Producing and Sharing Research Artifacts

Terry Benzel (USC-ISI),
David Balenson & Laura Tinnel (SRI International)

Discussion Topics

- NSF interest in and incentives for research artifacts (Rob Beverly, NSF)
- Issues and challenges
  - Packaging, sharing, and reusing artifacts
- Artifact initiatives at conferences and workshops
  - ACSAC, Usenix Security, CSET, etc.
- Platforms available for sharing artifacts
  - GitHub, Zenodo, etc.
- NSF-funded SEARCCH Hub - community collaboration portal for collecting and sharing experimental artifacts & expertise
  - https://searcch.cyberexperimentation.org (info)
  - https://hub.cyberexperimentation.org (hub)

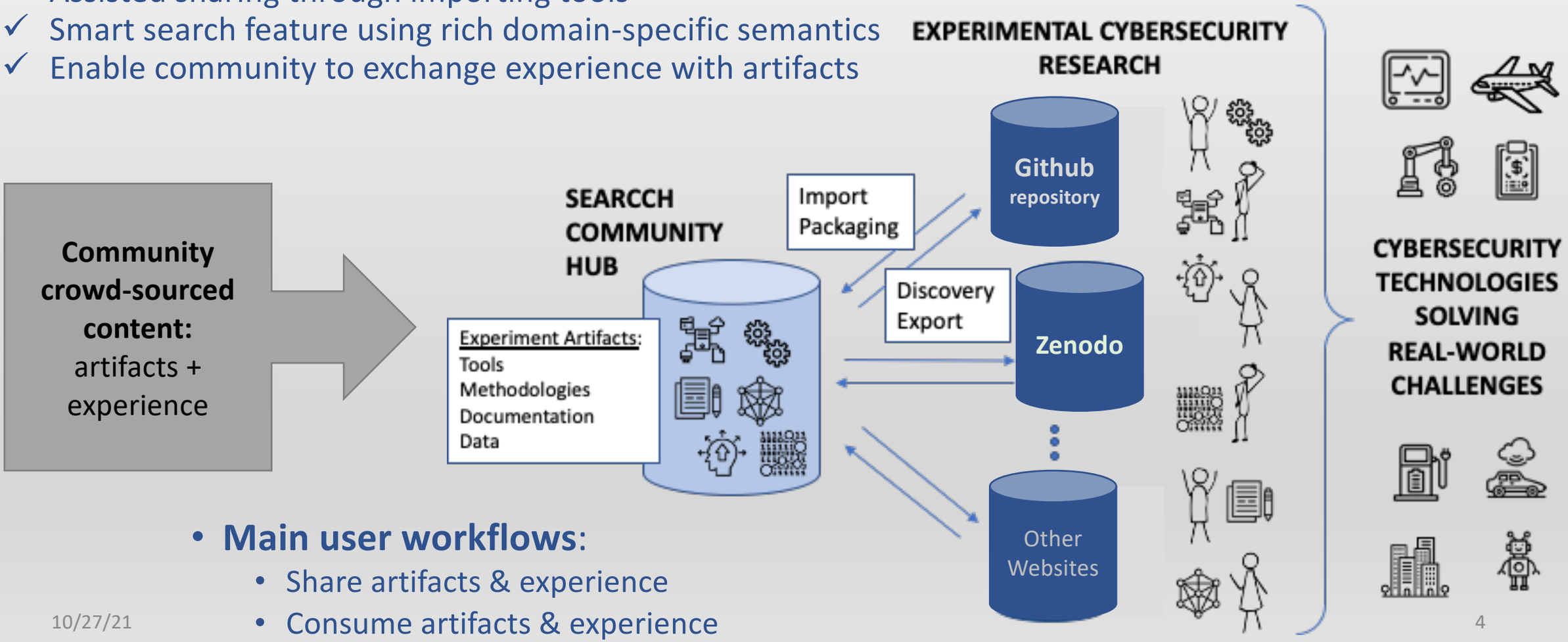
Capturing Experimental Results

Terry Benzel (USC-ISI),
David Balenson & Laura Tinnel (SRI International),
Eric Eide (U. Utah)

Artifacts-sharing Hub Concept of Operations

Collaborative, community-driven platform that lowers barrier to sharing and reuse

- ✓ Assisted sharing through importing tools
- ✓ Smart search feature using rich domain-specific semantics
- ✓ Enable community to exchange experience with artifacts



- **Main user workflows**:
  - Share artifacts & experience
  - Consume artifacts & experience

The Hub Stores Artifact Metadata

The SEARCCH Hub does not store artifacts directly; rather it

- stores a rich metadata representation of artifacts,
- enables researchers to quickly vet artifact relevance to their work and then access actual artifacts in their native location

Artifact Title, Description, and Author(s)
Subject Descriptor / Research Domain
Research Questions and Hypothesis
Methodology
Metrics
Dataset
- Type (several options plus freestyle entry)
- Time of collection
- How/where it was collected

Source Code - any script, research product, traffic generator, simulation, etc.
- Description
- Role in the experiment (e.g., research code, simulator, orchestration code, etc.)
- Language
- Dependencies
- How long it runs
- Any special memory, CPU, hardware, OS requirements

Publication
- Type (e.g., journal article, conference, whitepaper, blog post, technical report, thesis (MS/PhD), book, instructions (installation, use), citation)
- Where published
- Year of publication
- References

Executable - specific binaries used in experiment
- Type
- Purpose
- Supporting Information
- Visuals
- Supplements
- Tutorial

Organization - metadata at the collection level
- Type (e.g., company, academia, government)
- Name
- Group

System Environment
- Testbed
- Resources

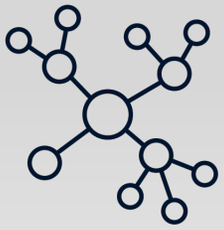
License
- Type
- Restrictions

Domain (aka., Research Applications)
- Current
- Potential

Note: source code details, if captured in the source's README file will show up in the hub as part of the text description.

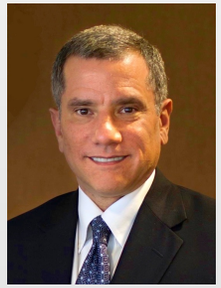
Fundamental Research Design Question

- Determine how to best represent cybersecurity experiment artifacts and the relationships between them and develop an optimized data model that facilitates the efficient artifact discovery
  1) Manually cataloged cybersecurity artifacts to better understand existing artifact features and the breadth of artifacts
  2) Performed automated “mining” of cybersecurity related artifacts from Zenodo as test subjects
  3) Implemented a general artifact "importer tool"
- Once fully operational, we expect most of the hub's catalog to come from user contributions, not automated mining



SEARCCH Collaborative Team PIs


Terry Benzel, Jelena Mirkovic
USC-ISI
Marina Del Rey, CA
benzel@isi.edu,
mirkovic@isi.edu

Laura Tinnel, David Balenson
SRI International
Arlington, VA
laura.tinnel@sri.com,
david.balenson@sri.com

Eric Eide
U. Utah
Salt Lake City, UT
eeide@cs.utah.edu

Tim Yardley
U. Illinois Urbana-Champaign
Urbana, IL
yardley@illinois.edu