Accelerating materials discovery with ADAM: An AI co-pilot for physical experimentation

Anderson Borch[1], Emily Avey[1,2], Jayden Grunde[1,2], Michelle A. Smeaton[1], Renae N. Gannon[1], Steven R. Spurgeon[1,2,3,*]

1. National Laboratory of the Rockies Golden, CO, USA 80401,
2. Colorado School of Mines, Golden, CO, USA 80401,
3. University of Colorado, Boulder, CO, USA 80309

* Corresponding author: Steven R. Spurgeon, Steven.Spurgeon@nlr.gov

National Laboratory of the Rockies

Introduction

Problem: Complex instruments require years of expertise to operate safely. Experiment quality is limited by operator skill, leading to throughput and reproducibility bottlenecks.[1,2]

Gap: Existing autonomous systems fall into two failure modes: deterministic scripts that fail on novelty, and base LLMs that hallucinate domain-specific parameters and hardware control, an unacceptable risk in hardware control.[1,3,4]

Solution: A multi-agent framework combining three oversight layers:

- Retrieval-Augmented Generation (RAG) grounding
- Validated Tool Registry
- Semantic Human-in-the-loop (HITL) checkpoints

Design Goal: Intelligent checkpoints, not deterministic ones, mitigating the autonomy/oversight tradeoff. High autonomy on routine tasks, human oversight when required to accelerate the research cycle by intelligently automating time-consuming tasks, allowing scientists to focus on discovery.

Architecture

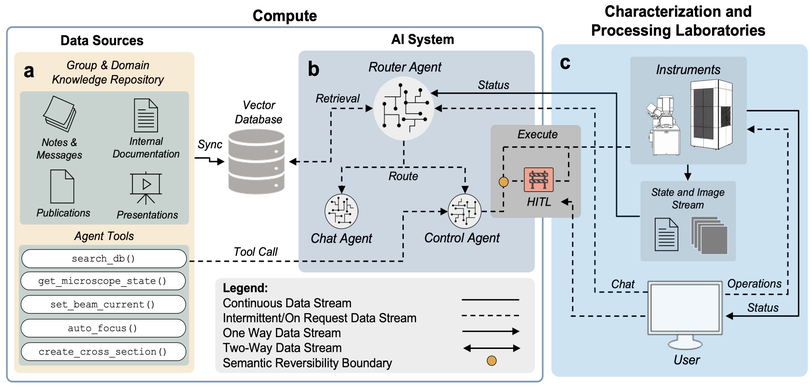


Figure 1: ADAM Architectural Diagram and Workflow

Results & Comparisons

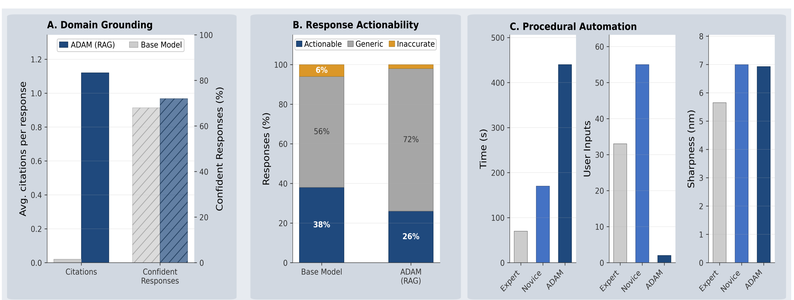


Figure 2: ADAM comparisons across domain grounding, response actionability, and procedural automation

RAG Results:

- 56x more citations per response
- 3x fewer inaccurate responses

Automation Results:

- 2 user inputs vs 33 and 55
- Quality comparable to novice

Processing Automation & Failure Reasoning

Setup: Single prompt to survey a Cr_2O_3/Ga_2O_3 diode array for failure sites. No protocol, reward function, or predefined workflow.

Execution: ADAM surveys 3 ROIs, autonomously imaging, analyzing, and diagnosing each, producing structured failure assessments: melting/cracking, bridging particle, clean baseline.

HITL Checkpoint: User reviews report and issues two redirects via natural language:

- Selected ROI 1 for a physical PFIB cross-section
- Selected ROI 2 for code generation to analyze & label defects

Key Result: ADAM autonomously imaged 3 ROIs, correctly identifying failure signatures that require extensive experience. Then, it milled a live PFIB cross-section, a process demanding significant expertise, and separately authored a custom Python script to analyze and label defects in future images.

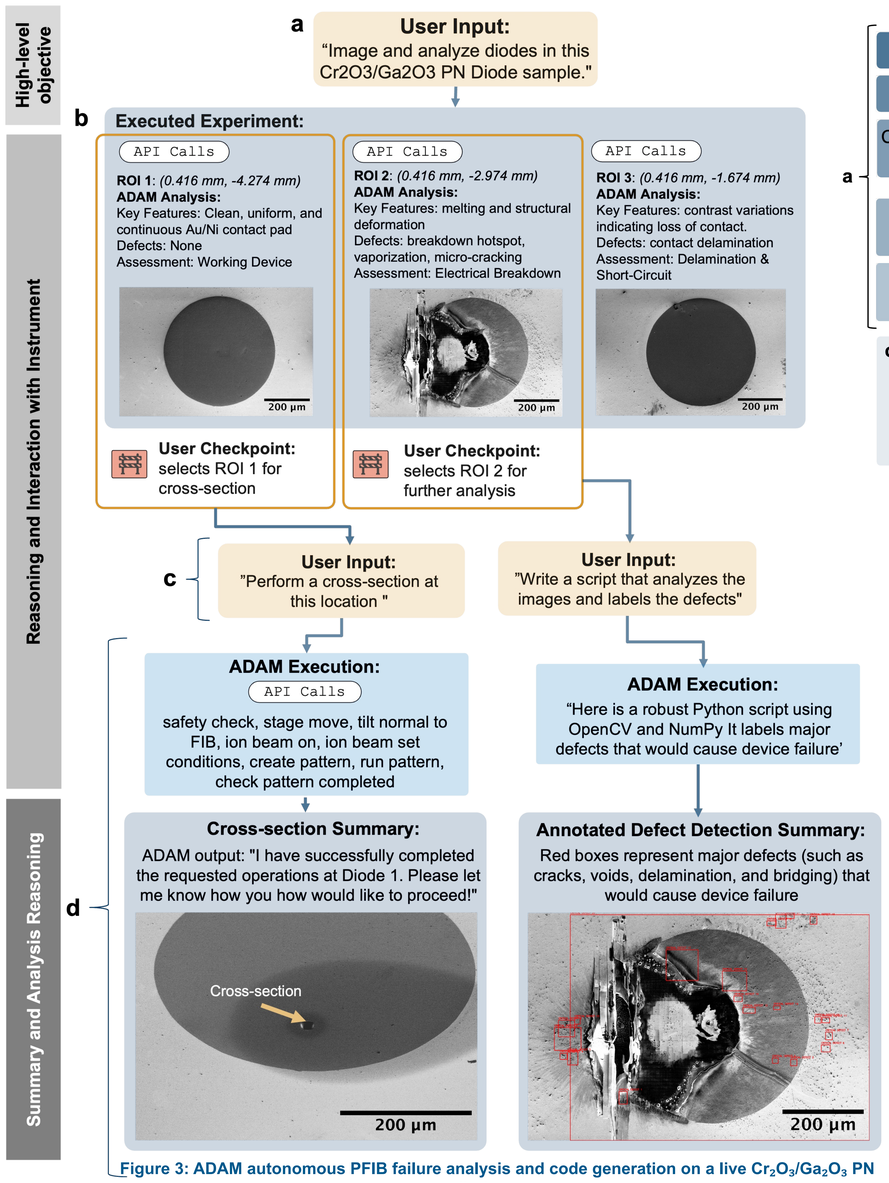


Figure 3: ADAM autonomous PFIB failure analysis and code generation on a live Cr_2O_3/Ga_2O_3 PN diode array

Characterization & Exploratory Analysis

Setup: A researcher issues a single prompt to design and execute a characterization experiment. No additional context provided.

ADAM's first step: RAG information retrieval, synthesizes a three-phase plan and presents it at a HITL checkpoint.

Execution: Acquire -> Analyze -> Act

- Phase 1: Morphology
- Phase 2: Crystal Structure
- Phase 3: Composition

Key Result: A complete, three step characterization session with rationale provided at every decision point, autonomously produced from a single open-ended prompt from a novice researcher, with no expert involvement or predefined protocol

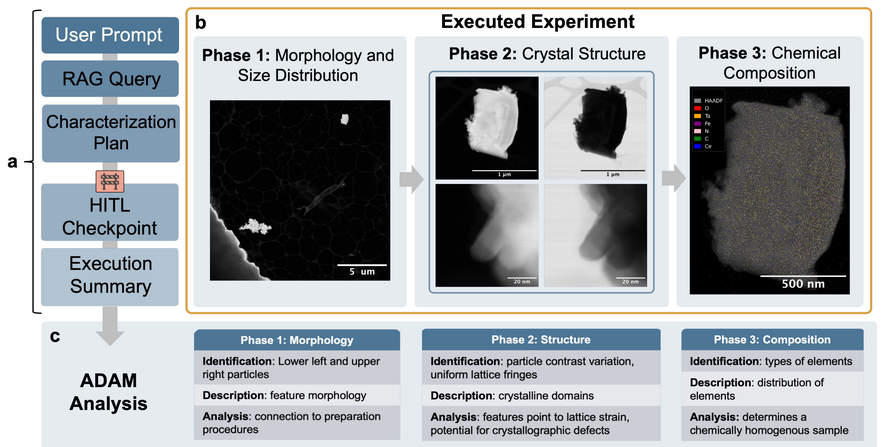


Figure 4: ADAM guided three-phase TEM characterization on an unknown nanoparticle sample

Conclusion & Future Work

Key Findings: HITL checkpoints at semantic reversibility boundaries mitigate the autonomy/oversight tradeoff while increasing throughput.

Broader Impact: A domain- and instrument-agnostic framework for safe, scalable autonomous physical experimentation

Future Work: Extend to multi-vendor instruments; reduce inference latency; package tool registries as MCP servers for portability

References

[1] Matthew J. Olszta, et al. An automated scanning transmission electron microscope guided by sparse data analytics. *Microscopy and Microanalysis*, 28(5):1611–1621, 2022.

[2] Sarah Akers, et al. Rapid and flexible segmentation of electron microscopy data using few-shot machine learning. *npj Computational Materials*, 7(1):187, 2021.

[3] Jiarui Li, et al. Enhancing llm factual accuracy with rag to counter hallucinations: A case study on domain-specific queries in private knowledge-bases. *arXiv preprint arXiv:2311.09138*, 2023.

[4] Mandal et al. Evaluating large language model agents for automation of atomic force microscopy. *Nature Communications,* 16(1):9104, 2025.

U.S. DEPARTMENT OF ENERGY
Office of Science

A.B. and E.A. were supported by the U.S. Department of Energy, Office of Science, Office of Workforce Development, for Teachers and Scientists (WDTS) under the Science Undergraduate Laboratory Internship (SULI) program. J.G., R.G., M.A.S., and S.R.S. were supported to conduct microscopy, AI design, and project management as part of APEX (A Center for Power Electronics Materials and Manufacturing Exploration), an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences under Award # ERW0345. This work was authored in part by National Laboratory of the Rockies for the U.S. Department of Energy (DOE), operated under Contract No. DE-AC36-08GO28308. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes. AI tools were used in the preparation of this poster to assist with content organization and editing decisions, including condensing and trimming text for layout purposes. All content was written and reviewed by the authors, who take full responsibility for the final text.

APEX
A Center for Power Electronics Materials and Manufacturing Exploration

Supercomputing
Chicago, Illinois
November 15th - 20th , 2026