C4I

COALITION FOR INTEGRITY

Coalition for Integrity: White Paper Series

May 2026

AI IN CORPORATE INTEGRITY AND COMPLIANCE: USE CASES, GOVERNANCE AND THE ROLE OF HUMAN JUDGMENT

Lead. Learn. Impact.

In collaboration with C4I Supporter

Clifford Chance

The Coalition for Integrity is a non-profit, non-partisan 501(c)(3) organization. We work across business, government, and civil society to strengthen the systems that uphold accountability, transparency, and trust to support the foundations of credible governance and responsible business.

The information contained in this report has been reviewed and assessed using sources believed to be reliable at the time of preparation

In 2021, the **Coalition for Integrity** issued a paper entitled *Using Machine Learning for Anti-Corruption Risk and Compliance*.[1] Its central argument was that anti-corruption and compliance teams were contending with rapidly expanding volumes of enterprise data, and that machine learning had real potential to help surface risk signals and detect patterns, and to allow humans to focus effort where it matters most. While this continues to be true, the use of AI for integrity and anti-corruption work has expanded significantly within many companies and enterprises since 2021, as has the regulatory and standard setting landscape in which practitioners operate. That paper reflected a landscape in which AI use at scale was only feasible for large, well-resourced organizations. That gap is now narrowing as AI is becoming ubiquitous. Vendor-built, configurable tools now make AI-assisted compliance increasingly accessible to mid-sized and small organizations. This paper is written with that broader audience in mind.

C4I's 2021 paper emphasized that before developing or acquiring any machine learning solution for integrity and compliance work, companies first need to articulate the specific challenge they are seeking to address. What exactly is the integrity problem they are trying to solve? Why are existing approaches insufficient? And what business and governance benefits do integrity teams want to achieve? These questions remain the starting point.

The prior paper proposed a structured framework for constructing and transforming data, and then for training, evaluating and ultimately deploying models. This framework continues to be relevant. Importantly, however, the paper emphasized that these technical activities cannot be separated from governance. It noted that the use of machine learning in an integrity and compliance context raises fundamental ethical questions and governance challenges around responsible design and use, lawful purpose, privacy, access to sensitive data, and cybersecurity.

This report explores these dynamics and brings the discussion up to date, providing an overview of the ways in which AI is being used in corporate integrity and compliance activity, as well as describing the main regulatory and standard-setting frameworks currently governing those uses.[2] A wider shift may also be underway as developments in AI reshape the broader risk environment in which integrity and compliance teams operate. There are now questions about whether companies are making sufficient use of available AI tools, how regulatory expectations may evolve, and whether more sophisticated technology-enabled threats are emerging. At the same time, there is increasing convergence around the expectation that AI systems must remain subject to meaningful human oversight, particularly in high-risk or high-impact use cases.

---

1 http://www.coalitionforintegrity.org/wp-content/uploads/2021/04/AI_report_2021_postprinting.pdf.

2 This report is intended as a governance-oriented overview of emerging AI use cases in integrity and compliance and of the principal regulatory and standards frameworks shaping corporate practice as regards AI. It does not constitute legal advice and does not attempt to provide an exhaustive analysis of jurisdiction-specific obligations, which remain fact-dependent and evolving. Organizations should assess their particular situations, regulatory footprint and risk profile in consultation with appropriate legal and technical advisors.

USE OF AI FOR CORPORATE INTEGRITY

How the technology landscape has evolved

In the span of three or four years, AI technology underpinning corporate integrity and compliance has evolved from tools designed to identify risk signals and detect patterns, to systems that can pre-populate templates and standardized formats, and more recently to generative AI capable of interpreting, drafting and summarizing text as well as crafting escalation and approval memos. Most recently, we have been seeing the emergence of agentic AI tools that can execute defined workflows.

Many companies are no longer choosing between which AI technology to adopt. Instead, they are combining multiple technologies within a single solution. As AI systems have become more sophisticated in their capabilities, the importance of governance has only increased. The critical point is that governance controls need to be embedded by design rather than bolted on later.

This principle is reflected in the major frameworks for responsible AI that we will discuss later: NIST's AI Risk Management Framework,[3] ISO's Artificial Intelligence Management System standard,[4] the OECD's intergovernmental recommendations on AI,[5] as well as in the EU AI Act.[6] While the frameworks are different in focus and legal force, they all recognize the foundational principle that as AI systems scale and become autonomous, governance must evolve in parallel.

The technology journey

The earliest large-scale deployments of technology in the integrity space were in **sanctions screening, name matching and transaction monitoring**. Machine learning (ML) and natural language processing (NLP) tools have helped teams process far greater volumes of data than manual review can realistically achieve, drawing from both internal operational data systems and external data sources. However, there are limitations. These tools are highly dependent on data

---

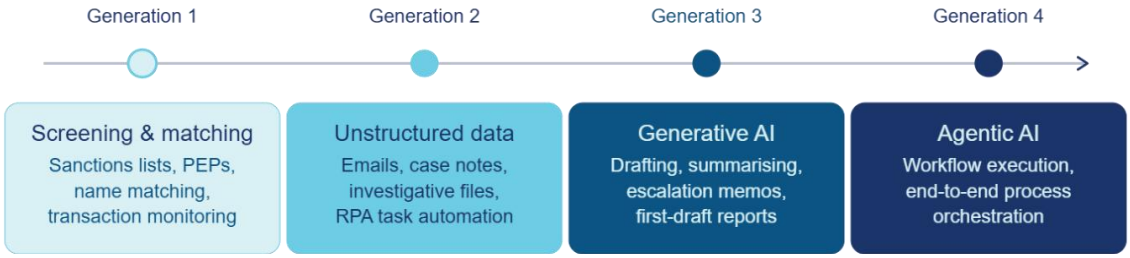
3 National Institute of Standards and Technology (NIST), *Artificial Intelligence Risk Management Framework (AI RMF 1.0),* NIST AI 100-1 (Jan. 2023), available at https://www.nist.gov/itl/ai-risk-management-framework. Companion guidance NIST, *AI Risk Management Framework: Playbook*, NIST AI 100-2 (2023, updated periodically), available at https://www.nist.gov/itl/ai-risk-management-framework/ai-rmf-playbook.

4 International Organization for Standardization / International Electrotechnical Commission, *ISO/IEC 42001:2023, Information technology — Artificial intelligence — Management system — Requirements (2023),* available at https://www.iso.org/standard/42001.

5 Organisation for Economic Co-operation and Development, *Recommendation of the Council on Artificial Intelligence (2019; updated 2024),* available at https://oecd.ai/en/ai-principles.

6 European Union, *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), Official Journal of the European Union, OJ L, 12 July 2024*, available at https://eur-lex.europa.eu.

quality. They can also generate significant noise and tend to struggle when data is incomplete and fragmented.



This paper describes these dynamics as a sequence, not a fixed chronology, where each generation layers on rather than replacing the last.

Next came tools capable of working with **unstructured data**. These systems can analyze emails, case notes and investigative files, which has resulted in much broader coverage as well as faster identification of relevant issues across datasets. Robotic process automation (RPA) technologies also began to **reduce repetitive administrative tasks** such as moving data between integrity systems, pre-populating standard integrity forms and questionnaires, and assembling investment committee or audit review packs to support integrity-related decision-making.

The more significant shift over the last three years has been the rise of generative AI (Gen AI). Large language models (LLMs) can **draft summaries, structure lists, propose follow-up questions, and produce first drafts of escalation and approval memos**. That said, LLMs generate responses by predicting language patterns. Generally speaking, they cannot independently verify whether the responses are accurate. Without retrieval and citation features, their responses may also lack defensibility, which is a critical weakness in the integrity and compliance context. Even with retrieval, the outputs may still require human verification.

The most recent development is the arrival of agentic AI systems (Agentic AI) capable of **executing elements of defined workflows**. These AI systems can retrieve records, assemble supporting materials, open or update cases, and even route items for review. They can also support and coordinate processes across workflows. Such systems need to operate within carefully constructed parameters and boundaries defined by the company. Yes, they can perform defined activities that previously required significant manual effort, but they should not make high-impact judgment calls on sensitive or high-impact matters. The key point is that agentic AI systems need to be carefully calibrated and monitored by humans.

Accountability expectations

Over time, the distinctions between these different AI technologies and systems are blurring. Increasingly, companies and third-party vendors are integrating multiple forms of AI within a

single solution and separate technologies are becoming layered. From a governance perspective, this means more complexity. Critically it means that companies must ensure that they are explicit about accountability and have robust governance control structures in place with the right human oversight at the right handoff points.

> "AI can dramatically extend the reach of an ethics and compliance program, but only if humans stay firmly in control of the decisions it informs."
> *Justin Ross, Vice President & Chief Compliance Officer,* Sysco

To meet accountability expectations, companies must clearly define:

- who is responsible for AI-supported systems and activities
- what approvals are required and from whom
- what thresholds should trigger escalation and handoffs and when
- how decisions are reviewed, and
- how both recommendations and underlying evidence are recorded and audited.

In thinking about accountability, it is helpful to have regard to the following anchor points:

1. **AI should not replace human judgment on high-risk matters:** While AI's potential is real and can transform integrity and compliance functions, it should not replace human judgment on high-impact or sensitive decision-making. Companies must carefully design parameters to ensure that high-risk judgment calls are made by humans so that decisions can have regard to context, nuance and the company's risk appetite. This is particularly important to the work of integrity and compliance teams as they help companies navigate reputational exposure, regulatory scrutiny, and enforcement risk.

2. **Integration and configuration:** Companies do not need to build these AI systems from scratch. Sanctions and screening platforms, third-party risk tools, adverse media systems, case-management solutions, governance risk and compliance (GRC) platforms, and, increasingly, agentic workflow systems are now widely available in the market. Generally speaking, the challenge for a company is therefore not about designing their own AI tools and systems but about carefully configuring vendor tools so that they align with the company's decision-making hierarchies and risk appetite, and so that they embed human-in-the-loop controls at the correct points and integrate across company workflows.

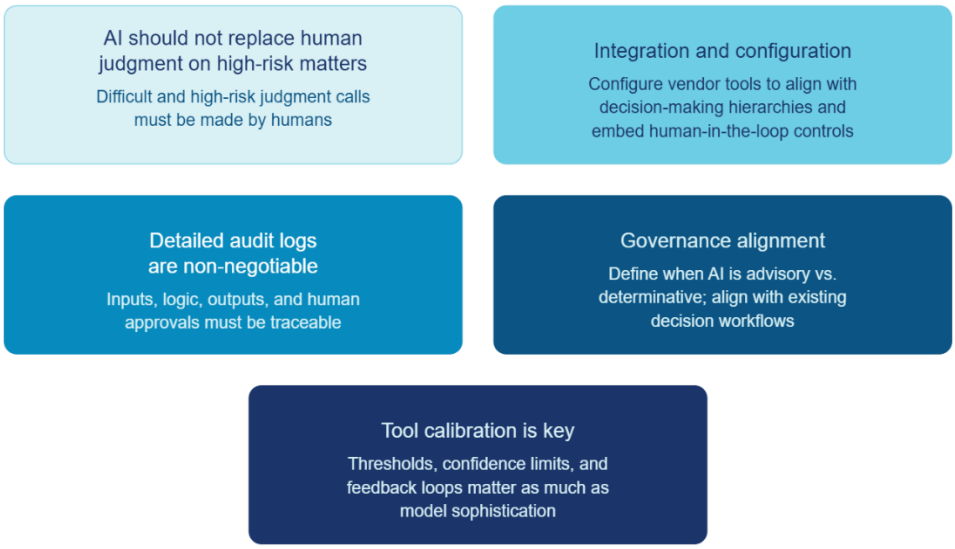
3. **Detailed audit logs are non-negotiable:** Auditability and traceability should be built into system design, not added later. AI systems must preserve a clear record of the inputs that have been used, the logic and reasoning that have been applied, the outputs produced, and the points at which humans have intervened to authorize and approve outcomes. This means that integrity risk signals, alerts, prioritizations, escalations, recommendations, and reports all need to be supported by traceable source citations, data lineage, records of human review, timestamps, and documented rationales. Some companies are actively engaged in the design

and development of audit logs, which are frequently more detailed and expansive than standard vendor logs.[7]

4. **Governance alignment:** In defining governance parameters, companies should seek to distinguish between AI systems designed to enhance productivity and AI systems that aim to augment judgment and decision-making. Relatively low-risk productivity-enhancing activities may not require extensive governance and oversight. Indeed, this risks blurring low-risk use cases with higher-stakes activities where AI systems need to align carefully with existing decision workflows in a way that can withstand audit review, litigation, or regulatory inquiry.

> "Not all AI carries the same risk. Productivity tools can operate with relatively light oversight, but systems that influence judgment and decision-making require much stronger governance to ensure they are reliable, free from undue bias, and are used in ways that align with company standards and risk tolerance."
> *Jeffrey Eglash, experienced chief business integrity officer*

5. **Tool calibration is key:** Thresholds, confidence limits and feedback loops increasingly matter. Risk-tiering, periodic recalibration, and structured human overrides are all essential to managing false positives, false negatives, missed exposures and data gaps – and this is another area in which companies are increasingly devoting time and attention and where human inputs are critical.



[7] Note that audit logs must also balance data protection, confidentiality and legal privilege, particularly where log material could be discoverable. Detailed system logging may increase exposure in litigation or regulatory discovery, and companies should align AI logging practices with legal hold policies and privilege frameworks.

Core use cases

Many of the AI capabilities referred to above are already being embedded into corporate integrity and compliance functions. Each of the following use cases maps to one or more established elements of an effective anti-corruption and compliance program, from risk-based due diligence to monitoring, auditing and documentation. These are areas in which regulators and enforcement authorities have long expected companies to maintain appropriate controls and oversight, even though the sophistication and scale of those measures will vary depending on the size, risk profile, and resources of the company.

- **Risk-based screening and watchlist matching.** For several years now, this has included sanctions and enforcement list screening, politically exposed person (PEP) identification and the handling of transliteration and aliases. More recent enhancements include calibrating AI systems by risk tier (e.g., more sensitive thresholds for high-risk jurisdictions) and refining the systems through human analyst feedback loops.

- **Adverse media triage**. AI has also been used for several years now to scan large volumes of multi-language media to identify relevant negative reporting. AI can now collate findings thematically and by materiality. Another enhancement has been to focus the triage on identifying only new developments to remove duplication and recycled content. More sophisticated implementations involve sentiment analytics and scoring.

- **Third-party and counterparty due diligence**,[8] including on vendors, suppliers, distributors, joint-venture partners and agents. This is a more recent development with AI being used to analyze questionnaires and filings, and surface missing documentation; it can update risk scores as new information becomes available. Some companies have also started to use AI systems to pre-populate forms, track and follow up on attestations and documentation gaps, and to assemble document packs for escalation or investment decisions.

- **Payments and procurement analytics.** AI tools are being used by some corporate compliance teams to detect anomalies that may signal fraud, bribery or collusion. This can include anomaly detection (e.g., invoice splitting, round-dollar payments analysis), as well as network analytics to flag suspicious patterns and bid-rigging indicators (e.g., recurring intermediaries, repeat subcontractors, and shared ownership structures).

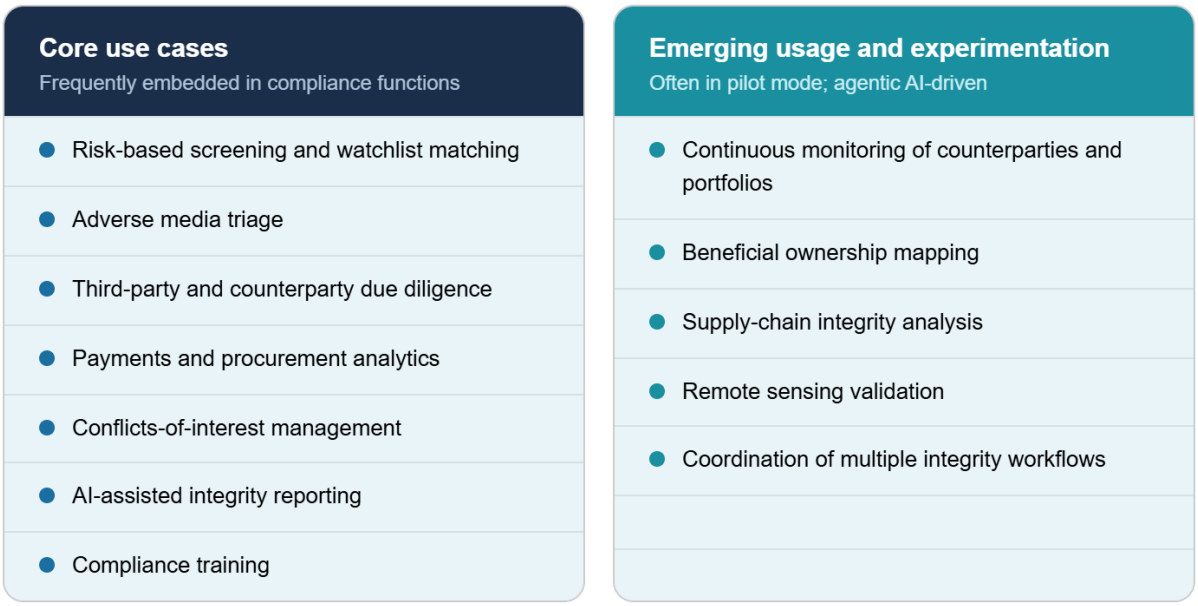
- **Conflicts-of-interest management**. AI is increasingly used to support and manage corporate declaration-of-interest and similar programs, while graph-based techniques are starting to be applied by some compliance teams to identify undisclosed links between employees, vendors and counterparties.

---

[8] Enhanced due diligence on third parties is a well-established expectation in FCPA enforcement guidance, OECD anti-bribery recommendations, and the UK Bribery Act's adequate procedures defense.

- **AI-assisted integrity reporting**, including the drafting of due-diligence summaries, investigation notes or monitoring reports, can improve consistency and reduce drafting churn, but with mandatory citations and human approval requirements built in.

- **Compliance training.** AI tools are being used by compliance teams to convert scripts, standing operating procedures (SOPs), or policy text into avatar-led training videos. These can be quickly revised, updated when rules change, and rolled out globally in multiple languages to help ensure consistent messaging across a distributed workforce. As a result, in some companies AI-generated video is now being used not only for annual compliance training, but also for just-in-time compliance reminders, leadership messages, and change-management communications.



Emerging usage and experimentation

A second group of use cases is starting to emerge in the integrity space, often in pilot mode and particularly as companies start thinking about the capabilities of agentic AI workflows. Examples include:

- **Continuous monitoring of counterparties and portfolios.** AI is starting to allow companies to address emerging integrity signals as they arise, rather than refreshing due diligences on fixed one-year, two-year or three-year review cycles. However, to sound a cautionary note, continuous monitoring can also result in new blind spots, such as over-reliance on the outputs from AI systems.

- **Beneficial ownership mapping**. AI is starting to offer more advanced mapping tools such as sophisticated graph analytics to identify beneficial ownership (e.g., repeated patterns across jurisdictions).
- **Supply-chain integrity analysis**. AI is increasingly being used to map multi-tier supply chains, as well as touchpoints with sanctioned jurisdictions, restricted sectors or territories, and high-risk intermediaries. These tools are also increasingly integrating shipping, registry, customs and contractual data.
- **Remote sensing validation.** Satellite imagery or drone data can assist in substantiating the existence of assets, site activity, or environmental compliance, particularly in sectors such as infrastructure, extractives and agriculture. These tools can be particularly useful in fragile markets where security concerns may prevent the collection of on-the-ground intelligence.
- **Coordination of multiple integrity workflows**. The newest emerging use cases involve using agentic AI to help orchestrate discrete repetitive tasks that were previously manual, such as retrieving documents, updating cases, assembling evidence bundles, preparing escalation memos, and routing those memos to the right people for human review. Critically, however, these agentic AI systems require clear guardrails and configuration rules to ensure that they do not unintentionally automate high-impact judgment calls.

STANDARD SETTING FRAMEWORKS

The governance landscape: main components

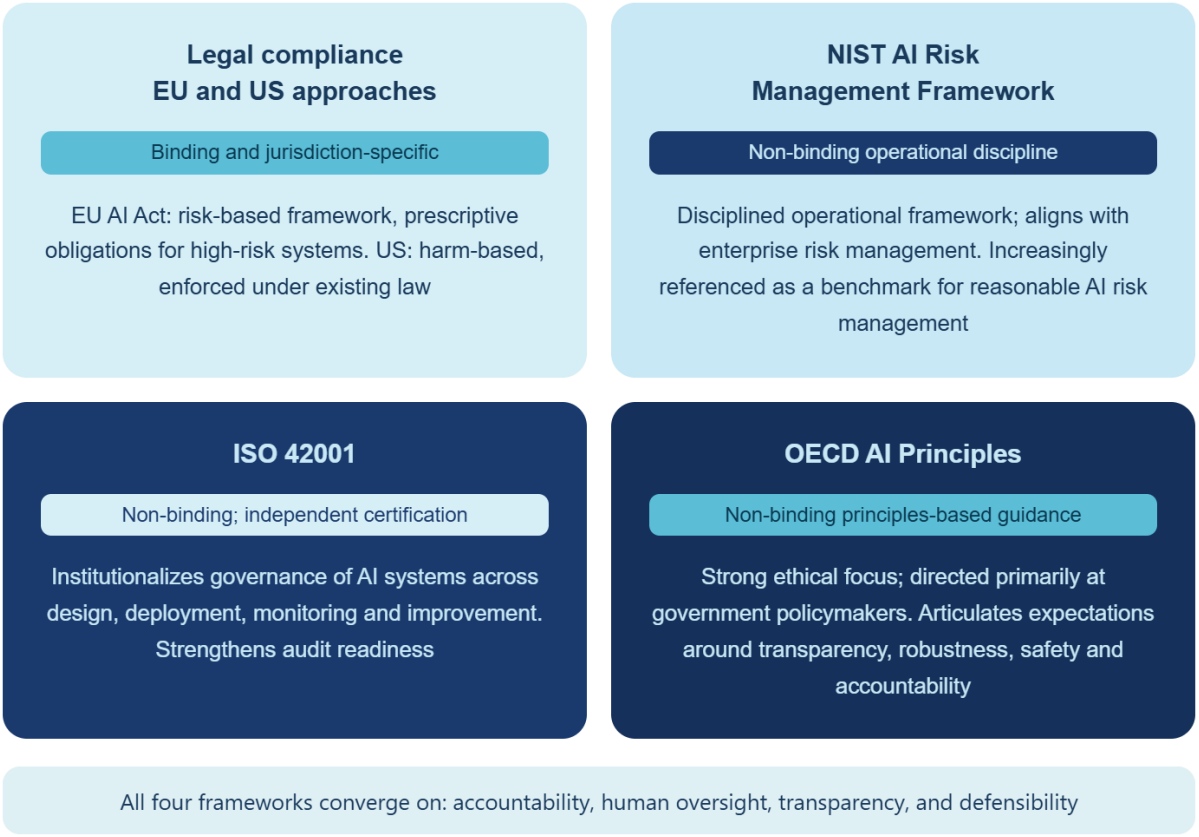
As noted already, the technical sophistication of AI models increases the need for strong governance discipline. Currently, AI governance is layered across four separate but interlocking components:

- **Legal compliance** (e.g. EU, US or other jurisdiction) - binding legal obligations.
- **NIST AI Risk Management Framework -** non-binding standards focused on promoting operational discipline.
- **ISO 42001 -** non-binding assurance that governance is embedded institutionally.
- **OECD -** normative principles for the responsible use of AI.

> "We've seen significant governance challenges before. While AI is different in scale it is not different in principle. The key is embedding it into existing enterprise risk frameworks from the outset and not as an afterthought, with clear accountability and real cross-functional ownership."
> Ceri Lawley, Former Chief Compliance Officer, International Finance Corporation

These components are continually evolving. For example, the OECD is shifting its focus from high-level principles towards implementation-related guidance, including approaches to risk

classification and expectations around incident reporting.[9] COSO has also issued specific guidance in early 2026 on applying its Internal Control–Integrated Framework to generative AI ("Achieving Effective Internal Control Over Generative AI").[10]



Looking at the four main governance models in turn:

1. Legal compliance: EU and US approaches

Legal compliance is both **binding and jurisdiction-specific**. Companies need to identify where they operate, what AI systems they have in each jurisdiction, and how these systems interact

---

9 OECD, *Towards a Common Reporting Framework for AI Incidents* (28 Feb. 2025), available at https://oecd.ai/en/ai-publications/towards-a-common-reporting-framework-for-ai-incidents. *See also* OECD AI Incidents Monitor, available at https://oecd.ai/en/incidents.

10 COSO, *Achieving Effective Internal Control Over Generative AI* (23 Feb. 2026), available at https://www.coso.org/generative-ai; This voluntary guidance maps COSO's five established internal control components onto AI-specific governance practices: Control Environment, Risk Assessment, Control Activities, Information and Communication, and Monitoring Activities. In doing so, it provides organizations with a practical roadmap for managing generative AI risks within their existing enterprise risk management and internal control structures.

with regulatory obligations. In practice, this means mapping AI use cases against applicable regulatory regimes to determine what the relevant requirements are and what regulators expect.

Focusing principally on the EU and the US:

The **EU AI Act** is the primary horizontal AI regulation in the EU. It establishes a risk-based framework for AI systems applicable across member states, categorizing AI systems into prohibited uses, high-risk systems and systems subject to transparency requirements, with differing requirements for each category. The Act entered into force on 1 August 2024, and its obligations are phasing in over time. First, the Act's prohibited practices (e.g. social scoring systems and manipulative AI) have applied since 2 February 2025. Second, the core obligations for high-risk AI systems are scheduled to become fully enforceable from August 2026, although this timeline may be extended through further legislative amendments.[11] As of the end of 2025, no public enforcement actions or fines under the Act's new provisions have been reported.

Many larger companies are currently assessing whether particular AI use cases fall within the Act's high-risk categories. This requires a careful and fact-specific analysis, and for AI systems designated as high-risk, the Act imposes prescriptive obligations around risk-management, documentation and data-governance. In addition, the Act establishes a separate framework for general-purpose AI (GPAI) models, including large language and foundation models, under Articles 53 to 55.[12] Providers of GPAI models must meet baseline obligations relating to model safety, technical documentation and transparency, including specific disclosure requirements for generative AI outputs. Importantly, GPAI models are not automatically classified within the Act's risk tiers; instead, the EU has adopted a voluntary Code of Practice for GPAI providers to support compliance and responsible development of these widely-used models.[13]

However, where a GPAI model is subsequently integrated into a high-risk application (e.g. credit scoring or recruitment), the deployer or integrator assumes the role of provider of the high-risk system and must ensure full compliance with the Act's high-risk AI requirements.

The EU Act is often described as an **ex-ante approach**, meaning that a key focus is to regulate AI system design before any harm occurs. In practice, this means that providers of high-risk AI systems must, prior to placing a system on the market, implement a documented risk management system and complete a conformity assessment (or, where applicable, a self-certification), demonstrating that the system meets all applicable requirements. However, the

---

11 Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (the "EU AI Act"), Arts. 113 and 185 (Official Journal of the European Union, L series, 12 July 2024).

12 EU AI Act, Arts. 53–55 (obligations for providers of general-purpose AI models).

13 European Commission / European AI Office, The General-Purpose AI Code of Practice (published 10 July 2025), available at https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai. The Commission and the AI Board confirmed the Code as an adequate voluntary tool to demonstrate compliance with the AI Act

Act's obligations do not end at market placement. Providers must also perform ongoing post-market monitoring of the AI system, maintain detailed performance logs, and report serious incidents or malfunctions to the relevant national authorities.[14] Compliance under the Act is therefore a continuous lifecycle obligation rather than a one-time certification exercise.

The Act also interacts closely with other legal regimes, notably GDPR and EU sector-specific financial regulation, consumer law and product safety frameworks.

In contrast, the **US has no single AI statute equivalent to the EU AI Act**. Instead, AI systems are governed through existing federal laws, including consumer protection law, anti-discrimination law, deceptive or unfair practices standards, and sectoral regulation in heavily regulated sectors such as financial services and healthcare, as well as through state-level legislation. In practice, U.S. federal oversight of AI remains distributed across multiple agencies and is shaped by evolving enforcement and administrative policy priorities. Rather than operating under a single federal AI statute, enforcement relies on a patchwork of preexisting authorities exercised by agencies such as the FTC, DOJ, EEOC, and, in the financial-services context, the CFPB. These agencies address AI-related misconduct under existing consumer-protection, civil-rights, employment, and sector-specific laws, using primarily an ex-post, harm-based enforcement model rather than a coherent pre-market approval or conformity-assessment system. The result is a multi-layered compliance environment in which legal defensibility turns on how existing laws are interpreted and enforced, often against a backdrop of shifting administrative priorities. This stands in contrast to the EU's more formalized ex-ante regime.

The US landscape is also influenced by Presidential Executive Orders. In January 2025, the Trump Administration signaled a shift in federal AI policy priorities from the Biden October 2023 executive order on AI governance[15] and in July 2025 introduced America's AI Action Plan.[16] This focuses on accelerating infrastructure, federal procurement processes, international competitiveness and deregulation rather than on specific federal-level AI regulation. In December 2025, the White House issued a further executive order, *Ensuring a National Policy Framework for Artificial Intelligence*, which made federal preemption a central theme by arguing

---

[14] EU AI Act, Arts. 9 (risk management system), 43 (conformity assessment procedures), 72 (post-market monitoring) and 73 (reporting of serious incidents).

[15] Executive Order 14110, *Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence* (Oct. 30, 2023), available at https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence; Executive Order 14179, *Removing Barriers to American Leadership in Artificial Intelligence* (Jan. 23, 2025), available at https://www.whitehouse.gov/presidential-actions/2025/01/removing-barriers-to-american-leadership-in-artificial-intelligence/.

[16] The White House, *America's AI Action Plan* (July 2025), available at www.whitehouse.gov/wp-content/uploads/2025/07/Americas-AI-Action-Plan.pdf

that a patchwork of state AI laws threatens innovation and interstate commerce and by calling for a minimally burdensome national standard that would displace conflicting state requirements.[17]

In March 2026, the Administration built on that position by unveiling a National AI Legislative Framework, which urged Congress to adopt a comprehensive federal framework across issues such as child protection, infrastructure, intellectual property and free speech, while also clarifying selected areas where generally applicable state laws might continue to operate.[18]

These developments sharpen the federalism question in U.S. AI governance: even without a single enacted federal AI statute, the policy direction of travel is currently toward stronger federal primacy and possible preemption of inconsistent state rules, which complicates the legislative landscape for companies operating across jurisdictions. It is important to note, however, that executive action does not displace existing statutory enforcement and federal agencies continue to exercise authority under relevant federal laws. In many sectors, regulators and enforcement agencies continue to expect organizations to maintain defensible governance and risk-management controls even in the absence of a single federal AI statute.

Meanwhile, state-level AI regulation remains active but fluid. States such as Colorado, California, Utah, Illinois and New York have enacted or are developing AI-specific or AI-adjacent laws, though the content and effective dates of these frameworks continue to evolve. In the US, companies therefore need to be focused on aligning AI systems with existing federal laws, while at the same time monitoring evolving state-level laws that impose substantive AI governance requirements separately from federal frameworks.

The **territorial reach of both the EU and US regimes is broad**. For example, the EU AI Act applies not just to companies established within the EU but to multinationals headquartered outside the EU that offer AI-enabled products or services to EU customers or whose AI systems affect individuals or entities within the EU. This design mirrors the EU's broader regulatory approach in areas such as privacy and data protection. Likewise on the US side, companies may face AI-related exposure under US federal and state laws based on where they conduct business or where affected individuals reside, rather than where a company is headquartered.

---

17 Executive Order, *Ensuring a National Policy Framework for Artificial Intelligence* (Dec. 11, 2025), available at www.whitehouse.gov/presidential-actions/2025/12/eliminating-state-law-obstruction-of-national-artifical-intelligence-policy.

18 The White House, *President Donald J. Trump Unveils National AI Legislative Framework* (Mar. 20, 2026), available at https://www.whitehouse.gov/releases/2026/03/president-donald-j-trump-unveils-national-ai-legislative-framework; The White House, *A National Policy Framework for Artificial Intelligence: Legislative Recommendations* (Mar. 2026), available at https://www.whitehouse.gov/wp-content/uploads/2026/03/03.20.26-National-Policy-Framework-for-Artificial-Intelligence-Legislative-Recommendations.pdf.

DIMENSION	EU	US
Core structure	Single AI-focused regulation	No single AI statute. Overlapping layer of existing federal and state laws
Regulatory requirements	Risk-tier classification with prescriptive obligations, including documentation and data governance	Harm-based enforcement under existing laws; defensibility-driven
Enforcement trigger	System classification under statute; ex ante approach regulating design before harm occurs	Legal exposure under consumer, civil rights, and sector-specific laws
Governance focus	Formal compliance with statutory classifications, documentation and conformity assessment obligations	Demonstrable and defensible fairness, oversight, and risk mitigation

2. NIST: Operationalizing AI risk management

NIST's AI Risk Management Framework is not law so does not create legal obligations. Instead, it is a **disciplined operational framework** that can be aligned with a company's enterprise risk management architecture and mapped to the company's internal policies, procedures, controls, performance metrics, and audit requirements. The Framework is particularly relevant in the integrity and compliance context because it focuses on operational issues that boards, management teams, regulators, and other stakeholders increasingly expect companies to address (e.g., governance structures, human oversight, documentation and explainability, risk mapping, testing, monitoring, and measuring how performance drift or emerging exposures are detected and corrected).

While not legally binding, the Framework is increasingly being used as a practical benchmark for operationalizing responsible AI governance and risk management practices. Its primary adopters were US companies with strong enterprise risk management frameworks, as well as federal contractors. Adopters increasingly include multinational companies seeking a common operational framework across jurisdictions, and companies seeking credible governance approaches that can evolve alongside rapidly developing regulatory expectations.

3. ISO 42001: Institutionalizing and proving governance

Like the NIST AI Framework, ISO 42001 is not a legal requirement. Rather, it is a management system standard that **institutionalizes the governance of AI systems** across design, deployment, monitoring and improvement. It is often described as strengthening audit readiness.

Importantly, it also provides for independent certification – which allows boards, regulators and stakeholders to obtain assurance that a company's AI governance would be capable of surviving events such as staff changes, vendor changes and AI model updates.

Adoption of ISO 42001 is in the early stages in many sectors but is increasing among regulated institutions seeking external assurance or certification. This includes banks, insurers, pensions funds, and other institutions selling into regulated markets where there are strong external audit expectations. Multinationals often adopt ISO 42001 over the top of a NIST or similar framework, particularly where they hold other ISO certifications (e.g., ISO 27001 for cybersecurity risk management or ISO 9001 for quality management systems), or where they see certification as offering strategic or investor value add. Other adopters include development finance institutions (DFIs), as well as non-profits and NGOs.

4. OECD AI Principles

The OECD AI Principles are primarily directed at government policy makers. They are **non-binding principles-based guidance with a strong ethical focus**, articulating the expectations that public stakeholders increasingly apply when judging AI trustworthiness (e.g., respect for human rights, transparency, robustness, safety and accountability). The Principles have been endorsed by over 40 countries, including the United States and all EU member states, and were subsequently adopted by the G20, giving them significant normative weight beyond the OECD membership.[19] The main adopters are policymakers in OECD member states that use the AI Principles to inform national AI strategies, as well as international organizations such as the World Bank and UN.

Convergence on foundational standards

While the above frameworks have materially different operational approaches, enforcement mechanisms and liability exposure, they nonetheless converge around a common set of governance themes and expectations. These convergence points can be traced to specific provisions across the frameworks, including the NIST AI RMF's core functions (Govern, Map, Measure, Manage),[20] ISO 42001's management system clauses,[21] the EU AI Act's high-risk

---

19 OECD, Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449 (adopted 22 May 2019, amended 3 May 2024), available at https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449; G20 Ministerial Statement on Trade and Digital Economy, Tsukuba, Japan (June 2019), endorsing the OECD AI Principles.

20 National Institute of Standards and Technology (NIST), Artificial Intelligence Risk Management Framework (AI RMF 1.0), NIST AI 100-1 (26 January 2023), available at https://www.nist.gov/artificial-intelligence.

21 International Organization for Standardization, ISO/IEC 42001:2023, Information technology — Artificial intelligence — Management system (2023), available at https://www.iso.org.

obligations,[22] the OECD AI Principles,[23] and COSO's internal control components as applied to generative AI.[24] For convenience, the common themes can be grouped as follows:

1. **Governance, accountability and human control**: All the frameworks point to the need for a clearly defined AI governance structure, including designated senior ownership, delegation of authority for threshold setting, escalation and oversight, and the need to be explicit about roles. Related to this, oversight requires human reviewers and decision-makers to understand system capabilities and limitations. Importantly from a board perspective, there needs to be confidence that high-impact and sensitive decisions are not unintentionally being automated and that oversight mechanisms are tested periodically.

2. **Lawful and ethical deployment**: AI systems need to operate within clearly defined boundaries, with clear articulation of what they are designed to do and not do. First and foremost, they must align with the applicable legal regimes. Additionally, however, ethical standards reinforce these boundaries by emphasizing proportionality and in varying degrees respect for individual rights.

3. **Transparency, explainability and defensibility**: All the frameworks converge around the principle that transparency and explainability need to be built into AI systems by design. AI systems need to preserve detailed logs so that companies can reconstruct and defend the decision pathway on specific matters if requested to do so by decision-makers, auditors, regulators or counterparties.

4. **Fairness, bias management and responsible data stewardship**: The frameworks also converge around the principle of fairness, and the requirement for active bias monitoring, including reviews of uneven and unequal impacts. This intersects directly with data stewardship when datasets involve sensitive personal information. To meet stewardship expectations, there is also convergence around the need for robust access controls, secure handling of data, and defined retention policies, together with clear oversight of how vendors and service providers handle their data.

5. **Continuous monitoring and improvement**: A final recurring theme is that AI system performance must be measured and recalibrated on a continuous basis with a view to improving output over time. This translates into the need for stress testing, as well as protocols for addressing non-performance and drift. There also needs to be periodic

---

22 EU AI Act, Title III (High-Risk AI Systems), including Arts. 6–49.

23 OECD, Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449 (adopted 22 May 2019, amended 3 May 2024), available at https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449.

24 COSO, *Achieving Effective Internal Control Over Generative AI* (2026), available at https://www.coso.org/generative-ai.

reporting to ensure that AI systems and related processes evolve dynamically in the face of shifts in underlying data sources.

Concluding Remarks

The trajectory of AI in integrity and compliance functions is clear, both in terms of increasing technical capability and in how these capabilities are being integrated into day-to-day workflows.

Several years ago, AI focused on screening and adverse media triage, but this is now evolving into AI systems capable of working across structured and unstructured data, that support sophisticated pattern recognition and signal detection, and that are increasingly capable of orchestrating defined components of end-to-end workflows in the agentic AI space.

Increasingly, risk signals can be surfaced and assessed in near real time. However, AI is not just an accelerator. It is also an augmenter, surfacing signals that human reviewers cannot reasonably be expected to find, which materially enhances both coverage and consistency. As a result, AI is changing the nature of integrity and compliance work, with teams doing less repetitive data gathering and assessment and having more time to focus on interpretation, validation and decision-making.

The most pressing question facing boards and their integrity and compliance teams is how to integrate these capabilities into their company's workflows in a way that preserves human judgment, accountability and defensibility. The second pressing question is what thresholds and confidence limits are appropriate, when and how human handoffs should happen, and how boards and other stakeholders can obtain assurance that systems are operating responsibly.

The governance frameworks that companies should have regard to in addressing AI adoption differ in legal force, structure and detail, as well as in jurisdictional reach and scope of application. The EU AI Act creates binding and prescriptive statutory obligations in defined contexts. US law governs AI in a multi-layered manner through existing sectoral regulation and harm-based doctrines. The NIST AI Framework is voluntary and provides operational risk discipline. ISO 42001 offers management-system assurance through a credible industry-level certification. And the OECD AI Principles articulate norms focused on government policymakers rather than enforceable rules.

Despite structural differences, the frameworks all converge around core foundational principles and expectations. These frameworks converge around the need for clear governance and accountability, human oversight, transparency and defensibility, ethical deployment including bias monitoring, robust disciplined data stewardship, and continuous monitoring and recalibration. These principles provide a foundation, but they do not resolve the broader questions that are now emerging.

Looking ahead, integrity and compliance functions are likely to face increasing pressure not only in how they use AI, but also whether they are making sufficient use of available tools to identify

and manage risk effectively, particularly in the context of ongoing cost and efficiency constraints. This may be particularly relevant for companies whose compliance programs are at an earlier stage of development or which operate with smaller compliance teams. For these companies, when deployed responsibly and with appropriate oversight, AI tools may help extend coverage, improve consistency, and reduce manual burden, allowing limited compliance resources to be directed toward higher-value review, investigation, and decision-making activities.

While regulators have to date focused primarily on the responsible use of AI, expectations may evolve over time that more advanced technologies can prevent or mitigate some compliance failures. At the same time, the broader risk environment is increasingly being shaped by technology-enabled threats, suggesting that organizations will need to adopt more sophisticated approaches to risk identification and monitoring. In parallel, the composition of compliance functions is evolving, with increasing demand for data, analytics, technology and cyber expertise alongside traditional compliance and integrity skills.

Taken together, these developments point to a more integrated, technology-driven model for integrity and compliance teams, but critically, one in which human judgment remains central.

Schedule: Market / Practice Insight

The following sources are illustrative of current market practice, adoption trends and emerging use cases in artificial intelligence relevant to integrity and compliance functions.

Adoption and Enterprise Use

- McKinsey & Company
  The State of AI (latest edition)
  https://www.mckinsey.com
- IBM
  Global AI Adoption Index (latest edition)
  https://www.ibm.com
- Microsoft and LinkedIn
  WorkLab, Work Trend Index, and AI at Work
  https://www.microsoft.com/worklab/work-trend-index

Technology Trends and Generative/Agentic Capabilities

- Gartner, Inc.
  Top Strategic Technology Trends (latest edition)
  https://www.gartner.com/en/information-technology/insights/top-technology-trends
- Accenture
  Reinventing Together with Generative AI (and related research)
  https://www.accenture.com

Data Quality, Performance and Limitations

- Massachusetts Institute of Technology (MIT) Sloan Management Review
  Research and articles on AI, data quality and model performance
  https://mitsloan.mit.edu/ideas-made-to-matter
- Stanford Institute for Human-Centered Artificial Intelligence
  AI Index Report (latest edition)
  https://aiindex.stanford.edu

Governance and Responsible AI Practice

- World Economic Forum
  AI Governance Alliance outputs and agentic AI discussions
  https://www.weforum.org

ACKNOWLEDGEMENT

This report was prepared and published by C4I. Lead authorship was provided by Ceri Lawley, Former Chief Compliance Officer, International Finance Corporation and C4I Board Member, in collaboration with Clifford Chance, with editorial support from Amy Selzer, President and CEO, C4I.

We would like to recognize and thank the following individuals for their support and insights:

- William Lanier, Associate, Brian Yin, Associate, and Peter Isajiw, Partner, Clifford Chance
- Lucinda Low, C4I Board Chair
- Jeff Eglash, C4I Board Member
- Justin Ross, Vice President & Chief Compliance Officer, Sysco, C4I Board Member
- Jonathan Ashtor, Co-Chair Global AI Group, Paul Weiss

While Coalition for Integrity benefited greatly from the work and advice provided by the foregoing persons, this report, including its conclusions and recommendations, represents the views of the Coalition for Integrity and does not necessarily reflect the views of those that provided advice, time, and services to the report.