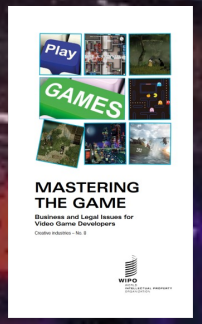
AI, Law and Governance
Lecture 3

Dr Michaela MacDonald
Lecturer, The School of Electronic Engineering and Computer Science
Queen Mary University of London

Queen Mary University of London
Centre for Commercial Law Studies
MORE THAN JUST A GAME



Data and AI

175 zettabytes

The volume of data produced in the world is expected to grow from 33 zettabytes in 2018 to 175 zettabytes in 2025 (one zettabyte is a thousand billion gigabytes)

Data – A core element of AI

- Structured / processed data (tagged, classified, integrated with metadata)
- Unstructured / raw data
- Metadata
- Free data / proprietary data
- Issues of accessibility, transparency, authenticity, control
- Data not protectable as such
  - Presentation or collection of data potentially protected by copyright, database rights, trade secrets
  - Contracts and license agreements

Data and AI in practice

- Spotify uses AI algorithms that estimate personal preferences based on listening habits. Every time you listen to music, the system notes your preferences and suggests songs similar to the genres you enjoy. Moreover, it enhances business efficiency by creating targeted advertisements.
- Amazon uses algorithms to analyse customer browsing behaviour, purchase history, and even the time spent looking at specific items, personalising recommendations to encourage further purchases.
- IBM's Watson uses AI to analyse vast amounts of medical data, from research papers to patient records. It assists doctors in diagnosing diseases and suggests treatments. For instance, Watson was used to treating rare forms of cancer by analysing genetic data to identify potential treatments.

Question

Discuss potential privacy risks and challenges in these scenarios.

Privacy risks and challenges

- Data Breaches
- Algorithmic Bias
- Privacy-preserving AI
- Informed Consent
- Third-party Data Sharing
- De-identification and Re-identification
- Global Regulatory Compliance
- Privacy vs. Utility Trade-off
- Lack of Transparent Explanations
- Surveillance and Facial Recognition

AI and privacy laws

- GDPR, e-privacy regulation, CCPA (California Consumer Privacy Act), Cybersecurity Law (China)
- GDPR – top-bottom principles-based legislation
- Introduced new rights and responsibilities that ensure greater accountability, transparency and control in the way AI applications are used and deployed with regards to personal data
- New rights for data subjects (the right for access, rectification, erasure, to object, not to be subject to automated decision-making or right to data portability) = new obligations for organisations

GDPR – Scope I

- The GDPR regulates the processing of ***personal data***
- ***Personal data*** is any information relating to an identified or identifiable natural person ('data subject')
- ***Identifiable natural person*** is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person
- Relates to living individuals only

Neural Data on Trial: Chile's Supreme Court addresses the First Global Neurorights Case

Micaela Mantegna · Follow

8 min read · Aug 17, 2023

GDPR – Scope II

- ***Special categories of personal data*** is subject to a stricter regime

1. Racial or ethnic origin
2. Political opinions
3. Religious or philosophical beliefs
4. Trade union membership
5. Genetic data
6. Biometric data for the purpose of uniquely identifying a natural person
7. Data concerning health
8. Data concerning a natural person’s sex life or sexual orientation

GDPR – Principles

- Principles-based regulation
- The EU has adopted similar risk-based safeguarding and information obligations in respect of telecommunication networks and payment services, as well as under the NIS Directive and the e-Privacy Directive
- Lawfulness, fairness and transparency
- Purpose limitation
- Data minimisation
- Accuracy
- Storage limitation
- ***Integrity and Confidentiality*** – ensures appropriate security of personal data, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures
- Accountability

GDPR Article 22

Article 22

1. The data subject shall have the right not to be subject to a decision based **solely** on automated processing, including profiling, which produces **legal effects** concerning him or her or **similarly significantly** affects him or her.

Paragraph 1 shall not apply if the decision:

- is necessary for entering into, or performance of, a contract between the data subject and a data controller;
- is authorised by Union or Member State law to which the controller is subject and which also lays down suitable measures to safeguard the data subject's rights and freedoms and legitimate interests; or
- is based on the data subject's explicit consent.

Key concepts

- What does ‘solely’ automated mean?

= Solely means a decision-making process that is totally automated and excludes any human influence on the outcome

A factory worker’s pay is linked to their productivity, which is monitored automatically. The decision about how much pay the worker receives for each shift they work is made automatically by referring to the data collected about their productivity.

Key concepts

- What types of decision have a legal or similarly significant effect?

= A decision producing a **legal effect** is something that affects a person's legal status or their legal rights. For example, when a person, in view of their profile, is entitled to a particular social benefit conferred by law, such as housing benefit.

= A decision that has a **similarly significant effect** is something that has an equivalent impact on an individual's circumstances, behaviour or choices.

A social security process which automatically evaluates whether an individual is entitled to benefit and how much to pay is a decision 'based solely on automated processing' for the purposes of Article 22(1).

GDPR safeguards

- Counterfactuals
- Verification
- Testing and outputs analysis
- Ongoing sampling

CAIDP FTC Complaint against OpenAI

- The FTC complaint identifies issues including:
  - Enhanced risk to cybersecurity
  - Enhanced risk to **data protection and privacy**
  - Enhanced risk to children's safety
  - Failure to conduct independent risk assessment prior to deployment
  - Failure to establish independent risk assessment throughout AI lifecycle
  - Failure to accurately describe data source
  - Failure to disclose data collection practices regarding users
  - False advertising regarding reliability
  - Lack of transparency in outputs produced
  - Replication of bias in protected categories

Federal Trade Commission Act (US)

- Consumer Protection Regulations
- FTC is an independent federal agency and the most important regulatory authority for consumer protection issues
- Section 5 forbids unfair and deceptive trade practices
- The FTC has now brought over 50 information security cases

FTC – Scope

- ***Unfair***
  1. Causes or likely to cause substantial harm/injury to consumer
  2. Consumer cannot reasonably avoid the harm
  3. There is not a benefit to the practice that outweighs the harm
- ***Deceptive***
  1. Representation or omission likely to mislead the consumer
  2. Not reasonable from the perspective of the consumer
  3. Affects consumer's decision; harm as otherwise, likely another decision

Privacy Law (China)

- **The Decision on Strengthening Online Information Protection**, effective from December 28, 2012 (Decision)
- **National Standard of Information Security Technology – Guideline for Personal Information Protection within Information System for Public and Commercial Services**, effective from February 1, 2013 (Guideline)
- **PRC Cybersecurity Law,** effective from June 1, 2017
- **National Standard of Information Security Technology – Personal Information Security Specification**, effective from May 1, 2018 (PIS Specification)
- **PRC Civil Code,** effective from January 1, 2021
- **Personal Information Protection Law** (PIPL), effective from 1 November 2021
- **Measures for the Management of Generative Artificial Intelligence Services** (Draft for Comment) – April 2023
- Decision has the same legal effect as law, while Guidelines and PIS Specifications are references to best practice

China's first facial-recognition lawsuit comes to an end with new ruling and new questions about the fate of individuals' data

Alternative approaches

- WIPO consultation on AI and IP proposes new *sui generis* right for data
- Rationale:
  - New significance that data assumed as a critical component of AI
  - To facilitate the development of new types of data
  - Appropriate allocation of value in relation to data
  - Fair competition
- What would be the impact and benefit of such new right?
- How would it apply to synthetic data – new type of data generated and used by AI systems

Alternative approaches

- Data trusts
  - An alternative approach to top-bottom regulation approach that empowers data subjects to truly control the access and use of their data
  - A legal structure that provides for an independent stewardship of data
  - Data is held on behalf of the data subjects and for their benefit
  - Potential to align privacy, IPRs and contractual limitations
- Synthetic data
  - Synthetic data is artificial data that is generated from original data and a model that is trained to reproduce the characteristics and structure of the original data
  - Synthetic data and original data should deliver very similar results when undergoing the same statistical analysis

Thank you!