

# AI Contextual Framework: A Zoning Approach to Ethical AI Deployment

Yao Xie

University College Dublin  
toyaoxie@gmail.com

## Abstract

Drawing on the real-world pressing need for ethical AI governance, this extended abstract proposes a contextual framework using a zoning approach to guide AI deployment: adaptation zones, non-adaptation zones, and hybrid zones. Beyond its practical implications, this framework also raises a deeper societal question: As we advance toward an era of superintelligence and pervasive automation, is a fully automated society truly compatible with human nature and long-term well-being?

## Introduction

In the context of current technological trends, humanity is increasingly moving towards greater digitalization and automation. The integration of artificial intelligence (AI) across sectors such as healthcare, education, finance, and transportation demonstrates a clear shift towards reliance on algorithms and machines for decision-making and operations. This trend suggests a possible future where AI systems play a central role in all aspects of life, potentially leading to a fully automated society. As AI continues to evolve and integrate across various sectors, an essential question arises: Should humanity transition to a fully automated society where AI systems operate independently across all domains? Or would a semi-automated society, where human evaluation and involvement remain central in key areas, or even prohibit automation in certain contexts, be more beneficial for human well-being, societal evolution, and even the future of the planet?

This perspective piece argues that despite AI offering considerable opportunities for technological efficiency, economic growth, and scientific advancement, a fully automated society risks eroding fundamental human values, particularly in areas requiring vulnerability, care, and moral judgment. To safeguard well-being at both individual and societal levels, there is a pressing need for a governance framework that clearly distinguishes between contexts where AI enhances human flourishing and where it undermines it.

Informed by foresight methods used to anticipate long-term societal impacts of emerging technologies, this article introduces the AI Contextual Framework, a zoning-based

model designed to map the ethical boundaries of AI deployment. The framework recognizes that AI's ethical acceptability is context dependent. It categorizes AI deployment into three zones: Adaptation Zones, where AI may be applied broadly with minimal risk to human dignity; Non-Adaptation Zones, where AI must not replace human judgment or agency; and Hybrid Zones, where AI may support but never supplant human decision-making.

Adaptation Zones involve domains that are data-heavy, rule-based, and low in interpersonal complexity, such as predictive maintenance and supply chain optimization. In such contexts, AI can improve efficiency without compromising human well-being. Conversely, Non-Adaptation Zones encompass domains where human presence remains ethically indispensable, such as psychotherapy, palliative care, areas requiring emotional intelligence and moral judgment that AI cannot replicate. Between both lie Hybrid Zones, where AI functions as a valuable copilot in areas like medical diagnostics and educational assessment, where it can augment human capabilities while respecting established boundaries.

The framework complements existing AI ethics paradigms by addressing the foundational question of appropriate deployment contexts, aligning with the UNESCO Recommendation on the Ethics of AI by promoting human rights, dignity, sustainability, and well-being. This zoning model introduces an anticipatory method to map emerging ethical boundaries using foresight approaches such as scenario exploration and value mapping. The framework initiates an evolving ethical cartography and functions as a conceptual map. As contexts shift, this model invites continued iteration through participatory governance, interdisciplinary dialogue, and ecological consideration.

From a governance perspective, the zoning framework embodies the principle of subsidiarity, offering high-level ethical guidance while enabling contextual adaptations across diverse cultural and institutional settings. Through participatory deliberation to strengthen contextual legitimacy and public trust, it further highlights the importance of fostering human connection and participatory engagement. AI systems must not erode social bonds but instead reinforce opportunities for meaningful human interaction and community participation. Zoning decisions should involve multi-level collaborations among stakeholders at local, na-

tional, and international levels to ensure inclusivity in their contexts.

This conceptual framework emphasizes that zoning is not about restricting progress, but about guiding innovation in ways that serves humanity and planet's best interests. Ethical AI begins with understanding not only how to build, but also where to build, and where not to. In the end, we must reflect again: do we really need a fully automated society for humanity?

### Future Work

This extended abstract presents a derivative strand of my PhD research project, which aims to develop a comprehensive, human-centered framework integrating artificial intelligence and human well-being. The overarching project introduces the Systematic AI Well-being Stack, a multi-dimensional blueprint that seeks to redefine mechanisms of well-being, bridge interdisciplinary perspectives on AI and well-being, and identify the meta-capabilities individuals need to sustain well-being in the era of AI.

This abstract focuses specifically on the governance-level context, presenting a contextual zoning framework to guide ethical AI deployment. It represents one foundational component within the broader architecture. Further development of this framework could investigate concrete mechanisms for implementing and updating zoning decisions, such as identifying appropriate governance bodies, legal instruments, or institutional processes. Clarifying boundary ambiguity and addressing edge cases, for example, how disputes over zone classification might be resolved, could enhance the framework's robustness. Additionally, future research may explore ways to conceptualize and, where feasible, measure the erosion of fundamental human values to strengthen the framework's analytical utility. Examining the scalability of the framework across diverse cultural, legal, and economic contexts could also support its relevance as a globally applicable ethical cartography. A further consideration is exploring potential incentive structures, whether regulatory, reputational, or market based, may help encourage organizational compliance. Empirical case studies will be especially valuable for testing and refining these dimensions.

### Acknowledgments

The author, Yao Xie, gratefully acknowledges the guidance, review, and editorial support of her supervisor, Prof. Walter Cullen. She also thanks the anonymous reviewers for their thoughtful and constructive feedback on potential extensions of this work. Appreciation is extended to the AAAI/ACM Conference on AI, Ethics, and Society for recognizing and sponsoring her participation.

### References

Batool, A.; Zowghi, D.; and Bano, M. 2025. AI Governance: A Systematic Literature Review. *AI Ethics*, 5: 3265–3279.

Floridi, L.; Cowl, J.; Beltrametti, M.; Chatila, R.; Chazerand, P.; Dignum, V.; Luetge, C.; Madelin, R.; Pagallo, U.; Rossi, F.; Schafer, B.; Valcke, P.; and Vayena, E.

2018. AI4People — An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. *Minds and Machines*, 28(4): 689–707.

Jobin, A.; Ienca, M.; and Vayena, E. 2019. The Global Landscape of AI Ethics Guidelines. *Nature Machine Intelligence*, 1(9): 389–399.

Lee, L. A.; Durakovic, B.; and Yeralan, S. 2025. AI Ethics: Contextual Frameworks and Domain-Specific Concerns. *Heritage and Sustainable Development*, 7(1): 1–14.

Mittelstadt, B. 2019. Principles Alone Cannot Guarantee Ethical AI. *Nature Machine Intelligence*, 1(9): 501–507.

Mussnug, A. M. 2024. Technology as Uncharted Territory: Contextual Integrity and the Notion of AI as New Ethical Ground. arXiv preprint arXiv:2412.05130.

Papagiannidis, E.; Mikalef, P.; and Conboy, K. 2025. Responsible Artificial Intelligence Governance: A Review and Research Framework. *The Journal of Strategic Information Systems*, 34(2): Article 101885.

UNESCO. 2021. Draft Text of the Recommendation on the Ethics of Artificial Intelligence. Intergovernmental Meeting of Experts Related to a Draft Recommendation on the Ethics of Artificial Intelligence. Document code SHS/IGM-AIETHICS/2021/JUN/3 Rev.2.

Vyhmeister, E.; Castane, G.; Östberg, P. O.; and et al. 2023. A Responsible AI Framework: Pipeline Contextualisation. *AI Ethics*, 3: 175–197.

World Health Organization. ??? WHO Foresight: Monitoring Emerging Technologies and Building Futures Thinking. World Health Organization.

Yudkowsky, E. 2008. Artificial Intelligence as a Positive and Negative Factor in Global Risk. In Bostrom, N.; and Cirkovic, M. M., eds., *Global Catastrophic Risks*, 308–345. Oxford, UK: Oxford University Press.