

## Week 2 Seminar Notes

### Task 1:

Read the paper “*A practical guide to Responsible Artificial Intelligence*” then explain:

1. What are the potential problems or risks with AI use?
2. What are customers’ potential major concerns of AI?
3. What are the major concerns of business leaders?
4. Use the Discussion board below to add your own notes and answers/thoughts

#### Notes:

The article “*A practical guide to Responsible Artificial Intelligence (AI)*” outlines several potential problems, risks, and concerns associated with AI use from the perspectives of consumers, business leaders, and the broader societal implications.

#### 1. Potential Problems or Risks with AI Use:

**Biased Decision-Making:** AI can perpetuate or even exacerbate biases present in the data it was trained on, leading to unfair or discriminatory outcomes.

**Interpretability and Explainability:** There's a challenge in understanding and explaining AI decisions, especially when AI systems are used in critical applications where transparency is essential.

**Robustness and Security:** AI systems need to be resilient and secure against manipulation or attacks, which can be particularly challenging as AI systems become more complex and autonomous.

**Displacement of Human Workers:** There's a concern about AI systems displacing human workers, potentially leading to unemployment and social unrest.

**Amplification of Inequality:** The uneven access to AI technology could widen the gap between the "haves" and "have-nots", exacerbating social and economic inequalities.

**Physical Threats:** In certain applications, such as autonomous weapons or vehicles, AI poses direct physical threats to human safety.

#### 2. Customers’ Potential Major Concerns:

**Privacy and Data Use:** Customers are concerned about how their data is being used by AI systems and whether companies are handling their information responsibly and ethically.

**Bias and Fairness:** Consumers worry about AI systems being biased against them, especially in sensitive areas like loan approvals, job applications, and law enforcement.

**Lack of Transparency:** There's a desire for services that are tailored to individual needs, but this comes with a demand for transparency and understandability in how AI makes decisions affecting them.

### 3. Major Concerns of Business Leaders:

- **Trust and Explainability:** Business leaders recognize the need for AI-based decisions to be explainable and understandable to earn the trust of stakeholders, including customers, regulators, and their own employees.
- **Alignment with Organizational Values:** There's a concern about ensuring AI use aligns with the company's core values and ethical standards.
- **Compliance and Ethical Considerations:** Ensuring AI solutions comply with current and future regulations and ethical guidelines is a major concern, including addressing issues like privacy, bias, and security.
- **Impact on Business Strategy:** Leaders are focused on how AI will change their business models and operations in the near future and how to integrate AI responsibly to leverage its benefits while mitigating risks.

The report emphasises the importance of addressing these concerns through a comprehensive framework that includes governance, ethics and regulation, interpretability and explainability, robustness and security, and bias and fairness to deploy AI responsibly.

## Task 2

*Search for another responsible AI frameworks to compare with the presented RAI framework.*

### 1. AI Ethics Frameworks: National and International Guidelines

Organisations	Guidelines/Frameworks
World Economic Forum (2020)	<a href="#">How to manage AI procurement in public administration</a>
UNESCO	<a href="#">Recommendation on the ethics of AI</a>
EU	<a href="#">EU AI Act</a>
G20 (2023)	<a href="#">G20 AI Principles</a>
Canada (2021)	<a href="#">The Digital Charter Implementation Act, 2022</a>
Singapore Government (2020)	<a href="#">The Model AI Governance Framework</a>
UK Government (2020)	<a href="#">Guidelines for AI Procurement</a>
Department of Industry, Science and Resources, Australia (2019)	<a href="#">Australia's Ethics Principles</a>
Australian Government Architecture	<a href="#">Adoption of AI in public sector</a>

## 2. AI Ethics Frameworks: Academic's and Practitioner's Guidelines

Authors	Guidelines/Frameworks
Ashok et al., 2019	<a href="#">Ethical framework for AI and Digital technologies</a>
Jobin et al., 2019	<a href="#">The global landscape of AI ethics guidelines</a>
Richardson (2021)	<a href="#">Best Practices for Government Procurement of Data Driven Technologies</a>
Jacobs & Mulligan (2022)	<a href="#">The Hidden Governance in AI</a>
Coglianese & Lampmann (2021)	<a href="#">Contracting for Algorithmic Accountability</a>
Rubenstein (2022)	<a href="#">Retooling the Acquisition Gateway for Responsible AI</a>

## 3. AI Ethics Frameworks: Common Themes



## 4. AI Ethics and Policy News by Casey Fiesler: <http://bit.ly/ai-ethics-news>