SYSTEMS ENGINEERING RESEARCH CENTER

AIRC | ACQUISITION INNOVATION RESEARCH CENTER

SE IN THE ERA OF HUMAN-MACHINE TEAMING ROADMAP FOR AI AND SE

Tom McDermott, Systems Engineering Research Center

NBC NEWS

Complaint claims Tesla's 'Full Self-Driving' software caused crash

14 November 2021

US safety regulator opens investigation into Tesla Autopilot following crashes with parked emergency vehicles

U.S. auto regulators have opened a preliminary investigation into Tesla's Autopilot advanced driver assistance system, citing 11 incidents in which vehicles crashed into parked first responder vehicles while the system was engaged. The Tesla vehicles involved in the collisions were confirmed to have either have had engaged Autopilot or a feature called Traffic Aware Cruise ... Continue reading

TechCrunch 0

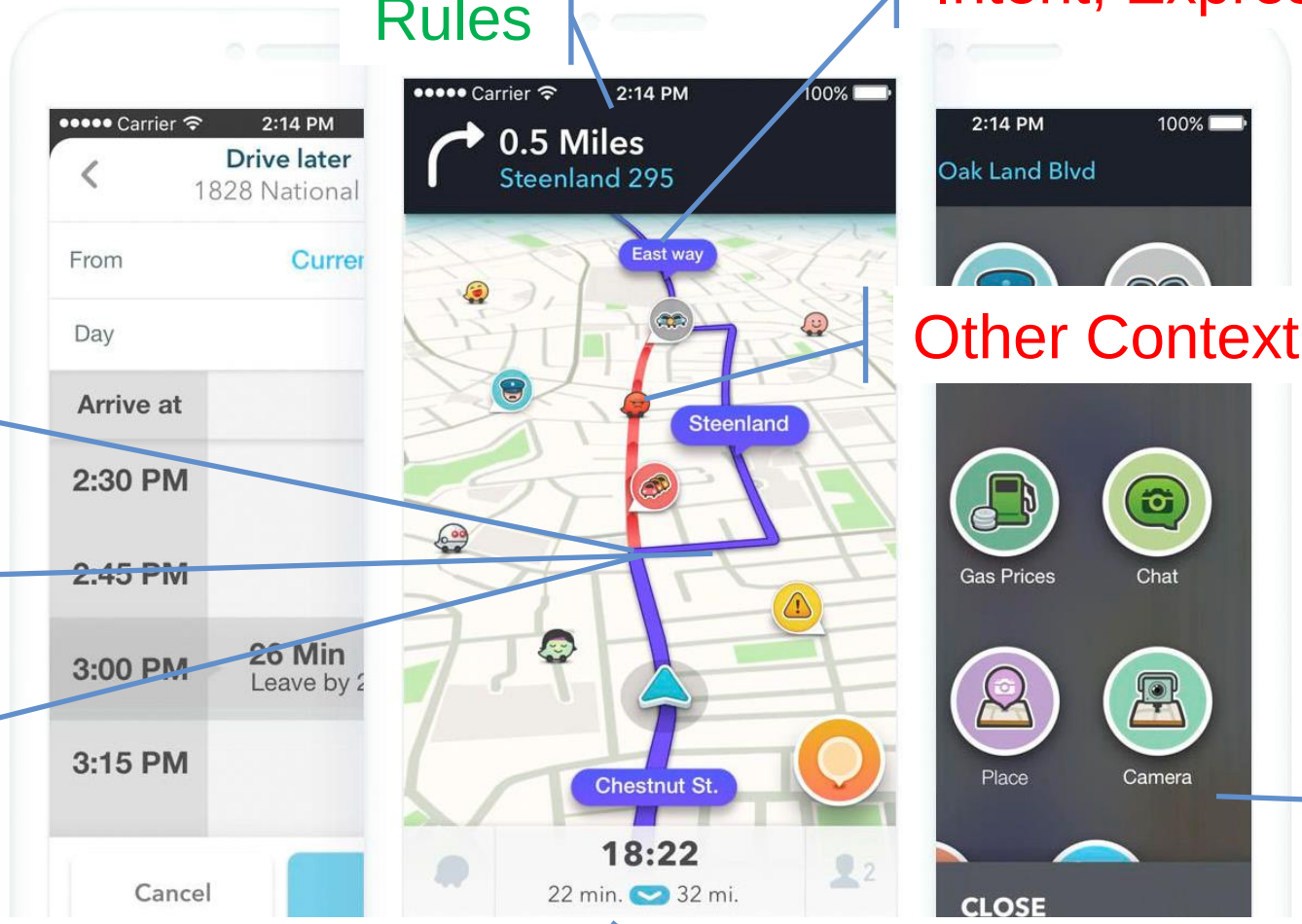
Distributed Autonomy	Human-Machine Teaming and Co-learning

NY Times photo

There are 9.1 driverless car crashes per million miles driven. Regular vehicles have a rate of 4.1 crashes per million miles driven. Fewer severe injuries are caused by self-driving cars. (carsurance.net/insights/self-driving-car-statistics)

Transfer of Authority between human and machine remains a concern.

WAZE Human-Machine Teaming



Rules

Intent, Expressives

Other Contextual Information

Action

Authorities

Order

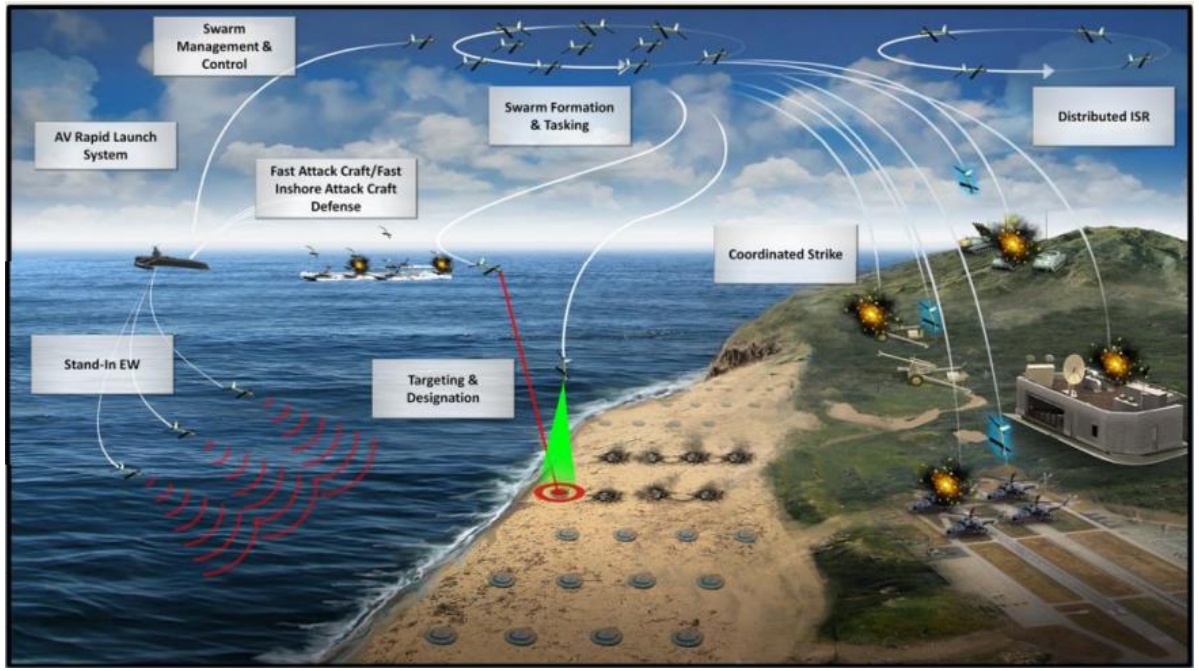
Interpretation of Information

Effect

Waze image: www.cnet.com/roadshow/news/google-assistant-waze-easier-reporting-less-distraction/

Hierarchical Control of Distributed Autonomous Human-Machine Teams

- Stochastic decision processes
- Controlled by both machine agents and humans
- Ideally leverage the distinct capabilities of each
- Must address the challenge of transferring control quickly, safely, and smoothly back-and-forth between the agent and the human
- Can be viewed as hierarchical levels of control using non-hierarchical distribution of information

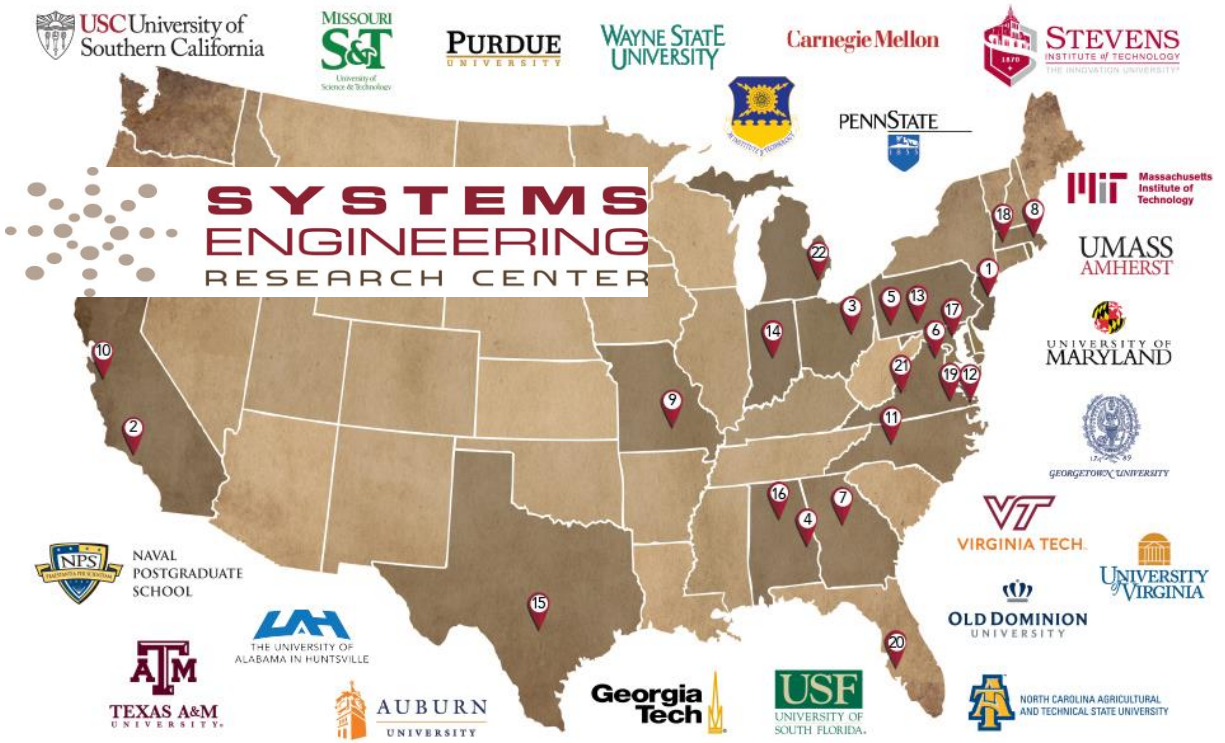


Office of Naval Research, Code 30 overview briefing

the Future of Systems Engineering

- How do we prepare the future systems engineering process in a world where humans and machines co-adapt to evolve a complex mission in response to dynamic operational conditions?

INCOSE FuSE
Future of Systems Engineering

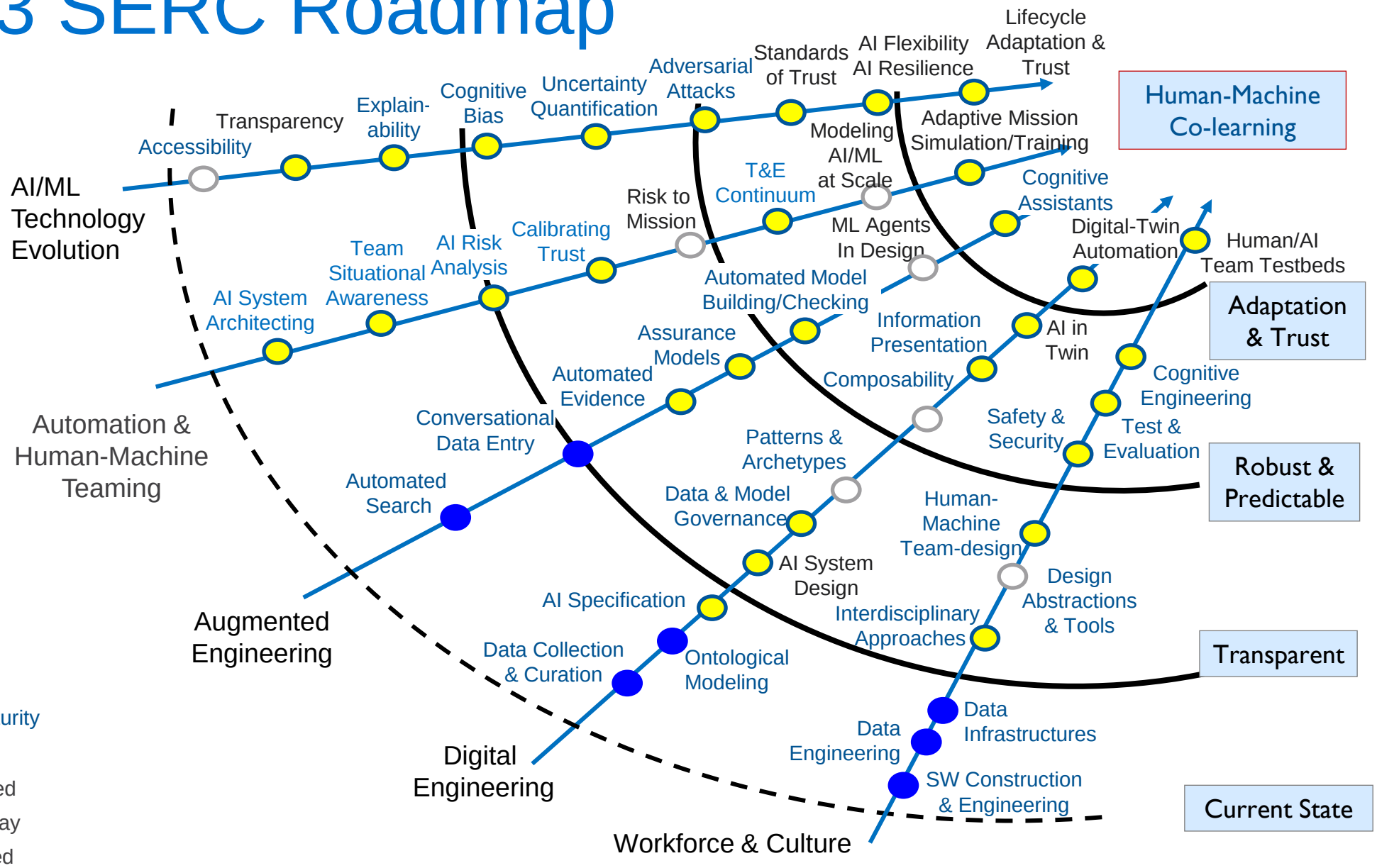


Why Model?

- Most accidents/mission failures will be caused by errors in interpretation of information by either the human or the machine
- Leading to errors transfer of control or authority made in the planning process
- Underlying concept of human informational transfer has subjectivity
  - **Intent**
  - **Rules**
  - **Authorities**
  - **Other Contextual Information**
- Desire a Systems Engineering approach to address both information design and control mechanization across layers of hierarchy

- Consistently used in hierarchical control structures
- Lack of multi-disciplinary research

2023 SERC Roadmap



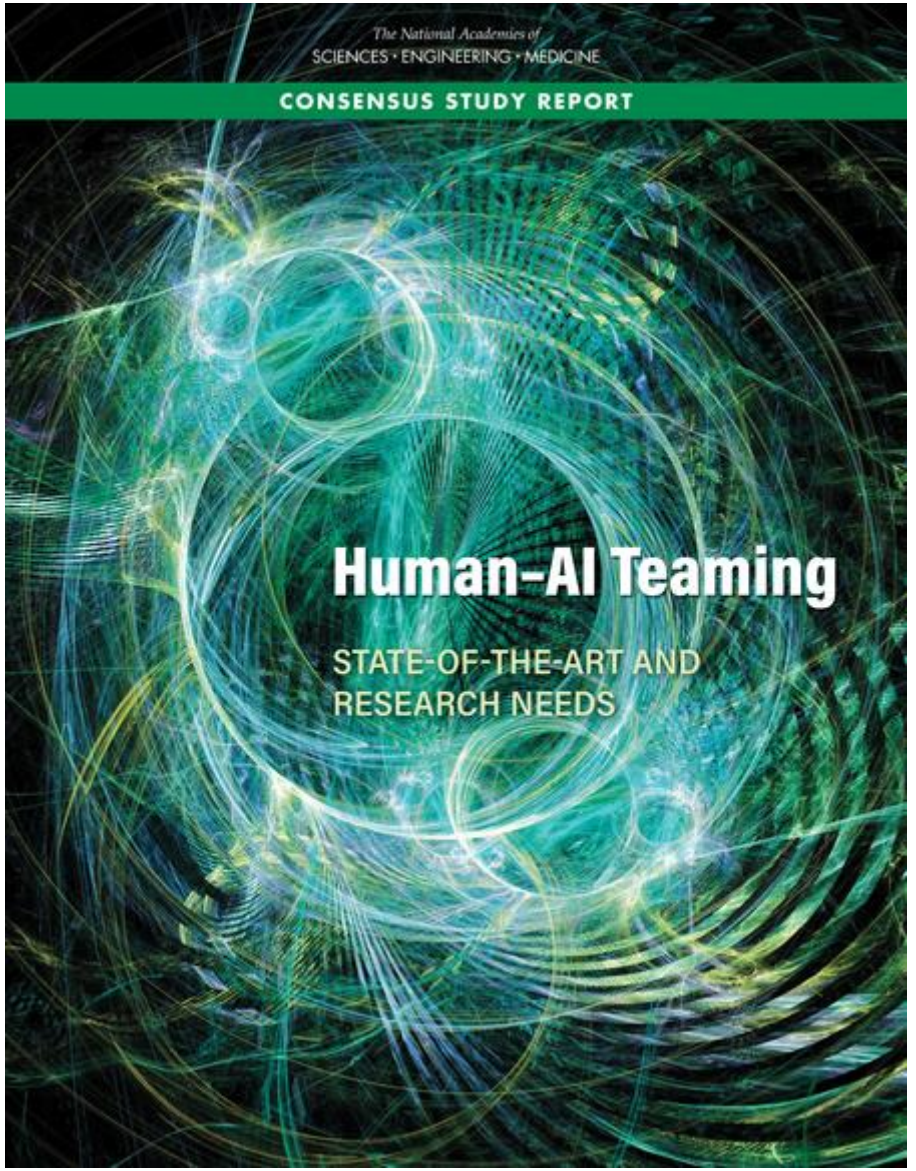
SERC Capability Maturity

- Transitioned
- Transition Needed
- Projects Underway
- Research Needed

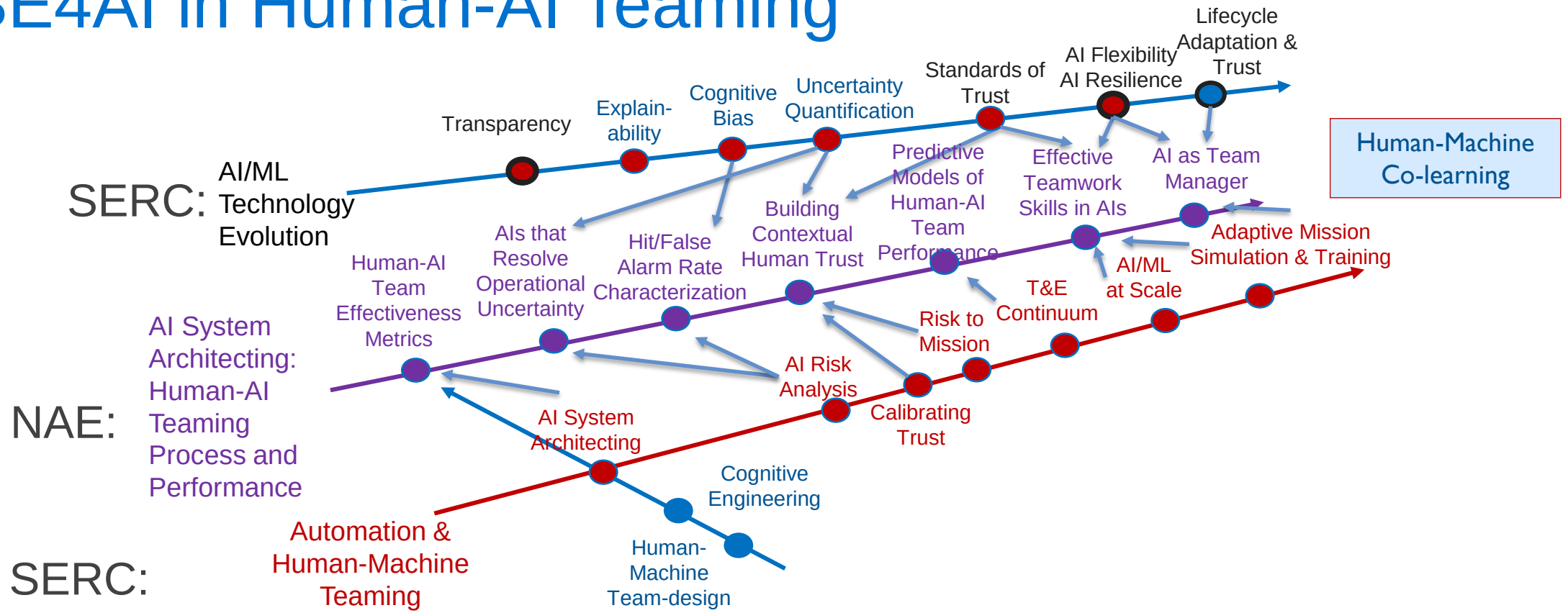
SE/HSI Objectives

Significant value in considering the human and AI as a team

- Long-term, distributed, and agile human-AI teams through improved team assembly, goal alignment, communication, coordination, social intelligence, and the development of a new human-AI language – **AI System Architecting**
- Methods for improving human situational awareness of AI systems
- Improved AI system transparency and explainability
- Interaction mechanisms and strategies within the human-AI team
- Advance understanding of how broader sociotechnical factors affect trust in human-AI teams
- Better understand the interdependencies between human and AI decision-making biases, how these evolve over time, and methods for detecting and preventing bias
- What, when, why, and how to best train human-AI teams
- **Advances in HSI processes and measures**



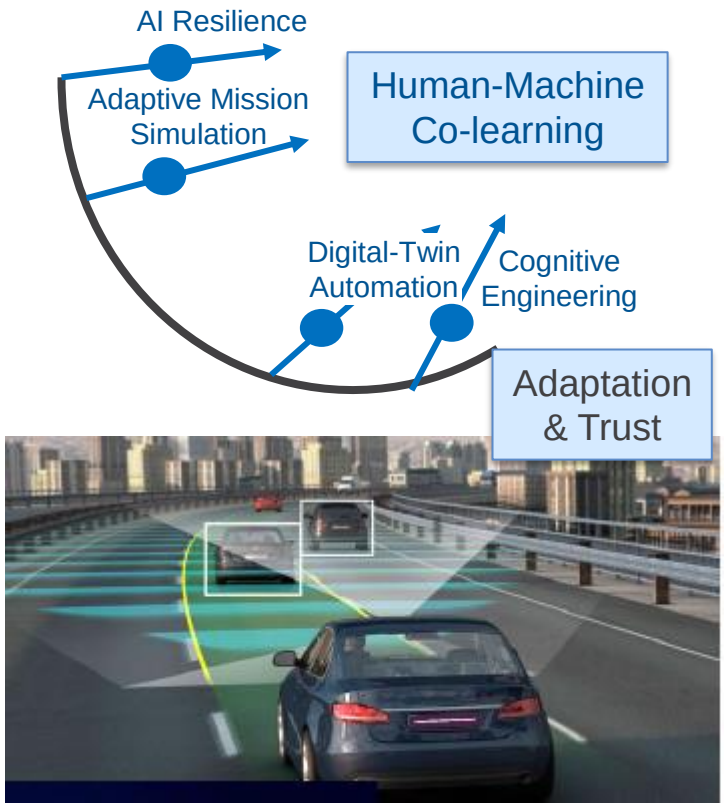
SE4AI in Human-AI Teaming



- Long-term, distributed, and agile human-AI teams through improved team assembly, goal alignment, communication, coordination, social intelligence, and the development of a new human-AI language – AI System Architecting
- What, when, why, and how to best train human-AI teams
- Advances in HSI processes and measures

Human-Machine Co-learning

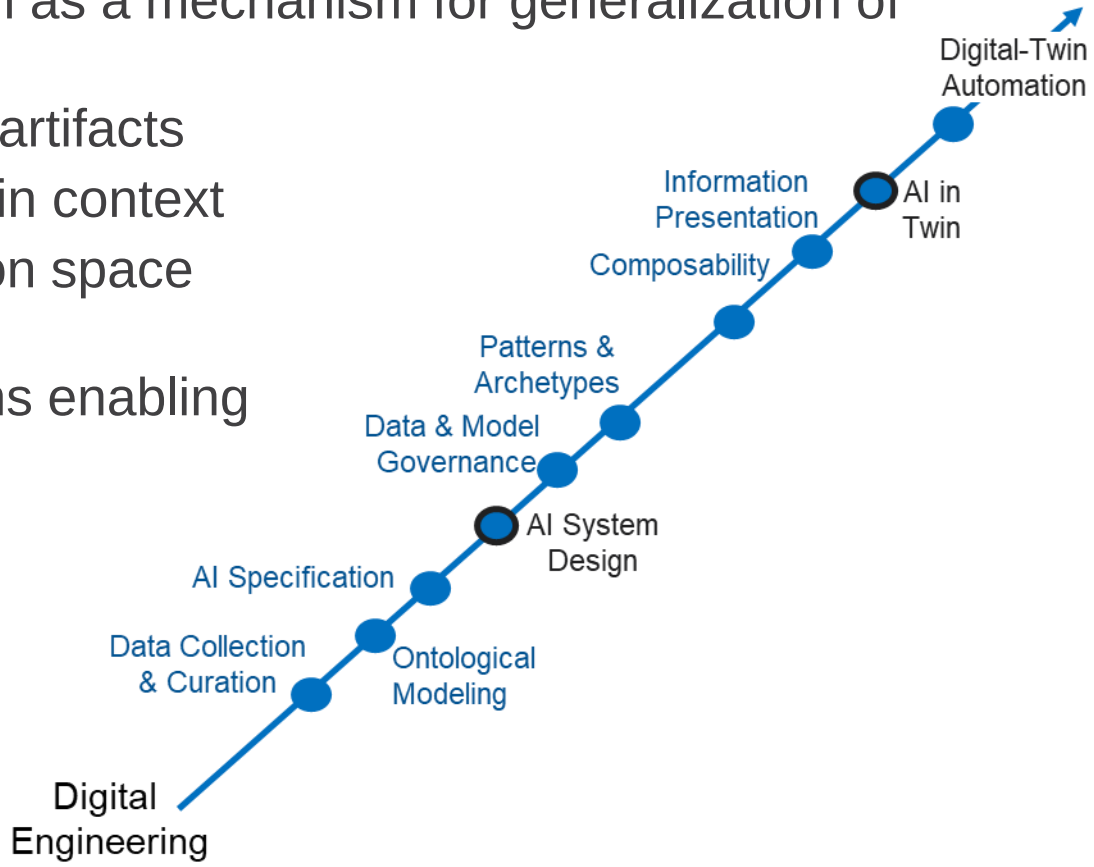
- Adaptive Cyber-Physical-Human Systems – modeling of cyber-physical systems as influenced by humans
  - Adaptive Mission Simulation – Simulation and training that supports non-static objectives (pick-up games)
  - AI Resilience – AI systems that self-adapt to changing operational boundaries while maintaining rigorous safety and security and policy constraints



AI Enabled Digital Engineering

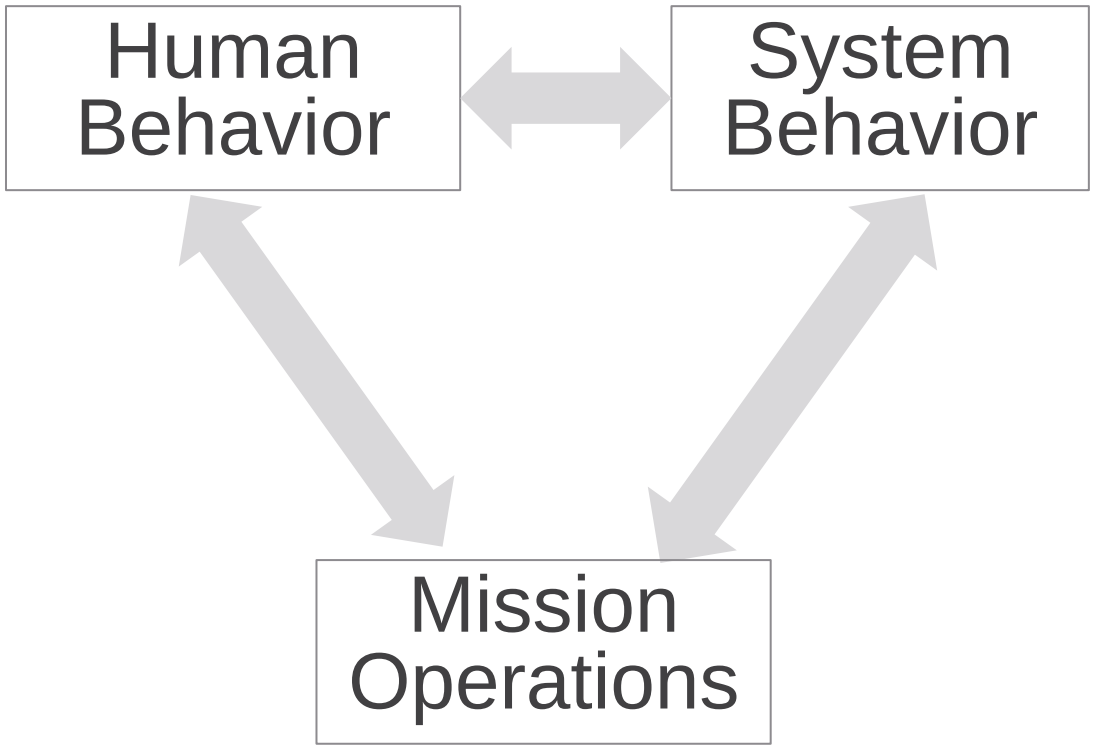
- Data Collection and Curation - data collection, management, curation and governance
- Ontological Modeling – schematic representation to semantic representation
- Specification – what will be allocated to the machine, in both product and process
- System Design for AI Performance - System design as a mechanism for generalization of AI performance factored into design activities
- Patterns and Archetypes – learning from modeling artifacts
- Composability – training and evaluating for design in context
- Information Presentation – representing the decision space for human understanding and learning
- AI in the Digital Twin - New uses of AI in digital twins enabling new functional and performance value

- Digital Twin Automation – real-time continuous learning from real system and shadow simulations

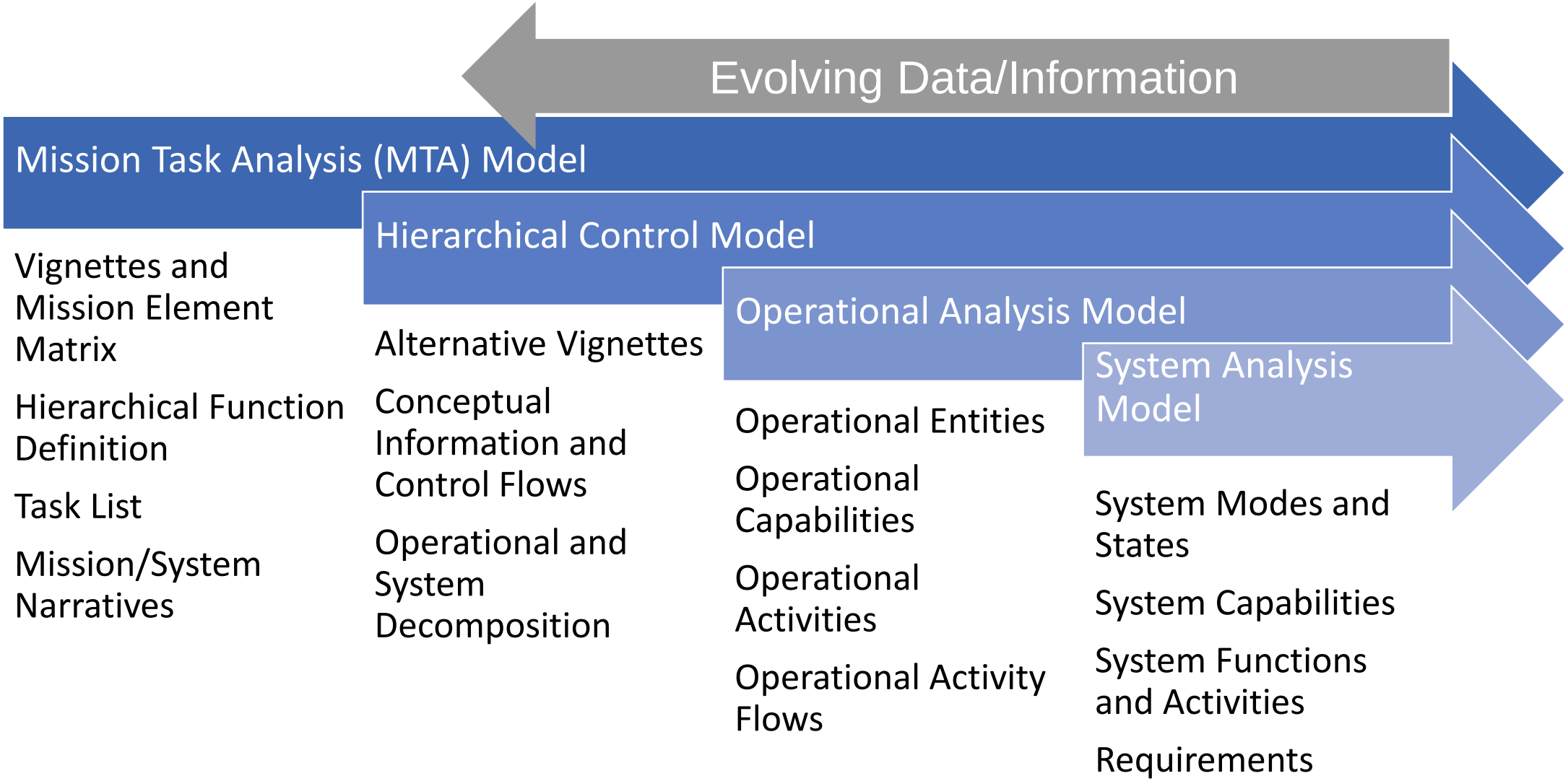


Need for Models

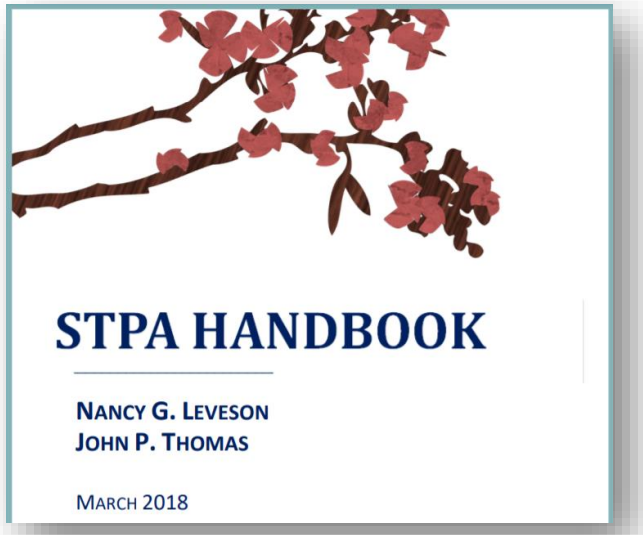
- Representing Human Behavior
  - Complexity
  - Human error
- Integrating Human and CPS
  - Need appropriate languages
  - Simulate learning over time
- Representing Uncertainty
  - Sensors monitor decision processes as well as the mission
  - Loosely coupled networks of events

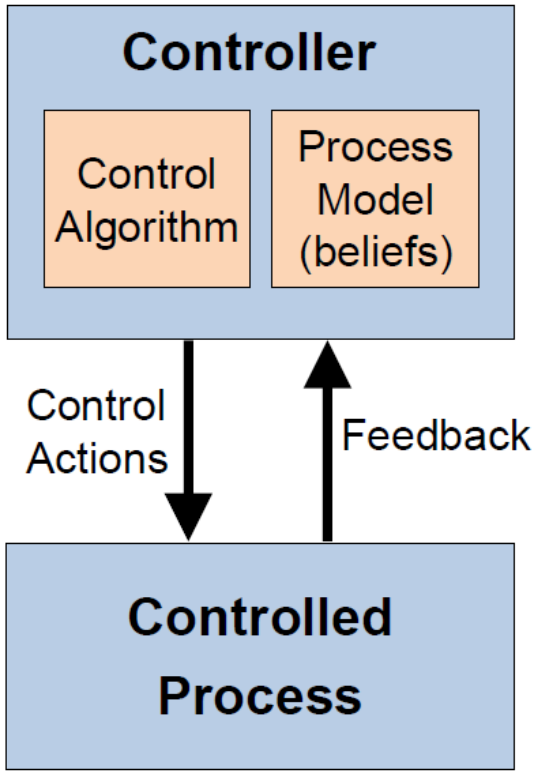


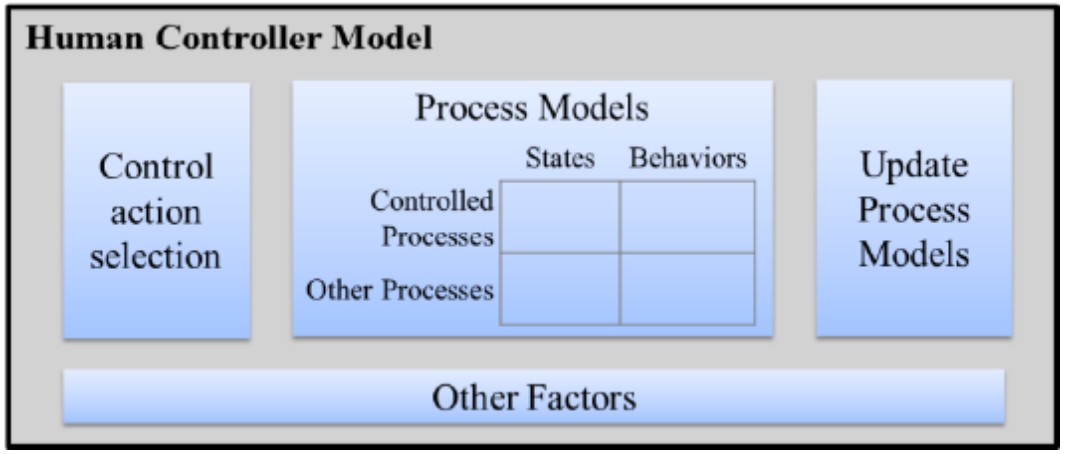
Example overall modeling flow



Control Hierarchies – System Theoretic Process Assessment

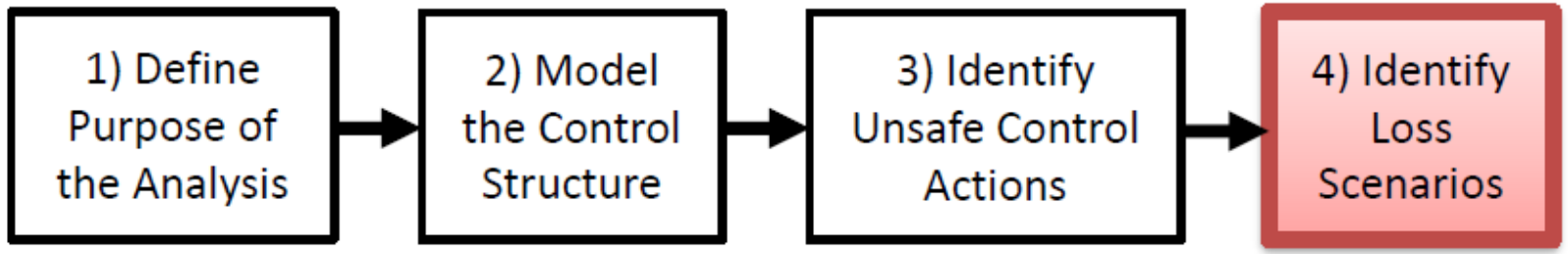






Enhancing Human Factors Analysis with STPA. Dr. John Thomas, MIT

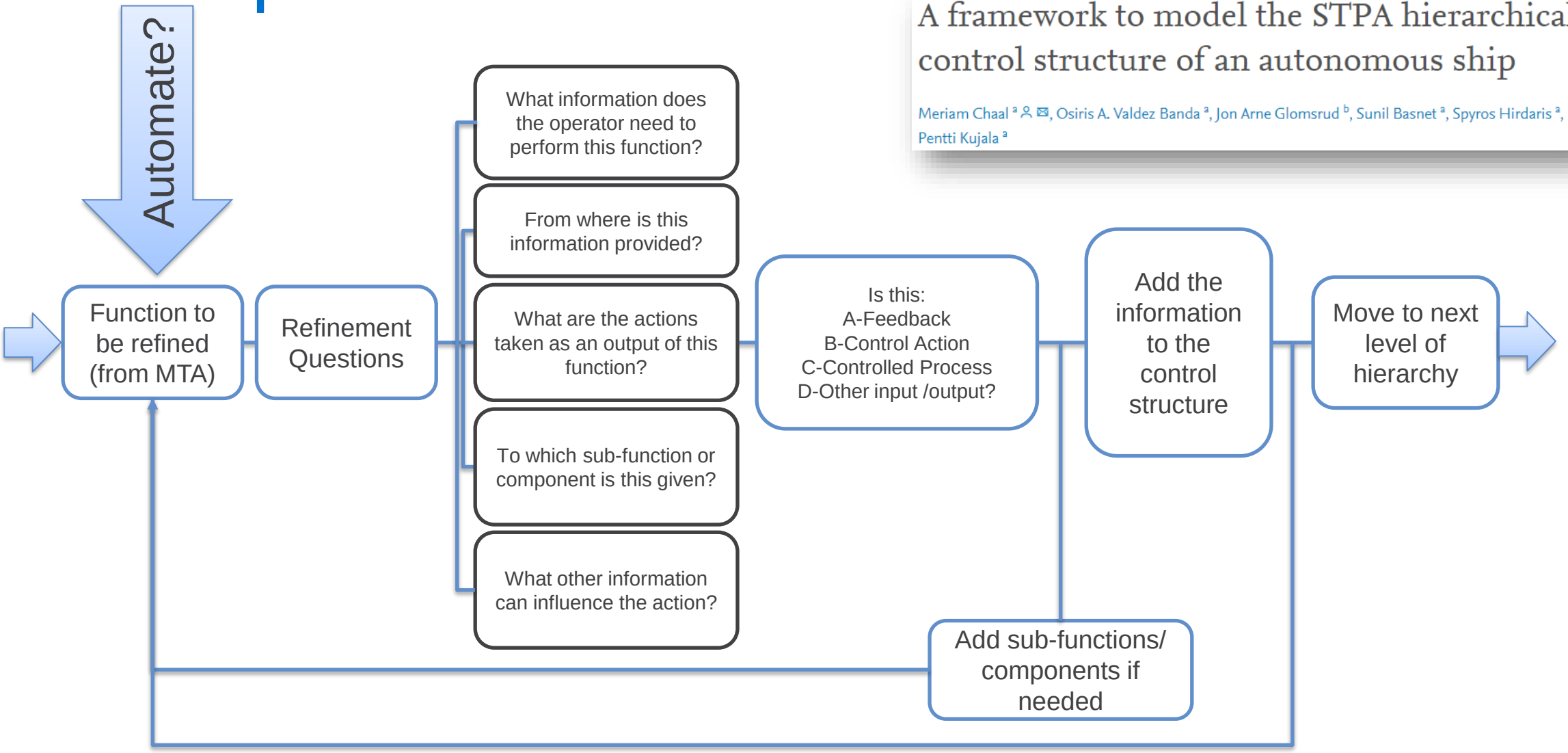
STPA



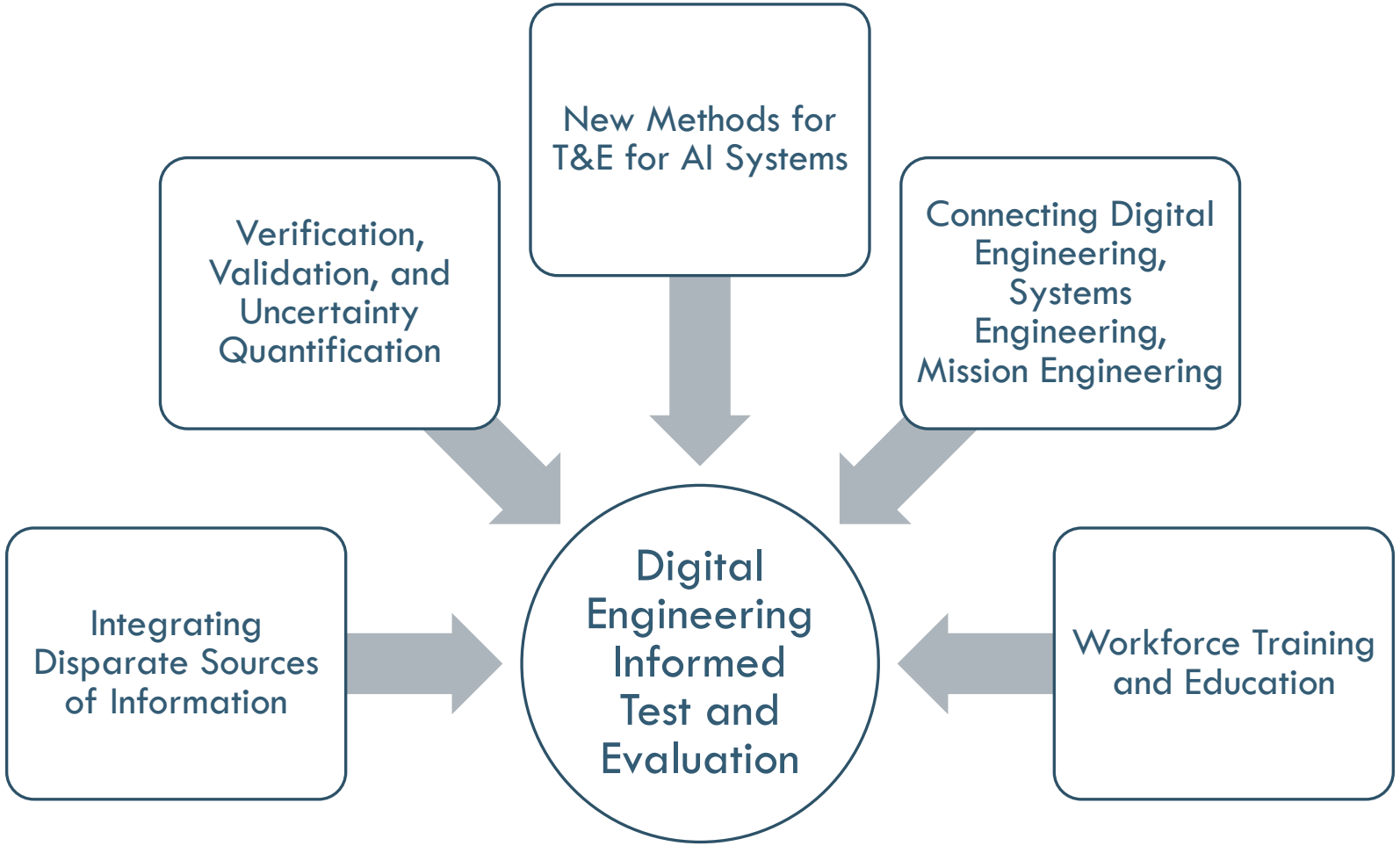
Decomposition Process

A framework to model the STPA hierarchical control structure of an autonomous ship

Meriam Chaal [a], Osiris A. Valdez Banda [a], Jon Arne Glomsrud [b], Sunil Basnet [a], Spyros Hirdaris [a], Pentti Kujala [a]

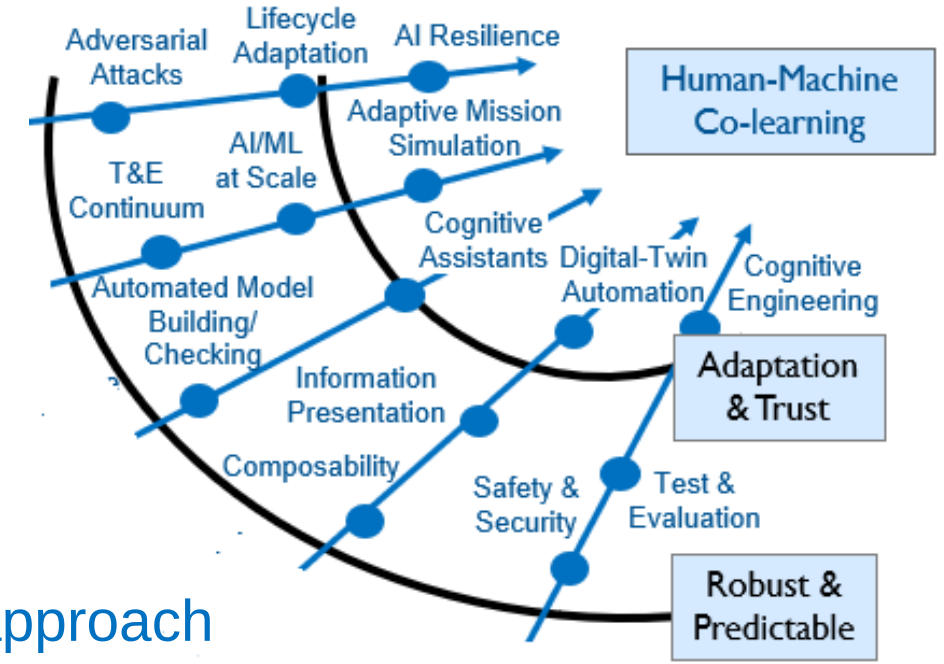


Transformation Building Blocks



Challenges for Test & Evaluation of AI

- Testing & Evaluation is a continuum
  - Information accumulates over time across varying operating envelopes
  - does not end until the system retires
- All AI areas need testbeds
- Operational relevance is essential
- Data Management is foundational
- AI systems require a probabilistic risk-based approach
- Previous test metrics apply, but may have different interpretations
  - Task & mission level performance, course of action, non-functional requirements
- An expanded definition of external context is necessary
- The T&E workforce and culture must evolve



Freeman, L. (2020), Test and Evaluation for Artificial Intelligence. INSIGHT, 23: 27-30.