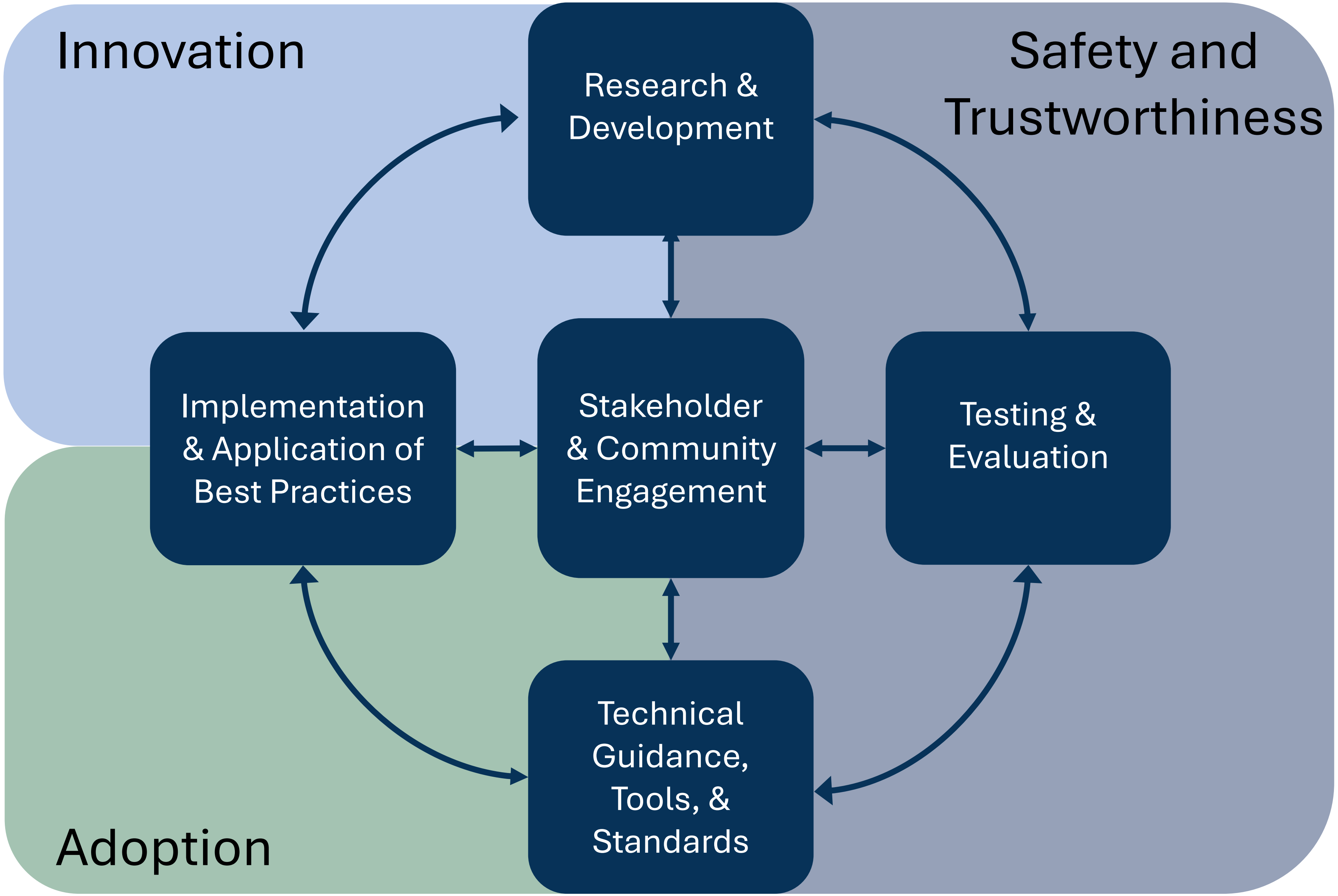
Intersections of AI Work at NIST

Dr. Laurie Locascio
October 29, 2024

NATIONAL INSTITUTE OF
STANDARDS AND TECHNOLOGY
U.S. DEPARTMENT OF COMMERCE

NIST AI Ecosystem



Innovation
Safety and Trustworthiness
Research & Development
Implementation & Application of Best Practices
Stakeholder & Community Engagement
Testing & Evaluation
Technical Guidance, Tools, & Standards
Adoption

NIST has three key pillars for promoting AI innovation

NIST AI Innovation Lab

U.S. AI Safety Institute

AI in Measurement Science & Services

NIST AI Innovation Lab (NAIIL)

Credit: NIST

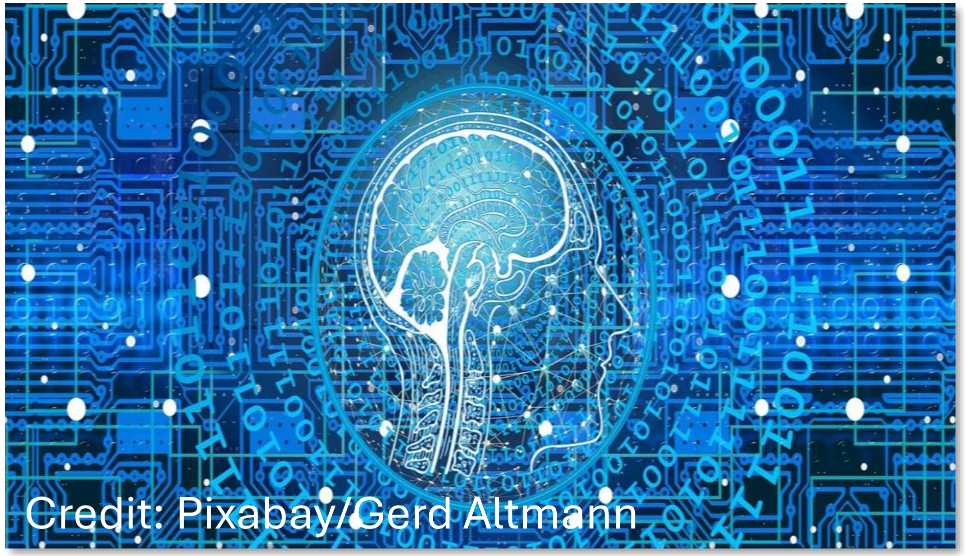
Goals & Impacts

- Provide a foundation for the AI community to evaluate and test AI in ways that will improve its functionality and trustworthiness
- Develop measurement science capabilities to accelerate AI innovation

Focus Areas & Example Efforts

- Fundamental research, including through the AI Cooperative Research Center with Carnegie Mellon University
- Research evaluation programs, including Assessing Risks and Impacts of AI and the Generative AI challenge problem
- Standards development and technical requirements for trustworthy AI, including the NIST AI Risk Mangagement Framework
- Convenes public and private sector stakeholders, such as the recent Unleashing AI Innovation Symposium

U.S. AI Safety Institute (AISI)


Credit: Pixabay/Gerd Altmann

Goals & Impacts

- Advance the science of AI safety and address the risks posed by advanced AI systems
- Develop the testing, evaluations, and guidelines that will help accelerate the safe development of AI here in the United States and around the world

Focus Areas and Example Efforts

- Evaluate highly capable AI models, including through agreements with Anthropic and OpenAI regarding AI safety research, testing, and evaluation
- Release guidance on safe AI development, including draft guidelines on managing misuse risk for dual-use foundation models
- Building a thriving and enduring global ecosystem of AI safety, including the upcoming inaugural convening of International Network of AI Safety Institutes in San Francisco

AI in Measurement Science & Services

Credit: J. Stoughton/NIST

Goals & Impacts

- Integrate AI tools, guidance, and frameworks into measurement science and measurement services to advance adoption of AI systems in domain-specific environments
- Accelerate the development, commercialization, and deployment of critical and emerging technologies (CETs)
- Revitalize U.S. advanced manufacturing

Focus Areas & Example Efforts

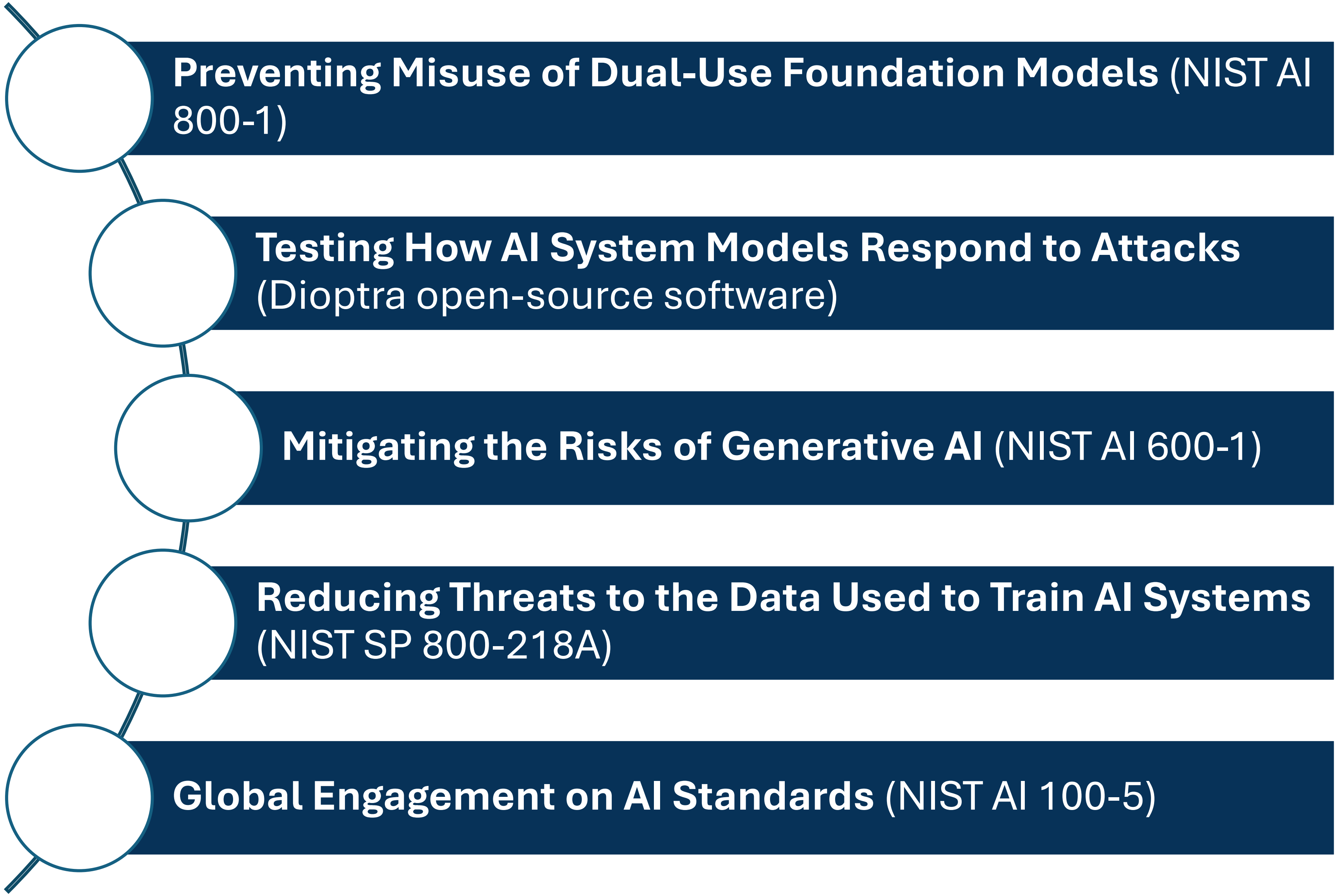
Intramural NIST research exploring:

- AI to improve advanced manufacturing processes and operations
- AI for material science innovation such as in the NIST Biofoundry and through JARVIS
- Leveraging AI in communications
- New devices and architectures for AI
- Other CETs
- Measurement service modernization

Promoting collaboration to accelerate new AI capabilities through CHIPS R&D and Manufacturing USA opportunities

Executive Order on the Safe, Secure and Trustworthy Development of AI

NIST, through NAIIL and AISI, successfully delivered five products in response to 2023 Executive Order on AI.



Preventing Misuse of Dual-Use Foundation Models (NIST AI 800-1)

Testing How AI System Models Respond to Attacks (Dioptra open-source software)

Mitigating the Risks of Generative AI (NIST AI 600-1)

Reducing Threats to the Data Used to Train AI Systems (NIST SP 800-218A)

Global Engagement on AI Standards (NIST AI 100-5)

National Security Memorandum (NSM) Taskings

The NSM tasked AISI with multiple deliverables, including to:

- Serve as the primary point of contact with the private sector for testing frontier AI models and communicating risk mitigation measures.
- Coordinate with other agencies on classified CBRN testing.
- Test at least two frontier models pre-deployment (within 180 days)
- Issue detailed guidance for AI developers on how to evaluate and manage risks to safety, security, and trustworthiness (within 180 days)
- Develop or recommend technical benchmarks for AI capabilities related to public safety or national security risks (within 180 days)
- Coordinate with NSA/DOE on cyber, nuclear, and radiological test planning
- Submit the first annual report on AISI activities to APNSA (within 270 days)

Public-Private Partnerships

NIST leads, convenes, and participates in public-private partnerships across the AI ecosystem.

U.S. AI Safety Institute Consortium

AI for Resilient Manufacturing Institute

Engineering Biology Research Consortium

CHIPS AI/AE

Building NIST's AI Community

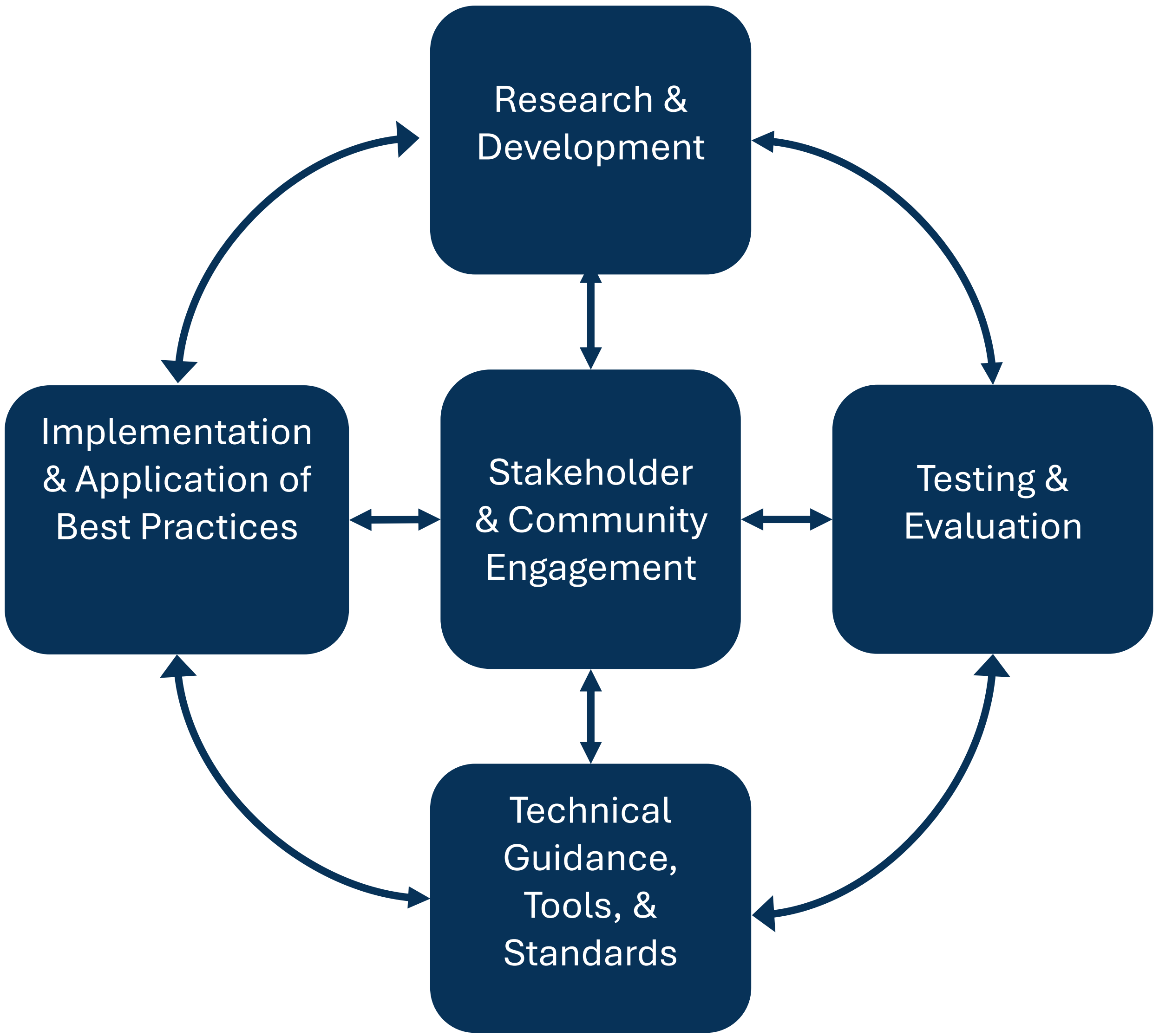
NIST
NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY
UNITED STATES DEPARTMENT OF COMMERCE
100 Bureau Drive

Credit: J. Stoughton/NIST

AI @ NIST Day

AI for Advancing NIST Measurements Workshop

What’s Next for NIST?


Research & Development
Implementation & Application of Best Practices
Stakeholder & Community Engagement
Testing & Evaluation
Technical Guidance, Tools, & Standards

NIST will continue working throughout the AI ecosystem, focusing next on:

- Build on our work, including the Executive Order
- New guidance, including companion documents to the AI RMF
- Request for Information related to responsible development and use of chem-bio AI models
- Public workshops for its Assessing Risks and Impacts of AI (ARIA) program
- Inaugural Convening of International Network of AI Safety Institutes
- Continue public-private partnership efforts
- Exploring convergence of AI with other CETS

Discussion

- What future applications and capabilities for AI can be envisioned for either NAIIL, AISI, and/or AI application in CETs?

- How is the rapid pace of the AI development changing how U.S. industry is looking to NIST for AI R&D, guidelines, and testing?