INTERNATIONAL DIRECTORS SUMMIT 2022

ICDM Institute of Corporate Directors Malaysia

Avoiding Ethical Missteps in the Digital & AI World

azharazib@rawonion.com.sg

‘Can You’ vs ‘Should You’ …

Today’s business dilemma

‘Can You ?’ is usually a legal question ..
Compliance to law, like Privacy Law … Do you have consent to use the data.

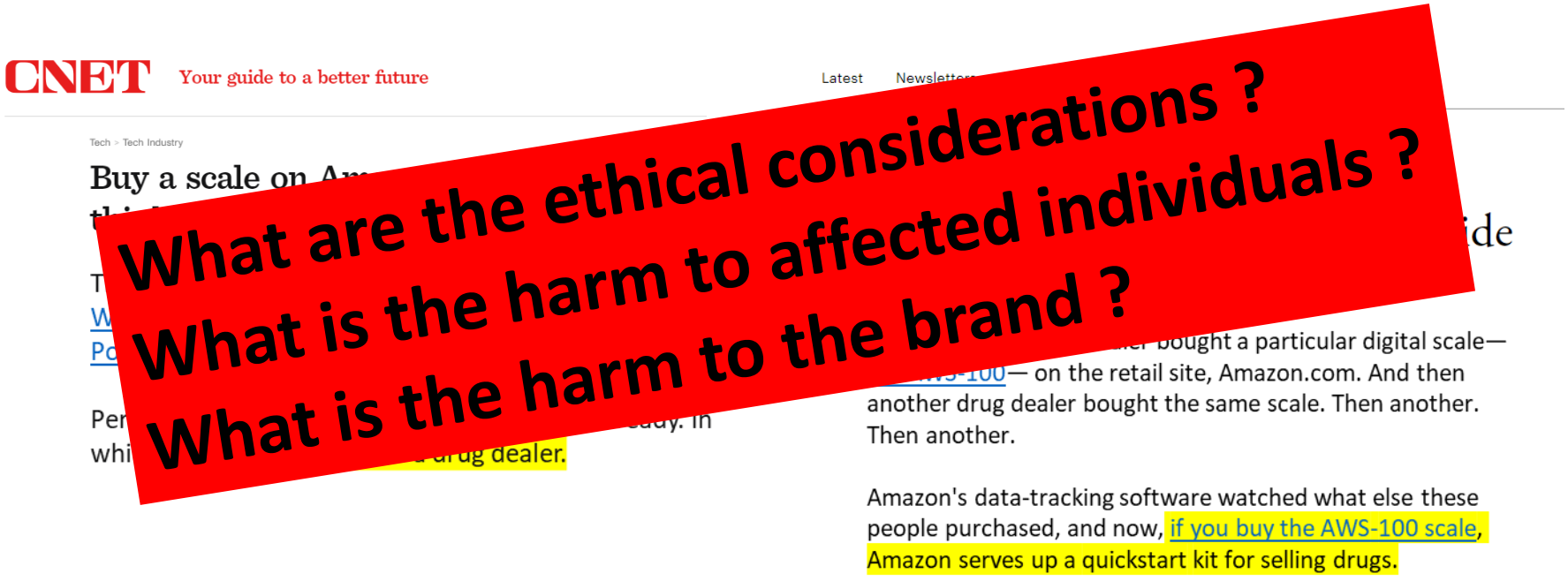
‘Should You ?’ is usually the ethics question ..
Yes, you are legally allowed to, but should you use this data.

This is not about data security or protection from cyber attack.
This is about (potential) HARM
to individuals affected by the use of this data

For this presentation, we will use the term Artificial Intelligence or AI to broadly include all techniques that use data for automated decision making .

Everybody makes mistakes,

including the AI–enabled algorithm



Like other technologies, **AI aims to increase human productivity.**

However, unlike earlier technologies, **some aspects of autonomous predictions or decisions made by AI may not be fully explainable.**

As AI technologies can make decisions that affect individuals, or have a significant impact on society, markets or economies, **organisations should consider a governance model for the use of AI which is consistent with existing ethical principles.**

A Compilation of Existing AI Ethical Principles

1. **Accountability**: Ensure that AI actors are responsible and accountable for the proper functioning of AI systems and for the respect of AI ethics and principles, based on their roles, the context, and consistency with the state of art.

2. **Accuracy**: Identify, log, and articulate sources of error and uncertainty throughout the algorithm and its data sources so that expected and worst-case implications can be understood and can inform mitigation procedures.

3. **Auditability**: Enable interested third parties to probe, understand, and review the behaviour of the algorithm through disclosure of information that enables monitoring, checking or criticism.

4. **Explainability**: Ensure that automated and algorithmic decisions and any associated data driving those decisions can be explained to end-users and other stakeholders in non-technical terms.

5. **Fairness**:

a. Ensure that algorithmic decisions do not create discriminatory or unjust impacts across different demographic lines (e.g. race, sex, etc.).

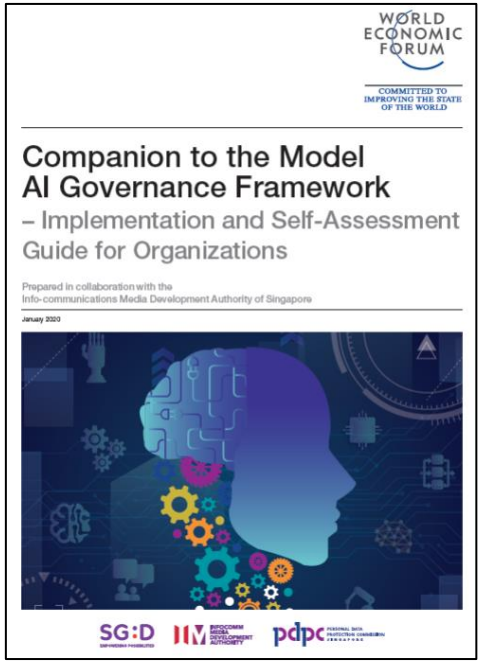
b. To develop and include monitoring and accounting mechanisms to avoid unintentional discrimination when implementing decision-making systems.

c. To consult a diversity of voices and demographics when developing systems, applications and algorithms.

Source includes the Institute of Electrical and Electronics Engineers ("IEEE") Standards Association's *Ethically Aligned Design* (https://standards.ieee.org/industry-connections/ec/ead-v1.html) …etc

AI and Automated-Decision-Making Ethics is about

Governance of Data Use and Data Processing

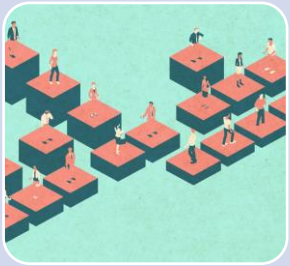


Regulators and various government agencies and industry groups are introducing **AI Governance Frameworks** to help organisations assess and manage risk and harm.

The Framework is

Algorithm-agnostic	Technology-agnostic	Sector-agnostic	Scale- and Business-model-agnostic
• It does not focus on specific AI or data analytics methodology. • It applies to the design, application and use of AI in general.	• It does not focus on specific systems, software or technology, and will apply regardless of development language and data storage method	• It serves as a baseline set of considerations and measures for organisations operating in any sector to adopt. • Specific sectors or organisations may choose to include additional considerations and measures or adapt this baseline set to meet their needs.	• It does not focus on organisations of a particular scale or size. It can also be used by organisations engaging in business-to-business or business-to-consumer activities and operations, or in any other business model.

Framework focuses primarily on four broad areas:



INTERNAL GOVERNANCE STRUCTURES & MEASURES

- Clear roles and responsibilities in your organisation
- SOPs to monitor and manage risks
- Staff training

HUMAN INVOLVEMENT IN AI-AUGMENTED DECISION MAKING

- Appropriate degree of human Involvement
- Minimise the risk of harm

OPERATIONS MANAGEMENT

- Minimise bias in data and model
- Risk-based approach to measures such as explainability, robustness and regular tuning

STAKEHOLDER INTERACTION & COMMUNICATION

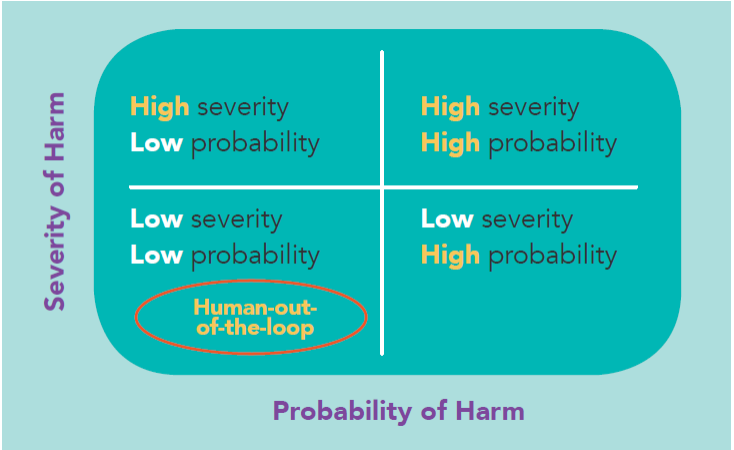
- Make AI policies known to users
- Allow users to provide feedback, if possible
- Make communications easy to understand

DETERMINING THE LEVEL OF HUMAN INVOLVEMENT

Human-in-the-loop suggests that human oversight is active and involved, with the human retaining full control and the AI only providing recommendations or input.

Human-out-of-the-loop suggests that there is no human oversight over the execution of decisions. The AI system has full control without the option of human override.

Human-over-the-loop (or human-on-the-loop) suggests that human oversight is involved to the extent that the human is in a monitoring or supervisory role, with the ability to take over control when the AI model encounters unexpected or undesirable events



OPERATIONS MANAGEMENT

Data Understanding

Understanding the lineage of data: This means knowing where the data originally came from, how it was collected, curated and moved within the organisation, and how its accuracy is maintained over time.

Ensuring data quality: Organisations are encouraged to understand and address factors that may affect the quality of data : accuracy, completeness, taxonomy compliance.

Minimising inherent bias: The Model Framework focuses on inherent bias in datasets (selection, measurement bias), which may lead to undesired outcomes such as unintended discriminatory decisions.

Different datasets for training, testing, and validation

STAKEHOLDER INTERACTION AND COMMUNICATION

Making sure that **consumers are aware** that the products or services that they are considering are AI-enabled.

Providing information so that **consumers know how the AI-enabled features are expected to behave** during normal use.

Consider giving consumers **the option to opt-out of automated decision making**.

The Framework is based on two high-level guiding principles

that promote trust in AI and understanding of the use of AI technologies:

Organisations using AI in decision-making should ensure that the decision-making process is explainable, transparent and fair.

Although perfect explainability, transparency and fairness are impossible to attain, organisations should strive to ensure that their use or application of AI is undertaken in a manner that reflects the objectives of these principles as far as possible. This helps build trust and confidence in AI.

AI solutions should be human-centric.

As AI is used to amplify human capabilities, the protection of the interests of human beings, including their well-being and safety, should be the primary considerations in the design, development and deployment of AI

Some Hard Questions to Ask

1. *Do you have clear roles and accountabilities when it comes to AI related facilities ?*
    a) *Who defines business rules and decision criteria*
    b) *Who approves data sources*
    c) *Who/how do we test AI rules*
2. *Have you thought about the role of the human in the AI augmented process and the risk associated with each option*
3. *Do you have a structured way to score risk/harm associated to this AI augmented process*
4. *Do you truly understand the lineage of data you plan to use to drive the AI*
5. *Do you have an easy way to tell if there is a data quality problem*
6. *Have you defined data quality with regards to the AI augmented process*
7. *Do you plan to let the customer know that some processes are AI augmented*
8. *Have you considered customer reactions and mitigation*
9. *Will be offering customers an option to opt-out of AI augmented processing*

Key Takeaways

- Automated decision making and AI Ethics has caught the attention of lawmakers. *(Many privacy laws have provision for automated decision making and/or direct marketing – EU GDPR, CN PIPL, PH DPA)*

- Government Agencies, AI Users and Tech Firms must collaborate. *(Get it right from the gate. Opportunity to influence AI use at industry level and not caught off guard when rules are implemented)*
- Organisations planning to deploy AI need clear guidelines *(Establish/adopt a structured approach to assessing AI risks)*

Thank you

FREEBIES

The SG Model AI Governance Framework - full document

- https://www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files/Resource-for-Organisation/AI/SGModelAIGovFramework2.pdf

The SG Model AI Governance Framework – Implementation & Self-Assessment Guide

- https://www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files/Resource-for-Organisation/AI/SGIsago.pdf

AI failures - editorial

- https://www.jumpstartmag.com/ai-gone-wrong-5-biggest-ai-failures-of-all-time/
- https://www.cnet.com/tech/tech-industry/buy-a-scale-on-amazon-and-it-thinks-youre-a-drug-dealer/
- https://www.theatlantic.com/technology/archive/2014/04/the-unintentional-amazon-guide-to-dealing-drugs/360636/