REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
06/13/2023	Technical Report	

4. TITLE AND SUBTITLE

Cyber Analytic Landscape (CAL) Workshop #3 Summary Report

5a. CONTRACT NUMBER

W56KGU-21-F-0008

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

MITRE; JHU-APL

5d. PROJECT NUMBER

5e. TASK NUMBER

W56KGU22F0017

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

The MITRE Corporation
7515 Colshire Drive
McLean, VA 22102

8. PERFORMING ORGANIZATION REPORT NUMBER

PRS-23-1734; MTR230103

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of the Under Secretary of Defense, Research and Engineering

10. SPONSOR/MONITOR'S ACRONYM(S)

OUSD(R&E) ED, DTE&A

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

DISTRIBUTION STATEMENT A. Approved for public release: distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This paper discusses the key highlights and findings from the Cyber Analytic Landscape (CAL) Workshop #3 held on 31 January 2023 at MITRE in McLean, VA. The overarching goal for the CAL effort is to understand the current analytic landscape and advocate for efficient, repeatable, and trustworthy analytics and models. The CAL team identified three strategic objectives (SOs) to achieve this goal and mapped each SO to the workshop sessions that would address it.

15. SUBJECT TERMS

Computer Security; Information Security (General); resilience; Analytic Characterization Framework. ACF

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE		21	Susan Carpenito
					19b. TELEPHONE NUMBER *(Include area code)* 781-271-7646

MITRE

MTR230103
MITRE TECHNICAL REPORT

Cyber Analytic Landscape (CAL) Workshop #3 Summary Report

Sponsor: OUSD(R&E) ED, DTE&A
Dept. No.: N222
Contract No.: W56KGU-21-F-0008
Project No.: 101074.23.401.D32A.P10

McLean, VA

Controlled By: OUSD(R&E) ED, DTE&A

April 2023

Prepared by:
MITRE and JHU-APL

Table of Contents

1 Introduction

This paper discusses the key highlights and findings from the Cyber Analytic Landscape (CAL) Workshop #3 held on 31 January 2023 at MITRE in McLean, VA. This full-day workshop included about 20 people in person, with about 35 virtual attendees throughout the day. The workshop was recorded through Microsoft Teams. This summary report includes information from the CAL Workshop #3 briefs, as well as key discussions and questions from both in-person and online attendees.

2 CAL Background

The overarching goal for the CAL effort is to *understand the current analytic landscape and advocate for efficient, repeatable, and trustworthy analytics and models*. The CAL team identified three strategic objectives (SOs) to achieve this goal and mapped each SO to the workshop sessions that would address it. Figure 1 below captures each SO and the FY23 workshop(s) at which it is planned to be discussed.

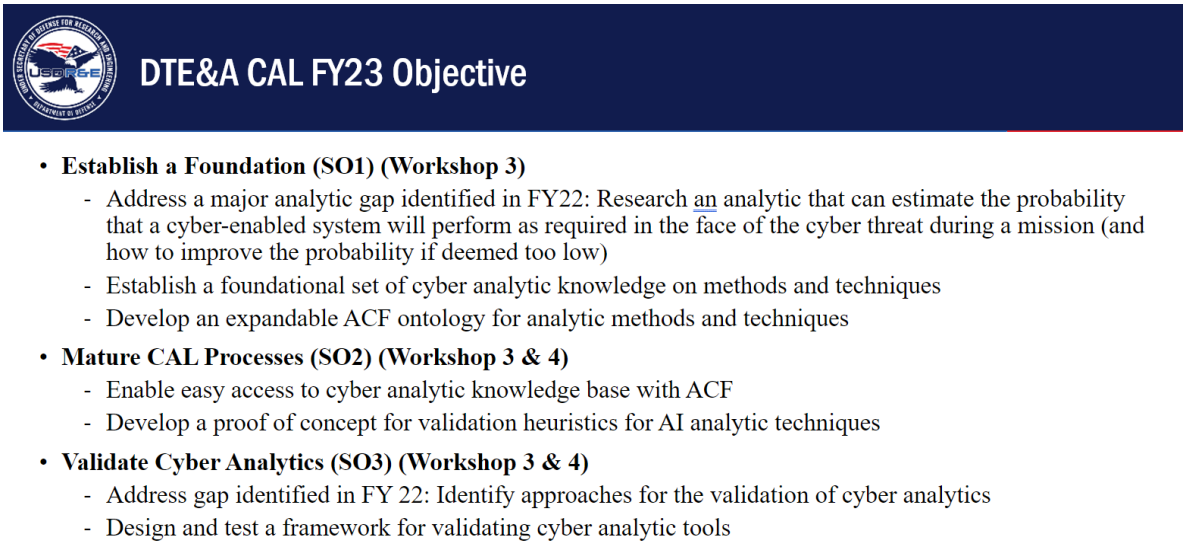

Figure 1: DTE&A Sponsored CAL FY23 Objectives Mapped to Workshops

The CAL background was discussed briefly during Workshop #3 and is discussed further below. The remainder of this report focuses on the specific agenda topics addressed during Workshop #3.

3 CAL Workshop #3

3.1 Opening Remarks and Workshop Outcomes

Ms. Sarah Standard kicked off the workshop with an introduction and brief opening remarks on CAL and its key challenges. She discussed that gaps such as validation and probabilistic estimation approaches are being explored this year based on prior findings during last year's work.

Ms. Standard reiterated her motto: "rules, before tools." She is urging the community as a whole to better define cyber analytic requirements up front to ensure cyber tools are efficiently designed and developed, and can interoperate with other existing tools.

Finally, she emphasized that this effort is incredibly large and invited other partners/stakeholders to assist in solving the challenges presented throughout the cyber analytic landscape.

To kick off the workshop, the CAL team leads outlined the above overarching CAL goal (see CAL Background), and its three SOs to level-set the audience on CAL's strategic vision and ultimate goals. Refer to Table 1 for the full agenda.

Table 1: CAL Workshop #3 Agenda

Topic	Facilitator	Duration (minutes)	Start time
Introduction	Ms. Sarah Standard	10-15	0900
FY22 CAL Summary	Dr. Tom Llanso / Mr. William Barnum	30	0915
CAL Workshop #2 Recap	Dr. Martha McNeil	15	0945
Blue sky brainstorming on probabilistic estimation approaches	Dr. Tom Llanso	45	1000
	BREAK	15	1045
Obtaining feedback on the Resilience Index probabilistic estimation approach	Dr. Martha McNeil	60	1100
	LUNCH	60	1200
Validation approaches for non-cyber threat analytics	Dr. Tom Llanso	45	1300
Update on Analytic Characterization Framework (ACF)	Mr. Peter Kaloroumakis	30	1345
Small Group Sessions Overview	Mr. Peter Kaloroumakis	15	1415
	BREAK	15	1430
Session 1 - Target audiences and their information requirements for ACF	Dr. Michael Smith	20	1445
Session 1 Out brief	Dr. Michael Smith	10	1505
Session 2 - What do we get right and wrong in cyber testing?	Mr. Peter Kaloroumakis	20	1515
Session 2 Out brief	Mr. Peter Kaloroumakis	10	1535
Session 3 - Notional taxonomies of analytical methods and critical feedback	Dr. Michael Smith	20	1545
Session 3 Out brief	Dr. Michael Smith	10	1605
Wrap up/Follow on actions	Ms. Sarah Standard	10	1615

3.1.1 Workshop Outcomes

The overall workshop was a success. The small group sessions were highlighted as a format to continue and expand for future workshops. They enabled more focused discussion.

The sponsor, as well as others in the audience, requested that the CAL team provide all attendees a full list of analytics that were captured in FY23. This would better inform the audience of what analytics exist, and the origin of why and how Pcyber was selected.

One observation during the workshop was a perceived disconnect between the ACF and the overarching CAL effort. Therefore, the sponsor requested a use case scenario from the ACF perspective that would highlight where and how the ACF integrates into the greater CAL strategic vision.

3.2 FY22 CAL Effort Recap

This section captures key notes and discussion points for the following agenda items: *"FY22 CAL Summary"* and *"CAL Workshop #2 Recap."*

FY 22 CAL Summary Highlights

Once the strategic vision and goals of CAL were briefed to the workshop audience, Dr. Llanso and Mr. Barnum then briefed the audience on the results from last year's efforts. They reviewed key statistics from the FY22 Analytic Data Gathering, which include their findings, as well as short-term and long-term recommendations. Highlights from the surrounding discussions included:

- Current FY23 work program is small compared to the larger challenge.
- Question about ongoing access and sustainment of CAL database.
- Lack of tools to evaluate trust is a gap.
- Various tools built for various purposes, not intended to integrate with one another.
- UML model and the related database from FY22 work are available for others and Ms. Standard would like us to share them with the group/those who are interested.
- Community struggles to agree on commonly used terminology.
- Large # of questions: Not always sure what questions to ask or how to use the answer.
- Large # of analytics: Low barrier to entry, therefore everyone has their own approach.
- Hypothesis: Varying competing hypotheses that are/can be at odds with one another.
- Analytic validation: Almost non-existent (used mostly on faith).
    - Analytics include processes, tools, methods, etc.
- Uncertainty: Unable to quantify certainty of results.
- Human footprint: Remains large even with analytic use.
- Is the issue the immaturity of the cost data or how generic it is?
    - Probably a little of both.
- Did you take into consideration what point of the acquisition lifecycle these were at?
- Did you consider trust and how to achieve trust of the system vs. tool?
    - This could be a stakeholder question to add to the catalog.
- Confluence of safety and security—does this affect how you evaluate the system?
- FY23 efforts are looking at estimating the probability the system will perform in the face of adverse cyber events.

CAL Workshop #2 Recap Highlights

Dr. McNeil briefed the workshop audience next on the June 2022 workshop topics and key observations. About 60 personnel attended that spanned various government organizations,

including Service components, FFRDCs, and other DoD organizations. Highlights from the surrounding discussions included:

- Human in loop is a big problem.
    - Dr. McNeil agreed and stated that "humans are really good at some things but not everything. We should use humans for things they are good at."
- Any consideration of cyber pipelines? Non-human consumers as output.
- Tools should consider who is receiving the output—human or cyber agent.
- Human in the loop—help or hinder?
    - Dr. McNeil: We should keep humans in the loop when it makes sense.
- Should analytics be thought of beyond just the humans using them?
    - Dr. McNeil: Yes, things that ingest this data (computer) should be considered.

3.3 Probabilistic Estimation Approaches

This section covers key notes and discussion points for the following two agenda items: *"Blue sky brainstorming on probabilistic estimation approaches"* and *"Obtaining feedback on the Resilience Index (RI) probabilistic estimation approach."*

Blue Sky Brainstorming Highlights

Dr. Llanso briefed the audience on one of the analytic gaps, Pcyber, previously identified during FY22 efforts. Pcyber is defined as: *What is the probability, PCYBER, that a cyber-enabled system will perform acceptably over a mission timeline in spite of adverse cyber events (e.g., malicious attack) that may occur?* The questions he then posed to the group were:

- *How can we estimate Pcyber in a way that can survive rigorous validation?*
- *If we're not happy with the value of Pcyber what can we do to raise the value?*

The floor was then opened to everyone in the audience and online to discuss ideas, concepts, or even potential answers to these questions. Highlights from this discussion include:

- Do we need different processes throughout the lifecycle?
- What are the right units of measurement? E.g., You have a 20% mission loss is easier for programs to understand than you are non-compliant with 45 metrics.
- Numerator vs. denominator for Pcyber.
    - What are these? How are these defined?
- It's not always the case that the adversary will turn the piece of equipment off.
- Requirements—just because you meet the requirements doesn't mean Pcyber is high.
    - Focused on mission critical and exploiting a vulnerability to turn something off.
    - This would show up in availability.
    - Authorizing Official should include cyber piece.

- Who is the consumer of Pcyber? Who is the consumer of the metric? E.g., security engineers, mission owners, program managers.
- USCYBERCOM is using a version of probability of kill (P-K) to evaluate cyber tools.
- Pcyber might be embedded in each term of the larger P-K equation.
- Cyber is an n-dimensional problem, "kill web" instead of "kill chain."
- Confluence of safety and security; can't be disentangled.
- Need to spell out the cyber requirements.
    - Facilitator response: (1) Pcyber could itself be a requirement (Pcyber must be at or above a given level for a given system/mission context) and (2) Pcyber provides a way of summarizing the value of a given proposed cyber architecture. If the architecture has an associated Pcyber value that is acceptable, then we can write requirements associated with that architecture.
- Requirements, components, attacks must be known before one can begin to estimate the probability of success.
- Threshold values for MEFs must include cyber.
- What is in the denominator of Pcyber? Facilitator response: We will cover the construction of the ratio, at least the way that we think of it, in the Resilience Index talk.
    - Denominator should span entire range of possibilities.
- When should we estimate Pcyber? Are there different processes for doing so at different points in the system lifecycle?
- To answer the stakeholder question Q1 posed on the slide, one must first define what "performing acceptably" means for the cyber system. Facilitator response: One way to do this is to define mission-related threshold values of metrics tied to the system's critical functions; performance then considers whether those minimum thresholds were met across the mission timeline.
- William (Data) Bryant suggested restating Q1 in terms of "mission loss" (likelihood x consequence).
    - Consider each system's contribution to the mission.
- The system under study could be a fleet of systems; e.g., many instances of the same type of aircraft.
- The goal of a cyberattack is not always to disable the system.
- Need to account for cyber in multiple factors of P-K.
- Could Pcyber be compared between systems? What are use cases and actions after knowing Pcyber?
- Cyber performance requirements are often stripped away under cost/schedule pressure.
- Factor effects that do not deny service into Pcyber.

Obtaining Feedback on RI Highlights

During this discussion, Dr. McNeil described one approach for estimating Pcyber called the Resilience Index (RI). RI is defined as: *The probability that a **cyber-enabled system** or **systems** will perform acceptably during a mission timeline despite **adverse cyber events (ACEs)***. This ongoing work originated from a conference paper submitted by both Dr. Llanso and Dr. McNeil, titled *"Towards an Organizationally Relevant Qualification of Cyber Resilience."* Dr. McNeil continued to brief the RI metric to the workshop audience. Notes and highlights from this discussion include:

- Where do we test RI? Response: Green diamonds shown on PowerPoint slide.
- System might change over time. Add new ACE type for changes to systems that are not well-communicated or well-understood.
- This is a promising approach—has quantitative rigor. Maybe a cyber requirement could be based on a threat profile from ACEs.
- Praised for using standard terminology—CNSS (criticality, MEFs, etc.). Data needed for probability distributions.
- Positive feedback on Pcyber, and the presentation.
- Structure makes sense. We have decades of data on hurricanes, reliability data. We don't have useful data on malicious attack. We need this data, but it can be hard to get (can't wait 30 years).
- Do you worry about overestimating RI if ACEs don't apply to a deployment environment?
    - Facilitator response: Might overestimate RI if include them all. Might create mitigations that work for test environment but not for deployed environment.
- Instead of using Gaussian + Monte Carlo, can you use generative functions from data gathered from an operational environment?
    - Facilitator response: Presumably to populate distributions used for simulation random variables.
- Any considerations given to using Monte Carlo analysis to generate synthetic data?
    - Facilitator response: As of right now, no; however, it has been noted.
- Obtain malicious attack data to drive simulation from an organization with a large network (e.g., university). Have an intern call up universities in the area.
- What we lack today are robust mission models.
- Mission space is the challenge. Some data in commercial world. DoD mission as defined by who (PM, actual ops).
- Big fan of BG/RG, but how can we recast the RI simulator to accept data from other risk tools?
    - Facilitator response: RI simulator can use risk data from any source as long as it comes in the right format and is semantically aligned. Probability Estimate: Single number.

- When do we test our estimate is valid?
  - Facilitator response: In the green diamonds would be a good opportunity.
- Should we add deployed configuration?
  - Within operator error?
  - As operated—differences in SW version that is deployed in various AoRs.
- Exposure and criticality = risk.
  - Exposure: If it has a greater # of relevant threats that are unmitigated.
  - Criticality: Asset has higher criticality if a greater number of highly weighted mission essential functions rely on the asset and greater number of data types are processed there.
- Standard deviation of Pcyber did not settle down to an acceptable range until about 1000 trials. This is not so much about edge cases, rather improving sample size.
- Maybe a cyber requirement could be based upon a threat profile that would be defined by existing threats (identified through ACDs?) against a current system?
- There was some discussion that MITRE has previously worked on data probability distributions and to investigate the MITRE contacts and prior products and research that may exist.
- Mitigations that are successful in tested environment might not be successful in operational environment.
- Dr. McNeil to look at edge cases for Monte Carlo sims/find the literature doc she came across last year during the survey … distribute appropriate levels of ACEs from the learned operational environments … how do we get something more representative of operational environments and not overestimate?
- Data inputs for validation of distribution.
  - We don't have useful truth data on malicious attack. That is the independent variable that we are most interested in to get truth data; we need to build the system, document if it was attacked/capture that data, and then 30 years later we'd know the truth data. How do we deal with the lack of truth data for the most important set of distributions?
- Were any considerations given for using Monte Carlo sims to generate synthetic data to represent real world data?
  - Facilitator response: No considerations yet, but we've noted this comment/question.
- Mission modeling. We lack robust mission models (BMD, air defense modeling, etc.). What does the degradation mean for the overall mission itself?
- Data in mission space associated with commercial world. Mission space as defined by PM vs. mission space defined by operations working the mission … they don't have ownership over all MRT-C. Can we use Monte Carlo sims in those areas we are more constrained in or don't have visibility into?

- How can we approach the challenge area without being specific to the analytic … how do we get away from specifying our favorite tools and instead focus on the rules/processes that a tool should be able to achieve?

3.4 Validation Approaches for Non-cyber Threat Analytics

Dr. Llanso posed two key questions, listed below, during this discussion for the audience to provide any thoughts and feedback on.

- *Question (1): What analytic approaches have been used in non-cyber threat domains to validate analytics?*
- *Question (2): What can we learn from them that might be transferable to the cyber domain?*

Key notes from Question 1 discussions include:

- Clarity needed for analytics:
  - Need to frame the problem in a way that has a tractable solution.
  - What cyber analytic is being used to identify an attack or method of attack? We don't "care" about how/when/where the attack occurred … We care about how we assess these cyber analytics? Analytic can be tool/process/method/etc.
  - Need a real definition of what we mean by "analytics" and type of analytic and the problem we're looking to solve.
  - Context of analytic (classes of analytics might need to exist).
  - Can the analytics be applied across the different platforms?
    - Can the ACF capture this?
  - What is the analytic tool or process?
    - If I have an MRI, I am interested in the analytics and related results used to validate the MRI; how do I know/trust these analytics and results?
    - M&S VV&A—bring in Dr. Fuzzy Wells (and the MBSE VV&A team) to understand MBSE processes and Melissa Wong (M&S DODM).
    - Various levels of accreditation—would this accreditation change from development to operational (above comment: accreditation is system-specific).
    - Really need a new accreditation, but could the V&V data be reused?
    - There's a difference between the physics (which is easier to model) and the cyber aspect.
  - MBCRA is an analytic.
    - Evaluate MBCRA—what were the results of the MBCRA vs. the results of the test? Does this show us how well the MBCRA was executed?
    - **IDA's MBCRA efficacy study currently underway**, looking at 30 different programs to understand how well the programs are executing MBCRAs.
    - Aircraft Survivability—use an MBCRA to build a probabilistic kill chain.

    - Different analytics for different platforms/objectives? An MBCRA is an analytic that spans multiple platforms. Do we really have a good understanding of the threat landscape? Known knowns, unknowns, etc.
    - Mr. Bryant suggested a way to validate MBCRA: Make predictions, test, compare results to predictions.

- Confidence:
    - What did you do to ensure confidence was valid?
    - Used previous models/warhead firing experience—compared real-world ops to model results. Were able to maintain integrity within the models.
    - Confidence due to previously relevant collected data.
- Joint Mission Effectiveness Model (JMEM):
    - Consider using JMEM as it should provide additional benefits.
    - JMEM for cyber—Cyber operational lethality and effectiveness tool (COLE); (validated model two weeks ago).
        - Is COLE already doing this? JHU-APL did look at COLE; Ms. Standard asked: Did we ask the right questions?
        - JMEM is validating COLE—but we are not sure who within JMEM?
        - Recommendation to look at JMEM COLE (Cyber Operational Lethality and Effectiveness).
- Verification, Validation, and Accreditation (VV&A):
    - What are we using M&S for? We need to VV&A against the intent of M&S purpose.
    - Validation is done best with real data. We need to collect real data.
    - This data collection is where the cyber space is unique compared to other "types" of things such as reliability, maintainability, etc.; cyber data is not physics based …
    - Accreditation—specific to each system scenario. To accredit an analytic that was "approved" for aircraft X, is it then "good enough" for CVN? How much more data might need to be collected?
    - What is the validation approach that should be followed when assessing analytics? (I.e., if this one analytic is used to describe/evaluate aircraft X; is that analytic still relevant to my CVN?)
- Survivability:
    - Intent is to not only be focused on survivability; we can expand aperture beyond threat analytics/non–cyber threat analytics.
    - Aircraft combat survivability—all physics based. Starts with data previously collected. Blowing up aircrafts is expensive—focus remains on M&S, and then you do spot checks to ensure the model is producing accurate results. Adding cyber in—you use same process with a few mods—we don't have history of data. Therefore, you need to do MBCRAs—identify the kill chains (what can occur) … then execute

component level testing (bench test; spot check via bench tests) ... then you do cyber testing at full system level—DOTE is working this now.

- Threats:
  - Some discussion about "we know all the threats that exist"—pushback from room on that.
  - Cyber world has access to all their threats—permutations of connecting threats together have significant implications and impacts to systems. DoD threats won't expose themselves until the very last moment when attack happens.
  - Can access many threats/exploits/vulnerabilities by going to the dark web.
  - Testers will have a hard time testing against adversaries who use zero day vulnerabilities.

Key notes from Question 2 discussions include:

- Dr. Llanso read Jon Bierce's quote on validation in the context of undersea warfare. This led to a discussion of the need for data/evidence to support a validation argument.
  - *"In USW model validation is always best done with real data. For physics/engineering data, we spend a lot of time/energy with the government to make platforms available for at-sea testing and will often create the instrumentation to collect the data. Those data sets are foundational to design and analysis. No other way to do validation. With human systems, e.g., sonar operators, there is a pretty rigorous way to set up testing to get data that will inform models. For tactical analysis, we have simulations built from performance data but operational platforms are in such high demand that it is difficult to do sea-test validation. We do it when we can though."*
- Commercial/Other Aircrafts:
  - Commercial airlines—can we use/tap into that? Can we assume it is a black-box system?
    - What about NASA?
    - Someone mentioned considering the NASA IV&V process.
  - Aircraft survivability models—based on history, theory, testing; models are spot-checked.
  - In the context of cyber risk in aircraft survivability context, we spot-checked models.
  - Mr. Bryant: (1) Build everything in M&S, 2) test components in isolation, 3) test the whole thing, (4) perform spot checks of model through a live fire simulation of the model at platform level.
  - Tom Andress mentioned that all analytics in surface-to-air missile testing are validated; models from prior generations of missiles can be checked once there is actual data.
- Verification, Validation, and Accreditation:
  - Models and simulations are usually "V-V+A"-ed for an intended purpose.

        - Verification and validation stages may be reusable.
        - Accreditation is suitability for a given purpose (is it good enough for my application).
        - Validation methods: empirical, testing.
    - If the analytic gets evaluated and validated for system x, does that validation transfer to system y or does it need to be revalidated?
        - Can the ACF capture this kind of validation?
- Acquisition considers these factors: safe, suitable, survivable, effective.
- Cyber is special—not physics-based. Facilitator response: Agreed. However, discussed analogy of biological systems; we do clinical trials to tease out effects of new drugs for humans (people are complex with no two exactly alike), so there is precedent in studying highly complex systems and making progress.
- Instrument systems with sensors to collect data.
- Broadly applicable analytic vs. specific analytic validation approach differs.

3.5 Update on Analytic Characterization Framework (ACF)

During this discussion, Mr. Kaloroumakis and Dr. Smith gave an update to the audience on the ACF. Key highlights and discussions include:

- Need a good definition of analytic. Cambridge Dictionary: "A process in which a computer examines information using mathematical methods in order to find useful patterns."
- Analytics we've looked at in the past don't involve AI or ML.
- Facilitator made distinction between engineering analytics and operational analytics.
- Multiple audience members had a strong desire to use ACF and a clear idea on how they would apply it; however, some groups were still confused on how to implement it as well as the value add from ACF.
    - There was a strong consensus that documenting the ACF use case scenarios would help mitigate this concern.
- One specific viewpoint was that the ACF is far too advanced right now in the analytic world to be of use currently.
- ACF must work for the analytics CAL looked at in FY22.
- Understand the variables that analytics would relate/integrate/etc. and what these variables mean.
- To evaluate an analytic, you need to understand (1) how analytic is implemented but also (2) what domain problem you are trying to solve verification and validation (V&V).
    - ACF captures what the analytic is and what it evaluates (V&V).
- Accreditation may eventually take in the (V&V) information.

- Need to cross-walk the previous analytics from last year (about 74) with ACF.
- Significantly different conceptualizations of cyber analytics.
    - Analytics of risk (overall, systemic, etc.) versus cyber analytics (e.g., cyber analytics for detection).
- For ACF, we need it to do what CAL needs, not let it become a beast of its own.
    - Can we apply the ACF to the taxonomy?
    - Not trying to do the VV&A for this (probably won't work) but want to get the taxonomy.

3.6 Small Group Sessions

This section covers key notes and discussion points from the three small group sessions. The small group sessions were facilitated by Mr. Peter E Kaloroumakis and Dr. Michael Smith. Each small group was given about 20 minutes to discuss two key questions within their group and capture any key highlights or discussions. After each session, one representative from each group out-briefed the key points to the entire workshop audience.

The in-person attendees were split up into about two separate groups with about five to six participants in each group. The virtual participants were separated into breakout rooms using the Microsoft Teams "Break out rooms" capability. There were three breakout rooms set up for the first session, and then two online breakout rooms for the second and third sessions. The decrease in rooms was due to reduced online attendance later in the day. Additionally, participants in each group changed with each topic, enabling everyone to work with a wide range of participants.

3.6.1 Session #1 Discussion Points

The first small group session covered the following two questions:

1. What are the key analytic development lifecycle roles for cybersecurity analytics?
2. What are the key information requirements for these roles and the analytic development lifecycle?

Discussion Notes

Question (1): What are the key analytic development lifecycle roles for cybersecurity analytics?

- Key roles recommended were systems engineer, developer, architect, software architect/engineer/developer, ISSE, ISSM, cyber engineer, threat analyst, criticality analyst, supply chain risk assessor, developmental and operational testers, data scientist, end users (hands on, decision makers), and tech support/training personnel to ensure the sustainment of an analytic.
- Many folks in the audience expressed interest in seeing the full list of 74 analytics identified during last year's efforts.
- A recurring theme of the roles aimed to ensure that cybersecurity is built in as early as possible. The goal would be to have analytics help ensure design at an early stage with a methodical and systematic process.

- Folks also expressed desire to have the entire systems engineering discipline, from concept development through O&S, involved in cyber analytics.

Question (2): What are the key information requirements for these roles and the analytic development lifecycle?

- Threat Centric:
  - Currently known threats
  - Existing threat profiles
  - Emerging threat projections
  - Red team against current list of threats & vulnerabilities
  - Pull CERT daily reports and profile prominent threats & test against requirements
- Requirements and Reference Artifacts:
  - Mission requirements
  - Cyber requirements of other systems
  - Requirements traceability
  - System requirement specification
  - System architecture documents
  - System design specifications
  - Cyber reference architecture(s)
- Cyber Activities
  - Privacy and security controls
  - Supply chain analysis
  - Source and quality of data (how do we know that it is good and repeatable)
  - Confidence of said source
- Other
  - Effectiveness, suitability, traceability, monitoring system in sustainment
  - Flexibility for measures regarding quality of data
  - Posture of the information environment (where system to be deployed) as well as other environmental aspects

The above small group discussions highlighted the need for ACF use cases for the CAL.

3.6.2 Session #2 Discussion Points

The second small group session covered the following two questions:

1. What are the approaches you have seen; what worked, what didn't?
2. How would you do cyber testing if money was no constraint?

Discussion Notes

Question (1): What are the approaches you have seen; what worked, what didn't?

- What has worked:
  - Working with cyber test ranges.
  - Implementing and executing iterative MBCRAs to scope test objectives.
    - One specific MBCRA called out was cyber table tops.
    - MBCRAs help to identify mission model ahead of time.
  - Requirements definition, mission modeling, user involvement.
- What has not worked:
  - Do not share data very well.
  - Do not measure what we are testing.
  - No recurring testing: Testing is too restricted during exercises, test for controls too limiting.
  - Don't articulate mission impact well, limited schedules.
  - Testers starting with zero knowledge—don't know requirements or mission; too much time learning the system; not including operational user; lack of requirements provided to testers and test requirements; system requirements for cyber not in system specs—not flowed down.
  - Red team:
    - Shortage of red teams.
    - No consistent behavior in testers.
    - Not enough sharing of tools, techniques, and information.
    - Does not always provide broader implications of findings.
    - Does not always provide the program office with what they need. "Program office fixes, then red team just side-steps the fix. The attack that is successful is a specific example of a broader problem that does not get addressed. Red team may understand the broader problem whereas the program office may not."
    - Communication breakdown—both red team and program offices. "Red team doesn't want to lay everything out and program office may not want to understand the problem, they just want to understand how to defend. PO deals with multitude of issues and may not have the bandwidth to tackle a broader spectrum problem or have resource barriers."
- Significantly different conceptualizations of cyber testing.
  - One group had the opinion that red teaming was the only acceptable cyber testing while another views it as ineffective. The MITRE D3FEND team is interested in creating a taxonomy of cyber testing to help differentiate among the concepts; this would be integrated with the ACF extension in D3FEND.

Question (2): How would you do cyber testing if money was no constraint?

- Test Environments:
  - Separate environment—sandboxed system, soft or hard environment, check things out early.
  - Ensure entire test environment is equipped with cyber tools.
  - Build tools that would characterize attack surfaces as opposed to identifying pathways.
  - Maximize digital twin utilization.
- Test Activities:
  - Continuous test and evaluation, including key activities such as:
    - Cooperative testing.
    - Automated red teaming.
    - Automated software testing.
    - Full spectrum software assurance testing, both static and dynamic.
    - Supply chain testing.
    - Destructive testing.
    - Field testing with systems currently in the field (provide real-world data back to test community).
    - System of systems testing.
    - Commercial cloud testing.
    - Test of critical infrastructure.
    - Use formal methods and document findings.
    - Regression and recovery testing.
    - Component testing with appropriate threats.
    - Test mission success despite threats.
    - Have system models up and running so testers can test live.
    - Measuring resilience of system to go along with functional testing.
  - Perform continuous research on the adversary.
  - Get rid of fee for T&E service construct.
  - Testers identify problems, what solutions exist, and provide this information to the program managers.
  - Incorporate a sponsor or PM to the process.
  - Funding available for both testing and fixing.
- Test Community:

- Improved trust and transparency between the program office and testers to have a more secure system.
- Work with engineering community to identify cyber requirements early.

Overall note from the discussions: Even if money was no object, time constraints would still pose challenge.

3.6.3 Session #3 Discussion Points

The final small group session covered the following two questions:

1. What do you see as the pros and cons of each notional taxonomy's organization for analytic technique characterization?
2. What do you see as the gaps in each taxonomy for the coverage of cyber analytic techniques that must be characterized?

Discussion Notes

Question (1): What do you see as the pros and cons of each notional taxonomy's organization for analytic technique characterization?

- Pros:
  - Good job on machine learning taxonomy; might prompt creation of new analytics.
  - Second taxonomy was preferred over the first by at least one group.
- Cons:
  - Discussed need for use cases for ACF as this would help inform opinion on taxonomies.
  - Need to define for audience the difference between taxonomy and ontology.
  - Need more information on how to evaluate taxonomies 1 and 2; what is a "good taxonomy"?
  - Need evolutionary framework; the framework should be more dynamic to capture updates in a more agile manner.
  - Feels like within this universe we have Flintstones world now, but taxonomy is starship—too far apart; are beyond where we are today.
  - Useful exercise: Try to bin tools reviewed in CAL into categories to test the taxonomies.
    - One working group representative believes they will all bin to one place.

Question (2): What do you see as the gaps in each taxonomy for the coverage of cyber analytic techniques that must be characterized?

- Dr. Llanso gave Mr. Kaloroumakis examples of analytic techniques that might be considered for inclusion.
- Understanding purpose and value measures of the ACF.
- Understanding who the intended users are of ACF.

4 Summary and Follow-on Actions

There were several action items captured throughout the workshop. This section highlights some of the key action items.

4.1 Develop an Integrated View of FY24 CAL Efforts

It was unclear to the audience how all the FY24 CAL efforts integrated into a single effort toward the CAL strategic visions. The FY24 efforts are focused on diverse pieces of the overall CAL problem space. The CAL team will work to create an integrated view of the FY24 efforts to show their dependencies and relationships.

4.2 Revisit ACF Use Cases

One main sponsor action from this workshop was to identify, develop, and deliver an updated use case scenario for ACF. This includes understanding the problem set that ACF wishes to mitigate or solve, and potential user community for the capability.

4.3 Provide List of 74 Analytics

During FY23 the CAL team documented 74 analytics. Participants in the workshops identified a need to know what analytics were documented and what information was captured on them to better support the discussions. Prior to CAL Workshop #4 a full list of analytics and documented information should be provided in the read-ahead materials.

4.4 Map CAL Analytics to ACF

Of the 74 analytics captured in FY23 none are currently mapped to the ACF. While the ACF is still in development, mapping the current analytics could be an opportunity to identify gaps and provide examples of ACF use cases with known analytics.

4.5 Identify Data Probability Distribution Subject Matter Experts

During the workshop discussion, it was brought up that reusable work might be available to support the “PCyber” analytic development. The CAL team will work to identify prior work across the DoD in this area.

4.6 Revisit Edge Cases for Monte Carlo Simulations

Prior literature was reviewed by the CAL team in FY23 on edge cases for Monte Carlo simulations. That literature will be reviewed to determine how we can apply more representative data to the “PCyber” Resilience Index approach.