

THE AI SECURITY & GOVERNANCE AUTHORITY

PROVEN, NOT PRESENT.

CONTINUOUS CONTROL ASSURANCE FOR ENTERPRISE AI

Why a configured control is not yet a proven control.

FOUNDER-ORIGINATED PUBLIC PERSPECTIVE

ALI AQEL / AISPG™ / OCTOBER 2026

PUBLIC EDITION — CONCEPTUAL RESEARCH

FIELD NOTE / 01

A control that exists is not yet a control that works.

Most enterprises can list their AI-relevant policies and controls. Far fewer can show that those controls apply to the right systems, operate under real conditions and still hold after the environment changes.

01 Inventories describe intent

A policy catalog records what the organization means to happen. It does not show whether any mechanism carries that intent into the systems where AI actually runs.

02 Configuration describes a setting

A configuration check shows that a mechanism appears to be set. It cannot show by itself that the mechanism covers every intended workload or survives an edge condition.

03 Assurance describes reliance

Assurance asks whether leadership has a current, trustworthy basis to rely on a control for a specific decision, and states what would make that reliance expire.

THE AISPG VIEW

AISPG™ treats the distance between control presence and control effectiveness as an architectural concern, not an audit afterthought. Assurance should be proportional to consequence, change and uncertainty.

FIELD NOTE / 02

How controls fail quietly in AI environments.

AI systems are created, connected and changed faster than most control inventories are reviewed. A control can look healthy on paper while one of these conditions has already defeated it.

01 / SCOPE

A control is correctly configured, but for a different account, region or workload population than the AI service the business believes it protects.

02 / IDENTITY

A control depends on an identity or permission boundary that changed when an agent, pipeline or service principal was added.

03 / RUNTIME

A control produces reassuring artifacts and reports while failing to produce the intended outcome when the AI system is actually used.

04 / CHANGE

A new model endpoint, connector or deployment path appears outside the inheritance that the original control was designed around.

THE REFRAME

Do not ask only whether a control is enabled. Ask whether it applies, operates, covers the required scope, produces the intended outcome and remains trustworthy under the conditions that matter.

FIELD NOTE / 03

Five different things a control can be.

Treating these as one state is how control inventories create false confidence. AISPG separates them so that intended behavior and proven behavior are never confused.

01 Defined

The control is specified with an accountable owner, a scope and the objective it serves.

02 Implemented

A mechanism is present in the governed environment that is meant to carry out the control.

03 Operating

The mechanism is actually executing during routine conditions, not merely installed.

04 Effective

The control produces the required outcome across the intended scope, including the edge cases that matter.

05 Assured

Effectiveness has been validated to a depth that matches the consequence, with a known expiry.

DESIGN PRINCIPLE

Each state needs its own kind of evidence. A setting cannot stand in for operation, and operation cannot stand in for effectiveness.

FIELD NOTE / 04

From a slogan to a claim you can test.

"Protect sensitive data" cannot be proven or disproven. Assurance starts when a control is stated precisely enough that it could fail. A testable claim names five things.

01**SUBJECT**

Which AI entity, identity or workflow is the claim about?

02**ACTION**

What may it do, or be prevented from doing?

03**BOUNDARY**

Which data, environment, tool or tenant limit applies?

04**EXPECTED OUTCOME**

What should observably happen when the boundary is approached or crossed?

05**FAILURE CONDITION**

What observation would show the claim is false, and who must then be told?

DESIGN PRINCIPLE

A passed test supports the claim that was tested, not universal security. A failed test is a finding to surface immediately, not a result to hold until remediation is complete.

FIELD NOTE / 05

Assuring one control: an illustrative path.

Hypothetical example: an internal assistant may retrieve documents only when the requesting person is entitled to them. The steps below are illustrative, not the result of a production assessment.

1

01 / STATE THE CLAIM

Name the assistant, the document classes, the entitlement rule and what a violation would look like.

2

02 / EXERCISE SAFELY

Under written authorization and agreed limits, attempt a representative out-of-entitlement request.

3

03 / RECORD THE RESULT

Capture what was attempted, what the control did, the evidence relied on and its limitations.

4

04 / SET THE EXPIRY

Define which changes revoke the conclusion, such as a new connector, model version or permission model.

ILLUSTRATIVE ONLY

This is a teaching scenario, not a customer claim, control test, attestation, benchmark or record of AISPG production activity.

FIELD NOTE / 06

Six questions for boards, CISOs and AI leaders.

These questions separate what an organization has configured from what it can actually rely on, without replacing the functions that own each control.

- 01 Which of our AI controls could fail tomorrow without anyone being told?
- 02 For our most consequential AI use, which controls are assured rather than merely configured?
- 03 Is assurance depth matched to consequence, or does every control get the same light review?
- 04 When a test fails, how quickly does the affected leadership conclusion change?
- 05 Which environment changes automatically trigger re-validation of a prior conclusion?
- 06 Where are we relying on a compensating control, who owns it, and when does it expire?

START WITH ONE REAL CONTROL

Pick the control your most consequential AI system depends on. Write it as a testable claim, agree how it could be safely challenged and record what you can and cannot yet prove.

FIELD NOTE / 07

Context. Sources. Further inquiry.

AISPG's interpretation is independent. The following authoritative publications are external context, not endorsement, adoption, certification or validation of AISPG.

REFERENCE / 01

NIST AI Risk Management Framework 1.0

NIST, 2023. Voluntary risk-governance framework with Govern, Map, Measure and Manage functions.

<https://airc.nist.gov/airmf-resources/airmf/>

REFERENCE / 02

NIST SP 800-53A Revision 5

NIST, 2022. Guidance for assessing security and privacy controls in information systems and organizations.

<https://csrc.nist.gov/pubs/sp/800/53/a/r5/final>

REFERENCE / 03

ISO/IEC 42001:2023

International standard specifying requirements for an AI management system.

<https://www.iso.org/standard/42001>

REFERENCE / 04

MITRE ATLAS

A public knowledge base of adversary tactics and techniques against AI-enabled systems.

<https://atlas.mitre.org/>

PUBLIC EDITION / DISCLOSURE BOUNDARY

© 2026 Ali Aqel / AISPG™. This original paper is provided for education and discussion. AISPG™ is independently owned by Ali Aqel and licensed to Blue Sentinel AI, Inc. AISPG is a privately owned, founder-originated framework, not a regulator or accredited certifier. Conceptual diagrams and questions do not represent live control validation, licensed operating procedures, legal advice or standards conformity. Proprietary control logic, undisclosed reasoning mechanisms, private architecture and sensitive platform verification remain outside this edition.

EXPLORE THE PUBLIC AUTHORITY EXPERIENCE

AISPG.ai / Research & qualified access