



PAPER

HR & L&D Leaders

Your Reskilling Budget Is Training for Yesterday

Your reskilling budget is training people for tasks AI already does better. What people leaders need to build instead: human capabilities that compound with AI, not compete against it.

C4AIL — Centre for AI Leadership
25 March 2026

c4ail.org • Centre for AI Leadership

Your Reskilling Budget Is Training for Yesterday

Paper 6: For People Leaders **Centre for AI Leadership (C4AIL)** - 2026

The Reskilling Paradox

Here is a number that should haunt every CHRO in the world: US firms spent \$40 billion on AI in 2024, and 95% of them saw zero measurable impact on their bottom line.

Now here is the number that should haunt you specifically: 82% of employees received no AI training at all in 2025 (Deloitte). Of the 18% who did, the vast majority completed prompt-writing courses - how to get the machine to produce what you want. Not how to verify what it produces. Not how to redesign work around it. Not how to build the human capabilities that AI cannot replace.

The reskilling gap is not small. Pearson and Faethm estimate it at \$1.1 trillion annually - the aggregate cost of workforces trained for roles that are being structurally transformed underneath them. That is not a training budget problem. That is a strategic misallocation of human capital at civilisational scale.

But the uncomfortable truth is this: even if you tripled your training spend tomorrow, you would be wasting most of it. Because the problem is not insufficient training. It is the wrong kind.

Standard AI reskilling produces better intellectual labourers - people who can draft faster, synthesise more, generate more polished output. But AI commoditises intellectual labour. It does intellectual labour better than any individual human, at near-zero marginal cost, 24 hours a day. Every dollar you spend making your people faster at the thing the machine already does is a dollar that feeds the machine you are supposed to be competing with.

Your reskilling budget is training for yesterday. The question is: what does training for tomorrow look like?

The Four Labours - What HR Needs to Understand

The first thing to understand is that "work" is not one thing. Every job, every role, every task in your organisation is a bundle of four fundamentally different types of labour. Getting this wrong is the reason most workforce planning is failing.

Intellectual Labour is strategy, synthesis, coding, writing, analysis, research - the production of ideas and documents. This is the category being commoditised. AI does it faster, cheaper, and at scale. Not perfectly - but well enough that the human version is no longer the competitive differentiator it was. **Physical Labour** is converging on the same trajectory, but on a 2-3 year lag. Robotics, autonomous systems, and warehouse automation are doing to physical tasks what language models did to knowledge tasks. For most white-collar workforce planning, this category matters less immediately - but it matters. **Accountability Labour** is the human monopoly. This is judgment, oversight, decision-making under uncertainty, and ownership of outcomes. It is the act of saying "I reviewed this, I stand behind it, and I will answer for it when it goes wrong." AI cannot do this - not because it lacks some future capability, but because accountability requires a person who can be held responsible. You cannot sue an algorithm for malpractice. You cannot hold a language model to account when the strategy it drafted fails. This labour type was always the real value in most senior roles. We just never measured it, because it was invisible next to the intellectual output.

Architectural Labour is the growth category. This is building the systems, templates, verification engines, and structured workflows through which AI operates. It is the difference between giving someone a chatbot and giving someone a guided process with built-in quality checks. Architectural labour is what the 5% of organisations seeing real AI returns have invested in. The other 95% have not - because their workforce planning does not even have a category for it.

Every job is a bundle of tasks across these four columns. Take a Compliance Officer. Before AI: roughly 40% intellectual (drafting policies, analysing regulations, writing reports), 10% physical (site visits, document handling), 30% accountability (signing off on compliance status, defending decisions to regulators), 20% coordination that is now architectural (building compliance frameworks, maintaining verification systems).

AI handles the 40% intellectual. The 30% accountability was always the real value - but your performance system measures the 40% being automated. Your annual review template rewards the person who produced the most reports. Not the person who caught the error that would have cost you a regulatory fine.

If your workforce planning does not distinguish between these four types, you are optimising for the wrong column.

The Five Roles – Your New Org Chart

Job titles describe reporting lines. They tell you who reports to whom and what budget code to charge. They do not tell you what kind of labour a person actually performs - and in the AI era, that is the only thing that matters.

The Five Roles describe labour functions, not positions on an org chart. A single person might hold one job title and perform across multiple roles. The roles are:

Floor Users make up 90-95% of your workforce. They work through AI-structured interfaces - guided workflows, templated processes, verification-enabled tools. Their job is not to be AI experts. Their job is to be domain experts who use AI-enhanced systems to do their actual work better. The Floor is not a demotion. It is where value is delivered. A senior claims adjuster working through a well-architected AI system produces more accurate, more consistent, more defensible output than the same person with a blank ChatGPT window. **Translators** bridge domain expertise and AI capability. They are the people who can look at a business process and say "this is where AI helps, this is where it hurts, and this is how we verify the output." The Translator does not need to write code. They need Minimum Viable Literacy - enough technical understanding to commission, interrogate, and challenge the systems they use. Lightcast and Revelio Labs data shows the Translator capability commands a 15-25% salary premium in 2025 - and the premium is growing as demand outstrips supply. **Architects** build the systems. They design the Logic Pipes - structured workflows that constrain AI behaviour, embed verification steps, and route output through quality checks before it reaches a human decision-maker. They build the templates, the specifications, the verification engines that make Floor Users effective. An Architect is not an AI prompt engineer. They are a systems designer who happens to work with AI as one of their materials. **Orchestrators** operate at the system-of-systems level. They design and govern the entire pipeline - how multiple AI-enabled workflows interact, how verification cascades across departments, how the organisation adapts when the AI capability shifts (which it does, quarterly). Orchestrators are developed from Architects over 2-3 years of progressive responsibility. They are never hired externally, because the role requires deep understanding of your specific organisational systems, not generic AI knowledge. **Trainers** maintain the human pipeline. They identify capability gaps, develop Floor Users into Translators, mentor Architects, and ensure the organisation does not lose the domain expertise that makes everything else work. The Trainer role is not a full-time position for most people - it is a capability that sits alongside another role. Your best practitioners are already Trainers. They just are not identified, resourced, or measured for it. The workforce planning question is not "how many AI experts do we need?" It is: "across these five roles, where are the gaps - and which gaps are costing us the most?"

Three Career Tracks, Not One Ladder

Traditional career progression assumes a single ladder: junior to mid to senior to manager to director. AI breaks this model, because the skills that make someone a good Floor User are not the skills that make someone a good Architect - and forcing everyone onto the Architecture Track wastes talent and creates miserable Architects.

The Depth Track: Floor User → Senior Floor User → Domain Expert. This is the choice to go deep - to become the person whose domain knowledge makes the AI system trustworthy. The senior claims adjuster who catches the edge case. The regulatory specialist who knows which AI output to trust and which to override. This track is not a consolation prize. It is where 70-80% of your value is generated. Depth Track professionals are compensated through domain-expertise premiums, not management responsibilities they never wanted.

The Architecture Track: Floor User → Translator → Architect → Orchestrator. This is the systems-building path. It requires a specific combination of domain knowledge, technical literacy, and systems thinking that not everyone has or wants.

Promotion on this track is portfolio-based: you advance by demonstrating that you have built systems that measurably improved output quality, not by accumulating years of service.

The Capability

Track: Any level → Trainer. This is the pipeline-maintenance path. It is available to anyone at any point - a senior Floor User who mentors juniors, an Architect who documents their methods, an Orchestrator who develops the next generation. Trainer capability is measured by the performance of the people they develop, not by training hours delivered.

What nobody is measured on, in any track: hours worked, volume of AI output, or number of prompts run. Those metrics measure the commoditised labour type. Measuring prompts-per-day is like measuring keystrokes-per-hour in the typing pool. It tells you nothing about value.

Performance Management That Measures the Right Thing

If your performance system was designed before 2023, it almost certainly measures intellectual output - reports produced, cases closed, code shipped, proposals written. AI makes all of those metrics meaningless as measures of individual contribution. The machine can produce infinite volume. Volume is no longer a differentiator.

Here is what performance measurement looks like when you measure the labour types that matter:

Floor Users are measured on validation accuracy (did they catch the errors?), interrogation quality (did they challenge the output with the right questions?), and domain-specific throughput (not raw volume, but verified output per unit time). A Floor User who produces 10 verified reports is more

valuable than one who produces 30 unverified reports - even though the dashboard shows the second person as 3x more productive. **Architects** are measured on template quality (do the systems they build actually improve Floor User output?), specification maintainability (can someone else maintain and update what they built?), and verification engine effectiveness (what is the false-positive rate? The false-negative rate? How often does bad output slip through?). **Orchestrators** are measured on system-level output (not individual productivity, but the productivity of the entire pipeline they govern), pipeline health (uptime, adaptation speed, error cascade rates), and business outcome verification (can you trace a measurable business result back to the system they designed?). **The compensation model follows the measurement model.** Domain-expertise premiums for Floor Users who go deep. System-output premiums for Architects whose templates improve measurable quality. Verified-business-outcome premiums for Orchestrators whose pipelines deliver traceable results. Each track has its own compensation logic, its own advancement criteria, and its own definition of "excellent."

What nobody is measured on: hours worked, prompts-per-day, "AI adoption score." Those are vanity metrics. Measuring them actively incentivises the wrong behaviour - generating more unverified output, which creates more rework, which consumes more senior review time, which degrades the very capability you are trying to build.

The Attitude Contradiction

If this sounds obvious, consider why most organisations resist it. The answer is structural, not intellectual.

In cybersecurity, the hiring mantra is "hire for attitude, train for skill." The attitude they mean - initiative, accountability, the willingness to push back when something looks wrong - is exactly the capability that AI deployment depends on. It is the capacity to say "this output looks right but my experience says otherwise" and act on that judgment rather than defer.

But the management structures those hires enter reward the opposite. Follow the process. Defer to the senior partner's conclusion. Do not slow down the team. Get your ticket count up. Organisations hire people who can challenge authoritative-sounding output - and then build cultures where challenging the output means challenging the person who approved it.

The "attitude" does not fade after two years. It is trained out by the performance system. And AI makes this lethal. The Eloquence Trap specifically exploits deference: accept the confident, well-structured output, do not question it, move on. An organisation that hires for independent judgment but measures compliance is building the exact workforce profile that the Eloquence Trap is designed to exploit.

This is why changing the performance system is not a nice-to-have. It is the structural prerequisite for every other AI capability investment. You can build the best Logic Pipes, the most rigorous

verification engines, the most sophisticated CAGE templates - and none of it matters if the management culture punishes the person who slows down to use them.

The Cost Model - It Is Less Than You Think

Here is the number that makes most CHROs pause: for a 500-person enterprise, the full Year 1 investment in this capability model is \$410,000-\$680,000.

That breaks down as:

- **Floor deployment:** \$130,000-\$210,000. Structured onboarding, guided workflow rollout, verification training for 90-95% of staff. This is not "AI training." It is workflow redesign with embedded capability development.
- **Translator development:** \$80,000-\$120,000. Identifying and developing the 5-8% of your workforce who have the domain expertise and technical curiosity to bridge the gap. Intensive, cohort-based, 6-month programmes.
- **Architect pipeline:** \$120,000-\$200,000. Building the small team (2-4% of workforce) who will design your Logic Pipes, verification engines, and structured AI systems. This is your highest-value investment per person.
- **Trainer capacity:** \$50,000-\$100,000. Identifying existing L4+ practitioners, giving them time and tools to develop others, building the internal capability pipeline that makes you self-sustaining.
- **Community infrastructure:** \$30,000-\$50,000. Internal knowledge-sharing platforms, practice communities, cross-team learning mechanisms. The connective tissue that prevents capability from being siloed.

Per employee, that is \$820-\$1,360. The Association for Talent Development's 2024 benchmark for average training spend per employee is \$1,283. You are within the existing budget envelope. This is not an additional cost. It is a reallocation from training that produces commoditised intellectual labourers to training that builds the capability AI cannot replace.

Now compare that to your AI infrastructure spend. Cloud compute, hardware, API licenses, platform subscriptions - for most enterprises, that number is 3-5x the human capability investment. For every dollar spent on the humans who use the machines, most organisations spend \$3-\$5 on the machines themselves.

And here is the punchline: MIT's NANDA Lab research shows that 95% of those machine investments fail to deliver measurable returns. Not because the machines do not work. Because the

human capability to use them effectively was never built. You bought a Formula 1 car and hired drivers who learned on go-karts.

The ROI case is not "spend more on training." It is "redirect existing spend from the column being automated to the columns that generate returns." The gap between 10-15% efficiency gains (tools-only deployment) and 25-30% gains (workflow redesign plus capability building) is your budget justification. For a 500-person enterprise, that 15-percentage-point gap typically represents \$2-4 million in annual productivity value. Against a \$410-680K investment. The payback period is measured in months, not years.

Where to Start Monday Morning

You do not need a two-year transformation programme. You need three actions this week.

Action 1: Run the Four-Column Task Decomposition on your top 10 roles. Take your 10 highest-headcount or highest-cost roles. For each one, map every major task into the four labour columns: Intellectual, Physical, Accountability, Architectural. Be honest about which column each task falls into. You will discover that 30-50% of what you currently measure and reward is intellectual labour that AI already handles. You will also discover that the accountability and architectural work - the work that actually differentiates your organisation - has no metrics, no development pathway, and no recognition in your current performance system. **Action 2: Find your existing L4+ practitioners.** They are already in your organisation. They are the people who have, without being asked, built their own verification methods, created their own structured workflows, developed their own quality checks for AI output. They are not in the "AI champions" programme - they are too busy doing actual work. They are your first Trainers. Identify them, resource them, and get out of their way. **Action 3: Stop measuring AI adoption rates. Start measuring verification quality.** The number of people using AI tells you nothing about value creation. The percentage of AI output that passes structured verification on the first attempt tells you everything. Track first-time-right rates. Track rework hours. Track the gap between AI-generated output and verified output. Those are the metrics that predict whether your AI investment will generate returns or generate workstop.

The gap between where you are and where the 5% are is not a technology gap. It is a capability gap. And capability gaps close with investment in people, not in platforms.

Your reskilling budget is not too small. It is pointed in the wrong direction. Turn it around.

This is Paper 6 in the C4AIL series, written for HR leaders, CHROs, and people strategists navigating AI workforce transformation. It draws on the capability framework detailed in "The Labour Architecture: Redesigning Work for the AI Age" (Whitepaper II) - available from C4AIL on request. Paper 1: "Why Your AI Investment Isn't Working" | Paper 2: "What the 5% Do Differently" | Paper 3: "Monday Morning: Where to Start" | Paper 4: "Building for Amplifiers" (for software engineers)

Contact: hello@c4ail.org | centreforaileadership.org