SWAMP
SOFTWARE ASSURANCE MARKETPLACE

WP004

June 24 2014

SWAMP Confidentiality and Privacy for User Data

Von Welch, SWAMP CSO
Indiana University

This document describes the policies and procedures in place in the Software Assurance Marketplace (SWAMP) to provide users with assurance regarding the confidentiality and privacy of their data (software assurance tools, software packages, generated assessment results, etc.).

Citation Information

MLA
Welch, Von. "SWAMP Confidentiality and Privacy for User Data." *Continuous Software Assurance Marketplace*, 24 June 2014. Web. <https://www.swampinabox.org/doc/SWAMP-WP004-Privacy.pdf>.

APA
Welch, V. (2014, June 24). SWAMP Confidentiality and Privacy for User Data [PDF file]. *Continuous Software Assurance Marketplace*. Retrieved from https://www.swampinabox.org/doc/SWAMP-WP004-Privacy.pdf

SWAMP Confidentiality and Privacy for User Data

For public dissemination
Von Welch, SWAMP CSO (vwelch@iu.edu)

Abstract

This document describes the policies and procedures in place in the Software Assurance Marketplace (SWAMP) to provide users with assurance regarding the confidentiality and privacy of their data (software assurance tools, software packages, generated assessment results, etc.).

Background and Terminology

The SWAMP[1] is a national resource for software assurance (SwA) and the research and development of SwA technologies. It is a freely available facility funded by DHS S&T to serve as both a research platform and a core component in the secure software development life-cycle. The vision of the SWAMP [1] is to be a foundation for a more secure and safer software ecosystem.

More concretely, the SWAMP is a facility for continuous software assurance operated at the Morgridge Institute for Research (MIR)[2]. It is operated by MIR, under funding from DHS, with a team including staff from the University of Wisconsin, University of Illinois and Indiana University. For purposes of this document, these respective parties at these institutions (MIR, DHS, UW, UI and IU) are collectively known as the *SWAMP Team*.

When using the SWAMP, users can upload to or generate different types of data:

- *Software packages*: Software to be assessed by the SWAMP.
- *Software Assessment Tools (SwA Tools)*: Tools for performing software assurance assessments.
- *Software Assessment Results (Results)*: Results of assessments of software packages.
- *User annotations on Results (Annotations):* SWAMP users can annotate their generated results to indicate false positives and otherwise indicate the level of seriousness of detected weaknesses.

[1] https://continuousassurance.org/ and https://www.swampinabox.org/doc/SWAMP-WP001-Capabilities.pdf [2]

https://morgridge.org

- *User information*: Data provided by SWAMP users when they register (email address, phone number, etc.).
- *Accounting information:* Data generated by the SWAMP as it is used. This includes accounting data (e.g. number of assessments done, which tools and software packages are used, number of failures encountered).
- *Audit Information*: Data for incident response and debugging (e.g. client IP address, logs about activities, software and tools accessed).

Collectively these three types of data - Software Packages, SwA Tools, and Results - are referred to in this document as *User Data*. The SWAMP Team considers the confidentiality and privacy of User Data a priority, and this document describes the principles, policies and security controls that stem from that priority. This document augments the SWAMP Privacy Policy [2] with additional detail. It stops short of sharing operational security details; for more information, users may contact the SWAMP security team as described in the "For More Information" section at the end of this document.

SWAMP Confidentiality and Privacy Principles

The SWAMP has three main principles related to the confidentiality and privacy of user data.

Principle One: Any software, analysis tools, or results you upload or create in the SWAMP are owned by you. The SWAMP asserts no rights to anything you upload or create in the SWAMP.

Any software or tool you upload to the SWAMP remains your property, and neither the SWAMP Team, DHS or the U.S. Government acquires any rights to it.

Any sharing of that data is yours to control. By default, data is private, and all sharing, as described in the subsequent principle, is controlled by you.

Any user information you provide at registration is also private and will not be shared by the SWAMP.

It is important to the SWAMP and its funders to understand to what degree and, in general terms, how the SWAMP is being used so efforts to improve it can be prioritized. To that end, any accounting information gathered will be aggregated and anonymized before sharing outside of the SWAMP with DHS and the public. To be transparent, the SWAMP will make any accounting reports public so that you can see what is being shared.

Detailed audit information about your use of the SWAMP is only used if we need to help you debug a problem or if we are investigating a security incident in the SWAMP.

Principle Two: The SWAMP provides mechanisms for sharing - you can share your SWAMP data or not as you see fit.

The importance of sharing is key to building the SwA community around the SWAMP. From the SWAMP Vision statement [1]: “By facilitating the sharing of tools, techniques, information, experiences and resources, the SWAMP will: 1) help advance the quality and adoption rate of SwA tools, 2) lower the threshold for using them, and 3) make it easier to interpret and use their output.”

However, that sharing is optional - the SWAMP provides an environment that allows for the degree of privacy or sharing that best fits the needs of its individual users. The fundamental principle of the SWAMP is that User Data belongs to the SWAMP user who uploaded or generated it, and that user is in control of with whom it is shared. Data can be kept private or shared as broadly as meets the needs of each user.

Sharing in the SWAMP is implemented through a coarse-grain *project* mechanism. Each project is owned by a SWAMP user who decides who is a member of the project. The owners of software, tools and results may choose to share which, if any, projects they allow to access their content.

The owners of software, tools and results may choose to make their content publicly accessible. The SWAMP user interface will explicitly warn a user before making something public in an effort to prevent it from being applied accidentally.

Principle Three: Strong trust between the users and the SWAMP is key to the delivery of our vision.

We have a number of security protections to ensure the implementation of our principles.

Privacy Policy

The SWAMP Privacy Policy [2] reflects and codifies the principles described in this document for User Data.

Contractual Agreements with DHS

The SWAMP’s contractual agreement with DHS stipulates the U.S. Government holds no rights to any User Data in the SWAMP.

Identity and Access Management

Secure access to the SWAMP is accomplished through the use of user accounts and an associated credential. The design of the SWAMP identity management system is overseen by staff from the University of Illinois and includes hashed passwords (no passwords are ever stored in the clear) and a strong password policy to prevent reverse engineering or guessing of passwords.

Access control in the SWAMP to User Data is controlled by the user owning the data and is based on projects, allowing controlled sharing to, e.g., groups of users from a user's company or research team. The owners of software, tools and results may choose to share which, if any, projects they allow to access their content.

As described subsequently, only the SWAMP Infrastructure team has privileged access to SWAMP infrastructure. That privileged access by the SWAMP Infrastructure team requires strong (two-factor) authentication to provide an extra layer of protection.

Separation of Duties

To provide additional protections for user data, the SWAMP Team implements separation of duties in several ways so that only limited individuals are privileged to access user data without the permission of the owner, and they are monitored by policy and technical means. Specifics include:

- Only members of the SWAMP Infrastructure team have privileged access to the SWAMP infrastructure needed to access data without permission of the owner. This limited subset of the SWAMP Team are all employees of MIR and undergo MIR's hiring process, including a criminal background check. They must agree to policies described in this document regarding accessing user data.
- The development team uses a separate development system to perform development and testing. Only the infrastructure team can deploy software into the production SWAMP environment that SWAMP users use.
- All accessing of User Data must be done with permission of the owner or under direct order from SWAMP management.
- Von Welch at Indiana University acts as the SWAMP Chief Security Officer (CSO). His role and the role of others in the SWAMP cybersecurity team at IU is to both ensure the development of a cybersecurity program for the SWAMP and continuously assess the performance of the MIR team in implementing it.

Secured Physical Environment

The SWAMP facility is located in the MIR data center and is physically isolated via electronic card key locks, accessible only to the MIR Infrastructure team.

Assessment Isolation

Each assessment is run in its own virtual machine to prevent privacy violations between assessments. Virtual machines are run in their own private isolated network environment to further isolate the assessment runs.

Secure Software Development

The software in the SWAMP is developed by a professional software development team at MIR. That team has received secure software development training and has a dedicated Software

Assurance Analyst whose job is to oversee security of the software development using a framework guided by OpenSAMM[3].

The SWAMP implementation is also continuously assessed by either the SWAMP itself or, for portions written in languages not yet supported by the SWAMP, other analysis tools to detect any defects. The SWAMP cybersecurity team also performs architectural reviews of the SWAMP design.

Secured Data Storage and Backups

All online SWAMP data is stored in the physically-controlled data center in MIR and is only directly accessible by the SWAMP Infrastructure team as described under Separation of Duties. Data backups are encrypted and stored off-site in the University of Wisconsin data center.

Network and Network Traffic Security

The SWAMP facility runs on a dedicated, firewalled network. Multiple network zones are used to isolate User Data on the network, isolating it from the portions users directly access. Each assessment is run in its own isolated network.

Network-based intrusion detection, using Suricata, and automated prevention systems are used to further protect the network following the SWAMP Intrusion Detection and Prevention Policy.

All network traffic to and from the SWAMP is encrypted, and the data within the SWAMP runs over networks dedicated to the SWAMP inside MIR's physically controlled facility. This means any User Data is protected in transit to the SWAMP and within the SWAMP as it is used in (or generated by) assessments.

Web Application Security

The SWAMP isolates user data and the main SWAMP application from the web front end in order to provide extra isolation and protection between these systems. Advanced web browser security functionality (cross-origin resource sharing) is used to provide this separation while maintaining a high level of usability, making the three systems function as one from the perspective of SWAMP users.

Auditing of the SWAMP Facility

The SWAMP undergoes regular internal vulnerability scanning. The cybersecurity team also regularly conducts reviews and audits of the infrastructure. It has also undergone red teaming from an independent third-party.

[3] http://www.opensamm.org/

Exceptions to Our Principles

The SWAMP Team will do their best to stand by their principles regarding the privacy of User Data, but the following exceptions, all of which must be authorized by a SWAMP Principal Investigator, are known to exist:

- We will share data if we are compelled by legal means (e.g., a warrant) and have used all reasonable recourse.
- If we have evidence that User Data violates our Acceptable Use Policy, agreed to by the user when they joined the SWAMP, we may examine it in order to determine its acceptability.

For More Information

For more information about SWAMP security, please contact the SWAMP CSO, Von Welch (vwelch@iu.edu).

To contact the SWAMP cybersecurity team regarding a suspected incident, please email security@continuousassurance.org - you may encrypt your email for privacy using GPG (Key id 0x739202FA, fingerprint 2793 A0A7 4340 7587 FC2A 160F FE83 C695 7392).

References

1. The Vision: Software Assurance Marketplace: A Transformative Force in the Software Ecosystem. https://www.swampinabox.org/doc/SWAMP-VISION-10.28.13.pdf
2. SWAMP Privacy Policy https://www.swampinabox.org/doc/SWAMP-Privacy-Policy.pdf