

Status: Canonical Version: 1.0 Effective Date: Upon publication Authority:  
Zenodo: 10.5281/zenodo.18031821

## Memorandum

**Subject:** AI Representation Risk and the Emerging Requirement for Audit-Grade Evidence

**Audience:** Board of Directors, Audit Committee, CFO, General Counsel

**Purpose:** Risk identification and governance posture, not technology strategy

---

## Executive Summary

Artificial intelligence systems now influence how organizations are represented to customers, partners, regulators, and the public. This influence often occurs **outside the organization's control, visibility, or ability to correct in real time.**

The primary risk is no longer technical error. It is **misrepresentation with potential legal, financial, and reputational consequences**, occurring without preserved evidence.

This memo outlines:

1. The nature of the emerging risk
  2. Why existing controls are insufficient
  3. Why audit-grade evidence is becoming mandatory
  4. A governance checklist for board-level oversight
- 

## 1. What Has Changed

Historically, AI risk discussions focused on internal use cases and model accuracy.

That framing is now outdated.

Today, AI systems:

- Compare companies and products
- Summarize financial and legal positions
- Frame risk, quality, and suitability
- Influence decisions upstream of formal disclosures

These representations increasingly occur:

- In third-party systems
- Without notification
- Without persistence
- Without reproducibility

The organization may be affected **without participating**.

---

## 2. Definition of the Risk

**Externally mediated representation risk** arises when:

*An AI system's interpretation of an organization becomes operationally relevant, despite the organization not owning, controlling, or being able to reliably observe that system.*

Key characteristics:

- The system is external
- The representation is consequential
- The output is transient
- The organization lacks evidence of what was said

This risk exists regardless of whether the AI system is “intelligent” or statistically accurate.

---

### **3. Why Existing Controls Do Not Cover This Risk**

#### **a) Accuracy controls are insufficient**

Model performance metrics do not address:

- Temporal drift
- Context sensitivity
- Category substitution
- Inconsistent framing across sessions

A statement can be accurate in isolation and still be misleading in context.

#### **b) Screenshots and anecdotes are not evidence**

In disputes, regulators and courts do not accept:

- Screenshots
- Vendor dashboards
- After-the-fact explanations

They accept:

- Time-stamped artifacts
- Reproducible methods
- Independent documentation

#### **c) Intervention increases exposure**

Attempting to “correct” AI outputs without a preserved record can:

- Create attribution ambiguity
- Trigger disclosure obligations

- Introduce manipulation claims
- Expand liability

Correction without evidence is not defensible diligence.

---

#### 4. Why Audit-Grade Evidence Is Becoming Inevitable

This is not a speculative trend. It follows a familiar pattern.

Every system that mediates value at scale eventually requires:

- Independent observation
- Standardized documentation
- Reproducibility
- Separation between operation and oversight

Examples:

- Accounting before GAAP
- Aviation before flight data recorders
- Financial markets before disclosure regimes

AI has reached the same inflection point:

- Consequence has outpaced control
  - Trust cannot be asserted without evidence
  - Oversight requires preserved context
- 

#### 5. The Shift Boards Should Recognize

AI governance is moving from a **capability era** to an **accountability era**.

DIMENSION	EARLIER FOCUS	EMERGING REQUIREMENT
Primary Question	How accurate is the system?	Can we prove what it represented?
Risk Type	Technical error	Legal and reputational liability
Evidence	None or anecdotal	Preserved, reproducible artifacts
Oversight Owner	IT / Innovation	Legal, Finance, Risk
Failure Mode	Hallucination	Misrepresentation

Boards should assume this shift is irreversible.

---

## 6. Governance Checklist for Directors

The following questions should be answerable without relying on vendors or assurances:

### 1. **Visibility**

Do we have any systematic way to observe how AI systems represent our organization externally?

### 2. **Evidence**

Can we reproduce what an AI system said about us at a specific point in time?

### 3. **Independence**

Is observation separated from intervention, optimization, or correction?

### 4. **Documentation**

Would preserved artifacts meet regulatory or litigation standards?

### 5. **Escalation**

Is there a defined process when material misrepresentation is detected?

## 6. Disclosure Readiness

Could we demonstrate diligence if asked by a regulator, insurer, or court?

If the answer to more than two of these is “no,” the organization is exposed.

---

## 7. Recommended Board Posture

This memo does **not** recommend:

- Modifying AI systems
- Influencing external models
- Making claims about correctness

It recommends:

- Treating AI representations as an external risk surface
- Prioritizing observability over optimization
- Preserving evidence before intervention
- Framing AI exposure as a governance issue, not an IT feature

The objective is not control.  
It is defensibility.

---

## Closing Note

AI adoption is accelerating faster than accountability frameworks.

Boards will not be judged on whether they predicted AI capabilities correctly.

They will be judged on whether they exercised reasonable oversight once consequences became foreseeable.

The question is no longer:

*“Is the AI smart?”*

It is:

*“Can we prove what it said about us, and can we prove we acted responsibly in response?”*