SYSTEMS ENGINEERING RESEARCH CENTER

AIRC | ACQUISITION INNOVATION RESEARCH CENTER

THE UPDATED SERC AI AND AUTONOMY ROADMAP

Tom McDermott, SERC CTO, Stevens Institute of Technology

FRAMING AI AND SE

Role for Systems Engineers in AI space

AI4SE and **SE4AI**

Focuses on **application of AI in support of systems engineering processes**, enabling enhanced decision-making, optimization, and efficient effort allocation.

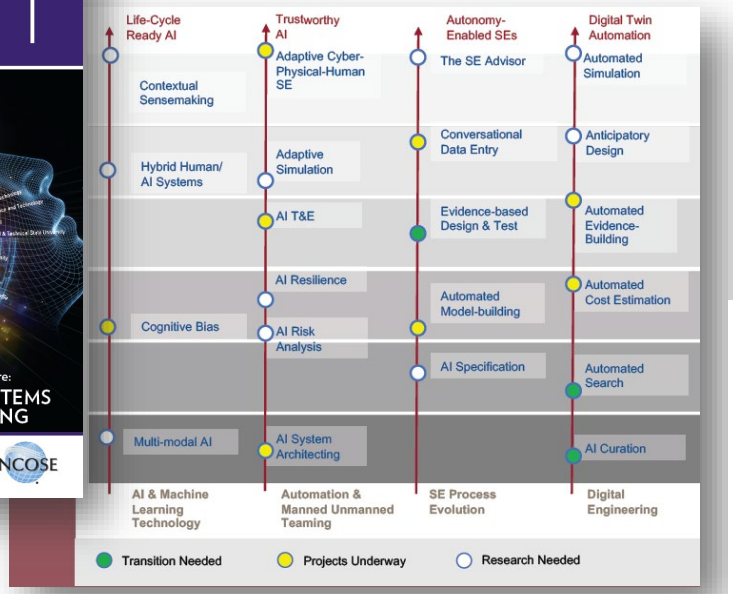
Focuses on **leveraging systems engineering principles to develop AIES that are safe,** robust, and efficient AI systems, while extending them in response to the nature of AI enabled systems.

SE4AI applies to AI4SE too, but types of AI tools tend to be different

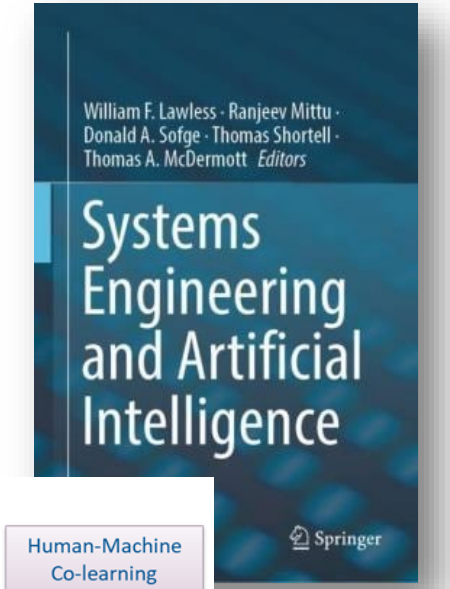
… and AI4SE might change what SEs do too.

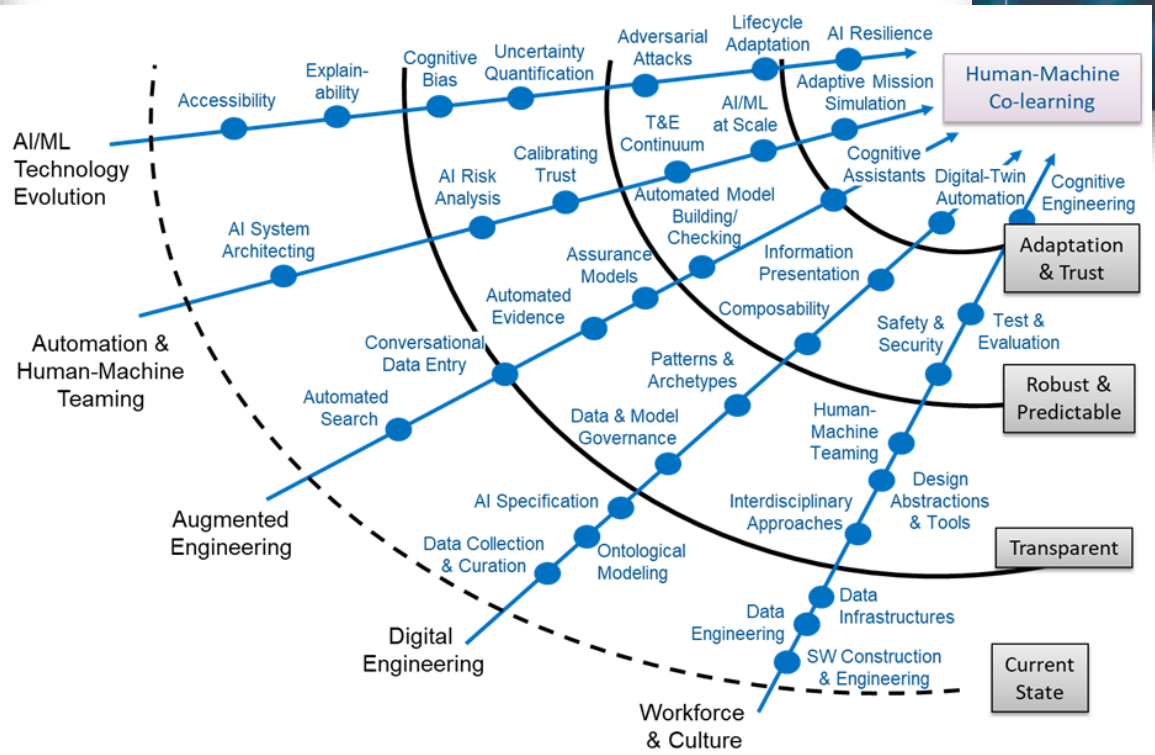
SERC ROADMAPS

INITIAL SERC AI ROADMAP



2021-22 AI ROADMAP

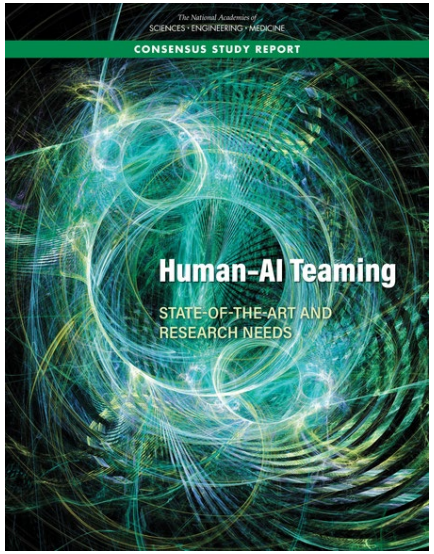


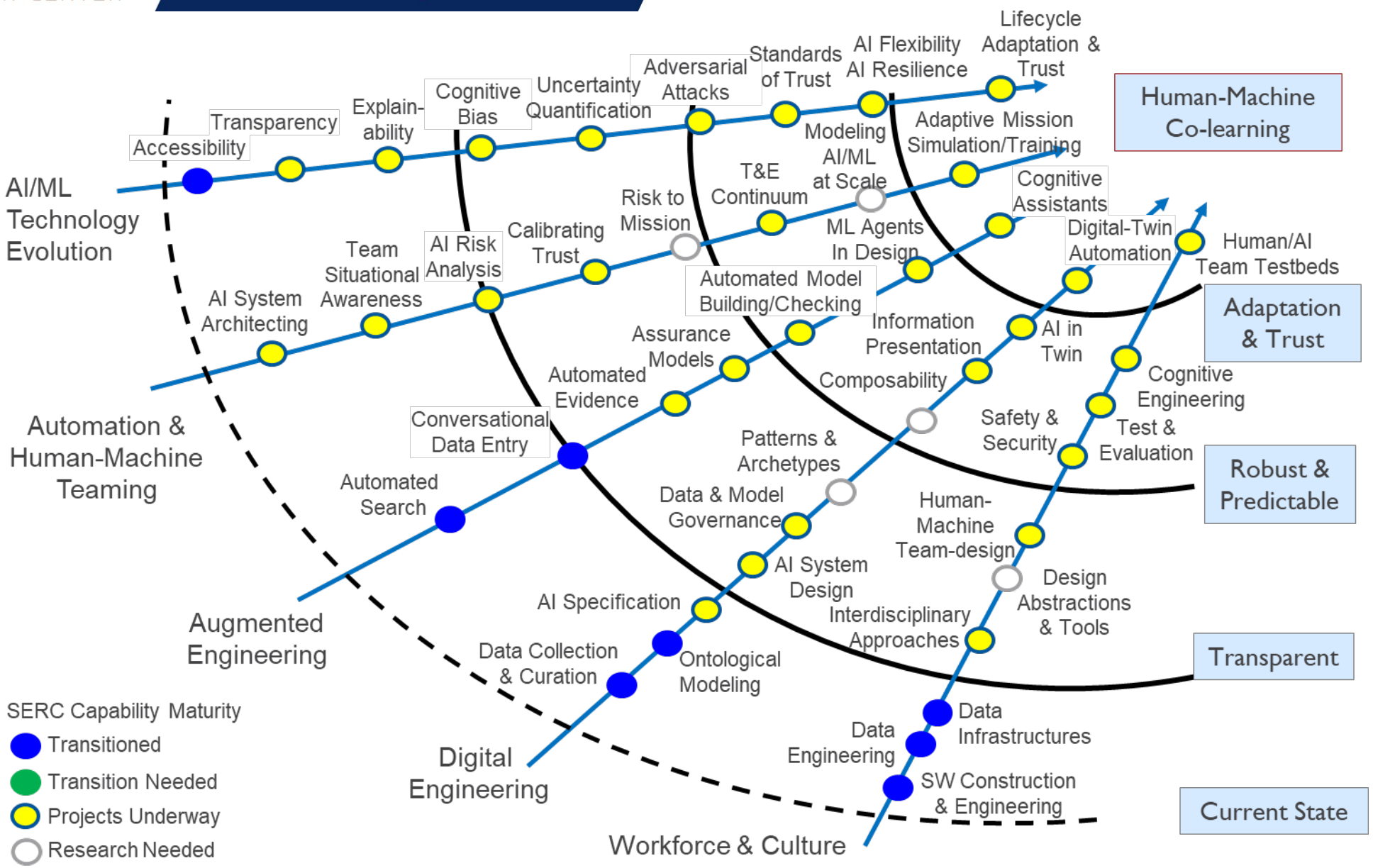


SYSTEMS ENGINEERING RESEARCH CENTER

AIRC ACQUISITION INNOVATION RESEARCH CENTER

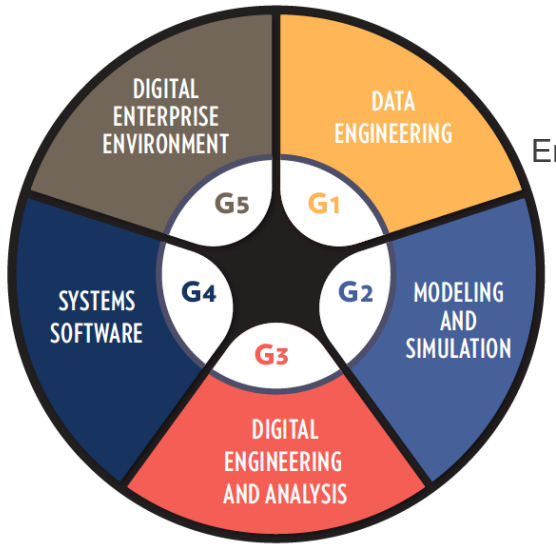
UPDATED 2024 ROADMAP

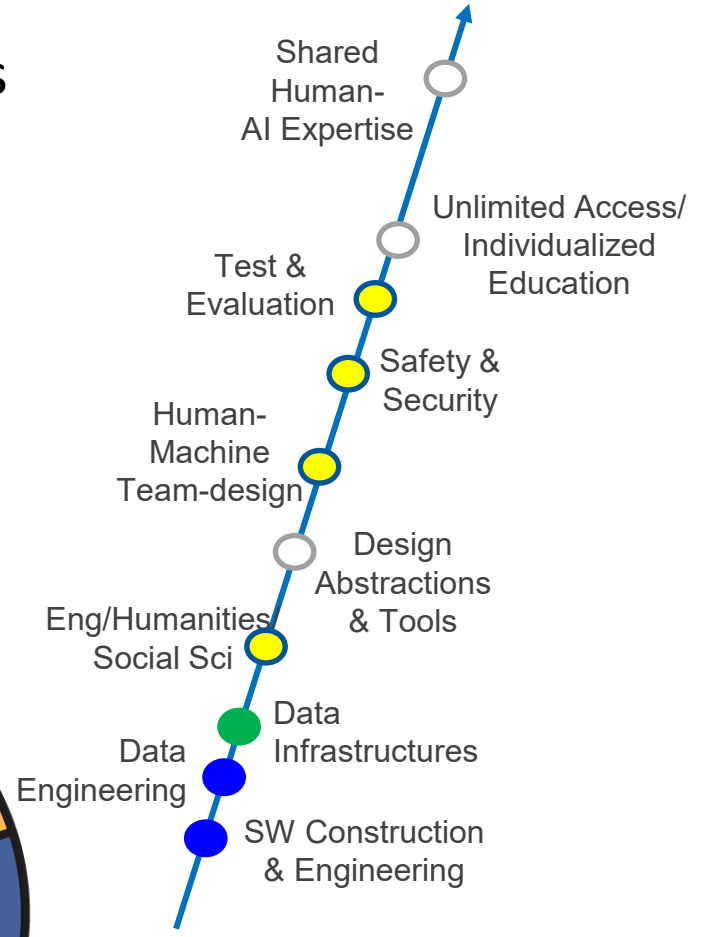




START WITH WORKFORCE AND CULTURE

- Integrating AI/ML experts with domain experts, all disciplines
- Trust: Engineering/Humanities/Social Sciences
- Evolving tools to align with design and disciplinary abstractions
- Human Systems Engineering: no longer a specialty discipline
- Threat models, safety, security, resilience, and other ‘ilities
- Evolving test and evaluation competency
- Fundamentally changing education





SERC DIGITAL ENGINEERING COMPETENCY FRAMEWORK

DIGITAL ENGINEERING TRANSFORMATION

Data Collection and Curation - data collection, management, curation and governance

Ontological Modeling – schematic representation to semantic representation

AI Specification – what will be allocated to the machine, in both product and process

AI System Design – system design as a mechanism for generalization of AI performance factored into design activities

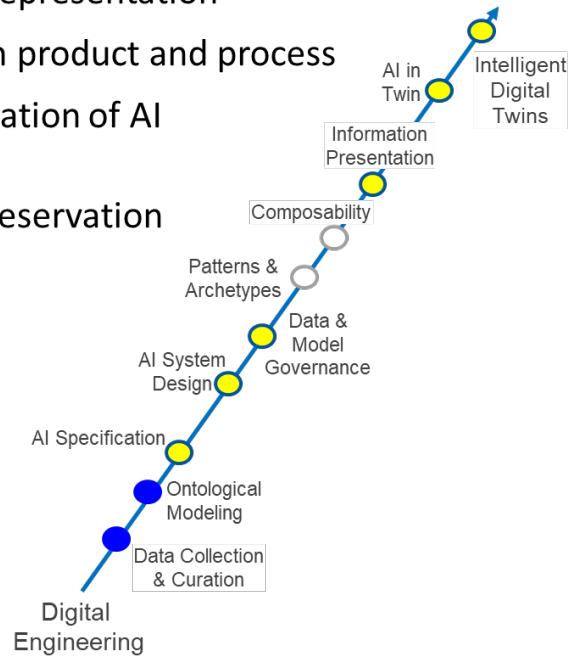
Data & Model Governance – Lifecycle management, control, preservation and enhancement of models and associated data

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

Digital Twin Automation –
AI in Intelligent Digital Twins: real-time continuous learning from real system and front-running simulations



Convergence of Data Science and Systems Engineering Disciplines

Models become central to defining complex systems of systems

Results in Product plus Digital Twins of Product

Human-Machine interfaces and Visualization of complex interrelationships

"AI4SE"

Maybe:

It's life, Jim, but not as we know it

Creating the perfect assistant…

What would you include?

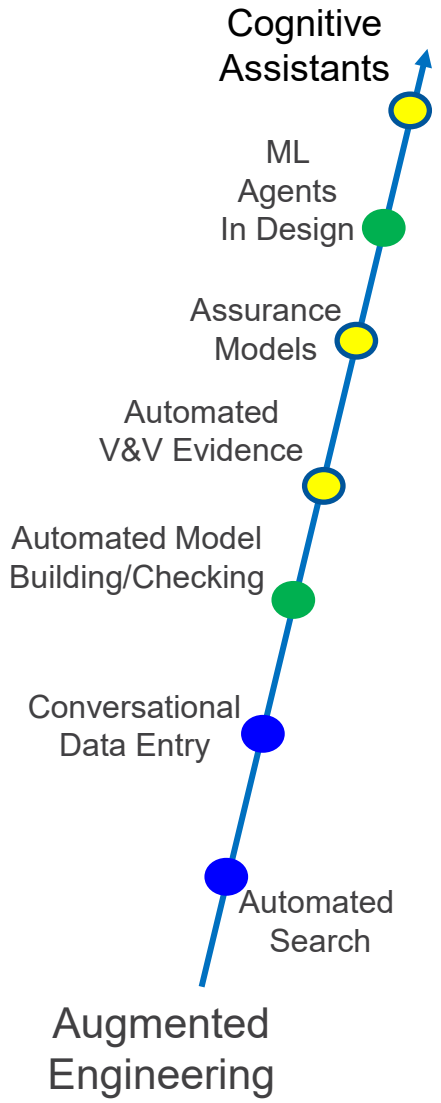
Would you make it sentient
(with rights, goals, desires, independent thoughts)

What should it know?

How should it behave?

A new kind of being:

- Neither human nor sentient
- No soul; no goals
- Vast knowledge and information
- Creativity of a sort
- Little (or no) judgment
- Great speed
- "Eager" to please
- Happy to re-do things many times



Barclay Brown, Living in a Generative World

NEW AREAS OF EMPHASIS IN SE4AI RESEARCH

Holistic view of the system of systems

Measurement of “ilities” (e.g., flexibility, resilience, trust)

Architecting / Human-system integration

Product platforms / evolvability of systems of systems

Lifecycle risk analysis

Linking “Design for X” “T&E” and lifecycle value.

Understanding human behavior as part of the system

Emergent system behavior

Building user Trust by understanding the Human AI system

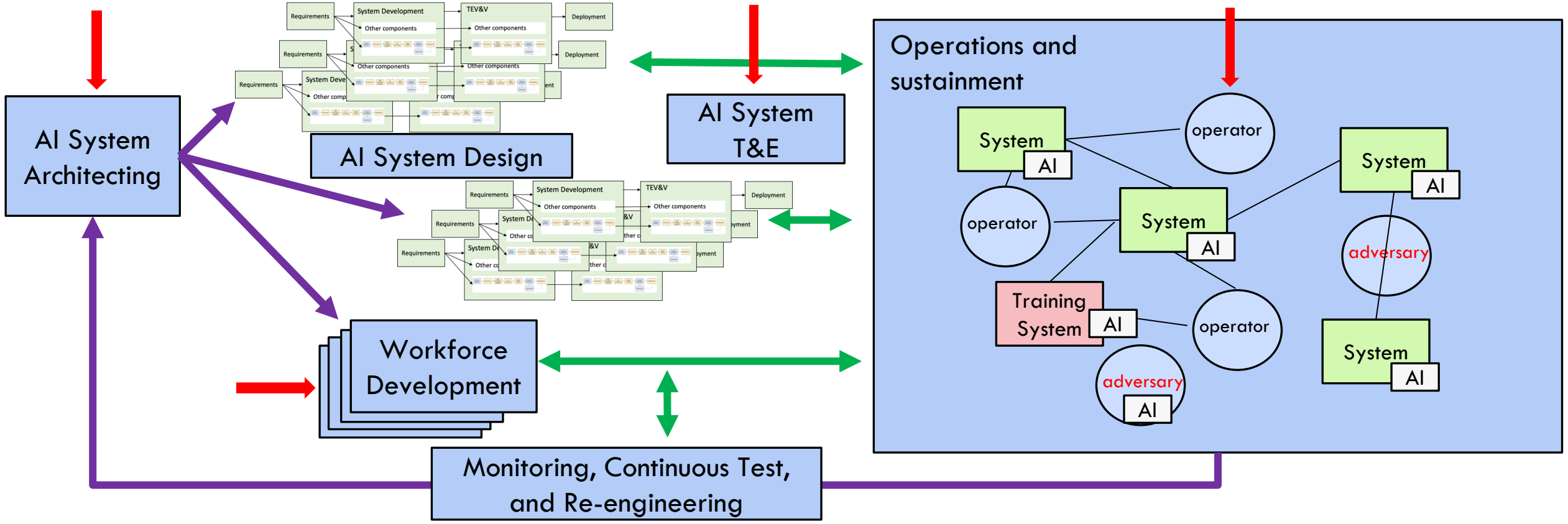
Architecting AI Systems for long-term trust: Linking task & function allocation, test and risk analysis and need for **systems** testbeds

T&E as a Continuum: what to test and how to interpret for AI Systems of varying complexity and embeddedness

AI Resilience: Strategies to mitigate disruptions / ensure acceptable behaviors and recoveries when failures occur

THE REAL WORLD OPERATES IN A SOCIO-TECHNICAL SYSTEMS ARCHITECTURE

Involving complex interactions among humans and systems that were not always intended to work together in a constantly changing environment.



AI in system context; Building user trust; Architecting for long-term trust; T&E as a continuum

WHAT MAKES YOU TRUST (OR NOT TRUST) "THE AI"?

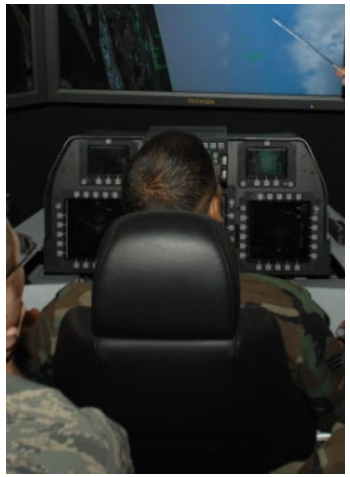
Developer

[1]

Accuracy:

If you're a computer scientist, you want to see the math of this specific algorithm or at least a visualization of the prediction.

Domain Expert



[2]

Agrees with me:

If you're a pilot flying in an engagement using your display image, you might want to see the system agree with you often enough.

End User

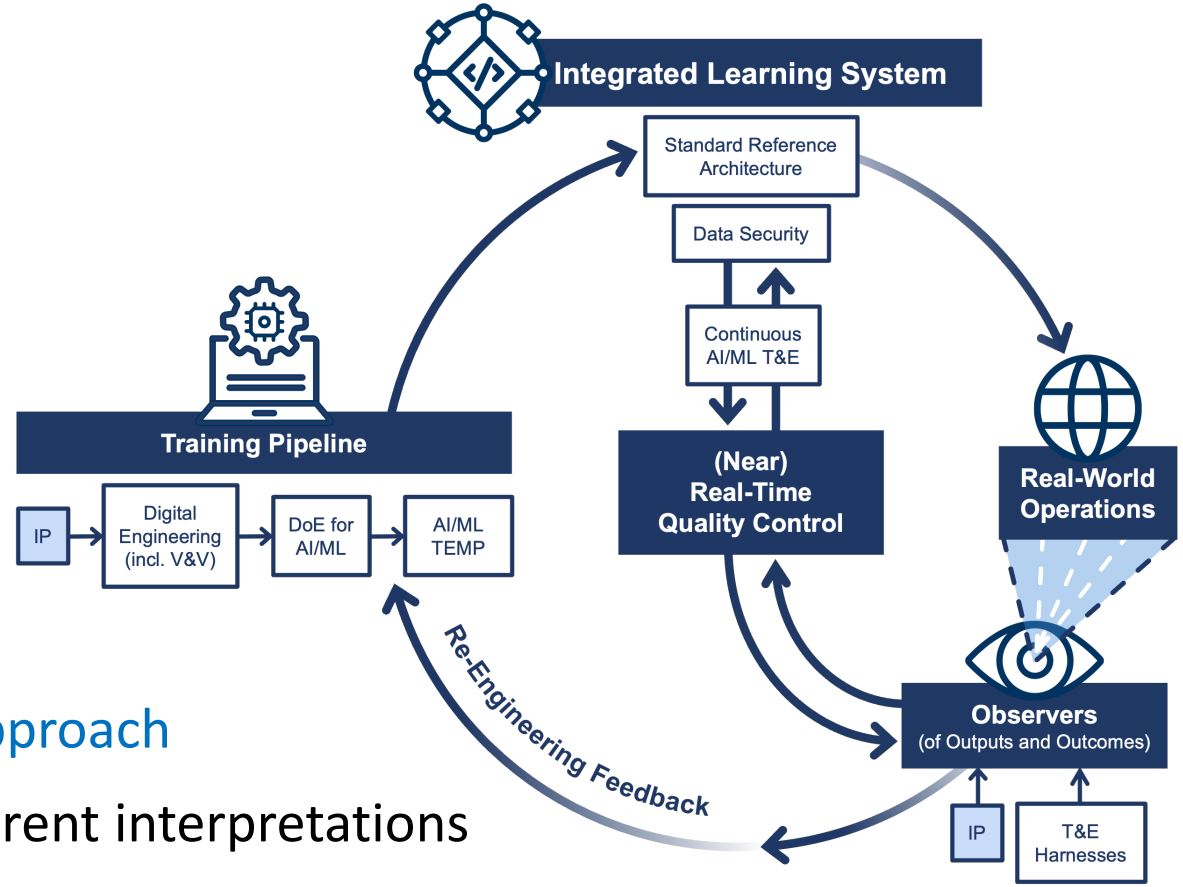
[3]

Trusted 3rd Party:

If you're an operational evaluator, you might want to certify its safety…and for commanders, not have created any international incidents!

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.

SERC 5TH ANNUAL AI4SE & SE4AI WORKSHOP

2023 SUMMARY REPORT

The conference theme, "Safer AI-Enabled Complex Systems: Responsible Deployment of AI through Systems Engineering," aims to foster discussions and insights on how systems engineering can support the development of robust and ethical AI systems, and how AI tools can in turn transform the practice of systems engineering.

https://sercuarc.org/event/ai4se-se4ai-workshop-2024/#dates