

# Artificial Intelligence and the Law

## Regulating AI

Dr. Adam Wyner

School of Law and Department of Computer Science

Swansea University, Swansea, United Kingdom

# What is it to Regulate AI?

- A presentation about *Responsible AI: A Global Policy Framework*:
- <https://www.itechlaw.org/ResponsibleAI>
- Based on *Responsible AI Policy Framework* summary:
- [https://www.itechlaw.org/sites/default/files/ResponsibleAI\\_PolicyFramework.pdf](https://www.itechlaw.org/sites/default/files/ResponsibleAI_PolicyFramework.pdf)

# Principles for Organisations and Governments

- Ethical purpose and societal benefit
- Accountability
- Transparency and explainability
- Fairness and non-discrimination
- Safety and reliability
- Open data and fair competition
- Privacy
- AI and intellectual property

# Principle 1

## Ethical Purpose and Societal Benefit

Organisations that develop, deploy or use AI systems and any national laws that regulate such use should (O\_should) require the purposes of such implementation to be identified and ensure that such purposes are consistent with the overall ethical purposes of beneficence and non-maleficence, as well as the other principles of the Policy Framework for Responsible AI.

If it is about what Governments should do, then we abbreviate to G\_should.

# Overarching Principles

- O\_should:
  - do so in a manner compatible with human agency and the respect for fundamental human rights;
  - monitor the implementation of such AI systems and act to mitigate against consequences of such AI systems (whether intended or unintended) that are inconsistent with the ethical purposes of beneficence and non-maleficence.
  - assess the social, political and environmental implications of such development, deployment and use in the context of a structured Responsible AI Impact Assessment that assesses risk of harm and, as the case may be, proposes mitigation strategies in relation to such risks.

# Work and automation

- O\_should:
  - provide opportunities for affected employees to participate in the decision-making process.
  - consider whether they can ensure that all possible occurrences can be pre-decided to ensure consistent behaviour. If not, can they confine the decision outcomes of an AI system to a reasonable, non-aberrant range of responses, taking into account the wider context, the impact of the decision and the moral appropriateness of “weighing the unweighable” such as life vs. life.
  - conduct a Responsible AI Impact Assessment to determine the net effects of such implementation.
- G\_should
  - monitor to determine where human workers are the most affected.
  - promote educational policies that equip all children with the skills, knowledge and qualities.
  - encourage the creation of opportunities for adults to learn new useful skills.
  - study the viability and advisability of new social welfare and benefit systems to help reduce socio-economic inequality.

# Environmental Impact

- O\_should:
  - assess the overall environmental impact of such AI systems
  - seek to promote and implement uses of AI systems with a view to achieving overall carbon neutrality or carbon reduction.
  - carry out research about how to reduce environmentally harmful resource waste and inefficiencies.
- G\_should
  - adjust regulatory regimes and/or promote industry self-regulatory regimes concerning market-entry and/or adoption of AI systems in a way that the possible exposure (in terms of 'opportunities vs. risks') that may result from the public operation of such AI systems is reasonably reflected.
  - ensure comprehensive and transparent investigation of such adverse and unanticipated outcomes of AI systems that have occurred through their usage.

# Weaponised AI

- Weaponised AI should respect the principles and standards of international humanitarian law on the use of weapons and human rights law.
- G\_should implement multilateral mechanisms to define, implement and monitor compliance with international agreements.
- Weaponised AI that selects and engages with targets should be under human control and oversight.
- O\_should inform their employees when they work on projects relating to weaponised AI.



# Weaponised Information

- O\_should, if they process internet content:
  - minimise the spread of false or misleading information where there is a material risk that such false or misleading information might lead to significant harm to individuals, groups or democratic institutions.
  - be encouraged to identify unlawful content.
  - provide a mechanism which flags potentially harmful content in a timely manner.
- G\_should:
  - provide guidelines to help Organisations identify prohibited content.
  - see that courts remain the ultimate arbiters of lawful content.

## Principle 2

### Accountability

O\_should respect and adopt the eight principles of this Policy Framework for Responsible AI (or other analogous accountability principles). In all instances, humans should remain accountable for the acts and omissions of AI systems.

# Accountability

- O\_should:
  - designate an individual or individuals who are accountable for the organisation's compliance with those principles.
  - make the designated individual known upon request
  - implement policies and practices to give effect to the principles by implementing
    - a “Responsible AI Impact Assessment” process;
    - “Responsible AI by Design” principles;
    - a complaints and inquiries procedure;
    - training about the organisation's policies and practices; and
    - explanations of the organisation's policies and procedures.

# Governments and Context

- G\_should, where they assess the potential for “accountability gaps” in existing legal and regulatory frameworks, adopt a balanced approach that encourages innovation while militating against the risk of significant individual or societal harm.
- Legal and regulatory frameworks should promote the eight principles of the Policy Framework for Responsible AI.
- G\_should not grant distinct legal personality to AI systems, as doing so would undermine the fundamental principle that humans should ultimately remain accountable for the acts and omissions of AI systems.
- Contextuality: the intensity of the accountability obligation will vary according to the degree of autonomy and criticality of the AI system.

# Principle 3

## Transparency and Explainability

O\_should ensure that, to the extent reasonable given the circumstances and state of the art of the technology, such use is transparent and that the decision outcomes of the AI system are explainable.

# Definitions

- Transparency is an obligation for organisations that use AI in decision-making processes to provide information regarding:
  - a) the fact that an organisation is using an AI system in a decision-making process;
  - b) the intended purpose(s) of the AI system and how the AI system will and can be used;
  - (c) the types of data sets that are used by the AI system; and
  - (d) meaningful information about the logic involved.
- Explainability is an obligation for organisations that use AI in decision-making processes to provide accurate information in humanly understandable terms explaining how a decision/outcome was reached by an AI system.

# Purpose

- preserve the public's trust in AI systems and provide sufficient information to help ensure meaningful accountability of an AI system's developers, deployers and users, and to demonstrate whether the decisions made by an AI system are fair and impartial.
- support the Accountability principle, the Fairness and Non-Discrimination principle, the Safety and Reliability principle and the Privacy, Lawful Use and Consent principles.

# Gradual and contextual approach

- The scope and intensity of the obligations of transparency and explainability will depend on the context of the decision and its consequences for the person subject to it.
- The scope and intensity of the obligations of transparency and explainability must balance the interests of the person subject to the decision and the interests of the organisation making the decision. The ultimate criteria shall be the reasonable expectations of a person subject to that type of decision.



# Transparency and explainability by design

- O\_should:
  - ensure that the system logic and architecture serves to facilitate transparency and explainability requirements.
- Users of AI systems and persons subject to their decisions must have an effective way to seek remedy in the event that organisations that develop, deploy or use AI systems are not transparent about their use.

# Technological neutrality

- The use of an AI system by a public or private organisation does not reduce the procedural and substantive requirements that are normally attached to a decision when the decision-making process is completely controlled by a human.

# Principle 4

## Fairness and Non-Discrimination

O\_should ensure the nondiscrimination of AI outcomes, and shall promote appropriate and effective measures to safeguard fairness

# Awareness and education

- Awareness and education on the possibilities and limits of AI systems is a prerequisite to achieving fairer outcomes.
- O\_should ensure that users are aware that AI systems reflect the goals, knowledge and experience of their creators, as well as the limitations of the data sets that are used to train them.

# Technology and fairness

- Decisions based on AI systems should be fair and non-discriminatory. They should be judged against the same standards as decision-making processes conducted entirely by humans.
- The use of AI systems should not exempt or attenuate the need for fairness.
- Users of AI systems and persons subject to their decisions must have an effective way to seek remedy in discriminatory or unfair situations generated by biased or erroneous AI systems and to obtain redress for any harm.

# Development and monitoring of AI systems

- AI development should be designed to prioritise fairness by addressing algorithms and data bias.
- O\_ should establish ethics boards and codes of conduct as well as adopt industry-wide standards and internationally recognised quality seals.
- AI systems should be independently reviewed and tested on a periodic basis.
- Particular attention should be paid to disadvantaged groups which may be incorrectly represented in the training data.

# A comprehensive approach to fairness

- Addressing fairness in AI use requires:
  - the close engagement of technical experts from AI-related fields with statisticians and researchers from the social sciences; and
  - a combined engagement between governments, organisations that develop, deploy or use AI systems and the public at large.
- The Fairness and Non-Discrimination Principle is supported by the Transparency and Accountability Principles.

# Principle 5

## Safety and Reliability

O\_should adopt design regimes and standards ensuring high safety and reliability of AI systems on one hand while limiting the exposure of developers and deployers on the other hand.



# Require or define explicit ethical and moral principles

- G and O\_ should define the relevant set of ethical and moral principles underpinning the AI system to be developed, deployed or used taking into account all relevant circumstances. A system designed to autonomously make decisions will only be acceptable if it operates on the basis of clearly defined principles and within boundaries limiting its decision-making powers.
- G and O\_ should validate the underpinning ethical and moral principles as defined periodically to ensure on-going accurateness.

# Standardisation of behaviour

- G and O\_ should recall that ethical and moral principles are not globally uniform but may be impacted e.g., by geographical, religious or social considerations and traditions. To be accepted, AI systems might have to be adjustable in order to meet the local standards in which they will be used.
- G and O\_ should consider whether all possible occurrences should be pre-decided in a way to ensure the consistent behaviour of the AI system, the impact of this on the aggregation of consequences, and the moral appropriateness of “weighing the unweighable” such as life vs. life.

# Ensuring safety, reliability and trust

- G\_should require and organisations should ensure that AI systems adhere to the ethical and moral principles, have been trained with data which are curated, and are as 'error-free' as practicable.
- G\_should adjust regulatory regimes and/or promote industry self-regulatory regimes for allowing market-entry of AI systems in order to reasonably reflect the positive exposure that may result from the public operation of such AI systems. Special regimes for intermediary and limited admissions to enable testing and refining of the operation of the AI system can help to expedite the completion of the AI system and improve its safety and reliability. (???)
- G\_should implement rules that ensure comprehensive and transparent investigation of such adverse and unanticipated outcomes of AI systems that have occurred through their usage.

# Facilitating technological progress at reasonable risks

- G\_should consider whether existing legal frameworks such as product liability should be adjusted for AI systems.
- G\_should support and participate in international standardisation bodies for the development and deployment of safe and reliable AI systems.

# Principle 6

## Open Data and Fair Competition

O\_should promote

- (a) open access to datasets which could be used in the development of AI systems and
- (b) open source frameworks and software for AI systems.

AI systems must be developed and deployed on a “compliance by design” basis in relation to competition/antitrust law.

# Supporting effective competition in relation to AI systems

- G\_should engage with international co-ordination (Organisation for Economic Co-operation and Development and the International Competition Network) to develop best practices and rigorous analysis in understanding the competitive impact of dataset control and AI systems on economic markets.
- G\_should regularly review competition law frameworks and the enforcement tools so they are sufficient and effective to ensure access to necessary inputs and foster adequate choice.

# Open data

- G\_should foster and facilitate national infrastructures necessary to promote open access to datasets, allowing free access for academic and research purposes and paid-for access for commercialised purposes.
- G-should support open data initiative and widen the understanding of the advantages to be gained from open access data, the structures through which datasets can be shared and exchanged, and the processes by which data can be made suitable for open access.
- G\_should ensure that the data held by public sector bodies are accessible and open, where possible. Private sector bodies should similarly support and promote open data within their industry sector.
- O\_should be encouraged to open up access to, and/or license, their datasets.

# Open source AI systems

- Organisations that develop AI systems are normally entitled to commercialise such systems as they wish.
- G\_ should advocate accessibility through open source to innovative AI systems which may be of particular societal benefit or advance the “state of the art”.
- Where an Organisation does not license an AI system as a whole, they ought to be encouraged to license as many of such re-usable components as is possible.



# Compliance by design with competition/antitrust laws

- O\_ should ensure consistency with the overarching ethos of subsisting competition/antitrust regimes to promote free and vibrant competition amongst corporate enterprises to the ultimate benefit of consumers.

# Principle 7

## Privacy

O\_should endeavour to ensure that AI systems are compliant with privacy norms and regulations, taking into account the unique characteristics of AI systems, and the evolution of standards on privacy.

# Finding a balance

- There is a conflict between the increasing use of AI systems to manage private data and regulatory protection to private data.
- G\_should regulate the privacy with respect to AI systems, acknowledging the characteristics of AI and not stifling innovation.
- O\_should analyse their current processes to identify whether they need be updated or amended to respect privacy.

# The operational challenges ahead for AI users

- O\_should make provision for alternative lawful bases for the collection and processing of personal data by AI systems.
- O\_should implement operational safeguards to protect privacy such as privacy by design principles.
- O\_should appoint an AI Ethics Officer to consider the ethics and regulatory compliance of their use of AI.

# AI as a tool to support privacy

- AI technologies could also be used to help organisations comply with privacy obligations.

# Principle 8

## AI and Intellectual Property

O\_should protect the rights in the resulting works through appropriate and directed application of existing intellectual property rights laws.

G\_should investigate how AI-authored works may be further protected, without seeking to create any new IP right at this stage.

# Supporting incentivisation and protection for innovation

- There must be incentivisation and protection for innovation if it is to attract investment and be brought to the greater good of society.
- O\_should be allowed to protect rights in works resulting from the use of AI, whether AI-created works or AI enabled works.
- There should not be overprotection, as this could prove detrimental to the ultimate goal of IP protection.

# Protection of IP rights

- At present IP laws are insufficiently equipped to deal with the creation of works by autonomous AI.
- O\_ should protect the rights in the resulting works. Where appropriate these steps should include asserting or obtaining copyrights, obtaining patents, when applicable, and seeking contractual provisions to allow for protection as trade secrets and/or to allocate the rights appropriately between the parties.



# Development of new IP laws

- G\_should be cautious with revising existing IP laws.
- G\_should explore the introduction of appropriate legislation to clarify IP protection of AI enabled as well as AI created works, without seeking to create any new IP right at this stage.
- When amending existing or implementing new IP laws, governments should seek adequately to balance the interests of all relevant stakeholders.
- G\_should reach a consensus about AI and IP rights to allow for unhindered data flows across borders and the rapid dissemination of new technologies.