A DIGITALXFORCE PERSPECTIVE

# Digital Trust in the AI Era

*Governing AI and Third-Party Risk Through Continuously Proven Trust*

**By Margareta Petrovic** • Chief AI Officer, DigitalXForce

## Executive Summary

Artificial intelligence has changed the physics of enterprise risk. Models retrain overnight. Autonomous agents act between audits. Third-party vendors carry data, code, and AI-driven decisions so deeply into the enterprise that "the perimeter" no longer describes where risk begins. Yet most governance, risk, and compliance programs still run on a model built for a slower world — annual questionnaires, point-in-time assessments, and a binary notion of trust that a vendor or a model either has or does not have.

This paper argues that the answer is not more controls. It is a different kind of trust. Zero Trust was the right answer for the identity and access era. Digital Trust — trust that is continuously validated with real-time evidence rather than declared once a year — is the requirement for the AI era. This is the framework Margareta Petrovic will walk through in her ISACA OC GRC Summit 2026 session, reinforced by a live platform demonstration.

## 1. The Three Assumptions That No Longer Hold

Traditional GRC, AI governance, and third-party risk management rest on three assumptions that the AI era has quietly invalidated. Recognizing them is the first step to designing a program that will survive the next three years.

***Assumption 1 — Risk can be assessed periodically.***

An annual TPRM questionnaire assumes the vendor on day 365 is materially the same vendor you assessed on day 1. For a static SaaS provider in 2015, that was roughly true. For a vendor whose product now embeds an AI model that retrains weekly on customer data, it is not. By the time the next annual review arrives, the model may have drifted, been fine-tuned on new data, or been replaced entirely — and your control assurance expired the day after you signed the attestation.

***Assumption 2 — Trust is binary and implicit.***

Traditional TPRM produces a verdict: approved or not approved. Traditional AI governance produces a sign-off: the model is in production. Neither reflects the reality that the trustworthiness of an AI system or a vendor is a continuous, observable property that rises and falls with signals — new CVEs, configuration drift, bias test results, data-handling changes, sub-processor additions. Treating trust as a stamp creates a dangerous gap between what the documentation says and what the live system is doing.

***Assumption 3 — Compliance equals safety.***

Every serious failure of the past decade — Equifax, SolarWinds, MOVEit, several high-profile AI bias incidents — involved organizations that were compliant on paper. Compliance frameworks are necessary but they describe a floor, not a ceiling. In the AI era, where regulatory frameworks (NIST AI RMF, ISO 42001, EU AI Act, SR 11-7, state AI laws) are evolving faster than most enterprises can adopt them, treating compliance as sufficient is an actively risky strategy.

> *"Zero Trust was the start. Digital Trust is the requirement for the AI era."*

## 2. From Zero Trust to Digital Trust

Zero Trust answered a specific question: how do we handle identity and access when the network perimeter has dissolved? Its answer — never trust, always verify — was the right one. But Zero Trust was designed for a world of humans, devices, and applications. It says very little about how to govern autonomous AI agents, how to continuously validate the trustworthiness of a vendor's AI model embedded in your stack, or how to tell a board that an emerging risk has been detected, contained, and measured in real time.

Digital Trust extends the Zero Trust principle from access decisions to every dimension of risk. Where Zero Trust asks "should this request be allowed?", Digital Trust asks a broader question: "can we prove, right now, that this system, vendor, or model is behaving within the bounds of the trust we granted it — and if it is not, how quickly can we detect and respond?"

Operationally, Digital Trust means four things:

- **Unified** — security, compliance, audit, third-party risk, AI risk, and operational resilience are collapsed from independent silos into a single continuously updated posture.
- **Continuous** — trust is measured in real time rather than asserted annually. Signals flow from the environment to the risk model continuously.
- **Evidence-based** — every trust claim is backed by observable evidence — telemetry, test results, configuration state — not a questionnaire response.
- **Board-communicable** — technical risk is translated into a score, a narrative, and an action that a board, regulator, or customer can use.

## 3. The Digital Trust Blueprint

The blueprint below is vendor-neutral. It describes the control domains and the continuous evidence signals any enterprise needs to instrument in order to move from periodic to continuous trust validation — whether delivered through a platform like DigitalXForce, built in-house, or assembled from best-of-breed components.

| Control Domain | Continuous Evidence Signals |
|---|---|
| **AI Model Governance** | Model inventory, version changes, drift metrics, bias test results, approval workflow state, usage telemetry |

| Control Domain | Continuous Evidence Signals |
|---|---|
| **Third-Party Risk** | Real-time vendor security posture, SBOM updates, breach signals, contract obligations, concentration risk |
| **Cyber Posture** | Control effectiveness telemetry, vulnerability status, identity and access signals, configuration drift |
| **Regulatory Compliance** | Framework mapping (NIST AI RMF, ISO 42001, SR 11-7, EU AI Act), evidence freshness, control coverage gaps |
| **Operational Resilience** | Dependency mapping, recovery objectives vs. actuals, incident signal correlation across AI and vendor layers |

The key design decision is the unified layer. Most enterprises already collect many of these signals somewhere. The breakthrough is treating them as inputs to a single trust posture, rather than as separate dashboards owned by separate teams. When AI governance, TPRM, and cyber can see each other's signals, the program becomes resilient in a way that no amount of additional controls in isolated towers can match.

## 4. The Digital Trust Score — Making Trust Board-Ready

Boards do not want a list of controls. They want a question answered: "are we trusted, and how do you know?" A Digital Trust Score is the board-facing artifact of the blueprint above — a normalized, time-series measure of trust posture that rolls up signals from AI models, third parties, cyber controls, and compliance state into a single number with a defensible decomposition.

The value of the Score is not the number. It is the conversation it enables. Instead of a compliance report that tells the board what happened last quarter, the Score tells them what is true right now, where it is trending, and what the top three drivers of movement are. When a vendor's posture drops or a model begins to drift, the Score moves before the incident — not after it.

> *"Rather than asking 'are we compliant?', leaders can now ask — and answer — 'can we prove trust right now?'"*

## 5. Industry Flavors: How This Plays Out

The Digital Trust framework is domain-agnostic, but the signals and regulatory context differ sharply by industry. The session at ISACA OC will include case vignettes from three sectors where the model is already being deployed.

| Financial Services | Healthcare | Government & Public Sector |
|---|---|---|
| PRIMARY RISK<br>AI decision transparency and model risk (SR 11-7, FRB SR 11-7 extensions for GenAI) | PRIMARY RISK<br>AI safety, algorithmic bias, PHI protection, HIPAA interoperability | PRIMARY RISK<br>Vendor trust, national risk exposure, FedRAMP and sovereign-data obligations |
| WHAT DIGITAL TRUST DELIVERS<br>Continuous model monitoring tied to credit, fraud, and trading decisions; real-time evidence for regulators and internal audit | WHAT DIGITAL TRUST DELIVERS<br>Bias and drift detection across clinical AI; continuous control mapping to HIPAA, FDA, and state AI laws | WHAT DIGITAL TRUST DELIVERS<br>Real-time trust signals across the vendor ecosystem; continuous authorization-to-operate (cATO) posture |

## 6. Three Moves for Monday Morning

The ISACA OC session closes with three moves any attendee can take back to their organization the same week, regardless of whether they have a platform budget or are starting from scratch.

- **Pick the two riskiest dependencies.** — Pick the AI model and the third party that, if compromised or changed tomorrow, would hurt the business most. Ask: what signals would tell us that is happening? Map the gap between those signals and what you actually collect.
- **Introduce one Digital Trust metric.** — Add a single trust score row to the board risk report. Even a crude version forces the team to aggregate signals that live in separate towers today, and creates the habit of continuous measurement.
- **Cut one assessment cadence in half.** — A quarterly TPRM review is better than annual; a monthly AI model posture check is better than quarterly; a weekly drift scan is better than monthly. Move one cadence by one step and measure what you find.

## A Final Word

The AI era will not be governed by the controls of the compliance era. It will be governed by enterprises that treat trust as an operational capability — instrumented, measured, and continuously proven. The organizations that build this capability first will not only be safer; they will be faster, because the answer to “are we trusted right now?” is also the answer to “can we ship this?”

That is the conversation the ISACA OC GRC Summit 2026 session is built to start — with a framework attendees can apply immediately, a blueprint they can adapt to their own environment, and a live demonstration of what continuous trust validation looks like when the signals actually flow.

ABOUT THE AUTHOR

Margareta Petrovic is Chief AI Officer at DigitalXForce. A pragmatic, results-driven executive with two decades of experience designing and delivering cybersecurity and risk programs for global enterprises across financial services,

healthcare, and critical infrastructure, she advises C-suites and boards on integrated cybersecurity architecture, AI governance, third-party risk, and compliance automation. She is based in Lake Forest, California.

### ABOUT DIGITALXFORCE

DigitalXForce is a next-generation, AI-native platform for Automated GRC and Enterprise Security Risk Posture Management (ESRPM). The company was named a leader in the IDC MarketScape 2025 assessment for worldwide GRC software vendors and a 2026 Global InfoSec Awards winner for AI-Powered Risk Management. Its X-ROC™ (Extended Risk & Operations Center) unifies security, compliance, audit, third-party risk, AI risk, and operational resilience into a continuously updated view of enterprise trust.