Compliance Institute
BETTER BUSINESS ETHICS

TRILATERAL RESEARCH
Ethical AI

IFCA
International Federation of Compliance Associations

Understanding the EU AI Act

10 th April 2024

compliance.ie

Welcome & Introduction

- Thank you for registering

- Questions
  - Please use the question box on the right of your screen to send the questions for our speaker

- Today’s session will be recorded and will be on our website later today

- The CPD code is noted below and will be sent out directly after this session has concluded

CPD CODE: 2024 - 0681

TRILATERAL
RESEARCH
Ethical AI

Understanding the EU AI Act

Rachel Finn (PhD), David Barnard-Wills (PhD), Sara Domingo Andres (LL.M)

WHO WE ARE

Ethical AI company based in UK/IE, est. 2004, 150+ personnel

Pioneering AI Solutions for

Pharma & Healthcare
Hospitality & Retail
Finance & FinTech
Manufacturing & Industry
Government & Defence
Law enforcement

Services in Responsible AI, Data protection, Cybersecurity, Innovation & Research

Ministry of Defence

University of Cambridge

The Alan Turing Institute

Department for Business Innovation & Skills

Sláinte Leanaí Éireann ag Sráid an Teampaill
CHI
Children's Health Ireland at Temple Street

GMCA Greater Manchester Combined Authority

unicef

West Yorkshire Police

Home Office

Bord Bia Irish Food Board

European Central Bank Eurosystem

ico. Information Commissioner's Office

UNHCR The UN Refugee Agency

NATO OTAN

V&A

World Health Organization

EBA European Banking Authority

[dstl] The Science Inside

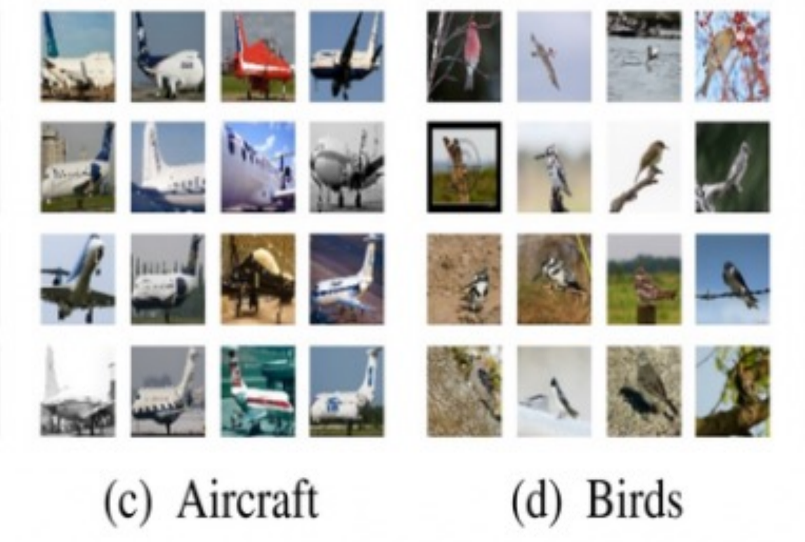
enisa European Union Agency for Cybersecurity

An Roinn Leanaí agus Gnóthaí Óige
Department of Children and Youth Affairs

AI Act concepts and basics

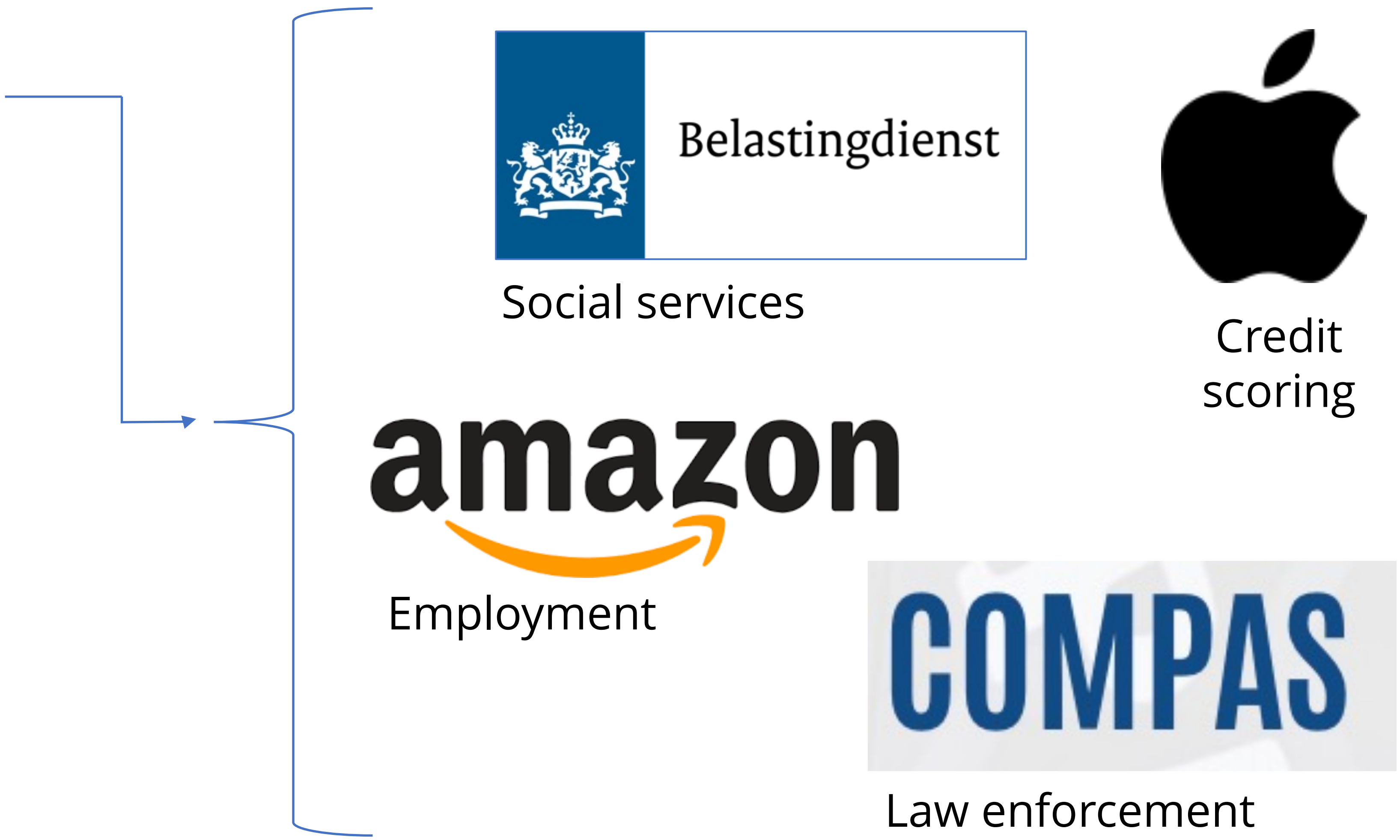
What is AI?

- **Artificial intelligence** (**AI**) systems are machine-based tools that perform tasks such as prediction, categorisation, pattern-identification, self-correction, or content-generation.
- AI systems based on **machine learning** (**ML**) are not programmed the "Good Old Fashioned" way, with specific instructions. They are "trained" on examples.
  - Images, text or previous decisions
  - Large data sets



Why does AI need to be regulated?

- Discrimination / bias
- Privacy violations
  - Combining multiple datasets
- Intellectual property concerns
  - Training data
- Hallucination



Social services

Credit scoring

Employment

Law enforcement

What factors make AI responsible?

1. Human Agency and Oversight
2. Technical Robustness and Safety
3. Privacy and Data Governance
4. Transparency
5. Diversity, Non-discrimination and Fairness
6. Societal and Environmental Well-being
7. Accountability

INDEPENDENT
HIGH-LEVEL EXPERT GROUP ON
ARTIFICIAL INTELLIGENCE
SET UP BY THE EUROPEAN COMMISSION

AI

THE ASSESSMENT LIST FOR
TRUSTWORTHY ARTIFICIAL
INTELLIGENCE (ALTAI)
for self assessment

The EU AI Act

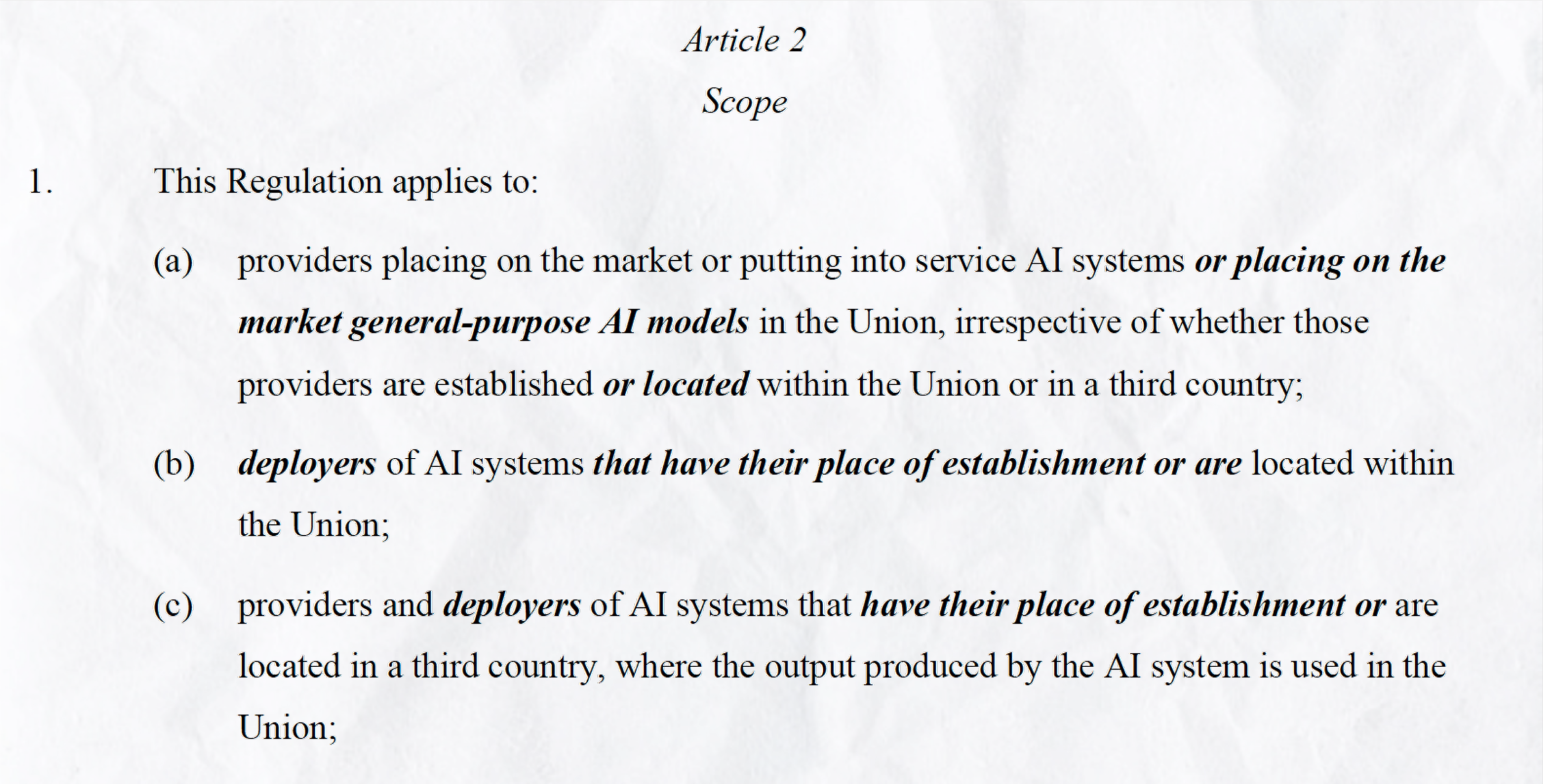
Introduction to the Act

- **First piece of comprehensive AI legislation in the world**
- 113 Articles, 180 Recitals and 13 annexes
- Clear rules for the governance and management of AI systems, including
  - strong consideration of **risk management** and **fundamental rights**.
  - **support innovation** by setting rules that create transparency and harmonisation for industry to support responsible development, deployment and marketing of AI tools.
- AI Act is likely to set a benchmark against which other countries will develop their own legislation

Technological scope

- Machine-based systems that:
  - operate with some level of autonomy,
  - that may be adaptive after they are deployed and
  - generate content, predictions, recommendations or decisions that influence the environment with which they are interacting. (Art 3.1)
- As such, it applies to tools like generative AI, digital advertising, facial recognition and other predictive tools.

Territorial scope



Article 2

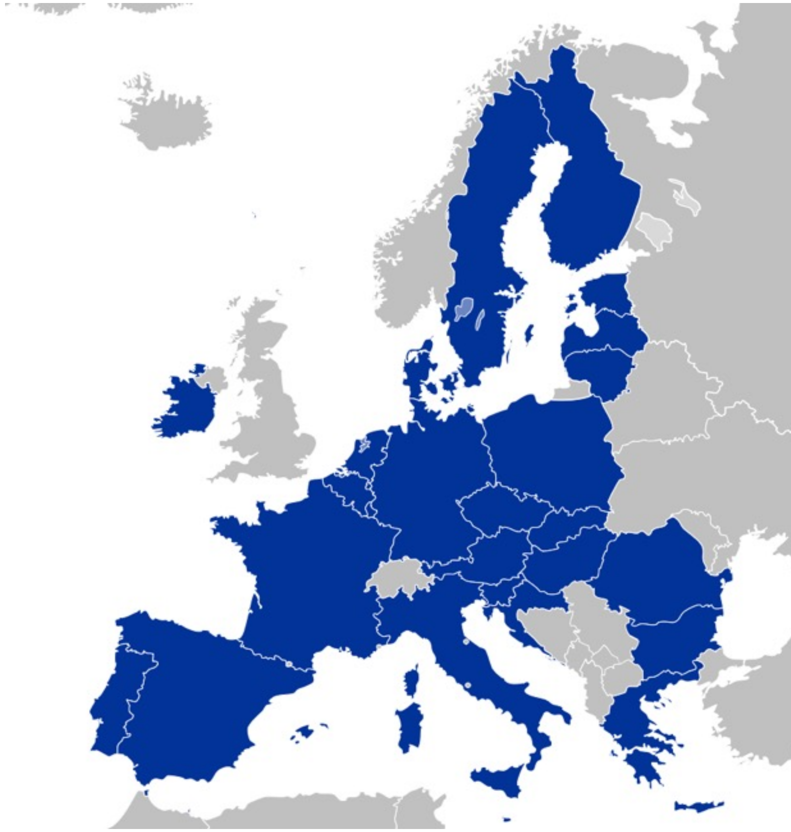
Scope

1. This Regulation applies to:

(a) providers placing on the market or putting into service AI systems ***or placing on the market general-purpose AI models*** in the Union, irrespective of whether those providers are established ***or located*** within the Union or in a third country;

(b) ***deployers*** of AI systems ***that have their place of establishment or are*** located within the Union;

(c) providers and ***deployers*** of AI systems that ***have their place of establishment or*** are located in a third country, where the output produced by the AI system is used in the Union;



Organisational scope

- The Act includes different obligations for different roles in the AI ecosystem:
  - Developers
  - Providers
  - Distributors
  - Deployers / users

- Out of scope
  - Scientific research and development
  - Research & development prior to placing a system on the market
  - Military or national security uses of AI by Member States

Risk management

Performance

Fundamental rights

protection

Transparency

AI literacy

What is the EU AI Act risk framework?

IFCA
International Federation of Compliance Associations

Compliance Institute
BETTER BUSINESS ETHICS

TRILATERAL RESEARCH
Ethical AI

Risk in the AI Act

- Categorising types of AI system by the level of risk to **health, safety and fundamental rights** is a big part of the philosophy of the Act.
  - Not about ‘existential AI risk’
  - But based on foreseeable, practical risks from prior technologies and research

- A huge part of the Act are the obligations on providers and deployers of high-risk systems
- Risk categories
  - Unacceptable risk - Prohibited AI
  - High risk AI – increased obligations
  - ‘Certain AI systems’ – specific cases, where managing risks requires particular transparency practices.
  - Systemic Risk - a particular type of risk posed by general purpose AI based upon their capability and market reach.
  - No explicit ‘low risk / minimal risk’ in the act

Prohibited AI systems (Art.5)

- **Manipulative** – AI systems that use subliminal, manipulative or deceptive techniques to materially distort people's behaviour and causing them to take decisions they otherwise wouldn't
- **Exploit vulnerabilities** – AI systems that exploit vulnerabilities due to age, disability, social or economic situation in a manner that causes significant harm
- **Social scoring** – AI systems that classify people based on social behaviour or personality leading to detrimental treatment in unrelated contexts or that is unjustified/disproportionate.
- **Offender prediction** - AI systems making predictions of likelihood of a person committing a criminal offence based solely on profiling or character traits
- **Untargeted scraping** – AI systems that create/expand facial recognition database through untargeted scraping of facial images from the internet or CCTV
- **Emotion recognition** – AI systems that infer people's emotions in workplace or educational institutions (medical or safety reasons are ok).
- **Biometric inferences** - AI systems that infer race, political opinions, trade union membership, religious or philosophical beliefs, sex life or sexual orientation from biometric data
- **Real time law enforcement biometrics in public places** – use of AI for real-time, biometric identification in publicly accessible space for the purposes of law enforcement
  - Except this one is very complex, and there are lots of carve-outs around crime types, urgency, threat to life, judicial authorization etc.

High Risk AI

Several ways for an AI system to count as **high risk:**

1) It is a **product** (or a safety system in a product) that is covered by a big list of EU product legislation (Annex I)

   And would have had to do conformity assessment under that legislation

2) It is intended for use in one of eight **domains** that are pre-identified as risky places to use AI (Annex III)

   and doesn't have an opt out

3) It is profiling natural persons

4) The provider has some other reason to believe that a particular AI in a particular context is a high risk to safety, health or fundamental rights.

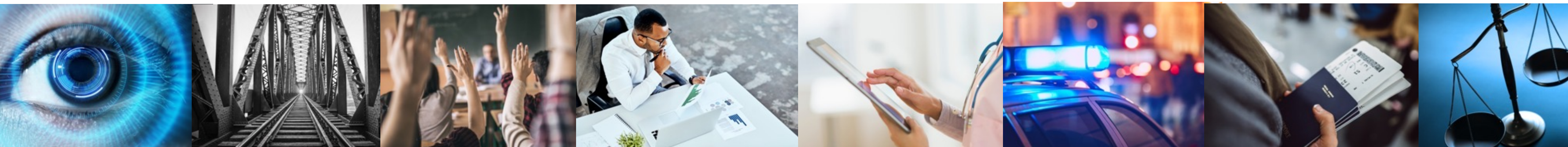
What types of products? (Annex I)

- Machinery
- Toys
- Recreational and personal watercraft
- Lifts
- Protective systems in explosive atmospheres
- Radio equipment
- Pressure equipment
- Cableways
- Personal protective equipment
- Appliances burning gases
- Medical devices
- Civil aviation
- 2, 3 wheeled vehicles
- Agricultural and forestry vehicles
- Marine equipment
- Railway equipment
- Motor vehicles and trailers

The details of a specific product and if it needs to do a conformity assessment come from those other piece of regulation

What types of use case ? (Annex III)

1. Biometrics
2. Critical infrastructure
3. Educational and vocational training
4. Employment, Workers Management and access to self-employment
5. Access to and enjoyment of essential private services and essential public services and benefits
6. Law enforcement
7. Migration, Asylum and border control management
8. Administration of justice and democratic processes



Obligations on high-risk AI systems & their providers

- Compliance with all applicable harmonisation legislation
- Risk management system
  - Inc identification of forseeable risks
- Data and data governance
  - Data quality requirements, bias assessment
- Procedures for the reporting of serious incidents
- Conformity assessment
- Quality management systems
- Automatic logging
- Corrective actions
- Cooperation with competent authorities

- **Technical documentation**
- **Transparency and provision of information to deployers**
- **Design for Human oversight**
- **Record keeping**
- **Accuracy, robustness and cybersecurity**
  - Model performance assessments
  - Security assessments
- **Fundamental rights impact assessments**
- **Registration with EU database**
- Post-market monitoring systems

‘Certain AI systems’

- Mixed bag of AI systems, where the risk is mainly managed using transparency requirements (Art. 50)
  - These can also be high risk, but needn’t be

- AI systems that
  - Directly interact with natural persons
  - Generate synthetic audio, image, video or text content
  - Recognise emotions
  - Categorise based upon biometrics
  - Generates deep fakes
  - Generate or manipulate news

General Purpose AI models **with systemic risk**

- General purpose AI models - Special type of AI model
  - Trained on very large amount of data
  - Capable of performing a wide range of tasks
  - Often fine-tuned for specific purposes or integrated into other AI systems
- Additional obligations on providers
- If a general purpose AI is so cutting edge, powerful, and with significant market reach it can present a **systemic risk**
- the potential for a widespread set of risks: e.g., lower barriers to entry for weapon design, offensive cyber capabilities, bias and discrimination, disinformation, harming privacy, impacts on democracy etc.
  - Additional requirements on providers of gen-purpose AI with systemic risks
  - Allowing for greater monitoring by the Commission
  - Commission can designate a GPAI as having systemic risk ex officio

What does the EU AI Act mean for generative AI?

IFCA
International Federation of Compliance Associations

Compliance Institute
BETTER BUSINESS ETHICS

TRILATERAL RESEARCH
Ethical AI

What is Generative AI?

- AI tools that can create new images, video, text, audio
  - Trained on huge volumes of existing images, video, text, audio, etc
  - Based upon a prompt from a user
  - Examples : ChatGPT, DALL-E, Midjourney, Gemini, Bard, Copilot
- Make AI tools very accessible in comparison other types of AI (e.g., machine learning)
  - Can work with unstructured data
  - Can ask questions and create content using normal human language
  - Automating of previously quite creative tasks (art, design, copywriting, coding).
- AI Act doesn’t explicitly mention generative AI
  - But it is practically there in several places
    - ‘Transparency requirements for certain AI systems’ (Art 50)
    - Exemptions from high risk (Art 6.3)
    - General purpose AI (Art 53)

Transparency requirements for ‘certain AI systems’

TRILATERAL RESEARCH
Ethical AI

- The AI Act create new additional transparency obligations for ‘certain AI systems’ (Art 50)

- Information should be provided to natural persons
  - in a clear and distinguishable manner
  - At the latest at the time of the first interaction or exposure

AI Systems that…	Transparency obligations	Article
Interact directly with people	• Providers must ensure persons are **informed they’re interacting with an AI.** • Unless it would be obvious to a reasonably well-informed person in that context	50.1
Generate synthetic audio, image, video or text content	• Providers must ensure these are **marked in a machine-readable format as artificial generated**. • Doesn't apply for basic editing	50.2
Generates or manipulate image, audio or video content constituting a deep fake	• Deployers must **disclose the content has been artificially generated**. • Limited disclosure requirements for art, satire, fiction, etc.	50.4
Generate or manipulate text published with the purpose of informing the public on matters of public interest (50.4)	• Deployers must **disclose the content has been artificially generated**. • Unless the AI content has undergone human review or editorial control and a natural or legal person holds responsibility	50.4

When Generative AI is ‘high-risk’

- Generative AI systems *might* be high risk
  - If they’re in a product covered by the legislation in Annex I
  - If they’re intend for use in the high-risk contexts Annex III
    - (E.g., education, democratic process, workplace management)

- But those Art 6.3 exemptions exclude a lot of generative AI applications
  - If the AI system is for a **narrow procedural task**
  - If the AI system is intended to **improve the result of previously completed human activity**

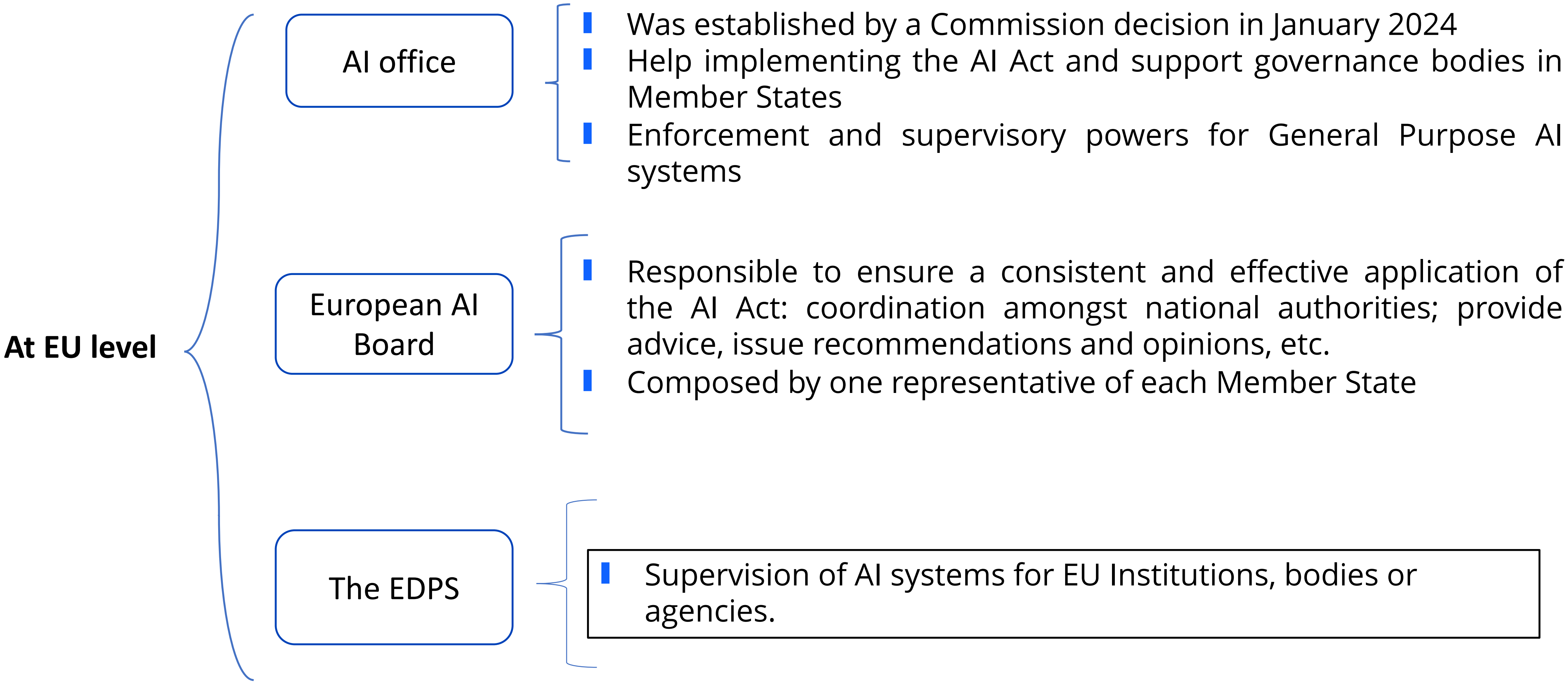
GenAI and General Purpose AI models

- Many generative AI tools are going to be based upon general purpose AI models
- Most of the obligations in the Act relating to general purpose AI are on **providers**
    - Including a policy to comply with EU copyright law

How will the AI Act be enforced?

Enforcement and supervisory authorities

Enforcement and supervisory mechanisms are as important as the approval of the AI Act.



At EU level

AI office

- Was established by a Commission decision in January 2024
- Help implementing the AI Act and support governance bodies in Member States
- Enforcement and supervisory powers for General Purpose AI systems

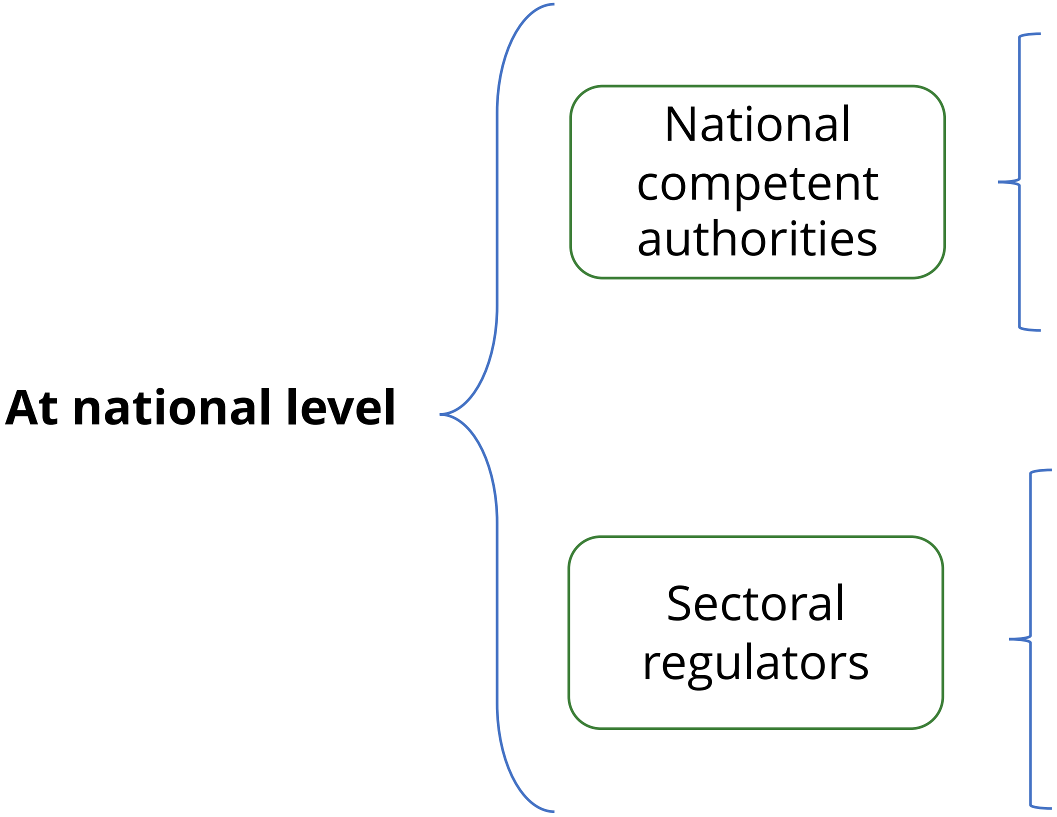
European AI Board

- Responsible to ensure a consistent and effective application of the AI Act: coordination amongst national authorities; provide advice, issue recommendations and opinions, etc.
- Composed by one representative of each Member State

The EDPS

- Supervision of AI systems for EU Institutions, bodies or agencies.

Enforcement and supervisory authorities



At national level

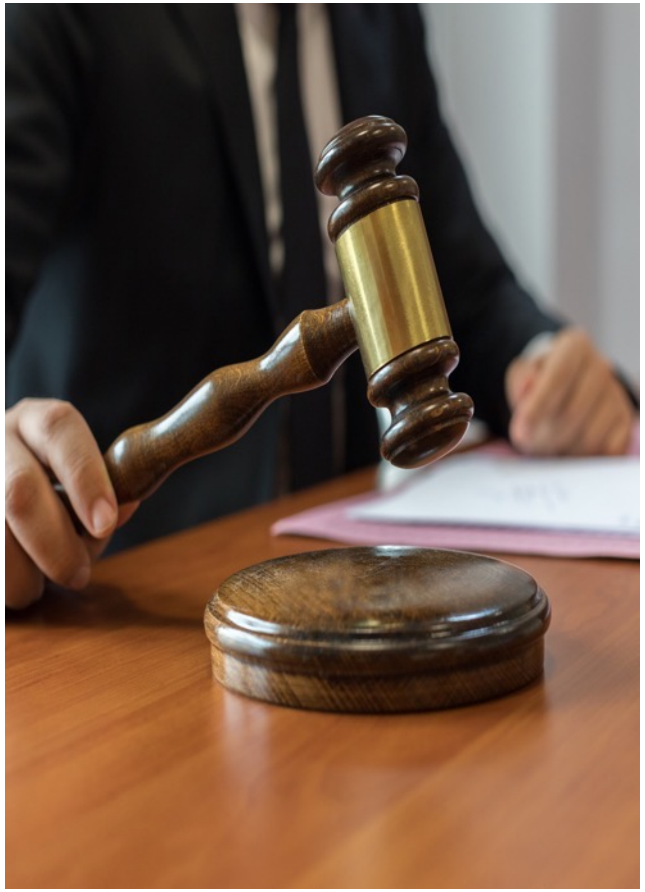
National competent authorities

Each Member State will need to designate at least one notifying authority and one market surveillance authority for the purposes of the AI Act (Article 70).

Sectoral regulators

For high-risk AI systems listed within Union harmonisation legislation, e.g. medical devices, aircrafts, cars: the enforcement authority will be the entity already in charge of regulating that technology. Although, Member States might designate other relevant authority (to be defined) (Article 74.3)

Penalties – Article 99



- For the use of **prohibited AI systems** fines may be up to **€35 million or 7%** of worldwide annual turnover, whichever is higher.
- Non-compliance with **other obligations**, fines may be up to **€15 million or 3%** of worldwide annual turnover, whichever is higher.
- The supply of **incorrect or misleading information** shall be subject to fines up to **€7.5 million or 1%** of the worldwide turnover, whichever is higher.
- **SMEs, including start-ups**, will be subject to the same, but up to the percentages or amounts that are lower.
- Each Member State shall lay down rules to decide to what extent administrative fines might be imposed on **public authorities**.
- The **EDPS** may impose fines on Union Institutions, bodies, offices and agencies (Article 100).

When does the AI Act come into force and be applicable?

Entry into force and applicability

Entry into force

- On the **20th day** following that of its publication in the Official Journal of the EU.
- Expected to be published in April → entry into force in **May 2024** (tbc).

Entry into force and applicability

Applicability

- **General** applicability → **24 months** from the entry into force → 2026 (tbc).
- **Prohibited practices** → **6 months** from the date of entry into force → November 2024 (tbc).

- General Purpose AI systems
- Appointment of national authorities
- Establishment of the AI office and the AI Board
- Penalties

} **12 months** from the date of entry into force → May 2025

- ➢ AI systems for safety components or products covered by **Union harmonisation legislation** (e.g. medical devices) → **36 months** from the date of entry into force → 2027 (tbc).

How can business prepare for the AI Act?

IFCA
International Federation of Compliance Associations

Compliance Institute
BETTER BUSINESS ETHICS

TRILATERAL RESEARCH
Ethical AI

IFCA International Federation of Compliance Associations Compliance Institute BETTER BUSINESS ETHICS TRILATERAL RESEARCH Ethical AI

Preparation

- Build an **inventory** of AI systems already in use and maintain it updated.
- Revise your **procurement guidance and contract management** policies to take account of AI features and functionalities to include responsible AI (e.g. responsible internal stakeholders, responsible AI providers).
- Use **privacy-by-design** and **ethics-by-design mechanisms** to ensure that any AI tools you are developing and implementing take the requirements of the legislation into consideration.
- Invest in **AI training** and awareness: AI literacy will be mandatory under the AI Act.
- **Keep up to date** with regulatory changes and forthcoming guidance.
- Consider the appointment of an **AI officer.**

Thank you!

IFCA
International Federation of Compliance Associations

Compliance Institute
BETTER BUSINESS ETHICS

TRILATERAL RESEARCH
Ethical AI

Compliance Institute

BETTER BUSINESS ETHICS

Questions & Answers

CPD CODE: 2024 - 0681

Thank You For Attending
Understanding the EU AI Act

A recording of this webinar and the CPD code will be available on our website later today.

CPD CODE – 2024 - 0681

Compliance Institute
BETTER BUSINESS ETHICS