JOHNS HOPKINS
APPLIED PHYSICS LABORATORY

Cybersecurity Automation and Threat Intelligence Sharing Best Practices

April 2021

ENABLING AUTOMATION IN SECURITY OPERATIONS

Increasing Automation Potential of Processes

Kimberly K. Watson

Many manual processes unknowingly rely on the operator or analyst to infer some fact from available information. This fact is often a key consideration for making an operational decision. While it may be trivial to provide the information via a user interface or set of dashboards, often it is very hard to do this accurately via automation. Accessibility, consistency, and relatability are the main reasons why a manual decision cannot be fully automated without modifications to processes or the resources that store the associated information.

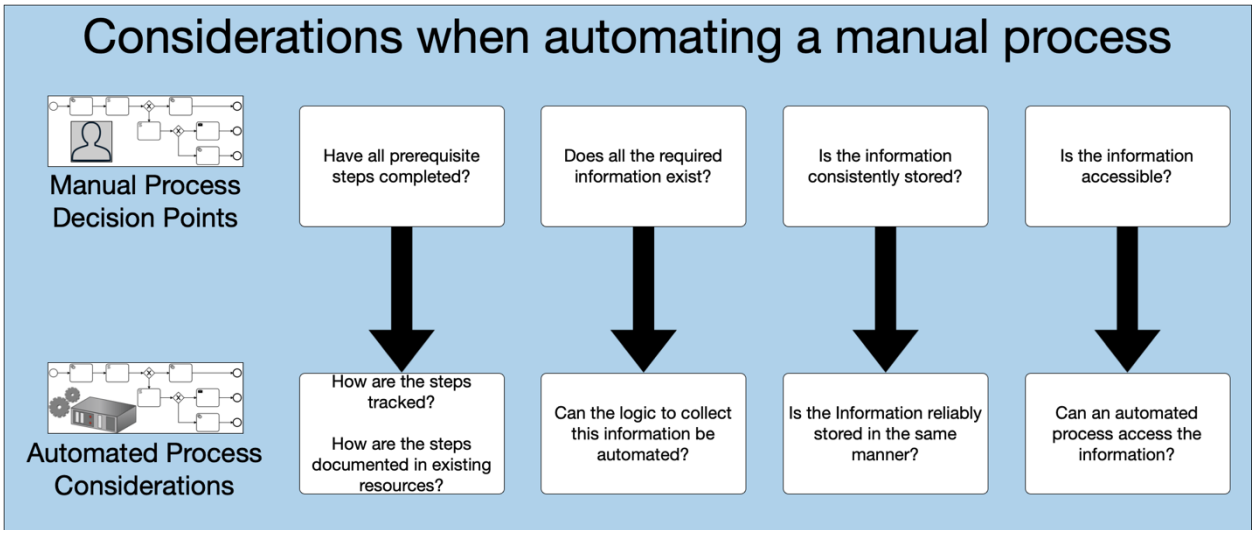

Figure 1 Considerations to Ensure a Process is Automatable

General Considerations

As organizations deploy Security Orchestration, Automation, and Response (SOAR) technologies, they encounter situations where processes are not developed and implemented in a manner that supports automation. Automation requires that there is a clear definition, consistent implementation, and a fixed process for determining states, results, and conditions that drive process steps and decisions.

Often completion of steps or stages in pre-requisite processes are not captured in any manner that automation can query or accurately determine with accessible information. As organizations implement SOAR capabilities to automate parts of a larger manual process, these types of issues need to be addressed to ensure alignment of the automated workflows' timing, resource accesses, and decision making with the manual process.

The other main consideration is that the information required for making a decision is not generated, stored, or accessible in a manner that supports automation. This is particularly true when the necessary information is stored in different resources accessible via different products and services.

When automating specific tasks, it is straightforward to determine what information is needed, where it resides, and how to access it. This is not always true when automating and orchestrating processes that execute a series of tasks, especially when the input for one task is based on the output of a different task. In these situations, it can be difficult to determine what information is available, at what point in the existing processes it is accessible, and with what level of consistency it is present in accessible resources. This results in automated workflows including incorrect or inappropriate logic that may not become obvious until operational.

In some cases, the data available to make a more consistent and appropriate decision exists, but are not always in the same resource, or able to be cross-referenced such that they can be consistently collected in an automated manner. For example, the name of domains associated with Intrusion Detection System (IDS) alerts may exist in a local database for certain types of alerts. For other alerts they may only be available in the packet capture (PCAP) that is stored in a data lake, but the ID associated with that information in the data lake is not contained in the database. Even when included in the database, the domain names may be in a format that has to be parsed or treated as text to find appropriate content. Often it may contain other information that is formatted in a manner similar to a domain.

While API access is used by analysts routinely to do investigations, often the existing service does not support automated access to the same information.[1] Local and custom applications are designed for analyst use instead of usage by automated processes, and commercial products might not be purchased with licensing that would allow automated access.

[1] Watson, K., "Cyber Threat Intelligence (CTI) Sharing Infrastructures", Feb 2021.

All of these different issues result in the conditions associated with making an operational decision being inconsistently defined, analyst dependent, and very hard to determine in a simple manner using automation.

Process and information management techniques and best practices can be used to make operational processes more amenable to automation and orchestration.

Information Management

The following best practices have helped organizations resolve issues with accessibility, consistency, and relatability of information to support operational decisions.

- Automated API access to all information in all associated resources
- Consistent location, application, and mechanism for cross-referencing different identifiers for the same or associated pieces of information. Make sure that the identifiers can be accessed, interpreted, and used in an automated manner to collect and correlate this information.
- Consistent (in format and existence of) historical data and metadata (e.g., timestamps) to support conditional logic, decision making, and measurements often required for automation of analytic processes and metric calculations.

Process Management

Instrumenting processes to make sure that specific fields associated with states or stages are documented in a consistent and accessible fashion is necessary if automation is being used in conjunction with a manual process. It must be possible for automated processes to poll or query existing resources, fork off existing processes, or apply static logic to existing information to determine when to execute different workflows. This is critical when the existence or completion of certain “gating” conditions of the manual processes are supposed to trigger execution or continuation of fully automated processes.

The following best practice has helped organizations resolve issues with making manual processes more capable of including or spawning automated workflows:

- Establish a consistently implemented mechanism to both denote and document execution of clearly defined stages of operational processes, to include any results or variables associated with the stage used by existing processes.
- Ticketing, tracking, and case management products often perform this task, but the existence of this information needs to be maintained in a manner that provides access to historical data, identifiers for cross-referencing with other systems, and structured context/enrichment for use by automated processes.

Conclusion

Enabling automation is a critical component of every organization that wishes to address the speed and scale of modern cyber attack. Without orchestrated automated response, it is often not possible to respond to cyber threat intelligence in a timeframe that enables network defense. However, organizations often find themselves struggling to orchestrate existing manual processes. The best practices identified in this guide will help each organization identify ways to make their processes and associated information management practices more conducive to cybersecurity orchestration.

Acknowledgement

This material is based upon work supported by the U.S. Department of Homeland Security / Cybersecurity and Infrastructure Security Agency under Grant Award Number DHS-19-CISA-128-SLT-001 (State, Local, Tribal, and Territorial Indicators of Compromise Automation Pilot).

Disclaimer

The views and conclusions contained in this document are those of the author and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the U.S. Department of Homeland Security / Cybersecurity and Infrastructure Security Agency.

✉ Kimberly Watson
Kimberly.Watson@jhuapl.edu