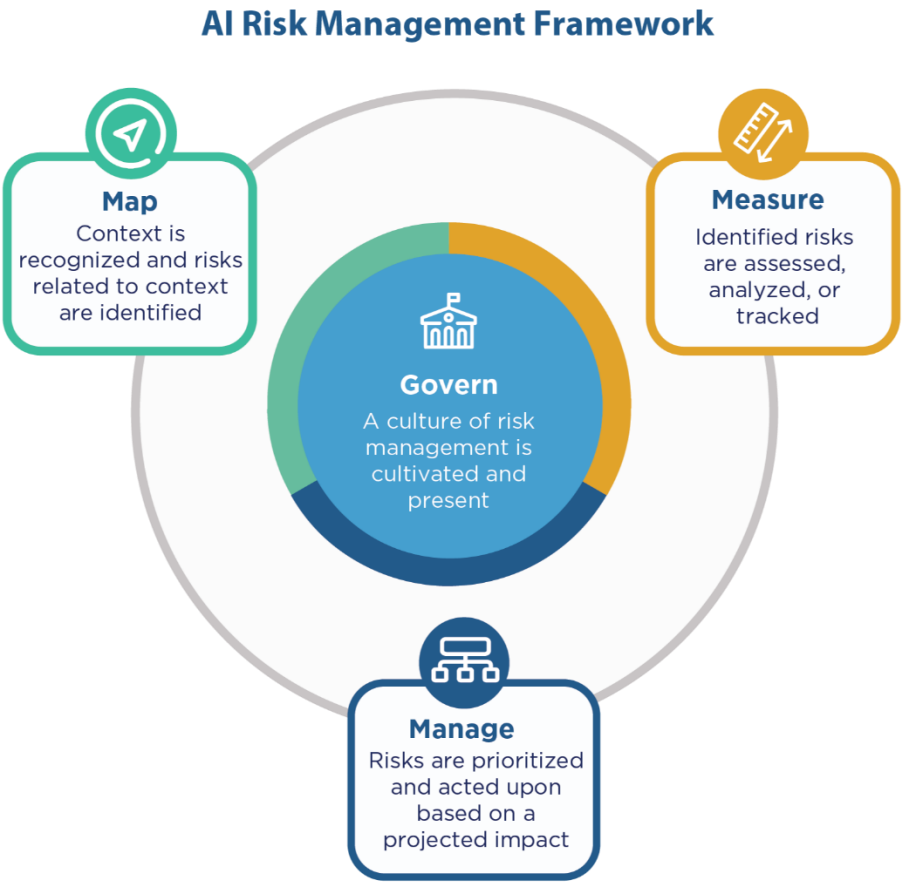

The **NIST AI Risk Management Framework (AI RMF)** is a voluntary, non-prescriptive guidance designed to help organizations manage risks associated with artificial intelligence systems while promoting trustworthy and responsible AI development and use.

The framework is structured around a "Core" composed of four interconnected functions: **Govern, Map, Measure,** and **Manage**. These functions are an iterative process, further broken down into categories and subcategories that outline desired outcomes and actions for managing AI risks across the entire AI lifecycle.

Core Functions of the NIST AI RMF

The four core functions help organizations cultivate a risk-aware culture and integrate AI risk management into their existing enterprise risk management processes.

- **Govern:** Establishes the foundational policies, processes, procedures, and accountability structures for managing AI risks across the organization. This function sets the organizational tone and ensures alignment with values, legal requirements, and strategic priorities.
    - **Categories:** Covers areas such as implementing transparent policies and procedures, establishing clear accountability structures (roles and responsibilities), prioritizing workforce diversity and inclusion, fostering a risk-aware culture, engaging with internal and external stakeholders, and managing third-party risks in the supply chain.
    - **Subcategories:** Provide specific, actionable guidance, such as documenting legal and regulatory requirements or ensuring senior leadership formally declares risk tolerances and delegates authority.

- **Map:** Focuses on contextualizing the AI system to identify and understand the potential risks and impacts throughout its lifecycle.
    - **Categories:** Involves establishing and understanding the context and purpose of the AI system, categorizing the system, defining capabilities and expected benefits, mapping risks across all components (including third-party data and software), and characterizing the impacts on individuals, communities, and society.
    - **Subcategories:** Detail activities like defining the AI system's mission and goals, documenting risk tolerance, and identifying potential negative impacts related to intended use or foreseeable misuse.
- **Measure:** Involves applying appropriate qualitative, quantitative, or mixed-method tools and metrics to assess, track, and monitor AI risks and the system's performance against trustworthiness characteristics.
    - **Categories:** Includes identifying and applying appropriate AI metrics and methods, evaluating AI systems for characteristics like fairness, security, and privacy, implementing mechanisms for tracking identified risks, and collecting feedback on measurement effectiveness.
    - **Subcategories:** Specify how to conduct rigorous testing, evaluate performance metrics (e.g., accuracy, error rates, bias), track emergent risks, and formally report test results to stakeholders.
- **Manage:** Prioritizes and implements the plans and resources to mitigate, respond to, and recover from identified AI risks.
    - **Categories:** Focuses on prioritizing and responding to risks based on assessments from the Map and Measure functions, developing strategies to minimize negative impacts while maximizing benefits, managing third-party risks, and creating ongoing monitoring, response, and communication plans.
    - **Subcategories:** Outline the need for sufficient resources, corrective action plans (e.g., model retraining), incident response protocols, and proactive communication plans for incidents.

For more detailed guidance and practical implementation examples, organizations can refer to the accompanying NIST AI RMF Playbook and other resources available on the official NIST AI Resource Center website.