Transforming the Testing and Evaluation of Autonomous Multi-Agent Systems: Introducing In-Situ Testing via Distributed Ledger Technology

Mr. Stuart Harshbarger
Chief Technology Officer
Assured Intelligence.ai

Rosa R. Heckle, PhD
Artificial Intelligence Principal MITRE Corporation.

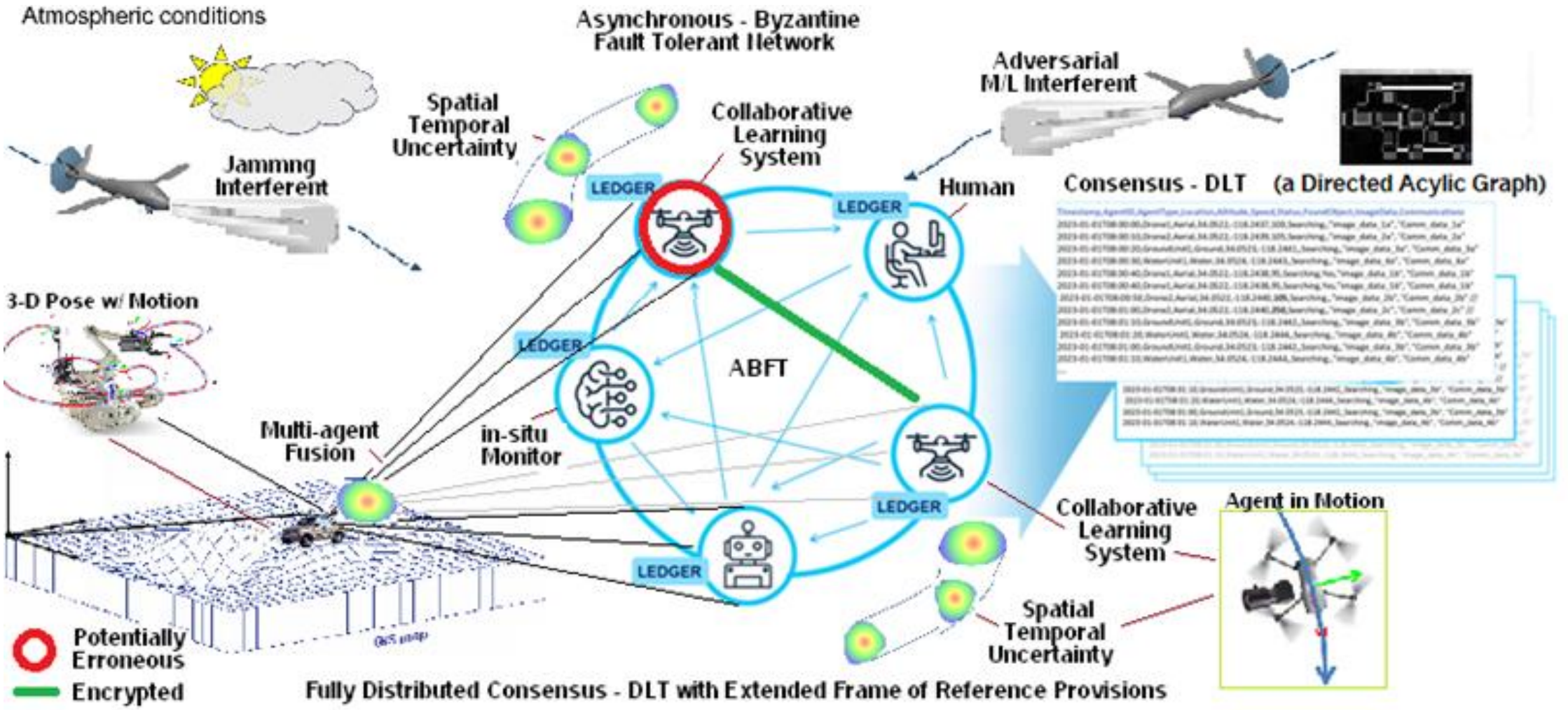
Michael Collins
Technical Director NSA Visiting Professors
National Security Agency

Toward Resilient AI-enabled Systems:

Our presentation explores the application of Consensus-based Distributed Ledger Technology (C-DLT) in the testing and evaluation of collaborative adaptive AI-enabled systems (CA2IS).

It highlights the potential of C-DLT to enhance real-time data collection, data validation, synchronization, and security while providing a trusted framework for AI model & parameter sharing, tiered access control, multi-agent data fusion, and (as emphasized for this topic area) novel methods for continuous monitoring and Test & Evaluation.

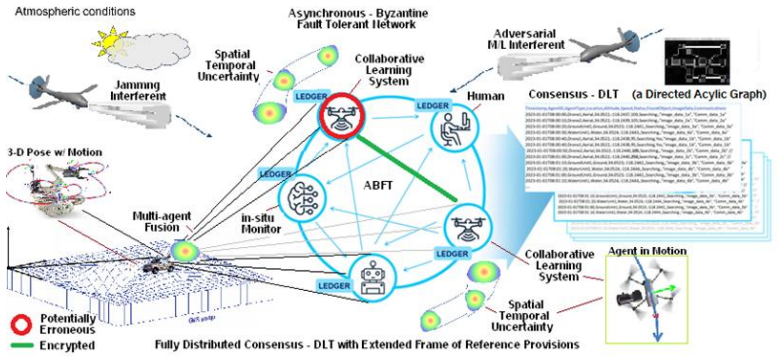
Continuous Learning and System Adaptation in Uncertain Environments:



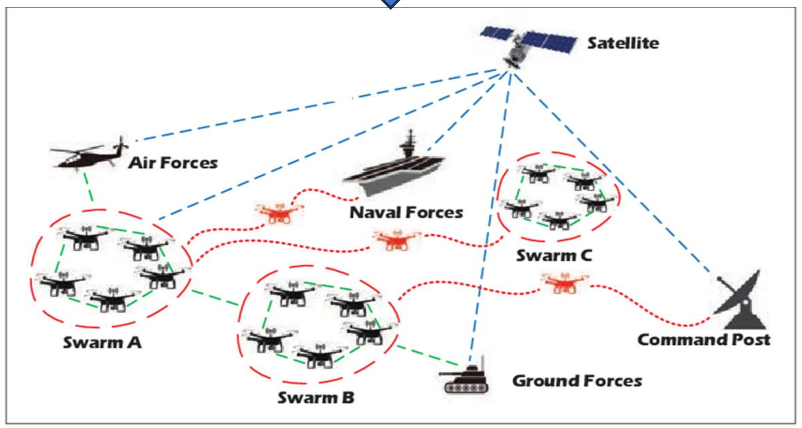
Beyond Traditional Testing Methods:

As autonomous multi-agent systems evolve, the underlying probabilistic foundation of increasingly AI/ML-enabled systems necessitates the adoption of new approaches to system Test and Evaluation (T&E) that extend traditional range-based & fixed scenario-based repetitive testing methods that have primarily evolved for testing of historically deterministic systems. New testing methods and capabilities are needed to assess the performance expectations of probabilistic systems for applications in highly complex and dynamic environments where both operating conditions and system performance may be constantly changing and continuously adapting. Ideally, T&E methods would seamlessly extend from developmental testing, through acceptance and validation testing, and into mission operations (of critical importance as autonomous weapons-capable systems emerge). Additionally, these provisions may provide a mechanism for enhancing system resilience in contested environments through the inherently distributed, time ordered, and redundant records of on-board status, diagnostic & behavioral indicators, as well as, inter-agent communications and data transactions. These indications may collectively provide for continuous assessment of system performance as an input to well-informed decision making – whether such decisions are made via an agenic AI, or human-in-the-loop, context.

Current Context



Future Context



A Cross-Community Partnership Initiative:

NSA Research MITRE DARPA US Naval Postgraduate School Other/DOD via CRADA	Resiliency	NSA Research US Naval Academy US Military Academy Vanderbilt University Assured Intelligence.ai MIT-Lincoln Laboratory	Real-World Applications
Formal Methods for Architectural Design	Vanderbilt University MITRE NSA Research MIT-Lincoln Laboratory	Enabling Technologies & Methods	NSA Research US Naval Academy Coast Guard Academy US Military Academy Vanderbilt University Assured Intelligence.ai MIT-Lincoln Laboratory

Fei Xiong, Aijing Li, et al, Published in IEEE Communications Magazine 21 August 2019
Engineering, Computer Science, An SDN-MQTT Based Communication System for Battlefield UAV Swarms | Semantic Scholar

Key Challenges for Testing CA2IS Systems:

Emergent and non-deterministic behavior: "same" inputs can lead to different outputs.

Scalability and complexity: difficult to test all scenarios.

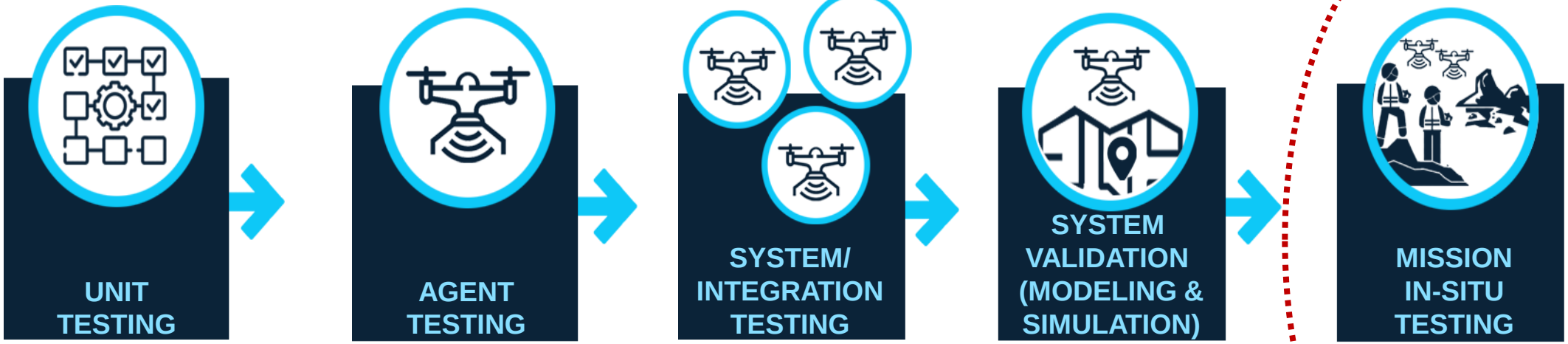
Adaptability and learning capabilities: causing changing behavior with passage of time

Dynamic and unpredictable environments: can influence their behavior

Interoperability and integration issues: agents developed by various organizations

Reproducibility of test scenarios: reproducing specific test cases or failures are difficult.

Development & Deployment Testing Progression:



Individual capabilities within an agent

- Object Detection & Recognition Capability
- Autonomous Navigation
- Gesture Recognition
- Path Planning & Optimization

A single agent with all integrated capabilities

Multiple agents working together

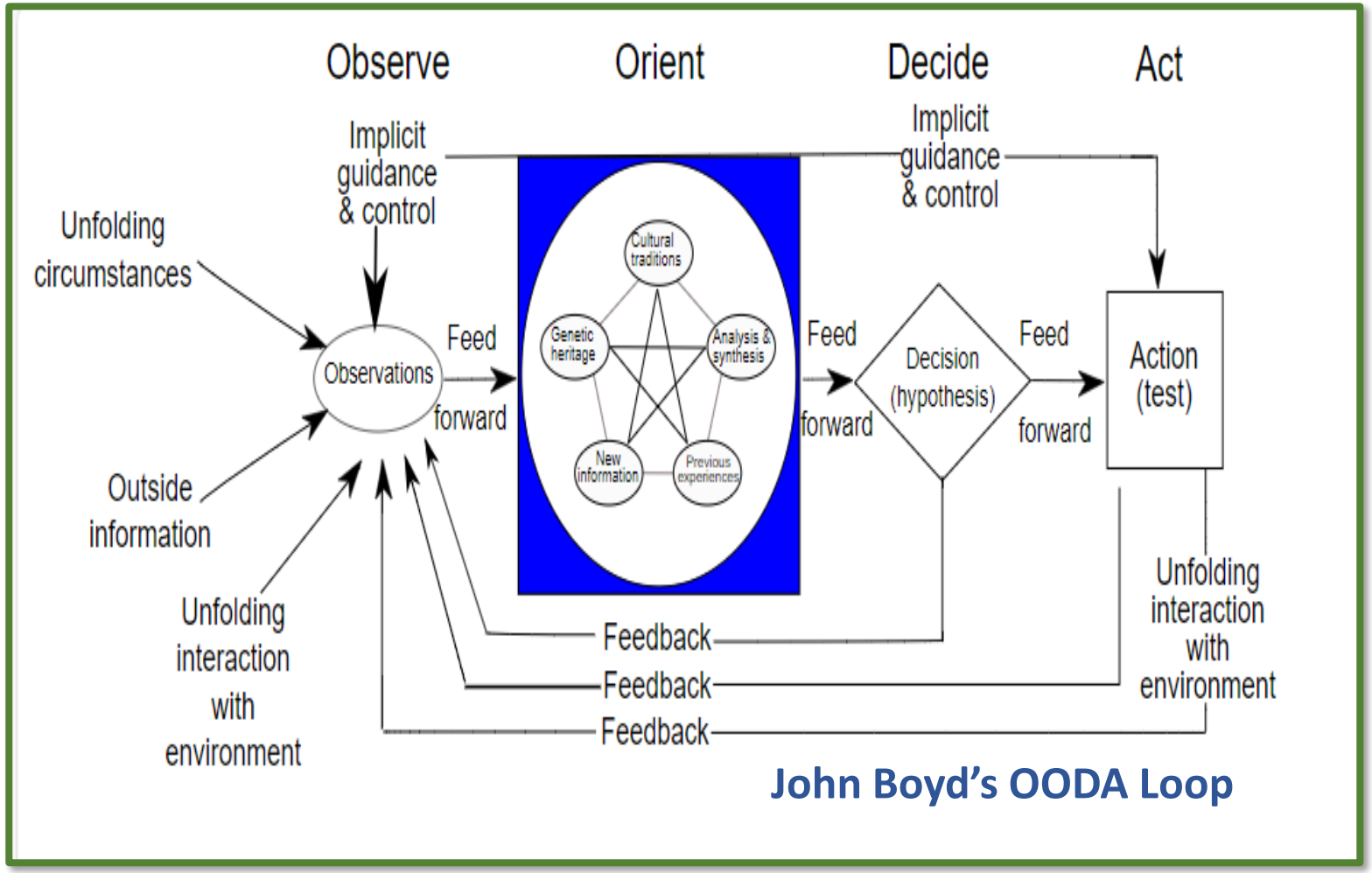
Systems Testing of All agents within virtual environments and scenarios

All agents in real-world conditions

Provides an accurate and comprehensive evaluation of system performance in its actual operational environment.

Exploring the Intricacies of Agent Perception and Decision-Making:

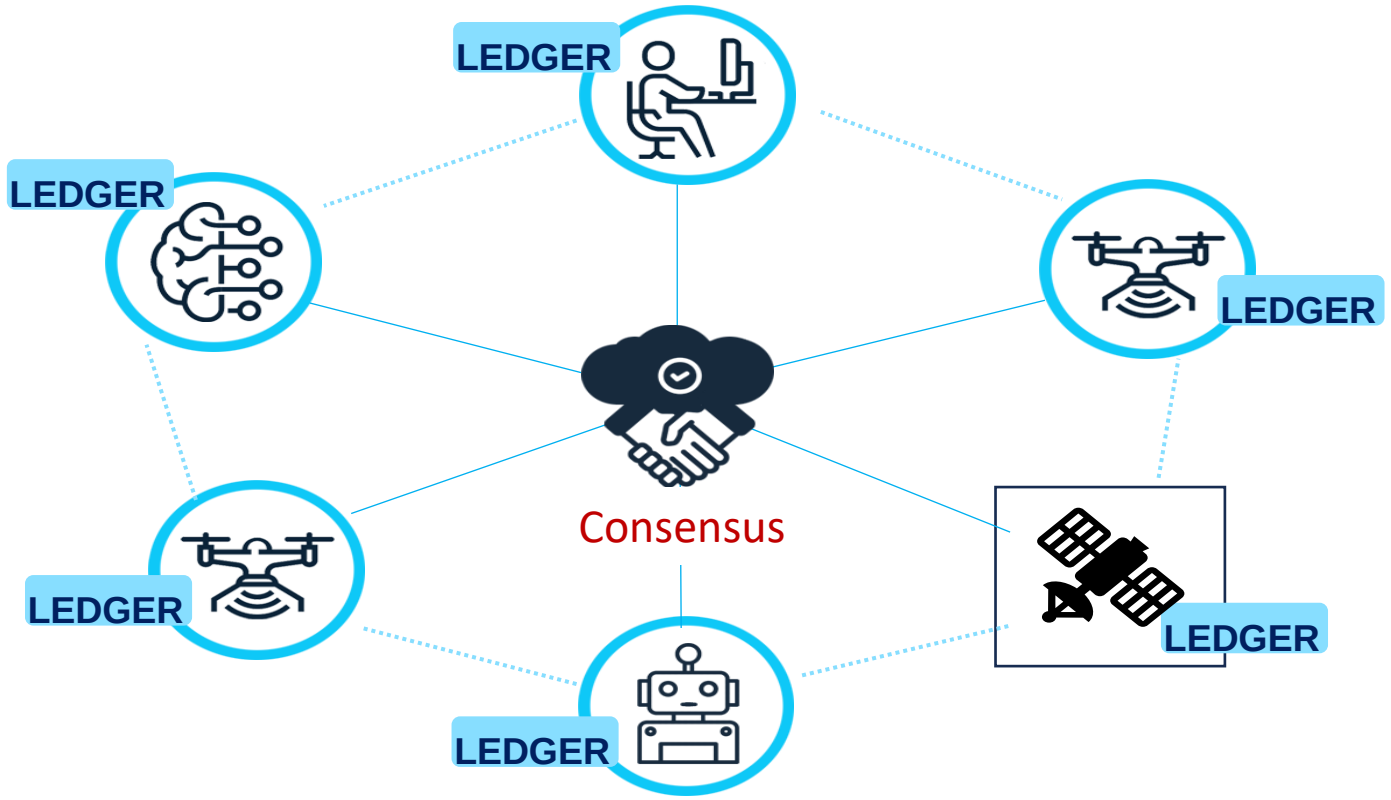
Testers need to understand the perceptions and decision-making processes of the agents;
how agents interpreted their environment, processed incoming data, and made their decisions
- and under what conditions/constraints may end-users predict and expect similar behavior?.



CULTURAL KSAs: Skill Development Using the OODA Loop > Air University (AU) > Article Display (af.edu)

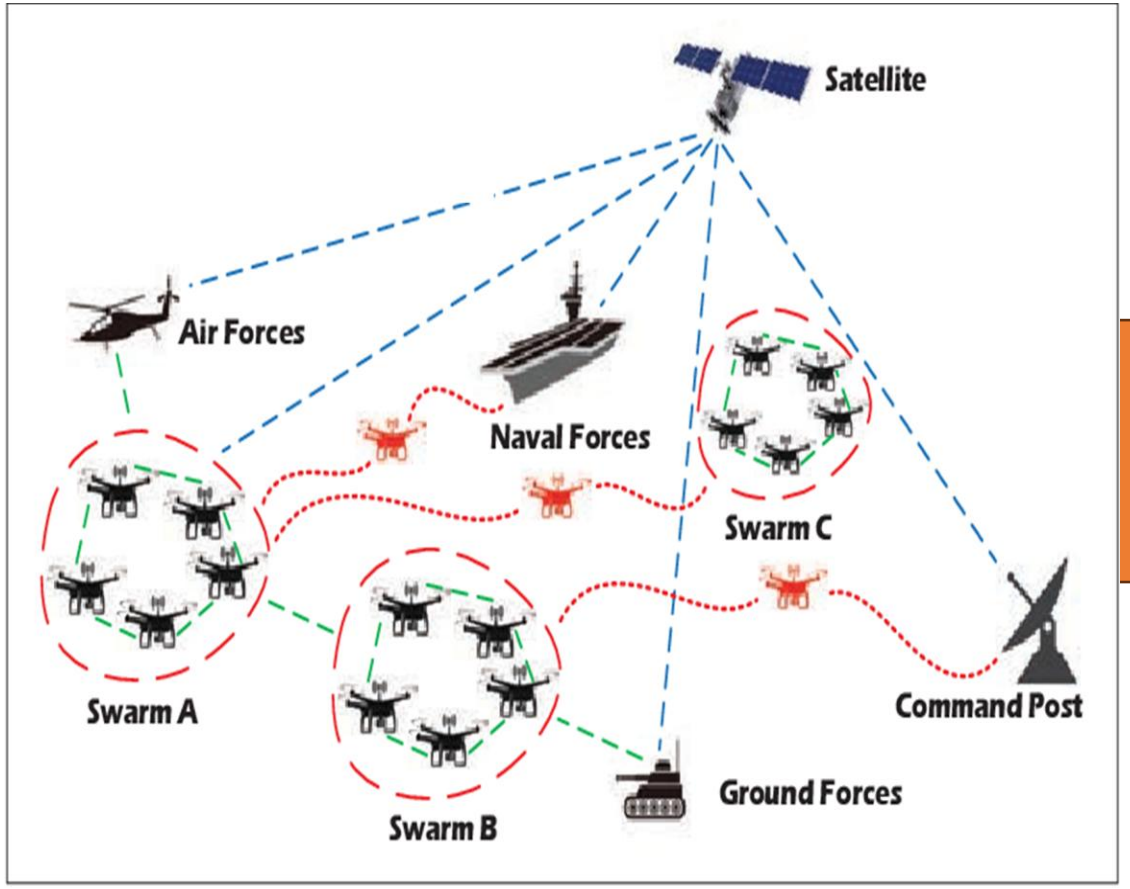
Consensus-based Distributed Ledger Technology (C-DLT):

Each agent records their data onto a ledger.



- **Real-time Data Collection and *Synchronization:*** Facilitates consistent and aligned data collection across multiple agents and sensors.
- **Enhanced Collaboration: S**ingle, trustworthy source of data for all stakeholders.
- **Traceability:** Provides a chronological record of all transactions
- **Consensus Mechanisms:** Ensures data consistency and synchronization across the network.
- **Decentralization: No** single point of failure.
- **Transparency and Immutability:** Ensures data integrity and accountability
- **Security:** Uses cryptographic techniques to secure data

Gathering Data During In-Situ Testing:

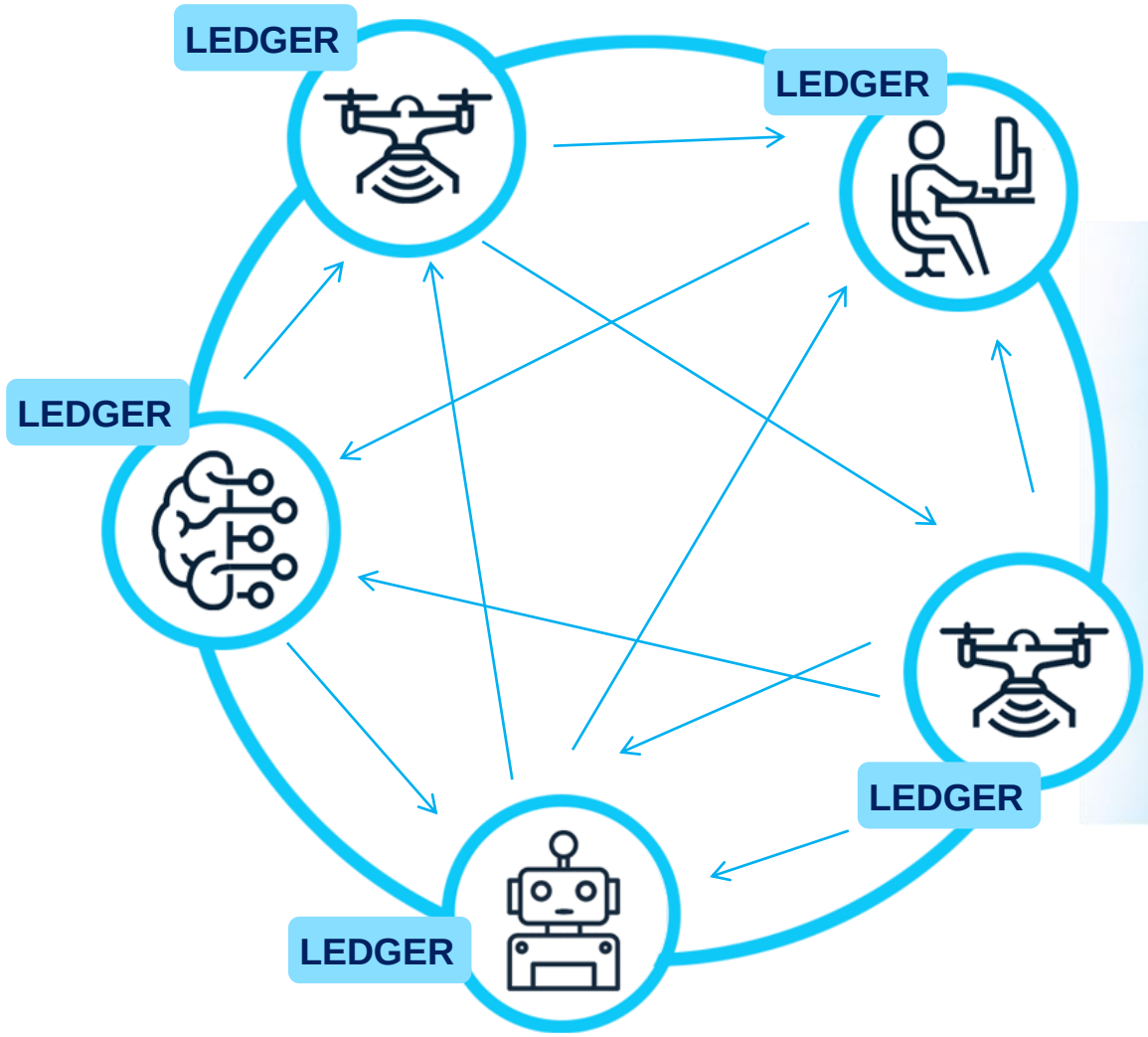


Fei Xiong, Aijing Li, et al, Published in IEEE Communications Magazine 21 August 2019
Engineering, Computer Science, An SDN-MQTT Based Communication System for Battlefield UAV Swarms | Semantic Scholar

C-DLT

- Real-World Data
- Data Synchronization
- Data Integrity
- Data Accessibility
- AI Analysis (e.g. Anomaly Detection)

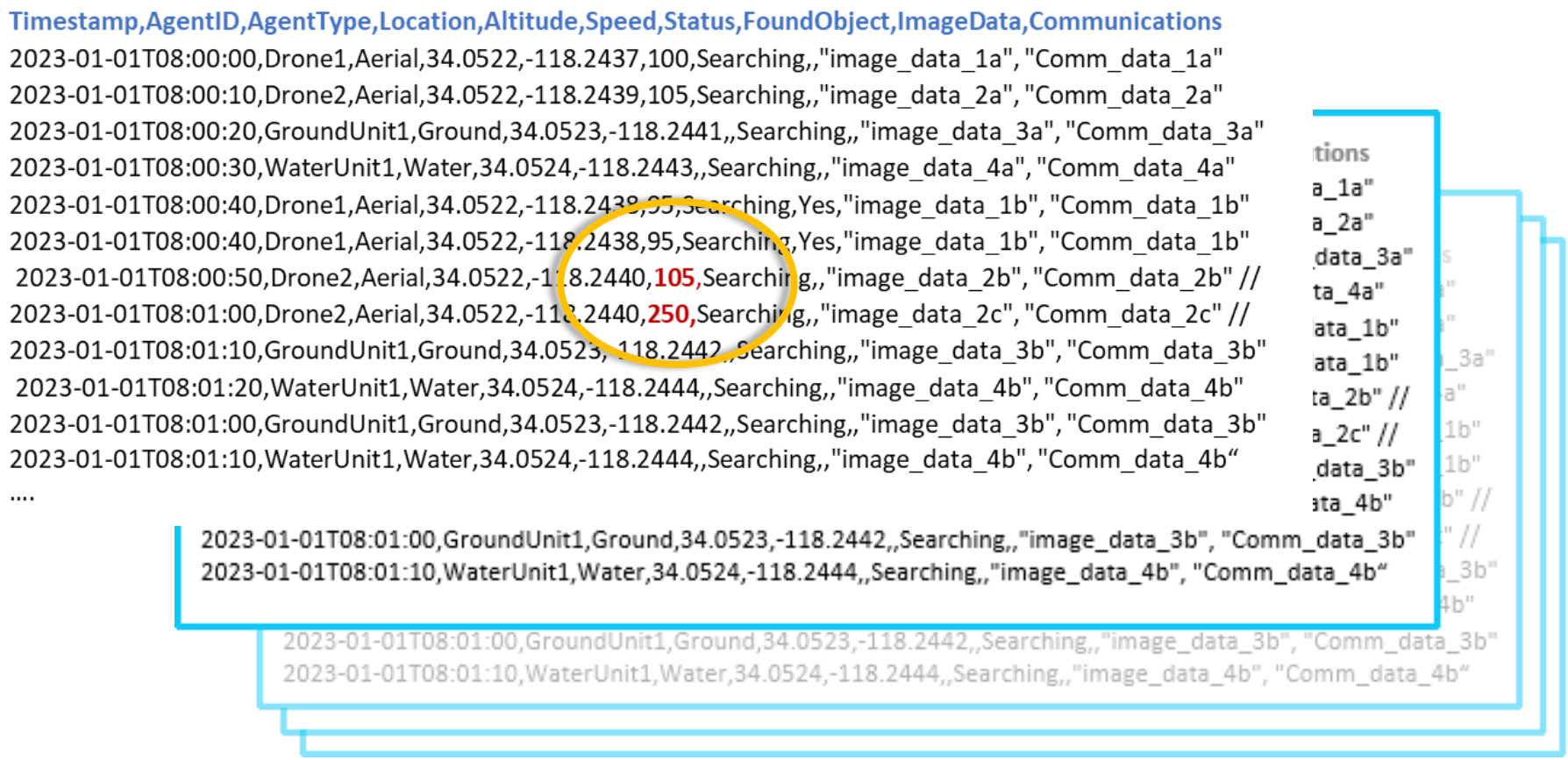
All Data is captured in the DLT:



Agents collect and log data on C-DLT (sensor data, environment, interactions, decision/actions.

```
Timestamp,AgentID,AgentType,Location,Altitude,Speed,Status,FoundObject,ImageData,Communications
2023-01-01T08:00:00,Drone1,Aerial,34.0522,-118.2437,100,Searching,,"image_data_1a", "Comm_data_1a"
2023-01-01T08:00:10,Drone2,Aerial,34.0522,-118.2439,105,Searching,,"image_data_2a", "Comm_data_2a"
2023-01-01T08:00:20,GroundUnit1,Ground,34.0523,-118.2441,,Searching,,"image_data_3a", "Comm_data_3a"
2023-01-01T08:00:30,WaterUnit1,Water,34.0524,-118.2443,,Searching,,"image_data_4a", "Comm_data_4a"
2023-01-01T08:00:40,Drone1,Aerial,34.0522,-118.2438,95,Searching,Yes,"image_data_1b", "Comm_data_1b"
2023-01-01T08:00:40,Drone1,Aerial,34.0522,-118.2438,95,Searching,Yes,"image_data_1b", "Comm_data_1b"
 2023-01-01T08:00:50,Drone2,Aerial,34.0522,-118.2440,105,Searching,,"image_data_2b", "Comm_data_2b" //
2023-01-01T08:01:00,Drone2,Aerial,34.0522,-118.2440,250,Searching,,"image_data_2c", "Comm_data_2c" //
2023-01-01T08:01:10,GroundUnit1,Ground,34.0523,-118.2442,,Searching,,"image_data_3b", "Comm_data_3b"
 2023-01-01T08:01:20,WaterUnit1,Water,34.0524,-118.2444,,Searching,,"image_data_4b", "Comm_data_4b"
2023-01-01T08:01:00,GroundUnit1,Ground,34.0523,-118.2442,,Searching,,"image_data_3b", "Comm_data_3b"
2023-01-01T08:01:10,WaterUnit1,Water,34.0524,-118.2444,,Searching,,"image_data_4b", "Comm_data_4b“
....
```

Monitor & Review Ledger Data and apply onboard tiny-AI and AI-enabled post-processing for Data Analysis



Example T&E Applications with extensibility to continuous life-cycle monitoring:

- e.g. The algorithm detects an anomaly in Drone2's speed, as it increased suddenly from 105 to 250 within a short time.

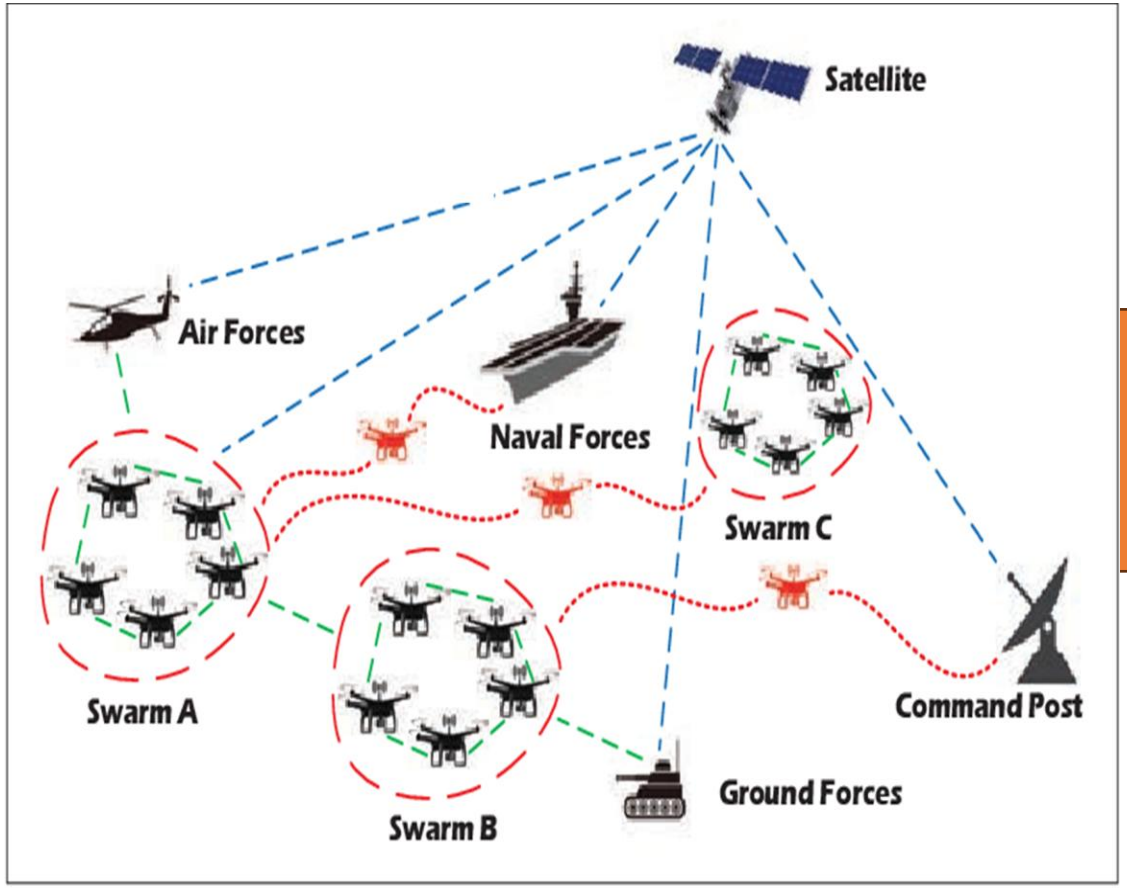
- e.g. Another anomaly detected is a communication delay between GroundUnit1 and other agents, indicating potential issues with the communication module or environmental interference.

Summary:

The proposed C-DLT framework addresses these considerations by enabling onboard and in-situ monitoring combined with system-wide record synchronizations to capture the real-world context and dynamics of inter-agent behaviors within a global frame of reference model (and common operating picture).

Additionally, the proposed method provides for recurring periodic testing of operational AI/ML-based systems, emphasizing and addressing the dynamic nature of collaborative adaptive Continuous Learning Systems (CLS) as they incorporate new training data, and adapt to new operational environments and changing environmental conditions. The performance trade-offs and T&E challenges that arise within this vision of CA2IS underscore the necessity for the proposed in-situ testing method.

Potential of In-situ C-DLT to enhance Modeling & Simulation:



C-DLT

- Real-World Data
- Data is Synchronized
- Data has Integrity
- Data is Accessible to All stakeholders
- AI Analysis (e.g. Anomaly Detection)

Modeling & Simulation

Challenges and Limitations:

Scalability: Increased number of agents and data volume can lead to increased latency* and higher computational and storage requirements

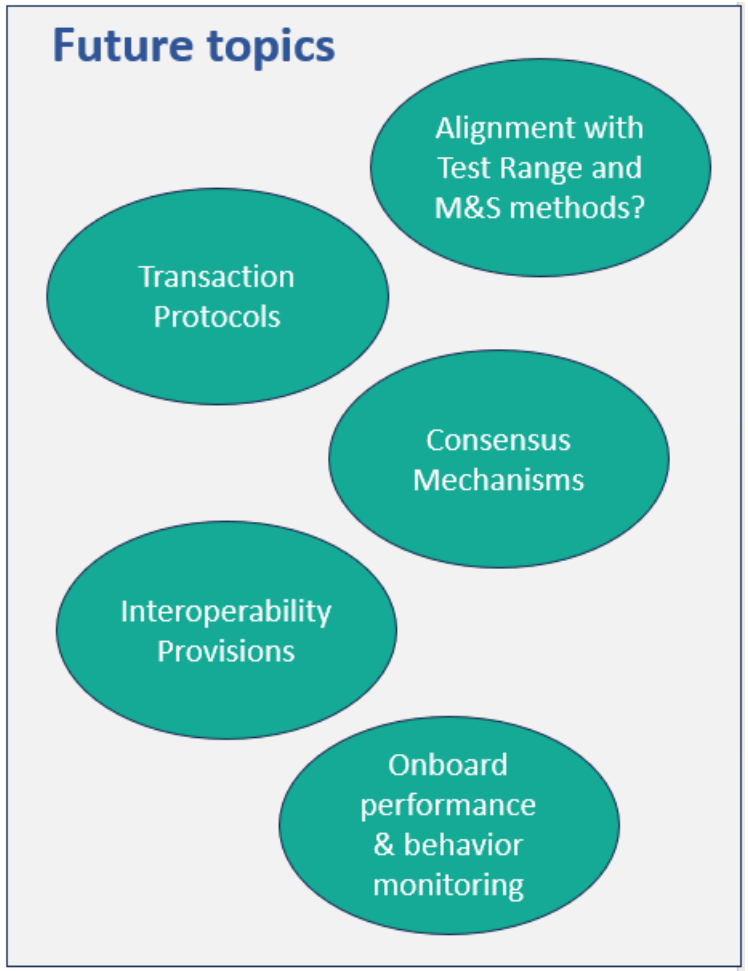
Privacy and Data Security: Ensuring confidentiality and privacy of sensitive data (e.g. proprietary technical info)

Standardization: Need for standardized protocols and data formats for communication.

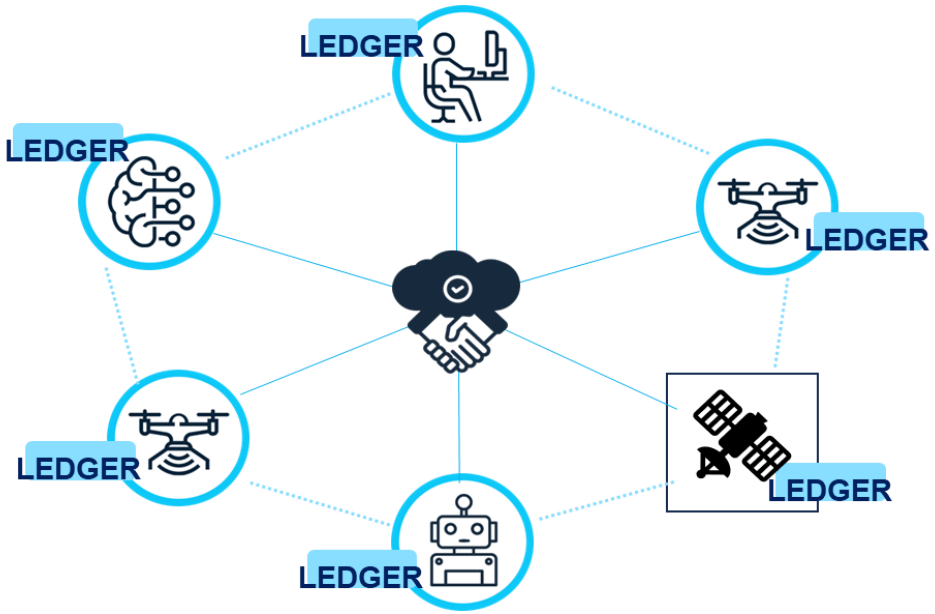
Complexity: Implementing and Managing Distributed Ledger Technology can be complex

Legal and Regulatory Issues: Legal and regulatory issues such as data ownership, liability, and compliance with data protection regulations.

*deemed not as critical in the envisioned T&E applications



Wrap-up:



The proposed in-situ T&E framework presents a promising solution for

BOTH

testing CA2IS

AND

for Increasing the operational Resilience & Performance of AI-enabled Systems

All agents in real-world conditions

Need further research and real-world applications to fully explore and validate

Questions?

For more detailed info: Journal Paper in International Test and Evaluation Journal

Transforming the Testing and Evaluation of Autonomous Multi-Agent Systems: Introducing In-Situ Testing via Distributed Ledger Technology

[Transforming the Testing and Evaluation of Autonomous Multi-Agent Systems: Introducing In-Situ Testing via Distributed Ledger Technology – International Test and Evaluation Association (itea.org)]

Recent & Enabling Research Findings (key topics):

1. S. Harshbarger & M. Collins, Advancing Resilient Collaborative Adaptive AI-Enabled Systems (CA2IS): A Community Partnership Initiative, Volume One: Overview, 2024
2. Excerpt of Recommendations from: Federal Cybersecurity Research and Development Strategic Plan Section 4.3 & Corresponding Framing, White House Office of Science & Technology Policy, National Science and Technology Council, USG, 2023
3. John James, K. Duncan, M. Collins, et al., An experimental framework for investigating Hashgraph algorithm transaction speed. In Proceedings on Blockchain-enabled Sensing, Association for Computing Machinery, 2020.
4. M. Shabbir, Xenofon Koutsoukos, et al., "Resilient Vector Consensus in Multi-Agent Networks Using Centerpoints", American Control Conference (ACC), Denver, CO, 2020.
5. J. Li, W. Abbas, M. Shabbir, and Xenofon Koutsoukos. Byzantine resilient distributed learning in multirobot systems. IEEE Trans-actions on Robotics, 2022.
6. Bhowmick & Koutsoukos, Resilient Peer-to-peer Learning based on Adaptive Aggregation, IEEE International Conference on Distributed Computing and AI, (Best Paper Award), 2024.
7. A. White, et al., US Patent # 12088569, entitled - Protocol Free Encrypting Device, Describing a method of paired encrypting devices to allow for communication of trusted data between trusted devices over an untrusted network, Provisional Application, January 2023.
8. K. Giammarco, M. (Misha) Novitzky, John James, S. Harshbarger, M. Collins, et al., "State machine execution traces for verifying and validating robot behaviors," Proc. SPIE 12544, Open Architecture/Open Business Model Net-Centric Systems and Defense Transformation, 2023.
9. K. Giammarco, M. (Misha) Novitzky, John James, S. Harshbarger, M. Collins, et al., "Behavior analysis of search and rescue operations employing human-machine teaming," Proc. SPIE 12544, Net-Centric Systems and Defense Transformation, 2023.
10. M. Kutzer, et al., Uncertainty Informed Position Estimation Using Multi-view Cameras. Journal of Intelligent & Robotic Systems, [Pending], 2024.
11. O'Brien, Dawkins, Galloway & Kutzer, Sensor Fusion for Multi-Robot Search and Rescue using the Unscented Kalman Filter. IEEE International Symposium on Safety, Security, and Rescue Robotics, [Pending], 2024.
12. O'Brien & Kutzer, Unscented Kalman Filtering for Localization using Range or Bearing Data, 14th International Conference on Control. IEEE, 2023
13. Civetta & Kutzer, Toward Position Approximation Using Asynchronous Multi-View Cameras: A 2D Investigation with Probabilistic Considerations. ASME International Mechanical Engineering Congress and Exposition, American Society of Mechanical Engineers, 2023.
14. S. Harshbarger, R. Heckle & M. Collins, Transforming the Testing and Evaluation of Autonomous Multi-Agent Systems: Introducing In-Situ Testing via Distributed Ledger Technology, The ITEA Journal of Test and Evaluation, Volume 45, Issue 1, 2024.
15. B. Teague, Z. Liu, et al., Network Localization and Navigation with Scalable Inference and Efficient Operation, IEEE, Transactions on Mobile Computing, Vol. 21, NO. 6, 2022.

Abbreviated Bibliography of Underpinning Works by collaborating partners.

Representative Snapshot of C-DLT data content (conceptual):

Timestamp	Agent ID	Event Type	Sensor Data	Position	AI/ML Model Version	AI/ML Model Score
10:00:01	Drone1	Deployed	N/A	(10,20)	v1.0	N/A
10:00:02	Drone2	Deployed	N/A	(15,25)	v1.0	N/A
10:00:03	Drone1	Scanning	Normal	(10,21)	v1.0	0.95
10:00:04	Drone2	Scanning	Normal	(15,26)	v1.0	0.96
10:00:05	Drone1	Scanning	Normal	(10,22)	v1.0	0.95
10:00:06	Drone2	Scanning	Normal	(15,27)	v1.0	0.96
10:00:07	Drone1	Scanning	Normal	(10,23)	v1.0	0.95
10:00:09	Drone2	Scanning	Normal	(15,28)	v1.0	0.96
10:00:10	Drone1	Scanning	Normal	(10,24)	v1.0	0.95
10:00:11	Drone2	No Signal	N/A	Unknown	N/A	N/A