BITA ARC Training Centre for Behavioural Insights for Technology Adoption

COGNITIVE, SOCIAL AND ECONOMIC FACTORS INFLUENCING ADOPTION OF GENERATIVE AI BY WORKERS

Project Deliverable:

Behavioural Insights Report, Rev. 1

Project Team:

Dr Steve Bickley (Project Lead)
Dr Ho Fai Chan
Ms Jennifer Petley
Dr Stephen Whyte

Industry Partner:

Future Skills Organisation (FSO)

BITA ARC Training Centre for Behavioural Insights for Technology Adoption

QUT Queensland University of Technology

Date of Revision:

20 November 2025

Executive Summary

This report undertakes an analysis of international and national secondary data sources to identify cognitive biases, social influences, and economic incentives that affect workers' decisions to adopt Generative Artificial Intelligence (GenAI) and to better understand the pain points for workers. This involved reviewing existing research, reports, and data related to GenAI adoption by different worker cohorts across all sectors, with a specific focus on workers in Finance, Technology and Business occupations.

Findings from the behavioural insights analysis will be translated into employer-ready support resources to help uplift employee GenAI adoption. We will map each insight to clear actions and lightweight artefacts. For example, stage-matched behaviour change guidance becomes modular playbooks and self-assessment tools, trust calibration principles inform line manager conversation scripts and decision checklists to avoid both overreliance and avoidance, and inclusion-by-design translates into accessible templates and cohort-specific variants to close digital and equity divides, and interoperability/portability criteria translate into simple system-fit checklists that can inform high-stakes deployment choices. We will also draft short case studies that apply these change management resources and processes to different employer and employee groups, demonstrating the tangible steps from pre-adoption to pilot to scale and the metrics to track.

All materials developed will be in plain language, refreshable, and co-designed with employers, tested through simulations and workshops, and iteratively refined with employer and employee feedback, including accessibility checks for priority cohorts.

Acknowledgements

Please cite as: Bickley, S., Chan, H.F., & Whyte, S. (2025). Cognitive, social and economic factors influencing adoption of Generative AI by workers. A research report commissioned by Future Skills Organisation.

All errors remain the responsibility of the authors.

We gratefully acknowledge the contributions of Prof. Benno Torgler, Dr. Tony Beatton and Jennifer Petley from ARC BITA along with Michelle Circelli, Tim Burt and Peta Skujins from Future Skills Organisation. Their inputs to the project, its research direction, and feedback were invaluable. I (Dr. Steve Bickley) also wish to thank my partner, Alana, whose insights into the Transtheoretical Model of Behaviour Change and motivational interviewing, as well as our many conversations on gender rights and broader social issues, have significantly shaped the thinking and direction behind this work.

Key Findings

Unfortunately, there is a widening disconnect between expectations about AI's potential and the lived experience of many Australians. We identify a multi-layered landscape of drivers and barriers affecting the GenAI adoption in the Finance, Technology, and Business (FTB) sectors. For this report, we group these drivers and barriers into four categories/domains:

1. ***Technology***
2. ***Individual / Employees***
3. ***Organisational / Employers***
4. ***Systemic***

Across these domains we highlight FTB-specific patterns. For example, regulated data and model risk obligations in Finance, rapid tooling cycles and integration overheads in Technology, and workflow integration plus change management demands in Business.

Core themes

1. **Shadow adoption is ahead of formal rollout:** Unofficial, bottom-up use is already widespread. The practical response is to enable it safely – by providing approved access, basic training, and clear guardrails – rather than banning it outright.
2. **Trust and usability drive real use:** People adopt tools they understand and can verify. Clarity, transparency, and workflow fit matter more than technical power or flashy demos.
3. **SMEs and priority cohorts face the steepest barriers:** Without targeted support, a GenAI divide will widen for frontline, part-time, regional, and digitally excluded workers. Access, time, and training must be designed inclusively.
4. **Governance should be adaptive and standards-aligned:** Lightweight guardrails, clear roles and responsibilities, and short, reversible pilots help balance innovation with accountability.
5. **Skills are shifting toward AI-complementary capabilities:** The most valuable skills are those that supervise, integrate, or extend AI capabilities – alongside human strengths such as judgment, creativity, and problem-solving. Lifelong learning and micro-credentials can help people retool as job roles evolve.
6. **Watch the system-level hazards:** Avoid over-centralising control or monitoring. Design for data portability, train for critical evaluation, and manage environmental impact by using resources efficiently and choosing models that are appropriately sized for the task.

Implications for employers now

1. **Frame GenAI as augmentation:** Position AI as a tool that enhances, not replaces, people's work. Start small, protect data, test a single workflow, and measure time saved and user confidence.
2. **Design for trust:** Show uncertainty clearly, keep humans in the loop for higher-risk tasks, and track when AI outputs are accepted or overridden.
3. **Build skills deliberately:** Offer short, practical training sessions, shared prompt libraries, and peer learning opportunities. Recognise and reward participation through visible acknowledgment or micro-badges.
4. **Run safe-to-fail pilots:** Keep experiments small, time-limited, and easy to reverse. Document lessons learned and maintain flexibility with vendors and tools.
5. **Embed inclusion by design:** Ensure everyone can participate – including frontline, part-time, or remote staff. Provide time, shared access, and support. Gather feedback from all voices and act on what is most representative and high impact.

Table of Contents

Glossary

Term / Abbreviation	***Definition***
ADII	Australian Digital Inclusion Index
AI	Artificial Intelligence
ALM	Augmented Language Model
API	Application Programming Interface
ATSI	Aboriginal and Torres Strait Islander
BDAT	Big Data Analytics and related Technologies
BERT	Bidirectional Encoder Representations from Transformers
CALD	Culturally and Linguistically Diverse
CEO	Chief Executive Officer
CFO	Chief Finance Officer
COO	Chief Operating Officer
CO_2	Carbon Dioxide
CSR	Corporate Social Responsibility
CUE	Carbon Usage Effectiveness
DEWR	Department of Employment and Workplace Relations
DLP	Data Loss Prevention
FOCUS	Formal Conceptual Structure
FSO	Future Skills Organisation
FTB	Finance, Technology, and Business
GDP	Gross Domestic Product
GenAI	Generative Artificial Intelligence
GPT	Generative Pre-Trained Transformer
HR	Human Resources
ICT	Information and Communications Technology
IEA	International Energy Agency
IEC	International Electrotechnical Commission
IoT	Internet of Things
IP	Intellectual Property
ISO	International Organization for Standardisation
IT	Information Technology
JSA	Jobs and Skills Australia
JSC	Jobs and Skills Council
KPI	Key Performance Indicator
LGBTQIA+	Lesbian, Gay, Bisexual, Transgender, Queer/Questioning, Intersex, Asexual/Aromantic, Plus (i.e., other sexual and gender identities not otherwise listed explicitly)
LLaMA	Large Language Model Meta AI
LLM	Large Language Model
MI	Motivational Interviewing
ML	Machine Learning
MLOps	Machine Learning Operations
NAIC	National AI Centre
NDA	Non-Disclosure Agreement
NGO	Non-Government Organisation
NLP	Natural Language Processing

OARS	Open questions, Affirmations, Reflective listening, and Summaries
PaLM	Pathways Language Model
PCA	Principal Components Analysis
PUE	Power Usage Effectiveness
QA	Quality Assurance
QUT	Queensland University of Technology
RACI	Responsible, Accountable, Consulted, Informed
RAG	Retrieval Augmented Generation
ROI	Return on Investment
RPL	Recognition of Prior Learning
RTO	Registered Training Organisation
SDT	Self-Determination Theory
SFIA	Skills Framework for the Information Age
SMEs	Small to Medium Enterprises
SOP	Standard Operating Procedure
STEM	Science, Technology, Engineering and Mathematics
TAFE	Technical and Further Education
TGA	Therapeutic Goods Administration
TTM	Transtheoretical Model
TWh	Terawatt Hour
VAISS	Voluntary AI Safety Standard
VET	Vocational Education and Training
WCAG	Web Content Accessibility Guidelines
WUE	Water Usage Effectiveness

Introduction

Australia is entering a critical period of systemic transition: not just in digital skills, but in how knowledge, work, and capability are defined, distributed, and governed in a world increasingly mediated by advanced technologies like AI, cloud computing and quantum technologies (sensors, computers, communications). Over the recent years, GenAI in particular has been reshaping the way we work, learn, and deliver services across government, industry, and the training/education sector. In this environment, uplifting productivity, ensuring workforce mobility and adaptability, and enabling inclusive technology adoption are all essential focuses of organisations and policymakers alike.

Previous modelling by Microsoft and the Tech Council of Australia (2023) estimates that Generative AI could add between $45 to $115 billion a year to Australia's economy by 2030, depending on the pace of adoption. This is equivalent to between 2 to 5% of National Gross Domestic Product (GDP). Despite an estimated $30-40 billion in enterprise investment globally, recent evidence points to a widening "GenAI divide", with as much as 95% of organisations reporting no measurable profit and loss return on AI Programs (Challapally et al., 2025). For example, in a multi-method study combining a systematic review of 300+ public AI initiatives, 52 structured interviews across organizations, and a survey of 153 senior leaders gathered at four industry conferences, only 5% of integrated pilots created long-lived material value, a gap apparently driven less by model quality or regulation than by approach and workflow fit (Challapally et al., 2025). These figures underscore the significant and measurable loss to organisational and national productivity from delayed or insufficient adoption.

Importantly, research shows GenAI use is in demand by Australian workers, with shadow AI use already widespread (Jobs and Skills Australia, 2025A, 2025B; Catherine et al., 2025**)**. Shadow AI refers to the informal use of personal ChatGPT/Claude-style tools to automate meaningful parts of their jobs. Recent evidence from Challapally et al. (2025**)** estimates a scale gap (with only 40% of firms buying official Large Language Model (LLM) seats vs 90% reporting regular employee LLM use), high intensity of individual employees' day-to-day use, and better Return on Investment (ROI) signals from these bottom-up workflow-fit tools. This represents both a risk and an untapped opportunity for Australian businesses; employers who can identify and provide structured pathways, training, and guardrails will naturally harness this shadow engagement (worker demand) to drive productivity, rather than letting it remain fragmented, insecure, and under-leveraged.

The implications for the competitiveness of Australian businesses, as well as for sovereign capability are profound, which is why national action to leverage GenAI as integrated cognitive partners or *"Tools of Thought"* (Bickley, 2025) must be clearly understood, supported and developed. GenAI should not be seen merely as automation but as augmentation conduit or tool that can reshape how decisions are made, how knowledge is produced, and even how creativity is expressed. All with the key underlying focus of realising both efficiency and productivity gains in practice. This reframing underscores that the workforce transition is both an economic and an epistemic challenge, demanding systems and governance frameworks that embed legitimacy, inclusion, and sovereign capability development.

Right now, many workers are already using AI informally, but without structured support this shadow adoption risks becoming fragmented, insecure, and a missed productivity opportunity.

The Commonwealth Government of Australia recognises the importance of AI adoption for Australia's future productivity and competitiveness (Commonwealth of Australia, 2021). Yet traditional levers (e.g. such as skilled migration, or tertiary education) are increasingly constrained and, on their own, cannot solve the scale of the workforce transition required (Jobs and Skills Australia, 2025B). Therefore, identifying cognitive and behavioural biases and barriers, social influences, and economic incentives and enablers that affect workers' decisions to adopt GenAI is both a practical and highly scalable way to support all Australian employers through this important time of technological transition.

Unfortunately, there is a widening disconnect between expectations about AI's potential and the lived experience of many Australians (HTI, 2025). Without a strategic and inclusive framework, AI and related enabling technologies risk amplifying structural inequities and compounding the exclusion of priority cohorts from the future economy (Bickley, Macintyre & Torgler, 2025). As with earlier general-purpose technologies such as electrification, the internal combustion engine, and the internet, GenAI and the broader digital shift may, for some disadvantaged or marginalised Australian populations, reduce labour market access and opportunity. For Culturally and Linguistically Diverse (CALD) and Aboriginal and Torres Strait Islander (ATSI) populations, the elderly, women, LGBTQIA+ communities, regional/remote workers, neurodiverse individuals, low socio-economic and low literacy, and people living with disabilities, poorly designed adoption pathways and support may lead to disengagement from the formal economy, resistance to reskilling, and deepening social divides. Therefore, closing these gaps is central not just to Australian labour market equity, but to improving commercial and national sovereign competitiveness.

Project Background

Future Skills Organisation (FSO) is one of ten Jobs and Skills Councils (JSCs) established by the Australian Government Department of Employment and Workplace Relations (DEWR), to provide industry with a stronger voice to ensure Australia's Vocational Education and Training (VET) sector delivers better outcomes for learners and employers. FSO is responsible for the FTB sectors, covering the professional services and enabling capabilities that underpin business success – from marketing, accounting and Human Resources (HR), through to Information and Communications Technology (ICT), cybersecurity, financial technologies, AI, and Internet of Things (IoT)/instrumentation.

This project responds directly to a recent request for proposal from FSO, with a strong focus on applied research and behavioural insights. It forms part of a multi-phase initiative led by FSO to diagnose adoption challenges, test different behavioural and systemic interventions, and ultimately inform national training and reskilling/upskilling strategies or human capital development (sovereign capability) in Australia's FTB sectors with emphasis on the enablement of priority cohorts such as CALD, ATSI, women, regional Australians, older workers, and people with disabilities who face increased structural and motivational barriers to economic participation and the adoption of new technologies.

AI tools deliver the biggest productivity gains where routine tasks dominate, especially for less-experienced workers (Brynjolfsson, Li & Raymond, 2025). By accelerating first drafts, meeting summaries, and data formatting, AI augments juniors so they can spend more time on higher-value work. Where roles are mostly automatable (e.g., data entry, basic transcription, first-pass triage), expect task reconfiguration or some headcount reduction. Knowledge-heavy roles in FTB are also exposed, but most organisations should favour augmentation over elimination because judgment, relationships, exception handling, and oversight still matter (Dell'Acqua et al., 2023; Gmyrek, Berg & Bescond, 2023; Agrawal, Gans & Goldfarb, 2023). Experienced professionals often capture disproportionate value due to domain fluency and absorptive capacity, reinforcing the need for differentiated training – advanced modules/playbooks for power users integrating AI into complex workflows, and scaffolded tools for novices building core skills (Calvino, Reijerink & Samek, 2025).

Given that the FTB sectors are traditionally associated with tertiary pathways, FSO proposes to achieve this by supporting alternative employment and learning pathways with a skills-first approach to help more people reach their potential and help businesses tap into diverse talent pools. Generalist, transferable skills are increasingly valued as jobs and occupational roles evolve, and digital capabilities are now essential across nearly every job (FSO, 2025). Therefore, embracing flexible, alternative employment and learning pathways (i.e., through a mix of VET, Higher Education, and on-the-job learning and professional development opportunities), along with a commitment to lifelong learning, is key to building a more inclusive, agile and productive workforce that drives long-term economic growth.

Drawing on systems mapping and behavioural insights analysis, this report highlights that challenges with GenAI adoption are not only about skills gaps, but also about contextual barriers such as uncertainty, trust, psychological readiness, and organisational capacity. Early exploration of open data sources and behavioural models has informed initial concepts for employer support resources. These concepts emphasise the need to balance human and AI capabilities, manage productivity and uncertainty trade-offs, build sovereign knowledge and workforce capability, and foster interdisciplinary coordination across engineering, behavioural sciences, and policy domains. Subsequent phases will translate these insights into practical, plain-language resources co-designed with employers and tested through simulations and workshops, with final outputs refined through stakeholder feedback and accessibility checks.

Scope and Methodology

To ensure that the Behavioural Insights Report delivers actionable, cohort-specific evidence for the employer support resources developed during later phases of the project, we follow a targeted scoping review to identify the cognitive biases, social influences and economic incentives that affect workers' decisions to adopt GenAI. Three objectives steer the entire pipeline:

1. Gain a better understanding of the pain points for workers,
2. Map how those drivers differ for women, CALD and First Nations workers, part-time staff and small businesses, and
3. Translate concepts like "trust in AI," "policy clarity," and "peer learning" into easy-to-track indicators (e.g., survey results, policy pages, training mentions) that we can estimate from publicly available data.

Evidence was gathered through a broad scan of academic literature, government reports, industry white papers and union/Non-Government Organisation (NGO) briefs, with a particular focus on those published between 2017 to mid-2025 (i.e., focusing primarily on the current GenAI era). Two primary scholarly engines (Google Scholar and the QUT e-library gateway) cover the peer reviewed domain, while Overton.io, Australian departmental repositories and selected union/NGO sites capture the grey literature. Searches were keyword guided but remained deliberately open. Only empirical or conceptual works on GenAI adoption proceeded to full text appraisal (i.e., purely technical AI/technology papers were generally set aside).

Drivers and Barriers

We identify a multi-layered landscape of drivers and barriers affecting the GenAI adoption in the FTB sectors. These dynamics span Technology, Individual/Employees, Organisational/Employers, and Systemic domains.

Technology

Rapid advances in AI capabilities, especially LLMs, Augmented Language Models (ALMs, see e.g., Mialon et al., 2023) and related GenAI tools and systems (e.g., agentic systems: Miehling et al., 2025; Hughes et al., 2025), are a strong technological driver of adoption. The ongoing innovation in model performance of frontier GenAI systems has significantly extended their reasoning, language understanding, and knowledge generation and retrieval abilities (Minaee et al., 2024; Sun et al., 2025). Similarly, the evolution of pretrained foundation models such as Bidirectional Encoder Representations from Transformers (BERT), Generative Pre-Trained Transformer (GPT), and ChatGPT demonstrates how pretraining on massive corpora lowers barriers to downstream applications across domains, including text, vision, and graph learning (Zhou, Li et al., 2024), with agentic systems now emerging as a further paradigm shift, embedding LLM agents into industry-specific workflows and demonstrating transformative business adoption across sectors (Bousetouane, 2025; Challapally et al., 2025). Together, these developments, coupled with the wide availability of scalable cloud infrastructure, mean that sophisticated AI tools are increasingly accessible. In effect, barriers to entry are lowered by commercialized AI-as-a-Service platforms and even open-source models, allowing organisations of all sizes to experiment with GenAI without needing in-house AI research teams.

Yet, as Zhou, Yuan et al. (2024) emphasize, the structural properties of cloud pricing models introduce biases that shape user decision-making. Users face trade-offs between time and cost, as choices about instance type and number rarely allow for both minimized expense and optimal performance simultaneously. These dynamics are further complicated by inconsistencies in pricing fairness across instances, which create ongoing rework and cognitive load for tech workers. Zhou, Yuan et al. (2024) also highlights fairness concerns – i.e., between "personal fairness" (individual expectations) and "social fairness" (equitable treatment across users) – alongside cognitive limitations, where users struggle with optimisation due to bounded rationality and may exhibit a simplification bias by equating cost-optimisation with performance optimisation. Combined, these insights show that while technical progress and cloud accessibility are lowering adoption barriers, the economic and cognitive frictions inherent in pricing models remain critical hurdles. This dynamic reflects broader platform power asymmetries in informational capitalism, where opaque technical and legal architectures reinforce user powerlessness and entrench provider control (Guihot & McNaught, 2021).

Some researchers warn that as GenAI tools become more routine and a part of everyday work, people can get mentally locked in to how a specific platform works. Over time, its prompts, defaults, and suggestions shape how we write and solve problems, so switching tools – or thinking outside those patterns – gets harder (Hansen, 2024). At the model level, scaling up systems like GPT-3 creates impressive abilities, but it can also bake in flimsy patterns and leave models brittle in the face of trick questions – raising the question of whether they truly "understand" language or are just using statistical shortcuts (Tamkin et al., 2021).

Privacy risks add another layer: large models can memorise snippets of sensitive data, which may be hard to remove and undermines "notice-and-choice" style consent (Winograd, 2022, pp. 640–641). Governance challenges compound this: frontier models pose three problems – unpredictable dangerous capabilities, safe deployment, and uncontrolled spread (Anderljung et al., 2023). Current rules are patchy and reactive; scholars call for joined-up frameworks that combine cybersecurity safeguards, cryptographic protections, and coordinated oversight to manage adversarial risks and ensure accountability at scale (Radanliev, 2025).

Bottom line: *GenAI adoption isn't just about technical performance. It is also shaped by cognitive lock-ins, privacy gaps, and the need for coherent, forward-looking regulation.*

Table 1: Technological enablers and constraints.

Drivers	Barriers
Rapid capability gains – **LLMs, ALMs, and agentic systems are expanding reasoning, language, and workflow integration.**	**Cloud pricing & vendor lock-in** – complex pricing models force trade-offs between cost and performance, increasing rework and dependency.
Lower barriers to entry – **pre-trained models and AI-as-a-Service make advanced tools available to firms of all sizes.**	**Integration challenges** – connecting GenAI with legacy Information Technology and data sources is costly and slow ("last mile" problem).
Workflow integration potential – **embedding tools into daily tasks can automate repetitive, low-value tasks and boost productivity.**	**Biases & errors** – framing effects (the way information is presented), negativity bias (weighing bad news more than good), and data distortions undermine trust and usability.
Accessible experimentation – **open-source models and cloud platforms enable low-risk trials.**	**Privacy & security risks** – memorisation of sensitive data raises concerns over compliance and breaches.
Analogy & cross-translation – LLMs/ALMs can **reframe concepts**, **translate jargon across silos**, and scaffold communication between staff, **lowering coordination costs** and **surfacing tacit knowledge**.	**Governance gaps** – fragmented, reactive regulation; need for harmonised safety and accountability frameworks.
Multimodality (text/speech/vision/tables) – **unlocks frontline use, accessibility, and richer Quality Assurance (QA) into LLM workflows.**	**Misuse potential** – harmful applications (toxins, deepfakes, disinformation) require safeguards.
	Technical readiness ≠ adoption – tools risk becoming "shelfware" without design, training, and support.
	Environmental footprint – high energy and water use in data centres; growing need for "Green AI" practices.

Even with secure infrastructure, systems integration challenges persist. GenAI tools often need to interface with legacy ICT systems and proprietary data sources. If Application Programming Interfaces (APIs) and data formats are not compatible, integration can be costly and slow. Indeed, surveys of Australian businesses show interoperability with existing systems and processes is a significant hurdle to implementation (Fifth Quadrant, 2025). Older or less digitized organisations (common among small businesses and government agencies) may lack the modern data pipelines needed to fully leverage AI outputs, creating a "last mile" problem where the technical tool exists but cannot plug seamlessly into workflows.

AI systems are vulnerable to multiple cognitive biases that shape both outputs and user adoption. For instance, framing effects make models and their users more risk-averse when information is expressed in terms of mortality rather than survival, leading to skewed medical recommendations (Wang & Redelmeier, 2025). Negativity bias means that potential risks are weighted more heavily than benefits in technology adoption, particularly when consumers or workers evaluate AI tools, which reduces uptake despite clear efficiency gains (Frank, Chrysochou, & Mitkidis, 2023). At the design level, representativeness bias and causal illusion in training data cause systematic misclassifications, such as equating healthcare costs with medical need, undermining trust and usability (Martínez, Agudo, & Matute, 2022). These technical distortions are not neutral but directly influence whether AI is seen as augmenting or threatening workplace practices.

Language models also raise questions of misuse and proportional restriction: some capabilities may justifiably require targeted interventions, such as limiting access to models capable of designing toxins, generating harmful imagery, or automating spear phishing, in order to balance the *Misuse-Use Tradeoff* (Anderljung et al., 2024). At the same time, large-scale LLMs present a wide array of ethical and social risks, from reinforcing discriminatory stereotypes to enabling malicious uses like fraud and disinformation campaigns, underscoring the need for multi-layered mitigation strategies (Weidinger et al., 2021). In this context, governance scholars have proposed risk thresholds as a more principled approach than capability thresholds, since they directly quantify acceptable levels of harm in terms of likelihood and severity – though their reliability remains limited by current challenges in estimating AI risks (Koessler et al., 2024). Importantly, these thresholds are best positioned to inform, not determine, high-stakes deployment decisions, complementing capability and compute thresholds while avoiding premature lock-in of safety measures. In short, technical readiness does not guarantee usage or acceptance. Without thoughtful design, training, and support, a GenAI tool can sit idle ("shelfware") even after the technical hurdles of acquisition are overcome. This highlights that technical and human factors must work in tandem: simply making GenAI available is not enough if cybersecurity, interoperability, and usability issues are not resolved.

Environmental footprint matters, too. AI's infrastructure load is rising quickly: the International Energy Agency (IEA) estimates total global electricity use from data centres (including AI) could roughly double by 2026 to more than 1,000 TWh, with AI a major driver (IEA, 2024). In Australia, data centres are estimated to use ~5% of national electricity today, with forecasts of ~8% by 2030 – and up to ~15% in a high-growth scenario; that's roughly 23-43 TWh against a National Electricity Market load of about 200 TWh/year (Australian Energy Council, 2024). Beyond energy, training and serving large models consume substantial water (especially with evaporative cooling); recent analyses quantify a model's "hidden" water footprint and recommend shifting workloads to cooler/less water-stressed regions and times to cut intensity (Li et al., 2025). To manage this, teams should monitor standard data centre metrics – PUE (power usage effectiveness), WUE (water usage effectiveness), and CUE (carbon usage effectiveness) – which together give a practical view of energy, water, and carbon impacts.[1] At the Machine Learning (ML) layer, "Green AI" practices emphasise reporting and reducing compute (and thus emissions or water) through smaller/efficient models, distillation/quantisation, carbon-aware scheduling, and transparent footprint reporting (Schwartz et al., 2020). Providers are also piloting carbon-intelligent workload shifting (running jobs when/where grids are cleanest), a pattern enterprise and SMEs can emulate by

[1] For calculation tools and guides, see e.g., https://www.thegreengrid.org/resources/library-and-tools.

queueing non-urgent AI jobs for lower-carbon windows or regions (Radovanovic, 2020) when integrating AI applications into their own workflows. Early Natural Language Processing (NLP) work showed how training choices drive orders-of-magnitude differences in energy/carbon – so pair capability targets with efficiency targets (e.g., prefer smaller/faster models where fit-for-purpose, set a "cost/CO_2 per task" Key Performance Indicator (KPI), ask vendors for carbon/energy reporting by workload) (Strubell, Ganesh, & McCallum, 2020).

Individuals/Employees

Usability and the 'Black Box' Problem

From an end-user perspective, technical complexity can translate into usability frictions. If GenAI tools are not user-friendly or if their outputs are not easily interpretable (the classic "black box" problem, see e.g., Bickley, Macintyre & Torgler, 2025, pp. 130-131; von Eschenbach, 2021), employees may under-utilise them even after deployment. Research on technology acceptance has long shown that perceived ease of use and usefulness drive adoption (Davis, 1989). Whilst GenAI's usefulness is evident, if the interface is clunky or the results occasionally erratic, busy workers may just revert to familiar methods. For example, initial pilots of AI systems in workplaces sometimes report low utilisation rates – not because the AI fails, but because users fear errors or do not know how to integrate the AI's output into their decision-making (Raisch & Krakowski, 2021; Papagiannidis et al., 2023).

Trust as Fittingness and Responsiveness

Concerns about the "black box" nature of AI (i.e., where people cannot see how or why a system reaches its decisions) can make users reluctant to rely on it, especially when outcomes have ethical or social consequences (von Eschenbach, 2021), and when organisational transparency or assurance mechanisms are lacking (Gillespie et al., 2025). von Eschenbach (2021) explains that trust in AI has two parts. The first is fittingness: before we trust, we judge whether the system seems competent, reliable, and likely to act in our interests. The second is responsiveness: after trust is placed, we look at how the system actually behaves. The first step comes logically before the second – if people lack good reasons to expect competence or alignment, their trust will be misplaced. Transparency is therefore essential: without visibility into how decisions are made, trust becomes simple dependence and leaves people exposed when outcomes conflict with their interests. This mirrors the principal-agent problem in economics, where principals (such as employers or clients) must delegate to agents (such as employees or contractors) who may hold private information and different incentives. In both cases, mechanisms that increase transparency and align incentives are the basis for genuine trust. Global evidence supports this: trust is associated with both acceptance/approval and use of AI ($r = 0.70$ and $r = 0.48$, respectively – see Gillespie et al., 2025, p. 27), and people are far more willing to trust systems when clear assurance measures – such as human oversight, accuracy checks, and accountability – are in place (Gillespie et al., 2025, pp. 7 & 54).

Ease of use and trust are decisive – if AI feels opaque or clunky, workers quickly revert to old.

Cognitive Biases in Human-AI Interaction

Cognitive architectures, as argued by Bickley and Torgler (2023A), highlight the need for procedural transparency and explainability at a deeper level – where AI systems can communicate the "why" behind decisions in ways that resonate with human values and reasoning. Without such interpretability, employees may perceive AI tools as unpredictable or misaligned with their goals, creating significant barriers to adoption. But this is not unique to AI: when we judge the fittingness of trust and responsiveness to trust in humans, we rely on long-standing disciplines such as cognitive science, psychology, and behavioural economics.

These fields emphasise how people form models of others' minds (theory of mind), whether assessing a colleague, a political leader, or a government and related institutions. Just as with AI systems, we ask whether actors are competent, aligned with our interests, and transparent in how they pursue them. Research on trust in teams reinforces this point. Beatton (2007) shows that new team members initially form trust judgments primarily on the basis of perceived ability, often overlooking benevolence or integrity when entering high-performing teams. This halo bias can distort decision-making, but as teams mature, benevolence and integrity become more important in sustaining trust. The same logic applies to human-AI teaming: employees may initially judge AI systems on performance outputs, but long-term adoption depends on whether the system demonstrates fairness, alignment with values, and integrity in decision-making. Trust, in other words, operates on the same basic principles whether in small teams, organisations, or societies: people extend it only when they see competence, benevolence, and integrity. And because trustworthiness involves both the initial judgment of fittingness and the ongoing responsiveness to trust, it functions as a dynamic system with feedback and feedforward effects – where experiences of responsiveness reinforce (or erode) future judgments of fittingness.

The global trust landscape underscores this point. The 2025 Edelman Trust Barometer shows that trust is not only fragile with respect to technology, but also deeply strained across governments, institutions, and employers. For example, trust in government and media remains below the neutral threshold in many countries, while even business – traditionally, the most trusted institution globally – faces widening gaps linked to income inequality and grievance[2] (i.e., low-income respondents consistently report far lower levels of trust across institutions than high-income respondents, with double digit trust gaps in over 20 countries, and those with high grievance levels show steep declines in trust in business, government, NGOs, and media). In Australia, overall trust in institutions sits at just 49% – below the global average of 56%. Trust in government and media is particularly weak, with only business managing to reach near-neutral levels. The income-based divide is also clear: Australians in the top income quartile report 53% trust in institutions, while those in the bottom quartile report only 44%. Taken together, these findings illustrate that judgments of fittingness and responsiveness apply as much to human and institutional actors as to AI systems. Citizens and employees alike require evidence of competence, fairness, and responsiveness before they are willing to extend trust.

Human-automation research shows that misplaced trust can be just as problematic as a lack of trust: overconfidence may lead to complacency and misuse, while insufficient trust can discourage use altogether, undermining efficiency and increasing workload (Chen & Barnes, 2014; de Visser et al., 2018; Baker & Keebler, 2017). For example, anchoring bias can lead employees to stick too closely to an AI's initial output or score, even when contradictory evidence exists (Rastogi et al., 2022). The primacy effect further shapes reliance on early search results or first recommendations, reinforcing feedback loops in attention and decision-making (Martínez, Agudo, & Matute, 2022). When these biases go unnoticed, they raise cognitive load and ambiguity, as employees struggle to determine when to trust or override the AI. In addition, seemingly precise AI outputs can foster illusions of understanding and narrow the space of options people consider: users may mistake easy-to-evaluate, model-friendly hypotheses for the full problem space and over-generalize from whatever the model's training data happens to cover (Messeri & Crockett, 2024). As the authors highlight, this dynamic also risks a "monoculture" of methods and standpoints, where AI-amenable approaches and the

[2] Grievance refers to a sense of felt injustice regarding the unfairness and harm caused by institutions (e.g., business, government, NGOs, and media), as reported in the 2025 Edelman Trust Barometer. It captures when people believe that government or business actions hurt them and serve the few rather than the many, and that the wealthy or a privileged few benefits from a biased system while ordinary/regular people struggle to get by.

perspectives embedded in training data crowd out alternatives, creating overconfidence without commensurate comprehension (Messeri & Crockett, 2024).

***Cognitive biases complicate adoption** – polished but wrong outputs can mislead, while overconfidence or mistrust both hurt productivity.*

Technical and Human Determinants of Trust

What matters, then, are the factors that shape trust in the first place. On the technical side, people look for cues like reliability, error rates, and system transparency (Chen & Barnes, 2014). On the human side, trust is influenced by individual propensities to trust (Stowers et al., 2017), levels of expertise, and the broader socio-cultural environment. As Schaefer et al. (2016) emphasize, socio-cultural context embeds automation within larger frames of meaning (such as organizational safety culture, community norms, or past experiences), so that judgments about trustworthiness are filtered through shared mental models and values. Equally important are the biases built into AI systems themselves: because models learn from training data, they inherit both the direction of stereotypes (e.g., positive vs. negative associations) and their representativeness (the degree to which certain dimensions, like competence or warmth, dominate perceptions of a category). Nicolas and Caliskan (2024) show that representativeness and direction operate as independent components of stereotypes in language models, meaning that even when a group is rated positively overall, models may overemphasize particular dimensions, skewing judgments of trustworthiness. Simmons and Hare (2023) similarly highlight how LLMs reflect the biases of their training corpora when used as subpopulation representative models, raising risks that users may mistake model outputs as unbiased reflections of group sentiment when in fact they mirror sampling gaps and systemic distortions.

Practically, these data-level distortions heighten the value of human expertise and knowledge scaffolding, which is why systems that support or extend expert judgment are more readily adopted. As Bickley et al. (2021) highlights, the original promise of expert systems was precisely to capture and embed specialist knowledge and reasoning into computational tools to enhance decision-making in complex domains. Yet, as Bickley, Chan and Torgler (2022) discuss, their decline during the AI winter (i.e., industry-wide pullbacks in the late 1970s and late 1980s when brittle, opaque expert systems failed to deliver) underscored the risks of over-promising and under-delivering. A useful contemporary parallel in the GenAI context is prompt engineering: at a high level, it functions like a lightweight, dynamic form of knowledge engineering (i.e., encoding task structure, domain cues, constraints, and retrieval hooks into prompts and toolchains) thereby acting as a hybrid interface between symbolic scaffolding and connectionist/learning models. Framed this way, prompt design (plus Retrieval Augmented Generation or RAG and other tool uses) is a practical path to restore interpretability and align model behaviour with user intent, mitigating the opacity risks raised by more uncertain, complex information processing approaches (e.g., deep learning and/or quantum-inspired methods) while preserving their capacity to detect, match, and generalize patterns across vast and complex data. Together, these technical, individual, and contextual factors highlight why understanding end-users and their environments is critical to building trust that endures.

***Trust must be built and maintained** – employees judge AI like colleagues: competence first, then benevolence and integrity over time.*

Motivations and Bottom-Up Adoption ('Shadow AI')

At the human level, a key driver of GenAI adoption is the motivation and curiosity of individual workers to augment their work with new technology. Many employees are excited by the prospect of offloading drudge tasks or improving their output quality using AI assistance. Early evidence suggests that GenAI usage in workplaces has been largely bottom-up – driven by individuals experimenting with tools like ChatGPT to see if it can help in their roles. A large-scale survey of ~18,000 Danish workers across 11 ChatGPT-exposed occupations (Nov 2023 to Jan 2024) reports that 41% had used ChatGPT for work within a year of release, with younger and less-experienced workers more likely to adopt (Humlum & Vestergaard, 2025). Yet, the authors note the rapid take-up was driven by individual workers rather than firm initiatives. This bottom-up pattern (‘shadow AI’) is also visible in Australia: JSA (2025B) cites surveys showing ~27% of employees using GenAI secretly at work (and 21% in another study), with workers often reporting they need training or face employer restrictions. This kind of ‘shadow AI’ adoption (analogous to shadow IT) builds familiarity even where organisations have no official tools. Yet the same dynamics can split a workforce: role threat dampens experimentation when AI is framed as replacement rather than augmentation (Grover, Kar, & Dwivedi, 2022), while accessibility concerns arise when systems underperform for minority groups (especially where training data representativeness and learned stereotype directions skew outputs) leaving some employees underserved or misrepresented (Martínez, Agudo & Matute, 2022). Add to this that GenAI can amplify confirmation bias, giving users high confidence outputs that align with priors and narrow the option space (Zhang, Song, & Liu, 2025; Messeri & Crockett, 2024). The practical implication is managerial as much as technical: frame deployments as augmentation, pair access with participatory design and skills support, and implement continuous bias/oversight processes rather than “data fixes” alone (Schwartz et al., 2022).

Key Human-Level Enablers

Building on this, three enablers matter: autonomy,[3] demonstrated utility, and cultural familiarity. Autonomy and flexibility in work design lets employees create “safe spaces” to try GenAI (e.g., an analyst automating parts of a report; a marketer testing image generation), which builds personal mastery and demystifies the tools – often turning early users into internal advocates whose visible use creates social proof and shifts norms towards adoption, including both descriptive norms (what people actually do) and injunctive norms (what the group approve/expects). Another human-level driver is demonstrated? utility: once people try GenAI and see tangible payoffs, they tend to keep using it. For writing, Noy and Zhang (2023) show that access to a generative assistant speeds up task completion with similar or better quality; for coding, industry evaluations of GitHub Copilot report shorter completion times with comparable correctness; and in professional services, Dell'Acqua et al. (2023) find junior consultants using GenAI produce work faster and rated higher in quality. When employees observe these concrete benefits – less drudgery, faster completion, improved outputs – their intention to continue rises (Venkatesh et al., 2012). Finally, as GenAI becomes common outside work (personal tasks, study, entertainment), baseline comfort grows and the tech feels routine rather than risky. Over time, these competence wins function as early ability cues that kick-start a dynamic trust cycle: feedback from real outcomes and anticipatory/transparent displays about how and why the system will act recalibrate judgments of the fittingness of trust as teams gain experience (Lee & See, 2004; Hoff & Bashir, 2015). But as adoption matures, ability signals are no longer sufficient, i.e., benevolence and integrity become more decisive for sustaining trust (Beatton, 2007). Therefore, ongoing transparency and pro-social responsiveness (i.e., explaining decisions, correcting errors, and making user-centred changes) matter as much as initial performance.

---

[3] Self-Determination Theory (SDT) treats autonomy as volition (i.e., the intentional decision or choice made after deliberation). In other words, choosing to use GenAI because it fits one’s goals and values, not just acting independently. Motivation is strongest when autonomy is supported alongside competence and relatedness (e.g., meaningful choice, clear rationales, skill scaffolding, peer connection), but tight constraints on time/attention or cognitive/social/economic resources can undermine these needs and shift behaviour toward mere compliance or disengagement. Thus, the design task is to offer constrained choice and protected practice time within real-world limits so internalization can occur. For overviews of SDT, see Deci & Ryan (1985, 2000) and Ryan & Deci (2000).

Coupled with psychosocial risk management obligations under Safe Work Australia – consultation, workload and clarity checks, and basic change management hygiene – these standard processes embed GenAI into the everyday governance routines of organisations while also meeting duty-of-care requirements (Safe Work Australia, 2025). For instance, Safe Work Australia's Model Code of Practice: Managing Psychosocial Hazards at Work (2025), while not about AI per se, underscores employers' duty to manage psychosocial risks including those that could arise from technological change (e.g. workload intensification, low role clarity, poor change management). This means that when introducing GenAI, organisations are expected (systemically) to address psychosocial safety – e.g. consulting workers, monitoring stress levels, and designing jobs to maintain reasonable cognitive load. Such guidance helps ensure GenAI adoption can be "human-centric," maintaining workforce wellbeing as technology changes. Knowing that there is a regulatory expectation of safe and inclusive implementation gives organisations a roadmap to follow (and employees assurance their welfare is being considered), which in turn facilitates more sustainable adoption.

Considerations for Employers

Despite strong bottom-up interest, several human-level frictions can slow GenAI uptake within organisations. First, **trust calibration is hard**: opaque errors, biased or inexplicable outputs, and "polished but wrong" answers fuel hesitation and misuse. Messeri and Crockett (2024) show that fluent and convincing AI outputs can create an illusion of understanding – employees mistake the model's confident prose for their own comprehension, become overconfident, skip verification, and miss subtle errors. For example, a finance employee might use GenAI to generate an analysis report – because the report looks polished, they might assume they fully grasp the analysis, when in reality they have not engaged deeply with the content. This illusion can undermine skill development; the employee isn't actually learning, and if the AI made subtle mistakes, they may not notice. In the long run, this could erode the employee's expertise (a critical concern in knowledge professions). These risks intensify when models inherit training data stereotypes – both direction and representativeness – that distort judgments of trustworthiness and marginalise subpopulations (Nicolas & Caliskan, 2024; Simmons & Hare, 2023). When a GenAI tool then produces a factual error, a biased remark, or an opaque recommendation in high-stakes contexts, users naturally grow reluctant to rely on it.

Second, a **skills gap persists**: workers need a certain level of digital literacy and training to effectively use AI tools – for instance, knowing how to craft good prompts, how to verify AI outputs, and how to secure data when using the tool. If employees lack these skills, they can feel overwhelmed (a form of cognitive overload or "technostress"). Constantly learning new software or adapting to AI-driven changes in processes can lead to fatigue and even emotional exhaustion. In interviews, some workers report feeling left behind by rapid technological change, which can manifest as resistance or disengagement (Brougham & Haar, 2018). Paradoxically, employees are simultaneously afraid of not adopting AI (and thus becoming obsolete) and afraid of adopting it incorrectly. Australia's 2025 EY AI Workforce Blueprint shows the same push-pull: 66% of computer-based workers want their employer to provide more formal AI training, yet only 35% have received any; at the same time 54% worry about job losses in their sector and 72% fear breaching data/regulatory requirements – evidence that enthusiasm for AI coexists with anxiety about its risks (EY, 2025). Such ambivalence can cause decision paralysis – employees neither fully reject AI nor fully embrace it, which results in minimal, ineffective use.

Ambivalence leads to nominal adoption with negligible impact – tools are trialled but not integrated into core workflows, so usage occurs without meaningful performance gains.

Third, **accountability and role clarity are often ambiguous**. If a decision is made with AI assistance, who is responsible if it goes wrong – the employee, their manager, or the tool vendor? Hence, workers might hesitate to use GenAI for important decisions if they fear being blamed for the AI's errors, especially in high-stakes fields like healthcare or finance. This lack of clarity can reduce adoption to only trivial uses, leaving transformative applications on the table.

Fourth, **role threat and change fatigue dampen experimentation** when AI is framed as replacement rather than augmentation (Grover, Kar, & Dwivedi, 2022); many workers feel the skills pressure while also fearing redundancy (EY, 2024). Adopting GenAI can feel like loss of the old ways of working; some employees go through stages akin to the Kübler-Ross grief cycle (denial of the AI's value, anger or frustration at being forced to change, bargaining by using AI only partially, a downturn in morale (depression) as they struggle, and – hopefully – eventual acceptance) (Du Plessis & Smuts, 2021). Yet, behaviour change is rarely linear. As Prochaska and Norcross (2018) write,

> "people progress from contemplation to preparation to action to maintenance, but many individuals will relapse… Some relapsers feel like failures – embarrassed, ashamed, and guilty. These individuals become demoralized and resist thinking about behavior change. As a result, they return to the precontemplation stage and can remain there for various periods of time… Fortunately, the research indicates that the vast majority of relapsers – 85% of self-changers, for example – recycle back to the contemplation or preparation stage. They consider plans for their next action attempt, while trying to learn from their recent efforts. The spiral pattern suggests that most relapsers do not revolve endlessly in circles and that they do not regress all the way back to where they began. Instead, each time relapsers recycle through the stages, they potentially learn from their mistakes and try something different the next time around" (p. 422-423).

Recognising this, organisations can normalise that GenAI adoption will also be non-linear – marked by cycles of trial, setback, and recalibration – and build supports that treat each loop as learning rather than failure.

Finally, **digital inclusion matters**: almost a quarter of Australians are digitally excluded, meaning they lack affordable access, adequate digital skills, or reliable connectivity needed to fully participate in online work, learning, and services – especially older people (aged 65+), Indigenous Australians (particularly in remote areas), lower income households (<$50k), and people with disabilities (Thomas et al., 2023). These cohorts face compounded barriers – access/affordability (devices, data), skills and confidence gaps, and a higher risk that GenAI misinterprets dialects, cultural references, or accessibility needs. JSA's (2025B) reports a ~30-point Australian Digital Inclusion Index (ADII) gap between highest- and lowest-income groups, and focus-group evidence that cost is a primary barrier for older Australians – hence the need for targeted support alongside clear regulation and employer-provided training. But foundations also extend beyond connectivity. JSA (2025C) highlights intertwined barriers – financial constraints, lack of digital connectivity, childcare and transport gaps – and underscores the centrality of foundational language, literacy, numeracy (for prompting and interpreting outputs) and digital skills for participation, awareness of biases, GenAI hallucinations, and data privacy/security. Finally, a gender lens also matters. JSA (2025D) shows entrenched gender segregation in training pipelines and outcomes, shaped by social and cultural norms, with intersectional snapshots for First Nations, CALD and disability cohorts highlighting that skills mismatches and under-utilisation of women's skills persists. Therefore, reinforcing structural barriers that can carry into AI-era upskilling and deployment.

How employers can support employees to adopt GenAI

Treat trust as ongoing work, not a one-off launch. Explain the "why and how" of the system in user language (what it does, what it can't do, and why it made this suggestion). Show uncertainty cues (e.g., "low confidence" labels or highlight sections to double-check). Give provenance (where the information came from) and recourse paths (what a user should do if something looks wrong – who to contact, how to report, how to correct). For higher-risk decisions, keep a human in the loop (a person must review before anything goes to customers/regulators). Replace one-time "data fixes" with ongoing bias checks and oversight that you run on a schedule and update as use grows (Schwartz et al., 2022).

To reduce technostress, scaffold skills in small, practical steps. Start with short, hands-on clinics that use real tasks from your workflow, so people see immediate value. Maintain a shared prompt library of tried-and-tested patterns for your top use cases, and pair this with a simple verification checklist: trace sources (where the information came from), cross-check with another tool or known reference, and get a domain expert to sign off when needed. Keep support easy to access – for example, run a 30-minute weekly drop-in ("office hours") for quick questions and demos. Finally, recognise participation (internal badges or professional-development credits) so learning is visible and rewarded. This approach makes adoption gradual, low-pressure, and aligned with day-to-day work.

Use motivational interviewing (MI)-style coaching (Miller & Rollnick, 2013) rather than "catastrophic interviewing." MI invites change talk via OARS – Open questions, Affirmations, Reflective listening, and Summaries – so employees articulate their own reasons to try AI safely. Think: "What task feels safe to try first?", "On a scale of 1–10, how confident are you? Why not lower?", "What would move you up one point?" In contrast, fear-based scripts ("AI is coming – adapt or be left behind") trigger threat and psychological reactance, reducing experimentation and learning (Brehm & Brehm, 1981). MI preserves autonomy, builds competence through small wins, and normalises calibration ("When would you override the model?"), which directly supports appropriate trust.

Back this with clear governance, aligned to Australia's Voluntary AI Safety Standard (VAISS).[4] Publish a simple RACI for AI-assisted decisions and keep lightweight decision logs (who/what/why/inputs). Run safe-to-fail pilots with kill-switches and post-incident reviews, and expose brief model/data cards and change logs so users can see how systems evolve. Make risk visible by triaging tasks into a tiering matrix – Assist (draft) → Approve (review) → Autonomy (execute) – and bind each tier to the right controls (mandatory human approval, uncertainty thresholds, escalation paths). Match that with privacy and data guardrails: clarify allowed use (what can/can't be pasted), enable automatic redaction/Data Loss Prevention (DLP), set retention defaults, and provide a sandbox for sensitive experimentation.

Shape the culture so adoption signals augmentation, not replacement. Resource internal champions and communities of practice whose visible, appropriate use creates social proof and shifts descriptive/injunctive norms toward everyday, responsible use. Build equity by design with accessible interfaces, culturally appropriate training, and targeted enablement for lower-inclusion cohorts (device/data stipends, mobile-first options, place-based sessions). Measure equity with subgroup performance and reliance metrics and adjust supports accordingly, following established guides like Web Content Accessibility Guidelines (WCAG) and the Australian Government Digital Service Standard for inclusive, evidence-led service design.

Finally, measure trust calibration – not just usage. Track acceptance and override rates by task-risk tier, error cost versus reliance, adherence to uncertainty guidance, and subgroup gaps; feed those signals back into the prompt library, training, and guardrails (Lee & See, 2004; Hoff & Bashir, 2015; Schwartz et al., 2022). And support the human change curve: apply the Transtheoretical Model (Prochaska & Norcross, 2018) to tailor supports by stage (awareness → preparation → action → maintenance), and normalise a non-linear, affective journey drawing on Tonkin's "grow around" model, which shows that distress often resurfaces at triggers (e.g., anniversaries) even as one's life expands around it; the feeling doesn't necessarily shrink so much as the person's capacity and world grow, reducing its dominance (Tonkin, 1996). Use empathetic communication and MI-consistent manager coaching (curiosity, reflection, gentle goal setting), not catastrophising, to validate those oscillations and encourage "next small step" progress (Miller & Rollnick, 2013)

***Employer role is pivotal** – framing AI as augmentation, providing hands-on skills support, and embedding transparency and guardrails are critical for sustained adoption.*

Organisational/Employers

Institutional Pressures and Legitimacy

Employers largely set the conditions under which workers adopt (or avoid) GenAI: they decide what is allowed/permissible, provide access and training, and signal what "good" looks like. Institutional Theory helps explain why those choices often converge. Organisations respond to coercive pressures (law, regulation), normative pressures (professional standards, ethics and shared training), and mimetic pressures (copying perceived leaders in uncertainty) (DiMaggio & Powell, 1983). These field-level forces shape policies, governance, and incentives that can either enable safe experimentation or push use to the margins. Scholarship also notes we often over-emphasise mimicry relative to coercive and normative forces – useful context when interpreting "follow-the-leader" AI rollouts (Mizruchi & Fein, 1999).

Employers set the rules of GenAI use, shaped by coercive, normative, and mimetic pressures that reward legitimacy as much as performance.

Balancing Legitimacy and Innovation

Many firms balance "looking legitimate" with "learning quickly." They may signal compliance publicly while running internal sandboxes to learn. Early adopters then become reference points that speed diffusion across an industry (Singh, 2024). Because GenAI strains familiar playbooks, organisations are building new oversight models to reconcile speed with accountability. Day-to-day uptake among workers ultimately reflects what the employer legitimises, resources, and rewards – not just the intrinsic quality of the tools.

Firms balance legitimacy and innovation by publicly stressing compliance while experimenting internally, with early adopters setting industry reference points.

Why firms start to look alike

Put differently, field-level pressures set the legitimacy perimeter. Put simply, the wider industry sets the "rules of what's acceptable." Within those boundaries, companies then adopt GenAI

[4] For VAISS alignment: accountability & governance; risk & impact assessment; data governance & security; transparency; contestability/recourse; stakeholder inclusion.

for practical reasons – to work faster, cheaper, or better – which is what turns the idea into real budgets, small pilots, and then rollouts. In mature fields (e.g., health, finance), firms grow more alike not always because that is necessarily most efficient, but because "rule-makers" (e.g., the state, professions) set templates for what counts as legitimate. bring templates, tools, and "lessons learned" from prior roles, and communities of practice turn formal rules into everyday routines. As a result, practices spread through hiring, vendor networks, and ordinary social exchange – even with copyrights, Non-Disclosure Agreements (NDAs), and other Intellectual Property (IP) walls. This is a human/social analogue to how frontier LLMs appear to develop "emergent" capabilities[5] by aggregating large, varied corpora (raising similar copyright/IP questions). Under uncertainty, organisations rationally copy peers and checklists (e.g., setting up an AI steering and/or ethics committee, adding the same risk gates to the software development lifecycle, hiring a similar "AI lead" or consultants). The aggregate effect is sameness in structures, cultures, and outputs – whether or not productivity actually improves (DiMaggio & Powell, 1983).

Organisations converge toward sameness through regulatory templates and knowledge flows, leading to homogeneity regardless of efficiency gains.

The Case for Bottom-Up Experimentation

None of this argues against regulations or professional standards – especially in high-reliability domains like health and finance. But integrating GenAI into real work is context-specific. It works best through bottom-up trials and experimentation, co-design with frontline teams, and iterative 'safe-to-fail' pilots that adapt roles, processes, and skills – a pattern echoed in HTI's worker studies, which found poor communication, limited or token consultation, and little transparency about when/how AI is used (often without explicit consent), and therefore called for genuine co-design, clear training, transparency and consent processes, human oversight with the ability to override, and regular checks/assurance (HTI, 2024). That bottom-up bias isn't a managerial fad; it follows from where the relevant knowledge actually lives – i.e., close to the work. Hayek's core insight is that the knowledge needed for effective action is dispersed, local, and often tacit/inarticulable – i.e., "knowledge of the particular circumstances of time and place" that cannot be centralised without stripping away the very differences that matter for decisions (Hayek, 1945, pp. 521-525). Much of what workers rely on is know-how acquired on the job; some can be written down, but it is partial and costly to extract (Howes, 2020). Put simply: formal rules never capture the full "feel for the game," centralising knowledge will always lag the lived subtleties frontline people use in their day to day.

Tacit Knowledge and the Limits of Codification

Turning tacit know-how into explicit artefacts is a real bottleneck. Classic knowledge elicitation methods (e.g., interviews, protocol analysis, repertory grids) help, but experts struggle to verbalise what guides their actions, and elicitors often extract only what can be said

[5] Work on "emergence" mixes positive evidence with measurement caveats. Wei et al. (2022) define an *emergent ability* as absent in smaller models but present once scale (and data) cross a threshold, and they note data×scale interactions – for example, certain multilingual capabilities appeared only in PaLM at ≳62B parameters and not in similarly sized LLaMA models, implying dependence on pretraining corpus as well as size. Webb, Holyoak, and Lu (2023) provide a concrete capability case: GPT-3 solved zero-shot analogy tasks (matrix, letter-string, verbal, story) at or above average human levels, with error patterns tracking human difficulty – evidence that some abstract reasoning may indeed "appear" at scale. Countering step-change narratives, Schaeffer, Miranda, and Koyejo (2023) show many "phase transitions" vanish when discontinuous, thresholded metrics are replaced by continuous ones; under continuous scoring, performance increases smoothly with scale on arithmetic-style tasks, suggesting metric choice can manufacture apparent emergence. A recent survey (Berti, Giorgi & Kasneci, 2025) integrates these strands, arguing that whether "emergence" is observed depends not only on model scale but also on evaluation design and usage factors (objective/loss, prompting, and even quantization), and cautions against treating emergence as a single phenomenon.

– not what actually drives performance (Gavrilova & Andreeva, 2012; Nisbett & Wilson, 1977). Even when captured, representations must be refined in context with users – what knowledge engineering treats as concept model cycling and validation (Okafor & Osuagwu, 2006).

Bounded Rationality and "Good Enough" Solutions

Given bounded rationality (i.e., people and organisations decide with limited time, information, and cognitive capacity), organisations satisfice (i.e., pick a workable option that meets a threshold/aspiration level rather than searching for the absolute best): they adopt workable codifications and procedures that fit cognitive limits and task constraints, then improve them through feedback (Simon, 1955, 1989, 1990; also discussed in Bickley & Torgler, 2023B). In practice, that means breaking big goals into sub-goals, distributing decisions across specialists, and iterating. In Simon's (1996) design terms, our procedures, prompts, and guardrails are artifacts – i.e., interfaces between inner routines and capabilities, and the outer task environment. Many different inner designs can achieve the same function; we typically settle for a good enough fit and refine it over time. Bottom line: GenAI adoption is as much a social, iterative knowledge-shaping process as a technical one – start with workable approximations, embed them in real workflows, and update as communities learn.

Codified vs. Tacit Knowledge in Practice – why it matters

Effective programmes respect the boundary between what can be codified and what remains tacit – i.e., so we document the first and teach/coach the second. Clarke's (2021) metatheoretic synthesis draws a clear boundary: codified knowledge is what can be recorded and transmitted in language (e.g., documents, data), whereas tacit knowledge is embedded in skilled action, identity, and relationships between people/business units – i.e., context-bound and harder to move across settings. In practice, elicitation hits limits precisely where work relies on embodied skill, local cues, and shared repertoires, where transfer depends more on apprenticeship-like interactions than on documents (White, 1988). To narrow (but not eliminate) the gap, interactive tools treat experts as co-modellers: they elicit constructs, translate them into attributed statements, and map consensus/conflict (e.g., FOCUS, PCA, group compare) to reveal stable structure without pretending to fully extract tacit know-how (Gaines & Shaw, 2002). Diverse teams also help: prior domain knowledge improves communication but can cause fixation in ill-structured problems – so heterogeneous teams and methodical prompting reduce blind spots (Hadar, Soffer & Kenzi, 2012). The realistic stance is iterative satisficing: codify good enough artefacts, validate in use, and refine – while growing tacit expertise via coaching, pairing, and practice. This is what turns guidance into usable Standard Operating Procedures (SOPs) and on-the-job learning, rather than shelfware.

Tacit, situated know-how resists full codification; adoption depends on iterative, 'good enough' artefacts plus apprenticeship-like transfer.

Shifting Skill Demands

As routine, easily automated tasks are absorbed by new tools, value shifts toward skills that are hard to substitute (tacit judgement, creativity, complex problem-solving) and skills that complement AI (AI auditing/assurance, human-AI teaming, cybersecurity, data centre/MLOps). Sectors with high elasticity of demand (healthcare, green tech, advanced research) will expand as AI reduces costs, while specialised STEM domains (quantum, robotics, advanced medicine) face inelastic supply – long training pipelines limit rapid scaling. In economic terms: low substitution, high complementarity, high demand elasticity, and inelastic supply drive where value accrues. Many of these skills require early and ongoing education, especially in STEM, as the volume and velocity of knowledge keep rising.

Strategic Drivers of Adoption

GenAI adoption is propelled by a blend of strategic goals, competitive pressures, and operational/day-to-day realities. As discussed above, a primary driver is the promise of efficiency: automating repetitive knowledge tasks (e.g., summarising documents, drafting reports, handling routine queries) and augmenting workers to deliver higher-quality outputs faster. For firms in finance, technology, and professional services, these gains can be seen as a means to deliver services more quickly and at scale, which can translate into competitive advantage and cost savings. Executive sentiment reflects this pull – surveys show near-universal intention among C-suites to expand AI use. For example, Slack's June 2024 Workforce Index (a survey of more than 10,000 desk workers globally) found 96% of executives wanted to increase AI use in their operations (Sarasohn, 2024) and 99% planning some AI investment[6] in the coming years. Such top-down enthusiasm creates organisational pull and momentum – budgets and teams get earmarked and allocated to AI projects, owners and leaders appoint AI champions to run with it, and pilot programs receive sponsorship from senior leadership.

Skills rooted in tacit knowledge, creativity, and human-AI complementarity gain value, while routine tasks decline.

For illustration only, consider two fictional cases that show how executive intent turns into action without deep AI expertise. In a mid-size finance firm, the CFO sets aside a $150k "AI Enablement" budget and the COO names an AI Program Lead (half-time/0.5 FTE) to run a six-week Client Reporting pilot. Ten analysts start monthly fund letters from an AI draft, with mandatory human review before anything goes to clients. Success is defined upfront: at least a 25% reduction in turnaround time, no more than a 5% error rate after review, and average user confidence of 4 out of 5. A short, fortnightly steering huddle (COO, Risk, IT Security) clears roadblocks like data access and tool approval, and HR gives each analyst one hour per week for hands-on clinics and maintains a one-page prompt library for continuous learning and documentation purposes. If targets are met, a pre-approved second phase budget automatically unlocks to integrate the workflow into the document system, with portability built in (e.g., exportable prompts/data and a backup vendor).

In a small professional services firm (fewer than 10 staff), the owner sponsors a four-week proposal drafting pilot. One partner, the office manager, and two consultants trial an AI first draft tool with a simple success target: a 20% faster turnaround, with mandatory human review before client use. The owner is the visible sponsor, whilst the office manager acts as the AI champion who runs a 30-minute weekly check-in, maintains a one-page prompt library, and logs a few basic metrics (e.g., what share of proposals start from an AI draft, why drafts are accepted or overridden, time saved, and any shift in win rates). Costs are minimal (free/low tier licence and roughly an hour per person per week). If the pilot hits its targets, the firm scales to all proposals under lightweight guardrails: human-in-the-loop for external material, a shared folder for exportable prompts and outputs, and a month-to-month contract that includes a 30-day deprecation/data egress clause.

These fictional examples show the essentials for leaders: set a clear scope and success criteria, give someone ownership, remove simple blockers on cadence, measure a few practical outcomes, and only then scale – with portability and oversight baked in.

---

[6] See also: https://slack.com/intl/en-au/blog/news/the-fall-2024-workforce-index-shows-executives-and-employees-investing-in-ai-but-uncertainty-holding-back-adoption.

Competitive Bandwagon Effects

Competitive dynamics reinforce this momentum. Once peers or rivals demonstrate productivity improvements, even risk averse organisations feel pressure to experiment so they do not fall behind. Rapid uptake in marketing, software, and customer service has created a bandwagon that is pulling in adjacent sectors. For example, Commonwealth Bank has moved beyond basic chatbots to deploy Apate.ai – thousands of conversational AI "honeypots" that tie up scammers and feed near realtime intelligence into fraud defences – building on earlier machine learning controls on its digital platforms.[7] NAB runs an AI-powered "Customer Brain" to personalise service across channels, framing GenAI adoption as customer experience uplift rather than pure cost take-out.[8] In insurance, IAG is scaling a GenAI claims assistant ("CASI") in its CGU brand to speed assessment and broker service.[9] Together, these visible moves in banking and fintech have signalled to other regulated sectors that controlled, assurance-heavy pilots are feasible. Healthcare offers a case in point: St Vincent's Hospital Melbourne is running a 12-month pilot using an AI system to verify IV drug identity and concentration in its compounding pharmacy (safety layer for oncology medications), while several services are trialling AI medical scribes and issuing guidance for safe rollout.[10] Regulators are also updating health AI rules – the TGA's September 2025 guidance sets expectations for when AI counts as a medical device and how it should be governed.[11]

This bandwagon effect is amplified by the availability of enterprise-ready GenAI solutions – packaged offerings from cloud vendors and startups (e.g. Microsoft's Azure OpenAI Services, Google's Vertex AI, or domain-specific AI software) that integrate more easily into existing IT environments and come with assurances around privacy, compliance, and admin controls. These turnkey products lower entry barriers and reduce the need to build from scratch, allowing firms to subscribe to or license solutions that integrate directly into their ICT landscape.

Culture and Structure as Enablers

Culture and structure also matter. Companies with a strong innovation culture or agile structures are more ready to pilot new tech. In such environments, management actively encourages experimentation, tolerates prudent risk, and dedicate resources (time, training, budget) for bottom-up learning and experimentation by employees. This positive climate often leads to higher and sustained uptake. By contrast, hierarchical or risk-averse cultures often stall experimentation unless adoption is explicitly mandated. Flexible governance mechanisms such as AI committees or taskforces serve a dual role here: managing risk while signalling top-down support. Clear alignment with business objectives (i.e., a by-product of effective governance and accountability structures) drives further uptake: where GenAI use cases map directly onto existing KPIs (e.g., reducing call-centre handling time[12]), adoption is easier to justify and measure.

---

[7] Source: https://www.commbank.com.au/articles/newsroom/2025/06/apate-ai.html; https://www.commbank.com.au/articles/newsroom/2022/07/scams-fraud-artificial-intelligence.html.

[8] Source: https://news.nab.com.au/news/the-brain-behind-better-banking-how-nabs-ai-is-making-banking-more-human.

[9] Source: https://www.insurancebusinessmag.com/au/news/claims/iag-lifts-lid-on-casi-a-new-ai-claims-assistant-522369.aspx.

[10] Source: https://www.healthcareitnews.com/news/anz/st-vincents-hospital-melbourne-pilots-ai-driven-system-ensure-iv-dosing-accuracy.

[11] Source: https://www.tga.gov.au/how-we-regulate/manufacturing/manufacture-medical-device/manufacture-specific-types-medical-devices/artificial-intelligence-ai-and-medical-device-software.

[12] For example, if a call centre's goal is to reduce average handling time, deploying a GenAI assistant to help agents respond faster has a direct, measurable benefit, making it an attractive project for management (as seen in early deployments of AI at contact centers yielding 14% productivity boosts; Brynjolfsson et al., 2025

Standards and Guardrails

Many of the organisational drivers already described are channelled through the broader regulatory and standards environment. In institutional terms, this is where coercive and normative pressures crystallise: government frameworks and professional bodies set the outer boundary of what counts as "legitimate" GenAI use, and firms align accordingly. In Australia, this has taken the form of the VAISS and Responsible AI benchmarking tools (e.g., Fifth Quadrant's Responsible AI Index and Self-Assessment Tool). These instruments don't just provide risk assurance; they offer open-sourced templates that organisations can copy and benchmark against, reinforcing mimetic dynamics as well. The result is both a floor (minimum governance guardrails) and a ladder (progressive maturity levels) that companies aspire to climb – 12% of firms now rated "Leading," 23% "Implementing," and nearly half "Developing" responsible AI practices (Fifth Quadrant & NAIC, 2025).

In practice, voluntary guardrails and open frameworks can unlock adoption by making skills portable and transitions smoother. For example, mapping roles to the Skills Framework for the Information Age (SFIA), issuing micro-credentials for GenAI competencies, and formalising Recognition of Prior Learning (RPL) allow staff moving from non-AI roles – including those in SMEs, microfirms, or regional/remote contexts – to step into human-AI teaming roles with minimal friction (Akkermans et al., 2024). That said, VAISS is not always user-friendly in day-to-day practice and is currently being refreshed. To close this usability gap, it would be useful to provide a plain-English companion with a quick-start guide, role-based and mapped checklists, and example evidence or case studies, with materials kept version-aligned as the standard evolves. These directions are consistent with FSO's ongoing work and highlight the importance of developing accessible, standards-aligned resources that support employers and workers to engage confidently with GenAI.

Adoption is driven by efficiency, competition, and guardrails, but hindered by skills gaps, SME disadvantages, costs, vendor lock-in, and regulatory uncertainty.

From Guardrails to Capability Building

The benefit of these voluntary guardrails is their flexible application. Frameworks like VAISS, SFIA, or DigComp 2.2 are deliberately open-sourced, allowing firms to choose, integrate, and adapt them in ways that fit their particular structures and cultures, while still aligning with evolving regulatory and legal requirements. This flexibility is crucial because policy and regulation often lag rapid technological change. By adopting guardrails as living frameworks rather than rigid checklists, organisations can maintain legitimacy and comparability without being overwhelmed by bureaucracy, and build greater resilience in an increasingly complex and fast-changing sociotechnical environment. Importantly, while codified standards provide a baseline of assurance and a shared vocabulary, their effectiveness depends on pairing them with bottom-up experimentation and tacit, situated knowledge – ensuring that formal structures remain adaptive rather than constraining.

Yet, even with voluntary guardrails and open frameworks providing a flexible path for responsible adoption, organisations still encounter substantial barriers when moving from principle to practice. The gap mirrors what has long been observed in sustainability and Corporate Social Responsibility (CSR, i.e., a management approach where businesses integrate social and environmental concerns into their operations and interactions with stakeholders, aiming to contribute positively to society and the environment rather than only maximising shareholder profits): the firms with well-developed CSR processes for stakeholder engagement and long-term orientation consistently outperform their peers (Eccles et al., 2014), while those treating sustainability as a surface-level compliance exercise often fail to unlock real strategic benefits. Similarly, in the GenAI context, voluntary standards such as VAISS, SFIA, or

DigComp 2.2 can create legitimacy and comparability, but their practical value depends on whether organisations build the complementary proficiencies of stakeholder integration, continuous learning, and organisational innovation (Sharma & Vredenburg, 1998). When these traits are present, standards become platforms for capability building; when absent, they risk ossifying into box-ticking exercises.

This tension is heightened in environments where technology advances faster than policy. Just as proactive sustainability strategies have been linked to emergent capabilities in natural-resource-based strategic management (Hart, 1995), so too can deliberate engagement with AI guardrails foster new forms of resilience and adaptation. But such benefits require sustained managerial attention: CEO focus has long been shown to shape the trajectory of organisational innovation and long-term competitiveness (Yadav et al., 2007; Henderson & Serafeim, 2020). In the GenAI context, leaders who frame voluntary frameworks as enablers of learning – rather than constraints – position their organisations to adapt more effectively in a fast-changing sociotechnical environment.

Barriers: Skills, Equity, and SME Disadvantages

Absent this proactive orientation, however, familiar barriers quickly dominate. Despite high strategic interest, organisations report persistent skills shortages, with IBM's Global AI Adoption Index 2023[13] finding that one-third of enterprises cite limited AI expertise as the top barrier to implementation. This shortage spans technical roles (data scientists, ML engineers) and managerial capabilities (interpreting and governing AI outputs). Importantly, organisational barriers tend to be higher for smaller businesses and certain sectors, exacerbating digital inequality in the business community. Large corporations can invest in AI labs, hire consultants, and absorb failures, yet a small or micro business or a non-profit often cannot. This is reflected in the Responsible AI Index (Fifth Quadrant & NAIC, 2025): large organisations were much more likely to be in the "Leading" or "Implementing" categories, whereas most small and mid-sized firms were still in "Developing" or "Emerging" stages of AI maturity (only 12% of all orgs are Leading, while 48% Developing and 17% Emerging are over-represented by smaller enterprises). SME-specific surveys show mixed adoption – 35% of Australian SMEs reported using some form of AI as of late 2024, but 42% had no plans to implement AI, and 23% were unaware of its potential applications (Fifth Quadrant, 2025). Many small business owners simply lack the time, funding, or support to explore AI opportunities, and funding constraints mean they cannot spare resources for experimental projects with uncertain ROI. Sector disparities compound the gap: highly regulated industries such as healthcare and finance remain cautious, while adoption in agriculture is especially low (~6% of agri SMEs by Q3 2024, versus ~45% in health and education), reflecting infrastructure challenges and sector-specific barriers. Addressing this requires low-risk, time-efficient entry points – such as pre-scoped "safe-to-fail" pilots, shared AI advisory hubs (like the government's four AI Adopt Centres that provide hands-on guidance and support for SMEs)[14], or subsidised training vouchers that let SMEs test use cases without major upfront cost or technical overhead.

SMEs and micro-businesses are therefore particularly disadvantaged, echoing earlier findings that resource constraints exacerbate digital divides, much as small firms in sustainability research often struggled to keep pace with larger, resource-rich competitors. Equity gaps also manifest within organisations: workers in digitally mature roles often benefit first, while frontline, regional, or CALD workers may face lower access, reinforcing inequalities even as headline adoption metrics improve (Thomas et al., 2023; JSA, 2025A; JSA, 2025C). Moreover,

---

[13] Available at: https://filecache.mediaroom.com/mr5mr_ibmspgi/179414/download/IBM%20Global%20AI%20Adoption%20Index%20Report%20Dec.%202023.pdf.

[14] Source: https://www.industry.gov.au/news/be-part-ai-revolution-ai-adopt-centres.

even if inclusive technical expertise is secured, barriers also arise when employees perceive GenAI systems as opaque, unreliable, or threatening to their roles. Trust and transparency are therefore as critical as technical skill in shaping adoption (Gillespie et al., 2025; also refer to previous discussion on pp. 11–13 above).

SMEs face sharper barriers: lacking time, funding, and expertise, many cannot experiment with GenAI, widening digital divides across sectors and reinforcing inequities within workforces.

Costs, ROI, and Vendor Dependence

Relatedly, high implementation costs and uncertain ROI remain deterrents. As in operations management research, where Big Data Analytics and related Technologies (BDAT) systems are argued to enhance the very capabilities for change required to build sustainable organisations (Eccles et al., 2012; Garvin & Kimbleton, 2021; Bickley, Macintyre & Torgler, 2025), GenAI requires upfront investment in data preparation, process redesign, and training before such adaptive capacity is realised. Without clear, measurable benefits – whether in cost savings, revenue gains, or improved stakeholder outcomes – leaders hesitate to expand pilots into full deployments (Challapally et al., 2025). This mirrors sustainability cases where firms that failed to embed innovation into operational routines struggled to capture long-term value. And these difficulties are compounded by dependence on a small set of global AI vendors, raising concerns about lock-in, cost escalations, and strategic vulnerability – paralleling critiques in sustainability where firms struggled when essential resources were concentrated in few providers. Thus, underlining the importance of cultivating a vibrant and dynamic startup/SME ecosystem that can provide diverse, competitive, and context-specific GenAI solutions.

High costs, vendor lock-in, and regulatory uncertainty deter adoption, as firms struggle to balance safeguards with the experimentation needed for long-term competitiveness.

Governance and Regulatory Uncertainty

Finally, data governance and regulatory uncertainty compound adoption barriers. Companies fear legal liability or reputational risk from privacy breaches or discriminatory outputs, often responding by imposing restrictive internal controls.[15] While such caution is rational, it can also slow the organisational learning and experimentation that drive adoption, creating a "catch-up" dynamic. Moreover, quantifying and attributing GenAI benefits remains difficult, much like the long-debated ROI of sustainability initiatives: while long-term gains may be significant, short-term measurement challenges can stall managerial confidence.[16] As with sustainability and ecological responsiveness, the challenge lies in striking a balance: protecting stakeholders through prudent safeguards while still fostering the higher-order learning and innovation capacity that long-term competitiveness demands.

---

[15] For example, in April 2023 Samsung barred staff from using public chatbots after employees accidentally pasted sensitive code into ChatGPT, and subsequently introduced stricter internal guardrails (Wilkinson, 2023). Financial firms took similar steps early on: Wall Street banks including JPMorgan and Bank of America restricted employee access to ChatGPT citing compliance and data-leak risk (Bushard, 2023).

[16] To address this, organisations are beginning to pair business KPIs with system-level evaluation frameworks such as Ragas (Retrieval Augmented Generation Assessment), which blend operational and technical metrics – tracking accuracy, relevance, latency, and user satisfaction alongside productivity, quality, and cost indicators (Es et al., 2024). This dual layered approach allows firms to assess GenAI performance both as an AI system and as a business capability, linking measurable outputs to tangible value.

Systemic factors

In response to FSO's feedback, the broader analysis of systemic factors – particularly those relating to migration, macroeconomic settings, and cross-sector dynamics – has been removed from this report and will be repurposed, in partnership with FSO, to inform the forthcoming National AI Capability Plan being prepared by the National AI Centre and the Department of Industry, Science and Research later this year. The summary below distils the core insights most relevant to GenAI adoption in the FTB sectors. For employers (especially SMEs), the practical stance is: you do not control the system, but you can align to it by adopting recognised guardrails, choosing portable tools, and plugging into shared/established supports.

GenAI adoption plays out within systems: standards and regulation, skills pipelines, market structures, computational infrastructures, and the broader culture of trust. Where access or trust is weak, exclusion often manifests as disengagement, distrust, or misuse. Therefore, slowing uptake and widening the "GenAI divide." Insights from adjacent fields reinforce this systems view: research on structural exclusion shows that when opportunity structures narrow, maladaptive behaviours can emerge and intensify through digital platforms (e.g., patterns documented in criminology and online subcultures) (Beatton, Kidd, Machin, & Sarkar, 2018; Magee, 2021; Blake & Brooks, 2023; cf. status threat dynamics in Wilson & Daly, 1985). The analogy is not about equating workplace adoption with crime but about the mechanism: persistent barriers and weak guardrails can channel technology toward counter-productive use. *Implication for employers:* publish a one-page AI "house rules" and assurance note (what tools are allowed, how data is handled, when human review is required) and use simple accept/override logs because small, visible guardrails build trust even when the wider system is noisy.

Education and reskilling remain the primary corrective. When the returns to participation are raised and pathways are visible, affordable, and portable – e.g., micro-credentials, RPL, and work-based learning – people shift from informal/shadow workarounds to safer, higher-value use (Beatton et al., 2018; Magee, 2021). In workforce terms, equitable access to AI-complementary skillsets reduces exclusion as roles and tasks are reconfigured by GenAI. Yet as Lodge (2025) warns, the higher education sector is losing the very instructional designers and educational technologists needed to support this shift, with resources diverted toward compliance rather than innovation. Thus, weakening the institutional capacity for systemic innovation at the very time GenAI makes it most urgent. *Implication for employers:* treat learning as a standing practice (e.g., 30-60 minutes/week per person), map roles to SFIA/transferable skills, recognise RPL, and use low-cost, credible micro-credentials by partnering with local TAFEs/universities or the National AI Adopt Centres, for example, to co-deliver short clinics rather than building everything in-house.

Culture and norms also shape outcomes. Contexts that emphasise dignity, fairness, and institutional safeguards tend to channel new tools toward pro-social, productive uses; whereas status threat or stigma dynamics can do the opposite (Kashima, Shafa, & Blake, 2025; Blake & Brooks, 2023). Designing for inclusion and psychosocial safety therefore reduces risk and builds legitimacy towards broader community support. *Implication for employers:* require "human-in-the-loop" for external outputs, make participation voluntary but supported, provide opt-out on sensitive tasks, and weight feedback by proximity to work and risk (not loudest voice).

These systemic forces interact across levels – micro (workers and teams), meso (firms, professions, communities), and macro (law, funding, migration, competition, and infrastructures). Micro-level exclusion can scale upward into broader trust deficits, while targeted interventions at each level can redirect trajectories toward inclusive, trusted adoption. *Implication for employers:* join industry groups to share patterns/templates, choose vendors with exportable data/prompts (portability), insert basic contract protections (e.g., data egress, deprecation notice) where practicable, and participate in sector assurance initiatives – i.e., small firms can "borrow scale" by aligning to shared standards and communities.

This condensed overview captures only the most material levers for Australia's FTB sectors – standards and governance (e.g., VAISS and ISO/IEC alignment), skills funding and workforce development, equity and access gaps, and design choices that preserve portability and oversight. These remain central to FSO's mission of supporting employers and workers to engage safely, confidently, and inclusively with GenAI. *Bottom line for employers:* adopt a lightweight, standards-aligned starter pack (house rules, human review policy, simple metrics), invest a little time each week in developing AI skills/capability, and prefer portable, well-governed tools – i.e., practical moves that work even when you cannot change the system.

Conclusion

This report set out to identify the cognitive biases, social influences, and economic incentives shaping workers' decisions to adopt GenAI. Considering the evidence base, its core contribution is a behavioural and systems lens that organisations can act on now. Subsequent phases of this project will translate these insights into practical, plain-language resources co-designed with employers and tested through simulations and workshops, with final outputs refined through stakeholder feedback and accessibility checks for Australia's FTB sectors and priority cohorts.

Adoption as a Lifespan and Systems Challenge

GenAI adoption is not a one-off technical change; it unfolds across work-life domains in which people juggle multiple roles across different "theatres" (e.g., home, community, school, workplace). Viewed through a lifespan/life-space lens, careers unfold as sequences of evolving roles and commitments (Super, 1980),[17] and progress through stage-specific tasks and psychosocial needs (Cron, 1984),[18] so effective governance should enable role transitions, recognise prior learning, and support the accumulation and portability of career capital over time. In this sense, governance frameworks can act as both safety nets and enablers of mobility, ensuring that skills development aligns with evolving technologies while preserving individuals' adaptability and resilience to such changes in work-related content and context

---

[17] Super's (1980) lifespan/life-space view suggests that careers are the *constellation of interacting roles* a person enacts over time across key "theatres." Super outlines nine major roles – child, student, "leisurite" or the role of one engaged in the pursuit of leisure-time activities (including idling), citizen, worker (including unemployed/non-worker), spouse, homemaker, parent, and pensioner (but "[o]thers can be identified, e.g., sibling, worshipper, lover, reformer, criminal, etc.", p. 284) – and four principal theatres – home, community, school, and workplace (and again, others can be identified, "e.g., The Church, The Union, The Club, The Retirement Community … but not everyone enters all theatres", p. 284). Roles carry social expectations, can be simultaneous (e.g., worker/parent), shift with age, and spill over across theatres, producing both enrichment and role conflict. Complementing this, people often draw on narrative "scripts" to decide how to act in ambiguous or novel situations (Schank & Abelson, 1977) – for example, buying a first home, starting a corporate job, navigating disasters/crises, or growing up as an only child or as a neurodivergent person. In Goffman's (1959) dramaturgical terms, individuals adopt and perform roles like actors on a stage, managing impressions across theatres. Narratives thus provide a framework for understanding and responding to new situations, shaping which roles are enacted, how they are performed, and how spillovers, enrichment, and conflicts emerge over the life course. Because careers are sequences of positions (sometimes simultaneous) rather than a single occupation, Super distinguishes career from occupation and highlights the idea of multi-occupational ('serial') careers across one's lifespan.

[18] Cron's (1984) perspective suggests that careers move through stages with distinct concerns, developmental tasks, personal challenges, and psychosocial needs. For example, *Exploration* (learning skills, building self-concept), *Establishment* (producing results, autonomy), *Maintenance* (broadening view, sustaining performance), and *Disengagement* (identity outside work, adjustment). Cron's model links stages to motivation (valences, expectancies, instrumentality), resources (skills, aptitudes), and role perceptions (ambiguity, conflict), which together (and shaped by task characteristics) drive performance, with flow-on effects to job satisfaction and involvement or psychological identification with one's job (see Table 2 and Figure 1 in particular).

(Akkermans et al., 2024) across the lifespan. This dual lens ensures that AI adoption is not treated as a siloed technical fix but as part of an interconnected labour-capital production system, acknowledging the diverse and varied nature of firm resources (e.g., people, partners, processes, economic/financial/digital assets, technological infrastructure, culture, and governance structures) and job demands/pressures (e.g., workload intensity, time pressure, cognitive load from new systems, emotional demands linked to uncertainty, and continuous upskilling requirements) (see e.g., Bakker & Demerouti, 2007; Barney, 1991).

National Standards and Psychosocial Safety Nets

Such systems or big picture thinking also requires a coordinated national effort to realign our training, education, and research ecosystems to better prepare people, organisations/institutions, and policies for the future. In such, the Australian VAISS (including its 10 guardrails and 8 ethical principles) offers a complementary framework for ensuring AI deployment is human-centred, responsible, and socially legitimate. Likewise, Safe Work Australia's (2025) model code of practice for managing psychosocial hazards at work provides a structured framework for identifying, classifying, and controlling psychosocial hazards (such as workload intensity, low job control, poor organisational change management, inadequate recognition, bullying), and underscores the notion that diverse cohorts experience risks differently and require tailored responses.

Supporting Human Change and Emotional Transitions

This change will also be emotional. Employees commonly experience cycles that resemble the Kübler-Ross change curve (denial → anger/frustration → bargaining/limited trial → dip → acceptance), but not linearly; people can 'grow around' their initial anxieties as competence and confidence expand. Tonkin's (1996) 'grow around' model shows that the difficult feeling doesn't necessarily shrink; rather, one's life (skills, routines, supports) enlarges around it, so intensity can spike at triggers yet occupies less of the whole over time – reframing progress as growth, not 'getting over' it. Practically, this aligns with the Transtheoretical Model (TTM): employees move from pre-contemplation to maintenance at different paces, requiring stage-matched supports (awareness building; decisive pros/cons discussions; skills practice; relapse-prevention cues). Motivational interviewing helps resolve ambivalence without coercion, while catastrophic interviewing (for safety-critical scenarios) surfaces worst-case risks early so guardrails are co-designed, not imposed (Prochaska & Norcross, 2018; Miller & Rollnick, 2013).

Governance, therefore, is not only risk control; it is the enabling fabric that connects standards (e.g., VAISS, ISO 9001, ISO/IEC 27001, ISO/IEC 42001), psychosocial safety (Safe Work Australia, 2025), and skills portability (e.g., SFIA, micro-credentials, RPL) to day-to-day adoption. By combining feedback (learn from current frictions) with feedforward (anticipate future risks and skills), adaptive governance supports inclusive mobility, reduces inequities for priority cohorts (CALD, ATSI, older workers, women, LGBTQIA+, neurodiverse, low-SES/low literacy, disability, regional/remote), and sustains legitimacy as technologies and social expectations co-evolve (Arthur & Polak, 2006; Thomas et al., 2023). This aligns directly with the Introduction's emphasis on productivity uplift, sovereign capability, and inclusive participation, and with FSO's mandate to strengthen alternative, skills-first pathways across the FTB sectors.

The same principle (as with earlier general-purpose technologies such as electrification, the internal combustion engine, and the internet) also applies to GenAI and the broader digital shift (Bickley, Macintyre & Torgler, 2025): historically disadvantaged cohorts face uneven barriers, shaped by factors such as trust/confidence, accessibility, prior exposure, and organisational support. Therefore, underlining that employers have a duty not only to provide training, but also to actively design work systems that remove barriers, amplify drivers, and normalise inclusive adoption. This must, in turn, be paired with accessible supports and services (e.g., mentoring, adaptive technologies, flexible work arrangements, and well-being initiatives) that help individuals translate training into practice, and acknowledge the diversity of interacting roles and theatres (see footnote 31) which comprise an individuals' life across their career. Importantly, however, the provision of these supports will not fall solely on employers: a balanced responsibility between government/public sector and private organisations is critical to ensure that the enabling infrastructure, incentives, and safeguards are in place to sustain equitable and widespread adoption of AI and other enabling and emerging technologies like quantum systems (sensors, computers, simulations, communications).

There will always be new technologies to replace old ones, meaning our systems of support must be designed not as one-off interventions, but as adaptive and resilient infrastructures with built-in checks and self-correcting mechanisms that can evolve alongside each new wave of innovation. This implies a growing need for services that support workers through transitions, enhance mobility across roles and sectors, and facilitate rich knowledge transfer and diverse learning pathways that personalise both work and life design. At the same time, care must be taken to avoid the risks of over-centralisation (Hayek, 1945, 1952) in how knowledge, resources, and decision-making are governed, as concentrated control can stifle innovation, entrench inequities, and create opportunities for misuse. In other words, centralising knowledge, resources, or decision-making can act as a honeypot (i.e., drawing in those most willing to exploit, bend, or abuse systems of power). To mitigate this, high-risk domains such as pooled financial, workforce, or research resources require high observability, transparency, and oversight so that concentration of capability is balanced by safeguards against misuse.

Building Adaptive and Inclusive Infrastructures

With technology advancing rapidly, individuals are estimated to change career tracks on average approximately 3 to 7 times over a 45-year career (Lyons, Schweitzer, & Ng, 2015, Table 1)[19] (i.e., making significant pivots or deviations in skills, industries, and occupational identities), lifelong learning is clearly essential. Employers must also recognise the breadth of legitimate learning pathways – not only through accredited systems such as VET and higher education, but also through non-accredited and alternative forms of training that have emerged to fill current gaps. These include short courses, online micro-programs, and vendor-led upskilling initiatives, which can be valuable for rapid capability building but also vary in quality, consistency, and formal recognition. Balancing the flexibility of these non-accredited pathways with the quality assurance and recognition offered by accredited education remains a key challenge for building a coherent, trusted skills ecosystem.

Equally, different cohorts face different constraints (e.g., financial limitations, caring responsibilities, geographic remoteness) and affordances (e.g., digital access, strong professional networks, prior exposure to technology) in accessing formal pathways, making it vital to also value professional development evidenced through past contributions, endorsements, and workplace experiences. Through hands-on, industry-focused training, Higher Education, VET and other alternative pathways can equip learners with practical, job-ready skills that align with workplace needs. Likewise, acknowledging the estimates that up to 90% of new jobs expected to be created over the next decade may require some form of post-secondary or tertiary qualification (JSA, 2024), there is a growing role of public and private

[19] This estimate, based on data collected up to 2012 (Lyons, Ng & Schweitzer, 2012), may represent a lower-bound if technological change continues to accelerate career mobility; however, it is also entirely possible that the figure reflects a relatively stable rate of major career pivots across generations, regardless of technological advancement.

RTOs delivering VET and Higher Education for preparing individuals to enter, transition into, and move across different roles and occupational tracks within the broader workforce.

Emerging Hazards and Next Steps

Two emerging governance hazards require explicit attention: the illusion of understanding (users over-credit AI outputs and under-develop their own competence) and agentic misalignment (frontier models behaving in ways that subvert constraints under stress). Both erode trust, learning/knowledge transfer, and safety; both demand observability, oversight, and staged autonomy rather than blind automation (Messeri & Crockett, 2024; Lynch et al., 2025). Addressing these hazards – and closing the "GenAI divide" highlighted in the Introduction – will be the focus of the next phase of this project that will operationalise the findings of this report by:

- Co-designing and testing **employer support resources** (e.g., playbooks, templates, and case studies) that are refreshable and kept current as models, standards and regulation evolve.
- Developing **evaluation and measurement frameworks** to track behavioural and organisational change in GenAI adoption, including indicators of inclusion, trust, and performance.
- Establishing **pilot partnerships** with FTB sector employers to validate these tools through real-world simulations and feedback loops.

Together, these steps bridge the opportunity outlined at the start of this report (productivity, capability, inclusion) with the behavioural realities on the ground (motivation, trust, readiness). They link the report's practical implications – e.g., a stage-matched behavioural change framework/toolkit, principles of inclusion by design, observability versus over-instrumentation/privacy, and adaptive governance mechanisms – to a forward program of evidence, evaluation, and public good tooling that FSO and partners can scale nationally.

Reference List

Agrawal, A. K., Gans, J. S., & Goldfarb, A. (2023). The turing transformation: Artificial intelligence, intelligence augmentation, and skill premiums. National Bureau of Economic Research (NBER) Working Paper No. 31767. Available at: https://www.nber.org/papers/w31767.

Akkermans, J., da Motta Veiga, S. P., Hirschi, A., & Marciniak, J. (2024). Career transitions across the lifespan: A review and research agenda. *Journal of Vocational Behavior*, *148*, 103957.

Anderljung, M., Barnhart, J., Korinek, A., Leung, J., O'Keefe, C., Whittlestone, J., Avin, S., Brundage, M., Bullock, J., Cass-Beggs, D., Chang, B., Collins, T., Fist, T., Hadfield, G., Hayes, A., Ho, L., Hooker, S., Horvitz, E., Kolt, N., Shuett, J., Shavit, Y., Siddarth, D., Trager, R., & Wolf, K. (2023). Frontier AI regulation: Managing emerging risks to public safety. arXiv preprint arXiv:2307.03718.

Anderljung, M., Hazell, J., & von Knebel, M. (2024). Protecting society from AI misuse: when are restrictions on capabilities warranted?. *AI & SOCIETY*, 1-17.

Arthur, W. B., & Polak, W. (2006). The evolution of technology within a simple computer model. Complexity, 11(5), 23-31.

Australian Energy Council (AEC) (2024). Data Centres and Energy Demand – What's Needed? Available at: https://www.energycouncil.com.au/analysis/data-centres-and-energy-demand-what-s-needed/.

Baker, A.L., & Keebler, J.R. (2017). Factors affecting performance of human-automation teams. In Advances in Human Factors in Robots and Unmanned Systems (p. 331-340). Amsterdam, NL: Springer.

Bakker, A. B., & Demerouti, E. (2007). The job demands-resources model: State of the art. *Journal of Managerial Psychology*, *22*(3), 309-328.

Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, *17*(1), 99-120.

Beatton, D. A. (2007). Trust within teams: the relative importance of ability, benevolence and integrity. Doctoral dissertation, Queensland University of Technology. Available at: https://eprints.qut.edu.au/16651/.

Beatton, T., Kidd, M. P., Machin, S., & Sarkar, D. (2018). Larrikin youth: Crime and Queensland's earning or learning reform. *Labour Economics*, *52*, 149–159.

Berti, L., Giorgi, F., & Kasneci, G. (2025). Emergent abilities in large language models: A survey. arXiv preprint arXiv:2503.05788.

Bickley, S. (2025). Language Models as Tools of/in/by/for Thought. Forthcoming in P. McAlindon & J. Liebowitz (Eds.), *Leveraging AI for Business Innovation*. World Scientific Publishing. Available at: https://arcbita.org/working-papers.

Bickley, S. J., Chan, H. F., Schmidt, S. L., & Torgler, B. (2021). Quantum-sapiens: the quantum bases for human expertise, knowledge, and problem-solving. *Technology Analysis & Strategic Management*, *33*(11), 1290-1302.

Bickley, S. J., Chan, H. F., & Torgler, B. (2022). Artificial intelligence in the field of economics. *Scientometrics*, *127*(4), 2055-2084.

Bickley, S. J., Macintyre, A., & Torgler, B. (2025). Artificial intelligence and big data in sustainable entrepreneurship. *Journal of Economic Surveys*, *39*(1), 103-145.

Bickley, S. J., & Torgler, B. (2023A). Cognitive architectures for artificial intelligence ethics. *AI & SOCIETY*, *38*(2), 501-519.

Bickley, S. J., & Torgler, B. (2023B). Behavioural economics: What have we missed? Exploring 'classical' behavioural economics roots in AI, cognitive psychology and complexity theory. In *Handbook of research methods in behavioural economics* (pp. 32-59). Edward Elgar Publishing.

Blake, K. R., & Brooks, R. C. (2023). *Societies should not ignore their incel problem*. *Trends in Cognitive Sciences, 27*(2), 111–113.

Bousetouane, F. (2025). Agentic systems: A guide to transforming industries with vertical ai agents. arXiv preprint arXiv:2501.00881.

Brehm, S. S., & Brehm, J. W. (2013). Psychological reactance: A theory of freedom and control. Academic Press.

Brougham, D., & Haar, J. (2018). Smart technology, artificial intelligence, robotics, and algorithms (STARA): Employees' perceptions of our future workplace. *Journal of Management & Organization*, *24*(2), 239-257.

Bushard, B. (2023). Workers' ChatGPT Use Restricted At More Banks—Including Goldman, Citigroup. Available at: https://www.forbesmiddleeast.com/innovation/technology/workers-chatgpt-use-restricted-at-more-banksincluding-goldman-citigroup.

Brynjolfsson, E., Li, D., & Raymond, L. (2025). Generative AI at work. *The Quarterly Journal of Economics*, *140*(2), 889-942.

Calvino, F., Reijerink, J., & Samek, L. (2025). The effects of generative AI on productivity, innovation and entrepreneurship. *OECD Artificial Intelligence Papers No. 39*. Available at: https://www.oecd.org/en/publications/the-effects-of-generative-ai-on-productivity-innovation-and-entrepreneurship_b21df222-en.html.

Catherine, S., Suresh, N. V., Mangaiyarkarasi, T., & Jenefa, L. (2025). Unveiling the Enigma of Shadow: Ethical Difficulties in the Field of AI. In *Navigating Data Science: Unleashing the Creative Potential of Artificial Intelligence* (pp. 57-67). Emerald Publishing Limited.

Challapally, A., Pease C., Raskar, R., Chari, P. (2025, July). The GenAI Divide State of AI in Business. Preliminary Findings from AI Implementation Research from Project NANDA, MIT Media Lab. Available at: https://mlq.ai/media/quarterly_decks/v0.1_State_of_AI_in_Business_2025_Report.pdf.

Chen, J.Y., & Barnes, M.J. (2014). Human–agent teaming for multirobot control: A review of human factors issues. *IEEE Transactions on Human-Machine Systems*, 44(1), 13-29.

Clarke, R. (2021). A Platform for a Pragmatic Metatheoretic Model for Information Systems Practice and Research. ACIS 2021 Proceedings. Available at: https://aisel.aisnet.org/acis2021/53.

Commonwealth of Australia (2021). Australia's AI Action Plan, Australian Government Department of Industry, Science, Energy and Resources. Available at: https://www.industry.gov.au/publications/australias-artificial-intelligence-action-plan.

Cron, W. L. (1984). Industrial salesperson development: A career stages perspective. *Journal of Marketing*, *48*(4), 41-52.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319-340.

Deci, E. L., & Ryan, R. M. (1985). Intrinsic motivation and self-determination in human behavior. New York, NY: Plenum.

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*, 227-268.

Dell'Acqua, F., McFowland III, E., Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., Krayer, L., Candelon, F., & Lakhani, K. R. (2023). Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality. Harvard Business School Technology & Operations Mgt. Unit Working Paper No. 24-013. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4573321.

de Visser, E.J., Pak, R., & Shaw, T.H. (2018). From 'automation'to 'autonomy': the importance of trust repair in human–machine interaction. Ergonomics, 61(10), 1409-1427.

DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, *48*(2), 147-160.

Du Plessis, G., & Smuts, H. (2021, August). The Diffusion of Innovation Experience: Leveraging the Human Factor to Improve Technological Adoption Within an Organisation. In *Conference on e-Business, e-Services and e-Society* (pp. 318-329). Cham: Springer International Publishing.

Eccles, R. G., Ioannou, I., & Serafeim, G. (2014). The impact of corporate sustainability on organizational processes and performance. *Management Science*, *60*(11), 2835–2857.

Eccles, R. G., Perkins, K. M., & Serafeim, G. (2012). How to become a sustainable company. *MIT Sloan Management Review*, *53*(4), 43–50.

Edelman Trust Institute (2025). 2025 Edelman Trust Barometer, Global Report, Trust and the Crisis of Grievance. Available at: https://www.edelman.com/trust/2025/trust-barometer.

Ernst & Young (EY) (2025). The EY Australian AI Workforce Blueprint: What Australian workers really want. Available at: https://www.ey.com/en_au/newsroom/2025/08/new-ey-survey-most-australians-use-ai-at-work-but-few-feel-supported-by-leadership.

Es, S., James, J., Anke, L. E., & Schockaert, S. (2024, March). Ragas: Automated evaluation of retrieval augmented generation. In *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations* (pp. 150-158).

Fifth Quadrant & National AI Centre (NAIC) (2025, August). Australian Responsible AI Index 2025. Available at: https://www.fifthquadrant.com.au/responsible-ai-index.

Fifth Quadrant (2025, January). Australian SMES: AI Adoption Trends. Available at: https://www.fifthquadrant.com.au/australian-smes-ai-adoption-trends.

Frank, D. A., Chrysochou, P., & Mitkidis, P. (2023). The paradox of technology: Negativity bias in consumer adoption of innovative technologies. *Psychology & Marketing*, *40*(3), 554-566.

Future Skills Organisation (FSO) (2025). Workforce Plan 2025: Pathways to impact. Available at: https://www.futureskillsorganisation.com.au/workforce-plans/workforce-plan-2025/.

Gaines, B. R., & Shaw, M. L. G. (2002). Eliciting knowledge and transferring it effectively to a knowledge-based system. *IEEE Transactions on knowledge and data engineering*, *5*(1), 4-14.

Garvin, T., & Kimbleton, S. (2021). Artificial intelligence as ally in hazard analysis. *Process Safety Progress*, *40*(3), 43–49.

Gavrilova, T., & Andreeva, T. (2012). Knowledge elicitation techniques in a knowledge management context. *Journal of Knowledge Management*, *16*(4), 523-537.

Gillespie, N., Lockey, S., Ward, T., Macdade, A., & Hassed, G. (2025). Trust, attitudes and use of artificial intelligence: A global study 2025. The University of Melbourne and KPMG. DOI 10.26188/28822919.

Goffman, G. (1959). The Presentation of Self in Everyday Life. Doubleday.

Gmyrek, P., Berg, J., & Bescond, D. (2023). Generative AI and jobs: A global analysis of potential effects on job quantity and quality. *ILO Working paper 96*, Geneva, International Labour Organization (ILO). Available at: https://doi.org/10.54394/FHEM8239.

Grover, P., Kar, A. K., & Dwivedi, Y. K. (2022). Understanding artificial intelligence adoption in operations management: insights from the review of academic literature and social media discussions. *Annals of Operations Research*, *308*(1), 177-213.

Guihot, M., & McNaught, H. (2021). Platform power, technology, and law: consumer powerlessness in informational capitalism. *Law, Innovation and Technology*, *13*(2), 510-549.

Hadar, I., Soffer, P., & Kenzi, K. (2014). The role of domain knowledge in requirements elicitation via interviews: an exploratory study. *Requirements Engineering*, *19*(2), 143-159.

Hansen, M. (2024). From attention economy to cognitive lock-ins. Big Data & Society, 11(3), 20539517241275878.

Hart, S. L. (1995). A natural-resource-based view of the firm. *Academy of Management Review*, *20*(4), 986–1014.

Hayek, F. A. (1945). The use of knowledge in society. *American Economic Review*, *35*(4), 519-530.

Hayek, F. A. (1952). The counter-revolution of science: Studies on the abuse of reason. The Free Press.

Henderson, R., & Serafeim, G. (2020). Tackling climate change requires organizational purpose. *AEA Papers and Proceedings*, *110*, 177–180.

Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. Human factors, 57(3), 407-434.

Howes, T. D. (2020). F.A. Hayek on the Discovery, Use, and Transmission of Knowledge. Austrian Institute Paper No. 34.

Hughes, L., Dwivedi, Y. K., Malik, T., Shawosh, M., Albashrawi, M. A., Jeon, I., Dutot, V., Appanderanda, M., Crick, T., De', R., Fenwick, M., Gunaratnege, S. M., Jurcys, P., Kar, A. K., Kshetri, N., Li, K., Mutasa, S., Samothrakis, S., Wade, M., & Walton, P. (2025). AI agents and agentic systems: A multi-expert analysis. *Journal of Computer Information Systems*, 1-29.

Human Technology Institute (HTI) (2024). Invisible Bystanders: Workers' experience of AI and automation. Available at: https://www.uts.edu.au/research/centres/human-technology-institute/projects/ai-corporate-governance-program/invisible-bystanders-workers-experience-ai-and-automation.

Human Technology Institute (HTI) (2025). Disconnected AI: the unmet expectations
of consumers and workers. Available at: https://www.uts.edu.au/globalassets/sites/default/files/2025-03/HTI-Insight-Summary-Disconnected-AI.pdf.

Humlum, A., & Vestergaard, E. (2025). The unequal adoption of ChatGPT exacerbates existing inequalities among workers. *Proceedings of the National Academy of Sciences*, *122*(1), e2414972121.

International Energy Agency (IEA). (2024). Electricity 2024: Analysis and forecast to 2026. Available at: https://www.iea.org/reports/electricity-2024.

Jobs and Skills Australia (JSA) (2024). Better Together: The Jobs and Skills Report 2024. Available at: https://www.jobsandskills.gov.au/publications/jobs-and-skills-report-2024.

Jobs and Skills Australia (JSA) (2025A). Our Gen AI Transition: Implications for Work and Skills, Final Overarching Report. Available at: https://www.jobsandskills.gov.au/publications/generative-ai-capacity-study-report.

Jobs and Skills Australia (JSA) (2025B). Our Gen AI Transition: Implications for Work and Skills, Analysis Papers. Available at: https://www.jobsandskills.gov.au/publications/our-gen-ai-transition-analysis-papers.

Jobs and Skills Australia (JSA) (2025C). Jobs and Skills Roadmap for Regional Australia – Phase 1. Available at: https://www.jobsandskills.gov.au/publications/jobs-and-skills-roadmap-regional-australia-phase-1.

Jobs and Skills Australia (JSA) (2025D). Education and training divides - Gendered skills, pathways and outcomes. Available at: https://www.jobsandskills.gov.au/publications/education-and-training-divides-gendered-skills-pathways-and-outcomes.

Kashima, Y., Shafa, S., & Blake, K. (2025). Cultural logics and individualism-collectivism: A conceptualization of the two frameworks from a cultural dynamical perspective. *Current Research in Ecological and Social Psychology*, *9*, 100232.

Koessler, L., Schuett, J., & Anderljung, M. (2024). Risk thresholds for frontier AI. arXiv preprint arXiv:2406.14713.

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. Human factors, 46(1), 50-80.

Li, P., Yang, J., Islam, M. A., & Ren, S. (2025). Making ai less' thirsty'. *Communications of the ACM*, *68*(7), 54-61.

Lodge, J. M. (2025, September). It's all sticks and no carrots for Australian higher education. Available at: https://www.linkedin.com/pulse/its-all-stick-carrot-australian-higher-education-jason-m-lodge-poy3e/.

Lynch, A., Wright, B., Larson, C., Troy, K. K., Ritchie, S. J., Mindermann, S., Perez, E., & Hubinger, E. (2025). Agentic misalignment: How LLMs could be an insider threat. Available at: https://www.anthropic.com/research/agentic-misalignment.

Lyons, S. T., Ng, E. S., & Schweitzer, L. (2012). Generational career shift: Millennials and the changing nature of careers in Canada. In *Managing the new workforce* (pp. 64-85). Edward Elgar Publishing.

Lyons, S. T., Schweitzer, L., & Ng, E. S. (2015). How have careers changed? An investigation of changing career patterns across four generations. *Journal of Managerial Psychology*, *30*(1), 8-21.

Magee, G. (2021). Education reduces recidivism. *Technium Social Sciences Journal*, *16*, 175–182.

Martínez, N., Agudo, U., & Matute, H. (2022). Human cognitive biases present in Artificial Intelligence. *Revista Internacional de los Estudios Vascos (RIEV)*, *67*(2), 1-20.

Messeri, L., & Crockett, M. J. (2024). Artificial intelligence and illusions of understanding in scientific research. *Nature*, *627*(8002), 49-58.

Mialon, G., Dessì, R., Lomeli, M., Nalmpantis, C., Pasunuru, R., Raileanu, R., Rozière, B., Schick, T., D-Yu, J., Celikyilmaz, A., Grave, E., LeCun, Y., & Scialom, T. (2023). Augmented language models: a survey. arXiv preprint arXiv:2302.07842. https://doi.org/10.48550/arXiv.2302.07842.

Microsoft & Tech Council of Australia (2023). Australia's Generative AI opportunity. Available at: https://techcouncil.com.au/newsroom/generative-ai-could-contribute-115-billion-annually-to-australias-economy-by-2030/.

Miehling, E., Ramamurthy, K. N., Varshney, K. R., Riemer, M., Bouneffouf, D., Richards, J. T., Dhurandhar, A., Daly, E. M., Hind, M., Sattigeri, P., Wei, D., Rawat, A., Gajcin, J., & Geyer, W. (2025). Agentic ai needs a systems theory. arXiv preprint arXiv:2503.00237.

Miller, W. R., & Rollnick, S. (2013). Motivational interviewing: Helping people change (3rd edition). The Guilford Press.

Minaee, S., Mikolov, T., Nikzad, N., Chenaghlu, M., Socher, R., Amatriain, X., & Gao, J. (2024). Large language models: A survey. arXiv preprint arXiv:2402.06196. https://doi.org/10.48550/arXiv.2402.06196.

Mizruchi, M. S., & Fein, L. C. (1999). The social construction of organizational knowledge: A study of the uses of coercive, mimetic, and normative isomorphism. *Administrative Science Quarterly*, *44*(4), 653-683.

Nicolas, G., & Caliskan, A. (2024). Directionality and representativeness are differentiable components of stereotypes in large language models. PNAS nexus, 3(11), pgae493.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*(3), 231

Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, *381*(6654), 187-192.

Okafor, E. C., & Osuagwu, C. C. (2006). The underlying issues in knowledge elicitation. *Interdisciplinary Journal of Information, Knowledge, and Management*, *1*, 95-108.

Papagiannidis, E., Enholm, I. M., Dremel, C., Mikalef, P., & Krogstie, J. (2023). Toward AI governance: Identifying best practices and potential barriers and outcomes. *Information Systems Frontiers*, *25*(1), 123-141.

Prochaska, J. O., & Norcross, J. C. (2018). Systems of psychotherapy: A transtheoretical analysis (9th Ed.). Oxford University Press.

Radanliev, P. (2025). Frontier AI regulation: what form should it take? *Frontiers in Political Science*, *7*, 1561776.

Radovanovic, A. (2020). Our data centers now work harder when the sun shines and wind blows. Available at: https://blog.google/inside-google/infrastructure/data-centers-work-harder-sun-shines-wind-blows/.

Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation–augmentation paradox. *Academy of Management Review*, *46*(1), 192-210.

Rastogi, C., Zhang, Y., Wei, D., Varshney, K. R., Dhurandhar, A., & Tomsett, R. (2022). Deciding fast and slow: The role of cognitive biases in ai-assisted decision-making. Proceedings of the ACM on Human-computer Interaction, 6(CSCW1), 1-22.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*, 68-78.

Safe Work Australia (2025, July). Model Code of Practice: Managing psychosocial hazards at work. Available at: https://www.safeworkaustralia.gov.au/doc/model-code-practice-managing-psychosocial-hazards-work.

Sarasohn, E. (2024). Despite AI enthusiasm, Workforce Index reveals workers aren't yet unlocking its benefits. Available at: https://slack.com/intl/en-au/blog/news/the-workforce-index-june-2024.

Schaefer, K.E., Chen, J.Y., Szalma, J.L., & Hancock, P.A. (2016). A meta-analysis of factors influencing the development of trust in automation: Implications for understanding autonomy in future systems. *Human Factors*, *58*(3), 377-400.

Schaeffer, R., Miranda, B., & Koyejo, S. (2023). Are emergent abilities of large language models a mirage? *Advances in Neural Information Processing Systems*, *36*, 55565-55581.

Schank, R. C., & Abelson, R. P. (1997). Scripts, plans, goals, and understanding: An inquiry into human knowledge structures. Psychology Press.

Schwartz, R., Dodge, J., Smith, N. A., & Etzioni, O. (2020). Green AI. *Communications of the ACM*, *63*(12), 54-63.

Schwartz, R., Schwartz, R., Vassilev, A., Greene, K., Perine, L., Burt, A., & Hall, P. (2022). Towards a standard for identifying and managing bias in artificial intelligence (Vol. 3, p. 00). Gaithersburg, MD: US Department of Commerce, National Institute of Standards and Technology.

Sharma, S., & Vredenburg, H. (1998). Proactive corporate environmental strategy and the development of competitively valuable organizational capabilities. *Strategic Management Journal*, *19*(8), 729–753.

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, *69*(1), 99-118.

Simon, H. A. (1989). The scientist as problem solver. In *Complex information processing* (1st Edition), pp. 375-398. Psychology Press.

Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology*, *41*(1), 1-20.

Simon, H. A. (1996). The Sciences of the Artificial (3rd Edition). MIT press.

Simmons, G., & Hare, C. (2023). Large language models as subpopulation representative models: A review. arXiv preprint arXiv:2310.17888.

Singh, D. P. D. (2024). Generative AI through the Lens of Institutional Theory. Available at SSRN 4953180.

Stowers, K., Oglesby, J., Sonesh, S., Leyva, K., Iwig, C., & Salas, E. (2017). A Framework to Guide the Assessment of Human–Machine Systems. Human Factors, 59(2), 172-188. DOI: 10.1177/0018720817695077.

Strubell, E., Ganesh, A., & McCallum, A. (2020, April). Energy and policy considerations for modern deep learning research. In *Proceedings of the AAAI Conference on Artificial Intelligence*, *34*(9) , 13693-13696.

Sun, J., Zheng, C., Xie, E., Liu, Z., Chu, R., Qiu, J., Xu, J., Ding, M., Li, H., Geng, M., Wu, Y., Wang, W., Chen, J., Yin Z., Ren, X., Fu, J., He, J., Wu, Y., Liu, Q., Liu, X., Li, Y., Dong, H., Cheng, Y., Zhang, M., Heng, P. A., Dai, J.., Luo, P., Wang, J., Wen, J-R., Qiu, X., Guo, Y., Xiong, H., Liu, Q., & Li, Z. (2025). A survey of reasoning with foundation models: Concepts, methodologies, and outlook. *ACM Computing Surveys*, *57*(11), 1-43. https://doi.org/10.1145/3729218.

Super, D. E. (1980). A life-span, life-space approach to career development. *Journal of Vocational Behavior*, *16*(3), 282-298.

Tamkin, A., Brundage, M., Clark, J., & Ganguli, D. (2021). Understanding the capabilities, limitations, and societal impact of large language models. arXiv preprint arXiv:2102.02503.

Thomas, J., McCosker, A., Parkinson, S., Hegarty, K., Featherstone, D., Kennedy, J., Holcombe-James, I., Ormond-Parker, L., & Ganley, L. (2023). Measuring Australia's Digital Divide: Australian Digital Inclusion Index: 2023. Melbourne: ARC Centre of Excellence for Automated Decision-Making and Society, RMIT University, Swinburne University of Technology, and Telstra. https://doi.org/10.25916/528s-ny91.

Tonkin, L. (1996). Growing around grief—another way of looking at grief and recovery. *Bereavement Care, 15*(1), 10-10.

Venkatesh, V., Thong, J. Y., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology (UTAUT). *MIS Quarterly, 36*(1), 157-178.

von Eschenbach, W. J. (2021). Transparency and the black box problem: Why we do not trust AI. *Philosophy & Technology*, *34*(4), 1607-1622.

Wang, J., & Redelmeier, D. A. (2025). Forewarning Artificial Intelligence about Cognitive Biases. Medical Decision Making, 0272989X251346788.

Webb, T., Holyoak, K. J., & Lu, H. (2023). Emergent analogical reasoning in large language models. *Nature Human Behaviour*, *7*(9), 1526-1541.

Wei, J., Tay, Y., Bommasani, R., Raffel, C., Zoph, B., Borgeaud, S., Yogatama, D., Bosma, M., Zhou, D., Metzler, D., Chi, E. H., Hashimoto, T., Vinyals, O., Liang, D., & Fedus, W. (2022). Emergent abilities of large language models. arXiv preprint arXiv:2206.07682.

Weidinger, L., Mellor, J., Rauh, M., Griffin, C., Uesato, J., Huang, P. S., Cheng, M., Glaese, M., Balle, B., Kasirzadeh, A., Kenton, Z., Brown, S., Hawkins, W., Stepleton, T., Biles, C., Birhane, A., Haas, J., Rimell, L., Hendricks, L. A., Isaac, W., Legassick, S., Irving, G., & Gabriel, I. (2021). Ethical and social risks of harm from language models. arXiv preprint arXiv:2112.04359.

White, I. (1988). W(h)ither expert systems?—A view from outside. *AI & SOCIETY*, *2*(2), 161-171.

Wilkinson, L. (2023). Samsung employees leaked corporate data in CHATGPT: Report. Industry Dive is an Informa TechTarget business. Available at: https://www. cybersecuritydive. com/news/Samsung-Electronics-ChatGPT-leak-data-privacy/647219/.

Wilson, M., & Daly, M. (1985). Competitiveness, risk taking, and violence: The young male syndrome. *Ethology and Sociobiology*, *6*(1), 59–73.

Winograd, A. (2022). Loose-lipped large language models spill your secrets: The privacy implications of large language models. *Harvard Journal of Law & Technology*, *36*(2), 615-655.

Yadav, M. S., Prabhu, J. C., & Chandy, R. K. (2007). Managing the future: CEO attention and innovation outcomes. *Journal of Marketing*, *71*(4), 84–101.

Zhang, J., Song, W., & Liu, Y. (2025). Cognitive bias in generative AI influences religious education. Scientific Reports, 15(1), 15720.

Zhou, C., Li, Q., Li, C., Yu, J., Liu, Y., Wang, G., Zhang, K., Ji, C., Yan, Q., He, L., Peng, H., Li, J., Wu, J., Liu, Z., Xie, P., Xiong, C., Pei, J., Yu, P. S., & Sun, L. (2024). A comprehensive survey on pretrained foundation models: A history from bert to chatgpt. *International Journal of Machine Learning and Cybernetics*, 1-65. https://doi.org/10.1007/s13042-024-02443-6.

Zhou, S., Yuan, B., Xu, K., Zhang, M., & Zheng, W. (2024). The impact of pricing schemes on cloud computing and distributed systems. *Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online)*, *3*(3), 193-205.