

AI, Law and Governance Lecture 2

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



Question

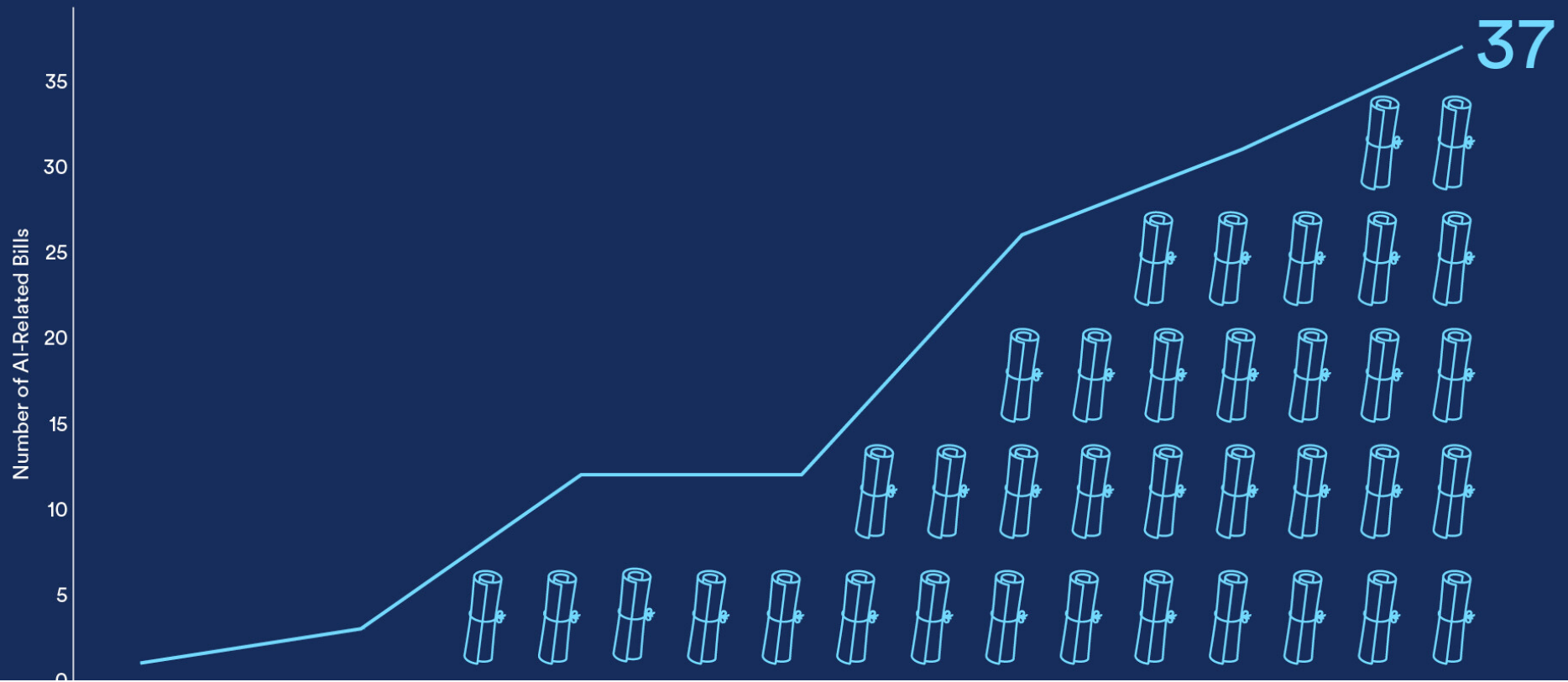
What do you think the main challenges and risks are that AI governance should address?

Go to www.menti.com and use the code **4707 9903**

Challenges and risks

- Based on report Worldwide AI Laws and Regulations, emerging laws and regulations pertain to:
 - the use of facial recognition and computer vision,
 - operation and development of autonomous vehicles,
 - issues of AI-relevant data privacy,
 - challenges arising from conversational systems and chatbots,
 - the emergence of the possibility of lethal autonomous weapons systems (LAWS),
 - concerns around AI ethics and bias,
 - aspects of AI-supported decision making,
 - the potential for malicious use of AI **Technology for profit** vs **technology for the benefit of the society**

Number of AI-Related Bills Passed into Law Globally



AI Governance

- Permissive versus prohibitive laws / no regulation / self-regulation, regulatory guidance / standards / best practice
- Application of hard law and soft law
- Plurality of governance models
- Evaluating their efficiency and suitability with regards to a specific industry and application
- Different approaches across US, China, UK and EU
- How much regulation is enough?
- AI Ethics

Ethics

- Ethics refer to a set of moral principles that govern a person's (entity's) behaviour or the conducting of an activity
- Determines what is right and wrong
- Task: Give an example of a moral principle

AI Ethics

AI Ethics address a broad range of concerns, from privacy and data protection to bias and fairness in AI systems. With AI increasingly integrated into various sectors, including healthcare, finance, and law, ethical considerations become crucial in ensuring these technologies benefit society without infringing on individual rights or exacerbating inequalities. The intersection of AI with intellectual property rights, particularly in creative industries, poses novel challenges, as does the environmental impact of AI technologies.

AI Ethics

- Set of guidelines – soft law
- Number of initiatives looking at providing guidance on how to design, develop and use AI system now and in the future – well over hundred different sources
- There is a broad agreement out there that AI should be ethical and there is an ongoing debate with regards to standards and practices are required to design, develop and deploy ethical AI
 - *House of Lords Select Committee on AI*
 - *White House report*
 - *Singapore Model AI Governance Framework*
 - *China's Guidelines on AI ethics*
 - *Individual companies' ethical guidelines – Google, Tencent, etc.*

Google's AI Principles



1. Be socially beneficial



2. Avoid creating or reinforcing unfair bias



3. Be built and tested for safety



4. Be accountable to people



5. Incorporate privacy design principles

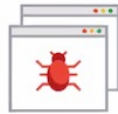


6. Uphold high standards of scientific excellence

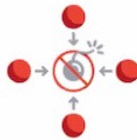


7. Be made available for uses that accord with these principles

Applications we will **not** pursue



likely to cause
overall harm



weapons or
those that direct
injury



surveillance
violating
internationally
accepted norms



purpose
contravenes
international law
and human rights



From Principles to Practice



Responsible

- Work with human rights experts
- Educate and empower customers and partners
- Open development and sharing research



Accountable

- Invite customer feedback
- Ethical Use Advisory Council
- External ethics review of research papers



Transparent

- Strive for model explainability
- Customers control of their data and models
- Clear disclosure of terms of use



Empowering

- Build AI apps with clicks not code
- Free AI education with Trailhead
- Deliver AI research breakthroughs



Inclusive

- Test models with diverse data sets
- Conduct Consequence Scanning Workshops
- Build inclusive teams



AI ethical dilemmas in practice

- Amazon attempted to leverage the HR teams using their AI recruiting and hiring tool. It allows organizations to take thousands of resumes and then select the top 5. It noticed that the system was rating women as candidates for software development and other technical positions differently than men.
- “Apple card,” the AI system of Apple, was offering significantly different interest rates and credit limits to different genders. It is giving large credit limits to men as compared to women.
- Law enforcement agencies using AI to predict where crimes might occur or who might commit them.

EU approach to AI

- General AI strategy
- Excellence in AI
 - Invest into research and innovation
 - Enable commercialisation of AI research
 - The EU Cybersecurity strategy
 - The Digital Services Act and the Digital Markets Act, and the Data Governance Act
- Trust in AI
 - Legal framework for AI to address fundamental rights and safety risks specific to the AI systems;
 - Rules to address liability issues related to new technologies, including AI systems
 - Legal framework for AI

The AI Act – Risk-based approach

4 categories

- **Unacceptable risk:** AI systems considered a clear threat to the safety, livelihoods and rights of people **will be banned**. This includes AI systems or applications that manipulate human behaviour to circumvent users' free will (e.g. toys using voice assistance encouraging dangerous behaviour of minors) and systems that allow 'social scoring' by governments.
- **High-risk** AI systems will be subject to **strict obligations** before they can be put on the market

The AI Act – Risk-based approach

- **High-risk:** AI systems identified as high-risk include AI technology used in:
 - **Critical infrastructures** (e.g. transport), that could put the life and health of citizens at risk;
 - **Educational or vocational training**, that may determine the access to education and professional course of someone's life (e.g. scoring of exams);
 - **Safety components of products** (e.g. AI application in robot-assisted surgery);
 - **Employment, workers management and access to self-employment** (e.g. CV-sorting software for recruitment procedures);
 - **Essential private and public services** (e.g. credit scoring denying citizens opportunity to obtain a loan);
 - **Law enforcement** that may interfere with people's fundamental rights (e.g. evaluation of the reliability of evidence);
 - **Migration, asylum and border control management** (e.g. verification of authenticity of travel documents);
 - **Administration of justice and democratic processes** (e.g. applying the law to a concrete set of facts).

The AI Act – Risk-based approach

- **Limited risk**, i.e. AI systems with specific transparency obligations:
When using AI systems such as chatbots, users should be aware that they are interacting with a machine so they can take an informed decision to continue or step back.
- **Minimal risk**: The legal proposal allows the free use of applications such as AI-enabled video games or spam filters. The vast majority of AI systems fall into this category.

US approach to AI

- The American AI initiative (2018) ; White House report on the AI Initiative (2020)
- Implementation of the OECD principles of AI
- National Defense Authorization Act (2021)
- Algorithmic Accountability Act (2022) (not passed)
- 'Blueprint for an AI Bill of Rights: Making automated Systems Work for the American People' (2022)
 - Five principles: safe and effective systems, algorithmic discrimination protections; data privacy; notice and explanation; and human alternatives, consideration, and fallback
- *State v Loomis*

The UK

- House of Lords Select Committee on Artificial Intelligence Report (2018)
- National AI Strategy (September 2021) sets out a 10-year plan to make the UK a 'global AI superpower'
 - Developing a pro-innovation regulatory and governance framework that protects the public
 - Examples include the new **AI Standards Hub**, the **AI Assurance roadmap**, and **Algorithmic Transparency Recording Standard**
- AI Action Plan (July 2022) outlines the activities being taken by each government department to advance the National AI Strategy

Singapore

- The Personal Data Protection Commission (PDPC) established in 2013, the main authority' on data protection, including the governance of AI
 - Discussion Paper on Artificial Intelligence and Personal Data (2018)
 - Model Framework (1st ed 2019, 2nd ed 2020)
 - Two principles:
 - a) decisions made by AI systems should be “explainable, transparent, and fair,” and
 - b) AI systems should be human-centric
- A flexible, pro-active approach to risk management aimed at building trust

The role of standards

Standard is...

= established or widely recognised as a model of authority or excellence (a standard reference work)

- Standards are to be developed to provide clear guidelines when designing AI systems
- Following such standards leads to a presumption of conformity, bringing down the cost for compliance and uncertainty
- Professional codes / accreditations

Key points

- Ethical guidelines describe high-level ethical principles, tenets, values, or other abstract requirements for AI development and deployment
- They need to be ‘translated’ into mid- or low-level design requirements and technical fixes, governance frameworks, and developer codes of ethics – standards will be essential
- AI has the potential to shape nearly every aspect of human life, it is therefore crucial that the industry fosters the public’s trust

Thank you!