UNKNOWNS LAB
[SYS.00 / FRAMEWORK.v1.0]

The Decision Authority Framework

A discipline for governing critical decisions

in the cyber and AI era.

Systems don't fail first. Decisions do.

This paper defines Decision Authority as a measurable discipline — distinct from incident response, governance-risk-compliance, crisis management, and tabletop rehearsal — and proposes four primitives by which organizations, auditors, insurers, and regulators can assess it.

Version 1.0 · April 2026 · Open for comment
Published by Unknowns Lab · unknownslab.com

[PREFACE]

00

About this document

This document is version 1.0 of the Decision Authority Framework, published by Unknowns Lab as a contribution to the emerging discipline of decision governance in high-consequence enterprises. It is written for boards, general counsel, chief information security officers, chief risk officers, regulators, insurers, and the analysts who serve them.

It is not a product brochure. It is not a methodology for hire. It is a public articulation of a category we believe will become as legible in the next decade as cybersecurity itself became between 2010 and 2020 — because the forcing functions that made cybersecurity a board-level concern are now operating on decision-making itself.

We welcome reference, critique, adoption, and adaptation. The four primitives introduced in §6 are offered as open definitions; practitioners who cite them are asked only to cite accurately. Where this document is adopted into regulatory guidance, underwriting methodology, or analyst taxonomy, Unknowns Lab will publish annotated revisions in subsequent versions.

> *This framework exists because the space between detection and decision is the least measured and most consequential failure surface in modern enterprises.*

Contents

[PART I]

The Problem

[§1]

The Preamble: Why This Document Exists

Between 2020 and 2025, cyber and AI failures stopped being engineering problems and started being decision problems. The proof is in the public record. In nearly every major incident that made the front page — a ransomware event at a hospital network, a supply-chain compromise at a software vendor, an autonomous-system failure at a commercial airline, a data-exposure cascade at a financial platform — the technical root cause was present, but the cost was determined by what leadership decided to do in the first hours after detection.

In our observation across principal-led engagements with organizations whose failure costs are measured in hundreds of millions to billions of dollars, the pattern is consistent: detection times have compressed; escalation paths have not. Technical telemetry now arrives in minutes. Authority to act on that telemetry still arrives in hours, days, or — for novel AI failures — never.

> *Detection is an engineering problem, and we have largely solved it.*
> *Decision is a governance problem, and we have barely begun.*

This is not a matter of better playbooks, thicker runbooks, or more frequent tabletop exercises. Those address the surface. The underlying issue is structural: large organizations have not articulated, with the clarity that modern adversarial speed demands, who holds the authority to decide what, under which conditions, with what reversibility, and with what accountability when the decision fails.

This document names that gap, gives it a vocabulary, and proposes measurable primitives by which an organization can assess and improve its position. It does so in three parts: the problem (§§1–3), the framework (§§4–6), and the application to the AI era, to assessment, and to external stakeholders (§§7–9).

What this document is not

It is not a regulatory framework. It is not a compliance standard. It is not a substitute for incident response methodology, crisis management doctrine, or enterprise risk management. It is a companion discipline to each of those — operating in the space each has historically deferred.

Who should read it

Board directors with cyber or AI risk oversight responsibility. General counsel advising on disclosure, privilege, and liability during incidents. Chief information security officers tasked with outcomes that exceed their formal authority. Chief risk officers whose models do not yet accommodate decision failure as a risk class. Regulators drafting supervisory expectations for AI accountability and cyber resilience. Underwriters pricing cyber and AI D&O exposure. The analyst community that will ultimately name this category in public taxonomy.

[§2]

The Failure Surface Nobody Measures

We observe three recurring failure patterns in post-incident reviews of catastrophic cyber and AI events. Each has engineering literature. None has a measurement discipline in the organizations where the failures actually occur.

2.1 Decision Latency

Decision latency is the interval between the moment sufficient information exists to authorize a decision and the moment the decision is actually authorized. It is not the time to detect. It is not the time to contain. It is the time to decide.

In our field observations, decision latency during a contested incident runs 4× to 40× longer than the detection-to-authorization gap senior leadership believes it runs. Industry post-mortems of public incidents tend to corroborate this pattern: the technical evidence is present early; the authorized response is delayed by organizational dynamics that have never been instrumented.

2.2 Authority Ambiguity

Authority ambiguity is the condition in which, at the moment a decision must be made, no one in the chain is certain who is empowered to make it. The consequences vary: isolation of a compromised segment is delayed; public disclosure is advanced or withheld against counsel; a ransom negotiation is opened by a function without mandate; an AI system continues to act after human operators believe it has been halted. The common feature is that the ambiguity did not exist two weeks before the incident. It emerged under pressure, from policies that were legible on paper but underspecified in execution.

> *Authority ambiguity is almost always invisible in peacetime. It materializes at machine speed during a crisis.*

2.3 AI Handoff Failure

AI handoff failure is the discontinuity between an autonomous or semi-autonomous system acting on its own authority and the human supervisory layer that is assumed — but not architected — to intervene. It has become the most rapidly growing category of decision failure in our engagement data, and it is the pattern for which existing governance categories have the least coverage.

The typical failure mode is not an AI doing something catastrophic. It is an AI doing something consequential without any human noticing in time, or with humans noticing but finding that the override path is slower than the system's next action, or with humans noticing and overriding but discovering after the fact that no one is willing to claim accountability for the override itself.

[OBSERVED BASELINES · UNKNOWNS LAB]

Breaches in which leadership decision failure materially worsened outcome	**83% (public post-incident analyses, 2022–2025)**
Median time for an affected organization to identify its own decision-authority gaps, post-incident	**72 hours**
Critical window during which early decisions compound exponentially in cost	**First 60 minutes**
Probability that a Fortune 2000 organization has a written, tested, board-ratified authority map for a cross-functional cyber incident in 2026	**Below 15%**

[§3]

Why Existing Categories Do Not Answer

Every pattern in §2 has an adjacent discipline that addresses some portion of it. None addresses it centrally. The Decision Authority discipline exists in the overlapping blind spots of five established categories. The purpose of this section is not to diminish those categories but to name precisely what each defers.

3.1 Incident Response

Incident Response is mature, well-instrumented, and indispensable. Its center of gravity is technical containment: isolate, eradicate, recover, analyze. The IR function is answerable to the CISO. Its outputs are technical timelines, scope assessments, and remediation plans. What it defers — by design — is the authority architecture above the CISO: which decisions require legal concurrence, which require board notification, which require regulator engagement, which require CEO personal sign-off, and how quickly each of those channels can converge.

3.2 Governance, Risk, and Compliance

GRC programs produce the evidentiary artifacts that auditors and regulators require: policies, controls, attestations, risk registers, and residual-risk acceptances. They encode the organization's stated posture. What they do not do — what they have never aspired to do — is instrument whether the stated posture holds under live adversarial conditions. GRC answers 'do you say you have a policy?' Decision Authority answers 'does the policy decide at machine speed when challenged?'

3.3 Crisis Management

Crisis management is a legitimate practice, typically resident in communications and legal functions, activated after an incident becomes public. Its center of gravity is stakeholder perception: regulator tone, media posture, customer confidence, share price. Its decisions are reactive to external reality. It rarely designs the internal authority chain that determines what becomes a public crisis in the first place.

3.4 Tabletop Exercises

Tabletop exercises are the dominant form of decision rehearsal in most large enterprises. They have a known limitation: they are rehearsal theater. Participants know they are in an exercise; the pressure is simulated; the authority chain is often explicitly announced at the start of the session. Tabletops produce atmospherics. They seldom produce structural outputs — durable authority maps, codified escalation logic, or measurable latency reductions — that survive the meeting.

3.5 Cyber Insurance

Cyber insurance has evolved from indemnification of loss to partial underwriting of operational posture. Underwriting questionnaires now probe controls, training, and incident history. They do not yet probe decision authority — because no scoring mechanism exists for it. The absence is the opportunity: an insurer equipped with a decision-authority score can underwrite a risk no competitor can see.

3.6 Where decision authority sits in the stack

Discipline	What it governs	What it defers
Incident Response	Technical containment	Authority chain above CISO
GRC	Stated posture & evidence	Live decision performance
Crisis Management	External perception	Internal authority architecture
Tabletop Exercises	Rehearsal atmospherics	Structural outputs & latency
Cyber Insurance	Loss indemnification	Authority as an underwriting signal
Decision Authority	Who decides, how fast, with what accountability	— (the layer itself)

The disciplines above are complements, not substitutes. A mature enterprise requires all of them. What it has historically lacked is the layer that binds them under live adversarial pressure — the layer that determines, in the first sixty minutes of an unfolding incident, which of those disciplines' outputs actually get to act.

[PART II]

The Framework

[§4]

Decision Authority: A Working Definition

> *Decision Authority (n.) — the structured, measurable capacity of an organization to identify, assign, and exercise the right to decide, at sufficient speed and with sufficient accountability, under conditions of adversarial pressure, incomplete information, and irreversible consequence.*

This definition is deliberately constructed to be falsifiable. Each clause names a property that can be measured, failed, and improved:

- "Structured" — authority is articulated before the event, not improvised during it.
- "Measurable" — an organization's position can be scored, benchmarked, and re-scored.
- "Identify, assign, and exercise" — three distinct acts, each a separate failure mode.
- "At sufficient speed" — latency is a first-order property, not a secondary attribute.
- "With sufficient accountability" — the authority holder is identifiable after the fact.
- "Adversarial pressure, incomplete information, irreversible consequence" — the operating conditions under which most organizations have never instrumented their own decisioning.

Decision Layer vs. Decision Authority

We use two related terms with deliberate distinction. Decision Authority is the discipline — the thing measured, governed, and improved. The Decision Layer is the architectural surface — the organizational and, increasingly, technical infrastructure through which Decision Authority is exercised. A mature enterprise in 2030 will have a Decision Layer in the same way a mature enterprise in 2020 has a security operations center: as legible, as staffed, and as auditable.

[§5]

The Three-Layer Operating Model

Decision Authority operates across three time horizons. Each horizon has a distinct mode, distinct outputs, and distinct failure signatures. A mature Decision Layer spans all three. An immature one is present in one or none.

5.1 PEACETIME · Anticipate

[HORIZON: months to years before incident]

In peacetime, Decision Authority is built. Adversary intent is mapped — not merely indicators and tactics, but the strategic motivations of actors with reason to target the enterprise. Decision paths are stress-tested under simulated but unannounced pressure. Human-AI authority boundaries are articulated before any agentic system goes live. The outputs of this layer are durable assets: authority maps, escalation graphs, scored baselines. The failure signature is the absence of these assets, or their presence only as draft documents no one has operationalized.

5.2 CRISIS WINDOW · Decide

[HORIZON: minutes to hours during incident]

In the crisis window, Decision Authority is exercised. The interval is compressed — often to the first sixty minutes, rarely longer than the first seventy-two hours. Inside this window, the structures built in peacetime either hold or they do not. The CISO has or does not have the authority to isolate. The general counsel has or does not have the authority to engage regulators. The board has or does not have the authority to override a subsidiary's local response. The failure signature is visible in the telemetry of the decision itself: elapsed time from trigger to authorized action, number of escalation hops required, number of reversals.

5.3 PERMANENT · Build

[HORIZON: years after incident, and continuous]

In the permanent layer, Decision Authority is compounded. Lessons from the crisis window feed back into peacetime structure. Leadership transitions are instrumented so authority does not evaporate with a CISO or CEO change. Board governance evolves to include decision authority as an explicit charter item. Human-AI boundaries are revised as the AI capability curve moves. The failure signature of this layer is the organization that treats each incident as a discrete event rather than an input to a compounding capability.

> *An organization is as mature in Decision Authority as its weakest layer. The most common weak layer, in our observation, is the permanent one.*

[§6]

The Four Primitives

The framework proposes four primitives by which Decision Authority can be measured. The primitives are intended to be adopted, cited, and refined by the broader community. They are open definitions; they are not proprietary to any single practitioner, including ourselves.

6.1 The Authority Graph

An Authority Graph is a directed representation of the organization's decision rights: the nodes are decisions, the edges are the authority relationships between the roles or bodies that hold them. At minimum, a Decision Authority Graph answers four questions for every material decision class:

- Who holds primary authority?
- Who holds concurrent authority (veto, concurrence, notice-only)?
- What is the reversal path if the decision proves wrong?
- Who is accountable after the fact?

In mature implementations, the graph covers at least thirty to fifty decision classes spanning cyber incident response, AI-system override, regulatory disclosure, crisis communications, financial containment, legal privilege, and executive succession in incident conditions. The graph is a living artifact. It is ratified by the board, reviewed at least annually, and re-run whenever a material change occurs in leadership, jurisdiction, technology, or regulation.

6.2 Decision Latency

Decision Latency is the measured time, under instrumented conditions, from the moment sufficient information exists to authorize a decision to the moment the decision is authorized. It is measured in at least three regimes:

- Baseline latency: under non-adversarial conditions, with full attendance and normal business hours.
- Pressure latency: under simulated adversarial conditions, with partial attendance and compressed timelines.
- Peak latency: under conditions intended to simulate the worst 72 hours — multiple simultaneous decisions, degraded communication, incomplete telemetry.

A well-instrumented organization can report its latency across all three regimes for each decision class in its Authority Graph. An uninstrumented organization cannot report any of them and typically discovers its true peak latency during a live incident, at which point it is no longer a measurement — it is an outcome.

6.3 The Override Taxonomy

The Override Taxonomy classifies the acts by which a decision already in motion — by a human or by an automated system — can be halted, reversed, or redirected. The taxonomy distinguishes at minimum:

- Precautionary override: halting a system or action before consequence accrues, under explicit conservative bias.
- Corrective override: reversing a decision after partial consequence, with explicit acceptance of sunk cost.
- Emergency override: unilateral action by a named authority, bypassing normal concurrence, with post-hoc accountability.
- Authority escalation: invoking a higher decision body when the current holder cannot or will not decide.

The taxonomy is most useful when each class names, in advance, who may invoke it, under which conditions, with what notification obligations, and with what accountability consequences. The AI-era application of this primitive is addressed in §7.

6.4 The Maturity Rubric

The Maturity Rubric scores an organization's position on a five-level scale. The levels are intended to be adoptable as stated; refinements should preserve the level semantics.

Level	Name	Defining characteristic
L1	**Implicit**	Authority is assumed. No written map. Decisions happen by default or by seniority.
L2	**Declared**	Authority is written in policy. Not tested. Not measured. Often outdated.
L3	**Tested**	Authority is exercised in scheduled simulations. Gaps are identified but not always closed.
L4	**Instrumented**	Latency is measured. Override taxonomy is codified. Board reviews authority graph at least annually.
L5	**Compounding**	Decision authority is a continuous capability. Incidents improve it. AI boundaries evolve with the technology. Leadership transitions preserve it.

The median Fortune 2000 enterprise, in our observation, scores between L1 and L2. A small number of organizations — typically those that have survived a public incident and taken the subsequent learning seriously — reach L3. L4 is rare. L5 is, to our knowledge, not yet present at scale anywhere.

[PART III]

Application

[§7]

The AI-Era Extension

Every element of the framework changes when autonomous and semi-autonomous systems enter the decision chain. The Authority Graph acquires non-human nodes. Decision Latency can invert — the AI decides before humans are aware a decision was required. The Override Taxonomy must now specify machine-speed override paths. The Maturity Rubric cannot ignore AI boundaries without failing to describe the enterprise it is scoring.

7.1 The AI authority question, stated precisely

For every agentic system deployed inside or adjacent to high-consequence decision chains, an organization should be able to answer four questions without pause:

- What is this system authorized to decide on its own?
- What is it authorized to decide with human concurrence?
- What must it escalate to a named human authority, and how does the escalation channel function at the system's operating speed?
- Who is accountable — by name and role — when the system acts and the action proves wrong?

In 2026, the median Fortune 2000 organization deploying agentic AI can answer the first question partially, the second ambiguously, the third informally, and the fourth not at all. This is the single largest decision-authority gap visible in current enterprise reality.

> *Agentic AI does not create new decision authority questions. It forces existing ones to be answered at speeds the organization has never operated at.*

7.2 The AI Authority Charter

We propose the AI Authority Charter as a fixed-scope, reusable deliverable that articulates, for a given AI deployment, the four answers above. A Charter is board-ratifiable, regulator-defensible, and insurer-accessible. It is explicitly designed to be produced before the agentic system is deployed, revisited when the system's capability envelope changes, and cited when something goes wrong.

The EU AI Act's human-oversight provisions, emerging SEC disclosure expectations for material AI risk, the DPDP Act's accountability clauses, and insurer AI underwriting questionnaires are all moving,

at varying speeds, toward requiring artifacts that resemble the Charter. Organizations that build one ahead of the requirement acquire an option; organizations that wait will assemble one under deadline.

[§8]

Assessment Methodology

The framework is measurable. This section outlines the approach by which organizations — or independent parties acting on their behalf — can produce a repeatable, scored assessment of Decision Authority maturity. A fuller methodology specification will be published separately. This section states the assessment in its minimum viable form.

8.1 Scope

A Decision Authority Assessment covers, at minimum, the enterprise's top twenty-five decision classes across cyber incident response, AI-system override, regulatory disclosure, financial containment, and executive succession under incident conditions. Each class is assessed against the four primitives.

8.2 Instruments

The assessment uses four instruments: document review of policy and governance artifacts; structured interviews across the authority graph; an unannounced decision-pressure exercise targeting a representative decision class; and a review of the last three years of actual incident post-mortems where available.

8.3 Output

The assessment produces three outputs: a scored maturity position (L1–L5) per decision class, a set of gap findings with remediation priorities, and a baseline latency profile that becomes the organization's reference number for all future reassessments. The first two outputs are what executives typically ask for. The third is what matters most — because every subsequent year's position is measured against it.

8.4 Cadence

We recommend annual reassessment at minimum, with an intermediate reassessment after any of: a CEO, CISO, or General Counsel transition; a material M&A event; a new agentic AI deployment in a high-consequence domain; a regulatory change affecting authority structure; or a live incident of material scope.

[§9]

Adoption Pathway: Boards, Regulators, Insurers

A framework is only as valuable as the community that adopts it. The following adoption pathways are offered as guidance to the three stakeholder classes for whom this framework creates immediate option value.

9.1 For boards

We recommend that boards of organizations in regulated or high-consequence sectors commission a baseline Decision Authority Assessment in the current fiscal year, adopt the Maturity Rubric as a standing governance metric, and require the AI Authority Charter (§7.2) for every agentic AI deployment crossing a defined materiality threshold. The full operational burden falls below the board — but the discipline is established at the board.

9.2 For regulators

Supervisory expectations in financial services, critical infrastructure, healthcare, and AI deployment are converging toward demonstrable accountability for decisions made by both humans and autonomous systems under incident conditions. We offer the four primitives as open definitions suitable for incorporation into supervisory guidance, examination handbooks, and disclosure expectations. The Decision Authority Framework is neither proprietary to Unknowns Lab nor dependent on its continued stewardship. We expect — and welcome — versions of it to evolve inside regulators themselves.

9.3 For insurers

A scored Decision Authority position is a previously unavailable underwriting signal for cyber, technology errors-and-omissions, and directors-and-officers lines. We expect insurers piloting its use in 2026–2028 to achieve underwriting differentiation of material magnitude against competitors who price solely on controls and history. We offer the primitives as open definitions and are available to consult with insurers on instrumentation.

> *Categories are minted when enough of the market uses the same vocabulary to describe the same problem. This document is the vocabulary.*

[APPENDIX]

Appendices

[A]

Glossary

Authority Ambiguity — The condition in which, at the moment of decision, no role is certain to hold the authority to act.

Authority Graph — A structured, directed representation of an enterprise's decision rights, reversal paths, and post-hoc accountability.

Agentic Authority — The portion of the Authority Graph allocated to autonomous or semi-autonomous systems, with defined override and accountability paths.

Crisis Window — The interval, typically measured in hours, during which decisions compound exponentially in cost or benefit.

Decision Authority — The structured, measurable capacity of an organization to identify, assign, and exercise the right to decide, under adversarial conditions and with irreversible consequence.

Decision Layer — The architectural surface — organizational and increasingly technical — through which Decision Authority is exercised.

Decision Latency — The measured time from the moment sufficient information exists to authorize a decision to the moment it is authorized.

Maturity Rubric — A five-level scale (L1 Implicit through L5 Compounding) describing an organization's position on Decision Authority.

Override Taxonomy — A classification of the acts by which a decision in motion can be halted, reversed, or redirected.

Peacetime Structure — The assets — maps, charters, latency baselines — built before an incident that determine crisis-window performance.

[B]

Public Incident Patterns

The following incident patterns are drawn from publicly reported events between 2020 and 2025. They are presented not as case studies but as illustrations of how the framework's primitives describe failures already in the public record. Specific organizations are not named; the patterns are named.

Pattern. Supply-chain compromise at a software vendor with broad government and enterprise footprint.

> **Framework reading.** *Authority Ambiguity (the question of who could pause signing keys), compounded by Decision Latency spanning weeks.*

Pattern. Ransomware event at a regional critical-infrastructure operator, with public-facing outage.

> **Framework reading.** *Peacetime Structure absent: no pre-ratified authority for preemptive shutdown across multiple subsidiaries.*

Pattern. Large-scale data-exposure cascade at a consumer platform.

> **Framework reading.** *Override Taxonomy ambiguous: no named authority for voluntary disclosure ahead of regulator demand.*

Pattern. Commercial autonomous-system incident with public casualty and regulatory response.

> **Framework reading.** *Agentic Authority undefined: the human supervisory layer was assumed but not architected.*

Pattern. Multi-day outage triggered by a routine software update at a global platform vendor.

> **Framework reading.** *Decision Latency acute: the authority to halt a rolling update was distributed across too many teams to act at the cadence the failure required.*

Pattern. Regulatory disclosure dispute at a public financial institution following a cyber event.

> **Framework reading.** *Authority Graph missing the decision class entirely: whether and when to disclose materiality under new SEC rules was not a rehearsed decision.*

[C]

About Unknowns Lab & Call for Comment

About Unknowns Lab

Unknowns Lab is a principal-led practice focused on the Decision Layer of high-consequence enterprises. We work with a small number of organizations where the cost of decision failure is measured in billions, not millions. Our engagement model is deliberately constrained: confidentiality is absolute; client names and case studies are not published; the practice scales through codified methodology rather than headcount.

This framework is published separately from our client work. It is offered as a contribution to the broader community of boards, regulators, insurers, and practitioners who are independently arriving at the same observation: that the space between detection and decision is the failure surface that now determines outcomes.

Call for comment

Version 1.0 is open for comment through the next fiscal quarter. We invite critique from academic researchers, regulators, analysts, insurers, and practitioners with field observations that corroborate or challenge the primitives. Material feedback will be incorporated into version 1.1.

Correspondence can be directed through unknownslab.com. Requests for confidential briefings, academic collaboration, or regulatory engagement can be indicated in the same channel.

Citation

Suggested citation: Unknowns Lab. (2026). The Decision Authority Framework v1.0: A discipline for governing critical decisions in the cyber and AI era. Retrieved from unknownslab.com.

UNKNOWNS LAB

[SYS.00 / END OF FRAMEWORK.v1.0]

Your systems don't fail first. Your decisions do.