Modern A.I. Expert Systems for Active Defense

By

Shawn Riley, Chief Visionary Officer, DarkLight, Inc.

Michael Forgione, Network Defense Operator, R9B

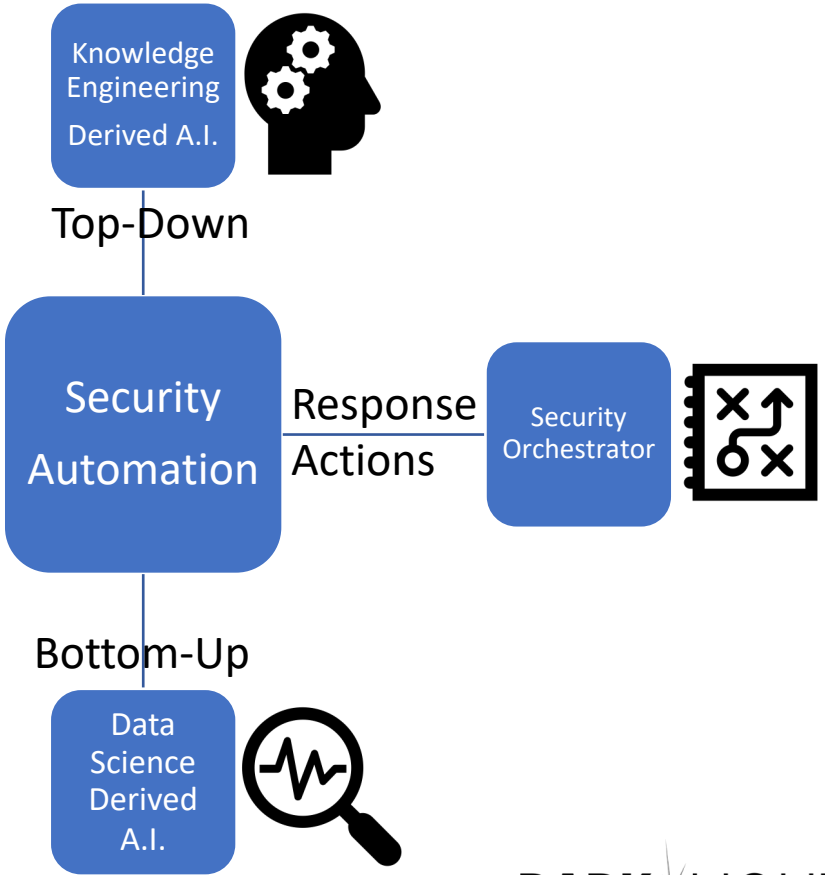
© 2019 DarkLight, Inc.

Three Key Security Automation Technologies Similar Value Propositions But Different Focuses

- **Knowledge Engineering Derived A.I.**
  - Focuses on integrating knowledge and mimicking how human expert's apply the knowledge
  - Applies deterministic reasoning to automate the verification of tentative hypotheses. (White Box)

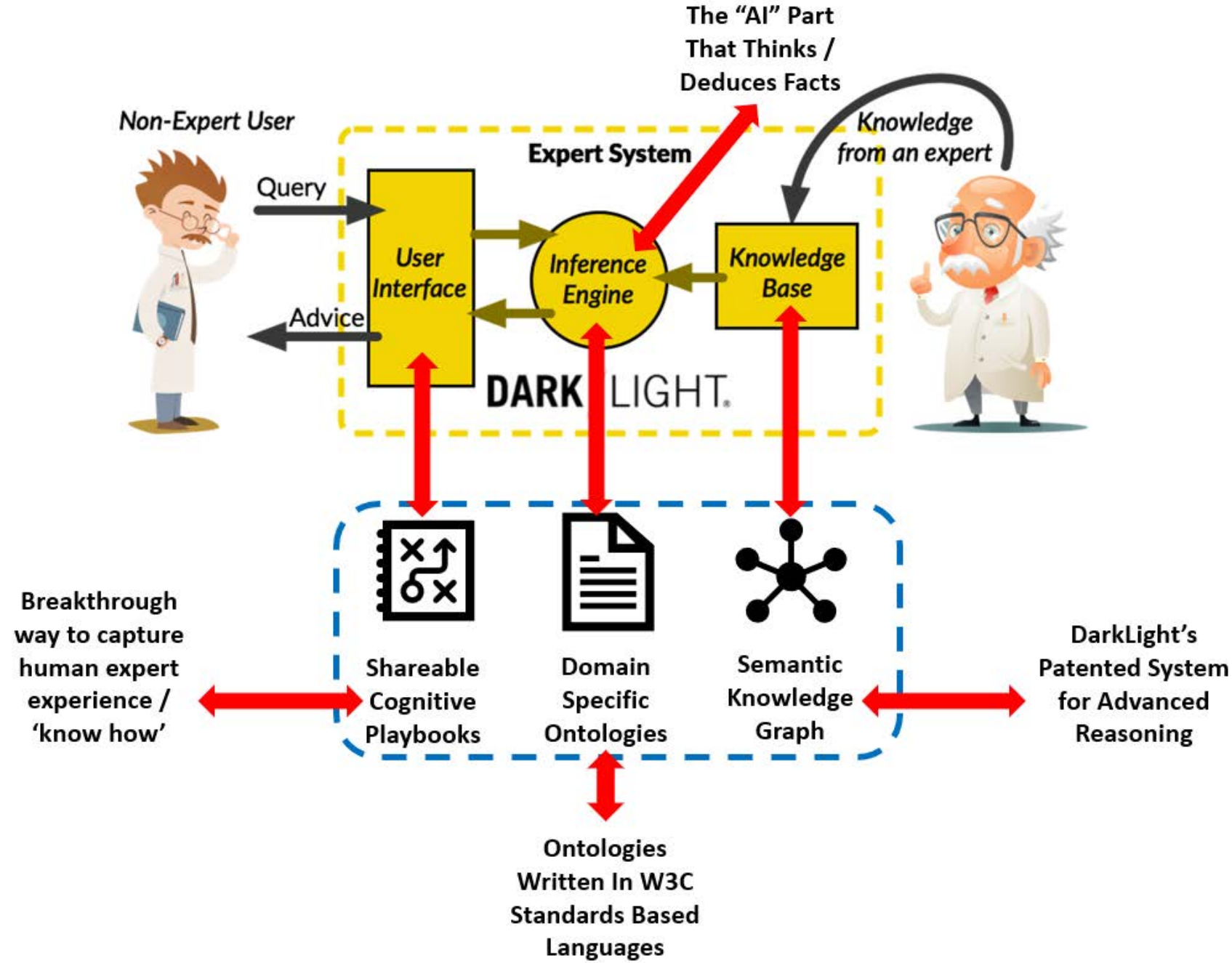
- **Security Orchestrator**
  - Focuses on integrating technology and automating mechanistic response actions.

- **Data Science Derived A.I.**
  - Focuses on the data for classification, clustering, and prediction
  - Applies probabilistic reasoning to produce tentative hypotheses (Black Box)
  - Saves the human from having to dig through all the data sets to find the patterns of interest.



DARK LIGHT.ai
Active Defense Expert System

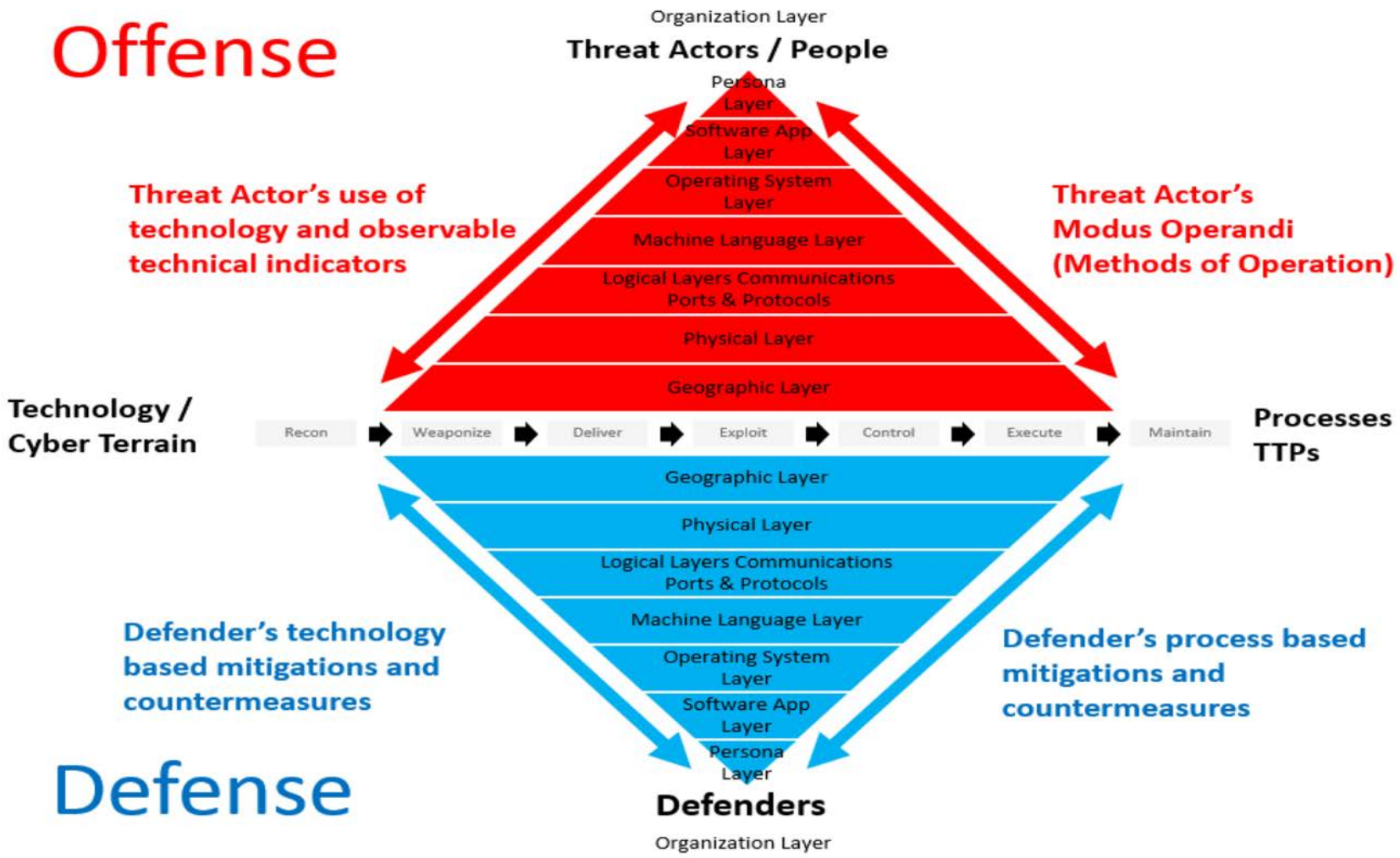
Integrate Automate Validate **Explain**



DARK LIGHT.ai
Active Defense Expert System

Integrate Automate Validate **Explain**

Organized Adversary & Defender Knowledge



Adversary Contextual Knowledge

Transactions & Activity

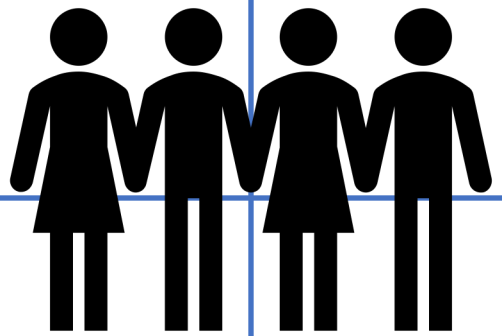
Enterprise Contextual Knowledge

DARK LIGHT.ai
Active Defense Expert System

Integrate Automate Validate **Explain**

Cyber Defense Analyst
Work Role ID: PR-CDA-001
Uses data collected from a variety of cyber defense tools (e.g., IDS alerts, firewalls, network traffic logs) to analyze events that occur within their environments for the purposes of mitigating threats.
Category:
Protect and Defend
Specialty Area:
Cyber Defense Analysis

Cyber Defense Incident Responder
Work Role ID: PR-CIR-001
Investigates, analyzes, and responds to cyber incidents within the network environment or enclave.
Category:
Protect and Defend
Specialty Area:
Incident Response



Mimicking These Roles
NICE Cybersecurity Workforce Framework

Threat/Warning Analyst
Work Role ID: AN-TWA-001
Develops cyber indicators to maintain awareness of the status of the highly dynamic operating environment. Collects, processes, analyzes, and disseminates cyber threat/warning assessments.
Category:
Analyze
Specialty Area:
Threat Analysis

Cyber Crime Investigator
Work Role ID: IN-INV-001
Identifies, collects, examines, and preserves evidence using controlled and documented analytical and investigative techniques.
Category:
Investigate
Specialty Area:
Cyber Investigation

DARK LIGHT.ai
Active Defense Expert System
Integrate Automate Validate **Explain**

Comparing Symbolic AI & Non-symbolic AI

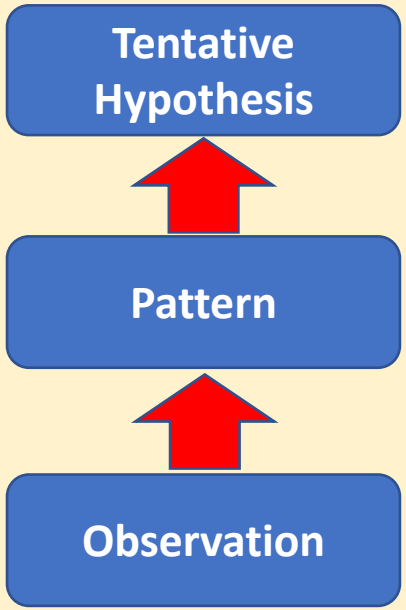
Artificial Intelligence (AI)

Symbolic AI	Non-symbolic AI
Knowledge Engineering	Data Science
Expert System	Machine Learning
Cognitive Playbooks & Ontologies	Algorithms & Models
Deductive Inference & Deterministic Reasoning	Inductive Inference & Probabilistic Reasoning
Validation of Hypotheses & Explanation	Predictions & Tentative Hypotheses
Transparent & Explainable	Black Box

DARK LIGHT.ai®
Active Defense Expert System
Integrate Automate Validate **Explain**

Machine Learning focuses on prediction, based on *known* properties learned from the training data. **Inductive Reasoning** uses patterns to arrive at a conclusion (conjecture). **Note:** A conclusion derived through inductive reasoning is called a hypothesis and is always less certain than the evidence itself. In other words, the conclusion is ***probable***.

Property Graphs

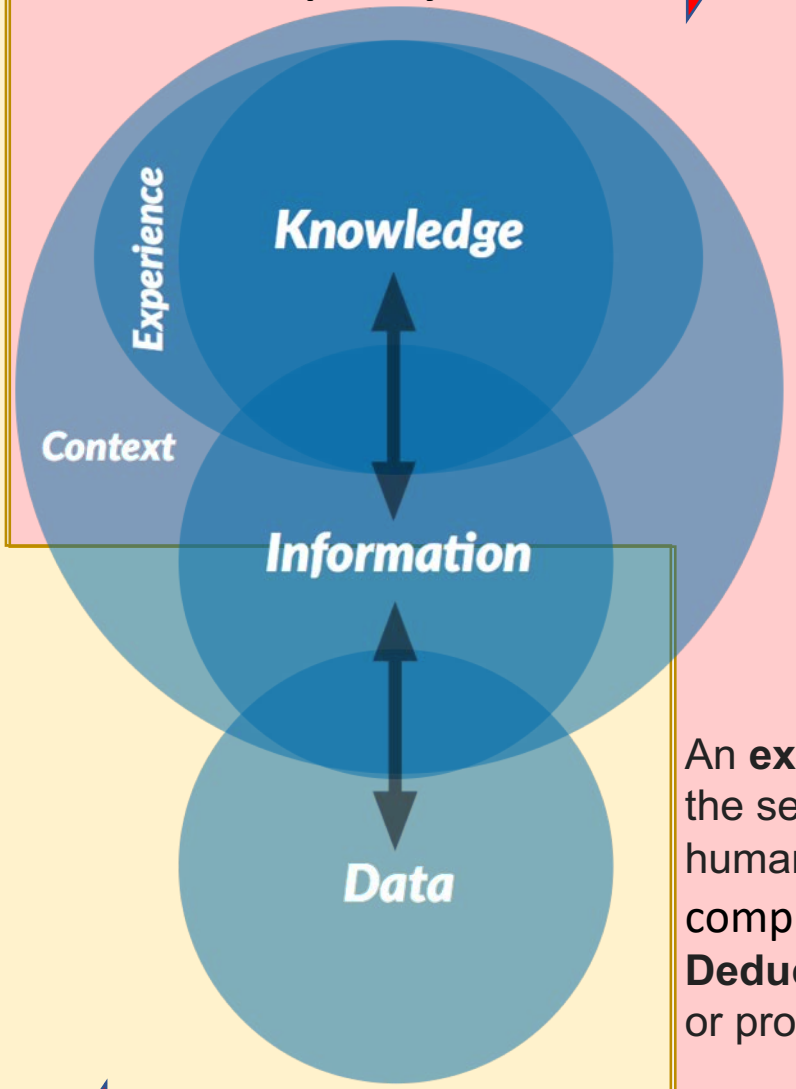


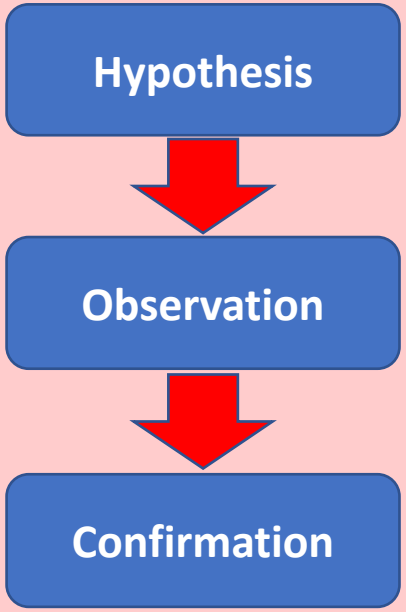
Inductive Reasoning
Bottom-Up Approach
Specific to Generalization

Machine Learning
Predictive Analytics
Scoring Engines

Knowledge Representation And Reasoning Expert System

Deductive Reasoning
Top-Down Approach
General to Specific



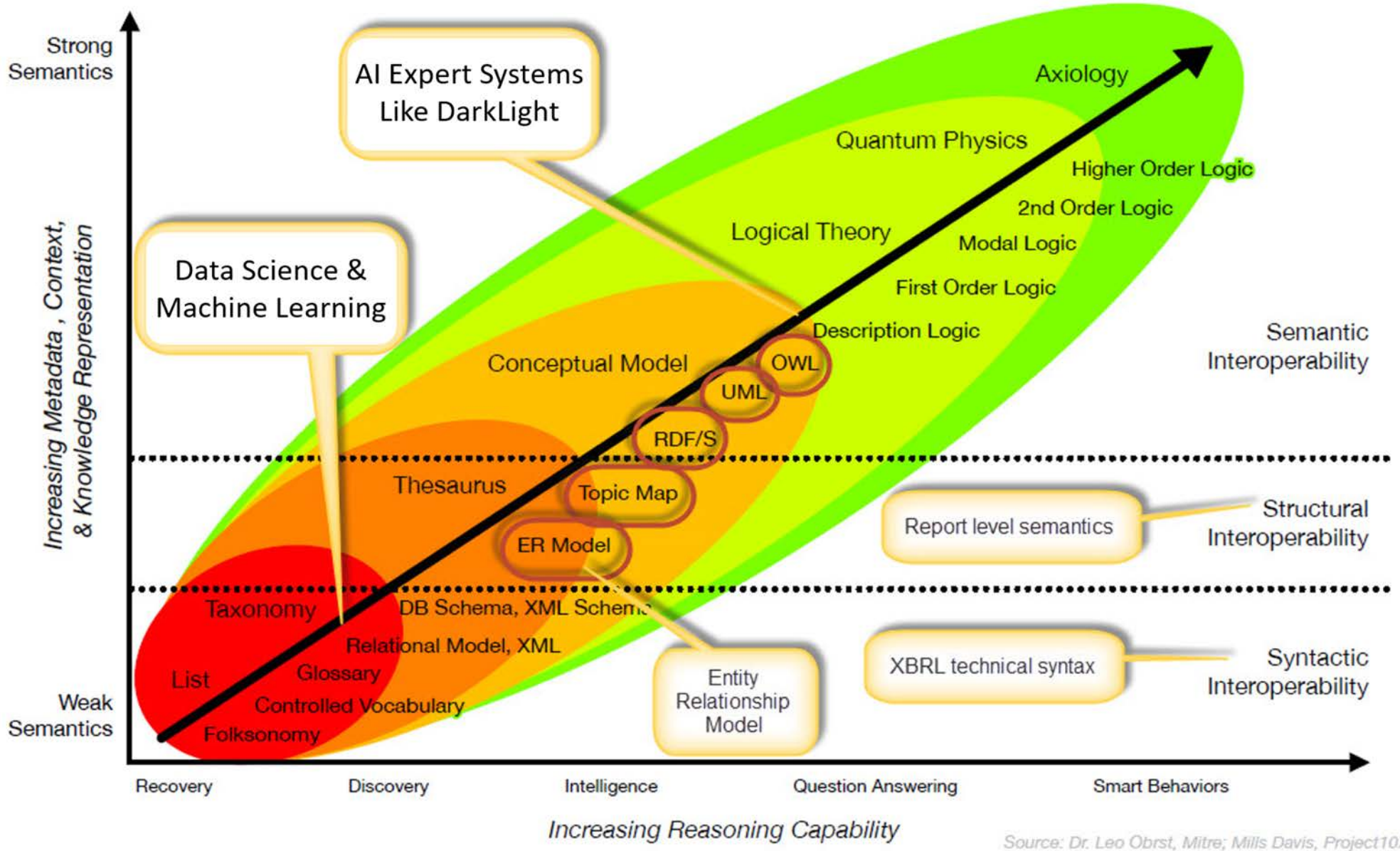


Knowledge Graphs

An **expert system** is an A.I. system that emulates the sense-making and decision-making ability of a human expert. Expert systems are designed to solve complex problems by reasoning about knowledge. **Deductive Reasoning** uses facts, rules, definitions or properties to arrive at a conclusion.

DARK LIGHT.ai
Active Defense Expert System

Integrate Automate Validate **Explain**



Source: Dr. Leo Obrst, Mitre; Mills Davis, Project10X

DARK LIGHT.ai
Active Defense Expert System

Integrate Automate Validate **Explain**

Axioms: assertions (including rules) in a logical form that together comprise the overall theory that the ontology describes in its domain of application.

Axiom 1
For every intrusion event there exists an adversary taking a step towards an intended goal by using a capability over infrastructure against a victim to produce a result.

Axiom 2
There exists a set of adversaries (insiders, outsiders, individuals, groups, and organizations) which seek to compromise computer systems or networks to further their intent and satisfy their needs.

Axiom 3
Every system, and by extension every victim asset, has vulnerabilities and exposures.

Axiom 4
Every malicious activity contains two or more phases which must be successfully executed in succession to achieve the desired result.

Axiom 5
Every intrusion event requires one or more external resources to be satisfied prior to success.

Axiom 6
A relationship always exists between the Adversary and their Victim(s) even if distant, fleeting, or indirect.

Axiom 7
There exists a sub-set of the set of adversaries which have the motivation, resources, and capabilities to sustain malicious effects for a significant length of time against one or more victims while resisting mitigation efforts. Adversary-Victim relationships in this sub-set are called persistent adversary relationships.

http://www.activeresponse.org/diamond-model-axioms/

DarkLight Uses W3C OWL2 (DL) Ontologies

Enterprise & Sensor-based Ontologies

CAPEC

MITRE ATT&CK

Structured Threat Information Expression (STIX) v2

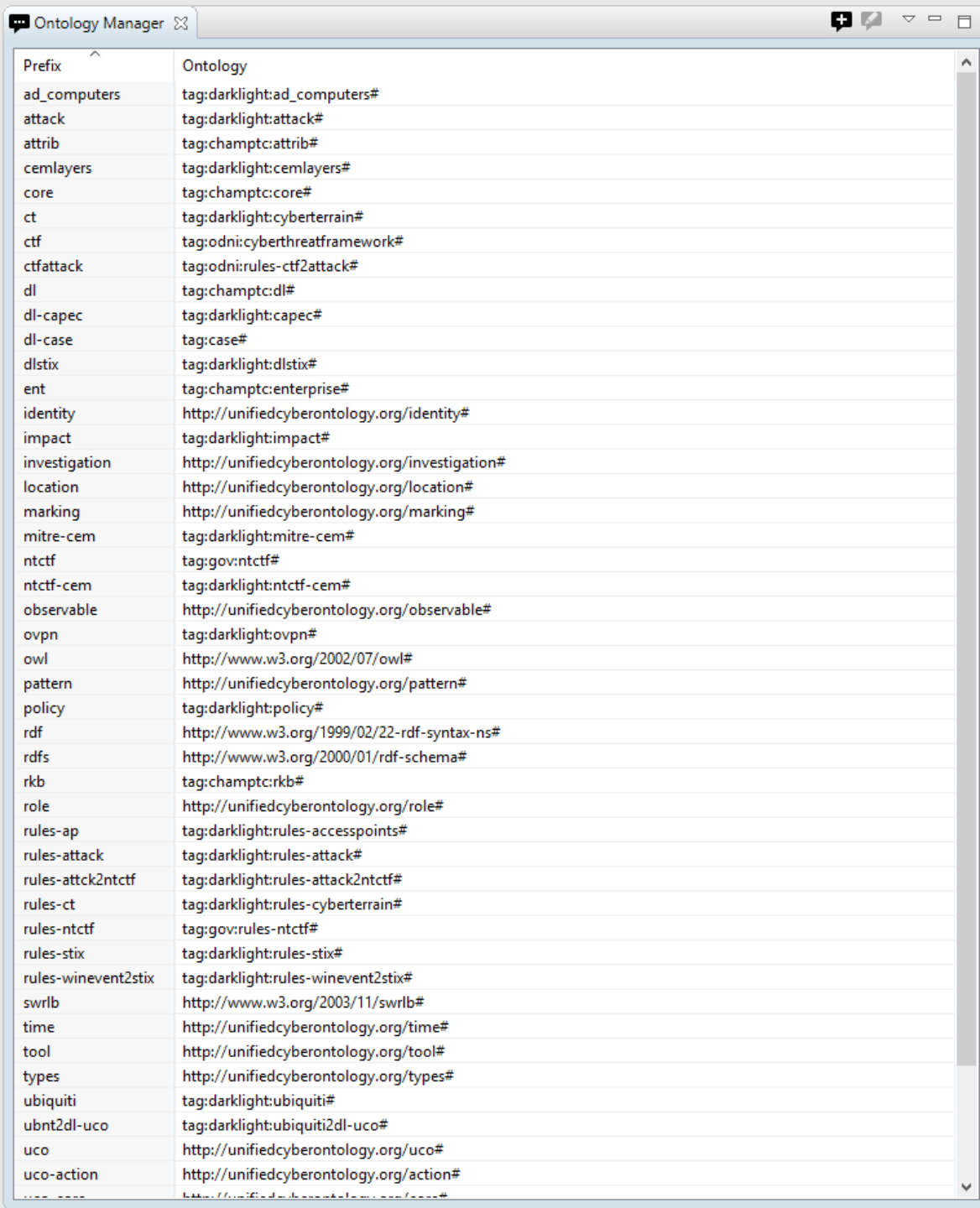
ODNI Cyber Threat Framework

NSA/CSS Technical Cyber Threat Framework

Cyber-investigation Analysis Standard Expression (CASE)

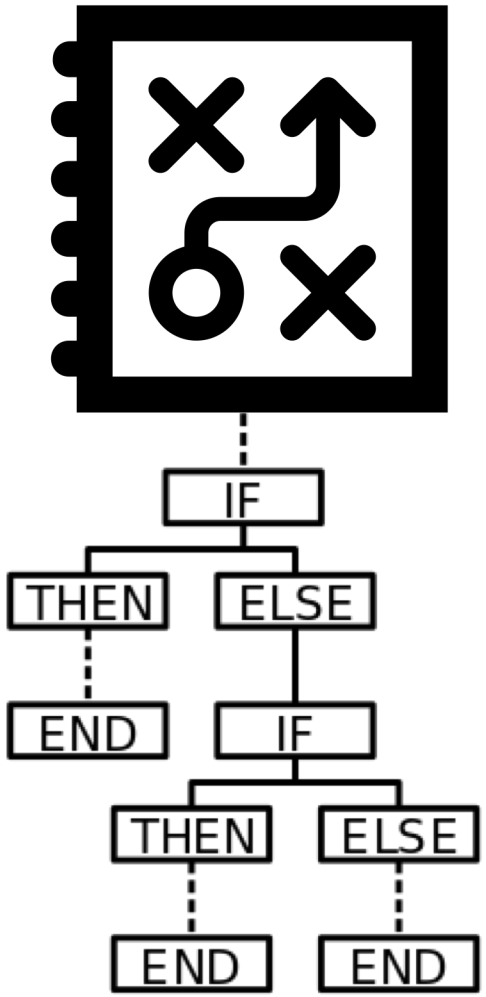
Unified Cyber Ontology

SWRL Inference Rules Ontologies



Prefix	Ontology
ad_computers	tag:darklight:ad_computers#
attack	tag:darklight:attack#
attrib	tag:champtc:attrib#
cemlayers	tag:darklight:cemlayers#
core	tag:champtc:core#
ct	tag:darklight:cyberterrain#
ctf	tag:odni:cyberthreatframework#
ctfattack	tag:odni:rules-ctf2attack#
dl	tag:champtc:dl#
dl-capec	tag:darklight:capec#
dl-case	tag:case#
dlstix	tag:darklight:dlstix#
ent	tag:champtc:enterprise#
identity	http://unifiedcyberontology.org/identity#
impact	tag:darklight:impact#
investigation	http://unifiedcyberontology.org/investigation#
location	http://unifiedcyberontology.org/location#
marking	http://unifiedcyberontology.org/marking#
mitre-cem	tag:darklight:mitre-cem#
ntctf	tag:gov:ntctf#
ntctf-cem	tag:darklight:ntctf-cem#
observable	http://unifiedcyberontology.org/observable#
ovpn	tag:darklight:ovpn#
owl	http://www.w3.org/2002/07/owl#
pattern	http://unifiedcyberontology.org/pattern#
policy	tag:darklight:policy#
rdf	http://www.w3.org/1999/02/22-rdf-syntax-ns#
rdfs	http://www.w3.org/2000/01/rdf-schema#
rkb	tag:champtc:rkb#
role	http://unifiedcyberontology.org/role#
rules-ap	tag:darklight:rules-accesspoints#
rules-attack	tag:darklight:rules-attack#
rules-attck2ntctf	tag:darklight:rules-attack2ntctf#
rules-ct	tag:darklight:rules-cyberterrain#
rules-ntctf	tag:gov:rules-ntctf#
rules-stix	tag:darklight:rules-stix#
rules-winevent2stix	tag:darklight:rules-winevent2stix#
swrlb	http://www.w3.org/2003/11/swrlb#
time	http://unifiedcyberontology.org/time#
tool	http://unifiedcyberontology.org/tool#
types	http://unifiedcyberontology.org/types#
ubiquiti	tag:darklight:ubiquiti#
ubnt2dl-uco	tag:darklight:ubiquiti2dl-uco#
uco	http://unifiedcyberontology.org/uco#
uco-action	http://unifiedcyberontology.org/action#

Security Orchestration - Acting Playbooks



+

API

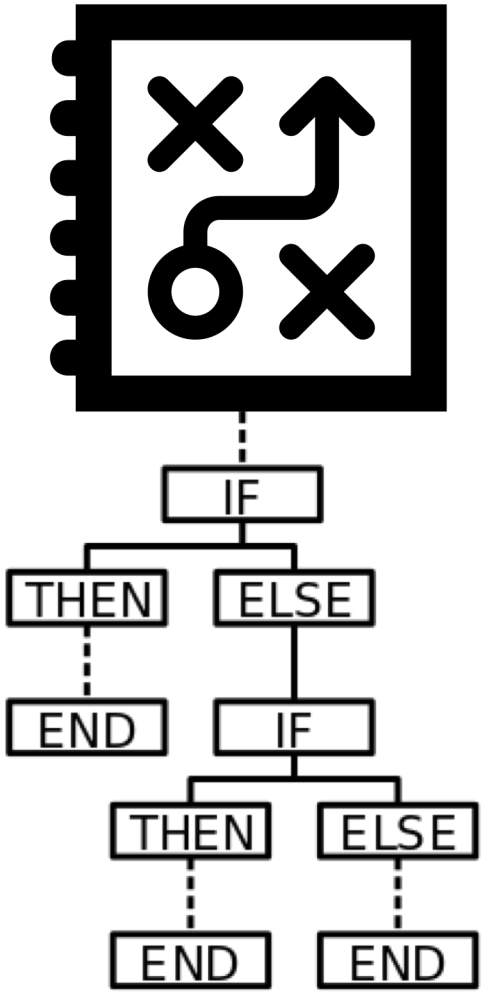
Application Programming Interface

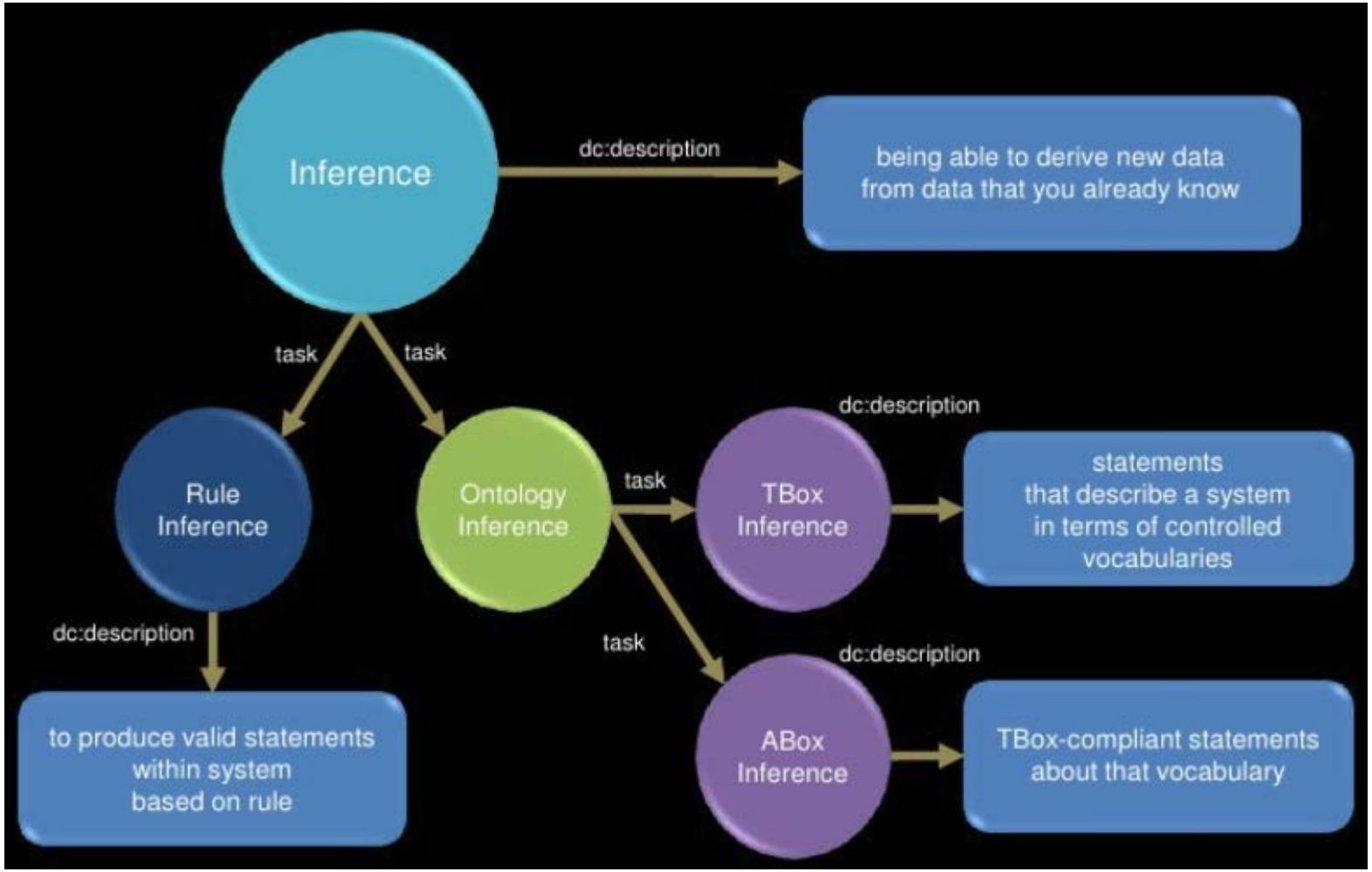
OpenC2

DARK LIGHT.ai
Active Defense Expert System

Integrate Automate Validate **Explain**

Knowledge Engineering AI - Cognitive Playbooks





DARK LIGHT.ai
Active Defense Expert System

Integrate Automate Validate **Explain**

Automating Argument-Driven Inquiry

The playbooks capture the human analyst's cognitive experience in applying the knowledge from the knowledge-base and can automate the Claim Evidence Reasoning framework.

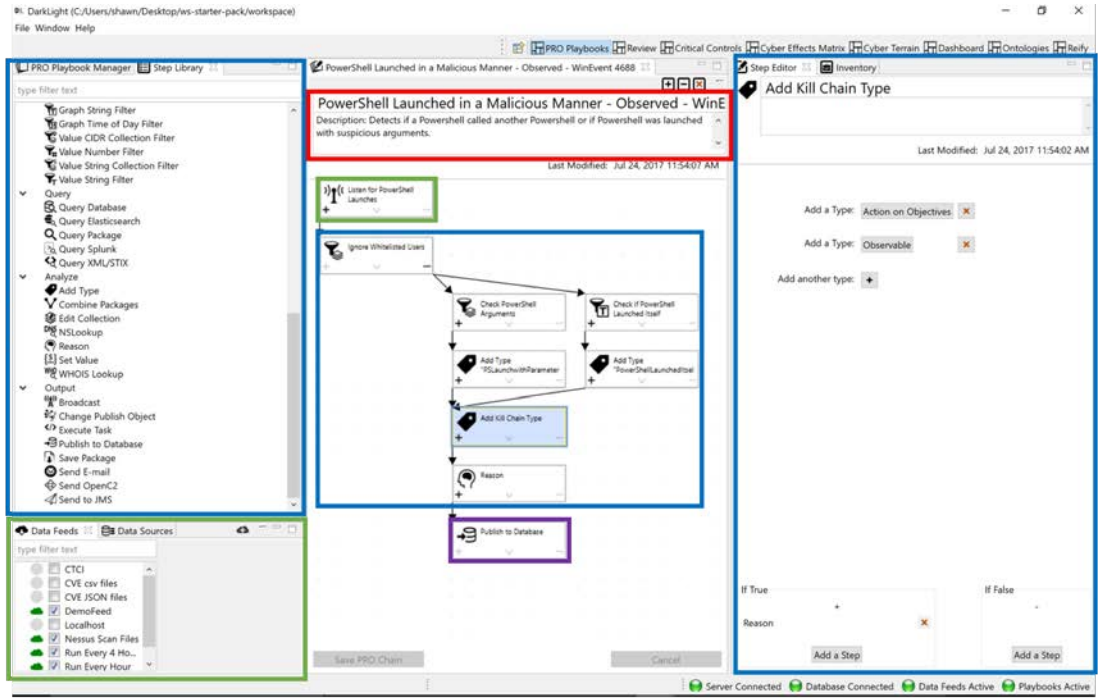
Cognitive Playbooks, their ontologies (knowledge models), and reify configuration are all sharable through import and export features allowing communities of trust to share knowledge and experience.

Claim

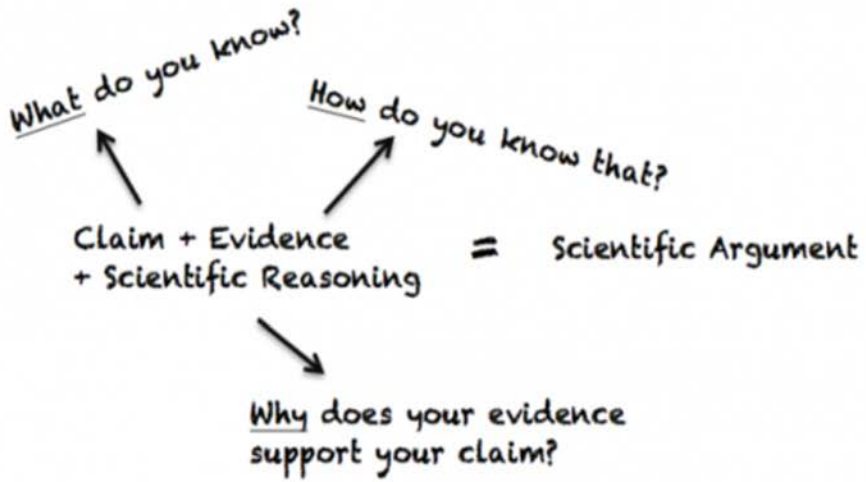
Evidence

Reasoning

CER Explanation Published to Memory



Typical Use Case: Argument-Driven Inquiry



Formal Description Logics

Predicate Logic – Statements made with Subject, Predicates, and Objects to form a Semantic Graph supporting logical arguments.

Definition – What is it? How should we define it?

Fact – Did it happen? Does it exist?

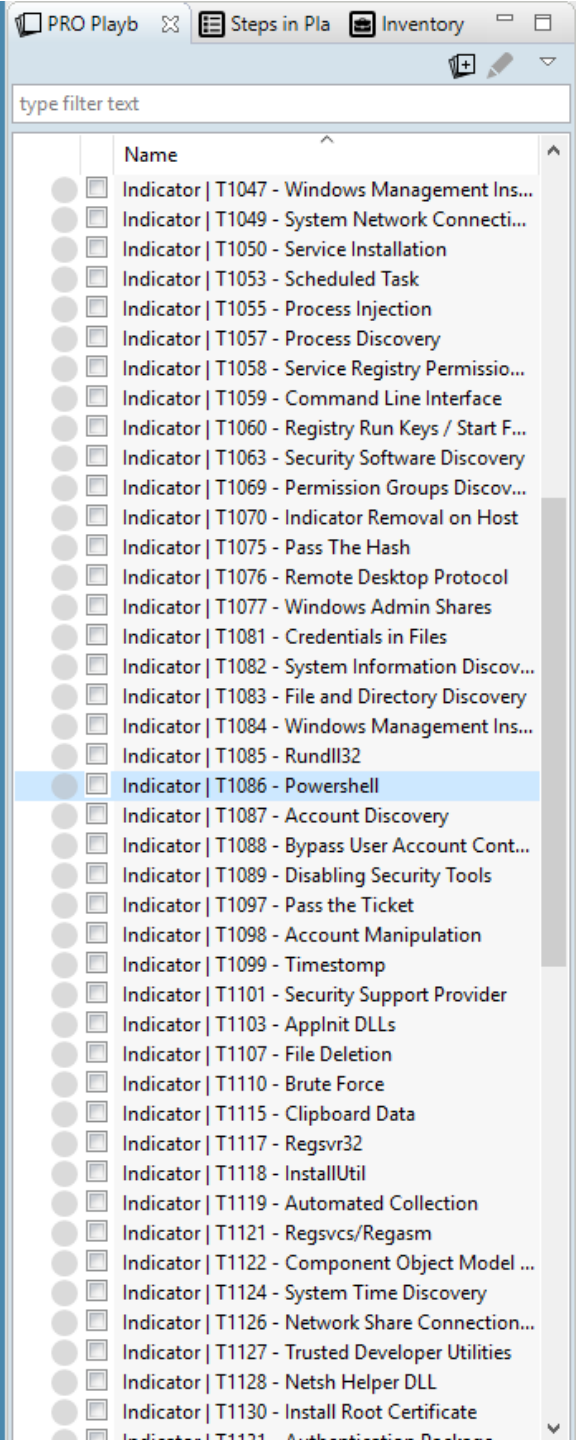
Value – Is it good or bad? What should be our criteria for deciding?

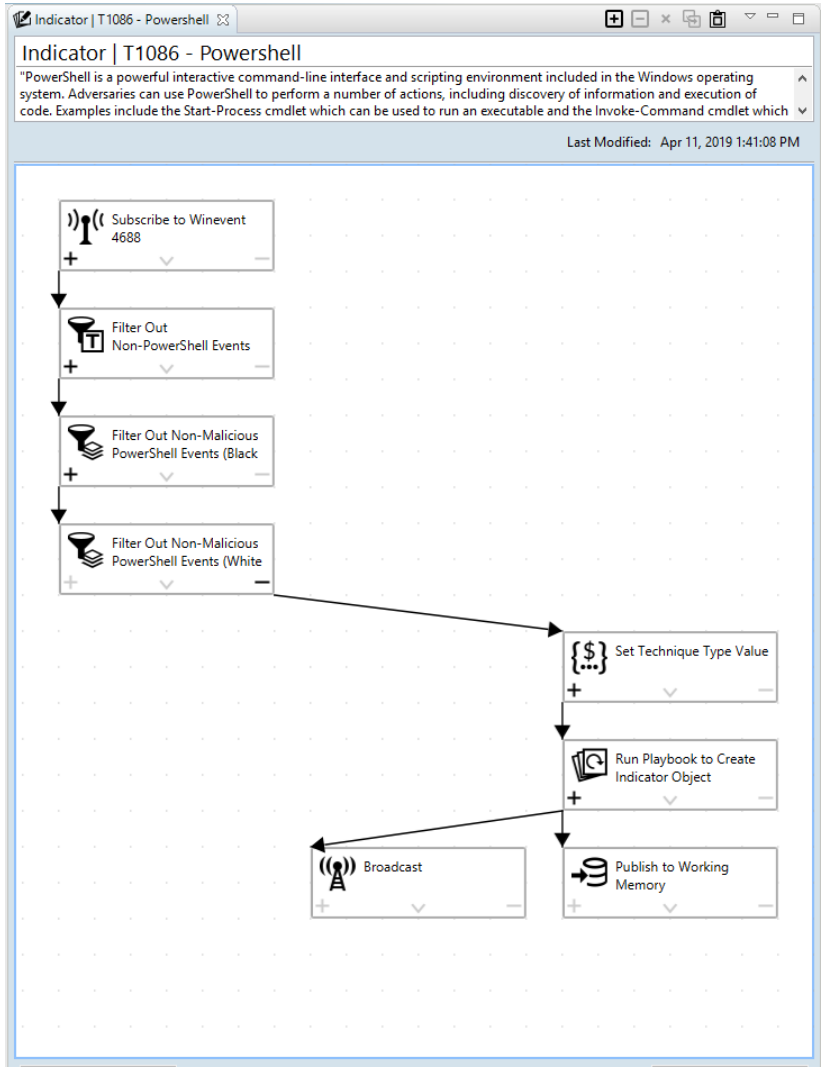
Policy – What should we do about it? What should be our future course of action?

Cause & Effect - What caused it? What are its effects?

DARK LIGHT.ai
Active Defense Expert System
Integrate Automate Validate **Explain**

Cognitive Playbook w/Sub-Playbooks





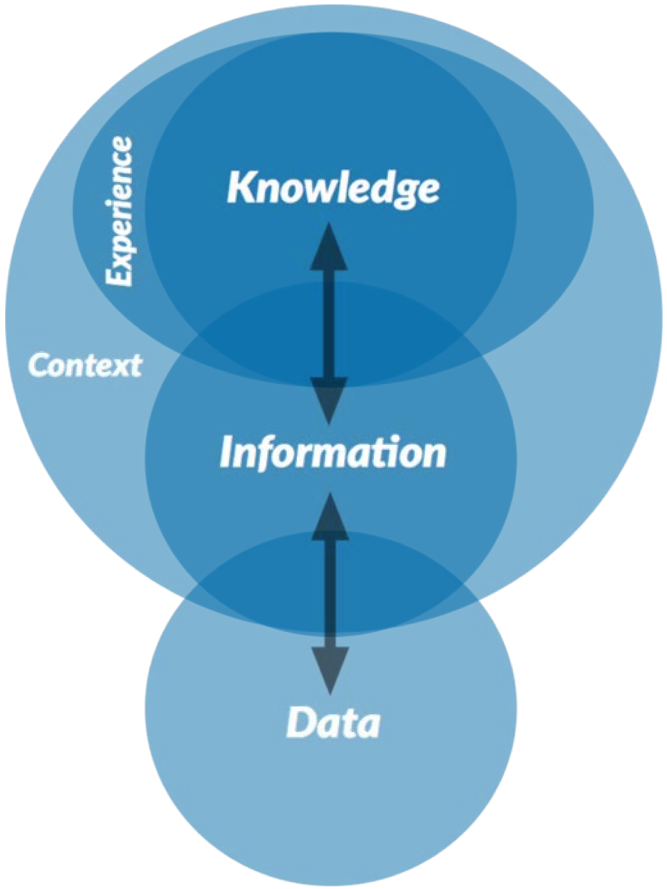
Adversary Behavior Indicator Playbook (Ex: ATT&CK T1086 PowerShell)

- Sub-Playbook – Create Indicator and Containers
  - Sub-Playbook – Active Directory Attribution
    - Sub-Playbook – Technique Pre-Attribution Separator
      - Sub-Playbook WinLogBeat Detail Extraction
      - Sub-Playbook Sysmon Detail Extraction
    - Sub-Playbook – Employee Attribution
    - Sub-Playbook – Device Attribution
  - Sub-Playbook – Create Impact Object (based on NIST)

DARK LIGHT.ai
Active Defense Expert System

Integrate Automate Validate **Explain**

Automating Logical Argumentation



Inferred Facts

Domain Specific Knowledge

Asserted Facts

Providing Context

Observed Facts

Sensor Observations

Logical Argument

STIX Threat Intel Report
CASE Cyber-investigation

Decision-making Support
OpenC2 Command & Control

DARK LIGHT.ai
Active Defense Expert System

Integrate Automate Validate **Explain**

Inference - a conclusion reached on the basis of evidence and reasoning.

Organization & Industry Knowledge

Acting

Decision Making

Assertion - a confident and forceful statement of fact or belief

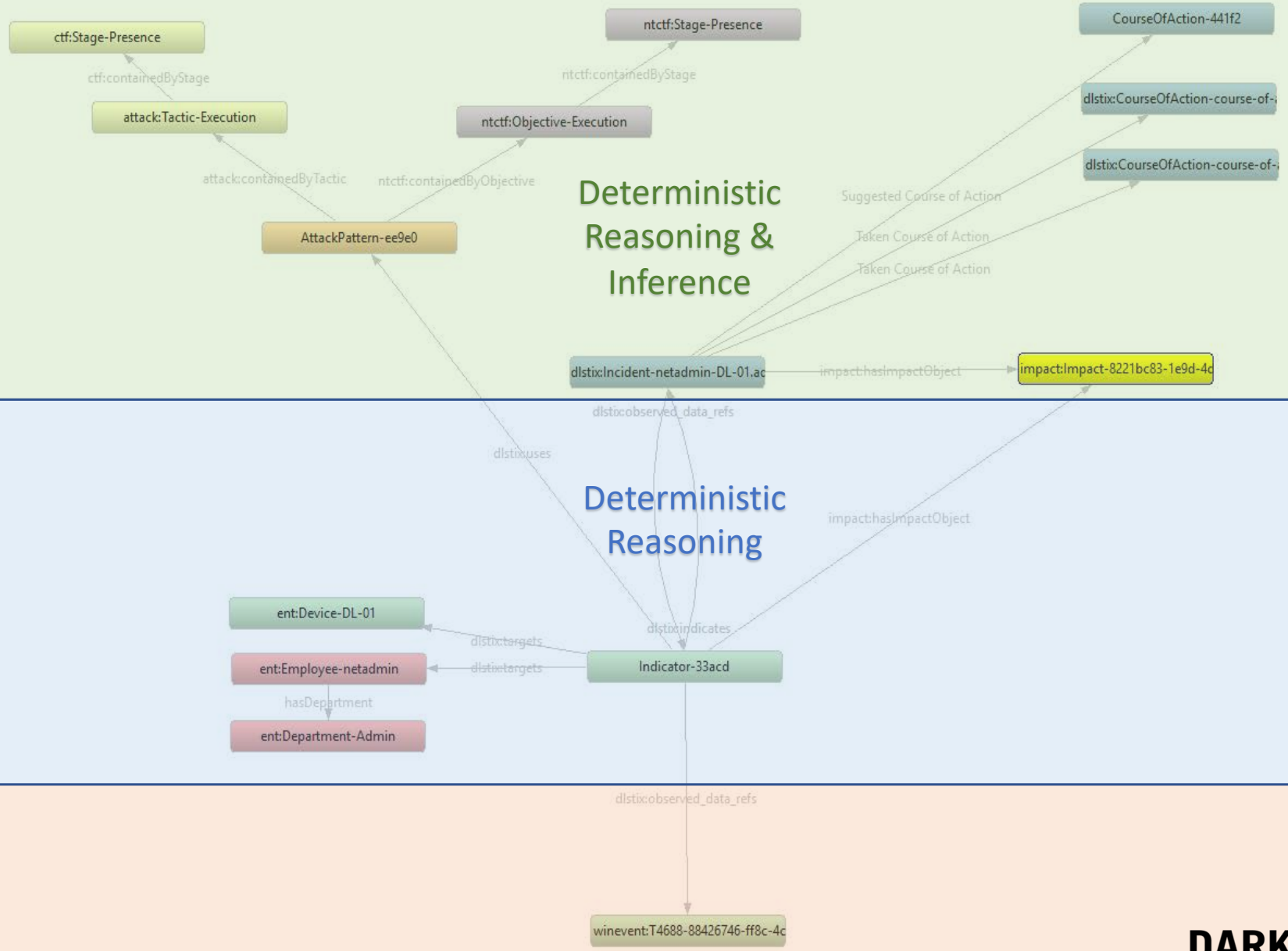
Contextual Information

Sense Making

Structured Observation from Sensor Output

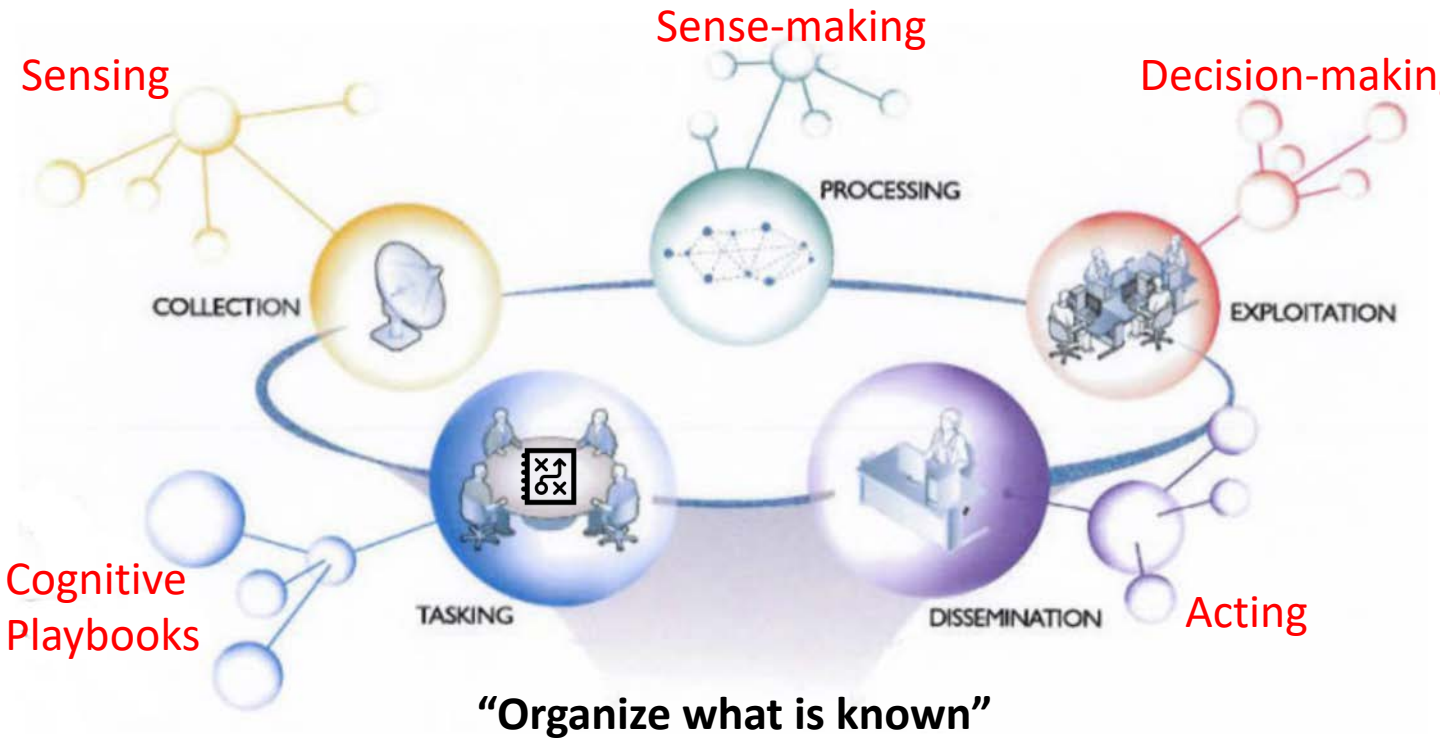
Evidence

Sensing



DARK LIGHT.ai
Active Defense Expert System

Object-Based Production & Activity-Based Intelligence



"Organize what is known"

"Discover the unknown unknowns"

Object-based production (OBP) and **activity-based intelligence (ABI)** are related IC analysis methodologies that rapidly integrates data from multiple sources to discover relevant patterns, determine and identify change, and characterize those patterns to create decision advantage and drive the sensing, sense-making, decision-making, and acting of the cyber OODA loop in the cyber environment. Activity-Based Intelligence promotes a **deductive approach to analytic reasoning** which reduces the space of potential outcomes by eliminating the impossible. **Note:** *ABI can be automated with cognitive playbooks with symbolic AI!!!*

DARK LIGHT.ai
Active Defense Expert System

Integrate Automate Validate **Explain**

Lockheed Martin Intelligence-Driven Defense Course of Action Matrix

Fundamentally, ***this approach is the essence of intelligence-driven defense*** that bases cyber mission assurance decisions and measurements on a detailed understanding of the adversary as they move through the cyber attack lifecycle.

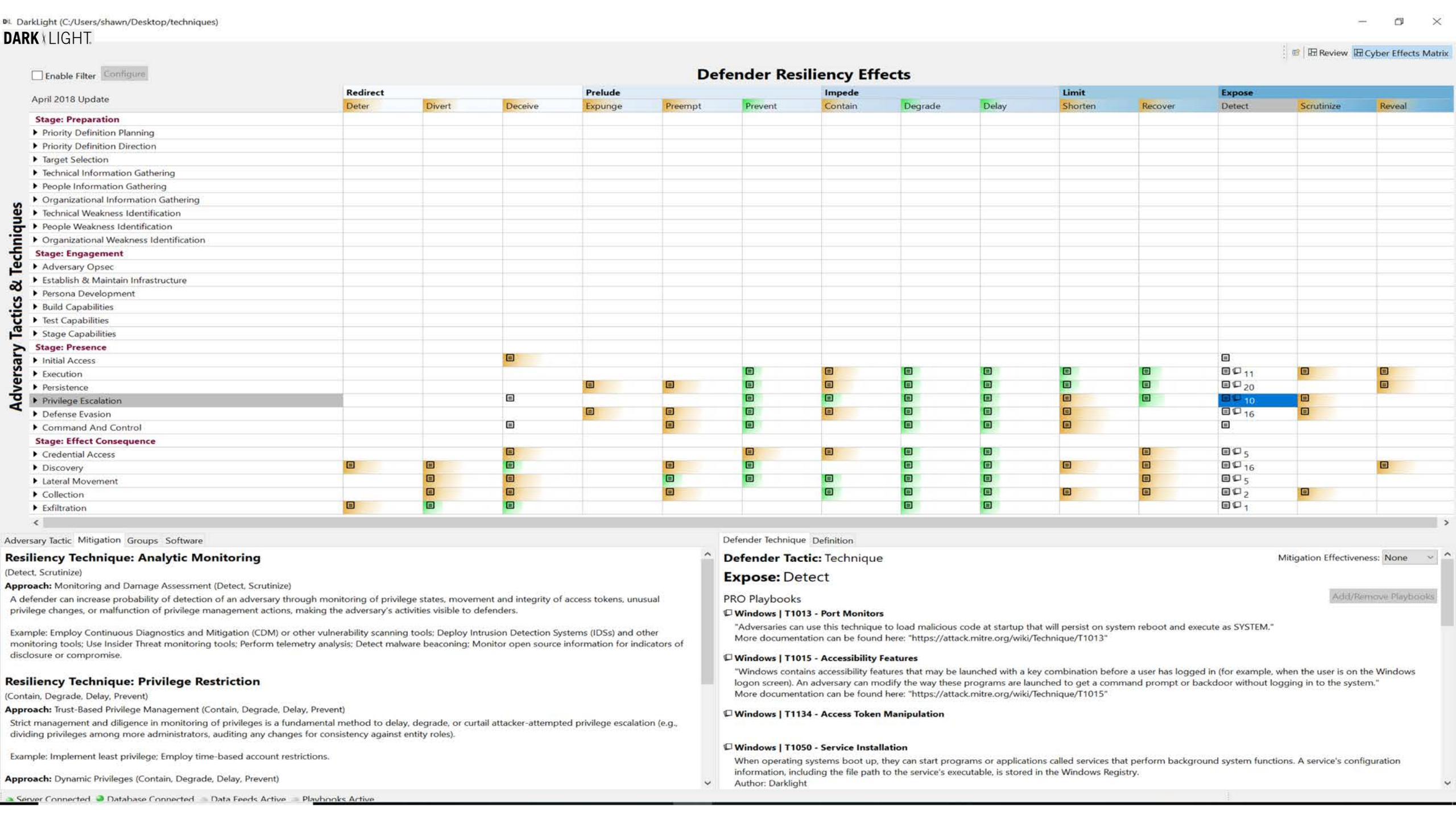
Phase	Detect	Deny	Disrupt	Degrade	Deceive	Destroy
Reconnaissance	Web analytics	Firewall ACL				
Weaponization	NIDS	NIPS				
Delivery	Vigilant user	Proxy filter	In-line AV	Queuing		
Exploitation	HIDS	Patch	DEP			
Installation	HIDS	"chroot" jail	AV			
C2	NIDS	Firewall ACL	NIPS	Tarpit	DNS redirect	
Actions on Objectives	Audit log			Quality of Service	Honeypot	

DARK LIGHT.ai
Active Defense Expert System
Integrate Automate Validate **Explain**

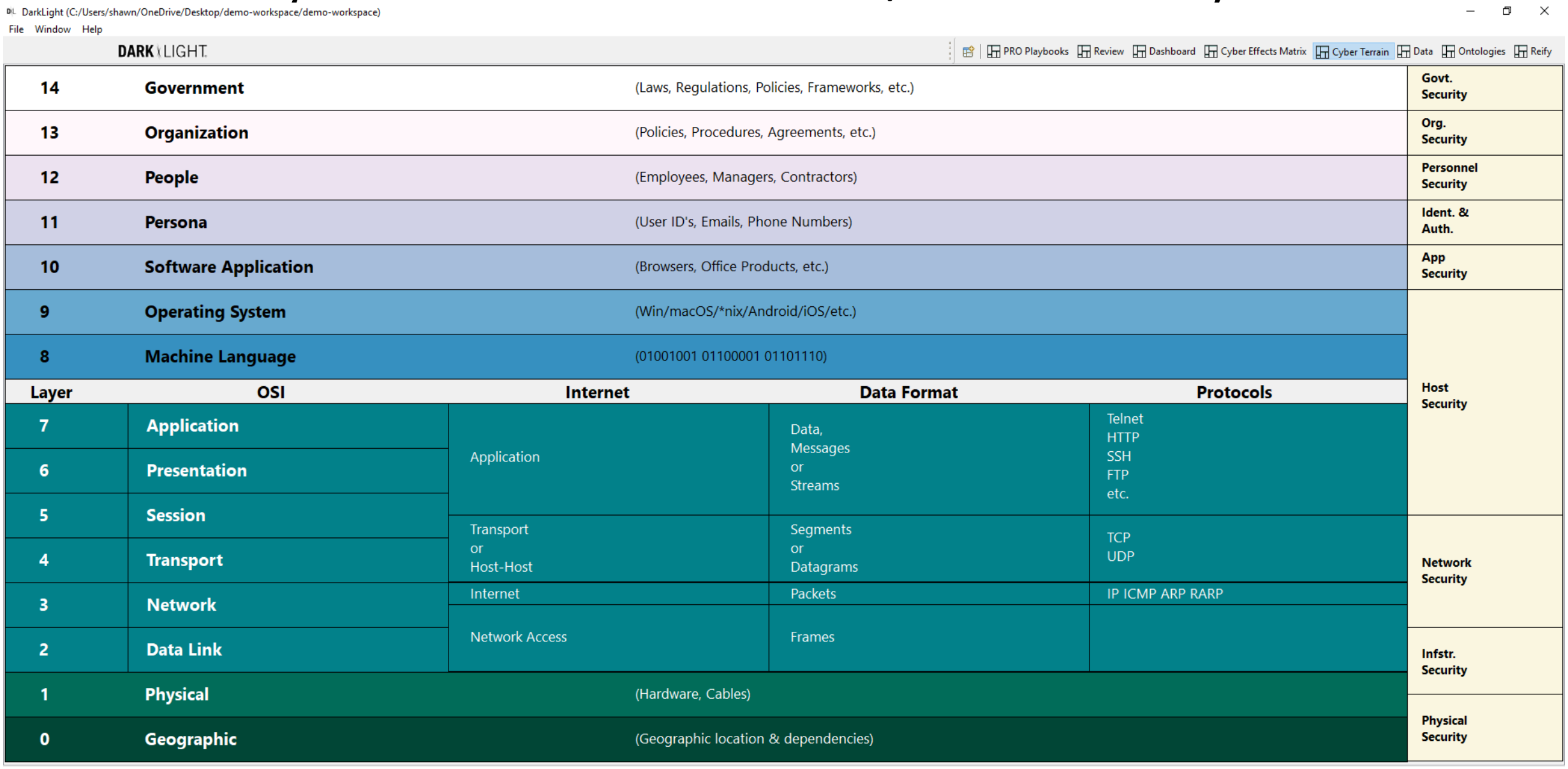
Cyber Resiliency Effects on Adversary Activities

- Deter, divert, and deceive in support of redirect;
- Prevent, preempt, and expunge in support of preclude;
- Contain, degrade and delay in support of impede;
- Shorten and recover in support of limit; and
- Detect, reveal, and scrutinize in support of expose

- NIST 800-160 vol 2 DRAFT


DarkLight (C:/Users/shawn/Desktop/techniques)
DARK LIGHT
Review
Cyber Effects Matrix
Enable Filter
Configure
Defender Resiliency Effects
April 2018 Update
Redirect
Prelude
Impede
Limit
Expose
Deter
Divert
Deceive
Expunge
Preempt
Prevent
Contain
Degrade
Delay
Shorten
Recover
Detect
Scrutinize
Reveal
Adversary Tactics & Techniques
Stage: Preparation
Priority Definition Planning
Priority Definition Direction
Target Selection
Technical Information Gathering
People Information Gathering
Organizational Information Gathering
Technical Weakness Identification
People Weakness Identification
Organizational Weakness Identification
Stage: Engagement
Adversary Opsec
Establish & Maintain Infrastructure
Persona Development
Build Capabilities
Test Capabilities
Stage Capabilities
Stage: Presence
Initial Access
Execution
Persistence
Privilege Escalation
Defense Evasion
Command And Control
Stage: Effect Consequence
Credential Access
Discovery
Lateral Movement
Collection
Exfiltration
11
20
10
16
5
16
5
2
1
Adversary Tactic
Mitigation
Groups
Software
Resiliency Technique: Analytic Monitoring
(Detect, Scrutinize)
Approach: Monitoring and Damage Assessment (Detect, Scrutinize)
A defender can increase probability of detection of an adversary through monitoring of privilege states, movement and integrity of access tokens, unusual privilege changes, or malfunction of privilege management actions, making the adversary's activities visible to defenders.
Example: Employ Continuous Diagnostics and Mitigation (CDM) or other vulnerability scanning tools; Deploy Intrusion Detection Systems (IDSs) and other monitoring tools; Use Insider Threat monitoring tools; Perform telemetry analysis; Detect malware beaconing; Monitor open source information for indicators of disclosure or compromise.
Resiliency Technique: Privilege Restriction
(Contain, Degrade, Delay, Prevent)
Approach: Trust-Based Privilege Management (Contain, Degrade, Delay, Prevent)
Strict management and diligence in monitoring of privileges is a fundamental method to delay, degrade, or curtail attacker-attempted privilege escalation (e.g., dividing privileges among more administrators, auditing any changes for consistency against entity roles).
Example: Implement least privilege; Employ time-based account restrictions.
Approach: Dynamic Privileges (Contain, Degrade, Delay, Prevent)
Defender Technique
Definition
Defender Tactic: Technique
Mitigation Effectiveness: None
Expose: Detect
PRO Playbooks
Add/Remove Playbooks
Windows | T1013 - Port Monitors
"Adversaries can use this technique to load malicious code at startup that will persist on system reboot and execute as SYSTEM."
More documentation can be found here: "https://attack.mitre.org/wiki/Technique/T1013"
Windows | T1015 - Accessibility Features
"Windows contains accessibility features that may be launched with a key combination before a user has logged in (for example, when the user is on the Windows logon screen). An adversary can modify the way these programs are launched to get a command prompt or backdoor without logging in to the system."
More documentation can be found here: "https://attack.mitre.org/wiki/Technique/T1015"
Windows | T1134 - Access Token Manipulation
Windows | T1050 - Service Installation
When operating systems boot up, they can start programs or applications called services that perform background system functions. A service's configuration information, including the file path to the service's executable, is stored in the Windows Registry.
Author: Darklight
Server Connected
Database Connected
Data Feeds Active
Playbooks Active

Cyber Environment w/Terrain Layers



14	**Government**	(Laws, Regulations, Policies, Frameworks, etc.)			**Govt. Security**
13	**Organization**	(Policies, Procedures, Agreements, etc.)			**Org. Security**
12	**People**	(Employees, Managers, Contractors)			**Personnel Security**
11	**Persona**	(User ID's, Emails, Phone Numbers)			**Ident. & Auth.**
10	**Software Application**	(Browsers, Office Products, etc.)			**App Security**
9	**Operating System**	(Win/macOS/*nix/Android/iOS/etc.)			**Host Security**
8	**Machine Language**	(01001001 01100001 01101110)			
Layer	**OSI**	**Internet**	**Data Format**	**Protocols**	
7	**Application**	Application	Data, Messages or Streams	Telnet HTTP SSH FTP etc.	
6	**Presentation**				
5	**Session**	Transport or Host-Host	Segments or Datagrams	TCP UDP	**Network Security**
4	**Transport**				
3	**Network**	Internet	Packets	IP ICMP ARP RARP	
2	**Data Link**	Network Access	Frames		**Infstr. Security**
1	**Physical**	(Hardware, Cables)			**Physical Security**
0	**Geographic**	(Geographic location & dependencies)			

DarkLight Cyber Effects Matrix Use Cases

- **Describe an intrusion chain of events based on the technique used from start to finish with a common reference of the cyber attack lifecycle and what layers of the cyber terrain are used by each technique**
- **Identify commonalities between adversary tradecraft (TTPs & Tools used), as well as distinguishing characteristics to support adversary attribution**
- **Determine “coverage” of a set of enterprise defensive capabilities to have different effects on adversary behavior across the cyber attack lifecycle to support gap analysis of sensors, actuators, and analytics that have can have a resiliency effect on the adversary activity**
- **Conduct analyses of alternatives between CND capabilities such as vendor ‘bake-offs’**
- **Prioritize development and/or acquisition efforts for CND capabilities**
- **Connect cyber defense mitigations, weaknesses, and adversaries**
- **Red Team – The cyber effect matrix provides a common language for describing the adversary objectives and actions at different stages of the cyber attack lifecycle and supports red team using the cyber attack lifecycle as a guide for pentesting and automated breach and attack simulation tools (adversary emulation).**
- **Blue Team – The cyber effects matrix provides a common languages for describing the effects of cyber mission assurance decisions on adversary actions. This includes enterprise defensive capabilities, mapping cyber threat indicators, incidence response playbooks, SOC runbooks (processes), any automated courses of action, etc.**
- **Purple Team – The cyber effects matrix provides a common language to describe both the stages, objectives, and actions of the adversary as they move through the cyber attack lifecycle as well as a common language for stating the effect(s) of cyber mission assurance decisions in the context of adversary activity.**
- **Understanding Threat Use Cases By Log Source Type**

IACD BRIEF

MAY 2019

DARK LIGHT.ai
Active Defense Expert System

Integrate Automate Validate **Explain**

CORPORATE OVERVIEW

R9B is a global cybersecurity leader founded on the principles of technical innovation, tailored solutions and professional excellence.

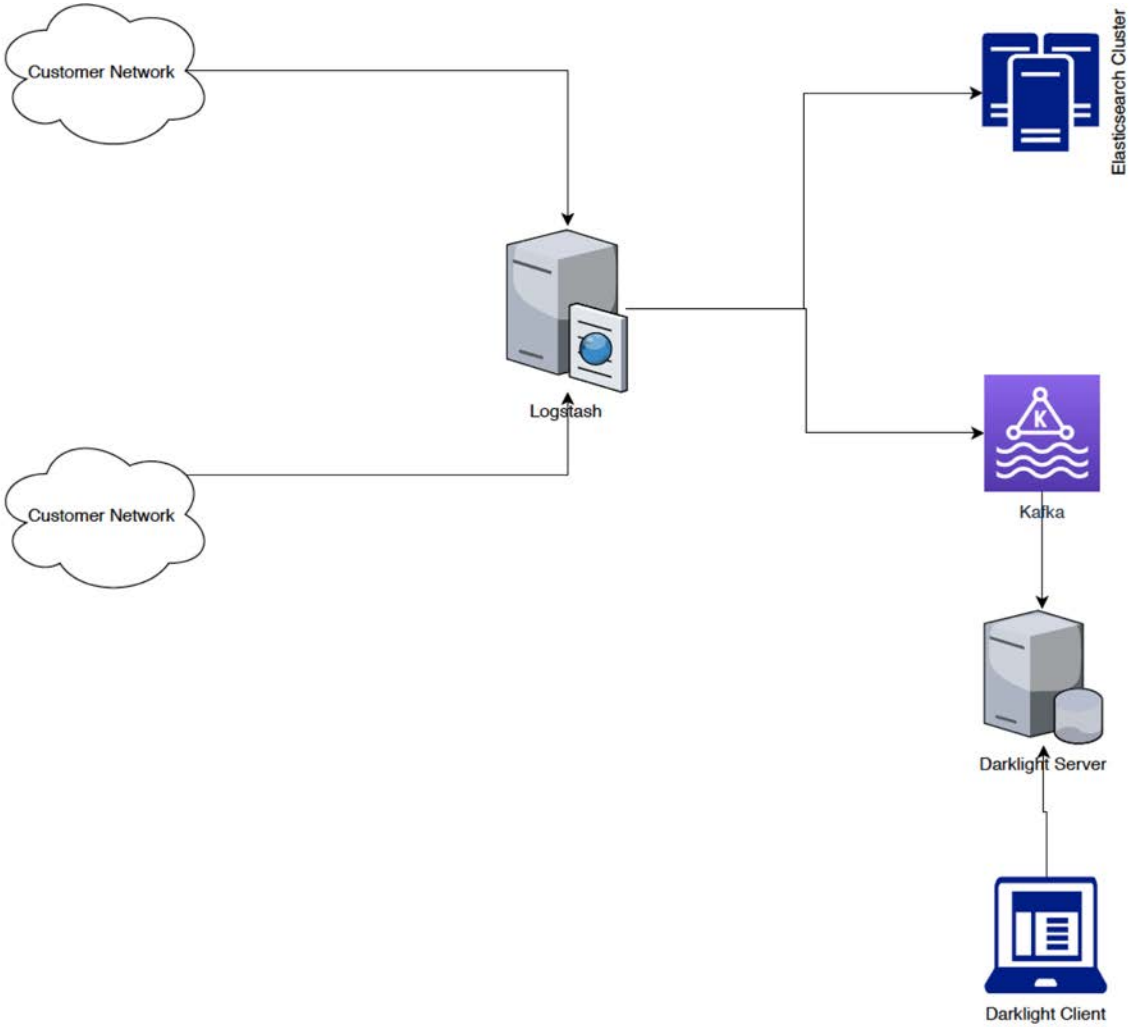
We provide customers all over the world the services, products and training necessary to deal with today's cybersecurity threats and challenges.

Our staff of certified Cybersecurity operators, product engineers, and threat analysts provide unparalleled vision, expertise, and experience tailored to meet each client's business context and mission needs.

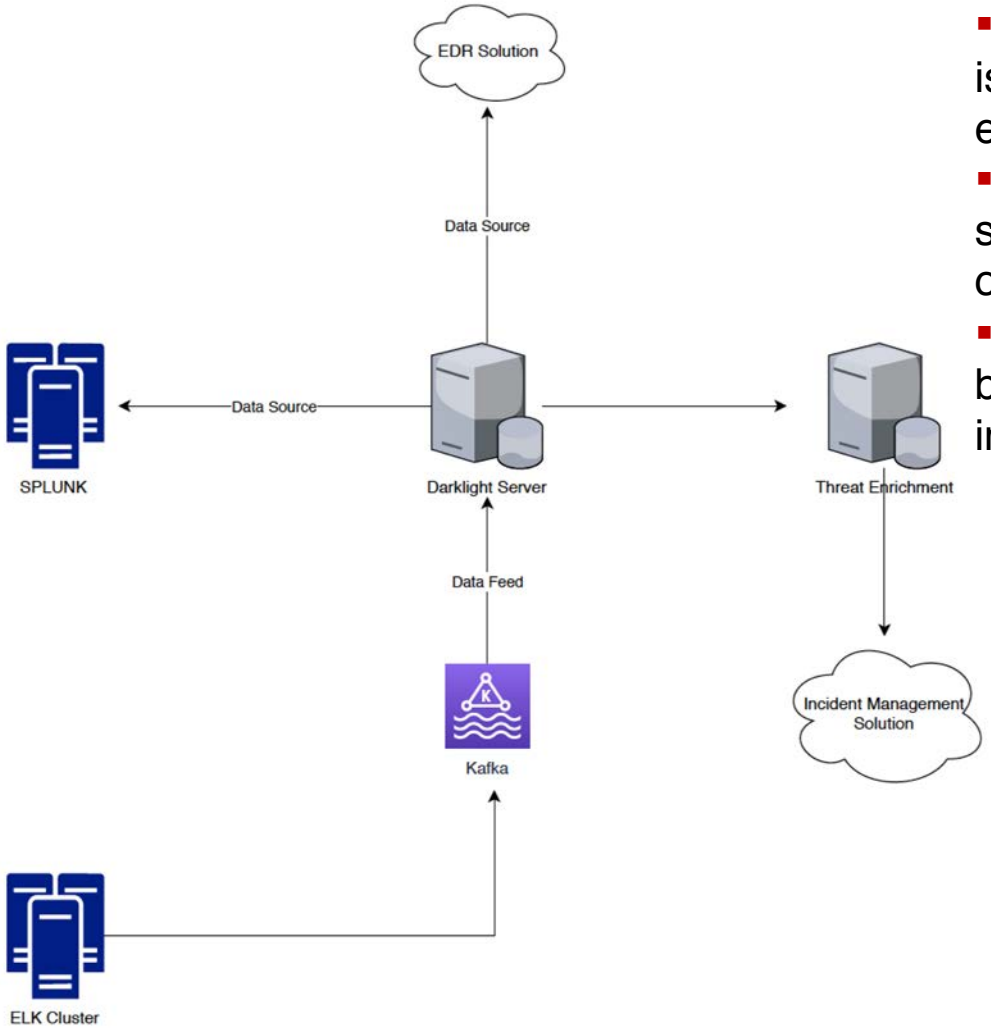
THE PROBLEM WITH SIEM

- Depends on signatures and very basic behavioral analysis algorithm to find attacks
- False positives can be overwhelming and cause Alert Fatigue
- Engineering resources have to be dedicated to maintain and optimize SIEM
- Analysts have to be trained to understand the data and reaction time is slow

Architecture



Architecture



- Pipeline approach leverages automation to enrich and pre-process data
- As DarkLight finds indicators, output is sent to platforms for data enrichment
- From data enrichment, output is then sent to FreshService with DarkLight output and additional information
- Result: faster incident investigation because analyst has all available information

ARCHITECTURE OVERVIEW

- Configuration of DarkLight-1 instance is to receive all data from Logstash for network and process it
  - Kafka receives data from Logstash and queues for DarkLight server to process in real-time
  - Allows for all logs that match specific parameters to be processed but with greater reach for development
- Configuration for DarkLight-2 instance is to query Splunk and pull data to process it.
  - DarkLight reaches into Splunk and runs a query to find matching information.
    - Ex. search EventCode=4625 | fields *
- Once matching data is found from query, pulls data back then processes it through playbooks to match overall techniques such as Brute Force

REALTIME KAFKA FEED

- Logstash is sending data directly to Kafka
- Approximately 100GB a day being processed
- DarkLight is processing the data at line speed
- So what…?
  - Processing time is not stalled at any point for the pipeline
  - DarkLight continuously runs
  - As data is fed in, it is processed to look for key indicators of Techniques being used
  - Outputs information to multiple areas such as threat enrichment platforms, incident management platforms, and messaging platforms.

MDR Use Case

Managed Detection and Response (MDR) brings a requirement to be able to see exactly what is happening on the endpoint.

Faster detection = faster response

Automation using DarkLight speeds detection

Goal: DarkLight direct integration with hunt platform further reduces reaction time

Road Map

- Integrate further external resources for better context and data enrichment
- Bring enriched data into DarkLight for ability to infer further based on context
- Integrate DarkLight with Orion V2 directly for automated endpoint collection

HUMAN-LED. TECHNOLOGY-ACCELERATED.

ROOT9B.COM