CERT | Software Engineering Institute
Carnegie Mellon University

Big Data: Learn to Reap the Benefits

Science of Cybersecurity

Could the data your organization collects help you identify potential cyberattacks or show you where to invest to prevent such attacks? Are you unsure about how to gather data to help understand your cybersecurity posture? Do you have massive amounts of data to analyze but don't know where to start?

Leveraging our experience and connections, we can help you achieve your data analysis goals and answer those questions with confidence.

Who We Are

Our Science of Cybersecurity (SOCS) Team is part of the CERT Division of Carnegie Mellon's Software Engineering Institute (SEI), world-recognized leaders in cybersecurity.

We study and solve problems with widespread cybersecurity implications, research security vulnerabilities in software products, contribute to long-term changes in networked systems, and develop cutting-edge information and training to help improve cybersecurity.

Who We Work With

We collaborate with high-level organizations such as US-CERT, Intelligence Advanced Research Projects Activity (IARPA), the NSA, the U.S. Secret Service, and the Department of Defense.

As part of our collaborations with organizations like yours, we have

- developed early-detection techniques for observing and predicting cyberattacks
- analyzed and classified malware attacks
- provided test and evaluation support for research initiatives across a range of problem domains
- developed a wide range of tools to decrease analyst workload, including an automated cyberattack clustering tool and a semi-automated malware classification algorithm
- cleaned, analyzed, and communicated findings from both large and small data stores
- collaborated with teams to design their research studies, ensuring that they get the data they need
- provided advice on how to implement big data tools and methods and incorporate them into existing processes and procedures

Key Capabilities

We develop and help others use machine-learning tools to extract insights from large amounts of data.

We establish mechanisms and processes for handling and managing big data.

We provide broad-based statistical expertise to help others design and plan their research and experiments.

We help identify, generate, acquire, and use synthetic data.

Our Applied Research Areas

We are prepared to assist you in the following applied research areas:

Data Management & Analysis

- developing and applying best practices in managing and maintaining very large data sets
- analyzing and extracting value from large or small data sets
- providing training for analysts on optimal techniques for extracting insights from their data sets
- operationalizing insights gained from analyses through cutting-edge statistical and machine learning techniques

Research Test & Evaluation Support

- providing technical support to agencies performing cybersecurity research
- obtaining data for research studies and working with research teams to ensure program success
- devising statistically sound study-scoring metrics to ensure participants don't "game the system" and win via technicalities

Our Products & Services

CERT products and services enable us to assist you by

- classifying malware attacks from available data using active learning techniques developed in collaboration with Carnegie Mellon University
- helping gather data and configure it so that you can analyze it and build predictors
- analyzing data sets to detect trends, clustering, correlations, and hidden patterns
- helping you analyze data sets to link attacks with their sources
- providing statistical experimental design and power analysis to help researchers extract the data needed to achieve their objectives

big data (noun)

extremely large data sets that may be analyzed computationally to reveal patterns, trends, and associations

About Our Team

We on the SOCS Team can help you with your data analysis challenges. Our team members are experts in statistical and machine learning techniques, experimental design, big data management, and other related areas.

We have extensive experience in

- managing and analyzing both very small and extremely large data sets
- creating realistic synthetic data for research studies where real data is minimal or nonexistent
- creating ongoing automated analytical processes
- detecting—using test and evaluation support—malicious human activity on your computer networks, to yield useful insights

We can undertake work for you in any of these areas or help position your team to conduct its own work.

Get Started Today

The SOCS Team is ready to help your organization. Engage with us to support your data management and analysis efforts. Contact us to discuss how we can help you reap the benefits of big data in your organization.

About

For nearly 30 years, the CERT Division of the SEI at Carnegie Mellon University has been a leader in cybersecurity. Originally focused on incident response, we have expanded into cybersecurity areas such as network situational awareness, malicious code analysis, secure coding, resilience management, insider threats, digital investigations and intelligence, workforce development, DevOps, forensics, software assurance, vulnerability discovery and analysis, and risk management.

Contact Us

Software Engineering Institute
4500 Fifth Avenue, Pittsburgh, PA 15213-2612

Phone: 412.268.5800 | 888.201.4479
Web: www.sei.cmu.edu | www.cert.org
Email: info@sei.cmu.edu