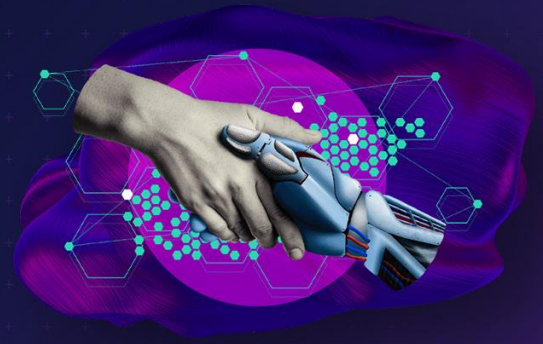




# AI Pact

## Organisations' commitments



Please note that the pledges presented in this document are in draft form and are subject to modification. These pledges have been drafted by the AI Office and have been shared with the relevant stakeholders in the AI Pact network to gather feedback and insights. As such, the final version of the pledges will be shaped by the input received by stakeholders. We appreciate your understanding and contributions to this process.

Artificial Intelligence (AI) is a transformative technology with numerous beneficial effects. Yet, its advancement brings also potential risks. In light of this, the European Union has adopted the **first-ever comprehensive legal framework on AI worldwide, the AI Act**<sup>1</sup>. The majority of rules of the AI Act (for example, some requirements on the high-risk AI systems) will apply at the end of a transitional period (i.e. the time between entry into force and date of applicability). In this context and within the framework of the AI Pact, **the AI Office calls on all organisations to proactively anticipate and implement some of the key provisions of the AI Act, with the aim of mitigating the risks to health, safety and fundamental rights as soon as possible.**

By taking part in this initiative, **organisations agree to make three 'core' commitments** and may decide to strive to meet all the other commitments mentioned, or to select specific ones, depending on, for instance, their area of activity.

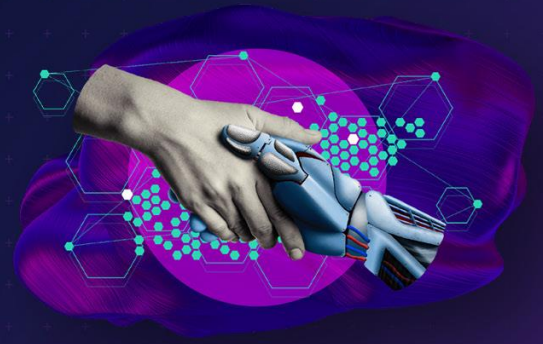
**The commitments are mostly focusing on transparency obligations and requirements for AI systems that are likely to classify as high-risk under the AI Act**<sup>2</sup>. Furthermore, depending on their position in the AI value chain, some organisations may not be able to meet themselves the commitments. In these cases, such organisations may also declare their intention to contribute to the best of their ability to the fulfilment of the commitments. When joining the initiative, organisations will have the opportunity to communicate how they intend to implement the commitments in their specific context to the general public.

Taking these considerations into account, for the period starting now and ending with the entry into application of the relevant provisions of the AI Act, **all organisations participating to this initiative agree to make the following 'core' commitments:**

- **adopt an AI governance strategy to foster the uptake of AI in the organisation and work towards future compliance with the AI Act;**
- **carry out a mapping of AI systems developed or used in areas that would be considered high-risk under the AI Act;**
- **promote awareness and AI literacy of staff and other persons dealing with the deployment of AI systems on their behalf, taking into account their technical knowledge, experience, education and training and the context the AI systems are to be used in, and considering the persons or groups of persons affected by the use of the AI systems.**

<sup>1</sup> [Regulation - EU - 2024/1689 - EN - EUR-Lex \(europa.eu\)](#)

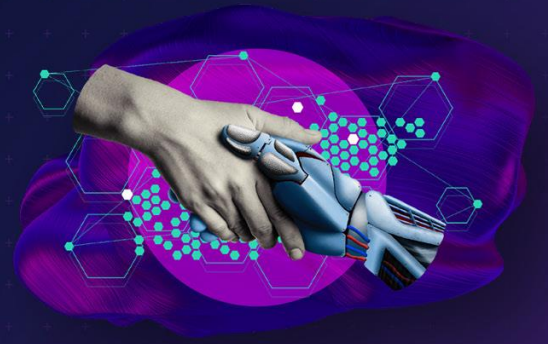
<sup>2</sup> Requirements on General Purpose AI models will mostly be dealt with by the Codes of Practice and should not be included in the commitments of the organisations taking part in the initiative.



In addition, the organisations may commit to make best efforts to meet or contribute to meeting the following:

- for **organisations that develop AI systems**, where relevant in light of future obligations:
  - **put in place processes to identify possible risks to health, safety and fundamental rights that could follow from the use of relevant AI systems throughout their lifecycle;**
  - **develop policies to ensure high-quality training, validation and testing datasets for relevant AI systems;**
  - **when developing all or certain AI systems, implement logging features to allow traceability;**
  - **inform deployers about how to appropriately use relevant AI systems, their capabilities, limitations and potential risks;**
  - **implement concrete measures to ensure human oversight over decisions that are recommended or taken by relevant AI systems;**
  - **implement policies and processes aimed at mitigating risks associated with the use of relevant AI systems, in line with the relevant obligations and requirements envisaged in the AI Act, to the extent feasible;**
  - **design AI systems intended to directly interact with individuals so that those are informed, as appropriate, that they are interacting with an AI system;**
  - **to the extent feasible, design generative AI systems so that AI-generated content is marked through technical solutions, such as watermarks and metadata identifiers;**
  - **provide means to clearly and distinguishably label AI-generated content, including image, audio or video constituting deep fakes<sup>3</sup>, as well as AI-generated text published to inform the public on matters of public interest, unless the text has been subject to human review and a natural or legal person holds editorial responsibility for its publication;**
- for **organisations that deploy AI systems**, where relevant in light of future obligations:
  - **carry out a mapping of possible risks to fundamental rights of persons and groups of individuals that may be affected through the use of relevant AI systems;**
  - **implement concrete measures that ensure human oversight over decisions that are recommended or taken by relevant AI systems;**
  - **clearly and distinguishably label deep fakes as well as AI-generated text published to inform the public on matters of public interest, unless the text has been subject to human review and a natural or legal person holds editorial responsibility for its publication;**
  - **ensure that individuals are informed, as appropriate, when they are directly interacting with an AI system;**
  - **inform with clear and meaningful explanations individuals when a decision made about them is prepared, recommended or taken by relevant AI systems;**
  - **when deploying relevant AI systems at the workplace, inform workers' representatives and affected workers.**

<sup>3</sup> AI-generated or manipulated image, audio or video content that resembles existing persons, objects, places, entities or events and would falsely appear to a person to be authentic or truthful.



The participating organisations consent to publicly share the commitments they intend to meet and to report on the outcome of the implementation of these commitments twelve months after the publication of their commitments.

