

Sharing Expertise and Artifacts for Reuse through Cybersecurity Community Hub (SEARCCH)

Presented to Workshop on Cyber Experimentation and the Science of Security, CESoS '21

Terry Benzel, USC-ISI November 10, 2021

This material is based upon work supported by the National Science Foundation under Grant Numbers 1925773, 1925616, 1925588, 1925564. 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.



SEARCCH Collaborative Team





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

SRI International[®]



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



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





Our Community's Challenges & Needs



- Sharing of repeatable, reproducible, and reusable artifacts in cybersecurity experimentation
 - Can greatly enhance one's ability to build upon the work of others
 - Helps in comparing solutions
- Sharing artifacts can be difficult and time-consuming
- Finding relevant experiments and artifacts can be challenging and time-consuming
- We need:
 - Broad sharing of experiment artifacts
 - Solution that facilitates rapid and open community sharing and reuse

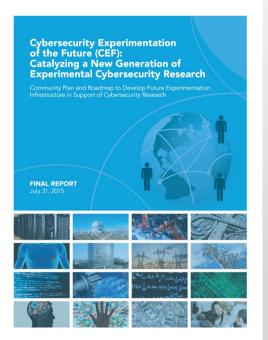


https://www.business2community.com/leadership/8-keys-innovation-mindset-0882548



CEF Study and Community Engagement

- SEARCCH is motivated by the conclusions of the NSF-funded Cybersecurity Experimentation of the Future (CEF)
- Community-based study groups and subsequent community engagement workshops
- Feedback indicating strong interest in community infrastructure that facilitates sharing and reuse of experimental designs, methodologies, tools, and artifacts



https://www.cyberexperimentation.org/



SEARCCH Hub Concept of Operations

SEARCCH

HUB

Experiment Artifacts:

Methodologies Documentation

Tools

Data

COMMUNITY

-

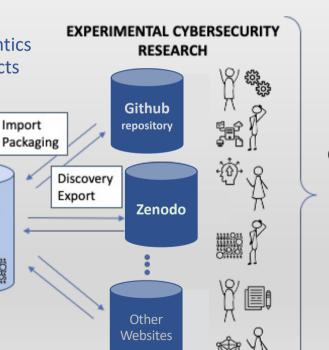
Collaborative, community-driven platform that lowers barrier to sharing and reuse

- 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







CYBERSECURITY TECHNOLOGIES SOLVING REAL-WORLD CHALLENGES





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
 Actual artifacts are then accessed in

their native location

Artifact Title, Description, and Author(s) Subject Descriptor / Research Domain 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.*

Dependencies

Publication

- Type (e.g., journal, conference blog, tech report)
- Where published
- Year of publication
- References

Executable - specific binaries used in experiment Organization - metadata at the collection level

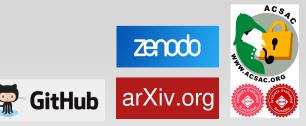
- Type (e.g., company, academia, government)
- Name
- Group
- License
- Type
- Restrictions

Fundamental Research Design Question



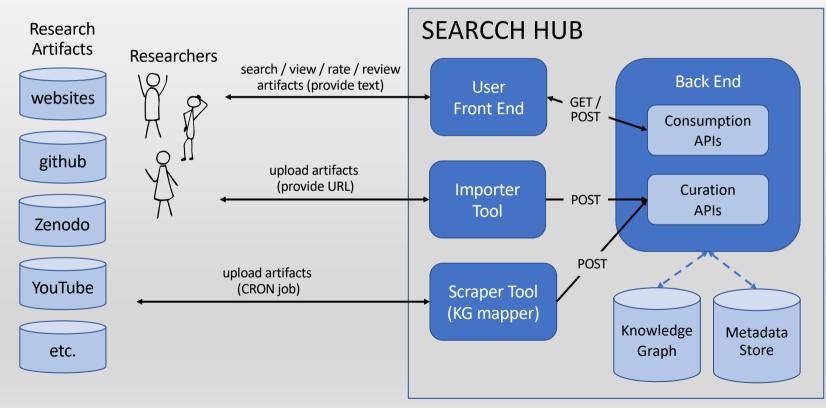
- Determine how to best represent cybersecurity experiment artifacts and the relationships between them and develop an optimized <u>data model</u> that facilitates the efficient artifact discovery
 - Manually cataloged cybersecurity artifacts to better understand existing artifact features and the breadth of artifacts
 - Performed automated "mining" of cybersecurity related artifacts from Zenodo as test subjects
 - 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





Hub High-Level Architecture





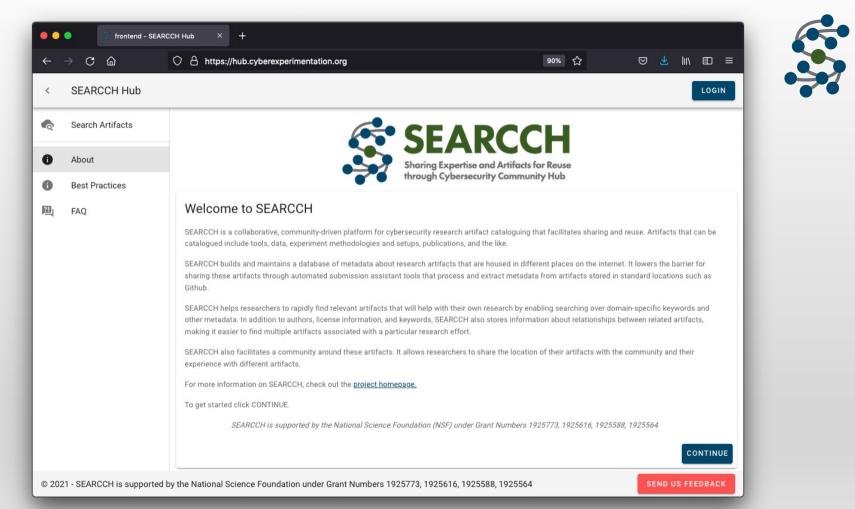


Current Hub Features & Capabilities

- Search Artifacts
- For current artifacts
 - View
 - Favorite
 - Review and rate
 - See other reviews
- Favorite Artifacts
- Submit Artifact
- Manage Account
- Best Practices
- FAQ



SEARCCH Hub Demo (Screenshots)

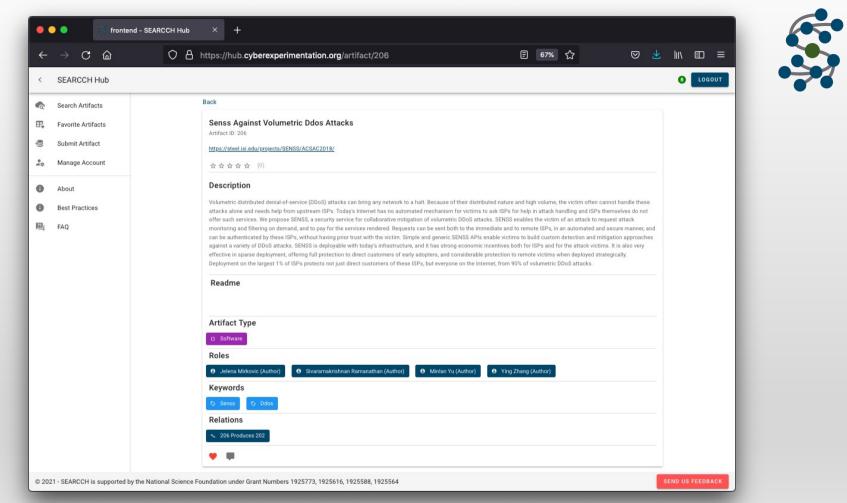


o C a	○ A https://hub.cyberexperimentation.org 90% ☆ ♡ 业 III\ □ =
SEARCCH Hub	LOGIN
Search Artifacts	For more information on SEARCCH, check out the project homepage.
	To get started click CONTINUE.
About	SEARCCH is supported by the National Science Foundation (NSF) under Grant Numbers 1925773, 1925616, 1925588, 1925564
Best Practices	CONTINUE
FAQ	
	Current Features SEARCCH has five major functions. Four may be accessed using the left-hand navigation menu. A summary of each follows. Search Artifacts. Users may perform keyword searches to find artifacts of interest. Favorite artifacts. Users may click on the heart icon on an artifact to add it to a favorites list. Favorited artifacts are displayed on the Favorite Artifacts menu for quick recall. Submit artifact. Users may submit their own artifacts to the SEARCCH catalog. Artifacts published on supported sites may be automatically processed by import assistant tools. Manage Account. Users may add information about themselves such as their research interests and institution affiliation. They may also access the list of their own artifacts, artifacts they have rated, and their favorites. The fifth function is reviewing artifacts. Users may provide reviews for an artifact when viewing it. Presently, reviews consist of a 1 to 5 star rating and a comment. Ratings and reviews are visible to the community and used by others to help them decide whether to invest their time in trying to use a specific artifact. Please provide comments and report bugs using the SEND US FEEDBACK button at the bottom right hand side of the page.

	🛇 🔒 https://hub.cyberexperimentation.org/login?code=a8dd458b23c973552c4f&state=fTX 🛛 🛚 🕲 🏠	?
SEARCCH Hub		>
	Login	
	GITHUB LOGIN	

\rightarrow C @	O A https://hub.cyberexperimentation.org/search		
SEARCCH Hub			
Search Artifacts	CEA	DCCH	
Favorite Artifacts		RCCH	
Submit Artifact	Sharing Experti through Cybers	se and Artifacts for Reuse ecurity Community Hub	
Manage Account	Search		
About	DDoS		×٩
Best Practices	Advanced Select advanced filters for your of	uery	^
FAQ	Artifact types dataset(+1 others)		•
	SEARCH		
		i 7 8 9 10 >	
21 - SEARCCH is support	ed by the National Science Foundation under Grant Numbers 1925773, 1925616, 1925588, 19:	25564	SEND US FEEDBACK

\rightarrow C \textcircled{a}	O A https://hub.cyberexperimentation.org/search		S 7 III ₪ =
SEARCCH Hub			R LOGOUT
Search Artifacts			
Favorite Artifacts		SEARCCH	
Submit Artifact	S S	haring Expertise and Artifacts for Reuse nrough Cybersecurity Community Hub	
Manage Account	Search		
About	DDoS		×۹
Best Practices	Advanced Select advanced	l filters for your query	^
FAQ	Artifact types		
	dataset(+1 others)		•
		4 5 6 7 8 9 10 5	
		4 5 6 7 8 9 10 >	{} Software
	Senss Against Volumetric Ddos Attacks		
	0 reviews 会会会会		
	Volumetric distributed denial-of-service (DDoS) attacks can bring any netw these attacks alone and needs help from upstream ISPs. Today's Internet do not offer such services. We propose SENSS, a security service for colla attack monitoring and filtering on demand, and to pay for the services ren	has no automated mechanism for victims to ask ISPs for help in att borative mitigation of volumetric DDoS attacks. SENSS enables the	ack handling and ISPs themselves victim of an attack to request
	manner, and can be authenticated by these ISPs, without having prior trus mitigation approaches against a variety of DDoS attacks. SFNSS is deploy		
21 - SEARCCH is supported	d by the National Science Foundation under Grant Numbers 1925773, 192561	6, 1925588, 1925564	SEND US FEEDBACK



. . .

S frontend - SEARCCH Hub

○ A https://steel.isi.edu/projects/SENSS/ACSAC2018/

SENSS Against Volumetric DDoS× +

♥ 👱 🛝 🗊 ≡

53

ARTIFACTS FOR ACSAC 2018

This page provides access to code and data used in the paper SENSS Against Volumetric DDoS Attacks, S. Ramanathan, J. Mirkovic, M. Yu and Y. Zhang, Proceedings of ACSAC 2018.

The following files and workflows describe how to reproduce graphs in Figure 5:

1. AS topology:

C 🏠

- Location: topology folder.
- Workflow: download AS relationships information from CAIDA and run convert_caida.pl script. Or to use the existing CAIDA topology from May 1, 2017 run run.sh

2. Simulation code:

- Location: figure5 folder holds all the scripts and data.
- Simulation files: Three main simulation files are:
 - 1. topdeploy.pl, which simulates top deployment strategy, with uniform attackers and clients
 - 2. randomdeploy.pl, which simulates random deployment strategy, with uniform attackers and clients
 - 3. realdeploy.pl, which simulates top deployment strategy with realistic attackers and clients

Each file takes three arguments - the *type* of simulated attack (sig, nosig, cross), the *limit* of how many deployment points to start with (usually 1, unless you are breaking the run into several sub-runs to parallelize it) and an optional argument *alpha*, which controlls the projected percentage of collateral damage and is betwen 0 and 100 (usually 5).

• Considerations for running: The code for each subfigure is in a separate script, e.g., run5a.sh will produce data for the Figure 5a. It takes a long time to produce one figure, because the code is very CPU intensive. On a t2.2xlarge EC2 instance it takes about 5 days to complete run5a.sh. You can reduce this time if you sacrifice accuracy, i.e., if you run fewer trials. The parameter ENOUGH controls the min number of trials for each data point (currently 1,000) and the parameter PRECISION controls the max allowed change between means in consecutive runs for the same datapoint (currently 0.01). Making ENOUGH parameter smaller and PRECISION bigger will reduce the number of runs, but it will potentially produce less stable results.

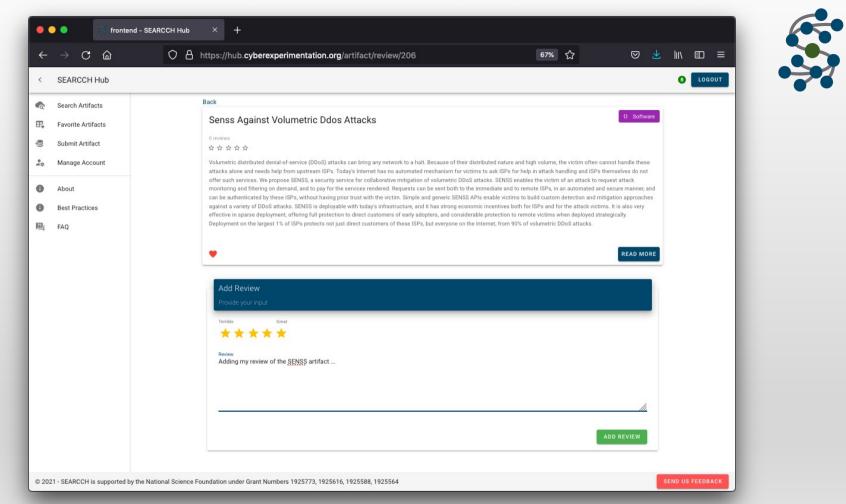
A good way to run the code is to run each script on a separate, high-CPU machine, and then transfer the results to one central place for processing. You can also break runs into sub-runs, which you can run on separate machines. This can be done by using e.g., limit 1 on one machine, and limit 200 on the second machine, with all the other parameters being the same. So:

perl topdeploy.pl sig 1

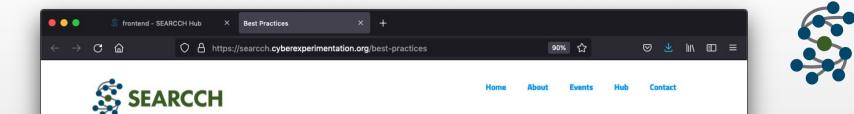
runs all trials on one machine starting with 1 deployment point (and then trying 2, 5, 10, ..., 10000), and

perl topdeploy.pl sig 200

 \sum



SEARCH Hub © court Search Artifacts Favorite Artifacts Submit Artifact Expension of the second o	\rightarrow	Cú	<u>a</u>	A https://hub.cyberexperimentation.org/import	80% 🖒	♥ ⊻ II\ ◻ ≡
Favorite Artifacts Submit Artifact Manage Account About Best Practices FAQ Durk URL Attifacts stored on unsupported sources may be manually imported. Click here to start a manual import. Imported Artifacts	SE	EARCCH	Hub			
Submit Artifact Submit Artifact <tht< td=""><td>Se</td><td>earch Artifa</td><td>acts</td><td>CEA</td><td>DCCL</td><td></td></tht<>	Se	earch Artifa	acts	CEA	DCCL	
Submit Artifact Image Account About Artifact Import Assistant Best Practices Supported artifact locations are: GitHub, ACM digital library, IEEE Xplore, USENIX web site publication, arXiv, Papers With Code, Zenodo, and openly-accessible generic git repositories. FAQ URL Artifacts stored on unsupported sources may be manually imported. Click here to start a manual import. Imported Artifacts	Fav	avorite Artifa	facts			
About Artifact Import Assistant Best Practices Supported artifact locations are: GitHub, ACM digital library, IEEE Xplore, USENIX web site publication, arXiv, Papers With Code, Zenodo, and openly-accessible generic git repositories. FAQ Enter the supported URL for your artifact: URL START IMPORT Artifacts stored on unsupported sources may be manually imported. Click here to start a manual import. Imported Artifacts	Su	ubmit Artifa	act	through Cybersed	and Artifacts for Reuse curity Community Hub	
About Supported artifact locations are: GitHub, ACM digital library, IEEE Xplore, USENIX web site publication, arXiv, Papers With Code, Zenodo, and openly-accessible generic git repositories. Best Practices Enter the supported URL for your artifact: FAQ URL Artifacts stored on unsupported sources may be manually imported. Click here to start a manual import. Imported Artifacts	ф Ma	anage Acco	ount	Artifact Import 🕐		
Best Practices FAQ URL URL URL URL Artifacts stored on unsupported sources may be manually imported. Click here to start a manual import. Imported Artifacts	Ab	pout			ite publication, arXiv, Papers With Code, Zer	odo, and openly-accessible
FAQ URL Artifacts stored on unsupported sources may be manually imported. <u>Click here</u> to start a manual import. Imported Artifacts	Be	est Practice	es		,,	
Artifacts stored on unsupported sources may be manually imported. <u>Click here</u> to start a manual import. Imported Artifacts	E FA	Q		Enter the supported URL for your artifact:		
Imported Artifacts				URL		START IMPORT
				Artifacts stored on unsupported sources may be manually imported. Click here to sta	rt a manual import.	
				A set a set of the		



Best Practices for Submitting Research Artifacts to SEARCCH

Here are some recommendations for preparing and submitting research artifacts to the SEARCCH hub in a way that will more meaningfully aid the cybersecurity research community.

The SEARCCH hub supports the following types of artifacts:

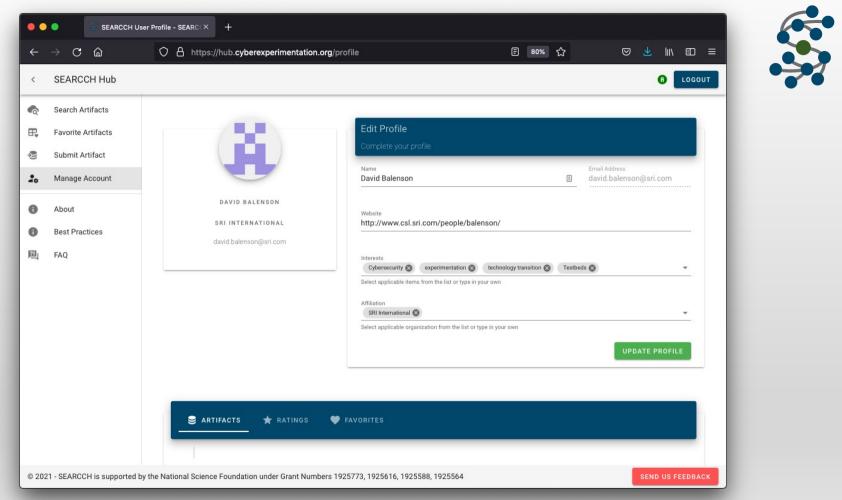
- software
- datasets
- publications
- presentations
- other

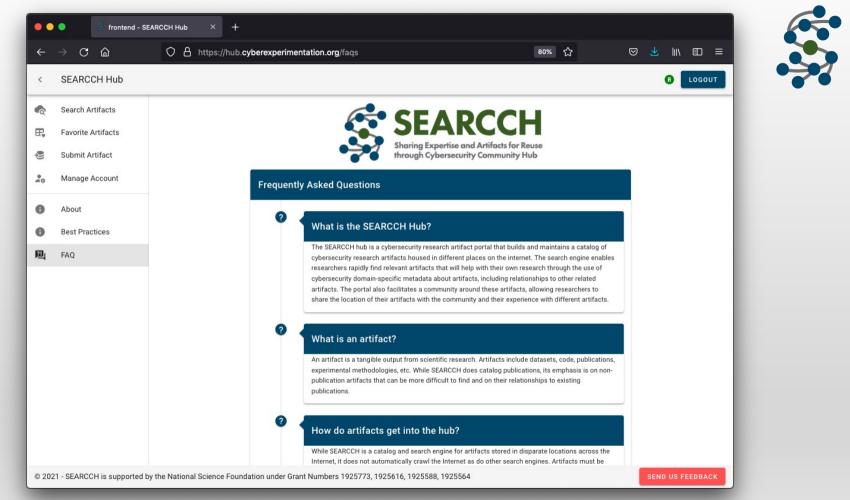
A small amount of curation planning in several key areas will improve the quality of submissions. These are:

- 1. determine the best artifact structure for your body of work (e.g., software vs data),
- 2. determine the relationships between the artifacts and potentially with others that are already in the system,
- 3. document, package, and publish code and/or datasets,
- 4. determine what metadata values should be used to best facilitate searches by other researchers.

Click on a link to learn more about each area.

Once you completed these planning steps, you are ready to import a set of artifacts into the SEARCCH hub.







SEARCCH Importer Tool

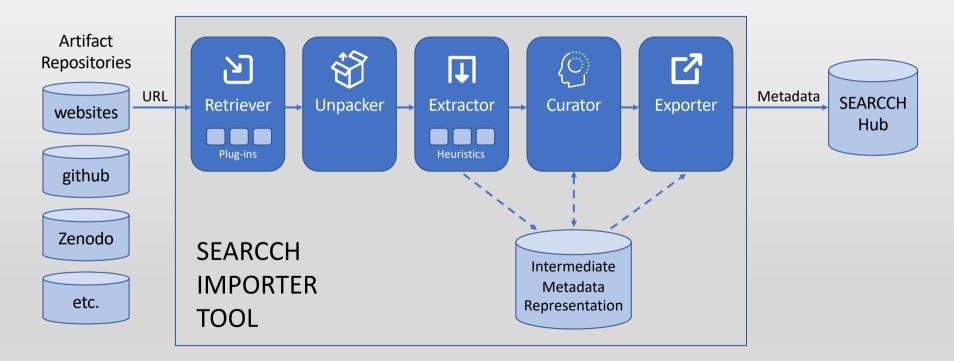
SEARCCH Importer Tool



- Python application that partially automates the task of creating the metadata that describes an artifact
 - Input: publicly accessible location of the artifact to be imported, e.g., a URL or DOI
 - Output: metadata to be stored within the SEARCCH Hub
 - <u>Configuration file</u>: default metadata values, user credentials, etc.
- Allows metadata to be manually edited prior to being exported to the hub
- Also partially automates the maintenance of existing metadata within the hub, when an artifact has evolved or changed location
- Can be used either (1) as a standalone command-line tool, or (2) a backend for a web form or other interface to help hub users import artifacts

Importer High-Level Architecture





Importer Command-line Usage



Usage: searcch-importer [-h] [-d] [-c CONFIG_FILE] {artifact.delete, artifact.export, artifact.import, artifact.list, artifact.publish, artifact.show, db.check, db.upgrade, metadata.add, metadata.delete, tag.add, tag.delete}

Subcommands

artifact.delete artifact.export	Delete an artifact. Export an artifact. Must be published.
artifact.import	Import an artifact from a URL.
artifact.list artifact.publish	List artifacts matching filter parameters. Publish an artifact.
artifact.show	Show artifact details.
db.check db.upgrade	
metadata.add	Add a metadata pair to an unpublished artifact (adds a new curation).
metadata.delete	Deletes a metadata pair from an unpublished artifact (adds a new
tag.add	Add a tag to an unpublished artifact (adds a new curation).
tag.delete Deletes	a tag from an unpublished artifact (adds a new curation).

Optional arguments:

-h, --help Show this help message and exit.
-d, --debug Enable debugging log level.
-c CONFIG FILE, --config-file CONFIG FILE Path to config file.

cu:



Community Building and Next Steps



SEARCCH Project Thrusts and Tasks

Thrust	Task	Description
Technology	Hub	Community collaboration portal for collecting and sharing experimental artifacts
	Artifacts import	Provide structure for shared artifacts as well as tools that facilitate content packaging for sharing
	Artifacts storage	Provide persistence mechanisms for content
	Artifacts discovery and export	Provide tools that facilitate rapid content identification and extraction
	Experiment design support	Provide hub-integrated tools to help researchers design sound experiments using hub artifacts
Data collection	Curate content	Build and use tools to harvest external artifacts to populate hub
Community building	Outreach	Recruit new collaborators from the community and keep participants informed
	Engagement	Actively involve community in requirements, design, and testing of hub

SEARCCH Community Engagement

Actively involve community in requirements, design, and testing of hub

- Poster at NSF SaTC PI Meeting in November 2019
- Talk at FABRIC Virtual Community Workshop in April 2020
- Poster, short talk, and BoF at IEEE S&P in May 2020
- Joint ResearchSOC and SEARCCH Panel at CSET Workshop in August 2020
- BoF at Usenix Security Symposium in August 2020
- BoF at ACSAC in December 2021
- Alpha test in Winter 2021
- Beta test in Summer/Fall 2021

Planning additional briefings and engagement events

Poster and "Artifact Party" at ACSAC 2021













Questions for the Community



- Hub user experience
 - What elements of an Amazon-like user model would you like to see?
 - What additional features would be needed? Which existing features are not needed?
 - What features should be changed and how?
 - What are your top 3 priorities for hub features?
- Content consumption
 - What kind of artifacts do you most need? What is hard to find?
 - Where do you currently go to find artifacts?
 - What was hard about adopting artifacts from others? What would make it easier?
 - What information would you need to decide to use a specific artifact (methodologies, tools, documentation, data)?
- Content contribution
 - Have you shared any experiment artifacts with the community?
 - If so, what was your experience in packaging and uploading? What worked? What was hard?
 - What kinds of tools would make sharing artifacts easier and/or faster?

Next Steps



- Launch the SEARCCH hub in December 2021
 - Implementation of the hub framework and basic set of features
 - Prepopulated content
 - Artifact import tool
- Expand the base implementation
 - Provide semantically rich, knowledge-based searching
 - Curate artifact collections and further seed the hub
 - Provide automated tools to assist with metadata extraction and insertion
 - Enable ratings and discussions around artifacts
- Continue to conduct community outreach and engagement activities
 - Poster and "Artifact Party" at ACSAC 2021
 - Additional engagement events to encourage sharing and reuse of artifacts

SEARCCH and CESoS'21



 SEARCCH supports CESoS'21's goal to lay the foundations for developing rigorous, generalizable approaches to defend against and thwart the growing security threats posed to Cyber-Physical Systems



- SEARCCH enables and supports the transfer and sharing of cybersecurity experimentation expertise and artifacts for large-scale experiments running on NSF scientific infrastructure
- We invite members of the CESoS community to actively participate in planned SEARCCH community engagement activities and to contribute experiment artifacts and expertise to the SEARCCH hub

Contact Us

Follow us on Twitter: @SEARCCH_Hub Visit us on the web: <u>https://searcch.cyberexperimentation.org</u>



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

> USC Viterbi Information Sciences Institute



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

SRI International°



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

THE UNIVERSITY OF UTAH



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



