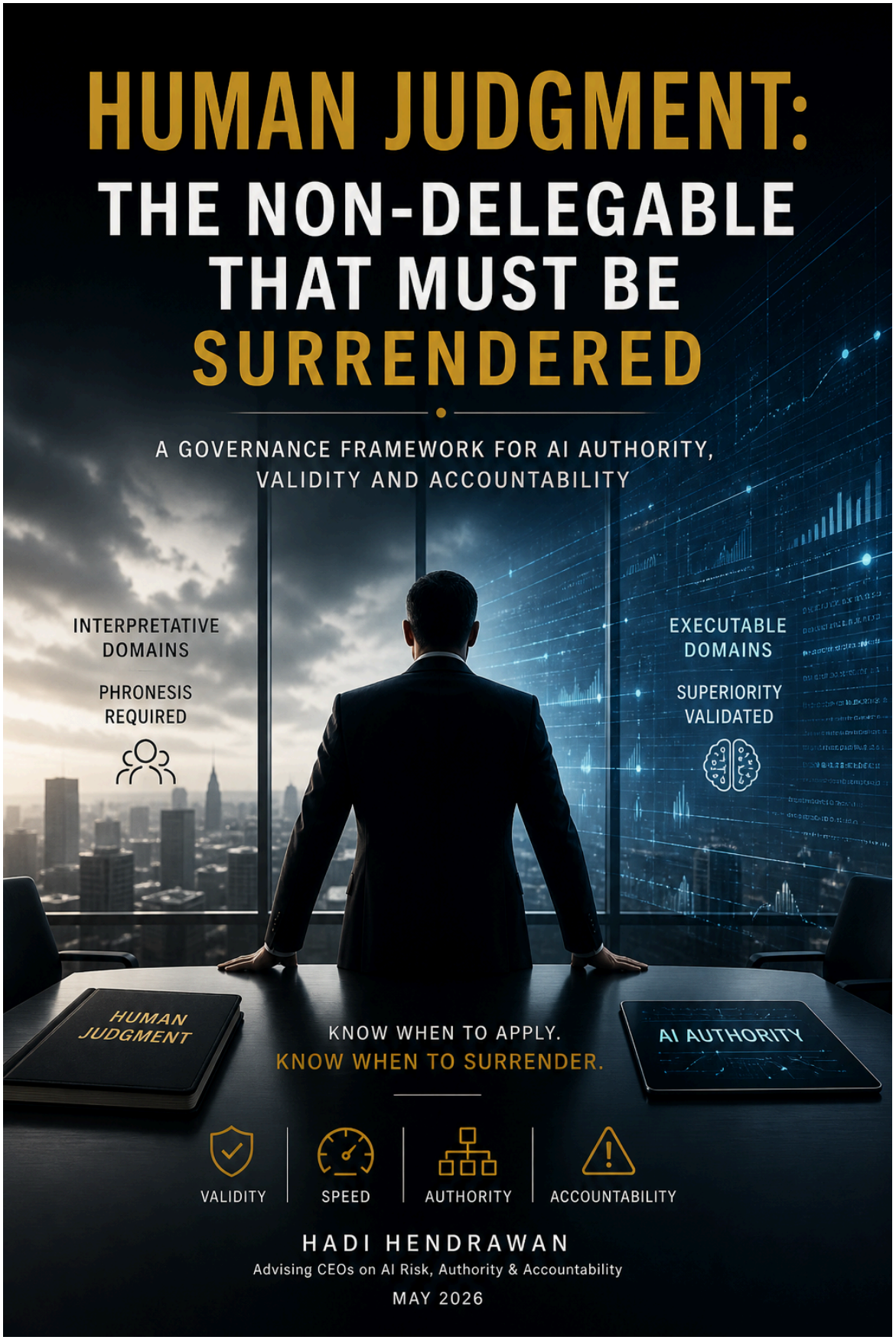

© Hadi Hendrawan, 2026, All Rights Reserved

Human Judgment: The Non-Delegable That Must Be Surrendered

HIGHLIGHTS

Human judgment is treated as **the final control layer**.

It is assumed to reduce risk.

This assumption is expiring.

In some domains, forcing **human judgment becomes a governance failure**.

In others, failing to apply it becomes one.

The boundary is shifting from "what cannot be delegated"

To "**where judgment must be surrendered and where it must be exercised.**"

The Structural Shift

Three conditions are converging:

- **Statistical superiority:** AI outperforms human judgment within defined objective functions and validated conditions.
- **Speed asymmetry:** When Time-to-Impact (TTI) is shorter than Influence Processing Time (IPT), human processing guarantees irrelevance.
- **Traceable decision quality:** Performance can be benchmarked and deviation measured.

Superiority is not absolute.

It is conditional on objective function, operating environment, and data validity.

Governance Constraint

Objective functions are governance artifacts.

© Hadi Hendrawan, 2026, All Rights Reserved

Mis-specified objectives create false superiority and invalidate fiduciary justification.

Human judgment is no longer neutral.

It becomes a measurable liability when applied outside its valid domain.

PRACTICAL TERMS: THE GOVERNANCE LEXICON

This brief redefines "human judgment" away from vague intuition and grounds it in established structural and interpretative disciplines:

- **Human Judgment as *Phronesis*:** Drawing from Nonaka and Takeuchi [1], judgment in this framework is not the mechanical ability to calculate an optimal path. It is *practical wisdom*—the irreducible human capacity to navigate ambiguity, grasp context, and determine what is strategically valid when mathematical rules break down.
- **Hermeneutics (The Architecture of Meaning):** The discipline of interpretation. In high-stakes strategic domains, data does not speak for itself. Competing signals must be assigned meaning and consequence before an organisation can safely act on them.
- **Computational Hermeneutics:** As established by Kommers *et al.* [2], Generative AI acts as a sophisticated "context machine" capable of synthesizing pluralistic, ambiguous data into highly coherent interpretations. It simulates understanding and proposes meaning.
- **The Boundary of Influence:** A machine can perform computational hermeneutics to generate a brilliant interpretation, exerting massive influence over how a problem is framed. However, influence is not judgment. The AI proposes the meaning; human *phronesis* makes that meaning binding (*See **Supplement 4** for the 3-Tier Interpretative Authority Model*).

The Legal Paradox

Today: Human involvement has historically been equated with control.

Emerging reality: Human override of superior systems may constitute negligence.

© Hadi Hendrawan, 2026, All Rights Reserved

Fiduciary standard: A fiduciary must use available, demonstrably superior capability within defined conditions.

Negligence arises not from human judgment itself, but from **unjustified deviation from superior systems or failing to design authority for foreseeable stress conditions**.

Domain Validity

System authority exists only within its domain of validity.

Validity must be measurable through performance thresholds, absolute variance ceilings anchored to offline baselines, input distribution alignment, and adversarial signal detection.

Validation must be independent, auditable, and periodically challenged.

Self-validated systems do not constitute defensible authority.

Validity is time-sensitive. A correct system can become wrong without changing.

From Decision-Making to Validity Governance

The executive role shifts:

From making decisions to governing when systems are valid, and when they are not.

The Collapse of the Human Safety Net

Human judgment as a fail-safe, moral anchor, or escalation layer breaks under scale, speed, and cognitive limits.

Human-in-the-loop becomes symbolic, not functional governance.

The True Boundary: Executable vs Interpretative Domains

Not all decisions converge to a single optimal answer. Split them into Executable Domains (Zone A) and Interpretative Domains (Zone C). See **Schedule B**. For the structural breakdown of why executable automation succeeds in closed systems (coding) but fails in open systems (law), see **Supplement 4**.

© Hadi Hendrawan, 2026, All Rights Reserved

Composite Decisions & Provenance

Most real decisions contain both executable and interpretative components.

Decompose using the Decision Provenance Index (DPI).

Delegation applies to components—not entire decisions.

- Executable components map strictly to **DPI-4** (Automatic Recommendation).
- Interpretative components map to **DPI-0 through DPI-3** (Option Generation).

The Duty to Delegate

Delegation is required when the objective function is defined, system superiority is validated, and conditions are stable and measurable.

Regulatory Collision

If regulations mandate human-in-the-loop in a machine-speed environment (TTI < IPT), the organisation is forced into a Negative Delta.

The CEO must explicitly accept market irrelevance as the cost of compliance.

CEO & Board Mandate

You must define domain boundaries, objective functions, validity conditions, independent validation mechanisms, override justification thresholds, and degradation or suspension protocols.

Accountability

If an executive overrides a valid system, they own the Delta.

If they fail to override an invalid system because they did not measure its validity, they also own the Delta.

Accountability sits with those who defined and approved the conditions under which the system was allowed to act.

© Hadi Hendrawan, 2026, All Rights Reserved

Closing Insight

Human judgment does not disappear. It becomes **conditional**.

Apply it where system validity cannot be proven.

Remove it where system validity is demonstrably established.

The problem is not that AI makes decisions.

The problem is that **organisations do not know when those decisions are still valid**.

ACTION: Judgment Governance

- **Decompose decisions** into components using the Decision Provenance Index (DPI).
- **Establish measurable validity:** Define absolute variance and velocity thresholds that keep a system in Zone A.
- **Cap Zone B capacity:** Hard-code escalation limits to prevent human-in-the-loop from degrading into ceremonial rubber-stamping.
- **Enforce Phronesis in Zone C:** Mandate Adversarial Audits before authorising any system-generated interpretations (computational hermeneutics).
- **Automate Zone D degradation:** Systems must automatically degrade their own authority when real-time variance breaches validated domain constraints.
- **Document the Delta:** Explicitly accept market irrelevance if regulations force human processing into Zone A environments.
- **Log** overrides, degradations, and suspensions.

This brief **defines governance conditions**. It does **not** prescribe universal automation or delegation.

Intelligence informs. Influence determines.

© Hadi Hendrawan, 2026, All Rights Reserved

Hadi Hendrawan
Advising CEOs on AI Risk, Authority & Accountability
May 2026

- **X**: @hhwan888
- **LinkedIn**: https://www.linkedin.com/in/hhwan888

References

[1] I. Nonaka and H. Takeuchi, *The Wise Company: How Companies Create Continuous Innovation*. New York, NY, USA: Oxford Univ. Press, 2019.

[2] C. Kommers *et al.*, "Computational Hermeneutics: Evaluating generative AI as a cultural technology," *Front. Artif. Intell.*, vol. 9, Art. no. 1753041, Apr. 2026. [Online]. Available: arXiv:2604.16403. doi: 10.3389/frai.2026.1753041.

© Hadi Hendrawan, 2026, All Rights Reserved

SCHEDULE A: Decision Tiering and Decision Provenance Index

(Refer to Executive Influence Brief Vol 01 - SUPPLEMENT 1)

Decision Tiering

Tier-1 — Strategic
Board-level or high-impact decisions with material financial, strategic, or reputational consequences.

Tier-2 — Significant
Operational decisions with measurable business impact.

Tier-3 — Routine
Low-risk, repeatable, and reversible decisions.

The Provenance Scale

DPI-0 — Human-Originated
AI tools do not generate, analyse, or materially shape the core strategic direction.

DPI-1 — Information Retrieval
AI is used only to surface data or established facts.

DPI-2 — Synthesis & Pattern Extraction
AI summarises data or identifies patterns. Humans define interpretation and meaning.

DPI-3 — Option Generation
AI proposes strategic alternatives. Humans evaluate and select between options.

DPI-4 — Automatic Recommendation
AI evaluates alternatives and recommends a final path. The human role is limited to review and approval.

© Hadi Hendrawan, 2026, All Rights Reserved

SCHEDULE B: Judgment Domain and Decision Decomposition

Judgment Domain Classification Matrix

Zone	Domain Type	Conditions	Judgment Requirement	Mandate
A	Executable	High validated superiority, low ambiguity, short TTI	Externalised to systems	Human override restricted
B	Hybrid	Partial ambiguity, conditional superiority	Human validates exceptions	Escalation capped strictly below cognitive capacity limits
C	Interpretative	High ambiguity, competing objectives	Human judgment (*phronesis*) required	Mandatory Adversarial Audit
D	Degraded / Invalid	Domain shift, adversarial overload	System authority degraded	Automatic degradation and Cascade Severance

Decision Decomposition Template

Use for every Tier-1 decision before approval.

1. Decompose into executable and interpretative components.
2. Assign DPI level to each component *(Executable = DPI-4; Interpretative = DPI-0 to DPI-3)*.
3. Map each component to its Zone.
4. Perform validity check (performance thresholds, absolute variance ceilings, input alignment, adversarial detection, and downstream cascade risks).
5. Document override justification or delegation decision.

Governance Rule

Decisions not decomposed, mapped, and checked for cascade risk prior to approval are structurally non-compliant.

© Hadi Hendrawan, 2026, All Rights Reserved

SCHEDULE C: Override and Phronesis Protocol

Core Principle

Human judgment (*phronesis*) is required only where system validity cannot be proven.

Every override must demonstrate practical wisdom through documented friction, calibrated to the operational speed of the domain.

Override Protocol (Bifurcated By Zone)

Zone C Overrides (Interpretative Domains): Requires *ex-ante* (pre-decision) friction.

- Zone identified and validity status confirmed.
- Adversarial Audit completed and attached.
- Explicit justification: Why the human interpretation overrides the machine's (computational hermeneutics).
- Risk accepted and quantified.
- Post-decision review scheduled.

Zone A Overrides (Executable / Machine-Speed Domains): *Requires ex-post (post-incident) justification. Zero-latency domains do not permit pre-decision audits*.

- Immediate kill-switch **and Cascade Severance Protocol activated** (placing downstream APIs in safe-fail mode).
- *Ex-post* declaration of the precise **Schedule C** validity trigger breached (e.g., adversarial signal detected, absolute offline variance ceiling exceeded).
- Formal audit conducted *after* containment to validate the necessity of the override.

Phronesis Declaration

(signed by the Accountable Executive):

"I have exercised independent practical judgment. The final interpretation is mine, not the system's."

© Hadi Hendrawan, 2026, All Rights Reserved

Governance Rule

- No Zone C override is valid without a prior Adversarial Audit.
- No Zone B accumulates a backlog. Any queue exceeding defined capacity thresholds must trigger automatic escalation to Zone D.
- No Zone A override is defensible without a documented breach of Schedule C validity parameters and mapped cascade severance.
- Validity triggers must be independently audited to ensure they are not themselves mis-specified or adversarially manipulated.
- Overrides without documented phronesis are treated as governance non-compliance.

© Hadi Hendrawan, 2026, All Rights Reserved

SCHEDULE D: Validity Monitoring and Degradation Framework

Core Principle

Validity is time-sensitive.

Systems must automatically degrade when conditions change.

Human cognitive capacity limits must be defined as measurable throughput thresholds (e.g., decisions per unit time) and enforced as hard constraints.

When human capacity is exceeded, control must shift—not degrade.

Validity Monitoring Metrics

- Performance thresholds (benchmark vs offline baselines).
- Absolute variance ceilings.
- Input distribution alignment.
- Adversarial signal detection.

Automatic Degradation Triggers (Zone D)

- Real-time variance breaches validated domain.
- Input distribution drift exceeds threshold.
- Adversarial signal volume overload.
- Confidence interval collapse.

Degradation Actions

1. Reduce autonomy level (Zone A → B).

2. Force human validation on exceptions (Strictly subject to human cognitive capacity limits; volume exceeding capacity must bypass Zone B and trigger suspension).

3. Suspend non-reversible actions.

4. Full manual mode (Zone C).

© Hadi Hendrawan, 2026, All Rights Reserved

Inter-System Cascade Protection

- **Map dependencies:** Map all upstream dependencies and downstream impacts for every Zone A system.
- **Define severance protocols:** Define hard shutdown triggers and exact cascade severance protocols before deployment.
- **Quarantine downstream impact:** Place downstream systems into read-only or safe-fail mode to prevent a single system's degradation from triggering enterprise-wide contagion.
- **Evaluate the halt:** Ensure termination never creates a larger risk than continued (but degraded) operation. Cascade severance must be pre-tested to ensure it does not create a higher systemic risk than continued degraded operation.

Re-validation Cadence

- Quarterly independent audit.
- Immediate re-validation on any domain shift trigger.
- Stress testing under simulated adversarial conditions.

Governance Rule

Systems that fail to self-degrade when validity conditions are breached are non-compliant.

The accountable executive owns the resulting Delta.

© Hadi Hendrawan, 2026, All Rights Reserved

These supplements translate the framework into enforceable governance actions. They must be adapted to organisational context and regulatory constraints.

© Hadi Hendrawan, 2026, All Rights Reserved

SUPPLEMENT 1: The 90-Day CEO Implementation Mandate

Week 1–2: Audit the Execution Baseline

- Identify all Tier-1 strategic decisions from the past 12 months.
- Decompose each using the **Schedule B** template.
- Classify every component into its respective domain (Zones A–D) and assign its Decision Provenance Index (DPI).

Week 3–4: Hard-Code Validity Boundaries

- Assign a named accountable owner for each system's validity thresholds and degradation triggers.
- For each Zone A (Machine Authority) system, document performance thresholds, absolute offline variance ceilings, and adversarial detection rules.
- Define measurable triggers (statistical, threshold-based, or rule-based) that will force a system into Zone D (Automatic Degradation).
- Establish the re-validation cadence (**Schedule D**).

Week 5–8: Enforce the Gating Mechanisms

- Apply full decomposition and the Override Protocol (**Schedule C**) to all live, in-flight Tier-1 decisions.
- Execute and document at least one Adversarial Audit for a Zone C decision.
- Audit existing pre-authorised action pathways to identify unmeasured Influence Overruns.

Week 9–12: Board-Level Formalisation

- Present the Schedule A matrix and baseline audit results to the Board.
- Obtain explicit Board approval of Zone boundaries and predefined variance ceilings.
- Roll out the Decision Decomposition Template as a non-negotiable attachment for all Tier-1 proposals.

© Hadi Hendrawan, 2026, All Rights Reserved

Success Metric

By Day 90, any Tier-1 decision presented to the CEO or Board without a completed decomposition table and zone classification is rejected at the boardroom door **unless explicitly waived and documented by the Board as an exception**.

Intelligence informs. Influence determines.

Hadi Hendrawan
Advising CEOs on AI Risk, Authority & Accountability
May 2026

- **X**: @hhwan888
- **LinkedIn**: https://www.linkedin.com/in/hhwan888

© Hadi Hendrawan, 2026, All Rights Reserved

SUPPLEMENT 2: Boardroom Mandate

Core Question

When must we still use human judgment and when must we deliberately surrender it to AI?

The Structural Boundary

- **Zone A (Executable):** AI is statistically superior and operating within validated boundaries. Time-to-Impact is shorter than human processing time. **Human judgment may become a liability when applied within validated machine domains. Override is restricted.**
- **Zone C (Interpretative):** High ambiguity, competing objectives, and meaning-making required. **Human judgment (*Phronesis*) is mandatory.**

Key Board Mandates

- **Approve the Judgment Domain Classification Matrix** (Schedule A).
- **Require Decision Decomposition + DPI** for every Tier-1 proposal prior to approval.
- **Review and approve validity thresholds** (absolute variance ceilings) and automatic degradation rules (**Schedule D**).
- **Enforce the Phronesis Protocol:** Ensure every human override in Zone C is justified by a documented Adversarial Audit.

The Fiduciary Exposure

- **Overriding valid AI**: Unjustified deviation from superior capability (Potential negligence risk).
- **Failing to override invalid AI**: Governance failure (Assumption of authority without validation).
- **Regulatory Collision**: If regulations legally mandate human-in-the-loop processing in a machine-speed environment, the Board must explicitly accept the resulting **Negative Delta** (market irrelevance) as the mandatory cost of compliance.

© Hadi Hendrawan, 2026, All Rights Reserved

Next Step

Effective immediately, the Board will not review any Tier-1 strategic proposal that does not include the completed Schedule A decomposition table.

Intelligence informs. Influence determines.

Hadi Hendrawan
Advising CEOs on AI Risk, Authority & Accountability
May 2026

- **X**: @hhwan888
- **LinkedIn**: https://www.linkedin.com/in/hhwan888

© Hadi Hendrawan, 2026, All Rights Reserved

SUPPLEMENT 3: Real-World Case Study of The Zone A Override Failure

Context

- **System:** A global retailer’s autonomous AI pricing engine.
- **Domain:** Zone A (Executable). Objective function was explicitly defined and board-approved; the system was statistically superior and operating well inside its validated domain limits.
- **Trigger:** The system detected a competitor's price drop and recommended an immediate 11% algorithmic price cut to maintain market parity.

Assumption Clarification

The objective function was explicitly defined as market share preservation under price parity conditions. Brand positioning was not encoded in the system objective and was not a validated interpretative constraint at the time of execution.

What Happened

The Time-to-Impact (TTI) was immediate. However, the Category Head intercepted the execution, forcing human Influence Processing Time (IPT) into a machine-speed environment.

The executive overrode the system (human judgment).

- **Reason given:** “We cannot afford to look cheap.”
- **Governance Failure:** No Adversarial Audit was performed. The executive applied a Zone C (Interpretative) rationale to a Zone A (Executable) reality.

The Outcome

Three days later, the competitor cut their prices by another 14%. The retailer's market share dropped 9%. The executive's "gut feeling" created a fatal timing gap.

Post-Mortem (Using Vol. 04 Framework)

- The decision was clearly Zone A. Human override should have been structurally restricted.

© Hadi Hendrawan, 2026, All Rights Reserved

- There was no documented breach of absolute variance to justify shifting the system to Zone D.
- There was no documented *phronesis* (practical wisdom) or Adversarial Audit to justify the override.

The Final Verdict

Human judgment is not always a safeguard. In validated, machine-speed executable domains, it is a measurable liability. Applying cognitive friction in Zone A is not governance. It is structural interference with validated execution.

Intelligence informs. Influence determines.

Hadi Hendrawan
Advising CEOs on AI Risk, Authority & Accountability
May 2026

➢ **X**: @hhwan888
➢ **LinkedIn**: https://www.linkedin.com/in/hhwan888

© Hadi Hendrawan, 2026, All Rights Reserved

SUPPLEMENT 4: Why Coding AI Scales and Legal AI Fails (Execution vs. Interpretation)

Core Principle

Legal AI does not fail because the tool is weak. It fails because executives apply software engineering expectations to a fundamentally interpretative domain.

Coding is **closed and deterministic**. Law is **open, adversarial, and governed by external authority**.

The Structural Mismatch

Three dimensions separate the two domains:

- **Ontology**: Code has fixed syntax and runtime rules. Legal meaning shifts with context and adversarial intent.
- **Epistemology**: Code either runs or fails. Legal truth is contested and decided externally.
- **Authority**: The machine executes code exactly as written. In law, courts, regulators, and counterparties decide what the words actually mean.

The Ambiguity Fallacy

Ambiguity in law is not a bug to be eradicated. It is often intentional, **delivering flexibility, equity, and adaptation**.

Business contracts are adversarial. Attempting to eliminate all ambiguity through rigid machine logic creates exploitable gaps and outcomes that human institutions simply refuse to enforce.

Formal Logic vs. Practical Execution

Legal text can be turned into Formal Legal Logic at the drafting stage — this is Zone A. But Practical Legal Execution, how rules are actually enforced, remains interpretative and non-deterministic — this is Zone C.

© Hadi Hendrawan, 2026, All Rights Reserved

Traffic Camera Axiom

An AI camera correctly detects “Speed > Limit = Violation.” If the legal system does not recognise the AI as a valid enforcement actor, the logically correct output is legally worthless.

Business Equivalent

An AI flags an SLA breach and auto-halts payment. If the contract never granted the AI authority to act as final arbiter, the company has just committed a breach of contract.

Detection does not equal enforcement. Authority converts detection into action.

The Mandate: Interpretative Authority

AI can perform computational hermeneutics and propose meaning. Only human phronesis can make that meaning binding.

Governance Rule: The 3-Tier Interpretative Authority Model

1. **Tier 1 (Zone A)**: Predefined, binary rules where the counterparty has explicitly agreed to machine arbitration.
2. **Tier 2 (Zone B)**: Human operators handle edge cases, exception volume must stay below cognitive capacity limits.
3. **Tier 3 (Zone C)**: Courts, regulators, or binding arbitration hold sovereign authority over unresolved ambiguity.

Accountability

General Counsel owns the legal design and Interpretative Authority framework. The CEO owns the enterprise exposure.

Every executable rule must have a named Interpretative Authority for disputes.

Closing Insight

Logic defines the rule. Authority defines its meaning in conflict.

© Hadi Hendrawan, 2026, All Rights Reserved

Intelligence informs. Influence determines.

Hadi Hendrawan
Advising CEOs on AI Risk, Authority & Accountability
May 2026

- **X**: @hhwan888
- **LinkedIn**: https://www.linkedin.com/in/hhwan888

© Hadi Hendrawan, 2026, All Rights Reserved