AI Research Day 2026

Poster Abstracts

Theme 1: AI in Education

Poster #1

AutoSCORE: Enhancing Automated Scoring with Multi-Agent Large Language Models via Structured Component Recognition

Yun Wang

Automated scoring plays a crucial role in education by reducing the reliance on human raters and offering scalable and immediate evaluation of student work. While large language models (LLMs) have shown strong potential in this task, their use as end-to-end raters faces challenges such as low accuracy, prompt sensitivity, limited interpretability, and rubric misalignment, which hinder practical implementation. To address the limitations, we propose AutoSCORE, a multi-agent LLM framework enhancing automated scoring via rubric-aligned Structured COmponent REcognition. With two agents, AutoSCORE first extracts rubric-relevant components from student responses and encodes them into a structured representation (i.e., Scoring Rubric Component Extraction Agent), which is then used to assign final scores (i.e., Scoring Agent). This design ensures that model reasoning follows a human-like grading process, enhancing interpretability and robustness. We evaluate AutoSCORE on four benchmark datasets from the ASAP benchmark, using both proprietary and open-source LLMs (GPT-4o, LLaMA-3.1-8B, LLaMA-3.1-70B). Across diverse tasks and rubrics, AutoSCORE predominantly improves scoring accuracy, human-machine agreement (QWK, correlations), and reduces error metrics (MAE, RMSE) compared to single-agent baselines, with particularly strong benefits on complex, multidimensional rubrics, and especially large relative gains on smaller LLMs. These results demonstrate that structured component recognition combined with multi-agent design offers a scalable, reliable, and interpretable solution for automated scoring.

Poster #2

A Framework for Designing Science and Engineering Practice Activities for an AI-Supported Learning Environment

Zhaoji Wang

This study addresses the challenges of applying existing curricula in artificial intelligence (AI)–involved learning environments, where classroom interactions are fundamentally transformed. In such contexts, interactions are no longer limited to traditional teacher–student exchanges but become triangular, involving teachers, students, and AI agents. Correspondingly, the sources of evidence for learning shift from teacher observations and final products to more process-oriented data, including logged student actions and dynamic exchanges among participants. Moreover, teachers and AI agents increasingly work in concert to provide adaptive scaffolding tailored to students' needs. These changes create a growing demand for traceable and structured (i.e., machine-readable) records of learning processes, an expectation that most existing curricula were not originally designed to accommodate.

Poster #3

LabAgent: A Generative AI Tutor for Scaffolding Scientific Inquiry

Zipei Zhu

The implementation of AI-supported scientific inquiry learning presents challenges for both students and teachers. For students, it requires mastering complex skills such as formulating testable questions and interpreting data. For teachers, providing individualized, real-time guidance to every student in a busy classroom is often infeasible due to time constraints and the difficulty of monitoring each student's progress. To address these challenges, this study introduces a digital learning environment that integrates a physics simulation with a generative AI tutor, LabAgent. The goal of LabAgent is to deliver scalable, personalized scaffolding that supports middle school students throughout the full scientific inquiry cycle, aligned with the Science and Engineering Practices (SEPs) outlined in the Next Generation Science Standards.

Poster #5

Generalizable and Efficient Automated Scoring with a Knowledge-Distilled Multi-Task Mixture-of-Experts

Luyang Fang

Automated scoring of written constructed responses typically relies on separate models per task, straining computational resources, storage, and maintenance in real-world education settings. We propose UniMoE-Guided, a knowledge-distilled multi-task Mixture-of-Experts (MoE) approach that transfers expertise from multiple task-specific large models (teachers) into a single compact, deployable model (student). The student combines (i) a shared encoder for cross-task representations, (ii) a gated MoE block that balances shared and task-specific processing, and (iii) lightweight task heads. Trained with both ground-truth labels and teacher guidance, the student matches strong task-specific models while being far more efficient to train, store, and deploy. Beyond efficiency, the MoE layer improves transfer and generalization: experts develop reusable skills that boost cross-task performance and enable rapid adaptation to new tasks with minimal additions and tuning. On nine NGSS-aligned science-reasoning tasks (seven for training/evaluation and two held out for adaptation), UniMoE-Guided attains performance comparable to per-task models while using $\sim$6$\times$ less storage than maintaining separate students, and $87\times$ less than the 20B-parameter teacher. The method offers a practical path toward scalable, reliable, and resource-efficient automated scoring for classroom and large-scale assessment systems.

Poster #7

A Framework for Human-AI Q-Matrix Refinement: A NeuralCDM Evaluation

Ying Zhang

Q-matrices are binary mappings that link assessment items to the knowledge components or misconceptions they test. They are essential for Cognitive Diagnosis Models (CDMs) to produce actionable student profiles. The Problem: Expert-crafted Q-matrices are subjective, hard to scale, and difficult to validate. Data-driven methods lack semantic interpretability.

Our Solution: A Generate-and-Validate framework where LLMs generate candidate Q-matrices and NeuralCDM provides empirical evaluation, connected by lightweight expert feedback.

Poster #11

Teacher PD on AI Integration: A Qualitative Study of Programs and Pedagogies

Zixian Jiang

This study investigates how current AI professional development (AIPD) programs support teachers' integration of artificial intelligence in education. Using qualitative document analysis, we examined 24 publicly available AIPD programs identified through systematic web search and screening. A total of 748 documents were analyzed using coding frameworks based on UNESCO (2022), Celik (2023), and Dogan et al. (2025), focusing on AI tools' purposes, levels of AI integration, teacher knowledge (i-TPACK), and pedagogical design features.

Findings show that AIPD programs primarily emphasize instructional design, student engagement, and AI literacy, with most content positioned at the introductory (Acquire) level. Teacher knowledge development is concentrated on technological (i-TK) and technological pedagogical knowledge (i-TPK), with limited attention to content-specific (i-TCK) and integrated (i-TPACK) knowledge. While ethical considerations are commonly included, pedagogical approaches rely heavily on active learning and lectures, with minimal coaching support.

These results highlight the need for more content-specific, inquiry-based, and transformative AIPD designs to better support meaningful AI integration in education.

Poster #12

ArguAgent: A Dialogue System for Scaffolding Argumentative Discourse with AI

Jennifer Kleiman

Getting students to argue well with evidence, not just opinions, is hard to teach. The teacher might break the class up into twelve discussion groups but can only help one group at a time. ArguAgent gives teachers a way to run multiple small-group argumentation discussions simultaneously, with an AI co-teacher monitoring each group and stepping in when students get stuck, go off-topic, or skip over evidence. A five-tier analysis pipeline tracks what's happening in each conversation, from basic turn-taking up through argument quality and intervention decisions, and delivers targeted scaffoldng autonomously The discussion system is integrated with the GENIUS platform and has begun testing in classrooms.

Poster #28

Engineering Miniature Smart City Testbed for Multi-Vehicle Experiments for Future Educational Purposes

Nikhita Raghavan

Autonomous vehicles are increasingly becoming imperative to the future of transportation and robotics. Utilizing artificial intelligence and advanced robotic sensing technologies, the integration of autonomous systems has the potential to increase efficiency and scalability specifically in autonomous vehicles. The Duckietown project, founded at MIT in 2016, serves as a small scale simulation to understand real world integration. The testbed components consist of scaled highway systems complete with traffic lights, stop signs, multi-robot interaction, road signs, and various road topologies, such as three-way intersections and four-way intersections. It serves as an accessible source for understanding autonomous vehicles and is simultaneously utilized for educational purposes. The research study intends to elevate the experiment with the Moorebot Scouts, a collection of small-scale robots that utilize ROS (Robot-Operating System) to perform given tasks. Performance is evaluated across varying conditions, such as the ability to maintain lane position at increasing speeds and responsiveness to obstructions. Using these features, the study aims to create various experiments showcasing the wide range of capabilities of this project highlighting multi-robot interactions around the course. Outside of experimentation, future work could investigate the implications of its usage in K-12 robotics education; introducing students from elementary to high school to fundamental computer science concepts through a hands-on learning model. While showcasing the different capabilities the project strives to introduce the integration of various types of robots and how that can be utilized to bring a novel approach to fulfill Duckietown's overarching goal of computer science accessibility and education.

Theme 2: Ethics of AI for Human Society

Poster #17

Output-Level Biases in LLM Regional Storytelling: Lexical Patterns across Africa, Western Asia, and South Asia

Shuborno Anwar

Large Language Models (LLMs) are increasingly utilized in daily life. However, past research has shown that algorithmic systems regularly reproduce societally held biases. This study researches identity-based biases in LLM-generated storytelling across global geographic regions, focusing on three: Africa, Western Asia, and South Asia.

Through using chi-square keyness analyses on a corpus of differing models' outputs, we find statistically significant lexical patterns in each model. Our results show that the queried LLMs produce model-specific consistent vocabularies that reinforce gender roles, region-based stereotypes, and biases towards Western, Educated, Industrialized, Rich, and Democratic (WEIRD) societies. Moreover, the results suggest that the shown societal biases operate at the token level, reflecting profound structural patterns in training data.

Poster #21

Collective Robot Accountability via Layered BT Explanations (CORAL)

Moneshk Rai Bajaj

The Collective Robot Accountability via Layered BT Explanations (CORAL) automatically generates and evaluates structured natural-language explanations of multi-robot mission executions. It helps human operators diagnose failures in multi-agent behavior tree deployments by tracing causal chains across multiple robots.

Operators today receive either raw execution logs or high-level status flags. Neither reveals which robot genuinely caused a failure, how that failure propagated across the team, or whether a generated explanation is factually faithful to what the system actually did.

The system operates in three phases: per-agent execution logging, multi-agent causal attribution, and explanation evaluation via MERIT, measuring temporal accuracy, causal fidelity, and cross-robot coherence against ground truth derived from simulation logs.

Poster #27

Extracting Expert Biases Using Inverse Reinforcement Learning

Joseph George

As robots improve and are increasingly brought into collaborative environments with humans, the ability for robots to be able to adapt to individual human biases in task performance becomes essential for quickly creating and maintaining safe and effective working environment. Traditionally, Inverse Reinforcement Learning has not been able to disentangle task dynamics from agent-specific behavior. This causes difficulties in generalization when the robot must work with various partners.

To this end, we propose a framework for extracting expert biases from demonstrations by factoring the task state and agent state before IRL training.

Theme 3: AI and Future of Health & Work

Poster #4

ConTraIRL: Factorized Contrastive Abstractions for Transferable IRL

Yikang Gui

ConTraIRL addresses the challenge of transferring rewards in IRL to environments with unseen combinations of dynamics and goals. It learns disentangled latent representations for dynamics and goal factors, enforces their independence, and aligns them over time to capture task progression. The reward is defined by similarity to expert behavior in these latent spaces and learned via contrastive objectives. Experiments on continuous control benchmarks show improved reward transfer under compositional settings.

Poster #13

A Pipeline for 3D Gaussian Splatting via Meta Quest 3

Yulong Wang

This project presents VReconstruct, an end-to-end pipeline for high-fidelity 3D Gaussian Splatting (3DGS). The workflow integrates Meta Quest 3 spatial sensors for guided scene capture—using Unity coordinates as a spatial priori—with automated processing on NVIDIA B200 GPU clusters for rapid reconstruction. The results are viewed through a dedicated Unity-based desktop application optimized for real-time rendering and navigation.

Poster #14

AD-GPT: Large Language Models in Alzheimer's Disease

Ziyu Liu

Alzheimer's disease (AD) research produces extensive genomic and clinical data, yet general large language models (LLMs) often generate inaccurate or superficial outputs. We introduce AD-GPT, a domain-specific framework for reliable information retrieval and synthesis of AD-related knowledge.

We integrated curated genomic resources, including cis-eQTL and sQTL data across 13 brain regions from GTEx, genomic location information from NCBI, and gene function annotations from OMIM, together with approximately 150{,}000 AD-related publications from NCBI's PubMed. AD-GPT adopts a retrieval-augmented generation (RAG) workflow with task-specific database partitioning, a BERT-based query router, and fine-tuned Llama models, augmented with router and context verifiers to validate task assignment and evidence relevance, supporting three tasks: genetic information retrieval, association study reasoning, and general AD-related knowledge synthesis.

AD-GPT consistently outperformed strong baseline LLMs in evidence-grounded evaluation metrics across all tasks, including factual consistency, citation validity, and instruction-level faithfulness. Task-specific retrieval and stacked routing improved evidence grounding and substantially reduced hallucination in complex AD-related queries.

Poster #15

CORAL: Covariance-Guided Resource Adaptive Learning for Efficient Edge Inference

Zaki Indra Sukma

For deep learning inference on edge devices, hardware configurations achieving the same throughput can differ by 2× in power consumption, yet operators often struggle to find the efficient ones without exhaustive profiling. Existing approaches often rely on inefficient static presets or require expensive offline profiling that must be repeated for each new model or device. To address this problem, we present CORAL, an online optimization method that discovers near-optimal configurations without offline profiling. CORAL leverages distance covariance to statistically capture the non-linear dependencies between hardware settings, e.g., DVFS and concurrency levels, and performance metrics. Unlike prior work, we explicitly formulate the challenge as a throughput-power co-optimization problem to satisfy power budgets and throughput targets simultaneously. We evaluate CORAL on two NVIDIA Jetson devices across three object detection models ranging from lightweight to heavyweight. In single-target scenarios, CORAL achieves 96% – 100% of the optimal performance found by exhaustive search. In strict dual constraint scenarios where baselines fail or exceed power budgets, CORAL consistently finds proper configurations online with minimal exploration.

Poster #18

EMSTT-AD: Explainable Multimodal Static–Temporal Transformer for Alzheimer's Disease Progression

Zeliang Sun

Alzheimer's disease (AD) is a progressive neurodegenerative disorder marked by long-term cognitive decline and pronounced clinical heterogeneity. Accurately modeling cognitive trajectories and disease subtypes from longitudinal multimodal data remains challenging due to irregular visit intervals, missing observations, and complex interactions among heterogeneous biomarkers. Existing methods often rely on implicit temporal assumptions or ad hoc multimodal fusion strategies, limiting their capacity to capture long-range disease dynamics. We propose EMSTT- AD, an Explainable Multimodal Static–Temporal Trans- former for AD progression that explicitly models interactions between static and dynamic multimodal features under irregular temporal sampling. EMSTT-AD employs structured cross-modal representations to integrate heterogeneous modalities and leverages a temporal transformer to capture long-term dependencies across clinical visits, enabling robust prediction of disease subtypes. In addition, we leverage feature attribution techniques to quantify the contributions of individual biomarkers and incorporate prompt-based large language models (LLMs) to generate patient-specific clinical reports informed by the learned representations and predictions of EMSTT-AD. Experiments on data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) demonstrate that EMSTT-AD consistently out-performs benchmark approaches in diagnostic classification, highlighting its effectiveness for modeling AD progression from longitudinal multimodal datasets.

Poster #20

Are we there yet!? Optimality and Uncertainty in Open Multi-Agent Systems

Daniel Redder

Modeling open systems, where agents, tasks, and agent capabilities change over time is exceptionally challenging. Fundamentally this is because of large sparse and unbounded action and observation spaces. In my recent work, I created a taxonomy of openness delineating between bounded and unbounded worlds. Here I discuss what it means for a policy to be optimal in an open system, how we can determine convergence in these systems, and I show how a belief update can be used to predict the presence of returning model components.

Poster #22

Multi-Modal AI Pathology: Bridging Whole-Slide Imaging and Spatial Transcriptomics for Cancer Biomarker Discovery

Jayant Parashar

Precision medicine in oncology requires integration of clinical, molecular, and pathological data, yet current AI-pathology tools rely on convolutional neural networks trained against expert annotations that are inherently limited by subjective interpretation, coarse labeling, and poor reproducibility. This seed project addresses these limitations by uniting whole-slide imaging (WSI) with CosMx Spatial Molecular Imaging to generate molecularly defined, pixel-level ground truth for AI-based pathology. Leveraging shared biobank infrastructure across the University of Georgia's School of Medicine Spatial Molecular Imaging Center and the College of Veterinary Medicine's Department of Pathology, the project links high-resolution histology with spatial transcriptomic "satellite views" to train biologically grounded models that achieve superior accuracy, interpretability, and cross-species generalizability. A central innovation is an interactive expert-annotation environment (Aim #2) that couples WSI and molecular data streams in real time, enabling pathologists to validate and iteratively refine AI outputs. By embedding expert supervision directly into model development and spanning human and veterinary samples, the framework establishes a reproducible foundation for next-generation digital pathology and advances UGA's leadership in precision health.

Poster #26

Zeroth-Order Optimization for Memory-Efficient Fine-Tuning LLMs

Harshitha Attaluri

The expanding size of large language models (LLMs) creates challenges for memory use during subsequent model training because gradient-based optimization processes require both backpropagation work and storage of activation data. The research examines memory-efficient zeroth-order optimization techniques which allow gradient-free fine-tuning through forward-pass loss evaluations that estimate update directions. The research starts with MeZO as a baseline which proves that zeroth-order optimization can successfully adapt LLMs without requiring direct gradient calculations. The research progresses through three stages which begin with LOZO, LABA and AZURE because these three methods enhance optimization accuracy through their implementation of low-rank perturbations and adaptive subspace learning. The full-space perturbations of MeZO decrease memory requirements yet they result in high-variance updates which occur throughout high-dimensional parameter spaces. The system uses low-rank subspace updates, LABA improves the basis adaptively, and AZURE further enhances efficiency through online ϵ-greedy subspace adaptation. The system provides more resource-efficient options for LLM fine-tuning through its combination of these methods.

Poster #32

Grounding Clinical AI Simulation: A Trajectory-Driven Physiology Engine for Multi-Agent ED Training

Xinliang Li

Multi-agent emergency department simulations offer a promising testbed for clinical AI training, but existing approaches rely on either rule-based physiology engines that lack realism or LLM-driven physiology that is difficult to validate. We present a trajectory-driven physiology engine grounded in MIMIC-IV-ED, designed to enable rigorous evaluation of clinical reasoning in simulated ED environments. The core contribution is a library of 1,040 medication effect profiles derived from 425,000 ED stays, calibrated via a two-step process: first, deconfounding raw vital-sign changes by subtracting estimated disease-drift per chief complaint; second, back-calculating peak drug effect magnitude from median observation timing to correct for measurement sampling bias. Each profile encodes a triangular temporal response curve (onset, peak, offset) across multiple vital dimensions, with adverse event logic that checks patient allergies, drug interactions, and overdose thresholds against ground-truth patient state. To isolate the physiology engine as the testable variable, clinical treatment plans are fixed as trajectories — ordered sequences of medications and orders assigned quality grades (gold standard, suboptimal, harmful, negligent) — while LLM agents retain full autonomy over assessment, communication, and explanation. An information model separates PatientGroundTruth from DiscoveredInfo, ensuring adverse events fire regardless of team awareness. We demonstrate the system with an LVOT obstruction scenario and outline an evaluation protocol in which the same patient is run under all four trajectory quality arms, with the hypothesis that trajectory grade produces statistically distinguishable vital-sign outcomes.

Poster #34

Domain-Specific RAG for Georgia EV Supply Chain Intelligence

Sreeja Muvva

This project introduces an end-to-end EV supply-chain intelligence framework for Georgia, combining domain-specific web corpus creation, county-aware geospatial retrieval-augmented generation (RAG), and structured LLM benchmarking. A Tavily-based pipeline curates credible research data from noisy web searches. The GeoSpatial RAG system integrates structured supplier data and geospatial tools to support supplier discovery, county-level analysis, and supply-gap detection with citation-backed responses. Multiple local LLM pipelines were benchmarked on validated EV supply-chain questions, demonstrating that structured-first RAG outperforms no-RAG baselines for domain-specific QA. Overall, the approach shows that grounded data, spatial retrieval, and evidence-based evaluation yield more reliable and useful supply-chain intelligence systems.

Poster #35

Quest-Based Motion Capture for 3D Human Reconstruction

Ethan Nash Carroll

This paper presents an early prototype pipeline for capturing dynamic human movement with a Meta Quest 3 headset and reconstructing three dimensional body geometry for later extended reality playback. The system exports raw headset camera data, decodes Quest image buffers into standard RGB frames, and performs remote inference with SAM 3D Body on a laboratory graphics processing unit server. We report a working proof of concept in which a Quest derived frame is successfully reconstructed into a full body output. We then position MoCat Studio as the next stage for sequence level temporal tracking and motion consistency. Rather than claiming a complete end to end motion capture system, this work establishes feasibility for a headset centered workflow that combines consumer mixed reality hardware with modern human mesh recovery methods.

Poster #36

Predicting Child Cognition Using Machine Learning Methods in the Future of Families Study

Praveen Rangavajhula

This study examines how early life adversity and family context predict child cognitive outcomes using longitudinal data from the Future of Families and Child Wellbeing Study. We compare linear models and random forests across ages 3, 5, and 9. Predictive performance is modest but improves at later ages. Random forest models capture more complex patterns and achieve the strongest performance when using a broader set of predictors.

Poster #37

Automated Sickling Dynamics Analysis Using an agentic pipeline for Real Time Cell Progression Monitoring in Sickle Cell Disease

Suqiang Ma

This poster presents a an agentic pipeline for real-time analysis of red blood cell sickling dynamics in sickle cell disease, integrating YOLO-based detection, EfficientNet-B3 classification, and Siamese-network-based tracking to monitor morphological changes under hypoxia. Using high-resolution microscopy data from HbSS blood samples, the system achieves accurate multi-class classification and dynamic tracking of cell polymerization, enabling automated and quantitative assessment of disease progression and therapeutic efficacy.

Theme 4: AI for 3F - Farm, Food & Forests

Poster #6

Validation of USDA TreeMap: Mean Accuracy vs Spatial Accuracy in Southeastern Forests

Jiayi Tu

The USDA TreeMap provides 30-meter forest attribute maps widely used for biomass and carbon estimation, yet existing validations rely on RMSE and bias metrics that measure mean accuracy without assessing spatial predictive skill — whether predicted high and low values correspond to their correct locations. Using FIA field plots in Talladega National Forest, Alabama, this study compared TreeMap 2016 and 2022 across key forest attributes, employing the Concordance Correlation Coefficient ($CCC = \rho \times Cb$) to separate mean accuracy from spatial accuracy. We find that near-zero bias can arise from error cancellation rather than true predictive skill, and that improving mean accuracy can paradoxically destroy spatial accuracy — a pattern driven by variance compression from increased false-positive tree detection in the 2022 version. Scale-dependent analysis further confirms that TreeMap is suitable for regional totals but not spatial mapping. We recommend CCC decomposition as a standard evaluation framework, and provide version-specific guidance for operational use.

Poster #8

Deep Learning-Driven Individual Tree Detection in LiDAR Point Clouds

Weixuan Liu

Due to factors such as overlapping canopies, varying tree heights, and dense undergrowth, tree identification often faces challenges related to the underrepresentation of understory trees, as laser-based methods may fail to effectively penetrate canopies in dense, heterogeneous natural forests. This study introduces a novel approach to tree identification and segmentation in natural forest environments using Light Detection and Ranging (LiDAR) point clouds, thereby circumventing conventional clustering techniques such as DBSCAN and K-Means. A deep learning model was developed to enhance the efficiency and accuracy of detecting lower sections of tree stems. The Tree Stem Identification (TSI) model utilizes multiple samples from lower stem sections as primary training inputs. Using the predicted lower stem sections from the TSI model as seed points, the Individual Tree Segmentation (ITS) model efficiently identifies and segments tree structures above them. By leveraging these seed points, the proposed method streamlines tree identification and enables accurate delineation of individual trees, even in complex forest environments.

Poster #9

Cyber-Physical Digital Twin for Robotics in Poultry Farms

Jonathan Porter Squires

Integrating robotics into precision animal farming is increasingly important but also challenging. To accommodate the risks and costs associated with experimenting with robotics on real poultry farms, we create a simulated poultry house digital twin to facilitate robotics experiments and testing, and to assist with user interactions. To ensure the environment is realistic, we are creating a behavior tree-driven autonomous decision-making system tailored to mimic the time budgets of animal and robotic behaviors recorded in literature and qualitative behavior from empirical data. Specifically, we integrate this large-scale simulation with previous research on predicting poultry death in order to allow training of models related to collected poultry mortalities. We stress test the simulation setup with our robotics solutions for collecting mortalities and path planning objectives. In the future, we plan to conduct experiments with our digital twin to optimize the robotic algorithms before deploying to real farms as well as adapting the deployed robotic systems based on changes in environmental conditions.

Poster #24

Enhancing 3D Reconstruction Using Multi-Objective Loss Functions in Neural Radiance Fields for Broad-scale Outdoor Scenes

Cameron Halaby

This study investigates the application of Neural Radiance Fields (NeRF) to generate continuous 3D renderings from UAV-captured data. The research specifically explores whether integrating a multi-objective loss function, combining traditional pixel-level metrics with Learned Perceptual Image Patch Similarity (LPIPS), can improve the structural and perceptual quality of reconstructions. By evaluating four experimental conditions that varied input scope, frequency, and weight scheduling, we assessed the trade-of s between mathematical pixel accuracy and human-centric visual fidelity. Our results demonstrate that periodically applying full-image LPIPS calculations (every ten iterations) significantly enhances edge preservation and structural coherence without sacrificing pixel-level precision. Additionally, though basic forms of NeRF involve computationally-heavy operations (e.g., sampling thousands of points along each ray out of millions of rays in a single image), our research explores methods to reduce these demands while maintaining high-quality 3D reconstructions. Use of perceptual similarity reduces the amount of ray sampling required in NeRF processing, creating a model that requires significantly less image overlap, geometric computation, and processing time, leading to more efficient data collection and richer feature reconstructions for UAV platforms, with the long-term goal of supporting onboard reconstruction capabilities for CubeSat missions. These findings suggest that prioritizing perceptual logic allows for more photorealistic and reliable 3D models, ultimately providing a more robust framework for monitoring real-time changes in complex physical environments.

Poster #25

Terrain-Aware Neural Radiance Fields for Geospatial Reconstruction from Nadir-Only UAV Imager

Evan Hammam

Unmanned aerial vehicle surveys with nadir imagery support environmental monitoring, change detection, and terrain analysis, but these applications require georeferenced point clouds with reliable elevations. Neural Radiance Fields learn density and appearance from imagery, yet nadir-only capture leaves geometry underconstrained along the vertical axis. This thesis introduces Terrain-Gated Neural Radiance Field, an extension of Depth-Nerfacto that learns a terrain surface and suppresses density below it. Training used radiometrically calibrated, pansharpened multispectral imagery acquired in Fairbanks, Alaska in May 2024, November 2024, and April 2025 with a MicaSense RedEdge-P sensor on a Freefly Alta X unmanned aerial vehicle. Evaluation compared Nerfacto, Depth-Nerfacto, and Agisoft Metashape Multi-View Stereo using rendering metrics, Light Detection and Ranging (LiDAR)-referenced geometric metrics, and overlap-reduction tests. Terrain-Gated Neural Radiance Field matched Depth-Nerfacto in peak signal-to-noise ratio, structural similarity, and learned perceptual image patch similarity while producing compact point clouds and fewer below-ground outliers. Under full overlap, it achieved a mean elevation root mean square error of 1.82 meters against LiDAR, compared with 1.98 meters for Depth-Nerfacto, while Multi-View Stereo remained strongest at 0.55 meters. Reduced overlap degraded geometry more than rendering quality, and Terrain-Gated Neural Radiance Field remained the most stable model. These results show that terrain-aware density gating improves nadir-only neural radiance field geometry without sacrificing rendering quality and that independent geometric evaluation is necessary for geospatial mapping.

Poster #29

Valuation of Pine Stands Under Uncertainty

Chaz Merritt

Stand-level forest planning is commonly conducted with deterministic growth and yield models that assume a single, idealized trajectory of stand development. However, forest stands evolve under substantial uncertainty arising from stochastic growth variation and disturbance events such as pests, wildfire, and wind damage. This thesis develops a modular stand simulation framework that augments an established deterministic loblolly pine (Pinus taeda) growth model with stochastic process noise and probabilistic disturbance regimes, and uses Monte Carlo simulation to quantify how uncertainty alters economic valuation.

In this work, we evaluate a set of predefined, interpretable management prescriptions under both deterministic and stochastic dynamics. For each prescription and risk regime, repeated simulations generate a distribution of discounted cash flows and corresponding Net Present Value (NPV) and Land Expectation Value (LEV). Outcomes are summarized using measures of central tendency and dispersion as well as downside risk metrics, including Value at Risk (VaR), Conditional Value at Risk (CVaR) to demonstrate tail-risk.

Results show that deterministic valuation can systematically mischaracterize stand value under uncertainty, particularly when low-probability, high-severity disturbances are possible. Across increasing disturbance frequencies and severities, the stochastic valuation distributions exhibit substantial tail risk and increased variability relative to deterministic baselines, and financially preferred rotation ages and prescriptions may shift accordingly. Collectively, these findings demonstrate the importance of distributional valuation in long-horizon timber management and provide a practical framework for comparing prescription robustness under stochastic disturbance regimes.

Poster #30

H-IPP: Altitude-Adaptive Informative Path Planning with Multiple UAVs

Swaraj Nistane

We present H-IPP, an altitude-adaptive information-gathering framework for multi-UAV mapping of unknown non-stationary spatial fields under sensing uncertainty. The key challenge is the trade-off between coverage and measurement quality: higher altitudes provide a larger field of view, while lower altitudes improve sensing fidelity and spatial resolution. To address this, the environment is modelled using a heteroscedastic Gaussian Process with an Attentive Kernel, which captures non-stationary spatial correlations and learns local lengthscales to represent complex environmental structure. Multi-UAV coordination is achieved through weighted Voronoi partitioning driven by the learned lengthscales, enabling sensing effort to be distributed according to spatial information density. A two-stage discrete optimization jointly selects waypoint and altitude using a multiplicative utility function that combines coverage, expected information gain, and variability-aware exploitation. Beyond accurate field reconstruction, the learned GP hyperparameters are also used to derive a semantic variability map, enabling unsupervised structure-aware region segmentation and supporting downstream heterogeneous robotic deployment.

Poster #38

Generative Artificial Intelligence for Planning Scenario Visualization: A Framework and Benchmark for Controllable Street-Level Image Synthesis

Qian Cao

The rapid advancement of Generative AI (GenAI) creates new opportunities for spatial planning. However, most existing applications focus on macro-scale overhead perspectives, leaving street-level visualization constrained by strict physical and spatial fidelity challenges. We introduce a controllable GenAI-human interactive framework for producing spatially faithful and realistic street-level imagery that reflects envisioned planning scenarios. Using solar farm planning as a test case, it leverages vision-language models to integrate human inputs (e.g., textual intent, scribbles) and existing status photos, synthesizing photorealistic outcomes. To empirically evaluate this, we constructed a large- scale street-view imagery dataset of before-and-after solar panel installation across the United States. We comprehensively evaluated 11 model settings covering five open-sourced diffusion models and one commercial vision foundation model to test fidelity and realism. Our study illustrates how a human-in-the-loop GenAI pipeline can support visualizing proposed environmental interventions and offers a benchmark for conditioned GenAI visualization applications.

Poster #39

AI-Assisted Breeding for Anthracnose Resistance in Pepper (Capsicum spp.)

Dr. Deepa Jaganathan

Anthracnose, caused by multiple Colletotrichum species, is one of the most destructive diseases of pepper (Capsicum spp.), leading to substantial yield losses under both field and post-harvest conditions globally. Progress in breeding for resistance is constrained by limitations in phenotyping, as disease assessment relies primarily on subjective visual scoring that is labor-intensive, inconsistent, and poorly suited for quantitative analyses. Recent advances in artificial intelligence (AI), particularly deep learning–based computer vision, provide an opportunity to develop objective, scalable, and high-throughput approaches for disease quantification. The present study aims to develop a convolutional neural network (CNN)-based framework for automated detection and severity quantification of anthracnose in pepper. Seven bell pepper accessions (UGA-CAPS-1, UGA-CAPS-254, UGA-CAPS-259, UGA-CAPS-315, UGA-CAPS-318, UGA-CAPS-361, and UGA-CAPS-444) were grown under controlled greenhouse conditions. Fruits were harvested at green or red maturity stages based on accession-specific readiness and subjected to anthracnose inoculation. Standardized image acquisition was performed before inoculation and at defined intervals following inoculation, including 0, 48, 72, 96, and 144 hours, to capture disease progression dynamics. A structured image dataset representing multiple stages of lesion development has been generated, and image preprocessing and systematic file organization is underway. Once image preprocessing and annotation are completed, CNN-based model development will be initiated for automated lesion detection and quantitative disease severity estimation. The resulting digital phenotypes offer reliable and reproducible measurements for integration with genomic and environmental data. This study establishes an AI-driven phenotyping pipeline for pepper anthracnose, addressing limitations of manual scoring and enabling more efficient, data-driven selection. Future work will scale this approach from the initial seven accessions to the complete UGA-Capsi-core collection, enabling the use of AI-derived phenotypes for genome-wide association studies and marker discovery to identify candidate genes underlying anthracnose resistance. In the long term, it supports the development and deployment of anthracnose-resistant cultivars and enables AI-driven decision-support tools for growers.

Theme 5: AI for Cyber & Societal Security

Poster #10

When Decentralized Training Outperforms Centralized: Insights from Competitive Sparse-Reward MARL

Lingze Zeng

We present the first learning-based solution to the Partial Information Target Defense Game, a zero-sum pursuit-evasion problem under partial observability and sparse rewards. Contrary to cooperative MARL findings, we show that Decentralized Training outperforms Centralized Training, achieving a higher defender success rate, and attribute this to gradient interference in centralized critics within competitive settings.

Poster #16

Edge-AI for Safe and Efficient Autonomous Vehicles

Zhipeng Bao

Autonomous vehicles frequently encounter decision impasses in complex, ambiguous environments. While Vision-Language Models (VLMs) demonstrate strong reasoning capabilities, their traditional reliance on single-frame processing fails to capture essential temporal dynamics. Furthermore, high computational demands severely limit their deployment on resource-constrained edge hardware. This project proposes an efficient, edge-deployable multi-frame VLM framework to bridge these gaps. We first curate a dataset of continuous video clips featuring real-world and simulated traffic impasses. Next, we fine-tune a lightweight VLM, applying temporal token pruning and quantization techniques to minimize redundancy and enable rapid edge inference. The optimized model is integrated as an onboard advisory module. When native planning fails, it activates to analyze temporal context and passenger input, generating actionable recovery strategies. This approach significantly enhances the safety, interpretability, and resilience of autonomous systems during critical edge cases.

Poster #19

Modeling LLM-Driven Action Discovery in Cyber Adversaries

Aditya Shinde

Recent developments in generative AI and large language models are creating a new dimension of cyber warfare. Adversaries can now leverage LLMs to discover novel vulnerabilities, generate new exploit code, and develop attack techniques. This fundamentally changes the nature of cyber threat — attackers can dynamically expand their capabilities during an ongoing campaign. Current decision-making frameworks for modeling such adversaries, including Markov Decision Processes and their extensions, are insufficient because they assume a fully specified and fixed action space known to the agent from the outset. They can model uncertainty about outcomes but not unawareness of capabilities.

We introduce a novel framework to model attackers that can expand this set irreversibly by querying an LLM oracle. We instantiate this framework in a simulated environment, where an attacker agent progressively discovers new exploit techniques through an LLM while learning to act via actor-critic methods over a heterogeneous graph neural network. The graph represents both network hosts and known attack actions, growing structurally as new actions are discovered.

Poster #23

DCL-Sparse: Distributed Range-only Cooperative Localization of Multi-Robots in Noisy and Sparse Sensing Graphs

Atharva Sagale

This work presents a novel approach to range-based distributed cooperative localization (DCL) for robot swarms in GPS-denied environments, relying solely on inter-robot range measurements, specifically addressing the limitations of current methods in noisy and sparse settings where the geometric non-rigidity of the sensing graph creates flipping (suboptimal) effects in the localization outcomes. We propose a robust multilayered localization framework (DCL-Sparse) that utilizes distributed 1-hop shadow edges (S1-Edge) to address the non-rigidity problem and improve localization convergence in sparse and noisy sensing graphs. Our approach leverages the advantages of distributed localization methods, enhancing scalability and adaptability in large robot networks.

We establish theoretical conditions for the new S1-Edge that ensure solutions exist even in the presence of noise, thereby validating the effectiveness of the new shadow edge localization. Extensive simulation and real-world experiments confirm the superior performance of our method compared to state-of-the-art techniques, resulting in a reduction of up to 93% in the localization error in DCL. These experiments demonstrate substantial improvements in localization accuracy and robustness to sparse graphs. DCL-Sparse increases the localizability of large multi-robot and sensor networks, offering a powerful tool for high-performance and reliable operations in challenging large-scale environments.

Poster #31

Spatial Intelligence Models for Robotics and Environmental Reasoning

Sai Krishna Ghanta

Environmental monitoring still faces major limitations in understanding hidden spatial processes such as Wi-Fi coverage, gas spread, fire, humidity, temperature, and sound. These phenomena are often invisible, high-dimensional, uncertain, and continuously changing, making them difficult to model, predict, and reason about using current sensing and foundation-model pipelines alone. These challenges become even more important in robotics, where agents must not only interpret such hidden processes but also act on them safely and effectively in real-world environments. In this work, we present an embodied reasoning framework that combines language, perception, and spatial process modeling to support understanding, prediction, and decision-making of robots in complex environments. Our approach converts raw multi-modal observations into a unified learnable representation. Building on this representation, we incorporate world models for counterfactual reasoning and future state prediction, enabling the system to answer queries, explain hidden spatial behavior, and support planning under uncertainty. Experimental results show clear improvements over text-only baselines across multiple reasoning tasks involving hidden and invisible phenomena. Overall, this work highlights a promising direction for environmental intelligence and opens the door to broader applications in robotics, smart buildings, disaster response, environmental monitoring, and other systems that must reason about unseen spatial dynamics.

Poster #33

Human-AI Tactical Teams in VR Environments

John Frericks

As artificial intelligence becomes increasingly embedded in combat and military contexts, there is a need to shift perspectives from seeing AI as merely a tool to treating it as a collaborative teammate. But to make these human-AI teams (HATs) successful, we have to understand how trust forms, breaks, and calibrates over time. Traditional testbeds often lack the real-world complexity needed to study this effectively. To bridge that gap, this research places users in a high-fidelity virtual reality (VR) environment alongside a fully autonomous, LLM-powered robot. The primary goal is to examine how the robot's transparency, physical form factor, and trust-building behaviors impact human reliance in high-stakes situations. In the VR simulation, participants are tasked with searching virtual rooms to gather intelligence clues, analyzing those clues to deduce a target city, and then collaborating with their AI partner to discuss their findings and finalize a tactical decision. To test the environment, we conducted a pilot study with 42 undergraduate psychology students using Meta Quest Pro headsets. We tracked a variety of metrics, ranging from their propensity to trust AI to their cognitive load and perceived team viability. Interestingly, we found that users' trust in the robot declined across two separate missions, and this erosion of trust appeared to lower their mission accuracy during the second phase. These early insights have helped us refine the VR experience for better data collection, setting the stage for a full-scale study launching in Summer 2026.