

Test and Evaluation for Insider Risk Programs (IRMP)

Bob Ditmore

Software Engineering Institute
Carnegie Mellon University
Pittsburgh, PA 15213

Document Markings

The following markings **MUST** be included in work product when attached to this form and when it is published. For purposes of blind peer review, markings may be temporarily omitted to ensure anonymity of the author(s).

Copyright 2022 Carnegie Mellon University and IEEE.

This material is based upon work funded and supported by the Department of Defense under Contract No. FA8702-15-D-0002 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center.

NO WARRANTY. THIS CARNEGIE MELLON UNIVERSITY AND SOFTWARE ENGINEERING INSTITUTE MATERIAL IS FURNISHED ON AN "AS-IS" BASIS. CARNEGIE MELLON UNIVERSITY MAKES NO WARRANTIES OF ANY KIND, EITHER EXPRESSED OR IMPLIED, AS TO ANY MATTER INCLUDING, BUT NOT LIMITED TO, WARRANTY OF FITNESS FOR PURPOSE OR MERCHANTABILITY, EXCLUSIVITY, OR RESULTS OBTAINED FROM USE OF THE MATERIAL. CARNEGIE MELLON UNIVERSITY DOES NOT MAKE ANY WARRANTY OF ANY KIND WITH RESPECT TO FREEDOM FROM PATENT, TRADEMARK, OR COPYRIGHT INFRINGEMENT.

This material may be reproduced in its entirety, without modification, and freely distributed in written or electronic form without requesting formal permission. Permission is required for any other use. Requests for permission should be directed to the Software Engineering Institute at permission@sei.cmu.edu.

Carnegie Mellon® and CERT® are registered in the U.S. Patent and Trademark Office by Carnegie Mellon University.

DM22-1154

Issues with current approach

- Risky to test on an operational environment
 - Can disrupt organizations mission
 - Overwhelm network resources
 - Spike resource utilization
 - Can impact insider risk analysts
 - Potentially flood analyst with false positives
 - Cause alert fatigue
 - Potentially cause analysts to miss true positives due to excessive 'noise'

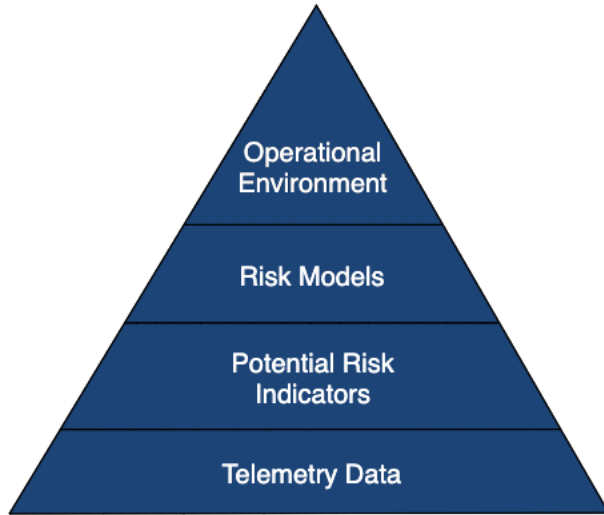


What are we trying to test

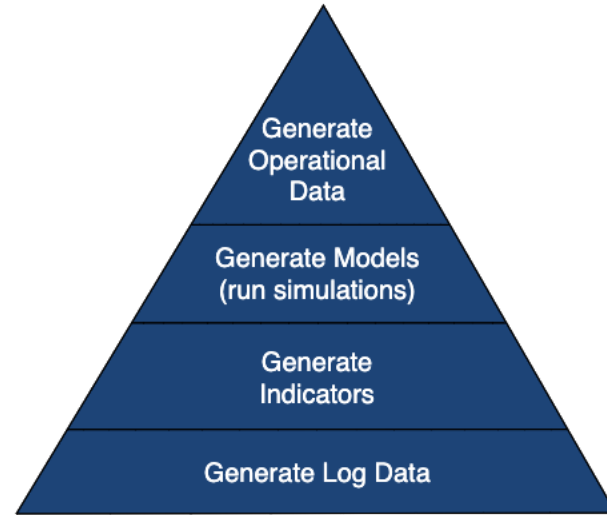
- At the sensor (data source) level and analytic level
 - Will my indicator work?
 - Will it indicate something an analyst can action?
 - Are my analytics effective and efficient?
 - What data/fields does my analytic need?
 - Can it produce timely results?
 - Is there a better sensor for my analytics?
 - Will my AI/ML algorithm find scenario X?
 - Based on scenarios in the IRMP's scope
 - Using the data sets available to the application



Questions can be asked a different levels



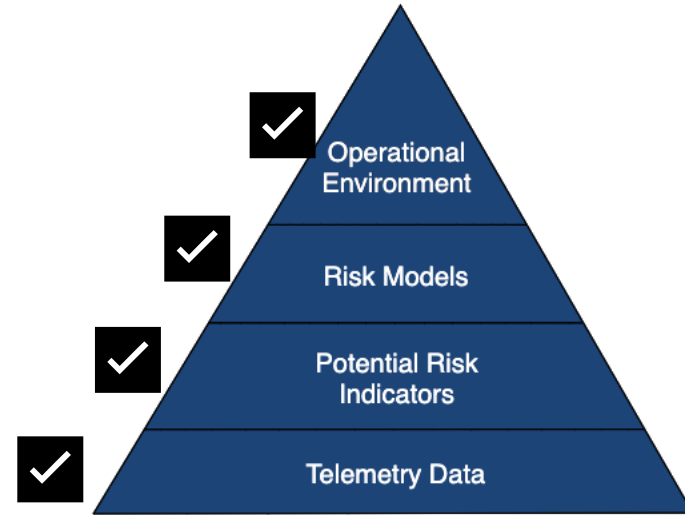
Testing Level



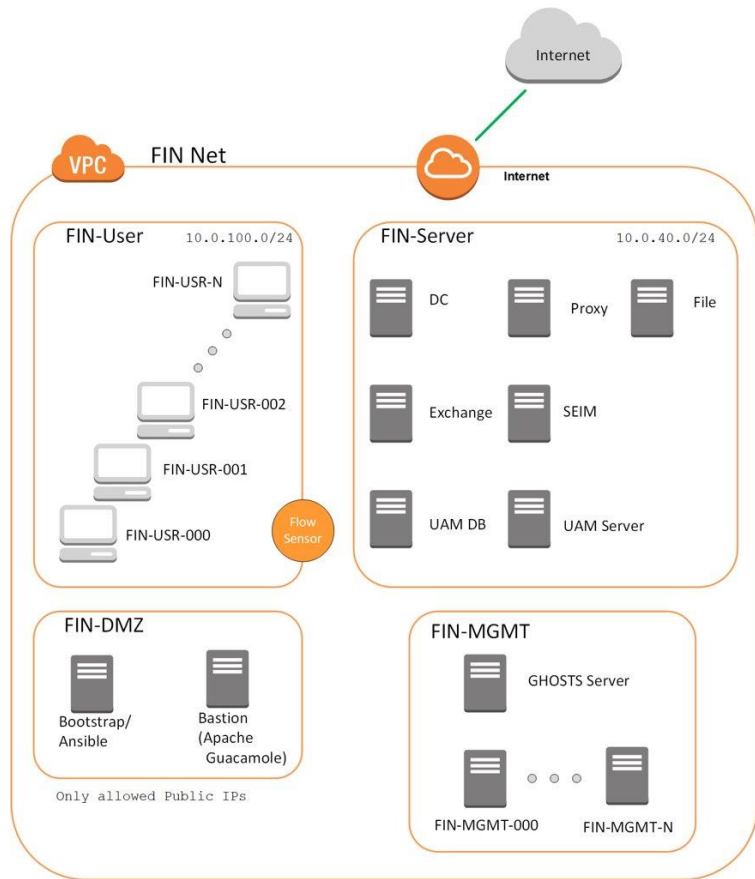
Generation or Collection Level

Our approach – NeedleStackCloud

- Infrastructure as Code (IaC)
 - Rapid setup and teardown of environments
 - Went from days/weeks to minutes
 - Easy to what-if different environments
 - Easy to scale environments
 - Easy to run multiple environments simultaneously
- Cloud Based approach
 - Provides scaling at a reasonable cost
 - Bonus: ability to share environments
- Non-player Character (NPC) Approach
 - Ground Truth



NeedleStackCloud



ANSIBLE

<https://www.ansible.com>



HashiCorp

Terraform

<https://www.terraform.io>



Apache Guacamole™

<https://guacamole.apache.org>



<https://github.com/cmu-sei/GHOSTS>

Contact Information

Bob Ditmore

Team Lead – Insider Risk, CERT Division

Software Engineering Institute

Carnegie Mellon University

rmditmore@sei.cmu.edu

<https://www.sei.cmu.edu/our-work/insider-threat/index.cfm>

