Panel Proposal for PAAISS-2025

Title: Algorithmic Authority vs. Human Oversight: Who Should Make the Final Call?

Abstract:

As Artificial Intelligence (AI) systems become more embedded in our daily lives, making decisions in healthcare, education, criminal justice, and finance. As a result, we're faced with a fundamental question: who should have the final say, humans or machines? While algorithms can process vast amounts of data quickly and appear impartial, they often reflect the biases of their training data, lack transparency, and make it difficult to trace accountability when things go wrong. On the other hand, relying solely on human judgment can bring its own challenges, like inconsistency, delays, and resistance to change. The panel will explore real-world situations where AI systems have conflicted with ethical standards or legal norms, such as in predictive policing, medical diagnosis, automated hiring systems, and risk assessment tools (risk scoring). This discussion aims to interrogate how policy, design, and governance can or should reaffirm human oversight in an increasingly automated world. Bringing together voices from research, technology, law, ethics, and public policy, this panel will unpack what responsible oversight really means in the age of intelligent systems, and how we can design AI that supports, rather than replaces, human judgment. This discussion will benefit researchers, policy-makers, and technologists.

2. Topic Description: Relevance to the AI Community and Conference Theme

This topic directly aligns with the PAAISS-2025, diving into one of the most pressing and unsettled questions in AI ethics and policy: who should be accountable for decisions made with or by AI systems? s technology continues to evolve and AI systems take on more autonomous roles, we're faced with growing concerns about trust, accountability, bias, and the limits of oversight. This panel will spark cross-disciplinary conversation about how we can uphold human responsibility and agency, even as machines become more capable decision-makers.

3. Panel Format

Interactive panel discussion and debate (60 minutes)

- 5-minute opening statements per panelist
- 30-minute moderated discussion
- 15-minute audience Q&A and real-time polling

4. List of Panelists

a. Prof Topside E. Mathonsi, Associate Professor and HoD, Tshwane University of Technology, South Africa. A researcher in Cybersecurity and Wireless Systems. He has authored and co-authored over 100 publications, including books, book chapters, journal articles, and conference papers. He presented at the AI Expo Africa 2024 on "The Impact of Generative AI on Cybersecurity," where he also participated in a panel on "Security & Privacy in the Age of AI".

b. Dr. Ayanda Mbele, Senior Researcher, Institute for Human-Centered AI, South Africa. An AI ethicist with a background in philosophy and machine learning. Her research focuses on ethical frameworks for AI accountability and the moral implications of decision automation.

c. Prof. Daniel Chen, Professor of Law and Data Science, European University Institute. Specializes in legal responsibility and algorithmic governance. Leads projects on AI in judicial decision-making and regulatory models for machine-assisted judgment.

d. Dr. Lindiwe Mokoena, Policy Advisor, UNESCO AI & Ethics Initiative. Works on global policy development for responsible AI. Brings insight into international frameworks and human rights considerations in algorithmic governance.

e. Mr. Rafael Díaz, AI Systems Engineer, OpenGovTech, Brazil. An applied technologist who builds AI tools for public sector use, including risk assessments and automated decision-support systems. Brings a systems design perspective.

5. Moderator Information

Moderator: Dr. Naledi Masinga, Associate Professor and Director of the Centre for Technology and Society, University of Johannesburg. Dr. Masinga has extensive experience moderating cross-disciplinary panels on emerging technology, ethics, and policy. She has facilitated high-level discussions at AI Expo Africa, UNESCO forums, and various national policy dialogues. Her moderation style is inclusive, dynamic, and focused on actionable insights.

6. Proposed Questions or Key Discussion Points

1. In what contexts should AI systems be allowed to make final decisions without human intervention?
2. How do we mitigate automation bias, especially in high-stakes or marginalized communities?

3. What legal or ethical frameworks can ensure accountability when an algorithm "goes wrong"?
4. Can we design systems that meaningfully incorporate human oversight without undermining AI's efficiency?
5. What role should humans play when decisions carry serious consequences? / Is it feasible or desirable to require a "human-in-the-loop" in all critical decision systems?
6. How do cultural and political contexts affect perceptions of algorithmic authority?
7. How do we strike the right balance between efficiency and ethics?