# Poster: SPHERE: A Community Testbed for Reproducible Cybersecurity and Privacy Research

Jelena Mirkovic*, David Balenson*, Erik Kline*, David Choffnes†, Daniel Dubois†, Luis Garcia‡,
Geoff Lawler*, Joseph Barnes*, Yuri Pradkin*, Christopher Tran*, Lincoln Thurlow*,
Terry Benzel*, and Alba Regalado*
* USC Information Sciences Institute, Email: mirkovic, balenson, kline,
glawler, jdbarnes, yuri, ctran, lincoln, benzel, alba@isi.edu
† Northeastern University, Email: choffnes@ccs.neu.edu, d.dubois@northeastern.edu
‡ University of Utah, Email: la.garcia@utah.edu

***Abstract*—The cybersecurity and privacy research community increasingly depends on realistic, reproducible experimentation to validate results, compare approaches, and build upon prior work. However, much existing research is conducted in isolated or ad-hoc environments that are difficult to reproduce, reuse, or evaluate consistently. To address these challenges, we present the Security and Privacy Heterogeneous Environment for Reproducible Experimentation (SPHERE), a public research infrastructure funded by the National Science Foundation that provides shared, user-configurable hardware, software, and networking resources tailored to cybersecurity and privacy research. SPHERE integrates heterogeneous experimental environments through a secure control substrate, supports multiple user communities through specialized portals, and enables reproducibility through built-in services and Representative Experimentation Environments. Together, these capabilities position SPHERE as a community resource for transparent, scalable, and sustainable cybersecurity and privacy experimentation.**

## I. Introduction

Cybersecurity and privacy threats increasingly impact daily life, critical infrastructure, and global industries, with high-profile attacks targeting governments, infrastructure operators, research institutions, and key economic sectors worldwide. As societies grow more dependent on interconnected digital systems, the scope of what must be protected continues to expand while adversaries rapidly adapt. Despite this urgency, much cybersecurity and privacy research is conducted in isolated or ad-hoc environments using private datasets, bespoke testbeds, or narrowly scoped experimental setups, making it difficult to reproduce results, compare findings, or build systematically upon prior work.

The research community has increasingly recognized the need for shared, rich, and representative research infrastructure that can serve a global audience and support rigorous, reproducible experimentation. Community discussions, including the *Cybersecurity Artifacts Workshop* [1] and the *Cybersecurity Experimentation of the Future 2022 Workshop* [3], have emphasized the limitations of piecemeal experimentation and the importance of common platforms that integrate diverse hardware, realistic network conditions, and support for artifact sharing and reuse.

SPHERE is based upon work supported by the National Science Foundation under award number 2330066. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

## II. Motivation

Advancing cybersecurity and privacy research requires experimentation that is realistic, repeatable, and comparable across studies. Many research questions depend on complex interactions among software, hardware, networks, users, and adversarial behaviors that cannot be fully captured through code and datasets alone. As a result, even when artifacts are shared, reproducing results often requires substantial manual effort, specialized expertise, or access to unavailable infrastructure.

Existing approaches frequently rely on narrowly tailored testbeds or one-off configurations optimized for individual studies, but difficult to generalize or extend. This fragmentation challenges researchers attempting to validate or build upon prior work, educators seeking realistic teaching environments, and artifact evaluation committees tasked with assessing reproducibility under constrained timelines. These challenges motivate the need for shared infrastructure that integrates diverse, user-configurable resources with built-in support for reproducibility, artifact reuse, and safe execution of security experiments.

## III. SPHERE Overview

To address these needs, USC Information Sciences Institute, Northeastern University, and the University of Utah are building the Security and Privacy Heterogeneous Environment for Reproducible Experimentation (SPHERE) with support from the National Science Foundation Mid-Scale Research Infrastructure program. SPHERE is designed as a community resource providing researchers, educators, and reviewers with access to diverse, user-configurable computing, networking, and device environments tailored to cybersecurity and privacy research [2].

SPHERE integrates heterogeneous hardware, software, and network resources through a secure, centrally managed control substrate that supports realistic experimentation while protecting the broader Internet. Resources are organized into

enclaves reflecting common research needs, including general-purpose computing, embedded computing, machine learning, Internet of Things, cyber-physical systems, and programmable networking, and can evolve over time as research priorities change.

SPHERE provides multiple user portals that expose the same underlying infrastructure through interfaces tailored to different expertise levels and workflows, supporting exploratory and mature research, education, human user studies, and artifact evaluation. Built-in services capture experiment configuration, execution, and workflow information to facilitate sharing, reuse, and reproducibility. The platform also supports the deployment of mature research artifacts as Representative Experimentation Environments (REEs), enabling published results to remain accessible, executable, and reusable by the community. In addition to research, SPHERE supports education, workforce training, cybersecurity exercises, and rigorous test and evaluation. These capabilities are exposed through services tailored to distinct user communities.

## IV. SPHERE Architecture and Capabilities

SPHERE is a modular, extensible research infrastructure that integrates heterogeneous computing, device, and networking resources through a secure, centrally managed control plane. This architecture enables realistic experimentation while maintaining strong isolation and safety guarantees appropriate for cybersecurity and privacy research. All resources are accessed through a common substrate, providing consistent experiment lifecycle management across hardware and interfaces.

SPHERE supports flexible network and execution policies that balance realism with protection of the platform and the broader Internet, enabling isolated experiments, controlled external interaction, and safe handling of potentially malicious software. Instrumentation and logging services capture experiment topology, configuration, and execution context, reducing user burden and improving repeatability and transparency. Together, these architectural choices provide a flexible and consistent foundation for reproducible cybersecurity and privacy research.

## V. SPHERE User Communities

SPHERE supports multiple cybersecurity and privacy communities through shared infrastructure that adapts to different goals, expertise levels, and workflows.

**Researchers** use SPHERE to conduct experimental research requiring realistic hardware, configurable networks, and safe execution of security-sensitive workloads, supporting activities such as systems and network security, measurement studies, IoT and CPS security, and machine learning in security contexts.

**Teachers and Students** use educational interfaces and curated environments to integrate hands-on experimentation into courses, assignments, demonstrations, and capture-the-flag exercises, without requiring specialized local infrastructure.

TABLE I
SPHERE Resource Categories and Enabled Research

| Resource Category | Enabled Research |
|---|---|
| General-Purpose Compute | Application, system, and network security experimentation; measurement studies; large-scale experiments; human user studies; trustworthy computing research |
| Machine Learning and GPU Resources | Security with machine learning in the loop; evaluation of ML-based defenses and attacks; reproducibility of machine learning security experiments |
| Internet of Things Devices | IoT security and privacy studies; behavior analysis of consumer and enterprise devices; experimentation with heterogeneous, real-world IoT ecosystems |
| Cyber-Physical and Industrial Control Systems | Critical infrastructure security; industrial control system experimentation; realistic CPS threat modeling and defense evaluation |
| Embedded and Edge-Computing Platforms | Edge and embedded system security; private and trustworthy edge computing; blockchain and federated learning in resource-constrained environments |
| Programmable Networking and SmartNICs | Programmable network security; software-defined networking (SDN) security; in-network measurement, detection, and mitigation mechanisms |

**Paper Authors and Artifact Evaluation Committees** use SPHERE to package and evaluate artifacts in a common execution environment, reducing setup effort and variability, while improving transparency and confidence in experimental results.

**Industry and Government** use SPHERE to test-drive new security solutions that they consider deploying. Small businesses working in security and privacy can use SPHERE to demo their new products to customers or to stress-test products and evaluate them in realistic settings.

## VI. REEs and Getting Started

Representative Experimentation Environments (REEs) host mature research artifacts as long-lived, executable environments, enabling reuse, comparison, education, and extension of published work.

SPHERE supports the creation and long-term hosting of REEs. Current REEs span multiple domains, including security paper artifacts, censorship measurement platforms, reconstructed datasets, and automotive CAN bus simulation environments.

New users can get started with SPHERE through multiple entry points tailored to different audiences. Researchers, educators, students, and reviewers can request access, explore available portals, and deploy experiments or REEs without building custom infrastructure. Documentation and onboarding materials support rapid progression from initial exploration to reproducible experimentation.

## References


[1] D. Balenson, J. Mirkovic, E. Eide, L. Tinnel, T. Benzel, D. Emmerich, and D. Johnson, "Cybersecurity artifacts workshop – report," https://bit.ly/CyberArtifactsWkshp2022, 2022.

[2] J. Mirkovic, B. Kocoloski, and D. Balenson, "Enabling reproducibility through the SPHERE research infrastructure," in *Usenix ;login Magazine*, 2024.

[3] J. Mirkovic, D. Balenson, S. Ravi, L. Garcia, and T. Benzel, "Cybersecurity Experimentation Workshop – 2022 – Report," https://bit.ly/CyberExperWkshp2022, 2022.

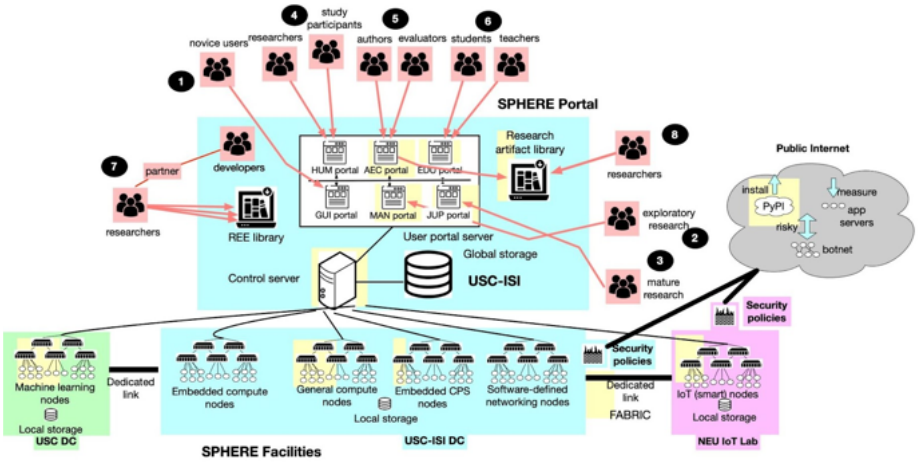
# SPHERE: A Community Testbed for Reproducible Cybersecurity and Privacy Research

Jelena Mirkovic, David Balenson, and Erik Kline (USC-ISI), David Choffnes and Daniel Dubois (Northeastern University), Luis Garcia (U. Utah), Geoff Lawler, Joe Barnes, Yuri Pradkin, Christopher Tran, Lincoln Thurlow, Terry Benzel, and Alba Regalado (USC-ISI)

## At-a-Glance

**SPHERE** is a public research testbed funded by the National Science Foundation and built by USC-ISI, Northeastern University, and the University of Utah
- Provides access to diverse, user-configurable hardware, software, and networking resources through six specialized user portals
- Enables integrated, representative, and reproducible cybersecurity and privacy experimentation that allows researchers to build directly on the work of their peers
- Supports a broad range of activities beyond research, including education, workforce training, cybersecurity exercises, and rigorous test and evaluation

## Architecture and Capabilities




SPHERE Resource Categories and Enabled Research

| Resource Category | Enabled Research |
|---|---|
| General-Purpose Compute | Application, system, and network security experimentation; measurement studies; large-scale experiments; human user studies; trustworthy computing research |
| Machine Learning and GPU Resources | Security with machine learning in the loop; evaluation of ML-based defenses and attacks; reproducibility of machine learning security experiments |
| Internet of Things Devices | IoT security and privacy studies; behavior analysis of consumer and enterprise devices; experimentation with heterogeneous, real-world IoT ecosystems |
| Cyber-Physical and Industrial Control Systems | Critical infrastructure security; industrial control system experimentation; realistic CPS threat modeling and defense evaluation |
| Embedded and Edge-Computing Platforms | Edge and embedded system security; private and trustworthy edge computing; blockchain and federated learning in resource-constrained environments |
| Programmable Networking and SmartNICs | Programmable network security; software-defined networking (SDN) security; in-network measurement, detection, and mitigation mechanisms |

- **Diverse hardware** to support diverse research needs (nearly 90% of today's publications)
- **User portals** supporting exploratory research, novice users, mature users, as well as education, human user studies, and artifact evaluation
- **Flexible security and execution policies**, including full isolation for risky experiments
- **Reproducibility support**, incl. user action logging, artifact packaging and verification
- **Libraries of artifacts** including REEs and others with easy reuse

**Citation:** Jelena Mirkovic, David Balenson, Brian Kocoloski. Enabling Reproducibility through the SPHERE Research Infrastructure. USENIX ;login: Online, USENIX Association. December 16, 2024. https://www.usenix.org/publications/loginonline/enabling-reproducibility-through-sphere-researchinfrastructure

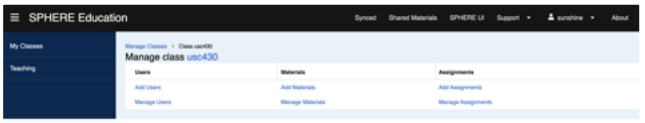
## User Communities

**RESEARCHERS**
- Conduct realistic experiments using diverse, user-configurable hardware and networks
- Safely execute security-sensitive workloads, including malware and adversarial behaviors
- Scale experiments from prototypes to larger deployments
- Package and share complete experimental environments to support validation and follow-on research

**TEACHERS AND STUDENTS**
- Integrate hands-on cybersecurity experimentation into undergraduate and graduate courses
- Use curated environments and intuitive interfaces without requiring local infrastructure
- Deploy labs, assignments, demonstrations, and capture-the-flag exercises at scale
- Align education with modern research practices and real-world systems

**PAPER AUTHORS AND ARTIFACT EVALUATION COMMITTEES**
- Package experimental artifacts, incl. code, data, and workflows, in common environment
- Evaluate artifacts using shared, centrally managed infrastructure
- Reduce setup effort and variability across reviewers
- Improve transparency, repeatability, and confidence in experimental results

**INDUSTRY AND GOVERNMENT**
- Test-drive and evaluate security and privacy solutions prior to deployment
- Demonstrate and stress-test systems in realistic experimental environments
- Compare approaches using shared, reproducible setups
- Collaborate with researchers to validate and mature technologies


## REEs

**Representative Experimentation Environments** hosted as mature, long-lived research artifacts and available as community-accessible infrastructure resources
- Enable reuse, comparison, education, and extension of published work
- Preserve realistic experimental setups beyond the lifetime of individual projects

**Current REEs** include security paper artifacts, CensorLab, reconstructed datasets, and an automotive CAN bus simulation environment

See the **Call for REEs** and **virtual internship information** on the SPHERE website

## Getting Started

**Request access** to SPHERE resources tailored to meet your goals and expertise
- Explore multiple user portals for research, education, and artifact evaluation
- Deploy existing experiments or REEs without building custom infrastructure
- Use documentation and onboarding materials to quickly experiment

**Learn more and get started: https://sphere-project.net**

SPHERE is based upon work supported by the National Science Foundation under grant number 2330066. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.