Ethical AI in Information Governance: Shaping the Future of Information Management

KENNETH ATUMA

LECTURER(ASSISTANT PROFESSOR) IN LIBRARY AND ARCHIVES STUDIES

THE UNIVERSITY OF MANCHESTER

DATE & TIME: TUESDAY 1 APRIL 2025 | 2:00PM
VENUE: HYATT REGENCY, ATLANTA, GEORGIA, UNITED STATES OF AMERICA

AI+IM GLOBAL SUMMIT

MANCHESTER
1824
The University of Manchester

INTRODUCING THE SPEAKER

❖ His is a Knowledge and Information Management Specialist, Scholar, and Consultant at the University Manchester, United Kingdom

❖ Bachelors' degree in Lib & Info. Science, M.Sc. In Information Management &Preservation, PhD in Information Studies (In view)

❖ His has previously taught in two academic institutions of higher learning in Nigeria prior to joining the University Manchester.

❖ His area specialisation and research interest are Archives, Records Management, and Information Governance and Compliance, AI in information Management.

AI+IM
GLOBAL SUMMIT

Do you think AI is a necessary evil?
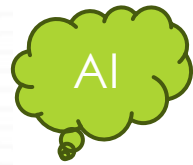
What is Ethical AI

❖ Ethics is a set of moral principles which help us discern between right and wrong. AI ethics is a multidisciplinary field that studies how to optimize the beneficial impact of artificial intelligence (AI) while reducing risks and adverse outcomes.

❖ It can also mean the principles and guidelines that govern the responsible development and use of artificial intelligence. It ensures AI systems are fair, transparent, accountable, and aligned with human values.

❖ The ethics, anchors on the key principle which are: **Fairness** ,**Transparency**, **Accountability**, **Privacy & Security** and **Beneficence**

The big question

European Fincial Review

How do you implement ethical AI Framework in your Information Management Practice?

AI+IM GLOBAL SUMMIT

Objectives and extent of coverage

In this session, we shall:

- Understand the principles of **ethical AI** in information governance.

- Explore **regulatory frameworks** from the EU, US, and the Global South.

- Identify key **ethical risks** (bias, privacy, transparency).

- Discuss **best practices** for implementing ethical AI.

- Engage in **interactive case studies** and Q&A.
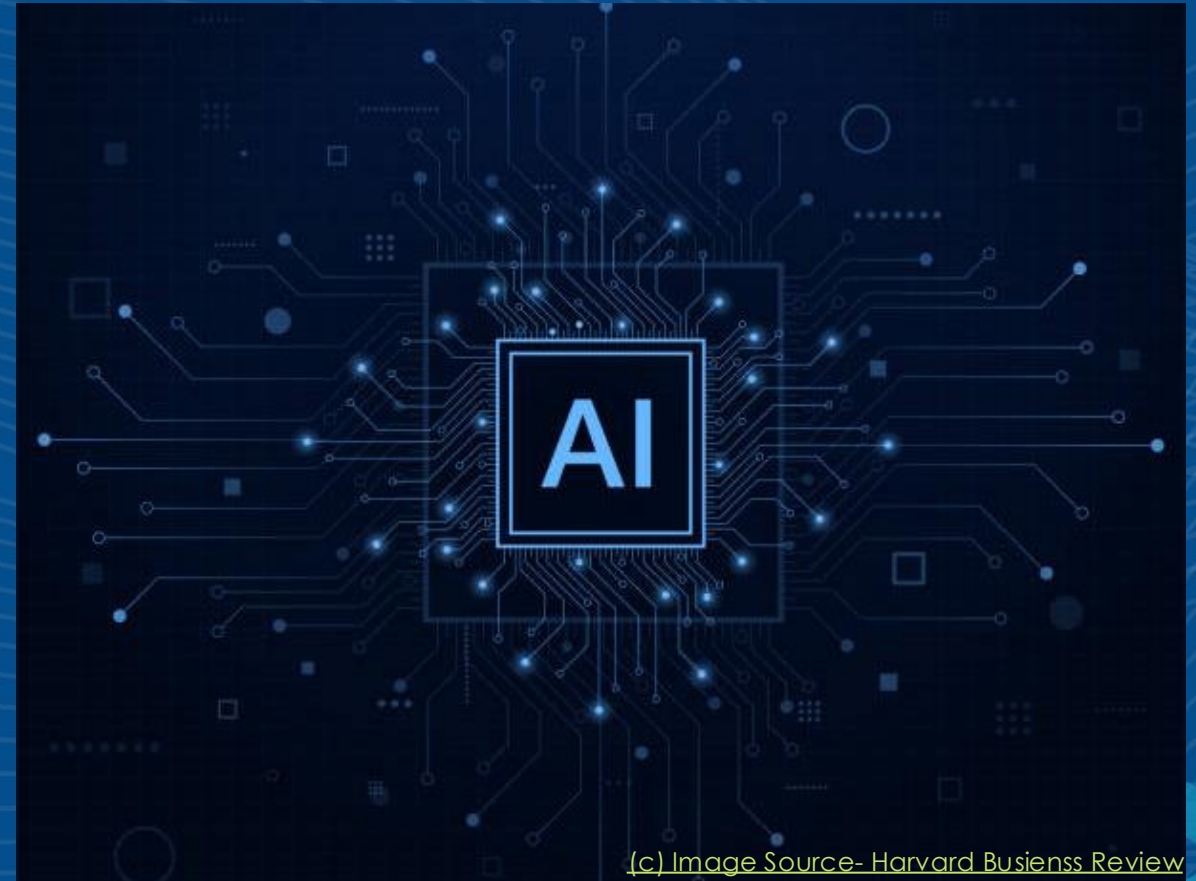
(c) Image Source- Harvard Busienss Review

Live Poll Q-one

Let us know how familiar you are with AI in information governance?

Click on this link to acess the live poll

The Role of AI in Information Governance

The Role of AI in Information Governance

- AI is transforming **records management, compliance, and decision-making** (Floridi & Cowls, 2019).

- Automation enhances **efficiency, accuracy, and scalability** (Mittelstadt, 2019).

- However, Ethical challenges arise when AI impacts **data privacy, accountability, and fairness** (Binns, 2018).

(c) Image Source- Harvard Busienss Review
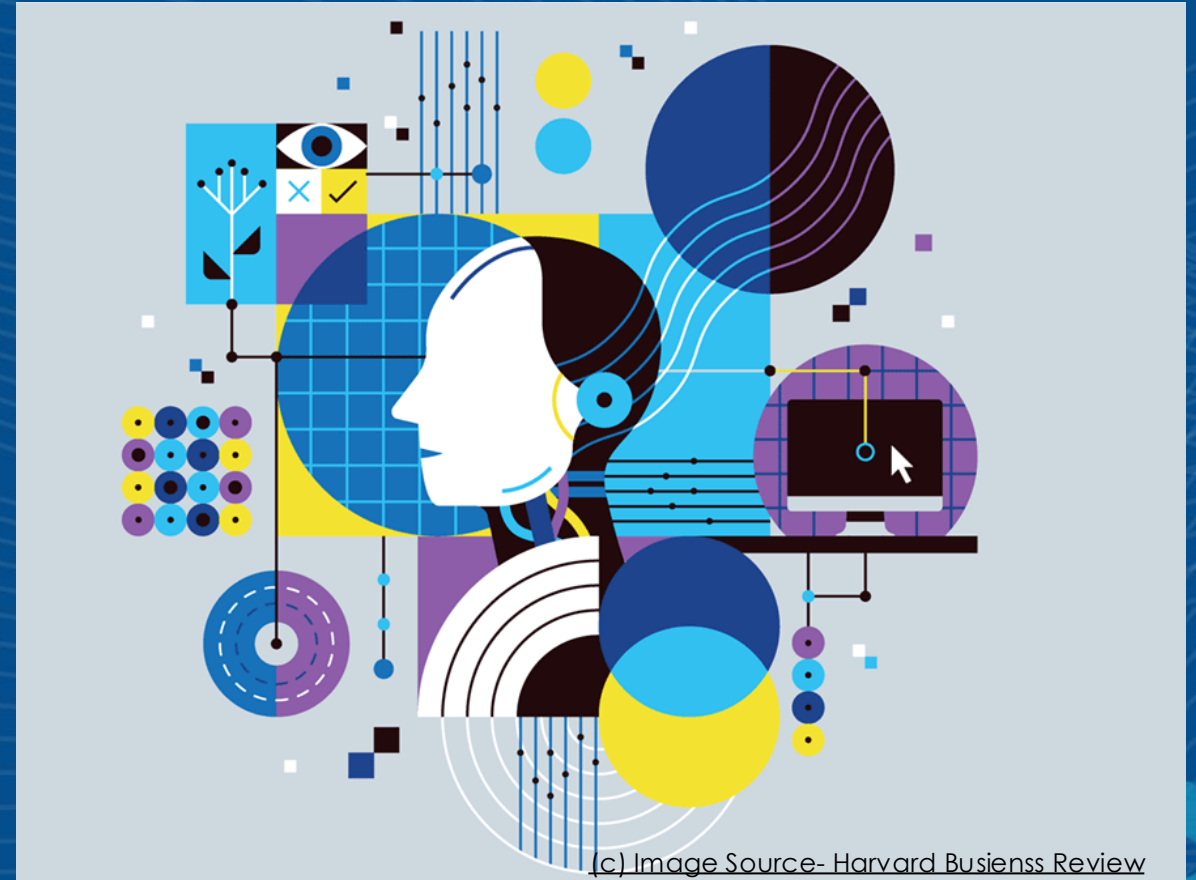
Here are the Key Questions we need to Consider

1. How does AI enhance information governance?

2. What ethical concerns emerge when AI processes sensitive data?

3. How can organisations balance AI innovation with ethical responsibility?

Principles of Ethical AI

Four pillars of ethical AI

Transparency

AI decisions should be explainable and understandable (**doshi-velez & kim, 2017**

Fairness

AI models must minimize bias and ensure equitable treatment (**kamiran et al., 2012**).

Accountability

Organizations should take responsibility for ai-driven outcomes (**raji et al., 2020**).

Privacy

AI laws and respect user rights (**European commission, 2021**).

Can you think of your organisation as a possible place where one of these principles could violated?

Live Poll Q-two

Let us know your thought on the biggest ethical risk when using AI in information governance.

Click on this link to acess the live poll

AI+IM
GLOBAL SUMMIT

- **Ethical Challenges**

  - **Bias in AI models**

    - AI systems can reflect and reinforce societal biases (**Binns, 2018**).

  - **Privacy violations**

    - AI can misuse personal data, violating legal frameworks (**European Commission, 2021**).

  - **Transparency concerns**

    - Black-box AI models lack explainability (**Doshi-Velez & Kim, 2017**).

  - **Regulatory compliance**

    - AI must align with evolving legal standards (**NIST, 2022**).

  - **Loss of human oversight**

    - AI should support, not replace, human decision-making (**Raji et al., 2020**).

AI+IM
GLOBAL SUMMIT

Global AI Law and Policy Tracker

This map shows which jurisdictions are in focus and covered by this tracker. It does not represent the extent to which jurisdictions around the world are active on AI governance legislation.

© Global AI Law and Policy Tracker. IAPP.

Our focus of discussion

United States of America. The United Kingdom. European Union. Nigeria. Peru. Columbia. Egypt

US AI Regulation in focus(1)

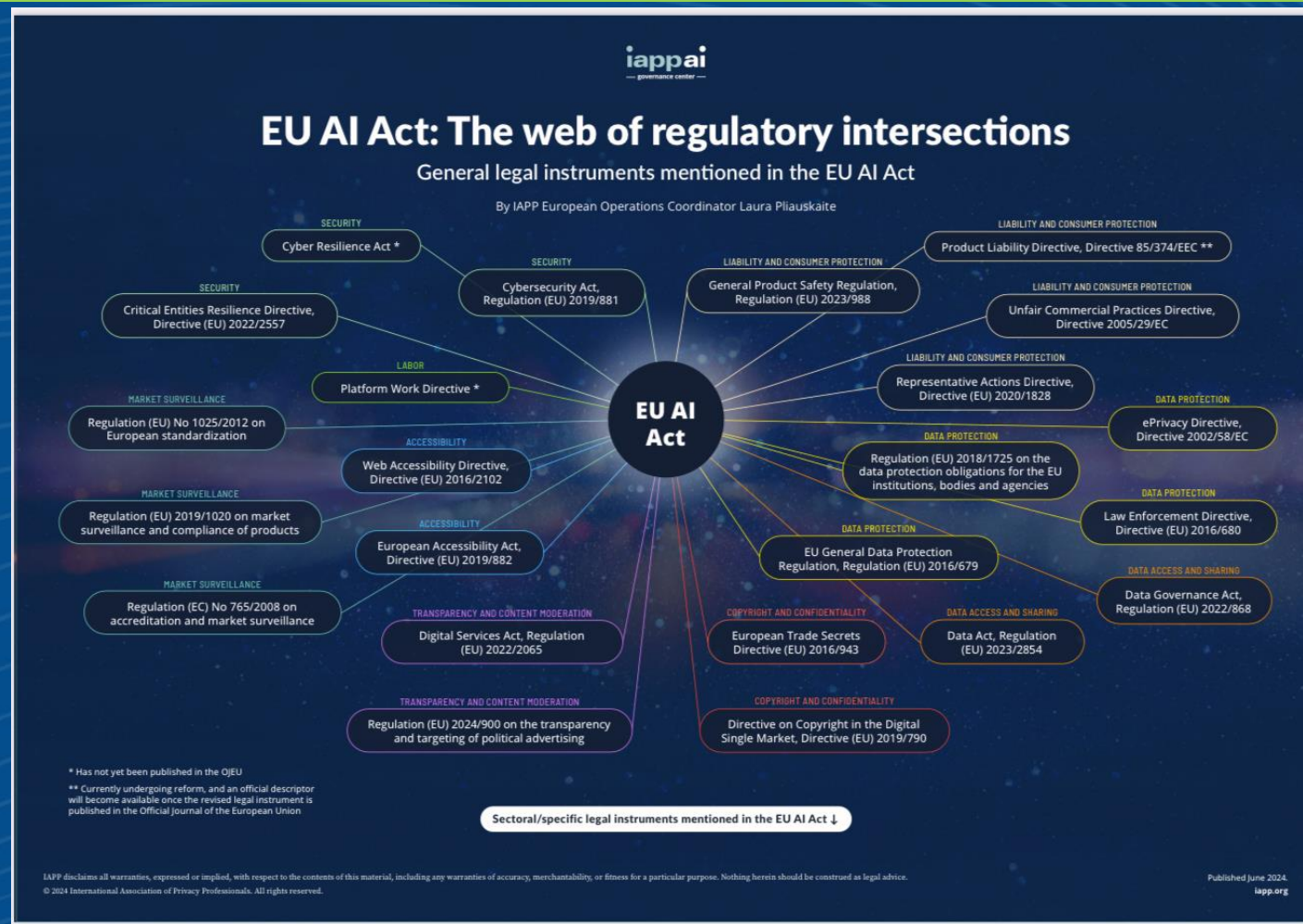|	Specific AI governance law or policy	Relevant authorities	Other relevant laws and policies	Wider AI context

U.S. (FEDERAL)	The U.S. has been active in providing guidance to government organizations and private businesses while introducing legislation to target specific issues, such as deepfakes or discrimination. In October 2023, President Joe Biden signed the Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, which put forth guidelines for AI use in federal agencies, including accountability and transparency requirements. This has mirrored efforts in Congress to promote and control the use of AI domestically.
	
	The U.S. has been active in many of the multilateral agreements on AI as well, for example by signing onto the Council of Europe's Framework Convention and promoting rulemaking at the U.N. At the state level, several bills have been passed, such as the Colorado AI Act, and several bills that regulate AI in specific sectors, such as House Bill 3733 in Illinois. While not an exhaustive list, the following federal laws and policies could place a compliance or regulatory burden on private businesses:
	
	• Executive orders:
	  - Maintaining American Leadership in AI
	  - Promoting the Use of Trustworthy AI in the Federal Government
	  - The Safe, Secure, and Trustworthy Development and Use of AI
	• Acts and bills:
	  - AI Training Act [IN FORCE]
	  - National AI Initiative Act (Division E, Sec. 5001) [IN FORCE]
		→ Office of Science and Technology Policy
		→ National AI Initiative Office
		→ Federal Trade Commission
		→ Consumer Financial Protection Bureau
		→ Department of Justice
		→ Equal Employment Opportunity Commission
		→ National Institute of Standards and Technology
			→ FTC Act, Section 5 [IN FORCE]
			→ Fair Credit Reporting Act [IN FORCE]
			→ Equal Credit Opportunity Act [IN FORCE]
			→ Title VII of the Civil Rights Act [IN FORCE]
			→ Americans with Disabilities Act [IN FORCE]
			→ Age Discrimination in Employment Act [IN FORCE]
			→ Fair Housing Act [IN FORCE]
			→ Genetic Information and Nondiscrimination Act [IN FORCE]
			→ American Privacy Rights Act [DRAFT]
			→ Health Equity and Accountability Act [DRAFT]
				• The U.S. is a party to the OECD's AI principles. See the OECD's Policy Observatory.
				• The U.S. participated in the 2023 U.K. AI Summit, which led to the Bletchley Declaration.
				• The U.S. also adopted UNESCO's Recommendation on the Ethics of AI.
				• As part of the G7, the U.S. endorsed the 11 Hiroshima Process International Guiding Principles for Advanced AI systems.
				• In general, the U.S. approach to AI governance has been slow and incremental, seeking to preserve civil and human rights for Americans throughout AI deployment, as well as mobilize international collaboration which upholds democratic values and mutual advancement.
				• See the U.S. AI Safety Institute.
				• U.S. Senate Committee on the Judiciary's Subcommittee on Privacy, Technology and the Law held a hearing on the legislation of AI.
				• The Bipartisan Senate Working Group on AI, led by Sen. Chuck Schumer, D-N.Y., has released a roadmap for AI policy. This document highlights the need to ensure enforcement of existing rules, tackle current threats not covered by legislation, such as the use of deepfakes in elections, prepare for long-term threats of AI use, and create a federal privacy legal framework.
				• The Singapore VerifyAI initiative known as "crosswalk" was unveiled at the inaugural U.S.-Singapore Dialogue on Critical and Emerging Technologies. The crosswalk links IMDA's AI Verify with the U.S. NIST's AI Risk Management Framework.

US AI Regulation in focus(2)

	Specific AI governance law or policy	Relevant authorities	Other relevant laws and policies	Wider AI context

U.S. (FEDERAL), continued	- AI in Government Act (Division U, Sec. 101) [IN FORCE]
	- AI CONSENT Act [DRAFT]
	- AI Transparency in Elections Act [DRAFT]
	- Algorithmic Accountability Act [DRAFT]
	- Content Origin Protection and Integrity from Edited and Deepfaked Media Act [DRAFT]
	- Digital Platform Commission Act [DRAFT]
	- NO FAKES Act of 2024 [DRAFT]
	- Protect Elections from Deceptive AI Act [DRAFT]
	- Protecting Consumers from Deceptive AI Act [DRAFT]
	
	• Nonbinding frameworks:
	  - Blueprint for an AI Bill of Rights
	  - National Institute of Standards and Technology AI Risk Management Framework
	  - Guidance for Regulation of AI Applications
	• Government initiatives:
	  - Voluntary Commitments from Leading AI Companies to Manage the Risks Posed by AI
	  - TTC Joint Roadmap on Evaluation and Measurement Tools for Trustworthy AI and Risk Management
	  - Congressional AI effort of Sen. Charles E. Schumer, D-N.Y.
	  - National Security Commission on AI
	  - Bipartisan legislative framework for AI announced by U.S. Sens. Richard Blumenthal, D-Conn., and Josh Hawley, R-Mo.
				• Along with the U.S., the U.K., Israel and the EU have has signed onto the Council of Europe's Framework Convention on AI and human rights, democracy and the rule of law.

European Union AI Regulation in focus

United Kingdom AI Regulation in focus

Specific AI governance law or policy	Relevant authorities	Other relevant laws and policies	Wider AI context
The U.K. government proposed a context-based, proportionate approach to regulation and will rely on existing sectoral laws to impose guardrails on AI systems. In July 2024, during the King's speech, the government affirmed its intention to regulate AI, although no firm legislation had been introduced under the new administration. Currently, the U.K. relies on existing sectoral laws to impose guardrails on AI system. The following resources are available for policy guidance: • A pro-innovation approach to AI regulation. • Algorithmic Transparency Recording Standard Hub. • AI Standards Hub, a new U.K. initiative dedicated to the evolving and international field of standardization for AI technologies. • Guide to using AI in the public sector by the U.K. government. • The Government Digital Service and the Office for AI's guide on understanding AI ethics and safety. • The Centre for Data Ethics and Innovation's AI Governance research report. • Guidance on the AI auditing framework from the Information Commissioner's Office. • ICO and Alan Turing Institute's Explaining decisions made with AI.	→ Office for AI → Information Commissioner's Office → Digital Regulation Cooperation Forum → Financial Conduct Authority → AI Council → Department for Science, Innovation and Technology	→ Equality Act [IN FORCE] → U.K. General Data Protection Regulations and Data Protection Act [IN FORCE] → Consumer Protection Act [IN FORCE] → Financial Services and Markets Act [IN FORCE] → Consumer Rights Act [IN FORCE] → National Security and Investment Act [IN FORCE] → Copyright, Designs and Patents Act [IN FORCE] → Advanced Research and Invention Agency Act [IN FORCE] → National Cyber Security Centre's Assessing intelligent tools for cyber security [IN FORCE]	• The U.K. is a party to the OECD's AI principles. See the OECD's Policy Observatory. • In 2023, the country hosted the AI Summit, which led to the Bletchley Declaration. • The U.K. also adopted UNESCO's Recommendation on the Ethics of AI. • As part of the G7, the U.K. endorsed the 11 Hiroshima Process International Guiding Principles for Advanced AI systems. • Specific action items include launching a national AI research and insights program, developing a diverse AI workforce, enabling better data availability, creating a national strategy for AI in health and social care, applying AI systems to climate change mitigation, piloting an AI standards hub to coordinate with global AI standardization, and developing a cross-government standard for algorithmic transparency. • The Centre for Data Ethics and Innovation published a Roadmap to an Effective AI Assurance Ecosystem, which is also part of the National AI Strategy. Further, the CDEI created an AI Assurance Guide as a companion to the roadmap. • See the U.K. AI Safety Institute. • Along with the U.S., the U.K., Israel, and the EU have has signed onto the Council of Europe's Framework Convention on AI and human rights, democracy, and the rule of law.

Nigeria & Peru AI Regulation in focus

	Specific AI governance law or policy	Relevant authorities	Other relevant laws and policies	Wider AI context
NIGERIA	In April 2024, Nigeria hosted a workshop to devise a national AI strategy, where Minister of Communications, Innovation and Digital Economy Bosun Tijani stated to the country's goal is to become a key player in global regulation and development of AI. In August 2024, the country released its draft national AI strategy, which recognizes the benefits and risks of widespread adoption of AI. Nigeria plans to address the ethical issues of using AI while embracing it as a driver of socioeconomic growth.	→ Federal Ministry of Communication, Innovation and Digital Economy → National Information Technology Development Agency → Nigerian Communications Commission	→ Nigeria Data Protection Regulation → Nigeria Data Protection Act	• Nigeria adopted UNESCO's Recommendation on the Ethics of AI. • Nigeria participated in the 2023 U.K. AI Summit, which led to the Bletchley Declaration. • In 2020, the Nigerian Communications Commission released a research paper on the ethical and societal impacts of AI.
PERU	Peru has drafted legislation around the use of AI, including law 3814, which would promote the use of AI "in favor of the economic and social development of the country." The law includes the following principles: • Risk-based security standards. • Multi-stakeholder approach. • Internet governance. • Digital society. • AI privacy. Peru also developed a National AI Strategy that aids in the promotion, development and adoption of AI in the country. The first draft includes a roadmap, goals, definitions and external context examples to further develop the strategy.	→ Secretariat of Government and Digital Transformation → Presidency of the Council of Ministers → National Directorate of Intelligence → Superintendence of Banking, Insurance and Pension Fund Administration → Ministry of Justice and Human Rights → National Authority for the Protection of Personal Data → National Authority for Transparency, Access to Public Information and Protection of Personal Data	→ Supreme Decree No. 157-2021-PCM [IN FORCE] → Supreme Decree No. 003-2013-JUS [IN FORCE] → Personal Data Protection Law No. 29733 [IN FORCE] → Law of Transparency and Access to Public Information [IN FORCE] → Finance Regulation for Information Security and Cybersecurity [IN FORCE] → Cyber Defense Law No. 30999 [IN FORCE] → Law 30096 on Computer Crime [IN FORCE] → Financial sector Cybersecurity Framework [IN FORCE] → Copyright Law, Legislative Decree 822 [IN FORCE]	• Peru is a party to the OECD's AI principles. See the OECD's Policy Observatory. • Peru also adopted UNESCO's Recommendation on the Ethics of AI. • See the National Digital Transformation Policy for 2030.

Columbia and Egypt Regulation in focus

	Specific AI governance law or policy	Relevant authorities	Other relevant laws and policies	Wider AI context
COLOMBIA	Colombia has various policies addressing AI governance, including the following: • AI Expert Mission. • AI National Strategy Policy. • AI in the Public Sector.	→ Administrative Department of the Presidency of the Republic → CAF → Ministry of Information and Communication Technologies → Ministry of National Education → Administrative Department of Science, Technology and Innovation → National Planning Department → Superintendence of Industry and Commerce → AI Task Force	→ Personal Data Protection Law [IN FORCE] → Habeas Data Law, Law 1266 amended by Law 2157 of 2021 [IN FORCE] → Decree 338 [IN FORCE]	• Colombia is a party to the OECD's AI principles. See the OECD's Policy Observatory. • Colombia also adopted UNESCO's Recommendation on the Ethics of AI. • Colombia published an Ethical Framework that reiterates best practices, suggestions and recommendations on how best to integrate ethical principles with the use of AI in projects primarily for the benefit of the public sector entities. • An AI Task Force was created in partnership with the CAF to bolster AI progress.
EGYPT	Egypt's National AI Strategy focuses on four pillars: • AI for government. • AI for development. • Capacity building. • International activities. The country's other initiatives include an AI roadmap and Charter for Responsible AI.	→ National Council for AI → Ministry of Communications and Information Technology	→ Law No. 151 of 2020 on the Protection of Personal Data [IN FORCE] → Law No. 175 of 2018 Regarding Anti-Cyber and Information Technology Crimes [IN FORCE] → Telecommunication Regulation Law, Law No. 10 of 2003 [IN FORCE] → Law No. 82 of 2002 on the Protection of Intellectual Property Rights [IN FORCE]	• Egypt is a party to the OECD's AI principles. See the OECD's Policy Observatory. • Egypt also adopted UNESCO's Recommendation on the Ethics of AI. • Egypt chaired several meetings for the Arab AI Working Group, which allows representatives from Arab countries to discuss AI strategies. See the group's chair election, second meeting and third meeting. • See the Applied Innovation Center. • The Senate Education Committee stressed the urgency of issuing a document to evaluate the ethics and control of AI in Egypt.

Global AI Law and Policy Tracker · IAPP · iapp.org

11

Global Regulatory Landscape for AI Ethics

Comparison of AI Regulations

Region	Key Regulation	Focus Areas
EU	AI Act (European Commission, 2021)	Risk-based classification, compliance obligations
US	NIST AI Risk Management Framework (NIST, 2022)	AI trustworthiness, bias mitigation
Global South	Various emerging laws	Data sovereignty, ethical AI deployment

- **Understanding different regulatory perspectives helps organizations ensure compliance.**

Now let us discuss how AI deployment apply to your own organisation

"Assuming your organisation integrated AI-powered document classification system, unintentionally discriminates against certain minority groups when flagging documents for review."

Here are the questions for us to discuss:

▶ What ethical risks do you identify in this case?

▶ How should the organisation address bias and transparency issues?

▶ What steps can be taken to prevent such incidents in the future?

Share insights with us.
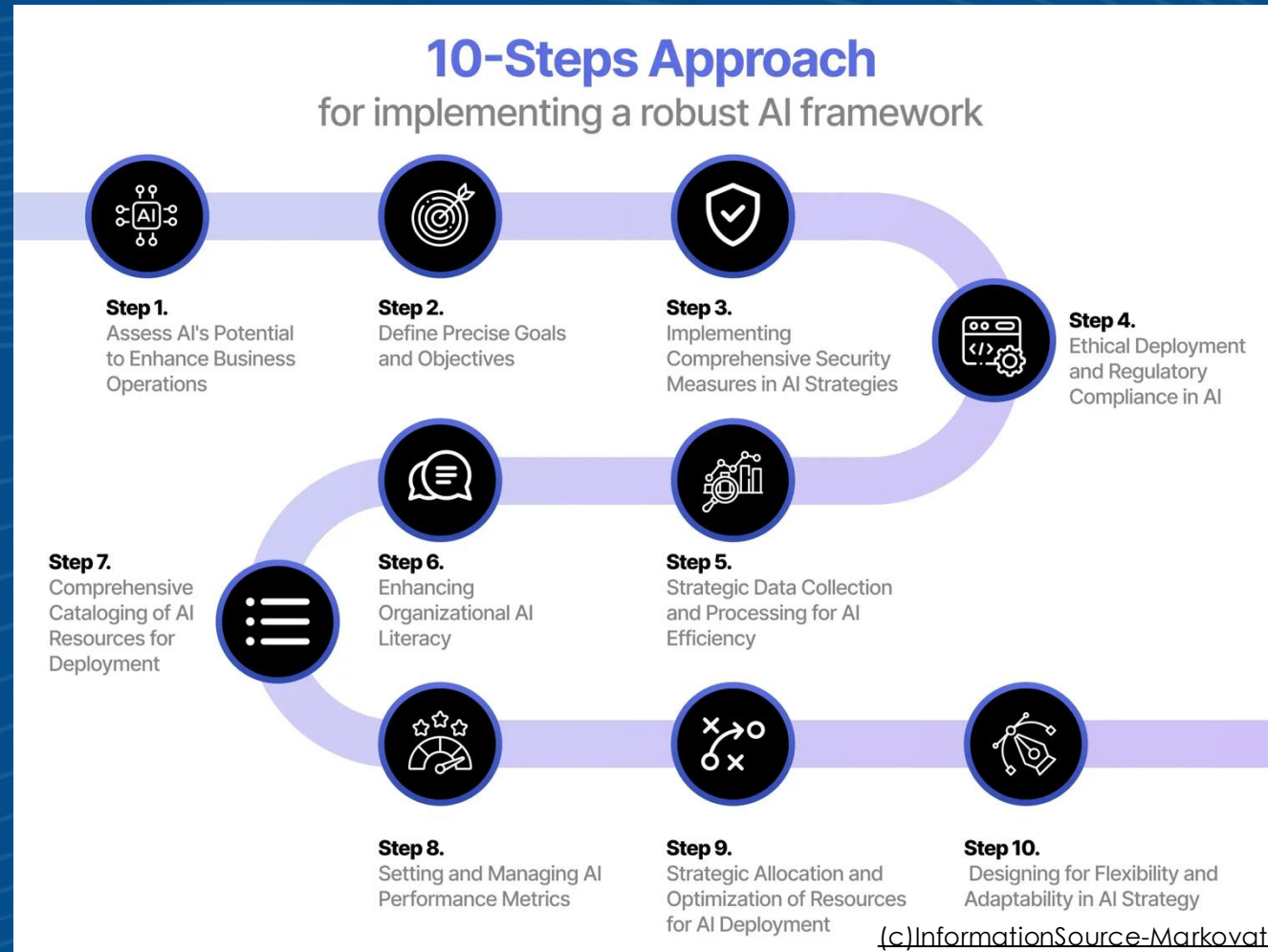
AI+IM
GLOBAL SUMMIT

Best Practices for Ethical AI Implementation

❖ **ACTIONABLE STRATEGIES**

  ❖ Conduct **AI Ethics Audits** to assess risks .

  ❖ Ensure **diverse datasets** to minimize bias.

  ❖ Adopt **explainable AI models** for transparency.

  ❖ Implement **privacy-preserving techniques** data anonymisation.

  ❖ Align AI use with **regulatory frameworks** to avoid legal risks.

Creating a Robust AI Strategy Framework for organisational success

10-Steps Approach
for implementing a robust AI framework

Step 1.
Assess AI's Potential to Enhance Business Operations

Step 2.
Define Precise Goals and Objectives

Step 3.
Implementing Comprehensive Security Measures in AI Strategies

Step 4.
Ethical Deployment and Regulatory Compliance in AI

Step 7.
Comprehensive Cataloging of AI Resources for Deployment

Step 6.
Enhancing Organizational AI Literacy

Step 5.
Strategic Data Collection and Processing for AI Efficiency

Step 8.
Setting and Managing AI Performance Metrics

Step 9.
Strategic Allocation and Optimization of Resources for AI Deployment

Step 10.
Designing for Flexibility and Adaptability in AI Strategy

(c)InformationSource-Markovate

AI+IM
GLOBAL SUMMIT

Discussion of Risks

"As with all powerful technologies, AI must be handled with great responsibility to manage the risks and harness its potential for the betterment of society." Hendrycks et al (2023)- Center for AI Safety)

RISK

Bias	→	AI can reinforce existing societal inequalities (binns, 2018).
Privacy violations	→	misuse of personal data (European commission, 2021).
Transparency issues	→	lack of clear explanations for ai-driven decisions (doshi-velez & kim, 2017).
Regulatory non-compliance	→	adapting to evolving laws (nist, 2022).
Human oversight	→	the danger of blindly trusting ai decisions (raji et al., 2020

Live Poll Q-Three

Give us your thought and opinion on the preparedness of your organisation to address AI ethics challenges.

Click on this link to acess the live poll

In Summary

▶ AI in governance requires ethical oversight (Raji et al., 2020).

▶ Regulations vary by region compliance is crucial (European Commission, 2021; NIST, 2022).

▶ Address bias, privacy, and transparency concerns proactively (Kamiran et al., 2012).

▶ Continuous assessment and improvement are necessary (Floridi & Cowls, 2019).

For your further Reading

- European Commission (2021). **AI Act.**

- NIST (2022). **AI Risk Management Framework.**

- Floridi & Cowls (2019). **A Unified Framework of Five Principles for AI in Society.**

- Mittelstadt (2019). **Principles Alone Cannot Guarantee Ethical AI.**

- Raji et al. (2020). **Closing the AI Accountability Gap.**

- Kamiran et al. (2012). **Fairness in Machine Learning.**

Thank you all joining this session !
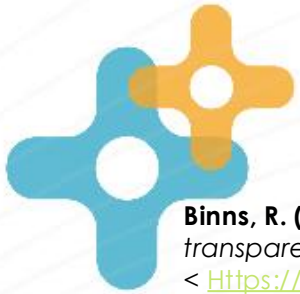
References

Binns, R. (2018). Fairness in machine learning: lessons from political philosophy. *Proceedings of the 2018 conference on fairness, accountability, and transparency*, 149–159.
< Https://proceedings.Mlr.Press/v81/binns18a.Html >Assessed 16 February 2025

Doshi-velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. *Arxiv preprint arxiv:1702.08608.*
< Https://arxiv.Org/abs/1702.08608 > Assessed 10 February 2025

Hendrycks, D., Et al. (2023). An overview of catastrophic AI risks.
< Https://www.Safe.Ai/ai-risk >March 17, 2025

European commission. (2021). Proposal for a regulation laying down harmonised rules on artificial intelligence (artificial intelligence act).
< Https://eur-lex.Europa.Eu/legal-content/EN/TXT/?Uri=celex%3a52021pc0206 >Assessed 10 February 2025

Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard data science review, 1*(1).
Https://hdsr.Mitpress.Mit.Edu/pub/l0jsh9d1

Global AI law and policy tracker.
<Https://iapp.Org/media/pdf/resource_center/global_ai_law_policy_tracker.Pdf> Assessed 10 February 2025

Kamiran, F., Calders, T., & Pechenizkiy, M. (2012). Data preprocessing techniques for classification without discrimination. *Knowledge and information systems, 33*(1), 1–33.
Https://link.Springer.Com/article/10.1007/s10115-011-0463-8 Assessed 10 February 2025

Mittelstadt, B. D. (2019). Principles alone cannot guarantee ethical AI. *Nature machine intelligence, 1*(11), 501–507.
< Https://www.Nature.Com/articles/s42256-019-0114-4 >Assessed 10 February 2025

National institute of standards and technology (NIST). (2022). AI risk management framework.
< Https://www.Nist.Gov/itl/ai-risk-management-framework >. Assessed 18 February 2025

Raji, I. D., Et al. (2020). Closing the AI accountability gap. *Proceedings of the 2020 conference on fairness, accountability, and transparency.*
< Https://dl.Acm.Org/doi/10.1145/3351095.3372873> Assessed 10 February 2025

Sharma, R. (N.D.). Creating ai strategy framework: the blueprint for ai success.
< Https://markovate.Com/blog/ai-strategy-framework/ > Assessed 10 February 2025

THANK YOU!

Kenneth Atuma
kenneth.atuma@manchester.ac.uk