**Prompt: Read this document, and 1) tell me what the authoring is trying to make me believe; 2) tell me what they are really saying between the lines; 3) How would they do on the FANTT Scale (Fairness, Accountability, Non-maleficence, transparency and trust)?**

**Name:** Ethical Governance of Artificial Intelligence

**Issuer:** Institute for AI International Governance (AIIG), Tsinghua University

**Country of Issuer:** China

**Date of Issue:** Ongoing

**Sector:** Ethics and Technology

**Principles:** ethical governance of artificial intelligence, global AI governance frameworks, AI governance ethics; **Values:** fairness, justice, security, privacy, inclusivity, shared responsibility, openness, collaboration

**URL:** Ethical Governance of Artificial Intelligence

**1) What the Authoring Is Trying to Make You Believe**

The authors are trying to convince the reader that AI is a powerful force that needs ethical guidance to ensure it benefits society. They want you to believe that:

* AI offers significant opportunities for human advancement and societal improvement.
* There are potential risks associated with AI's development and use, including misuse, overuse, and underuse.
* Ethical principles are essential for navigating these opportunities and risks.
* A multi-stakeholder approach is crucial for creating a "Good AI Society."
* The recommendations provided in the document offer a solid foundation for achieving this goal.

**2) What They Are Really Saying Between the Lines**

Between the lines, the authors are also conveying a few subtle but important messages:

* **A sense of urgency:** The language used emphasizes that AI is not a future problem but a present concern requiring immediate attention.
* **A call for proactive measures:** The authors stress the need to move beyond simply identifying ethical principles to implementing concrete actions and policies.
* **A recognition of complexity:** The document acknowledges the inherent tensions and trade-offs in AI development and the need for ongoing dialogue and adaptation.
* **A European perspective:** While advocating for global principles, the authors frame their recommendations within a European context, hinting at specific regional priorities and values.

**3) How Would They Do on the FANTT Scale?**

Here's an assessment of the document based on the FANTT scale:

* **Fairness:** The document explicitly addresses fairness, particularly in the context of justice. It calls for AI to promote justice, eliminate discrimination, and ensure equitable access to benefits. It also acknowledges the risk of bias in AI systems and the need to protect social support systems. So, it seems to score well on fairness.
* **Accountability:** Accountability is a central theme, especially concerning the explicability of AI systems. The authors emphasize the need for mechanisms to understand AI decision-making processes, assign responsibility for AI actions, and provide redress for harm caused by AI.
* **Non-maleficence:** The document demonstrates a strong focus on non-maleficence. It dedicates significant attention to identifying and mitigating the potential harms of AI, such as privacy violations, misuse of AI, and erosion of human autonomy.
* **Transparency:** Transparency is closely linked to explicability in the document. The authors advocate for making AI systems understandable and providing clear explanations of how they work, especially in critical decision-making contexts.
* **Trust:** Building public trust in AI is a key objective of the document. The authors believe that explicability, accountability, and ethical guidelines are essential for fostering trust and encouraging the adoption of AI technologies.

Overall, the document performs well on the FANTT scale. It comprehensively addresses the ethical considerations surrounding AI and emphasizes the importance of fairness, accountability, non-maleficence, transparency, and trust in developing and deploying AI systems.

# PHD KSAs for AI Governance

* Query successful

Here is a summary of what you might learn in a conversation with a PhD student studying AI Governance, based on the reference text:

A PhD student studying AI Governance would likely discuss the challenges and principles surrounding the regulation and ethical use of Artificial Intelligence (AI). They might cover topics such as:

* **The importance of AI Governance:** AI is a key part of the ongoing digitalization process at government, economic, and social levels, offering innovation and growth potential. However, it also presents challenges like data-based discrimination, lack of traceability, and privacy concerns.
* **Key principles of AI Governance:** These include putting people first, ensuring regulatory conditions for AI development, transparency, traceability, accountability, safety, actively shaping AI governance, and involving stakeholders.
* **The need for a balanced regulatory approach:** This approach should foster innovation while ensuring responsible application, legal certainty, and addressing risks.
* **The role of transparency and traceability:** AI-based decision-making should be recognizable, and the functioning and purpose of AI systems should be disclosed.
* **The importance of accountability and safety:** Liability must be clearly defined when using AI, and AI systems must be designed to be safe, robust, and resilient.
* **International cooperation in AI Governance:** This involves engaging with international organizations to develop global standards and norms for AI use, while also considering developments in regions like the EU.
* **The necessity of stakeholder involvement:** This includes governments, the private sector, civil society, and experts from various fields, to ensure responsible AI use.
* **Specific guidelines for education, research, and innovation:** These guidelines emphasize the importance of digital skills, promoting open AI systems, and ensuring AI research considers societal and environmental impacts.
* **The applicability of existing legal frameworks:** AI development and use are subject to national and international laws, including those related to fundamental rights, data protection, intellectual property, and liability.
* **The role of data protection:** The revised Data Protection Act includes measures that impact AI, such as those related to profiling, genetic and biometric data, privacy by design, impact assessments, and automated decisions.

Overall, the conversation would likely provide a comprehensive overview of the multifaceted nature of AI Governance, encompassing ethical, legal, and policy-related dimensions.