Coactive Emergence as a Sensemaking Strategy for Cyber Defense

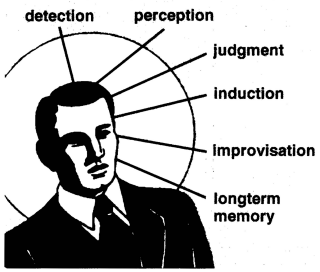
Jeffrey M. Bradshaw[1], Marco Carvalho,[2] Larry Bunch[1], Tom Eskridge[1], Paul Feltovich[1], Robert R. Hoffman[1], Matt Johnson[1]

[1]Florida Institute for Human and Machine Cognition (IHMC)

[2]Department of Computer Science, Florida Institute of Technology (FIT)

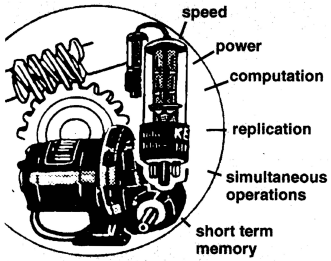
Paul Fitts: HABA-MABA Chart

HUMANS SURPASS MACHINES IN THE:



- Ability to detect small amounts of visual or acoustic energy
- Ability to perceive patterns of light or sound
- Ability to improvise and use flexible procedures
- Ability to store very large amounts of information for long periods and to recall relevant facts at the appropriate time
- Ability to reason inductively
- Ability to exercise judgment

MACHINES SURPASS HUMANS IN THE:

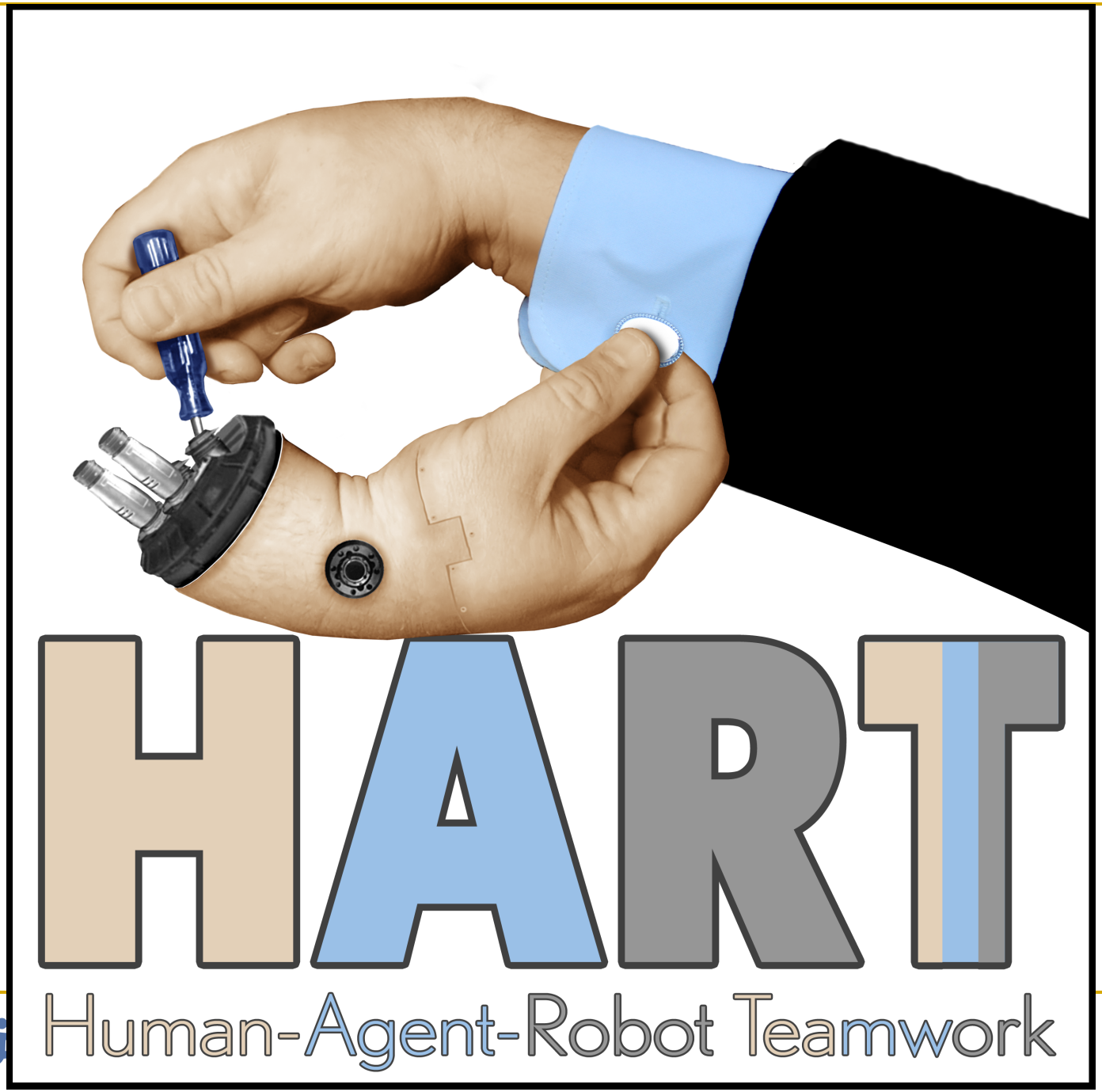


- Ability to respond quickly to control signals, and to apply great force smoothly and precisely
- Ability to perform repetitive, routine tasks
- Ability to store information briefly and then to erase it completely
- Ability to reason deductively, including computational ability
- Ability to handle highly complex operations, i.e., to do many different things at once.

ihmc

Woods and Hoffman: An "Un-Fitts List"

Machines	
Are constrained in that:	**Need people to:**
Sensitivity to context is low and is ontology-limited	Keep them aligned to context
Sensitivity to change is low and recognition of anomaly is ontology-limited	Keep them stable given the variability and change inherent in the world
Adaptability to change is low and is ontology-limited	Repair their ontologies
They are not "aware" of the fact that the model of the world is itself in the world	Keep the model aligned with the world
People	
Are not limited in that:	**Yet they create machines to:**
Sensitivity to context is high and is knowledge- and attention-driven	Help them stay informed of ongoing events
Sensitivity to change is high and is driven by the recognition of anomaly	Help them align and repair their perceptions because they rely on mediated stimuli
Adaptability to change is high and is goal-driven	Effect positive change following situation change
They are aware of the fact that the model of the world is itself in the world	Computationally instantiate their models of the world


HART
Human-Agent-Robot Teamwork

Emergence

- American Heritage Dictionary: “the process of coming into view or becoming exposed” (e.g., identification, characterization, and interpretation of cyber threats)
- Commonly used to describe a process whereby complex system behavior arises from interactions of much simpler components or primitives (e.g., Perrow “normal accidents”)

Teamwork and Taskwork are Separable

Soccer Taskwork:

- Kicking to a target
- Dribbling, tackling
- Tracking the ball, goal …

Soccer Teamwork:

- Allocating players to roles
- Synchronizing tactics
- Sharing relevant information
- …….

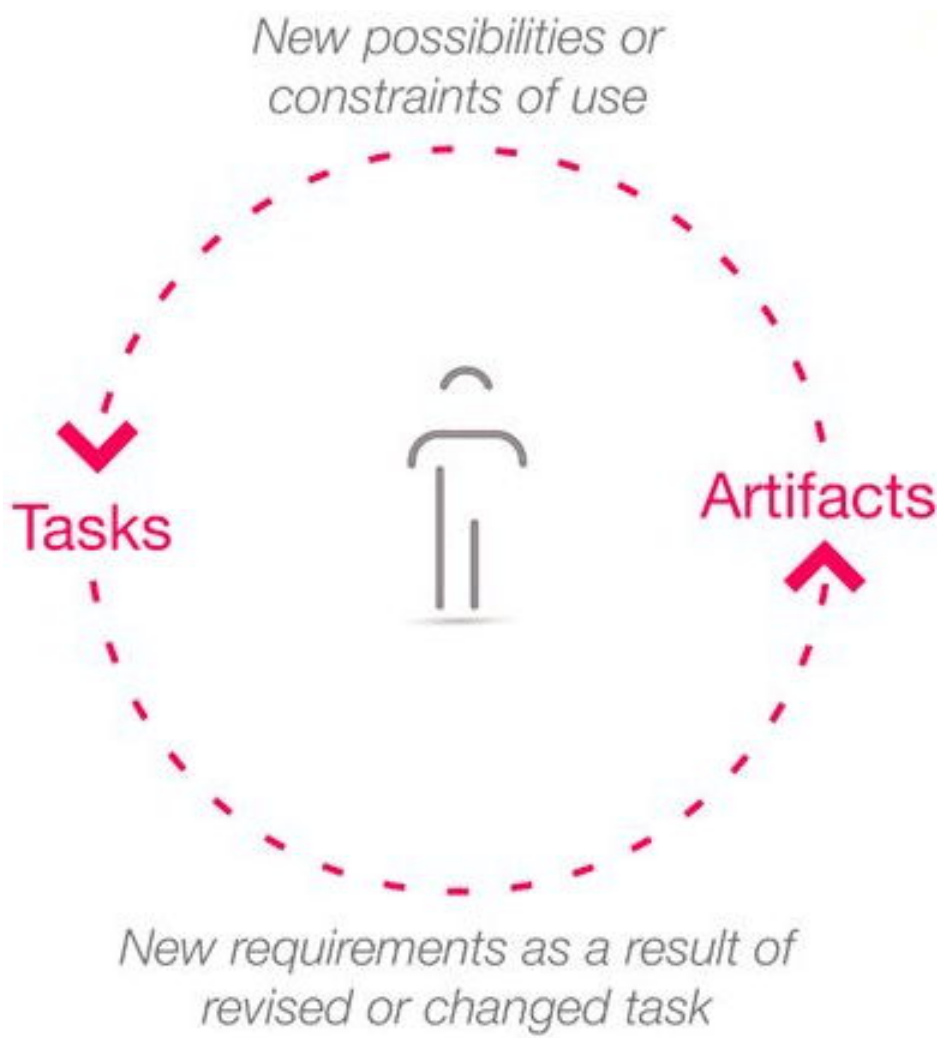
Requirements for Resilient Human-Agent Teamwork

- Observability
- Directability
- Interpredictability
- Learning
- Multiplicity

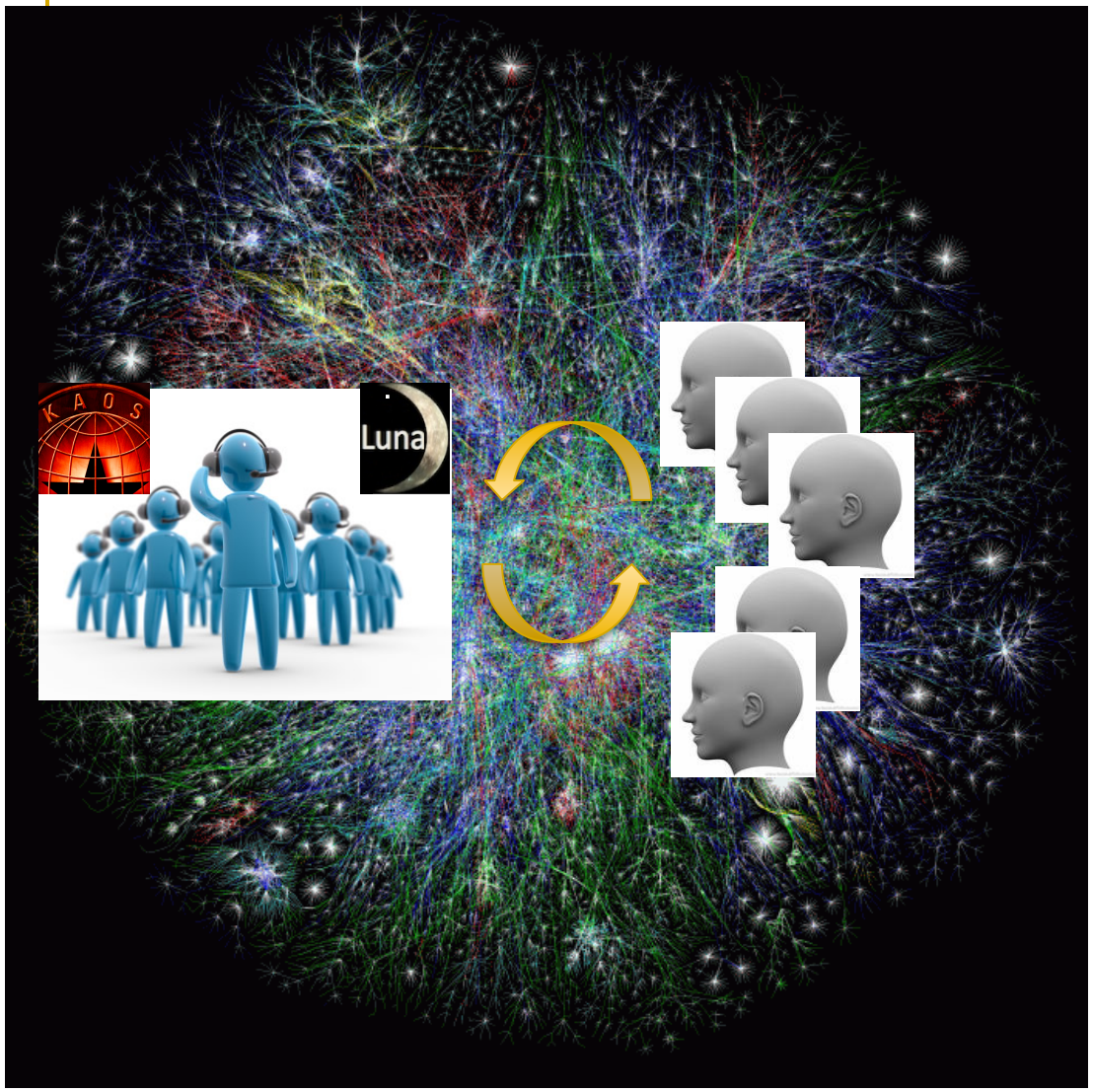
CyberSA Research Questions

- How can help be provided to analysts with tedious, time-consuming, complex, and high-tempo tasks?
- What breakthrough capabilities can improve both threat detection and overall sensemaking?
- What can be done to facilitate teamwork and information-sharing among analysts, and collaboration between humans and machines?
- Can today's stovepiped tools be integrated into a context-sensitive, task-aware capability?
- How can we architect an organically-extensible system that automatically scales to dynamically-varying computing and network resources, and that accommodates future analytic innovations that cannot now be anticipated?
- How can we build resilient systems capable of ensuring mission continuity, even when under attack or experiencing failures?

Carroll: Task-Artifact Cycle

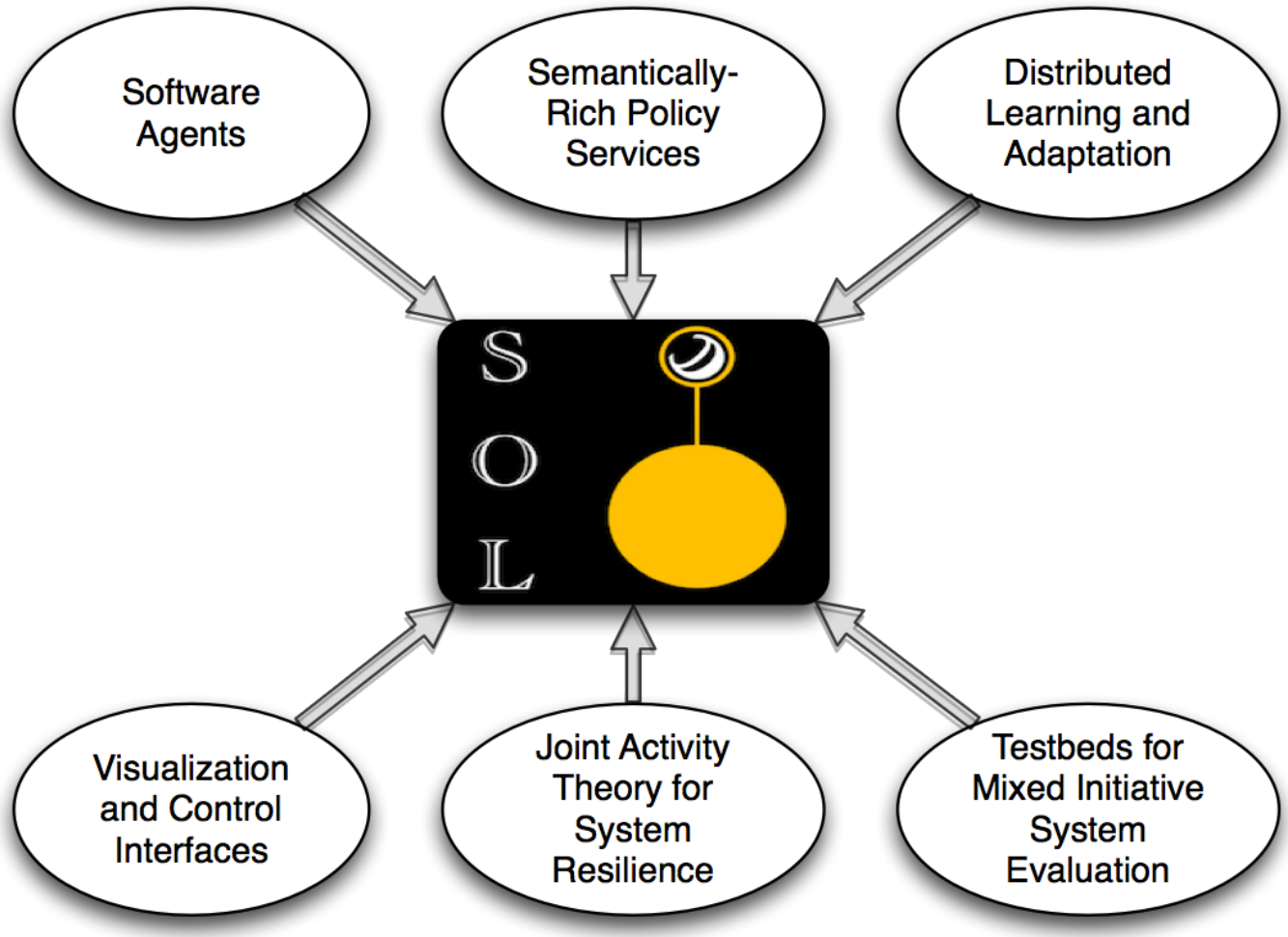


Key Capabilities for Distributed Sensemaking



- *OZ-Inspired Visualizations* based on principles of human perception and cognition, and designed around what people need to know in a given context
- *Luna Software Agents* based on an understanding of joint activity and biologically-inspired learning, and designed to assist team members with a variety of tasks
- *An Adaptive, Policy-Governed System Architecture* designed to ensure directability, robustness, survivability, and resilience

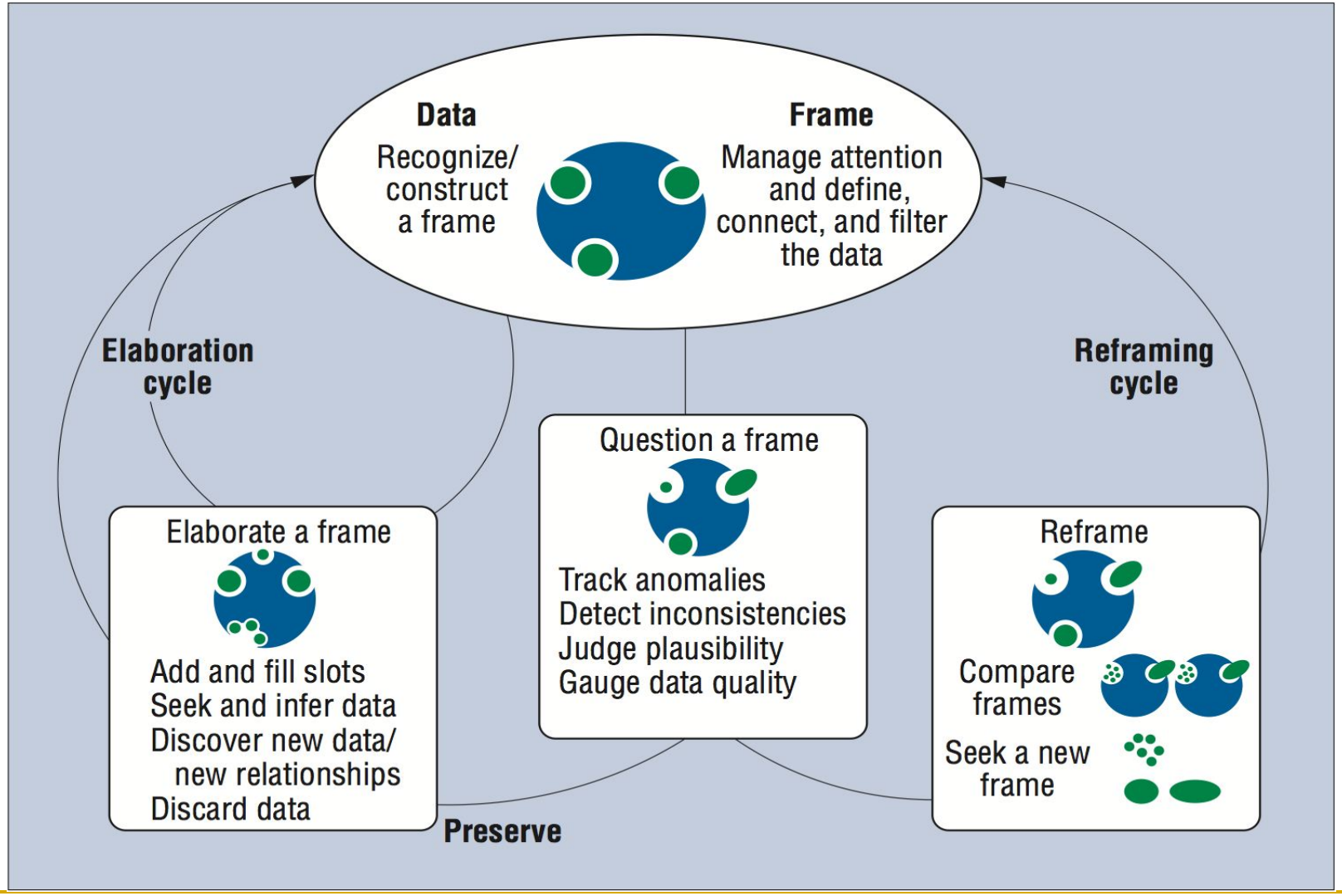
Research Contributions to Sol



Sensemaking in Cyber Defense Analysis

- Well understood:
  - Ways to help analysts counteract lines of reasoning that might lead to misconceptions
- Less understood:
  - Impact of new forms of visualization and automation on the sensemaking process
  - How to design such tools in light of what is known about sensemaking
- Our emphasis is on putting questions about the role and benefits of computer interaction with people in center stage

The Data/Frame Theory of Sensemaking



Role of Constraints in Emergence

- All emergent systems are bounded by structural and environmental constraints that “frame” the data
- These constraints define the “wiggle room” within which emergent phenomena may develop

Second-Order Emergence

- Some complex systems are capable of exhibiting "second-order" emergence
- In such systems, the constraints that govern emergence also are subject to change
- Such dynamic constraints "wiggle the wiggle room" within which emergent phenomena may develop (i.e., reframing)
- The environmental influences that prompt second-order changes are, in the general case, indifferent to the objectives of the system

Coactive Emergence

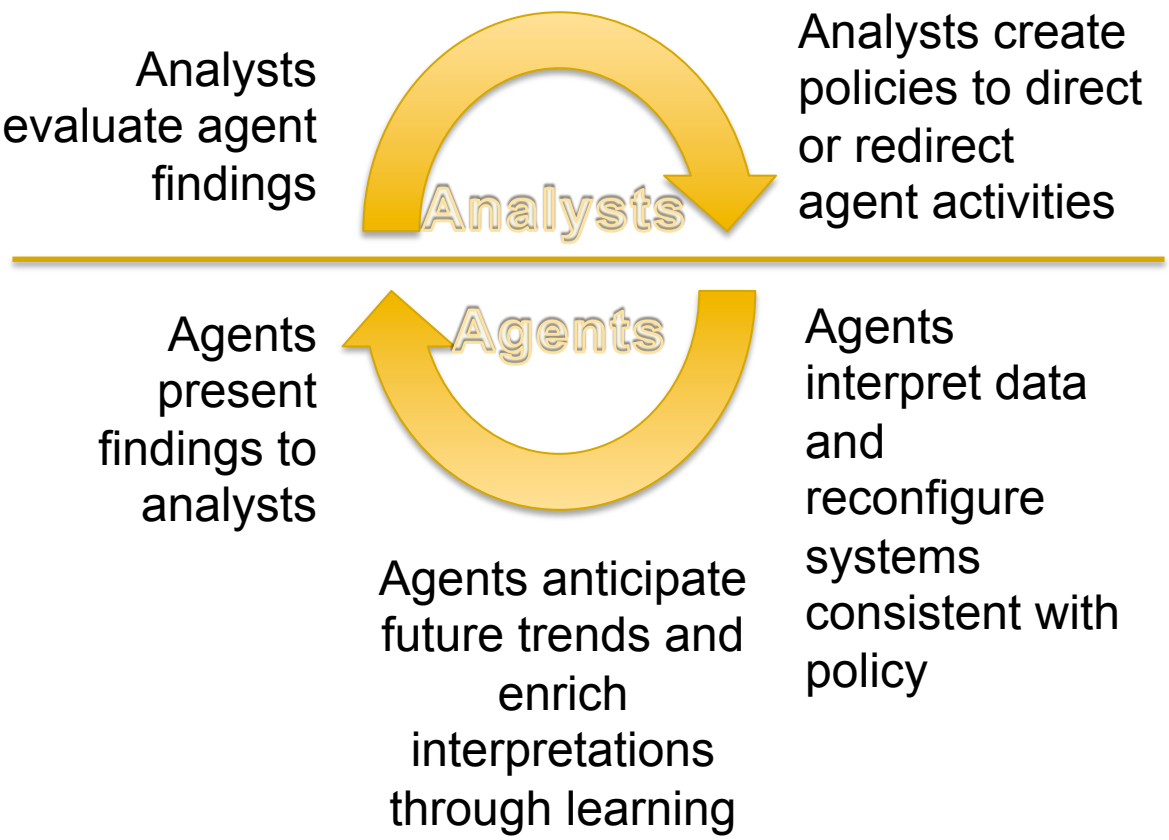
- Coactive emergence is a form of second-order emergence
- Unlike ordinary second-order emergence, it involves multiple participants who deliberately seek to influence the direction of adaptations in ways that converge on mutually-shared objectives (convergent reframing)
- “Coactive” emphasizes the explicit aim of synergy in joint development of a common threat understanding by humans and software agents working interdependently

Adversarial Modeling

- Adversaries engage in a similar coactive process involving intentional adaptation to friends and to foes (Branlat, Woods, et al.)
- Adaptations to foes seek to disrupt any activity thought to be useful to one's adversary
  - Disrupt sensemaking through causing adversary uncertainty about frame/reframing (deception)
  - Disrupt adversary action through protecting targets or making them unpredictable
- Example: Moving Target Defense, which seeks to present attackers with an uncertain and unpredictable target

Sol

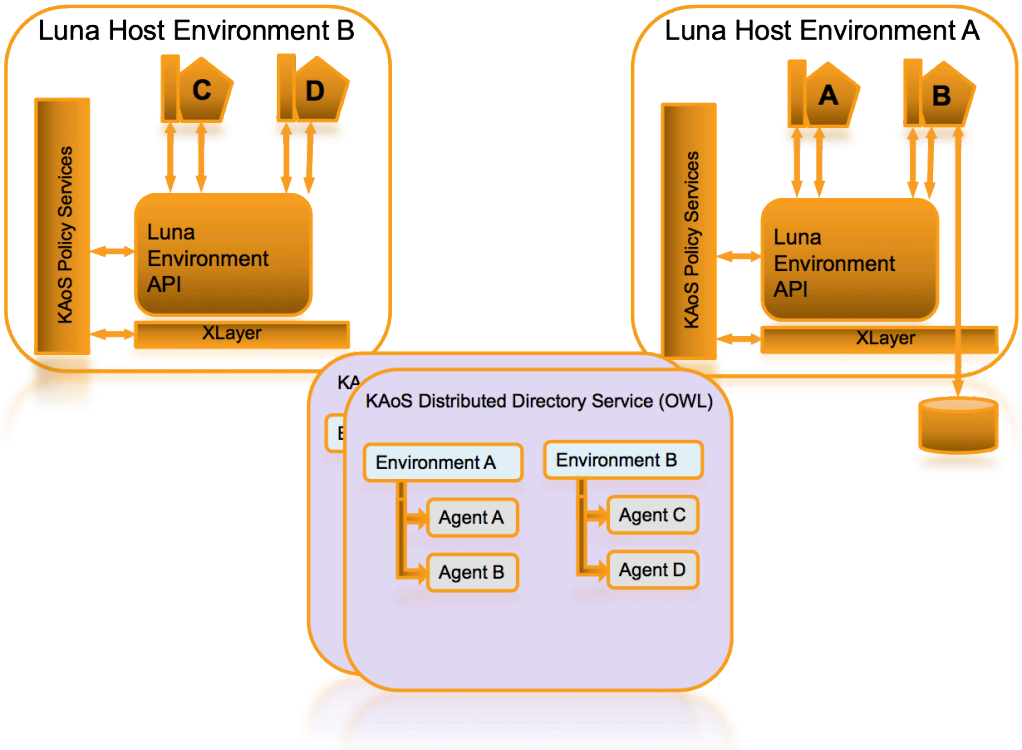
Coactive Emergence



Luna Agent Platform

Luna Agent Framework Objectives

- Lightweight and Efficient
- Comprehensive Policy-Governed Operation
- Controlled Hosting Environment
- Agent Persistence
- State Mobility
- Dynamic Load-Balancing
- Adaptive Resource Management for Scalability
- Resource Awareness and Control
- Flexible Agent Messaging
- Full Semantic Representation
- Automatic Generation of OWL Ontologies
- Easy to Learn, Use, and Maintain
- Robust Implementation



Semantically-Rich KAoS Policies

Credentials

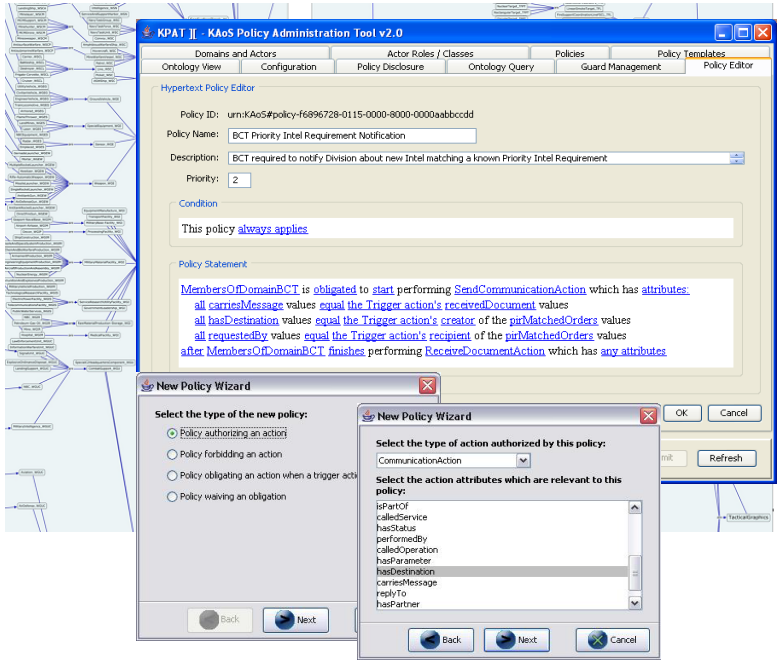
- IHMC researchers participated in the DARPA DAML program that resulted in the standardization of the Web Ontology Language (OWL) and were the first to use it for policy representation and reasoning
- IHMC researchers have been highly involved in the professional community as organizers and presenters

Selected Accomplishments

- A decade of R&D maturing the *KAoS Policy and Domain Services framework*, and its application in dozens of sponsored research projects
- Facilities for analysis, deconfliction, enforcement

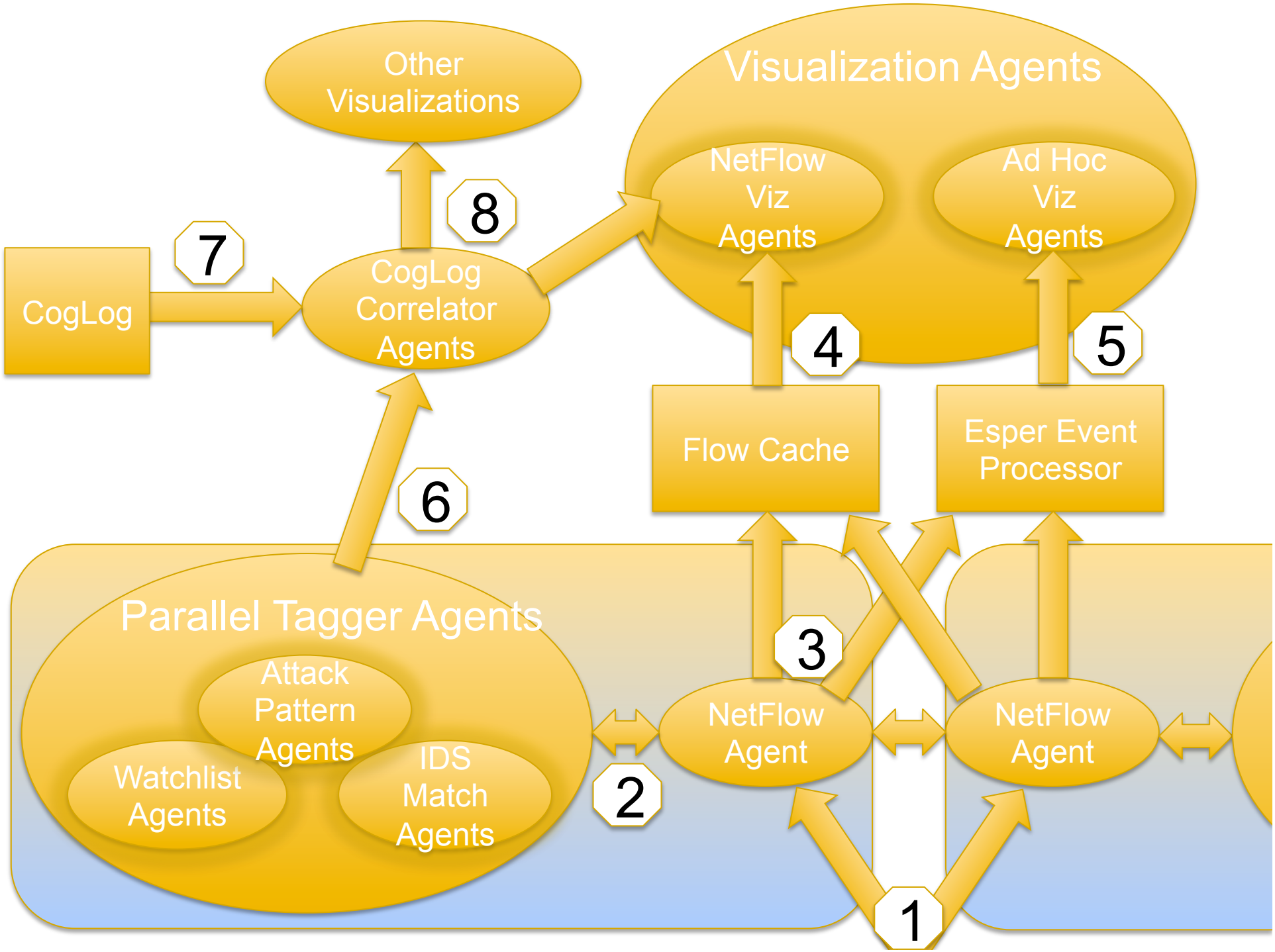
Application to Project

- With the combination of software agents, semantically-rich policies, and sensitivity to context in our net watch framework, we aim to provide a powerful means to deal with complex, high-tempo event monitoring, pattern detection, and intervention.
- Institutions and individuals will use a point-and-click hypertext interface to specify policies defining agent and other system behavior.
- Policies will assure that security and other important constraints on agent behavior will be enforced.

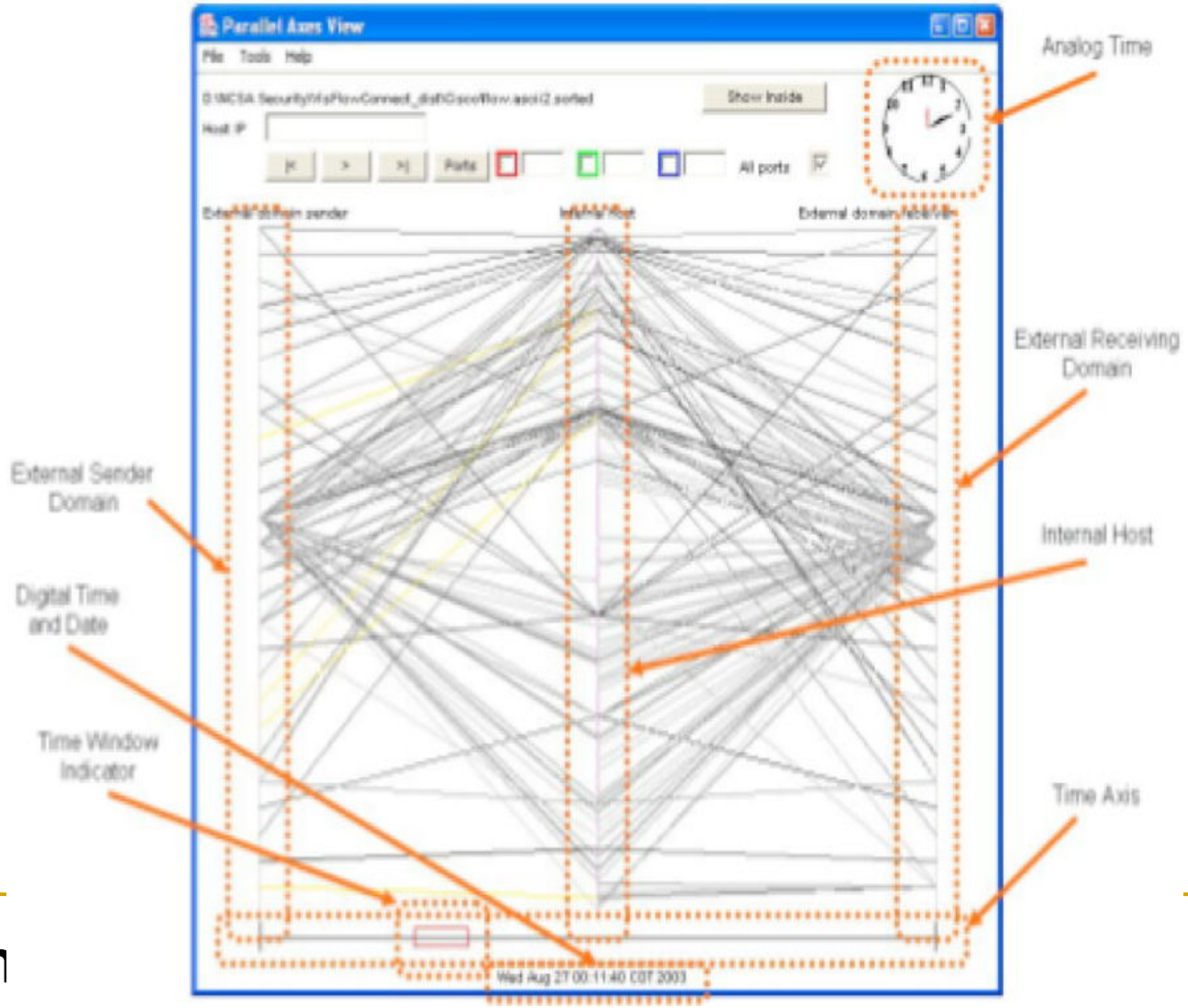


Benefits

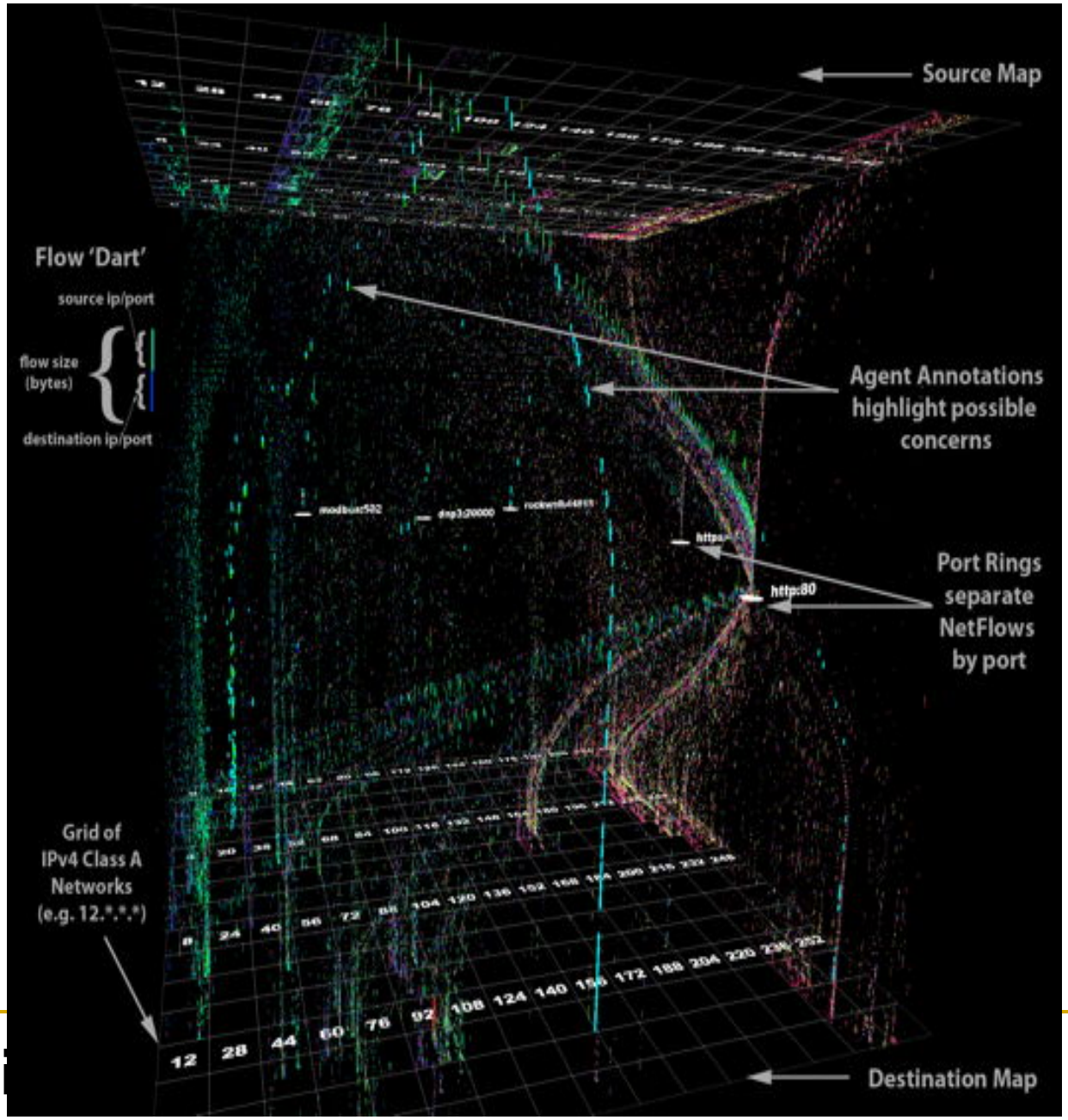
- The application of policy to regulate agent and system behavior makes interaction as natural and effective as possible, while providing assurance that security and other important operational concerns are fully respected.


Other Visualizations
Visualization Agents
NetFlow Viz Agents
Ad Hoc Viz Agents
8
7
CogLog
CogLog Correlator Agents
4
5
Flow Cache
Esper Event Processor
6
Parallel Tagger Agents
Attack Pattern Agents
Watchlist Agents
IDS Match Agents
3
2
NetFlow Agent
NetFlow Agent
1

VisFlowConnect Parallel Coordinates Display

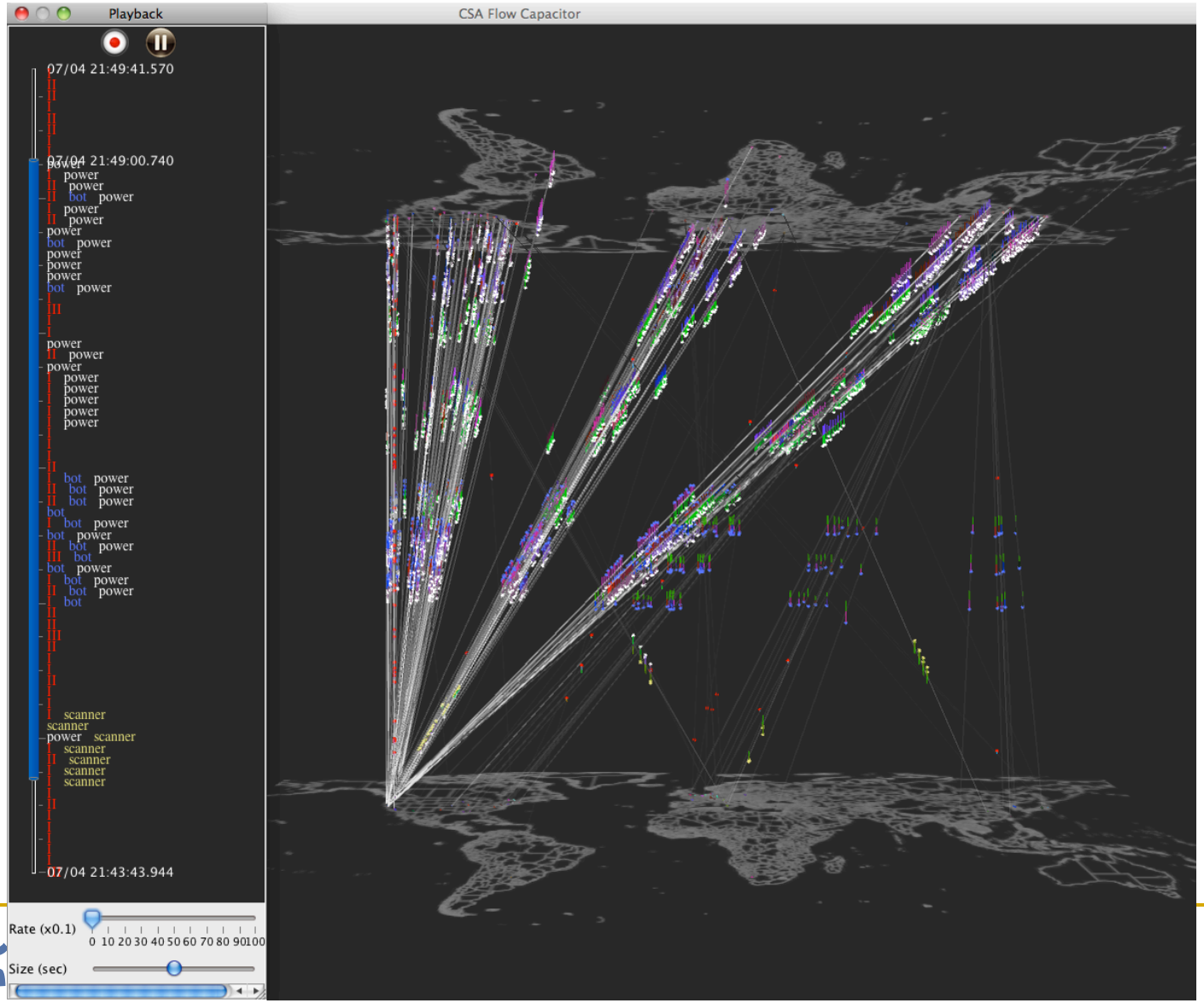


Flow Capacitor (Patent Pending)

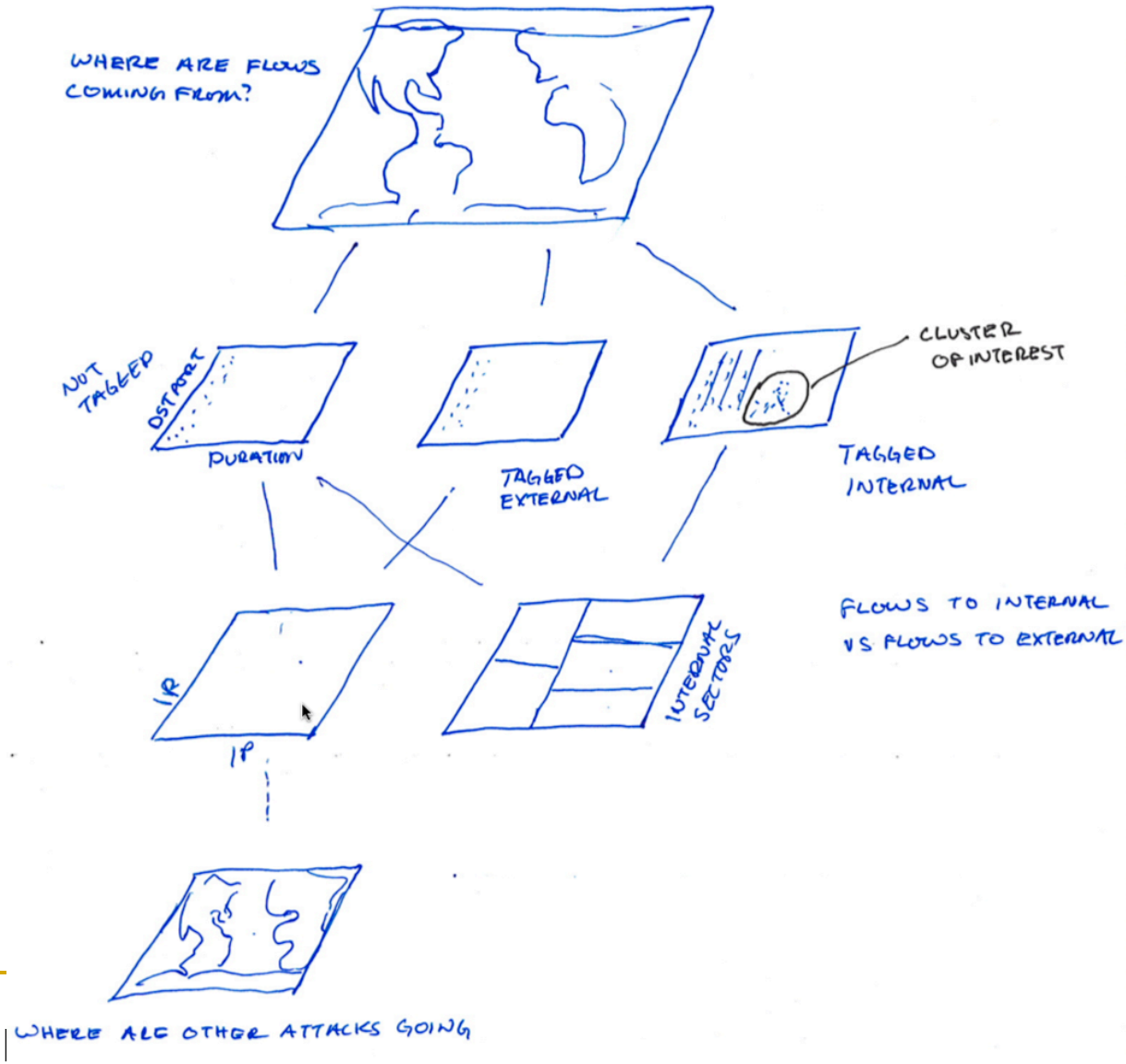


Agents Augment Human Pattern Recognition (Patent Pending)

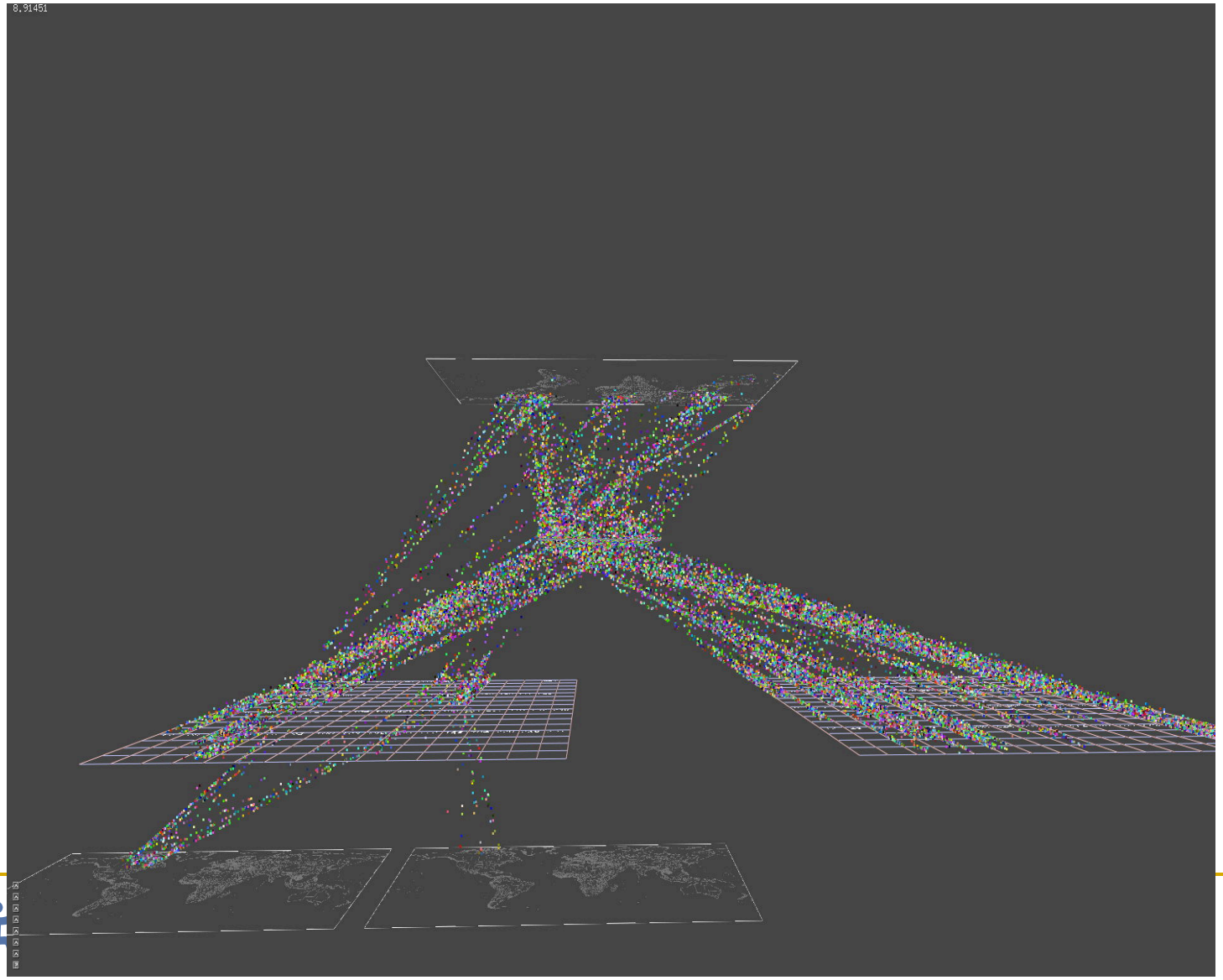
… by learning new threat patterns and presenting them to the analyst for validation



Parallel Coordinates 3D Observatory (PC3O)



PC3O Prototype



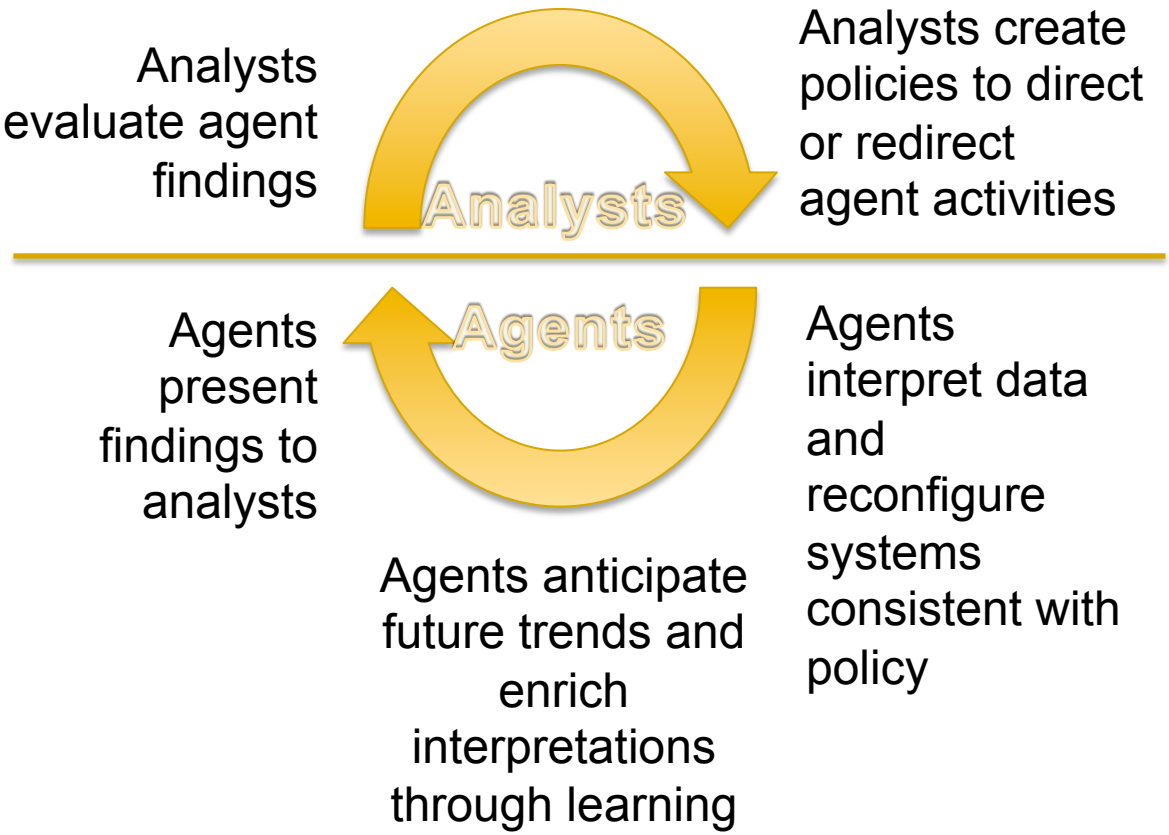
Hypotheses for Future Evaluation Studies

- Increased speed of co-evolution of tools and strategies
- Increased range, richness, and utility of threat models by considering perspectives of multiple analysts and software agents
- Wider range of adaptations through policy-based second-order changes
- Ability to explore hypotheses in both collaborative and competitive modes

Examples of Potential Issues

- Can end user customization tools be made sufficiently usable?
- Stability vs. adaptability

Summary and Contact Information



Jeff Bradshaw, IHMC, jbradshaw@ihmc.us, www.ihmc.us/groups/jbradshaw/
Marco Carvalho, FIT, mcarvalho@fit.edu, http://research.fit.edu/hcdi/people/carvalho.php