Re-view ethical frameworks

Future-proof your strategy around the AI lifecycle

KPMG Smart Government
Catalyse digital progress

Executive Summary

AI holds the power to change the world. Across government organisations, it has revolutionised the way employees deliver services, and the way citizens interact with government.

As an emerging, relatively new technology, AI draws scepticism from many citizens, employees, and government officials. The potential for algorithmic bias, and the cybersecurity risks of any technology, hold many organisations back from fully embracing AI.

Why is ethical AI so important?

With the potential power and scale of AI, and its ability to make autonomous decisions based on evolving algorithms, ensuring AI and the algorithms it relies on are built upon an ethical foundation is vital.

Unfair use of AI can be harmful for individuals, entire demographics, environments, communities, and society as a whole.

Unethical AI can cause biased hiring decisions, privacy violations when facial detection is used for surveillance, and the potential ethical consequences of implementing predictive policing, to name just a few examples.

Trust is the cornerstone of any modern government. Unethical AI puts trust at risk.

Managing bias across the entire AI lifecycle

Challenging bias is a key responsibility of any government organisation. It is the government's remit to provide equitable services to each citizen — this forms the foundation of trust.

[2] Anushka Jain, "UK Government Orders Probe Into Bias In Medical Devices, Artificial Intelligence Tools", 23 November 2021

How can we mitigate bias in AI?

- It is important to think about the primary purpose for using data and who could potentially be affected by it
- Implement tools for continuous monitoring and governance of AI to ensure algorithms are equipped to learn from data for scale, without developing bias in the long-term
- Continuously assess the impact of unfair data processing — implement human intervention periodically
- Use bias mitigation techniques (such as the AI Fairness 360 library) to tackle any bias you discover

Building transparency into the foundations of AI

There will be a varying level of understanding of AI across any government organisation and among its citizens. But it's important to build transparency into your AI operations, making information on how and why AI is essential to government service delivery.

As the systems are so complex, many citizens and colleagues will never understand fully how AI and automated systems reach conclusions — this is a 'black-box' model.

Thanks to emerging Explainable AI (XAI) techniques, it is possible to open up black-box AI to certain extents without scarifying model performance and accuracy. This is the foundation of understanding how decisions are made and assessing whether they are fair.

AI and the law: A grey area

The Data Protection Act 2018 is UK's implementation of the General Data Protection Regulation (GDPR). It controls how an individual's personal information is used by organisations, businesses or the government.

Without proper legislation, government organisations should have stringent processes and policies in place for safeguarding sensitive data across all technology systems.

This could mean restricting the access to sensitive information on certain systems, and controlling which colleagues across the organisation come into contact with said data, as a start.

The human touch

Human oversight is necessary to maintain control over AI and maintain trust at a stakeholder level. But several questions should be asked for an effective framework to be put in place:

- Which decisions should remain in the human realm?
- Why and how were certain use cases chosen as candidates for AI?
- Will the results of AI algorithms impact live (e.g. hiring), or other objects (e.g. asset register)? If impacting live, the higher ethical standard should be in place, even if we need to sacrifice model accuracy for more transparency.
- Why did the team, or the feature engineering algorithm, choose the features they chose, or exclude what they excluded?
- How do we measure and demonstrate success or explain failures?
- Why did the algorithm do what it did, and who was responsible for the outcome?

As systems become ever more powerful, decisions and blame cannot be focused squarely on the algorithm, and AI should not be used to unduly influence or manipulate thoughts and behavior.

Some or all of the services described herein may not be permissible for KPMG audit clients and their affiliates or related entities.

kpmg.com/uk

CREATE | CRT14789B | July 2023

Document Classification: KPMG Public