Georgia Tech Research Institute

T&E of Autonomous Systems: Challenges and Opportunities of Autonomy Vulnerability T&E

Don Strausberger

Georgia Tech Research Institute

March 2017

Purpose

Drive thought/discussion in the Test Community towards innovative solutions to evaluate vulnerability in the context of autonomous cyber-physical system T&E

Agenda

- Example
- Definitions and Terminology
  - Automation/Autonomy
  - Cyber and Cyber Physical Systems
  - What is autonomy vulnerability?

- CPS Decomposition/Vulnerabilities/Mitigation Solution
- A-CPS Decomposition/Vulnerabilities/Mitigation Solution
- Key Takeaways
- Summary

CPS: Cyber-Physical System
A-CPS: Autonomous Cyber-Physical System

Example - Autonomous-CPS Sea Vessel

Problem –How will autonomy execute the mission, and as necessary, safely navigate COLREGS?

Real-World Scenario

Human Captain

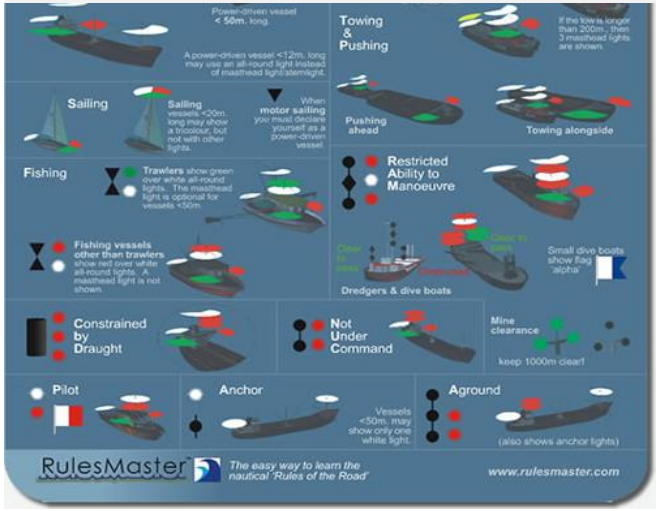
replaced by

Autonomy

- **Perceive**
- **Comprehend**
- **Project**
- **Decide**

World Model

COLREGS



Perceive
Are the other Captains aware of my presence?
What are the vessels current status?

Comprehend
Are they acting in a cooperative, uncooperative, unaware, or adversarial manner?

Project
What is their tactical intent?
What is their strategic intent?
How can I gather more information?
How can I improve the outcome?
What are the options? What is the cost and risk?

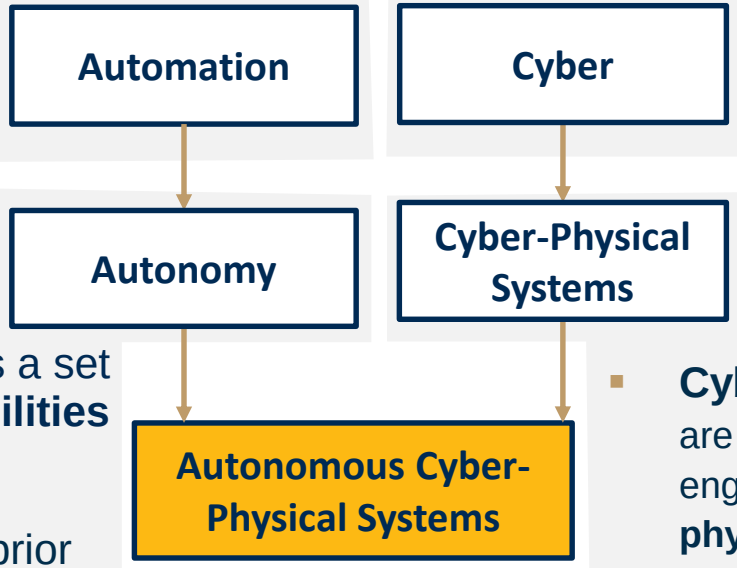
Vulnerabilities could exist at all levels of autonomy.

World Model
Is it expected to encounter fishing boats this hour of the day, this time of the year, in this geographic area?

Origins of Autonomous Cyber-Physical Systems

- **Automation*** The system functions with no/little human operator involvement; however, the system performance is **limited to the specific actions** it has been designed to do. Typically these are **well-defined tasks** that have **predetermined responses**.
  - **Rule-based responses**

- **Cyber**: Relating to or involving computers or computer networks…



- **Autonomy* -**The system has a set of **intelligence-based capabilities** that allow it to **respond** to situations that were **not pre-programmed** or anticipated prior to deployment. Autonomous systems have a degree of **self-government** and self-directed behavior.
  - **Decision-based responses**

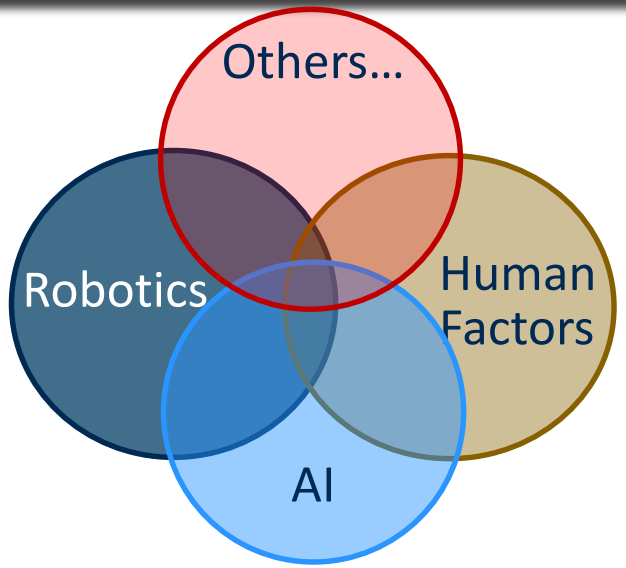
*As defined in the DoD Autonomy Community of Interest (COI) Test and Evaluation Verification and Validation Working Group Technology Investment Strategy 2015-2018

- **Cyber-physical systems (CPS)** are smart systems that include engineered interacting networks of **physical and computational** components. These **highly interconnected and integrated** systems provide new functionalities to improve quality of life and enable technological advances in critical areas… **(NIST)**

Autonomy Terminology

Georgia Tech Research Institute

Different disciplines have developed different terminology in closely related areas



Terminology does not map one-to-one across disciplines, however they can be associated with the OODA loop to provide a common reference.

Goal-based
Rule Based

Perception
Comprehension
Projection

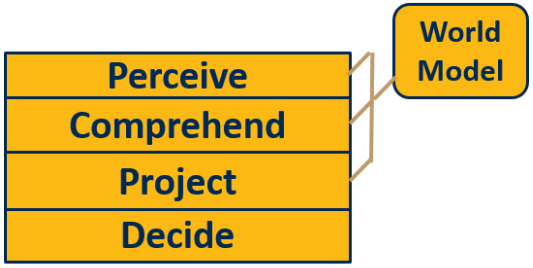
Beliefs
Desires
Intent

Reflective
Deliberative
Reactive

Situational Awareness
Situational Understanding
Situational Assessment

A Model of Autonomy

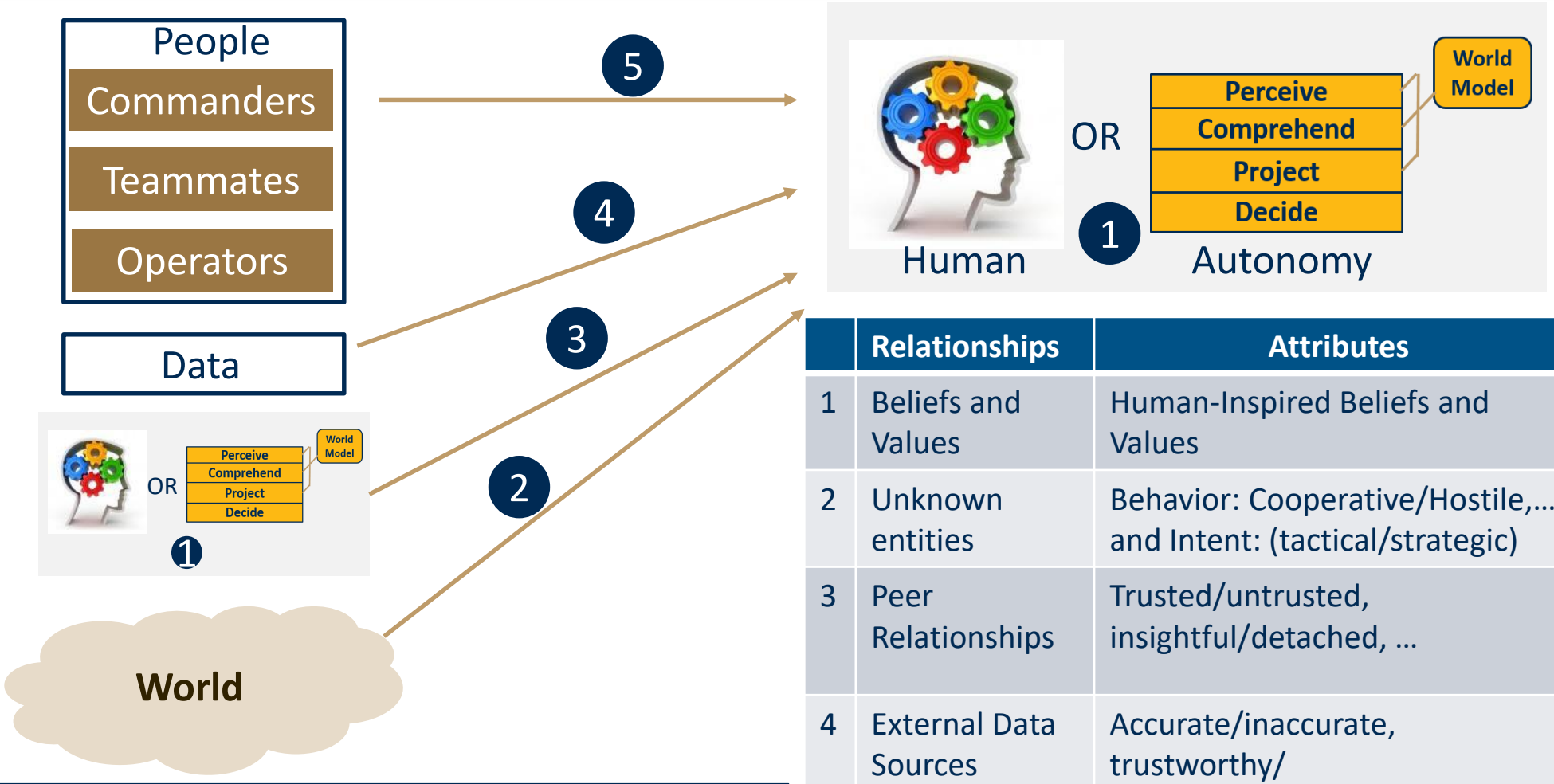
- **Perception**[1]
  - **Identification, state, and attributes of relevant objects in a scene**
- **Comprehension**[1]
  - **Understanding, ordering, and relevance of what was perceived (i.e. "cause and effect", correlation, etc.)**
- **Projection**[1]
  - **Assessing possible courses of action and potential outcomes**
- **World Model**
  - **Past and present beliefs – example: our "world model" of our daily commute allows us to decide if a long line at a traffic light is "normal" or if an event (i.e. accident) has occurred. Our world model grows with experience.**



Autonomy vulnerability is the result of as *past and present* external *influences* that *unduly alter* the *current and future* world model, perception, comprehension, and projection abilities of the autonomy, which may result in changes to current and future *decisions*.[2]

1 - M. R. Endsley and D. J. Garland, *Situation awareness analysis and measurement*: CRC Press, 2000.
2 - Authors definition

Many Paths to Influence Decision Making

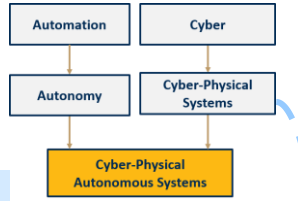


	Relationships	Attributes
1	Beliefs and Values	Human-Inspired Beliefs and Values
2	Unknown entities	Behavior: Cooperative/Hostile,… and Intent: (tactical/strategic)
3	Peer Relationships	Trusted/untrusted, insightful/detached, …
4	External Data Sources	Accurate/inaccurate, trustworthy/ Dated/current Biased/unbiased
5	Support Team	Uncertain/confident, experienced/inexperienced, engaged/uninvolved

Vulnerabilities may result from:
1. Altering the information.
2. Altering its temporal, spatial, or geo-temporal attributes.
3. Altering the attributes of relationships of one or more sources of information in the decision-making process.

A Cyber Physical System (CPS) Decomposition

Georgia Tech Research Institute



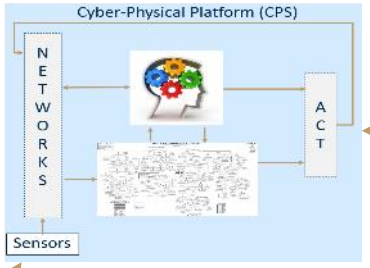
People

Commanders

Teammates

Operators

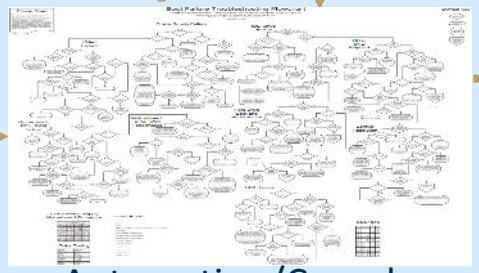
Data



NETWORKS

Cyber-Physical System (CPS)

NETWORKS



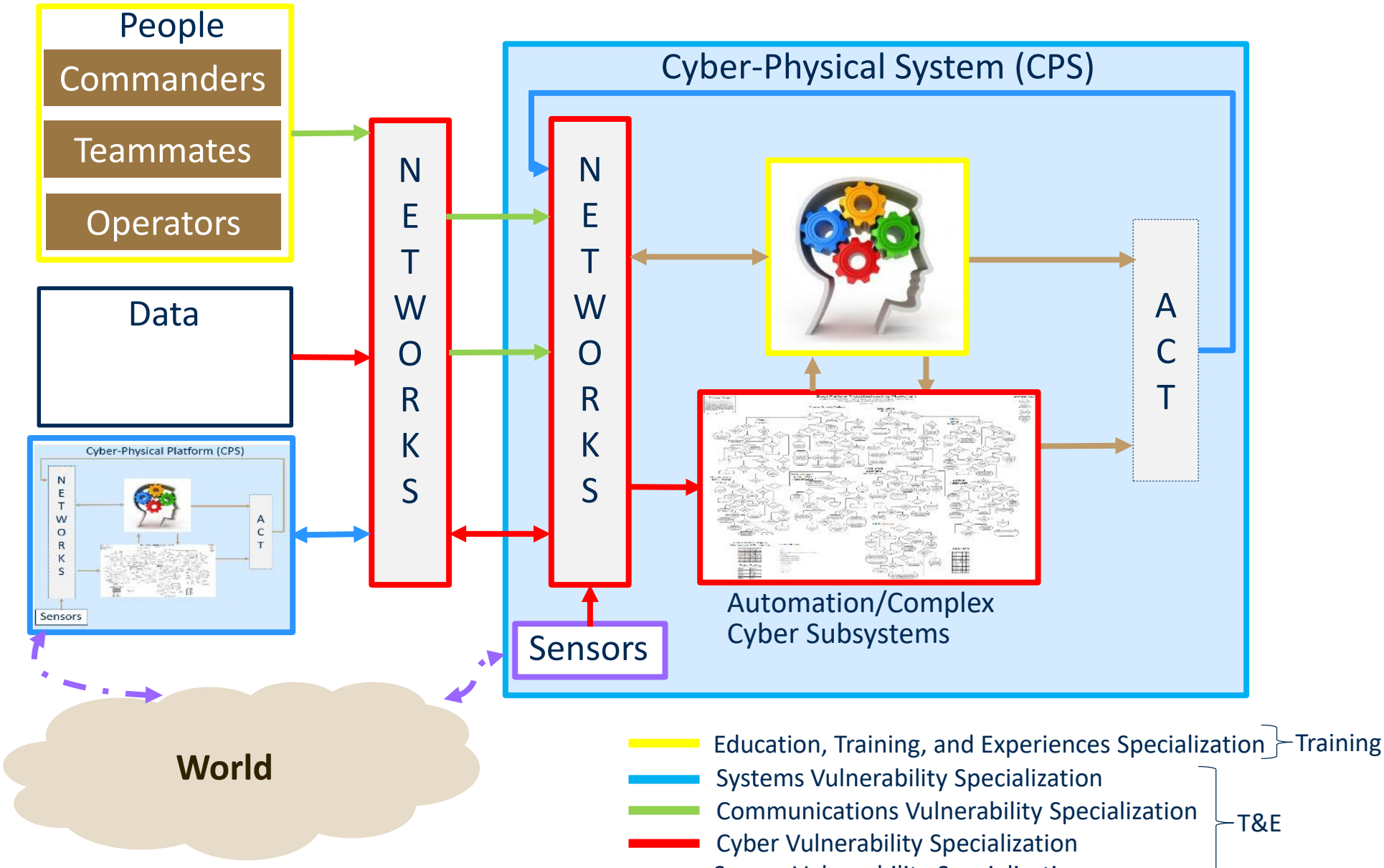
Automation/Complex Cyber Subsystems

ACT

Sensors

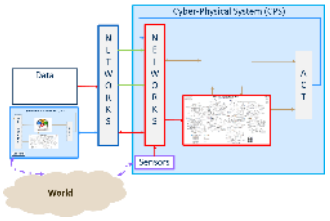
World

Vulnerability Mitigation in a CPS



Addressing CPS Vulnerabilities

Solved today with two complementary paths, each rooted in different foundations.

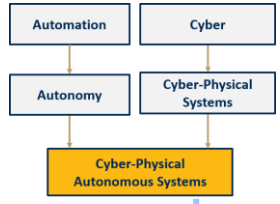


People
Commanders
Teammates
Operators

Problem	**Scenario with Technical Solution**	**Mission ("Commanders Intent") with Warfighter Solution**
Scope	**Domain/task specific sensing, perception, and decision making**	**Warfighter-mission perception, comprehension, projection, and decision making**
Solution	**Solved with evolving T&E Capabilities, Design of Experiments to meet complex SoS demands.** Systems Vulnerability Specialization Communications Vulnerability Specialization Cyber Vulnerability Specialization Sensor Vulnerability Specialization	**Solved with established warfighter training, and evolving T&E-enabled infrastructure (JMETC, etc.), and operationally relevant test and training environments.** Education, Training and Experiences Specialization
Foundation	**"Scientific Method"** **Laws of physics, Statistics, …**	**US Warfighter Values and Beliefs, Attributes of information: Trust & Confidence; Attributes of others (humans): Intent, Trust, and Confidence; Expectations of self/of others**

An Autonomous CPS Decomposition

Georgia Tech Research Institute



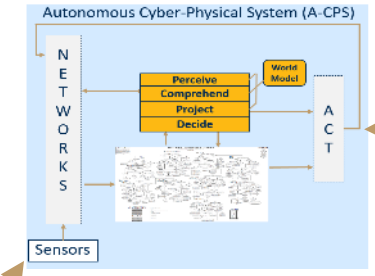
People

Commanders

Teammates

Operators

Data



NETWORKS

Autonomous Cyber-Physical System (A-CPS)

NETWORKS

Perceive

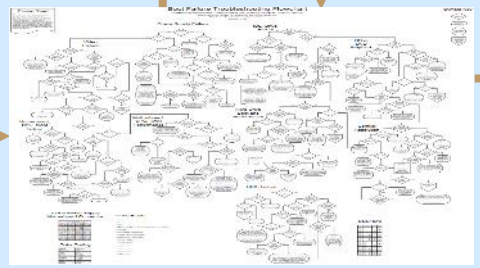
Comprehend

Project

Decide

World Model

ACT

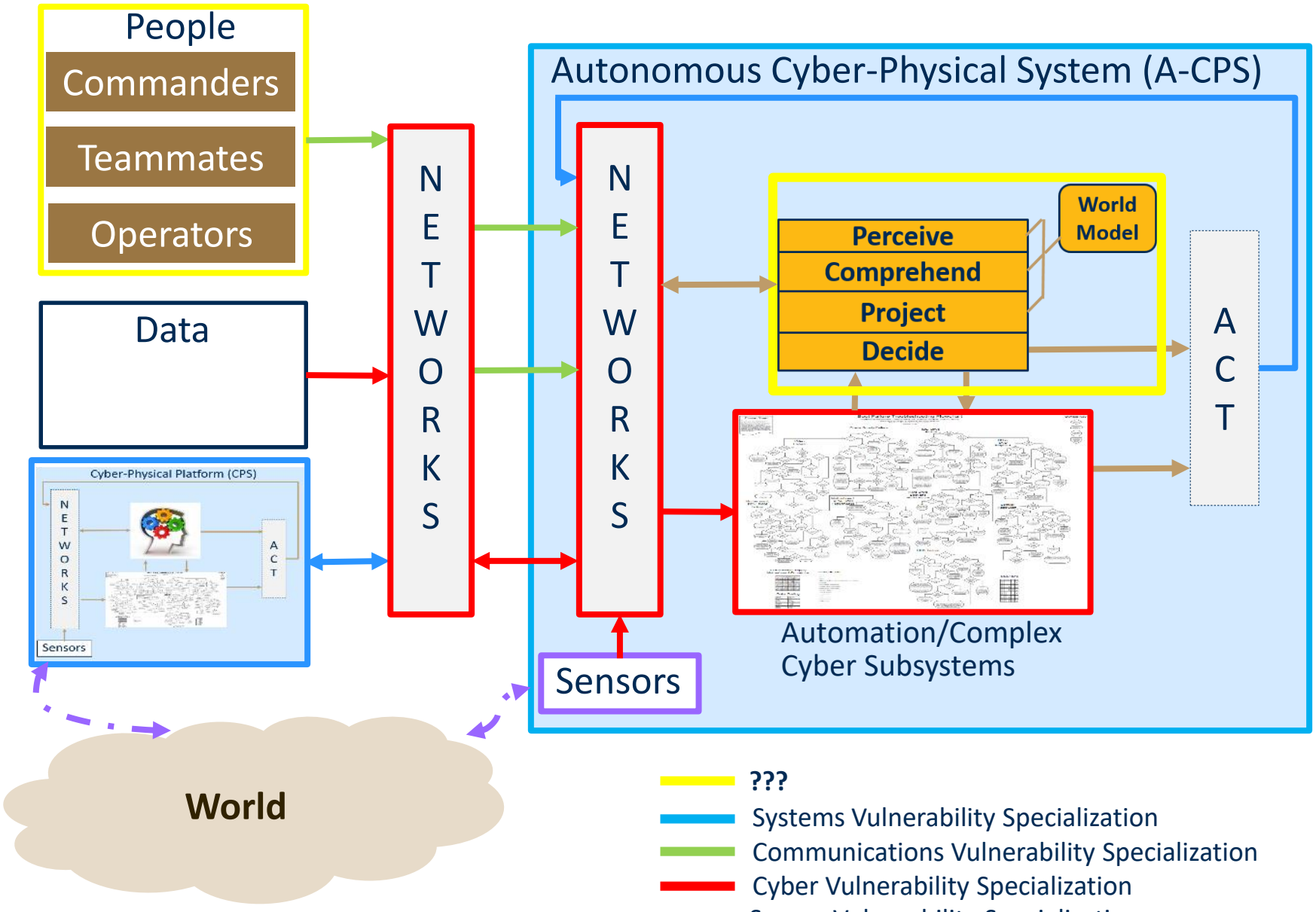


Automation/Complex Cyber Subsystems

Sensors

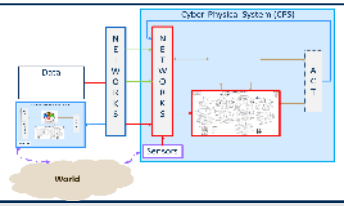
World

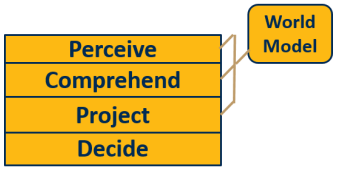
Vulnerability Mitigation in an A-CPS



Addressing A-CPS Vulnerabilities

While the problem and scope do not change, a *new foundation* must be established and an appropriate *solution* must be formed.





Problem	**Scenario with Technical Solution**	**Mission ("Commanders Intent") with Warfighter Solution**
Scope	**Domain/task specific sensing, perception, and decision making**	**Warfighter-mission perception, comprehension, projection, and decision making**
Solution	**Solved with evolving T&E Capabilities, Design of Experiments to meet complex SoS demands.** Systems Vulnerability Specialization Communications Vulnerability Specialization Cyber Vulnerability Specialization Sensor Vulnerability Specialization	**Do established warfighting training solutions apply? i.e. do we seek to train our autonomy and establish standardized evaluations for it to "pass" and become "qualified"?** **Do technical solutions "port" over? i.e. Should/ can we define "autonomy attack surfaces"?**
Foundation	**Foundation: "Scientific Method"** **Laws of physics, Statistics, …**	**Does the autonomy have Values and Beliefs? What attributes does the autonomy give to unknown entities, to peers, to data sources, to support team? How are these attributes formed?**

Key Takeaways

- **The human operator/controller decision-making in today's Cyber Physical System exists because of his/her foundation of:**
  - ***Human-inspired* US warfighter Values and Beliefs**
  - ***Human-inspired* abilities to attribute trust and confidence in information**
  - ***Human-inspired* abilities to attribute intent, trust and confidence in others**
  - ***Human-inspired* abilities to establish expectations of self and others**
- **Today's vulnerabilities of the decision making process are assessed and mitigated based upon this *human-inspired* foundation**
- **This foundation naturally enables today's decomposition into complimentary "technical" solutions and "training" solutions**

Key Takeaways

- **The autonomy decision-making in tomorrow's Autonomous Cyber Physical System has a very different foundation:**
  - ***Machine-based* Warfighter Values and Beliefs**
  - ***Machine-based* abilities to attribute trust and confidence in information**
  - ***Machine-based a*bilities to attribute intent, trust and confidence in others**
  - ***Machine-based* abilities to establish expectations of self and others**
- **Tomorrow's vulnerabilities of the decision making process will need to be assessed and mitigated based upon these *machine-based abilities***
- **This (significantly different) foundation limits today's "training-based" solutions from solving tomorrow's autonomy decision-making vulnerability**

Summary

- **Assessing vulnerability of the autonomy in an Autonomous-Cyber Physical System is a significant challenge rooted in foundational changes in transferring decision making from a human to a machine.**
- **Solutions may come from a combination of the following:**
  - **Carrying forward and applying established (communication, cyber, sensor, etc.) vulnerability domain T&E toolsets and methodologies to the complex Cyber-Physical System/SoS – *"leverage"***
  - **Expanding established vulnerability domain T&E methodologies such as attack surface identification, attack vector generation, etc., to evaluate autonomous decision making - *"build"***
  - **Adaptations to today's training-centric solution *"change"***

Questions/Comments?

- **Contact Information**
  - **donald.strausberger@gtri.gatech.edu**