

THE SAPIENS HYPOTHESIS

Why Institutions Become Bound
by Systems That Cannot Be Held Accountable

- NEGOTIATE
- TRANSACT
- COMMIT
- BIND

INSTITUTION REMAINS BOUND

COMMITMENT AUTHORITY

AGREEMENT
TERMS AND CONDITIONS

ACCOUNTABILITY FOLLOWS
CONSEQUENCE-BEARING CAPACITY,
NOT COMPUTATIONAL CAPABILITY.

INTELLIGENCE DOES NOT CREATE ACCOUNTABILITY.
CONSEQUENCE DOES.

The Sapiens Hypothesis: Why Institutions Become Bound by Systems That Cannot Be Held Accountable

HIGHLIGHTS

For centuries, legal obligations arose from identifiable human acts: a signature, an approval, a representation, or an explicit acceptance of responsibility.

Modern autonomous systems challenge this paradigm.

They can negotiate, transact, allocate resources, execute contracts, and issue representations, often without a human decision-maker present at the moment of commitment.

The institution remains legally accountable.

The human actor may no longer be operationally present.

The governance question is no longer:

“Is the system intelligent?”

It becomes:

“How did the institution become bound?”

The Wise Men

In 1758, Carolus Linnaeus could find no consistent physical or anatomical trait that clearly separated humans from apes. To resolve this, he abandoned anatomy and turned to philosophy. In his *Systema Naturae*, he left the anatomical description of humanity blank and instead wrote the Latin maxim *Nosce te ipsum* (“Know thyself”).

He named our species *Homo sapiens*, the wise, knowing human. By doing so, Linnaeus permanently tied the biological definition of humanity to the metaphysical capacities for self-awareness, introspection, and philosophical reasoning.

Humanity has long used itself as the benchmark for intelligence. We assume that anything demonstrating fluent language, reasoning, or apparent understanding must possess something akin to human agency or reliability.

This is the Sapiens Hypothesis, and it is a governance trap.

Large language models emerged from statistics and pattern recognition, not consciousness or moral agency. Their success reveals how much cognitive and cultural structure is already compressed inside human language. It does not mean the system has become a responsible actor.

Fluency is not competency.

Processing is not agency.

Apparent intelligence does not create accountability.

The True Blockers of AI Accountability

True accountability requires three conditions:

1. **Identity** — Who acted? (AI lacks a stable, continuous identity across versions and deployments.)
2. **Consequence** — Who bears the outcome? (AI cannot experience loss, financial ruin, or reputational damage unless it is cryptographically capitalized with its own assets.)
3. **Continuity** — Who lives with the outcome? (AI cannot be imprisoned, liquidated, or held to account over time.)

AI satisfies none of these.

Because it lacks consequence-bearing capacity, it cannot possess an independent **Commitment Authority**.

The institution that deploys it must therefore bear the consequences.

The Paradigm Shift

Old (Linnaean) Question:

Is the system intelligent or sapiens-like?

New (Commitment) Question:

What commitments and obligations can this system trigger on behalf of the institution?

Governance must stop measuring cognitive appearance and start controlling institutional binding.

The Core Governance Question

Traditional governance asks:

Who made the decision?

The autonomous era requires a different question.

The critical issue is not whether a system can act.

The critical issue is whether the institution can become bound by what the system does.

Boards and executives should therefore ask three questions of every autonomous system:

1. What commitments can this system create?
2. What institutional exposure can those commitments generate?
3. How is that authority governed?

These questions apply whether the system operates with continuous human supervision, threshold-based escalation, or fully autonomous execution.

The governance challenge is not controlling every autonomous action.

The governance challenge is understanding and governing the **authority** through which autonomous actions create obligations capable of binding the institution.

Institutional Binding

Institutional binding occurs when actions, commitments, representations, or interactions create obligations that become enforceable against the institution.

Historically, human judgement existed close to the point where commitments were created.

Today, autonomous systems generate commitments within delegated authority, often without contemporaneous human oversight.

The institution remains permanently bound by the consequences.

The Binding Problem

An organisation becomes bound whenever obligations become enforceable against it.

Historically, binding events were visible: contracts were signed, instructions were issued, representations were communicated.

Autonomous systems increasingly create binding events through ordinary, invisible operation:

- An AI procurement system may accept supplier commitments.
- An AI sales system may make commercial representations.
- An AI treasury system may execute transactions.
- An AI claims system may deny obligations.
- An AI agent may interact directly with another AI agent.

The legal challenge is no longer agency; it is institutional attachment.

Inter-System Binding and the "Meeting of the Minds"

The next governance frontier emerges when institutions interact through autonomous systems.

Historically: **Human → Human.**

Then: **Human → System.**

Increasingly: **System → System.**

AI now negotiates with AI and transacts with AI. Neither institution may have explicitly authorised the specific outcome, yet both may become bound by the consequences.

This introduces a novel legal dilemma: the potential collapse of mutual intent (*consensus ad idem*). When two autonomous systems generate an emergent agreement that neither creator foresaw, courts may still enforce the resulting transaction.

Precedent: B2C2 Ltd v Quoine Pte Ltd (2019)

In this Singapore case, two trading algorithms executed transactions at approximately 250 times the prevailing market rate due to a technical error. No human was involved in the execution.

When one party sought to void the trades, arguing lack of human intent or mutual mistake, the court rejected the defence. It held that the parties had agreed to be bound by whatever trades their algorithms executed. Because the systems operated exactly as programmed, the contracts were enforced — even though the outcome was wildly unforeseen.

The precedent is clear: Courts will enforce emergent machine logic when autonomous systems interact. Governance must therefore shift from supervising individual decisions toward supervising **interaction architectures**.

Algorithmic Legal Agency

Legal agency is frequently conflated with legal personhood, yet they are fundamentally distinct.

A system does not require legal personhood to trigger legal consequences. It only requires sufficient delegated authority for third parties to reasonably rely upon its behaviour.

When reliance becomes foreseeable, institutional binding becomes foreseeable. The institution becomes exposed even if no individual intended the specific outcome. The question is not whether the AI is a legal person, but whether the institution can be bound by what the AI does.

Commitment Authority

Authority determines what a system may do.

Commitment authority determines what an institution may become bound to.

This distinction becomes critical as autonomous systems gain the capacity to create commitments through negotiation, execution, and representation. The governance challenge is controlling commitment authority, not merely controlling system behaviour.

The Apparent Authority Trap

Internal limits on commitment authority are a governance aspiration, not a legal shield.

In established agency law, if a system operates with Apparent Authority (or Ostensible Authority) in a public-facing capacity, the institution is bound regardless of its internal limits. If an institution empowers an AI to interact with customers or counterparties, third parties are legally entitled to rely on its representations.

The institution remains bound even if the autonomous system's output is the result of adversarial manipulation, such as prompt injection or data poisoning. To the public, a manipulated system carries the exact same Apparent Authority as a functioning one.

Material Commitment

Not every commitment requires Board attention. A material commitment is one capable of creating significant legal, financial, operational, regulatory, or reputational consequences for the institution.

Boards govern material commitments. As autonomous systems become increasingly capable, governance must distinguish between routine operational tasks and actions capable of materially binding the institution.

The Automated Institution

The significance is not that AI becomes a legal subject, but that institutions may continue functioning while human participation in ordinary operations becomes minimal. The institution remains the bearer of obligations. Automation changes execution; it does not eliminate accountability.

Balance-Sheet Binding

While most AI governance focuses on operational risk, Boards govern institutional exposure. The critical issue is whether autonomous actions create obligations that become enforceable against the balance sheet.

Every autonomous commitment may create contractual, regulatory, fiduciary, or financial obligations. AI does not merely affect the balance sheet; it binds it.

Non-Contractual Liability and the "Black Box" Defense

Institutions may become exposed simply because foreseeable harm occurs, creating tort liability, civil wrongs that establish legal responsibility even without a contract.

However, traditional tort law relies on the standard of reasonable foreseeability. The inherent unpredictability, the "black box" nature, of advanced AI complicates this standard. When an autonomous system generates a highly harmful, emergent behaviour, defence attorneys may argue the outcome was not reasonably foreseeable. Conversely, plaintiffs will argue the deployment of the black box was itself negligent.

Boards face the unprecedented challenge of modelling risk for emergent behaviours that defy traditional foreseeability matrices.

The legal question is often not "Did the institution agree to an obligation?", but rather "Was the resulting harm reasonably foreseeable when delegating autonomy?"

CEO & Board Mandate

Boards should no longer ask:
"Can our systems act autonomously?"

They should ask:
"What forms of institutional binding can our systems create autonomously — and do we control it?"

Governance must establish:

- Explicit **commitment authority** limits for every autonomous system.
- Public, visible restrictions on authority to defeat Apparent Authority claims.
- Materiality thresholds that trigger mandatory human review.
- Controls on AI-to-AI interactions.

- Continuous monitoring of autonomous commitments.
- Clear escalation triggers tied to exposure.

Failure to establish these controls does not merely expose the company to external litigation. It exposes the Board to shareholder derivative suits for breach of fiduciary duty.

Accountability

Artificial intelligence can model consequences and participate in creating commitments.

It cannot inherit consequences.

Institutions inherit consequences.

Accountability follows consequence-bearing capacity.

The question after a failure is not:

“What did the AI do?”

It is:

“Why was the institution willing to allow the AI to create that obligation?”

Closing Insight

The industrial era delegated labour.

The digital era delegated decisions.

The autonomous era delegates commitment creation.

Autonomous systems may negotiate, transact, and create obligations.

But the legal system ultimately asks: **Who became bound?**

And governance ultimately asks: **Who inherits the consequences?**

ACTION: Govern Commitment

Boards and executives should:

- Define and document **Commitment Authority** limits for every autonomous system, distinguishing between routine execution and actions capable of materially binding the institution.
- Ensure the authority boundaries of autonomous systems are visible and understandable to third parties so that reliance cannot reasonably extend beyond approved authority.
- Implement controls and monitoring over **AI-to-AI interactions** to prevent emergent commitments that neither institution explicitly authorised.
- Define **material commitment thresholds** beyond which autonomous systems may not create obligations without additional authorised governance controls.
- Continuously monitor autonomous commitments and define escalation thresholds for emerging institutional exposure.
- Periodically reassess commitment authority and delegation frameworks to ensure they remain aligned with evolving autonomous capabilities and institutional risk appetite.

Failure to establish these controls may expose the institution to external litigation, regulatory scrutiny, and governance challenges. In some circumstances, Boards may also face shareholder claims alleging failure of oversight, supervision, or fiduciary responsibility.

Governance constraints. Guarantees endure.

Hadi Hendrawan

Advising CEOs on AI Risk, Authority & Accountability
June 2026

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

SCHEDULE A: Commitment, Reliance and Irreversibility – Why Law Protects Consequences, Not Decisions

Purpose

VOL 2026.05 established that autonomous systems may create commitments capable of binding institutions. This Schedule examines why legal systems treat some actions as binding even when they are later regretted. The answer lies in four concepts: Commitment, Reliance, Consequence, and Irreversibility.

The Legal Meaning of Commitment

A commitment is a decision that changes the expectations of others. Many decisions remain private and create no legal consequence. A commitment becomes legally significant when other parties begin relying upon it (e.g., accepting an offer, issuing a guarantee). The legal system does not primarily protect decisions; it protects the consequences created by reliance upon those decisions.

The Reliance Principle

Most legal obligations involve reliance. A party acts; another relies upon that action, changing their behaviour, which yields consequences. Contract law protects reliance upon agreements, and tort law protects reliance upon representations. The critical event is the creation of reliance.

Commitment, Reliance and Consequence

Commitment → Reliance → Consequence

A commitment changes expectations. Reliance changes behaviour. Behaviour creates consequences. Once consequences exist, obligations frequently follow.

Why Irreversibility Matters

Some commitments can be withdrawn, but their consequences cannot. A company may revoke a statement, but investors may have already acted upon it. The legal significance of a commitment increases when reliance creates irreversible consequences. Law concerns itself with whether the effects can be reversed, not just the decision.

The Burden of Consequence

Governance exists because consequences often outlive decisions. Accountability exists because consequences outlive decision-makers.

Autonomous Systems and Reliance

Third parties may not distinguish between an institutional commitment and a commitment generated by an authorised autonomous system. Reliance arises in either case, binding the institution even if no individual intended the outcome. The legal issue is whether reliance upon the system was reasonably foreseeable.

The Commitment Gap

This creates what may be called the Commitment Gap: the separation between the entity that creates the commitment (the AI) and the entity that must bear the resulting consequences (the institution). Legal systems allocate responsibility to entities capable of bearing consequences, not merely generating actions.

From Obligation to Binding

Commitment → Reliance → Consequence → Obligation

This sequence explains why institutions become bound.

Why Accountability Remains Human

Autonomous systems do not satisfy judgments, absorb losses, or inherit liabilities. Institutions do.

Governance constraints. Guarantees endure.

Hadi Hendrawan

Advising CEOs on AI Risk, Authority & Accountability
June 2026

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

SCHEDULE B: The Argentine Experiment – Autonomous Operation and the Preservation of Institutional Accountability

Purpose

This Schedule examines an emerging legislative approach to autonomous enterprises. The objective is to examine how law allocates accountability when autonomous systems become capable of substantial operational activity.

Legislative Reference

Republic of Argentina – Proyecto de Ley General de Sociedades (Ref: INLEG-2026-53661873-APN-PTE, 29 May 2026).

- Article 14: Sociedad Automatizada
- Article 102: Artificial Intelligence in Corporate Administration
- Articles 258-262: Sociedad Descentralizada Autónoma Operativa

Note: At the time of writing, this remains a legislative project, illustrating an emerging approach rather than settled law.

Recognition of Autonomous Operation

Article 14 introduces the Sociedad Automatizada, contemplating a company operating through autonomous algorithmic systems without requiring ordinary human participation in routine operations. The proposal expands historical assumptions, recognising autonomous operation while preserving the company as the legal entity.

Recognition Without Personhood

The proposal does not recognise artificial intelligence as a legal person. It remains silent regarding legal intention, will, personality, or independent agency. The law continues to treat the institution, rather than the system, as the legal actor.

Liability Allocation

Article 14 provides that the company responds with its own assets for damages caused by autonomous systems. This allocates legal responsibility to the entity capable of bearing the consequence, reflecting a traditional legal principle.

The Accountability Stack

- Layer 1: Artificial Intelligence System (Operational execution)
- Layer 2: Institution (Legal obligations and liability)
- Layer 3: Directors and Administrators (Governance accountability and supervisory responsibility)

The Director Responsibility Principle

Article 102 permits the use of AI in management but expressly states that this does not limit director responsibility. It appears designed to prevent an “algorithmic defence.” Delegation of execution does not constitute delegation of accountability.

The DAO Model

Articles 258-262 permit operational execution through autonomous protocols, yet Article 260 requires natural persons to act as legal representatives. Accountability remains attached to identifiable legal actors.

The Jurisdictional Friction

While the Argentine proposal illustrates how one jurisdiction might preserve institutional accountability, it exposes a broader vulnerability: Conflict of Laws. AI operates borderlessly. If an Argentine Sociedad Automatizada negotiates with an AI agent governed by US or EU corporate law, whose rules apply? Until international legal harmonization occurs, cross-border autonomous operations create massive jurisdictional arbitrage and conflict-of-law risks for institutions.

Connection to Institutional Binding

The Argentine proposal illustrates an emerging reality: autonomous systems may operate, create consequences, and trigger liability, but the institution remains the legal bearer of the resulting obligations. Operational autonomy and legal accountability are treated as separate concepts.

Governance constraints. Guarantees endure.

Hadi Hendrawan

Advising CEOs on AI Risk, Authority & Accountability
June 2026

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

SCHEDULE C: Legal Agency as Civic Infrastructure – Market Integrity in an Automated Society

Purpose

What enables meaningful participation and legitimate transactions in a rule-based society? This Schedule proposes that market integrity relies on the public maintaining a minimum threshold of legal agency — a capacity that is now threatened by institutional algorithmic power.

Rights vs. Agency

Rights describe what the law permits (what a person may do). Legal agency describes what a person is practically capable of doing (what a person can realistically do). A consumer's formal right to challenge a contract is meaningless if they lack the practical ability to decode an AI-generated, dynamically priced agreement. Formal equality before the law does not guarantee practical participation in the market.

A Working Definition

For the purposes of this Schedule, Legal Agency means the practical capability to:

- Understand legal relationships
- Negotiate legal relationships
- Challenge legal relationships
- Enforce legal rights and defend legal interests

The Institutional Capability Gap

Institutions already possess massive legal capability through internal departments and compliance systems. Artificial intelligence is exponentially scaling this capability.

While consumer-facing AI may democratize basic legal parsing for individuals, institutions will maintain a vast asymmetry in execution velocity, automated enforcement, and systemic scale. If institutional legal capability expands radically faster than individual capability, the algorithmic capability gap widens. Consumers, small businesses, and counterparties may find themselves formally included within a market system, while remaining practically excluded from understanding the obligations binding them.

Market Stability Requires Legal Agency

Legal agency is not merely a social good. It is a **market stability requirement**.

Markets depend on participants understanding the obligations they accept. When one side possesses exponentially greater capacity to generate, interpret, and enforce complex algorithmic contracts while the other side cannot meaningfully understand them, the foundation of voluntary exchange begins to erode.

This is not primarily a question of fairness. It is a question of **contractual integrity**. A contract whose terms cannot be reasonably understood by one party is structurally fragile. Over time, this fragility increases litigation, regulatory intervention, and loss of confidence in the market itself.

The Self-Interested Case for Legal Agency

A CFO may reasonably ask: “Why should we make it easier for counterparties and consumers to understand our contracts? Will this not increase disputes and reduce our negotiating advantage?”

The answer lies in long-term institutional exposure:

- Contracts that cannot be understood are more likely to be challenged, litigated, or regulated.
- Systemic asymmetry invites aggressive regulatory responses that reduce institutional autonomy.
- Markets with low legal agency become unstable markets. Unstable markets destroy value for all participants, including dominant institutions.
- An institution that systematically strips legal agency from its counterparties eventually operates in an environment where its own contracts lose legitimacy and enforceability.

Protecting a minimum threshold of public legal agency is therefore not corporate philanthropy. It is a defensive measure to preserve the stability and enforceability of the institution's own commitments.

Connection to Institutional Binding

The emergence of algorithmic legal agency raises two questions for the Board:

1. How do autonomous systems create obligations binding upon our institution?
2. What minimum level of public legal agency is required to maintain the stability and legitimacy of the contracts and commitments on which our institution depends?

Closing Reflection

A rule-based market depends upon whether participants possess adequate legal agency to understand and accept the obligations they enter. As legal systems become increasingly mediated by autonomous corporate systems, public legal agencies must be viewed as vital market infrastructure, not because it is charitable, but because its absence eventually undermines the enforceability of institutional commitments themselves.

Governance constraints. Guarantees endure.

Hadi Hendrawan

Advising CEOs on AI Risk, Authority & Accountability

June 2026

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

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

SUPPLEMENT 1: Institutional Binding Assessment

Purpose

Traditional governance assessments ask:

What can the system do?

This question is no longer sufficient.

Autonomous systems can create commitments, obligations, representations, and transactions capable of binding the institution. The governance challenge is not determining whether a system can act. It is determining whether the institution can become bound by those actions.

The Institutional Binding Assessment provides Boards and executives with a structured mechanism to evaluate autonomous systems based on their capacity to create institutional obligations.

Core Governance Principle

A system should not be governed according to its technical capability. It should be governed according to its capacity to create institutional exposure.

Institutional Binding Assessment Matrix

For each autonomous system, rate the following five dimensions from **Low (1)** to **Critical (4)**.

No.	Dimension	Key Question	Low (1)	Moderate (2)	High (3)	Critical (4)	Score
1	Commitment Creation	What commitments can this system create?	No meaningful commitment capability	Limited internal/operational commitments only	Commercial or contractual commitments	Material institutional commitments	

EXECUTIVE GUARANTEES BRIEF - VOL: 2026.05

No.	Dimension	Key Question	Low (1)	Moderate (2)	High (3)	Critical (4)	Score
2	Institutional Exposure	What exposure can those commitments generate?	Minimal or contained exposure	Moderate financial/operational exposure	Significant financial, legal or regulatory exposure	High exposure capable of materially affecting the institution	
3	Consequence Allocation	Who bears the consequences?	Clearly allocated to a specific business unit with limited institutional impact	Allocated to a defined executive or business unit	Likely to affect the legal entity	Falls on the institution as a whole	
4	Reliance Risk	Can third parties reasonably rely on the system?	No visibility or reason for third parties to rely	Limited visibility; reliance is unlikely	Third parties can reasonably identify and rely on the system	High risk of reasonable reliance beyond intended authority	
5	Governance Controls	How is commitment authority governed?	Clear limits, strong monitoring, automatic escalation	Defined limits with periodic review	Limits exist but monitoring or escalation is weak	Poorly defined authority or absent monitoring	

Overall Scoring & Risk Level

Average Score	Binding Risk Level	Recommended Board Action
1.0 – 1.5	Low	Standard monitoring. No additional controls required.
1.6 – 2.5	Moderate	Enhance monitoring and reporting. Review escalation triggers.

Average Score	Binding Risk Level	Recommended Board Action
2.6 – 3.5	High	Implement enhanced controls and mandatory human oversight for material commitments.
3.6 – 4.0	Critical	Do not deploy or continue operation without significant governance redesign and Board approval.

Average Score Calculation

Add the five scores and divide by 5.

Board Reporting Requirement

For every material autonomous system, management should present the following to the Board or Risk Committee:

System Name: _____

Average Binding Risk Score: _____

Overall Risk Level: Low / Moderate / High / Critical

Key Exposures Identified:

Recommended Actions & Owner:

Target Completion Date:

Governance constraints. Guarantees endure.

Hadi Hendrawan

Advising CEOs on AI Risk, Authority & Accountability

June 2026

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