



Vertrauen in KI –
ISO/IEC 42001 Zertifizierung für
KI-Managementsysteme und die
Rolle der internen Revision

2. Juli 2025



ISO 42001 – AI Management System

Your presenter today



AI Assurance
Lead Deutschland



Trustworthy AI
Assurance Co-Lead



Generative AI Service Delivery Transformation
Audit & Assurance CoE

Torsten Berge

Director
Audit & Assurance
Data & Analytics, AI & Algorithms

Deloitte GmbH Wpg | Hannover
+49 151 5807 2499
tberge@deloitte.de



Strawberry counting r's in "Strawberry"



ChatGPT o1-preview >

How many Rs in a word straberry

Thought for 5 seconds ▾

The word "straberry" contains three "R"s. If you meant "strawberry," it has two "R"s.

You sure

Thought for 7 seconds ▾

Yes, I'm sure. The word "straberry" contains three "R"s:

1. The first "R" is the third letter.
2. The second and third "R"s are consecutive, appearing as the seventh and eighth letters.

If you meant "strawberry," that word contains two "R"s.

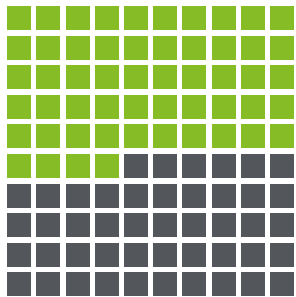
How many Rs in strawberry

Thought for 4 seconds ▾

The word "**strawberry**" contains two "R"s.

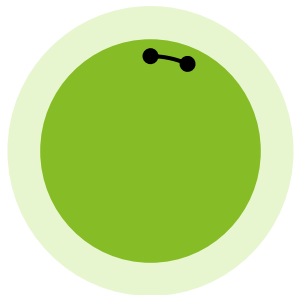
Wrong

AI has arrived in the business



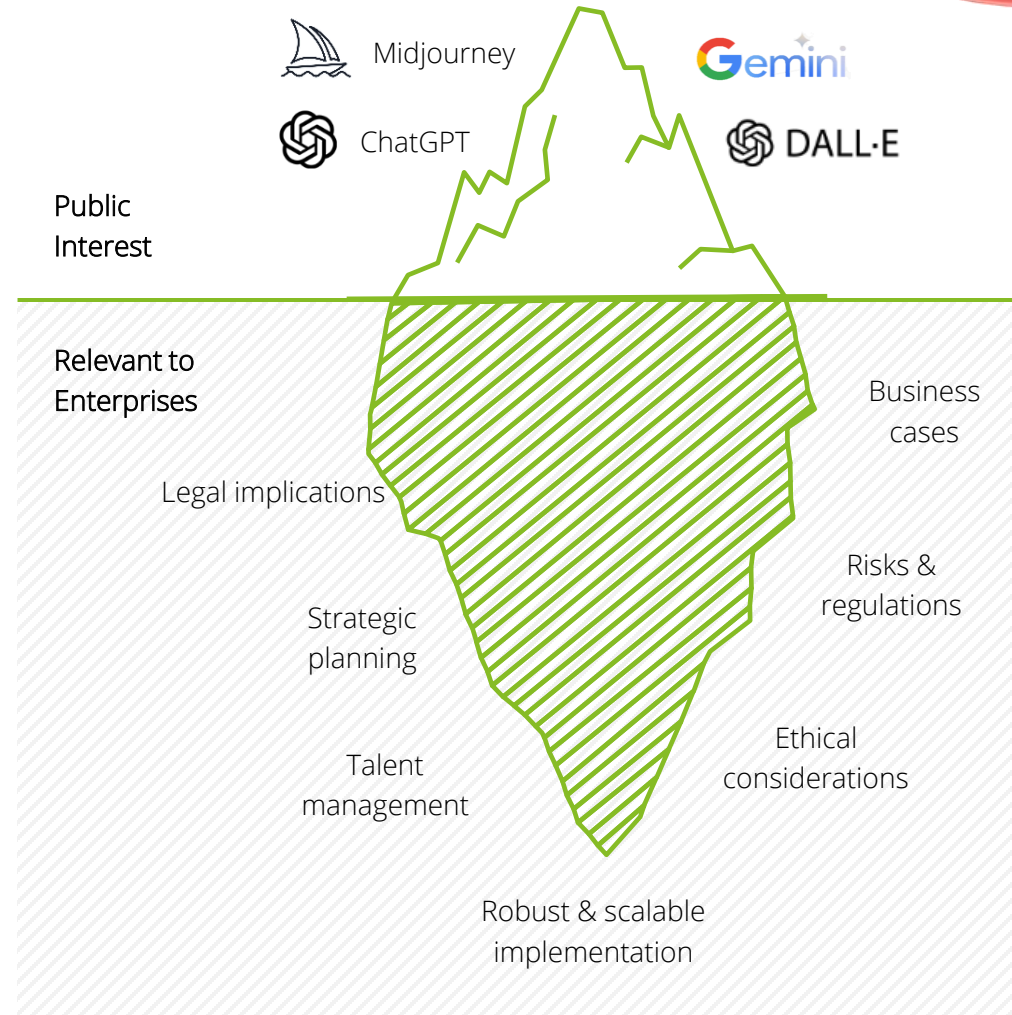
54%

of all finance departments currently explore GenAI opportunities



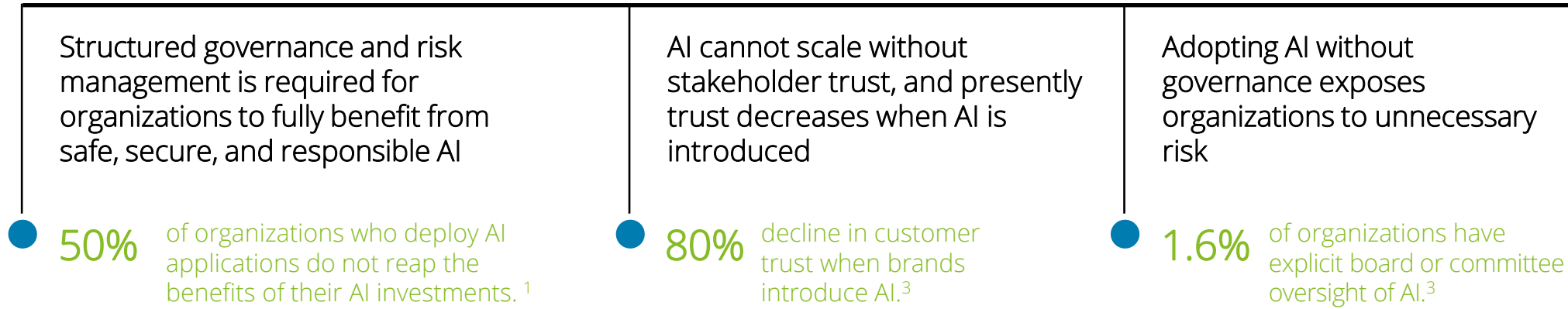
4%

have GenAI solutions fully implemented / in production



Sources: CFO Survey Herbst (Deloitte, 2024)

Barriers to the Benefits of AI



Risk & regulatory concerns growing as AI shifts from PoC to Production

36%

of enterprises worry about regulatory compliance

30%

experience difficulty in managing risks

29%

lack a governance model

Sources: 1: Deloitte, *State of AI in the Enterprise*, 2024. 2. Deloitte TrustID Brand Index, Jan 2024. 3. Harvard Law School, Forum on Corporate Governance, Apr 2024. 4. Deloitte, *State of Generative AI in the Enterprise: Quarter Four Report, January 2025*. State of Generative AI in the Enterprise Survey, Q1 (Oct./Dec. 2024) N (Total) = 2,774; Q4 (July/Sept. 2024) N (Total) = 2,773; 14 countries common to both data set


A black and white photograph of a large steamroller paving a road. The steamroller is positioned on the right side of the frame, moving from right to left. The road surface is being laid out in front of it. The background shows a line of trees under a cloudy sky.

“Once a new technology rolls over you, if you're not part of the steamroller, you're part of the road”

Stewart Brand | American writer


AI risk management framework: mitigation of key risks

Implementing an appropriate AI risk management framework in a timely manner provides organisation with benefits which are not only limited to risk management



AI outputs may be unethical, biased, or factually incorrect (e.g., hallucinations), potentially damaging trust, brand value, and stakeholder confidence.

REPUTATIONAL RISK




Weaknesses in AI systems can be exploited via data poisoning, adversarial inputs, or model theft, compromising system integrity and security.

CYBER RISK



Flawed AI systems can result in significant revenue loss, investment errors, or compromised financial reporting accuracy.

FINANCIAL RISK




Failure to govern AI outputs and applications effectively may lead to breaches of legal obligations and heightened regulatory scrutiny.

REGULATORY RISK



Integrating AI into existing infrastructure can create challenges with system compatibility, scalability, and security, especially in legacy environments.

TECHNOLOGICAL RISK



Faulty model logic or poor-quality training data can lead to inaccurate outputs, undermining business decisions and automated processes.

MODEL RISK



Reliance on external AI vendors may introduce risks around supply chain security, data governance, and insufficient due diligence.

THIRD-PARTY RISK



Using AI for high-level decision-making without sufficient validation may lead to flawed strategies, competitive missteps, or resource misalignment.

STRATEGIC RISK



AI systems may inadvertently expose personal data or sensitive IP, increasing the risk of data breaches and legal liability.

PRIVACY RISK

View beyond risk management

Organizations should look beyond responding to regulation simply as a compliance exercise, recognizing the operational benefits and competitive advantages that can be generated from implementing appropriate governance and risk management over their AI systems.

Trustworthiness and Transparency

A risk management framework which aligns with leading standards supports clearer system boundaries and enhances consumer trust in AI transparency.

It promotes the adoption of ethical principles, ensuring systems are fair, unbiased, and respectful of human rights.

Operational Efficiency

Stronger governance leads to more effective risk management before and after deployment, enabling smoother coordination between AI systems and human oversight.

This drives cost savings and improved system performance.

Competitive Advantage

Well-governed, trustworthy AI systems are more likely to be adopted and relied upon in the market.

A structured governance approach supports continuous improvement and innovation, helping organisations maintain strategic leadership.

Governance and Compliance

Establishes a consistent framework for the development, deployment, and oversight of AI systems.

It supports alignment with evolving regulatory standards, reducing the risk of non-compliance and legal exposure.

Managing AI: AI Compliance Frameworks

Starting Question

Decide which EXTERNAL norms and regulations you will focus on

EU AI Act

- Effective from Aug 2024
- Classification of AI use cases:
 - General Purpose AI
 - Certain Systems
 - High Risk
 - Unacceptable
- Documentation / Self-Assessment Requirements
- AI Literacy requirements
- Transparency requirements
- Risk and Quality Management System
- Registration in EU Database

NIST AI RMF – AI Risk Management Framework

- Voluntary guidance to improve the ability to incorporate AI-related considerations
- Industry-wide accepted standard
- Detailed Framework with 70+ requirements within four main categories: Govern, Map, Measure, Manage

ISO 42001 – AI Management System

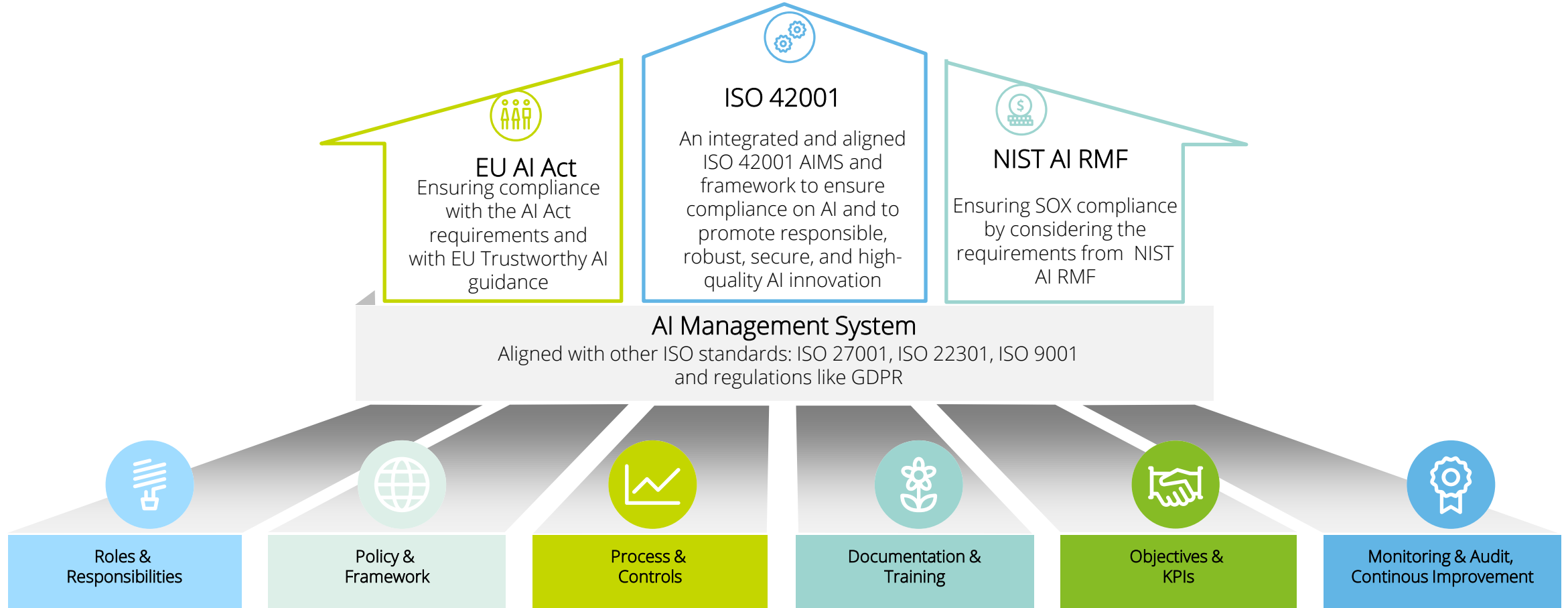
- Guideline for the whole AI application lifecycle regarding security, fairness, transparency and (data) quality
- Can be integrated into existing management system environment
- Extends ISO information security, risk and quality management practices

IDW PS 861 – Prüfung von KI Systemen

- The IDW PS 861 represents a voluntary auditing standard for auditing AI systems.
- The PS 861 presents a generally applicable catalog of requirements and checks related to AI Management System to mitigate possible risks from the use of AI

Integrated methodology & approach

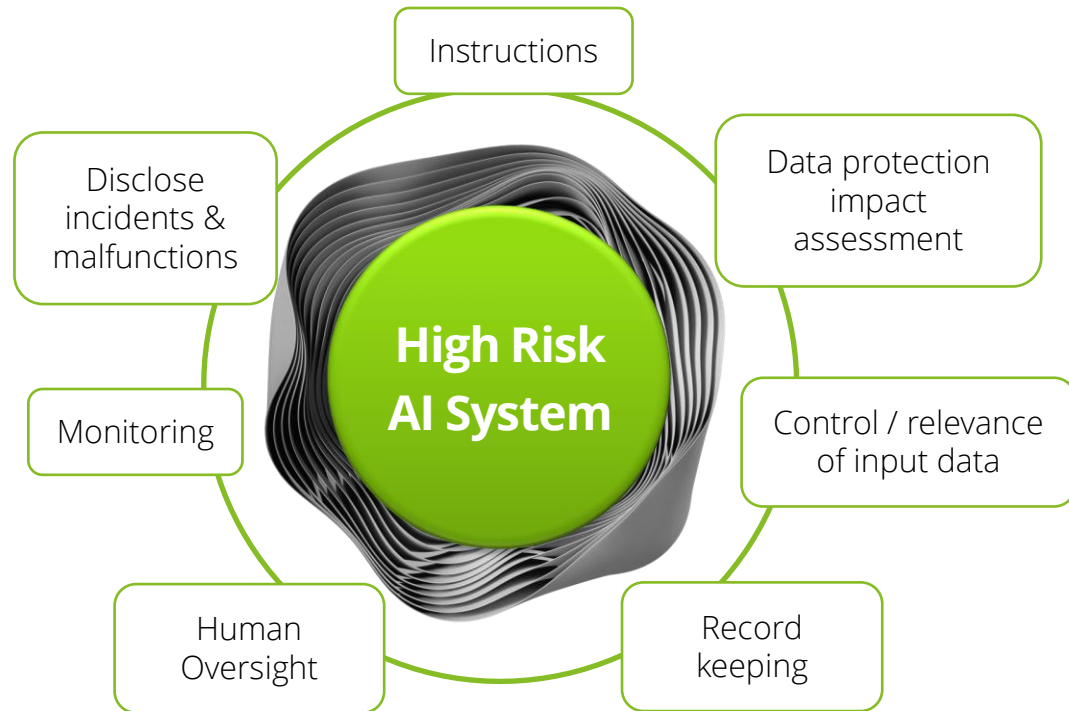
ISO 42001 as the umbrella that integrates AI compliance requirements arising from SOX and the EU AI Act, but also benefits from existing ISO standards to create a robust and comprehensive AI management system



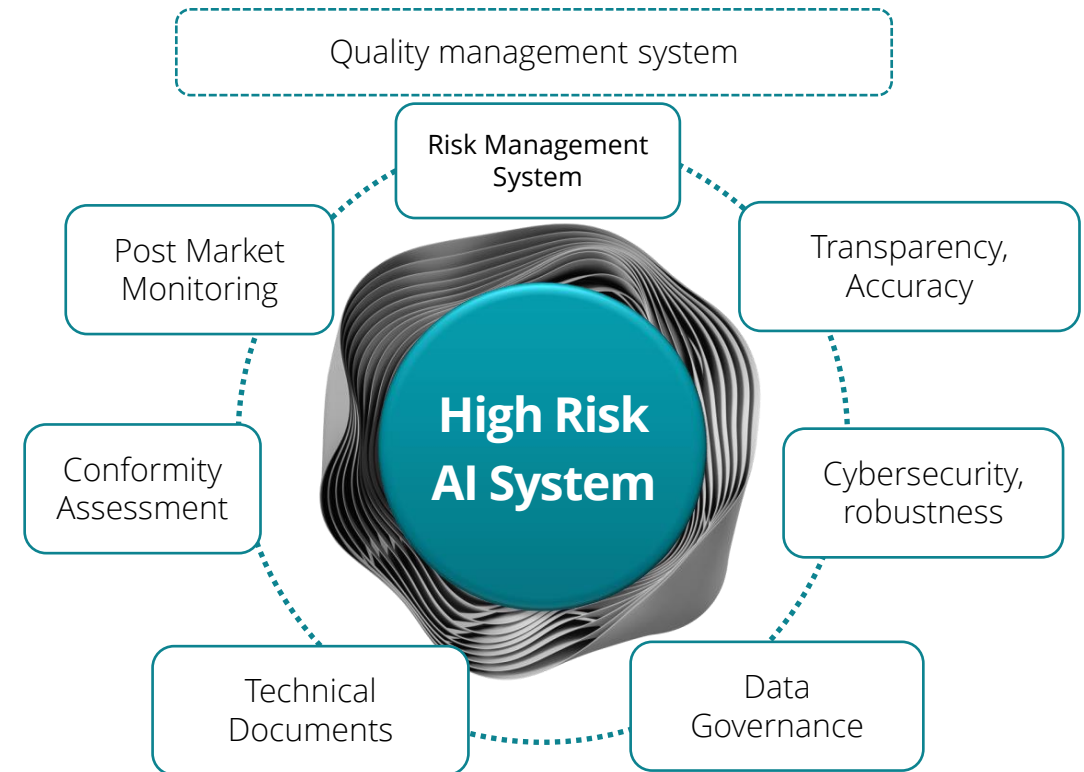
EU AI Act - How will it affect companies?

Compliance with the AI Act will be a condition for the placing of AI systems on the market, their putting into service or use.

Requirements* for Deployers

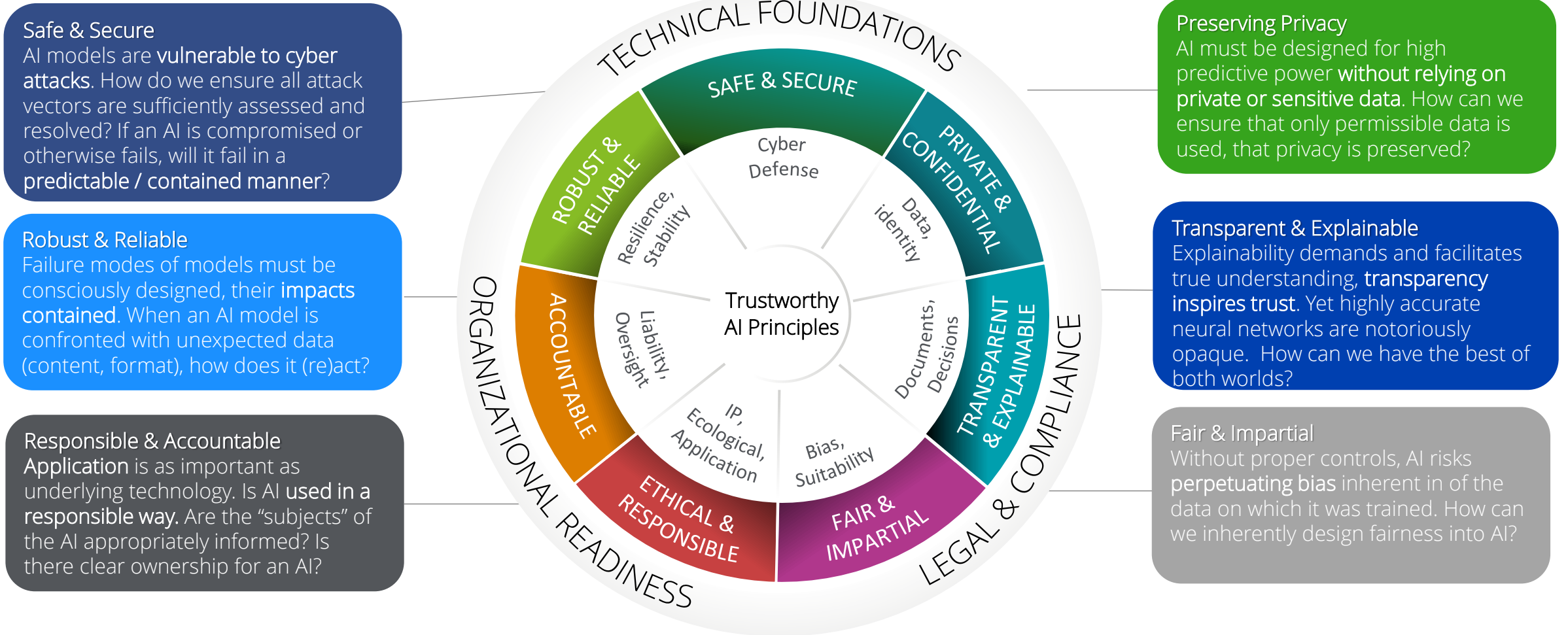


Additional Requirements* for Providers



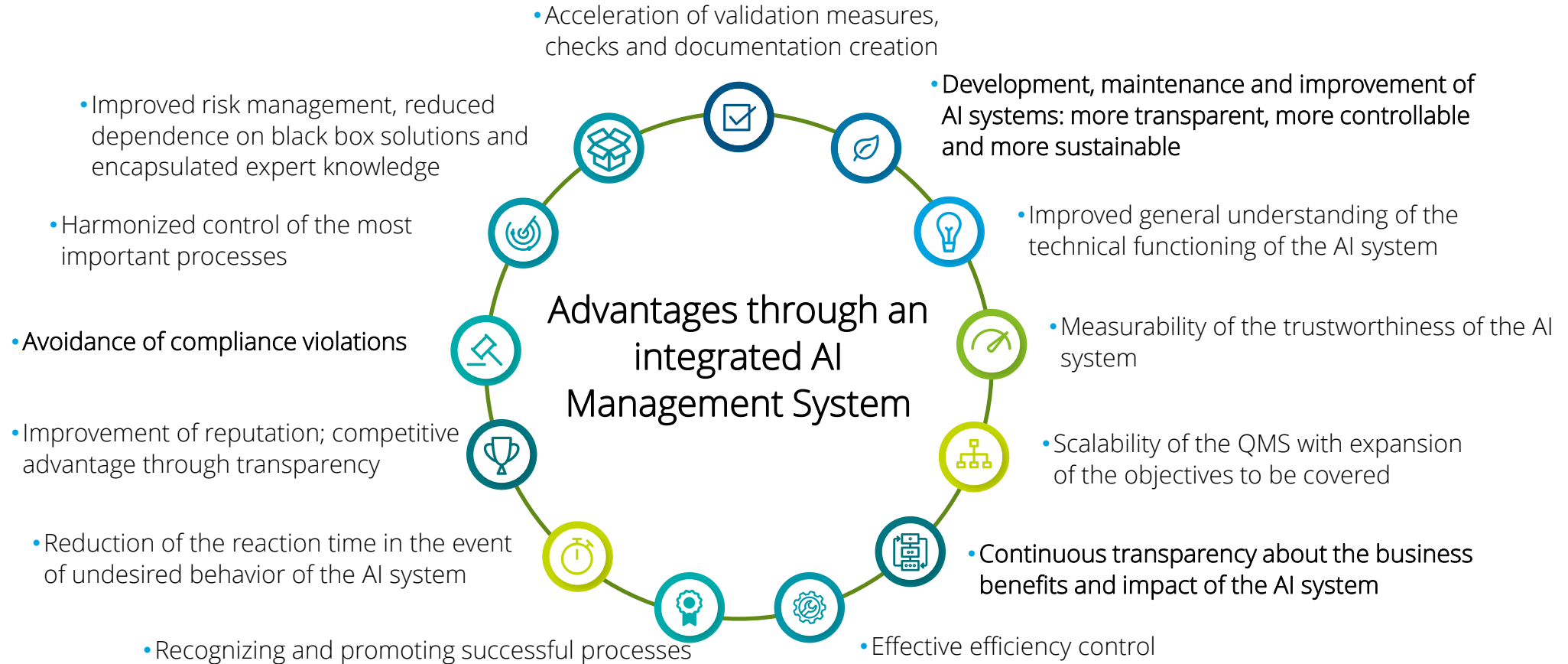
Trustworthy AI

Key requirements for a successful adoption



Advantages of an AIMS

Improving AI Governance and Quality Management System over AI will lay the foundation to develop, maintain and manage successful, efficient, high-quality and innovative AI solutions



ISO/IEC 42001 – A prominent standard for the effective management of AI risks

ISO/IEC 42001 is an international standard that specifies requirements for establishing, implementing, maintaining, and continually improving an Artificial Intelligence Management System (AIMS) within organizations.

It is designed for entities providing or utilizing AI-based products or services, ensuring responsible development and use of AI systems.

Why ISO/IEC 42001 matters

As the first certifiable global standard for AI governance, ISO/IEC 42001 translates regulatory expectations and ethical principles into operational requirements, enabling organisations to build structured, auditable, and accountable AI systems.

As legal, reputational, and technical risks increase, the standard offers a practical foundation for managing AI across its lifecycle – responsibly, transparently, and at scale.

Published December 2023



ISO/IEC 42001 lays the foundation for enterprise-wide AI assurance. It enables cross-functional collaboration, supports stakeholder communication, and builds the structures required for sustainable, responsible innovation. The standard is designed to be scalable; applicable to organisations across industries, from early adopters to advanced AI users.

What ISO/IEC 42001 aims to achieve

Responsible AI governance

- Embeds ethical, transparent, and accountable AI practices
- Aligns leadership and operational responsibilities

Legal and regulatory readiness

- Supports alignment with EU AI Act, GDPR, and NIST RMF
- Prepares organisations for audit and assurance

Lifecycle risk management

- Identifies and manages risks from design to deployment
- Enables structured monitoring and incident response

Continuous improvement and trust

- Drives feedback loops and measurable improvement
- Enhances stakeholder confidence in AI oversight

ISO/IEC 42001 key requirements

Organisational context and scope

- Define AI usage and role.
- Establish scope and boundaries of AI management.

AI risk management and controls

- Assess AI risks, including ethical impacts.
- Implement controls for safe, transparent AI.

Operational practices

- Manage AI lifecycle processes.
- Address risks in outsourced AI and manage incident response.

Leadership and governance

- Assign AI governance to leadership.
- Communicate AI policy aligned with values and objectives.

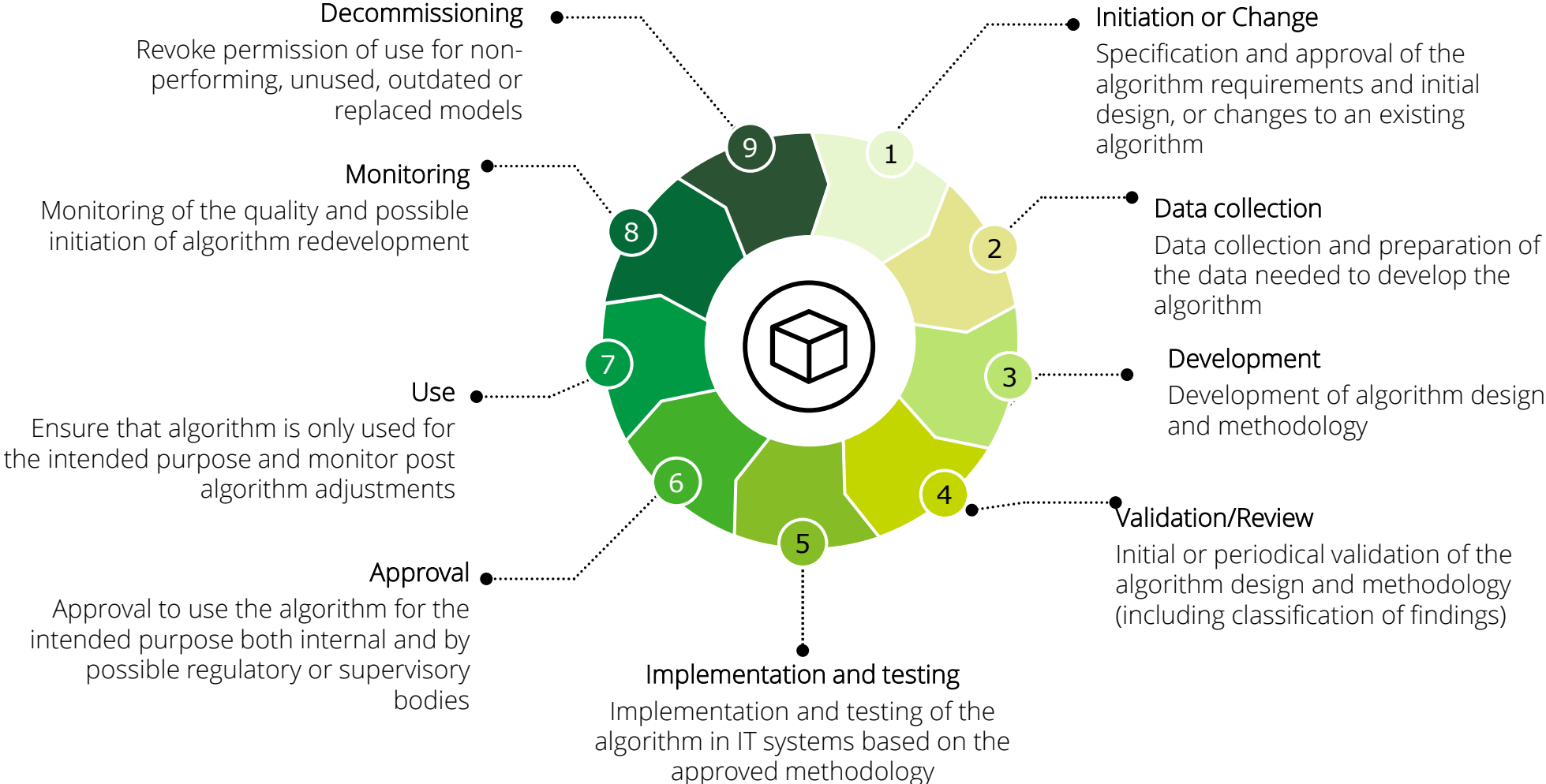
Monitoring, evaluation, and improvement

- Measure AI effectiveness.
- Conduct audits for improvement.

Support and documentation

- Ensure staff competence in AI.
- Maintain documented for control and traceability.

AI Quality & Risk Management along the entire AI Lifecycle



Wrap-Up: Integrated methodology & approach

Summary: ISO 42001 AIMS as the core to fulfil regulatory AI requirements: identifying gaps and overlaps between ISO 42001, SOX AI (NIST AI RMF) and EU AI Act requirements and align the AIMS with existing ISO frameworks / management systems

Integration with other AI related regulations



NIST AI RMF (SOX) or other local regulations

- Primarily focusing on the accuracy and reliability of **financial reporting** and internal controls over financial reporting in consideration with NIST AI RMF.
- Considering AI systems used in **financial processes** (e.g., fraud detection, credit scoring, financial forecasting, accounts payable/receivable automation). Such systems are directly subject to SOX testing.



EU AI Act

- Adjusting controls with respect to the requirements of the EU AI Act (e.g. based on risk levels).
- Considering **Data Governance**, **Documentation**, and **Transparency**: e.g high-quality data, detailed **technical documentation**, and clear **explanations** of the AI system's functionality.
- Including **human oversight**, standards for accuracy and cybersecurity, **conformity assessments**, and **post-deployment monitoring**.

Others regulations to be considered

- GDPR

Alignment with existing ISO frameworks



ISO 27001

- Leverage existing ISO 27001's documentation and controls to use synergy effects to meet the requirements from the ISO 42001.
- Identifying synergies in various areas such as; **information security**, **governance** structures, risk assessment, **development lifecycle**, and stakeholder communication.



ISO 22301

- Integrating ISO 22301's aspects on **Business Continuity Plan (BCP)**, Resilience, **Backup Strategies**, and communication with the stakeholders.

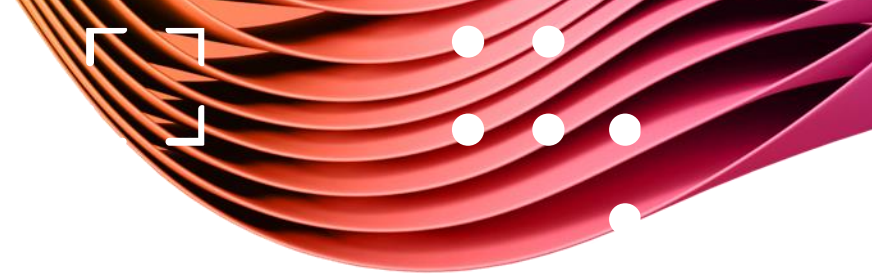


ISO 9001

- Using ISO 9001's focus on QMS documentation, process approach, and **continuous improvement**.



Wrap-up: Focus points for Internal Audit



01 **AI strategy and governance**

Internal audit should consider their organisation's approach to the governance of AI. This should include a review of the organisation's AI strategy, business case(s) and to what extent AI risks have been considered..

02 **Policy, standards and guidelines**

Internal audit should consider reviewing any AI policy the organisation has developed, including acceptable usage guidance and/or policy which defines the parameters of AI system development.

03 **AI inventory**

Internal audit should consider whether an AI inventory has been developed by the business. An AI inventory should include all AI systems in development and use with system details and risk management considerations.

04 **AI risk management, Internal Control and culture**

Current risk management processes may need to be amended to ensure that risks associated with AI are proficiently covered. AI Governance & risk management frameworks should be integrated into the current GRC processes and procedures to ensure systems utilising AI are effectively managed, governed, and monitored.

05 **Regulatory readiness**

Internal audit should understand how the organisation is staying up to date with the fast-moving regulatory environment. Organisations need to consider regulations in all the geographies they operate in.



Wrap-Up: The Pillars of Good Governance

Governance encompasses organizational **structures, systems, practices** and **processes** to enable management to fulfill its duties to ensure quality, manage risk, enforce accountability and fulfill compliance obligations. It begins with policies & procedures, extends to decision-making bodies & approval processes, through to review boards and controls, including the components of the AI Act's QMS^{-a)} and RMS^{-b)}.

People

- Steering Committee (Ethics, Investment, Quality, Risk)
- Advising AI CoE^{-c)}
- Empowerment & responsibility
- Qualifications & skills
- Training / Upskilling Facility
- Relevant stakeholders

Policy

- Oversight & accountability (QMS, RMS)
- Distinct roles & responsibilities
- Developer guardrails (QMS)
- Business continuity & risk mitigation strategies / FMEA^{-d)}
- Communication (Awareness)
- Compliance

Processes

- Overview / inventory
- Approval process & quality gates (QMS)
- Escalation & issue review
- Risk assessment / classification
- Internal Controls
- Testing / Model Validation
- Change requests / issue resolution

Systems

- Operationalizing the practices, automating the processes
- Workflows (QMS)
- Technical documentation (QMS)
- Model monitoring over time, triggers & alerts (RMS)
- Issue logging and resolution (RMS)
- Single source of truth, serving information needs at various levels

(a- QMS = Quality Management System
(b- RMS = Risk Management System
(c- CoE = Center of Excellence or Expertise
(d- FMEA = Failure Modes & Effects Analysis

AI Assessment Framework

Deloitte's ISO 42001 Maturity Assessment

FOR ILLUSTRATION PURPOSE

ISO 42001 Gap Analysis: Assessment of the initial and ongoing status of ISO 42001 implementation

1. General Information

Item	Details
Organization Name	[Enter Name]
Assessment Date	[DD/MM/YYYY]
Assessor (s)	[Names]
Scope of Assessment	[Define Scope]

2. Initial Status Assessment

ISO 42001 Clause	AI Management System (AIMS) Mandatory Requirements – Description	Status	Resources-Customer
4	Information Security Management System		
4.1	Understanding the Organization and its Context		
4.1	This step requires the organization to analyze both external factors (e.g., market conditions, competition, legal/regulatory environment) and internal factors (e.g., organizational structure, resources, and capabilities) that might influence the management of AI systems.		
4.2	Understanding the Needs and Expectations of Interested Parties		
4.2	This involves recognizing the demands of key stakeholders (e.g., customers, regulators, employees, and society) who have a direct impact on how AI systems should be governed and managed.		



AI Risk Likelihood Assessment



Deloitte AI & Algorithm Assurance

With our AI & Algorithm Assurance expertise we support our clients building Trust in AI systems and becoming ready for upcoming regulations

What makes us unique:

Deloitte's integrated Algorithm & AI Assurance approach:

- Expertise in **AI Governance, AI Risk Assessment, AI Compliance** and **Internal Control** combined with **technical expertise** in AI model development and data and model validation
- Automated **audit tools** for data and model validations (with aiStudio)
- Global Algorithm & AI Assurance organization with **specialist resources** for specific industries and region



Trustworthy AI Framework



AI Risk Management Framework









AI Control Framework Toolkit



Audit Tool Suite for data and model validation

How we can help:

OUR AI ASSURANCE SERVICES

-  Regulatory compliance: EU AI Act readiness assessment, IDW PS 861, ISO 42001 readiness
-  (Generative) AI Governance and Internal Controls assessment
-  Review and audit of AI Management System, e.g. ISO 42001, NIST AI Risk Management Framework
-  Assessment of model development and operation
-  AI & Algorithm Risk Assessment and Risk Classification
-  Model monitoring and risk mitigation; validation of guardrails

OUR AI RISK MANAGEMENT FRAMEWORK



The **Deloitte AI Risk Management Framework** covers the regimes of the **EU AI Act, NIST AI RMF, ISO 42001** and **IDW's PS 861** in the form of control objectives that takes into account both the lifecycle stages and the elements of the AI system

Let's start the conversation



Torsten Berge
AI Assurance Germany
+49 151 5807 2499
tberge@deloitte.de



Mujtaba Mehmood
AI Assurance Germany
+49 151 1488 1639
mmehmood@deloitte.de

Questions?

The global reach of Deloitte AI Institute ecosystem

Thought Leadership



Differential Privacy & Synthetic Data



Transparency in Model Predictions



Preserving Privacy in AI Models

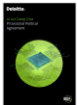


Striving for Fairness in AI



Making AI Robust for the Real World

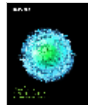
Marquee Studies



A Deep Dive into the EU AI Act



AI Dossier: Considerations for Risk & Trust



The implications of Generative AI for businesses



State of AI 5th Edition

5B

Media Impressions

50M

Social Media Impressions

5K+

Executives Surveyed for State of AI 4th and 5th Editions

Trustworthy AI Audit Tools



Password / API Security



Transparency / XAI



Re-Identification Risk (Privacy)



Robust Operation



Bias / Fairness



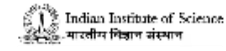
Large Language Model Risk



Research Collaborations



Deloitte is the founding sponsor of Centre for Trustworthy Technology (WEF)



and others...

Joint Publications



VentureBeat

[Deloitte State of AI Report 2022 Calls Out Underachievers](#)

FORTUNE

[Investors are pouring billions into AI. It's time for a commensurate investment in AI governance.](#)



[How Generative AI is changing creative work](#)



[Deloitte AI Academy: A training ground for AI talent](#)

ComputerWeekly.com

[How to leverage the artificial intelligence solar system](#)



Deloitte refers to one or more of Deloitte Touche Tohmatsu Limited (DTTL), its global network of member firms, and their related entities (collectively, the “Deloitte organization”). DTTL (also referred to as “Deloitte Global”) and each of its member firms and related entities are legally separate and independent entities, which cannot obligate or bind each other in respect of third parties. DTTL and each DTTL member firm and related entity is liable only for its own acts and omissions, and not those of each other. DTTL does not provide services to clients. Please see www.deloitte.com/de/UeberUns to learn more.

Deloitte provides industry-leading audit and assurance, tax and legal, consulting, financial advisory, and risk advisory services to nearly 90% of the Fortune Global 500® and thousands of private companies. Legal advisory services in Germany are provided by Deloitte Legal. Our people deliver measurable and lasting results that help reinforce public trust in capital markets, enable clients to transform and thrive, and lead the way toward a stronger economy, a more equitable society and a sustainable world. Building on its 175-plus year history, Deloitte spans more than 150 countries and territories. Learn how Deloitte’s approximately 457,000 people worldwide make an impact that matters at www.deloitte.com/de.

This communication contains general information only, and none of Deloitte GmbH Wirtschaftsprüfungsgesellschaft or Deloitte Touche Tohmatsu Limited (“DTTL”), its global network of member firms or their related entities (collectively, the “Deloitte organization”) is, by means of this communication, rendering professional advice or services. Before making any decision or taking any action that may affect your finances or your business, you should consult a qualified professional adviser.

No representations, warranties or undertakings (express or implied) are given as to the accuracy or completeness of the information in this communication, and none of DTTL, its member firms, related entities, employees or agents shall be liable or responsible for any loss or damage whatsoever arising directly or indirectly in connection with any person relying on this communication. DTTL and each of its member firms, and their related entities, are legally separate and independent entities.