Senate Select Committee on Adopting Artificial Intelligence (AI)
ANSWERS TO QUESTIONS ON NOTICE
Department of Industry, Science and Resources
Inquiry into Adopting Artificial Intelligence (AI)
16 August 2024

AGENCY/DEPARTMENT: Department of Industry, Science and Resources

TOPIC: EU fact sheet

REFERENCE: Question on Notice (Hansard, Page 70-71)

QUESTION No.: 3

Senator SHOEBRIDGE: The other thing that is clear is the degree of urgency—that pretty much everybody has been communicating. I don't think we can wait for perfection. I'm assuming that's not your goal, to wait for perfection. Do we agree that it's not the goal to wait for perfection, but rather to act with as much alacrity as we can?
Ms Wilson: I think that's borderline asking me for an opinion. What I will say is that the government has been really clear: it is going to regulate the application of AI in high-risk settings. We are working that through, and we are absolutely informed by what is happening in the EU. It's still early days in the EU. In fact, I don't know whether you've been able to see it, but the EU did release a fact sheet yesterday. It's excellent. I'm very happy to provide that to the committee.
CHAIR: That would be helpful, thank you.

ANSWER

See attached PDF.

Artificial Intelligence – Questions and Answers*

Brussels, 1 August 2024

Why do we need to regulate the use of Artificial Intelligence?

The EU AI Act is the world's first comprehensive AI law. It aims to address risks to health, safety and fundamental rights. The regulation also protects democracy, rule of law and the environment.

The uptake of AI systems has a strong potential to bring societal benefits, economic growth and enhance EU innovation and global competitiveness. However, in certain cases, the specific characteristics of certain AI systems may create new risks related to user safety, including physical safety, and fundamental rights. Some powerful AI models that are being widely used could even pose systemic risks.

This leads to legal uncertainty and potentially slower uptake of AI technologies by public authorities, businesses and citizens, due to the lack of trust. Disparate regulatory responses by national authorities would risk fragmenting the internal market.

Responding to these challenges, legislative action was needed to ensure a well-functioning internal market for AI systems where both benefits and risks are adequately addressed.

To whom does the AI Act apply?

The legal framework will apply to both public and private actors inside and outside the EU as long as the **AI system** is placed on the Union market, or its use has an impact on people located in the EU.

The obligations can affect both providers (e.g. a developer of a CV-screening tool) and deployers of AI systems (e.g. a bank buying this screening tool). There are certain exemptions to the regulation. Research, development and prototyping activities that take place before an AI system is released on the market are not subject to these regulations. Additionally, AI systems that are exclusively designed for military, defense or national security purposes, are also exempt, regardless of the type of entity carrying out those activities.

What are the risk categories?

The AI Act introduces a uniform framework across all EU Member States, based on a forward-looking definition of AI and a risk-based approach:

- **Unacceptable risk**: A very limited set of particularly harmful uses of AI that contravene EU values because they violate fundamental rights and will therefore be banned:
  - **Exploitation of vulnerabilities of persons, manipulation and use of subliminal techniques;**
  - **Social scoring** for public and private purposes;
  - **Individual predictive policing** based solely on profiling people;
  - **Untargeted scraping** of internet or CCTV for facial images to build-up or expand databases;
  - **Emotion recognition in the workplace and education institutions**, unless for medical or safety reasons (i.e. monitoring the tiredness levels of a pilot);
  - **Biometric categorisation** of natural persons to deduce or infer their race, political opinions, trade union membership, religious or philosophical beliefs or sexual orientation. Labelling or filtering of datasets and categorising data in the field of law enforcement will still be possible;
  - **Real-time remote biometric identification in publicly accessible spaces by law enforcement,** subject to narrow exceptions (see below).
- The Commission will issue guidance on the prohibitions prior to their entry into force on 2 February 2025.

- **High-risk**: A limited number of AI systems defined in the proposal, potentially creating an adverse impact on people's safety or their fundamental rights (as protected by the EU Charter of Fundamental Rights), are considered to be high-risk. Annexed to the Act are the lists of high-risk AI systems, which can be reviewed to align with the evolution of AI use cases.
- These also include safety components of products covered by sectorial Union legislation. They will always be considered high-risk when subject to third-party conformity assessment under that sectorial legislation.
- Such high-risk AI systems include for example AI systems that assess whether somebody is able to receive a certain medical treatment, to get a certain job or loan to buy an apartment. Other high-risk AI systems are those being used by the police for profiling people or assessing their risk of committing a crime (unless prohibited under Article 5). And high-risk could also be AI systems operating robots, drones, or medical devices.
- **Specific transparency risk**: To foster trust, it is important to ensure transparency around the use of AI. Therefore, the AI Act introduces specific transparency requirements for certain AI applications, for example where there is a clear risk of manipulation (e.g. via the use of chatbots) or deep fakes. Users should be aware that they are interacting with a machine.
- **Minimal risk**: The majority of AI systems can be developed and used subject to the existing legislation without additional legal obligations. Voluntarily, providers of those systems may choose to apply the requirements for trustworthy AI and adhere to voluntary codes of conduct.

In addition, the AI Act considers **systemic risks** which could arise from **general-purpose AI models,** including **large generative AI models**. These can be used for a variety of tasks and are becoming the basis for many AI systems in the EU. Some of these models could carry systemic risks if they are very capable or widely used. For example, powerful models could cause serious accidents or be misused for far-reaching cyberattacks. Many individuals could be affected if a model propagates harmful biases across many applications.

How do I know whether an AI system is high-risk?

The AI Act sets out a solid methodology for the classification of AI systems as high-risk. This aims to provide legal certainty for businesses and other operators.

The risk classification is based on the intended purpose of the AI system, in line with the existing EU product safety legislation. It means that the classification depends on the function performed by the AI system and on the specific purpose and modalities for which the system is used.

AI systems can classify as high-risk in two cases:

- If the AI system is embedded as a safety component in products covered by existing product legislation (Annex I) or constitute such products themselves. This could be, for example, AI-based medical software.
- If the AI system is intended to be used for a high-risk use case, listed in an Annex III to the AI Act. The list includes use cases from in areas such as education, employment, law enforcement or migration.

The Commission is preparing guidelines for the high-risk classification, which will be published ahead of the application date for these rules.

What are examples for high-risk use cases as defined in Annex III?

Annex III comprises eight areas in which the use of AI can be particularly sensitive and lists concrete use cases for each area. An AI system classifies as high-risk if it is intended to be used for one of these use cases.

Examples are:

- AI systems used as safety components in certain **critical infrastructures** for instance in the fields of road traffic and the supply of water, gas, heating and electricity;
- **AI systems used in education and vocational training,** e.g. to evaluate learning outcomes and steer the learning process and monitoring of cheating;
- **AI systems used in employment and workers management** and access to self-employment, e.g. to place targeted job advertisements, to analyse and filter job applications, and to evaluate candidates;
- **AI systems used in the access to essential private and public services** and benefits (e.g. healthcare), **creditworthiness evaluation** of natural persons, and risk assessment and pricing in relation to **life and health insurance**;

- AI systems used in the fields of **law enforcement**, migration and **border control**, insofar as not already prohibited, as well as in administration of **justice** and **democratic processes**;
- AI systems used for **biometric identification, biometric categorisation and emotion recognition,** insofar as not prohibited.

What are the obligations for providers of high-risk AI systems?

Before **placing a high-risk AI system on the EU market** or otherwise putting it into service, providers must subject it to a **conformity assessment**. This will allow them to demonstrate that their system complies with the mandatory requirements for trustworthy AI (e.g. data quality, documentation and traceability, transparency, human oversight, accuracy, cybersecurity and robustness). This assessment has to be repeated if the system or its purpose are substantially modified.

AI systems that serve as safety components of products covered by sectorial Union legislation will always be deemed high-risk when subject to third-party conformity assessment under that sectorial legislation. Moreover, all biometric systems, regardless of their application, will require third-party conformity assessment.

Providers of high-risk AI systems will also have to **implement quality and risk management systems** to ensure their compliance with the new requirements and minimise risks for users and affected persons, even after a product is placed on the market.

High-risk AI systems that are deployed by public authorities or entities acting on their behalf will have to be **registered in a public EU database**, unless those systems are used for law enforcement and migration. The latter will have to be registered in a non-public part of the database that will be only accessible to relevant supervisory authorities.

To ensure compliance throughout the lifecycle of the AI system, market surveillance authorities will conduct regular audits and facilitate post-market monitoring and will allow providers to voluntarily report any serious incidents or breaches of fundamental rights obligations that come to their attention. In exceptional cases, authorities may grant exemptions for specific high-risk AI systems to be placed on the market.

In case of a breach, the requirements will allow national authorities to have access to the information needed to investigate whether the use of the AI system complied with the law.

What would be the role of standardisation in the AI Act?

Under the AI Act, high-risk AI systems will be subject to specific requirements. European harmonised standards will play a key role in the implementation of these requirements.

In May 2023, the European Commission mandated the European standardisation organisations CEN and CENELEC to develop standards for these high-risk requirements. This mandate will now be amended, to align with the final text of the AI Act.

The European standardisation organisations will have until the end of April 2025 to develop and publish standards. The Commission will then evaluate and possibly endorse these standards, which will be published in the EU's Official Journal. Once published, those standards will grant a "presumption of conformity" to AI systems developed in accordance with them.

How are general-purpose AI models being regulated?

General-purpose AI models, including **large generative AI models**, can be used for a variety of tasks. Individual models may be integrated into a large number of AI systems.

It is crucial that a provider of an AI system integrating a general-purpose AI model has access to all necessary information to ensure the system is safe and compliant with the AI Act.

Therefore, the AI Act obliges providers of such models to **disclose certain information to downstream system providers.** Such **transparency** enables a better understanding of these models.

Model providers additionally need to have policies in place to ensure that that they **respect copyright law** when training their models.

In addition, some of these models could pose **systemic risks**, because they are very capable or widely used.

Currently, general purpose AI models that were trained using **a total computing power of more than 10^25 FLOPs** are considered to pose systemic risks. The Commission may update or

supplement this threshold in light of technological advances and may also designate other models as posing systemic risks based on further criteria (e.g. number of users, or the degree of autonomy of the model).

Providers of models with systemic risks are obliged to **assess and mitigate risks**, **report serious incidents**, **conduct state-of-the-art tests and model evaluations** and ensure **cybersecurity** of their models.

Providers are invited to collaborate with the AI Office and other stakeholders to develop a Code of Practice, detailing the rules and thereby ensuring the safe and responsible development of their models. This Code should represent a central tool for providers of general-purpose AI models to demonstrate compliance.

Why is 10^25 FLOPs an appropriate threshold for GPAI with systemic risks?

FLOP is a proxy for model capabilities, and the exact FLOP threshold can be updated upwards or downwards by the Commission, e.g. in the light of progress in objectively measuring model capabilities and of developments in the computing power needed for a given performance level.

The capabilities of the models above this threshold are not yet well enough understood. They could pose systemic risks, which is why it is reasonable to subject their providers to the additional set of obligations.

What are the obligations regarding watermarking and labelling of the AI outputs set out in the AI Act?

The AI Act sets transparency rules for the content produced by generative AI to address the risk of manipulation, deception and misinformation.

It obliges providers of generative AI systems to mark the AI outputs in a machine-readable format and ensure they are detectable as artificially generated or manipulated. The technical solutions must be effective, interoperable, robust and reliable as far as this is technically feasible, taking into account the specificities and limitations of various types of content, the costs of implementation and the generally acknowledged state of the art, as may be reflected in relevant technical standards.

In addition, deployers of generative AI systems that generate or manipulate image, audio or video content constituting deep fakes must visibly disclose that the content has been artificially generated or manipulated. Deployers of an AI system that generates or manipulates text published with the purpose of informing the public on matters of public interest must also disclose that the text has been artificially generated or manipulated. This obligation does not apply where the AI-generated content has undergone a process of human review or editorial control and where a natural or legal person holds editorial responsibility for the publication of the content.

The AI Office will issue guidelines to provide further guidance for providers and deployers on the obligations in Article 50 which will become applicable two years after entry into force of the AI Act (on 2 August 2026).

The AI Office will also encourage and facilitate the development of Codes of Practice at Union level to streamlining the effective implementation of the obligations related to the detection and labelling of artificially generated or manipulated content.

Is the AI Act future-proof?

The AI Act sets a legal framework that is responsive to new developments, easy and quick to adapt and allows for frequent evaluation.

The AI Act sets result-oriented requirements and obligations but leaves the concrete technical solutions and operationalisation to industry-driven standards and codes of practice that are flexible to be adapted to different use cases and to enable new technological solutions.

In addition, the legislation itself can be amended by delegated and implementing acts, for example to review the list of high-risk use cases in Annex III.

Finally, there will be frequent evaluations of certain parts of the AI Act and eventually of the entire regulation, making sure that any need for revision and amendments is identified.

How does the AI Act regulate biometric identification?

The use of **real-time remote biometric identification in publicly accessible spaces** (i.e. facial recognition using CCTV) for law enforcement purposes is prohibited. Member States can introduce exceptions by law that would allow the use of real-time remote biometric identification in the

following cases:

- Law enforcement activities related to 16 specified very serious crimes;
- Targeted search for specific victims, abduction, trafficking and sexual exploitation of human beings, and missing persons; or
- The prevention of threat to the life or physical safety of persons or response to the present or foreseeable threat of a terror attack.

Any exceptional use would be subject to **prior authorisation by a judicial or independent administrative authority** whose decision is binding. In case of urgency, approval can be granted within 24 hours; if the authorisation is rejected all data and output must be deleted.

It would need to be preceded by **prior fundamental rights impact assessment** and should be **notified to the relevant market surveillance authority and the data protection authority**. In case of urgency, the use of the system may be commenced without registration.

The use of AI systems for **post remote biometric identification** (identification of persons in previously collected material) of persons under investigation requires **prior authorisation** from a judicial authority or an independent administrative authority, as well as notification to the relevant data protection and market surveillance authority.

Why are particular rules needed for remote biometric identification?

Biometric identification can take different forms. Biometric authentication and verification i.e. to unlock a smartphone or for verification/authentication at border crossings to check a person's identity against his/her travel documents (one-to-one matching) remain unregulated, because they do not pose a significant risk to fundamental rights.

In contrast, biometric identification can also be used remotely for example to identify people in a crowd which can significantly impact privacy in the public space.

The accuracy of systems for facial recognition can be significantly influenced by a wide range of factors, such as camera quality, light, distance, database, algorithm, and the subject's ethnicity, age or gender. The same applies for gait and voice recognition and other biometric systems. Highly advanced systems are continuously reducing their false acceptance rates.

While a 99% accuracy rate may seem good in general, it is considerably risky when the result can lead to the suspicion of an innocent person. Even a 0.1% error rate can have a significant impact when applied to large populations, for example at train stations.

How do the rules protect fundamental rights?

There is already a strong protection for fundamental rights and for non-discrimination in place at EU and Member State level, but the complexity and opacity of certain AI applications (‘black boxes') can pose a problem.

A human-centric approach to AI means to ensure AI applications comply with fundamental rights legislation. By integrating accountability and transparency requirements into the development of high-risk AI systems, and improving enforcement capabilities, we can ensure that these systems are designed with legal compliance in mind right from the start. Where breaches occur, such requirements will allow national authorities to have access to the information needed to investigate whether the use of AI complied with EU law.

Moreover, the AI Act requires that certain deployers of high-risk AI systems conduct a fundamental rights impact assessment.

What is a fundamental rights impact assessment? Who has to conduct such an assessment, and when?

Providers of high-risk AI systems need to carry out a risk assessment and design the system in a way that risks to health, safety and fundamental rights are minimised.

However, certain risks to fundamental rights can only be fully identified knowing the context of use of the high-risk AI system. When high-risk AI systems are used in particularly sensitive areas of possible power asymmetry, additional considerations of such risks are necessary.

Therefore, deployers that are bodies governed by public law or private operators providing public services, as well as operators providing high-risk AI systems that carry out credit worthiness assessments or price and risk assessments in life and health insurance, shall perform an assessment of the impact on fundamental rights and notify the national authority of the results.

In practice, many deployers will also have to carry out a data protection impact assessment. To avoid substantive overlaps in such cases, the fundamental rights impact assessment shall be conducted in conjunction with that data protection impact assessment.

How does this regulation address racial and gender bias in AI?

It is very important to underline that AI systems **do not create or reproduce bias**. Rather, when properly designed and used, **AI systems can contribute to reducing bias and existing structural discrimination**, and thus lead to more equitable and non-discriminatory decisions (e.g. in recruitment).

The **new mandatory requirements for all high-risk AI systems will serve this purpose**. AI systems must be **technically robust** to ensure they are fit for purpose and do not produce biased results, such as false positives or negatives, that disproportionately affect marginalised groups, including those based on racial or ethnic origin, sex, age, and other protected characteristics.

High-risk systems will also need to be **trained and tested with sufficiently representative datasets** to **minimise the risk of unfair biases** embedded in the model and ensure that these can be addressed through appropriate bias detection, correction and other mitigating measures.

They must also be **traceable and auditable**, ensuring that appropriate **documentation is kept**, including the data used to train the algorithm that would be key in ex-post investigations.

Compliance system before and after they are placed on the market will have to ensure these systems are **regularly monitored** and **potential risks are promptly addressed**.

When will the AI Act be fully applicable?

The AI Act will apply two years after entry into force on 2 August 2026, except for the following specific provisions:

- The prohibitions, definitions and provisions related to AI literacy will apply 6 months after entry into force on 2 February 2025;
- The rules on governance and the obligations for general purpose AI become applicable 12 months after entry into force on 2 August 2025;
- The obligations for high-risk AI systems that classify as high-risk because they are embedded in regulated products, listed in Annex II (list of Union harmonisation legislation), apply 36 months after entry into force on 2 August 2027.

How will the AI Act be enforced?

The AI Act establishes a two-tiered governance system, where **national authorities** are responsible for overseeing and enforcing rules for AI systems, while the EU level is responsible for governing general-purpose AI models.

To ensure EU-wide coherence and cooperation, the **European Artificial Intelligence Board** (AI Board) will be established, comprising representatives from Member States, with specialised subgroups for national regulators and other competent authorities.

The **AI Office**, the Commission's implementing body for the AI Act, will provide strategic guidance to the AI Board.

In addition, the AI Act establishes two advisory bodies to provide expert input: the **Scientific Panel** and the **Advisory Forum**. These bodies will offer valuable insights from stakeholders and interdisciplinary scientific communities, informing decision-making and ensuring a balanced approach to AI development.

Why is a European Artificial Intelligence Board needed and what will it do?

The European Artificial Intelligence Board comprises **high-level representatives of Member States** and the European Data Protection Supervisor. As a key advisor, the AI Board provides guidance on all matters related to AI policy, notably AI regulation, innovation and excellence policy and international cooperation on AI.

The AI Board plays a crucial role in ensuring the smooth, effective and harmonised implementation of the AI Act. The Board will serve as the forum where the AI regulators, namely the AI Office, national authorities and EPDS, can coordinate the consistent application of the AI Act.

What are the penalties for infringement?

Member States will have to lay down effective, proportionate and dissuasive penalties for

infringements of the rules for AI systems.

The Regulation sets out thresholds that need to be taken into account:

- **Up to €35m or 7%** of the total worldwide annual turnover of the preceding financial year (whichever is higher) for infringements **on prohibited practices or non-compliance** related to requirements on data;
- **Up to €15m or 3%** of the total worldwide annual turnover of the preceding financial year for **non-compliance with any of the other requirements** or obligations of the Regulation;
- **Up to €7.5m or 1.5%** of the total worldwide annual turnover of the preceding financial year for the **supply of incorrect, incomplete or misleading information** to notified bodies and national competent authorities in reply to a request;
- For each category of infringement, the threshold would be the lower of the two amounts for SMEs and the higher for other companies.

The Commission can also enforce the rules on providers of general-purpose AI models by means of fines, taking into account the following threshold:

- **Up to €15m or 3%** of the total worldwide annual turnover of the preceding financial year for **non-compliance with any of the obligations** or measures requested by the Commission under the Regulation.

EU institutions, agencies or bodies are expected to lead by example, which is why they will also be subject to the rules and to possible penalties. The European Data Protection Supervisor will have the power to impose fines on them in case of non-compliance.

How will the General-Purpose AI Code of Practice be written?

The drawing-up of the first Code follows an inclusive and transparent process. A Code of Practice Plenary will be established to facilitate the iterative drafting process, consisting of all interested and eligible general-purpose AI model providers, downstream providers integrating a general-purpose AI model into their AI system, other industry organisations, other stakeholder organisations such as civil society or rightsholders organisations, as well as academia and other independent experts.

The AI Office has launched a call for expression of interest to participate in the drawing-up of the first Code of Practice. In parallel to this call for expression of interest a multi-stakeholder consultation to collect views and inputs from all interested stakeholders on the first Code of Practice is launched. Answers and submissions will form the basis of the first drafting iteration of the Code of Practice. From the start, the Code is therefore informed by a broad array of perspectives and expertise.

The Plenary will be structured in four Working Groups to allow for focused discussions on specific topics relevant to detail out obligations for providers of general-purpose AI models and general-purpose AI models with systemic risk. Plenary participants are free to choose one or more Working Groups they wish to engage in. Meetings are conducted exclusively online.

The AI Office will appoint Chairs and, as appropriate, Vice-Chairs for each of the four Working Groups of the Plenary, selected from interested independent experts. The Chairs will synthesise submissions and comments by Plenary participants to iteratively draft the first Code of Practice.

As the main addressees of the Code, providers of general-purpose AI models will be invited to dedicated workshops to contribute to informing each iterative drafting round, in addition to their Plenary participation.

After 9 months, the final version of the first Code of practice will be presented in a closing Plenary, expected to take place in April, and published. The closing Plenary gives general-purpose AI model providers the opportunity to express themselves whether they would envisage to use the Code.

If approved, how does the Code of Practice for general-purpose AI model providers serve as a central tool for compliance?

At the end of the Code of Practice drafting process, the AI Office and the AI Board will assess the adequacy of the Code and will publish their assessment. Following that assessment, the Commission may decide to approve a Code of Practice and give it general validity within the Union by means of implementing acts. If by the time the Regulation becomes applicable, the Code of Practice is not deemed adequate by the AI Office, the Commission may provide common rules for the implementation of the relevant obligations.

Providers of general-purpose AI models may therefore rely on the Code of Practice to demonstrate

compliance with the obligations set out in the AI Act.

As per the AI Act, the Code of Practice should include objectives, measures and as appropriate, key performance indicators (KPIs).

Providers adhering to the Code should regularly report to the AI Office on the implementation of measures taken and their outcomes, including as measured against key performance indicators as appropriate.

This facilitates enforcement by the AI Office, which is underpinned by the powers given to the Commission by the AI Act. This includes the ability to conduct evaluations of general-purpose AI models, request information and measures from model providers, and apply sanctions.

The AI Office will, as appropriate, encourage and facilitate the review and adaptation of the Code to reflect advancements in technology and state-of-the-art.

Once a harmonised standard is published and assessed as suitable to cover the relevant obligations by the AI Office, compliance with a European harmonised standard should grant providers the presumption of conformity.

Providers of general-purpose AI models should furthermore be able to demonstrate compliance using alternative adequate means, if Codes of Practice or harmonised standards are not available, or they choose not to rely on those.

Does the AI Act contain provisions regarding environmental protection and sustainability?

The objective of the AI proposal is to address risks to safety and fundamental rights, including the fundamental right to a high-level environmental protection. The environment is also one of the explicitly mentioned and protected legal interests.

The Commission is asked to request European standardisation organisations to produce a standardisation deliverable on reporting and documentation processes to improve AI systems' resource performance, such as reduction of energy and other resources consumption of the high-risk AI system during its lifecycle, and on energy efficient development of general-purpose AI models.

Furthermore, the Commission by two years after the date of application of the Regulation and every four years thereafter, is asked to submit a report on the review of the progress on the development of standardisation deliverables on energy efficient development of general-purpose models and asses the need for further measures or actions, including binding measures or actions.

In addition, providers of general-purpose AI models, which are trained on large data amounts and therefore prone to high energy consumption, are required to disclose energy consumption. In case of general-purpose AI models with systemic risks, energy efficiency furthermore needs to be assessed.

The Commission is empowered to develop appropriate and comparable measurement methodology for these disclosure obligations.

How can the new rules support innovation?

The regulatory framework can enhance the uptake of AI in two ways. On the one hand, increasing users' trust will increase the demand for AI used by companies and public authorities. On the other hand, by increasing legal certainty and harmonising rules, AI providers will access bigger markets, with products that users and consumers appreciate and purchase. Rules will apply only where strictly needed and in a way that minimises the burden for economic operators, with a light governance structure.

The AI Act further enables the creation of **regulatory sandboxes** and **real-world testing,** which provide a controlled environment to test innovative technologies for a limited time, thereby fostering innovation by companies, SMEs and start-ups in compliance with the AI Act. These, together with other measures such as the additional **Networks of AI Excellence Centres** and the **Public-Private Partnership on Artificial Intelligence, Data and Robotics,** and access to **Digital Innovation Hubs** and **Testing and Experimentation Facilities** will help build the right framework conditions for companies to develop and deploy AI.

Real world testing of High-Risk AI systems can be conducted for a maximum of 6 months (which can be prolonged by another 6 months). Prior to testing, a plan needs to be drawn up and submitted to the market surveillance authority, which has to approve the plan and specific testing conditions, with default tacit approval if no answer has been given within 30 days. Testing may be subject to unannounced inspections by the authority.

Real world testing can only be conducted given specific safeguards, e.g. users of the systems under real world testing have to provide informed consent, the testing must not have any negative effect on them, outcomes need to be reversible or disregardable, and their data needs to be deleted after conclusion of the testing. Special protection is to be granted to vulnerable groups, i.e. due to their age, physical or mental disability.

What role does the AI Pact play in the implementation of the AI Act?

Initiated by Commissioner Breton in May 2023, the AI Pact aims to enhance engagement between the AI Office and organisations (Pillar I) and to encourage the industry's voluntary commitment to start implementing the AI Act's requirements ahead of the legal deadline (Pillar II).

In particular, under Pillar I, participants will contribute to the creation of a collaborative community, sharing their experiences and knowledge. This includes workshops organised by the AI Office which provide participants with a better understanding of the AI Act, their responsibilities and how to prepare for its implementation. In turn, the AI Office can gather insights into best practices and challenges faced by the participants.

Under Pillar II, organisations are encouraged to proactively disclose the processes and practices they are implementing to anticipate compliance, through voluntary pledges. Pledges are intended as 'declarations of engagement' and will contain actions (planned or underway) to meet some of the AI Act's requirements.

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.

More than 700 organisations have already expressed their interest in joining the AI Pact initiative, further to a call launched in November 2023. A first information session was held online on 6 May, with 300 participants. The official signing of the voluntary commitments is planned for autumn 2024. There will be a workshop on the AI Pact in the first week of September.

What is the international dimension of the EU's approach?

AI has consequences and challenges that transcend borders; therefore international cooperation is important. The AI Office is in charge of the European Union international engagement in the area of AI, on the basis of the AI Act and the Coordinated Plan on AI. The EU seeks to promote the responsible stewardship and good governance of AI in collaboration with international partners and in line with the rules-based multilateral system and the values it upholds.

The EU engages bilaterally and multilaterally to promote trustworthy, human-centric and ethical AI. Consequently, the EU is involved in multilateral forums where AI is discussed – notably G7, G20, the OECD, the Council of Europe, the Global Partnership on AI and the United Nations – and the EU has close bilateral ties with e.g. Canada, the US, India, Japan, South Korea, Singapore, and the Latin American and Caribbean region.

**Updated on 01/08/2024*

QANDA/21/1683

Press contacts:

Thomas Regnier (+32 2 29 9 1099)
Patricia Poropat (+32 2 298 04 85)

General public inquiries: Europe Direct by phone 00 800 67 89 10 11 or by email