Disruption of Economic Theories in the Era of Generative AI:
Performance, Accountability, and Reward

Michal Kubicek
Kubicek.AI / Mediatoring.com s.r.o., Ostrava, Czech Republic
michal@kubicek.ai

Working Paper · Preliminary Draft · March 2026
This paper has not been peer reviewed.

JEL Classification: J24, J31, O33, D23, J23, B41, C45

Keywords: generative AI, labor economics, accountability, cognitive work, automation, human capital, EMA framework, EMMA framework, TRCL, responsible cognitive labor

About this paper. This is an idea paper, not an academic study in the traditional sense. It contains no research question amenable to conventional empirical testing, nor a hypothesis to be formally verified. It proposes two working frameworks — EMA and EMMA — and advances a working proposition on responsible cognitive labor. Both are proposals for further inquiry, not finished theories. The paper acknowledges its own limitations: the key concepts are not sufficiently operationalized for formal testing, the predictions are directional rather than precise, and the frameworks are intended to name a tension, not to resolve it definitively. What the paper does offer is a precise identification of a problem that most debates about AI circumvent: that the contemporary economy is severing the relationship between performance, accountability, and reward.

Abstract

Generative artificial intelligence disrupts a foundational assumption underlying most economic theories of labor: that performance, accountability, and reward are mutually intertwined. This idea paper problematizes that assumption and advances three working concepts for its analysis. The EMA framework (Empathy–Mechanics–Assistance) classifies labor activities according to the degree of human irreplaceability. The EMMA framework (Evaluation–Monitoring–Modification–Arbitration) describes what a responsible human worker concretely does in an environment where AI generates a substantial portion of outputs previously delivered by human effort alone. The Thesis on Responsible Cognitive Labor (TRCL) connects both frameworks into a working proposition: the value of human labor should be proportional to the degree of accountability assumed, not to the quantity of time expended or effort exerted. This proposition contains an empirically testable prediction: where accountability is institutionally anchored, human labor will retain its economic value longer than where such anchoring is absent. The paper acknowledges its limitations — the concepts are not yet sufficiently operationalized for formal testing and the predictions are directional rather than precise — and offers itself as a basis for further research rather than a finished theory.

1. Introduction

If generative AI is capable of performing a substantial portion of cognitive work, what exactly does the human worker do, what should they be paid for, and why?

The question of human labor in an environment of generative AI has three distinct dimensions. The first is technological: which cognitive tasks can today be delegated to a machine. The second is organizational: how firms integrate these capabilities into work processes. The third is economic: how, as a result of these changes, value is redistributed among workers, capital, and the owners of technology. This paper focuses primarily on the third dimension. Technological progress is its premise, not its subject.

This question touches everyone who uses AI at work and wonders whether what they do still constitutes "labor" in an economic sense. It concerns managers who do not know how to evaluate an employee who "merely" reviews AI outputs. It concerns policymakers who must decide whether and how to tax AI-generated productivity. It is not, however, a question this paper can answer with finality. What it offers is a framework for thinking about it.

This paper contributes to a nascent strand of research at the intersection of labor economics and AI studies that moves beyond questions of employment levels to examine the structural decoupling of performance from accountability and reward. The argument proceeds as follows. Section 2 surveys recent empirical evidence on AI adoption and its effects on labor. Section 3 reviews classical economic theories and identifies, for each, a single assumption now under pressure from generative AI. Sections 4 and 5 introduce the EMA and EMMA frameworks as working tools for mapping the labor landscape. Section 6 develops the TRCL thesis and examines its normative and empirical implications. Section 7 addresses the invisibility of accountability as a structural market problem. Section 8 confronts the thesis with market realities. Section 9 sketches the broader societal dimension. Section 10 briefly considers the Central European context. Section 11 concludes.

2. What the Data Show

Generative AI is no longer an experiment. It is a global economic phenomenon with measurable effects. The figures cited below do not prove the validity of any single theory — but they suggest that traditional intuitions about labor, reward, and expertise are under significant pressure. The following subsections examine both the global and the local Czech context.

According to Bick, Blandin, and Deming (Federal Reserve Bank of St. Louis / NBER), in August and November 2024, 26.5 percent of employed Americans used generative AI for work; 22.9 percent had used it at least once in the preceding week and 9.0 percent used it every working day. According to the PwC Global Workforce Hopes and Fears Survey 2025, 54 percent of workers had used AI in their role over the preceding twelve months, but only 14 percent do so daily. It is precisely the daily users who show markedly better outcomes than irregular users: 92 percent of them report higher productivity, compared with 58 percent among less frequent users, and 52 percent report salary growth.

Controlled experiments consistently show a 15 to 50 percent reduction in the time required for cognitive tasks. Noy and Zhang (2023) show that access to ChatGPT for professionally relevant writing tasks among 453 college-educated professionals led to an average time reduction of 40 percent and a quality improvement of 18 percent. Brynjolfsson, Li, and Raymond (2025) find approximately 15 percent average productivity growth in the customer support division of a Fortune 500 firm, with gains substantially higher for less experienced and lower-performing workers — a result commonly interpreted as skill compression.

The BIS Working Paper (Aldasoro et al., 2026), based on data from more than 12,000 non-financial firms in the EU and the United States, estimates that AI adoption in European firms increases labor productivity by approximately 4 percent, without adverse short-run effects on employment.

The Anthropic Economic Index, based on privacy-preserving analysis of 1 million conversations from Claude.ai and 1 million records from the first-party API, shows that on

Claude.ai in November 2025 augmentation — situations in which the tool is used as an assistant while the human remains the primary decision-maker — exceeded automation in a ratio of 52 to 45 percent. The remaining three percent of conversations are classified in the index's methodology as "neither": neither augmentation nor automation. These are typically uses in which Claude functions more as an information source or conversational partner, or where the nature of the task did not permit reliable classification according to the criteria applied (degree of human oversight, nature of input and output, task repetitiveness, and so on). The report simultaneously notes that, relative to earlier periods, automation remains elevated and the long-term trend may point toward further growth. Across reports, 49 percent of occupations show AI use for at least one-quarter of their tasks in Claude data — though this is not equivalent to operational AI deployment for one-quarter of tasks across those occupations.

EY (2025) found that 88 percent of employees use AI at work, but only 5 percent do so in advanced ways that transform the way they work. So-called basic users issue isolated, one-off prompts — typically for information retrieval, document summarization, or translation. The AI tool thus functions more as an intelligent search engine or a rapid assistant for discrete subtasks. The most common basic uses are information search (54 percent of employees) and document summarization (38 percent). Advanced users, by contrast, engage AI throughout the entire thinking process: they treat it not as a simple automation device but as a colleague, coach, and intellectual partner, combining multiple tools simultaneously, with AI present across the full arc of problem-solving. As a result, they gain approximately one and a half additional productive days per week. Yet only 28 percent of organizations are at all capable of converting AI deployment into high-value-added results. The paradox deepens in the domain of training: employees with more than 81 hours of AI training per year save 14 hours per week, but are simultaneously 55 percent more likely to leave the organization, because their skills have become highly sought on the external labor market. Organizations thus face a dilemma in which the more they invest in developing AI competencies, the more they risk losing precisely their most valuable employees. A further paradox is so-called shadow AI: between 23 and 58 percent of employees (depending on sector) bring their own, unapproved AI tools to work — which testifies to motivation and initiative, but simultaneously creates security and compliance risks.

The Czech Context

According to Wiedermann et al. (2025), by 2035 generative AI will significantly affect nearly every second job (43 percent of the labor market), and for 11 percent of positions (approximately 600,000 people) the impact will be fundamental and disruptive. The study estimates a need to relocate 355,000 individuals from disappearing to newly created positions, of whom 67,000 will require extensive retraining. Adaptation to generative AI could bring the Czech economy additional GDP growth of 0.3 percent annually, raising the total estimated GDP growth to 2.5 percent.

These data demonstrate that conventional intuitions can no longer be straightforwardly applied. A working hour is no longer a reliable unit of value. Experience is no longer a guarantee of superior performance. Productivity no longer correlates with headcount.

The empirical literature on generative AI currently paints a reasonably consistent picture. The technology increases productivity primarily among less experienced workers, accelerates the performance of routine cognitive tasks, and partially commoditizes certain types of expertise. These three effects together create tension within the traditional relationship between performance and reward. If a less experienced worker supported by AI can achieve output comparable to that of an experienced specialist, the length of one's career ceases to be a reliable indicator of economic value. If AI generates enormous quantities of high-quality outputs at minimal human-labor cost, the working hour ceases to function as a reliable unit of value. And if most people use AI only superficially while a narrow group benefits dramatically from it, performance and reward also diverge. These shifts inspire the remainder of this paper.

3. Classical Economic Theories from the Perspective of the AI Era

The following overview does not aspire to serve as a comprehensive exposition of the existing economic traditions. It is a deliberately simplified map: for each theory, I identify one assumption now under pressure from the advent of generative AI. A reader with knowledge of economic thought may justifiably object that each of these traditions is more sophisticated than the way it is presented here. That is correct. The purpose is not to refute these theories, but to show that their foundational concepts were formed in a context that is now changing.

Smith: Labor as a Measure of Value

Adam Smith used labor as a measure of value primarily because in the pre-industrial economy it constituted a relatively stable unit of comparison. Labor was a universal input into production. Smith was nevertheless aware that market price is the outcome of the relationship among labor, capital, and land, and he distinguished between "labour commanded" and "labour embodied." It is therefore imprecise to say that Smith viewed labor as the sole source of value — rather, he treated it as the most reliable comparative unit.

Generative AI destabilizes this framework in a manner different from industrial machinery. While the steam engine replaced physical labor, language models intervene directly in the domain of cognitive production. The result is a situation in which a small quantity of human labor can activate enormous computational capacity. Labor thus ceases to be a reliable measure of economic value, because its quantity no longer corresponds to the scale of output.

Marx: Living Labor as the Source of Value

Marx's analysis of capitalism rests on the assumption that new value arises only from living labor. Machines, in his account, merely transfer value previously embodied in them. Generative AI complicates this relationship. Language models are not merely tools that increase the productivity of labor; in many situations they perform the cognitive activity itself. A one-time investment in model development can generate a practically unlimited volume of outputs with minimal requirement for human labor.

Marx's analysis faithfully captures the logic of nineteenth-century industrial production, in which incremental output was directly tied to incremental living labor. The digital economy — and generative AI as its most extreme case — fundamentally violates this mechanism rather than merely complicating it. Tools such as spreadsheet software, content management systems, or template engines had already demonstrated that a one-time development investment can generate value for millions of users without a proportional engagement of additional labor. The marginal cost of replicating a digital output approaches zero, while the value to the recipient remains real and measurable. Generative AI does not push this phenomenon to an extreme — it renders it so starkly visible that it can no longer be ignored.

Critical authors such as Durand (2020) and Srnicek (2017) are more valuable in this regard precisely because they do not retrofit new reality to an old framework, but instead describe the intrinsic logic of scalability and monopolization of intangible assets: once infrastructure is built, its outputs can be reproduced at scale with very low additional costs — something industrial production never permitted.

Marginalism: Marginal Utility and Scarcity

Marginalist economics shifted attention from labor to subjective value and marginal productivity. The price of a good reflects its marginal utility and availability. Generative AI, however, alters the very structure of scarcity. Text, code, illustrations, or analyses that formerly required hours of skilled labor can today be produced in seconds. Scarcity therefore migrates away from output production toward other factors: reputation, trustworthiness, accountability for decisions, or access to distribution channels.

Simultaneously, the theory of marginal productivity holds that wages should correspond to the marginal product of the worker. If a worker with AI has an order-of-magnitude higher product, wages should rise accordingly. In practice, however, firms tend to absorb increased AI productivity into margins and reinvestments rather than into the wages of those performing the labor.

It may be objected that AI is merely a tool, and like any tool it increases worker productivity — a backhoe operator earns more than a worker with a pickaxe without this undermining the theory of marginal productivity. The structural problem with generative AI lies elsewhere, however: tools had previously amplified physical or manual performance, while cognitive output — text, analysis, code, judgment — was the domain in which the expert premium of a worker was unambiguously attributable to their education, experience, and considered judgment. That is precisely where the theory of marginal productivity worked best. Generative AI enters this domain with zero marginal cost of replication and a near-zero entry threshold, because outputs that formerly required years of practice are now accessible upon mere registration with a chatbot interface.

The firm pays for the worker's labor and captures the output of the worker-with-AI, while the increment of value is generated by a tool whose value was created and is owned by someone else. This is not a market lag or normal tool dynamics; it is the transfer of the expert premium into an asset the worker does not own. The data further suggest that a low technical entry threshold does not imply a low threshold for transformative use (EY, 2025). The appropriation of the value increment therefore does not affect all workers equally — on the contrary, it tends to deepen differentiation within the labor force itself.

Keynes: The Problem of the Rate of Adjustment

Marx addressed where value comes from; marginalism addressed how it is distributed; Keynes adds a third dimension: what happens to those who are excluded from the increment, and for how long. His argument about technological unemployment was not merely a prediction of job losses. It was above all a problem of pace: technology can increase productivity faster than new forms of work are created.

Generative AI updates this problem in the domain of cognitive labor and knowledge-based occupations — which had seemed, up to this point, largely unaffected and resistant to technological encroachment. While previous waves of automation primarily struck routine manual tasks, the current wave affects analytical and creative activities. The result need not be immediate unemployment, but rather the gradual erosion of traditional career pathways. As has become apparent, firms in the early years following the introduction of public AI models visibly curtailed hiring of junior workers in an effort to outsource their tasks to machines (PwC, 2025; Wiedermann et al., 2025).

Keynes's rate-of-adjustment problem thus falls, in the case of generative AI, not uniformly across the entire workforce, but disproportionately on those who are entering occupations for the first time and who have not yet had the opportunity to build the expert premium that would protect them.

Veblen: Value as a Social Construct

Veblen argued that the economic value of labor is not merely a technical question of productivity but also a question of social recognition. Society decides which activities it regards as valuable and which it does not. Generative AI renders this social dimension of labor visible. If a large share of production can be delegated to a machine, society will need to redefine which human activities merit economic recognition. The question of labor thus becomes a question of social contract.

Empirical studies (e.g., Zhang & Gosline, 2023) show that the perception of value is deeply intertwined with the identity of the creator. Research has confirmed the existence of "human favoritism": even when AI outputs are judged to be of higher quality, the mere awareness of human authorship leads people to express greater satisfaction and willingness to pay. Value therefore resides not only in the text itself but in the ritual of the human creative act.

Human Capital Theory: Investment Under Pressure

Becker and Schultz conceived of education as the most important investment in human capital, enhancing productivity and market value, future earnings, and further benefits. Generative AI does not negate this framework but transforms its internal logic. It reduces the premium for experience (skill compression) while simultaneously creating a new meta-competency that does not fit neatly into traditional categories of human capital. Experience therefore does not lose value as such — it migrates toward the domains of judgment, interpretation, prioritization, accountability, and social coordination. The key meta-competency that is emerging is AI management capability: the ability to purposefully organize work with AI, to assign tasks, verify outputs, and connect them to concrete objectives. What changes is not the existence of human capital itself, but the relationship between education, experience, and productivity — a relationship that is now less stable and less linear.

Empirical data indirectly confirm this hypothesis, though with a paradoxical outcome. Brynjolfsson, Li, and Raymond (2025) found that the benefits of generative AI flow disproportionately to less experienced workers — AI in effect transfers the patterns of the most productive employees to those who had not previously possessed them. Similarly, Noy and Zhang (2023) showed that the greatest productive leap was recorded by workers with a lower baseline output quality. In other words, AI compresses the differences between workers with varying levels of human capital — the very effect Becker anticipated as the result of education is here produced by a tool. This challenges the traditional logic of the return on educational investment: if AI equalizes outcomes, the market premium for years of accumulated expertise also declines.

At the same time, BCG (2025) shows that managers and leaders use generative AI more than twice as frequently as rank-and-file employees, suggesting that the ability to work with AI at a strategic level — precisely the meta-competency defined above — correlates with position and experience rather than with mere technical proficiency. EY (2025) reinforces this interpretation by finding that the key predictor of advanced AI use is not technical skill, but the ability to connect AI outputs to organizational goals — precisely what has traditionally characterized an experienced worker at a senior level. Human capital does not therefore cease to be relevant; it changes its content. What matters is no longer what a person knows, but how they can organize the work of systems that know more.

Task-Based Models and the Question of Reinstatement

Contemporary labor economics had, even before the advent of generative AI, offered a framework directly relevant to this paper. Autor, Levy, and Murnane (2003) formulated the task-based approach, which shifted attention from occupational categories to specific tasks.

They distinguish in particular between routine and non-routine tasks, and show that computerization displaces labor where procedures can be formalized into explicit rules, while for non-routine problem-solving and complex communication it functions rather as a complement to human labor. Their argument therefore does not point toward the extinction of entire occupations, but toward a transformation in the composition of tasks within occupations and toward growing relative demand for college-educated labor.

Generative AI complicates this framework, because it impinges precisely on the portion of non-routine cognitive activities that earlier literature treated as difficult to formalize. Writing, analysis, information synthesis, and the formulation of arguments no longer appear to lie beyond the reach of automation — they look rather like a further band of tasks for which the boundary between complementing and substituting human labor is shifting. In this sense, the original task-based predictions may be a conservative lower bound rather than an upper ceiling of automation pressure.

An even more direct confrontation with the theses of this paper is offered by the work of Acemoglu and Restrepo (2018, 2019). They distinguish between two types of technological change: displacement effects, whereby automation transfers previously human tasks to capital and displaces human labor, and reinstatement effects, whereby technological change creates new tasks in which labor again finds application. Historically, technological waves relied on a combination of both mechanisms, and their impact on employment and the labor share of income depended on whether the creation of new tasks could offset the automation of existing ones. Acemoglu and Restrepo simultaneously argue that in recent decades displacement effects have intensified and reinstatement effects have weakened, such that automation is no longer balanced by new task creation to the same degree as before.

This asymmetry is significant for the normative thesis of this paper: if reinstatement is systematically slower or weaker than displacement, then the thesis advanced below — that human value resides in accountability — describes rather a protective reservoir for a minority of workers than a general alternative to the devaluation of labor.

The common denominator of all these cited economic theories is an assumption that performance, accountability, and reward are more or less interconnected. Whoever works more bears greater responsibility and is better compensated. Generative AI disrupts this interconnection. Performance migrates to the machine. Accountability becomes diffuse. For some types of activity, reward converges toward those who own the model, not those who oversee its outputs.

Consider how a surgeon who uses an AI-generated diagnosis assumes full legal, professional, and moral accountability for the decision: if the diagnosis proves wrong, the surgeon bears the consequences, not the model's developer. In this case, it is precisely accountability that allows the worker to maintain a dominant share of the reward. The patient pays for the physician's accountability and the final decision, not for the use of AI. The model owner collects only a fraction in the form of a licensing fee.

In another scenario, a firm deploys an AI chatbot for customer service without continuous human oversight. Accountability is blurred among the platform owner, the model provider, and, at best, an operator who occasionally reviews outputs. The customer pays for the service; the lion's share of value is captured by the firm and the model owner; the human worker — if one exists in the process at all — receives the minimum.

Between these poles lies most of the real economy: the lawyer who reworks and signs an AI-generated analysis; the architect who modifies and stamps an AI-generated design; the journalist who transforms an AI-researched brief into an article under their byline. In each of these cases, the ratio of reward between worker and model owner differs — and depends on how much observable, institutionally anchored accountability the worker assumes.

From this spectrum a prediction emerges: in occupations with high institutional accountability (medicine, law, auditing), human labor will retain a substantial portion of reward even under massive AI deployment. In occupations with low or unclear accountability (generic content, routine data analysis), reward will migrate to the owners of models and platforms. The middle band — where accountability exists but is not formally anchored — will be the terrain on which the direction of the labor economy is determined. This decoupling and disruption is the core of the problem that the remainder of this paper attempts to name.

4. Knowing Where to Strike: The Expert Premium

The well-known anecdote of the mechanic who repairs a car with three blows of a hammer and charges one hundred dollars — one dollar for the blows and the rest for knowing where to strike — or the story of Henry Ford's engineer marking the generator with a piece of chalk and charging ten thousand dollars, captured precisely the logic of the twentieth-century economy. What was paid for was not the physical or cognitive act itself, but the scarcity of judgment, the depth of experience, and the ability to diagnose correctly.

Generative AI disrupts this logic. Part of what was formerly scarce know-how becomes almost instantly accessible, at minimal or zero cost to the recipient, and at fully mass scale. The only barrier is access to the internet and a device through which one can communicate with AI tools by text or voice.

A language model can propose a course of action, draft a technical procedure, and offer solution variants within seconds. This dramatically lowers the cost of the "cognitive semi-finished product." It does not, however, mean the extinction of human value — it merely relocates that value to new domains: final accountability for outcomes, reputation, and the willingness to bear the consequences of outputs and the decisions that flow from them. Three main scenarios open before us:

Collapse to Commodity Level

In domains where AI can cover the entire journey from prompt to result without critical risk, the price tends toward zero. Expertise becomes widely available infrastructure — much as electricity or internet connectivity transformed from a competitive advantage into a basic business prerequisite. Translation, basic graphic design, routine data analysis, and generic copywriting have already embarked on this trajectory. Human labor in these areas does not vanish entirely, but loses its expert premium: it survives only where it is cheaper or more convenient than configuring an AI tool.

Market Bifurcation (Human-in-the-Loop)

A cheap, purely machine-generated output for mass use emerges alongside premium human oversight for situations requiring validation and quality control. The human here functions as a checkpoint within the process: approving, rejecting, correcting. They are part of the loop, but they do not govern the loop. Accountability for the overall result remains distributed among the system, the firm, and the operator — and precisely this dispersal is why the human-in-the-loop model does not itself resolve the problem of the decoupling of performance and reward. This pattern closely resembles the job polarization described by Autor, Katz, and Kearney (2006), now occurring within knowledge occupations rather than between them. A human serving as a checkpoint is more easily replaceable, and their reward tends to converge toward the wage of an operator, not an expert.

Redefinition of Value (Responsible Arbitration)

The third scenario goes further than the checkpoint within the loop. Here the human does not stand within the process but above it: deciding whether the output should be produced at all, how it should be situated in context, and who will personally bear the legal or professional risk for it. The difference from human-in-the-loop is not that the human "reviews" — it is that they own the decision and bear its consequences. A human-in-the-loop operator at a bank verifies that an AI-generated client analysis conforms to a format specification. An arbitrator decides whether to recommend an investment to the client on the basis of that analysis — and is professionally liable for doing so. Less will be paid for finding the answer and more for the capacity to determine whether the answer is correct and ethically defensible in the given context.

The remainder of this paper develops the third scenario — not as a forecast, but as a proposal.

Choosing the third scenario as the analytical starting point requires justification. It is neither a prediction nor a normative imperative — it is a hypothesis about the conditions under which space opens for institutional solutions. The first scenario — commodity collapse — is probably the dominant trajectory where outputs are not bound to accountability or reputational risk: generic texts, routine analyses, simple code. In these domains, the pressure toward commoditization undoubtedly exists and will likely grow. The second scenario — market bifurcation — is descriptively accurate for regulated professions: medicine, law, and financial advice already exhibit this bifurcation between automated products and premium human guarantees. The third scenario — the redefinition of value through accountability — is analytically the most interesting, yet simultaneously the most poorly conceptualized. Existing economic theories offer tools for describing labor as performance, productivity, or human capital, but lack adequate concepts for labor as curation, evaluation, and assumed accountability for a machine-generated output. Before the third scenario can be analyzed in any meaningful way, this conceptual apparatus must be proposed. The two frameworks that follow are the author's proposals — not finished theories, but working tools for structuring debate.

5. The EMA Framework: A Map of Labor Domains

The preceding sections have been primarily descriptive and critical — mapping what the empirical evidence shows and identifying which assumptions of classical theories are under pressure. The following two sections shift register: they move from observation to proposal, advancing two working frameworks as conceptual tools for the subsequent argument.

EMA (Empathy–Mechanics–Assistance) is a working framework for approaching AI that classifies activities according to the degree of human irreplaceability. It answers the question of where labor is located — in which domain.

The three components of EMA do not stand on the same ontological plane: Empathy is a type of human capacity; Mechanics is the routine quality of a task; Assistance is a mode of collaboration. This intentional mixing of levels has its own rationale, however. EMA is not a classification of human abilities, nor a classification of tasks. It is a classification of delegation decisions. Every work task can be understood as a decision about whether it should be performed by a human, a machine, or a combination of the two. EMA therefore does not describe the world of labor as such, but the manner in which labor is divided between human and technology.

The Empathy (E) domain is the domain of the human with their experiences and lived understanding, where conscience, ethics, intuition, trust, and the ability to read between the lines are decisive. It encompasses decision-making under uncertainty, leading people, facilitating conflict, and bearing accountability for difficult choices. This domain remains human not because AI is incapable of simulating empathy, but because simulated empathy and genuine empathy are not the same — and the recipient can tell the difference. It may be objected that current models increasingly resemble humans in their empathic approach; they remain, nonetheless, only a more refined version of what Joseph Weizenbaum discovered in the late 1970s with ELIZA, his program simulating a Rogerian psychotherapist.

The Mechanics (M) domain is the domain of machine efficiency: routine tasks, structured information, data volumes, repeatable procedures. This is where the human should not waste time or energy. If AI operates on pure mechanics, that is an asset; if the firm holds it back there, human burnout follows.

The Assistance (A) domain is the domain of collaboration. The human provides direction and tasks; AI generates proposals, structures content, expands cognitive space; the human evaluates, combines, and refines outputs. It is important not to succumb to idealization: a large part of what counts as "assistance" in firms today is not creative partnership but cheap acceleration of production, routine checking, and distributed risk. The tension between meaningful collaboration and AI-assisted pseudo-expertise is something EMA names but does not itself resolve.

6. The EMMA Framework: Operations of Accountability

EMA says where labor is located. It does not say what exactly the human does when they are in the Assistance layer or at the intersection of domains. EMMA (Evaluation–Monitoring–Modification–Arbitration) supplies this dimension.

As with EMA, an acknowledgment applies: the four components of EMMA are not equivalently defined. Evaluation is assessment of output; Monitoring is procedural oversight; Modification is iterative activity; Arbitration is strategic decision-making. The acronym is simpler than the reality. The boundary between Evaluation and Monitoring is not sharp in practice, and Arbitration pervades all other dimensions. EMMA is a working tool for structuring debate, not a taxonomy with crisp borders.

Evaluation (E): the capacity to assess the quality, relevance, and correctness of AI outputs; to recognize hallucinations, biases, and errors. Evaluation is more valuable the higher the potential impact of an erroneous output.

Monitoring (M): continuous oversight of the process — is the AI operating within its parameters? Do outputs remain consistent with objectives? Is the model drifting?

Modification (M): iterative refinement — formulating the prompt, sharpening instructions, making final corrections to the output. This requires understanding of both the domain and the capabilities of AI.

Arbitration (A): the strategic decision — when and where to deploy AI, which tasks to delegate, how to allocate human cognitive resources where they generate the highest value.

The relationship between EMA and EMMA: EMA is a map of domains (where am I?); EMMA is a set of operations (what am I responsibly doing here?). Together, they describe two dimensions of human labor in an AI environment — space and activity.

7. The Thesis on Responsible Cognitive Labor (TRCL)

The EMA and EMMA frameworks are observational and descriptive: they map where labor is located and what a person responsibly does there. This section moves from description to proposal. Building on both frameworks, it advances the paper's central analytical claim, which is referred to here as the Thesis on Responsible Cognitive Labor (TRCL). This is not a theory in the strict sense — it lacks formal definitions of key concepts, observable predictions derived from first principles, and empirical verification. It is an orienting proposition advanced as a basis for further inquiry.

TRCL rests on three premises. First: in an economy permeated by AI, the production of output does not in itself constitute the source of human economic value; that source is the responsible decision about what should be created, how the output should be evaluated, and how it should be integrated into a broader context. Second: human labor is shifting from execution to curation — managing, evaluating, and assuming accountability for AI outputs. Third: the value of human labor should be proportional to the degree of accountability assumed, not to the quantity of time expended or effort exerted.

The concept of "accountability" is relatively broad and requires clarification. Bovens (2007) distinguishes forms of accountability according to the nature of the forum to which an actor is answerable — legal, professional, administrative, or social. For the purposes of this paper, four layers are relevant: legal accountability (who is suable when harm occurs), professional accountability (who stakes their expert reputation), moral accountability (who bears the ethical consequences), and reputational risk (who loses credit when things fail).

The choice of accountability as the foundational principle of economic value is not self-evident and merits explicit defense. Several alternative candidates present themselves. The scarcity of skills is a traditional market mechanism: whoever can do what others cannot will be paid more. Reputation functions as a signal of past quality and reduces information asymmetry in markets where value is difficult to measure. Contextual knowledge — deep understanding of a specific organization, client, or situation — is a type of know-how to which AI typically lacks access.

Each of these principles, however, encounters a specific problem in the environment of generative AI. The scarcity of skills erodes rapidly: what was scarce (writing, synthesis, coding) ceases to be scarce as AI outputs are commoditized. Reputation is easily mimicked or falsified where output cannot be distinguished from a human's. Contextual knowledge remains valuable, but is difficult to institutionally operationalize — the market can compensate it, but cannot easily verify it.

Accountability, by contrast, possesses a property that sets it apart: it is institutionally anchorable. Mechanisms exist — licenses, contractual guarantees, legal liability for damage — that render accountability visible, assign it to a specific subject, and allow it to be priced. The point is not that accountability is more valuable than scarcity or reputation — it is that it is institutionally more robust and less susceptible to erosion by AI commoditization. That is precisely why TRCL places it at the center.

For accountability to function as a source of economic value, a mechanism must exist that translates it into price. In practice, this means institutions that make accountability visible: professional licenses, legal liability for outputs, reputational systems, or contractual guarantees. Where such institutions exist — in medicine, law, or financial advising, for instance — human oversight of AI can be economically valuable, because some specific person assumes observable accountability. Conversely, where such institutions are absent, cheap automation will probably dominate. TRCL is therefore not merely a conceptual proposal — it contains an empirically testable prediction: where accountability is institutionally anchored, human labor will retain its value longer than where it is not.

8. The Invisibility of Accountability

The abstract thesis about accountability as a source of value runs into a concrete market mechanism that undermines it: the invisibility of accountability. Markets are typically adept at pricing what is visible, comparable, and immediately attributable — speed, volume, cost, output. Accountability is a different type of value. It often does not manifest in the ordinary course of affairs, but in the fact that an error, damage, collapse, loss of trust, or legal problem does not occur. As long as nothing goes wrong, accountability looks like "inactivity," overhead, or superfluous caution. In that sense it is genuinely invisible — until it is breached. In some sectors, accountability is partially visible even before any breach. This is the case wherever licenses, audits, reputation, insurance, compliance requirements, fiduciary duties, or personal legal liability exist. There, markets at least partially price the capacity to bear risk before any failure occurs. In general, however, we lack a unified metric that could directly and reliably express the value of accountability.

The craftsman was historically paid for time and materials, and this logic survived into the knowledge economy. He had, moreover, the advantage of transparency: a plane, wood shavings, and hours at the workbench were observable. Physical labor was its own proof. Programmers, architects, or copywriters lost this immediate transparency, but the basic nexus remained: their price came gradually to reflect the value of the outcome (success fees, project rates, licensing payments), yet know-how remained in the calculation as a proxy for time and experience. Even where outcomes were invoiced, the price rested on an implicit estimate of how much skilled labor the outcome had required. The result was derived from the effort invested, and the client understood this relationship as fair.

Generative AI splits this relationship into two parts that begin to move independently. On one side stands the value of the output to the client — unchanged or even increased. On the other side stands the time invested in production — which may be compressed from a month to days or hours. Traditional compensation logic says: less time means lower reward. Yet the client receives the same or greater value. This separation of output value from production cost is not new to economic theory — standard microeconomics holds that price should reflect value to the customer, not the producer's costs. But the labor market for knowledge work was organized for decades on precisely the opposite basis: what was invoiced was hours differentiated by the level of the worker's knowledge, not outcomes as such.

The worker now faces three strategies, each carrying different risks.

They can charge the same amount for a shorter period and remain silent about having used AI — but risk that the client will next time estimate the project as requiring less time and offer less, or that AI use will be discovered and call into question both the legitimacy of the price and the quality of the delivered work. This risk is not merely speculative and operates on two distinct levels. On the side of output, Altay and Gilardi (2024) demonstrated that the mere label "AI-generated" reduces perceived accuracy and trustworthiness of content, even when that content is true and of high quality — and the same effect emerged for human outputs incorrectly labeled as AI-generated. Retroactive discovery can thus retroactively devalue work that was substantively correct. On the side of the relationship, Schilke and Reimann (2025), across a series of thirteen experiments, showed that disclosure of AI use reduces trust in a specific worker or supplier across professional contexts; in an experiment simulating a client relationship with a graphic designer, a simultaneous decline in willingness to rehire the supplier was recorded. The negative impact is not dampened even when the evaluator themselves uses AI, and is substantially stronger when discovered by a third party than when proactively disclosed — retroactive discovery thus carries higher reputational costs than upfront transparency.

They can reduce their price proportionally to the time saved — and thereby signal to the market that their reward is a function of hours, not expertise, thereby accelerating the commoditization of their own labor. This strategy is deceptively honest but contains a hidden logical trap: if the worker reduces their price, they confirm to the client that the previous price was derived from time intensity, not from the value of the outcome. The next project the client will appraise with the same optic — asking for a time estimate, not inquiring about the quality of judgment. The worker thus voluntarily enters a competition they cannot win in the long run: AI will always be faster and cheaper. The data suggest that this dynamic is already operating at the level of the labor market: Hosseini and Lichtinger (2025), in an analysis of 62 million workers across 285,000 American firms, found that following generative AI adoption, junior employment in adopting firms declined by nearly 8 percent compared with non-adopting firms, with the decline driven by slower hiring rather than layoffs, while senior employment remained virtually unchanged. Firms are therefore not responding to AI by replacing expensive seniors with AI-equipped juniors — they are simply ceasing to hire juniors altogether. The logic of the hourly rate facilitates this process on both sides of the market.

Or they can reformulate what they sell: not a month of labor, but accountability for the outcome, domain knowledge, the capacity to assign AI the correct problem, and the ability to recognize a bad solution. This strategy is aligned with what EY (2025) identifies as the key predictor of advanced AI use — not technical proficiency, but the ability to connect AI outputs to organizational goals and to bear accountability for them. BCG (2025) further shows that managers and leaders with greater experience use AI more than twice as frequently as rank-and-file employees at a strategic level, indicating that this meta-competency correlates with level of accountability rather than technical skill. Reformulating what the worker sells is not, however, merely a marketing repositioning — it requires genuine change in how the work is performed and how it is communicated. The worker must be able to demonstrate what decision they made, why they rejected certain AI outputs, and what specifically they added beyond what the machine generated. Without this documentation, the reformulation remains an empty gesture that the client has no way to verify, and the market, as shown above, gravitates toward what is visible and comparable.

This third strategy is precisely what TRCL describes as desirable. It runs into a fundamental information asymmetry, however: the client cannot see inside the process. They do not know whether the worker spent the shortened time in careful architecture, iteration, and evaluation of outputs, or whether they copied a prompt and left the result unchecked. The resulting artifact — code, document, analysis — looks the same in both cases until its weaknesses manifest. Accountability becomes observable only through failure, not through success. And what is not observable in advance is difficult to price.

This asymmetry is not a technical detail — it is a structural condition under which the market naturally gravitates toward the first or second scenario, not the third. TRCL therefore requires more than good intentions on the part of workers: it requires mechanisms that make accountability visible before failure. Concretely, this involves three layers. Professional regulation: extending licensing systems to domains where AI outputs carry real risk, and explicitly embedding an obligation to evaluate AI outputs as part of the professional standard in fields where licenses already exist. Contractual architecture: standardized clauses defining who is responsible for evaluating AI outputs, to what extent, and with what consequences upon failure. Auditability of process: technical and organizational requirements for recording how an output was produced, who evaluated it, and what steps it went through. Without these instruments, accountability will remain economically mute.

9. What Does the Market Say?

Every normative thesis must be confronted with descriptive reality.

Accountability does not create market value on its own. Willingness to pay does. A chatbot (ChatGPT, Gemini, Copilot, Claude) or an agentic AI tool (Claude Cowork, OpenAI Codex, Comet Browser) without human oversight can generate massive market value because millions of people pay for it. Without institutional mechanisms — licenses, guarantees, legal accountability — accountability remains invisible and therefore economically irrelevant.

The market today tends in the opposite direction. Firms use AI for the "juniorization" of the workforce: replacing expensive seniors with AI-equipped juniors and externalizing accountability. The metaphor of the "curator" is appealing, but real practice often looks different: cheap acceleration of production, routine checking, degradation of expertise.

"Curator" may be only a new managerial buzzword. To say that the human is shifting from executor to curator may obscure a reality in which accountability is blurred and diluted rather than concentrated. If we wish to examine the matter realistically, we must admit the possibility that the market may not move in the direction of TRCL, but perhaps in precisely the opposite direction — toward the devaluation of accountability, its loosening, and a cheap AI-supported pseudo-expertise.

It may be objected that the market will ultimately find its way to accountability on its own. If firms systematically degrade evaluation, they will eventually pay for it — through failures, damages, or loss of trust. The feedback mechanisms of the market in this sense do not refute TRCL but rather confirm it: accountability will ultimately be reflected in the cost of its absence. In this reading, TRCL describes a market optimum toward which the market moves on its own. Institutional anchoring is then not the precondition for reaching this optimum, but the precondition for the path to it being shorter and less destructive. The market recovered after the Industrial Revolution too, but only after decades whose costs were borne by specific people. And some failures are irreversible or externalized: if a physician approves an AI output without genuine evaluation and causes harm, the market will price this — but the patient first.

Laws, contractual frameworks, and compensation systems are not substitutes for market logic — they are instruments for shortening the transition period and for distributing who bears its costs.

10. The Social Contract

Even if TRCL functions as a framework for those who still have cognitive work, it offers zero answer for those who have lost it. There will be a large number of people who become economically redundant, unemployed, and possibly unemployable (Harari, 2017) — not because they are incapable, but because people working with AI will be many tens of percent more productive.

Moreover, the primary revenues from AI-driven change will not go to the state — they will go to firms. Firms capture the increased productivity. The state attempts to finance safety nets by taxing an ever-narrower base. This is a mathematically rather unstable situation.

This structural problem is not unprecedented. History shows that technological change has always led to the rewriting of the social contract — not automatically, but as the result of political conflict and institutional innovation. The Industrial Revolution gave rise to labor law, trade unions, and the welfare state — but only after decades of social tension had demonstrated that the market would not resolve distributional outcomes on its own. The digital economy gave rise to data and platform regulation — but only after it became clear how far market failure could go in the domains of privacy and market power. Generative AI will probably trigger a similar process, but with one crucial difference: the rate of change is substantially higher than in previous technological waves. Institutions may lag behind the market more than before.

From the standpoint of redistribution of revenues, roughly three types of intervention circulate in discussion. Taxation of AI-generated productivity — whether in the form of an automation tax, a tax on corporate profits flowing from AI acceleration, or an extension of the tax base toward returns on capital — is technically feasible but politically demanding, since firms operate transnationally and tax competition is intense. Labor law reform represents another possibility, though it requires not so much an extension of protection as a reconsideration of the logic of compensation. Existing compensation systems are built on measurable performance: hours worked, volume delivered, or position held. None of these metrics captures the difference between the worker who copied a prompt and left the result unchecked, and the one who spent an hour in careful iteration and stakes their professional reputation on the outcome. Law provides protection, but cannot price accountability as an independent component of the value of labor. New forms of redistribution, such as unconditional basic income, represent a more radical answer: decoupling part of the standard of living from labor market participation. The debate about UBI in the context of AI is not new (Van Parijs & Vanderborght, 2017), but the empirical evidence on its macroeconomic effects remains limited.

The question therefore is not only whether AI will increase productivity, but also who will be entitled to its revenues and who will bear the risks. TRCL as a principle can help define where human contribution remains economically justified — but it does not, of itself, resolve the distribution of revenues to those whose labor is displaced entirely. Naming this gap is a precondition for any serious policy debate.

11. The Central European Context

Small and medium-sized enterprises form the backbone of the Czech, Polish, and Slovak economies — in the Czech Republic, over 99 percent of business entities and nearly 60 percent of the workforce. When it comes to AI adoption, they face distinctive challenges: less capital, the absence of internal specialists, but also a different structure of work.

Labor in SMEs is less specialized and more dependent on individual accountability. The owner of a small firm naturally does everything EMMA describes: evaluation, monitoring, modification, and arbitration. While large firms will probably use AI to centralize decision-making, SMEs may use AI to reinforce individual autonomy. Empirical investigation of this paradox represents a promising research direction.

12. Conclusion

This idea paper does not propose a finished theory. It proposes a naming of a problem that most debates about AI circumvent: that the contemporary economy is severing the relationship between performance, accountability, and reward.

The EMA framework helps identify where an activity is located. The EMMA framework helps identify what a person responsibly does there. The Thesis on Responsible Cognitive Labor connects both frameworks into a working proposition: the value of human labor should reside in the degree of accountability assumed. And where accountability is institutionally anchored, this value will probably be retained longer.

All three concepts are proposals for discussion, not finished theories. EMA and EMMA are working tools with blurred boundaries. TRCL is a conceptual proposition with one testable prediction. The core of this paper is not in the frameworks themselves, but in the naming of a tension that most debates avoid: the problem is not only the automation of labor, but the decoupling of performance, accountability, and reward.

References

Acemoglu, D., & Restrepo, P. (2018). The race between man and machine: Implications of technology for growth, factor shares, and employment. American Economic Review, 108(6), 1488–1542. https://doi.org/10.1257/aer.20160696

Acemoglu, D., & Restrepo, P. (2019). Automation and new tasks: How technology displaces and reinstates labor. Journal of Economic Perspectives, 33(2), 3–30. https://doi.org/10.1257/jep.33.2.3

Aldasoro, I., Gambacorta, L., & Pál, R. (2026). AI adoption and firm-level productivity in Europe (BIS/EIB Working Paper). Bank for International Settlements. https://www.bis.org/publ/work1325.htm

Altay, S., & Gilardi, F. (2024). People are skeptical of headlines labeled as AI-generated, even if true or human-made, because they assume full AI automation. PNAS Nexus, 3(10), pgae403. https://doi.org/10.1093/pnasnexus/pgae403

Anthropic. (2026). Anthropic Economic Index: Reports IV. https://www.anthropic.com/research/anthropic-economic-index-january-2026-report

Autor, D. H., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. The Quarterly Journal of Economics, 118(4), 1279–1333. https://doi.org/10.1162/003355303322552801

Autor, D. H., Katz, L. F., & Kearney, M. S. (2006). The polarization of the U.S. labor market. American Economic Review, 96(2), 189–194. https://doi.org/10.1257/000282806777212620

BCG. (2025). AI at Work: Momentum Builds, but Gaps Remain. Boston Consulting Group. https://www.bcg.com/publications/2025/ai-at-work-momentum-builds-but-gaps-remain

Becker, G. S. (1964). Human capital: A theoretical and empirical analysis, with special reference to education. University of Chicago Press.

Bick, A., Blandin, A., & Deming, D. (2025). The rapid adoption of generative AI (Working Paper No. 2024-027C). Federal Reserve Bank of St. Louis / NBER. https://www.nber.org/system/files/working_papers/w32966/w32966.pdf

Bovens, M. (2007). Analysing and assessing accountability: A conceptual framework. European Law Journal, 13(4), 447–468. https://doi.org/10.1111/j.1468-0386.2007.00378.x

Brynjolfsson, E., Li, D., & Raymond, L. (2025). Generative AI at work. The Quarterly Journal of Economics, 140(2), 889–942. https://doi.org/10.1093/qje/qjae044

Durand, C. (2020). Techno-féodalisme: Critique de l'économie numérique. La Découverte.

EY. (2025). Work Reimagined Survey 2025. Ernst & Young Global Limited. https://www.ey.com/en_gl/insights/workforce/work-reimagined-survey

Harari, Y. N. (2017, February 24). The rise of the useless class. Ideas.ted.com. https://ideas.ted.com/the-rise-of-the-useless-class/

Hosseini, S. M., & Lichtinger, G. (2025). Generative AI as seniority-biased technological change: Evidence from U.S. résumé and job posting data. SSRN Working Paper. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5425555

Keynes, J. M. (1930). Economic possibilities for our grandchildren. In Essays in persuasion (pp. 321–332). London: Palgrave Macmillan UK.

Marx, K. (1867). Das Kapital: Kritik der politischen Oekonomie. Verlag von Otto Meissner.

Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative AI. Science, 381(6654), 187–192. https://doi.org/10.1126/science.adh2586

PwC. (2025). Global Workforce Hopes & Fears Survey 2025. https://www.pwc.com/gx/en/issues/workforce/hopes-and-fears.html

Schilke, O., & Reimann, M. (2025). The transparency dilemma: How AI disclosure erodes trust. Organizational Behavior and Human Decision Processes, 188, 104405. https://doi.org/10.1016/j.obhdp.2025.104405

Smith, A. (1776). An inquiry into the nature and causes of the wealth of nations. Strahan & Cadell.

Srnicek, N. (2017). Platform capitalism. Polity Press.

Van Parijs, P., & Vanderborght, Y. (2017). Basic income: A radical proposal for a free society and a sane economy. Harvard University Press.

Veblen, T. (1899). The theory of the leisure class: An economic study of institutions. Macmillan.

Weizenbaum, J. (1976). Computer power and human reason: From judgment to calculation. W. H. Freeman.

Wiedermann, T., Švejcar, J., Makarava, U., Nohejl, K., Krejsová, K., & Starý, D. (2025). Budoucnost českého pracovního trhu [The future of the Czech labor market]. BCG, MPSV, Aspen Institute Central Europe. https://www.bcg.com/publications/2025/czech-republic-budoucnost-ceskeho-pracovniho-trhu

Zhang, Y., & Gosline, R. (2023). Human favoritism, not AI aversion: People's perceptions (and bias) toward generative AI, human experts, and human-GAI collaboration in persuasive content generation. SSRN. https://doi.org/10.2139/ssrn.4453958