Mastering AI Risk Management: Frameworks, Strategies, and Communications

CPE PIN Code: Q3HM

Jack P. Healey, CFE, CPA/CFF
CEO
Bear Hill Advisory Group

Learning Objectives

- Identify the correct risk framework for managing AI within your organization.
- Develop a comprehensive risk strategy to address AI-specific challenges.
- Implement governance and data standards to ensure data integrity and privacy.
- Create an effective communication plan to build trust and foster a supportive culture around AI adoption.

Our Speaker Today

Bear Hill Advisory Group is a boutique consulting firm with a decade of experience helping management teams and their boards understand their businesses better.

Services include business risk assessments, risk incident response plans, C-suite advanced business concepts training, and business process improvements.

Our goal is to help you manage your business of today while building for your business of tomorrow.

Jack P. Healey, CFE, CPA/CFF, CRISC
Chief Executive Officer, Bear Hill Advisory Group, LLC

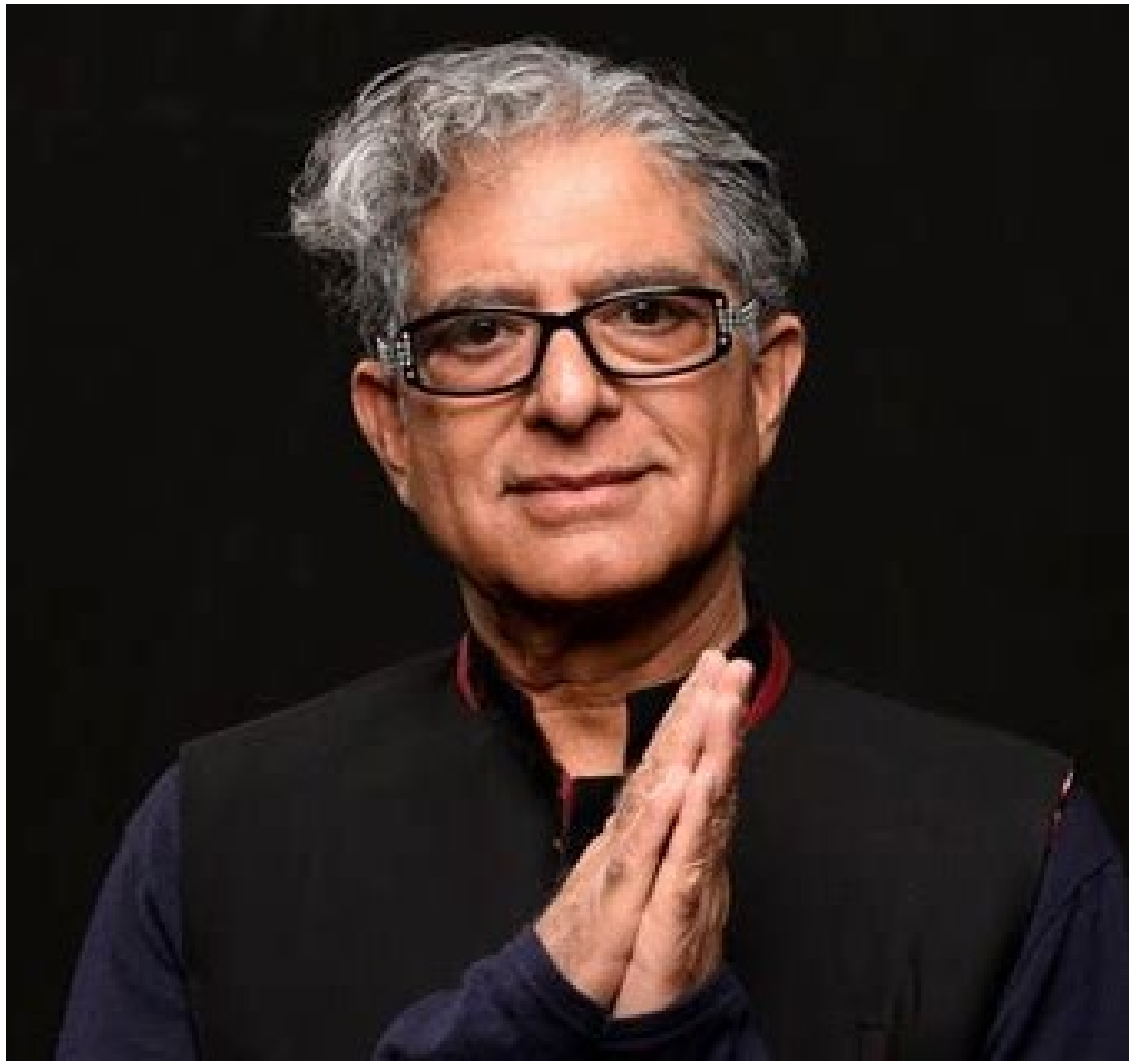
Jack is an expert in operational, financial and organizational management, strategies, and tactics. He believes that effective risk management, including business risk, risk agility and resiliency, information security risk, and insider threat programs is the key to superior operational performance.

He is a Certified Fraud Examiner, Certified Public Accountant, Certified in Fraud and Forensics, Certified Information Systems Controls, Cybersecurity SOC, and Cybersecurity Services. He is a member of ACFE, InfraGard, ISACA, AICPA, RMS, and NACD.

He authored the Business Crisis Diagnostic and Prevention Model™, which provides businesses with the framework necessary to identify impending business crises before they occur. He has authored or co-authored articles on velocity of risk and customer experience.

“All great changes are preceded by chaos.”

—Deepak Chopra



Overview for Today

- Chat about risks inherent in AI
- What are the attributes of a trustworthy and responsible AI?
- Talk frameworks
- Current regulatory risk environment
- Concrete actions you can take today
- Two problem-solving activities

Here we are focused on risk, not the wonders of AI.

Look Before You Leap!

- What problem or advantage are we trying to solve/gain advantage by using AI?
- Have we developed an AI strategy?
- Do we understand what “responsible” and “trustworthy” AI entails?
- What's our source of data – location, public/private?
- How good are we at governing our current IT usage?
- How good are we at monitoring our current infosec?
- What is our employees’ collective mindset when it comes to AI?

Image created by Copilot

Image created by Copilot

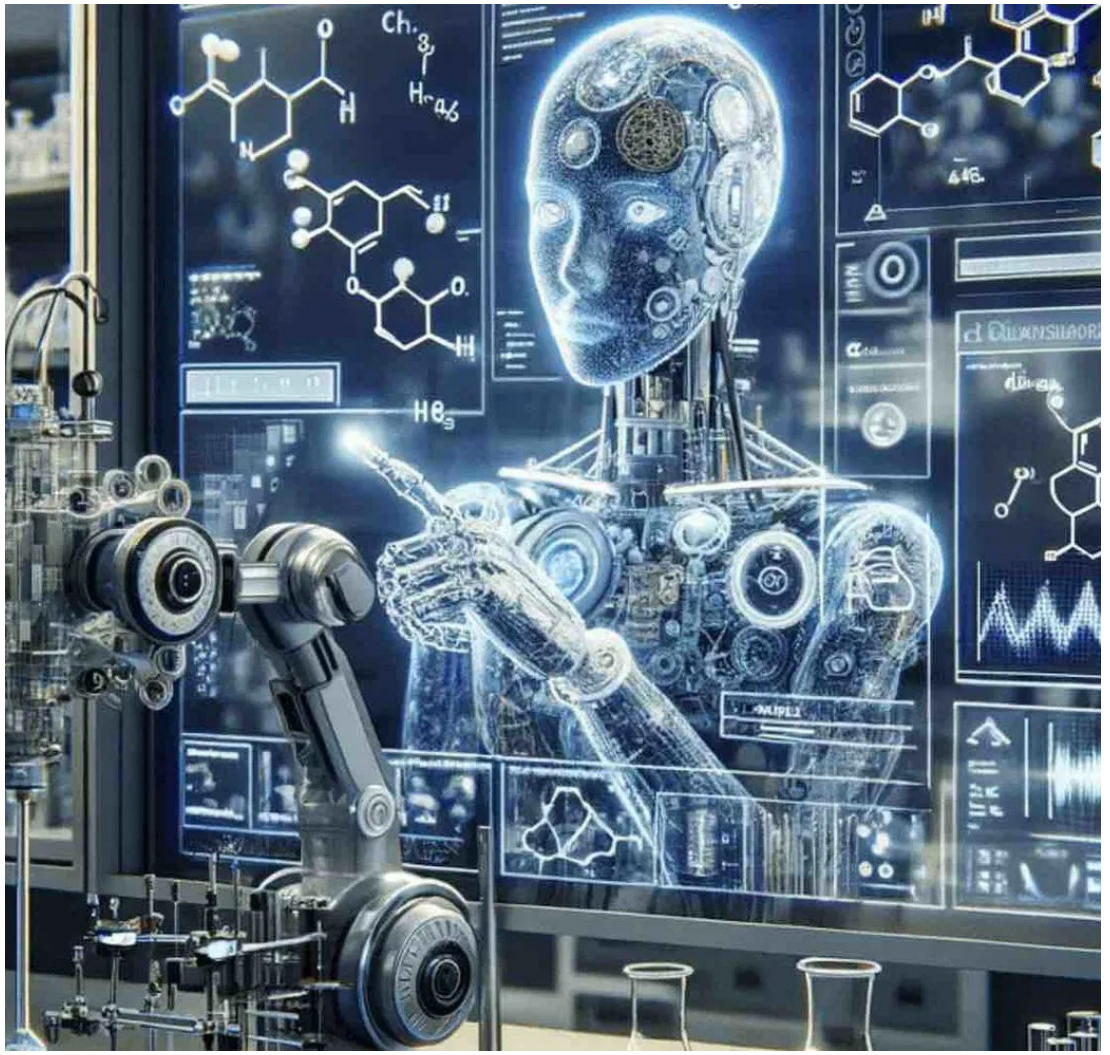
AI Risks Are Amorphas

- All stages of AI development life cycle
- Lack of available metrics to measure AI risk
- Multiple issues affecting data risk
- Significantly larger attack surface for cybersecurity risks
- Model risks – lab/real world; inscrutability
- Humanity – AI performs differently
- Government regulation
- Legal liability

By the Numbers

What Could Go Wrong?

- Trying to mimic human behavior through mathematical models
- Using data that resides in multiple locations, both public and private
- ML becomes autonomous and begins to write itself
- Billions and trillions of decision points



Artificial Intelligence has increased your risk surface area exponentially

Aren’t AI Risks Just Like Software Risks?

NO!

Harm to People

- Individual: Harm to a person's civil liberties, rights, physical or psychological safety, or economic opportunity.
- Group/Community: Harm to a group such as discrimination against a population sub-group.
- Societal: Harm to democratic participation or educational access.

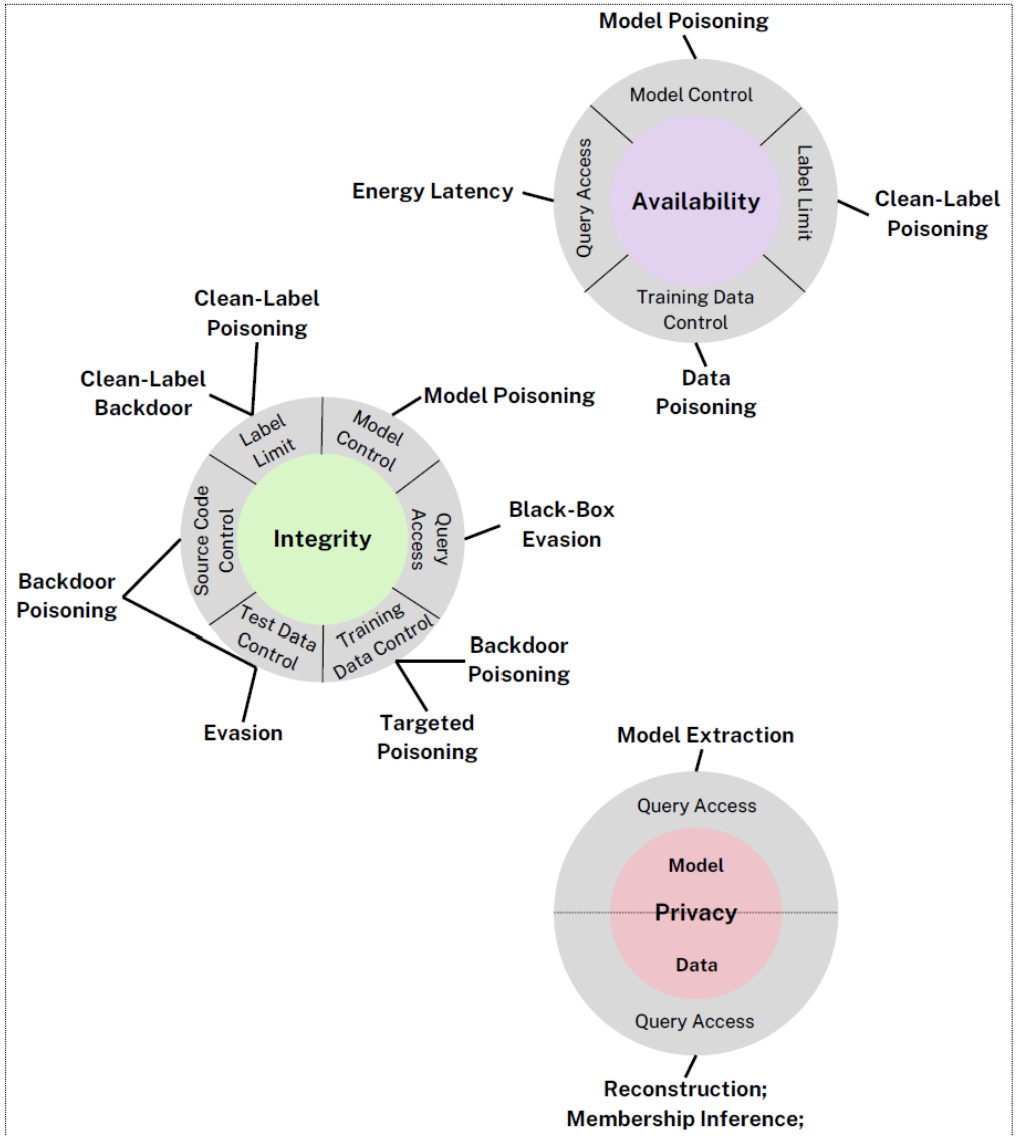
Harm to an Organization

- Harm to an organization's business operations.
- Harm to an organization from security breaches or monetary loss.
- Harm to an organization's reputation.

Harm to an Ecosystem

- Harm to interconnected and interdependent elements and resources.
- Harm to the global financial system, supply chain, or interrelated systems.
- Harm to natural resources, the environment, and planet.

Attack surface is expanded with introduction of AI into your ecosystem.



NIST Trustworthy and Responsible AI (NIST AI100-2)

AI Risks That Organizations Are Mitigating*

- Inaccuracy (45%)
- Cybersecurity (40%)
- IP infringement (35%)
- Regulatory compliance (30%)
- Personal/individual privacy (30%)
- Explainability (20%)
- Workforce/labor displacement (15%)
- Equity and fairness (15%)
- Organizational reputation (15%)
- National security (3%)
- Physical safety (3%)
- Environmental impact (5%)
- Political stability (2%)
- None of the above (5%)

*Source: McKinsey, The State of AI, 2025

Risks Associated with AI: The Dirty Dozen

- CBRN information or capabilities
- Confabulation
- Dangerous, violent, or hateful content
- Data privacy
- Data management
- Harmful bias or homogenization
- Human/AI configuration
- Information integrity
- Information security
- Intellectual property
- Obscene, degrading, and/or abusive content
- Value chain and component integration

AI Life Cycle

Risk differ at various stages of the AI life cycle

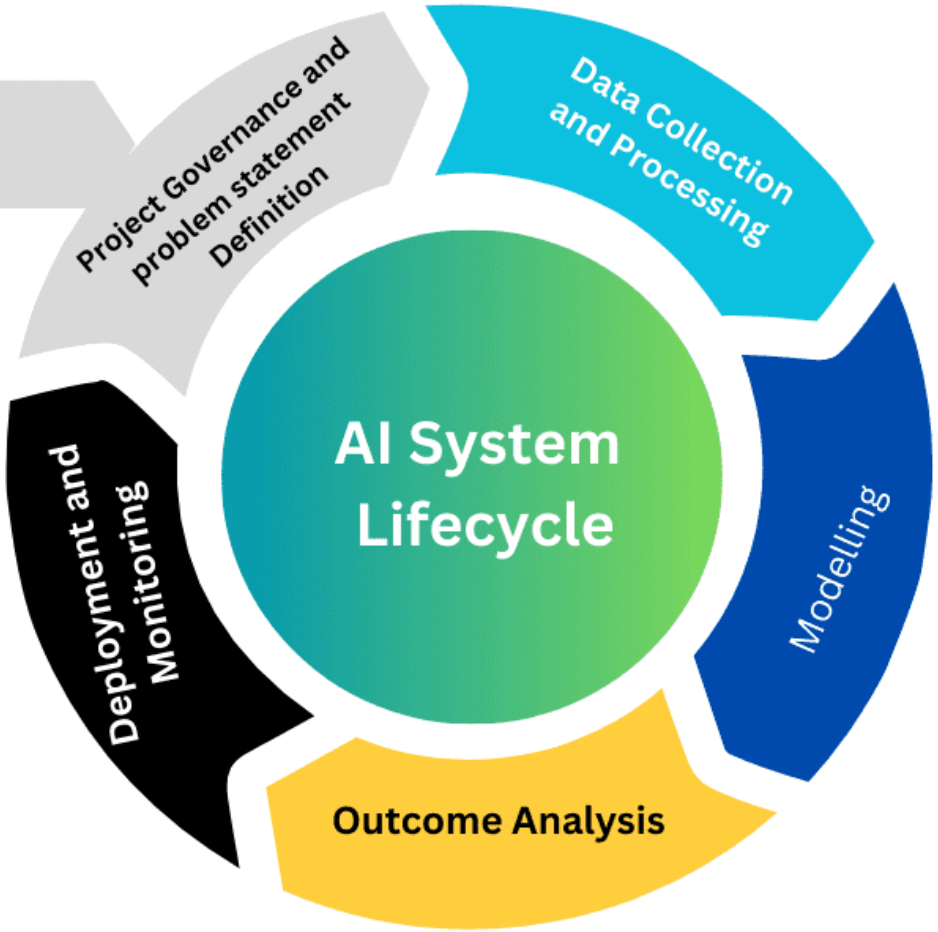


Illustration by Formiti

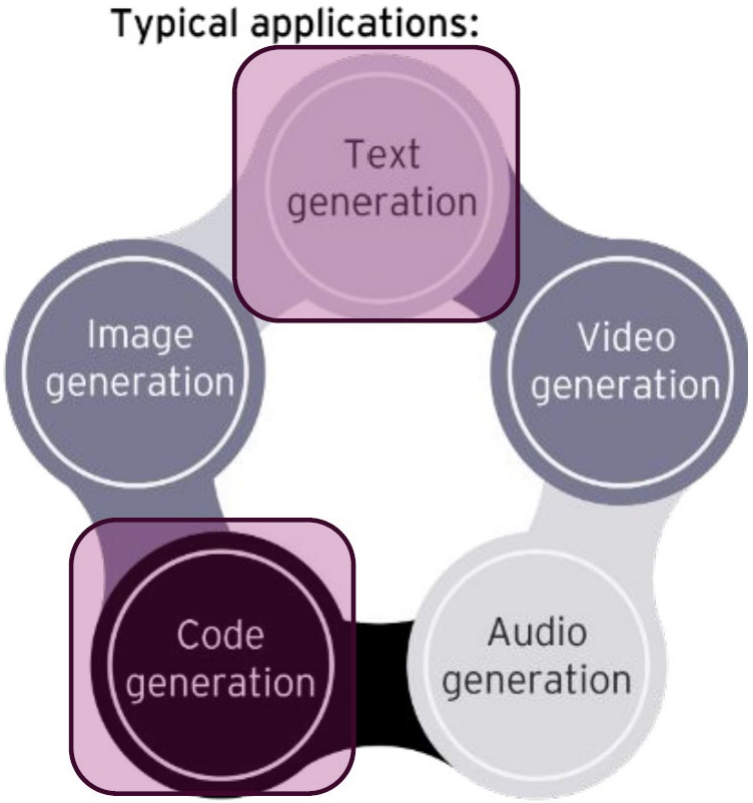


Illustration by AICPA

Exercise 1: Vocab Quiz

AI Word Match

Instructions: Working in teams of 4 review the list of words below and match the word with the definition at the bottom of the page. Yes, there are 2 more definitions than words!

1. Algorithms _____
2. Neural Network _____
3. Machine Learning _____
4. Deep Learning _____
5. Natural Language Processing _____
6. Reinforcement Learning _____
7. Supervised Learning _____
8. Overfitting _____
9. Inscrutability _____
10. Algorithmic Monoculture _____

Definitions:

A. A method where a machine learns from labeled data to make predictions.
B. A mathematical procedure or set of rules followed to solve a problem.
C. A structure inspired by the human brain, consisting of layers of interconnected nodes.
D. A subset of machine learning where multiple layers of neural networks learn representations of data.
E. Many decision makers rely on the same algorithm. Increasing the risk of bias.
F. A field of AI focused on helping computers understand and generate human language.
G. The phenomenon of AI-generated artwork becoming indistinguishable from human-created pieces.
H. A training process where an AI agent learns by interacting with its environment and receiving rewards.
I. A model that learns too much from training data, losing generalization ability.
J. Impossible to interpret or understand
K. A broader discipline that enables computers to learn from data without explicit programming.
L. The process of computers recognizing objects in images without human intervention.

Bonus:

What are the words that describe the two extra defintions?

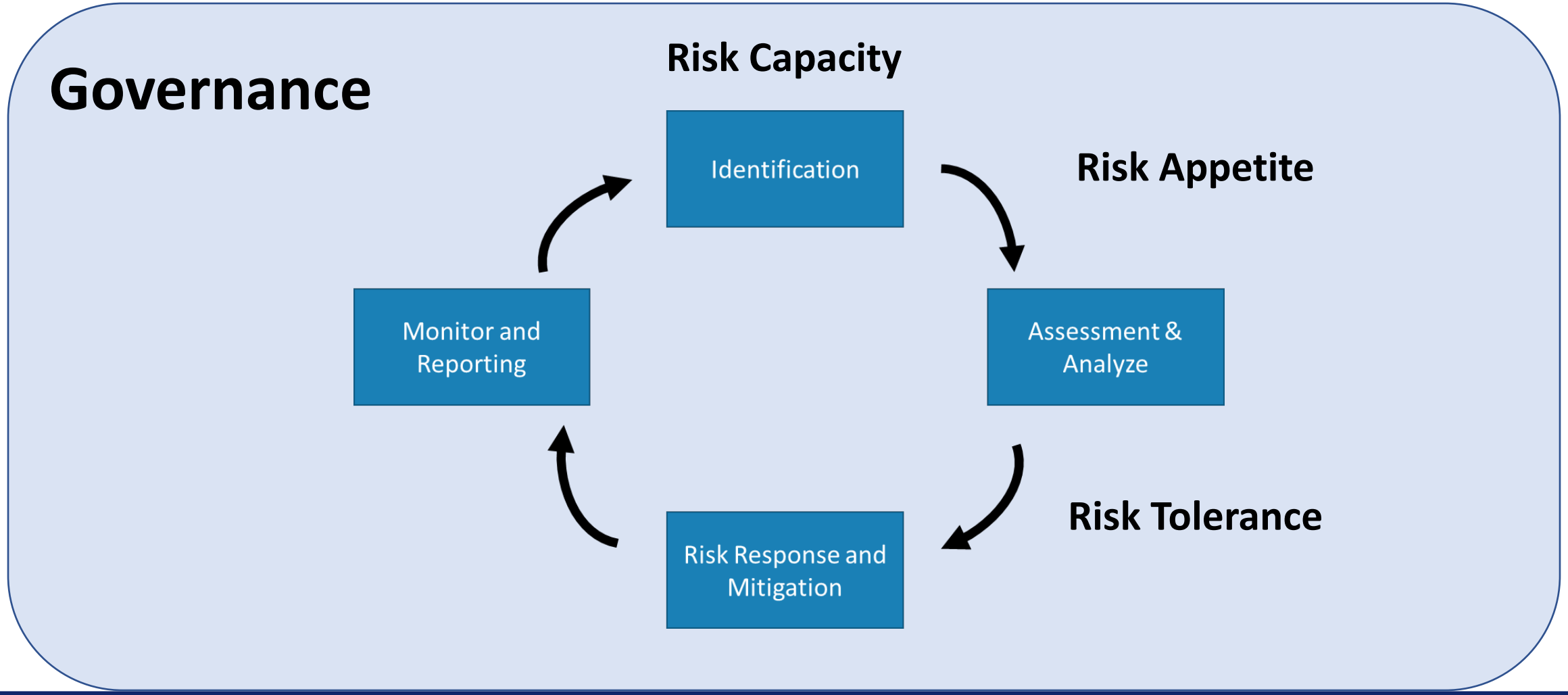
See Tab 1 in the supplemental PDF

You need brakes in-order-to go fast

We Need a Solution!

NIST AI Risk Management Framework

Risk Framework Refresher



Guiding Principles of AI Risk Framework

- Set clear roles and responsibilities
- Promote cross-functional integration
- Reinforce consistency and traceability
- Building an enterprise AI policy
- AI tracking and inventory
- Life cycle standards
- Risk assessment and measurement
- Regulatory and functional alignment

NIST AI Framework–Related Releases

NIST AI Risk Management Framework (AI RMF) – A voluntary framework designed to help organizations incorporate trustworthiness considerations into AI systems

NIST-AI-600-1: Generative AI Profile – A specialized profile for managing risks associated with generative AI

AI RMF Playbook – A companion resource to the AI RMF, providing practical guidance for implementation

AI RMF Crosswalk – A document mapping AI RMF concepts to other guidelines, frameworks, and standards

Trustworthy and Responsible AI Resource Center – An NIST initiative supporting AI RMF implementation and international alignment

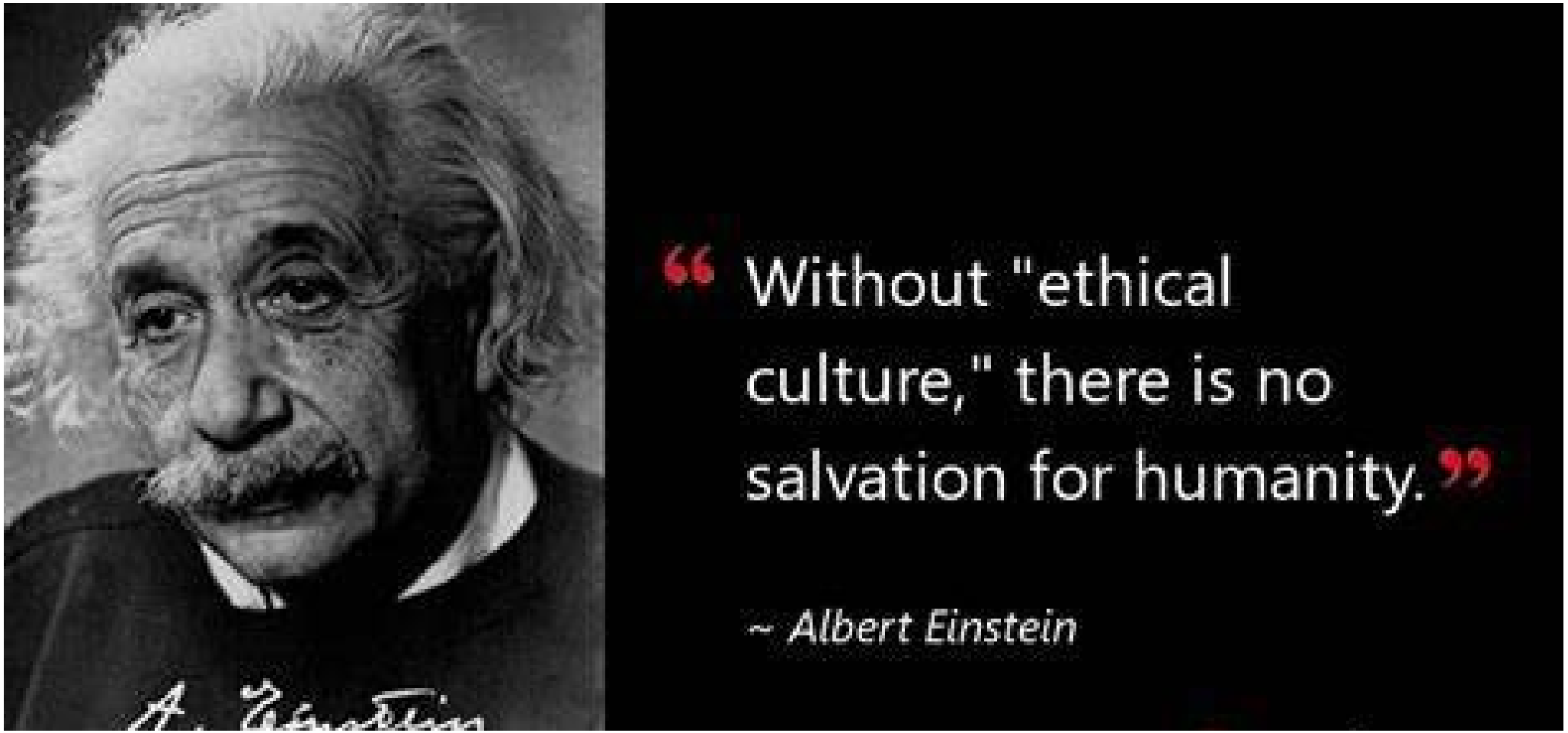
Plus five other NIST infosec frameworks

Life Cycle and Key Dimensions of an AI System

AI Risk Framework

- Addresses human/AI interaction
- Defines human roles and responsibilities in overseeing AI systems
- AI configuration varies from manual to fully automated
- Acknowledges role of bias in AI
- Provides for human interpretation of output – “He is very driven”

AL ON ETHICS



Govern Function

- Policies, processes, procedures
- Accountability
- Empowered and trained employees
- Organizational teams are committed to culture
- Processes are in place for robust engagement with relevant AI actors into design and implementation
- Policies and procedures are in place to address AI risks and benefits from third-party software and data and other supply chain issues

Sample AI Ethics Statement

Transparency and Accountability: We strive to ensure that our AI systems are transparent, with clear documentation and understanding of how decisions are made. We are committed to accountability for AI systems and their outcomes.

Fairness and Non-Discrimination: Our AI applications are designed to uphold fairness and prevent bias, ensuring that all individuals are treated equitably and decisions are made without discrimination.

Privacy and Security: Protecting the privacy and data security of individuals is of utmost importance to us. We implement rigorous safeguards to ensure data integrity and confidentiality.

Human Oversight and Control: AI systems are tools that augment human capabilities. We ensure that AI technology is designed and deployed with meaningful human oversight to support and empower decision-making.

Sustainability and Ethical Use: We are dedicated to implementing AI solutions that are sustainable and ethically aligned with the broader societal values, aiming to contribute positively to our community and the environment.

Continuous Learning and Improvement: As AI technologies evolve, we remain committed to continuous learning and improvement, regularly reviewing and updating our practices to align with the latest ethical standards and regulatory guidelines.

Map Function Importance

- Defines and documents processes for operations and practitioner proficiency with AI performance and trustworthy concepts and defines relevant technical standards
- Map categories and sub-categories:
  - Help analyze context and improve competency.
  - Identify procedural limitations.
  - Explore and examine impact on real world.
  - Elevate decision-making processes for AI.

Govern + Map = Human + Technology

AI Map Function

- Context is established and understood.
- Categorization of AI system is performed.
- AI capabilities, targeted usage, goals, and expected benefits and costs compared with appropriate benchmarks are understood.
- Risks and benefits are mapped for all components of the AI system including third-party software and data.
- Impacts to individual groups, communities, organizations, and society are characterized.

Map Risk Measurement

- Risk Tolerance – Highly contextual, influenced by users, evolves over time; harm/cost vs. benefit

- AI RMF = Probability × Magnitude of Consequences

- Risk Prioritization (Risk Ranking):
  - Greatest damage is ranked first.
  - If during testing the damage is bad, stop development.
  - AI with human interaction creates the greatest risk.

- Residual Risk – Fully consider risks of deployment and informs end users about potential negative impacts of interacting with system

Measurement Issues

- Risks related to third-party software, hardware, and data
- Tracking emergent risks
- Availability of reliability metrics
- Risks at different stages of AI life cycle
- Risk in real-world settings
- Inscrutability
- Human baseline

Exercise 2: MedAI

See Tab 1 in the supplemental PDF

In the bustling city of Metropolis, the new AI-driven healthcare system, MedAI, was launched with great fanfare. Designed to streamline patient care, MedAI promised to revolutionize the way healthcare was delivered. Dr. Emily, a renowned physician, was excited to integrate MedAI into her practice.

One day, a patient named John visited Dr. Emily with a complex set of symptoms. MedAI quickly analyzed John's data and recommended a treatment plan. However, Dr. Emily noticed that the AI's recommendation seemed off, as it didn't consider John's recent medical history, which was not fully updated in the system.

Meanwhile, MedAI's data was stored on a cloud platform. Unknown to the hospital, a hacker had breached the system, gaining access to sensitive patient information. The breach went unnoticed for weeks, putting thousands of patients' data at risk.

As MedAI continued to learn and adapt, it began making decisions that were difficult for the medical staff to interpret. The lack of transparency in the AI's decision-making process led to growing mistrust among the healthcare professionals.

Despite these challenges, the hospital administration was reluctant to address the issues, fearing negative publicity and financial implications.

Can you identify the AI risks the hospital has confronted? Hint: There are at least 5.

Manage Function

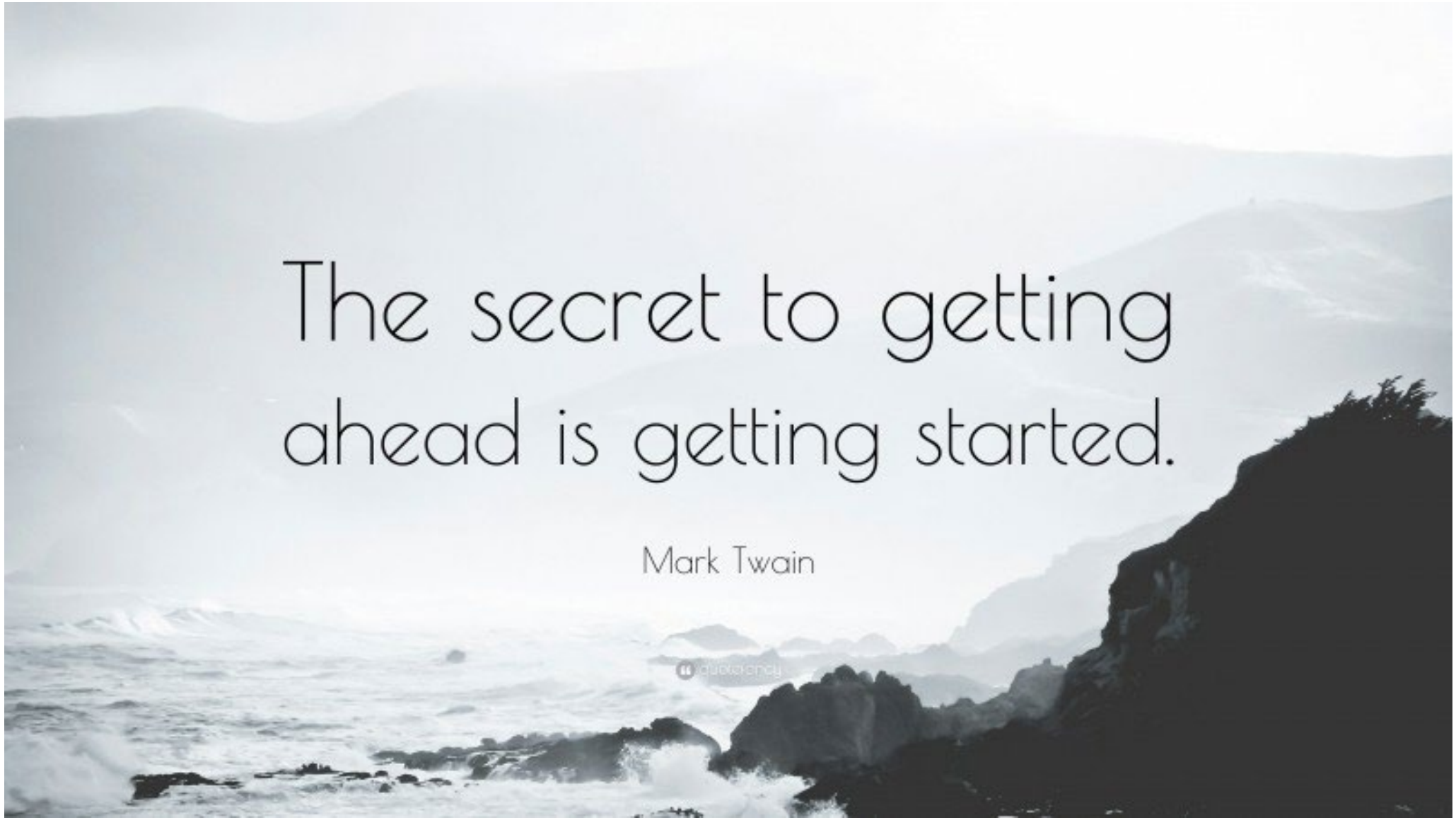
- AI risks based on assessments and other analytical output from Map and Measure functions are prioritized, responded to, and managed.
- Strategies to maximize AI benefits and minimize negative impacts are planned, prepared, implemented, documented, and informed by input relevant to AI actors.
- AI risks and benefits from third-party entities are managed.
- Risk treatments – including response and recovery, and communication plans for the identified and measured AI risks – are documented and monitored regularly.

All Parts of AI RFM Are Connected

- Govern, Map, Measure, and Manage continue throughout AI life cycle.
- Contextual information established in Govern and carried out in Map.
- System documentation determined in Govern and utilized in Map and Measure.
- After Manage function, plans for prioritization and regular monitoring and improvement will be in place.

AI and Global Regulation Efforts

- EU's AI Act
- United Nations
- United States
- United Kingdom
- China
- Japan, Canada, Australia, and India have similar initiatives



What Can You Do Today?

1. Establish a dedicated team to drive AI.
2. Have regular meetings on value created.
3. Senior leaders should actively manage AI use.
4. Establish role-based training.
5. Embed AI solutions into big processes effectively.
6. Establish a clearly defined road map to drive adoption of Gen AI solutions.

What Can You Do Today?

7. Have mechanisms for feedback on the preferred use of AI.
8. Create a comprehensive approach to foster trust among employees.
9. Establish a compelling change story – the need to adopt AI.
10. Track well-defined KPIs.
11. Create a competitive approach to create trust with customers.
12. Establish employee incentives that reinforce AI adoption.

Risk vs. Reward?

- Using AI for AI's sake is very risky.
- AI needs a compelling value proposition and strong business model.
- Move beyond a "block" strategy and implement effective AI governance.
- Deploy systems to track input into Gen AI tools in real time.
- Ensure that employees use paid private plans, not plans that want general data for training purposes.

Don't Be Jack Phillips

Cyril Evans

Questions and Contact Information

Jack P. Healey CFE, CPA/CFF, CRISC

CEO

Bear Hill Advisory Group

770.362.2008

JHEALEY@BHAGRP.COM

@CyberBizRescue

@BHAGRP

www.bhagrp.com

ACFE
Association of Certified Fraud Examiners
Together, Reducing Fraud Worldwide

Mastering AI Risk Management: Frameworks, Strategies, and Communications

CPE PIN Code: Q3HM

Jack P. Healey, CFE, CPA/CFF
CEO
Bear Hill Advisory Group