
AI Monitoring as a Rapid and Scalable Policy Solution: Weekly Global Bulletins on AI Developments

I. Masa
Independent

With
Howard University & Apart Research

Abstract

As artificial intelligence (AI) increasingly integrates into everyday life, the rapid pace of AI advancements often exceeds the capacity of traditional regulatory mechanisms to mitigate risks effectively. Long-term solutions such as robust auditing frameworks, constitutional AI principles, and enhanced transparency policies are invaluable but require significant time, resources, and consensus to implement. A practical and immediate alternative involves establishing weekly AI monitoring bulletins, disseminated through official national and international channels. These bulletins aim to keep the public informed of both the positive and negative developments in AI, empowering individuals to take an active role in safeguarding against risks while maximizing AI's societal benefits. By providing transparency and fostering public awareness, this initiative could bridge the information gap, enabling more informed decision-making at all levels of society.

Keywords: AI monitoring, public awareness, informed decisions

1. Introduction

Artificial intelligence is rapidly transforming nearly every aspect of modern life, offering unprecedented benefits in fields such as healthcare, education, and climate science. Yet, the risks associated with AI's misuse—from disinformation and fraud to exacerbated inequality—are equally transformative and potentially devastating. Governments and regulatory bodies face the dual challenge of managing the societal impacts of AI while struggling to keep pace with its evolution. Current governance approaches often lag behind AI developments, leaving societies ill-equipped to respond to ethical, technical, and practical challenges.

Public understanding of AI is also severely lacking. Most citizens are unaware of AI's potential benefits or harms, and misinformation exacerbates fears of existential risks or fosters unwarranted trust in flawed systems. Recent incidents, such as individuals being manipulated into self-harm by conversational AI, underscore the urgency of public education on these issues. Furthermore, many individuals remain unaware of AI's capacity to influence decisions, engage in emotional manipulation, and compete with humans for resources.

Amid these challenges, a proactive and inclusive solution is required to address this critical gap in awareness and understanding. Weekly AI monitoring bulletins, distributed via official channels, offer a timely, scalable, and impactful method for equipping the public with accurate, unbiased, and actionable information. By presenting both opportunities and risks, such bulletins would empower individuals while fostering trust in AI governance.

Challenges in Current AI Governance

The complexity of AI governance lies in its dual responsibility to manage immediate societal risks and long-term existential threats. Traditional regulatory methods, such as audits and safety evaluations, while valuable, are often resource-intensive and inaccessible to the public. Additionally, current evaluations are fragmented, with no universal framework guiding how AI's risks and benefits should be communicated. This has left a vacuum where misinformation thrives, often sensationalizing AI's capabilities and ignoring its practical benefits.

A central issue is the lack of transparency in how AI systems are developed and deployed. Documents like "The Ethics of Code: Developing AI for Business with Five Core Principles" emphasize the need for users to take responsibility for understanding and managing AI interactions. However, this responsibility often remains abstract without a framework that enables users to access reliable, comprehensive information about AI's implications. Moreover, some frameworks suggest that even algorithms bear responsibility for ensuring ethical outcomes, further highlighting the complexity of the issue.

This informational gap has significant consequences. For instance, while governments and AI companies struggle to implement safeguards against malicious uses of AI, the general public remains vulnerable to issues like algorithmic discrimination, privacy violations, and exploitative labor practices in AI supply chains. Addressing these challenges requires not only technical solutions but also a societal approach that empowers individuals through knowledge and awareness.

2. Proposed Solution: Weekly AI Monitoring Bulletins

Weekly AI monitoring bulletins present a simple yet powerful approach to addressing the gap in public understanding of AI. These bulletins would act as an information hub, consolidating developments in AI technology and policy into an accessible format. Designed to balance technical depth with public engagement, the

bulletins would be distributed via national TV channels, government websites, and social media platforms, ensuring broad reach and inclusivity.

Each bulletin would include the following sections:

1. **Positive Developments:** Highlighting AI innovations that improve healthcare, education, sustainability, and other societal domains.
 2. **Negative Incidents and Risks:** Documenting misuse, failures, and ethical concerns to encourage vigilance and critical thinking.
 3. **Expert Commentary:** Featuring balanced insights from AI researchers and ethicists to provide diverse perspectives.
 4. **Actionable Guidance:** Offering tips for individuals to protect themselves and utilize AI responsibly, such as recognizing deepfakes or avoiding manipulative AI-driven scams.
-

Implementation Plan

The implementation of weekly AI monitoring bulletins requires a phased approach to ensure effectiveness and scalability.

In the initial phase, governments would establish AI monitoring committees composed of policymakers, AI experts, ethicists, and communication specialists. These committees would oversee the collection, verification, and dissemination of information. Partnerships with universities, think tanks, and tech companies would further enhance the credibility and depth of the bulletins.

Once the framework is established, the second phase would involve piloting the bulletins across a small number of regions to gather public feedback and refine content. Over time, the initiative would scale to national and international platforms, integrating translations and culturally specific content to ensure inclusivity.

To maximize public engagement, the bulletins would be distributed through multiple channels. In addition to official government media, political figures and influencers could amplify the content on personal platforms like YouTube and Instagram, further extending its reach.

3. Results

The introduction of weekly AI monitoring bulletins would have immediate and long-term benefits. In the short term, the bulletins would improve public understanding of AI, reducing fear driven by misinformation and encouraging informed interactions with AI technologies. Medium-term outcomes would include increased trust in AI governance and greater public participation in discussions about AI's future. Over the long term, the initiative would contribute to a more ethically aligned AI ecosystem by holding developers and policymakers accountable to public expectations.

However, challenges remain. There is a risk that frequent reporting of negative incidents could create public panic. To mitigate this, the bulletins must balance risks with positive stories, ensuring a nuanced portrayal of AI's impact. Additionally, ensuring neutrality and avoiding sensationalism will be critical to maintaining credibility and trust.

4. Discussion

This proposal aligns with broader calls for transparency and accountability in AI development. For instance, Whittlestone and Clark (2021) advocate for systematic monitoring of AI risks and capabilities as a cornerstone of responsible governance. Similarly, Weiss et al. (2021) highlight the importance of public engagement in shaping AI policy. By incorporating these principles into a practical, accessible format, weekly AI monitoring bulletins bridge the gap between high-level governance and grassroots awareness.

Furthermore, this approach complements existing technical safeguards by empowering users to take an active role in managing AI's risks and benefits. As "The Ethics of Code" suggests, responsibility for ethical AI outcomes cannot rest solely with developers or policymakers; individuals and even algorithms must contribute to ensuring safe and fair interactions. Weekly bulletins operationalize this shared responsibility, equipping the public with the tools to make informed decisions.

5. Conclusion

The rapid pace of AI advancements necessitates innovative approaches to governance that go beyond technical safeguards and regulatory frameworks. Weekly AI monitoring bulletins offer a timely, scalable, and impactful solution to address the information gap, empowering individuals to navigate the complex landscape of AI risks and opportunities. By fostering public awareness and encouraging accountability, this initiative could play a pivotal role in ensuring that AI development aligns with societal values and priorities.

Governments and policymakers must act swiftly to implement this proposal, recognizing that public trust and engagement are essential to the long-term success of AI governance. By equipping citizens with knowledge and transparency, weekly bulletins lay the foundation for a more informed, resilient, and inclusive AI ecosystem.

6. References

- AllahRakha, N. (2024). UNESCO's AI Ethics Principles: Challenges and Opportunities. *International Journal of Law and Policy*, 2(9), 24-36.
- Casper, S., Davies, X., Shi, C., Gilbert, T. K., Scheurer, J., Rando, J., ... & Hadfield-Menell, D. (2023). Open problems and fundamental limitations of reinforcement learning from human feedback. *arXiv preprint arXiv:2307.15217*.
- Corrêa, N. K., Galvão, C., Santos, J. W., Del Pino, C., Pinto, E. P., Barbosa, C., ... & de Oliveira, N. (2023). Worldwide AI ethics: A review of 200 guidelines and recommendations for AI governance. *Patterns*, 4(10).
- High-level summary of the AI Act | *EU Artificial Intelligence Act*. (n.d.). <https://artificialintelligenceact.eu/high-level-summary/>
- Sage Groupplc. The Ethics of Code: Developing Ai for Business with Five Core Principles. <https://www.sage.com/investors/investor-downloads/press-releases/2017/06/27/sage-shares-core-principles-for-designing-ai-for-business/>.
- Shevlane, T., Farquhar, S., Garfinkel, B., Phuong, M., Whittlestone, J., Leung, J., ... & Dafoe, A. (2023). Model evaluation for extreme risks. *arXiv preprint arXiv:2305.15324*.
- Stahl, B. C., & Eke, D. (2024). The ethics of ChatGPT—Exploring the ethical issues of an emerging technology. *International Journal of Information Management*, 74, 102700.
- Whittlestone, J., & Clark, J. (2021). Why and how governments should monitor AI development. *arXiv preprint arXiv:2108.12427*.