

Building Trustworthy AI with ISO 42001 and the NIST AI Framework

JULY 31, 2025

3:00 PM CEST



Hafiz Sheikh Ahmed

Cyber Governance Assessor, Advisor, Strategist and PECB
Titanium Trainer

#GlobalLeadingVoices



Global Trust in AI — Key Statistics (2025)

54%

Wary of AI

of people globally are wary
about trusting AI systems

72%

Accept AI Use

despite concerns about safety
and societal impact

82%

Heard of AI

but **49% are unclear** about
how it's used

68%

Use AI Apps

yet **41% don't realize** AI
powers them



⚠ When Trust Is Absent: High-Stakes Failures

🛡 Healthcare – Sepsis Detection Failure

- An AI model designed to detect sepsis missed **over two-thirds of actual cases** in hospital settings.
- While it performed well in training, it failed in real-world deployment due to lack of robustness and explainability.
- **Lesson:** Without trustworthy AI, even life-saving tools can become liabilities.

👮 Predictive Policing – Bias Amplification

- AI systems used in law enforcement disproportionately targeted minority communities.
- Biased training data led to **discriminatory outcomes**, undermining civil liberties.
- **Lesson:** Fairness and transparency are non-negotiable in public sector AI.

💼 Hiring Algorithms – Gender Discrimination

- Applicant tracking systems favored male candidates over female ones due to biased historical data.
- Lack of oversight and explainability led to **systemic exclusion**
- **Lesson:** Trustworthy AI must be inclusive and auditable.

Meet Hafiz

PECB Certified Titanium trainer, Advisory Board Member, Public Speaker, Mentor, Writer, Advisor

Distinguished in

- Information Security
- Governance, Risk, Compliance
- Data Privacy
- Artificial Intelligence
- IT Service Management
- Business Continuity

Current Role(s)

- Lead Assessor & Auditor
- Advisor & Strategy Consulting
- vCISO

Entrepreneur

Co-founder and CIO,
AZAAN Cybertech Consulting

Director & Advisor,
Cyberverse Pty Ltd

Director & Trainer,
NEXTGEN KNOWLEDGE

Key Achievements

- ❑ PECB Certified Titanium Trainer
- ❑ PECB Certified Trainer of the year, 2022
- ❑ CISO of the Year 2021, 2022
- ❑ Top 100 South Asian Innovative Business Leaders, Victoria (Australia)
- ❑ Bx Business Excellence Awards – 2024, 2025
- ❑ '40 under 40' Business Elite Award 2024, Australia



PECB

Agenda

- Global trust in AI – Key Statistics
- Global Landscape of AI Regulations
- ISO/IEC 42001 - AI Management System
- NIST AI Risk Management Framework
- ‘Trustworthiness’ in AI
- AI Governance Model – GRAIT
- Building Trustworthy AI vs Bringing Trustworthy AI
- Closing Remarks
- Call to Action



Disclaimer

- ✗ Not intended as a technical deep dive into AI
- ✗ This is not a sales pitch or marketing gimmicks.
- ✗ Not a training on ISO/IEC 42001 or the NIST AI Risk Management Framework.
- ✓ Focus is on how AI frameworks synergistically support building and bringing trustworthy AI solutions



Why Trust Matters

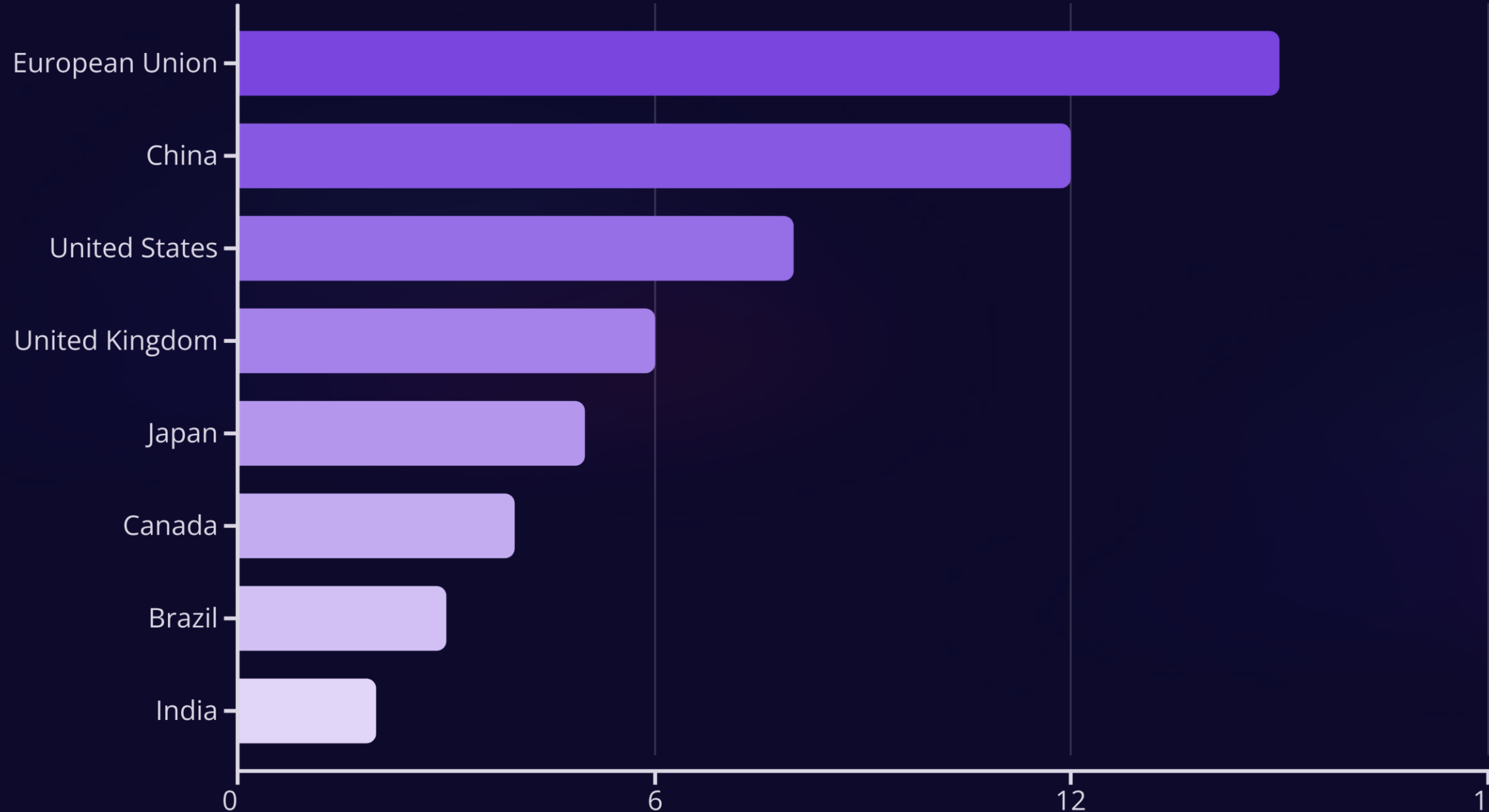
Customer Confidence: 70% of consumers say trust influences their buying decisions

Regulatory Pressure: EU AI Act, GDPR, NIST RMF, and APRA guidelines

Reputational Risk: AI misuse and data breaches can cause irreversible brand damage

Global Landscape of AI Regulations

The proliferation of Artificial Intelligence has prompted governments worldwide to develop new frameworks and legislation to manage its risks and ensure responsible adoption.



Introducing the NIST AI Risk Management Framework (AI RMF)

Voluntary Framework

Released January 2023 by the US Department of Commerce as a non-regulatory approach to AI governance

Risk-Based Methodology

Focuses on identifying, assessing, and mitigating potential harms from AI systems

Practical Guidance

Provides actionable steps for implementing trustworthy AI practices



NIST AI RMF emphasizes continuous learning, adaptation, and improvement of AI systems throughout their lifecycle

NIST AI RMF – Trustworthiness Principles

Principle	Description
Valid & Reliable	AI performs as intended
Safe	No harm to users or society
Secure & Resilient	Withstands threats and recovers
Explainable	Outputs are understandable
Accountable	Clear ownership and oversight
Privacy-Enhanced	Protects sensitive data
Fair	Minimizes bias and discrimination

NIST AI RMF: Core Functions & Categories

Govern

Establish policies, responsibilities, and organizational culture

- AI ethics committee formation
- Leadership accountability

Manage

Mitigate risks, monitor, and respond to incidents

- Feedback loops integration
- Incident response protocols



Map

Identify AI risks, context, and potential impacts

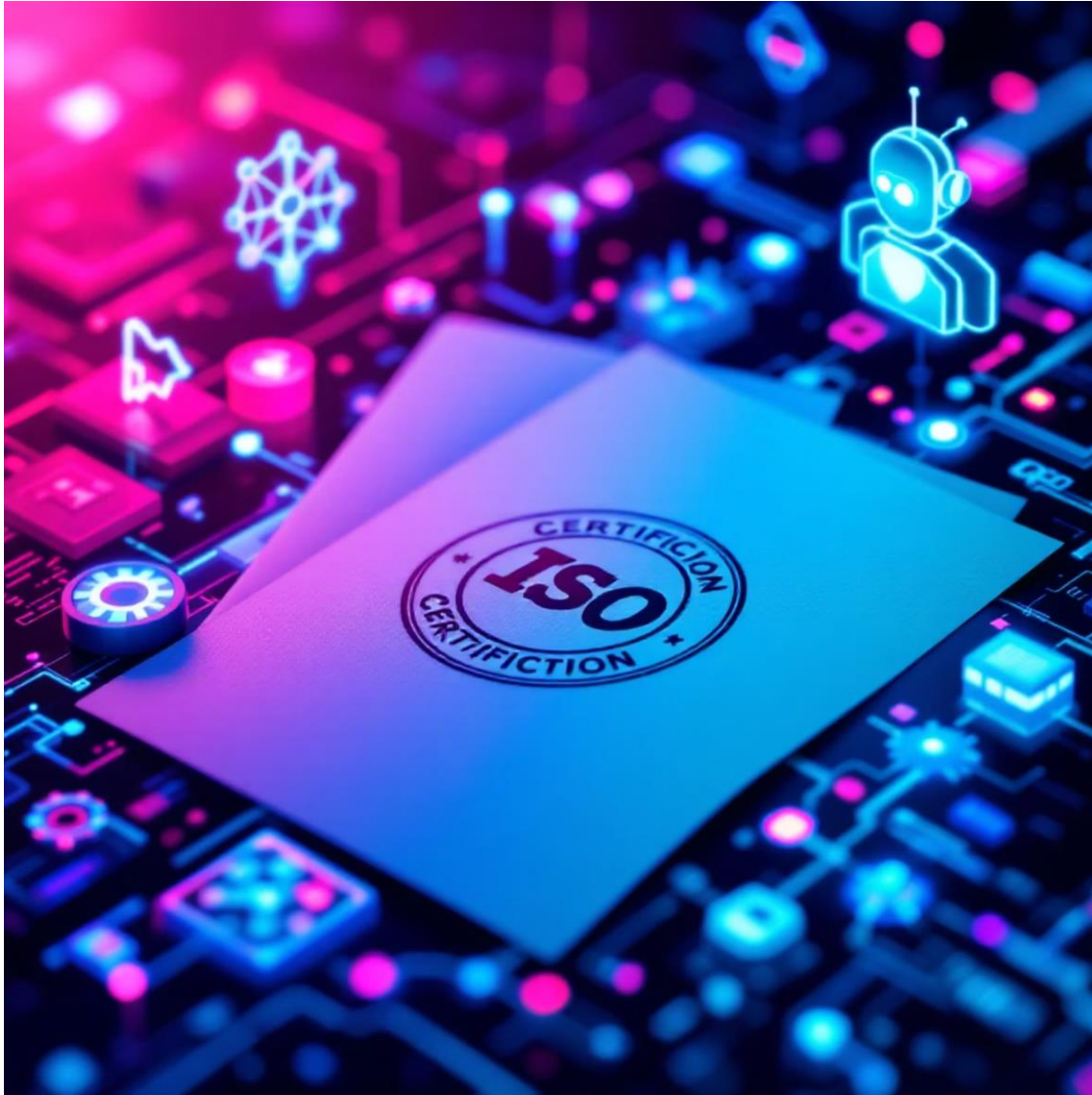
- Data source analysis
- Stakeholder impact assessment

Measure

Evaluate AI risks and trustworthiness

- Fairness metrics implementation
- Accuracy and performance testing

Introducing ISO/IEC 42001: AI Management System



- First International Standard

Published December 2023 by ISO/IEC, establishing the global benchmark for AI governance

- Comprehensive Framework

Covers the entire AI lifecycle from planning to retirement with auditable processes

- Certification Path

Enables organizations to obtain formal certification of their AI management practices

- Organizational Focus

Addresses governance processes rather than prescribing specific technical implementations

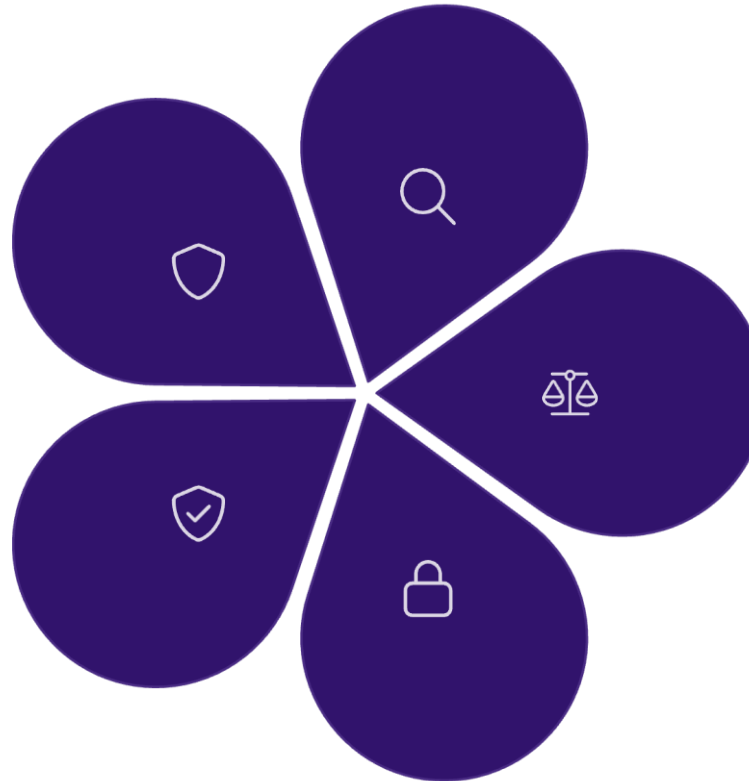
Key Principles & Controls

Accountability

Clear assignment of roles and responsibilities for AI impacts throughout the organization

Safety & Robustness

Measures to ensure resilience against adversarial attacks with 50+ specific controls



Transparency

Documented AI decisions and implementation of explainability measures

Fairness

Structured approaches to identify and mitigate bias in AI systems

Privacy

Robust data protection protocols throughout the AI lifecycle

The standard provides a structured approach covering AI planning, operation, and performance evaluation phases



Ethics and Security in AI

□ Trustworthiness in AI: A Unified Definition

Trustworthy AI refers to systems that are *ethically designed, securely deployed, transparently operated, and continuously monitored* to ensure they are fair, reliable, and aligned with societal values

4

The word
'trustworthiness'

Global AI Standards & Regulations



ISO/IEC 42001: AI Management System Standard

- **Focus:** Establishes a certifiable AI governance system
- **Trustworthiness Pillars:**
 - **Ethical AI principles:** fairness, accountability, transparency
 - **Lifecycle governance:** from design to decommissioning
 - **Stakeholder engagement:** inclusive oversight
 - **Continuous improvement:** monitoring and refinement



EU AI Act: Legally Binding Regulation

- **Focus:** Risk-based regulation of AI systems
- **Trustworthiness Pillars:**
 - **Risk categorization:** Unacceptable, High, Limited, Minimal
 - **High-risk AI requirements:**
 - Data quality and bias testing
 - Human oversight
 - External audits and transparency disclosures
 - **Accountability:** Fines up to €35M or 7% of global turnover for non-compliance

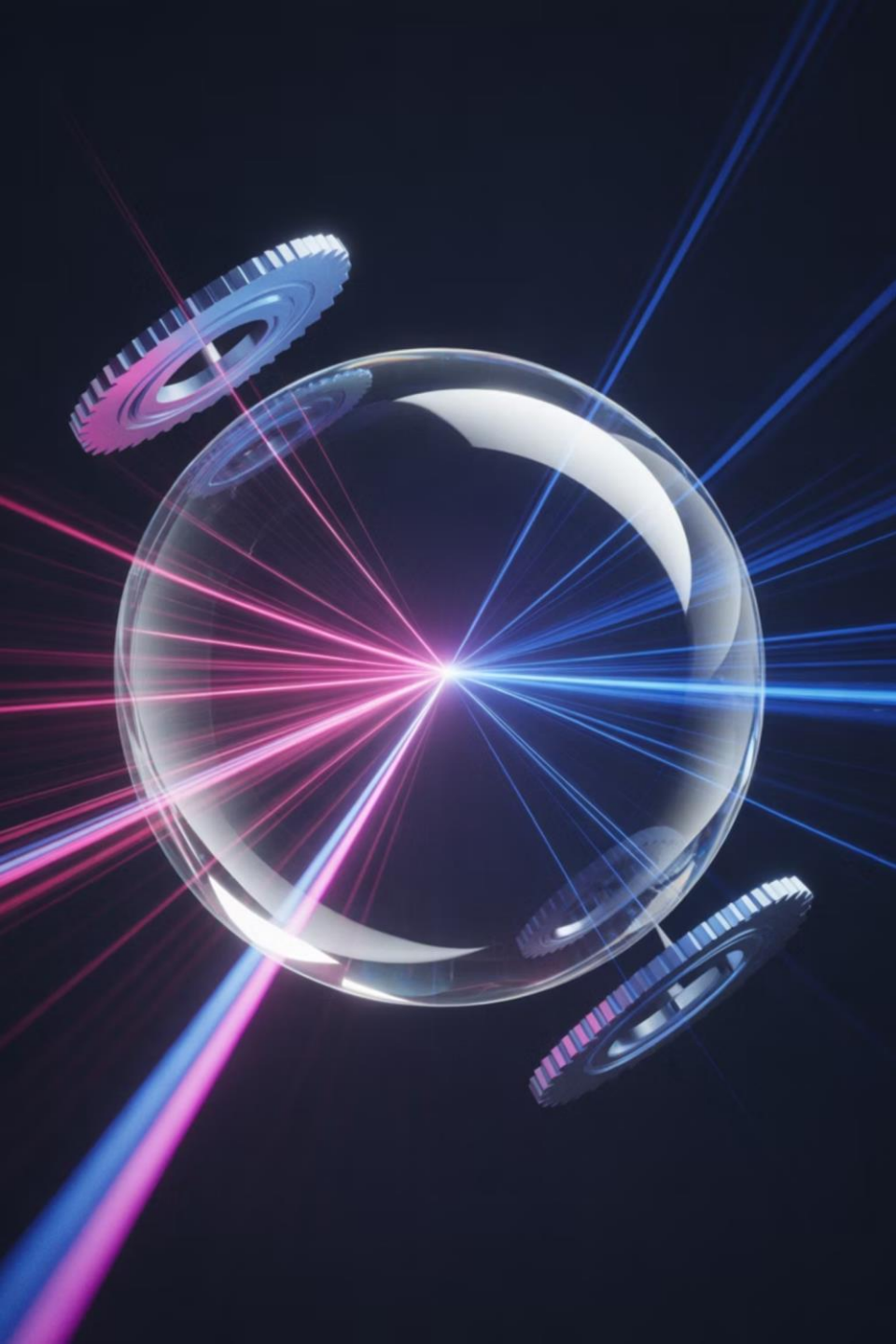


NIST AI RMF: Voluntary Risk Management Framework

- **Focus:** Flexible, principle-based guidance
- **Trustworthiness Pillars:**
 - **Map:** Understand context and risks
 - **Measure:** Assess fairness, reliability, and explainability
 - **Manage:** Mitigate risks with controls
 - **Govern:** Embed policies and accountability

Common Threads Across All Three

Principle	ISO/IEC 42001	EU AI Act	NIST AI RMF
Fairness	Required	Mandatory for high-risk	Core trustworthiness trait
Transparency	Embedded in lifecycle	Mandatory disclosures	Encouraged throughout
Accountability	Governance roles	Legal liability	Policy and oversight
Security & Privacy	ISMS integration	GDPR alignment	Risk mitigation controls
Explainability	Lifecycle documentation	Required for high-risk	Measured and managed



□ Strategic Insight

Together, these frameworks define **trustworthiness** not just as a technical attribute, but as a **governance outcome**. It's about building systems that earn confidence—through **design, documentation, and demonstration**.



AI Governance

■ What Is GRAIT?

- **GRAIT** is a conceptual framework that integrates traditional governance, risk, and compliance (GRC) with the emerging imperatives of **artificial intelligence oversight** and **digital trust**.
- It's designed to address the unique challenges posed by AI systems, such as algorithmic bias, explainability, data provenance, and ethical deployment.



Key Components of GRAIT



Governance

Establishing policies, accountability structures, and oversight mechanisms for AI systems.



Risk

Identifying and mitigating risks related to AI, including security, bias, and operational failures.



AI

Ensuring responsible development, deployment, and lifecycle management of AI technologies.



Trust

Building stakeholder confidence through transparency, fairness, and ethical use of AI.

Why GRAIT Matters

AI is no longer optional

—it's embedded in decision-making, automation, and customer interactions.

Trust is fragile

—misuse or poor governance of AI can erode public confidence and invite regulatory scrutiny.

Regulations are evolving

—frameworks like ISO/IEC 42001:2023 and the EU AI Act demand proactive governance.



⚙️ How GRAIT Aligns with Industry Trends

AI TRiSM (Trust, Risk & Security Management)

A Gartner-endorsed approach that supports GRAIT by enforcing fairness, reliability, and data protection in AI deployments.

KPMG's Trusted AI Framework

Offers ethical design and governance strategies that mirror GRAIT's principles.

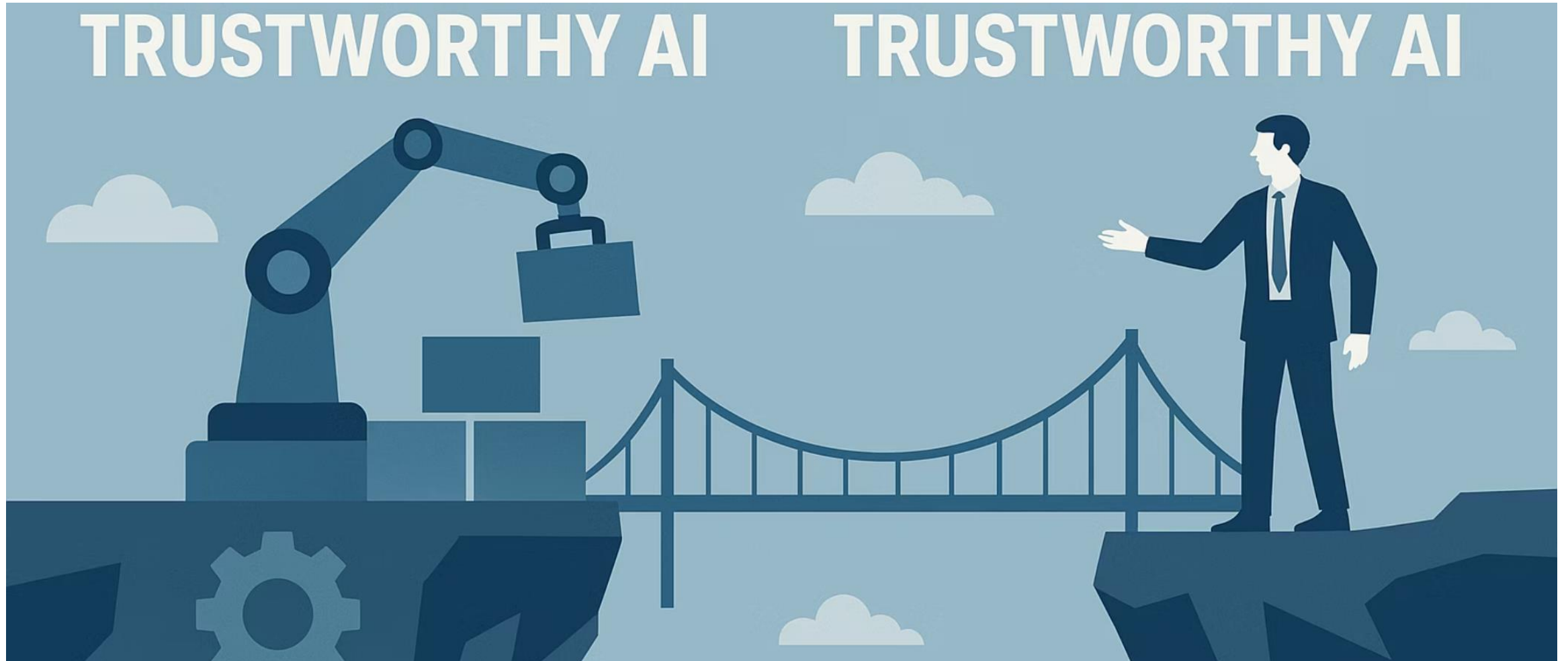
Open-Source Ecosystems

Platforms like Red Hat emphasize transparency and auditability, which are central to trust in AI governance.

□ Mapping to International Standards

GRAIT Pillar	ISO/IEC 42001:2023	ISO/IEC 27001:2022	NIST AI RMF
Governance	AI policy roles & leadership (6.1.2)	Governance & Org Structure	Govern Function: Organizational responsibilities
Risk	AI risk treatment planning (6.2.3)	Risk Management processes	Map Function: Contextual Risk Identification
AI Oversight	Development lifecycle checks (7.2)	Asset & Supply Chain controls	Measure & Manage Functions
Trust	Transparency and impact evaluation (8.2.2)	Security awareness & training	Govern & Map: Trustworthiness Attributes

Building Trustworthy AI vs Bringing Trustworthy AI



From Building to Bringing Trustworthy AI

⌘ Building Trustworthy AI

This is the **foundation**. It involves:

- Designing systems with fairness, transparency, security, and explainability
- Applying frameworks like **ISO/IEC 42001** and **NIST AI RMF** to guide development
- Ensuring governance, lifecycle management, and continuous monitoring are in place

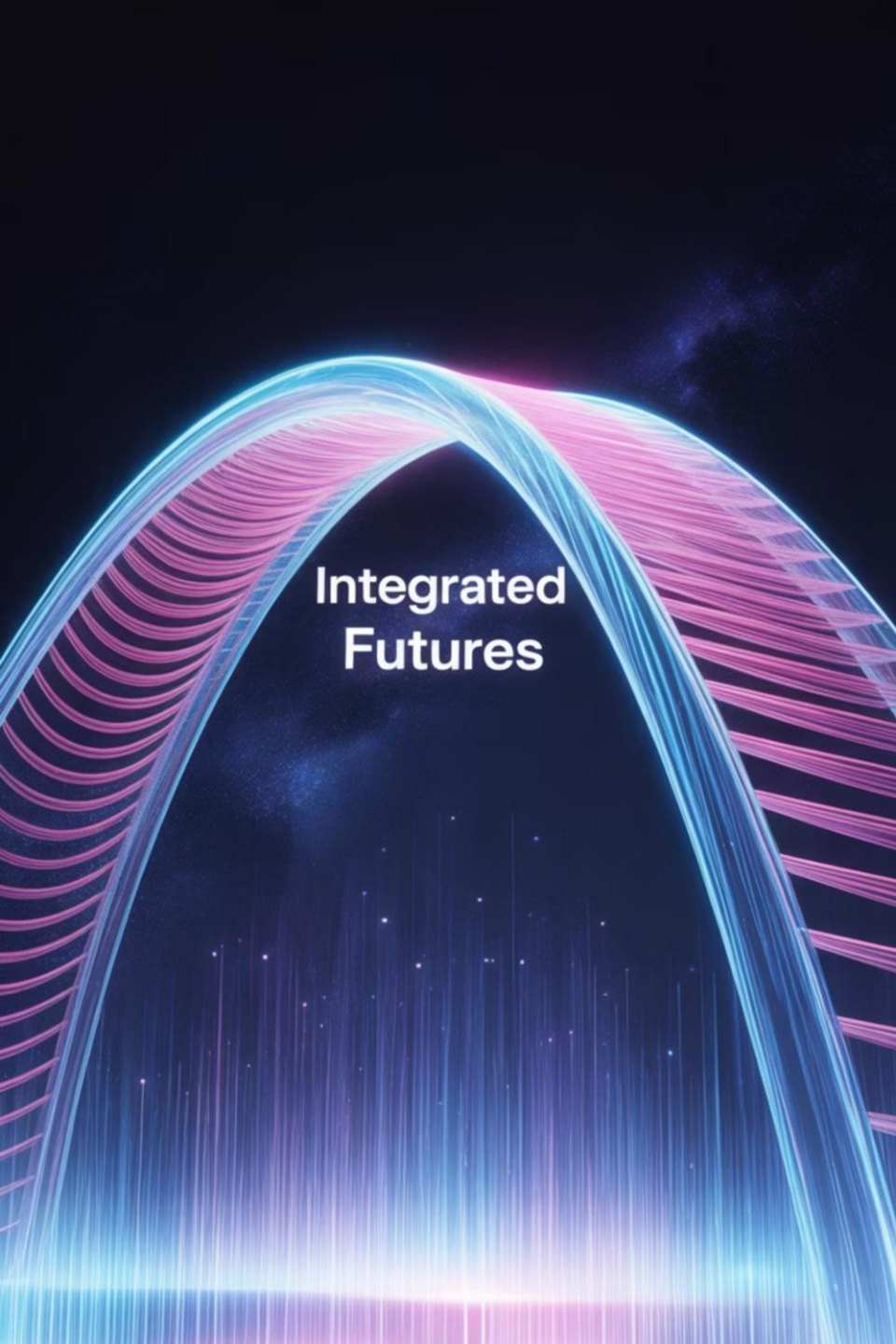
Without building trustworthy AI, there's no substance to stand on — you can't “bring” what doesn't exist.

🌐 Bringing Trustworthy AI

This is about **communication, credibility, and operationalization**:

- Translating technical trust into **stakeholder confidence**
- Demonstrating trustworthiness through audits, certifications (like ISO 42001), and disclosures
- Embedding responsible AI practices into **external perception**, partnerships, and regulation

Even the most ethical systems can be met with skepticism if trust isn't **brought to life** through outreach, transparency, and user engagement.



Crossing the Trust Bridge in the Age of AI

“Trust in AI isn’t declared—it’s demonstrated. And that demonstration begins long before we launch a product or policy.”

Executives and regulators alike are standing at a pivotal threshold: the bridge between AI potential and AI assurance.



On one side lies the engineering of ethical systems—built with the rigor of ISO/IEC 42001, the guidance of NIST AI RMF, and the backbone of secure, transparent, and explainable design. This is the **structure** of trust: precision, accountability, lifecycle governance.



But trust isn’t earned by infrastructure alone. It’s activated when we **bring** that integrity to light—through operational transparency, credible assessments, and clear communication.



This is the **walk across the bridge**—where governance frameworks become public confidence, and responsible innovation becomes sustainable adoption.



□ So Which Is More Important?

They're **interdependent**. But if we must prioritize:

Building trustworthy AI is the non-negotiable base. Without it, “bringing” trust is marketing spin.

Bringing trustworthy AI is what transforms internal governance into public trust — it's the **activation layer**.

Think of it like building a secure bridge (ISO/NIST) and then inviting people to walk across it. You need both — but **no bridge, no trust**.

Closing Remarks

AI market projected to reach **\$738B by 2030**, with a **CAGR of 38.1%**.

As adoption accelerates, trust becomes the foundation for successful implementation.


- ✓ Trustworthy AI isn't just a compliance checkbox—it should be a market differentiator.
- ✓ When we **build** it right, and **bring** it with integrity, we elevate brand equity, stakeholder assurance, and long-term resilience.





- ✓ A bridge without strong foundations collapses.
- ✓ Policies must start with technical trustworthiness and extend across real-world impact.

“Trust isn’t built on good intentions—it’s engineered through structured frameworks and continuous assurance.”

An overhead view of a business meeting around a dark blue wooden table. A man in a dark suit is shaking hands with another person whose arm and hand are visible at the bottom. A woman in a light blue suit sits across from them, looking at a document with a bar chart. On the table are a glass of water, a clipboard, glasses, and a coffee cup.

**“Trustworthiness must leave the lab.
Bring it forward—to regulators, users,
and partners who seek proof over
promise.”**

Call to Action

Executives:

“Trust is our differentiator. Let’s embed it early and bring it forward visibly.”

Regulators:

“Let’s co-create guardrails that turn principles into practice.”

Level Up Your Career with **PECB Skills**

Unlock unlimited learning with 3,000+ short 15-minute courses
in Cybersecurity, Information Security, Artificial Intelligence,
Business Continuity, Auditing & Compliance,
Digital Transformation & more.

Why Choose PECB Skills?

- Global training, local flexibility
- Short, practical courses for busy professionals
- Learn from top industry experts
- Original PECB product
- Earn CPD credits and get a certificate of completion

Get full access for only \$236 (Regular price: \$295)

Get 20% off Today!

Exclusive Discount for PECB Anniversary
Offer valid until August 5th

Start your 14-day free trial today!

Visit: growth.pecb.com/skills/

Or contact us at: skills.marketing@pecb.com

PECB Skills



#LevelUpinMinutes

Get In Touch



✉ hafiz.ahmed@cyberverset.net.au

✉ hahmed@nextgenknowledge.ai

in <https://www.linkedin.com/in/adnanahmed16/>



THANK YOU