Distilling Knowledge Work

PETER ORMEROD*

Generative artificial intelligence has begun converting records of work into the work itself. Employers have long recorded their workers, but a century of surveillance produced only archives that couldn't perform anything. Now, though, AI systems trained on those records are reproducing the judgment they contain. Pairing records with processing transforms expertise: memorialized as work data and analyzed by a capable AI model, the capacity to perform knowledge work is becoming alienable and reproducible.

This Essay names this phenomenon the distillation of knowledge work, and it shows how this process has ignited a recursive conflict over who will cede and who will capture wealth in the age of AI.

Within firms, employers capture workers' judgments and corrections as proprietary training assets. But recording and processing need not occur in the same hands: the firm that surveils its workers typically rents AI processing capacity—the still—from a frontier AI lab. In search of a viable business model, the labs are attempting to burrow into the place where work is represented—the place where work data accrues, market opportunities become visible, and the intelligence can be monetized. So the enterprises eagerly distilling their workers' expertise are simultaneously threatened by the labs. It's little wonder, then, that the capital class is increasingly adopting the rhetoric of the proletariat.

Who will distill whom? This Essay shows how inherited legal endowments will prove instrumental to workers' dispossession, whereas the contest between firms and the labs is likely to be decided either by private ordering among the powerful or by one side enlisting the state. The labs' unconditional victory, however achieved, could summon a regulatory paradigm older than antitrust. But whatever happens in the contest above, the worker below holds no claim to any version of the learning loop her judgment fills.

The monetization of data about people defined the prior era of information capitalism. The monetization of work itself will define the next.

---

* Associate Professor of Law, Villanova University Charles Widger School of Law.

DRAFT: AUG. 2026
CONTACT: peter.ormerod@law.villanova.edu

CONTENTS

INTRODUCTION

In early 2026, an unassuming product announcement by the frontier artificial intelligence (AI) lab Anthropic triggered a trillion-dollar stock selloff concentrated among enterprise software vendors—a decline from which many have not recovered.[1] A few months later, the social-media company Meta undertook two related initiatives. It announced a comprehensive surveillance effort to capture its employees' keystrokes, mouse movements, and screens to generate data for its AI models,[2] and it also began reassigning engineers en masse to the drudgery of labeling that and other training data.[3] In response to employee outrage, one senior executive wrote, "There is no option to opt-out" of the surveillance,[4] and his lieutenant explained that data-labeling "transfers aren't optional."[5] A few months after that, the CEO of Microsoft—a company that has invested billions of dollars in Anthropic's arch-rival OpenAI and that holds the largest outside stake in it—penned an essay that inveighed against the labs.[6] He asserted that "every company across every sector is ceding value to a few [AI] models," and warned that "[t]here is no societal permission for an AI future that hollows out entire industries."[7]

The untrained eye might conclude that these are disparate episodes at distinct technology companies of different vintages. But in fact they are three episodes in the same battle—one that implicates the fundamental question of who will cede and who will capture wealth in the age of AI. This Essay outlines the contours of this conflict and describes the dynamics driving it. In doing so, it articulates a framework for understanding how the battle will be fought and for evaluating its combatants' prospects.

---

1. Chibuike Oguh, Danilo Masoni, and Medha Singh, *Selloff wipes out nearly $1 trillion from software and services stocks as investors debate AI's existential threat*, REUTERS (Feb. 4, 2026), reuters.com/business/media-telecom/global-software-stocks-hit-by-anthropic-wake-up-call-ai-disruption-2026-02-04.
2. Katie Paul & Jeff Horwitz, *Meta to start capturing employee mouse movements, keystrokes for AI training data*, REUTERS (Apr. 21, 2026), reuters.com/sustainability/boards-policy-regulation/meta-start-capturing-employee-mouse-movements-keystrokes-ai-training-data-2026-04-21.
3. Eli Tan, *Before Mass Layoffs, Meta Reassigns 7,000 Workers to Focus on A.I.*, N.Y. TIMES (May 18, 2026), nytimes.com/2026/05/18/technology/meta-reassigns-7000-employees-ai.html.
4. Kalley Huang, Eli Tan, and Kate Conger, *Meta's Embrace of A.I. Is Making Its Employees Miserable*, N.Y. TIMES (May 8, 2026), nytimes.com/2026/05/08/technology/meta-ai-employees-miserable.html.
5. Varsha Bansal, *Meta is rapidly reorganizing its workers' jobs around AI: 'Transfers aren't optional'*, THE GUARDIAN (May 19, 2026), theguardian.com/technology/2026/may/19/meta-jobs-ai-transfers.
6. Satya Nadella (@satyanadella), *A frontier without an ecosystem is not stable*, X (June 14, 2026), x.com/satyanadella/status/2066182223213293753 [hereinafter Nadella, *Frontier*]; Ashley Capoot, *OpenAI completes restructure, solidifying Microsoft as a major shareholder*, CNBC (Oct. 28, 2025), cnbc.com/2025/10/28/open-ai-for-profit-microsoft.html.
7. Nadella, *Frontier*, *supra* note 6.

Generative AI has created a novel method for transforming data into profit. In the prior, social-data era—of smartphones, social-media apps, personalized feeds, and targeted advertising—technology companies collected massive amounts of data about human activity and wrung predictive insights from population-level patterns of behavior.[8] Those predictive insights allowed them to engage in what I've called individualized differentiation—a data processor's offering of personalized products, services, experiences, and prices to its counterparties.[9] Those individualized offerings enabled the appropriation of surplus value for the firms that target advertisements, personalize prices, precisely tune risk profiles, and algorithmically modulate wages.[10] Though there can be little question those techniques will endure and extend into the AI era,[11] there is also something new afoot.

It has become increasingly clear that the companies that train and deploy frontier AI models see an acute business opportunity in knowledge work. OpenAI's ChatGPT Work and Anthropic's Claude Cowork are agentic AI systems built on an infrastructure pioneered for software development.[12] They are more than just chatbots. They enable the augmentation and automation of vast amounts of white-collar work performed on computers: from building and maintaining databases to drafting emails and documents to crafting slide decks and beyond.[13]

These systems are changing both the object and purpose of data processing. The paradigmatic input has changed from data about people to the data generated by and representing knowledge work: the judgments, corrections, procedures, feedback, and workflows that constitute human expertise. And the paradigmatic output has evolved from differentiated offerings to the replication and execution of work itself. Whereas social data yielded prediction value, work data yields execution value—the value of being able to do the work itself; whereas prediction value was monetized through differentiation, execution value is monetized by substitution. This process, I'll contend, is the distillation of knowledge work.

Employers have long recorded work, but a century of worker surveillance

---

8. *See* Amanda Parsons & Salomé Viljoen, *Valuing Social Data*, 124 Columbia L. Rev. 993 (2024).

9. *See* Peter Ormerod, *Regulating Data Monetization*, 13 Tex. A&M L. Rev. 1003 (2026) [hereinafter Ormerod, *Monetization*].

10. *See id.* at 1026–29.

11. *See, e.g.*, OpenAI, *New ways to buy ChatGPT ads* (May 5, 2026), openai.com/index/new-ways-to-buy-chatgpt-ads/.

12. OpenAI, *ChatGPT is now a partner for your most ambitious work* (July 9, 2026), openai.com/index/chatgpt-for-your-most-ambitious-work [hereinafter OpenAI, *Partner*]; Anthropic, *Claude Cowork Overview*, claude.com/docs/cowork/overview (last visited July 21, 2026).

13. *See* OpenAI, *Partner*, *supra* note 12; Anthropic, *Get started with Claude Cowork*, support.claude.com/en/articles/13345190-get-started-with-claude-cowork (last visited July 21, 2026).

produced only archives that could not perform anything.[14] Increasingly powerful AI systems, however, are now enabling employers' conversion of those records into the work itself.[15] Knowledge-work distillation has thus destabilized the prior equilibrium and ignited a two-level conflict over the environment where work is represented.[16] Within firms is the first level: employers seek to capture workers' processes and tacit knowledge as proprietary training assets.[17] Across firms is the second: the frontier-model providers supply the software scaffolding—context systems, agents, and evaluation infrastructure—through which AI performs work.[18] From the privileged position of selling intelligence, the labs gain the ability to commoditize their customers' expertise, collect rent on their workflows, and ultimately enter their markets.[19]

These are not two distinct conflicts occurring in parallel. Instead, they are the same conflict recurring at two successive levels of the economy: the enterprises distilling their workers' expertise are simultaneously having their own value siphoned away by their vendor—and potential competitor.[20] Value at both levels ultimately traces to the same source—human judgment enacted in work—but the workers who enact it are unlikely to capture it.[21] And the recursion between levels suggests the generalizability of a lesson from the social-data era: the ruthless logic of information capitalism means that value-extractive data processing eventually comes for everyone and everything.[22]

This Essay makes three contributions. Part 1 names and theorizes knowledge-work distillation, thereby extending the value taxonomy of the social-data era to the work-data present. Part 2 demonstrates that the two conflicts—within firms, between labor and management; and across firms, between the frontier labs[23] and incumbent enterprises—are in fact the same conflict. Commentators have so far treated them separately,[24] but the same structure recurs at both levels: though the data originates with the party being distilled, everything else—the environment where that data accrues, the data processing power, access to buyers, and favorable legal entitlements—belongs to the distilling party. That

---

14. *See* notes 55–58 and accompanying text, *infra*.
15. *See* Part 1.2.3, *infra*.
16. *See* Part 2, *infra*.
17. *See* Part 2.1, *infra*.
18. *See* Part 2.2, *infra*.
19. *See* Part 2.2.2, *infra*.
20. *See* Part 2.3, *infra*.
21. *See* Part 3.1, *infra*.
22. *See* Part 3.2, *infra*.
23. I refer herein to the likes of Anthropic and OpenAI as "frontier labs" not because they resemble a research laboratory in a meaningful sense, but rather to clearly distinguish between them and their enterprise opponents.
24. *See, e.g.*, Salomé Viljoen, Danielle Li, and Zoë B. Cullen, *Knowledge Guilds* (Jan. 12, 2026), danielle.li/assets/docs/KnowledgeGuilds.pdf; Mark McCarthy, *What happens when AI companies compete with their customers?*, BROOKINGS (Mar. 12, 2026), brookings.edu/articles/what-happens-when-ai-companies-compete-with-their-customers.

symmetry raises the ultimate question of who will distill whom, and Part 3 answers it with the legal lessons of the prior era: law will prove instrumental to workers' dispossession, the enterprise contest is wide open, and the labs' unconditional victory could summon a regulatory paradigm older than antitrust.

1. TWO WAYS TO EXTRACT VALUE FROM DATA

As the economy has transitioned from industrial to informational, profit-maximizing firms have pioneered techniques and business models for monetizing information. This Part first surveys the dominant approach of the past generation and then turns to how AI has created a new opportunity.

1.1 THE SOCIAL-DATA APPROACH

The proliferation of networked information technologies in the past 50 years has made the collection and processing of vast quantities of data trivially inexpensive.[25] Companies have long sought to profile their counterparties to leave less money on the table. And law has on occasion risen up in response: the first American consumer privacy statute was enacted in 1970, and it was precipitated by growing concern about the dossiers of information that credit profilers were amassing in secret databases.[26]

The runaway success of digitized technology rapidly accelerated that instinct. Companies in the past 20 years have pushed far beyond the risk profiling of credit reports, today collecting and processing untold quantities of information about human behaviors to pursue profit.[27] How do firms convert that social data into money? Amanda Parsons and Salomé Viljoen have explained that social data "produces value from its capacity to materialize and store traces of human activity," which confers the ability "to catalog, analyze, aggregate, and mine such traces for insight and, in turn, to use such insights to apprehend, predict, and modify behavior."[28] They coined the term "prediction value" to refer to "the value of being able to infer or more accurately predict future actions or effects" by systematically cultivating social data.[29]

More recently, I've explained that companies' dominant strategy for converting prediction value into exchange value—that is, monetizing data—is individualized differentiation.[30] The predictive insights gleaned from processing so-

---

25. *See generally* SHOSHANA ZUBOFF, THE AGE OF INFORMATION CAPITALISM (2019).
26. *See* Consumer Credit Protection Act, Pub. L. No. 91-508, 84 Stat. 1127 (1970) (codified as amended in scattered sections of 15 U.S.C.); CHRIS JAY HOOFNAGLE, FEDERAL TRADE COMMISSION PRIVACY LAW AND POLICY 271 (2016).
27. *See, e.g.*, ZUBOFF, *supra* note 25, at 8–12.
28. Parsons & Viljoen, *supra* note 8, at 1009–10.
29. *Id.* at 1013.
30. *See* Ormerod, *Monetization*, *supra* note 9, at 1026–29.

cial data allow firms to offer products, services, experiences, and prices that are customized, personalized, tailored, or otherwise differentiated between different counterparties—be they customers, users, suppliers, or employees.[31] These activities, I've argued, produce a differentiation premium: the difference between a one-size-fits-all offering and one that reflects the processor's ability to extract added surplus from apprehending and predicting their counterparty's willingness to buy or sell.[32] Concretely, the differentiation premium is the difference between the price a platform charges for a targeted versus untargeted advertisement or the difference between health insurance policies that do or do not discriminate based on preexisting conditions.[33]

An increasing number of online experiences are individually differentiated in this way—TikTok's "For You" page, Spotify's algorithmically generated playlists, Character.AI's companion chatbots, and so forth[34]—because doing so has proven incredibly profitable.[35]

1.2 THE WORK-DATA APPROACH

The rise of generative AI since ChatGPT debuted in November 2022 has precipitated a new means for converting information into profit.

1.2.1 FROM CHATBOTS TO AGENTS

Understanding how requires a brief technical primer. The software that frontier labs like OpenAI and Anthropic supply has evolved beyond a simple chatbot, but the large language model ("LLM" or simply the "model") remains at the center of their offerings.[36] An LLM is first pre-trained using vast quantities of data and computing power; that model is then post-trained and refined through a series of techniques, including reinforcement learning from human feedback (RLHF).[37] Everything that goes into or out of the models is reduced to a token—in a language model, a token is approximately a word or a four-letter portion thereof.[38] The model is served to users after training, and the users feed it

31. *Id.*
32. *See id.* at 1027.
33. *Id.*
34. *See id.* at 1027–28; Brett M. Frischmann & Peter Ormerod, *Regulating Manipulative Design Is Not Preempted by CDA 230 or the First Amendment*, 75 EMORY L.J. 1101, 1109–14 (2026).
35. *See* Ormerod, *Monetization*, *supra* note 9, at 1025 (recounting recent profits from targeted advertising).
36. *See, e.g.*, Khalid Abdelaty, *Anthropic vs. OpenAI: The Two AI Giants Compared*, DATACAMP BLOG (Mar. 28, 2026), datacamp.com/blog/anthropic-vs-openai.
37. *See* Ashish Vaswani et al., *Attention Is All You Need*, 31st Conference on Neural Information Processing Systems (2017), arxiv.org/abs/1706.03762; Toloka Team, *Complete guide to RLHF for LLMs: How human feedback shapes modern AI*, TOLOKA (Feb. 24, 2026), toloka.ai/blog/what-is-rlhf.
38. *See, e.g.*, OpenAI, *Learn about language model tokenization*, platform.openai.com/tokenizer

inputs (e.g., by asking it questions) and the model responds by producing outputs, a computational process known as inference.[39] The models are not, however, trapped in amber: the labs employ an ongoing feedback loop, further post-training their models using cues, signals, and feedback from users about what makes a useful output.[40]

Unlike other computer programs, the models operate probabilistically, rather than deterministically; their outputs are generated by repeatedly asking: given the training and everything that's come before, what token is most likely to come next?[41] The leading models have now been trained on virtually all of the publicly available text on the internet, much that is not, and vast quantities of synthetically generated data from other language models.[42] LLMs today are thus a distillation of a vast amount of human knowledge rendered in text.

Two follow-on developments are salient here. First, a "reasoning" or "thinking" model is one that employs additional inference-time compute to work through a problem and verify its outputs before presenting a final answer to the user.[43] This additional computation tends to improve performance on multi-step tasks, thereby laying the foundation for the second development—the harness.[44] A harness is the software that surrounds the model and provides it with additional functions and tools, including instructions, memory, additional output validation, permissions, and an execution environment.[45] Pairing a model and harness produces what the labs refer to as an "agent"—an integrated software system that is capable of taking complex, multi-step actions autonomously on behalf of the user.[46] In practical terms, the model is the agent's reasoning engine, whereas the harness is the surrounding machinery that gives it tools and

---

(last visited July 21, 2026).

39. *See, e.g.*, Nivetha Suruliraj, *What is LLM inference?*, IBM, ibm.com/think/topics/llm-inference (last visited July 21, 2026).

40. *See, e.g.*, Cen (Mia) Zhao et al., *Agent-in-the-Loop: A Data Flywheel for Continuous Improvement in LLM-based Customer Support*, EMNLP 2025 Industry Track submission (Paper #305) (Oct. 9, 2025), arxiv.org/abs/2510.06674v1.

41. *See, e.g.*, Matthew Burtell & Helen Toner, *The Surprising Power of Next Word Prediction: Large Language Models Explained, Part 1*, GEORGETOWN CSET (Mar. 8, 2024), cset.georgetown.edu/article/the-surprising-power-of-next-word-prediction-large-language-models-explained-part-1.

42. *See, e.g.*, Allen Pike, *LLMs Aren't Just "Trained On the Internet" Anymore* (May 31, 2024), allenpike.com/2024/llms-trained-on-internet/; Bartz v. Anthropic PBC, 787 F. Supp. 3d 1007, 1016 (N.D. Cal. 2025) (describing Anthropic's practice of acquiring millions of print books, "stripp[ing] the books from their bindings, cut[ing] their pages to size, and scann[ing] the books into digital form" for purposes of training their LLMs).

43. *See, e.g.*, Alyona Vert., *What is test-time compute and how to scale it?*, TURING POST (Feb. 5, 2025), turingpost.com/p/testtimecompute.

44. *See* Shubham Parashar et al., *Inference-Time Computations for LLM Reasoning and Planning: A Benchmark and Insights* (Feb. 18, 2025), arxiv.org/abs/2502.12521v1.

45. *See, e.g.*, Databricks Staff, *What is an AI Agent Harness?*, DATABRICKS DATA + AI FOUNDATIONS BLOG, databricks.com/blog/ai-harness (last visited July 21, 2026).

46. *See, e.g.*, Anthropic, *Measuring AI agent autonomy in practice* (Feb. 18, 2026), anthropic.com/research/measuring-agent-autonomy.

the ability to take action.

The labs' agentic systems were initially most useful for software development: OpenAI's Codex and Anthropic's Claude Code are software-development tools that streamlined software engineers' workflows.[47] No longer did those engineers need to use an LLM-based autocomplete or to copy and paste code from a chatbot's output; now they had an application that would write the code, run and test it, incorporate it into existing files, and ultimately deploy it.[48]

But it didn't take long before the labs took the infrastructure undergirding these coding agents and set them on other forms of knowledge work. By abstracting away some of the technical complexity from the user, the same agentic systems began autonomously building and maintaining spreadsheets, drafting and circulating documents, and crafting slide decks.[49]

1.2.2 THE PRIOR EQUILIBRIUM

These systems' proliferation in white-collar workplaces has fundamentally unsettled a long-term equilibrium between labor and management. Scholars and commentators have long identified a gap between what workers know about how work is done and the firm's ability to memorialize and formalize those practices.[50] This "tacit knowledge" is the byproduct of people knowing more than they can tell.[51] Tacit knowledge resists capture by firms because there is a great deal of day-to-day know-how that cannot be elicited via conventional interviews or written procedures.[52] Workers pass on tacit knowledge mostly by showing how things are done, rather than telling.[53]

By way of illustration, consider a senior litigation associate reviewing a junior associate's brief. The senior associate knows more than she can tell. Though her firm has a style manual and a brief bank, neither has ever produced a winning filing. What she knows—which arguments to lead with; what a particular judge will tolerate; when an aside should reside in the text, move to a footnote, or be cut altogether—she conveys by redlining and returning the draft. The junior associate learns by observing and internalizing her edits, not by being told.

This tacit knowledge, both legally and practically, has long resided with the worker: black-letter law in all fifty states provides that an employee's general

---

47. *See* Anthropic, *Claude Code Docs Overview*, code.claude.com/docs/en/overview (last visited July 21, 2026); OpenAI, *Introducing Codex* (May 16, 2025), openai.com/index/introducing-codex.

48. *See, e.g.*, George William Amalan, *Coding With AI in 2026: From Autocomplete to Autonomous Engineering*, CodeToDeploy (Feb. 15, 2026), medium.com/codetodeploy/coding-with-ai-in-2026-from-autocomplete-to-autonomous-engineering-a601ad1b01ba.

49. *See* note 12 and accompanying text, *supra*.

50. *See, e.g.*, Viljoen et al., *supra* note 24, at 9–14.

51. *See* Michael Polanyi, The Tacit Dimension (1966).

52. *See id.*; Viljoen et al., *supra* note 24, at 13–14.

53. *See* Polanyi, *supra* note 51.

knowledge, skill, and experience are not the employer's trade secrets, and former employees retain the right to use those general skills, knowledge, training, and accumulated expertise at a new employer.[54] When our senior associate leaves for a rival firm, her knowledge, skills, and experience go with her.

The ambition to capture the labor process is a century old. Frederick Winslow Taylor's scientific management was an explicit program of "gathering together all of the traditional knowledge which in the past has been possessed by the workmen" and reducing it "to rules, laws, and formulæ."[55] Frank and Lillian Gilbreth filmed workers' motions at fractions of a second to extract the "one best way" to perform any task.[56] Those methods succeeded against manual work—motion could be recorded, and machines could reproduce motion—but they stalled at judgment, which could be neither filmed nor articulated.[57] The 1980s' "expert systems" were the attempt to employ the same playbook on cognition, and they collapsed because experts could not translate their reasoning into the if-then rules the systems required—and, sensing the stakes, many experts sabotaged the effort—and the brittle rules that emerged could not reliably perform outside typical cases.[58]

Elicitation was not the employer's only impediment, though. Even what they could record could not actually perform the work. Our senior associate's firm has always had records—a document-management system holding all her past briefs and redlines—but those records could not themselves produce the brief. The worker's retention of capacity was thus overdetermined, and a century of intensifying surveillance could not automate her away.

1.2.3 KNOWLEDGE-WORK DISTILLATION

Agentic AI systems are destabilizing this arrangement. Because the systems learn from their users' demonstrations, comparisons, corrections, rankings, and other feedback, tacit knowledge that previously resisted capture by management is becoming increasingly legible. AI is thus succeeding where the interview-based expert systems failed, because LLMs learn like apprentices do—

---

54. *See* Camilla A. Hrdy, *The General Knowledge, Skill, and Experience Paradox*, 60 B.C. L. Rev. 2409 (2019).

55. Frederick Winslow Taylor, The Principles of Scientific Management 36 (1911)

56. Frank Gilbreth, Motion Study (1911).

57. Harry Braverman, Labor and Monopoly Capital 79 (1974).

58. Mark A. Musen & Johan Van der Lei, *Of Brittleness and Bottlenecks*, 7 Machine Intelligence & Pattern Recognition 335 (1988) (the knowledge-acquisition bottleneck); Edward Feigenbaum, Pamela McCorduck, and H. Penny Ni, The Rise of the Expert Company (1988); Viljoen et al., *supra* note 24, at 12–14 (vividly recounting this history); *see also id.* at 14 ("Am I supposed to give this damn system MY knowledge so I can be replaced by a computer? No way!") (quoting Rafe Brena, *What Happened with Expert Systems?*, Towards Data Science (Apr. 27, 2024), towardsdatascience.com/what-happened-with-expert-systems-aad399eab180).

from showing. Consider our senior associate: the junior associate's draft is now the agent's draft, and her redline is no longer only pedagogy. It is a demonstration pair—the model's output and the expert's correction—that shares the same learning loop that makes RLHF so effective. Every markup can teach the system what a good brief looks like to her. It's worth emphasizing that this remains true whether or not AI precipitates large-scale job displacement. Even if these systems merely increase workers' productivity, the systems' integration into workflows and the resulting datafication of how work is performed will enable someone else to harvest those productivity gains.

An increasing chorus of scholars has noted how these systems thus threaten workers. Ifeoma Ajunwa contends that deployment of AI in the workplace constitutes "captured capital—that is, the coercive collection and use of worker data to facilitate workplace automation."[59] Salomé Viljoen, Danielle Li, and Zoë B. Cullen argue that labor generates "knowledge data," records of workers' expertise, which firms are capturing and using to train AI systems.[60] This produces "knowledge capital": knowledge that was traditionally possessed by workers is becoming capital possessed by firms.[61]

Those accounts have focused on the capture and named the resulting asset. This Essay, by contrast, separates those stages—the capture and its conversion—because the same party need not conduct both.[62] That separation shows that the conflict need not stop at the firm's edges. I'll therefore use distinct terminology to refer to these processes and phenomena, which together constitute a framework that exposes a three-stage, multi-party process of datafying human judgment, using a machine to replicate it, and bringing the replica to market.

"Work data" consists of records that encode how work is performed. It contains traces of enacted human judgment, just as social data contains traces of human activity.[63] Work data is generated at every scale—from an individual's demonstrations and corrections (the senior associate's individual redlines) to an organization's evaluations, workflows, and institutional memory (the firm's brief bank and analyses of what persuades which judges).

"Execution value" is the value of the capacity to perform work. Just as prediction value is the value of being able to predict a counterparty's behavior,[64] execution value is the value of being able to do the work itself. That value has always been monetized through wages and mobility by the workers who hold the capacity to perform. And much of that value, in the form of tacit knowledge, has

59. Ifeoma Ajunwa, *AI and Captured Capital*, 134 YALE L.J. F. 372 (Jan. 31, 2025) (internal quotation marks omitted).
60. *See* Viljoen et al., *supra* note 24, at 1.
61. *See id.* at 1–2.
62. Indeed, as Part 2 will show, the final stage—monetizing the distillate—may belong to yet another party.
63. *See* Parsons & Viljoen, *supra* note 8, at 1009–10.
64. *See id.*

long resisted capture by firms; literally embodied in human expertise, it has been inalienable.[65]

Recording alone never threatened that arrangement. Employers have long surveilled their workers—the call captured for quality assurance and keystroke logging—but nothing could turn a transcript into the performance of work.[66] Data processing disrupts that settled ground, since an AI system trained on records of work can increasingly reproduce the judgment they contain. It is therefore the recording *paired* with the processing that changes the capacity's state: memorialized as work data and analyzed by a capable model, the ability to perform work becomes alienable and reproducible. Our senior associate's years of archived briefs were inert; joined to a stream of her corrections and processed by an AI model, they become an engine that drafts in her hand. And because work data accumulates where work leaves a record—the digitized memorialization of judgments, corrections, and workflows—the entity supplying that environment, which I'll call the representational layer, possesses valuable raw material.

The use of AI systems to transform records of work into its performance—and thereby to appropriate execution value from those who previously performed the work—is the "distillation of knowledge work." In its paradigmatic form, knowledge-work distillation is the AI-enabled breach of the bargain between labor and capital.[67] The wage bargain compensated the outputs of work, but the capacity to perform it remained with the worker: even what a firm could codify could not perform the work, and the tacit remainder simply could not be captured. Distillation reaches the capacity itself, sweeping in both what workers could have written down and what they never could have, and enabling the appropriation of execution value by those who do not perform the work.

In short, then, records of work (work data) are the raw material—the mash; the AI system is the still; the distillate is the capacity to perform, concentrated and portable; the execution value is the price that distillate will fetch; the distillate is monetized where the work it replaces is sold; and those who run the apparatus and control the point of sale capture that value. "Distillation" is, after all, machine learning's own name for the act of absorbing one model's capabilities into another—a practice that the leading frontier labs' terms of service forbid, even as the human-to-machine equivalent is the entire enterprise.

In the social-data approach, business practices motivated by prediction value modulate the terms of a transaction between a data processor and its counterparty. But knowledge-work distillation is monetized by substitution: the replacement—threatened or realized—of those who perform the work, which allows another to capture value previously tied to the performer. Substitution,

65. *See* POLANYI, *supra* note 51.
66. *See* Part 1.2.2, *supra*.
67. *See* Viljoen et al., *supra* note 24, at 10, 13.

however, requires more than the distillate: the substitute must find a buyer, though that requirement goes unnoticed within the firm because the employer is that buyer. Concretely, the firm now needs fewer associates, or the same associates produce three times the output and the firm stands to gain much of the difference. And note what has happened to the senior associate: when she leaves, the trade-secret carve-out still protects her right to take her general skill and experience to a rival, but a distilled copy of her skills stays behind. Knowledge-work distillation is thus a distinct form of data monetization from the prior approach in both its object and its purpose, and it's only made possible by increasingly capable AI systems.

Many software companies have now realized there is an enormous business opportunity in providing the software tools that enable and facilitate the distillation of knowledge work. But recording and processing rarely happen in the same hands: the firm that records its workers typically cannot distill them without help and thus relies on vendors to supply the models and the environments. In our litigation example, the drafting, corrections, and evaluations went through a model and software environment the firm leases from an AI provider. The firm that distills its workers' knowledge therefore does not necessarily keep what it extracts. And who distills whom is the subject of the next Part.

2. RECURSIVE CONFLICT OVER KNOWLEDGE-WORK DISTILLATION

The ability to distill knowledge work has ignited conflict at two successive levels of the economy.

2.1 INTRA-FIRM DISTILLATION OF WORKERS

In April 2026, Meta began installing new comprehensive tracking software on U.S.-based employees' computers.[68] This tool, called the Model Capability Initiative (MCI), captured all mouse movements, clicks, and keystrokes, and it regularly took screenshots to capture the content displayed on the screen.[69] The purpose, an internal memo explained, was to aid the company's foundering efforts to develop its own frontier AI model by providing valuable training data about how humans interact with computers.[70] Responding to employee outrage, the company's CTO flatly explained: "There is no option to opt-out."[71] MCI was part of a broader effort to reorganize the company around AI. Around the same time, management laid off ten percent of the workforce[72] and began compulsorily reassigning huge portions of the engineers who remained to a training-data-

68. *See* Paul & Horwitz, *supra* note 2.
69. *Id.*
70. *Id.*
71. *See* Huang et al., *supra* note 4.
72. *Id.*

labeling initiative it called Agent Data Optimization.[73] As a result of these initiatives, internal morale hit "rock bottom"—employees derided senior leaders with expletive-laden tirades, media leaks accelerated, and workers described their jobs in stark terms like "soul-crushing" and "literally the gulag."[74] Some attempted to organize opposition: over 1,600 employees signed an internal petition to halt the changes[75]—even in the face of termination risk and what one employee referred to as Meta's "extreme culture of fear"[76]—and others abroad began the formal process of unionizing.[77] As of this writing, MCI has collapsed under its own weight—"paused" not by law or by the petition, but instead after it came to light that any employee could access the MCI-derived databases.[78]

The drama at Meta is but one illustration of the ongoing, protracted struggle—previewed in Part 1.2.2—over management's efforts to capture and replicate workers' contributions to the enterprise. AI systems, however, are fundamentally shifting the ground in management's favor because of their ability to distill knowledge work.

Sensing the opportunity, many other executives have undertaken efforts to use AI to exert ever-greater pressure on their workforces. For example, Vercel is a cloud-based platform for software developers to build and distribute websites and applications.[79] Senior leadership at Vercel assigned three engineers to shadow the top performer in a sales department of ten employees.[80] After documenting that top performer's every move for six weeks, the engineers used an AI model to build a software agent that later allowed the company to reduce the department to a single employee who oversees the AI agents.[81] As one Vercel vice president put it, "Modeling after the top-performing employees has always been a standard business practice. The difference now is that technology lets us accelerate it."[82]

---

73. *See* Tan, *supra* note 3.
74. Marc Vartabedian, *Toxic mix of chaos and drudgery turns Meta's AI unit into a real-world hell: 'Soul-Crushing'*, N.Y. Post (June 15, 2026), nypost.com/2026/06/15/business/toxic-mix-of-chaos-and-drudgery-turns-metas-ai-unit-into-a-real-world-hell-soul-crushing.
75. *See* Katie Paul, *Meta lays out details of May 20 restructuring in internal document*, Reuters (May 18, 2026), reuters.com/world/meta-lays-out-plans-may-20-layoffs-restructuring-internal-document-says-2026-05-18; Diana Corker, *Meta's Employee Tracking Program Triggers Privacy and Trust Crisis*, HR Digest (June 25, 2026), thehrdigest.com/metas-employee-tracking-program-triggers-privacy-and-trust-crisis.
76. Bansal, *supra* note 5.
77. *Id.*
78. Paresh Dave, *Meta Pauses Employee-Tracking Program Following Internal Data Leak*, Wired (June 22, 2026), wired.com/story/meta-pauses-employee-tracking-program-following-internal-security-breach.
79. Lakshmi Varanasi, *Vercel trained an AI agent on its best salesperson. Then it cut the 10-person team down to 1.*, Bus. Insider (Oct. 27, 2025), businessinsider.com/ai-agent-entry-level-sales-jobs-vercel-2025-10.
80. *Id.*
81. *Id.*
82. *Id.*

Researchers have found that AI can increase the productivity of existing workers, though those effects are unevenly felt. A recent study of the introduction of AI to customer-support call centers found that while access to AI assistance increased worker productivity in the aggregate, its distributive effects were stark.[83] The AI's training leaned heavily on the transcripts generated by top-performing workers, but those workers obtained little benefit in return: they became a tiny bit quicker but the quality of their work declined, whereas the least experienced and least skilled employees dramatically improved both the speed and quality of their performance.[84] AI-enabled distillation thus benefited the business but provided little or no corresponding benefit to the workers responsible for those gains.

Of course, not every worker is similarly exposed. Sectors dominated by physical labor and those that emphasize rare craftsmanship are unlikely to experience widespread disruption anytime soon.[85] But there can be little doubt that white-collar knowledge work performed on computers faces an uncertain future, with software engineering being a particularly hard-hit early bellwether.[86] Executives in Silicon Valley—especially among the frontier labs—have been warning about labor disruption for years,[87] and public opinion polls consistently show widespread fear and anxiety about what AI means for the future of work.[88]

For their part, some workers appear ready to resist management's efforts to capture their judgments and workflows. Zoë Cullen, Danielle Li, and Shengwu Li found that workers report considerable discretion over the breadth and depth of the data they supply about their work and that most claim they possess both the ability to withhold that data and supply more.[89] In a follow-on experiment, the same researchers found that workers' awareness that their data would be used to train AI systems decreased their willingness to supply such data.[90] These findings together indicate that workers have at least some control over the cap-

83. Erik Brynjolfsson, Danielle Li, Lindsey Raymond, *Generative AI at Work*, 140 QUARTERLY J. ECON. 889 (Feb. 2025), academic.oup.com/qje/article/140/2/889/7990658.
84. *Id.*
85. *See, e.g.*, Callie Holtermann, *Mom, Dad, I Want to Be a Welder*, N.Y. TIMES (July 13, 2026), nytimes.com/2026/07/13/style/gen-z-trade-school-careers-ai-college.html.
86. Ana Altchek, *AI hit software engineers first. Here's what they want you to know.*, BUS. INSIDER (Apr. 19, 2026), businessinsider.com/software-engineers-lessons-white-collar-works-ai-disruption-2026-4.
87. *See, e.g.*, Jim VandeHei & Mike Allen, *Behind the Curtain: A white-collar bloodbath*, AXIOS (May 28, 2025), axios.com/2025/05/28/ai-jobs-white-collar-unemployment-anthropic; Jasmine Sun, *Silicon Valley Is Bracing for a Permanent Underclass*, N.Y. TIMES (Apr. 30, 2026), nytimes.com/2026/04/30/opinion/ai-labor-work-force-silicon-valley.html.
88. *See* Jason Lange & Courtney Rozen, *Half of Americans fear AI could put someone in their household out of work, Reuters/Ipsos poll finds*, REUTERS (June 10, 2026), reuters.com/business/world-at-work/half-americans-fear-ai-could-put-someone-their-household-out-work-reutersipsos-2026-06-10/.
89. Zoë Cullen, Danielle Li, and Shengwu Li, *Labor as Capital: AI and the Ownership of Expertise* (March 30, 2026), zoecullen.com/assets/docs/Labor_as_Capital_WP.pdf.
90. *Id.*

ture side of the distillation equation and that some may try to wield that control to undermine AI-motivated surveillance, echoing the expert-system sabotage of the 1980s.[91]

But knowledge workers' resistance faces a steep uphill climb for two independent reasons. For one, selling expertise is a prisoner's dilemma, since the holdouts face the ongoing risk that others will supply the data in their stead. Indeed, there is now a burgeoning market for selling expertise directly to the model developers.[92] And second, the legal backdrop for their resistance is particularly unfavorable. Upon surveying the legal landscape governing worker surveillance amid an increasingly digital workplace, Ifeoma Ajunwa, Kate Crawford, and Jason Schultz concluded that expanding technological abilities threaten to leave American workers at the mercy of constant, ongoing monitoring by their employers.[93] Matthew Bodie has similarly explained that employee data is generally governed by two distinct legal regimes—privacy law and intellectual property law—that together leave workers "virtually enmeshed within their businesses—without control, ownership, or regulatory protection."[94]

In short, the history of worker surveillance and law's role in facilitating it suggests that little stands in the way of employers committed to capturing their employees' data. Those employers may find, however, that dynamics other than legal obstacles counsel against the rapid, unfettered distillation of their workers. Meta, for instance, is in the unusual position of having a great deal of high-quality training data readily available and also in pursuit of developing a frontier model. Few others occupy that position. Other firms' distillation efforts invariably require them to rent the still from the labs developing the cutting-edge models. That means the conflict over management's impulse to distill its workers does not stop at the boundary of the firm, because the labs themselves appear increasingly committed to distilling not just labor but entire industries.

2.2 INTER-FIRM DISTILLATION OF ENTERPRISES

The past year has revealed that the frontier AI labs are locked in an existential struggle against the very firms that buy their intelligence, and the software companies that constitute the incumbent representational layer are the first, most visible front of that war.

91. *See* note 58 and accompanying text, *supra*.
92. Lora Kelley, *The Work of Helping A.I. Destroy Work*, N.Y. TIMES (July 10, 2026), nytimes.com/2026/07/10/business/ai-white-collar-jobs.html.
93. Ifeoma Ajunwa, Kate Crawford, and Jason Schultz, *Limitless Worker Surveillance*, 105 CAL. L. REV. 735 (2017).
94. Matthew T. Bodie, *The Law of Employee Data: Privacy, Property, Governance*, 97 IND. L.J. 709, 711 (2022).

2.2.1 THE LABS' SHIFTING STRATEGY

In late January 2026, Anthropic announced new functionality in its Cowork application. Recall that Cowork is Anthropic's agentic software offering for knowledge work—its application that autonomously works with the user's files and other work data.[95] The new functionality, called Plugins, allows users to "tell Claude how you like work done, which tools and data to pull from, [and] how to handle critical workflows."[96] The legal plugin, for example, showed that an AI system could independently manage contract review, non-disclosure-agreement triage, and compliance briefings—all tasks that had long required expensive, specialized enterprise software.[97]

Investors' response to Cowork Plugins was unmistakable: the stock of three legal enterprise software providers—Thomson Reuters, RELX, and Wolters Kluwer—dropped between 13 and 16 percent.[98] This episode precipitated what analysts have called the SaaSpocalypse: widespread investor fear that the frontier labs stand to subsume many software as a service (SaaS) companies—those that sell costly, niche software to enterprises on a per-seat basis.[99] Cowork Plugins triggered a broad stock sell off, with software companies specifically falling further and recovering far later than the broader market.[100]

Analysts reasoned that the per-seat licensing business model is fundamentally threatened by the prospect of autonomous AI agents performing repetitive, computer-based tasks rather than humans.[101] For one, fewer humans doing the work will mean there are fewer people who need software licenses.[102] But more fundamentally, it was increasingly murky why enterprises would need SaaS companies at all.[103] After all, why not just go straight to a frontier lab and have its

95. *See* note 12 and accompanying text, *supra*.
96. Anthropic, *Customize Cowork with plugins* (Jan. 30, 2026), claude.com/blog/cowork-plugins.
97. *See id.*
98. Rob Hales, *Thomson Reuters, RELX, and Wolters Stocks Crushed After Anthropic Debuts Claude Legal Plug-In*, MORNING STAR (Feb. 4, 2026), morningstar.com/stocks/reuters-relx-wolters-stocks-crushed-after-anthropic-debuts-claude-legal-plug-in.
99. Lynn Doan & Carmen Reinicke, *What's Behind the 'SaaSpocalypse' Plunge in Software Stocks?*, BLOOMBERG (Feb. 4, 2026), bloomberg.com/news/articles/2026-02-04/what-s-behind-the-saaspocalypse-plunge-in-software-stocks.
100. *See, e.g.*, Peter Cohan, *The Stock Market's Biggest Winners And Losers Of 2026 So Far*, FORBES (June 8, 2026), forbes.com/sites/investor-hub/article/stock-market-winners-losers-2026.
101. Eric Maltiel & Josh Sandberg, *Per-Seat Software Pricing Isn't Dead, but New Models Are Gaining Steam*, BAIN & CO. (Oct. 9, 2025), bain.com/insights/per-seat-software-pricing-isnt-dead-but-new-models-are-gaining-steam.
102. *See* Ben Thompson, *Microsoft and Software Survival*, STRATECHERY (Feb. 3, 2026), stratechery.com/2026/microsoft-and-software-survival.
103. *See* Alistair Barr, *AI coding tools upend the 'buy versus build' software equation and threaten the SaaS business model*, Bus. Insider (June 25, 2025), businessinsider.com/ai-coding-tools-buy-versus-build-software-saas-netlify-bolt-2025-6.

AI build a bespoke software solution for the customer's unique uses—solutions designed from the ground up for AI agents?[104] The SaaSpocalypse was thus a concrete signal that investors expect frontier labs to displace incumbent enterprise-software providers.

It's not just legacy enterprise software companies at risk. There is also an entire ecosystem of AI-native startups that are built on top of the frontier labs' AI models.[105] In the legal space, for example, Harvey and Legora are two companies whose businesses exist at the application layer, the industry's term for software's user touch-point.[106] Harvey and Legora serve as an intermediary between the labs' models and customers' specialized end-use: in the case of Harvey and Legora, AI software for legal professionals that is built atop the underlying AI models. Theirs is an unenviable position because of their ongoing dependence on programmatic access (via what's called an API) to the labs' AI models.[107] Anthropic and OpenAI's movement into the application layer thus puts the likes of Harvey and Legora—and an untold number of other AI startups—at the same risk as legacy SaaS businesses: that the labs will come for their customers directly and cut them out of the equation altogether.

What, you may be wondering, is driving the labs into the application layer, ever closer to end users, especially since doing so promises to destroy the labs' own commercial customers?[108] It increasingly appears that the business of developing and serving frontier models may not be a sound one.[109] For one, the costs are proving astronomical.[110] Those costs include large fixed ones, like any software business.[111] But what sets AI apart from the economics of other software is its marginal costs—the ongoing cost of electricity necessary to produce inference tokens.[112]

---

104. *See* Vaibhav Bhatnagar, *How is AI Replacing SaaS? What's Next for Enterprise Software*, HEXWARE BLOG (Mar. 13, 2026), hexaware.com/blogs/how-is-ai-replacing-saas-whats-next-for-enterprise-software.
105. *See* Tim Tully et al., *2025: The State of Generative AI in the Enterprise*, MENLO VENTURE CAPITAL (Dec. 9, 2025), menlovc.com/perspective/2025-the-state-of-generative-ai-in-the-enterprise.
106. Tim Dotan, *Harvey & Legora in a Land-Grab Race for Dominance of Legal AI*, NEWCOMER (Apr. 1, 2026), newcomer.co/p/harvey-and-legora-in-a-land-grab.
107. Hence Harvey has announced an effort to train its own models. *See* Rhys Dipshan, *Harvey Announces Development of Custom Legal-Specific AI Models*, LAW.COM (June 18, 2026), law.com/legaltechnews/2026/06/18/harvey-announces-development-of-custom-legal-specific-ai-models.
108. *See generally* McCarthy, *supra* note 24.
109. *See, e.g.*, *id.*
110. In 2025, McKinsey estimated that the cumulative AI-specific data-center capital expenditures by 2030 would total $5.2 trillion. *See* Jesse Noffsinger et al., *The cost of compute: A $7 trillion race to scale data centers*, MCKINSEY QUARTERLY (Apr. 2025), mckinsey.com/industries/technology-media-and-telecommunications/our-insights/the-cost-of-compute-a-7-trillion-dollar-race-to-scale-data-centers.
111. *See, e.g.*, Peter Ormerod, *Privacy Law's Incumbency Problem*, 58 U.C. DAVIS L. REV. 179, 185–86, 212–13 (2024) (describing the economics of software development).
112. *See, e.g.*, Ben Thompson, *Who's Afraid of Chinese Models?*, STRATECHERY (July 20, 2026),

To develop and release a new cutting-edge model thus requires a huge amount of investment, but there is growing concern that those fixed costs cannot be recouped by the labs because their models may be commodities.[113] Over time, other companies and providers release models that compete with OpenAI's and Anthropic's leading offerings.[114] These competitors' models are not just cheaper than the frontier labs'; they are also open-weight, which affords their users considerable additional flexibility and customization.[115] While competitors' models have always lagged the frontier in performance, it's not clear that many customers need the premier models at the bleeding edge, and open-weight's promise of independence from the labs is itself alluring.[116] In other words, developing and serving the model may be a commodity business—one where many offerings converge in capability, making it all but impossible for any one provider to distinguish itself from competitors.[117] Commodity markets are by definition fiercely competitive, and that competition drives prices down to marginal costs.[118] In short, the labs appear increasingly concerned that they need to fundamentally change their business—finding a way to distinguish themselves from cheaper, open alternatives—in order to recoup their fixed costs and thereby attain profitability.[119]

The possibility that the model layer will not provide a durable advantage is thus the underlying dynamic that explains the labs' movement into the application layer, which, in the context of knowledge work, is what I've been calling the representational layer—the place where work is represented. And burrowing into that layer and thus the day-to-day workflows of enterprise customers is—as the legacy SaaS sector illustrates—a highly lucrative and sticky opportunity, because doing so ensures that the labs' customers' switching costs rise precipitously, thereby reducing the risk that the customers will flee to alternative providers.

2.2.2 FROM MASH TO MARKET

Motive alone is not sufficient. But the labs also have the opportunity.

Stated in terms of this Essay's framework, the labs own the still and they've been leasing it to their customers, but business imperatives have led them to reconsider that strategy. After all, why lease the still when another opportunity—

---

stratechery.com/2026/whos-afraid-of-chinese-models [hereinafter Thompson, *Chinese Models*].

113. *See* McCarthy, *supra* note 24.

114. *See id.*

115. *See* Thompson, *Chinese Models*, *supra* note 112.

116. *See, e.g.*, Rebecca Bellan, *The real AI race may no longer be at the frontier*, TECHCRUNCH (July 14, 2026), techcrunch.com/2026/07/14/the-real-ai-race-may-no-longer-be-at-the-frontier-open-models-hugging-face.

117. *See* McCarthy, *supra* note 24.

118. *See* Thompson, *Chinese Models*, *supra* note 112.

119. *See* McCarthy, *supra* note 24.

obtaining the mash, producing the distillate, and selling it directly—is far more lucrative? Doing so requires three things. The labs need the mash: records of how work is performed. They need to know which work is worth distilling: where the margins lie among the endless variety of their customers' workflows. And they need the point of sale: because distillation is monetized by substitution, the labs must reach the customers of those they seek to replace. Supplying the layer where work is represented—context systems, agents, evaluations, and interfaces—achieves all three at once: work data accrues in the environment its supplier controls; demand becomes visible in the aggregate; and the environment itself becomes the channel through which the distillate is deployed and sold. Whoever controls the representational layer is thus positioned to control capture, processing, and commercialization. It promises to be a highly differentiated, difficult-to-dislodge, and extremely profitable place to be.

Take the mash first. The labs do not have direct, unfettered access to either the inputs (prompts) or outputs (responses) of their commercial and API customers. The labs have maintained that by default they do "not use your inputs or outputs from our commercial products . . . to train our models."[120] This policy is in stark contrast to their policies for consumer accounts, which involve the default sharing of inputs and outputs for training purposes.[121] But that's hardly the end of the matter, because the labs have both direct and indirect means for obtaining the raw work data for processing.

The labs have several distinct ways that they do obtain direct access to enterprises' usage data, notwithstanding their formal policies. For one, the labs offer their customers rich incentives to provide the data. OpenAI, for example, offers millions of free tokens each day to enterprise customers that agree to share input and output data.[122] For another, there is an enormous amount of enterprise usage that is governed by the permissive consumer policy rather than the restrictive commercial one. Termed "shadow AI" usage, this occurs when employees use their personal ChatGPT and Claude accounts for work purposes. One study found that workers from over 90 percent of surveyed companies reported regularly using their personal chatbot accounts for daily work tasks,

120. Anthropic, *Is my data used for model training?*, privacy.claude.com/en/articles/7996868-is-my-data-used-for-model-training (last visited July 21, 2026) [hereinafter Anthropic, *Commercial Data Policy*]; *see also* OpenAI, *Sharing feedback, evaluation and fine-tuning data, and API inputs and outputs with OpenAI*, help.openai.com/en/articles/10306912-sharing-feedback-evaluation-and-fine-tuning-data-and-api-inputs-and-outputs-with-openai (last visited July 21, 2026) [hereinafter OpenAI, *Commercial Data Policy*] ("By default, we don't use any inputs or outputs from our products for business users, including ChatGPT Business, ChatGPT Enterprise, and the API, to improve our models.").

121. *See, e.g.*, OpenAI, *How your data is used to improve model performance*, help.openai.com/en/articles/5722486-how-your-data-is-used-to-improve-model-performance (last visited July 21, 2026) ("ChatGPT, for instance, improves by further training on the conversations people have with it, unless you opt out.")

122. *See* OpenAI, *Commercial Data Policy*, *supra* note 120.

typically without approval from, oversight by, or knowledge of their IT departments.[123] Finally, enterprises can't help but share the most valuable data of all: corrections and explicit feedback on outputs. Every time an enterprise user presses the thumbs up/down button, flags a bad response, reports a bug, or responds to the labs' frequent feedback solicitations, Anthropic proceeds to "store the entire related conversation, including any content, custom styles, conversation preferences, or model settings, . . . for up to 5 years."[124] As discussed above,[125] this correction and feedback data has outsized importance to the labs because it encodes human judgment through showing the models how to improve—of a kind with what the labs use to post-train their models using RLHF.

The labs also have an indirect, substitutive means for obtaining valuable data for training. As flagged earlier, there is a burgeoning industry committed to generating valuable expertise data and selling it directly to AI providers.[126] The startup Mercor is a notable illustration.[127] Every day, the company pays 30,000 contractors more than $4 million to generate expert data for AI training.[128] The labs no longer need "armies of low-paid workers, often overseas, to do rote tasks like tag images of cars or transcribe audio."[129] Rather, they now seek "mathematicians to annotate proofs, lawyers to mark up briefs and professors to grade essays," and that is precisely what Mercor offers.[130] Research has suggested that this data is incredibly useful: a lab founded by a former OpenAI senior executive demonstrated that specially training a cheap, open-weights model with a curated dataset of expert-inflected insights dramatically improved its performance—above that of frontier general-purpose models and at a fraction of their cost.[131] The distillation of high-margin expertise requires high-quality data, and the labs have ready access to it outside the constraints imposed by their customers.

The labs thus use these channels to obtain the mash for their distillery, but that alone is not enough. They also need the representational layer, which provides visibility into which enterprise workflows are ripe for the taking. Here too, the labs are moving aggressively to supply that environment directly to end

123. Aditya Challapally, *The GenAI Divide: State of AI in Business 2025*, MIT NANDA, at 8 (July 2025), mlq.ai/media/quarterly_decks/v0.1_State_of_AI_in_Business_2025_Report.pdf.

124. Anthropic, *Commercial Data Policy*, *supra* note 120. OpenAI's policy is similar, describing content shared as "the conversation up to that point (including inputs, outputs, and files uploaded)." OpenAI, *Commercial Data Policy*, *supra* note 120.

125. *See* Part 1.2.3, *supra*.

126. *See* note 92 and accompanying text, *supra*.

127. *See generally* Jacob Brazeal & Derek Shimozawa, *When you go from $2 million a month to $2 million a day*, MERCOR (May 8, 2026), mercor.com/blog/when-you-go-from-2-million-a-month-to-2-million-a-day.

128. *See* Kelley, *supra* note 92.

129. *Id.*

130. *Id.*

131. Sarah Su et al., *Learning to Replicate Expert Judgment in Financial Tasks*, THINKING MACHINES (June 30, 2026), thinkingmachines.ai/news/learning-to-replicate-expert-judgment-in-financial-tasks.

enterprises.

Days after Anthropic announced Cowork Plugins, OpenAI unveiled Frontier, "a new platform that helps enterprises build, deploy, and manage AI agents that can do real work."[132] This platform "gives agents the same skills people need to succeed at work: shared context, onboarding, hands-on learning with feedback, and clear permissions and boundaries."[133] The announcement detailed specific partnerships with companies like HP, Intuit, State Farm, Uber, Cisco, and T-Mobile. In OpenAI's own telling, enterprises are "overwhelmed with the disconnected systems and governance spread across clouds, data platforms, and applications."[134] That fragmentation makes the work of AI agents particularly difficult, hence the new platform for providing a unified software environment that ties disparate pools of work data together so AI agents can frictionlessly pull from them all. In a similar vein is Anthropic's creation of Ode, a "new AI services company" in partnership with financiers like Blackstone and Goldman Sachs.[135] This venture will involve applied AI engineers "identify[ing] where Claude can have the most impact, build custom solutions, and support . . . mid-sized companies across sectors . . . over the long term."[136]

The labs' own research shows that direct relationships with knowledge workers supply clear visibility into how their agentic systems are being used. Anthropic, for instance, recently published a paper that sampled 1.2 million Claude Cowork sessions from a period of less than three weeks, allowing the lab to study in detail how workers are using their software for knowledge work.[137]

These initiatives position the labs to observe demand and control the channel through which distilled capacity may be deployed and sold, thereby threatening the SaaS and other enterprise-software businesses that stand between the labs and their customers. In sum, the labs' motive (business necessity) and opportunity (work data, processing power, and market visibility) are together pushing them to aggressively target the representational layer—the ground on which they ultimately monetize execution value.

---

132. OpenAI, *Introducing OpenAI Frontier* (Feb. 5, 2026), openai.com/index/introducing-openai-frontier.

133. *Id.*

134. *Id.*

135. Anthropic, *Building a new enterprise AI services company with Blackstone, Hellman & Friedman, and Goldman Sachs* (May 4, 2026), anthropic.com/news/enterprise-ai-services-company; Rebecca Bellan, *Anthropic, Blackstone bet the next trillion-dollar AI business is implementation, not just models*, TechCrunch (July 15, 2026), techcrunch.com/2026/07/15/anthropic-blackstone-bet-the-next-trillion-dollar-ai-business-is-implementation-not-models.

136. Anthropic, *Building*, *supra* note 135.

137. Anthropic, *How people are using Claude Cowork* (July 7, 2026), claude.com/blog/how-people-are-using-claude-cowork.

2.2.3 A LEGAL RATCHET

Little legally stands in the labs' way.

The labs occupy a position of unique legal privilege. Indeed, one way to understand their business is as an intellectual-property-law arbitrage: their legal efforts to have training declared fair use have met considerable success,[138] and yet the labs turn around and protect their own models using other intellectual-property protections like patents and trade secrets.[139]

IP is hardly their only legal tool. As noted above, they also forbid the distillation of their own models and contractually bar their customers from using outputs to train competitors. Those restrictions produce uneven effects: Chinese labs readily disregard the labs' terms of service, which no doubt aids their efforts to catch the frontier.[140] But domestic open-weight labs are not similarly privileged. They're left using the Chinese models to train their own, and the downsides of making a copy of a copy are evident in the performance gap between the Chinese open-weight models and the leading domestic ones.[141]

Some legal scholars have suggested the labs' terms are unenforceable.[142] But the social-data era taught that though "mark[ing] data flows with indicia of ownership" may fall short of a good-against-the-world property right,[143] what actually "matters to the construction of power is not the legal category but its practical effect."[144] The labs have illustrated that their terms are not empty threats. Anthropic, for instance, has repeatedly cut off customers' access to its models, including to a coding startup in April 2025 because of rumors it was in talks to be acquired by OpenAI.[145]

---

138. *See, e.g.*, *Bartz*, 787 F. Supp. 3d 1007; Kadrey v. Meta Platforms, 788 F. Supp. 3d 1026 (N.D. Cal. 2025).
139. *See, e.g.*, Sidney Wright, *From Patents to Privacy: The Strategic Turn Toward Trade Secrets in the AI Era*, BERKELEY TECH. L.J. BLOG (Dec. 30, 2025), btlj.org/2025/12/from-patents-to-privacy-the-strategic-turn-toward-trade-secrets-in-the-ai-era.
140. *See, e.g.*, John Liu, *What is China's Kimi K3 and why is the US so rattled by it?*, CNN (July 23, 2026), cnn.com/2026/07/23/tech/china-ai-moonshot-kimi-explainer-intl-hnk.
141. *See* Keach Hagey & Berber Jin, *Mira Murati's AI Startup Releases First Model in Bid to Loosen AI Giants' Grip*, WALL ST. J. (July 15, 2026), wsj.com/tech/ai/mira-muratis-ai-startup-releases-first-model-in-bid-to-loosen-ai-giants-grip-e042bb2b; Thompson, *Chinese Models*, *supra* note 112.
142. *See* Peter Henderson & Mark A. Lemley, *The Mirage of Artificial Intelligence Terms of Use Restrictions*, 100 IND. L.J. 1327 (2025).
143. JULIE E. COHEN, BETWEEN TRUTH & POWER: THE LEGAL CONSTRUCTIONS OF INFORMATION CAPITALISM 62–63 (2019).
144. Amy Kapczynski, *The Law of Information Capitalism*, 129 YALE L.J. 1460, 1502 (2020).
145. Maxwell Zeff, *Anthropic co-founder on cutting access to Windsurf: 'It would be odd for us to sell Claude to OpenAI'*, TECHCRUNCH (June 5, 2025), techcrunch.com/2025/06/05/anthropic-co-founder-on-cutting-access-to-windsurf-it-would-be-odd-for-us-to-sell-claude-to-openai; see also Kyle Robison, *Anthropic Revokes OpenAI's Access to Claude*, WIRED (Aug. 1, 2025), wired.com/story/anthropic-revokes-openais-access-to-claude; Grace Kay, *How xAI Went From Chasing Anthropic to Powering It*, THE INFORMATION (June 5, 2026), theinformation.com/articles/xai-went-chasing-anthropic-powering.

To be sure, these efforts can fall short. Model distillation by open-weight Chinese firms is proving extremely difficult to halt.[146] But the labs' executives look elsewhere to achieve what they can't programmatically. They have leaned heavily into national-security concerns about China and lobbied the U.S. government fiercely for protection.[147] Those efforts have met at least some halting success in the second Trump administration's erratic China policies, with announcements about tightening and loosening export controls for powerful AI chips constantly whipsawing back and forth.[148]

The labs' inbound data is thus free from most legal constraints, but law—both de facto and de jure—is helping protect their outbound data.[149]

2.2.4 THE ENTERPRISES' RESPONSE

The incumbents aren't going without a fight. And unlike the social-data era, the opposition this round is organized, wealthy, and a potent political-economic force.

The CEO of Microsoft, Satya Nadella, has been a vociferous critic of the labs' new business strategy, a particularly cruel twist of fate given that Microsoft was instrumental to OpenAI's and ChatGPT's initial success.[150] He's written that "the political economy will simply not tolerate . . . a world where every company across every sector is ceding value to a few models that eat everything they see."[151] "There is no societal permission," he continued, "for an AI future that hollows out entire industries."[152] In a follow-on essay detailing some of the many

---

146. Anthropic, *Detecting and preventing distillation attacks* (Feb. 23, 2026), anthropic.com/news/detecting-and-preventing-distillation-attacks; Deepa Seetharaman & Fabiola Arámburo, *OpenAI says China's DeepSeek trained its AI by distilling US models, memo shows*, REUTERS (Feb. 12, 2026), reuters.com/world/china/openai-accuses-deepseek-distilling-us-models-gain-advantage-bloomberg-news-2026-02-12.
147. Dario Amodei & Matt Pottinger, *Trump Can Keep America's AI Advantage*, WALL ST. J. (Jan. 6, 2025), wsj.com/opinion/trump-can-keep-americas-ai-advantage-china-chips-data-eccdce91; Dario Amodei, *On DeepSeek and Export Controls* (Jan. 2025), darioamodei.com/post/on-deepseek-and-export-controls; Cecilia Kang, *Silicon Valley's A.I. Lobbying Reaches a Fever Pitch*, N.Y. TIMES (May 13, 2026), nytimes.com/2026/05/13/technology/ai-lobbying-washington-openai-anthropic.html.
148. *See, e.g.*, Julia Shapero, *Trump faces growing pushback over China chip reversal*, THE HILL (July 29, 2025), thehill.com/policy/technology/5426103-trump-china-ai-chip-policy; Zack Whittaker, *The US government's Anthropic models ban was never about an AI jailbreak*, TECHCRUNCH (June 15, 2026), techcrunch.com/2026/06/15/the-us-governments-anthropic-models-ban-was-never-about-an-ai-jailbreak; Justin Hendrix, *Anthropic's Mythos Recall and the White House's Missing AI Safety Playbook*, TECH. POLICY PRESS (June 13, 2026), techpolicy.press/anthropics-mythos-recall-and-the-white-houses-missing-ai-safety-playbook.
149. This is a pitch-perfect replay of the social-data era's treatment of behavioral surplus. *See* COHEN, *supra* note 143, at 48–49; *see also* ZUBOFF, *supra* note 25.
150. *See* note 6 and accompanying text, *supra*.
151. *See* Nadella, *Frontier*, *supra* note 6.
152. *Id.*

ways that enterprise customers are leaking data to the labs through "exhaust," he called for something akin to a reverse patent: while a patent protects a seller from giving away valuable information about their product or service, the AI era needs a legal instrument for protecting the buyer of intelligence from disclosing to the seller the valuable "proprietary knowledge [the buyer] must reveal to make that intelligence useful."[153]

In a similar register was a remarkable interview that Alex Karp, the CEO of Palantir, gave to CNBC.[154] In it, he laced into OpenAI and Anthropic's token-based billing, stressed data-leakage worries, and echoed Nadella's political-economic concerns: "What the technical customers want is control over their compute, their models, their data stack, and their [competitive position]. They want to know they own the means of production, and it's not being transferred to someone else."[155] To that end, Palantir recently published a report that details the 15 things every company must do to guard against the labs; it begins by assuring them: "[Y]ou have complete agency over your AI model usage."[156] These defensive strategies include ensuring zero data retention by the labs, staying model agnostic, owning the feedback loop, and verifying vendors' hardware.[157]

Incumbent firms are doing more than just talking about it. They're also attempting to counteract each of the labs' recent moves. For one, incumbents are trying to withhold the mash by tightening firewalls,[158] insisting on on-premises model deployment,[159] and seeking contractual restrictions on how the labs may use their leaking data.[160] For another, the incumbents are trying to

---

153. Satya Nadella (@satyanadella), *The Reverse Information Paradox*, X (July 12, 2026), x.com/satyanadella/status/2076323181154230284 [hereinafter Nadella, *Paradox*].

154. Samantha Subin, *Palantir's Karp bashes OpenAI, Anthropic token model: 'Something has gone completely wrong'*, CNBC (July 1, 2026), cnbc.com/2026/07/01/palantir-karp-open-ai-anthropic-tokens.html;

155. *See id.*; *see also* Palantir (@PalantirTech), X (July 1, 2026), x.com/PalantirTech/status/2072326189079757277.

156. *Institutional Sovereignty in the Age of AI*, PALANTIR, at 3 palantir.com/assets/xrfr7uokpv1b/7BF74dqccPeVFMHRmy7FO3/2a33ff9b4f9e11ba904445e637095960/Palantir_-_Institutional_Sovereignty_in_the_Age_of_AI.pdf (last visited July 21, 2026) [hereinafter Palantir, *Sovereignty*].

157. *See id.*

158. *See* Stan Kaminsky, *Taming shadow-AI on corporate devices*, KASPERSKY DAILY (June 10, 2026), kaspersky.com/blog/how-to-detect-disable-shadow-ai-in-enterprise/55952; Moveo AI Team, *Companies Banning ChatGPT (2026): The Enterprise Security List*, MOVEO BLOG (May 15, 2026), moveo.ai/blog/companies-that-banned-chatgpt.

159. *See* Nadella, *Paradox*, *supra* note 153 ("[E]nterprises need a real trust boundary for their human capital and token capital to compound. It is where an organization's data, traces, evals, adapted weights, and memory accumulate and improve together. And it is a hard boundary across which nothing crosses, not even the intelligence exhaust, without consent.").

160. *See, e.g.*, Anthropic, *I have a zero data retention agreement with Anthropic. What products does it apply to?* (June 9, 2026), privacy.claude.com/en/articles/8956058-i-have-a-zero-data-retention-agreement-with-anthropic-what-products-does-it-apply-to.

build their own stills: training their own models,[161] building private output evaluations and reinforcement-learning environments in-house,[162] and substituting the labs' proprietary models with good-enough open-weight alternatives from China.[163] They're also defending the point of sale by adopting model-agnostic architectures and multi-model harnesses.[164] And they've started to organize, publicly releasing a joint open letter that called for more development of domestic open-weight models and that specifically defended distillation.[165]

Nadella and Karp, of course, have their own self-motivated reasons for sounding the alarm. Their companies, after all, are the incumbent representational layer that the frontier labs are threatening, and Microsoft specifically bills on a per-seat basis.[166] And their strategy—owning and controlling data—is just the same as the workers' sabotage of expert systems and their withholding of tacit knowledge from memorialization into legible work data.[167] So the executives' campaign against the labs should be understood as combatants in the inter-firm distillation war drawing battle lines. And those battle lines should sound particularly familiar, since they are also the workers' complaints with different nouns.

2.3 ONE RECURSIVE CONFLICT

It is striking to witness the capital class adopt the posture and rhetoric of the proletariat. The CEO of one of the world's most valuable companies is penning essays that draw explicit parallels between AI and globalization's decimation of blue-collar work and that say "the political economy will simply not tolerate . . .

---

161. *See* Edited Transcript, *TRI.TO - Thomson Reuters Corp. at Bank of American Info. & Bus. Services Conf.*, at 5 (Mar. 12, 2026), investors.thomsonreuters.com/static-files/2955885b-0556-4d3b-a64a-f136a159ecb8 [hereinafter Thomson Reuters, *Investor Day Transcript*]; Dipshan, *supra* note 107; Ben Thompson, *An Interview with Microsoft CEO Satya Nadella About Finding Core Competencies*, STRATECHERY (June 4, 2026), stratechery.com/2026/an-interview-with-microsoft-ceo-satya-nadella-about-finding-core-competencies [hereinafter Thompson, *Nadella Interview*].

162. Nadella, *Frontier*, *supra* note 6 ("Private evals should capture whether a model is actually improving against outcomes that matter to the business (not just external benchmarks!). Private reinforcement learning environments should let models grow stronger on real traces from inside the organization.").

163. Amelia Charles & Harpreet Arora, *Open-weight models surge to 29% of volume, price per token flattens*, VERCEL BLOG (July 13, 2026), vercel.com/blog/ai-gateway-production-index-july-2026; Kai Nicol-Schwarz, *Chinese AI models are gaining ground with U.S. companies as OpenAI, Anthropic costs surge*, CNBC (July 7, 2026), cnbc.com/2026/07/07/chinese-ai-models-costs-us-openai-anthropic.html.

164. *See* Thompson, *Nadella Interview*, *supra* note 161; Thomson Reuters, *Investor Day Transcript*, *supra* note 161.

165. *See Open Weights and American AI Leadership* (July 24, 2026), images.nvidia.com/pdf/Open-Weights-and-American-AI-Leadership.pdf (last visited July 25, 2026).

166. *See* Palantir, *Sovereignty*, *supra* note 156; Microsoft, *Find the best Microsoft 365 plan for your organization*, microsoft.com/en-us/microsoft-365/enterprise/microsoft-365-plans-and-pricing (last visited July 21, 2026).

167. *See* notes 58, 89–91 and accompanying text, *supra*.

an AI future that hollows out entire industries";[168] his counterpart who heads a right-wing defense-and-surveillance contractor likewise asserts that enterprises are losing faith that "they own the means of production."[169] This Essay's central claim—that we're witnessing a single conflict over knowledge-work distillation that recurs at two distinct levels of the economy—renders these seemingly mind-bending reversals sensible.

That conflict, in this Essay's terms, is over who will obtain the mash, run the still, and ultimately monetize the distillate. Both management and the frontier labs are scrambling to control that process from beginning to end, and the symmetry becomes most apparent by answering the same five questions at each level: (1) Who *performs* the work and thereby generates the work data? (2) Who supplies the *environment* where the work data accrues? (3) Who *processes* the data? (4) Who is the *market* where the distilled execution value is monetized? (5) What *law* allocates control?

At the first level, within the firm: The performers are the workers who generate work data through their redlines, corrections, and the captured traces of enacted human judgment. The environment is the employer's—the document system and monitored tools. The processing occurs on a model the employer leases from the labs. The market is the employer: as the buyer of labor, it monetizes the distillate by hiring fewer workers and capturing the gains of those who remain. And the law plants control with the employer: little restricts its surveillance and capture, and the trade-secrecy carve-out protects the employees' tacit knowledge only in its embodied state.

At the second level, between firms: The performers are the enterprises who generate work data through their people's prompts, model evaluations, corrections, and workflows. The environment, increasingly, belongs to the labs—offerings and initiatives like ChatGPT Work and Frontier, Claude Cowork and Ode. The processing is the labs' own. The market is contested: the enterprises have traditionally sold execution value to their customers, but the labs are attempting to cut them out. And the law favors the labs, as little restricts their data-acquisition efforts, and their own adhesion terms govern outbound data flows.

Beyond the symmetry of these answers between distiller and distillee, the two levels share two other things in common. First is the atomic unit: the correction, which is RLHF fuel at the first level and enterprise exhaust at the second. Second is what fills the legal void: self help, which is the worker's withholding and sabotage at the first level and the enterprises' control efforts at the second.

What differs between them is therefore not the structure of the conflict but who holds the bargaining power. The workers are given "no option to opt-out"

168. Nadella, *Frontier*, *supra* note 6.
169. Subin, *supra* note 154.

from the data capture,[170] whereas the enterprises are engineering what those workers cannot—private evaluations, model-agnostic architectures, and negotiated zero-data-retention agreements.[171] That disparity produces a remarkable entitlement asymmetry: Nadella claims that "what you create should belong to you,"[172] but that's a principle that surfaces at only one level. The workers whose corrections fill the enterprises' feedback loops—including Microsoft's—appear nowhere in it. What *they* create, in Nadella's telling, belongs to their employer.

* * *

Revisit our senior litigation associate for a moment. She generates enormous value through her redlines for both her firm and the lab that can get its hands on them, but everything else belongs to someone else. She has no claim to the firm's use of her redlines to fine-tune the models they use, and she has no visibility into how the labs might benefit from them—though she should know the labs are willing to pay Mercor a pretty penny for redlines just like hers.[173] Her tacit knowledge, expertise, and judgment will thus continue to be monetized by someone else long after she leaves the firm. Whether those who so eagerly distilled her will themselves be successfully distilled by another is an open question—unlike her, the enterprises have the means to resist, and the next Part takes up their prospects. But either way, the senior associate's fate remains the same: if her firm holds its ground, her expertise enriches her employer; if it loses, her expertise enriches the labs. So while the contest above her is undecided, her dispossession is overdetermined.

3. LESSONS & IMPLICATIONS

Lessons from the recent past shed light on how each level of the conflict is likely to resolve. The same legal arrangements that advantaged data processors in the social era of data monetization are likely to pay similar dividends in the work-data era. But it would be too hasty to conclude the labs' ultimate victory is a fait accompli. And if the labs really do land a decisive blow against the enterprises, then it becomes difficult to resist the conclusion that a new paradigm altogether might govern.

3.1 LAW & WORKERS' DISPOSSESSION

Taking seriously the lessons from the social-data era counsels that law will not only fail to stop the distillation of workers by their employers but it will prove

170. Huang et al., *supra* note 4.
171. *See, e.g.*, Palantir, *Sovereignty*, *supra* note 156.
172. Nadella, *Paradox*, *supra* note 153; *see also* Palantir, *Sovereignty*, *supra* note 156, at 3 ("Sovereignty is your alpha. It is your ownership of the value you create and your freedom to pursue new opportunity; a set of economic rights and the ability to expand them.").
173. *See* notes 127–130 and accompanying text, *supra*.

essential to enabling it.

Consider the worker-protection proposals on offer. Viljoen, Li, and Cullen argue for the enactment of institutions of collective data governance, arguing that such knowledge guilds can secure workers' ability to benefit from their automated expertise and thereby share in the gains of that automation.[174] Ajunwa advocates for creating a licensing regime for workers' data and for requiring firms to pay into a guaranteed-income fund.[175] Those writing before the rise of generative AI have similarly focused on policy reforms that restrict how data is collected and used, and they've sought to empower employees in negotiations with management over such practices.[176]

The social era of data monetization provides little reason for optimism about such approaches because they require policymakers to act inconsistently with the past three decades of practice. What little data governance the past generation has produced is a set of meek consumer privacy laws enacted in half the states that focus almost exclusively on the data collection side of the ledger.[177] The data processors have had an outsized hand in shaping these statutes, so they tend to legitimate—rather than constrict—their drafters' collection and processing practices.[178]

Labor law is similarly grim. The same collective-action problems that allowed the rise of gig work, algorithmic wages, and oppressive digital surveillance are likely to persist for knowledge workers.[179] Contractors like Mercor's are excluded from the National Labor Relations Act altogether.[180] And though the Act nominally covers professional employees, its exclusions track the very attributes that make a knowledge worker worth distilling: the Supreme Court has held that professionals who exercise "ordinary professional or technical judgment" in directing less-skilled colleagues may be excluded as supervisors, and that faculty whose collective judgment shapes institutional policy are excluded as managerial employees.[181] So the more judgment and authority a worker wields, the weaker her claim to the Act's protections. Progressive Executive Branch interpretations of existing authorities face a hostile federal judiciary and die on the vine after presidential-administration turnover.[182] State-level restric-

---

174. Viljoen et al., *supra* note 24, at 32–33.
175. Ajunwa, *supra* note 59, at 397–401.
176. *See, e.g.*, BRISHEN ROGERS, DATA AND DEMOCRACY AT WORK: ADVANCED INFORMATION TECHNOLOGIES, LABOR LAW, AND THE NEW WORKING CLASS 397–401 (2023); Bodie, *supra* note 94, at 743–51.
177. *See* Ormerod, *Monetization*, *supra* note 9, at 1008–13 (surveying these laws); *id.* at 1013–26 (criticizing their approach).
178. *See id.* at 1005 nn.13–14 and accompanying text.
179. *See, e.g.*, Zephyr Teachout, *Algorithmic Personalized Wages*, 51 POL. & SOC'Y 436 (2023); Veena Dubal, *On Algorithmic Wage Discrimination*, 123 COLUM. L. REV. 1929 (2023).
180. *See* 29 U.S.C. § 152(3).
181. *See id.* § 152(11); NLRB v. Kentucky River Community Care, 532 U.S. 706 (2001); NLRB v. Yeshiva University, 444 U.S. 672 (1980).
182. *See, e.g.*, Ryan LLC v. FTC, 739 F. Supp. 3d 496 (N.D. Tex. 2024); NLRB General

tions on automated decisionmaking in the workplace face the same stiff headwinds that have bedeviled consumer privacy laws.[183]

They also repeat the social era's error of narrowly focusing on the legal regime that governs data cultivation.[184] Consent, ownership, deletion, retention are all tools that double down on data privacy's failed control paradigm.[185] Even if such approaches meet more success in the work era, they will not reach the still or the market—the places that Part 2 showed the value is accruing. As scholars of law and political economy have said, these tilted scales are not the natural order of things—they are socially constructed legal endowments.[186]

Workers' best hopes therefore lie beyond the law. Meta's ham-fisted attempts at work-data capture were, after all, not halted by law or by workers acting collectively, but rather by its architects' own fecklessness.[187] There is also genuine uncertainty about exactly how much knowledge work can be fully distilled by the LLM architecture. Code must compile, a harsh true-or-false judgment on the models' outputs.[188] This binary evaluation of software's fidelity renders its development uniquely vulnerable to probabilistic AI, a vulnerability that much other knowledge work does not share.[189] But even if the AI bulls are wrong and technological progress stalls, the existing tools promise to fundamentally reshape knowledge work in the twenty-first century.[190] So even the optimistic case deprives the worker of self-determination; she remains at the whims of the re-

---

Counsel, *Electronic Monitoring and Algorithmic Management of Employees Interfering with the Exercise of Section 7 Rights*, MEMO. GC 23–02 (Oct. 31, 2022), *rescinded by* NLRB General Counsel, *Rescission of Certain General Counsel Memoranda*, MEMO. GC 25–05 (Feb. 14, 2025).

183. *See, e.g.*, Lucas Wright et al., *Null Compliance: NYC Local Law 144 and the challenges of algorithm accountability*, FAccT '24: PROCEEDINGS OF THE 2024 ACM CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY (June 5, 2024), doi.org/10.1145/3630106.3658998; Ridhi Shetty, *CDT Urges California to Undo Changes Weakening Proposed Rules on Automated Decision-Making*, CENTER FOR DEMOCRACY & TECH. (June 3, 2025), cdt.org/insights/cdt-urges-california-to-undo-changes-weakening-proposed-rules-on-automated-decision-making.

184. *See* Ormerod, *Monetization*, *supra* note 9, at 1018–26; Parsons & Viljoen, *supra* note 8, at 1071–73.

185. *See, e.g.*, Salomé Viljoen, *A Relational Theory of Data Governance*, 131 YALE L.J. 573, 597–600 (2021).

186. *See, e.g.*, COHEN, *supra* note 143, at 48 ("Understood as processes of resource extraction, the activities of collecting and processing personal data require an enabling legal construct."); Kapczynski, *supra* note 144, at 1499 ("Control over data and its production is socially and also legally determined.").

187. *See* Dave, *supra* note 78.

188. *See, e.g.*, Zeynep Tufekci, *The Unstoppable Force of A.I. Hype Is Meeting One Immovable Fact*, N.Y. TIMES (June 30, 2026), nytimes.com/2026/06/30/opinion/ai-agents-steal-jobs-employment.html ("L.L.M.s can do many things with astounding proficiency, but they can't do the vast majority of human jobs without skidding into disaster here and there. . . . The exceptions to that rule are jobs that occupy formal or verifiable domains. Coding is one such job. It relies on a structured, formal language that can be tested in real time.").

189. *See id.*

190. *Id.*

searchers' uncertain attempts to produce so-called "general" or "super" intelligence.

3.2 INTER-FIRM UNCERTAINTY

By contrast, considerable uncertainty pervades the conflict between enterprises and the labs.

Most existing public law is unlikely to play much of a role. Decades of precedent foreclose the conclusion that Anthropic's API cutoffs violate the Sherman Antitrust Act,[191] and Neo-Brandeis antitrust scholars have convincingly shown how the consumer-welfare standard blinds antitrust law to competitive harms in the information age.[192] Cheap and abundant intelligence is therefore likely to register as welfare gain while rents shift among actors.[193] Ex post enforcement actions are particularly ill-suited to the fast-moving present,[194] and ex ante control over model releases by the second Trump administration appears fueled by grievance rather than reasoned debate about who should gain.[195] To the extent existing authorities have entered the equation, it's been an incumbent wielding an outdated and ambiguous computer-hacking statute to prevent AI scraping.[196]

So the contest is more likely to be decided by private ordering among the powerful. Incumbents with enough resources and sufficient bargaining power are best positioned to withstand the labs' distillation efforts. Microsoft and Palantir are the rare few able to build sovereign architectures, sealed feedback loops, and model-agnostic infrastructures while insisting upon zero-data-retention agreements.[197] The winners of the social-data era are hard at work attempting to defend their existing moats by dedicating every cent of free-cash flow and then some to the AI infrastructure buildout—an insurance policy for a perpetually compute-constrained future.[198] Nadella's reverse-patent idea is best understood as entitlement entrepreneurship addressed to contract and technical architecture rather than a serious policy proposal directed to Congress.

AI startups and smaller legacy enterprises lack the ability to decide their

191. *See* Verizon Communications Inc. v. Law Offices of Curtis V. Trinko, 540 U.S. 398 (2004).
192. *See* Peter Ormerod, *Disciplining Mechanisms: Governing Data Markets with Competition and Regulation*, 28 YALE J. LAW & TECH. 308, 311–315 (2026).
193. *See, e.g.*, Ryan Chapman, *Antitrust in the Age of AI: Is the Consumer Welfare Standard Equipped to Address the Rise of Generative Artificial Intelligence?*, ABA ANTITRUST L. NEWSLETTER (2026), *available at* papers.ssrn.com/sol3/papers.cfm?abstract_id=7039359.
194. *See* Elettra Bietti, *Experimentalism in Digital Platform Markets: Antitrust and Utilities' Convergence*, 2024 U. ILL. L. REV. 1277, 1326–38 (2024).
195. *See, e.g.*, Whittaker, *supra* note 148; Hendrix, *supra* note 148.
196. *See* Amazon.com Service LLC v. Perplexity AI, Inc., No. 25-cv-09514-MMC, 2026 WL 658407 (N.D. Cal. Mar. 11, 2026), *stayed by* Amazon.com Service LLC v. Perplexity AI, Inc., No. 26–1444, Docket Entry 18.1 (9th Cir. Mar. 16, 2026).
197. *See* notes 156–164 and accompanying text, *supra*.
198. *See* Isabel Juniewicz, *Hyperscaler capex is on trend to outpace their cash inflows by the end of 2026*, EPOCH AI (June 16, 2026), epoch.ai/data-insights/hyperscaler-capex-vs-cash-flow.

own fate. Licensing deals between the labs and publishers like News Corp. and Axel Springer suggest that powerful and concentrated market actors are the prisoner's dilemma defectors most capable of extracting concessions;[199] small and diffuse authors' best hope is a class-action settlement check.[200] Collective institutions for compensating those whose work made the models possible were once considered.[201] But the labs' executives can read opinion polls like anyone else and have recently pivoted to assuring the populace the future won't be so bad rather than advocating for broad compensation mechanisms that presuppose widespread displacement.[202]

But it's also possible that the state intervenes on behalf of one of the combatants. Both sides cloak their asks in national-security garb: For the labs, it's apocalyptic warnings about the dangers of Chinese open-weight models;[203] any resulting restrictions would also threaten the domestic open-weight ecosystem built upon those models. The prospect of a macroeconomic crisis if there's no one left to pay for all those data centers might also help.[204] For the enterprises, it's calling for policy interventions to cultivate a "strong AI ecosystem" that will "drive American sovereignty and prosperity" by "giving Americans greater control over the technology they rely on."[205] Both possibilities remind us that "[l]egal institutions . . . are enlisted to help produce the profound economic and sociotechnical transformations that we see all around us."[206]

3.3 A NEW PARADIGM

Suppose briefly that the strongest legacy enterprises cannot even fight the labs to an uneasy détente—either because the enterprises lose outright or because

---

199. Katie Robertson, *OpenAI Strikes a Deal to License News Corp Content*, N.Y. TIMES (May 22, 2024), nytimes.com/2024/05/22/business/media/openai-news-corp-content-deal.html.
200. Bartz v. Anthropic PBC, No. 24-cv-05417, Docket Entry No. 680 (N.D. Cal., July 20, 2026) (granting final approval for class-action settlement between Anthropic and book authors).
201. *See* Bartz v. Anthropic PBC, No. 24-cv-05417, Docket Entry No. 563, Attachment No. 2 (Exhibit 40) (N.D. Cal., Jan. 21, 2026) (containing a memorandum authored by Anthropic CEO Dario Amodei in 2021 wherein he advocated for "compensat[ing] data/content producers for their labor, not at a flat rate, but with a fraction of the profits from the model produced").
202. Katherine Bindley, *Big Tech Has Suddenly Flipped on the AI Jobs Wipeout Scenario*, WALL ST. J. (July 5, 2026), wsj.com/tech/ai/ai-workers-tech-ceos-job-losses-afc71e15.
203. *See* Amrith Ramkumar & Tina Li, *Top American AI Execs Sound Alarm on Chinese Models*, WALL ST. J. (July 20, 2026), wsj.com/tech/ai/top-american-ai-execs-sound-alarm-on-chinese-models-3c74f8c1.
204. Dominic O'Connell, *The AI data centre boom built on a mountain of hidden debt*, THE TIMES (June 19, 2026), thetimes.com/business/technology/article/data-centre-boom-cft98jgf5.
205. *See Open Weights*, *supra* note 165, at 1–2.
206. COHEN, *supra* note 143, at 2; *see also* KATHARINA PISTOR, THE CODE OF CAPITAL 4 (2019) ("[M]ost observers treat law as a sideshow when in fact it is the very cloth from which capital is cut.").

the state forecloses their open-weight exit. Questions abound in any such scenario where workers and enterprises alike face rapid, unrelenting distillation by the labs.

Some scholars, in response to the data processors' definitive victory in the social-data era, have sought to resuscitate the policymaking playbook of a bygone era.[207] The networks, platforms, and utilities (NPU) toolkit does not look to competition to tame outsized market actors but rather seeks to overtly bring them to heel using access tools like non-discrimination, interoperability, and structural separation; economic levers like price caps and rate-of-return regulation; and tight transparency and oversight.[208] Some have pressed the well-established supply-side arguments in favor of casting AI as a natural monopoly that necessitates the NPU toolkit.[209] But it also seems undeniable that there will be untold productive demand for cheap, abundant intelligence on a meter—a second, and if anything, better reason for the public-utility approach.[210] And while turning to that toolkit raises as many questions as it answers,[211] it appears increasingly urgent that we answer them.[212]

These are not premature pipe dreams. The likes of Microsoft, Palantir, and IBM are already publicly petitioning for policymakers to address the labs' foreign distillation concerns using "targeted legal and commercial frameworks rather than sweeping restrictions" on "legitimate model-development techniques" that "play an important role in AI innovation"[213]—the seed of an access mandate in the tradition of interconnection duties and compulsory licenses.

So the present moment may be creating the same conditions that previously produced the NPU toolkit, as utility regulation has historically arrived when organized commercial interests were the aggrieved parties. *Munn v. Illinois* grew as much out of the Chicago Board of Trade's war with the grain elevators as out

---

207. *See* MORGAN RICKS, GANESH SITARAMAN, SHELLEY WELTON, AND LEV MENAND, NETWORKS, PLATFORMS, AND UTILITIES LAW AND POLICY (2022) [hereinafter Ricks et al, *NPU Casebook*]; Morgan Ricks, Ganesh Sitaraman, Shelley Welton, and Lev Menand, *Introduction to the Symposium on Networks, Platforms, and Utilities: Law and Policy*, YALE J. REG. NOTICE & COMMENT BLOG (Jan. 17, 2023), yalejreg.com/nc/symposium-networks-platforms-utilities-00.

208. *See* Ricks et al., *NPU Casebook*, *supra* note 207, at 24–31.

209. *See* Tejas N. Narechania & Ganesh Sitaraman, *An Antimonopoly Approach to Governing Artificial Intelligence*, 43 YALE L. & POLICY REV. 95 (2024).

210. *See* BRETT M. FRISCHMANN, INFRASTRUCTURE: THE SOCIAL VALUE OF SHARED RESOURCES (2012); *see also* Thibault Spirlet, *Sam Altman says AI will eventually be sold like electricity and water — by companies like OpenAI*, BUS. INSIDER (Mar. 13, 2026), businessinsider.com/sam-altman-ai-utility-electricity-water-openai-2026-3 (quoting OpenAI CEO Sam Altman: "We see a future where intelligence is a utility like electricity or water and people buy it from us on a meter and use it for whatever they want to use it for").

211. *See* Julie E. Cohen, *Public Utility for What? Governing AI Datastructures*, 28 YALE J. L. & TECH. 135 (2026); Elettra Bietti, *Data is Infrastructure*, 26 THEORETICAL INQUIRIES L. 55 (2025).

212. *See* Peter Ormerod, *Intelligence as Utility* (draft on file).

213. *Open Weights*, *supra* note 165, at 3.

of a farmers' revolt;[214] farmers, ranchers, and stockmen banding together proved instrumental to the utility-style regulation of stockyard infrastructure;[215] and the Kingsbury Commitment arose in part from the threat that AT&T posed to independent operators.[216] Public utility law, then, is what happens when capital large or small needs protection from infrastructure. So Nadella's forecast—that "the political economy will simply not tolerate it"[217]—understood in the right light, might be true whether the enterprises save themselves through private ordering or the state must arrive to save them. Alas, even if the state arrives, it does so summoned by capital and in service of saving Microsoft. Our senior associate can only hope to be a third-party beneficiary.

CONCLUSION

This Essay has shown that we are witnessing the opening skirmishes of a new conflict over how prosperity in the twenty-first century will be generated and who stands to capture it. My past work asked how law governs the monetization of data about people; this Essay has named what comes next: the monetization of work itself.

Both conflicts share in common that powerful commercial interests will use law and technology alike to appropriate great wealth for themselves. This time around, however, vast swaths of capital—rather than mere behavioral surplus—are subject to contest. And while the ultimate resolution of that conflict has not yet been written, there is one thing that is certain: in the age of information capitalism, data processing ultimately comes for everyone.

---

214. *See* Edmund W. Kitch & Clara Ann Bowler, *The Facts of* Munn v. Illinois, 1978 Sup. Ct. Rev. 313, 324–29 (1978); *id.* at 325 ("In [the Board of Trade's] position as speculators they wanted more: the power to inspect the contents of the elevators. This the warehousemen resisted, and to get it the Board of Trade turned to the State legislature. The 1871 legislation that provided interior inspection also contained the rate section litigated in *Munn*.").

215. *See* Luke Herrine, *Regulating Cutthroat Business*, 103 N.C. L. Rev. 1573, 1586 (2025) (citing Elizabeth Sanders, Roots of Reform: Farmers, Workers, and the American State, 1877–1917, at 173 (1999)).

216. *See* Richard John, Network Nation: Inventing American Telecommunications 340–69 (2010); *id.* at 354 ("Like most small independent[phone operators], [John H.] Wright dreaded Bell's legal and financial clout and cast a wary eye at its repeated incursions into the independents' financially precarious toll-line business. Since 1911 Wright had lobbied the Justice Department to prosecute Bell and was delighted when [Attorney General James Clark] McReynolds reopened the Sherman Act lawsuit that [former Attorney General George W.] Wickersham had shelved.").

217. Nadella, *Frontier*, *supra* note 6.