

EDUCATION

North Carolina State University | Raleigh, NC **August 2024 – May 2028 (Expected B.S.)**
Honors: Caldwell Fellow, ICML 2026 Silver Reviewer, Dean's List (all semesters)

PUBLICATIONS

- Zhao, X., Zhao, C., **Mengesteab, M.**, Wong, C.-W., Zhan, M., Tan, Z., Chen, T. "Automating Content Analysis with Multiple LLM Agents: Impacts of Agent Attributes and Human-AI Collaboration" *Social Science Computer Review*, 2026

PROFESSIONAL EXPERIENCE

Software Engineering Intern | SAS — Advanced Analytics R&D **May 2026 – August 2026**

- Built Go static-analysis PR impact tooling to map a 300k+ LoC service through custom graph data structures, tracing PR diffs to impacted APIs and covering tests to drive smarter CI/CD test routing
- Implemented call tracing across pods in K8s for dynamic analysis of test coverage in event driven architecture
- Extended agentic pull request review system in Go and Python with heterogeneous model routing
- Developed a knowledge base architecture that amortizes review costs by caching learnings across PRs
- Created React and TypeScript impact visualization to map PR blast radius across services, enabling reviewers to identify affected components at a glance

Deep Learning Undergraduate Researcher | NCSU **December 2025 – Present**

- Developing evaluation harnesses for multi-agent content analysis systems, investigating how inter-rater reliability measures decompose when applied to agent systems vs. human experts
- Researching memory and context management architectures for multi-agent systems
- Translating social science methodology and literature into agent system design constraints
- Co-authorship: updated codebase, verified reliability and statistical significance claims, and contributed to interpretation of results during rebuttal response
- Assisted in work on federated learning, LoRA fine tuning, conformal prediction, and inference verification

Software Engineer | NCSU Liquid Rocketry Lab **September 2024 – January 2026**

Programming Subteam

- Designed Redis pub/sub telemetry architecture for sub-millisecond latency
- Developed mission control dashboard with drag-and-drop analytics using React
- Led implementation of Rust-based fault detection system for critical flight data

Software Engineering Intern | Ultisim **August 2023 – March 2024**

- Built 3D simulation environments for education and healthcare using Unity and React Three Fiber
- Addressed technical challenges such as integrating real-time data feeds into Digital Twin systems
- Developed QuestSim, an immersive historical learning platform for middle and high school students

PROJECTS

Kiln **July 2026**

- Built a fused Triton top-k/top-p sampling kernel for 131K-token vocabularies, running 16.4× faster than PyTorch eager and 13.4× faster than `torch.compile` on an NVIDIA L40S
- Designed an exact fp16/bf16 keyspace-histogram algorithm and optimized it through Nsight Compute profiling, converting the worst benchmark from 0.25× to 3.2× `torch.compile` performance
- Implemented Triton and CUDA C++ RMSNorm kernels with PyTorch autograd, FakeTensor, and dynamic-shape support; validated correctness with 2,897 randomized comparisons, 115 CUDA tests, `opcheck`, and Compute Sanitizer
- Froze op semantics in a written contract with an executable reference backing ~2,900 randomized property tests; benchmarked against a predeclared case matrix with median/IQR reporting and committed raw replicates

Sigil **May 2026**

- Built a two-key trust protocol combining Apple App Attest with a biometric-gated Secure Enclave key to authorize high-trust actions from legitimate human clients
- Implemented challenge-based enrollment and server verification of certificate chains, app identity, nonces, replay counters, P-256 signatures, and payload hashes

- Packaged protocol into a developer-facing framework with enrollment, key registration, action signing, device recovery flows, and server verification quickstarts.

MemBench

May 2026

- Built a reproducible benchmark comparing full-context, recursive-compaction, BM25 RAG, and structured-memory agents on 50 stratified LongMemEval questions with approximately 110K-token histories
- Found BM25 RAG reached 78% accuracy versus full context's 80% at 1/18th the token cost, while structured memory fell to 60% due primarily to memory-maintenance failures
- Developed resumable parallel evaluation, trace caching, maintenance-cost accounting, bootstrap analysis, and an interactive failure explorer; measured 99.1% agreement between independent graders across 112 shared items

Flotilla

March 2026

- Built a Kubernetes-native Go control plane and 1.4 MB Rust edge agent for signed software rollouts across a simulated 300-device fleet, supporting health gates, drift correction, and automatic rollback
- Designed reproducible chaos experiments comparing big-bang, wave, and canary rollouts under partitions, latency, packet loss, and device failures; found canaries reduced peak failures from 43 to 6 under healthy visibility but could promote bad releases during partitions
- Implemented idempotent reconcilers, Ed25519 artifact verification, Redis-backed device state, Prometheus metrics, and a React fleet-convergence dashboard

RaterLab

February 2026

- Built an evaluation framework measuring agreement and correlated errors across 10 LLM raters, 8 annotation attributes, 6 human-labeled datasets, and 6,100+ model ratings
- Implemented nominal, ordinal, and multi-label Krippendorff's alpha with item-level bootstrap confidence intervals, cross-group reliability, and systematic-error decomposition, validated against reference implementations
- Found LLM ensembles agree with each other more than trained humans do (α 0.73 vs. 0.52) while tracking human judgments worse, with systematic errors comprising 51% of ensemble mistakes on gold-labeled tasks

Meridian

August 2025

- Built multimodal RAG system for student government, indexing 50+ hours of meetings and 2,000 pages of documents
- Audio: Custom OpenAI-compatible whisper.cpp wrapper and pyannote diarization for meeting recording, transcription, and diarization, alongside ChromaDB and SQLite for indexing and storage.
- Web: PGVector embeddings + Postgres full-text hybrid search. Used small models for summarization and context-aware retrieval over legal and governance materials

Recall

July 2025

- Built a privacy-first iOS agent providing fully on-device search and source-cited chat over emails, messages, documents, and notes
- Ported all-MiniLM-L6-v2 to Core ML with a custom WordPiece tokenizer and integrated 384-dimensional vector search through a C-bridged sqlite-vec extension
- Implemented semantic chunking, concurrent ingestion, WAL-mode persistence, and Apple Foundation Models for semantic retrieval and full-thread expansion

Syllabase

December 2024

- Architected graph-based course requisite system using Neo4j and Next.js, serving 2000+ NCSU courses
- Implemented GraphQL API layer for efficient data querying (and RateMyProfessor integration)
- Deployed scalable architecture using Docker containers on AWS EC2 with CI/CD pipeline

Universitrack

July 2023

- Architected full-stack platform (React/Go/PostgreSQL) serving 100+ active users
- Implemented custom RAG system for personalized essay feedback using LangChain
- Built automated recommendation engine processing academic and extracurricular data

DineRight

December 2024

- Reimagined NCSU dining hall menus with nutritional filtering, allergen tracking, and meal planning
- Built with Next.js, Express, and Scrapy for automated menu data collection

Quizless

May 2023

- Reimagined study tools with GPT integration, enabling auto-generated questions and smarter quizzes
- Led to feature on panel discussion for EdTech and AI in education at SparkNC, July 2023

LEADERSHIP EXPERIENCE

Student Senate President | North Carolina State University

March 2026 – Present

Previously: Student Senator

September 2024 – March 2026

- Representing 40,000+ students to university administration and Board of Trustees
- Member of the Board of Visitors, University Council, Fee Review Committee, Tuition Review and Advisory Committee, and liaison to Faculty and Staff Senate
- Coordinating ballot access protection across the UNC system
- Led digital infrastructure modernization including digital ID systems and cloud printing
- Authored Senate Transparency and Safety Act; led transparency reforms in Senate financial processes

President | Code Black

May 2025 – Present

Previously: Vice President

December 2024 – May 2025

- Leading organization for Black students in CS, managing \$20k+ budget and corporate partnerships
- Redesigned programming around annual career fair, now serving 100+ students
- Restructured corporate partner model to prioritize sustained mentorship over one-off engagements
- Established cross-organizational partnerships with LA/CSC and WiCS

Conference Speaker | Spark NC

July 2023

- Spoke on panel about future of generative AI in education to 100+ students and professionals

SKILLS

Languages: Python, Go/Golang, JavaScript/TypeScript, SQL, Rust, Swift, C/C++, Java, C#

Frameworks/Technologies: PyTorch, Triton, CUDA, Kubernetes, React, Next.js, Node.js, Express, Redis, Prometheus, Neo4j, GraphQL, Docker, AWS, Core ML, SwiftUI, PostgreSQL, MySQL, SQLite, ChromaDB

Relevant Coursework: Neural Networks and Deep Learning (CSC 525), Data Structures and Algorithms (CSC 316), Operating Systems (CSC 246), Software Engineering (CSC 326), Linear Algebra (MA 305)