Big Data, Artificial Intelligence and Privacy

Stephen Kai-yi Wong, Barrister
Privacy Commissioner for Personal Data, Hong Kong
Big Data and AI

- Banking
- Education
- Retail
- Transportation
- Government Operations
Big Data and AI

- driving forces of “Smart City”
Big Data and AI

- what are the privacy implications?
Big Data and AI

Sensors

RFIDs

Cookies

Cameras

Big Data & AI
Big Data and AI

- our digital footprints (social media data, instant messages, email...) create rich sources of Big Data to fuel AI algorithms
Big Data & AI

- AI algorithms find correlations between data to make predictions and generate new knowledge
Data Risks and Challenges
(1) Covert Data Collection

- massive and ubiquitous data collection from multiple sources
- tracking online and offline
- data subject may be unaware of data collection and use
- Meaningful notice & consent?
(1) Covert Data Collection

- your digital virtual assistants may be eavesdropping....
(2) Unexpected Data Use

- companies analyse innocuous data to predict intimate and sensitive data
- correlations (not causality)
- users may be surprised by predictions
(2) Unexpected Data Use

- researchers use algorithms to analyse “likes” to infer sensitive personal particulars including religious and political beliefs, race and sexual orientation
(3) Re-identification

- linkage of datasets destroys anonymity

- US researchers successfully combined anonymised movie ratings dataset with public movie ratings dataset and re-identified movie subscribers, unveiled personal preference, political affiliation and sexual orientation
(4) Profiling, Unfairness and Discrimination

- profiling to infer or predict individuals’ preferences, health, work performance, credit worthiness, propensity to crime...
(4) Profiling, Unfairness and Discrimination

- loan institutions lowered people’s credit limits based on poor repayment histories of other people shopping at same stores
- fairness?
(4) Profiling, Unfairness and Discrimination

- search engine generates more advertisements associated with arrest records when users input “black-sounding” names than “white-sounding” names
- internalise societal biases
(4) Profiling, Unfairness and Discrimination

- transport app profiles its customers and denies them of services if they appear to be law enforcement agents
(5) Unpredictability

- self-evolving logic
- not following logic of engineers
- DeepMind’s AI agent learns to play 49 classic video games with minimal human input
(6) Transparency

- may not be able to specify data collection purposes for Big Data
- low transparency
- “black box” algorithms, opaque and complicated
(6) Transparency

- what if AI Judge gives you unfavorable ruling without reasons?
- what if you are denied bank account opening by anti-money laundering assessment programme without reasons?
- civil rights and freedom
(7) Filter Bubble

- news feeds and video recommendations may be extremely customised
- “filter bubble” or “echo chamber effect”
(8) Control by Big Data and AI

- people controlled by Big Data and AI
- “Big Nudging”: Big Data and Nudging
- no democratic process
- top-down controlled, totalitarian society
Privacy-based Solutions
Accountability

- “Protect, Respect Personal Data”
- top management cultivates respect for privacy within organisations
- adopt measures to protect data
- Privacy Management Programme
- “Privacy by Design” & “Privacy by Default”
Accountability

Privacy Management Programme

- paradigm shift from compliance to accountability
- consider privacy at each stage of the data lifecycle
Accountability

- Privacy Management Programme: A Best Practice Guide

Accountability

“Privacy by Design” & “Privacy by Default”

- embed privacy in all business processes rather than an afterthought
- protect privacy right and maximise value of data
EU General Data Protection Regulation 2018:

- stringent consent requirements
- increased accountability of data controllers
- reveal logic behind automated decisions
Transparency

- **Two “Ts” – Transparency and Trust**
- transparency builds up trust
Transparency

- explain **what data is collected and purposes of use**
- explain **logic and rationale behind decision**
Meaningful Choices and Control

- allow individuals to object to profiling
- allow individuals to object to decisions which may significantly affect them
Code of Ethics for AI

- guide the use of AI
- ensure fair and just processing
- harm to liberty and freedom should be justified
Protect, Respect Personal Data

- Personal Data (Privacy) Ordinance is technology neutral; principle-based
- balance between privacy and free flow of information
- keep individuals informed of collection and use of data, and obtain meaningful express consent
39th International Conference of Data Protection and Privacy Commissioners

Thank You!