© SEP 2026 | IRE Journals | Volume 10 Issue 3 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I3-1722865

Racing to the Bottom: Impact of the Global AI Race on Ethical and Responsible AI Development

SAHAJVEER SINGH BHATIA
School of Science, Computing and Emerging Technologies, Swinburne University of Technology, Melbourne, Australia

Abstract- ***As one of the most pressing influences over the future course of AI development in the twenty-first century, the AI Race can be defined as a contest among nations for dominance in the domain of AI research and development. Rooted in geopolitical competition among world powers, increased business competition among firms, and an iterative and provisionally adaptive culture, the AI Race introduces systematic incentives that oppose responsible AI. This essay analyzes the nature of the AI Race and its effects on responsible AI, with special reference to the eight AI Ethics Principles proposed by the Department of Industry, Science and Resources in Australia [4]; UNESCO's global Recommendation on the Ethics of Artificial Intelligence [10]; and Hagendorff's [6] meta-analysis of AI ethics guidelines. Specifically, issues of importance considered in this study include: the degradation of safety and reliability under competitive pressure timelines; the deepening problem of bias and unfairness due to insufficiently audited AI development pipelines; increasing AI opaqueness resistant to human monitoring and control; decreased privacy due to competitive pressure for data acquisition; and the expanding chasm between the capabilities of rapidly evolving AI and regulation governing its use. The study draws upon a structured review of authoritative and peer-reviewed primary sources and secondary literature published up until 2026, thus offering analysis at a crucial moment of inflection within the AI Race. It is argued that the AI Race represents a multidimensional threat to responsible AI, which cannot be sufficiently addressed merely via the implementation of ethical principles. Although the eight AI Ethics Principles offered by Australia are a substantive set of provisions, compatible with similar guidelines worldwide, Hagendorff's [6] analysis suggests that voluntary approaches fail to generate any changes in industry. Moreover, the findings of Deloitte's [3] survey demonstrate that organizations in Australia fall behind their global peers in terms of AI governance maturity, while Darden's [7] analysis predicts a narrow window of opportunity over the next five years for building ethical infrastructure in AI before systems are integrated into essential social/governmental functions. Five practical recommendations are made based on the analysis: introducing a mandatory tiered AI risk system (similar to the EU AI Act); tying government AI procurements to compliance with Australia's ethics principles; promoting Australia's role as a leading player in international AI standards setting; investing in AI governance capacity within Australia; and considering a paradigm shift in thinking about ethical AI as critical national infrastructure.***

Keywords - ***Artificial intelligence, responsible AI, AI ethics, AI governance, AI race.***

I. INTRODUCTION

In this paper, the importance of artificial intelligence (AI) in terms of its increasing significance in economic, political, and social aspects is discussed. AI serves as a powerful and rapidly progressing tool in decision-making in crucial areas like predictive diagnostics of patients or autonomous finance management systems. As a result, it influences people's rights, equality and welfare. Nevertheless, despite all the advances, the level of ethical governance of AI technologies remains relatively low. Therefore, the question of fairness, accountability, transparency, and harm prevention in relation to technology policies arises.

The topic under discussion is as follows: What are the impacts of the AI Race on the development of ethical and responsible AI systems? The concept of the "AI Race" refers to competition between nation-states, corporations, and organizations for reaching and retaining their dominance in AI development. This competition concerns not only investments but also the involvement of talent, resources, models, and fast deployment. Even though the results of the AI Race contribute to major innovations, they create several obstacles when developing AI responsibly.

The main idea of this paper lies in the assumption that the AI Race poses a threat to the development of ethical AI systems due to various reasons associated with it, including geopolitical rivalry, business interests, and the desire to innovate. This issue is manifested through such factors as lack of testing for safety, erosion of privacy, opacity, and neglecting the need for governance for the sake of human welfare. On the other hand, the sense of urgency of this process also leads to international cooperation and setting norms of behavior regarding ethics in AI.

Australia's national AI ethics framework, consisting of eight principles developed by the Department of Industry, Science and Resources [4], is used as a foundation for this analysis. Moreover, the paper utilizes such documents as the Recommendation on the Ethics of Artificial Intelligence [10] and a meta-analysis by Hagendorff [6]. Other sources include industry reports for 2026 as well as relevant research. This study employs a structured literature and policy review, positioned as a conceptual and policy-oriented analysis rather than an empirical study. Sources include official government documents on AI, which involve the Department of Industry, Science and Resources' AI Ethics Principles and Guidance for AI Adoption. They are supplemented by peer-reviewed sources, such as Hagendorff's [6] analysis, and authoritative secondary sources of data by such organizations as CSIRO, UNESCO, Deloitte, Atlantic Council, and the Darden School of Business. In this way, the range of sources covers technical, economic, geopolitical, and social factors affecting the AI Race.

This report includes the following sections: Section 2 - definition of the AI Race and identification of its drivers; Section 3 - discussion of its impact on responsible AI development; Section 4 - consideration of Australia's reaction to its own ethical AI framework; and finally, Section 5 - conclusions and recommendations.

II. UNDERSTANDING THE AI RACE

The AI Race is defined as the global competition for AI supremacy. Unlike previous technology contests such as the twentieth-century space race, the AI race involves multiple parties, multi-levels, and extensive integration into civil and defense infrastructure. It includes state-level strategic investments, mainly by the United States, China, and the European Union, and private competition between hyperscale and AI laboratories, such as Google DeepMind, OpenAI, Meta AI, and Anthropic.

In assessing the human capital element of the AI race, the Voronoi Global AI Brain Race Report [11] highlights the concentration of elite AI researchers. A few elite organizations and technology firms control a disproportionately large percentage of global AI expertise. In addition to providing a competitive edge, this also ensures that the AI systems developed align with the values of their developers. As the developers of the most advanced AI systems operate in a relatively narrow community, it implies that the ethical assumptions underlying the systems would be similarly limited.

In examining the geopolitics of AI competition in 2026, the Atlantic Council [1] outlines eight dimensions in which AI is likely to affect global power dynamics. They include AI's influence on military strategy, trade, surveillance, and democracy resistance. These findings suggest that the AI race transcends the purely economic aspect to become a fight for normative frameworks shaping the creation and use of AI technologies globally. The geopolitical nature of the race is critical for responsible AI development because national security considerations can often undermine transparency and justice principles. These geopolitical stakes, however, are only part of the picture - private sector dynamics introduce an equally powerful set of pressures that push against responsible development.

In the private sector, the competition to build better AI technology relies on the idea of the advantages associated with being first to do so. Companies that get their hands on functional AI solutions will be able to acquire more clients and create their own data monopoly that will discourage others from entering the competition. Competitive logic encourages speed at the expense of prudence. According to The Conversation [9], developers are often tempted not to consider possible risks when they work on creating

new technology, seeing the process as one that is slowed down by ethical reviews.

The Darden [7] report provides a different perspective on the problem. The "move fast and break things" mindset acceptable for regular software turns extremely dangerous when applied to AI algorithms that make serious decisions about the credit score, treatment plan, job prospects, or criminal record of a person. The price of mistakes and inefficiencies is much higher in this case than when the technology has been designed ethically from the very beginning.

III. IMPACTS OF THE AI RACE ON RESPONSIBLE AI

The AI Race threatens responsible AI across several interconnected dimensions, the most immediate of which is the erosion of safety and reliability. One of the core principles stated in Australia's AI Ethics Principles [4] is "Reliable and Safe". That is, an AI system should function reliably and safely at all stages of its life cycle, with the relevant alternatives being available should any faults occur. Both features of the AI Race contradict this requirement, namely, the tendency towards accelerated development and testing and deployment in situations when the AI system lacks necessary validation.

According to Hagendorff [6], one of the persistent trends in AI ethics is the separation between the principles laid down in ethical guidelines on the one hand and practices on the other. In addition to that, even if certain organizations pursue ethical goals, their principles are typically not enforced; thus, the gap is caused by a failure to provide relevant measures. The practice, called "ethics washing", may be exacerbated because of competitive pressures, which can make companies sacrifice ethical considerations for the sake of faster deployment.

CSIRO [2] emphasizes the need for a comprehensive approach to responsible AI design and development. Responsible use of data, models, technologies, etc. involves a full-fledged approach that includes ethics at all stages of life cycle of AI products. The race mentality often results in the preference for fast development and rapid deployment without considering the necessity of monitoring activities.

These pressures on safety do not operate in isolation they are compounded by equally serious structural threats to fairness and non-discrimination.

AI systems should be designed in accordance with the Australian Fairness and Inclusivity principles so as not to perpetuate any form of bias against people or specific communities [4]. However, in the current competitive environment in which AI operates, the principles have structural vulnerabilities for the following two reasons: biases in training data and deployment priorities.

Firstly, AI systems that use historical data as a foundation reproduce and often magnify societal biases. In cases where there is pressure to deploy such systems before conducting extensive testing and auditing to eliminate these biases, then the problem will persist within production systems as well. According to Hagendorff [6], the principle of fairness is one of the most discussed in the ethical AI documents worldwide, but simultaneously, it is also one of the least implemented - many guidelines discuss the problem but do not provide practical solutions.

Secondly, the logic of the AI race implies deploying the system initially in areas where it promises more value and profit for an organization, including financial services, advertising, and improving employees' productivity. In turn, vulnerable and underserved communities will receive no benefits from such technology. According to UNESCO [10], these guidelines violate the recommendations in the Ethics of Artificial Intelligence because, while deploying, organizations fail to ensure equity and protection of vulnerable groups.

"Transparent and Explainable" and "Contestable" are two of Australia's eight AI ethics principles that specifically focus on ensuring the use of AI in a way that can be understandable, challengeable, and contestable by humans [4]. Both principles suffer from enormous pressure in connection with the AI Race. Most frontier AI systems - such as large

language models, multimodal AI, and autonomous agents - are usually opaque: their algorithms and processes of inference are difficult to interpret, even for the AI developers themselves.

If businesses compete to introduce frontier AI models, there will likely be little business-related incentive for them to develop explainability technology, which slows down the process of AI deployment without giving the firm any competitive advantage. As Hagendorff [6] observes, the issue of explainability is among the most ignored aspects when companies establish ethical policies regarding the use of AI in their operations. Consequently, businesses may be at risk of developing a structural opacity in the post-deployment stage, as described by Darden [7]. Beyond transparency, the competitive race for data creates a parallel and equally serious threat to individual privacy.

“Human, Societal, and Environmental Wellbeing” is another AI ethics principle stipulated by Australia [4] that requires businesses to create AI systems accounting for their effect on society. Nevertheless, the race to develop frontier AI systems means that benefits go only to businesses, whereas the adverse impacts of such deployment - in particular, disinformation campaigns, replacement of workforce, and erosion of privacy - affect wider society. According to The Atlantic Council [1], one of the top global risks of 2026 was the possibility of using AI for information warfare purposes.

The “Privacy Protection and Security” principle [4] imposes an obligation on the AI systems to safeguard privacy rights throughout the data life cycle. This principle is challenged by the very nature of the competition for AI systems’ creation, because creating a model of this kind requires a significant amount of data. Also, there are additional motivations to neglect data protection policies to get the necessary data faster.

According to UNESCO [10], the development of AI systems must comply with already established human rights legislation, such as the right to privacy. In addition to this point, UNESCO suggests implementing data governance frameworks that would protect people from being surveyed, profiled, and having their personal information misused. Nevertheless, the competitive race puts pressure in the reverse direction, as companies with looser regulations can develop bigger models trained using more abundant amounts of data than their competitors who do not ignore privacy rights. Underlying all these specific harms is a broader structural problem: the growing gap between the pace of AI development and the regulatory frameworks meant to govern it.

The other major issue that can be highlighted as resulting from the AI race in relation to responsible AI is the widening gap between the capabilities of AI and regulatory structures put in place to regulate it. According to Darden [7], the regulatory environment has been found to be disjointed and out-of-step. Specifically, the EU AI Act represents the most comprehensive horizontal AI regulatory regime to date, with its provisions entering into application progressively from 2025 onward. Elsewhere, the situation is less developed, as is the case in Australia where the regulation process appears to be more voluntary.

The Australian ethical approach to regulating AI [4] is not mandatory but relies on voluntary use by companies of eight high-level principles. While being a very well-developed initiative, it has also been recognized by Hagendorff [6] in his general discussion of AI regulations that voluntary frameworks of this kind which lack enforcement provisions, audits, and legal liability regimes have repeatedly proved themselves ineffective in influencing the behaviour of companies in the industry. Furthermore, the CSIRO [2] emphasizes that AI adoption in Australia does not meet consistent criteria.

It is apparent from the survey conducted by Deloitte in 2026 [3] that Australian companies trail behind their counterparts in other countries in terms of harnessing the full potential of AI. Indeed, according to the survey, both AI adoption and its governance maturity remain weaker in Australia than in other countries. There are two important points here. Firstly, Australian companies might start adopting AI

systems without developing necessary internal capabilities to do so ethically, given the competitive environment. Secondly, it is doubtful whether a voluntary framework of the kind can change the behavior of international companies who develop AI systems. It is within this global governance vacuum that Australia's own ethical framework must be evaluated.

The eight principles of AI ethics for Australia, including human, societal, and environmental wellbeing; human-centred values; fairness; privacy protection and security; reliability and safety; transparency and explainability; contestability; and accountability [4], represent a solid ethical framework. The Digital Transformation Agency [5] places those principles in the context of the technological policy of Australian government, thus indicating a commitment of the entire governmental body to the responsible development of AI.

Importantly, it must be acknowledged that those principles are consistent with the international framework proposed by UNESCO [10]. Nonetheless, it should be noted that as argued by Hagendorff [6], through meta-analysis, having the principles in place is not sufficient as the key issue is their implementation into technical measures, procurement policies, auditing procedures, and regulatory frameworks that can withstand competition during the ongoing AI race.

The Guidance for AI Adoption [8] represents an important step beyond principles because it provides a detailed tool set to evaluate risks in the application of AI. Nevertheless, considering the extensive development of frontier AI that takes place overseas, mainly in US laboratories and Chinese scientific institutions, it needs to be determined whether the document itself can affect the ethics of AI adopted by Australian companies from overseas providers. Yet to characterize the AI Race as exclusively harmful would be analytically incomplete - competitive pressures have also, in certain respects, driven progress toward responsible AI.

To view the AI race as an exclusively negative development regarding responsible AI would be analytically flawed. Positive dynamics have also resulted from the development of the competitive dynamic thus far. The competition pressures have led to investments in research into the safety and alignment of AI - not only out of moral consideration but also due to the commercial and reputational necessity associated with this field. Major labs have invested heavily in such research, at least partially because of the pressure resulting from the visibility created by the competition.

In a related vein, international competition is a driver of the formation of standards and benchmarks of interoperability and ethical conduct. Each nation wants its own model to be recognized as exemplary, and thus a body of soft and hard international law has been established by such efforts, even though this law might not be perfect at the present point in time. According to the Atlantic Council [1], AI governance has become another dimension of great power competition whereby countries seek to sell their models of governance alongside their technological achievements.

According to Darden [7], the following five years offer a "window of opportunity". On one hand, this window is not so distant that action should be delayed; on the other hand, it is not so close that the risks cannot be addressed before becoming entrenched. The implication is that not only is the current race harmful, but there is an opportunity here to harness it for good purposes.

IV. CONCLUSION AND RECOMMENDATIONS

This report argues that the AI Race - a period marked by significant geopolitical competition, commercial first-mover advantage, and fast-paced innovation - is a considerable danger to the establishment of ethical and responsible AI systems. Evidence of harm includes rushed safety tests and inadequate monitoring after deployment; reinforcement of biases by poorly audited datasets; increasing opacity of AI systems that diminishes human supervision; decreased privacy due to pressure for data acquisition; and growing discrepancy between the

capabilities of global AI systems and the rules meant to regulate them.

While Australia's AI Ethics framework is largely adequate and consistent with international best practices, its voluntary nature limits its impact. Hagendorff's [6] analysis shows that voluntary guidelines without enforcement mechanisms have generally struggled to produce systematic changes in industry practice. This means that fundamental reformation of Australia's system for regulating AI technology is needed.

Three primary conclusions follow from this analysis. Firstly, ethical AI cannot be reached merely by establishing principles because it needs enforceable technical standards, mandatory assessment of risks of high-risk AI applications, and liability regulations that would incentivize compliance. Secondly, Australia's countermeasures to the AI Race must be dual - they must develop governance infrastructure for AI adoption domestically and shape global norms, standards, and regulation governing how frontier AI systems are built. Thirdly, although the AI Race is a detrimental phenomenon in its current form, it creates urgency that should be capitalized on - the next five years are critical for making decisions about the architecture of governance for AI technology.

Based on these conclusions, the report provides the following recommendations:

- Mandatory risk tiering: Australia should create a legal framework for the risk assessment and classification of AI applications like the one provided by the European Union's AI Act. It should mandate conformity assessment, transparency, and human supervision in the application of high-risk AI systems in medicine, finance, law, and employment.
- Mandatory procurement standards: Any purchases of AI systems by the Australian government, regardless of whether they come from local or international producers, should require the producer to comply with Australia's eight principles of ethical AI. Such compliance can be verified by technical standards established together with CSIRO and relevant international standards organizations.
- International standards leadership: Australia should leverage its existing reputation for principled AI governance to play an active role in international standards development through ISO, IEEE, and the UN AI Advisory Body, ensuring that Australian values - including Indigenous data sovereignty and equitable access - are represented in global frameworks.
- Domestic AI governance capability: In view of Deloitte's [3] finding that Australian organizations lag behind their global counterparts in AI adoption, there is a need to build domestic capacity for responsible AI governance by providing training and infrastructure that makes possible responsible AI deployment even if using foreign AI systems that adhere to different ethical standards.
- Embedding ethics as infrastructure: According to Darden [7], Australia should think of ethical AI as essential infrastructure - invest in it with the same strategic seriousness as in cybersecurity, economic competitiveness, and other issues of national importance.

The AI Race will neither slow down on its own nor stop without policy intervention. However, the time when the ethical architecture of AI systems could still be shaped and integrated into their foundation is still available. Australia's AI Ethics Framework offers a solid basis for this process. What remains is to establish the necessary governance infrastructure to endow these principles with enough weight.

ACKNOWLEDGMENT

The author acknowledges the publicly available resources which were used in the preparation of this report:

- Australia's AI Ethics Principles [4]: Provided the primary normative framework against which the impacts of the AI Race were assessed. The eight principles formed the backbone of the analysis in Section 4.

- Hagendorff, T. [6]: Used as the core academic source for the critical meta-analysis of AI ethics guidelines. Informed the argument about the gap between principles and practice.
- CSIRO Responsible AI Explainer [2]: Provided an authoritative Australian institutional perspective on responsible AI practices and lifecycle governance.
- UNESCO Recommendation on the Ethics of AI [10]: Used to situate Australia's framework within the international ethical consensus and to assess the adequacy of

Australian principles against global standards.

- The Conversation [9] - AI developers and safety: Provided evidence for the claim that commercial AI development systematically deprioritises safety in competitive environments.
- Atlantic Council [1] - Eight ways AI will shape geopolitics: Used to document the geopolitical dimensions of the AI Race and their implications for AI ethics and governance.
- Darden Report [7] - Ethics is the defining issue for AI: Provided contemporary academic and institutional commentary on the urgency of ethical AI, the structural risks of delayed governance, and the distinction between AI ethics and ethical AI.
- Deloitte [3] - State of AI in the Enterprise: Provided empirical data on Australian organisations' AI adoption and governance maturity relative to global peers.
- Guidance for AI Adoption [8]: Provided the Australian Government's practical framework of core practices for adopting AI safely and responsibly across industry.
- Voronoi Global AI Brain Race Report [11]: Provided quantitative data on the global distribution of AI talent and its implications for who shapes the values embedded in frontier AI systems.
- Digital Transformation Agency [5]: Used to contextualise Australia's AI Ethics Principles within the broader wholeof-government digital strategy.

REFERENCES

[1] Atlantic Council, "Eight ways AI will shape geopolitics in 2026," *Atlantic Council*, 2026. [Online]. Available: https://www.atlanticcouncil.org/dispatches/eight-ways-ai-will-shape-geopolitics-in-2026/

[2] CSIRO, "Responsible AI explainer," *CSIRO*, Nov. 2023. [Online]. Available: https://www.csiro.au/en/news/All/Articles/2023/November/Responsible-AI-explainer

[3] Deloitte, "Australian organisations lag global peers in realising AI's transformational potential," *Deloitte Australia*, Feb. 2026. [Online]. Available: https://www.deloitte.com/au/en/about/press-room/australian-organisations-lag-global-peers-realising-ai-transformational-potential-120226.html

[4] Department of Industry, Science and Resources, "Australia's AI ethics principles," *Australian Government*, 2023. [Online]. Available: https://www.industry.gov.au/publications/australias-artificial-intelligence-ethics-principles

[5] Digital Transformation Agency, "Strategy: Australia's artificial intelligence ethics principles," *Australian Government*, n.d. [Online]. Available: https://architecture.digital.gov.au/strategy/australias-artificial-intelligence-ethics-principles

[6] T. Hagendorff, "The ethics of AI ethics: An evaluation of guidelines," *Minds and Machines*, vol. 30, no. 1, pp. 99–120, 2020. Crossref

[7] M. McCance, "Ethics is the defining issue for the future of AI. And time is running short," *Darden Report Online*, Jan. 22, 2026. [Online]. Available: https://news.darden.virginia.edu/2026/01/22/ethics-is-the-defining-issue-for-the-future-of-ai-and-time-is-running-short/

[8] National Artificial Intelligence Centre, "Guidance for AI adoption," *Department of Industry, Science and Resources, Australian Government*, Oct. 2025. [Online]. Available:

https://www.industry.gov.au/publications/guidance-ai-adoption

[9] The Conversation, "AI developers often ignore safety in the pursuit of a breakthrough," *The Conversation*, 2021. [Online]. Available: https://theconversation.com/ai-developers-often-ignore-safety-in-the-pursuit-of-a-breakthrough-so-how-do-we-regulate-them-without-blocking-progress-155825

[10] UNESCO, "Recommendation on the ethics of artificial intelligence," *UNESCO*, 2021. [Online]. Available: https://unesdoc.unesco.org/ark:/48223/pf0000381137

[11] Voronoi, "The global AI brain race report 2026," *Voronoi*, 2026. [Online]. Available: https://www.voronoiapp.com/technology/The-Global-AI-Brain-Race-Report-2026-7709