

DATA GOVERNANCE AS KEYSTONE FOR COMPLIANT AI AND DIGITAL TRUST

March 27th, 2024

Jan Henderyckx
Partner



Be the
Boss of AI

— AI: Happening here

Powered by DALL·E 3

Your Presenter



JAN HENDERYCKX,
Partner, BearingPoint



JANHENDERYCKX



JANHENDERYCKX

Regional leader FBLA Data & Analytics and AI
CoE leader for Data Governance
Practice leader Belgium

- Data Governance and Data Strategy
- AI
- Data Architecture And Strategy
- Data Quality
- Business Intelligence and Analytics
- Master- And Reference Data Management

Seminars and workshops
IRMUK, SAI, Adept Events, IT Works

Involvement in non-profit initiatives

- Past International Board member DAMA
- President of DAMA Belux Chapter

“CHALLENGING SYSTEMATIC PREJUDICES: AN INVESTIGATION INTO GENDER BIAS IN LARGE LANGUAGE MODELS”.

IRC AI

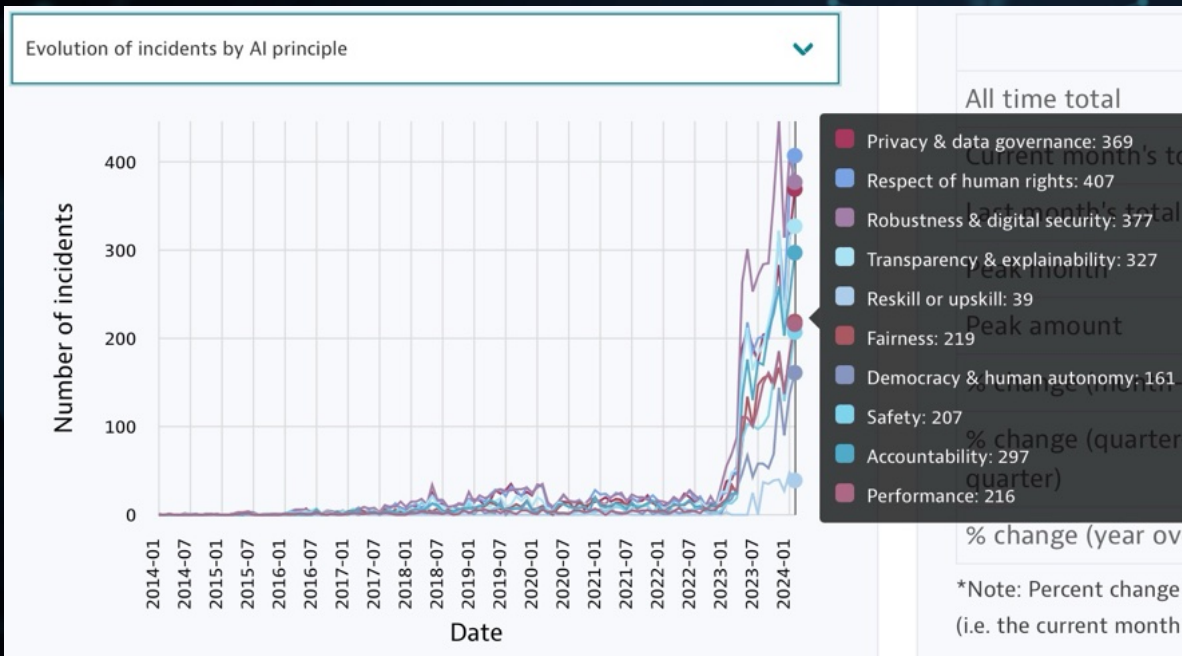
International Research Centre on Artificial Intelligence under the auspices of UNESCO

Is Google's Gemini chatbot woke by accident, or by design?

The tech giant's new artificial intelligence model veers black Vikings and Asian popes



Are we experiencing a “monkey with a hand grenade” moment?

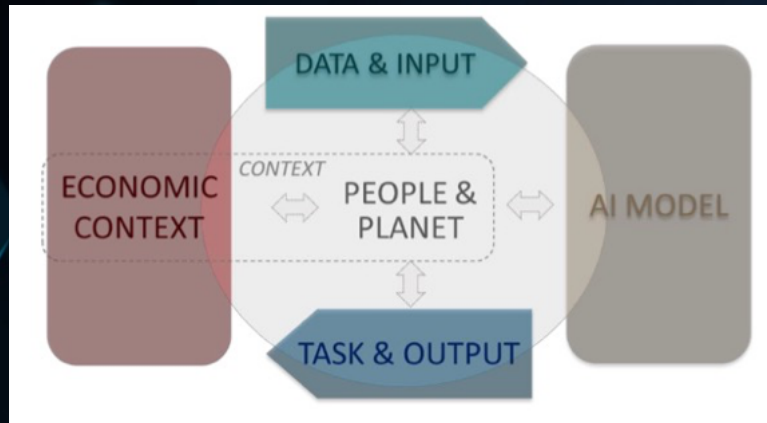


 OECD AI Incidents Monitor (AIM)



Powered by DALL-E 3

What could go wrong with the application of (GEN)AI?



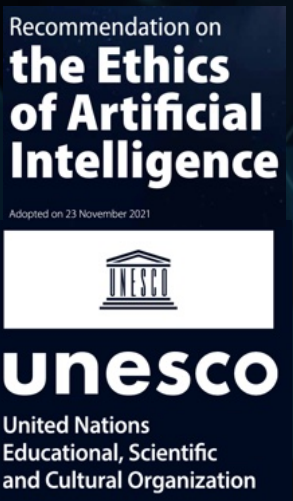
© OECD 2022

Key high-level dimensions of the OECD Framework for the Classification of AI Systems



NIST AI 100-1

The evolving landscape of regulations and standards shaping ethical data use.



AI Act



OCTOBER 30, 2023

Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence

BRIEFING ROOM • PRESIDENTIAL ACTIONS

Establishing the US approach to AI development and use

- **Principles:**

- Safety, security, and **trustworthiness** in AI development and use
- Economic growth, national security, and global competitiveness

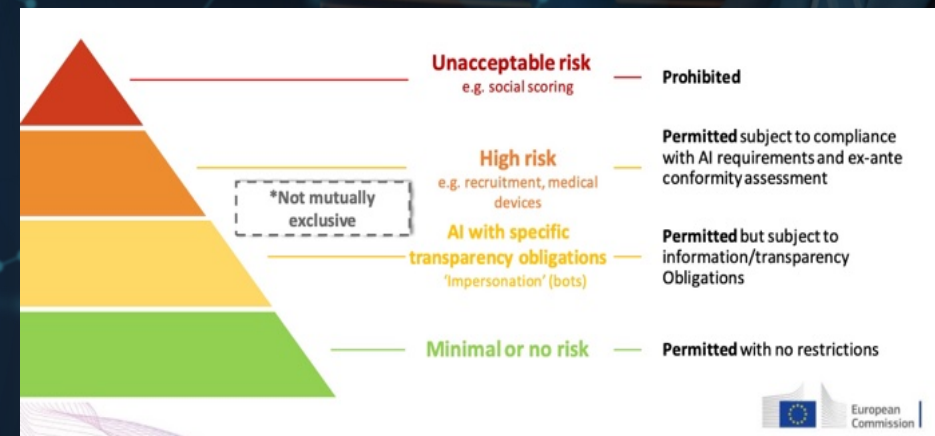
- **Key actions**

- Promoting research and development in AI
- Enhancing **access to data** and computing resources for AI development
- Fostering public **trust in AI and ensuring ethical and legal use**
- Establishing **guidelines for AI use** in collaboration with stakeholders
- Encouraging international cooperation in AI development

Powered by DALL·E 3

EU AI ACT

- December 2023, the European Parliament and the Council of the EU reached a political agreement on the AI Act it will most likely be formalised in April '24.
- The AI Act, fully applicable 2 years later, with some exceptions: prohibitions will take effect after six months, the governance rules and the obligations for general-purpose AI models become applicable after 12 months and the rules for AI systems - embedded into regulated products - will apply after 36 months.
- Penalties and fines: Sanctions for violations of the regulations.
- Prohibited practices: Certain AI practices deemed unacceptable will be banned.
- High-risk AI systems: Strict obligations and requirements for AI systems considered high-risk.
- Market surveillance and governance: Mechanisms for monitoring and enforcing compliance.
- European AI Board: Establishment of a regulatory body to support the implementation of the AI Act.



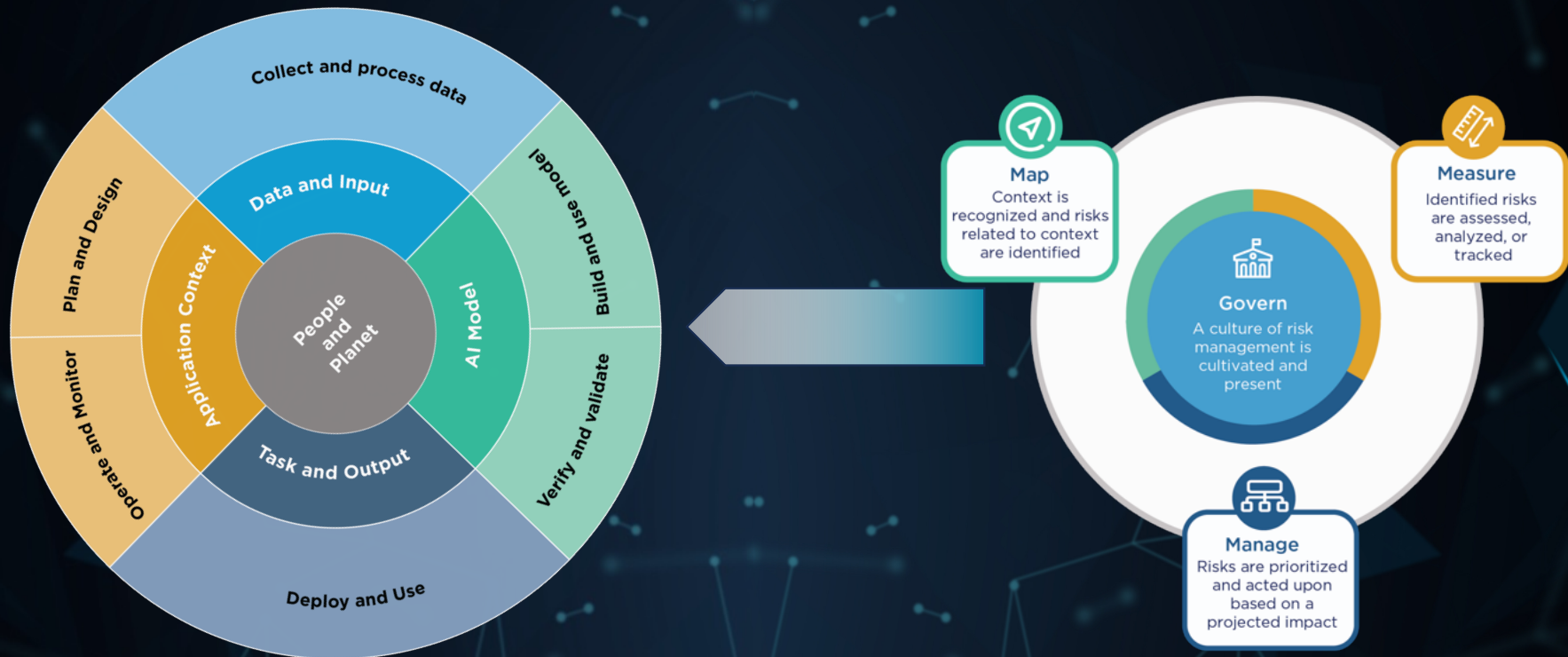
Powered by DALL·E 3

EU AI ACT

- **RISK ASSESSMENT:** Companies must assess the risk level of their AI systems based on the intended use and potential harm to individuals or society. High-risk AI systems will be subject to stricter requirements.
- **DATA QUALITY:** Companies must ensure that the data used to train and test their AI systems is relevant, representative, and free from errors or bias.
- **TRANSPARENCY:** Companies must provide clear and understandable information about their AI systems, including their purpose, functionality, and limitations.
- **HUMAN OVERSIGHT:** Companies must ensure that their AI systems are designed and used in a way that allows for effective human oversight, including the ability to intervene and override the system's decisions.
- **ROBUSTNESS AND ACCURACY:** Companies must ensure that their AI systems are robust, accurate, and secure, and that they perform consistently in different environments.
- **RECORD-KEEPING:** Companies must keep records of their AI systems' development, testing, and deployment, and make them available to regulatory authorities upon request.
- **CONFORMITY ASSESSMENT:** High-risk AI systems must undergo a conformity assessment before they can be placed on the EU market. This may involve independent third-party testing and certification.

Powered by DALL·E 3

Managing the “monkey with a hand grenade” AI risks Towards an AI Governance Framework



NIST AI 100-1

“Risk management refers to coordinated activities to direct and control an organization with regard to risk” (Source: iso 31000:2018).

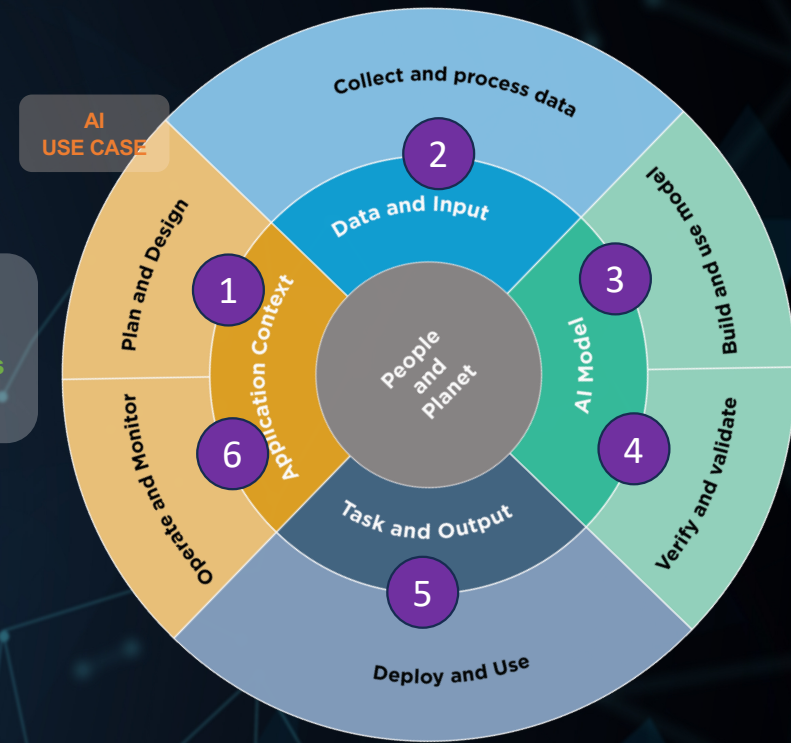
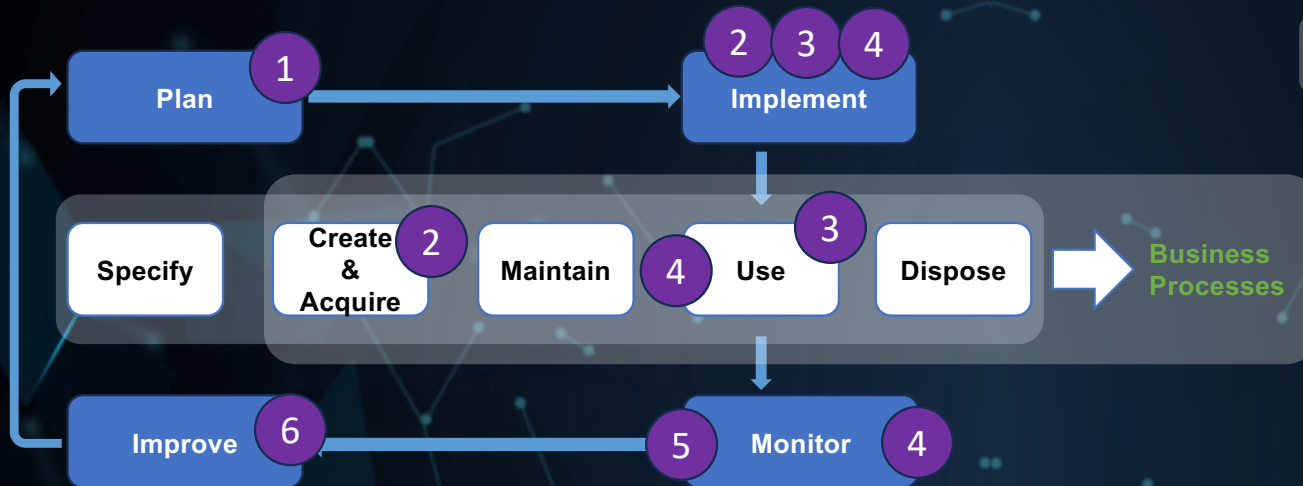
AI Lifecycle Activities

Handling the “monkey with a hand grenade” AI risks

- **Use Case Planning:**
 - What’s the objective of the use case? Are we allowed to do the processing? Are the use cases documented?
- **Use Case Design:**
 - Do we have the data to train the model? Are there any Data Privacy/IP rights constraints
- **Model Building:**
 - Are we using the most relevant model approach? Is the relevant data available in the right format?
- **Model Validation:**
 - Is the model fair and transparent? Is the data quality of the source data sufficient?
- **Model Deployment:**
 - Can we do proper model governance? Is the model secure?
- **Model Operations:**
 - Is there model drift? Is there human oversight? Can a person exercise their rights?
Is the full lifecycle documented?

Towards a mature AI Governance Framework

- How does a typical Data Governance Framework link to a regulatory compliant AI Governance Framework?



- Data Lifecycle
- Data Governance
- Business Process
- AI USE CASE

NIST AI 100-1

Towards a mature AI Governance Framework

Assure your AI is FAT Compliant:

Fair:

Do you respect privacy, IP and copyright law and is there no Bias and Discrimination in the model?

Accountability:

Implement appropriate and effective measures to ensure that principles are complied with

Demonstrate compliance of the measures upon request

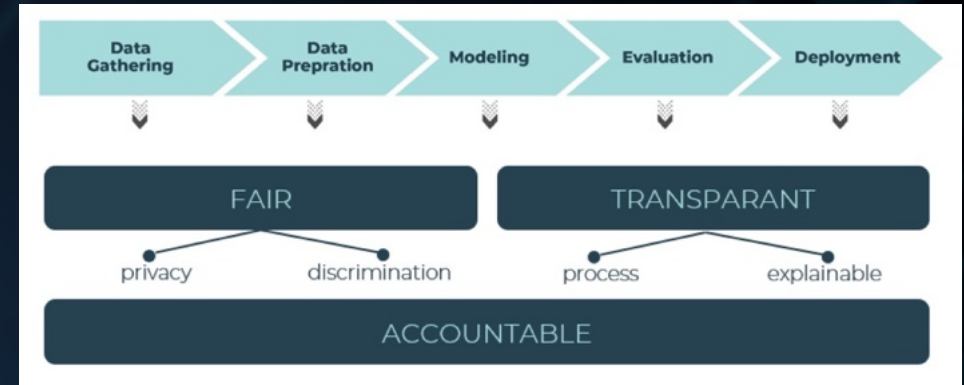
Recognize potential negative consequences.

Transparency:

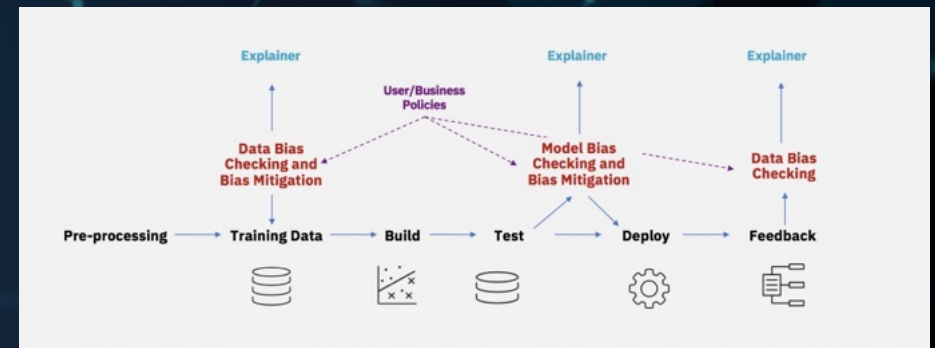
Extends beyond explaining decisions and encompasses all stages of a data science project.

Apply techniques such as CHAP or LIME.

Not necessitate revealing every detail to end-users to protect company secrets and privacy. Feature importance analysis can support this.



FAT Flow: a Data Science Ethics Framework David Martens (UNIVERSITY OF ANTWERP)



IBM: AI Fairness 360

Towards a mature AI Governance Framework

Assure your AI is FAT Compliant:

It's highly likely that you don't have the adequate infrastructure to assure your model deployment is error proof

Reduce the risks by:

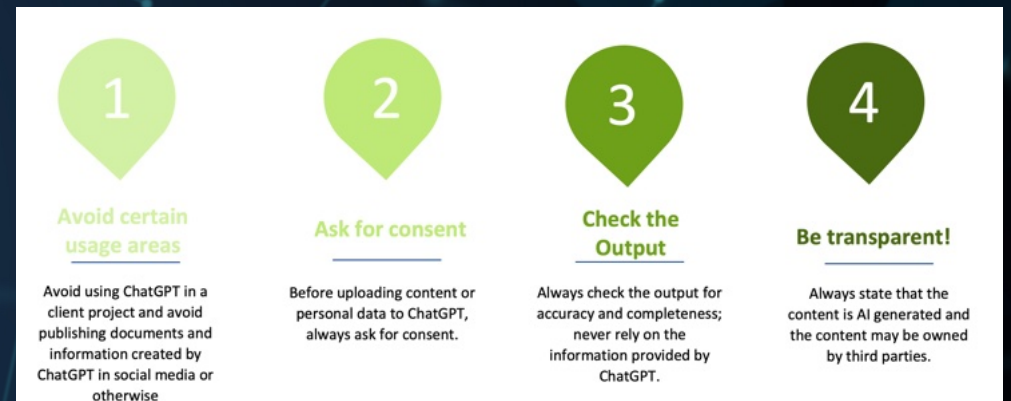
- **Putting a « red team » in place**
- **Adapt the model size to match the purpose (LORA, QLORA, RAG)**
- **Assuring there is sufficient AI literacy in the organisation**



Towards a mature AI Governance Framework

Apply critical thinking

- What are the Trust requirements for you data product?
 - Is statistical relevance (inferred or trained) appropriate for the use of your data product?
- We need to extend the notion of data literacy to include both the question and the answer side of the equation.
- Plan for a broad data (science) and AI awareness session in your organisation in the long run and GPT solutions immediately.



Data Quality

What are relevant dimensions to govern?

We have been using 6 data quality dimensions to cover the needs to govern our data quality framework.







They have been proven largely sufficient to cover DQ related risks including Data Privacy requirements.

Can they cover all aspects of AI Governance such as bias, fairness, ...?



Data Quality

What specific AI and machine learning quality dimensions should be observed?

-  **RELEVANCY:** The data should be relevant to the problem the AI model is trying to solve. Irrelevant data can introduce noise and negatively impact model performance.
-  **BALANCE:** In classification problems, the dataset should ideally have a balanced representation of each class. Imbalanced data can lead to biased models that perform poorly on the underrepresented class.
-  **ABSENCE OF BIAS:** The data should be free of biases to prevent the AI model from learning and perpetuating these biases. Biased data can lead to unfair or discriminatory outcomes.
-  **REPRESENTATIVENESS:** The data should be representative of the real-world scenarios where the AI model will be applied. If the training data is not representative, the model may not generalize well to unseen data.
-  **VARIABILITY:** The data should cover a wide range of scenarios and edge cases to ensure the model can handle a variety of inputs.
-  **FEATURE QUALITY:** The features or variables used in the model should be well-engineered, informative, and have minimal multicollinearity.

Towards a mature AI Governance Framework

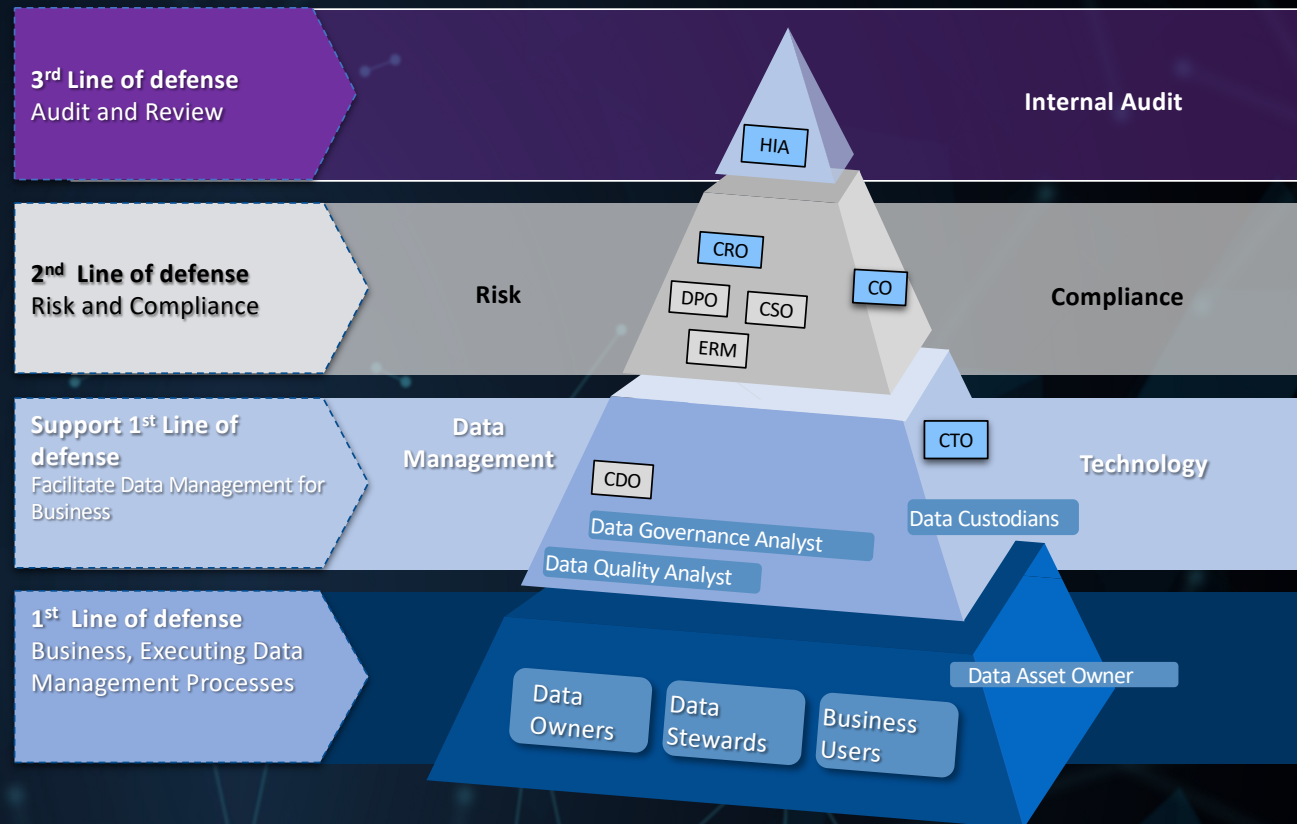
AI Observability

A comprehensive AI Governance Framework covers both:

- **The static metadata**
 - Use case description
 - Model documentation
- **The dynamic metadata**
 - The actual training dataset
 - The model test case and test results
 - The model usage data (prompts) and outcome

Extending the Data Governance Operating Model

- We need to extend the data governance framework with AI specific roles, policies, procedures and deliverables.
- AI Ethics and Trustworthiness need to be transparent.
- AI engineers must be aware of the constraints, business users need to understand the underlying principles and limitations.
- AI ethics need to be integrated in the data strategy.



Getting ready to manage (Generative) AI?

- ✓ **Include AI in your data strategy but recognize it's just a tool**
 - Actively consider the potential use case
- ✓ **Extend your Governance Framework to include AI aspects**
 - Data landscape
 - Ethical use
 - Legal considerations such as IP infringement
- ✓ **Put a strong focus on Data Literacy, don't miss the potential but also mitigate the risks**
- ✓ **Actively manage *Your* data (quality) throughout its full lifecycle**



"With great power comes great responsibility!"
Famous philosopher, Spider-Man

RESPECT HUMAN DIGNITY AND AUTONOMY:

GENAI should not be used to create fake news, deepfakes, or propaganda that can harm people's reputation, privacy, or freedom.

PROMOTE DIVERSITY AND INCLUSION:

GENAI should not generate content that is racist, sexist, or offensive to any group of people.

ENSURE TRANSPARENCY AND ACCOUNTABILITY:

GENAI should not be a black box that hides its logic, data, or outcomes from its users or regulators.

FOSTER TRUST AND COLLABORATION:

GENAI should not replace human workers, but rather augment their skills and creativity



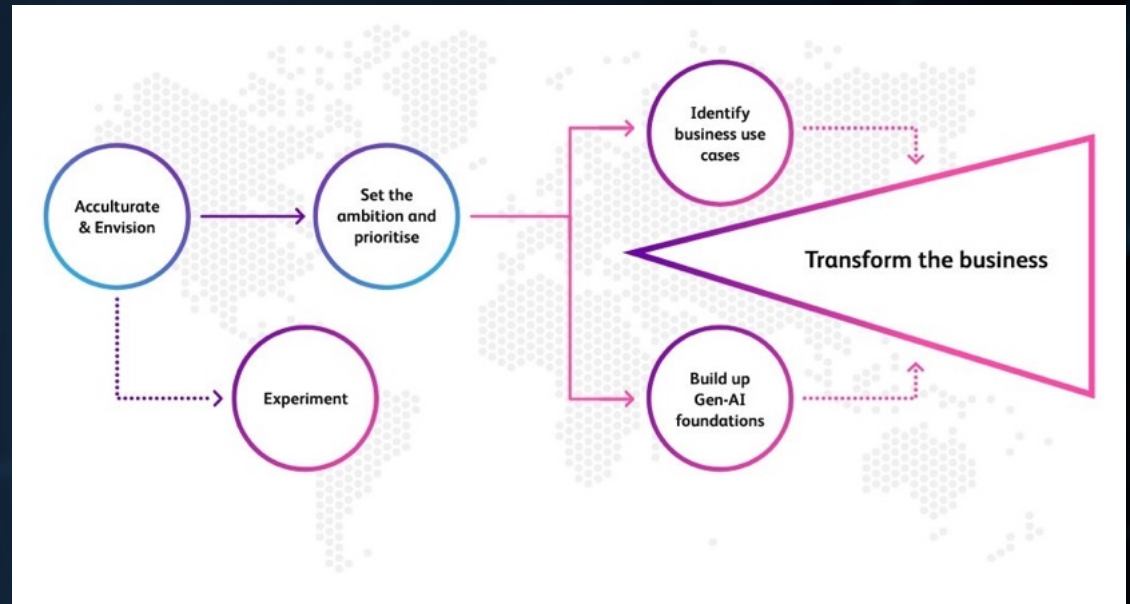
Is your organization ready?



Take our self assessment

Be the Boss of AI

— AI: Happening here





OECD.AI
Policy Observatory

OECD publishing

OECD FRAMEWORK FOR THE CLASSIFICATION OF AI SYSTEMS

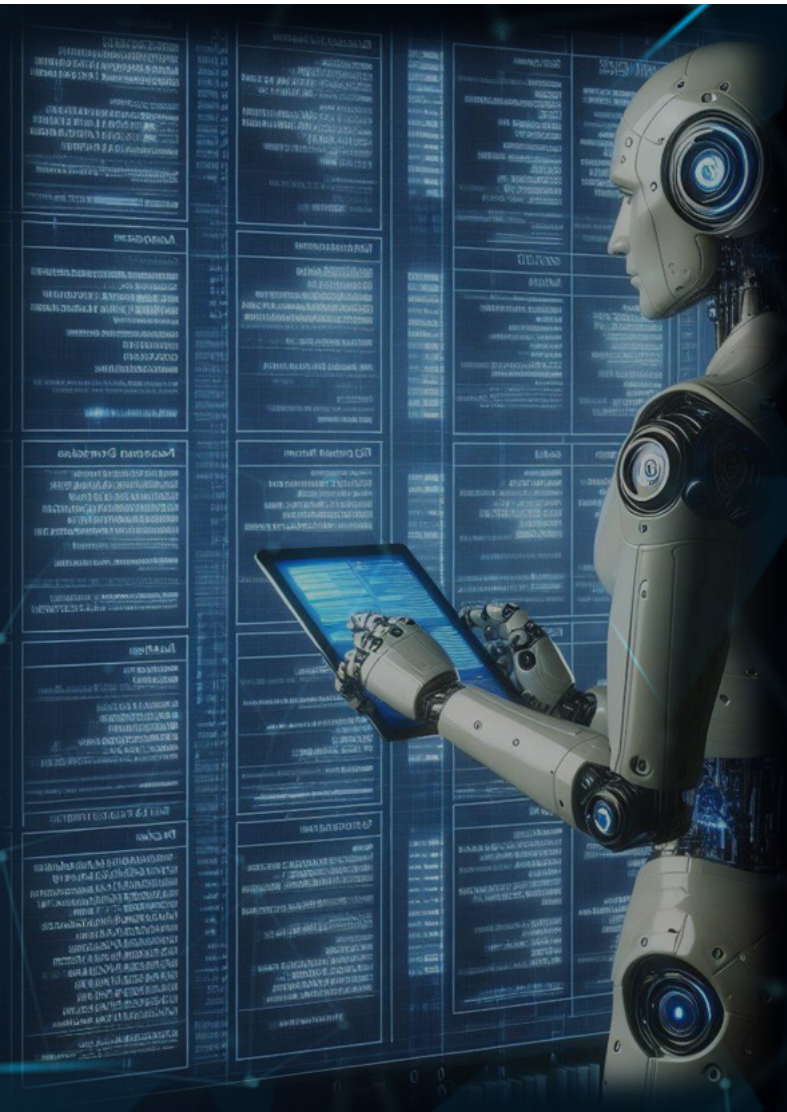
OECD DIGITAL ECONOMY
PAPERS

February 2022 No. 323



EUROPEAN
COMMISSION

<https://digital-strategy.ec.europa.eu/en/policies/expert-group-ai>





DATA GOVERNANCE AND MASTER DATA MANAGEMENT CONFERENCE EUROPE

11 - 14 March 2024 | London, UK

****Please score and comment on this session and speaker
in the event mobile app****



AI Best Practice

- Provide detailed model documentation
- Visualize the model's structure and flow
- Conduct feature importance analysis
- Apply interpretability techniques
- Generate explanations for model predictions
- Regularly validate and evaluate the model's performance
- Utilize open-source frameworks with built-in interpretability features
- Be transparent about data sources, biases, and limitations
- Consider ethical implications and communicate potential risks or biases

