Cybersecurity Innovation for CyberInfrastructure

Dan Massey
Kevin Thompson
Office of Advanced Cyberinfrastructure
National Science Foundation

August 2025

Cybersecurity Innovation for CyberInfrastructure (CICI)

Mission:

support trustworthy scientific discovery and innovation by enhancing the

security and privacy of ***scientific cyberinfrastructure***.

What is Scientific CyberInfrastructure (SCI)?

Tools, techniques, services, and datasets that enable scientific research

- *Advanced Networks*
  - High bandwidth/low latency networks connecting researchers to instruments, datasets, and computational resources.
- *Datasets and Repositories*
  - Datasets used in scientific research along with corresponding support for storage, management, and access control
- *Software Tools, Libraries, and Systems*
  - Software supporting scientific workflows including computational, modeling, and visualization tools
- *High Performance Computing*
  - Computational resources for processing, simulating, modeling, and visualizing scientific research challenges

What is Unique About Cybersecurity for Scientific CI?

- Scientific data and workloads differ from traditional network, storage, and computation.
- Science drivers operate at vast scales that push the limits of high performance computing;
  - Artificial Intelligence (AI) further increases these demands
- Scientific CI employs unique hardware and software with unique features and vulnerabilities
- Scientific CI is most effective when it enables collaborative, open, and reproducible results

CICI cybersecurity innovations are tailored for scientific cyberinfrastructure and enable trustworthy reproducible science

CICI Emphasis on Open Federated Datasets

- Problem: valuable datasets scattered across multiple locations known only to niche groups of researchers with varying degrees of data providence, integrity, and authenticity
    - Scientific use of AI dramatically exacerbates the challenge of finding and securing datasets
- CICI Solution: new Open and FAIR Dataset Sharing Plan requirements
    - CICI awardees submit an Open and FAIR Data Sharing Plan to publish datasets in open data federated system.
    - NSF and awardee identify common federated data system and set out annual goals for publishing datasets.
    - Progress toward the Open and FAIR dataset sharing plan in each annual report.
    - Aim is to ensure CICI funded datasets are widely visible and address integrity, provenance, and authenticity
    - Further aim is to support development and use of federated data systems
        - CICI has partnered with the Open Science Data Federation, supported by NSF under Grant Nos. 2030508 and 1836650

CICI

CICI Tracks

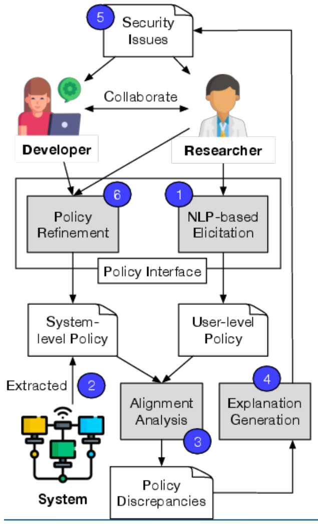
Usable and Collaborative Security for Science (UCSS)	7 awards in 2025, 3 year awards up to $600,000
Reference Scientific Security Datasets (RSSD)	2 awards in 2025, 3 year awards up to $600,000
Transition to Cyberinfrastructure Resilience (TCR)	6 awards in 2025, 3 year awards up to $1,200,000
Integrity, Provenance, and Authenticity for Artificial Intelligence Ready Data (IPAAI)	Introduced in 2025, 5 awards, 3 year awards up to $900,000

Usable and Collaborative Security for Science (UCSS)

- Enable secure and reliable ***collaboration*** among scientists
  - Collaboration across distributed scientific cyberinfrastructure with complex technical relationships between users, institutions, and information technology service providers.
- Ensure the ***usability*** of cybersecurity for domain scientists
  - Trade-off between research goals and security/privacy concerns specific to a scientific domain.
- Develop ***integrated cybersecurity environments*** including scientific cyberinfrastructure Security Operation Centers (SOCs) and secure computing enclaves.
  - SOC efforts focusing on small and under-resourced institutions that would facilitate the establishment of regional enclaves.

CloudSec: Collaborative Policy Alignment for Secure Scientific Computing Infrastructures

PI: Joe Stubbs (UT Austin) Co-PIs: Eunsuk Khang (CMU), Smruti Padhy (UT Austin)



- Collaborative security policy analysis for research cyberinfrastructure
- Understand abstraction gap between high-level user policies (natural language) and low-level system policies (formal language)
- Analyze gap using a combination of AI and formal methods.
- Identify vulnerabilities due to policy misconfiguration

Collaboration and Usability Challenge Addressed

- Researchers and project PIs have a high-level vision of the security requirements for their projects, but technical developers must implement security policies using low-level formal semantics.
- Collaboration between researcher and developer on security policies requires bridging the abstraction gap between high-level and low-level policy descriptions.

Technical Cybersecurity Solution

- Our approach utilizes a new cross-layer policy analysis based on the idea of *discrepancies*, which can be computed, clustered and treated using formal methods.
- We combine this cross-layer policy analysis with machine learning for translation of the high-level descriptions in natural language of the policy requirements to low-level system

Benefits to Scientific Cyberinfrastructure

- Development of a new "cross-layer" policy analysis with applications to many research infrastructures
- A novel interface for eliciting policy requirements and explanations from project stakeholders
- A workflow that facilitates collaborative policy development and refinement between researchers and technical developers

Risks Versus Potential For Advances

- Use of ML/LLMs presents risk (poor performance/hallucinations); Use of LLMs balanced by FM and can be augmented with more interface capabilities.
- Risk that results may not generalize beyond Tapis; Use of Tapis strikes a balance between traceable scale and general use.

Result Dissemination Plans

- New software will be integrated with the Tapis framework and evaluated using a controlled study for the difficulty of use, tool performance with respect to policy extraction, schema authoring, and policy analysis time. We will also measure tool correctness in terms of the number of discrepancies identified and policies modified/corrected.
- We will also measure impact to Tapis users; for example, the number of policies deployed by the system to Tapis projects.
- Results will be integrated into undergraduate and graduate courses at UT Austin and CMU.

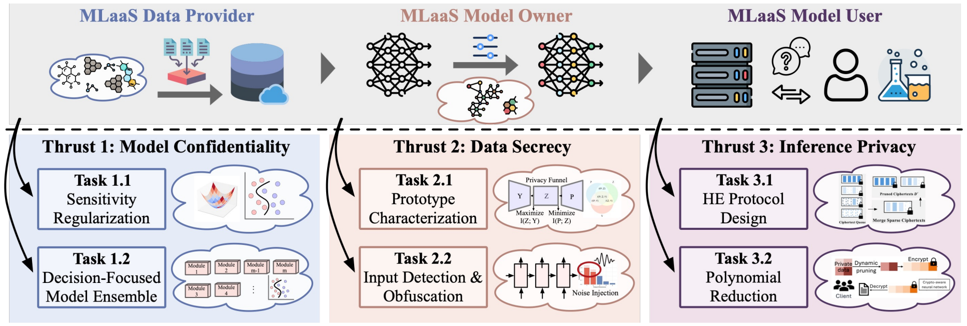
Programmatic Details

- 3 year project started on Sept, 2025.
- Led by UT Austin with CMU

This material is based upon work supported by the National Science Foundation under Grant No. 2530889

MLaaS: Secure Machine Learning as a Service for Collaborative Scientific Research

PI: Yushun Dong (Florida State University) Co-PIs: Neil Gong (Duke University), Mengxin Zheng (University of Central Florida)



Secure MLaaS framework that protects the three stakeholders in collaborative scientific computing environments.

Benefits to Scientific Cyberinfrastructure

- Strengthens security, integrity, and reproducibility of MLaaS-based scientific collaboration workflows.
- Enhances trust in collaborative CI for sensitive domains (healthcare, disaster response, genomics).

Risks Versus Potential For Advances

- **Risks**: Performance overhead from encryption; resistance to adoption if workflows disrupted; **Mitigation**: Efficiency-focused design; modular integration; continuous usability feedback loops.
- **Payoff**: A deployable, modular security framework that protects models, data, and inference in collaborative MLaaS environments.

Collaboration and Usability Challenge Addressed

- Model extraction attacks (MEAs) → theft of proprietary models.
- Model inversion attacks (MIAs) → leakage of sensitive training data.
- Inference data breaches → exposure of user queries and outputs.
- ## Technical Cybersecurity Solution
- Model Confidentiality: Sensitivity Regularization; Model Ensemble.
- Data Secrecy: Privacy Funnel; Suspicious Input Detection&Obfuscation.
- Inference Privacy: Efficient Encryption; Dynamic Polynomial Reduction.

Result Dissemination Plans

- Stolen model fidelity↓; reconstructed feature distance↑; encrypted inference latency↓; negligible task-relevant utility loss↓.
- Open-source software; Public datasets & benchmarks; Tutorials, user guides, workshops, etc.
- Audience: CI admins, domain scientists, research institutions.

Programmatic Details

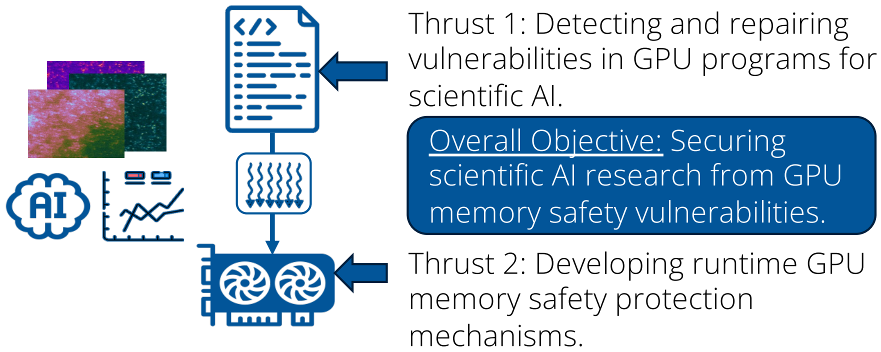
- 3 years project started on Jan. 2026
- Led by Florida State University and with Duke University and University of Central Florida

CICI UCCS

This material is based upon work supported by the National Science Foundation under Grant No. 2530786

SecGPU4AI: Securing GPU Computing for AI-Driven Scientific Workflows

PI: Yanan Guo (University of Rochester) Co-PIs: Zhenkai Zhang (Clemson University), Tong Geng (William Marsh Rice University)



Thrust 1: Detecting and repairing vulnerabilities in GPU programs for scientific AI.

Overall Objective: Securing scientific AI research from GPU memory safety vulnerabilities.

Thrust 2: Developing runtime GPU memory safety protection mechanisms.

Collaboration and Usability Challenge Addressed

- Provides missing GPU security solutions for AI-driven science to support safe collaboration across institutions.
- Delivers easy-to-adopt security tools usable by non-expert scientists.

Technical Cybersecurity Solution

- Develops a GPU-tailored fuzzing framework and lightweight protections through GPU system software modifications.
- Leverages proven CPU concepts, adapted to GPU execution models.

Benefits to Scientific Cyberinfrastructure

- For domain scientists: usable vulnerability fuzzing and repair tools.
- For CI providers: deployable runtime protection mechanisms.

Risks Versus Potential For Advances

- Risks: Protections may add runtime cost or vary across GPU drivers.
- Advances: Significantly strengthen the security, privacy, and reliability of AI-driven science.

Result Dissemination Plans

- Open-source release of tools with documentation.
- Collaboration with CI providers for deployment and integration.
- Responsible vulnerability reporting to AI software development teams.

Programmatic Details

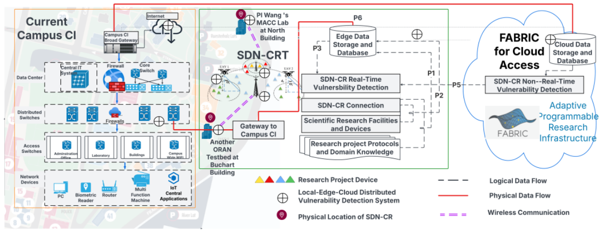
- 3-year project started on December 2025.
- Led by University of Rochester and with Clemson University and William Marsh Rice University.
- Unfunded collaborations with Tyler Sorensen (Microsoft Research)

This material is based upon work supported by the National Science Foundation under Grant No. 2530649

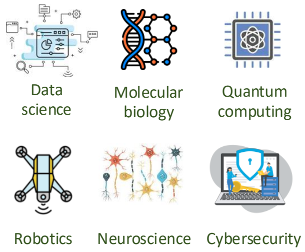
WRAP: Programmable Wireless Infrastructure with Formal Assurance for Cross-Campus Research

PI: Ying Wang (Stevens Institute of Technology) Co-PI: Juntao Chen (Fordham University)

Secure Programmable Wireless Cyberinfrastructure for Cross-Campus Research



Collaborative research areas



Implementation & Evaluation:

- Stevens and Fordham IT network pilot implementation and integration
- Digital twin–based wireless CI testbed
- Joint policy assurance and federated access for research workflows
- Usability evaluation via researcher engagement

Collaboration and Usability Challenge Addressed

- Collaborative research requires dynamic cross-campus wireless access, but the state of practice security approaches limit flexibility.
- WRAP combines open program infrastructure with formal assurance to align usability and compliance. Researcher-facing tools simplify configuration while IT teams maintain visibility and control.

Technical Cybersecurity Solution

- Formal assurance translates researcher goals into verifiable policies.
- Runtime anomaly detection ensures continuous, secure operation.
- Replicable architecture unites programmability, assurance, and usability.

Benefits to Scientific Cyberinfrastructure

- Enables secure, policy-compliant collaboration across campuses.
- Reduces friction between researcher agility and IT enforcement.
- Supports diverse workflows in genomics, AI, robotics, quantum, and more.
- Provides reusable open-source toolkits and training materials for adoption.

Risks Versus Potential For Advances

- **Risk**: Unforeseen domain regulations, cross-campus policy conflicts, and stakeholder priorities may limit seamless adoption.
- **Payoff**: Cross-domain researcher and IT engagement builds a replicable, evolving architecture with error resilience, sustainability, and scalability beyond campuses.

Result Dissemination Plans

- Open-source release of WRAP toolkits, policy models, and verification datasets.
- Deployment templates, dashboards, and training materials for replication at peer institutions.
- Publications, workshops, and cross-campus demos to engage both researchers and IT teams.

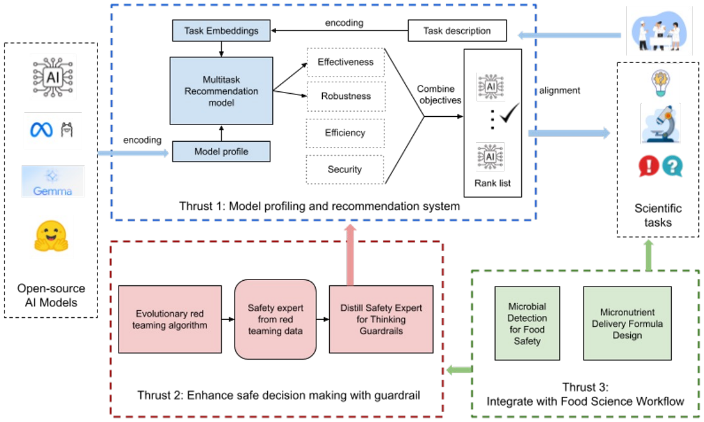
Programmatic Details

- 3 year project started on September, 2025
- Led by Stevens Institute of Technology and with Fordham University

This material is based upon work supported by the National Science Foundation under Grant No. 2530916

FoodGuard: Enabling a Safe and Directive Multi-modal Foundation Model Ecosystem for Food Science Research:UCCS

PI:Zhe Zhao (UC Davis) Co-PIs: Muhao Chen (UC Davis) Xin Liu (UC Davis), Nitin Nitin (UC Davis)


Open-source AI Models
Gemma
encoding
Task Embeddings
encoding
Task description
Multitask Recommendation model
Model profile
Effectiveness
Robustness
Efficiency
Security
Combine objectives
Rank list
alignment
Scientific tasks
Thrust 1: Model profiling and recommendation system
Evolutionary red teaming algorithm
Safety expert from red teaming data
Distill Safety Expert for Thinking Guardrails
Thrust 2: Enhance safe decision making with guardrail
Microbial Detection for Food Safety
Micronutrient Delivery Formula Design
Thrust 3: Integrate with Food Science Workflow

Collaboration and Usability Challenge Addressed

- AI Model Profiling and recommendation for scientific tasks.
- Guardrail for model safe usage in scientific collaboration.
- Integrate with critical food science applications.

Technical Cybersecurity Solution

- multi-task recommendation for AI model adaptation.
- Evolutionary redteaming and slow thinking guardrail.
- Practical and safe use of AI for food science workflow.

Benefits to Scientific Cyberinfrastructure

- Scientists can reliably use AI models for their tasks.
- Scale up use of AI with safety, security and robustness constraints.
- Food science researchers advance their critical research applications.

Risks Versus Potential For Advances

- Uncontrollably development of open source AI models for scientific advancement without robustness and safety check.
- This project will enable safety scaling of AI tools for scientific research.

Result Dissemination Plans

- A recommendation system that can provide AI model recommendation to food science researchers for their tasks.
- AI model ecosystem with model profiles evaluated on effectiveness, safety, robustness and efficiency for scientific collaboration.
- Research publications on the system, as well as datasets generated from the project.
- Food science research advancement with the help of the developed system.

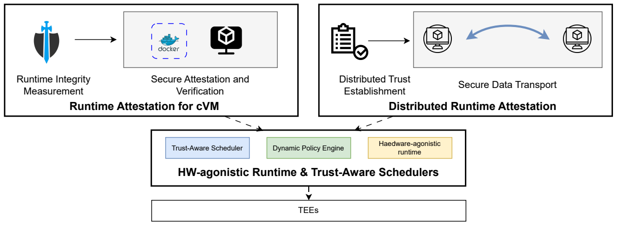
Programmatic Details

- 3 year project started on Sep 2025.
- Led by UC Davis.

This material is based upon work supported by the National Science Foundation under Grant No. 2531126

SafeSci-Tee: Advancing Security in TEE-Enabled Scientific Research Workflows: A Holistic Approach

PI: Wei Zhang (Uconn)


Runtime Integrity Measurement
docker
Secure Attestation and Verification
Runtime Attestation for cVM
Distributed Trust Establishment
Secure Data Transport
Distributed Runtime Attestation
Trust-Aware Scheduler
Dynamic Policy Engine
Haedware-agonistic runtime
HW-agonistic Runtime & Trust-Aware Schedulers
TEEs

What we're doing: Developing comprehensive security for scientific workflows (genomics, drug discovery, materials science) on untrusted HPC infrastructure

Why it matters:

- Scientific data increasingly targeted by attackers
- Multi-institutional collaborations need trust
- Current HPC lacks runtime integrity guarantees

Collaboration and Usability Challenge Addressed

- Workflows span multiple untrusted HPC nodes
- Sensitive genomic/proprietary data at risk
- No end-to-end integrity verification exists

Technical Cybersecurity Solution

- **Runtime attestation** for continuous VM/container integrity
- **Distributed attestation communication and scheduler** across workflow stages
- **Hardware-agnostic runtime** for Intel TDX, AMD SEV, ARM TrustZone

Benefits to Scientific Cyberinfrastructure

- Trustworthy execution of sensitive computations
- Secure multi-institutional data sharing
- Reproducible, verifiable scientific results

Risks Versus Potential For Advances

- Risk: Performance overhead from continuous attestation
- Mitigation: Hardware acceleration, optimized protocols, three-buffer system
- Without this: Scientific collaborations limited by security concerns, vulnerable to attacks that invalidate research
- Payoff: Accelerate scientific discoveries while protecting critical research integrity

Result Dissemination Plans

- Open-source software (Apache 2.0)
- Integration at UConn, LBNL HPC centers
- 3 new cybersecurity courses
- Publications in top security/HPC venues

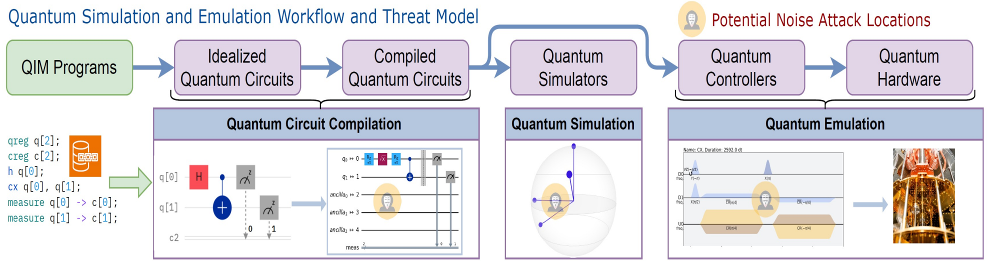
Programmatic Details

- 3-year project started Sep. 2025
- Led by University of Connecticut
- Partners: Lawrence Berkeley National Lab, NVIDIA
- Unfunded collaborations with national facilities

This material is based upon work supported by the National Science Foundation under Grant No. 2530909

QURE: Usable and Attack-Resistant Security Framework for Quantum Emulators

PI: Qian Wang (UC Merced) Co-PIs: Lin Tian (UC Merced), Yuntao Liu (Lehigh University)



```
qreg q[2];
creg c[2];
h q[0];
cx q[0], q[1];
measure q[0] -> c[0];
measure q[1] -> c[1];
```

Benefits to Scientific Cyberinfrastructure

- **Secure cloud-based quantum emulation** – ensures their emulation platforms are secure, reliable, and trusted by scientific users.
- **Trustworthy scientific results** – reduces the risk of adversarial noise or Trojan attacks corrupting quantum emulation outcomes and assures accurate and trusted results.
- **Collaborative security validation** – establishes a collaborative security operations center (SOC) to enhance security validation efficiency for the scientific community.

Risks Versus Potential For Advances

- **Noise attack detection** – distinguishing between natural noise and adversarial perturbations in quantum emulation is technically challenging
- **Cloud dependency** – evolving architectures of commercial quantum platforms may limit access to low-level error data needed for robust validation.
- **Foundational security framework** – establishes first-of-its-kind theory and tools for quantum emulator security, extendable to future quantum computing.

Collaboration and Usability Challenge Addressed

- Distinguishing **benign** noise and **adversarial** perturbations.
- Develop an **end-to-end security validation framework** for quantum emulation, integrating **compile-time** and **runtime** protections.
- Build a **user-friendly sign-off software tool** so that physicists can adopt security checks without needing deep cybersecurity expertise.

Technical Cybersecurity Solution

- **Categorizing and characterizing** compile-time and runtime threats (e.g., Trojan insertion, targeted noise injection)
- Developing **static** and **dynamic** threat detection mechanisms
- Use of **Quantum Ising Models as attack testbeds**, bridging theoretical modeling with real-world cloud quantum platforms.
- Establishing the **collaborative SOC** based on the web-of-trust model

Result Dissemination Plans

- Demonstrated ability to **detect and mitigate adversarial attacks** on quantum emulation with quantifiable improvements in fidelity and reliability.
- Establishment of a **collaborative SOC** with active participation from quantum physicists and cybersecurity experts.
- Publication of **attack datasets, benchmarks, and detection scripts** on publicly accessible repositories (e.g., GitHub, Zenodo).

Programmatic Details

- 3 year project started on January 2026
- Led by (UC Merced) and with (Lehigh University)
- Unfunded collaboration with Kent State and the other collab letter

CICI UCCS

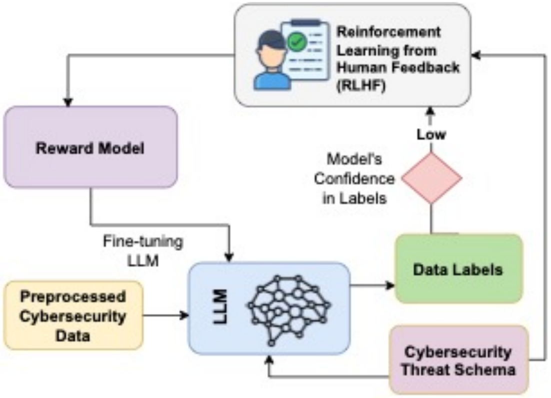
This material is based upon work supported by the National Science Foundation under Grant No. 2530705

Reference Scientific Security Datasets (RSSD)

- Data That Helps Understand Security of Operational Cyberinfrastructure
- Realistic Data for Development and Validation of Cybersecurity Research
- Collect and Disseminate Labeled Data for Scientific Cybersecurity AI Systems

LLMDaL: LLM-Driven Data Labeling for Training Machine Learning Models: RSSD

PI: Kemal Akkaya (Virginia Commonwealth University) Co-PIs: Yanzhao Wu, Julio Ibarra (Florida International University)


Reinforcement Learning from Human Feedback (RLHF)
Reward Model
Low
Model's Confidence in Labels
Fine-tuning LLM
Data Labels
Preprocessed Cybersecurity Data
LLM
Cybersecurity Threat Schema

The project uses Large Language Models (LLMs) to *convert unstructured network data into high-quality, labeled cybersecurity datasets*

Why it matters:

- Addresses dataset scarcity
- Reduces manual labeling costs
- Enhances ML-based threat detection for scientific cyberinfrastructures

Cybersecurity Innovation

- Fine-tunes open-source LLMs with AmLight network data and cybersecurity knowledge
- Builds LLM-powered Labeling Agent to automatically & accurately label network data for AI-based network solutions
- Earlier methods include *manual review, simulations, controlled setups*, and *honeypots*
    - These methods are time-consuming, could not scale with evolving threats, and lacked real-world complexity

Approach For Transitioning the Innovation

- The project employs a new LLM-based methodology to generate reliable labels for cybersecurity data by combining:
    - Retrieval-Augmented Self-Refinement, Expert Verification, and LLM Ensemble

Benefits to Scientific Cyberinfrastructure

- The project produces labeled datasets from AmLight, a real-world scientific network maintained at FIU
    - Relevant to the unique traffic and threat patterns of scientific cyberinfrastructures, which are typically not captured in simulated or generic datasets

Risks Versus Potential For Advances

- Risk:
    - LLM labeling can produce inconsistent or hallucinated results
    - Mitigated through self-reflection mechanisms, verification steps, and robust evaluation processes.
- Payoffs:
    - Scalable, high-quality labeled cybersecurity datasets generated from real-world scientific network
    - An LLM-based automated labeling approach that reduces cost and manual effort

Evaluating and Demonstrating Transition

- Main metric: *Accuracy*
    - Comparing LLM-generated labels against expert annotations on benchmark datasets and real-world incidents
- Other success metrics:
    - Ability to *identify emerging threat* landscapes
    - *Reduction in time and human effort* for data labeling
- The project will make its code, fine-tuned LLMs, and Labeling Agent publicly available under open-source license, along with the labeled datasets,
    - Researchers and organizations can easily replicate, build on, and improve the work

Programmatic Details

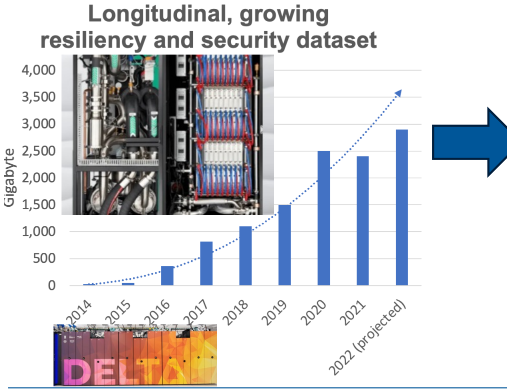
- 3-year project start on January 2026
- Led by Virginia Commonwealth University and Florida International University

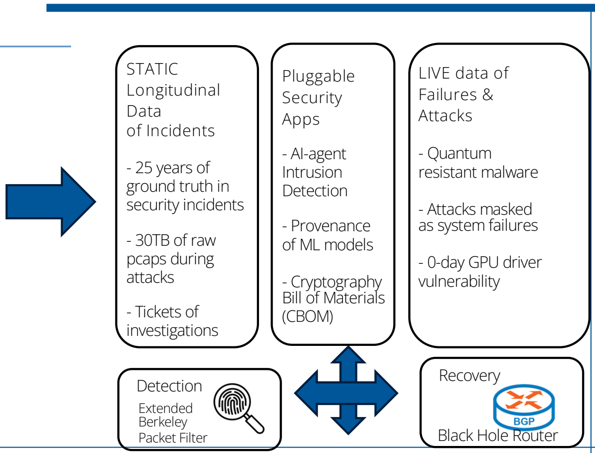
This material is based upon work supported by the National Science Foundation under Grant No. 2530965

NCSA | National Center for Supercomputing Applications

AICyberLake: Live Evaluations of Real-World Security Data Lake from National AI-Cyberinfrastructure

PI: **Phuong Cao**, NCSA/UIUC Co-PI: Ravishankar Iyer, CSL/UIUC


Longitudinal, growing resiliency and security dataset


STATIC Longitudinal Data of Incidents

- 25 years of ground truth in security incidents

- 30TB of raw pcaps during attacks

- Tickets of investigations

Pluggable Security Apps

- AI-agent Intrusion Detection

- Provenance of ML models

- Cryptography Bill of Materials (CBOM)

LIVE data of Failures & Attacks

- Quantum resistant malware

- Attacks masked as system failures

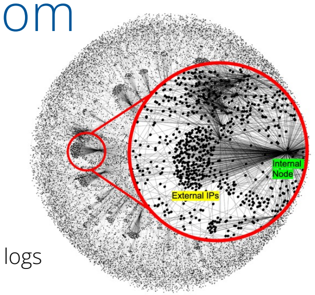
- 0-day GPU driver vulnerability

Detection

Extended Berkeley Packet Filter

Recovery

Black Hole Router



Scientific Security Dataset Needs

- Traditional security data: packet metadata (Zeek), syslogs, audit logs
- Our AICyberLake: heterogenous data sets: **AI models SBOM, GPU utilization/power/energy, job execution traces across supercomputing centers, quantum-resistant cryptographic metadata, federated authentication**, etc.

Technical Approach to Generating Datasets

- New fine-grained monitoring system, and privacy-preserving log streaming APIs.
- Enable a 360-view of U.S. scientific infrastructure AI jobs via Delta, high-fidelity analyses, interactive queries of the anonymized log stream on the fly to vetted research teams.

Benefits to Scientific Cyberinfrastructure

- Research community needs real-world dataset for evaluations
- SSH backdoors, GPU 0-day vulnerabilities, distribution shifts

Risks Versus Potential For Advances

- Adversaries aware of internal detection logics; PII leak possibilities.

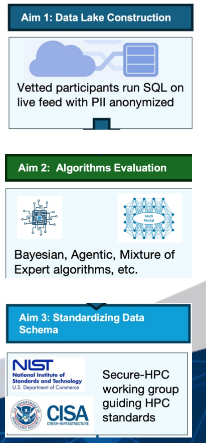
+ A new FAIR-Secure standard guiding Open-Science security research.

+ Robustness evaluation of novel security in real-world with open detection logics.

Publications in USENIX Security, IEEE Quantum Computing Engineering, and more.

Dataset Dissemination Plans

- A new FAIR-Secure standard in collaboration with NIST
- Direct data access with PI's support in understanding domain-data for limited teams
- API self-serving log accesses for broader community with higher degree of anonymity
- Interest form available at https://go.Illinois.edu/aicyberlake ;


Aim 1: Data Lake Construction

Vetted participants run SQL on live feed with PII anonymized

Aim 2: Algorithms Evaluation

Bayesian, Agentic, Mixture of Expert algorithms, etc.

Aim 3: Standardizing Data Schema

Secure-HPC working group guiding HPC standards
Multi Modal
NIST National Institute of Standards and Technology U.S. Department of Commerce
CISA CYBER+INFRASTRUCTURE

Programmatic Details

- 03 year project started on 08/2025
- Led by National Center for Supercomputing Applications (NCSA) at UIUC.
- Teams consuming the logs & collab. with ACCESS, NAIRR, SDSC, Argonne, and ORNL.

This material is based upon work supported by the National Science Foundation under Grant No. 2530738

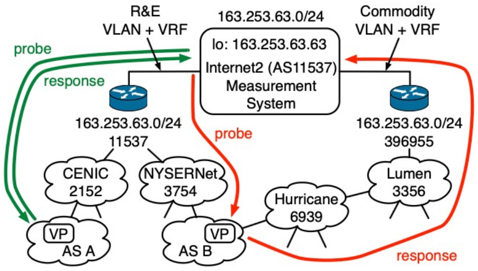
Transition to Cyberinfrastructure Resilience (TCR)

- Transition the Latest Cybersecurity Innovations to Scientific Cyberinfrastructure
- Advances the Cybersecurity Innovation By Providing Operational Validation
  - Promising innovate concept developed but not yet widely transitioned to operations
- Enhances the Scientific Cyberinfrastructure By Deploying Innovative Advances
  - Scientific cyberinfrastructure benefits from very latest cybersecurity tools and techniques

ROOTBEER: Routing Operations Observational Technology Building to Ensure Efficacy of Research

PI: kc claffy (CAIDA/UC San Diego) Co-PIs: Steven Wallace (Internet2), Matthew Luckie (CAIDA/UC San Diego)

Develop a security-focused routing observatory and operational support system to ensure routing policies align with the security and integrity goals of the U.S. science ecosystem.



Cybersecurity Innovation

- Leverage recent innovations in measurement and analytics
- Detect route leaks between R&E and commodity networks

Approach for Transitioning the Innovation

- Develop dashboard to operationalize findings
- Path to sustainability by Internet2

Benefits to Scientific Cyberinfrastructure

- Facilitate creation of robust routing security auditing framework.
- Safeguard integrity and effectiveness of U.S. scientific workflows.

Risks Versus Potential For Advances

- Rigorous independent assessment of network security properties
- Goal: help overcome long-standing barriers to improving routing security, including beyond R&E ecosystem

Evaluating and Demonstrating Transition

- Metrics of success: # of R&E address blocks registered in authenticated routing databases; reduction in routing anomalies; enhanced RPK validation coverage
- Present results at Internet2 Tech Exchange meetings, workshops.

Programmatic Details

- 3-year project started on October 2025
- Led by CAIDA/UC San Diego
- Unfunded collaborator Internet2

This material is based upon work supported by the National Science Foundation under Grant No. 2530871

GRISL: Protecting and Hardening Scientific Use Of Software Libraries With GRISL

PI: Justin Cappos (New York University) Co-PIs: Yuchen Zhang (New York University)

Project Overview

- Harden unsafe scientific libraries (C/C++) without changing scientists' code.
- Use lightweight userspace isolation to prevent crashes and data corruption.
- Deliver pre-hardened libraries with ~1% performance overhead.
- Improve reliability of HPC, AI/ML, and scientific workflows nationwide.

Benefits to Scientific Cyberinfrastructure

- Society needs science to be accurate. Libraries (e.g., ML/AI) often are performance optimized and have flaws.
- Significant reduction in crashes and silent data corruption in scientific workflows.
- Improved reproducibility and reliability of computational research results.

Risks Versus Potential For Advances

- **Risks:** Integrating isolation into complex, performance-critical libraries may expose unforeseen compatibility / performance issues. Adoption requires sufficient POC / momentum.
- **Payoff:** Increased reliability, safety, and reproducibility for HPC, AI/ML, and scientific computing. This framework can be applied to AI, etc. use across domains.

Cybersecurity Innovation

- High performance, backwards-compatible, legacy library isolation for HPC.
- Instead of: failures in one library can crash entire workflows or corrupt results.

Approach For Transitioning the Innovation

- Strong community relationships: build on 15+ years of secure software deployments (TUF, Uptane, in-toto, etc.)
- Open-source release with strong community engagement and maintenance plan.

Evaluating and Demonstrating Transition

- Metrics for success:
  - Substantial improvements to ~10 widely used libraries.
  - Minimal performance overhead (~1% target).
  - Reduction in crashes, data corruption, and debugging time reported by users.
- Integration into major HPC and cloud research platforms (JupyterHub, Open OnDemand, CloudLab, Chameleon). **Open-source implementation and community**. Prebuilt packages available via Conda, pip, and Spack for easy installation.

Programmatic Details

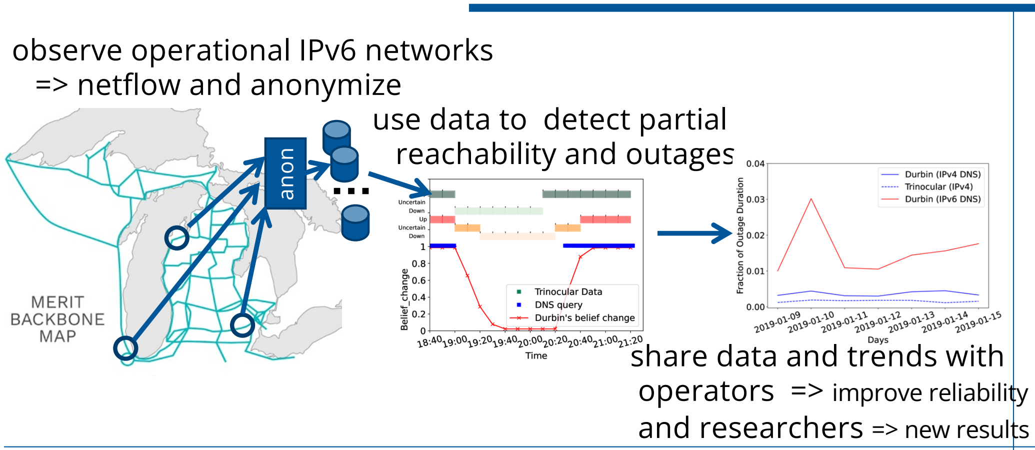
- **3-year project** starting **January 2026.**
- Led by **New York University (NYU).**
- Working with Intel, EngageLively, w/ plans to engage others moving forward.

This material is based upon work supported by the National Science Foundation under Grant No. 2531027

CICI TCR

BIPOD: Building a more Resilient IPv6 with Passive Outage Detection

PI: John Heideman (USC/ISI)   Co-PIs: Pierrette Renée Dagg (MERIT Network)



observe operational IPv6 networks
=> netflow and anonymize

use data to detect partial
reachability and outages

share data and trends with
operators => improve reliability
and researchers => new results

Cybersecurity Innovation

- observing network traffic to generate IPv6 data
- using this data to detect partial reachability and outages
- using those results to help R&E nets deploy and improve IPv6

Approach For Transitioning the Innovation

- sharing new IPv6 datasets with the research community
- sharing detection alerts with network operators

Benefits to Scientific Cyberinfrastructure

- new IPv6 datasets can be used by multiple groups
- info. about outages and reachability improve operational networks
- supports outreach about IPv6 to late adopters

Risks Versus Potential For Advances

- effects of anonymization and sampling vs. data utility

Evaluating and Demonstrating Transiton

- how many datasets are shared with the community?
- do operators find reachability results helpful?
- how many IPv6-late-adoptors come on board?

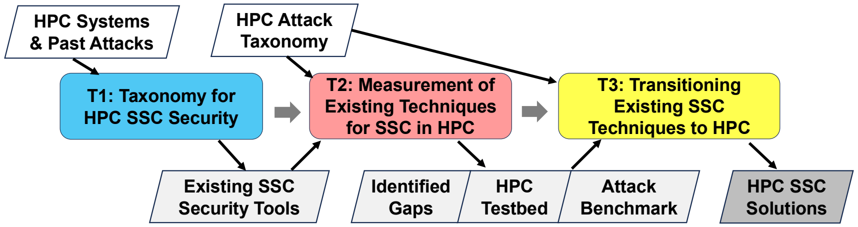
Programmatic Details

- **two-year** project started on 2025-08
- led by U. of Southern California / Information Sciences Institute, working with MERIT Network

This material is based upon work supported by the National Science Foundation under Grant No. 2530698

HPCSafeChain: Software Supply Chain Security in High-Performance Computing: Understanding, Evaluation and Transition

PI: Mu Zhang (University of Utah) Co-PIs: Heng Yin (UC Riverside), Xunchao Hu (DeepBits)



HPC Systems & Past Attacks

HPC Attack Taxonomy

T1: Taxonomy for HPC SSC Security

T2: Measurement of Existing Techniques for SSC in HPC

T3: Transitioning Existing SSC Techniques to HPC

Existing SSC Security Tools

Identified Gaps

HPC Testbed

Attack Benchmark

HPC SSC Solutions

- Scientific computing relies on open-source software vulnerable to software supply chain (**SSC**) **attacks** and **vulnerabilities**
- The **applicability** of SSC security research to HPC is **underexplored**, given HPC's limited C/C++ package management, rapid AI package evolution, and weaker installation controls.

Cybersecurity Innovation

- Novel Attack Taxonomy for HPC SSC
- First HPC testbed/benchmark for evaluating SSC security solutions
- Validating and transitioning SSC security techniques in HPC

Approach For Transitioning the Innovation

- Collaboration with real-world HPC centers: CHPC, HPCC, INL
- A student co-advised by the HPCC team and given cluster admin rights
- Use collaborators' feedback to verify the fidelity of our taxonomy, testbed, attacks and solutions

Benefits to Scientific Cyberinfrastructure

- HPC admins: identifying outdated, vulnerable packages
- HPC users: detecting compromised, malicious artifacts
- HPC (security) researchers: developing a realistic HPC security testbed integrated with a comprehensive attack suite benchmark.

Risks Versus Potential For Advances

- Risk 1: Completeness of attack surface exploration
- Risk 2: Challenges for admins/users in adopting SSC security techniques
- Mitigation: Prioritize high-profile attacks; focus on automation
- Payoff: A practical solution to counter existing threats, which severely compromise the integrity/confidentiality of scientific computing tasks and data.

Evaluating and Demonstrating Transiton

- Expert/community's feedback on our taxonomy, testbed and attack benchmark
- Evaluating applicability of SSC security solutions using our HPC SSC security testbed and attack suite

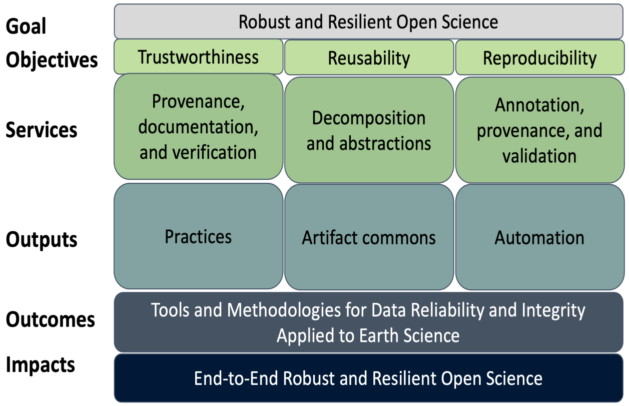
Programmatic Details

- 3-year project started on January 2026
- Led by University of Utah and with UC Riverside and DeepBits
- Unfunded collaborations with CHPC at UofU, HPCC at UC Riverside and Idaho National Laboratory

This material is based upon work supported by the National Science Foundation under Grant No. 2530911

SAFARI: Scientific Analytics, Forensics, and Reproducibility for Workflows in Cyberinfrastructures

PI: Michela Taufer (U. Tennessee, Knoxville) Co-PIs: Ewa Deelman (U. Southern California)


Goal
Robust and Resilient Open Science
Objectives
Trustworthiness
Reusability
Reproducibility
Services
Provenance, documentation, and verification
Decomposition and abstractions
Annotation, provenance, and validation
Outputs
Practices
Artifact commons
Automation
Outcomes
Tools and Methodologies for Data Reliability and Integrity Applied to Earth Science
Impacts
End-to-End Robust and Resilient Open Science

Safari integrates forensic data analytics into the Pegasus Workflow Management System to enhance trustworthiness, reusability, and reproducibility (TR&R) of scientific workflows. By modularizing workflows, embedding provenance and verification, and automating integrity checks, SAFARI ensures results are reliable, secure, and explainable across high-throughput computing (HTC) platforms, with a strong emphasis on Earth science.

Cybersecurity Innovation

- Embed **forensic data analytics** directly into workflow systems (Pegasus); ensure **trustworthiness, reusability, and reproducibility (TR&R)** of scientific workflows; and automate **data provenance, validation, and integrity checks** for end-to-end security.
- Modularize workflows into **containerized components** for explainability and reuse; and establish **forensic analytics as a core CI service**, setting new standards for open science.

Approach For Transitioning the Innovation

- Provide **interoperable building blocks** (artifacts, toolchains, automation, practices).
- Deliver **containerized workflow modules** for portability and reusability.
- Integrate **automated forensic verification frameworks** directly into Pegasus workflows.
- Train students and researchers through **hands-on forensic CI services**.
- Engage with communities (ACCESS, SC, PASC, AGU, AAAS) to build adoption pathways.

Benefits to Scientific Cyberinfrastructure

- Impact Earth science researchers (drought, wildfire, flood prediction), plus broader domains like bioinformatics, materials science, and physics.
- Establish trust in workflows by ensuring data integrity, reproducibility, and reusability—raising the standard for open and explainable science across HPC/HTC platforms.

Risks Versus Potential For Advances

- **Risks:** Integration of forensic analytics may introduce overhead or complexity; adoption requires cultural and technical shifts in workflow communities.
  - **Payoff:** Sets a new benchmark for trustworthy, reproducible scientific workflows; enables cross-domain reuse, reduces costs; and builds workforce expertise in secure and explainable CI.

Evaluating and Demonstrating Transiton

- Num. of workflows with embedded forensic analytics with trustworthiness (provenance integrity, validation accuracy) and reproducibility across HPC/HTC platforms. **(Metrics)**
- Adoption and reuse of components by Earth science and cross-domain users, with training outcomes measured by num. of students and early-career researchers engaged. **(Metrics)**
- Open-source forensic workflow modules via Pegasus WMS GitHub, with documentation, tutorials, and training through ACCESS affinity groups and workshops**. (Community Access)**
- Dissemination via SC, PASC, AGU, AAAS, and community platforms to ensure broad access and reproducibility**. (Community Access)**

Programmatic Details

- **3-year project** started on **October 2021**
- Led by **U. of Tennessee, Knoxville** with **U. of Southern California** (ISI)
- Unfunded collaborations with **Earth science research groups and ACCESS communities**

CICI TCR

This material is based upon work supported by the National Science Foundation under Grant No. 2512345

ForCORE- Fortifying Cyberinfrastructure Operations for Research and Education at B-CU

PI: Grace Kouadjo, co-PI: Kekeli Nuviadenu, Christopher Winlock, Mayra Martinez, and Dawn Eastmond
Bethune-Cookman University, Daytona Beach, Florida 32114

Project Overview

The ForCORE project aims to modernize Bethune-Cookman University's (B-CU) cyberinfrastructure by upgrading its network to 10 Gbps capacity, strengthening campus wide cybersecurity, and enabling resilient, high-performance connectivity for research, teaching, and collaboration. This includes replacing outdated infrastructure, implementing AI-driven firewalls, enhancing real-time threat monitoring, deploying behavior-based trust models, and improving disaster recovery systems.

It matters because B-CU'(s) current infrastructure is a critical bottleneck, hindering research productivity, limiting collaboration, and leaving the institution vulnerable to cyber threats. Enhanced cyberinfrastructure will empower faculty, students, and partners to participate in data-intensive, interdisciplinary research, improve educational outcomes, and position B-CU as a leader in secure, cutting-edge research.

Benefit to Scientific Cyberinfrastructure

Upgrading B-CU's cyberinfrastructure will directly benefit the national research community by:

- Enabling seamless large-scale data transfers, HPC access, and real-time collaboration across scientific disciplines.
- Providing a secure environment for sensitive research data, ensuring compliance with federal and state regulations.
- Serving as a replicable model for cyber-resilient infrastructure in under-resourced institutions, particularly HBCUs.
- If successful, the project will expand the university's ability to contribute to multi-institutional research networks (e.g., via Florida LambdaRail and Internet2), increase research throughput, and broaden participation in national STEM initiatives.

Risks versus Potential for Advances

• Integration challenges with legacy systems during phased upgrades. • Potential for temporary network disruptions during deployment. • Need for continuous adaptation to evolving cybersecurity threats.

Why funding is warranted:

• **Advances:** Secure, scalable infrastructure will support breakthrough research in public health, climate science, behavioral psychology, and more.
• **Impact:** Significant improvement in institutional competitiveness, faculty recruitment, and student preparedness for technology-driven careers.
• **Resilience:** Disaster recovery capabilities will safeguard research continuity during hurricanes and other crises.

Cybersecurity Innovation

What we're trying to do:
•Deploy AI-driven, behavior-based cybersecurity systems that dynamically adjust security levels based on user behavior and real-time threat monitoring.
•Implement adaptive authentication models that reduce unnecessary security friction without compromising data protection.

How it's done today:
•Most academic institutions rely on static, role-based access controls and multifactor authentication applied uniformly to all users. While secure, these measures often disrupt workflows and reduce usability, particularly in high-speed research environments.
•ForCORE's approach moves toward continuous, intelligent security adaptation that balances protection with user experience.

What's new:
•Integration of behavior-based trust models with AI-driven real-time monitoring in an operational academic network—rather than in a lab-only setting.
•Use of the upgraded 10 Gbps infrastructure as both a functional backbone and a live testbed for evaluating adaptive cybersecurity methods.

Why it might work:
•It directly addresses the dual challenge of maintaining strong security while minimizing workflow disruption.
•Leveraging live operational data will allow rapid iteration and refinement of security models, ensuring practical applicability and scalability to other institutions.

Evaluating and Demonstrating Transition

Metrics for success:

- Reduction in mean time to detect (MTTD) and mean time to respond (MTTR) to security incidents.
- Quantifying the number of successful attacks prevented by AI-driven firewalls
- Quantifying the number of automated security threat tasks(e.g. Intervention alerts triaged, quarantines applied, unauthorized devices contained)
- Measurable improvements in network performance (latency reduction, increased throughput, reduced congestion).
- User satisfaction scores regarding security usability and network reliability.
- Growth in number and scope of external research collaborations utilizing B-CU's network.

Community access to results:

- Publication of findings in cybersecurity and cyberinfrastructure journals/conferences.
- Making datasets publicly accessible through NSF supported public facing web hosts for research

Programmatic Details

Three year project starting on January 2026
Led by Bethune-Cookman University, Inc.

This material is based upon work supported by the National **Science Foundation under Grant No.** 2530949

Integrity, Provenance, and Authenticity for Artificial Intelligence Ready Data (IPAAI)

- ***Scientific AI Excels At Ingesting Data But Lacks An Effective Ability To Remove Data***

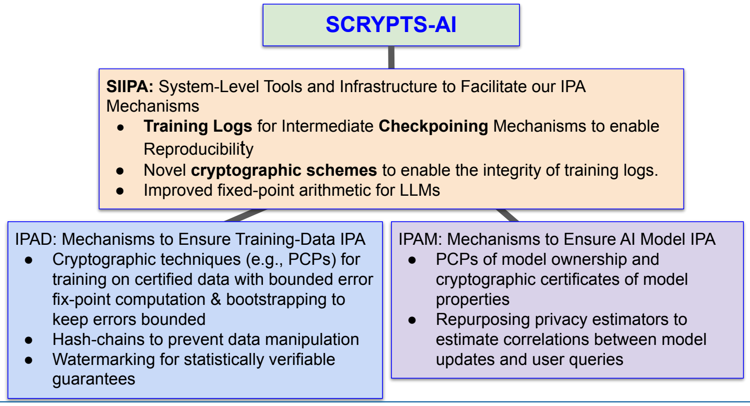
- Trustworthiness of AI Scientific Findings Depend on AI Data
  - Integrity – data is complete and not modified or altered (intentionally or unintentionally)
  - Provenance – data lineage is traceable and uses known collection and processing techniques
  - Authenticity – data comes from a trusted source and is genuine

- Develop Tools, Techniques, and Systems to Data Used In Scientific AI Has The Needed

Integrity, Provenance, and Authenticity

SCRYPTS-AI: Systems and CRYPtographic Tools for Scientific AI:IPAAI

PI: Teodora Baluta (Georgia Tech) Co-PIs: Vassilis Zikas (Georgia Tech), Rafail Ostrovsky (UCLA)



SCRYPTS-AI

SIIPA: System-Level Tools and Infrastructure to Facilitate our IPA Mechanisms
- **Training Logs** for Intermediate **Checkpoining** Mechanisms to enable Reproducibility
- Novel **cryptographic schemes** to enable the integrity of training logs.
- Improved fixed-point arithmetic for LLMs

IPAD: Mechanisms to Ensure Training-Data IPA
- Cryptographic techniques (e.g., PCPs) for training on certified data with bounded error fix-point computation & bootstrapping to keep errors bounded
- Hash-chains to prevent data manipulation
- Watermarking for statistically verifiable guarantees

IPAM: Mechanisms to Ensure AI Model IPA
- PCPs of model ownership and cryptographic certificates of model properties
- Repurposing privacy estimators to estimate correlations between model updates and user queries

Benefits to Scientific Cyberinfrastructure

- Design solutions for **collaboration across disciplines;** develop new insights and approaches for distributed, sensitive, private or noisy datasets using AI.
- **Accountability and transparency** in scientific discovery enables trusted collaboration when aggregating data from multiple parties and when training models on such data.
- **Enhance confidence** by (1) checkpointing and reproducibility of AI training logs, (2) data and model provenance tracking and (3) post-training property certification.

Risks Versus Potential For Advances

- Challenges: **large scale** generative models; **floating-point** arithmetic in reproducibility, cryptography **overhead;** need for **use-friendly toolchains**.
- **High impact** on scientific discovery via securely sharing distributed data; can become a success story to the bigger AI ecosystem and supply chain trust in machine learning.

Need For Integrity, Provenance, and Authenticity (IPA)

- Data privacy and security concerns ⇒ limited access to scientific datasets ⇒ potential for noisy/adversarial manipulation of AI data and models.
- Lack of accountability and transparency impacts the downstream AI uses, where users need confidence over the outputs of the AI model.

Approach For Achieving IPA in AI Datasets

- **SIIPA**: Reproducible AI model training with integrity, cryptographic schemes to enable integrity, improved schemes for fixed-point arithmetic for large models.
- **IPAD**: Develop cryptographic techniques and statistically verifiable watermarks for verifiable data provenance, differential privacy black-box estimation techniques.
- **IPAM**: Enable model ownership and model certification using PCPs, protection of user data privacy using black-box (differential) privacy estimation techniques.

Evaluating and Demonstrating IPA

- Systems-level tools and infrastructure, **end-to-end toolchains** to enhance integrity and provenance of training data and AI artifacts.
- **Mechanisms, implementations and benchmarks** on AI models with scientific datasets that come with guarantees which can be widely adopted by the research community.
- **Source code and generated artifacts** published on GitHub under an open-source license.

Programmatic Details

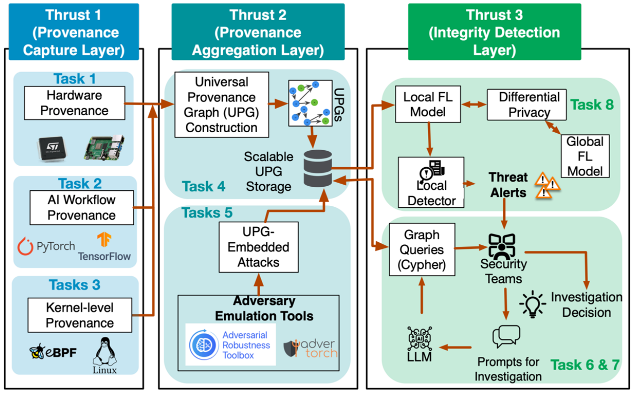
- 3-year project started on January 2026
- Led by (Georgia Tech) and with (UCLA)

This material is based upon work supported by the National Science Foundation under Grant No. 2531010

MLDL: Multi-Layer Data Provenance and Federated Learning for Securing Scientific AI Pipelines

PI: Wajih Ul Hassan (University of Virginia) Co-Pi: Aidong Zhang (University of Virginia)

- Build an end-to-end provenance infrastructure that tracks the full dataset lifecycle, enabling transparency, accountability, and trustworthy, reproducible AI-driven science



Need For Integrity, Provenance, and Authenticity (IPA)

- Multi-source datasets lack end-to-end traceability, reducing trust, reproducibility, and collaboration

Approach For Achieving IPA in AI Datasets

- Capture detailed data provenance across hardware, operating system, and application layers
- Unify this provenance into a scalable provenance graph with privacy-preserving federated anomaly detection

Benefits to Scientific Cyberinfrastructure

- Who cares: domain scientists (genomics, imaging, climate); HPC/data platform operators; security and incident response teams;
- Faster detection of tampering, pipeline drift, and integrity violations
- **Risks Versus Potential For Advances**
- Gaps in hardware visibility, provenance graph related overheads
- A new standard for verifiable AI data lifecycles that improves trust in scientific results across academia, healthcare, and government

Evaluating and Demonstrating IPA

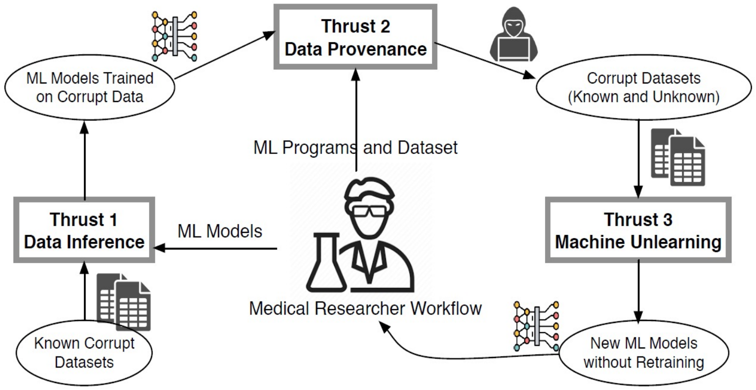
- Metrics of success: end-to-end traceability from hardware to model; low capture overhead; fast load and query; accurate and timely anomaly alerts; privacy preserved across sites
- Community access: open-source code and containers; documented APIs; starter datasets and reproducible scripts; tutorials and workshops
- **Programmatic Details**
- 3 years project, starts on 01/01/2026
- Led by the University of Virginia

CICI IPAAI

This material is based upon work supported by the National Science Foundation under Grant No. 2530655

Dprov: A Data Provenance Framework for Medical Machine Learning Research

PI: Yuan Tian (UCLA) Co-PIs: Aichi Chien (UCLA), Yanyan Zhuang (UCCS)


Thrust 2
Data Provenance
ML Models Trained on Corrupt Data
Corrupt Datasets (Known and Unknown)
ML Programs and Dataset
Thrust 1
Data Inference
ML Models
Thrust 3
Machine Unlearning
Medical Researcher Workflow
Known Corrupt Datasets
New ML Models without Retraining

Need For Integrity, Provenance, and Authenticity (IPA)

- Integrity – Detect public ML models trained on corrupt data.
- Provenance – Standardize efficient, reproducible dataset–model tracking.
- Authenticity – Remove patient or corrupt data from models effectively.

Approach For Achieving IPA in AI Datasets

- Membership inference-based methods for medical AI data provenance.
- Program analysis methods for data provenance in medical workflows.
- Machine unlearning scheme for multimodal clinical data.

Benefits to Scientific Cyberinfrastructure

- Drive new scientific discoveries in medical research.
- Advancing accountable medical AI research.

Risks Versus Potential For Advances

- Challenges to accurately and efficiently trace medical AI data.
- Drive new scientific discoveries in medical research.
- Enhance medical applications and bolster patient confidence.

Evaluating and Demonstrating IPA

- Metrics for success: accuracy of data provenance, efficiency of data provenance, privacy guarantee of machine unlearning
- The tools for medical AI provenance will be made publicly available.
- Plans to host workshops for better outreach to the community.

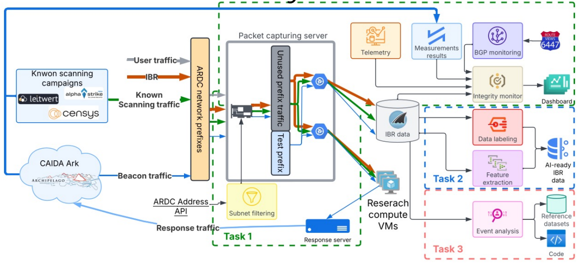
Programmatic Details

- 3 year project starting on November 2025
- Led by University of California Los Angeles and with University of Colorado Colorado Springs

This material is based upon work supported by the National Science Foundation under Grant No. 2531140

CANIS: Curated AI-ready Network telescope datasets for Internet Security

PI: Ka Pui (Ricky) Mok (CAIDA/UC San Diego) Co-PIs: kc claffy (CAIDA/UC San Diego)



CANIS is a suite of modules to improve the UCSD-NT infrastructure for acquisition, processing, & analytics of cybersecurity research workflows.

Need For Integrity, Provenance, and Authenticity (IPA)

- Increasing complexity of UCSD-NT threatens IPA of the data
- AI/ML tools require high-quality labeled datasets for training & evaluation

Approach to Achieve IPA in AI Datasets

- Strengthen telemetry of UCSD-NT, transparency to data consumers
- Create open-source labeling functions using traffic fingerprints

Benefits to Scientific Cyberinfrastructure

- Provide high-quality data that informs cyber threat intelligence to enhance security of scientific cyberinfrastructure
- improve performance (speed) and accuracy of event detection

Risks Versus Potential For Advances

- Not all malicious traffic has known fingerprints for labeling
- Obtaining ground-truth for reference datasets is a challenge

Evaluating and Demonstrating IPA

- Metrics of success: # of data users, publications, and AI models
- Publicly available data on CAIDA website, object storage (accessible from SDSC Expanse), Grafana dashboards
- Host tutorials to demonstrate the use of datasets for AI applications

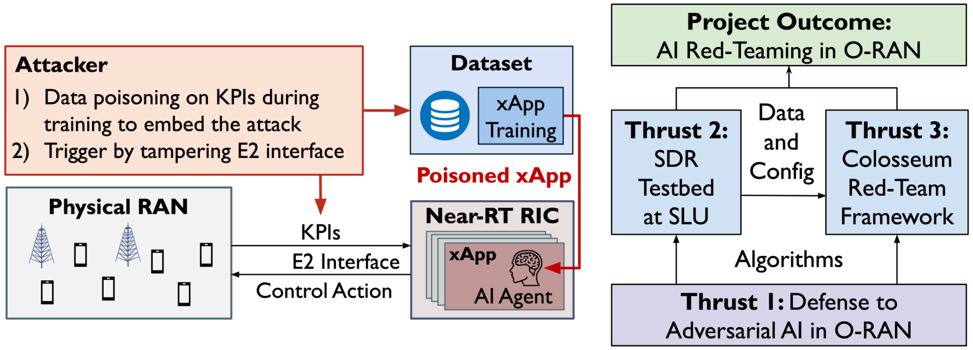
Programmatic Details

- 3-year project started on October 2025
- Led by CAIDA/UC San Diego
- Unfunded collaborators at TU Dresden, U Twente

This material is based upon work supported by the National Science Foundation under Grant No. 2531134

REPAIRT: Securing xApps in Open RANs with Reliable and Principled AI Red-Teaming

PI: Francesco Restuccia (Northeastern University) Co-PI: Flavio Esposito (Saint Louis University)


Attacker
1) Data poisoning on KPIs during training to embed the attack
2) Trigger by tampering E2 interface
Dataset
xApp Training
Poisoned xApp
Physical RAN
KPIs
E2 Interface
Control Action
Near-RT RIC
xApp
AI Agent
Project Outcome:
AI Red-Teaming in O-RAN
Thrust 2: SDR Testbed at SLU
Data and Config
Thrust 3: Colosseum Red-Team Framework
Algorithms
Thrust 1: Defense to Adversarial AI in O-RAN

Need For Integrity, Provenance, and Authenticity (IPA)

- Shared AI-native xApps in O-RAN enables adversarial ML
- No HW/SW CI to vet AI-native xApps before deployment

Approach For Achieving IPA in AI Datasets

- O-RAN specific HW/SW CI for effective and efficient AI red-teaming of O-RAN xApps at scale
- Algorithms for proactive & dynamic defense from adversarial AI

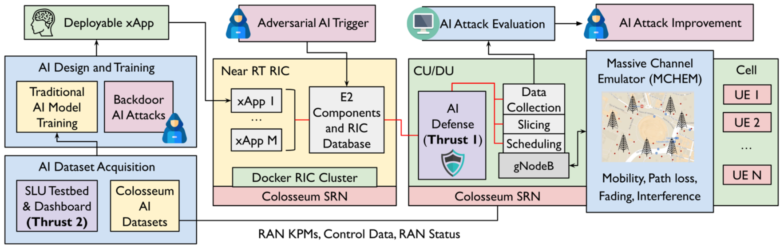
Benefits to Scientific Cyberinfrastructure

- Open-sourcing datasets and code to the O-RAN and AdvML research community
- Train next-gen experts and engage K-12 STEM students

Risks Versus Potential For Advances

- Risk: Integrating sophisticated AdvML in dynamic O-RANs
- Advance: Reproducible AI security, boosting trust in O-RAN, immense economic and societal benefit for U.S.

Evaluating and Demonstrating IPA



Programmatic Details

- 3 year project will start on January 1, 2026
- Led by Northeastern, SLU as subawardee

This material is based upon work supported by the National Science Foundation under Grant No. 2530896

NSF