TRUSTED AND SECURE AI
FROM DELOITTE AND IBM

Platinum Partner

Adoption of artificial intelligence (AI) and Generative AI (Gen AI) across commercial and government and public service (GPS) sectors is on the rise. Most organizations (94%) believe AI is critical to their success over the next five years[1], while 79% of public sector leaders indicate their agencies already use AI[2].

But without proper governance—which includes defined frameworks, strong underlying processes, and cutting-edge, flexible technical platforms—organizations and agencies can quickly find themselves exposed to significant risks due to untrustworthy AI-enabled services. About half (48%) of surveyed executives indicated "insufficient trust" as the top obstacle to expanding AI usage[2], while 78% of professionals believe more government regulation of AI is necessary[3].

From customers and citizens to workers and regulators, establishing trust across the AI spectrum is critical to harnessing its full potential and gaining a competitive edge.

Together, Deloitte and IBM can help organizations operationalize AI and GenAI governance, strategies, mitigation plans, and policy guardrails that cover technical, regulatory, ethical, and reputational considerations. Leveraging the foundational elements of trustworthy AI governance—comprehensive enterprise AI policies, full AI lifecycle tracking and standards, real-time risk measurement and mitigation, and alignment to regulation and legislation—organizations can establish trustworthy AI operations and outputs.

What are the threats to be mitigated by trustworthy AI?

Organizations see threats and risks from both internal and external threats. Bad actors present continual challenges from the outside while AI governance present their own evergreen challenges from within – spanning both AI inputs and more systemic, organizational-wide issues:

BIAS
if AI model training data is biased, then the outputs could also exhibit biases

HALLUCINATION
AI models can generate statements that are unfounded and factually false

PRIVACY VIOLATIONS
confidential information being used without consent

MALICIOUS INTENT
bad actors can target AI models; security requires constant vigilance

IP INFRINGEMENT
when training AI models, the data used could potentially include protected IP

Mitigating these risks is made more complex by emerging **regulatory compliance pressures:**

- As early as December 2024, **US federal agencies** will be required[4] to publicly disclose AI use case inventory, share and release AI code and data assets, and a list of other requirements intended to address and manage AI usage risks.
- Similarly, the **European Union's AI Act,** which entered into force in August 2024, includes requirements to address citizen health, safety, and fundamental rights when it comes to AI development and deployment in Europe.

[1] Gartner Press Release, "Gartner Poll Finds 55% of Organizations are in Piloting or Production Mode with Generative AI", Oct 2023.
[2] CSO "The Rising AI Tide: Achieving Digital Resilience", Feb 2024.
[3] Deloitte, State of Generative AI in the Enterprise, 2024.
[4] Office of Management and Budget "Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence", Mar 2024.

Plot a course forward by operationalizing trust

To help organizations and agencies navigate today's maze of external and internal threats, regulatory pressures, and enterprise risks, Trusted and Secure AI solutions from Deloitte bring together Deloitte and IBM capabilities to operationalize AI and GenAI strategy and transformation through:

DEFINED SERVICES AND PRINCIPLES

Deloitte's trustworthy AI and GenAI services—designed to help organizations build up the internal teams, resources, and guidelines they'll need to begin deploying AI safely and responsibly.

CUTTING-EDGE TECHNOLOGY

IBM's watsonx.governance—an AI lifecycle toolkit designed to help organizations transform the way they manage AI processes, risks, and control platforms to accelerate responsible, transparent, and explainable AI workflows.

Together, these underpinning services, capabilities, and technologies combine Deloitte's industry insights and deep AI and GenAI experience with IBM's watsonx.governance platform to empower teams to:

- **Establish an AI governance program,** enabled by technology
- **Tailor the platform implementation** based on client/industry/regulatory needs
- **Create an enterprise-wide inventory** of AI models and use cases across the deployment life cycle
- **Establish AI governance workflows**
- **Automate compliance** with preconfigured workflows and questionnaires tailored to client-specific AI policies and regulations

Teams leveraging capabilities provided by operationalized AI and GenAI governance can drive improvements across:

- **Risk management** to proactively address potential issues
- **Policy management** so teams can create and enforce policies that support compliance
- **Transparency** to ensure clear visibility into AI model usage across the enterprise
- **Model lifecycle tracking** for better control and auditability
- **Explainability** for AI model outputs, which fosters trust and understanding
- **Ethical monitoring** using advanced metrics that support responsible AI use

Gain a new perspective on trusting your AI

With more than 25 years of collaboration, deep industry experience, agile technology capabilities, and trustworthy insights, the Deloitte and IBM alliance can open a new window into how your team implements AI governance confidently—all supported by a "single pane of glass" that provides end-to-end AI and GenAI governance.

DELOITTE'S TRUSTWORTHY AI SERVICES

Marketing brand management & eminence

Governance structure & development

Guidelines & policy development

Controls development & assessment

Ethical framework development

Awareness training

AI tool design & requirement gathering

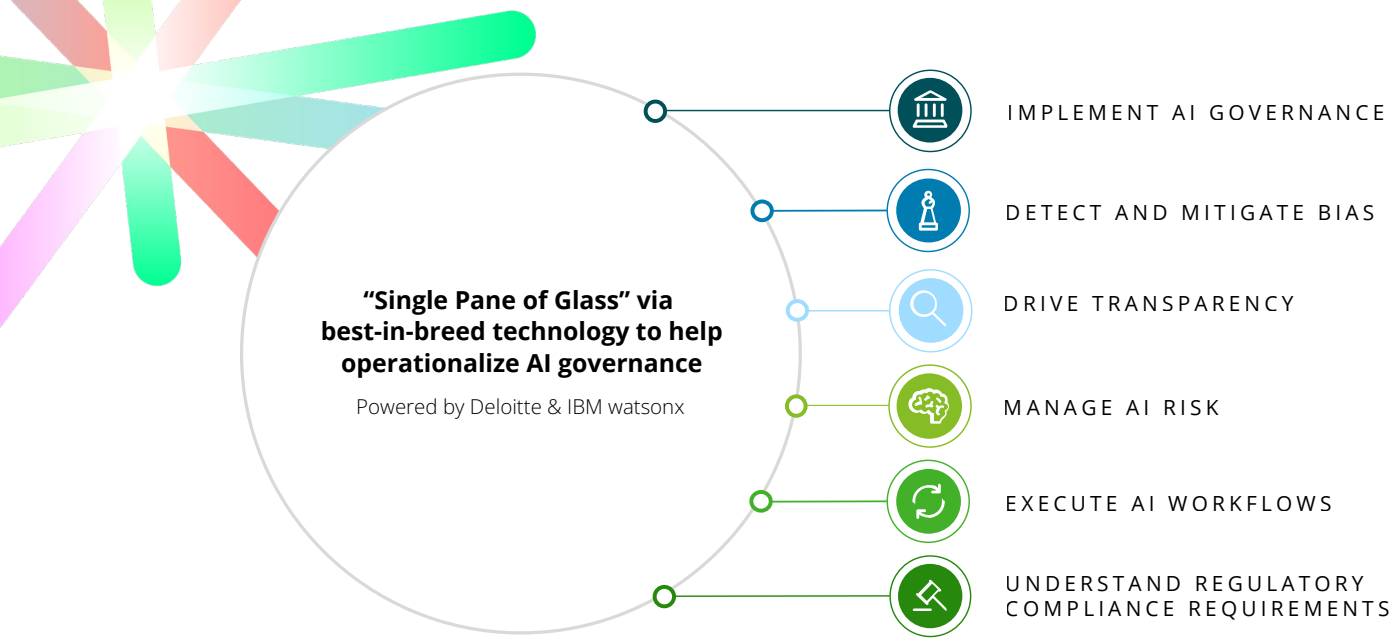
Monitoring & reporting dashboards

Data scrubbing

AI solution development lifecycle assessment

Algorithm risk assessment

Implementation risk assessment



DELOITTE

DEEP INDUSTRY EXPERIENCE
extensive industry knowledge and insights from decades of technology transformations, including tailored IP per industry

AGILE DESIGN & TECHNOLOGY CAPABILITIES
accelerating discovery, design, implementation for AI strategies and tools through tailored, tested approaches, such as Trustworthy AI

LEADING TRUST INSIGHTS & OFFERINGS
risk and controls identification and development, contextualization of trust and its link to business performance

RESPONSIBLE SCALABILITY ACROSS BORDERS
integration with ongoing activities and existing governance, risk, and compliance (GRCO) processes and technology, strength and power of a global network of resources

IBM

HISTORY OF INVESTMENT IN AI
IBM has invested in developing AI tools, including watsonx, and continues to roll-out new capabilities to enhance AI deployment and governance

SOFTWARE DEVELOPMENT EXPERIENCE
providing software that captures the full power of hybrid cloud and AI, increase productivity, reduce costs, and improve business outcomes

ECOSYSTEM OF STRATEGIC PARTNERS
bringing together the world's leading platform and infrastructure partners who add critical value

SECURITY AND PRIVACY BY DESIGN
commitment to embedding security and privacy into the design of products and services, aligned with NIST

INFRASTRUCTURE AGNOSTIC SOLUTION
supporting the deployment of 'AI anywhere' by integrating watsonx into any tech stack (including hybrid, multi-cloud environments)

Getting started

Deloitte offers multiple engagement options to kickstart your AI and GenAI governance journey, including:

- ***Trust assessments*** with a specific focus on tech enablement for AI governance
- ***AI governance workshops*** with interactive simulations to explore processes, approaches, and tools
- ***AI governance operating model design*** to establish an AI ethics board and evaluate AI risk and reward
- ***Reference architecture review and vendor analysis*** to evaluate tools and capabilities
- ***watsonx.governance implementation*** and configuration to operationalize AI governance

Let's discuss your path forward.

Bob Stradtman
Principal
Trustworthy AI Leader
Deloitte Transactions and Business Analytics LLP
rstradtman@deloitte.com
+1 571 882 8790

Derek Snaidauf
Principal
Trustworthy AI Leader
Deloitte Transactions and Business Analytics LLP
dsnaidauf@deloitte.com
+1 312 486 1282

Joe Conti
Managing Director
Trustworthy AI Leader
Deloitte & Touche LLP
joconti@deloitte.com
+1 212 436 7395