professional Service

Reviewing

- AAAI 2026 Main Track Reviewer
- ICML 2022 Reviewer (Top 10%)
- EMNLP 2024 Demo Track Reviewer

Teaching & Mentoring

- Guest Lecturer, DSE 697 Generative AI (Topic: Trustworthy AI), University of Tennessee (2025)
- Assistant Instructor, ORNL AI Summer Institute Tutorial (2024)
- Mentor, ORNL PCIP Undergraduate Summer Internship Program - 2 students (2024-2025)
- Peer Mentor, Bredesen Center Peer Mentoring Program (2022-2023)

Outreach

- Ijams River Rescue, Oak Ridge (2026)
- Hour of Code Volunteer, Little River Montessori School (2025)
- ORNL Traveling Science Fair Volunteer (2024-2025)
- Introduce Your Daughter to AI, ORNL (2018-2019)

Professional Affiliations

- Oak Ridge Postdoctoral Association (ORPA) - 2026 Executive Board Secretary 2025–Present
- Institute of Electrical and Electronics Engineers (IEEE) - Member 2022–Present
- Association for Computing Machinery (ACM) - Member 2017–Present