# 1. 0 Introduction

AI (Artificial Intelligence is one of the major technological innovations that occurred in the 21st century. It has brought promises for unprecedented progress and profound disruption. On one hand, AI systems have potential to unlock trillions of dollars in economic value, providing efficiency and innovation at a scale of what previously was unimaginable. From enhancing operational efficiency in the financial sector to creating new tools for scientific discovery, AI represents transformative force with the potential to reshape industries and improve human experiences [1]. However, with how rapid this technology being integrated, it brings in serious ethical concerns and risk of significant economic dislocation. The tension between this immense potential and the attendant risk has become a central topic of both global and national policy debates, reflecting a societal effort to navigate a future increasingly shaped by intelligent machines.

By way of distraction, public discourse emphasizes long-term conceptual risk, for example, the emergence of superintelligence. But the immediate and palpable reality of AI is making an impact on global labour markets. The current wave of AI, particularly generative AI (genAI), is unique in its capacity to automate cognitive tasks that form the bedrock of white-collar and mid-level professions [2]. The roles in law, finance, journalism, and administration which were once seen as isolated from potential automation are now increasingly defining characteristics of an AI-driven transformation. This transformation is not going to happen somewhere in the far future; it is happening now and directly affecting millions of workers-the fundamental assumptions of career stability and the value of human expertise as an avenue for achieving it. This report touches on some of this widespread, immediate change that calls for urgent ethical scrutiny and robust responses in policy.

The underlying research problem is that the existing debates at both ethical and policy levels on AI do not engage fully with the range of harms experienced by workers. The emphasis on job loss as the main adverse impact veils a plethora of other ethical issues, such as de-skilling, loss of autonomy and meaning in work, and the aggravation of economic inequality. Thus, the frameworks of governance put forward may not be enough to ensure the dignity and welfare of the workforce. To fill this gap, this report is guided by three primary research questions:

1. What are the main ethical issues that arise when AI displaces work, beyond simply loss of jobs?
2. How do proposed national frameworks, specifically those in Australia and Malaysia, address responsibility, fairness, reskilling, and stakeholder agency?
3. How should principles of Responsible AI be interpreted or modified to better protect workers?

To answer these questions, this report employs a comparative analytical methodology. It synthesizes evidence from foundational academic texts on AI ethics [3], national policy documents such as Australia's proposal for mandatory AI guardrails [4],[5] and Malaysia's National Guidelines on AI Governance & Ethics [6]. The empirical investigation is based on selected case studies from the UK, Malaysia, and Singapore to illustrate the actual impact of AI in real-world situations. This research deals with social and ethical as well as governance aspects of job destruction due to AI, leaving aside technical AI safety issues.

# 2.0 The Shifting Landscape of White-Collar Work

The labour market is an active restructuring by a deep force within the integration of AI into the workplace; it is not a collective process. Evidence from the field shows that opportunities are increasingly polarized in favour of highly skilled workers while causing some mid-level white-collar occupations to become less stable. The World Economic Forum’s *Future of Jobs Report 2025* forecasts that while roles like AI and Machine Learning Specialists will grow, the largest declines are anticipated in clerical and secretarial positions [7]. This trend is corroborated by academic reviews, which find that AI-driven automation is systematically contributing to the erosion of mid-skill roles in administration and customer service [8].

The scale of this shift is substantial. A McKinsey report estimates that, accelerated by generative AI (genAI), up to 30% of hours currently worked could be automated by 2030, creating a "hollowing out" of the middle-skill segment of the workforce [2]. While macroeconomic reports often project a net gain in jobs globally, this aggregate optimism masks a severe structural mismatch [7]. The skills required for new high-tech roles are vastly different from those possessed by displaced mid-skill workers, threatening to leave a significant portion of the workforce behind. Concurrent with this job transformation is an "AI-driven skills earthquake," as described by PwC, which found that the skills required for jobs highly exposed to AI are changing 66% faster than for other jobs [9].  This rapid obsolescence of existing skills is directly translating into wage polarization. The same PwC report identifies a significant wage premium of 56% for workers who possess specific AI-related skills [9]. This aligns with analysis from ISIS Malaysia, which predicts that occupations anchored in "human-edge" skills, such as complex judgment and creativity, will command higher wages, while workers in routine cognitive roles will face downward wage pressure [10].  This evidence points toward a future where AI acts as an accelerant of economic inequality, framing the challenge not merely as one of market efficiency, but as a pressing issue of distributive justice [11].

# 3.0 Ethical Harms Beyond Job Displacement

Even though employment transformations are discerned in quantitative terms through economic figures, deeper engagement and ethical analysis will further clarify and illuminate qualitatively these changes in the human life. The harms of AI-driven displacement extend far beyond the loss of a paycheck, touching upon fundamental aspects of human dignity, autonomy, and the search for meaning [3].  Academic literature identifies a cluster of psychological and professional harms, including skill degradation, dependency on AI, loss of professional autonomy, and pervasive "automation anxiety" [8]. When AI systems automate the core tasks of a profession, workers may find their hard-won expertise rendered obsolete, leading to a diminished sense of purpose [8].

These abstract harms are vividly illustrated by personal accounts. A report in *The Guardian* profiled a copywriter whose role was gradually reduced to merely proofreading content generated by ChatGPT before her eventual dismissal [12]. This instance boasts a vivid portrayal simultaneously showcasing deskilling and the painful absence of meaningful work. The proliferation of AI is also giving rise to new forms of algorithmic management, where AI systems are used for workplace surveillance, automated performance evaluation, and high-stakes decision-making in hiring and promotion [5]. This raises significant concerns related to fairness and justice, as algorithms trained on historical data can perpetuate and even amplify existing societal biases [13].  The Community and Public Sector Union (CPSU) in Australia has voiced concerns that these systems could be used to discriminate against workers, codifying biases under a veneer of objectivity [5].

Meanwhile in the industries where workers give free play to their creative faculties, the damages that exist there are particularly acute. Case studies of journalists, illustrators, and voice actors reveal a pattern of digital appropriation, where unique human skills, artistic styles, and even personal identities are treated as raw data to be scraped, synthesized, and redeployed without consent or compensation [12]. Such commodification of human experience raises some very serious ethical issues surrounding digital self-ownership, cultural integrity and other non-considerations, some of which are rather divergent from so-called classical issues of algorithmic bias.

# 4.0 A Comparative Analysis of AI Governance and Worker Protection

While countries navigate the implications of AI for their societies, new regulatory philosophies are emerging. Nowhere is the contrast in strengths, but more so weaknesses in addressing worker protections, clearer than when contrasting Australia's proposed mandatory, risk-based framework with Malaysia's voluntary, principle-based guidelines.

Australia is moving toward a mandatory framework that would impose legal "guardrails" on AI in "high-risk" settings, including requirements for testing, human oversight, and transparency [4].  However, this approach has drawn criticism for its narrow definition of "high-risk." Labor organizations like the CPSU have argued that the framework fails to provide certainty that employment is covered and have called for employment itself to be explicitly classified as a high-risk domain [5]. Whereas Malaysia has adopted a voluntary, principle-based approach.  Its National Guidelines on AI Governance and Ethics outline seven core principles, including Fairness, Accountability, and the Pursuit of human benefit, to encourage responsible AI practices among all stakeholders rather than mandating compliance [6]. The framework's strength lies in its holistic vision, but its voluntary nature raises questions about its efficacy.

Table 1: Comparative Overview of Australian and Malaysian AI Governance Frameworks

|  |  |  |
| --- | --- | --- |
| Attribute | Australia (Proposed Framework) | Malaysia (National Guidelines) |
| **Legal Status** | Mandatory (Proposed) | Voluntary |
| **Primary Focus** | Risk-Based (Preventing Harm) | Principle-Based (Guiding Conduct) |
| **Scope** | "High-Risk" Settings (Narrowly Defined) | All Stakeholders (Broad) |
| **Key Directives** | Mandatory Guardrails (e.g., Testing, Human Oversight, Transparency) | 7 Core Principles (e.g., Fairness, Inclusiveness, Human Benefit) |
| **Worker-Specific Provisions** | Vague; relies on employer discretion for consultation (per CPSU critique) [5] | General principle of "human benefit"; no specific worker mandates [6] |
| **Enforcement** | Strong, well-resourced regulator (Proposed) | No formal enforcement mechanism; relies on voluntary adoption |

The limitations of Australia's proposed model become apparent when tested against real-world scenarios. In the UK legal profession, AI is primarily being used as an augmentation tool to increase productivity, a use case that would likely not be classified as high-risk under the current proposal [14]. This contrasts sharply with the UK's creative sectors, where generative AI has led to acute job losses and systemic cultural harms [12]. A regulatory model that doesn't categorize an entire professional class undergoing systematic erosion as "high-risk" by definition puts forward a most grave limitation; this argument, therefore, necessitates a broader and more socially conscious definition for "risk." Malaysia's principle-based approach faces challenges of implementation. An ISIS Malaysia report estimates that 28% of the country's labor force is "highly exposed" to generative AI, with the finance sector being particularly vulnerable [10], [15]. In this context, the case of DBS Bank in Singapore provides a "best practice" model for a human-centric transition [16]. DBS has successfully deployed AI for efficiency gains while simultaneously making massive investments in its workforce, including an AI-powered career development platform and extensive upskilling courses [16]. The DBS case showcases that it is possible, but with huge proactive investments from the corporations. The key question for Malaysia is whether voluntary guidelines would encourage all companies to behave the way that DBS did or if the economic incentive to just cut costs would prevail. This raises an important governance gap: If the application of these principles cannot truly be upheld, they run the danger of becoming an "ethics wash," cloaking employers in spurious responsible behaviour without affording protection to workers.

# 5.0 Forging a New Compact: Modifying Responsible AI for Worker-Centric Governance

Such limitations exist in both the risk- and principle-based model of Responsible AI to require a rethinking of how these frameworks are applied to the world of work. Standard principles will need to be interpreted and expanded in consideration of workers' particular vulnerabilities, moving from a paradigm of reactive harm prevention to proactive dignity, agency, and equity-cantered approaches.

Therefore, a worker-centric perspective demands that one shift from standard conventional definitions of fairness, accountability, and transparency to having agency as a core principle of the worker.

* **From Fairness as Non-Discrimination to Fairness as Distributive Justice:** The principle of "Fairness" has predominantly focused on mitigating algorithmic bias [13].A worker-centric interpretation must expand this to include distributive justice: the fair sharing of the productivity gains generated by AI. It also requires **procedural justice**: ensuring the decision to deploy automation is made through a fair process that includes the voices of those affected.
* **From Accountability for Outputs to Responsibility for Transitions:** "Accountability" is often narrowly defined as assigning blame when an AI system malfunctions [13]. This must be broadened to a concept of shared  
  corporate and governmental responsibility for managing the entire workforce transition, including a proactive duty to fund and provide accessible, high-quality reskilling and upskilling programs.
* **From Transparency as Explainability to Transparency as Disclosure:** For a worker, a meaningful form of transparency is not just model explainability but **full and clear disclosure** about how AI is being used to monitor their activities, evaluate their performance, or contribute to employment decisions. This aligns with demands from labour unions for mandatory consultation [5].
* **Elevating Worker Agency as a Core Principle:** Finally, **Worker Agency** must be treated as a foundational principle. This would move beyond mere consultation to mandate genuine co-design and co-governance of workplace AI systems, granting workers and their representatives a formal role in the selection, implementation, and oversight of technologies that reshape their working lives [5].

Table 2: Proposed Modifications to Responsible AI Principles for Worker Protection

|  |  |  |  |
| --- | --- | --- | --- |
| Responsible AI Principle | Standard Interpretation (Tech-Centric) | Proposed Worker-Centric Interpretation | Rationale / Supporting Evidence |
| **Fairness** | Mitigating algorithmic bias in outputs. | **Distributive Justice:** Equitable sharing of productivity gains. **Procedural Justice:** Fair processes for automation decisions. | Addresses wage polarization and inequality documented by PwC [9] and academic studies [8]. |
| **Accountability** | Assigning blame for system failures and erroneous outputs. | **Responsibility for Transitions:** Proactive corporate and state duty to fund reskilling and social safety nets. | Responds to the large-scale displacement projected by WEF [7] and the need for upskilling [7]. |
| **Transparency** | Model explainability (XAI); making the "black box" understandable. | **Disclosure and Consultation:** Full disclosure of AI use in monitoring and decision-making; mandatory worker consultation. | Aligns with demands from labor unions like the CPSU for a stronger worker voice [5]. |
| **(New) Worker Agency** | (Often subsumed under "Human-in-the-Loop" or stakeholder engagement) | **Participatory Governance:** Formal rights for workers and their representatives to co-design and co-govern workplace AI systems. | Moves beyond passive oversight to active participation, ensuring technology serves human ends. |

# **6.0 Conclusion and Recommendations**

The rapid deployment of AI is not just transforming the terrain of white-collar work but is also engendering changes far beyond the mere displacement of jobs, foreseeing deep ethical harms of deskilling, worker autonomy erosion, and heightened inequality. In analyzing the frameworks for governance emerging in Australia and Malaysia, this report finds the current policy proposals deficient. Risk-based approaches narrowly define harm, thereby ignoring adverse economic and psychological effects, while voluntarist, principle-centered approaches lack the mechanisms of enforcement to make corporate behavior responsible. The most vital conclusion is that AI governance requires a paradigmatic change toward a worker-centered paradigm, which treats employees as relevant stakeholders with individual dignity and welfare being primary goals in the technological transition.

On this basis, we propose the following recommendations:

* For policymakers:
  + Adopt Hybrid Governance: Implement a regulatory model that combines high-level ethical principles with mandatory, enforceable guardrails.
  + Define Employment as "High-Risk": Explicitly classify the use of AI in all aspects of employment as a "high-risk" domain requiring stringent oversight.
  + Legislate Worker Rights: Enshrine in law the right of workers and their representatives to meaningful consultation and co-design in the deployment of workplace AI.
  + Invest in Social Infrastructure: Substantially increase public funding for lifelong learning ecosystems and strengthen social safety nets to support displaced workers.
* For Corporate Actors:
  + Invest in Workforce Transformation: Commit to strategic, long-term investment in comprehensive upskilling and reskilling programs, following proactive models like that of DBS Bank [16].
  + Embrace Participatory Design: Involve employees directly in the process of integrating AI into workflows to leverage their domain expertise and improve buy-in.
  + Commit to Equitable Gainsharing: Develop clear policies for sharing the productivity gains from AI with the workforce through improved wages, benefits, or working conditions.
* For Educational Institutions:
  + Reform Curricula for the AI Era: Shift the focus of education from automatable knowledge to the cultivation of durable "human-edge" skills, including critical thinking, creativity, and social-emotional intelligence [10].
  + Integrate Applied Ethics: Embed AI ethics deeply into all technology, business, and public policy curricula to equip future leaders to navigate complex socio-technical choice [3].

Artificial intelligence is not a deterministic force; its impact will be a direct consequence of the human choices we make today. The goal is not to halt progress, but to consciously steer it toward outcomes that are not only economically productive but also ethically sound and humanly fulfilling.

# References

[1] M. Chui *et al.*, "The economic potential of generative AI: The next productivity frontier," McKinsey & Company, Jun. 2023. [Online]. Available: <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier>.

[2] A. Madgavkar *et al.*, "A new future of work: The race to deploy AI and raise skills in Europe and beyond," McKinsey Global Institute, May 2024. [Online]. Available: <https://www.mckinsey.de/~/media/mckinsey/locations/europe%20and%20middle%20east/deutschland/news/presse/2024/2024%20-%2005%20-%2023%20mgi%20genai%20future%20of%20work/mgi%20report_a-new-future-of-work-the-race-to-deploy-ai.pdf>.

[3] E. Burton, J. Goldsmith, S. Koenig, B. Kuipers, N. Mattei, and T. Walsh, "Ethical considerations in artificial intelligence courses," *AI Magazine*, vol. 38, no. 2, pp. 22–34, 2017.

[4] Department of Industry, Science and Resources, "Introducing mandatory guardrails for AI in high-risk settings: proposals paper," Australian Government, 2024. [Online]. Available: <https://consult.industry.gov.au/ai-mandatory-quardrails>.

[5] Community and Public Sector Union, "Introducing mandatory guardrails for AI in high-risk settings," Submission to the Department of Industry, Science and Resources, Oct. 2024. [Online]. Available: <https://www.aph.gov.au/DocumentStore.ashx?id=bc4f52af-511d-45d5-954f-e4958bbf788d&subId=768956>.

[6] Ministry of Science, Technology and Innovation (MOSTI), "The National Guidelines on AI Governance & Ethics," MASTIC, Sep. 2024. [Online]. Available: <https://mastic.mosti.gov.my/publication/the-national-guidelines-on-ai-governance-ethics/>.

[7] World Economic Forum, "The Future of Jobs Report 2025," Geneva, Switzerland, Jan. 2025.

[8] A. A. Karangutkar, "The Impact of Artificial Intelligence on Job Displacement and the Future of Work," *International Journal of Advanced Research in Science Communication and Technology*, Jul. 2023.

[9]  PwC, "The Fearless Future: 2025 Global AI Jobs Barometer," Jun. 2025. [Online]. Available: <https://www.pwc.com/gx/en/issues/artificial-intelligence/ai-jobs-barometer.html>.

[10] C. Cheng and A. Jasmin, "Novel AI technologies and the future of work in Malaysia," ISIS Malaysia, Jul. 2025. [Online]. Available: <https://www.isis.org.my/2025/07/06/novel-ai-technologies-and-the-future-of-work-in-malaysia/>.

[11] J.P. Morgan Global Research, "Is AI already impacting job growth?," J.P. Morgan, Aug. 2025. [Online]. Available: <https://www.jpmorgan.com/insights/global-research/artificial-intelligence/ai-impact-job-growth>.

[12] A. C. Cassin, "‘One day I overheard my boss saying: just put it in ChatGPT’: the workers who lost their jobs to AI," *The Guardian*, May 31, 2025. [Online]. Available: <https://www.theguardian.com/technology/2025/may/31/the-workers-who-lost-their-jobs-to-ai-chatgpt>.

[13] S. Rawat, S. Kumar, and P. Kumar, "Ethics in Artificial Intelligence: Principle and Challenges," *ResearchGate*, Oct. 2023.

[14] Thomson Reuters, "How UK lawyers are redefining excellence in an AI-driven world," Sep. 4, 2025. [Online]. Available: <https://www.thomsonreuters.com/en-us/posts/legal/uk-lawyers/>.

[15] Nucamp, "Will AI replace finance jobs in Malaysia? Here's what to do in 2025," 2025. [Online]. Available: <https://www.nucamp.co/blog/coding-bootcamp-malaysia-mys-finance-will-ai-replace-finance-jobs-in-malaysia-heres-what-to-do-in-2025>.

[16] Singapore Economic Development Board (EDB), "How DBS, Southeast Asia’s largest bank, is capturing the full value of AI and machine learning in Singapore," Oct. 4, 2024. [Online]. Available: <https://www.edb.gov.sg/en/business-insights/insights/how-dbs-southeast-asias-largest-bank-is-capturing-the-full-value-of-ai-and-machine-learning-in-singapore.html>.