AI ETHICS AND GOVERNANCE: CASE STUDY FROM IBM

(SWIPE TO SEE)

NEW APPROACH TO AI ETHICS AND GOVERNANCE

IBM, one of the leading technology companies in the world, has taken a proactive approach to addressing AI ethics and governance

HOW IBM ADDRESSES THESE TOPICS?

IBM's AI Ethics and Governance Approach can be broken down into several key components:

1. **AI Principles**: IBM has developed a set of AI Principles to guide the development and use of AI technologies.
2. **AI Governance Framework**: IBM has established an AI Governance Framework to help organizations implement ethical AI practices. This framework includes guidelines, tools, and resources to help organizations assess and manage the ethical risks associated with AI.
3. **AI Ethics Board**: IBM has created an AI Ethics Board to provide independent oversight and guidance on AI-related ethical issues. The board is composed of experts in various fields, including technology, law, ethics, and social sciences.
4. **AI Education and Training**: IBM offers AI education and training programs to help professionals develop the necessary skills and knowledge to design, build, and manage ethical AI systems.
5. **AI Research**: IBM invests in AI research to advance the state of the art in AI technology while also exploring new ethical approaches to AI development and use.

IBM'S PRINCIPLES

AI Principles: IBM has developed a set of AI Principles to guide the development and use of AI technologies. These principles include:

- **Transparency**: AI systems should be transparent, allowing users to understand how they work and the factors that influence their decisions.
- **Responsibility**: AI developers and users must be held accountable for the outcomes of their AI systems.
- **Privacy**: AI systems should respect user privacy and ensure that personal data is protected.
- **Inclusion**: AI systems should be designed to benefit all people, regardless of their background or abilities.
- **Safety**: AI systems should be designed to minimize the risk of physical or digital harm.
- **Fairness**: AI systems should be fair and unbiased, avoiding discrimination based on race, gender, age, or other factors.

IBM'S PILLARS

- EACH PILLAR HAVE EXTENSIVE OPENSOURCE TOOLKIT
- https://www.ibm.com/impact/ai-ethics

Explainability
Good design does not sacrifice transparency in creating a seamless experience.

AI Explainability 360

Fairness
Properly calibrated, AI can assist humans in making fairer choices.

AI Fairness 360

Robustness
As systems are employed to make crucial decisions, AI must be secure and robust.

Adversarial Robustness 360

Transparency
Transparency reinforces trust, and the best way to promote transparency is through disclosure.

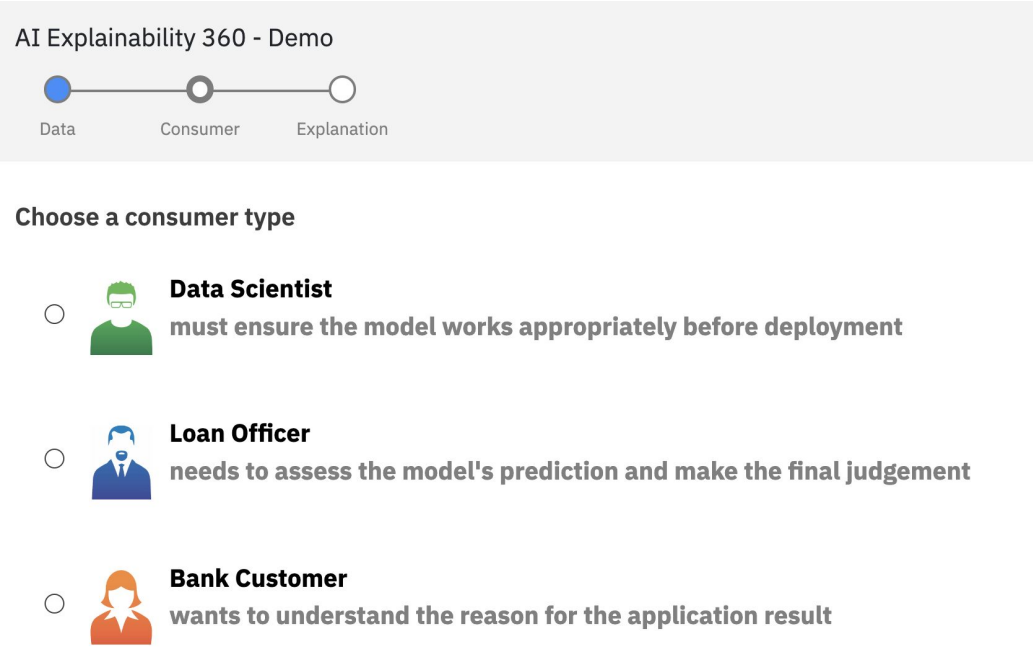
AI FactSheets 360

Privacy
AI systems must prioritize and safeguard consumers' privacy and data rights.

AI Privacy 360 toolkit

AI EXPLAINABILITY

- “The AI Explainability 360 Python package includes algorithms that span the different dimensions of ways of explaining along with proxy explainability metrics”
- “The AI Explainability 360 interactive demo provides a gentle introduction to the concepts and capabilities by walking through an example use case from the perspective of different consumer personas”



AI ETHICS AND GOVERNANCE: CASE STUDY FROM IBM

(SWIPE TO SEE)