WORKING PAPER / 0.1

Human-Indexed Work

A labor architecture for personally operated AI capability

HIRE THE HUMAN. MEASURE THE WHOLE SYSTEM.

A vendor-neutral proposal for portable operator layers, job-bounded capability evidence, clean-room transfer, and human accountability in agentic work.

15 AUGUST 2026

Independent working draft · not an accredited standard or legal instrument

Abstract

The dominant AI-labor debate assumes two units: a human worker and an AI system. It then asks which tasks move from the first unit to the second. This paper proposes a third unit: the **Human Capability Unit**, consisting of an identified natural person, a personally operated control layer, one or more AI-enabled capability modules, an employer-controlled work context, and licensed compute substrates. The company hires the human principal and evaluates the combined unit.

This is more specific than "augmentation," "copilot," or "human in the loop." The worker's computational capacity is not merely a generic tool assigned by the employer. It is partly a portable, cultivated way of working: interfaces, automations, evaluations, procedures, controls, and learned operating skill that the person can adapt across contexts. Its value cannot be inferred from the number of agents or sophistication of prompts. It must be demonstrated under job-relevant conditions, including variation, failure, recovery, transfer, and measured human attention.

The proposal is economically plausible but not self-executing. If companies own universally replicable agents and people contribute no scarce complement, labor demand may still contract. Human-Indexed Work preserves a meaningful role for employment only where human judgment, accountability, trust, adaptation, relationships, taste, or contextual authority remain scarce and where increased capability expands useful demand. This paper defines the institutional machinery needed to test that proposition honestly: ownership boundaries, a capability protocol, evidence receipts, a portable Capability Passport, field trials, and failure criteria.

1. The proposition

Modern hiring treats software competence as human capital. A designer is valuable partly because they can operate a design stack; a developer carries fluency in languages and systems; an analyst brings methods, templates, and judgment. Employers do not hire the software. They hire the person who can make it produce valuable work.

AI generalizes this pattern. The relevant asset will increasingly be neither "knowing a software package" nor possessing a generic chatbot. It will be a **personally operated computational system**: a network of tools, automations, prompts, agents, interfaces, tests, memory structures, and intervention procedures that a particular person knows how to build, tune, supervise, and adapt.

The unit of work therefore changes:

> *human + operator layer + AI execution + work context = verified capability*

The system may be highly automated. In that case, the person operates it through architecture: designing automations, revising controls, interpreting monitoring, and handling exceptions. It may be interactive. In that case, the person operates it through real-time decisions, approvals, controls, and feedback surfaces. In both forms, performance arises from the coupling. Separating the output into "what the human did" and "what the AI did" is often less useful than evaluating whether the combined system reliably accomplished the job.

2. What already exists—and what is missing

The surrounding pieces are emerging quickly. Microsoft describes human-agent teams and an "agent boss" role in which humans set direction and agents execute. Its 2025 Work Trend Index explicitly argues that firms will need to manage a human-agent ratio, while its 2026 report emphasizes humans directing work and owning outcomes as agents

take on execution. The International Labour Organization's 2025 occupational analysis finds transformation exposure broader than full automation exposure, reinforcing the importance of job redesign rather than assuming immediate whole-job substitution.

Risk and governance foundations also exist. NIST's AI Risk Management Framework organizes work around governing, mapping, measuring, and managing AI risk, and its generative-AI profile calls for differentiated responsibilities in human-AI configurations. The European Union's AI Act requires competent human oversight and monitoring for covered high-risk deployments. ISO/IEC 17024 provides a benchmark for certification of persons. W3C Verifiable Credentials 2.0 provides an interoperable model for issuer-holder-verifier claims.

None of those pieces, by itself, establishes the labor unit proposed here. "Agent boss" does not decide what belongs to the worker when employment ends. Human oversight does not measure an operator's leverage or recovery skill. Personnel certification normally assesses a person, not a continuously changing person-plus-system configuration. Verifiable credentials can carry a claim but do not define a valid capability assessment.

The missing synthesis is an institution that can answer five questions:

- What exactly is the combined unit being hired?
- Which parts travel with the person, stay with the employer, or remain licensed from vendors?
- How is its job-relevant performance tested without exposing private internals?
- How does an employer trust a capability claim as models, contexts, and risks change?
- Under what evidence would the model be rejected rather than promoted?

3. The Human Capability Unit

The Human Capability Unit has five layers.

3.1 Human principal

An identified natural person sets intent, interprets context, authorizes actions, intervenes, and answers for outcomes. This is not ceremonial accountability. The

operator must have enough competence, time, information, and authority to exercise real control.

3.2 Operator layer

This is the worker's practical craft: interfaces, prompts, procedures, automations, routing logic, quality gates, monitors, escalation rules, evaluations, and non-confidential preferences. It may be code, configuration, documentation, or embodied operational knowledge. The layer is valuable because its creator understands not only how to run it, but when not to trust it.

3.3 Capability modules

Modules execute bounded functions such as research, document production, customer triage, code repair, forecasting, or monitoring. A module declares its inputs, outputs, tools, autonomy level, constraints, failure modes, and escalation path.

3.4 Work context

The employer supplies data, credentials, policies, customer relationships, internal systems, and decision rights. This layer is engagement-specific and normally does not travel with the person.

3.5 Compute substrate

Foundation models, hosted services, tools, and infrastructure remain governed by provider licenses and technical constraints. A worker cannot truthfully claim to own a third party's model merely because their stack invokes it.

This layered model turns "I own my agent" from a slogan into a negotiable architecture. The person can own or license a portable method without claiming employer data or vendor infrastructure.

4. Ownership and the portability bargain

The central institutional problem is not technical. It is the boundary between portable human capital and employer property.

Today, employment agreements often assign inventions created within the scope of work to the employer. AI makes the boundary harder because a reusable personal system can absorb workplace examples, corrections, data, and procedures. A credible model cannot promise workers unrestricted portability. It needs a **three-domain separation**.

The operator-owned domain contains pre-existing and independently developed methods, generic orchestration patterns, personal interfaces, reusable evaluations, general working preferences, and the person's credential history. The employer domain contains confidential data, workplace credentials, proprietary integrations, customer information, internal policies, work product, and improvements that the governing agreement assigns to the employer. The provider domain contains foundation models, hosted systems, licensed datasets, and third-party tools.

Every engagement should produce a signed carry-in/carry-out schedule:

- what the worker brought;
- what the employer provided;
- what was jointly adapted;
- what must be removed or reconstructed at exit;
- what proof may be retained in redacted or aggregate form;
- what post-employment rights each party holds.

Portability must be demonstrated through clean-room transfer. The operator rebuilds the portable layer against synthetic or new-context data while an evaluator checks that former-employer content, credentials, and protected procedures are absent. The worker carries capability, not a copy of the old workplace.

This requires legal templates and jurisdiction-specific review. HIWP is not a declaration that labor law, trade-secret law, copyright, invention assignment, professional regulation, or data-protection obligations disappear. It is a technical and commercial structure for making those boundaries explicit.

5. Measuring the coupled system

A conventional résumé lists experience and tools. A generic AI certificate tests conceptual knowledge or prompting. Neither measures the proposed unit.

HIWP uses a job-bounded capability claim. A useful claim has the form:

> *This identified operator, using this declared class of system, can produce this outcome to this acceptance threshold, within these time, cost, attention, data, autonomy, and risk constraints.*

The protocol reports a performance vector rather than a single score:

- **Outcome quality:** blind or rubric-based review against professional acceptance criteria.
- **Reliability:** variance across repetitions, ordinary input variation, and time.
- **Operator leverage:** verified output per unit of human attention, not merely elapsed machine time.
- **Intervention skill:** detection, diagnosis, correction, rollback, and escalation.
- **Recovery:** severity and duration of failures and the ability to restore a safe state.
- **Transferability:** time and quality loss when adapting to a new context or substitute model.
- **Evidence integrity:** trace completeness, evaluator independence, reproducibility, and tamper resistance.
- **Risk compliance:** adherence to declared autonomy, privacy, security, and professional constraints.

Human attention is crucial. A workflow that appears autonomous but requires invisible checking, prompt repair, and cleanup is not equivalent to one that delivers the same quality with low intervention. Conversely, a highly interactive system may be excellent if judgment is the scarce value and machine execution compresses the routine work around it.

6. Assessment design

An assessment must test the operator-system coupling, not reward a memorized demo. HIWP v0.1 proposes six trial families.

The **baseline trial** uses representative tasks and known acceptance criteria. The **variation trial** changes normal inputs so brittle scripting cannot masquerade as capability. The **novel-task trial** stays within the role but requires adaptation. The **exception and recovery trial** injects missing, contradictory, or malformed inputs and

records whether the operator notices and restores control. The **adversarial trial** probes data leakage, unsafe tool use, prompt injection, false confidence, and policy evasion. The **substitution and transfer trials** change a material tool or context to measure whether the capability belongs to the operator layer or is just accidental dependence on one vendor or dataset.

Each run emits an evidence receipt. Sensitive inputs can be represented by hashes, controlled descriptors, and evaluator attestations. The receipt records the declared stack version, task version, evaluator, timing, human attention, interventions, output hash, quality result, failure events, and applicable risk controls.

The resulting Capability Passport is held by the worker and selectively disclosed. An employer need not see proprietary prompts. It sees what was tested, by whom, under what conditions, how recently, and with what result. Claims expire or require revalidation when material dependencies change.

7. Hiring and work design

Human-Indexed Work changes hiring from tool checklists to capability procurement.

The employer first defines a role slice in terms of outcomes, constraints, and risk. Candidates then demonstrate a relevant HCU rather than claim abstract AI literacy. A provisional assessment leads to a paid, supervised pilot. During employment, the employer provides contextual access through least-privilege credentials while the operator supplies or develops the operating layer under an explicit IP schedule. Production use is monitored against the same capability and risk claims used in assessment.

The employer still hires a person. The person may be substantially more capable, but the organization gains another independent source of judgment, accountability, adaptation, relationships, and computational execution. Hiring an additional capable operator adds another human principal and another execution tree.

8. Economics: why this might preserve or expand employment

The naive productivity model holds output constant. If one AI-enabled worker produces five times as much, the firm needs one-fifth as many workers. That outcome is possible, especially where demand is fixed and human contribution becomes negligible.

But competitive firms do not always hold output constant. Lower cost and greater capability can create new products, higher service levels, finer personalization, broader market reach, and activities that were previously uneconomic. If demand expands and human complements remain scarce, firms may employ more highly leveraged people rather than minimize headcount to the previous output level.

Human-Indexed Work strengthens this expansion mechanism in three ways. First, each hire adds distinct judgment and context, not merely another copy of the same model. Second, personal operator layers can produce heterogeneous strategies and products, supporting differentiation. Third, portable capability lowers the time required for skilled people to become productive while preserving an incentive for them to invest in better systems.

The model does not guarantee full employment. It is most plausible where:

- output demand is elastic or new output categories can be created;
- errors, legitimacy, relationships, or accountability matter;
- contextual variation rewards adaptation;
- the human can govern more execution than they could perform manually;
- multiple perspectives outperform one centralized operator;
- workers can capture enough return to justify investing in their stacks.

It is weakest where tasks are standardized, demand is capped, outputs are cheaply verifiable, and one centrally owned agent can replicate the entire job without meaningful human complementarity.

9. Risks and failure modes

The proposal could become exploitative if “bring your own AI” turns into unpaid infrastructure, permanent availability, surveillance, or an expectation that every employee personally finances models and tools. Employer benefit must be paired with compensation, security support, workload limits, and clear liability.

It could deepen inequality if only wealthy workers can build powerful stacks. Public education, employer-funded compute allowances, open reference implementations, accessible interfaces, and portable credentials are therefore not peripheral—they are adoption requirements.

It could produce credential theater. A glossy score could hide leakage, cherry-picked tasks, or extensive invisible labor. This is why receipts, variation tests, attention accounting, expiration, and evaluator independence matter.

It could weaken team learning if capability remains too personal. HIWP therefore separates portable personal methods from employer-funded shared improvements and allows explicit licensing or contribution agreements.

It could also centralize risk in the human principal. Accountability must track actual control. Employers and providers cannot delegate liability to a worker while withholding authority, observability, or the ability to stop the system.

10. A falsifiable field program

The proposal should begin with one 90-day pilot, not a universal certification launch. Select a low-to-moderate-risk role slice with frequent, objectively reviewable outputs. Recruit 8–20 experienced operators and measure three conditions: unaided or conventional tools, access to generic AI tools, and personally configured operator stacks.

Pre-register primary measures: accepted output per human hour, blind quality, serious-error rate, intervention burden, recovery time, adaptation time, and total cost. Require at least one model substitution and one clean-room transfer. Keep task authors and evaluators separated where practical. Publish negative findings.

The thesis is weakened if personalized stacks do not outperform generic AI access after full attention and maintenance costs; if gains disappear under realistic variation; if clean transfer is consistently infeasible; if employers see no value in independently verified capability; or if the worker's irreplaceable contribution falls below a meaningful threshold.

The thesis is strengthened if personally operated stacks produce durable gains, those gains survive transfer and substitution, operator skill predicts recovery and safe use, and firms treat an additional human principal as incremental productive capacity rather than redundant supervision.

11. The initial institution

The near-term product should be a protocol and assessment company, not an agent marketplace. An agent marketplace encourages buyers to separate automations from their creators. HIWP should instead provide:

- an open capability manifest and evidence-receipt format;
- a role-design toolkit for employers;
- assessment harnesses and independent evaluators;
- worker-held Capability Passports;
- clean-room portability audits;
- contractual schedules for operator, employer, and provider layers;
- field research on how human-plus-system capability affects hiring and demand.

Revenue can come from employer pilots, assessment fees, verification infrastructure, portability audits, and enterprise governance. The open protocol prevents lock-in; trustworthy assessment data, evaluator networks, role-specific benchmarks, and institutional adoption create defensibility.

Conclusion

AI may reinvent software, but that does not imply software becomes irrelevant. The craft moves upward. People will build and operate personalized computational systems whose behavior is partly automated and partly interactive. The labor market then needs a way

to recognize the combined capacity without erasing the person, exposing private internals, or confusing access to a model with competence.

Human-Indexed Work is a proposal for that missing layer. It says the organization should remain made of accountable human principals, each able to direct an expanding computational tree. Whether that structure preserves or increases employment is an empirical question. The immediate task is to build the protocol, test the coupled unit, make ownership portable without violating employer rights, and let evidence decide.

References

- Microsoft, “2025: The year the Frontier Firm is born,” 23 April 2025. https://www.microsoft.com/en-us/worklab/work-trend-index/2025-the-year-the-frontier-firm-is-born
- Microsoft, “Agents, human agency, and the opportunity for every organization,” 5 May 2026. https://www.microsoft.com/en-us/worklab/work-trend-index/agents-human-agency-and-the-opportunity-for-every-organization
- International Labour Organization, “Generative AI and Jobs: A Refined Global Index of Occupational Exposure,” 20 May 2025. https://www.ilo.org/publications/generative-ai-and-jobs-refined-global-index-occupational-exposure
- NIST, “Artificial Intelligence Risk Management Framework (AI RMF 1.0),” January 2023. https://www.nist.gov/itl/ai-risk-management-framework
- NIST, “Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile,” July 2024. https://www.nist.gov/publications/artificial-intelligence-risk-management-framework-generative-artificial-intelligence
- European Union, AI Act Service Desk, Article 14 (Human oversight) and Article 26 (Obligations of deployers). https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-14 and https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-26
- ISO, “ISO/IEC 17024:2026 — Conformity assessment — General requirements for bodies operating certification of persons.” https://www.iso.org/standard/17024
- W3C, “Verifiable Credentials Data Model v2.0,” W3C Recommendation, 15 May 2025. https://www.w3.org/TR/vc-data-model-2.0/
- OECD, “Artificial intelligence and the changing demand for skills in the labour market,” 10 April 2024. https://www.oecd.org/en/publications/artificial-intelligence-and-the-changing-demand-for-skills-in-the-labour-market_88684e36-en.html

WORKING DRAFT STATUS

This paper is designed to generate a real protocol and field trial. It is intentionally falsifiable. Names are provisional pending trademark and legal review; legal and regulated-sector implementation requires qualified counsel and competent authorities.