

Sharing Is Caring, But AI Don't Care – Assessing Guardrails for Sharing Data In Large AI Datasets

Rohan Massey
Ropes & Gray

Marc Groman
Groman Consulting Group

Patricia Martín-Marrero
Takeda

Sajai Singh
JSA Law

Stephen Burns
Bennet Jones

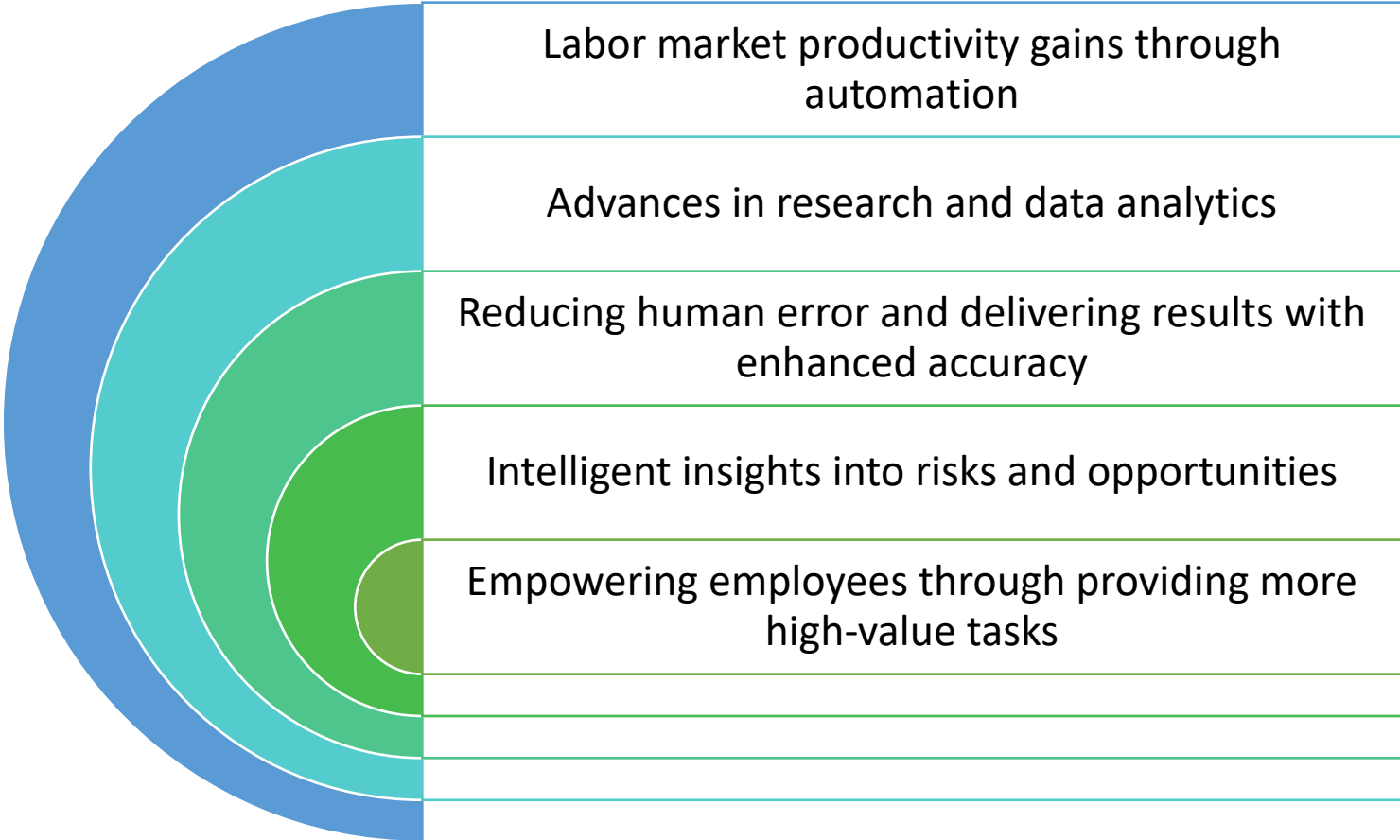
Agenda

- 1. Potential Societal Benefits of Responsible AI**
- 2. Legal and Regulatory Issues of AI**
- 3. Potential Harms and Risks**
- 4. Responsibility and Accountability v Balance and Context**
- 5. What Should We Not be Sharing and Why?**
- 6. Takeaways**

1. Potential Societal Benefits of Responsible AI

Potential Societal Benefits of AI

AI could contribute more than \$15 trillion to the global economy by 2030 (PwC)



Potential Sectoral Benefits of AI

1. Healthcare

- AI can automate tasks in healthcare and has the potential to make rapid, informed medical decisions.
- Potential to improve cost of care and health outcomes.
- Areas of focus for AI include digital biomarkers to monitor human behavior, detection of cancers and diseases, and interventions such as AI clinicians.
- UK government announced a £100 million fund to capitalize on AI's potential in life sciences and healthcare, and the NHS has launched an AI laboratory.

2. Environment

- AI can contribute against climate change by assessing climate patterns to forecast weather, identify which icebergs are melting and how fast, and help predict climate disasters. But this may be a double-edged sword given environmental implications associated with resource-heavy computing demands.

3. Development

- AI can help alleviate poverty, through improving access to education and healthcare, improve efficiencies in productivity (such as agriculture technology), and assist zero hunger targets through leveraging insights on resource allocation.
- Stanford Study: AI able to predict poverty based on satellite imagery with 81-99% accuracy.

National Interests

1. Public safety and defense

- AI can sift through data analytics in real time. The U.S. military's "Project Maven" deployed AI to identify abnormal activity through data captured by surveillance.

2. Protection against cyber threats

- AI can improve cyber-defense from state-sponsored cyber threat actors, through malware detection and analysis, vulnerability detection, and threat risk assessments.
- Google: AI tools 70% better at detecting malicious scripts, 300% more effective at identifying files that exploit vulnerabilities, and Google's detection and response team identified time savings of 51% when using generative AI.

3. International competition

- Significant investment into AI by nation states and competition.
- For instance, China's State Council announced it seeks to build a domestic AI industry worth \$150 billion by 2030.

Clearview AI used by Ukraine for national security

Clearview AI provided its platform to over 1,500 Ukrainian officials, who deployed it to identify more than 230,000 Russians in Ukraine, including Russian collaborators.

2. Legal and Regulatory Issues of AI

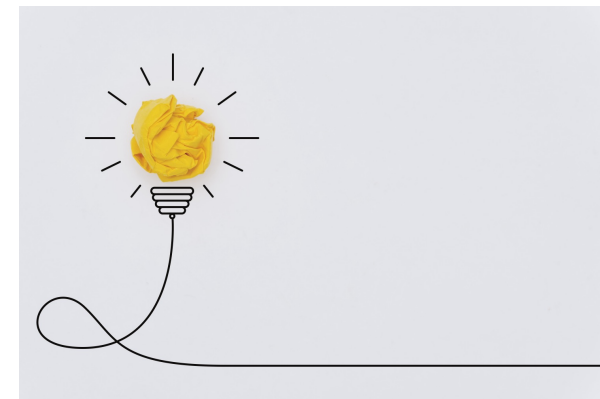
Legal ownership. The law may not provide a definition of the legal owner of AI-generated work. Ownership of work may be claimed by the AI developer rather than the person delivering the prompts.

Software licensing risks. If AI has been trained on licensed code, which is used to generate similar code in a new software, then the new software may need to take into account the licensed code.

Data scraping. AI models reliant on scraped data may infringe privacy rights of individuals whose personal information is used without their knowledge or consent (there are blurred lines between IP and privacy / data protection issues).

Infringement of IP rights in training AI. Training AI models may involve data which is protected and must respect laws on data/IP. In the UK, training of AI models with copyright works is only for non-commercial purposes.

IP concerns from output of AI. Getty Images Inc lawsuits against Stable Diffusion (AI image generating platform) demonstrates IP output risks. In this case, Stable was accused of training its AI with Getty's images, with the AI generating an output which varied Getty's watermark.



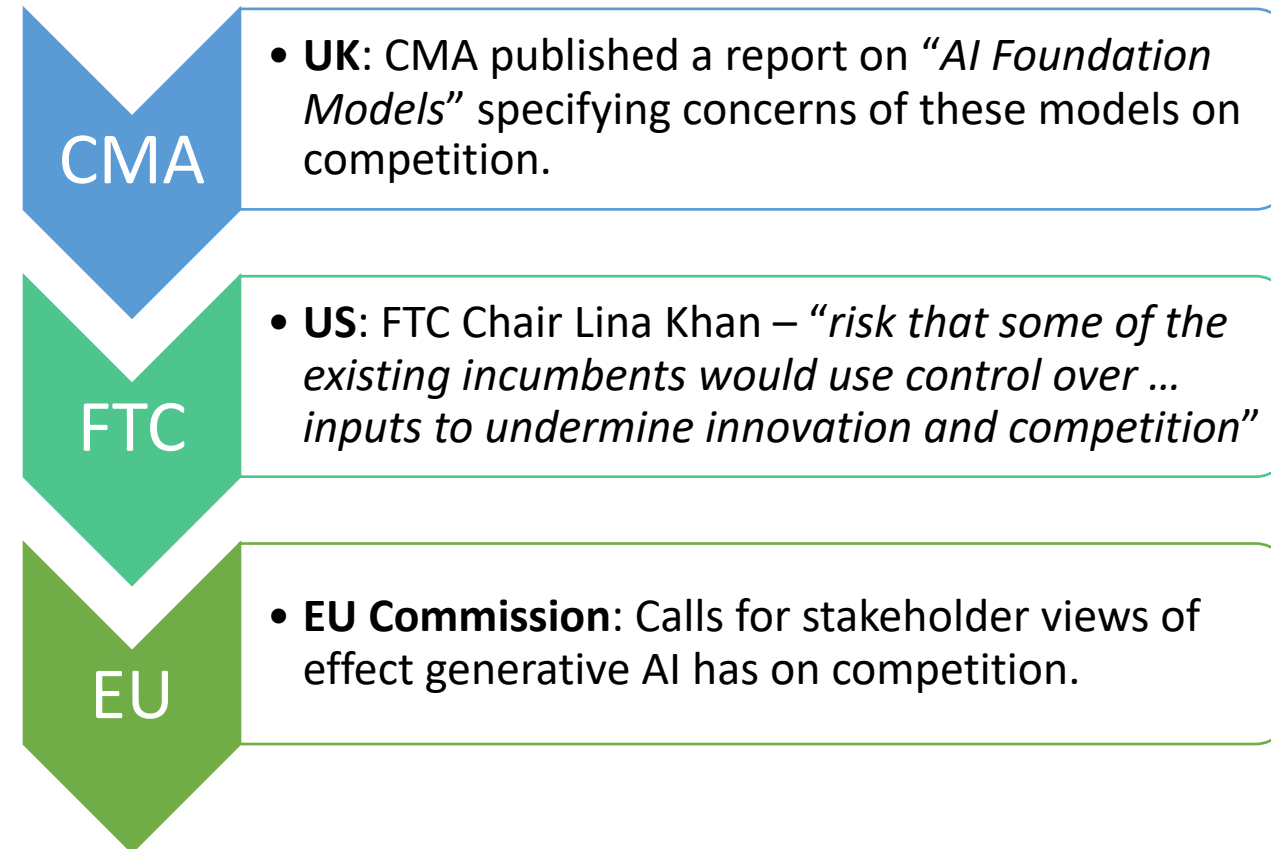
Issues

Production: AI tools can make decisions on output and production value and lead to undue influence resulting in unlawful agreements, such as on division of markets and customers with competitors.

Algorithmic pricing: Training AI algorithms on industry pricing practices can assist companies predict how their competitors are likely to set prices in the future.

Purchasing: AI can be used to make purchasing decisions, such as on volume and pricing. This can lead to price discrimination and collusion. AI may result in concentrated control of key inputs, unlawful collusion, and actions by dominant firms to disadvantage competitors, leading to a reduction in consumer choice, higher prices, and other cascading impacts. High barriers of entry disadvantage SMEs (demand for talent, data center costs).

Regulatory Approaches



Privacy and security are not the same thing and should be addressed separately.

AI models require access to large amounts of data, which may contain sensitive or personal information. If the data is not properly protected, it could be used for improper purposes or be accessed by unauthorized parties. **More than 2/3 of organizations rank data security risks as top AI concern.**

Protection of Sensitive information

- Privacy and security intrinsically linked.
- Data inserted into AI models may include financial data, contact information such as names and addresses, legal agreements, and confidential information such as medical records. This is vulnerable to exploitation by attackers. Organisations must ensure they have reasonable safeguards proportionate to the data, as required by law.
- Security of learning systems needs to be ensured - potential of attacks (data or model poisoning; creation of backdoors).
- If one participant sends carefully manipulated model updates, they can corrupt the performance of the trained global model on specific subtasks.

Data Retention Policies

- The complex mechanics of AI systems pose challenges in understanding when data ought to be deleted or is no longer required. Storing data for an increased period of time also leads to vulnerabilities from unauthorised access.

Data Sharing and Transfer Risks

- AI platforms may be used by multiple platforms or third parties. This poses data confidentiality risks, but also issues if data is transferred outside the jurisdiction.

Under GDPR, if AI is used to process personal data it can trigger the legal requirement to undertake a DPIA

- Without an understanding of the AI model, it is challenging for organisations to produce DPIAs compliant with regulations.

Loss of Privacy/Data Privacy

- Undisclosed use of underlying data to train AI systems - use of personal information in ways that are contrary to the contexts in which they are collected.
- Release, exposure or publication of non-public information or sensitive data about an individual or household by Generative AI models resulting in negative outcomes (expand dissemination of public but sensitive information).
- Accurate or inaccurate analytical insights impact outcomes or reputation.
- Comprehensive surveillance and data harvesting breed distrust, anxiety, and other mental health problems; chilling speech, protest, and worker organising.

Misinformation

- AI can be used by hostile nation states and threat actors to create content which fosters societal division and support for their regimes.
- World Economic Forum – *“AI-powered misinformation is the world’s biggest short-term threat.”*


Cyberwarfare

- AI can be used to develop advanced malware, identify system vulnerabilities for cyber attacks, and impersonate staff to access critical infrastructure.
- UK National Cyber Security Centre – *“Artificial intelligence will almost certainly increase the volume and heighten the impact of cyber attacks over the next two years”*

Heightened concerns in defense - aiding in the development of weapons

- Interpol Report on the 2023 Global Congress on Chemical Security — *“In just six hours, an Artificial Intelligence system was able to create tens of thousands of chemical compounds that could be used as chemical weapons”*

UK National Security and Investment Act regime – AI technologies can potentially be used for harmful purposes. Where a transaction involves a business developing AI, a **merger filing may be required on national security grounds, and government approval is required**



US Secretary of Homeland Security: *“AI can present transformative solutions for U.S. critical infrastructure, and it also carries the risk of making those systems vulnerable in new ways to critical failures, physical attacks, and cyber attacks”*

U.S. National Security Risks Include

Interruption of critical services or infrastructure

Interference with operation of global financial systems, supply chain, and other interrelated systems

Enhanced capabilities for espionage and surveillance

Greater analysis of data and data sets (combining data sets) to reveal information (citizens, military government etc.)

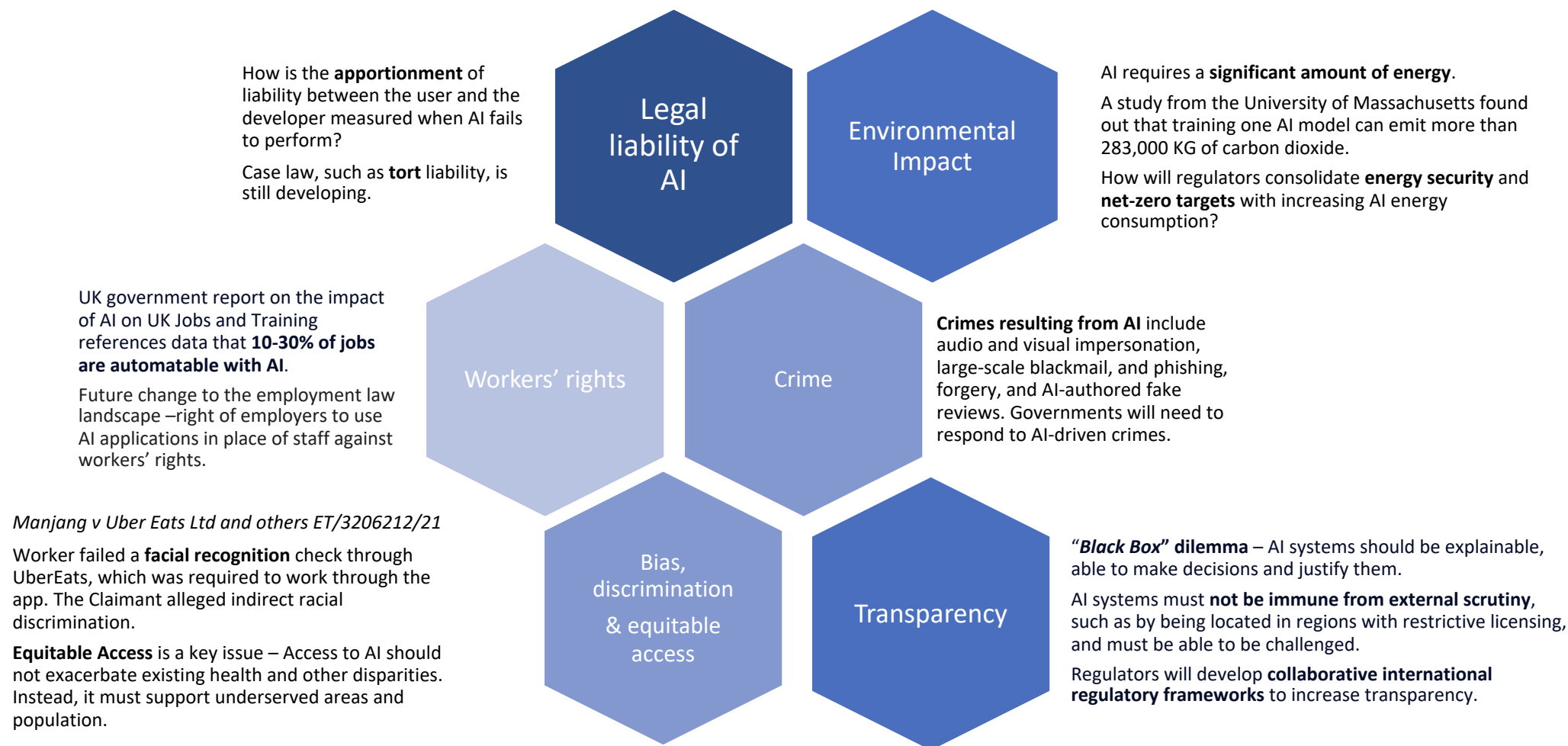
Inability to distinguish accurate information from false or modified

Cyberattacks by enhanced non-state actor(s)

Proliferation of WMDs and other destructive tools across nations and malicious non-state actors

Loss of transportation (air, marine, rail and road (autonomous or semi-autonomous vehicles etc.)

Election interference (misinformation and disinformation campaigns)



3. Potential Harms and Risks

What risks do we see today – legal / regulatory / societal (inc. national security)?

Existential Risk - Rise of the machines v. human society

Military Capabilities - AI-capable warfare using drones etc. v. Peace dividend (akin to the Cold War nuclear stand-off)

Workforce and Resourcing - potential to take away jobs as we know them today

Risk of Bias - threat of magnification of bias in training data v. identification of bias to avoid denial of schooling; financial freedom; jobs etc. to certain minorities, groups or non-traditional stakeholders

Security issues - Vulnerability to attack due to a lack of security thinking among the AI proponents



Risk to Society and Democracy

Loss of public confidence in the integrity of official content

Loss of trust and public confidence

- inability to establish authenticity and provenance of digital content
- inability to understand algorithms and AI

Democratic values are devalued

- decrease in participation in activities such as voting

Comprehensive surveillance and data harvesting breed distrust

Chilling effect on speech and expression -

- data collection, surveillance, ability to identify and reidentify data

Health Context – Presents Significant Risks from Accident or Error

AI-enabled systems can produce unintended harm if the system (i.e., chatbot response) leaves out part of a protocol, safety measures, or biosecurity requirements even if otherwise technically accurate

Unreliable predictions produce unintended outcome, which would translate to real-world harm if the biological agent is produced and escapes containment

AI system deviates from intended or expected functionality

What risks do we foresee?

AI lacks self-awareness

Hallucinations and the need for LLMs to “understand” the generated responses

Data poisoning

Unregulated data integration by consolidation of datasets

“Technology Debt”

Responsibility of Data Trusts

Need for comprehensive privacy-enhancing technologies

The need for regulation of Data Sharing Partnerships academic/commercial

“Unknown unknowns”

International Sharing Frameworks to Manage Risks: Example – U.S. National Institute of Standards and Technology (NIST)
“AI Risk Management Framework (AI RMF) 1.0” – guidance document for organizations developing AI systems

NIST Four Fundamental Principles of Risk Management

Governance. Structures, policies and processes in relation to risk management. Foster a risk management culture

Mapping. Understand how to frame and categorize risk, e.g., goals, benefits, costs.

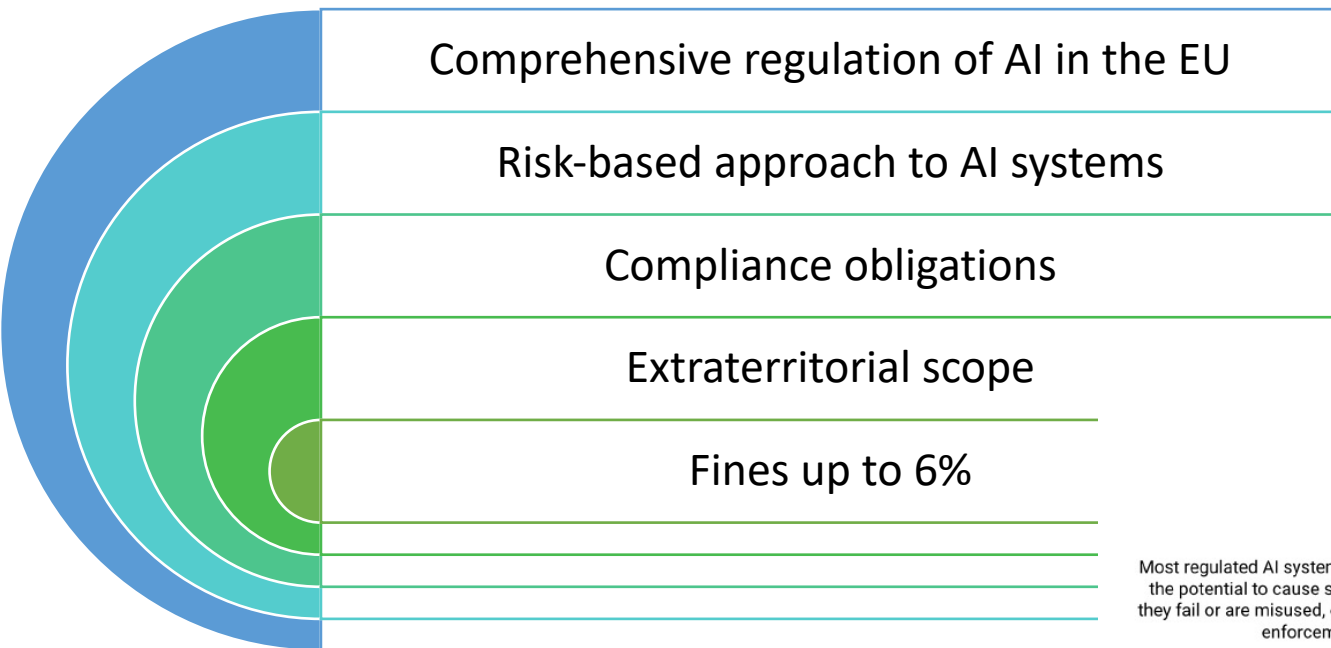
Management. Risk monitoring and prioritization of risks.

Management. Deploy broad qualitative and quantitative analysis to assess AI characteristics and impacts.

UK Regulation

- Proposal **does not target specific technologies** and focuses on context instead to avoid stifling innovation or placing undue burdens on businesses.
- **Principles-based regulatory regime** overseen by existing regulators.
- **No new laws or sanctions**
- **Sector-specific guidance** for organisations (e.g. ICO Guidance on AI and Data Protection, issued on 15 March 2023)
- April 2024, the UK has begun to explore new legislation to regulate artificial intelligence which would likely **put limits** on the production of large language models, the **general-purpose technology** that underlies AI products.



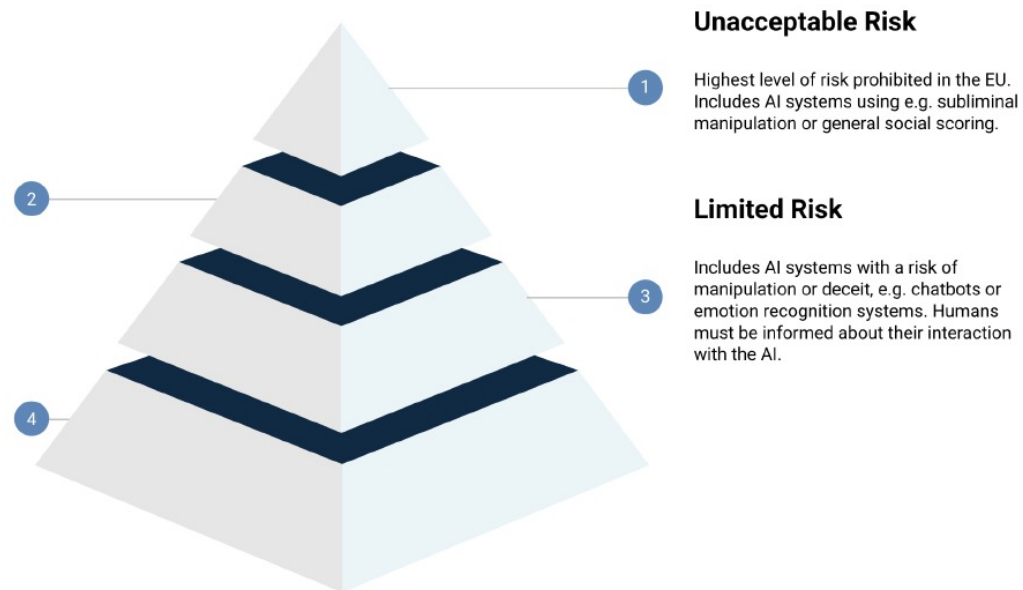


High Risk

Most regulated AI systems, as these have the potential to cause significant harm if they fail or are misused, e.g. if used in law enforcement or recruiting.

Minimal Risk

All other AI systems, e.g. a spam filter, which can be deployed without additional restrictions.



An AI model created for one purpose should not subsequently be used by a client for an entirely different - and potentially inappropriate – purpose

MOST critical issue?

Impacts every risk discussed

1

WARRANTIES AND REPRESENTATIONS

IP Rights. A Buyer may want to ensure it is acquiring the required rights to use and/or own the technology, such as rights over the source code. The Seller will need to warrant it has the relevant licenses or take ownership of creating the product.

Compliance with privacy regulations. AI models utilize and are reliant on a large volume of data, which may involve personal and sensitive personal data. User consent will often be required for processing data through AI models to take place, and whether practices are compliant with privacy legislation and regulation. The Seller ought to warrant data has been obtained and is processed lawfully.

Security. A Buyer will want to ensure robust cybersecurity policies have been in place, to safeguard data from unauthorized access or misuse.

Performance. The Buyer may want to ensure the Seller provides warranties on the AI's system performance and ensures targets are met.

Risk allocation. The Buyer should request provisions to mitigate risks of liability arising from the AI model, such as breach of contract and negligence.

2

INDEMNITIES

Damages. A Buyer may request an indemnity against damages awarded to third parties as a consequence of the AI model, such as from data breach.

IP infringement. The Seller to indemnify the Buyer if the AI model infringes against IP claims from third parties.

Defects. The Seller to indemnify the Buyer for losses sustained by defects in the AI model and/or misuse of data.

4. Responsibility and Accountability v Balance and Context

How should organizations be held accountable for harm?

Reporting and oversight

Reporting obligations monitor performance and allow for justification. Oversight of reporting is necessary for regulators to oversee and scrutinize conduct.

Compliance

Compliance to bind AI organisations to standards of accountability and ensure organisations maintain societal responsibility.

Enforcement

Corrective actions, prohibitions and financial sanctions must be available as an enforcement mechanism for violations and to act as a deterrent.

Fines

- Italy fines Clearview AI €20m for **biometric data scraping**
- France fined Google €50m for **targeted advertising** through AI
- U.S. fined Hello Digit \$2.7 million for a **faulty algorithm** in its AI processing system
- UK fines iTutor Group for **automated age discrimination**

Structure and procedures

Ethics

- UNESCO — *"The use of AI systems must not go beyond what is necessary to achieve a legitimate aim"*.
- Fairness, transparency and accountability as core principles behind the use of AI.
- Regular review for bias and discrimination. and a functional complaints procedure for users to follow and corrective action to be taken.

Legal

- Creation of a robust legal framework with adequate policies in relation to data privacy (e.g., GDPR) that are adapted to the organisation.
- Risk analysis of the AI platform with legal input on areas such as input and output risks of using AI, vendor risk assessments, and disaster recovery plans.
- Develop legal practices to respond to regulatory inquiries, complaints and investigations.

Security

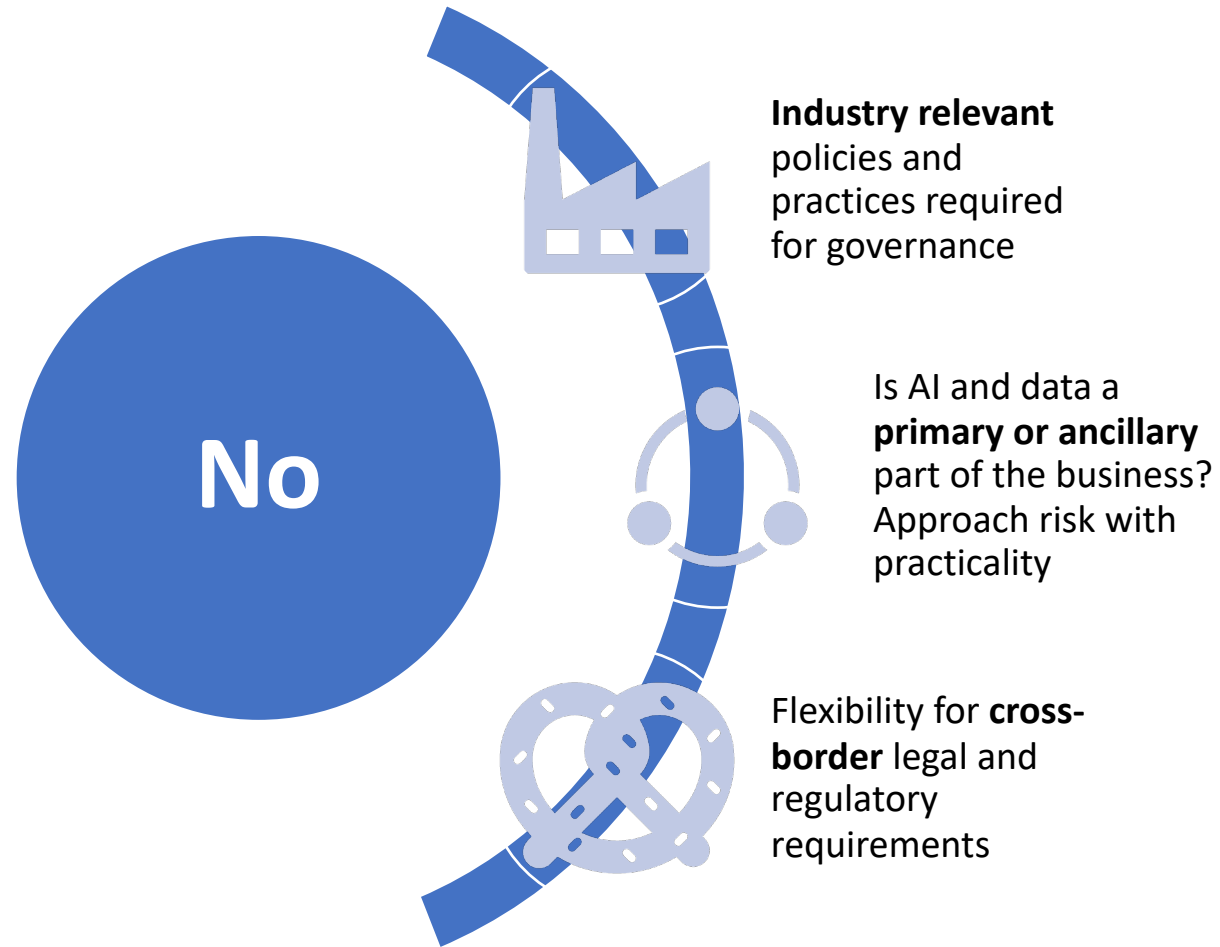
- Set up and review cyber-security policies and incorporate a robust cyber awareness training program.
- Implementation of physical, administrative and technical safeguards.

Other

- Review source code to understand the quality and security of the AI product.
- Consider insurance to mitigate risks from AI, such as cyber insurance or W&I insurance adapted to AI.

Cisco Study:
Only **14%** of
business
leaders are fully
prepared for an
AI transition.

One-size-fits-all regulatory and governance approach?



UN Secretary-General: "develop a governance model that is networked and adaptive"

Jurisdiction

- EU strict **risk-based approach**.
- UK and U.S. **principles-based approach**.
- Monitoring divergence/convergence of **upcoming global legislative frameworks** to AI (e.g., Brazil, China, India) and assessing applicability.

Area

- Analyse **focus points** based on industry. Examples include creative industries and IP rights, healthcare sector and ethical and bias concerns, and recruitment sector and transparency.
- Overriding **risks pervasive across industries**, such as data privacy.

Industry Size

- Identify actions which are **proportionate and reasonable to take against the risk**. A small manufacturing business using AI to assess production data and expansion opportunities will not need an as intensive compliance framework compared to a large pharmaceutical company.
- Consider **outsourcing the management risks to third parties for cost efficiency** or **carrying out due diligence on vendors** (e.g., cloud providers).

5. What should we not be sharing and why?

Considerations for data not to use in AI models

Confidential information

Intellectual property

Non-reviewed pricing and financial data

Value of combined data sets

Computing power + data from more sources
= Greater potential benefits / harm

Context is key

6. Takeaways

Questions & Contacts



Rohan Massey
Ropes & Gray
Rohan.massey@ropesgray.com



Marc Groman
Groman Consulting
gromanmarc@gmail.com



Sajai Singh
JSA Law
sajai@jsalaw.com



Stephen Burns
Bennet Jones
burnss@bennettjones.com



Patricia Martín-Marrero
Takeda Pharmaceuticals
Patricia.Martin@takeda.com