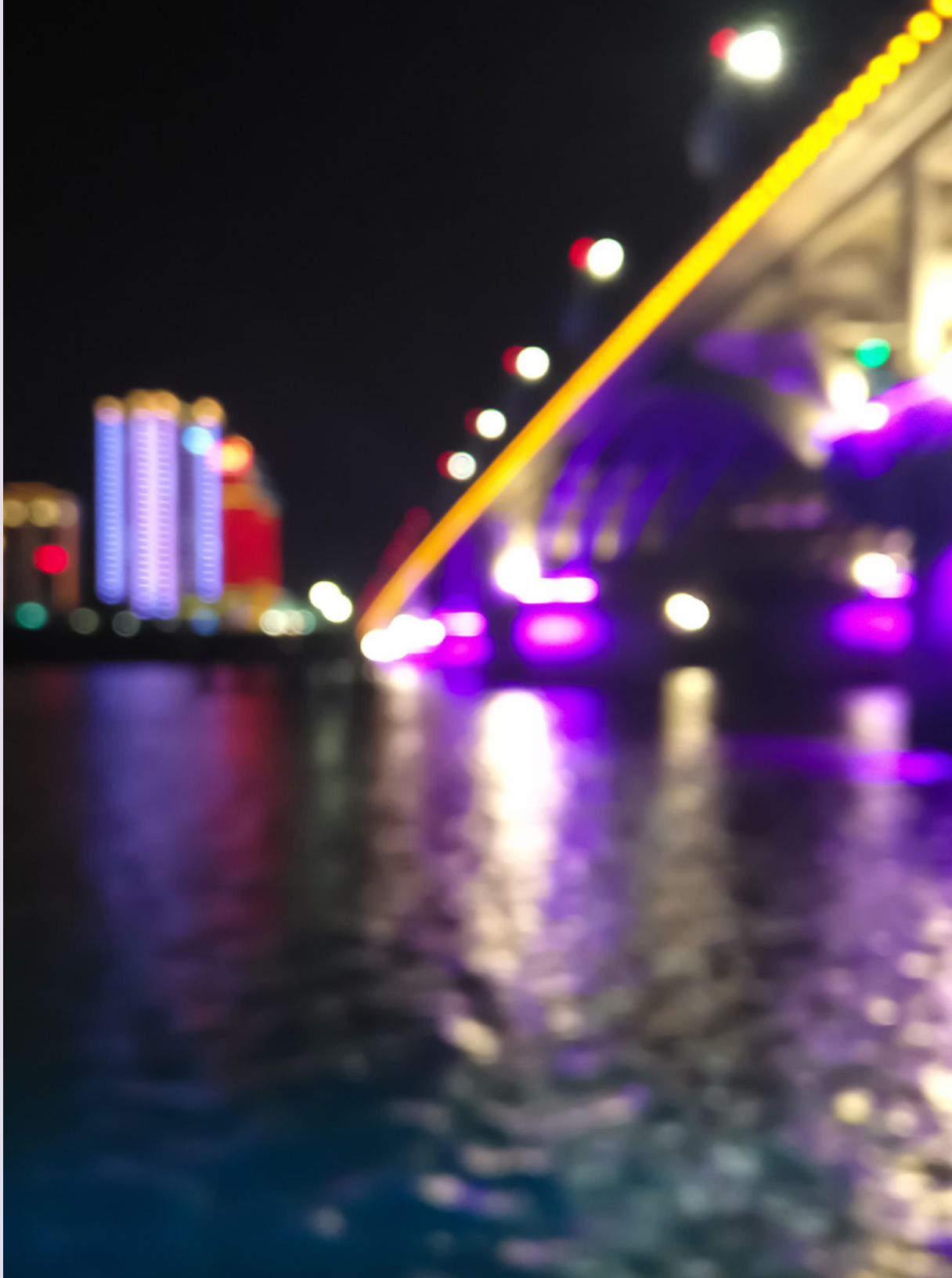
National AI Engineering Initiative

Carnegie Mellon University Software Engineering Institute



AI Engineering

An Emergent Discipline for Human-Centered, Robust and Secure, and Scalable AI

AI ENGINEERING IS CRUCIAL FOR SUCCESSFUL AI DEPLOYMENTS.

In contrast to the current rush to develop capabilities and tools, AI Engineering asks a different set of questions:

- How can AI help humans to achieve mission outcomes?
- What are the limits of AI systems in practice today?
- How can we ensure that ethical standards are upheld as AI systems are deployed?

AI Engineering provides a framework and tools to proactively design AI systems to function in environments characterized by high degrees of complexity, ambiguity, and dynamism.

The discipline of AI Engineering aims to equip practitioners to develop systems across the enterprise-to-edge spectrum, anticipate requirements in changing operational environments and conditions, and ensure that human needs are translated into understandable, ethical, and trustworthy AI.

85%

of AI deployments *will fail* through 2022.[1]

[1] Gartner. "Gartner Says Nearly Half of CIOs Are Planning to Deploy Artificial Intelligence." 2018.

AI ENGINEERING IS CRITICAL TO AI FOR US DEFENSE AND NATIONAL SECURITY.

As a federally funded research and development center, the SEI approaches AI Engineering through the lens of some of the toughest AI challenges: those posed by defense and national security missions. Through our work in developing AI mission capabilities, we gather insights that help build the emerging discipline for AI Engineering.

Featured Projects

xView 2 Challenge

A computer vision prototype for assessing building damage from overhead imagery after disasters

READ MORE ▶

Train, But Verify

New methods to identify, prioritize, and prevent adversarial attacks on ML algorithms

READ MORE ▶

Knowing When You Don't Know

New techniques to accurately estimate uncertainty in ML models, detect domain shift, provide efficient retraining mechanisms, and reason in the open world

READ MORE ▶

The xView 2 Challenge applies ML to help first responders detect and classify damage to structures.

Train, But Verify extends our ability to secure, test, evaluate, verify, and validate AI systems.

Knowing When You Don't Know enables AI systems to manage uncertainty in real-world situations.

AI ENGINEERING IS THE WAY TO HUMAN-CENTERED, ROBUST AND SECURE, AND SCALABLE AI.

AI Engineering combines the principles of systems engineering, software engineering, computer science, and human-centered design to create AI systems in accordance with human needs for mission outcomes. Through conversations with our partners, we've developed three pillars to guide our approach to AI Engineering.

1. Human-Centered
2. Robust and Secure
3. Scalable

HUMAN-CENTERED

Human-centered artificial intelligence (HCAI) is AI that's designed to work with, and for, people.

Developing and deploying human-centered AI is a formidable task. Challenges include the evolving nature of AI technology, operators and team members with varying levels of knowledge and expertise, rugged or unpredictable operational contexts, translation barriers between humans and machines, and the level of oversight needed to create and maintain ethical systems. The real-world environments in which we hope to embed AI systems are often highly complex, ambiguous, and dynamic. Humans—their behaviors, influences, decisions, and actions—are central to such environments. Therefore, as AI technology is developed for the DoD, harmonizing those systems with humans is crucial.

Areas of Opportunity

- Maintaining operational intent and mechanisms for adapting and evolving systems based on dynamic contexts and user needs
- AI organizational readiness and culture
- AI workforce development and digital transformation
- Mechanism and implementations of principles of AI ethics
- Integrated human-machine workflows
- Trust and confidence for human-machine systems

ROBUST AND SECURE

Robust and secure AI systems reliably operate at expected levels of performance, even in the presence of uncertainty, danger, or threat.

One of the biggest challenges facing the broad adoption of AI is having confidence that AI systems will work as expected when they are deployed outside of closely controlled development, laboratory, and test environments. AI technologies are often tasked with complex problems for which there are no guaranteed ways to achieve perfect solutions.

As a result, the goal must shift from achieving a perfect outcome to building confidence in AI throughout its entire lifecycle—from design to development to test to operations and around again as AI systems evolve. Developing new tools, processes, and practices for testing, evaluation, verification, and validation (TEV&V) is critical for building and deploying robust and secure AI Systems with confidence.

Areas of Opportunity

- How to build robustness and security into AI components and AI systems
- Beyond Accuracy: test, evaluation, verification, and validation (TEV&V) across the AI system lifecycle – model evaluation, model integration, model deployment, continuous monitoring
- Training and testing for adversarial robustness and AI red teaming
- Interdependent security policies for AI systems
- Understanding and designing for confidence and uncertainty in AI systems
- Analyzing, interpreting, and understanding AI system behaviors

SCALABLE

Scalable AI refers to the ability of algorithms, data, models, and computing infrastructure to accommodate the size, speed, and complexity requirements of mission needs.

Defense and national security missions are carried out in the midst of challenges that include scarcity of labeled data, the need to repurpose tools, and environments that may not support the computing infrastructure that is used for complex problems in industry or university settings. Scalable AI systems can adapt to these challenges to meet mission needs.

Areas of Opportunity

- Cost, size, weight, and power (CSWaP) constraints for AI systems from the data center to the edge
- Scaling AI capabilities to mission size, speed, and complexity
- Enabling algorithmic and computational innovations to advance AI capabilities
- Enterprise scalability for the design, development, deployment, operations, and evolution of AI systems
- Scalable oversight for the creation, curation, management, and maintenance of data and knowledge for AI systems

AI ENGINEERING TAKES ALL OF US.

The National AI Engineering Initiative is leading the development of the AI Engineering discipline.

The National AI Engineering Initiative, led by the Carnegie Mellon University Software Engineering Institute with sponsorship from the US Office of the Director of National Intelligence, has begun the work of building a professional discipline for AI Engineering. One of the most important steps is to create a community. Along with partners from defense and national security, industry, and universities, the SEI will guide the creation of a multi-year research and development roadmap and foster the development of AI systems that are human-centered, scalable, robust, and secure.

By building upon our collective resources, we will be able to fuel this initiative and unlock the full potential of AI for Defense and National Security—and thus achieve the vision of creating viable, trusted, and resilient AI systems. Our network of partners works to conduct and share research, shape the evolving state of the art, and define gaps and needed capabilities.

Join Us Today!

Who We Need

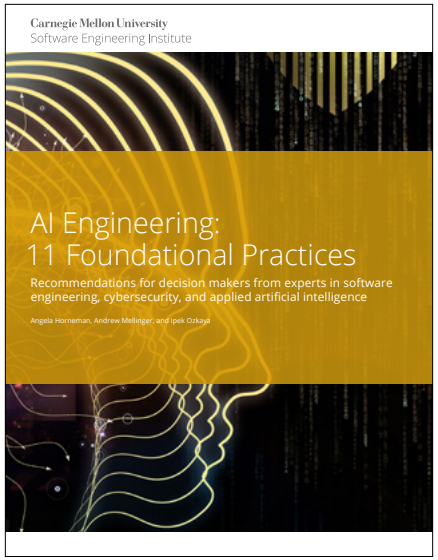
Collaborators:
Conduct and Shape Research

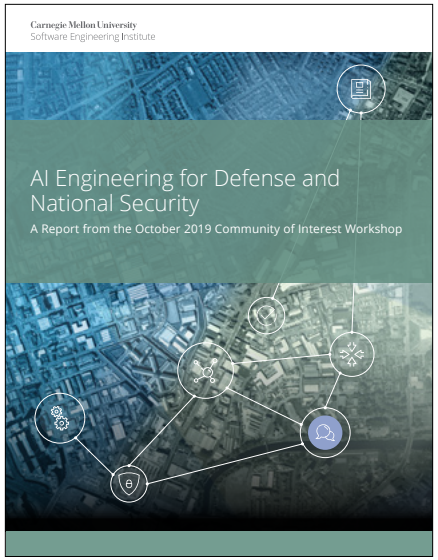
Advocates:
Spread the Word

Co-Funders:
Fuel Research and Operations

JOIN US ►

RESOURCES





AI Engineering: The National Initiative...

Defining AI Engineering

More AI Engineering Resources from the SEI

This collection includes papers, videos, presentations, and other publications related to the SEI's AI Engineering work.

SEE MORE ►

CONTACT

info@sei.cmu.edu