midwest architecture community collaboration

AI CLASSIFICATION OF THREAT ACTORS

NOVEMBER 9, 2023

MACC 2023: THE FUTURE OF AI AND SECURITY

Who is this presenting?

Ryan Hohimer
Associate Ontologist
Phone: (509) 430-6890
ryan.hohimer@semanticarts.com
https://www.semanticarts.com

- Electrical Engineer
- Knowledge Representation and Reasoning (KR&R) Zealot
- Object Oriented Programmer and Project Manager at Pacific Northwest National Laboratory for decades
- Founded a Knowledge Graph Cybersecurity company based on KR&R software he invented at the lab.
- Now: An Ontologist at Semantic Arts!

Classification with Artificial Intelligence

The Problem

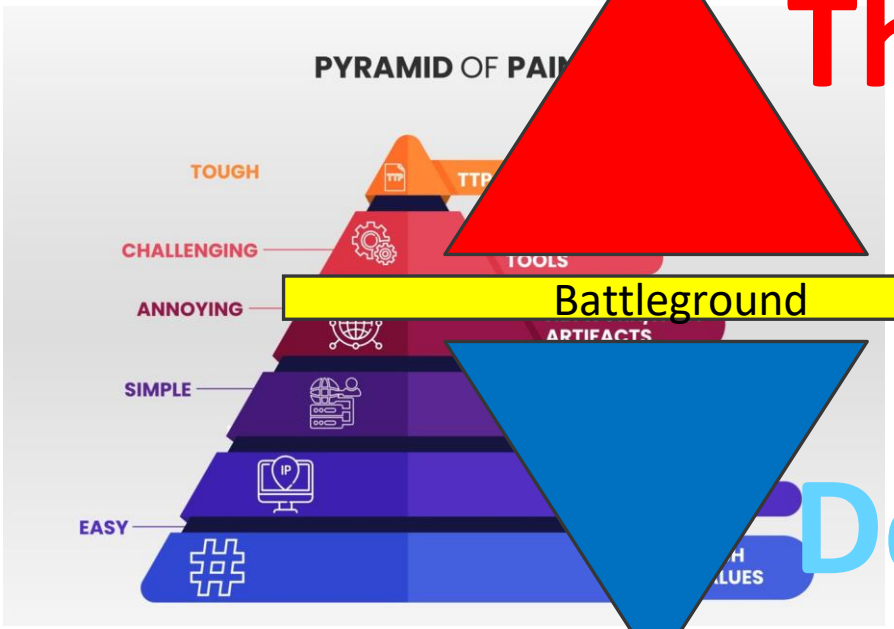
- Broad Range of Knowledge Required (Multiple Diverse Domains)
- Specialized Knowledge and Logic Required (Expert Humans)

The Right Automation (Tools) for the Job

- Data Centric Architecture (DCA)
- DCA empowering Artificial Intelligence (AI)

Use Case: Curious Healthcare Worker

Attackers, Defenders, and the Cyber Terrain



Threat Actors

Battleground

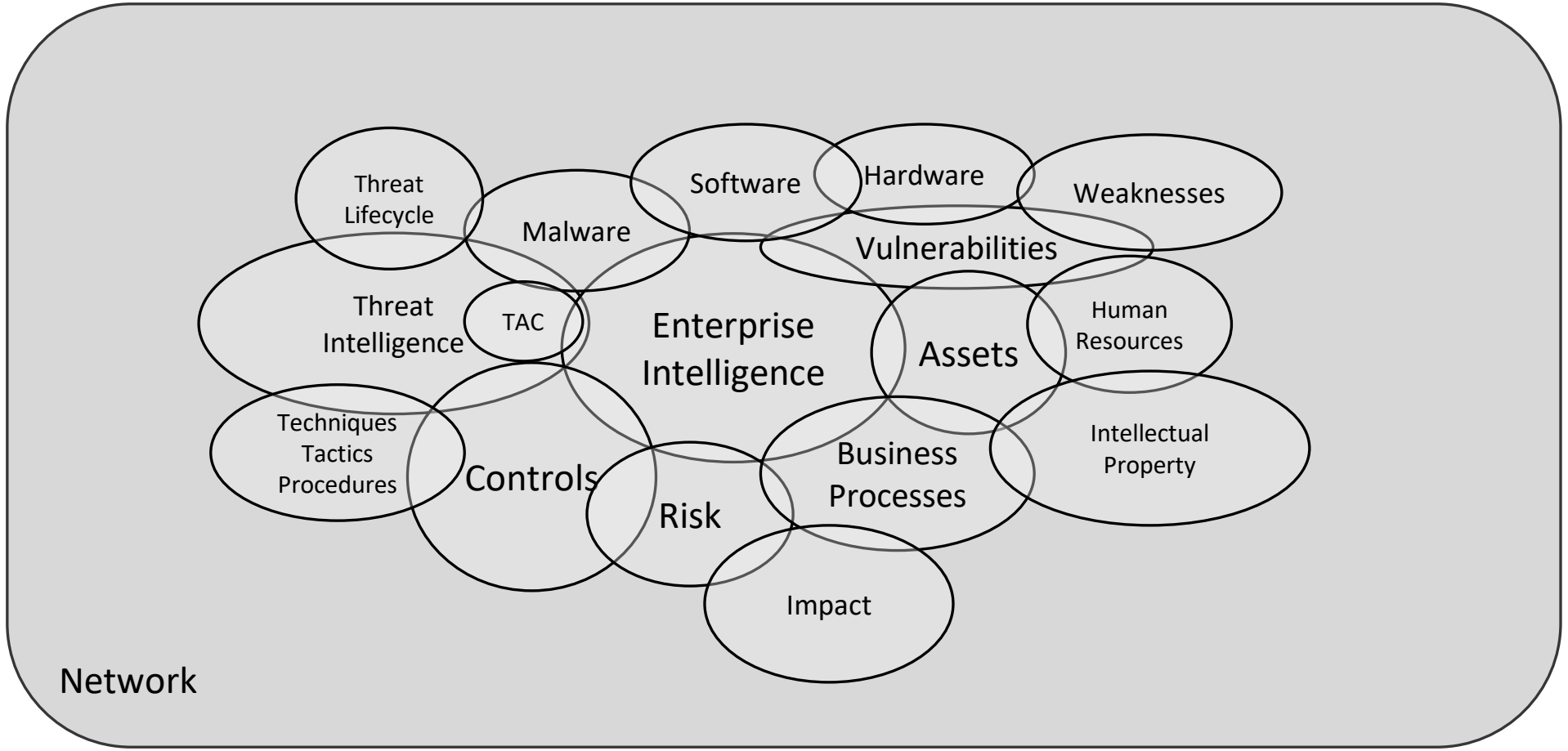
Defenders

If you know the enemy and know yourself, you need not fear the result of a hundred battles.

- Sun Tzu

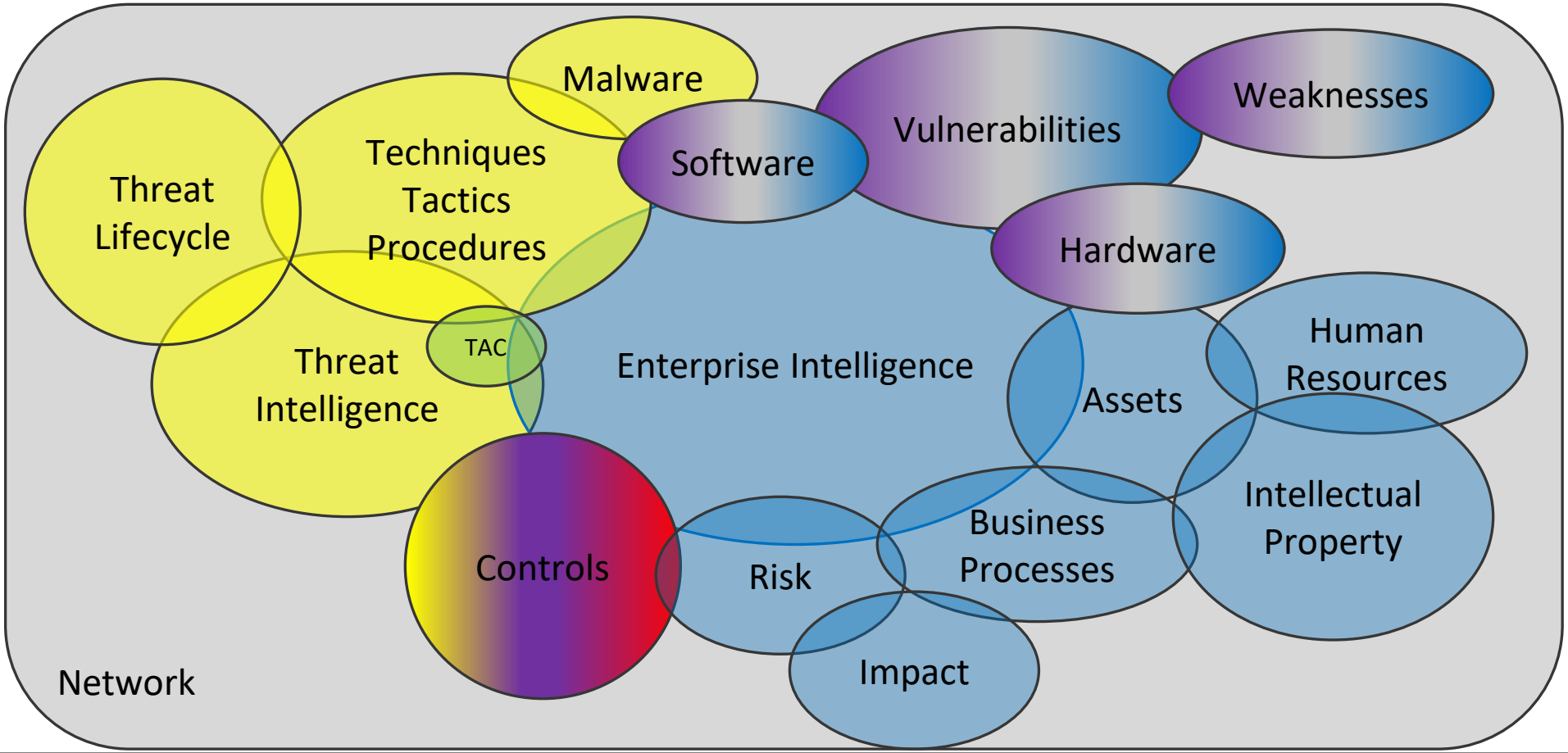
https://cyware.com/security-guides/cyber-threat-intelligence/the-concept-of-pyramid-of-pain-f358

Disparate Cyber Domains



Cyber Domains

Identify | Protect | Detect | Respond | Recover



Many Resources of models and data

ATT&CK®

CIS Center for Internet Security®
Creating Confidence in the Connected World.

OASIS OPEN

STIX

National Institute of Standards and Technology (NIST)
- Many publications on Cybersecurity
- Cybersecurity Framework

MITRE
- ATT&CK (Adversary Tactics, Techniques & Common Knowledge)

National Security Agency (NSA)
- Cybersecurity Lifecycle

Center for Internet Security (CIS)
- The CIS Critical Controls

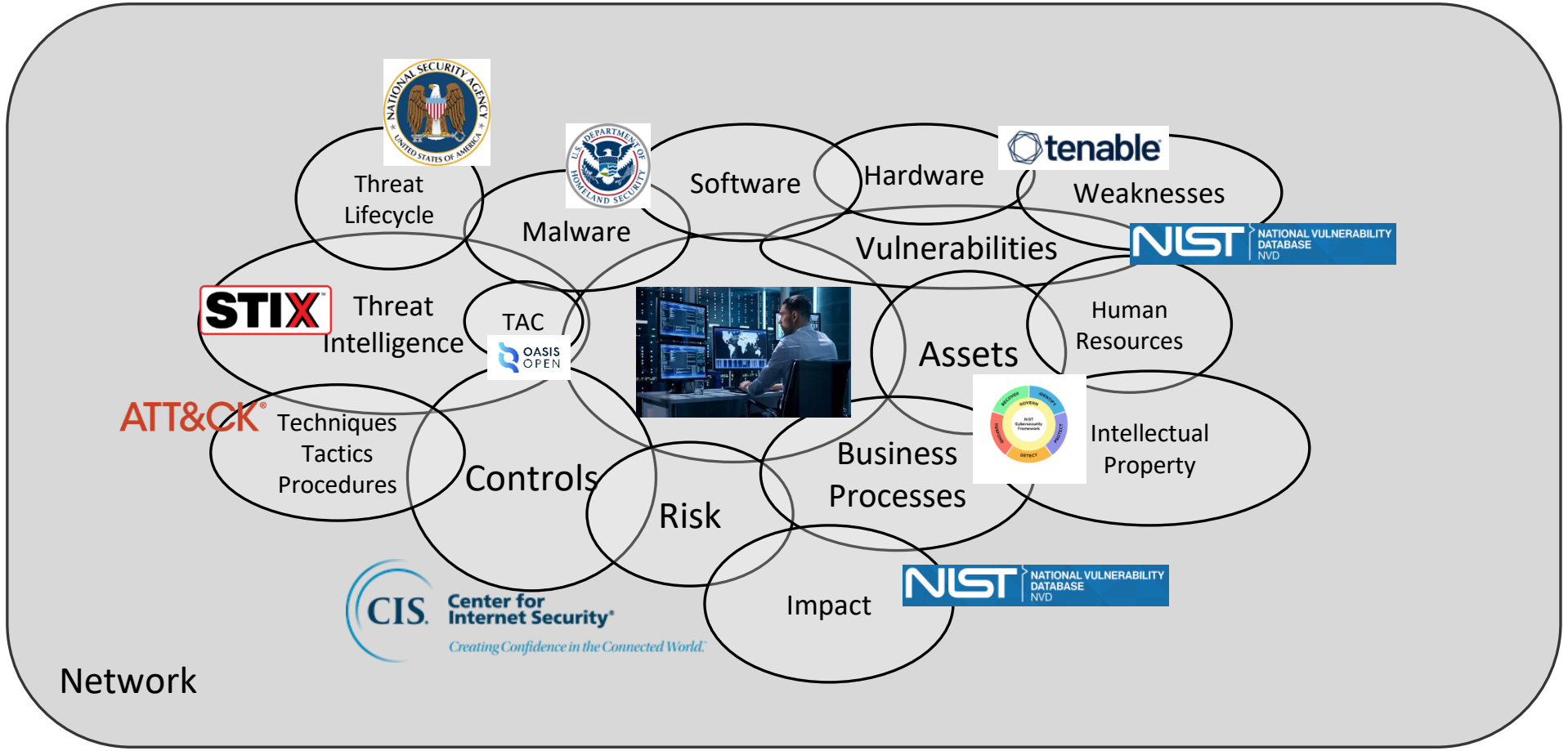
Department of Homeland Security

Organization of Advanced Structured Information Systems (OASIS)
- STIX & TAC

Human Cybersecurity Analysts

And of course … Vendors Galore

Human Readable – We Need democratized machine readable



Data Centric Architecture Foundational to Artificial Intelligence

Bringing in all together in a data fabric

"Semantic knowledge graphs are the underlying framework for the ability to seamlessly connect to, access, and query all data sources relevant to the enterprise. This capability includes sources internal and external to organizations, in any type of cloud setting, on-premises, or at the cloud's edge. The first way semantic knowledge graphs enable a uniform fabric across each of these environments, tools, and technologies is by furnishing a layer harmonizing the semantics between them." – Jans Aasman

Jans Aasman – "The Foundation of Data Fabrics and AI: Semantic Knowledge Graphs"

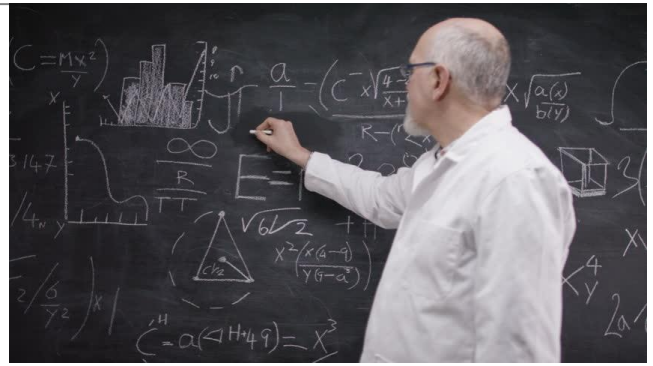
- https://www.datasciencecentral.com/the-foundation-of-data-fabrics-and-ai-semantic-knowledge-graphs/

AI = Mimicry of Human Intelligence

REAL INTELLIGENCE – The Very Smart People in High Demand

- They know the enemy
- They know the defender
- They know the cyber terrain (cyber theater)

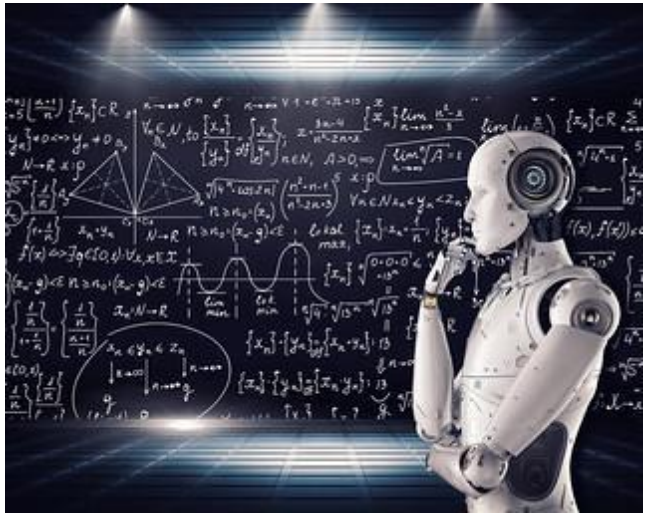
Mental Models



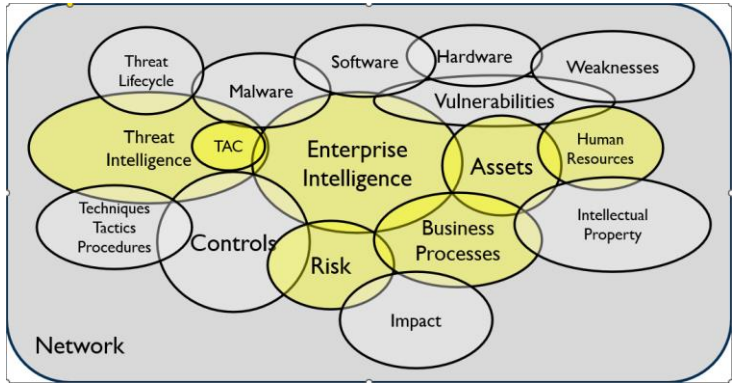
ARTIFICIAL INTELLIGENCE – Mimicry of Those Very Smart People

- They have a machine-readable form of the enemy knowledge
- They have a machine-readable form of the defender knowledge
- They have a machine-readable form of the cyber terrain knowledge

Ontology Models



USE CASE: AI Classification of Threat actors



This use case cuts across the multiple domains of the cyber terrain. Knowledge from one domain is not adequate.

A unified upper ontology of the cybersecurity domains enables an Enterprise Data Fabric. The Enterprise Data Fabric enables AI classification.

In this case, the definition of a Threat Actor in the context of *YOUR* enterprise.

The Cyber Terrain

The Cyber Terrain

Models of Cybersecurity

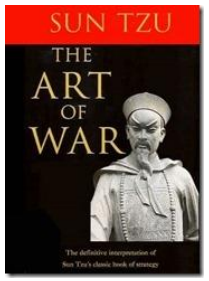
The talented cybersecurity expert has knowledge about multiple facets of the problem space. They hold mental models of how things work together. The cybersecurity community is composed of experts in multiple disciplines and multiple specialized models.

It is inconceiveable that a single model can adequately cover the whole of cybersecurity. The Cyber Terrain Ontology brings together a collection of ontologies that describe in greater detail the bigger picture of cyber situational awareness. It is an ontology alignment project that resolves ambiguity that can exist between the separate ontologies being integrated together.

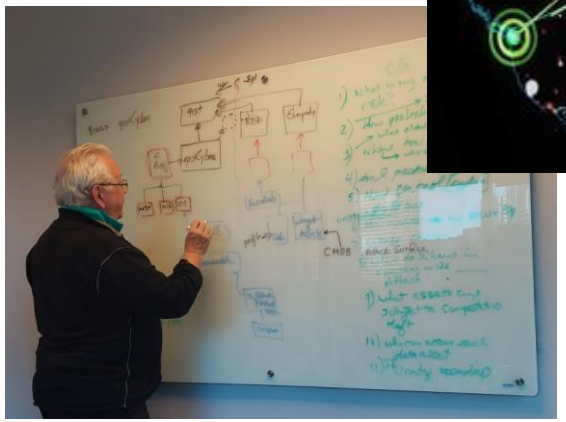
Essential Pieces to the Cybersecurity Puzzle

The Cyber Terrain contains:

- Adversary
- Defender
- Asset
- Vulnerability
- Risk
- TTP
- Controls



gist Cyber



semantic arts

Building on Common Knowledge and Standards

Analysts' logic is embedded into ontologies to support automation

Healthcare Threat Agent Ontology

← **Enterprise-specific** Threat Actor Knowledge
(new subclasses of TAs relevant to Enterprise)

Threat Agent Library Ontology

← General **Threat Actor defining characteristics**
(basis for being a subclass member)

Threat Actor Context Ontology

← **Classification** of Threat Actors
to provide context (allows classification)

STIX 2.1 Ontology STIX.OWL

← **Semantic graph** representation
of STIX specification

STIX 2.1 Specification

← STIX 2.1 foundation
(data exchange)

Files enabling automated intake of the ontologies listed above are available on the TAC TC GitHub

semantic arts

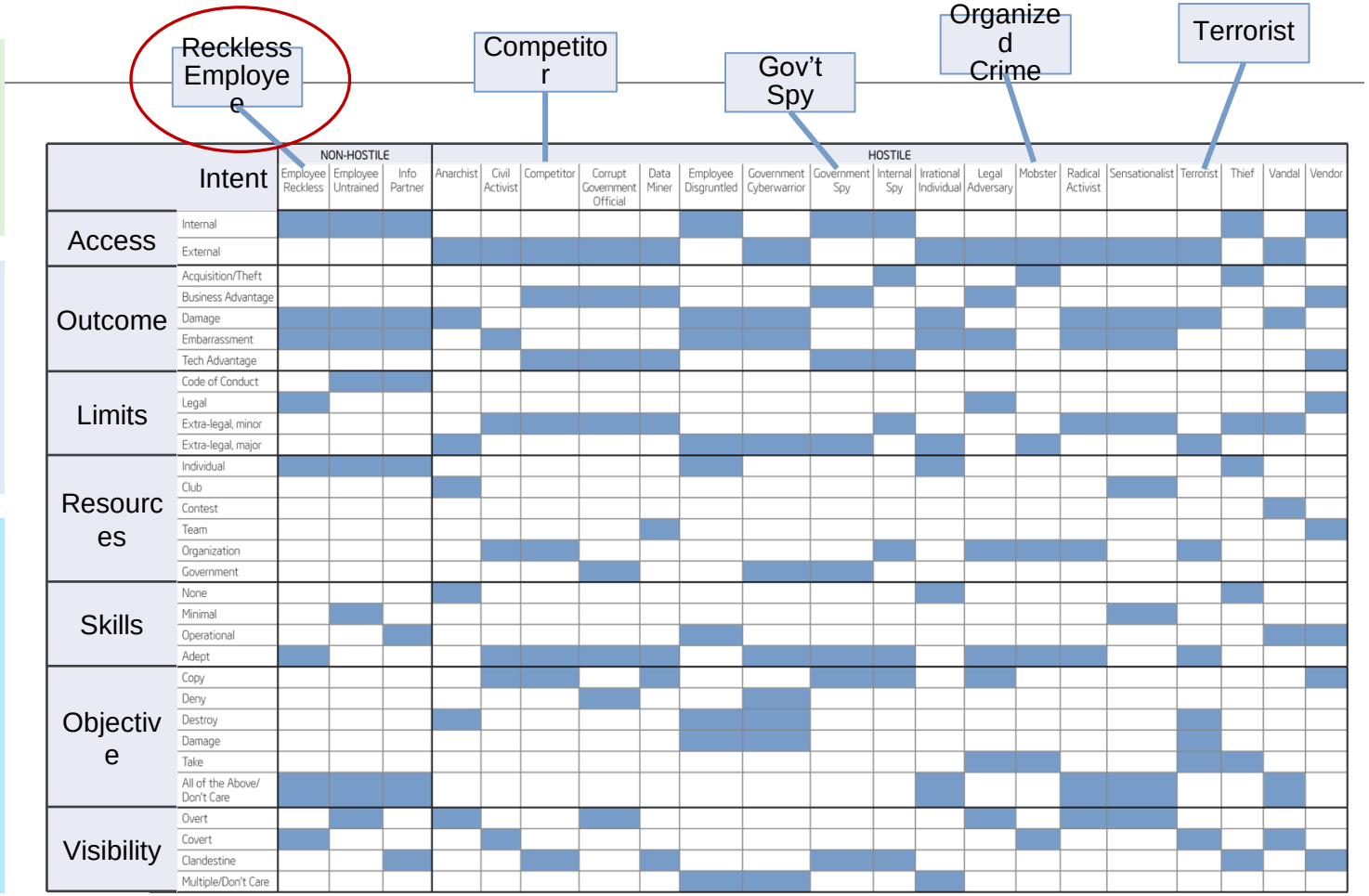
Defining the Characteristics of Threat Actors

Threat Actors can be **classified based on the set of defining characteristics** they possess.

STIX can *list* Threat Actor attributes like their objective and resources, **but interpretation is left to human analysts**.

This method **captures your internal expertise** in machine readable form **to facilitate**:

- Intelligence Automation
- Onboarding Experienced Analysts
- Training New Analysts



Intel Threat Agent Library http://dx.doi.org/10.13140/RG.2.2.30094.46406

Logic, not just Data

Classifying the Curious Healthcare Worker

- stix:StixObject
  - stix:StixCyberObservable
  - stix:StixDomainObject
    - stix:AttackPattern
    - stix:Campaign
    - stix:CourseOfAction
    - stix:Grouping
    - stix:Identity
    - stix:Incident
    - stix:Indicator
    - stix:Infrastructure
    - stix:IntrusionSet
    - stix:Location
    - stix:Malware
    - stix:Note
    - stix:ObservedData
    - stix:Opinion
    - stix:Report
    - stix:ThreatActor
      - tal:Anarchist
      - tal:CivilActivist
      - tal:Competitor
      - tal:CorruptGovernmentOfficial
      - tal:DataMiner
      - tal:GovernmentCyberWarrior
      - tal:GovernmentSpy
      - tal:InformationPartner
      - tal:InternalSpy
      - tal:IrrationalIndividual
      - tal:LegalAdversary
      - tal:Mobster
      - tal:RadicalActivist
      - tal:RecklessEmployee
        - hal:CuriousHealthcareWorker

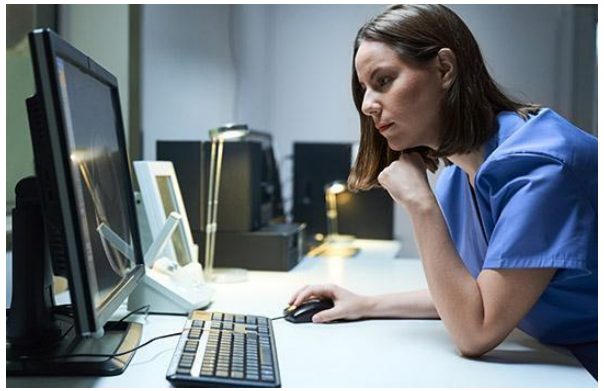
stix: StixObject

stix: StixDomainObject

stix: ThreatActor

tal: RecklessEmployee

hal: CuriousHealthcareWorker



The pace and volume of data is too overwhelming for human analysts. TAC enables automated analysis based on cybersecurity expert's knowledge and logic embedded in the graph

Description: tal:RecklessEmployee

Equivalent To

stix:ThreatActor
and (tac:categorizedBy value tal:AdeptSkills)
and (tac:categorizedBy value tal:AllDontCareObjective)
and (tac:categorizedBy value tal:CovertVisibility)
and (tac:categorizedBy value tal:DamageOutcome)
and (tac:categorizedBy value tal:EmbarrassmentOutcome)
and (tac:categorizedBy value tal:IndividualResources)
and (tac:categorizedBy value tal:InternalAccess)
and (tac:categorizedBy value tal:LegalLimits)

Description: hal:CuriousHealthcareWorker

Equivalent To

tal:RecklessEmployee
and (tac:categorizedBy value hal:EnjoymentObjective)
and (tac:categorizedBy value hal:PolicyViolationOutcome)
and (tac:categorizedBy value tal:PersonalStatisfaction)

Enabling Automated Knowledge Graph Analysis

STIX 2.1 JavaScript Object Notation (JSON) file

{JSON}

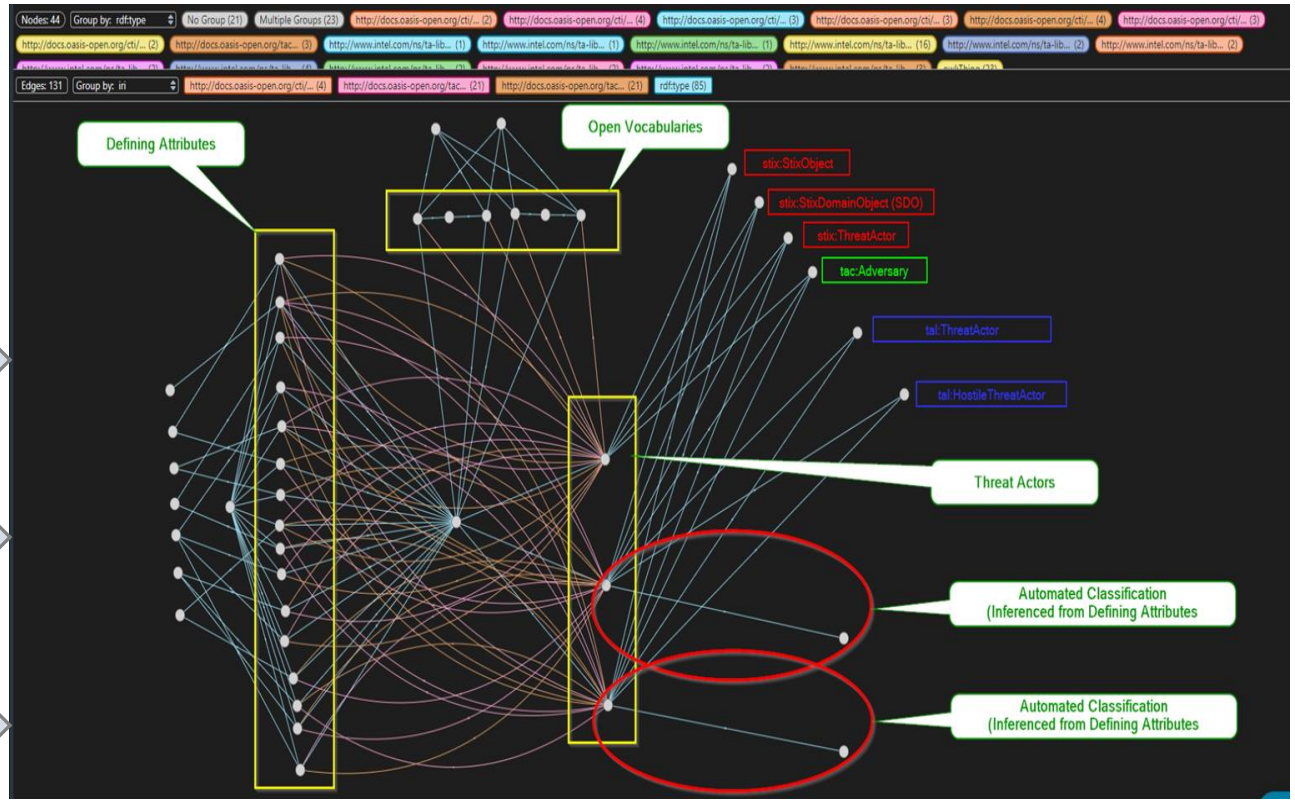
STIX 2.1 Ontology Web Language (OWL) file

Threat Actor Context Ontology OWL file

Threat Agent Library Ontology OWL file

Healthcare Threat Agent Ontology OWL file

Screenshot of a Knowledge Graph executing the logic to automatically infer the classification of Threat Actors via a Description Logics Reasoner.



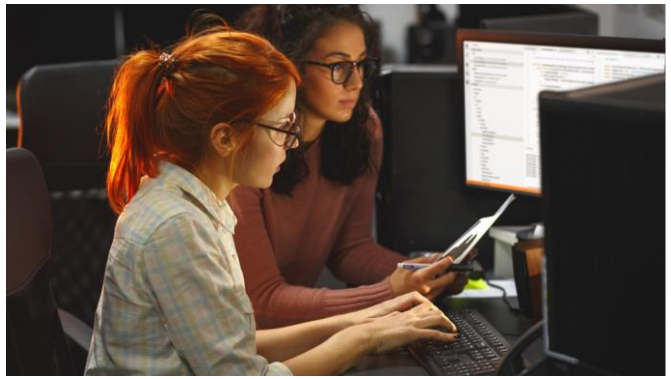
These OWL files are available on the TAC TC GitHub

And here's the Ronco Cordless Electric!

Threat Actor Descriptions

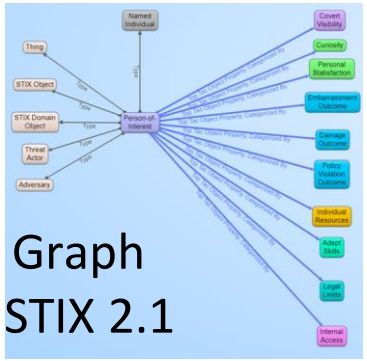
Defining Characteristics Lineup

Human Readable,
Machine Readable,
&
Automatable!



Capture Analyst Tradecraft

Prevent Braindrain, Accelerate Onboarding



Semantic Graph
Version of STIX 2.1

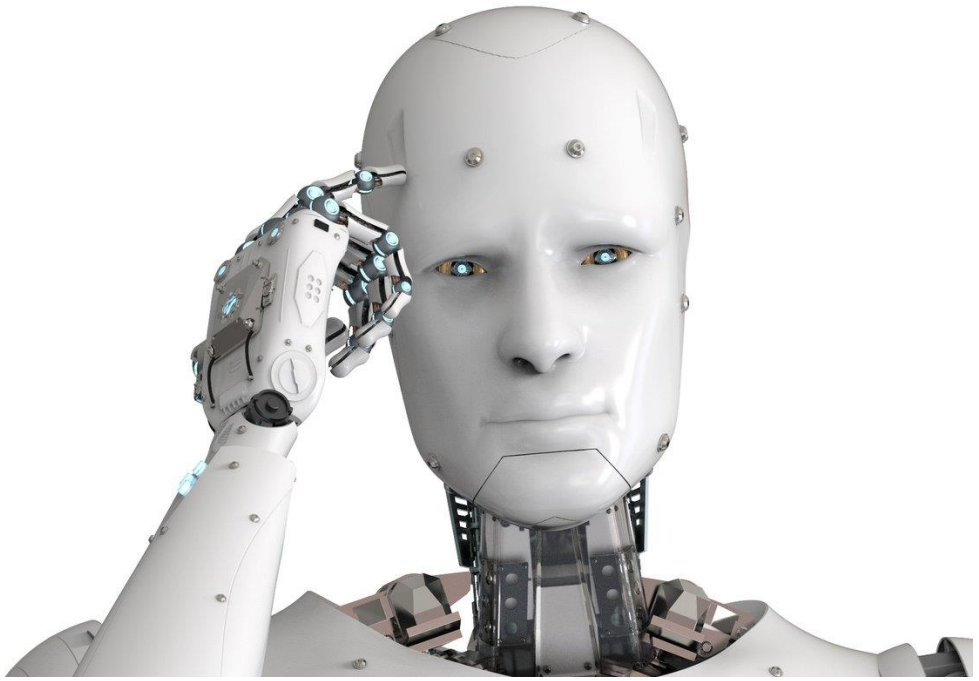
It's Not

TAC TC GitHub https://github.com/oasis-open/tac-ontology

Use Case: Recognizing Policy Violations

https://www.semanticarts.com/case-study/investment-bank-analyzing-employee-violations-to-investigate-culture-and-conduct/

Interest in Data-Centric Enablement of AI ?



Join us in the development of gist-Cyber!

Web: www.semanticarts.com

Phone: (970) 490-2224

Email:

- Ryan.Hohimer@semanticarts.com
- Steve.Case@semanticarts.com

Or

- info@semanticarts.com

semantic arts

MACC OVERVIEW

The Midwest Architecture Community Collaboration's (MACC) purpose is to bring all domains of architecture together to share information and techniques of interest to all of us. It is our shared belief that through collaboration, we can better understand and promote the significance of architecture to business success.