COMPLIANCE CONSIDERATIONS FOR EVALUATING AI GOVERNANCE IN DCGA PLATFORMS

THETALAKE

CONFIDENTIAL

SUBJECT

This brief sets out the perspectives financial services firms need to consider when evaluating AI governance in the context of digital communications governance and archiving (DCGA) platforms.

It covers two key aspects of AI governance for firms to evaluate and how these can be used in practice. Firstly it describes the features of DCGA platforms that provide firms with transparency into a vendor's AI models. Secondly it explains how DCGA vendors maintain oversight and governance of AI models and systems.

The brief also outlines the applicable standards and regulations in the US, UK and Canada and provides a set of recommendations that firms can use in their evaluation of platforms and vendors.

WHAT AI GOVERNANCE IS AND WHY IT IS REQUIRED

AI governance relates to the management, oversight and explainability of how AI systems function. For financial services firms, AI governance is a core compliance expectation, requiring institutions to be able to explain how AI-driven decisions are reached. Therefore, the deployment of DCGA platforms for supervision is predicated on the ability to provide these explanations to regulators and auditors. This transparency is therefore a critical component of firms' procurement process when assessing third-party vendors.

Most DCGA platforms now use AI features for the purposes of assisting in the analysis, summarization, search, and other aspects of oversight and management. As a result, some DCGA platforms include features that can explain the rationale behind AI decision-making. These features enable users to understand the specific logic that triggered an AI detection and provide a defensible method for firms to justify decisions and subsequent actions to regulators and auditors.

In addition, DCGA platforms should provide firms with transparency into their development and management of AI features throughout their entire lifecycle. This includes the controls to manage how AI models are built, trained, and overseen as well as assurance over model drift and performance. These governance practices allow firms to assess the level of oversight and assurance a vendor provides.

KEY AI GOVERNANCE CAPABILITIES, OVERSIGHT & CERTIFICATIONS

Explainability of Integrated AI Features

AI is leveraged in DCGA platforms to detect risks across multiple modalities, including voice, video, chat, email, and AI interactions. Explainability features within the platforms provide firms with the defensibility and comfort needed to understand exactly why a model identified specific content as a risk. These capabilities include:

Summarizing AI Decisions for Explainability
Where an AI classifier detects a potential compliance, data privacy, or security risk, the AI governance features explain the rationale behind the AI's decision-making. A detection annotation feature provides a plain-language rationale for the triggering of a particular AI risk detection directly within the platform, even when detections span multiple channels or platforms. For example, if a conversation triggers a collusion detection, the system might point to specific phrases or indicators as evidence of an attempt to obscure information. This summary allows human reviewers to easily verify the alert and defend the decision to regulators.

AI Performance Transparency and Metrics
DCGA platforms provide built-in audit reporting that tracks the performance of risk detections over time and generates statistics regarding the frequency of positive or negative alerts. These reports can be used as a continuous feedback loop to track false negatives and false positives. Furthermore, these metrics can be provided to internal and external stakeholders to demonstrate robust AI oversight.

Human-in-the-Loop Controls
Firms can retain full operational control by enabling or disabling AI risk detections at any time. Compliance teams can also manage the risk sensitivity of AI detections—categorizing them as risky, informational, or validations—which results in higher, lower, or neutral aggregated conversation risk scores. These risk scores allow compliance teams to control and prioritize which conversations are reviewed.

AI System Development and Management Explainability

In parallel with explainability features, the provision of robust, auditable protocols by vendors for the oversight and governance of their AI models and systems is paramount. Financial services firms should assess whether third-party service providers align with the following protocols and frameworks:

ISO/IEC 42001 for Artificial Intelligence Management Systems (AIMS)
ISO/IEC 42001 certification involves rigorous, independent auditing of how an organization establishes, operates, monitors, and maintains its AIMS, requiring auditable evidence across data governance, explainability, and incident management. Certified organizations must demonstrate how their AI systems function, what data they use, how that data is protected, and how decisions are made. The certification provides verifiable assurance of governance maturity and responsible AI development.

CSA STAR for AI Level 2
The STAR for AI Level 2 certification builds on the ISO/IEC 42001 framework by integrating the Cloud Security Alliance (CSA) AI Controls Matrix to deliver macro governance, granular security, and continuous validation to address bias mitigation, model risk management, algorithmic explainability, and training data privacy for AI.

Third-Party Assurance and Audits
Vendors can continuously assess AI vulnerabilities and systemic risks as part of independent third-party reviews such as SOC 2 Type II, ISO 27001, PCI DSS, and TruSight audits. These third-party audits provide consistent, repeatable, and measurable protocols that demonstrate compliance with articulated controls, including those critical to AI development processes, administrative access, and the software development lifecycle.

AI Model Performance and Continuous Improvement
AI models should be designed to be fair and objective in risk detection capabilities to mitigate systemic bias. This includes training models on thousands of examples without incorporating sensitive characteristics, such as age, race, religion, sexual orientation, or ethnicity. Continuous improvement and ongoing oversight are critical components of AI governance. Vendors focus on long-term performance by monitoring models post-deployment for data and concept drift, and by routinely retraining models to account for emerging patterns and feedback. Models are regularly refined to address evolving regulatory compliance expectations.

KEY AI GOVERNANCE REGULATIONS & GUIDANCE

The governance and oversight of AI is a key focus of regulators and national authorities.

The regulatory landscape is evolving with regulators worldwide publishing rules and guidelines to ensure the compliant and responsible use of AI in financial services firms. Some of the key pronouncements from global regulators relating to firms' use and oversight of AI include:

The SEC's FY 2026 Examination Priorities sets out Information Security and Operational Resiliency and Emerging Financial Technology as key areas it will focus on during examinations. It notes that the Division will continue to review registrant practices to prevent interruptions to mission-critical services and to protect investor information, records, and assets and that the "focus will be on training and security controls that firms are employing to identify and mitigate new risks associated with artificial intelligence (AI)". It will further "assess whether firms have implemented adequate policies and procedures to monitor and/or supervise their use of AI technologies.."

The CFTC's Letter No. 24-17 similarly states that its "Staff expects that CFTC-regulated entities will assess the risks of using AI and update policies, procedures, controls, and systems, as appropriate, under applicable CFTC statutory and regulatory requirements" clarifying that "as with all material system or process changes, a CFTC-regulated entity should ensure that its adoption of AI has been reviewed for compliance with the CEA and CFTC regulations."

FINRA's 2026 Annual Regulatory Oversight Report includes a new, dedicated section on "GenAI: Continuing and Emerging Trends" which reminds firms that "FINRA's rules—which are intended to be technologically neutral—and the securities laws more generally, continue to apply when firms use GenAI or similar technologies in the course of their businesses, just as they apply when firms use any other technology or tool." It explains that firms contemplating using GenAI tools and technologies may want to consider " Ongoing monitoring of prompts, responses and outputs to confirm the GenAI solution continues to perform as expected and results in compliant behavior."

The UK FCA's AI Update reinforces that it already has "a number of frameworks in place which are relevant to firms' safe use of AI" but will continue to closely monitor the adoption of AI across UK financial markets to identify material changes that impact on consumers and markets. This includes keeping under review if amendments to the existing regulatory regime are needed.

The Federal Reserve's Supervisory Letter SR 26-2 on Revised Guidance on Model Risk Management -- April 17, 2026 updates its Model Risk Management (MRM) guidance to address AI and complex third-party tools noting that an important element of model risk management is the validation of vendor

products, either by internal or outside parties. Sound practice includes developing an understanding of the vendor model, including its conceptual soundness, design, development data, and performance. Similarly, sound practice involves conducting ongoing monitoring and outcome analysis to assess whether vendor models are accurate, remain fit for purpose, and continue to be reliable.

In Canada, CSA Staff Notice and Consultation 11-348 outlines the existing requirements under securities law and guidance that apply to market participants' use of AI systems in capital markets. In the AI Governance & Oversight section it notes that market participants policies and procedures should be designed in a way that accounts for the unique features of AI systems and the risks they pose, including: having a "human-in-the-loop", where humans can effectively monitor the input and/or output of an AI system; adequate measures to mitigate the technological and operational risks related to the use of AI systems (e.g. initial and ongoing monitoring of cybersecurity risks, AI system bias, model drift and hallucinations); and the importance of data in the functioning of AI systems, including ensuring that data be accurate and complete, not include prohibited information and account for privacy considerations.

RECOMMENDATIONS TO CONSIDER

The following checklist can be used by financial services firms to evaluate the AI Governance aspects of DCGA platforms.

Explainability and Audit Readiness
Does the platform provide audit-ready summaries explaining the specific reasons that triggered an AI-driven risk detection?

ISO/IEC 42001 Compliance
Has the vendor established, implemented, and maintained an Artificial Intelligence Management System (AIMS) that is certified by an independent third party?

CSA STAR for AI Level 2
Has the vendor achieved the Cloud Security Alliance STAR for AI Level II certification, demonstrating global AI compliance, safety, and operational transparency?

Third-Party Framework Attestations
Does the vendor maintain third-party audits and attestations, including SOC 2 Type II and ISO 27001 that cover model development and administrative access?

Human-in-the-Loop Controls
Can compliance teams review AI decisions and adjust risk sensitivity (risky, informational, or validation) as well as enable or disable specific AI classifiers to maintain operational control?

Continuous Monitoring and Drift Detection
Does the vendor regularly monitor models post-deployment to measure data and concept drift?

AI Performance Transparency and Metrics
Does the DCGA platform provide built-in audit reporting that tracks the performance of risk detections over time?

Learn More Visit: **ThetaLake.com**