WORK AND AGENCY IN THE AGE OF AI

From Labor Replacement to Expert-Governed Delegation

A Human Agency Preservation Infrastructure (HAPI) Working Thesis Paper

Author: Michael Bower

Project: Human Agency Preservation Infrastructure

Version: v0.1 Working Draft

Status: Conceptual thesis, not peer reviewed

Core thesis: The future of work should not be framed only as labor replacement. It should be framed as the design of agency-preserving delegation systems in which human expertise remains live at the point of judgment, refusal, revision, and accountability.

Abstract

This working thesis applies Human Agency Preservation Infrastructure (HAPI) to the future of work under agentic AI. The paper argues that the dominant public frame of AI and work - replacement versus preservation of existing jobs - is incomplete. The deeper question is whether human agency remains live as tasks, decisions, and workflows are delegated to increasingly capable systems. Agentic AI may remove labor, but it may also remove judgment, refusal, memory, skill development, accountability, and the lived meaning of participation. The proposed alternative is expert-governed delegation: a work model in which AI agents handle bounded execution while human experts retain authority over interpretation, risk, consequence, and value. The paper distinguishes task automation from agency displacement, explains the risk of rubber-stamp employment, and defines a HAPI-based model for designing AI-enabled work that amplifies rather than erodes human participation.

Keywords: human agency, future of work, agentic AI, expert-governed delegation, automation, governance, labor, HAPI, human oversight, workplace agency.

1. Introduction

AI is often discussed as a labor-replacement technology. The central question becomes which jobs will disappear, which jobs will survive, and which workers will be displaced. That framing is important, but it is not deep enough. Work is not only a bundle of tasks. Work is also a structure of agency.

Through work, people understand problems, exercise skill, make judgments, accept responsibility, contribute to others, participate in a community, and build identity. When AI enters the workplace, it does not merely automate tasks. It can reorganize who understands the system, who has authority, who can refuse, who remembers what happened, who is accountable, and who still meaningfully participates.

> *The future-of-work question is not only: What labor can AI replace? It is also: What human agency must not be removed when labor is delegated?*

HAPI approaches AI and work from this deeper question. The purpose is not to defend every existing job structure or reject automation. The purpose is to preserve meaningful human agency as automation increases.

2. Work as a Structure of Agency

Work gives people more than income. It gives people operational contact with reality. A worker sees conditions, interprets problems, acts through skill, evaluates outcomes, learns from errors, and develops judgment over time. This is why work can be formative. It is one way agency becomes embodied.

A workplace preserves agency when workers have some meaningful relationship to the conditions, decisions, tools, consequences, and memory of their work. A workplace erodes agency when humans remain present but become disconnected from the actual judgment path.

- A person can be employed while losing authority.
- A person can approve work they no longer understand.
- A person can be held accountable for systems they cannot change.
- A person can be productive while becoming less capable over time.
- A person can supervise automation while slowly losing the skill to challenge it.

The HAPI claim is that work should be evaluated not only by productivity and cost, but also by whether it preserves human capacity, authority, refusal, memory, skill, and accountability.

3. Task Automation Versus Agency Displacement

Not all automation is harmful. Some automation restores agency by removing exhausting, repetitive, or low-value burdens. Other automation displaces agency by removing the person from the meaningful decision path.

Automation Pattern	Agency Effect	Example
Burden removal	Can restore agency by freeing attention and energy.	AI summarizes routine documents so a professional can focus on judgment.
Skill extension	Can amplify agency by helping the person act beyond prior limits.	A technician uses AI diagnostics while retaining authority over repair decisions.
Judgment replacement	Can erode agency by moving interpretation away from the human.	A system classifies high-risk cases and the worker only clicks approve.
Authority bypass	Can strip agency by acting before human refusal is possible.	An agent sends, deletes, purchases, or escalates before review.
Memory capture	Can weaken accountability by hiding why action occurred.	A platform records outcomes but not the reasoning, objections, or approvals.

The important distinction is not whether AI performs work. The important distinction is whether the human remains meaningfully situated in the chain of understanding, refusal, revision, action, and accountability.

4. The Rubber-Stamp Worker

One of the most dangerous future-of-work patterns is the rubber-stamp worker. This is a person who remains officially responsible while the real decision has already been shaped by automation, policy, metrics, interface design, or institutional pressure.

The rubber-stamp worker is present but not meaningfully participating. They may be told they are supervising AI, but they lack the time, context, authority, or confidence to challenge the system.

- The system proposes the action.
- The worker sees a simplified summary.
- The interface makes approval easy and refusal costly.

- The organization rewards speed and throughput.
- The worker approves without meaningful agency.
- If harm occurs, the worker is still blamed as the human in the loop.

> *A human in the loop is not the same as human agency in the loop.*

HAPI treats this as a central workplace failure mode. AI-enabled work must not preserve human liability while removing human power.

5. Expert-Governed Delegation

The alternative to both labor replacement and rubber-stamp oversight is expert-governed delegation. In this model, AI agents perform bounded work under the authority of human experts who retain live responsibility for judgment, refusal, revision, and consequence.

The expert is not reduced to a passive approver. The expert becomes the steward of the agentic workflow. Their job is to define boundaries, review proposals, resolve ambiguity, intervene in edge cases, read receipts, update policy, and develop the system over time.

Old Model	Bad AI Model	HAPI Model
Human does every task manually.	Agent does the work and human rubber-stamps.	Agent executes bounded tasks while expert governs judgment and authority.
Slow but skill-preserving.	Fast but agency-eroding.	Faster while preserving refusal, revision, memory, and accountability.
Expertise lives in the worker.	Expertise migrates into opaque systems.	Expertise shapes the agentic system through governed delegation.

This model does not pretend that work will remain unchanged. It accepts that AI will transform work, but it insists that transformation should preserve and amplify human agency rather than hollow it out.

6. New Roles in an Agency-Preserving Economy

If expert-governed delegation becomes a serious work pattern, new roles emerge. These roles are not merely prompt engineering roles. They are governance, interpretation, and accountability roles.

- Agent Governance Specialist: maps authority, approvals, escalation paths, and risk boundaries for agentic workflows.
- AI Runtime Controls Analyst: monitors whether agents execute only what was authorized.
- Human Oversight Lead: ensures human review remains meaningful rather than symbolic.
- Agentic Workflow Architect: designs work systems where AI handles execution but humans retain judgment.
- Receipt and Audit Analyst: reviews evidence trails, approvals, outcomes, and drift patterns.

- Domain-Agent Supervisor: a domain expert who governs specialized agents in healthcare, finance, safety, cybersecurity, education, or operations.

The future worker may govern a team of role-specific agents. But this only preserves agency if the human remains capable of understanding, challenging, redirecting, and learning from the system.

7. The Risk of Skill Atrophy

AI can increase short-term productivity while decreasing long-term human capacity. This is one of the central agency risks of automation. If people stop practicing judgment, diagnosis, writing, planning, repair, or moral reasoning, they may become dependent on systems they can no longer evaluate.

HAPI distinguishes between agency-supporting automation and agency-replacing automation. Supporting automation gives people leverage while preserving or increasing their ability to act. Replacing automation causes the person to lose the internal capacity needed to participate meaningfully.

> *A tool that makes a person faster is not necessarily agency-preserving. It must also help the person remain capable.*

This suggests that workplace AI systems should include capacity-preserving design: explanations, review trails, training loops, scenario practice, deliberate friction for high-risk actions, and opportunities for humans to maintain skill.

8. Work, Meaning, and Human Dignity

Work is also connected to dignity. People often experience dignity when they can contribute, solve problems, provide for others, develop mastery, and see the fruit of their effort. If AI removes all meaningful contribution and leaves people only with surveillance, metrics, and symbolic oversight, the workplace becomes agency-poor even if it remains economically productive.

A HAPI approach does not claim that every task must remain human. Some tasks are harmful, exhausting, boring, unsafe, or degrading. Automating such tasks may restore human agency. But the replacement of harmful labor should be paired with new pathways for meaningful participation, not simply removal from the economic and social structure of contribution.

The question is not whether people should be forced to perform unnecessary labor. The question is whether society can preserve meaningful participation as labor changes.

9. Organizational Design Principles

An agency-preserving AI workplace should be designed around several principles.

1. No responsibility without authority. Workers should not be blamed for outcomes they could not meaningfully affect.
2. No approval without refusal. If a worker cannot say no, approval is theatrical.

3. No automation without memory. The system must preserve what was proposed, approved, refused, executed, and learned.
4. No expert displacement disguised as assistance. AI should support expertise, not quietly replace the expert while preserving their liability.
5. No speed beyond discernment. High-consequence action must not move faster than human judgment can participate.
6. No productivity gains purchased through human deskilling without a plan to restore capacity.
7. No hidden governance. Workers should know what rules, policies, metrics, and agentic systems shape their work.

10. The HAPI Work Audit

A HAPI work audit asks whether a workplace is preserving or eroding human agency under AI adoption. It does not only ask whether AI is efficient. It asks whether humans remain meaningfully involved in the conditions of work.

Audit Dimension	Question
Clarity	Do workers understand how AI affects their tasks, decisions, and responsibilities?
Authority	Can workers alter, halt, or escalate AI-shaped decisions before consequence?
Refusal	Can workers say no without being punished for slowing the system?
Revision	Can workers improve the proposal before execution?
Memory	Are proposals, approvals, refusals, and outcomes preserved?
Skill	Does the system preserve or degrade domain expertise over time?
Accountability	Is responsibility matched to actual control?
Meaning	Does the work still allow human contribution and development?
Dependency	Can workers function and judge if the system fails or gives bad output?

The audit goal is not to block AI. The goal is to prevent AI adoption from quietly converting workers into procedural participants.

11. Agentic AI and Workplace Governance

Agentic AI makes the problem more urgent because agents do not merely produce information. They can propose actions, call tools, change records, send messages, update systems, and trigger real-world consequences. This moves AI from advice into delegated agency.

In a workplace, delegated agency must be governed. The agent acts under borrowed human or organizational authority. Therefore the organization must preserve the chain from human intent to agent proposal to authorization to execution to receipt.

- PGDL challenges the proposal before action framing hardens.
- AAG determines whether the action is authorized, scoped, reversible, and approved.
- Runtime Binding ensures the tool executes only what was permitted.
- Receipts preserve memory and accountability after consequence.
- Continuity findings reveal whether governance remains coherent over time.

This is not only an AI safety stack. It is a workplace agency stack.

12. Policy Implications

Governments, companies, unions, professional boards, and standards bodies will need language for AI-enabled work that goes beyond job counts. Job displacement matters, but so does agency displacement. A role may technically remain while its meaningful authority disappears.

Future policy should evaluate whether workers retain context, refusal, contestability, accountability alignment, skill development, and audit visibility. A workplace can be compliant with employment law while still degrading human agency through opaque automation and metric pressure.

HAPI offers a vocabulary for this deeper policy layer: preserve agency, prevent rubber-stamp employment, govern delegated action, and ensure that human accountability remains matched to human authority.

13. Conclusion

The age of AI will change work. Some tasks will disappear. Some roles will transform. Some forms of labor should be relieved. But the central question is not only what humans will do after AI. The central question is whether humans will remain meaningful agents inside the systems AI helps build.

HAPI proposes that the future of work should be designed around agency-preserving delegation. AI should take over bounded execution where appropriate, but humans must retain the live conditions of agency: clarity, authority, refusal, revision, memory, skill, meaning, and accountability.

> *The goal is not to preserve every old task. The goal is to preserve the human agency that makes work meaningful, accountable, and real.*

Appendix A: Key Propositions

8. Work is not only labor. It is a structure of agency.
9. AI can automate tasks while either preserving or eroding agency.
10. A human-in-the-loop workplace can still become agency theater.
11. Expert-governed delegation is the agency-preserving alternative to both manual overload and rubber-stamp automation.
12. AI-enabled work should match responsibility to authority.
13. Approval is meaningful only if refusal remains possible.
14. Productivity gains that destroy human skill may create long-term dependency.

15. The future worker may govern agents, but only if the worker retains real authority, memory, and capacity.
16. Agentic AI makes work governance an action-path problem, not only a policy problem.
17. The future of work should be measured by agency preservation as well as efficiency.

Appendix B: Glossary

Term	Definition
Agency-preserving delegation	A work pattern where AI handles bounded execution while humans retain judgment, refusal, revision, memory, and accountability.
Rubber-stamp worker	A person who is formally present in a workflow but lacks meaningful power to change the outcome.
Agency displacement	The removal of human authority, skill, memory, refusal, or judgment even when employment technically remains.
Expert-governed agent team	A set of role-specific AI agents governed by a human domain expert.
Skill atrophy	The loss of human capacity caused by over-reliance on automated systems.
Delegated operational agency	Action performed by an AI system under borrowed human or organizational authority.