ORGANISATIONAL GOVERNANCE & PHILOSOPHY OF TECHNOLOGY

Where Accountability Sits

Why the People Function Must Frame Work-System Enterprise AI

Simon Ives, FCPHR, MBA

Published November 2026 · simonives.com · 81 References

EXECUTIVE ABSTRACT

Cybersecurity and enterprise architecture are constraint specifications that an organisation can verify after a business problem is framed; human work design is constitutive of the problem frame itself. Drawing on philosophy of technology, accountability theory, sociotechnical research, enacted multi-jurisdictional employment and AI legislation, and comparative institutional analysis, this essay argues that in large, complex, regulated enterprises deploying work-system artificial intelligence, formal institutional accountability for framing work redesign, human-agent capacity allocation, performance incentives, workforce transition, and employment-risk alignment should sit in the enterprise people function. That mandate operates inside a six-domain binding partnership with Technology, Operations, Finance, and independent assurance mapped to the Institute of Internal Auditors Three Lines Model, ISO/IEC 42001:2023, and the NIST AI Risk Management Framework, conditioned on four measurable technical-readiness preconditions and an independently governed three-tier workforce materiality triage gate.

KEYWORDS: AI transformation · Work-system AI · Philosophy of technology · Philosophy of AI · Institutional accountability · Enterprise people function · Comparative institutional analysis · Sociotechnical governance · Three Lines Model

Suggested Citation (Harvard): Ives, S 2026, *Where accountability sits: why the people function must frame work-system enterprise AI*, simonives.com, November, <https://simonives.com/essays/where-accountability-sits.pdf>. *Additional citation formats (APA 7th and Chicago) appear at the close of this essay.*

Copyright & Licence: © 2026 Simon Ives. Licensed under Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0). You may share this essay in its original, unmodified form for non-commercial purposes with full attribution to the author and simonives.com.

WHERE ACCOUNTABILITY SITS

Contents

Introduction: Framing Before Building

Method, Boundaries, and Positionality

Chapter 1: Defining the Terrain, or Why "Technology Rollout" Is a Category Mistake

1.1 Fixing the Referents: Accountability, the People Function, and Work-System AI

1.2 *Techne*, *Politikē*, Autonomous Technique, and the Category Mistake

1.3 Seventy-Five Years of Sociotechnical Proof and Its Real Lesson

Chapter 2: The State of the Question, and the Capability-Mandate Paradox

2.1 Political Economy, Jurisdictional Contests, SHRM, and Critical HRM

2.2 The Executive Inflection Point: New Roles, Structural Mergers, and Binding Law

2.3 What Corporate Benchmarks Quietly Admit: The Capability-Mandate Paradox

Chapter 3: The Constructive Case for Placing Accountability in the People Function

3.1 Premise 1: The Binding Bottlenecks of Work-System AI Are Workforce Architecture Problems

3.2 Premise 2: Information Technology Governance Theory Rejects IT Primacy

3.3 Premise 3 and the Comparative Institutional Quintilemma: Honouring Hume, Global Law, and the Framing Asymmetry

3.4 Premise 4: Organisational Hysteresis, the Automation-Augmentation Paradox, and Technological Momentum

Chapter 4: Objections, Rivals, and the Capability-Mandate Paradox

4.1 Objection 1: The Technology-Led Case and the Capability-Mandate Paradox

4.2 Objection 2: Appointing a Dedicated Chief AI Officer

4.3 Objection 3: Structural Fusion (Moderna and AXA Germany)

4.4 Objection 4: Operations-Led Governance, CFO Capital Hurdles, Fissured Workforces, and Contingency

4.5 Objection 5: Purely Operational AI, Desktop Productivity Tools, and Bottlenecks

Chapter 5: The Architecture of Binding Partnership, Three-Lines Assurance, and Earned Authority

5.1 Legal-Rational Authority, Specific Knowledge, and Escaping "Rule by Nobody"

5.2 Six Decision Domains of the Binding Partnership

5.3 Workforce Materiality Triage: Independent Classification Preventing Bottlenecks and Tier-Shopping

5.4 Operational Classification Tests and Rebuilt Empirical Disconfirmation Standard

5.5 Four Preconditions: Converting Formal Authority into Real Authority

Conclusion: Accountability as Institutional Design

References

How to Reference This Essay & Licence

EXHIBITS

Exhibit 1: Enterprise Work-System AI Decision-Rights Constitution (Three Lines Mapped)

Exhibit 2: Three-Tier Workforce Materiality Triage Gate

INTRODUCTION

Framing Before Building

Cybersecurity and technical architecture are constraint specifications that an enterprise can verify after a business problem is framed; human work design is constitutive of the problem frame itself. That asymmetry explains why accountability for work-reconfiguring enterprise artificial intelligence (AI) belongs in the enterprise people function. Most executive committees still treat AI ownership as a territorial dispute between the Chief Human Resources Officer and the Chief Information Officer. That habit degrades a deep question of organisational philosophy and institutional design into a personality contest. Beneath the turf war sits an assignable structural decision that governs what an enterprise builds over the next decade: where formal accountability for **work-system AI** (generative and agentic systems that reconfigure human roles, judgement, capacity, incentives, and employment risk) must sit.

When a Strategy unit or business division frames an internal AI initiative around a 25% headcount reduction before work design enters the room, downstream employment checks arrive too late. Technology can inspect an already-framed workflow for cyber vulnerability, data lineage, and latency without distorting the underlying business problem. The people function cannot wait downstream to inspect human work as an afterthought. Once vendor shortlists, capital hurdle rates, and executive expectations lock around pure labour substitution, a downstream employment veto can block unlawful bias or negotiate redundancy packages; it can no longer reframe the initiative from labour-displacing automation into capability-building augmentation without triggering an adversarial board clash.

Contemporary corporate benchmarks illustrate how urgently that structural fault line presses on boards, even while exposing an acute readiness paradox. In *HR Monitor 2026*, McKinsey & Company (2026, pp. 3, 39) warns that HR will either step forward as "architect of AI transformation" shaping how human and agentic workforces operate together or watch IT absorb an expanding share of its responsibilities. Yet across the same survey of 1,303 HR decision-makers, long-horizon workforce planning, technical automation, and agentic fluency lag well behind that ambition (McKinsey & Company 2026, pp. 9, 33–36). Asking a Chief Executive Officer to place enterprise-wide work-system accountability inside an unreformed HR department invites justified scepticism.

This essay moves past consulting rhetoric to defend a disciplined, stand-alone thesis. **In large, complex, multi-jurisdictional enterprises deploying work-system AI, formal institutional accountability for framing work redesign, human-agent capacity allocation, performance incentives, workforce transition, and employment-risk alignment should sit in the enterprise people function (organisational HR strategy).** That structural mandate carries two conditions. First, the people function must operate inside a binding partnership with domain authorities across Technology (holding authority over cybersecurity, enterprise architecture, the unified data platform, and model safety), Operations (holding authority over frontline workflow execution and value realisation), Finance (aligning capital-allocation hurdles with redesigned capacity), and independent assurance mapped to the Institute of Internal Auditors (IIA) Three Lines Model.

Second, the people function must earn and sustain real authority alongside formal charter by building internal technical fluency and a logically isolated People Data domain inside the unified enterprise platform.

Method, Boundaries, and Positionality

My method combines conceptual analysis drawn from philosophy of technology and accountability theory with **comparative institutional analysis** (Komesar 1994; Williamson 1996). No single randomised trial compares people-led against technology-led enterprise AI structures across multi-year horizons; institutional design must therefore weigh imperfect real-world alternatives against explicit normative and empirical criteria. I triangulate across four source tiers: peer-reviewed scholarship in information systems, organisational theory, strategic human resource management (SHRM), and philosophy; enacted primary legislation and active court dockets across Europe, the United States, Chile, the United Kingdom, and Australia; published Harvard Business School teaching cases treated candidly as protagonist-framed secondary syntheses; and critically appraised corporate benchmarks whose commercial interests and sampling limits are disclosed upfront.

Four explicit boundaries govern this essay. First, the argument concerns *structure* (the assignable institutional location of decision rights, framing pen, and accountability) and never personal charisma. Second, scope covers large, regulated enterprises (typically 5,000 or more employees operating legacy platforms across multiple labour jurisdictions), excluding startups, small firms under 500 staff, and isolated backend upgrades that leave core work roles unchanged. Third, this essay separates **external customer-product AI strategy** (clinical diagnostics, credit underwriting, dynamic pricing, owned by the CEO, Corporate Strategy, and Commercial Product Leaders) from **internal work-system AI**, where the people function holds the framing locus. Fourth, intellectual honesty requires disclosing my own standpoint. I write as an enterprise HR technology strategy and AI governance practitioner. Just as readers should discount advisory reports for commercial self-interest, readers should weigh my argument knowing that my professional vantage point sits inside the people-technology domain. That standpoint supplies direct practitioner visibility into how governance fails in production; it confers no epistemic exemption, so every claim below rests on public evidence and visible inference.

The essay proceeds in five chapters. Chapter 1 establishes why treating work-system AI as a software rollout commits a category mistake, fixing the referent of "HR" cleanly onto the people function and tracing the philosophical and sociotechnical arc from Aristotle, Ellul, Ryle, and Vallor through seventy-five years of workplace scholarship. Chapter 2 maps the state of the question across political economy, professional-jurisdiction theory, formal-versus-real authority, SHRM, Critical HRM, and corporate benchmarks. Chapter 3 builds the constructive case across empirical bottlenecks, IT governance theory, an explicit two-track normative premise honouring Hume's Law, enacted multi-jurisdictional law, a five-option Comparative Institutional Quintilemma anchored in accountability theory, and organisational hysteresis. Chapter 4 steelmans and answers five structural rivals: Technology-led governance, the dedicated Chief AI Officer, structural HR-IT fusion, Operations-led models (including centralised operating systems, CFO capital hurdles, and fissured workforces), and the objection that operational or desktop productivity AI turns the people function into a bottleneck. Chapter 5 translates the thesis into a six-domain decision-rights

constitution aligned with the Three Lines Model, ISO/IEC 42001, and the NIST AI Risk Management Framework, paired with an independently governed materiality triage gate, operational and Lakatosian disconfirmation tests, and four readiness preconditions.

CHAPTER 1

Defining the Terrain, or Why "Technology Rollout" Is a Category Mistake

Treating work-system AI as a software rollout places a problem of human governance inside the wrong logical category. Before an executive committee can assign institutional accountability for enterprise AI, it must define both the object it is governing and the specific institutional referent behind the shorthand label "HR".

1.1 Fixing the Referents: Accountability, the People Function, and Work-System AI

Three conceptual confusions derail boardroom debates before analysis begins. The first confuses *leadership* with *structure*. Asking whether AI transformation should be "HR-led" or "technology-led" often invites talk of executive charisma or town-hall visibility. My concern is *structure*: which executive function holds formal institutional accountability, holds the initial framing pen, and owns specific decision rights when human work and intelligent software are redesigned together. Transformation itself is an emergent enterprise property with no single office address; *accountability* is an assignable institutional fact. Accountability is held or located, never "led".

The second confusion slides carelessly across three distinct referents of **"HR"**: (a) *organisational HR strategy* (the enterprise people function governing workforce allocation, capability, incentives, and employment governance, Dyer 1985, cited in Wright & Ulrich 2017, p. 46); (b) the legacy *HR department* administering payroll and policies; and (c) the *HR profession* as a career tribe. Throughout this essay, my thesis concerns the **enterprise people function** (the first referent). Whether a board fills the executive seat heading that function with a builder-literate career CHRO or with a Strategy and Data executive (as AXA Germany did in 2025), the institutional locus of accountability belongs to the people function. Holding that referent fixed removes departmental tribalism from the argument.

The third confusion collapses three distinct classes of enterprise AI under the single label "digital transformation". Information systems scholarship spent three decades studying how large enterprises absorb major software platforms such as enterprise resource planning (ERP) suites, naming that process *IT-enabled organisational transformation* (ITOT). In a comparative longitudinal study of a French university hospital implementing an electronic medical record and a Finnish industrial manufacturer shifting from hardware sales to digital services, Wessel et al. (2021) separated ITOT from external *digital transformation* (DT). In ITOT, technology supports an existing value proposition and reinforces established identity; clinicians keep practising medicine the same way while administrative workflows grow cleaner (Wessel et al. 2021, pp. 1, 38–39). In DT, digital technology redefines the external value proposition, alters core value-creating work, and forces a new organisational identity (Wessel et al. 2021, pp. 39–41). Wessel et al. show that transformation activities impose direct shocks on ground-floor work practices (metrics of success, professional knowledge, and internal power relations) that "if not attended to, may derail the whole transformation agenda" (2021, p. 38).

Precision requires separating Wessel et al.'s framework into three institutional categories: 1. **Pure ITOT:** Backend cloud migrations, database upgrades, and isolated operational tools that leave core human roles intact. Technology-led execution belongs cleanly here. 2. **External product and business-model DT:** AI systems that redefine what the enterprise sells to external customers (credit underwriting engines, clinical diagnostics, dynamic customer pricing). Strategic framing of these external propositions belongs to the CEO, Corporate Strategy, and Commercial Product Leaders. 3. **Work-system AI:** Generative and agentic systems that reallocate human judgement, reshape role boundaries, alter apprenticeship and career ladders, redesign performance incentives, or trigger statutory employment duties across the internal workforce, whether or not the firm's external business model changes.

Baiyere, Salmela and Tapanainen (2020) captured the internal structural break of work-system AI in a two-year ethnographic study of deep transformation. Traditional business process management relied on three ERP-era assumptions: static process modelling, infrastructural alignment to fixed workflows, and *procedural actors* following documented rules (Baiyere, Salmela & Tapanainen 2020, pp. 3–4, 36–37). Generative digital change breaks all three. Process manuals become obsolete faster than analysts can update them (p. 26). To survive, the case enterprise shifted to *light-touch processes*, *infrastructural flexibility*, and *mindful actors* "whose judgement would be critical in choosing appropriate actions, particularly in situations where appropriate models or guidelines were not available" (Baiyere, Salmela & Tapanainen 2020, p. 36). IT-enabled transformation resembles a cub growing into a faster lion; deep work-system change resembles a larva metamorphosing into a butterfly (p. 38).

Generative and agentic AI push that metamorphosis past prior software precedents. Makarius et al. (2020, p. 271) note that legacy ERP suites were "organization centered systems" requiring operators to bend their habits to deterministic code. AI systems behave as active collaborators that sense, learn, and adapt across non-routine tasks, creating *human-AI centred systems* whose durable value lies in *sociotechnical capital* (Makarius et al. 2020, pp. 263, 265–266). Treating work-system AI as pure ITOT places a governance problem inside the wrong logical category.

1.2 *Techne*, *Politikē*, Autonomous Technique, and the Category Mistake

Philosophy gives precise names to the structural error unfolding in corporate boardrooms while also exposing the boundaries of any simple departmental fix. In Book VI of the *Nicomachean Ethics*, Aristotle separated two distinct forms of reasoned capacity dealing with variable human affairs: *techne* (art, craft, or making) and *phronesis* (practical wisdom or acting). Making and acting belong to different kinds of activity: "neither is acting making nor is making acting" (Aristotle 2009, Book VI, ch. 4, 1140a). *Techne* aims at producing a sound artefact whose purpose sits outside the activity itself, such as a secure cloud architecture or a low-latency model pipeline (1140a). *Phronesis* is "a true and reasoned state of capacity to act with regard to the things that are good or bad for man", where good action itself is the end (Aristotle 2009, Book VI, ch. 5, 1140b).

Aristotle pushes this distinction one vital step further. He identifies *politikē* (political or legislative wisdom) as the architectonic, "most authoritative art and that which is most truly the master art", ordaining how subordinate crafts fit within the wider institutional order (Aristotle 2009, Book I, ch. 2, 1094a–b; Book VI, ch. 8, 1141b). That architectonic passage immediately invites a pointed reply: inside a corporation, the architectonic ruler is the CEO and board, never a single functional

officer. I accept that reading directly. The CEO and board hold supreme strategic *politikē* over enterprise purpose, capital allocation, and constitutional chartering. Because a CEO cannot personally architect workflows across hundreds of concurrent AI deployments (Weill & Ross 2004, p. 126), the board delegates architectonic framing over internal work systems to a horizontal institutional home.

I also avoid caricaturing corporate functions. Modern enterprise architects exercise practical judgement under real trade-offs, while routine HR operations (payroll tables, job-grading matrices, compliance checklists) consist heavily of administrative *techne*. Commercial enterprises also operate under capital-return pressures that constantly pull practical wisdom toward cost extraction. Even under those caveats, institutional charters shape what each function is staffed, measured, and rewarded to optimise. Technology's primary institutional charter centres on *techne*: resilient infrastructure, cyber defence, data integrity, and reliable software artefacts. Work-system AI requires *techne* at every turn, yet its governing institutional problem belongs to *phronesis*: deciding what judgement humans retain, how capability pipelines evolve, and what fair workplace order looks like under uncertainty.

Jacques Ellul sharpened this warning from two directions that cut equally against Technology primacy and unreformed HR. In *The Technological Society*, Ellul defined *technique* as "the totality of methods rationally arrived at and having absolute efficiency (for a given stage of development) in every field of human activity" (Ellul 1964, Note to the Reader). Modern technical civilisation runs on "the quest of the one best means in every field" until "technique has become autonomous; it has fashioned an omnivorous world which obeys its own laws" (Ellul 1964, ch. 1). Crucially, Ellul turned that critique onto human-relations management itself. In his analysis of *human techniques*, Ellul classified conventional personnel administration, vocational selection, and industrial psychology alongside mechanical engineering: "Among the most notable techniques developed and applied in this area are public relations and human relations, which have as their goal to associate, adapt, and integrate the human individual into the technical milieu in such a way that he will not suffer from it" (Ellul 1964, ch. 5). Ellul's boomerang destroys any complacent claim that handing AI governance to a legacy HR department automatically humanises technology. An unreformed HR department functioning as a change-management lubricator is itself an instrument of autonomous *technique*. Restraining autonomous *technique* inside a commercial firm requires a chartered people function constrained by genuine worker voice, statutory floors, and independent assurance.

Assigning a *phronesis* problem to an institutional home engineered for *techne* commits what Gilbert Ryle (1949) called a *category-mistake*: representing the facts of one logical type as if they belonged to another (1949, p. 26). His famous visitor is shown Oxford's colleges, libraries, playing fields, museums, scientific departments, and administrative offices, and then asks, "But where is the University?" (Ryle 1949, p. 26). As Ryle noted, the most consequential category-mistakes are made by highly competent practitioners who master familiar concepts in daily work yet allocate them to the wrong logical type when theorising (1949, p. 27). Taken strictly, Ryle's analogy cuts cleanly in both directions: transformation itself has no departmental address, because transformation emerges from the constitutional organisation of the whole firm. What *does* have an assignable address is **accountability** for framing the work system.

Hannah Arendt diagnosed the political sister of Ryle's error in *The Human Condition* as "the traditional substitution of making for acting" (Arendt 2013 [1958], Part V, sec. 31): the impulse to escape the negotiated, plural messiness of human affairs by treating institutional change as deterministic fabrication. Later, in *On Violence*, Arendt added an equally piercing warning for modern bureaucracies: complex committee hierarchies degenerate into **"rule by Nobody"**, the most tyrannical administrative form because no identifiable actor can be called to account when harm occurs (Arendt 1970, pp. 38–39). Between deterministic fabrication (IT-led rollout) and rule by Nobody (unowned cross-functional committees) sits assignable institutional accountability.

Shannon Vallor updates Aristotle's practical wisdom directly for contemporary AI. In *Technology and the Virtues*, Vallor argues that rapid technological change induces "acute technosocial opacity", a growing blindness that obscures how institutions identify and secure human flourishing (Vallor 2016, intro.). Static ethical checklists fail under acute technosocial opacity because workplace conditions shift faster than rulebooks anticipate. Emerging technology demands *technomoral wisdom*, Vallor's modern reconstruction of *phronesis* as "'complete virtue,' representing the successful integration of a person's moral habits, knowledge, and virtues in an intelligent, authentic, and expert manner" (Vallor 2016, ch. 6, sec. 6.13). Individual virtue alone cannot withstand quarterly commercial pressure; an enterprise deploying work-system AI needs an institutional constitution built to hold space for technomoral judgement before autonomous *technique* hardens into daily routine.

1.3 Seventy-Five Years of Sociotechnical Proof and Its Real Lesson

The price of letting engineering logic override work structure was documented seventy-five years before large language models existed. In 1951, Eric Trist and Ken Bamforth (a former miner with eighteen years underground) published their two-year field study of mechanised longwall coal-getting in British collieries. Above-ground mining engineers, some of whom had never worked underground (Pasmore et al. 2018, p. 2), imported factory mass-production logic into the pits: coal-cutters and conveyors stretched across 200-yard faces split into three rigid shifts, seven single-skilled roles, and five separate payment schemes (Trist & Bamforth 1951, pp. 9, 14–19). Output stagnated, absenteeism rose, and safety deteriorated.

Trist and Bamforth identified why engineering reports missed the collapse: "Accounts so far appearing have presented recent changes almost exclusively in engineering terms" (1951, p. 10). Their diagnosis of reductionist calculation speaks directly to modern AI rollouts: "The production engineer might write a simple equation: 200 tons equals 40 men over 200 yds. over 24 hours. But there are no solutions of equivalent simplicity to the psychological and social difficulties raised" (Trist & Bamforth 1951, p. 14). Splitting whole tasks across disconnected shifts fractured the self-regulating group ("Differentiation gives rise to the need for social as well as technological integration", p. 19). Inside a rigid technical system they could neither control nor trust, miners developed four adaptive defences: fragile informal cliques (p. 30), reactive individualism and pay-note secrecy (p. 31), mutual scapegoating across shifts where "nothing is resolved and no one feels guilty" (p. 32), and self-compensatory absenteeism (pp. 33–34), settling into an entrenched "norm of low productivity" (p. 23). A ripping team inside the same mine that retained a complete, intelligible group task kept high morale and low absenteeism, showing that the pathology stemmed from work structure instead of worker character (pp. 36–37).

Intellectual candour requires noting what Trist and Bamforth actually prescribed at Haighmoor. Their remedy was composite, self-regulating work groups designed jointly by miners and researchers; corporate personnel departments played no part in it. Read honestly, the Tavistock lineage supports **worker voice and autonomous work design** inside the governance frame, warning against paternalistic HR control just as strongly as against engineering reductionism.

Three decades later, Shoshana Zuboff (1988, pp. 9–11) bridged industrial mechanisation to enterprise computing in *In the Age of the Smart Machine*, contrasting technology's capacity to **automate** (substituting for human labour to cut cost and tighten control) with its capacity to **informate** (generating digital visibility over underlying processes that demands deeper human intellective skill). Wanda Orlikowski (2000, pp. 404–408) added a vital practice-based insight: work structures never sit fixed inside software specifications at procurement; they emerge dynamically as *technologies-in-practice* enacted by workers in daily use, meaning governance cannot end at initial procurement charters.

In contemporary algorithmic workplaces, Katherine Kellogg, Melissa Valentine and Angèle Christin (2020, pp. 366–372) synthesised how software platforms exercise workplace control across six mechanisms: **directing** (recommending and restricting), **evaluating** (recording and rating), and **disciplining** (replacing and rewarding). Reviewing post-2020 empirical AI studies, Sharon Parker and Gudela Grote (2022, pp. 1215, 1217) find that proactive work design remains severely neglected during digital deployments, widened by an exponential gap between technical velocity and stagnant organisational design (Pasmore et al. 2018, p. 7). When a bank introduced a substitutive AI loan-decision system without redesigning roles, senior loan consultants lost override authority, saw their specialist judgement sidelined, and turned to covert data manipulation to outsmart the self-learning algorithm until the machine adapted to their gaming (Parker & Grote 2022, pp. 1217–1218). Reviews of algorithmic management likewise find that automated scheduling, goal setting, and monitoring routinely intensify workload, silence worker voice, and erode autonomy when work design sits outside system implementation (p. 1217).

Louise Harder Fischer, Nico Wunderlich and Richard Baskerville (2023) sharpen that empirical pattern into the *sociotechnical reversal*. In classical sociotechnical systems, physical work came first and software supported it. Under datafication and algorithmic management, digital reality is engineered first and human work is forced to conform afterwards (Fischer, Wunderlich & Baskerville 2023, pp. 229–230). Their conclusion goes to the heart of institutional design: "Where the digital world is created first, and copied into the physical world, there is necessarily a sociotechnical imperative for humanistic outcomes in the digital world. Otherwise, the impact of the digital world on the humanistic outcomes in the physical world becomes accidental at best" (Fischer, Wunderlich & Baskerville 2023, p. 232).

Technology never dictates a single work design (Pasmore et al. 2018, p. 3). Whenever an enterprise lets engineering logic set the architecture first and leaves human work as downstream adjustment, value stalls into quiet compliance, workslop, and covert workarounds. Examining the prevailing academic, economic, and corporate discourse reveals why that failure persists.

CHAPTER 2

The State of the Question, and the Capability-Mandate Paradox

Enterprise AI governance suffers from a split intellectual inheritance in which political economists, sociologists of professions, organisational economists, management scholars, and corporate advisers diagnose the same institutional fracture from opposite sides of the wall. Surveying peer-reviewed scholarship alongside executive practice and corporate benchmarks reveals why neither pure Technology primacy nor unreformed HR ownership survives scrutiny.

2.1 Political Economy, Jurisdictional Contests, SHRM, and Critical HRM

The direction an enterprise gives to AI is an institutional choice before it is a technical constraint. In *Power and Progress*, economists Daron Acemoğlu and Simon Johnson (2023) show across a millennium of economic history that whoever holds institutional framing power decides whether new tools expand shared human capability or concentrate control while degrading work. They contrast pursuing **"machine intelligence"** as an end in itself with designing for **"machine usefulness"**, "meaning how machines can be most useful to humans", chiefly by complementing worker capabilities and creating new tasks (Acemoğlu & Johnson 2023, intro.). When institutional power remains concentrated inside technical and capital-optimising silos, firms slide predictably into **"so-so automation"** (automating tasks or shifting burdens onto workers and customers for negligible productivity gains, extracting greater effort while failing to raise shared prosperity) (Acemoğlu & Johnson 2023, intro. & ch. 9). Where an enterprise places framing accountability decides whether AI builds machine usefulness or scales so-so automation.

Sociology of professions explains why the resulting conflict between HR and Technology looks so bitter inside modern corporations. In *The System of Professions*, Andrew Abbott (1988) showed that professions operate as an interdependent ecology competing for **jurisdiction** over contested domains of expert work through claims of abstract knowledge. Generative and agentic AI disrupt an existing jurisdictional settlement: technologists claim jurisdiction over AI through engineering mastery over models and data pipelines, while HR claims jurisdiction through employment relations and organisational design. Whenever an incumbent profession fails to upgrade its abstract knowledge base as the underlying work shifts, adjacent professions absorb its jurisdiction (Abbott 1988). That ecological mechanism explains both the anxiety behind corporate HR benchmarks and why boards replace traditional CHROs with Strategy and Data executives when incumbent HR lacks technical fluency.

Organisational economics states the resulting governance trap formally. Philippe Aghion and Jean Tirole (1997, pp. 1–29) distinguished **formal authority** (the chartered right to decide) from **real authority** (effective control over decisions arising from superior information). When a board grants formal authority over AI transformation to a people function that lacks technical literacy and trusted data access, real authority stays inside Technology, because decision-makers lacking independent information rubber-stamp the proposals prepared by better-informed engineers. Formal charter without informational capacity is hollow.

Information systems and management literatures converge on that institutional fracture from opposite poles. On the technology side, Enholm et al.'s (2022) systematic review of 43 AI business-value studies in *Information Systems Frontiers* concludes that top managers "should thus take part in exploring AI technologies and not leave this solely to the technologists" (p. 1718). Realising enterprise value requires technical builders working alongside domain experts who understand workflow logic (p. 1718), backed by structural shifts in cross-departmental collaboration and data governance (p. 1719). Legacy IT governance breaks down under AI, Enholm et al. warn, "since the effects towards, as well as the interactions with humans shifts fundamentally" (2022, p. 1726), and unmanaged role displacement breeds distrust aimed at the units promoting deployment (p. 1727).

On the people side, four decades of Strategic Human Resource Management (SHRM) scholarship provide both a conceptual spine and necessary critical restraint. Lee Dyer's foundational distinction (1985, cited in Wright & Ulrich 2017, p. 46) separated *organisational HR strategy* (aligning workforce allocation, utilisation, and capability with enterprise goals) from *functional HR strategy* (running the HR department). Meta-analytic synthesis across 92 empirical studies by Combs et al. (2006, pp. 501–528) found that coherent bundles of high-performance work practices correlate positively with operational and financial performance (overall corrected correlation 0.20, explaining roughly 4% of performance variance). Yet Wall and Wood (2005, pp. 429–462) warned clearly against romanticising universalistic HRM causal claims without rigorous longitudinal evidence, a caution echoed by Wright and Ulrich's (2017, pp. 49, 60) warning against chasing "the newest, shiniest toy" without grounded theory.

Equally vital is the pluralist challenge from **Critical HRM** scholarship. Karen Legge (2005) exposed the enduring gap between HRM's developmental *rhetoric* (capability, commitment, mutual gains) and its commercial *reality* inside shareholder-owned firms, where HR departments frequently act as instruments of cost compression and compliance control. Taking Legge's critique seriously alongside Ellul's warning on *human techniques* removes any naive unitarist assumption that HR executives are unconflicted custodians of worker flourishing. Because a CHRO answers to the CEO and board on commercial outcomes, vesting work-system framing in the people function remains credible only when constrained by genuine worker consultation, statutory employment floors, and independent assurance separated from management delivery targets.

When those disciplinary streams meet real enterprise org charts, a glaring institutional gap opens. In *Workforce Ecosystems*, Altman et al. (2023) document how large organisations fragment their own labour and technology governance: "AI and automation are part of the tech organization (but rarely regarded as part of the broader workforce ecosystem). Meanwhile, HR oversees full- and part-time employees" (Altman et al. 2023, ch. 6). Procurement buys contingent contractors and managed-service vendors as if they were office supplies; Technology deploys software agents as if they were servers; HR administers permanent headcount inside legacy HRIS silos. No single institutional structure governs human-agent work as an integrated system.

2.2 The Executive Inflection Point: New Roles, Structural Mergers, and Binding Law

Across 2025 and 2026, three simultaneous pressures stripped boards of the luxury of leaving that siloed arrangement intact. First, C-suite experimentation accelerated. An unrefereed University of Oxford working paper analysing archival board records identified 258 organisations operating a chief-level AI executive by August 2026, spanning 127 dedicated Chief AI Officers and 134 hybrid

titles across dual-role and multi-spell observations (Schmitt 2026, p. 19), alongside high-profile structural experiments such as Moderna's merger of HR and digital technology (Bousquette 2025) and AXA Germany's creation of a board-level "People and Data Experience" portfolio while retaining separate core IT (AXA Konzern AG 2025).

Second, employment, workplace safety, and privacy law moved from voluntary AI principles to enforceable statutory obligations across Europe, California and other US states, Chile, the United Kingdom, and Australia. Chapter 3 examines these enacted statutory regimes and legislative trajectories in full; here, their significance for the state of the question is immediate. Governance accountability is now a board-level compliance imperative alongside an operating-model choice.

Third, enterprise executives face mounting evidence that capital expenditure on AI platforms is outrunning realised workflow value. When technical infrastructure matures faster than work redesign and skill architecture, boards turn to management consultancies and vendor research for an operating-model blueprint.

2.3 What Corporate Benchmarks Quietly Admit: The Capability-Mandate Paradox

Large-scale consulting surveys deserve critical scrutiny for two reasons: they capture current executive sentiment faster than multi-year peer-reviewed cycles, and they act as commercially interested artefacts showing how advisory markets frame the problem they sell into. McKinsey & Company's *HR Monitor 2026* (fielded in January 2026 and published in June 2026 across 1,303 HR decision-makers and 5,501 employees in ten countries) illustrates that tension clearly (McKinsey & Company 2026, pp. 3, 40–41).

Three methodological limits narrow what such corporate benchmarks can establish. First, advisory firms hold a direct commercial interest in the operating-model redesigns, skills taxonomies, and agentic transformations their reports prescribe. Second, McKinsey's survey sets its firm-size floor at 100 employees with no disclosed enterprise-size stratification in the public report (McKinsey & Company 2026, pp. 40–41). Pooling 150-person businesses with 80,000-person multinationals means that aggregate weakness indicators (such as no agentic awareness or pre-Ulrich generalist structures) may sit disproportionately inside the smaller-firm tail; those figures therefore serve as directional warning signals instead of precise large-enterprise measurements. Third, adding China, Belgium, and the Netherlands to the 2026 sample makes year-on-year comparisons directional instead of longitudinal (p. 3), and citing Moderna and AXA Germany on page 32 as aspirational blueprints supplies no independent outcome data.

Even under those caveats, *HR Monitor 2026* exposes a stark contradiction between advisory ambition and functional readiness. Alongside its normative call for HR to act as enterprise "architect of AI transformation", internal "lighthouse", and "copilot of the AI enterprise" (McKinsey & Company 2026, pp. 3, 37), its own survey exhibits document an acute maturity gap:

- **Short strategic horizons:** Although 62% of HR respondents claim organisation-wide workforce planning, only 11% forecast workforce needs three years or more ahead (7% three to five years; 4% beyond five years), while 62% look under twelve months ahead (McKinsey & Company 2026, pp. 9–10). Skills taxonomies exist in 85% of surveyed firms, yet only 22% use skill-based forecasting while 46% still plan headcount by legacy job titles (p. 10).

- **Fragmented technical maturity:** Half of surveyed HR functions still run classic three-pillar Ulrich models (32%) or pre-Ulrich generalist structures (18%), while only 15% report a strongly machine-powered HR operating model (pp. 32–33). Operational AI in core HR sits at 28% (with 37% stuck in pilot), concentrated in routine administration such as time tracking (33%) and employee data management (30%) (p. 35). HR practitioners lowered their own estimate of core HR automation potential from over 30% in 2025 down to 20% in 2026 after colliding with fragmented HRIS estates and inconsistent data architectures (pp. 32, 35). On agentic AI specifically, 18% of HR decision-makers report no awareness (rising to 34% in France, 30% in Belgium, and 28% in Germany), and only 10% report transformative impact (p. 36).
- **Workforce perceptual blind spots:** Surveyed HR leaders estimate employees receive 6.2 training days annually; employees report 3.4 days, with 24% receiving no training hours in the past year (pp. 20–21). While 20% of employees report receiving no formal performance feedback, HR believes the figure is 3% (p. 20). Meanwhile, 71% of employees expect AI to alter their job and skill requirements, yet formal generative AI training reaches just 25% of workers in Europe and 31% in the United States, where US coverage fell 14 percentage points from 2025 (pp. 21–22).

Here is the **capability-mandate paradox** stated both empirically and theoretically. Political economy and sociotechnical scholarship warn against leaving AI direction inside technical silos; Abbott (1988) shows how jurisdictional vacancies invite adjacent professions to absorb contested work; Aghion and Tirole (1997) show that granting formal authority without informational competence leaves real authority elsewhere; corporate benchmarks document widespread readiness gaps. To move past that impasse, I must establish why, despite HR's present capability deficit, formal accountability for framing work-system AI still belongs structurally in the enterprise people function.

CHAPTER 3

The Constructive Case for Placing Accountability in the People Function

Formal institutional accountability for framing work-system AI belongs in the enterprise people function because work design is constitutive of the transformation problem frame, whereas technical architecture and cybersecurity operate as verifiable constraints downstream of that frame. I defend that conclusion across four interlocking premises: the empirical nature of the enterprise AI bottleneck, the refusal of Information Technology governance theory to claim IT primacy, a comparative institutional quintilemma honouring Hume's Law through an explicit two-track normative premise and multi-jurisdictional law, and the organisational hysteresis of technological momentum. Throughout this chapter, I state openly what empirical evidence establishes directly (that work design sits inside the binding constraint) and where the step to the people function rests on comparative institutional reasoning.

3.1 Premise 1: The Binding Bottlenecks of Work-System AI Are Workforce Architecture Problems

When an enterprise deploys generative or agentic AI into live operations, the primary failure mode shifts rapidly from model accuracy to work design, skill transfer, and remuneration incentives. Two landmark field studies show how unevenly generative AI reshapes human work across skill tiers. First, Erik Brynjolfsson, Danielle Li and Lindsey Raymond (2025) tracked a staggered quasi-experimental rollout of a GPT-based conversational assistant across 5,172 customer support agents at a Fortune 500 enterprise software firm. Aggregate productivity (customer issues resolved per hour) rose 15% on average (Brynjolfsson, Li & Raymond 2025, p. 2). Behind that headline sat a substantial redistribution: less-skilled and less-experienced agents improved their hourly resolution rate by roughly 30%, compressing onboarding curves so substantially that treated agents with two months of tenure matched untreated agents with more than six months of tenure (pp. 2, 22). Customer sentiment improved, supervisor escalations dropped, and attrition fell among newer workers (p. 3).

Then the study surfaced a structural tension. While novice agents advanced rapidly, the firm's highest-skilled agents saw minimal speed gains alongside a small decrease in conversation quality (Brynjolfsson, Li & Raymond 2025, p. 2). The authors report directly: "In our data, top workers increase their adherence to AI recommendations, even though those recommendations marginally decrease the quality of their conversations" (p. 4). Because the AI model learns from human-generated conversations, fewer original contributions from top experts risk starving future iterations of the model of novel problem-solving patterns. As the authors concluded in the published *Quarterly Journal of Economics* edition, "compensation policies that provide incentives for people to contribute to model training could be important" (Brynjolfsson, Li & Raymond 2025, p. 915).

Second, Fabrizio Dell'Acqua et al. (2023; 2026) confirmed that this boundary problem extends beyond codifiable contact-centre scripts into elite knowledge work. In a preregistered field experiment across 758 Boston Consulting Group consultants performing 18 realistic consulting tasks, GPT-4 lifted output inside AI's reliable capabilities (+12.2% more tasks completed, 25.1% faster,

roughly 40% higher quality, with bottom-half consultants gaining 43% against 17% for top performers) (Dell'Acqua et al. 2023; 2026, pp. 403–423). Yet on a complex analytical task engineered to sit outside the "jagged technological frontier", consultants using AI were **19 percentage points** less likely to produce correct solutions than peers working without AI because unengaged users deferred to plausible-sounding model output (Dell'Acqua et al. 2023; 2026, pp. 403–423).

Pause on what unfolds philosophically and economically across both studies. Shannon Vallor identifies *moral attention* (perceiving the salient human features of a particular situation) and *prudential judgement* (refining character into practical wisdom in live choices) as the core perceptual habits sustaining practical wisdom (Vallor 2016, ch. 5, secs. 5.1–5.2). When operational throughput targets reward raw speed while an AI assistant supplies frictionless default text, even seasoned experts surrender prudential judgement to algorithmic convenience. Left unaddressed, that incentive distortion drags the enterprise into Acemoğlu and Johnson's (2023, intro. & ch. 9) *so-so automation*, eroding the human judgement that makes *machine usefulness* possible. While operational line leaders often set daily throughput metrics such as tickets per hour, Technology teams hold no mandate over job grades, variable remuneration architecture, senior specialist career ladders, or onboarding curricula. Repairing those structural incentives requires integrated people-and-operations work design beyond further model tuning.

A Harvard Business School teaching case of JPMorganChase's firm-wide generative AI rollout illustrates how those pressures surface inside a global bank (Bojinov, Lakhani & Lane 2025). After deploying its internal **LLM Suite** to 200,000 employees and building **Coach** for 3,000 Private Bank advisors, Chief Risk Officer Ashley Bacon warned explicitly against human-in-the-loop complacency: "We know that our employees learn to trust quickly what their tools can do. That can result in not triangulating them with other sources" (Bojinov, Lakhani & Lane 2025, p. 4, citing Dell'Acqua 2022). Private Banking relied on a multi-year "teaching hospital" apprenticeship pairing junior advisors with senior mentors (p. 10). Yet once **SmartBrief** automated routine preparatory tasks and **Coach** drove "the internal cost of expertise for basic information … to zero" (p. 14; Exhibit 5, p. 19), the case admitted candidly that "the effects of GenAI on workforce size remained an unknown, as were their implications for the apprenticeship model common to PB advisors and other JPMC jobs" (Bojinov, Lakhani & Lane 2025, p. 14). Because Harvard teaching cases are authored around selected protagonists (here CIO Lori Beer, CDAO Teresa Heitsenrether, AWM CEO Mary Erdoes, and CRO Ashley Bacon), the absence of HR from the 22-page narrative (despite Head of HR Robin Leopold sitting on the 15-member Operating Committee, Exhibit 1, p. 16) cannot establish that HR played no internal role behind the scenes. What that protagonist framing reveals clearly is how an elite bank's public governance narrative treated workforce architecture and apprenticeship pipelines as unresolved afterthoughts even as Asset and Wealth Management tied AI Day participation into performance reviews, tracked usage by group, and deployed **CoachAssist** to classify advisors' client calls as "great", "mediocre", or "requiring improvement" (pp. 5–6; Exhibit 5, p. 19).

Broader multidisciplinary reviews reinforce this diagnosis. Across 98 leading journals spanning HRM, information management, operations, and general management, Chowdhury et al. (2023) isolated three technical resource categories (data resources, technology infrastructure, and algorithmic transparency tools) alongside ten non-technical organisational and human categories centred on AI-employee integration, skills development, leadership, inter-team coordination, knowledge management, and ethical governance (Chowdhury et al. 2023, Table 2, pp. 15–17).

Technology investment alone rarely yields business gains because hard-to-imitate capability resides in human-machine integration (p. 19). In a companion framework for generative AI, Chowdhury, Budhwar and Wood (2024, pp. 6–7) name the core transformation task *re-institutionalisation*: restructuring norms, roles, policies, and work relations around human-AI symbiosis while guarding against skill obsolescence and workplace authoritarianism (pp. 4, 10).

Public-sector IT directors report the same boundary. Drawing on the European Commission's 1,295-case Public Sector Tech Watch dataset, Tangi, Rodriguez Müller and Janssen (2025) interviewed 14 integration experts across nine European public administrations (twelve IT directors or AI unit heads, no HR directors, Table 1, p. 5). Allowing for civil-service tenure differences and noting that these interviews establish only that the bottleneck sits outside IT (never by themselves naming which non-IT function should resolve it), their admission remains telling. One national IT director stated, "We are [at] a good level in aspects like data strategy and data privacy … when it comes to people and processes, there are bottlenecks that need to be addressed" (p. 7). Another manager overseeing 14 production AI applications with a 30-person unit reported that "65% of our time is dedicated to the maintenance and improvement of these applications", after dedicating daily staff attention for 18 to 24 months to bring a single conversational agent to proficiency (Tangi, Rodriguez Müller & Janssen 2025, pp. 7–8).

3.2 Premise 2: Information Technology Governance Theory Rejects IT Primacy

Advocates of Technology-led AI transformation often assume they have the weight of enterprise IT governance scholarship behind them. Canonical IT governance texts reject IT primacy directly, constructing institutional boundaries to prevent Ellul's autonomous *technique* (1964, ch. 1) from dictating organisational ends.

In 1993, Henderson and Venkatraman (1999 [1993], p. 473) established that "no single I/T application, however sophisticated and state of the art it may be, could deliver a sustained competitive advantage"; durable value comes from continuous organisational adaptation around technology. They warned against drawing a straight line from IT strategy directly to organisational restructuring: "the direct link between an I/T strategy and organizational infrastructure has no straightforward logic. One cannot and should not simply seek to identify and adopt the best available technology to restructure the organization or streamline the business processes" (Henderson & Venkatraman 1999, p. 481). Any Technology-led AI rollout that procures agent platforms first and expects roles and workflows to adjust afterward violates Henderson and Venkatraman's rule.

In 2004, Peter Weill and Jeanne Ross surveyed 256 large enterprises across 23 countries (averaging eight business units and 850 IT staff each, p. 58) for MIT Sloan's Center for Information Systems Research (CISR). They opened *IT Governance* with an explicit rejection of IT-led ownership:

> "The tight linkage between IT and organizational processes means that the IT unit cannot bear sole, or even primary, responsibility for the effective use of information and information technology. Getting more value from IT is an increasingly important organizational competency. Leaders throughout an enterprise must develop this competency." (Weill & Ross 2004, pp. 1–2)

Firms with above-average IT governance achieved 20% higher return on assets than poorly governed peers pursuing the same strategy (Weill & Ross 2004, pp. 2, 122), because top performers confined *IT Monarchy* to technical architecture and infrastructure standards (pp. 59–60) while assigning explicit non-IT accountability "for the organizational changes required to benefit from new IT capabilities" (p. 2). Ross, Weill and Robertson (2006, pp. viii, 8) reiterated that enterprise architecture is a business issue anchored in the firm's operating model. In June 2026, Sebastian, Weill, Haskamp and vom Brocke (2026, pp. 1–3) extended that logic in the *MIT CISR AI Decision Matrix*, splitting every enterprise AI decision into **Frame** (defining the problem, assumptions, constraints, and success criteria), **Act** (gathering information, evaluating options, executing approved steps), and **Learn** (monitoring outcomes, holding accountability, updating the system). For consequential decisions, MIT CISR assigns **Frame** and **Learn** to named business and domain owners while the central AI and data function supplies unified platforms and technical guardrails (Sebastian et al. 2026, pp. 2–3).

Candour requires confronting an immediate tension inside Premise 2 before moving on. Both Weill and Ross (2004) and Sebastian et al. (2026) reject IT primacy, yet both hand **Frame** and accountability to *business and domain owners* (and Weill and Ross's 2004 leverage table places CEOs, COOs, and business-unit leaders at the top without mentioning a CHRO). Premise 2 therefore eliminates Technology primacy cleanly while supporting Operations-led governance (examined in Chapter 4.4) just as readily as it supports the people function. Deciding *which* non-IT function should hold framing accountability for work-system AI requires Premise 3: our explicit normative grounds, the legal duties binding employers, accountability theory, and the Comparative Institutional Quintilemma.

3.3 Premise 3 and the Comparative Institutional Quintilemma: Honouring Hume, Global Law, and the Framing Asymmetry

In *A Treatise of Human Nature*, David Hume warned against sliding silently from descriptive propositions about what *is* to normative conclusions about what *ought* to be: "instead of the usual copulations of propositions, *is*, and *is not*, I meet with no proposition that is not connected with an *ought*, or an *ought not*" (Hume 2000 [1739–40], Book III, Part I, sec. I). Supplying an explicit normative major premise does not "resolve" or defeat Hume's Law; it **honours** Hume's discipline by putting our evaluative commitments on the table where readers can inspect or contest them.

The Two-Track Normative Premise (Intrinsic and Instrumental Grounds)

My normative major premise is explicit: *a large enterprise ought to design its AI governance constitution to preserve human prudential judgement, sustainable capability development, non-domination, and lawful dignity of work, even where short-term labour-cost extraction or deskilling remains strictly legal.* Because a shareholder-primacy CEO may challenge a purely ethical axiom on sight, I defend that premise on two independent tracks:

1. **Intrinsic political-philosophical grounds (Non-Domination):** Philip Pettit (1997) defines freedom as **non-domination**, living free from the arbitrary capacity of another agent to interfere in one's choices unchecked. Elizabeth Anderson (2017) applies that republican lens directly to corporate employment in *Private Government*, demonstrating that modern firms exercise sweeping, largely unaccountable governing authority over workers' daily autonomy, speech, and livelihoods. Algorithmic management amplifies private government by automating direc-

tion, evaluation, and discipline inside opaque code (Kellogg, Valentine & Christin 2020). Institutional governance that embeds worker voice, contestation, and transparent work design limits arbitrary domination inside the workplace.

2. **Instrumental value grounds (Sustainable Enterprise Capability):** Even a board focused strictly on long-term shareholder value faces the empirical failure modes documented in Section 3.1: unredesigned work breeds covert algorithmic gaming (Parker & Grote 2022), erodes expert tacit judgement that future AI models depend on (Brynjolfsson, Li & Raymond 2025, p. 915), degrades accuracy outside the jagged frontier (Dell'Acqua et al. 2026, pp. 403–423), destroys apprenticeship pipelines (Bojinov, Lakhani & Lane 2025, p. 14), and erodes discretionary effort when employees attribute cost-cutting motives to management practices (Nishii, Lepak & Schneider 2008, pp. 503–545). The thesis therefore survives even if a reader rejects the intrinsic republican premise and accepts only long-run enterprise value preservation.

The Enacted Multi-Jurisdictional Statutory Floor (and Looming Trajectory)

Statutes legally bind **the employer as a corporate legal entity**, never HR as a department; data protection laws designate Data Protection Officers and AI statutes designate "deployers". Yet across major jurisdictions, enacted positive law attaches binding duties directly to employment decisions, workplace monitoring, worker-representative notification, and AI literacy, creating a sprawling, actively conflicting compliance floor that an employer can discharge only through its horizontal people-governance levers alongside Legal, Privacy, and Safety:

1. **European Union Artificial Intelligence Act and German Codetermination:** Annex III, point 4 of Regulation (EU) 2024/1689 (with compliance dates under Article 113 of Regulation (EU) 2024/1689, as amended by the Digital Omnibus, Regulation (EU) 2026/1744, applying from 2 December 2027) classifies AI systems as high-risk whenever used for recruitment, promotion, termination, **allocating tasks based on individual behaviour or personal traits**, or monitoring and evaluating worker performance. Two operational duties inside the AI Act bind general employers even more directly: **Article 4** obliges providers and deployers to ensure a sufficient level of **AI literacy** across staff operating AI systems, and **Article 26(7)** mandates that before putting a high-risk AI system into service at the workplace, **employers shall inform workers' representatives and affected workers** in accordance with national labour law (European Parliament and Council of the European Union 2024a). (By contrast, Chapter III of the EU Platform Work Directive, Directive (EU) 2024/2831, applies specifically to digital labour platforms, serving only as a sector-specific regulatory template for wider algorithmic management, European Parliament and Council of the European Union 2024b.) In Germany, enacted labour law goes further still: § 87(1) No. 6 and § 90 of the *Betriebsverfassungsgesetz* (*BetrVG*) grant elected Works Councils a **mandatory codetermination veto** over introducing or applying *any* technical device objectively capable of monitoring employee behaviour or performance, regardless of management intent (Federal Republic of Germany 2001).
2. **California's Twin Binding Regimes, US State Patchwork, and Litigation Signals:** Positive US state regulation already locks employer accountability into force. The California Civil Rights Council's *Employment Regulations Regarding Automated-Decision Systems* (Title 2 CCR §§ 11008–11013, in force 1 October 2025) apply to hiring, screening, promotion, pay, and leave **even when a human remains in the loop**, mandate four-year ADS recordkeeping, tie proac-

tive anti-bias testing to legal defence, and classify any third-party vendor performing traditional employer functions through an ADS as an employer "agent" (California Civil Rights Council 2025). Companion California Privacy Protection Agency regulations (Title 11 CCR §§ 7001, 7150, 7200–7222, compliance by 1 January 2027) add pre-use notices, risk assessments, and opt-out or human-appeal rights covering hiring, **allocation or assignment of work**, compensation, and termination (California Privacy Protection Agency 2025). New York City Local Law 144 (enforced since 5 July 2023, annual independent bias audits, City of New York 2021), Illinois HB 3773 (in force 1 January 2026, anti-discrimination and zip-code proxy ban, State of Illinois 2024), and Colorado SB 26-189 (signed May 2026 repealing and replacing SB 24-205, effective 1 January 2027, State of Colorado 2026) add divergent local duties. Active US federal litigation such as *Mobley v. Workday, Inc.* (No. 3:23-cv-00770, N.D. Cal., conditionally certified as an ADEA collective action on 16 May 2025 with over 15,000 opt-in plaintiffs) illustrates plaintiff appetite on vendor-as-agent liability; even if court claims narrow or settle, California regulation already codifies employer-agent responsibility in positive law.

3. **Chile Law 21.719 (Latin American Automated-Decision Safeguards):** Chile's personal data reform (*Ley N° 21.719*, entering into force 15 October 2027 under *Ley N° 21.855*, fines up to 20,000 UTM) grants individuals under Articles 8 *bis*, 14 *ter*, and 15 *ter* the right to object to automated decisions and profiling with significant effects, disclosure of applied logic, mandatory data protection impact assessments, and preserved safeguards (transparency, explanation, human intervention, and review) even under employment-contract exceptions (Congreso Nacional de Chile 2024).

4. **United Kingdom Data (Use and Access) Act 2025 (Permission-Plus-Safeguards):** Section 80 (substituting Articles 22A to 22D into UK GDPR, in force 5 February 2026 under S.I. 2026/82) adopts a lighter *permission-plus-safeguards* model than Brussels, permitting solely automated non-special-category decisions provided employers guarantee information, representations, human intervention, and contestation (Parliament of the United Kingdom 2025).

5. **Australia's Enacted Multi-Layer Floor and Policy Trajectory:** Inside Australia alone, enacted positive law spans three distinct tracks: mandatory enterprise-agreement consultation terms for major workplace technological change under section 205 of the *Fair Work Act 2009* (Cth) (Commonwealth of Australia 2009); **disclosure-only** automated-decision transparency under Australian Privacy Principles 1.7 to 1.9 of the *Privacy and Other Legislation Amendment Act 2024* (Cth) (commencing 10 December 2026, Commonwealth of Australia 2024); and New South Wales's enacted *Work Health and Safety Amendment (Digital Work Systems) Act 2026* (assented 18 February 2026), imposing a primary workplace health and safety duty directly over algorithmic work allocation, workload strain, and digital monitoring (Parliament of New South Wales 2026). Separately from enacted law, the parliamentary policy trajectory points toward stricter corporate agent liability: in November 2026, the Joint Select Committee on Artificial Intelligence tabled *Artificial Intelligence – the Australian way*, issuing 46 recommendations spanning a risk-based statutory framework with penalties (Recommendation 3), review of AI workplace monitoring and WHS impacts (Recommendation 21), holding companies legally accountable for tasks performed by AI agents (Recommendation 22), early worker consultation (Recommendation 39), and employer investment in reskilling and redeployment (Recommendation 40) (Parliament of Australia, Joint Select Committee on Artificial Intelligence 2026).

The resulting compliance challenge is defined by **scale, expanding scope, and active conflict** across borders: ex-ante high-risk rules in Brussels and mandatory Works Council vetoes in Berlin sit alongside post-hoc contestation in London, human-in-the-loop ADS liability in California's Civil Rights Council, disclosure-only federal privacy in Canberra, and primary digital-work WHS duties in Sydney. As Chowdhury et al. (2023, p. 13) note, "It should not be acceptable that 'blame' for such inappropriate outcomes such as prejudice fall upon a 'mathematical model'."

Accountability Theory and the Comparative Institutional Quintilemma

To move legitimately from employer liability and workforce bottlenecks to an institutional locus without committing an argument from status-quo incumbency, I draw on **comparative institutional analysis** (Komesar 1994) and Oliver Williamson's **remediableness criterion** (1996, ch. 8): an existing institutional allocation should be altered only where a feasible alternative yields expected net institutional gains after accounting for real transition costs and specialist capabilities. I pair that economic lens with the **accountability literature**: Dennis Thompson's (1980, pp. 905–916) **"problem of many hands"** (where diffused collective governance makes it impossible to assign responsibility when harm occurs), Mark Bovens's (1998) analysis of organisational responsibility, and the AI **responsibility-gap** literature (Matthias 2004, pp. 175–183; Santoni de Sio & Mecacci 2021, pp. 1057–1084), which shows that self-adapting AI creates culpability, moral accountability, public accountability, and active-responsibility gaps unless meaningful human control is anchored in identifiable institutional roles.

Against our two-track normative premise, coupled employment levers, and accountability theory, a large enterprise faces **five** structurally distinct institutional options for governing work-system AI:

- **Option A (Split Framing from People Levers):** Leave job architecture, pay, and industrial relations in the people function while vesting work-system AI framing in Technology, a standalone CAIO, or autonomous business units with no upstream people gate. This splits problem framing from employment accountability, generating unredesigned jobs (Parker & Grote 2022, p. 1217) and unmonitored shadow performance metrics (Belizón & Kieran 2022, p. 21).
- **Option B (Absorb People Levers into Technology or a Transformation Office):** Strip job design, remuneration, industrial relations, and employment governance out of the people function and merge them into Technology or a new transformation office. Under Williamson's remediableness criterion, this fails on high transition cost, specialist capability dilution (Bousquette 2025, p. 2; Schmitt 2026, pp. 7–8), and proponent-gatekeeper collapse.
- **Option C (Parity-of-Gates Federation: Strategy or Operations Frames; People Function Holds Downstream Gate):** By apparent logical parity, if Technology need not lead overall transformation because exclusive technical-architecture authority plus a pre-deployment cybersecurity gate protects technical systems (detailed across six formal domains in Chapter 5), an executive might argue that granting the people function analogous downstream work-redesign ownership and a pre-deployment employment veto suffices. The symmetry breaks on **framing architecture**. Cybersecurity and cloud architecture are *constraint specifications* verifiable after a business problem is framed; work design and human-agent capacity are *constitutive of the problem frame itself*. When a Strategy or P&L unit frames an internal AI initiative around a 25% headcount reduction before the people function holds the pen, a downstream veto arrives after the commercial case, vendor shortlist, and executive expectations have hard-

ened. The people function can then block unlawful bias or negotiate redundancies, yet it cannot reframe *so-so automation* into *machine usefulness* without triggering an adversarial board escalation. (External customer-product AI strategy properly follows Option C; internal work-system AI cannot.)

- **Option D (CEO- or ExCo-Chaired Collective Co-Framing):** Establish no single functional locus; instead, a CEO- or Executive-Committee-chaired group frames every work-system AI initiative collectively, with the people function acting as an equal mandatory co-signer on the Work and Capacity Charter. This is the strongest rival because it sits upstream inside the frame, and Chapter 5 deliberately adopts a variant of it as a temporary transition safeguard (Months 1 to 18) and emergency fallback. As a **steady-state operating constitution**, however, collective co-framing is structurally second-best for three reasons. First, CEO bandwidth is strictly bounded: Weill and Ross (2004, p. 126) show CEOs achieve 2.1x governance leverage through periodic strategic chartering, never by mediating work design across dozens of concurrent enterprise AI programmes. Second, equal co-framing reproduces Thompson's (1980) problem of many hands and Arendt's (1970, pp. 38–39) **"rule by Nobody"**: when an apprenticeship pipeline quietly erodes or contested incentives stall, collective sign-off diffuses accountability across four executives. Third, when a Work and Capacity Charter is genuinely contested between margin reduction and capability preservation, co-equal pens produce veto-player stalemate (Tsebelis 2002) unless one named function holds primary drafting authority and accountability for the work system.
- **Option E (People-Function Framing Locus in Binding Partnership):** Retain horizontal people levers in the enterprise people function and vest primary upstream framing accountability for **work-system AI** in that function under CEO charter, bound symmetrically by hard Technology, Operations, Finance, and Three-Lines Assurance gates (specified in Chapter 5), conditional on closing its technical fluency gap.

Under Komesar's comparative institutional method, Option E is never claimed as a flawless ideal; it is the least defective feasible equilibrium available to a complex regulated enterprise, placing human work inside the initial frame while maintaining single-point accountability against Arendt's rule by Nobody.

3.4 Premise 4: Organisational Hysteresis, the Automation-Augmentation Paradox, and Technological Momentum

One final premise explains why framing accountability must sit upstream before deployment begins. In *The Imperative of Responsibility*, Hans Jonas (1984) argued that modern technology breaks classical ethics because predictive knowledge lags the causal reach of technical interventions, proposing a "heuristics of fear" that gives "the prophecy of doom priority over the prophecy of bliss" (Jonas 1984, pp. x, 26). Jonas reserved his absolute zero-miss rule ("we really must allow none", p. 31) for civilisational survival threats; importing that absolutism untouched into corporate management would be disproportionate.

What enterprise AI exhibits is **severe organisational hysteresis**: asymmetries of rebuilding cost and multi-year recovery lag. Risks cut symmetrically in both directions: prolonged governance paralysis carries competitive and solvency risk, and technical failures such as data exfiltration via prompt injection or poisoned data lineage are equally hard to undo. That symmetry justifies a

joint upstream assurance gate spanning Cybersecurity, Risk, Legal, and the people function. Once a system clears technical security, however, the remaining asymmetric hysteresis falls heavily on the human system: a delayed software sprint can be re-sequenced next quarter, whereas eroding a multi-year apprenticeship pipeline (Bojinov, Lakhani & Lane 2025, p. 14), deskilling mid-career specialists (Brynjolfsson, Li & Raymond 2025, p. 4), embedding disparate impact across screening (California Civil Rights Council 2025), or fracturing trust through covert monitoring (Parker & Grote 2022, p. 1217) takes years to rebuild.

This also illuminates what Sebastian Raisch and Sebastian Krakowski (2021, pp. 192–210) theorise as the **automation-augmentation paradox**: automating routine tasks and augmenting human experts are interdependent, persistent tensions that organisations must manage dynamically instead of treating as a binary choice. In Brynjolfsson, Li and Raymond's (2025) study, generative AI lifted novice productivity by 30% while marginally lowering elite conversation quality. Blind precaution would block a tool that helps lower-skilled workers advance; blind acceleration hollows out senior expertise. Managing that paradox requires deliberate upstream work redesign that captures novice acceleration while redesigning senior roles and incentives to protect expert judgement.

Jonas identified a second mechanism familiar to every AI governance practitioner: *automotive technological momentum*. Technologies launched "with short-term aims tend to make themselves independent, that is, to gather their own compulsive dynamics, an automotive momentum, by which they become not only, as pointed out, irreversible but also forward-pushing and thus overtake the wishes and plans of the initiators. The motion once begun takes the law of action out of our hands" (Jonas 1984, p. 32). Paired with Orlikowski's (2000, pp. 404–408) finding that *technologies-in-practice* crystallise rapidly in daily routines, once an agentic workflow embeds across an enterprise, extracting it becomes politically and operationally prohibitive. Accountability for human-work harms must sit inside the upstream problem frame and remain continuously active post-deployment.

That completes my constructive case. I now test my thesis against the five strongest rival objections and examine every point where they compel qualification.

CHAPTER 4

Objections, Rivals, and the Capability-Mandate Paradox

Five rival objections challenge my thesis that formal accountability for framing work-system AI should sit in the enterprise people function, and the strongest expose genuine vulnerabilities in current HR maturity and scope boundaries. I construct each rival at maximum strength before examining where it breaks down and where it forces a disciplined qualification of my thesis.

4.1 Objection 1: The Technology-Led Case and the Capability-Mandate Paradox

The most formidable counter-argument concedes that AI reshapes human work yet concludes that placing framing accountability in the people function fails on epistemic and operational grounds. A modern CIO or CTO is no 1995 server administrator. Senior technology executives run product operating models, employ human-centred design researchers, manage agile cross-functional squads, and understand data pipelines at depth. JPMorganChase's technology and risk organisation shows that ceiling at industrial scale: backed by a $17 billion annual technology budget and 63,000 technologists, CIO Lori Beer, CDAO Teresa Heitsenrether, and CRO Ashley Bacon built a unified cloud data mesh, a multi-model **LLM Suite** serving 200,000 employees on consumption pricing, 12-level hallucination guardrails, and mandatory control-framework registration for every AI model (Bojinov, Lakhani & Lane 2025, pp. 4, 7–9). Agentic AI systems fail in subtle technical ways (prompt injection, silent configuration drift, permission escalation across enterprise APIs, data-lineage contamination, and runaway metered token costs) that non-builders miss. Practitioners who have never built or red-teamed an AI workflow govern the *described* version of a system on a slide deck while missing how the *deployed* system behaves in production.

Organisational economics gives this objection formal teeth. Under Aghion and Tirole's (1997, pp. 1–29) distinction between **formal authority** and **real authority**, vesting formal framing rights in an HR function lacking technical fluency and integrated data leaves real authority inside Technology, because poorly informed executives inevitably rubber-stamp proposals drafted by better-informed engineers. Longitudinal peer-reviewed research turns that formal prediction into empirical warning. María Jesús Belizón and Sarah Kieran (2022) tracked the institutional legitimacy of HR Analytics over three years (2017 to 2020) across three large Irish-headquartered enterprises: Alpha (a telecoms firm with over 3,400 staff), Delta (a commercial bank with over 10,000 staff), and Gamma (a multinational retailer with over 65,000 staff). Using process theory across 22 multiparty interviews, they evaluated cognitive legitimacy (domain expertise), socio-political legitimacy (executive acceptance), and technological legitimacy (integrated data infrastructure) (Belizón & Kieran 2022, pp. 4, 6–7). Because HRM historically lacked a data-driven decision legacy (p. 4), siloed HR sponsorship and fragmented legacy HRIS estates stalled legitimacy inside Delta and Gamma. Senior HR leaders struggled to envision analytical value, both HR Analytics Leads left the field, and in both Delta (May 2019) and Gamma (February 2020), the CEO transferred HR Analytics out of HR into centralised Business Intelligence and Data Analytics (Belizón & Kieran 2022, p. 18). Scholars advocating Strategic HRM for generative AI likewise warn that "sourcing HR with the requisite skillset to effectively harness GAI poses another considerable challenge" and question "whether HR managers will have the resources and autonomy to make such a difference" (Chowdhury, Budhwar & Wood 2024, pp. 6–7). Combined with the directional readiness

gaps documented across corporate benchmarks in Chapter 2.3 (McKinsey & Company 2026, pp. 9, 33–36), the Technology-led objection lands hard: handing enterprise AI accountability to the function with the weakest historical data legitimacy looks indefensible on its face.

My reply: I concede two parts of this objection outright. First, placing work-system AI accountability inside an unreformed, siloed, administrative HR department will fail; both Aghion and Tirole's formal model and Belizón and Kieran's three-year study establish that cleanly. Second, core enterprise data infrastructure, cybersecurity, model architecture, and platform integration belong under Technology's authority, hosting governed enterprise data free of isolated legacy HR silos (Belizón & Kieran 2022, p. 17; Weill & Ross 2004, pp. 59–60).

Where the Technology-led case overreaches is in treating incumbent HR's current maturity deficit as proof that Technology should own work redesign, remuneration architecture, and employment compliance. Even the most people-oriented modern CIO holds no formal authority over job grading, enterprise bargaining, Works Council codetermination, industrial relations, or employment-law compliance. That is why Weill and Ross (2004, pp. 1–2) and Tangi, Rodriguez Müller and Janssen's twelve IT directors (2025, p. 7) explicitly reject placing primary organisational-change accountability inside IT.

Even more revealingly, Belizón and Kieran's closing finding shows the real-world cost of bypassing the people function. In their third round of interviews with business-unit managing directors and agile transformation leads, the authors found operational managers already using non-conventional people data (velocity, output quantity, and quality metrics pulled directly from software platforms) to evaluate, train, and discipline staff "without input from, or even the knowledge of, the HR function", warning that this shadow practice "presents a threat to the very role of HR within organisations and should not be over-looked" (Belizón & Kieran 2022, p. 21). Running unmonitored platform-generated employee metrics outside people-function governance creates immediate exposure under EU AI Act Annex III point 4(b) and Article 26(7), German *BetrVG* § 87(1) No. 6, California Civil Rights Council regulations, Chile Law 21.719, and NSW workplace safety law.

A critic will press one follow-up here, charging that requiring the people function to embed builder-literate people technologists and integrate into the enterprise data platform moves the goalposts and disguises a Technology-HR hybrid under an HR label. I reject that charge. Upgrading a function's informational capacity (converting formal authority into real authority in Aghion and Tirole's sense) preserves its governing institutional identity. High-reliability enterprises routinely embed data engineers and telemetry specialists inside horizontal governance functions such as Safety and Clinical Governance without turning them into IT departments. Anticipating the obvious counter-reply (that in high-reliability industries safety is a line accountability while the safety function acts as second-line assurance, which would relegate the people function to second-line advice) brings us straight back to the framing asymmetry established in Chapter 3.3. Physical safety operates primarily as an enforceable constraint checked against operational workflows; work design, job grading, and human-agent capacity are *constitutive of the work-system frame itself*. Two structural invariants keep an upgraded people-function locus distinct from a Technology-led model: its **governing objective function** (*phronesis* over human-agent capability, non-domination, and lawful dignity of work) and its **default burden of proof** (verifying work-system soundness before scaling).

4.2 Objection 2: Appointing a Dedicated Chief AI Officer

The second rival abandons both the people function and the CIO in favour of role creation: appointing a dedicated Chief AI Officer (CAIO) inside the Top Management Team. Marc Schmitt's (2026) unrefereed University of Oxford working paper supplies the strongest conceptual and empirical case for this model. Schmitt argues that earlier enterprise technologies (ERP, internet, cloud, mobile) were deterministic instruments of execution posing an *integration problem* that existing CIOs or CDOs could absorb (2026, pp. 9–10). AI layers three novel properties on top (distributed accountability for judgement, upstream constitutive governance over training data and optimisation objectives, and continuous post-deployment non-stationarity), creating a distinct *judgement-governance problem* (Schmitt 2026, pp. 10–11). Extending any incumbent C-suite mandate, Schmitt warns, triggers three failure modes: *governance lag*, *capability dilution*, and *legitimacy ambiguity* (pp. 7–8). Because local business units capture AI's upside while failure costs fall on the enterprise as a whole, Schmitt argues a dedicated CAIO is needed to internalise those accountability externalities (p. 12). His archival BoardEx/WRDS census shows the role is no passing fad: by August 2026, 258 organisations held a chief-level AI executive (spanning 127 dedicated CAIOs and 134 hybrid titles across dual-role and multi-spell observations), with a Kaplan-Meier median tenure of 37 months and 74% still in role at 24 months (Schmitt 2026, pp. 19, 22).

Historical research on Chief Digital Officers reinforces the coordination rationale. Analysing a panel of 913 S&P 500 and MSCI Europe companies across 7,318 firm-years (2010 to 2018), Sebastian Firk et al. (2021) found that transformation urgency (+29% probability from new digital entrants, $p < 0.001$) and coordination needs (+15% from firm diversification, +16% from lagging national digital infrastructure, $p < 0.05$) predicted centralised CDO appointments (pp. 18–19). Neither CIO presence ($p = 0.517$) nor CTO presence ($p = 0.371$) substituted for that trans-functional coordination need (Firk et al. 2021, pp. 18, 25).

My reply: Dialectical fairness requires avoiding an asymmetrical comparison. I cannot evaluate the people function in its upgraded, builder-literate form while evaluating Chief AI Officers solely by their under-endowed empirical average. I therefore test the CAIO at **both** its modal empirical baseline and its upgraded, board-endowed ideal.

At the **modal empirical baseline**, Schmitt's and Firk et al.'s own data expose severe structural fragility. Schmitt documents a stark **endowment deficit**: "Across the entire history of Execucomp, only nine named-executive-officer titles reference AI, and only 10% of dedicated-CAIO spells in BoardEx carry board-level seniority" (Schmitt 2026, p. 20). By Schmitt's own Proposition 3 (p. 16), a CAIO lacking board seat, budget, and firm-wide decision rights cannot internalise accountability externalities and instead "may worsen jurisdictional ambiguity by adding another partial claimant to contested territory." Schmitt also exposes a **governance rhetoric deficit**: across 141 appointment announcements, Value Creation (83%), Cross-Organisational Integration (72%), and Strategy (70%) dominate, while "genuine governance-and-risk language, responsible AI, ethics, oversight, compliance, appears in fewer than one in four announcements (23%)" (Schmitt 2026, pp. 23–24). Meanwhile, Firk et al. (2021) document classic **mimetic isomorphism** (DiMaggio & Powell 1983, pp. 147–160) driving CDO adoption: industry peer appointments strongly predict firm adoption (coefficient 3.223, $p = 0.002$, Firk et al. 2021, pp. 18, 26), with no established causal link to firm performance (p. 26) alongside heightened top-team complexity and legitimacy struggles (pp. 2, 5).

Now test the **upgraded, board-endowed ideal**: suppose a board appoints a fully empowered Chief AI and Transformation Officer seated on the Executive Committee with dedicated capital and explicit risk authority, comparing that ideal symmetrically against an upgraded, builder-literate people function. Even a board-endowed CAIO hits an immovable structural wall: the CAIO still holds **no institutional ownership over employment contracts, remuneration architecture, job grading, enterprise bargaining, Works Council codetermination, or statutory employer consultation**. To exercise those levers, the upgraded CAIO must either coordinate continuously through the people function (recreating Option C's reactive handoff from Chapter 3.3 while adding another C-suite seat) or strip employment levers out of the people function (recreating Option B's capability dilution and assurance collapse). Searching for a new collateral "AI office" sitting alongside HR, Technology, and Operations repeats Ryle's Oxford-visitor category-mistake (Ryle 1949, p. 26). Tellingly, Schmitt's 34-page C-suite census omits the CHRO completely while using "ranking a job candidate" (p. 10) and building global workforce AI fluency (p. 25) as flagship examples. A CAIO or Chief Data and AI Officer can strengthen the Technology pole of the partnership, yet even at full board endowment the role never replaces the people function as the permanent horizontal locus of work-system framing.

4.3 Objection 3: Structural Fusion (Moderna and AXA Germany)

The third rival proposes dissolving the HR-Technology boundary by merging both functions under one executive. Recent executive commentary frequently pairs Moderna's 2025 HR-IT merger with AXA Germany's "People and Data" board portfolio as twin blueprints for an AI-first enterprise (see McKinsey & Company 2026, p. 32). When thousands of custom GPTs operate beside human teams, proponents of structural fusion contend that consolidating HR and IT eliminates slow handoffs and unifies budget in a single stroke.

My reply: Primary reporting immediately separates that neat pairing, and classical contingency theory explains why full fusion breaks down at industrial scale.

In the *Wall Street Journal*, Isabelle Bousquette (2025) documented the real context behind Moderna's late-2024 decision to elevate CHRO Tracey Franklin to Chief People and Digital Technology Officer. Moderna had expanded rapidly during the pandemic from roughly 830 employees in March 2020 to more than 5,000 by May 2025, deploying over 3,000 tailored internal GPTs (including chained HR GPT workflows converting junior HR analyst roles into automated routing, Bousquette 2025, pp. 2–3). Three material facts rarely mentioned in advisory summaries shaped that merger: Franklin had first spent a full year partnering with then-CIO Brad Miller (who led Moderna's 2023 OpenAI alliance); Miller departed Moderna in February 2025; and in the same period Moderna cut 10% of its digital technology staff (about 50 jobs) and reported a $971 million first-quarter loss in May 2025 as vaccine demand fell (Bousquette 2025, pp. 2–3). Cost compression and an executive vacancy converged with AI ambition inside a compact, single-sector, 5,000-person biotech.

AXA Germany made a markedly different structural choice, and reading it cleanly under our fixed referent (*the enterprise people function*, established in Chapter 1.1) illuminates my thesis directly. Effective 1 January 2026, AXA Germany kept core IT infrastructure separate while expanding its executive board People portfolio into "People and Data Experience", appointing Dr Johannes Dick (who succeeded departing CHRO Sirka Laudon on 1 November 2025) to unite HR, Data, and AI at

board level (AXA Konzern AG 2025). Dr Dick is a career Strategy, Data, Organisation, and commercial-insurance executive, never a traditional career HR specialist. Viewed through Andrew Abbott's (1988) ecology of professions, AXA Germany illustrates two points simultaneously: where traditional HR practitioners fail to master data and AI, boards import technical leaders into the executive seat, yet **the institutional locus of accountability remains inside the People-and-Data board function**, kept distinct from core enterprise IT under German *BetrVG* codetermination realities. Neither Moderna nor AXA Germany yet carries independent longitudinal outcome validation.

For a large, multi-jurisdictional enterprise of 50,000 or more employees running legacy ERP, operational technology, unionised workforces, and multi-country safety regimes, Paul Lawrence and Jay Lorsch's (1967) foundational theory of **differentiation and integration** explains why full HR-IT structural fusion fails. Highly uncertain, complex environments require deep internal *differentiation* across specialist sub-units paired with formal *integrating mechanisms*, never the collapse of incompatible crafts under one manager. Full fusion introduces two severe failure modes:

1. **Capability dilution across incompatible masteries:** As independent CIO Christopher Desautel told the *Wall Street Journal*, formal HR-IT mergers remain rare "in part because they require unique skill-sets" (Bousquette 2025, p. 2). Cybersecurity, network resilience, ERP technical debt, and enterprise architecture demand deep engineering mastery; industrial relations, enterprise bargaining, remuneration design, and employment law demand deep people-governance mastery. Combining both under one executive at industrial scale risks sub-scaling whichever half sits outside the incumbent's career domain.
2. **Collapse of independent assurance:** Placing the executive who builds and scales enterprise AI platforms in charge of employment-risk and worker-privacy gates violates a foundational governance principle: gate-criteria authors must sit outside the reporting line of the proponent of the gated work. Intellectual honesty requires applying that same principle reflexively to my own thesis. A people-function-framed transformation would suffer the same proponent-gatekeeper conflict if the executive championing workforce productivity targets also chaired post-deployment risk assurance. Chapter 5 resolves that tension structurally through the Three Lines Model.

4.4 Objection 4: Operations-Led Governance, CFO Capital Hurdles, Fissured Workforces, and Contingency

The fourth rival contends that both the people function and Technology are corporate enabling functions divorced from frontline profit-and-loss accountability, that enterprise AI governance sits properly with Operations and Finance, or that no single locus fits every firm. Weill and Ross's empirical leverage index across 256 enterprises supplies strong ammunition: indexing CIO involvement at 1.0, the Chief Executive Officer delivers 2.1 times the governance leverage, the Chief Operating Officer delivers 1.7, and Business Unit Leaders deliver 1.6 (Weill & Ross 2004, pp. 125–126; no CHRO appears in the 2004 table). Altman et al. (2023, ch. 4) likewise place senior leaders and business-unit heads on the vertical central axis of workforce ecosystem orchestration. Operations-led governance also need not mean chaotic business-unit decentralisation: mature industrial enterprises operate **centralised COO / Enterprise Operating System models** (grounded in Lean and Toyota Production System standardisation) that unify frontline workflows across di-

visions. Two commercial realities intensify this objection: **CFO capital-allocation governance** (Investment Committees routinely underwrite AI business cases on committed FTE cost reductions, overriding unaligned HR charters) and the **fissured enterprise** (Weil 2014; Altman et al. 2023, ch. 6), where substantial shares of technology, contact-centre, and back-office work are executed by outsourced business-process outsourcing (BPO) vendors and managed-service contractors outside the firm's direct employment payroll. Contingency scholars reinforce the challenge: Henderson and Venkatraman (1999, p. 482) deny that any single alignment mode is universally superior, while Kaufman's critique cautions against universalistic HRM formulas (cited in Wright & Ulrich 2017, pp. 50, 55).

My reply: I welcome every element of this challenge because it clarifies scope boundaries and institutional realism. First, CEO chartering is indispensable, Operations must own frontline workflow execution and value realisation, and **external customer-product AI strategy** (clinical diagnostics, credit underwriting, dynamic pricing) belongs under CEO, Strategy, and Commercial Product leadership. I also accept explicit contingency boundaries: in a 200-person startup or a pure ITOT cloud migration leaving core roles untouched, founder-led or Technology-led execution works cleanly.

Inside a large, diversified, regulated enterprise deploying **work-system AI**, however, vesting primary framing in decentralised business units triggers *Feudal* fragmentation (Weill & Ross 2004, pp. 60–61). Belizón and Kieran (2022, p. 19) observed that degradation empirically when Alpha decentralised data and analytics into business units: unit managing directors admitted they lost the cross-enterprise diagnostic capability a centralised team had provided. Even under a **centralised COO / Enterprise Operating System**, the COO remains accountable primarily for operational throughput and quarterly margin. Under quarterly P&L pressure, expecting individual line managers to protect multi-year apprenticeship pipelines against short-term headcount-reduction incentives asks individuals to bear personal career penalties against institutional reward structures, predictably breeding negative employee motive attribution (Nishii, Lepak & Schneider 2008), covert algorithmic gaming (Parker & Grote 2022, p. 1218), and unmonitored employment-law exposure (Belizón & Kieran 2022, p. 21). Chapter 5 therefore binds the **CFO** directly into enterprise framing and pre-deployment capital gates (Domains 1 and 3) so Investment Committee hurdles evaluate redesigned capacity alongside efficiency.

David Lepak and Scott Snell's (1999, pp. 31–48) **Human Resource Architecture** provides a clear conceptual frame to resolve the **fissured-workforce paradox** (Weil 2014; Altman et al. 2023, ch. 6). Lepak and Snell distinguish core, firm-specific internal human capital from contracted and alliance-based labour, showing that distinct employment modes demand distinct governance instruments. When managed-service contractors or BPO personnel perform major enterprise workflows, the people function must never issue direct task-level work charters governing vendor employees, which would trigger statutory co-employment or joint-employer liability. Instead, the people function co-authors pre-deployment **vendor-procurement and data-boundary clauses** (formalised under Domain 3 in Chapter 5) alongside Procurement and Legal, obliging managed-service vendors deploying agentic AI on enterprise data to disclose algorithmic evaluation practices, prohibit unapproved worker surveillance on enterprise systems, and align service-level metrics with human-oversight and safety standards.

4.5 Objection 5: Purely Operational AI, Desktop Productivity Tools, and Bottlenecks

A fifth objection attacks people-function accountability from the factory floor and the daily desktop. Most enterprise AI deployments have little to do with employment policy. They consist either of physical-process and operational models (computer vision identifying oversized rocks on a conveyor belt for separation, automated network-incident log triage, cyberthreat code scanning, or Procter & Gamble's Enterprise Data Science and CIO-run "AI Factory" standardising MLOps pipelines for supply-chain, media, and retail forecasting models, Bojinov, Lakhani & Lefort 2025, pp. 1, 5–8) or of everyday desktop productivity assistants embedded inside enterprise suites such as Microsoft Office, P&G's internal **chatPG** (p. 12), or JPMorganChase's **LLM Suite**. Inserting people-function approval gates into ore-sorting sensors, model factories, or office software would throttle engineering speed, irritate frontline operators, and create a bureaucratic choke point.

My reply: This objection performs an essential service by forcing an operational scope boundary. Its force splits cleanly across two distinct classes of AI that executives routinely conflate.

On **Class A: purely operational and physical-process AI**, I concede the objection outright. Requiring a Work and Capacity Charter before a reliability engineer deploys computer vision to divert oversized rock on a conveyor belt, before JPMorganChase runs an automated cyberthreat model across four million annual code changes (Bojinov, Lakhani & Lane 2025, p. 9), or before Procter & Gamble's AI Factory automates MLOps deployment for media and supply-chain forecasting (Bojinov, Lakhani & Lefort 2025, pp. 5–8) would be pure bureaucratic drag. Those systems sit inside operational technology and ITOT. They belong entirely under Technology and Operations governance with **no HR gate**, subject to one clear tripwire: the people function and Safety enter if and only if the operational system alters worker workload intensity, cognitive strain, or work pace (triggering workplace health and safety duties such as Parliament of New South Wales 2026), logs operator latency for performance scoring, or eliminates a safety-critical monitoring role without redesigning exception handling.

On **Class B: enterprise-wide desktop productivity suites** (such as Microsoft 365 Copilot, **chatPG**, or JPMorganChase's **LLM Suite** across 200,000 desks), the "just a productivity tool" dismissal is a dangerous Trojan horse. Treating ubiquitous generative AI as an updated spellchecker confuses micro-level convenience with macro-level neutrality. Three systemic workforce effects appear rapidly at enterprise scale:

1. **Apprenticeship and capability pipeline erosion:** Removing routine preparatory work hollows out the foundational learning rungs through which junior staff historically built domain judgement, leaving multi-year mentorship pipelines unredesigned (as documented in Chapter 3.1's JPMorganChase Private Bank case, Bojinov, Lakhani & Lane 2025, pp. 10, 14).
2. **Shadow performance telemetry and automated evaluation:** Enterprise productivity suites generate granular behavioural logs by default. Without upstream people-function boundaries, adoption logs rapidly migrate into group leaderboards, appraisal tie-ins, and automated call-rating bots (Bojinov, Lakhani & Lane 2025, pp. 5–6, 19; Belizón & Kieran 2022, p. 21; Kellogg, Valentine & Christin 2020), triggering binding employment and codetermination obligations.
3. **Unowned capacity redeployment:** Freeing 15% to 20% of drafting and research time across tens of thousands of knowledge workers either dissipates into low-value workslop or releases aggregate labour capacity requiring an institutional decision on growth redeployment,

reskilling, or attrition-managed transition (Bojinov, Lakhani & Lane 2025, p. 15). Technology can provision licences; it holds no authority over capacity allocation or job architecture.

The operational resolution eliminates the bottleneck without surrendering governance. The people function never polices individual prompts or slows safe self-service sandbox access once Technology sets security controls; JPMorganChase's democratised **LLM Suite** and P&G's **chatPG** show how effectively broad sandbox access drives bottom-up experimentation (Bojinov, Lakhani & Lane 2025, pp. 6, 9–10; Bojinov, Lakhani & Lefort 2025, p. 12). Instead, the people function sets the **macro systemic rules** surrounding productivity AI (embedded surveillance prohibitions, redesigned apprenticeship pathways, capacity redeployment, and SaaS agentic capability-drift triggers) while staying completely out of pure Class A operational engineering. Chapter 5 formalises that boundary through an independently administered three-tier Workforce Materiality Triage gate.

CHAPTER 5

The Architecture of Binding Partnership, Three-Lines Assurance, and Earned Authority

Practical wisdom describes how leaders reason under uncertainty; it supplies neither an organigram nor a veto. Calling for "people-function framing in partnership with Technology and Operations" collapses into an empty corporate slogan unless I specify concrete decision rights, Three-Lines assurance boundaries, independent triage authority, veto-player escalation clocks, operational and empirical tests, and readiness entry criteria.

5.1 Legal-Rational Authority, Specific Knowledge, and Escaping "Rule by Nobody"

In *Economy and Society*, Max Weber (1978 [1922]) distinguished charismatic and traditional rule from **legal-rational authority**, in which institutional legitimacy derives from explicit, codified charters and assignable decision rights. Inside a complex enterprise, board-chartered decision rights provide Weberian legal-rational authority that prevents two symmetrical failures I have observed across enterprise AI governance: quiet technical coercion (when Technology ships locked interfaces without redesigning work, breeding shadow workarounds) and powerless persuasion (when unchartered HR advisers circulate slide decks while business units deploy unvetted tools anyway, Boudreau & Cascio 2017, cited in Belizón & Kieran 2022, p. 5). Just as importantly, assigning explicit framing ownership to a single executive function defeats Hannah Arendt's bureaucratic **"rule by Nobody"** (Arendt 1970, pp. 38–39) and Dennis Thompson's (1980) problem of many hands.

Organisational economics supplies the matching design rule for allocating those rights across specialist functions. Michael Jensen and William Meckling (1992, pp. 251–274) showed that efficient governance **collocates decision rights with specific knowledge** (because transferring tacit domain knowledge across silos is costly) while structurally separating **decision management** (initiation and implementation) from **decision control** (ratification and monitoring). Chartered decision rights cannot manufacture personal virtue on their own; they create the institutional space within which Shannon Vallor's *technomoral wisdom* (2016, ch. 6, sec. 6.13) can operate without requiring individuals to resist commercial incentives unarmed.

To make this constitution auditable against international governance standards, I map the six decision domains below explicitly onto the **Institute of Internal Auditors (IIA) Three Lines Model (2020)**, aligned with **ISO/IEC 42001:2023** (Artificial Intelligence Management Systems) and the **NIST AI Risk Management Framework (AI RMF 1.0)** (NIST 2023):

- **First Line (Framing and Delivery Management):** The People Function (work-system framing, capacity, incentives), Technology (architecture, data platform, cyber), Operations (frontline execution), and Finance (capital hurdles).
- **Second Line (Independent Management Oversight and Assurance):** General Counsel, Chief Risk Officer, Data Protection Officer, and Independent Information Risk Assurance, sitting outside both the CIO and CHRO delivery lines, authoring gate criteria, **administering the**

Materiality Triage Registry, holding Domain 3 stop-work authority, and **chairing Domain 6 post-deployment assurance**.

- **Third Line (Board-Independent Assurance): Internal Audit**, reporting functionally to the **Board Audit and Risk Committee** (independent of the CEO, who sponsors enterprise transformation), periodically auditing both First-Line compliance and Second-Line triage and gate integrity.

Adapting Weill and Ross's Governance Arrangements Matrix (2004, pp. 11–12, 59–60; note that "People-Led Federal" and "People-Led Duopoly" below explicitly adapt Weill and Ross's archetype terminology to work-system AI) and Sebastian et al.'s (2026, pp. 1–3) *Frame* / *Act* / *Learn* decision model, alongside Baiyere, Salmela and Tapanainen's (2020, p. 37) triad of *mindful actors*, *infrastructural flexibility*, and *light-touch processes*, I structure work-system AI across six non-overlapping decision domains.

5.2 Six Decision Domains of the Binding Partnership

Domain 1: Work-System AI Principles, Scope, and Capital Alignment (*Frame*, First Line)

- **Adapted archetype:** People-Led Federal (under CEO charter; external customer-product AI strategy stays under CEO, Corporate Strategy, and Commercial Product leadership).
- **Accountable owner:** Head of the People Function (CHRO / Chief People and Data Officer).
- **Hard veto:** Chief Executive Officer and Executive Committee (strategic alignment and enterprise capital envelope).
- **Binding partners:** Chief Financial Officer (aligning capital-hurdle metrics with redesigned capacity and capability preservation), CIO/CTO (technical feasibility), COO and Business Unit Leaders (value pools), General Counsel and Chief Risk Officer.
- **What this governs:** Enterprise principles for human-agent capacity design, classes of work remaining human-owned, capital business-case rules preventing unexamined headcount-only targets, and internal workforce identity evolution (Wessel et al. 2021, p. 41).

Domain 2: Work Redesign, Role Architecture, Capacity Allocation, and Incentives (*Frame and Learn*, First Line)

- **Adapted archetype:** People-Led Duopoly (People Function and Operations), subject to statutory worker consultation or codetermination (*BetrVG* § 87(1) No. 6; EU AI Act Article 26(7); *Fair Work Act 2009* s. 205).
- **Accountable owner:** Head of the People Function.
- **Hard veto:** Business Unit P&L Leader (**scoped strictly to operational continuity and physical/workplace safety**; commercial margin disputes over apprenticeship or capability design escalate to CEO/ExCo under a binding **15-business-day deadlock-breaking SLA**).
- **Binding partners:** Frontline operations managers, elected Works Councils, employee and union representatives (operationalising Trist and Bamforth's autonomous worker-voice lesson), HSE officers, People Analytics.

- **What this governs:** Task-level allocation between humans and agents, redesigned job grades, apprenticeship preservation, AI literacy programmes (EU AI Act Article 4), and remuneration structures rewarding senior experts for original judgement when algorithms offer shortcuts (Brynjolfsson, Li & Raymond 2025, p. 915; Parker & Grote 2022, p. 1220).

Domain 3: AI Procurement, Vendor Clauses, and Pre-Deployment Assurance Gate (*Frame*, First-Line Framers and Second-Line Gate)

- **Adapted archetype:** Dual-Specification Partnership with Second-Line Assurance Gate (audited by Third-Line Internal Audit).
- **Accountable framers (First Line):** Head of the People Function (work-design and employment-impact specification) and CIO (architecture and technical security specification).
- **Hard veto (Second Line): Independent Risk and Legal Assurance** (General Counsel, Chief Risk Officer, and independent Information Risk Assurance reporting outside both the CIO's and CHRO's lines).
- **Binding partners:** Chief Financial Officer (capital-case verification), Enterprise Procurement (embedding algorithmic-oversight, anti-surveillance, and non-co-employment clauses for BPO and managed-service vendors, per Lepak & Snell 1999 and Weil 2014), Data Protection Officer.
- **What this governs:** Embedding Parker and Grote's (2022, p. 1220) work-design criteria and *sociotechnical envelopment* into vendor selection and build gates before capital locks. Keeping CISO technical-security execution inside Technology while vesting final Domain 3 veto authority in Second-Line officers outside both delivery lines eliminates self-review: neither a CIO chasing deployment velocity nor a CHRO sponsoring workforce productivity targets marks their own homework. Contested Second-Line holds escalate to an Executive Risk Committee review within 15 business days.

Domain 4: Enterprise Architecture, Unified Data Platform, Cybersecurity, and Model Safety (*Act*, First Line)

- **Adapted archetype:** IT Monarchy / Technical Authority (Weill & Ross 2004, pp. 59–60).
- **Accountable owner:** Chief Information Officer / Chief Technology Officer / CISO.
- **Hard veto:** CISO and Chief Enterprise Architect (unconditional technical and cyber stop-work authority).
- **Binding partners:** People Technology Strategy (functional and semantic requirements), Business Unit IT.
- **What this governs:** Cloud infrastructure, identity and permission boundaries, API security, model evaluation, prompt-injection defence, metered cost telemetry, SaaS release-management drift controls, and operating the unified enterprise data platform hosting the people function's logically isolated People Data domain (Belizón & Kieran 2022, p. 17; Baiyere, Salmela & Tapanainen 2020, p. 41).

Domain 5: Frontline Workflow Execution and Business Value Realisation (*Act and Learn*, First Line)

- **Adapted archetype:** Operations-Led Execution (within chartered People and Technology guardrails).

- **Accountable owner:** Chief Operating Officer and Business Unit P&L Leaders.
- **Hard veto:** Head of the People Function (**stop-use trigger for unapproved shadow people-metrics, unlawful automated evaluation, or material breach of the signed Domain 2 Work and Capacity Charter**, per Belizón & Kieran 2022, p. 21) and CISO (cyber breach).
- **Binding partners:** Embedded HR Business Partners, agile AI delivery squads, frontline mindful actors.
- **What this governs:** Converting redesigned human-agent workflows into customer outcomes, throughput, and safety performance through Chowdhury et al.'s (2023, p. 25) hub-and-spoke ecosystem while blocking strictly lawful yet sociotechnically destructive *so-so automation* that violates the signed Work and Capacity Charter.

Domain 6: Continuous Post-Deployment Assurance and Algorithmic Sunset (*Learn*, Second Line, Audited by Third Line)

- **Adapted archetype:** Second-Line Independent Assurance (chaired by Enterprise Risk and Legal, **never co-chaired by HR or IT**, audited annually by Third-Line Internal Audit reporting to the Board Audit and Risk Committee).
- **Accountable owner:** Chief Risk Officer and General Counsel (Enterprise AI Assurance Committee; the People Function and Technology supply workforce-impact and drift evidence as audited First-Line inputs).
- **Hard veto:** General Counsel and Chief Risk Officer (statutory stop-use and model-decommissioning authority), with escalation to the Board Audit and Risk Committee where CEO sponsorship conflicts with Second-Line findings.
- **What this governs:** Mandatory control-framework registration before production (echoing JPMorganChase's control standard, Bojinov, Lakhani & Lane 2025, p. 4) and continuous auditing of Orlikowski's (2000) enacted *technologies-in-practice* in line with ISO/IEC 42001:2023 and NIST AI RMF 1.0 (Measure/Manage). Covers comparative output audits against drift and bias (Tangi, Rodriguez Müller & Janssen 2025, p. 8), EU AI Act Annex III and Article 26(7) records, German *BetrVG* works agreements, California four-year ADS logs, Chile Law 21.719 impact assessments, NSW digital-work WHS reviews, and Australian Privacy Act APP 1.7 to 1.9 inventories.

EXHIBIT 1

Enterprise Work-System AI Decision-Rights Constitution (Three Lines Mapped)

Domain	IIA Line & Archetype	Accountable Owner	Hard Veto Holder
1. Principles, Scope & Capital (*Frame*)	1st Line · People-Led Federal	People Function Head (+ CFO alignment)	CEO / Executive Committee
2. Work, Capacity & Incentives (*Frame*/*Learn*)	1st Line · People-Led Duopoly	People Function Head	BU P&L Leader (safety/continuity only; 15-day ExCo SLA)
3. Procurement & Assurance Gate (*Frame*)	1st Line Framers + 2nd Line Gate	People Function Head + CIO (Framers)	2nd Line Assurance (GC / CRO outside CIO & CHRO lines)
4. Architecture, Data & Cyber (*Act*)	1st Line · IT Monarchy	CIO / CTO / CISO	CISO & Chief Enterprise Architect
5. Frontline Execution & Value (*Act*/*Learn*)	1st Line · Operations-Led Execution	COO / BU Leaders	People Function Head (unlawful ADM / charter breach) & CISO
6. Post-Deployment Assurance & Sunset (*Learn*)	2nd Line (3rd Line Board-Audited)	CRO & General Counsel (HR/IT supply evidence)	General Counsel / CRO (Board Audit Committee escalation)

Source: Author's synthesis adapting Weill & Ross (2004, pp. 11–12, 59–60), MIT CISR AI Decision Matrix (Sebastian et al. 2026, pp. 1–3), IIA Three Lines Model (2020), ISO/IEC 42001:2023, and NIST AI RMF 1.0 (2023).

A clear-eyed governance designer must confront George Tsebelis's (2002) **veto-player theory**: multiplying actors with veto authority raises institutional status-quo bias and slows adaptation. This matrix deliberately limits veto friction through two mechanisms: first, the **Workforce Materiality Triage** below keeps high-volume operational and desktop productivity tools completely outside Domain 3's multi-veto path; second, binding **15-business-day escalation SLAs** force contested Tier 3 holds to rapid executive resolution rather than quiet committee burial. Because SLA effectiveness is an empirical claim rather than an axiom, Third-Line Internal Audit tracks cycle times and shadow-AI bypass rates annually for the Board Audit and Risk Committee.

5.3 Workforce Materiality Triage: Independent Classification Preventing Bottlenecks and Tier-Shopping

To operationalise Chapter 4.5 without inviting Technology tier-shopping or HR bottlenecking, **tier classification authority sits strictly with Second-Line Enterprise Risk and Legal via the Enterprise AI Intake Registry** (backed by periodic Third-Line Internal Audit sampling), never with the proposing IT squad or HR alone (adapting tiered risk-control intake such as JPMorganChase's differentiated risk pathway, Bojinov, Lakhani & Lane 2025, pp. 3–4):

EXHIBIT 2

Three-Tier Workforce Materiality Triage Gate

Materiality Tier	Examples	Governance Path	People-Function Role & Drift Tripwires
Tier 1: Pure Operational / OT AI	Rock-sizing conveyor vision; network-log triage; cyberthreat code scanners; P&G supply-chain MLOps	100% Technology & Operations under engineering and safety controls	**No HR gate** (lightweight 2nd-Line registry check confirming no workload-intensity/cognitive WHS strain, operator scoring, or safety-role elimination tripwire)
Tier 2: Broad Desktop Productivity AI	Microsoft 365 Copilot; internal **LLM Suite** or **chatPG** sandboxes; drafting assistants	Fast Technology self-service access inside security guardrails	**Macro rules only** (ban unapproved usage telemetry/call-scoring; govern apprenticeship redesign and capacity redeployment; **automated Vendor Release-Drift Trigger** requires 2nd-Line triage before SaaS autonomous agent features activate)
Tier 3: Work-System & Employment AI	Agentic workflow orchestration; hiring, promotion, pay, task-allocation, or performance-evaluation tools	Full Binding Partnership (Domains 1 to 6) + 2nd-Line Assurance Gate	**Primary framing locus** (upstream Work & Capacity Charter, incentive redesign, statutory conformity, post-deployment audits)

Source: Author's governance triage specification separating pure operational engineering (Tier 1) and broad desktop productivity tools (Tier 2) from work-system AI (Tier 3), administered independently by Second-Line Risk and Privacy under a 48-hour classification SLA.

5.4 Operational Classification Tests and Rebuilt Empirical Disconfirmation Standard

Two observable **operational classification tests** distinguish this people-function framing model from a generic committee or boilerplate governance theatre:

1. **The *Frame* pen test (before vendor selection and capital lock).** In a people-function-framed enterprise, no Tier 3 work-system AI programme clears Domain 3 capital approval until the people function holds the primary drafting pen on a co-signed **Work and Capacity Charter** (co-signed by Operations and Finance) specifying human-agent task splits, apprenticeship and senior-role evolution, funded reskilling pathways, and applicable statutory obligations. That single pen-holder rule also marks the concrete operational difference between Option E and Option D (CEO/ExCo collective co-framing) when a charter is contested.
2. **The dual-closure test (respecting accounting boundaries without stranding work redesign).** Financial accounting standards such as IAS 38 and ASC 350-40 govern when software capitalisation ceases and asset amortisation begins (once an asset is technically ready for intended use). While those standards never legally forbid organisations from tracking downstream workforce outcomes, corporate budgeting habits routinely treat capitalisation cut-off as project closure. Separating **Technical Capital Release** (owned by Technology and Finance at technical go-live) from **Transformation Benefit Closure** (owned by the People Function and Operations, closing only once redesigned roles, mindful-actor capability, aligned incentives, and baseline Second-Line bias audits verify in live operations) respects accounting rules without stranding work redesign.

Evaluating my thesis scientifically requires separating its **instrumental empirical claim** from its **deontic normative claim**. The normative claim (that workers merit protection against arbitrary algorithmic domination and preserve human dignity, Pettit 1997; Anderson 2017) is a philosophical commitment defended by argument, never disproved by a quarterly earnings report. The **instrumental empirical hypothesis** (that vesting work-system AI framing in the enterprise people function outperforms rival governance models on sustained value, capability retention, trust, and compliance) *is* empirically testable, provided we avoid immunising it via a tautological "Precondition-compliant" escape hatch.

I therefore specify **ex-ante observable compliance indicators** and **Lakatosian disconfirming evidence**: * **Ex-ante observable indicators** (measured before outcomes are known, avoiding post-hoc excuses): (a) a named, funded internal people-technology engineering team; (b) a logically isolated, access-logged People Data domain on the enterprise platform; and (c) a live Second-Line automated-decision inventory. * **Comparative test design:** Across a 36-month longitudinal sample of large, regulated enterprises deploying Tier 3 work-system AI, compare *all* enterprises formally vesting Domain 1 work-system framing in the people function against Technology-led, COO-led, and ExCo co-framed (Option D) peers, modelling the three ex-ante maturity indicators as explicit statistical moderators and controlling carefully for baseline management-quality selection effects. * **Lakatosian disconfirming evidence (short of binary refutation):** My institutional hypothesis would be substantially weakened, requiring structural revision, if (1) enterprises using Option D (ExCo collective co-framing) or centralised COO governance paired with strong statutory Works Councils match or outperform people-function-framed peers on mid-career capability retention and employee trust at lower governance cycle times, or (2) formal people-function framing shows persistent multi-year failure to retain builder-literate talent even where boards fund competitive technical remuneration.

5.5 Four Preconditions: Converting Formal Authority into Real Authority

Because Aghion and Tirole's (1997) gap between formal and real authority is real, no Chief Executive Officer should vest Domain 1 framing in an unreformed HR department unconditionally. The people function earns and retains real authority by meeting four preconditions:

1. **Build an internal, builder-literate people-technology core.** The people function must employ and train technologists who build, test, and red-team AI workflows alongside software engineers, closing the informational gap between policy slides and live model behaviour (Chowdhury et al. 2023, p. 21; Parker & Grote 2022, pp. 1216, 1221).
2. **Replace fragmented HR silos with a logically isolated People Data domain on the unified enterprise platform.** Workforce, operational, and financial analytics must integrate on Technology's unified enterprise platform so human-agent capacity modelling operates on trusted company-wide data (Belizón & Kieran 2022, pp. 17–19). Retiring legacy HR silos must never expose sensitive personal records (medical accommodations, disciplinary investigations, union membership, pay-equity deliberations) to unrestricted enterprise inference. The architecture requires a **purpose-built People Data domain** inside the enterprise platform enforcing formal data classification, a canonical ontology and semantic layer, strict role-based permissioning (RBP), statutory purpose-limitation controls, enhanced PII security, zero-trust retrieval-augmented generation (RAG) permission boundaries, and immutable access logging.

3. **Institutionalise Three-Lines assurance outside every proponent's line.** Second-Line gate authorship and Domain 6 oversight sit outside both CIO and CHRO reporting lines, backed by Third-Line Internal Audit reporting directly to the Board Audit and Risk Committee.
4. **Enforce a time-bounded readiness window (12 to 18 months) with a binding Month-19 Board fallback.** Where an enterprise's current HR department sits in the bottom maturity tier (lacking agentic fluency, governed data integration, and multi-year workforce planning, McKinsey & Company 2026, pp. 9–10, 32–36), the CEO should institute a formal **12-to-18-month CHRO-CIO Co-Chair transition** (a deliberate, temporary Option D safeguard) governed by explicit six-month milestones. Because 18 months of platform selection can trigger Jonas's *automotive momentum* before the people function matures, Second-Line Risk assurance holds binding Domain 3 veto authority from Day 1. If the people function still misses its ex-ante readiness milestones at **Month 19**, co-chairmanship terminates automatically and triggers a **Board People and Audit Committee intervention**: either re-architecting people-function executive leadership (following AXA Germany's precedent of importing technical leadership into the People-and-Data seat) or maintaining CEO-chaired cross-functional governance under Second-Line Risk veto and Third-Line Board audit until every precondition passes verification.

Weill and Ross (2004, p. 13) found empirically that the single strongest predictor of enterprise governance performance is the percentage of leaders who can accurately describe their governance model (averaging only 38% across studied firms, rising past 45% in top performers, pp. 13, 124). Making the six decision domains, Three-Lines assurance boundaries, independently administered materiality triage, and four preconditions explicit turns people-function framing from boardroom aspiration into enforceable Weberian legal-rational structure.

CONCLUSION

Accountability as Institutional Design

Where an enterprise locates accountability for work-system AI is a moral and strategic choice written into its organisational chart. Boards frequently treat that decision as a question of executive convenience or departmental turf. Across this essay I have argued that technical architecture and cybersecurity operate as verifiable constraints downstream of a business problem, whereas human work design is constitutive of the problem frame itself. Locating framing accountability determines whether an organisation builds Acemoğlu and Johnson's *machine usefulness* or slides into *so-so automation*, and whether human work remains a site of prudential judgement and non-domination or becomes a residual adjusted to autonomous *technique* (Ellul 1964, ch. 1; Pettit 1997; Anderson 2017; Acemoğlu & Johnson 2023, intro.).

Defaulting work-system AI to Technology treats a fundamental reconfiguration of roles, judgement, apprenticeship pipelines, and statutory duties as a software rollout. Seventy-five years of empirical research, spanning Trist and Bamforth's (1951) mechanised coal pits, Zuboff's (1988, pp. 9–11) *automate* versus *informate* divide, Brynjolfsson, Li and Raymond's (2025) generative AI contact centres, Dell'Acqua et al.'s (2026) jagged frontier, JPMorganChase's unredesigned apprenticeship pipeline (Bojinov, Lakhani & Lane 2025), and Tangi, Rodriguez Müller and Janssen's (2025) European IT directors, shows the recurring cost of that category mistake: unredesigned work, eroded prudential judgement, covert gaming, and stranded value. Yet vesting accountability unconditionally in an unreformed HR department invites a failure equally severe, documented by Belizón and Kieran's (2022) three-year legitimacy study, formalised by Aghion and Tirole's (1997) gap between formal and real authority, warned against by Ellul's indictment of *human techniques* (1964, ch. 5) and Legge's Critical HRM scepticism (2005), and echoed across global readiness benchmarks. Neither autonomous technical determinism nor technically illiterate, management-captured HR administration deserves the trust of a board.

The disciplined institutional answer for large, complex, regulated enterprises is placing **formal framing accountability for work-system AI in the enterprise people function**, operating inside a binding partnership with Technology, Operations, and Finance, backed by Three-Lines assurance (IIA 2020; ISO 2023; NIST 2023) and an independently administered Workforce Materiality Triage that keeps the people function out of pure operational engineering. Technology holds authority over resilient architecture, data platforms, and cyber safety; Operations owns frontline execution; Finance aligns capital hurdles with redesigned capacity; the people function holds the horizontal framing pen protecting space for *phronesis* and *technomoral wisdom* (Aristotle 2009, Book VI, ch. 5, 1140b; Vallor 2016, ch. 6, sec. 6.13) across work design, capacity allocation, remuneration integrity, and employer compliance. Explicit Weberian legal-rational decision rights, paired with statutory worker consultation and codetermination, defeat Arendt's bureaucratic "rule by Nobody" (Arendt 1970, pp. 38–39) and Thompson's (1980) problem of many hands while placing upstream restraint where organisational hysteresis and technological momentum demand it (Jonas 1984, pp. 31–32).

I close with a practical intra-organisational fairness heuristic adapted from John Rawls's *A Theory of Justice*. Rawls formulated the *veil of ignorance* to test principles for society's basic structure, ensuring that "no one is advantaged or disadvantaged in the choice of principles by the outcome of natural chance or the contingency of social circumstances" (Rawls 1999 [1971], p. 11; see also

pp. 118–123). Adapted strictly to test our *general institutional principles* (leaving the specific functional assignment to the comparative institutional reasoning of Chapters 3 and 5), imagine designing your enterprise's AI governance principles without knowing which seat you will occupy once the veil lifts: the Chief Executive Officer answering to conflicting regulators across Europe, California, Chile, and Australia; the Chief Information Officer defending system uptime and zero-trust data security; the head of the people function facing a statutory discrimination or employee-data breach; the senior specialist whose tacit knowledge trains tomorrow's model while her incentive scheme rewards blind compliance; or the frontline worker or outsourced contractor scheduled and rated by an opaque algorithm they cannot question.

Behind that veil, rational participants would reject leaving work redesign inside a build-incentivised software hierarchy, reject entrusting safety-critical platforms to an unmodernised administrative department, and reject relying on any unchecked executive alone without worker voice, statutory guardrails, and board-independent assurance. They would demand the foundational principles this essay has defended: upstream human-work framing, non-domination, uncompromising engineering authority, and genuinely separated assurance. Putting those principles to work inside a complex enterprise requires naming one accountable institutional home for the work-system frame. That home is the enterprise people function.

BIBLIOGRAPHY

References

Formatted in Harvard author-date style. Page references throughout the text correspond directly to the published editions, Versions of Record, accepted manuscripts (disclosed explicitly where used), Bekker lines, HBS case pagination, or statutory sections listed below.

Abbott, A 1988, *The system of professions: an essay on the division of expert labor*, University of Chicago Press, Chicago.

Acemoğlu, D & Johnson, S 2023, *Power and progress: our thousand-year struggle over technology and prosperity*, PublicAffairs, New York. *(Reflowable EPUB edition; cited by Introduction and Chapter.)*

Aghion, P & Tirole, J 1997, 'Formal and real authority in organizations', *Journal of Political Economy*, vol. 105, no. 1, pp. 1–29, DOI:10.1086/262063.

Altman, EJ, Kiron, D, Schwartz, J & Jones, R 2023, *Workforce ecosystems: reaching strategic goals with people, partners, and technologies*, MIT Press, Cambridge, MA. *(Reflowable EPUB edition; cited by Chapter.)*

Anderson, E 2017, *Private government: how employers rule our lives (and why we don't talk about it)*, Princeton University Press, Princeton, NJ.

Arendt, H 1970, *On violence*, Harcourt, Brace & World, New York.

Arendt, H 2013 [1958], *The human condition*, 2nd edn, University of Chicago Press, Chicago. *(Reflowable EPUB edition; cited by Part and Section.)*

Aristotle 2009, *The Nicomachean ethics*, trans. WD Ross, rev. L Brown, Oxford University Press, Oxford. *(Cited by standard Book, Chapter, and universal Bekker pagination.)*

AXA Konzern AG 2025, *Corporate governance and executive board announcement: Dr Johannes Dick succeeds Sirka Laudon and leads expanded People and Data Experience portfolio effective 1 January 2026*, press release, AXA Germany, Cologne, accessed 29 November 2026, <https://www.axa.de/presse/mediathek/pressemitteilungen>.

Baiyere, A, Salmela, H & Tapanainen, T 2020, 'Digital transformation and the new logics of business process management', *European Journal of Information Systems*, vol. 29, no. 3, pp. 238–259, DOI:10.1080/0960085x.2020.1718007. *(In-text page numbers verified against the authors' Copenhagen Business School accepted manuscript.)*

Belizón, MJ & Kieran, S 2022, 'Human resources analytics: a legitimacy process', *Human Resource Management Journal*, vol. 32, no. 3, pp. 603–630, DOI:10.1111/1748-8583.12417. *(In-text page numbers verified against the University of Limerick open-access early-view manuscript, pp. 1–28.)*

Bojinov, II, Lakhani, KR & Lane, D 2025, *JPMorganChase: leadership in the age of GenAI*, Harvard Business School Case 9-325-066 (rev. 28 April 2025), Harvard Business School Publishing, Boston.

Bojinov, II, Lakhani, KR & Lefort, A 2025, *Building an AI factory at Procter & Gamble*, Harvard Business School Case 9-625-015 (6 March 2025), Harvard Business School Publishing, Boston.

Bousquette, I 2025, 'Why Moderna merged its tech and HR departments', *The Wall Street Journal* (CIO Journal), 12 May (print edition 13 May 2025), pp. 1–4 (PDF print), accessed 28 November 2026, <https://www.wsj.com/articles/why-moderna-merged-its-tech-and-hr-departments-95318c2a>.

Bovens, M 1998, *The quest for responsibility: accountability and citizenship in complex organisations*, Cambridge University Press, Cambridge.

Brynjolfsson, E, Li, D & Raymond, L 2025, 'Generative AI at work', *The Quarterly Journal of Economics*, vol. 140, no. 2, pp. 889–942, DOI:10.1093/qje/qjae044. *(In-text page 915 verified against the published Version of Record; remaining page numbers verified against the authors' revised manuscript, arXiv:2304.11771v2, 7 November 2024, pp. 1–36.)*

California Civil Rights Council 2025, *Modifications to employment regulations regarding automated-decision systems*, California Code of Regulations, Title 2, Division 4.1, Chapter 5, Subchapter 2, §§ 11008–11013 (filed 27 June 2025; effective 1 October 2025), accessed 14 October 2026, <https://calcivilrights.ca.gov/wp-content/uploads/sites/32/2025/06/Final-Text-regulations-automated-employment-decision-systems.pdf>.

California Privacy Protection Agency 2025, *CCPA updates, cybersecurity audits, risk assessments, automated decisionmaking technology (ADMT), and insurance regulations*, California Code of Regulations, Title 11, §§ 7001, 7150, 7200–7222 (approved 22 September 2025; ADMT compliance by 1 January 2027), accessed 14 October 2026, <https://cppa.ca.gov/regulations/pdf/ccpa_updates_cyber_risk_admt_appr_text.pdf>.

Chowdhury, S, Budhwar, P & Wood, G 2024, 'Generative artificial intelligence in business: towards a strategic human resource management framework', *British Journal of Management*, vol. 35, early view pp. 1–12, DOI:10.1111/1467-8551.12824.

Chowdhury, S, Dey, P, Joel-Edgar, S, Bhattacharya, S, Rodriguez-Espindola, O, Abadie, A & Truong, L 2023, 'Unlocking the value of artificial intelligence in human resource management through AI capability framework', *Human Resource Management Review*, vol. 33, no. 1, art. 100899, DOI:10.1016/j.hrmr.2022.100899. *(In-text page numbers verified against the authors' Aston University accepted manuscript, pp. i, 1–25.)*

City of New York 2021, *Local Law 144 of 2021: Automated Employment Decision Tools*, New York City Administrative Code §§ 20-870 to 20-874 (enforcement commenced 5 July 2023), NYC Department of Consumer and Worker Protection, accessed 29 November 2026, <https://www.nyc.gov/site/dca/about/automated-employment-decision-tools.page>.

Combs, J, Liu, Y, Hall, A & Ketchen, D 2006, 'How much do high-performance work practices matter? A meta-analysis of their effects on organizational performance', *Personnel Psychology*, vol. 59, no. 3, pp. 501–528, DOI:10.1111/j.1744-6570.2006.00045.x.

Commonwealth of Australia 2009, *Fair Work Act 2009* (Cth), s. 205 (Enterprise agreements to include a consultation term), Federal Register of Legislation, accessed 29 November 2026, <https://www.legislation.gov.au/C2009A00028/latest/text>.

Commonwealth of Australia 2024, *Privacy and Other Legislation Amendment Act 2024* (No. 128, 2024), Australian Privacy Principles 1.7–1.9 (commencing 10 December 2026), Federal Register of Legislation; Office of the Australian Information Commissioner guidance accessed 13 October 2026, <https://www.oaic.gov.au/news/media-centre/new-resources-on-transparency-for-use-of-ai-and-automated-decision-making>.

Congreso Nacional de Chile 2024, *Ley N° 21.719 que regula la protección y el tratamiento de los datos personales y crea la Agencia de Protección de Datos Personales* (published 13 December 2024; entry into force amended to 15 October 2027 by *Ley N° 21.855*, published 23 November 2026), Articles 8 bis, 14 ter, and 15 ter, Biblioteca del Congreso Nacional de Chile, accessed 25 November 2026, <https://www.bcn.cl/leychile/navegar?idNorma=1209272>.

Dell'Acqua, F 2022, *Falling asleep at the wheel: human/AI collaboration in a field experiment on HR recruiters*, Laboratory for Innovation Science at Harvard working paper, Cambridge, MA.

Dell'Acqua, F, McFowland, E III, Mollick, ER, Lifshitz-Assaf, H, Kellogg, K, Rajendran, S, Krayer, L, Candelon, F & Lakhani, KR 2023, *Navigating the jagged technological frontier: field experimental evidence of the effects of AI on knowledge worker productivity and quality*, Harvard Business School Working Paper No. 24-013, September, DOI:10.2139/ssrn.4573321 (published Version of Record: Dell'Acqua et al. 2026, *Organization Science*, vol. 37, no. 2, pp. 403–423, DOI:10.1287/orsc.2025.21838).

DiMaggio, PJ & Powell, WW 1983, 'The iron cage revisited: institutional isomorphism and collective rationality in organizational fields', *American Sociological Review*, vol. 48, no. 2, pp. 147–160, DOI:10.2307/2095101.

Ellul, J 1964, *The technological society*, trans. J Wilkinson, Vintage Books / Alfred A. Knopf, New York (originally published in French 1954). *(Reflowable EPUB edition; cited by Note to the Reader and Chapter.)*

Enholm, IM, Papagiannidis, E, Mikalef, P & Krogstie, J 2022, 'Artificial intelligence and business value: a literature review', *Information Systems Frontiers*, vol. 24, pp. 1709–1734, DOI:10.1007/s10796-021-10186-w.

European Parliament and Council of the European Union 2024a, *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*, Articles 4, 26(7), 113, and Annex III, as amended by *Regulation (EU) 2026/1744 (Digital Omnibus)*, Official Journal of the European Union, OJ L 2026/1744, 24 July 2026 (Annex III employment high-risk obligations applying from 2 December 2027), accessed 14 October 2026, <https://eur-lex.europa.eu/eli/reg/2026/1744/oj>.

European Parliament and Council of the European Union 2024b, *Directive (EU) 2024/2831 of 23 October 2024 on improving working conditions in platform work (Platform Work Directive)*, Chapter III (Articles 7–15), Official Journal of the European Union, OJ L 2024/2831, 11 November 2024, accessed 29 November 2026, <https://eur-lex.europa.eu/eli/dir/2024/2831/oj>.

Federal Republic of Germany 2001, *Betriebsverfassungsgesetz (Works Constitution Act, BetrVG)*, as promulgated 25 September 2001 (BGBl. I p. 2518) and amended, §§ 87(1) No. 6 and 90, Gesetze im Internet, accessed 29 November 2026, <https://www.gesetze-im-internet.de/betrvg/>.

Firk, S, Hanelt, A, Oehmichen, J & Wolff, M 2021, 'Chief digital officers: an analysis of the presence of a centralized digital transformation role', *Journal of Management Studies*, vol. 58, no. 7, pp. 1800–1831 (early view pp. 1–32), DOI:10.1111/joms.12718.

Fischer, LH, Wunderlich, N & Baskerville, R 2023, 'Artificial intelligence and digital work: the sociotechnical reversal', *Proceedings of the 56th Hawaii International Conference on System Sciences (HICSS)*, pp. 226–235, DOI:10.24251/hicss.2023.027.

Henderson, JC & Venkatraman, N 1999 [1993], 'Strategic alignment: leveraging information technology for transforming organizations', *IBM Systems Journal*, vol. 38, nos 2 & 3, pp. 472–484 (reprinted from vol. 32, no. 1, 1993, pp. 4–16), DOI:10.1147/sj.382.0472.

Hume, D 2000 [1739–40], *A treatise of human nature*, eds DF Norton & MJ Norton, Oxford University Press, Oxford. *(Cited by universal Book, Part, and Section.)*

Institute of Internal Auditors (IIA) 2020, *The IIA's three lines model: an update of the three lines of defense*, Institute of Internal Auditors, Lake Mary, FL.

International Organization for Standardization (ISO) 2023, *ISO/IEC 42001:2023 Information technology: artificial intelligence management system*, ISO, Geneva.

Jensen, MC & Meckling, WH 1992, 'Specific and general knowledge, and organizational structure', in L Werin & H Wijkander (eds), *Contract economics*, Blackwell, Oxford, pp. 251–274.

Jonas, H 1984, *The imperative of responsibility: in search of an ethics for the technological age*, trans. H Jonas with D Herr, University of Chicago Press, Chicago.

Kellogg, KC, Valentine, MA & Christin, A 2020, 'Algorithms at work: the new contested terrain of control', *Academy of Management Annals*, vol. 14, no. 1, pp. 366–410, DOI:10.5465/annals.2018.0174.

Komesar, NK 1994, *Imperfect alternatives: choosing institutions in law, economics, and public policy*, University of Chicago Press, Chicago.

Lawrence, PR & Lorsch, JW 1967, *Organization and environment: managing differentiation and integration*, Division of Research, Graduate School of Business Administration, Harvard University, Boston.

Legge, K 2005, *Human resource management: rhetorics and realities*, anniversary edn, Palgrave Macmillan, Basingstoke.

Lepak, DP & Snell, SA 1999, 'The human resource architecture: toward a theory of human capital allocation and development', *Academy of Management Review*, vol. 24, no. 1, pp. 31–48, DOI:10.5465/amr.1999.1580439.

Makarius, EE, Mukherjee, D, Fox, JD & Fox, AK 2020, 'Rising with the machines: a sociotechnical framework for bringing artificial intelligence into the organization', *Journal of Business Research*, vol. 120, pp. 262–273, DOI:10.1016/j.jbusres.2020.07.045.

Matthias, A 2004, 'The responsibility gap: ascribing responsibility for the actions of learning automata', *Ethics and Information Technology*, vol. 6, no. 3, pp. 175–183, DOI:10.1007/s10676-004-3422-1.

McKinsey & Company 2026, *HR Monitor 2026: a comprehensive benchmark survey of workforce and HR trends across Europe, the United States, and China*, McKinsey & Company, June, accessed 28 November 2026, <https://www.mckinsey.com/capabilities/people-and-organization/our-insights/hr-monitor#/>.

Mobley v. Workday, Inc., No. 3:23-cv-00770 (N.D. Cal., Judge Rita F. Lin, conditional ADEA collective certification 16 May 2025; class certification motion 14 September 2026), CourtListener docket, accessed 14 October 2026, <https://www.courtlistener.com/docket/66831340/mobley-v-workday-inc/>.

National Institute of Standards and Technology (NIST) 2023, *Artificial intelligence risk management framework (AI RMF 1.0)*, NIST AI 100-1, US Department of Commerce, Gaithersburg, MD, DOI:10.6028/NIST.AI.100-1.

Nishii, LH, Lepak, DP & Schneider, B 2008, 'Employee attributions of the "why" of HR practices: their effects on employee attitudes and behaviors, and customer satisfaction', *Personnel Psychology*, vol. 61, no. 3, pp. 503–545, DOI:10.1111/j.1744-6570.2008.00121.x.

Orlikowski, WJ 2000, 'Using technology and constituting structures: a practice lens for studying technology in organizations', *Organization Science*, vol. 11, no. 4, pp. 404–428, DOI:10.1287/orsc.11.4.404.14600.

Parker, SK & Grote, G 2022, 'More than "more than ever": revisiting a work design and sociotechnical perspective on digital technologies', *Applied Psychology*, vol. 71, no. 4, pp. 1215–1223, DOI:10.1111/apps.12425.

Parliament of Australia, Joint Select Committee on Artificial Intelligence 2026, *Artificial Intelligence – the Australian way: In our interest. On our terms.*, Final Report (46 recommendations), Commonwealth of Australia, Canberra, 24 November, accessed 29 November 2026, <https://www.aph.gov.au/Parliamentary_Business/Committees/Joint/Artificial_Intelligence/ArtificialIntelligence/Report>.

Parliament of New South Wales 2026, *Work Health and Safety Amendment (Digital Work Systems) Act 2026* (NSW), assented 18 February 2026 (amending *Work Health and Safety Act 2011* (NSW), ss. 19(3)(c1) and 21A), SafeWork NSW / NSW Legislation.

Parliament of the United Kingdom 2025, *Data (Use and Access) Act 2025* (UK), c. 18, s. 80 (substituting Articles 22A–22D into UK GDPR; brought into force 5 February 2026 by S.I. 2026/82), accessed 14 October 2026, <https://www.legislation.gov.uk/ukpga/2025/18/section/80>.

Pasmore, W, Winby, S, Mohrman, SA & Vanasse, R 2018, 'Reflections: sociotechnical systems design and organization change', *Journal of Change Management*, online first pp. 1–19 (print 2019, vol. 19, no. 2, pp. 67–85), DOI:10.1080/14697017.2018.1553761.

Pettit, P 1997, *Republicanism: a theory of freedom and government*, Oxford University Press, Oxford.

Raisch, S & Krakowski, S 2021, 'Artificial intelligence and management: the automation–augmentation paradox', *Academy of Management Review*, vol. 46, no. 1, pp. 192–210, DOI:10.5465/amr.2018.0072.

Rawls, J 1999 [1971], *A theory of justice*, rev. edn, Belknap Press of Harvard University Press, Cambridge, MA.

Ross, JW, Weill, P & Robertson, DC 2006, *Enterprise architecture as strategy: creating a foundation for business execution*, Harvard Business School Press, Boston.

Ryle, G 1949, *The concept of mind*, Hutchinson, London (1951 reprint).

Santoni de Sio, F & Mecacci, G 2021, 'Four responsibility gaps with artificial intelligence: why they matter and how to address them', *Philosophy & Technology*, vol. 34, no. 4, pp. 1057–1084, DOI:10.1007/s13347-021-00450-x.

Schmitt, M 2026, *Strategic integration of artificial intelligence in the C-suite: the role of the Chief AI Officer*, working paper, University of Oxford, arXiv:2407.10247v4, 29 October, accessed 28 November 2026, <https://arxiv.org/pdf/2407.10247>.

Sebastian, IM, Weill, P, Haskamp, T & vom Brocke, J 2026, *MIT CISR AI Decision Matrix: designing decision rights for AI*, Research Briefing No. XXVI-6, MIT Center for Information Systems Research & University of Münster, 18 June, pp. 1–4, accessed 15 October 2026, <https://cisr.mit.edu/publication/2026_0601_AIDecisionMatrix_SebastianWeillHaskampVomBrocke>.

State of Colorado 2026, *Senate Bill 26-189: Consumer protections for artificial intelligence / automated decision-making technology* (signed May 2026, repealing and reenacting SB 24-205; effective 1 January 2027), Colorado General Assembly, Denver.

State of Illinois 2024, *House Bill 3773: Amendment to the Illinois Human Rights Act regarding artificial intelligence in employment* (Public Act 103-0804, signed August 2024; effective 1 January 2026), Illinois General Assembly, Springfield.

Tangi, L, Rodriguez Müller, AP & Janssen, M 2025, 'AI-augmented government transformation: organisational transformation and the sociotechnical implications of artificial intelligence in public administrations', *Government Information Quarterly*, vol. 42, art. 102055, pp. 1–12, DOI:10.1016/j.giq.2025.102055.

Thompson, DF 1980, 'Moral responsibility of public officials: the problem of many hands', *American Political Science Review*, vol. 74, no. 4, pp. 905–916, DOI:10.2307/1954312.

Trist, EL & Bamforth, KW 1951, 'Some social and psychological consequences of the longwall method of coal-getting: an examination of the psychological situation and defences of a work group in relation to the social structure and technological content of the work system', *Human Relations*, vol. 4, no. 1, pp. 3–38, DOI:10.1177/001872675100400101.

Tsebelis, G 2002, *Veto players: how political institutions work*, Princeton University Press, Princeton, NJ.

Vallor, S 2016, *Technology and the virtues: a philosophical guide to a future worth wanting*, Oxford University Press, New York. *(Reflowable EPUB edition; cited by Introduction, Chapter, and Section.)*

Wall, TD & Wood, SJ 2005, 'The romance of human resource management and business performance, and the case for big science', *Human Relations*, vol. 58, no. 4, pp. 429–462, DOI:10.1177/0018726705055032.

Weber, M 1978 [1922], *Economy and society: an outline of interpretive sociology*, eds G Roth & C Wittich, University of California Press, Berkeley.

Weil, D 2014, *The fissured workplace: why work became so bad for so many and what can be done to improve it*, Harvard University Press, Cambridge, MA.

Weill, P & Ross, JW 2004, *IT governance: how top performers manage IT decision rights for superior results*, Harvard Business School Press, Boston.

Wessel, L, Baiyere, A, Ologeanu-Taddei, R, Cha, J & Blegind Jensen, T 2021, 'Unpacking the difference between digital transformation and IT-enabled organizational transformation', *Journal of the Association for Information Systems*, vol. 22, no. 1, pp. 102–129, DOI:10.17705/1jais.00655. *(In-text page numbers verified against the authors' Copenhagen Business School accepted manuscript, pp. 1–41.)*

Williamson, OE 1996, *The mechanisms of governance*, Oxford University Press, Oxford.

Wright, PM & Ulrich, MD 2017, 'A road well traveled: the past, present, and future journey of strategic human resource management', *Annual Review of Organizational Psychology and Organizational Behavior*, vol. 4, pp. 45–65, DOI:10.1146/annurev-orgpsych-032516-113052.

Zuboff, S 1988, *In the age of the smart machine: the future of work and power*, Basic Books, New York.

HOW TO REFERENCE THIS ESSAY & LICENCE

HARVARD: Ives, S 2026, *Where accountability sits: why the people function must frame work-system enterprise AI*, simonives.com, November, <https://simonives.com/essays/where-accountability-sits.pdf>.

APA (7TH EDN): Ives, S. (2026, November). *Where accountability sits: Why the people function must frame work-system enterprise AI*. simonives.com. https://simonives.com/essays/where-accountability-sits.pdf

CHICAGO (17TH EDN): Ives, Simon. 2026. *Where Accountability Sits: Why the People Function Must Frame Work-System Enterprise AI*. simonives.com, November 2026. https://simonives.com/essays/where-accountability-sits.pdf.

© 2026 Simon Ives. Published at simonives.com. This essay is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International Licence (CC BY-NC-ND 4.0). Readers are free to copy and redistribute the material in any medium or format for non-commercial purposes, provided appropriate credit is given to Simon Ives and simonives.com, a link to the licence is included, and the material is not modified, adapted, or built upon.

