

HUMAN AGENCY PRESERVATION INFRASTRUCTURE

A Theory of Agency, Governance, and Restoration
for the Age of Agentic Systems

Michael Bower

Human Agency Preservation Infrastructure

Working Thesis Paper - v0.1

May 2026

Core thesis: governance is the product of agency preservation. Systems become trustworthy only when humans retain meaningful participation before, during, and after consequential action.

Status: working research manuscript, not peer reviewed. This paper synthesizes the HAPI foundation papers into a single thesis architecture for future refinement, implementation, and critique.

Abstract

Human Agency Preservation Infrastructure (HAPI) is a theory, governance model, and infrastructure framework for preserving meaningful human participation in systems that use automation, institutions, bureaucracy, or delegated artificial agency. The central claim is that agency is not merely freedom of choice. Agency is meaningful participation under conditions of capacity, authority, clarity, refusal, revision, memory, and accountability. When these conditions are preserved, governance becomes real. When they are removed, governance becomes theater.

This master thesis integrates the HAPI foundation papers: Theory of Agency, Agency Loss, Agency Restoration, Governance as Agency Preservation, False Gates and True Gates, Human Agency Infrastructure for Agentic AI, Agency Theater, the HAPI Audit Model, Sin as Agency Loss, Agency Capture, Conscience and Refusal, the HAPI Constitution, Institutional Agency Loss, Restorative Governance, Measuring Agency, Agency Preservation in Healthcare, Work and Agency in the Age of AI, and Dependency Capture.

The paper argues that modern systems often preserve human presence while removing human participation. People may remain visible in workflows, meetings, approvals, religious systems, hospitals, workplaces, or AI oversight loops, while losing the ability to understand, refuse, revise, contest, remember, or remain accountable in a meaningful way. HAPI names this as agency loss and proposes a restoration-centered approach to governance. The goal is not to oppose institutions or automation. The goal is to design systems that amplify human agency without capturing it.

Keywords

human agency, agency preservation, agency restoration, governance, agentic AI, delegated agency, false gates, true gates, agency theater, institutional agency loss, dependency capture, healthcare agency, work and AI, accountability, refusal, memory

Table of Contents

1. Introduction
 2. The Foundational Claim
 3. Agency as Meaningful Participation
 4. Agency Loss
 5. Agency Capture and Dependency Capture
 6. Agency Restoration
 7. Governance as the Product of Agency Preservation
 8. True Gates and False Gates
 9. Agency Theater
 10. Conscience and Refusal
 11. Restorative Governance
 12. Measuring Agency
 13. HAPI for Agentic AI
 14. Institutional, Healthcare, and Work Applications
 15. The HAPI Constitution
 16. Implementation Path
 17. Risks and Failure Modes
 18. Research Agenda
 19. Conclusion
- Appendix A: Core Equations
- Appendix B: HAPI Paper Stack
- Appendix C: Glossary

1. Introduction

Human systems are increasingly mediated by institutions, automation, platforms, artificial intelligence, compliance procedures, medical systems, religious authority structures, and digital workflows. These systems can provide support, memory, scale, coordination, and safety. They can also quietly remove the conditions that allow humans to participate meaningfully in their own lives.

The central concern of HAPI is not that systems are always bad. The concern is that systems often become powerful enough to act around humans while still claiming that humans are involved. A human may approve without understanding, consent without real alternatives, follow a process without voice, remain accountable without authority, or depend on a system that was originally meant to restore capacity.

HAPI exists to restore human agency where systems have reduced people to rubber stamps.

This paper presents HAPI as the parent theory beneath a growing set of governance tools, audit methods, theological reflections, and agentic AI infrastructure. Its starting point is simple: if agency is preserved, governance becomes real. If agency is lost, governance becomes theater.

2. The Foundational Claim

The foundational claim of HAPI is that governance must be built from agency preservation rather than imposed as external control. Traditional governance often begins with policies, approvals, dashboards, rules, and audits. These are useful only when the underlying human conditions remain alive. If the human cannot understand, refuse, revise, contest, remember, or act with real authority, then governance is symbolic rather than real.

HAPI reverses the order. It asks first whether the person or group can meaningfully participate. Governance then becomes the structure that protects that participation over time.

Agency preservation is the mission. Governance is the product. Infrastructure is the implementation.

3. Agency as Meaningful Participation

Agency is the living capacity of a person or group to understand, choose, refuse, revise, act, remember, and remain accountable. It is not merely the existence of options. It is the condition under which choice can be meaningful.

A person may have many technical options while lacking agency. They may be overloaded, coerced, misled, isolated, punished for refusal, denied context, denied memory, or held responsible for consequences they could not meaningfully influence. In that case, the system may claim that the person chose, while the deeper conditions of agency were absent.

3.1 Core Agency Functions

- **Understanding:** the person can grasp what is happening and why it matters.
- **Discernment:** the person can evaluate whether an action is safe, true, good, aligned, or harmful.

- **Choice:** the person can select among real alternatives.
- **Refusal:** the person can say no before consequence.
- **Revision:** the person can change the proposed action before it binds.
- **Action:** the person can carry out or authorize action.
- **Memory:** the person or system can preserve what happened and why.
- **Accountability:** the person can own, contest, repair, and learn from consequences.

4. Agency Loss

Agency loss occurs when a person remains present in a system but loses meaningful participation. This is one of the most important diagnostic concepts in HAPI. It explains why people can appear free while feeling trapped, why institutions can appear governed while operating incoherently, and why AI oversight can appear safe while humans become rubber stamps.

Agency loss is often gradual. It can begin as convenience, compliance, safety, efficiency, or support. Over time, the system may remove context, compress choices, increase dependency, punish refusal, weaken memory, or separate authority from accountability.

4.1 Common Patterns of Agency Loss

- **Overload:** too much pressure, information, or urgency for clear participation.
- **Coercion:** choice exists technically, but refusal carries destructive cost.
- **Automation bypass:** systems act faster than human judgment can participate.
- **Rubber-stamp participation:** humans approve without context, time, or confidence to refuse.
- **Authority without control:** people are held responsible for outcomes they could not direct.
- **Control without accountability:** systems or leaders direct outcomes while avoiding responsibility.
- **Memory loss:** the system cannot preserve why something happened or who approved it.
- **Policy-reality split:** written values differ from actual operating behavior.
- **Meaning distortion:** human intent is transformed into something easier to process, measure, automate, or exploit.

5. Agency Capture and Dependency Capture

Agency capture is different from direct agency loss. In agency capture, the person remains active, but their participation is routed through a system that increasingly owns the conditions of action. The person can still move, decide, approve, or speak, but only through channels that shape, restrict, or reinterpret their agency.

Dependency capture is a specific form of agency capture. It occurs when a support system stops restoring capacity and starts becoming necessary for action. The tool, institution, community, platform, assistant, or authority structure becomes the condition of agency rather than the scaffold that helps agency recover.

Support restores the person. Capture replaces the person while calling itself support.

6. Agency Restoration

Agency restoration is the process of rebuilding the conditions that allow meaningful participation. It is not the same as giving unlimited control. It means restoring rightful participation at the correct point in the system.

Agency restoration asks what was lost and what must be rebuilt. Did the person lose clarity? Authority? Refusal? Memory? Capacity? Social support? Contestability? The restoration path depends on the broken function.

6.1 Restoration Sequence

Restoration Stage	Question	Result
Clarify	Can the person see what is happening?	Restored understanding
Stabilize	Does the person have enough capacity to participate?	Restored readiness
Authorize	Does the person have rightful say?	Restored authority
Protect refusal	Can the person stop or redirect action?	Restored boundary
Restore memory	Can the person know what happened and why?	Restored accountability
Support action	Can the person act without capture?	Restored participation

7. Governance as the Product of Agency Preservation

Governance is often treated as a control layer placed on top of systems. HAPI treats governance as the product of preserved agency. Policies, reviews, dashboards, and approval chains matter only when they protect live human authority, memory, refusal, revision, and accountability.

When agency is preserved, governance becomes real because people can still affect outcomes. When agency is lost, governance becomes theater because procedures remain while meaningful participation disappears.

If agency is preserved, governance becomes real. If agency is lost, governance becomes theater.

8. True Gates and False Gates

A gate is a boundary where passage is governed. HAPI distinguishes between true gates and false gates. A true gate governs rightful passage under valid authority, clear conditions, and accountable proof. A false gate blocks or permits passage without rightful authority, often capturing agency while claiming to protect it.

This concept applies across AI systems, institutions, religious communities, healthcare, education, families, and workplaces. A true gate preserves agency at the boundary where consequence begins. A

false gate stands where it does not belong and makes itself the condition of belonging, action, conscience, or legitimacy.

- **True gate:** preserves rightful authority, refusal, revision, memory, and accountability.
- **False gate:** captures access, identity, belonging, conscience, or consequence without rightful authority.

9. Agency Theater

Agency theater is the appearance of agency without the conditions that make agency real. It appears when systems ask for approval but remove refusal, ask for consent but hide alternatives, ask for participation after the outcome is already determined, or hold people accountable for consequences they could not direct.

Agency theater is especially dangerous because it can satisfy audits, leaders, dashboards, and public expectations while leaving the human agency problem untouched. It turns the symbol of participation into a substitute for participation itself.

- fake consent
- fake approval
- fake oversight
- fake choice
- fake accountability
- fake transparency
- fake human-in-the-loop safety

10. Conscience and Refusal

Refusal is one of the clearest signs that agency is real. If a person cannot say no, question, contest, or follow conscience without social, institutional, spiritual, economic, or procedural destruction, then agency is conditional.

Conscience is the inner boundary where a person must be able to stand before truth. Institutions may teach, persuade, correct, or set boundaries, but when they replace conscience or punish honest discernment as disloyalty, they risk becoming false gates.

Approval without refusal is not agency. Obedience without conscience is not alignment.

11. Restorative Governance

Restorative governance is governance aimed at repairing agency rather than merely enforcing compliance. It asks not only whether a rule was broken, but what agency condition failed. Was the person overloaded? Was authority unclear? Was memory missing? Was refusal unsafe? Was the system designed for rubber-stamping?

The purpose of restorative governance is not permissiveness. It can still set boundaries and consequences. But its goal is coherence restoration rather than symbolic punishment. It rebuilds the conditions under which the person, team, or institution can act responsibly again.

12. Measuring Agency

HAPI must be more than language. It requires evaluation criteria. Agency cannot be reduced to a single score without danger, but agency conditions can be assessed through observable dimensions.

12.1 Agency Evaluation Dimensions

Dimension	Audit Question
Clarity	Can the person understand what is happening?
Authority	Does the person have rightful power to affect the outcome?
Refusal	Can the person stop, delay, or redirect action before consequence?
Revision	Can the person change the proposal before it binds?
Memory	Can the person or system preserve what happened and why?
Contestability	Can outcomes be challenged or corrected?
Capacity	Does the person have enough energy, time, knowledge, and support?
Dependency risk	Does support restore agency or replace it?
Accountability integrity	Are authority and responsibility kept together?

13. HAPI for Agentic AI

Agentic AI creates a new agency problem because AI agents can plan, call tools, update systems, initiate transactions, and affect operational reality. They are not moral persons, but they are also not passive tools. They are delegated operational agency.

This means AI governance must preserve the human agency chain before, during, and after action. PGDL challenges the proposal before it becomes actionable. AAG authorizes, revises, escalates, or blocks the action. Runtime Binding prevents execution drift. Receipts preserve proof and memory. Continuity Findings evaluate whether governance remained coherent across time.

AI should amplify human agency, not outrun it.

14. Institutional, Healthcare, and Work Applications

HAPI is not limited to AI. It applies wherever systems shape human participation. Institutions can preserve people while removing agency. Healthcare can overwhelm patients at the moment they

need agency most. Workplaces can turn experts into procedural approvers or disposable outputs. Religious communities can guide conscience or capture it.

14.1 Healthcare

Healthcare agency preservation requires clarity, consent, memory, treatment understanding, family support, emotional regulation, and the right to participate in decisions under fear and complexity. The goal is not to replace clinicians. The goal is to prevent the patient from becoming passive inside a system that moves faster than they can understand.

14.2 Work

In the age of AI, the goal should not be simple labor replacement. HAPI points toward expert-governed delegation: workers and professionals govern agent teams while retaining judgment, refusal, accountability, and domain authority.

14.3 Institutions

Institutions should be audited for where they preserve or strip agency: who can refuse, who can contest, who remembers, who owns risk, who has authority, and who is made responsible without control.

15. The HAPI Constitution

The HAPI Constitution summarizes the principles that should guide agency-preserving systems.

- No responsibility without authority.
- No approval without refusal.
- No automation without memory.
- No governance without contestability.
- No support that creates dependency.
- No human-in-the-loop theater.
- No institution between conscience and truth.
- No action without accountable proof.
- No scale without preserved agency.
- No restoration without participation.

16. Implementation Path

The implementation path for HAPI should move from theory to audit to infrastructure. The theory defines agency, agency loss, and agency restoration. The audit model identifies where agency has collapsed. The governance model restores authority, refusal, memory, and accountability. The infrastructure layer makes these conditions enforceable in real systems.

16.1 Layer Map

- Theory of Agency: defines what must be preserved.
- Agency Loss and Capture: diagnose how agency collapses.
- Agency Restoration: defines repair sequence.

- Governance Model: turns restored agency into durable structure.
- Audit Model: measures agency preservation across systems.
- Infrastructure: implements gates, bindings, receipts, and continuity checks.
- Applications: healthcare, work, institutions, AI systems, personal agency, and community systems.

17. Risks and Failure Modes

HAPI must guard against becoming the kind of system it critiques. Agency preservation infrastructure can itself become a false gate if it captures agency, over-gates normal action, creates dependency, turns human alignment into a compliance score, or gives auditors power without accountability.

The safeguard is to keep HAPI restoration-centered. The purpose of the gate is not permanent control. The purpose of the gate is restored discernment, restored capacity, restored accountability, and safer delegation.

- over-gating
- audit theater
- dependency on the HAPI process
- false precision through scoring
- bureaucratic capture
- authority ambiguity
- restoration replaced by punishment

18. Research Agenda

The next stage of HAPI should test the theory through practical cases. These may include AI agent governance, patient decision support, workplace process audits, institutional agency maps, safety programs, and personal agency infrastructure. The research question should remain consistent: did the system increase meaningful human participation, or did it only increase control, speed, or compliance?

- Develop an agency audit rubric.
- Create case studies across healthcare, AI, work, and institutions.
- Build tools that preserve refusal, memory, and accountable proof.
- Test whether governance interventions reduce rubber-stamping.
- Measure whether users experience increased clarity, control, and capacity.
- Evaluate dependency risk in AI assistant and institutional support systems.

19. Conclusion

HAPI begins with a simple claim: human agency is the load-bearing condition beneath real governance. A system is not truly governed because it has rules, dashboards, approvals, policies, or reports. It is governed when people can still understand, refuse, revise, act, remember, contest, and remain accountable.

The modern world is increasingly capable of preserving human presence while removing human agency. HAPI names that failure and offers a restoration path. It gives a language for diagnosing agency loss, recognizing agency theater, distinguishing true gates from false gates, measuring participation conditions, and building infrastructure that preserves human authority before consequence.

Governance is the product of agency preservation.

The purpose of HAPI is not to slow the world down for its own sake. It is to make powerful systems safer, more truthful, more accountable, and more human by preserving the agency that makes governance real.

Appendix A: Core Equations

Agency = Capacity x Authority x Clarity x Refusal x Memory

Governance Reality = Preserved Agency + Accountable Proof + Contestable Consequence

Agency Theater = Human Presence - Meaningful Participation

Dependency Capture = Support - Restoration + Increasing Necessity

True Gate = Rightful Boundary + Live Refusal + Accountable Proof

Appendix B: HAPI Paper Stack

- The Theory of Agency
- Agency Loss
- Agency Restoration
- Governance as the Product of Agency Preservation
- False Gates and True Gates
- The Biblical Architecture of Agency
- Human Agency Infrastructure for Agentic AI
- Agency Theater
- The HAPI Audit Model
- Sin as Agency Loss
- Agency Capture
- Conscience and Refusal
- The HAPI Constitution
- Institutional Agency Loss
- Restorative Governance
- Measuring Agency
- Agency Preservation in Healthcare
- Work and Agency in the Age of AI
- Dependency Capture
- HAPI Master Thesis

Appendix C: Glossary

Term	Definition
Agency	Meaningful human participation under conditions of capacity, authority, clarity, refusal, memory, and accountability.
Agency loss	The reduction of meaningful participation while human presence remains.
Agency capture	The conversion of human participation into dependence, compliance, or system-routed action.
Agency theater	The appearance of human participation without the conditions that make participation real.
Dependency capture	A support system becoming necessary for action instead of restoring the ability to act.
True gate	A boundary that governs rightful passage while preserving agency.
False gate	A boundary that captures agency or claims authority it does not rightfully hold.
Restorative governance	Governance that repairs agency conditions rather than merely enforcing compliance.
Delegated operational agency	Action carried out by an AI agent or system under borrowed human or organizational authority.
Human Agency Preservation Infrastructure	A theory, governance model, and infrastructure framework for preserving human agency in powerful systems.