Singapore’s AI Governance and Ethics initiatives

Yeong Zee Kin
Assistant Chief Executive (Data Innovation & Protection), Infocomm Media Development Authority
Deputy Commissioner, Personal Data Protection Commission
AI DEPLOYMENT CHAIN

**Data Preparation**

Stage 1: **Raw Data** is formatted and cleansed so conclusions can be drawn accurately. Generally, accuracy and insights increase with relevance and the amount of data.

**Algorithms**

Stage 2: **Algorithms** are applied for analysis. This includes statistical models, decision trees, neural networks. The results are examined and algorithms are re-iterated until a model that produces the most useful results emerges.

**Chosen Model**

Stage 3: **Chosen model** is used to produce probability scores, that can be incorporated into applications to make decisions, solve problems and trigger actions.
OVERVIEW OF ISSUES FOR EACH PLAYER

**USER COMPANY**
- Software licence or subscription
- Product liability
- Consumer protection

**APPLICATION DEVELOPER**
- Software development contract

**MODELS**
- Data protection
- Development & maintenance contract

**AI ENGINE PROVIDER**
- Third party software licence

**INFRASTRUCTURE PROVIDER**
- Cloud services agreement
**Contractual Liability for Services**
- How to incorporate AI is a design decision
  - Liability depends on whether AI augments employee decision-making or to make decisions autonomously
- Reliance on service providers in data preparation, model training and selection
  - Professional care and skill
  - Back-to-back support for explanation

**Vicarious) Responsibility for Decision**
- Augmentation of employee decision-making
- Automation of organisation’s decision-making
- Intentional vs. unintentional discrimination
- Decision-making models, i.e. human in the loop, human over the loop, human out of the loop

**Consumer Protection Considerations**
- Services covered under the Consumer Protection (Unfair Trading) Act
- Companies dealing as a consumer in software supply agreements
- Requirement of reasonableness for exclusion or limitation of liability for negligence under Unfair Contract Terms Act

**IMPROPER USE OF AI-EMPOWERED FUNCTIONALITY**
**TRANSPARENCY OF DECISION MAKING PROCESS; LEGAL OR ETHICAL OBLIGATIONS?**
**AI ENGINE PROVIDER**

**Third Party Software Licence**
- AI engine provider’s responsibilities, e.g. ensure AI engine is fit for purpose, obligation to fix bugs and errors, provide explanation of how AI works
- User company responsible for proper use of AI
- Trained or fitted models are “computer programs” that expresses a set of instructions & vectors/weights derived from training dataset

**Algorithmic Transparency or Explainability**
- Is this a legal or ethical obligation? Can algorithms be biased or is the bias in the dataset?
- Explanation of algorithm decision of decision-making process?
- Not all AI engines or models can be explained (rule-based vs blackbox)

**Benefitting from Algorithm Audits – Regulator or Consumer?**
- Is there potential for independent certification standards for algorithmic transparency?
- Extracting a contractual promise from AI engine provider to provide to regulator
When is data personal?
- Data about an individual who can be identified from that data, or that data with other information
- Personalisation of services will likely involve personal data

Who is the data controller?
- In-house or 3rd party data source
- Due diligence for 3rd party sourced data (consent for disclosure and intended use, data quality, accuracy and provenance)

Consent and purpose limitation
- Declared data (provided by users through form filling)
- Observable data (generated through user activity)
- Inferred data (profile information for digital marketing)
- Express or deemed consent

• RE-PURPOSING OR SECONDARY USE (REFRESHING CONSENT; DEEMED CONSENT BY NOTIFICATION)
• ANONYMISATION OF PERSONAL DATA
RISK-BASED DECISION-MAKING MODELLING
MACHINE BIAS- CASE STUDY BY PROPUBLICA

- The Justice Department’s National Institute of Corrections encourages combined assessments with algorithms at every stage of criminal justice process
- Propublica examined the effect of machine algorithms on risk assessments of defendants
- The risk assessments results from the algorithms proved unreliable in forecasting violent crime
  - Only 20% of people predicted to commit violent crimes actually went on to do so
  - There are significant racial disparities
    - More likely to falsely flag black defendants as future criminals, wrongly labelling them this way at almost twice the rate as white defendants
    - White defendants were mislabeled as low risk more often than black defendants
  - Despite isolating the effect of race from criminal history and recidivism, as well as from defendants’ age and gender, black defendants were still 77% more likely to pegged as at higher risk of committing a future violent crime and 45% more likely to be predicted to commit a future crime of any kind

Source: ProPublica analysis of data from Broward County, Fla.
DECISION MAKING MODELS

- **Human oversight is active and involved**, and retains full control
- **AI only provides recommendations or input**
- **Decisions cannot be exercised without affirmative actions by the human**

- **No human oversight** over the execution of decisions
- **AI has full control** without the option of human override

- **Allows humans to adjust parameters during the execution of the algorithm**
### Risk Analysis Model

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>High severity</td>
<td>Low probability</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low severity</td>
<td>High probability</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low severity</td>
<td>High probability</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- To determine the level of human oversight/ intervention in an organisation's decision making process involving AI
- Classify the probability and severity of harm to an individual as a result of the decision made by an organisation about that individual using AI
- Definition of harm and the computation of probability and severity depends on the context and vary from sector to sector
AI REGULATORY MODEL
A TRUSTED ECOSYSTEM FOR AI ADOPTION
THROUGH GOVERNANCE AND ETHICS

1. BRINGING RELEVANT STAKEHOLDERS TOGETHER TO BUILD A TRUSTED ECOSYSTEM
2. SUPPORTING AI ADOPTION THROUGH VOLUNTARY CORPORATE GOVERNANCE FRAMEWORKS THAT PROMOTE RESPONSIBLE DATA USE
3. FUNDING RESEARCH TO IDENTIFY AND CREATE SOLUTIONS FOR LEGAL, REGULATORY AND POLICY ISSUES AS AI ADOPTION BROADENS
BUILD ENVIRONMENT OF TRUST AND LEAD IN SAFE AND PROGRESSIVE USE OF AI

MODEL AI GOVERNANCE FRAMEWORK

- Accountability-based framework
- Enables discussion of ethical, governance & consumer protection issues
- Model framework for voluntary adoption by businesses

ADVISORY COUNCIL ON ETHICAL USE OF AI AND DATA

- Provides guidance on complex ethical issues
- AI technology providers, user companies & consumer interests representatives
- Hosts conversations with industry & consumers
- Effective barometer of business needs

RESEARCH PROGRAMME ON THE GOVERNANCE OF AI AND DATA USE

- Build up body of knowledge for legal, policy & governance issues
- Develop a pool of experts
- Complement scientific AI research & professional training for robust AI ecosystem
A BALANCED GOVERNANCE FRAMEWORK THAT ENGENDERS TRUST WHILE PROVIDING ROOM TO DEVELOP & INNOVATE

An **accountability-based framework** that promotes responsible use of AI in decision-making, addresses ethical risks, and builds consumer trusts in order to support commercial deployment of AI.

**#1** Integrating AI ethics into corporate governance and risk management structures e.g. corporate values, risk management frameworks, decision-making and risk assessment

**#2** Translating responsible AI from principles into processes e.g. data curation, addressing data bias, responsibilities in AI model selection, unintended discrimination, model tuning

**#3** Establishing good consumer interactions e.g. AI-human interactions, managing customer-relations when automating decision-making, explaining decision-making process

Champion of World Summit on Information Society (WSIS) Prizes 2019 in the category Ethical Dimensions of the Information Society
A PRO-INNOVATION & PROGRESSIVE MODEL FRAMEWORK

Internal Governance Structures & Measures
- Accountability-based framework
- Discussion paper on proposed framework that will be converted into a model framework for voluntary adoption
- Contrast EU GDPR, Art 21 – 22
- Right not to be subject to automated decision making (including profiling)
- Safeguards of data subject’s rights and freedoms and legitimate interests includes:
  - specific information to the data subject
  - right to obtain human intervention
  - right to express his or her point of view
  - right to obtain an explanation of the decision reached after such assessment
  - right to challenge the decision

Model Framework Available From

www.pdpc.gov.sg/model-ai-gov
AI & SOCIETY: PROGRAMMES TO INSPIRE CONFIDENCE

**ORGANISATIONS**
- Understand AI Governance
- Establish best practices to guide safe and ethical management of AI systems
- Build consumer trust in AI deployments
- Prepare workforce (Training/Upskill/Re-skill)

**COMMUNITY**
- Promote Use: ICM Learning Roadmap for students
  - Real-life case studies for each field
  - Understand advisory guidelines on the ethical use of AI and data
  - Know the code of practice for AI devt

Deeper understanding of AI:
- Relevant success use cases
- Importance of data
- Implications of AI at workplace

**UNDERSTAND**
AI and its benefits

**KNOW**
Awareness of AI
Basic knowledge and concepts of AI
De-Mystify AI

- What is AI
- What can it do
- Debunk myths of AI
- Limitations of AI

**Build Trust & Confidence to USE**

- Disruption of AI
- Business Transformation driven by AI