Cybersecurity R&D: Where is Technology Taking Us

David Hogue

Deputy Chief

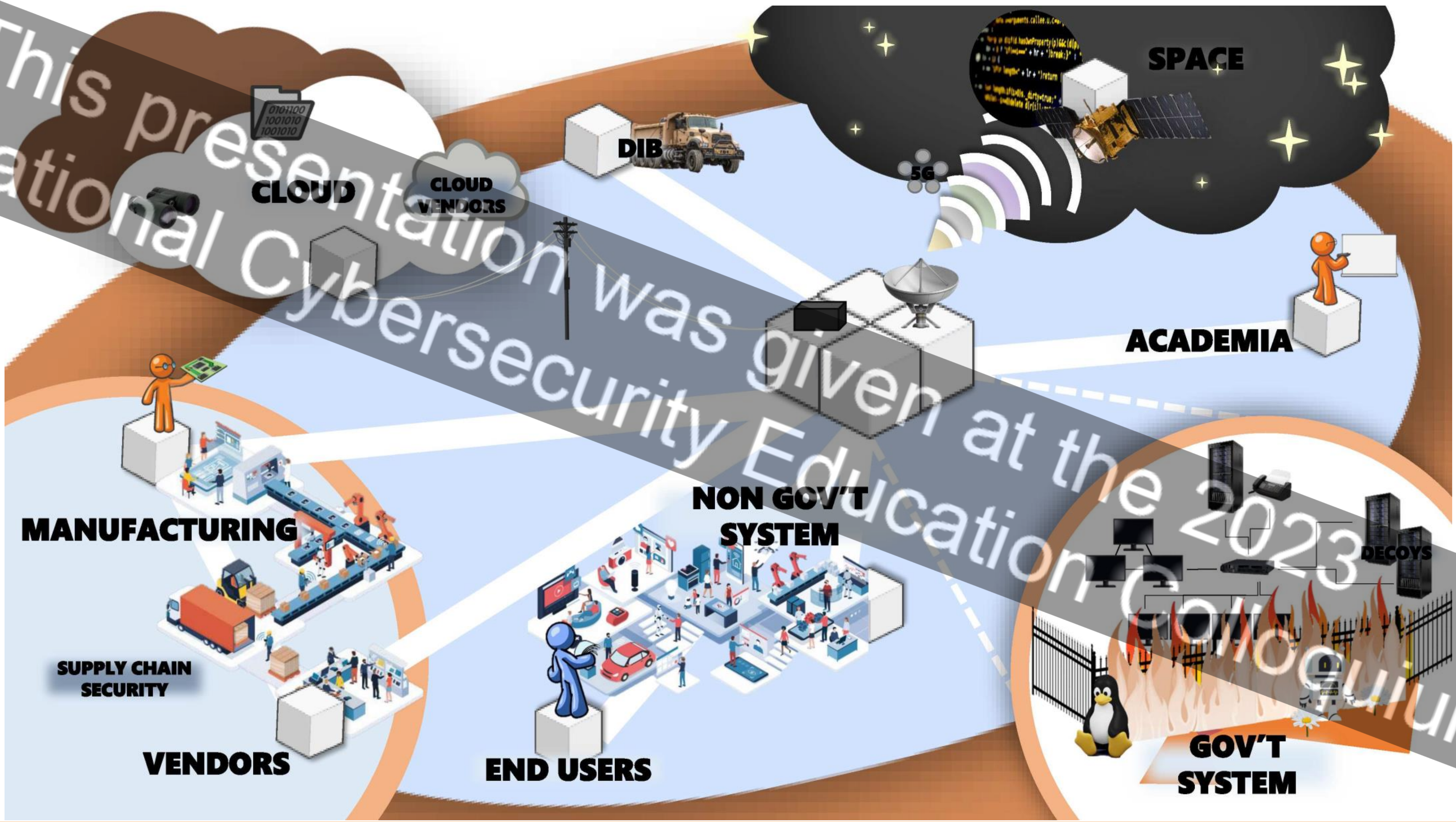
Laboratory for Advanced Cybersecurity Research

National Security Agency

About me…

10 Yrs Private Industry

Cybersecurity Landscape

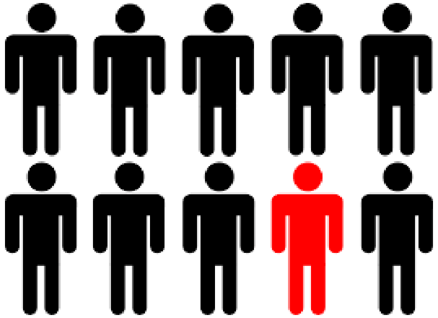


Threat Actors

Hostile Nation-States

Terrorist Groups

Insiders



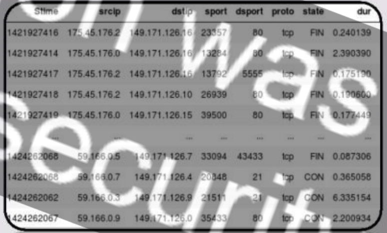
Organized Crime

Hackers

Attack Types

Ransomware

Insider Threat



Stime	srcip	dstip	sport	dsport	proto	state	dur
1421927416	175.45.176.2	149.171.126.16	23357	80	tcp	FIN	0.240139
1421927414	175.45.176.0	149.171.126.16	13284	80	tcp	FIN	2.390390
1421927417	175.45.176.2	149.171.126.16	13792	5555	tcp	FIN	0.175190
1421927418	175.45.176.2	149.171.126.10	26939	80	tcp	FIN	0.190600
1421927419	175.45.176.0	149.171.126.15	39500	80	tcp	FIN	0.177449
...	...	...	...	...	...	...	...
1424262068	59.166.0.5	149.171.126.7	33094	43433	tcp	FIN	0.087306
1424262068	59.166.0.7	149.171.126.4	20848	21	tcp	CON	0.365058
1424262062	59.166.0.3	149.171.126.9	21511	21	tcp	CON	6.335154
1424262067	59.166.0.9	149.171.126.0	35433	80	tcp	CON	2.200934

```
$char_name;
char_name = $name_info;
nction get_class($charid) {
$class_dect = mysql_query("select class
while($data = mysql_fetch_array($class_dec
$class_info = $data['class'];
$class_text = array(1 => "Warrior", 2 => "Paladin", 3
$class_color = array(1 => "#C79C6E", 2=> "#F58CBA",
global $class;
$class = array('class_txt' => $class_text[$class_info],
get_race($charid) {
dect = mysql_query("select race from character.chara
data = mysql_fetch_array($race_dect,MYSQL_ASSOC))
fo = $data['race'];
(1 => "Human", 2 => "Orc", 3 => "Dwarf"
```

timestamp	object	action	command_line	hostname	image_path	parent_image_path	pid	ppid	user
[illegible]	PROCESS	CREATE	NaN	[illegible]	Idle	Idle	[illegible]	[illegible]	NT AUTHORITY\SYSTEM
[illegible]	PROCESS	CREATE	NaN	[illegible]	System	Idle	[illegible]	[illegible]	NT AUTHORITY\SYSTEM
[illegible]	PROCESS	CREATE	[illegible]	[illegible]	[illegible]	System	[illegible]	[illegible]	NT AUTHORITY\SYSTEM
[illegible]	PROCESS	CREATE	NaN	[illegible]	MemCompression	System	[illegible]	[illegible]	NT AUTHORITY\SYSTEM
[illegible]	PROCESS	CREATE	NaN	[illegible]	Idle	Idle	[illegible]	[illegible]	NT AUTHORITY\SYSTEM
...	...	...	...	...	...	...	...	...	...
[illegible]	FLOW	START	NaN	[illegible]	[illegible]	NaN	[illegible]	[illegible]	NaN
[illegible]	FLOW	START	NaN	[illegible]	[illegible]	NaN	[illegible]	[illegible]	NaN
[illegible]	FLOW	START	NaN	[illegible]	[illegible]	NaN	[illegible]	[illegible]	NaN
[illegible]	FLOW	START	NaN	[illegible]	[illegible]	NaN	[illegible]	[illegible]	NaN
[illegible]	FLOW	START	NaN	[illegible]	[illegible]	NaN	[illegible]	[illegible]	NaN

Social Engineering

Phishing

Mal-advertising

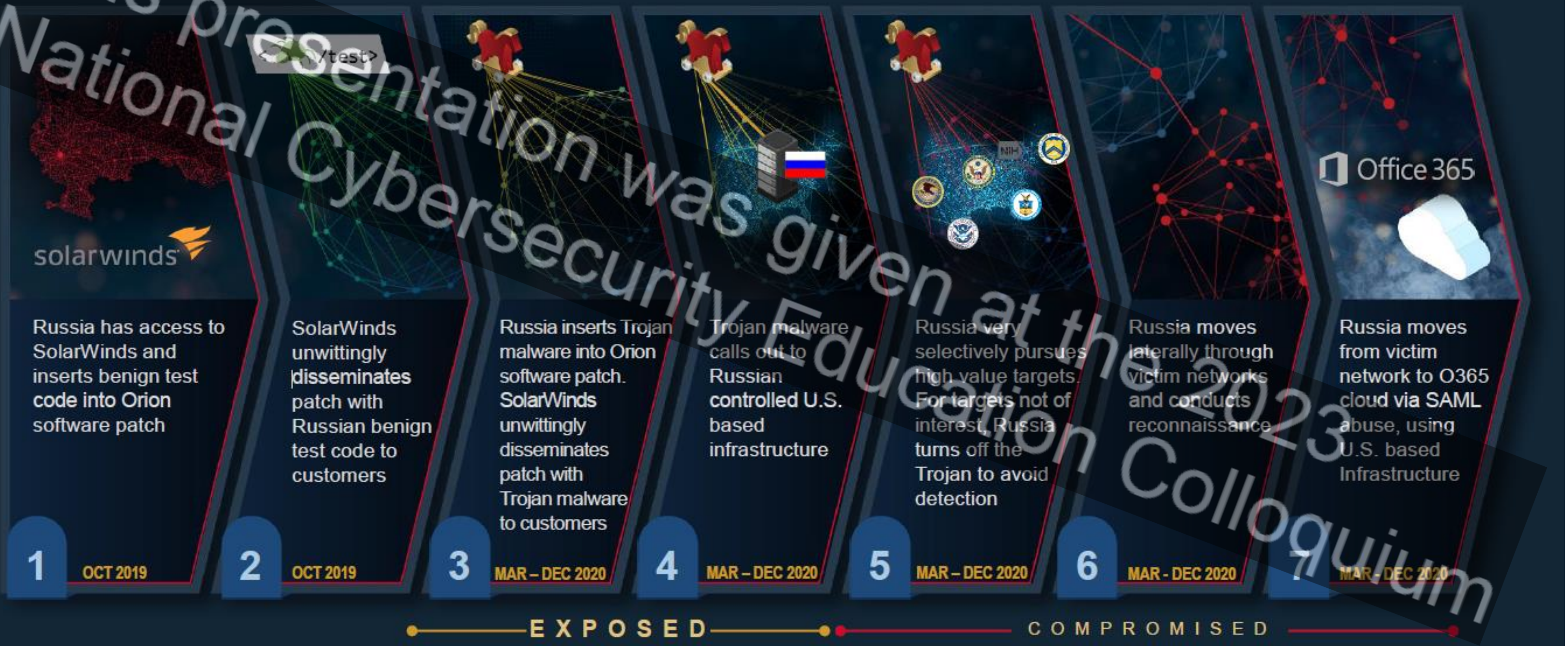
Supply Chain Attacks

DDOS

Zero-Days

SolarWinds: Intrusions Into the USG and Private Sector

The timeline below is based on industry analysis.



1 Russia has access to SolarWinds and inserts benign test code into Orion software patch
OCT 2019

2 SolarWinds unwittingly disseminates patch with Russian benign test code to customers
OCT 2019

3 Russia inserts Trojan malware into Orion software patch. SolarWinds unwittingly disseminates patch with Trojan malware to customers
MAR – DEC 2020

4 Trojan malware calls out to Russian controlled U.S. based infrastructure
MAR – DEC 2020

5 Russia very selectively pursues high value targets. For targets not of interest, Russia turns off the Trojan to avoid detection
MAR – DEC 2020

6 Russia moves laterally through victim networks and conducts reconnaissance
MAR - DEC 2020

7 Russia moves from victim network to O365 cloud via SAML abuse, using U.S. based Infrastructure
MAR - DEC 2020

EXPOSED

COMPROMISED

Impacts of Cyber Threats

Economic

- Lost Productivity
- Ransoms
- Identity Theft
- Intellectual Property Theft

National Security

- Compromised employees / increased insider threat
- Compromised systems & weapons
- Lost advantage in military & diplomatic missions

Cybersecurity Defense Mechanisms

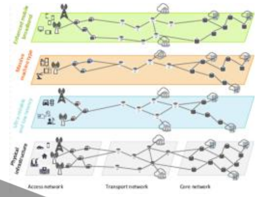
- Standards
- Zero Trust Design
- Supply Chain Risk Management
- Vulnerability Discovery

- AI for Cybersecurity
  - Threat Discovery
  - Autonomous Defense
- Cybersecurity for AI
- Cyberpsychology

Standards

Contributing to standards and closing vulnerabilities

Network Slicing



Standards & Open Source

3GPP A GLOBAL INITIATIVE

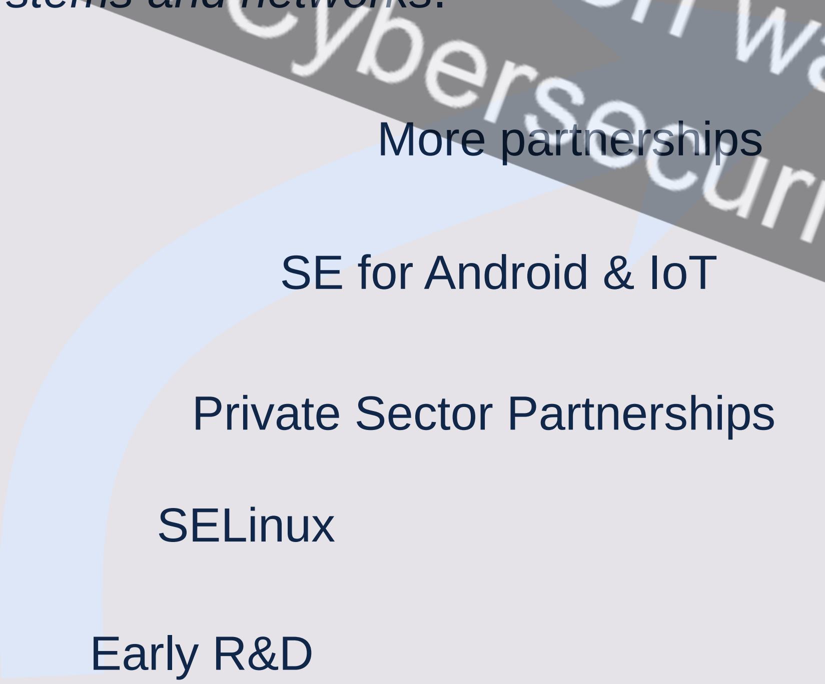
ONAP OPEN NETWORK AUTOMATION PLATFORM

O-RAN ALLIANCE

Advanced Security Topics

Trust Mechanisms

Platform and system security architectures and mechanisms that advance the security and assurance of computing systems and networks.



More partnerships

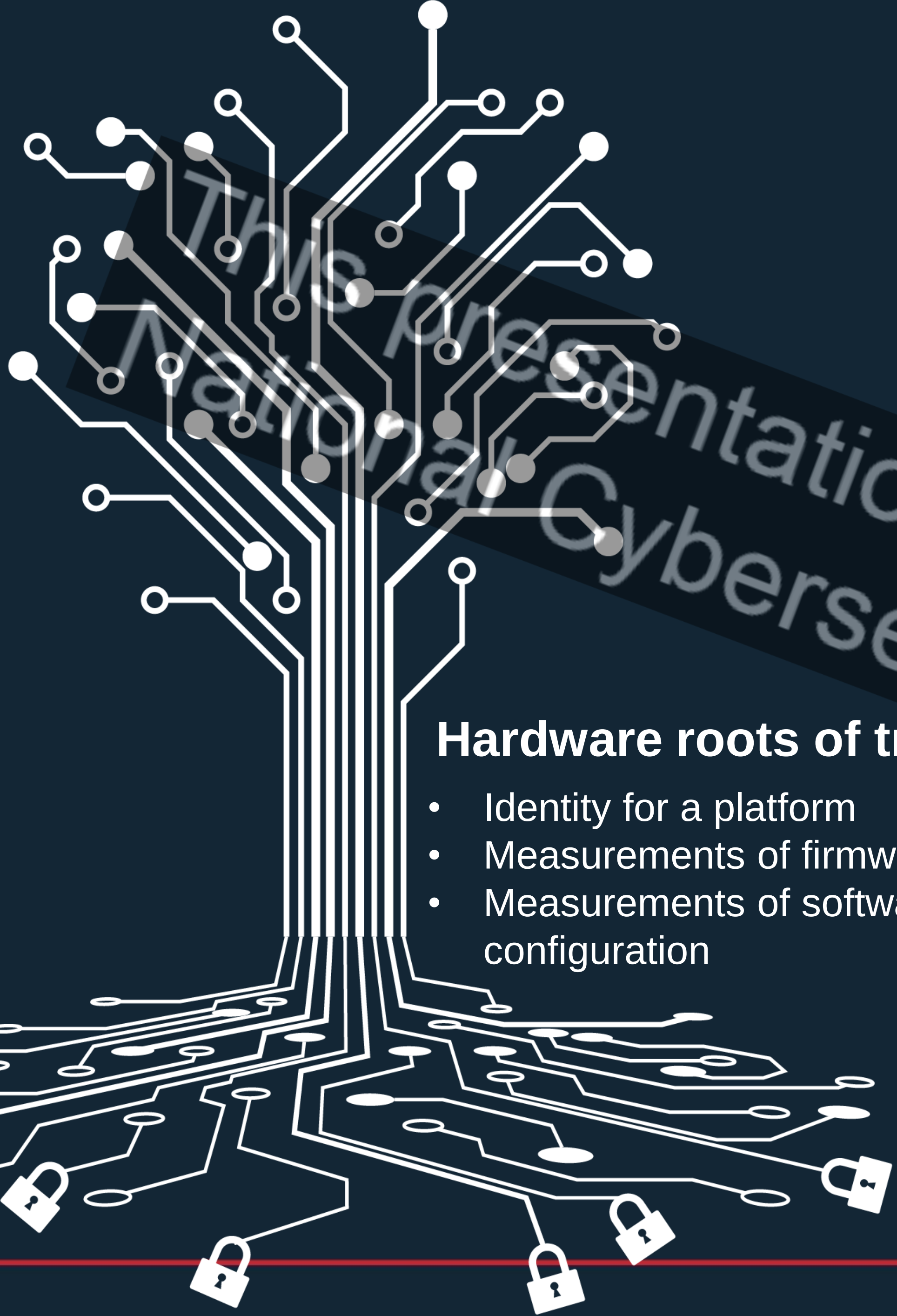
SE for Android & IoT

Private Sector Partnerships

SELinux

Early R&D

Trusted Computing



Hardware roots of trust provide

- Identity for a platform
- Measurements of firmware
- Measurements of software and configuration

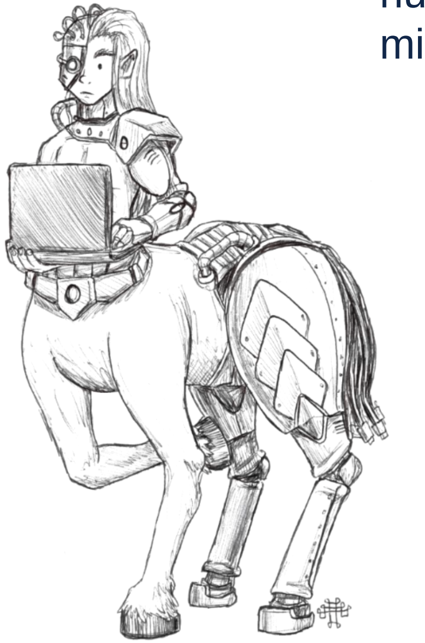
Confidential Cloud Computing

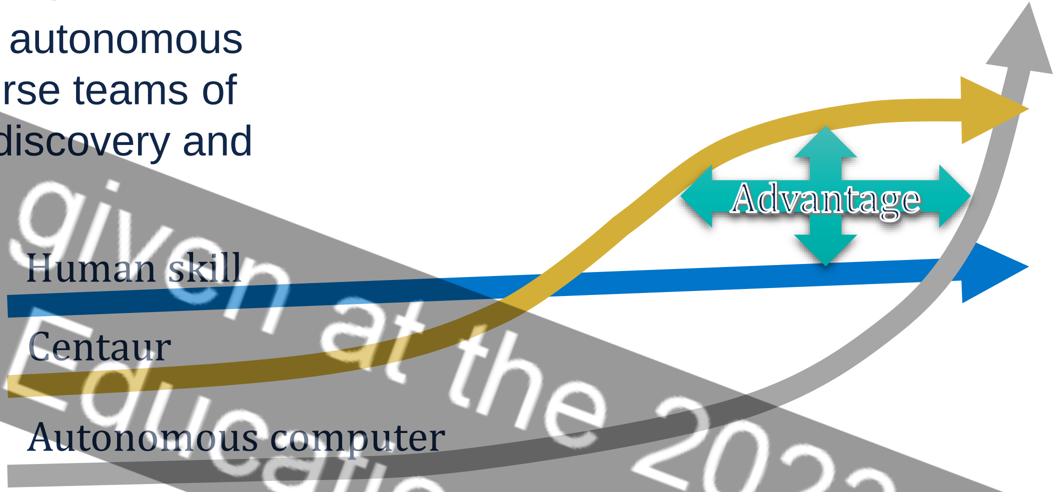
- Trusted Execution Environments
- Processor based encryption and isolation
- Protects sensitive information from other software
- Uses Attestation to confirm software and hardware configurations

Vulnerability Discovery

Achieve Maneuverability in Cyberspace

Researching tools and workflows around autonomous technologies working in concert with diverse teams of humans to enable software vulnerability discovery and mitigation at scale.





Integrate new advances in cyber autonomy

- Discovery of flaws
- Proof of vulnerability
- Automated patching
- Severity grading

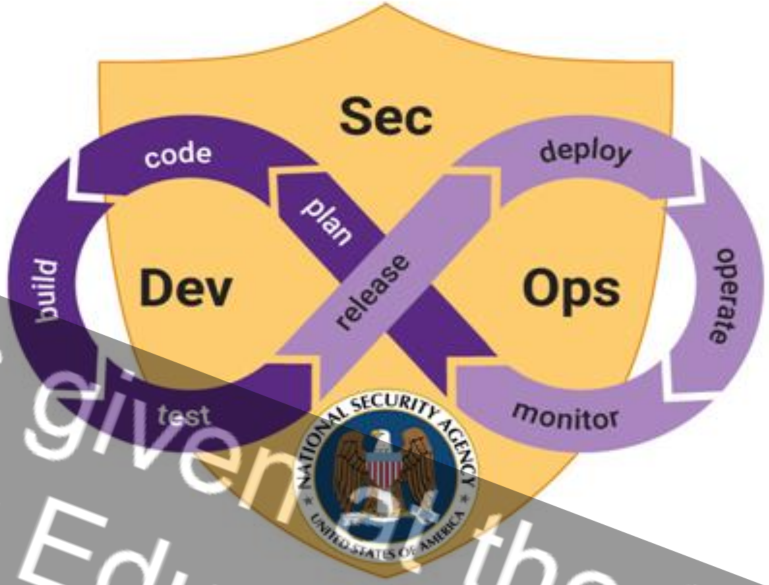
Achieve scalability and efficiency

DevSecOps

Create more trust in systems during design.

Ease the adoption of high confidence software development tools and pipeline

Cryptol



Ruby

Swift

> *“NSA advises organizations to consider making a strategic shift from programming languages that provide little or no inherent memory protection, such as C/C++, to a memory safe language when possible.”*
>
> Cybersecurity Information Sheet, November 2022

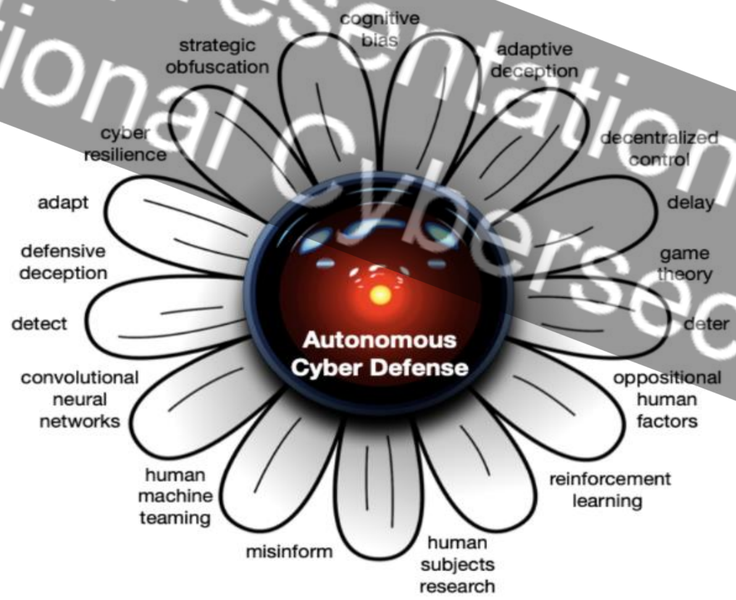
Cyberpsychology: Deception for Cyber Defense

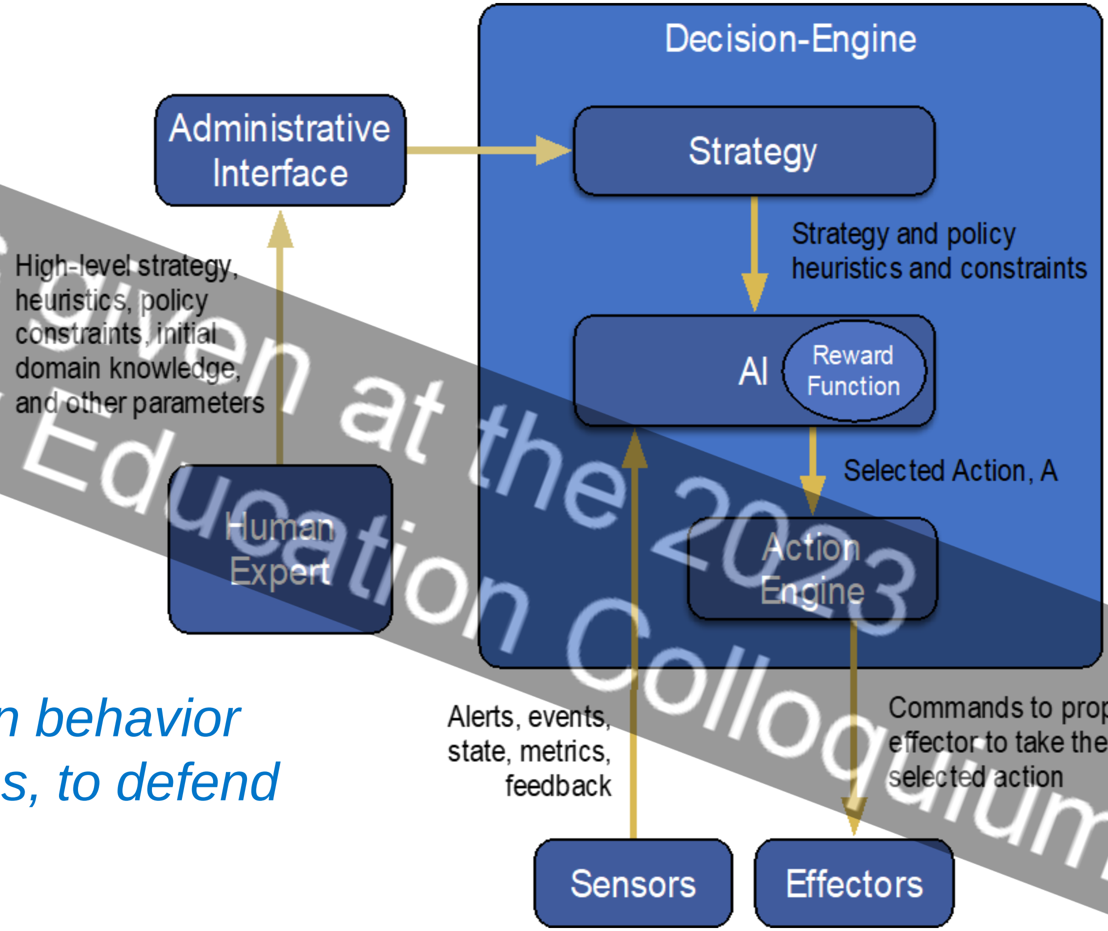
GOAL: Rebalance asymmetric nature of cyber defense

- Attacker only knows what is perceived through observation
  - Computers **unintentionally** reveal to an attacker more information than we desire
  - System owner can control what is revealed to the attacker
- Cyber deception plays on an attacker's cognitive bias and cognitive load to:
  - Control what an attacker knows about the network
  - Influence their behavior
  - Increase the workload of the attacker
  - Decrease the workload of the defender

Employ cyber deception to confuse, frustrate, delay, and deter attacker.

Adaptive Deception for Cyber-Defense




Decision-Engine
Administrative Interface
Strategy
Strategy and policy heuristics and constraints
AI
Reward Function
Selected Action, A
Action Engine
High-level strategy, heuristics, policy constraints, initial domain knowledge, and other parameters
Human Expert
Alerts, events, state, metrics, feedback
Commands to proper effector to take the selected action
Sensors
Effectors

Adaptive deception for defense, informed by human behavior science and powered by trustworthy AI technologies, to defend against cyber-attacks at machine speed and scale.

Cybersecurity, Artificial Intelligence (AI), and Machine Learning (ML)

AI for Cybersecurity

Speed

- Analysis in minutes versus weeks/months
- System response faster than human response

Scale

- Reasoning over large data sets
- Recognition of patterns that humans cannot even describe

Cybersecurity for AI

(Unsafe at any) Speed

- Trusting poorly-designed AI products could be disastrous
- Cyber decisions will increasingly be driven by auto-derived models

Scale

- Opaque modern AI models are becoming ubiquitous
- All stages in the AI pipeline can be attacked

AI/ML for Cybersecurity

TODAY

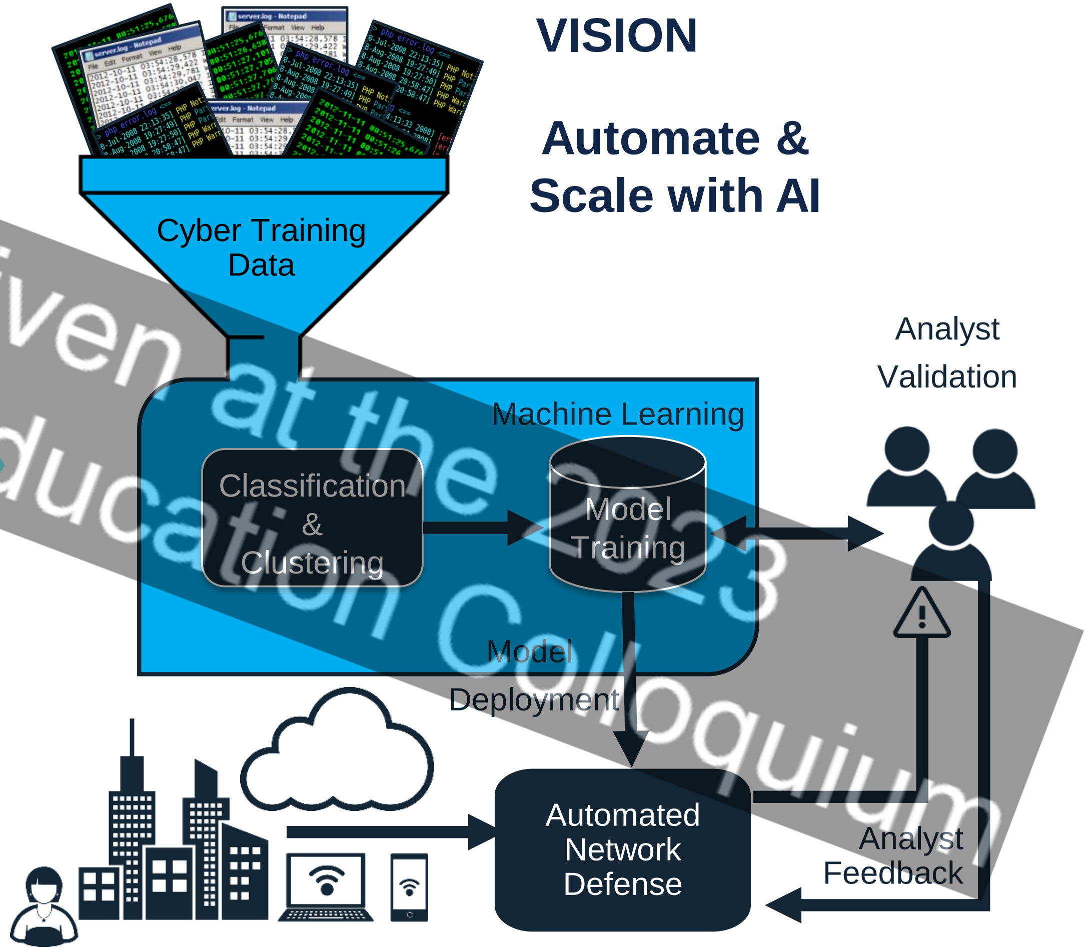
Human-Driven Response

Slow, Subject to Data Overload

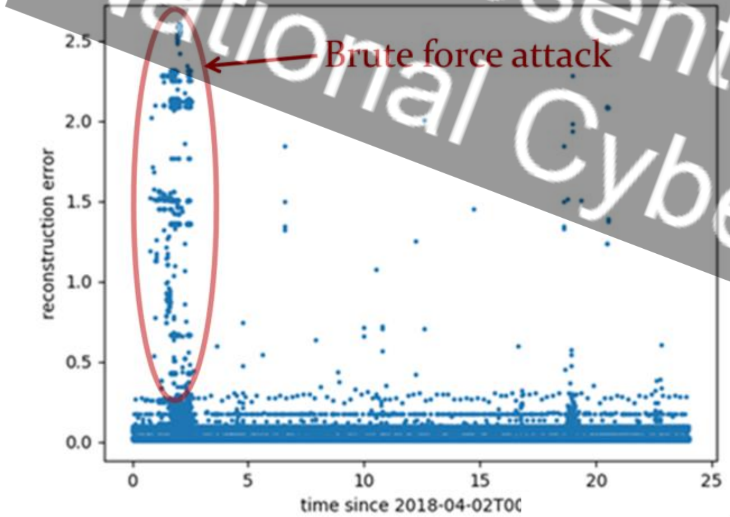
Research Question:
Can AI/ML improve the quality and speed of cyber incidence detection, response and mitigation?

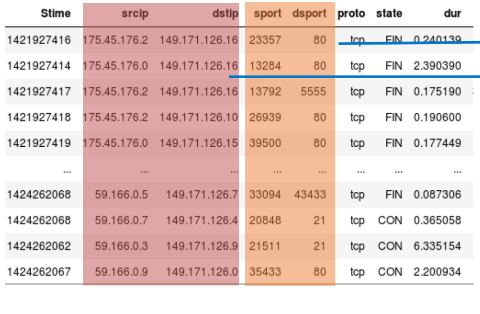
VISION

Automate & Scale with AI

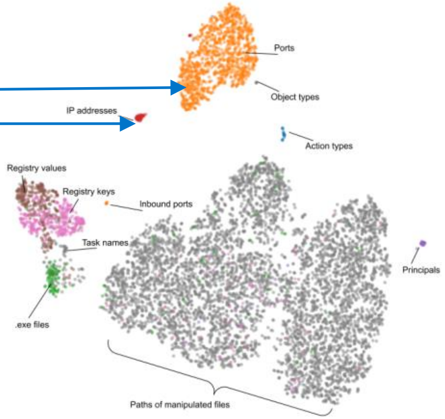


Anomaly Detection using AI/ML





Stime	srcip	dstip	sport	dsport	proto	state	dur
1421927416	175.45.176.2	149.171.126.16	23357	80	tcp	FIN	0.240139
1421927414	175.45.176.0	149.171.126.16	13284	80	tcp	FIN	2.390390
1421927417	175.45.176.2	149.171.126.16	13792	5555	tcp	FIN	0.175190
1421927418	175.45.176.2	149.171.126.10	26939	80	tcp	FIN	0.190600
1421927419	175.45.176.0	149.171.126.15	39500	80	tcp	FIN	0.177449
...	...	...	...	...	...	...	...
1424262068	59.166.0.5	149.171.126.7	33094	43433	tcp	FIN	0.087306
1424262068	59.166.0.7	149.171.126.4	20848	21	tcp	CON	0.365058
1424262062	59.166.0.3	149.171.126.9	21511	21	tcp	CON	6.335154
1424262067	59.166.0.9	149.171.126.0	35433	80	tcp	CON	2.200934



- Early deep learning prototypes detect real attack
- Later prototypes model APT and defender decision calculus and dynamics.
  - Cross-disciplinary approaches: multi-agent RL, co-evolutionary computation, and game theory.

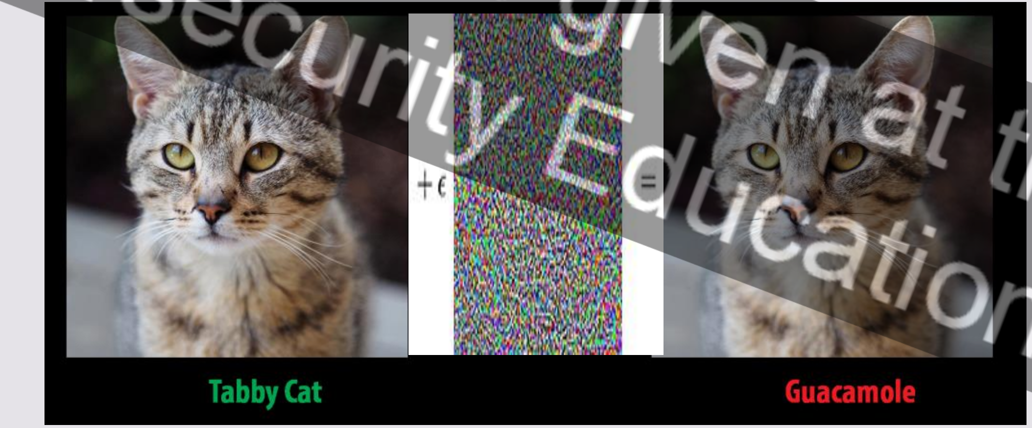
Cybersecurity for AI

Research Question : Can we secure our AI/ML models from attack?

AI/ML CHALLENGES

- Sufficiency of training data
- Model drift
- Reliability and security
- Explainability
- Model training on streaming data
- Multi-modal data fusion (e.g. events and content)

Adversarial Attack Tailored "Background Noise"



Model Evasion

TYPES OF ATTACKS

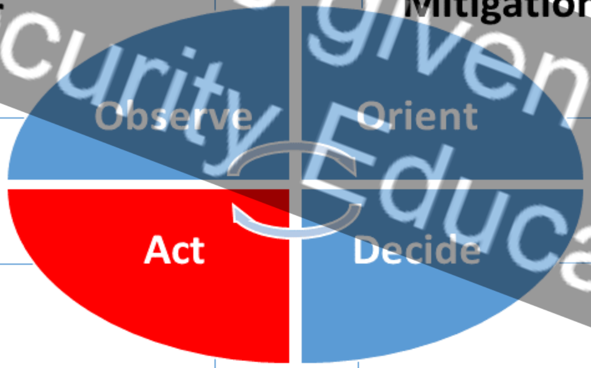
- Data Poisoning
- Model Evasion
- Modeling Stealing
- Inversion

Pre-Emptive Mitigation against ML Attack

Model:

- Decision Support User
- ML Defender
- Adversary

- Simulate Attack
- Develop and Test Mitigation



- Modify Behavior

- Deploy and Evaluate Mitigation

Cyberpsychology at the Intersection

Identify individual and group differences that relate to attackers' behavior and susceptibility to influence

Information Environment

Determine effectiveness of persuasive messaging, cyber defense strategies, and tactics that <u>influence adversary behavior</u>

Discover patterns of cyber behavior

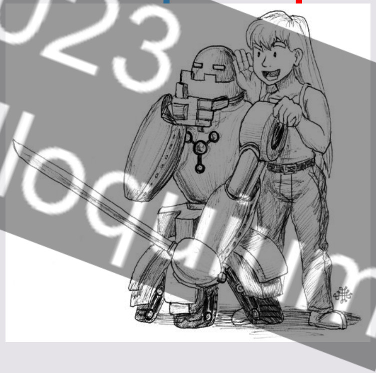
Cyber Environment

Cyber mitigations and response options

Adversary

***Research Question**: Can we apply psychological science to disrupt and frustrate cyber attackers progress and advance defenders' success?*

And inform research in most effective cybersecurity defense strategies



Defender

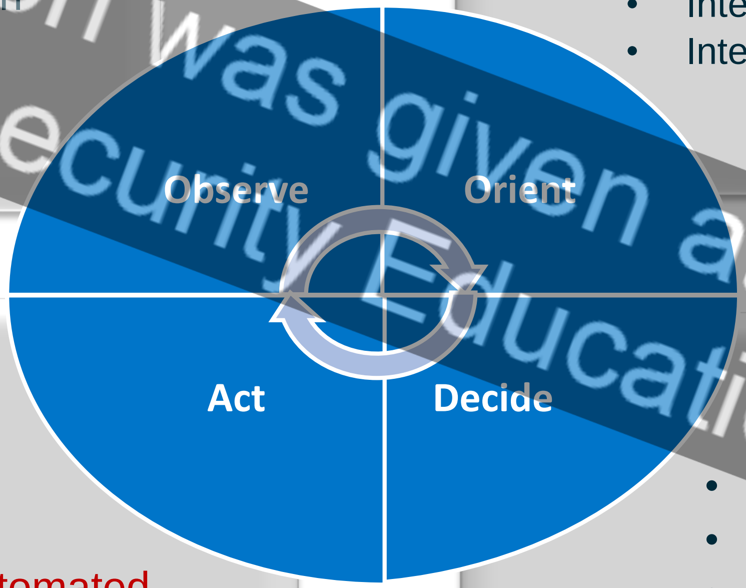
Towards Autonomous Cyber Defense

- Dynamic and adaptive data collection
- Exploratory data analysis
- Unsupervised pattern recognition

Example: Reinforcement Learning (RL) agents for adaptive sensor placement; Unsupervised machine learning for anomaly detection

- Data fusion and enrichment
- Contextualization
- Human – Machine Teaming
  - Integrated feedback
  - Interpret patterns

Example: Machine Learning & cyber-deception for prioritizing cyber alerts



Response
- Planning
- Orchestration
- Execution

Example: Security Orchestration & Automated Response (SOAR) tools for implementing cyber response, i.e., editing permissions, disabling services, or disconnecting devices

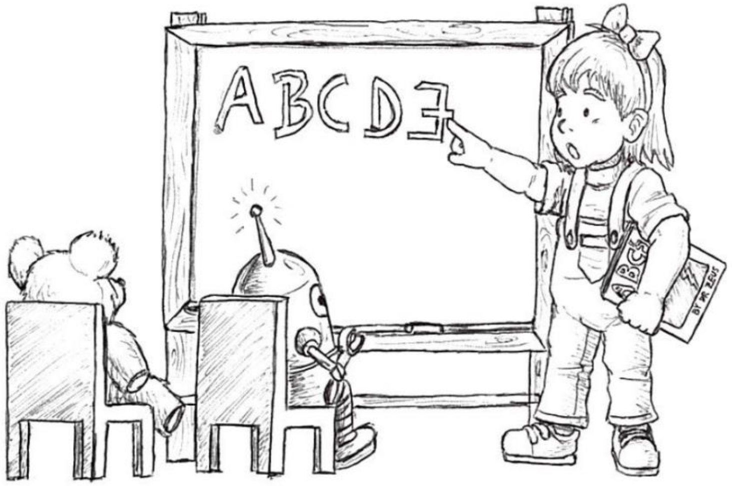
- Reasoning
- Impact analysis
- Strategy analysis and selection
- Explore response space
- Response selection
- Human – Machine Teaming

Example: AI planning and/or RL agents for reasoning and response selection

Transformational Research:
Neuro-Symbolic AI for Cybersecurity

(U) "The next decade of AI research will likely be defined by efforts to incorporate existing knowledge, push forward novel ways of learning, and make systems more robust, generalizable, and trustworthy."…"For example, neuro-symbolic research is combining symbolic manipulation with neural networks."

National Security Commission on AI, Final Report, 2021



Thank You!