onetrust

FUTURE OF PRIVACY FORUM

Conformity Assessments Under the proposed EU AI Act: A Step-By-Step Guide

Authors: Katerina Demetzou, Vasileios Rovilos

Editors: Gabriela Zanfir-Fortuna, Rob van Eijk, Andrew Clearwater, Alexis Kateifides

Copyeditor: Alexander Thompson

E-BOOK | NOVEMBER 2023

Table of Contents

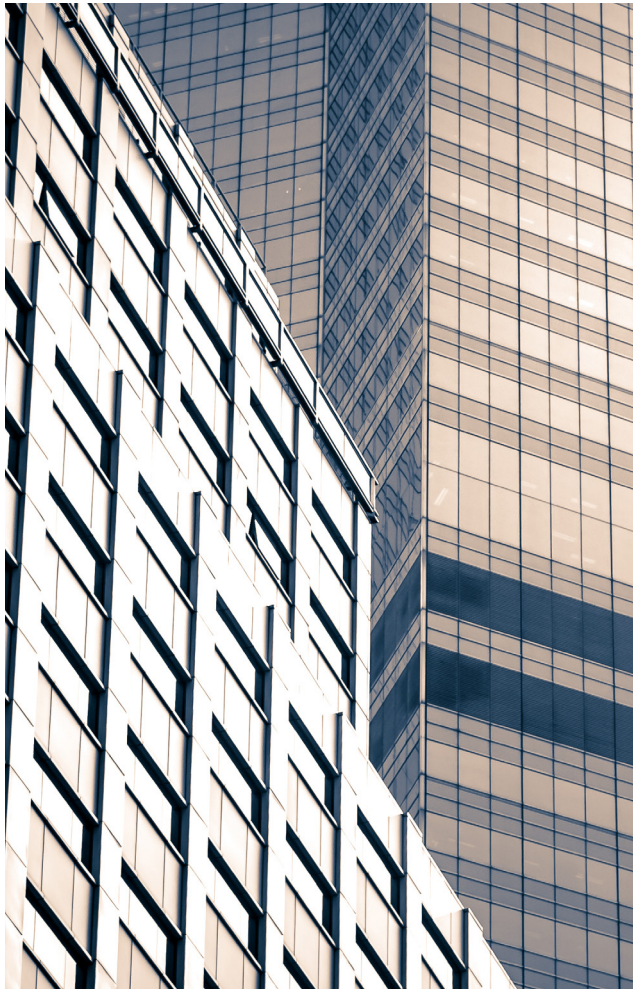
I. Introduction to the Guide

This Guide explains what a Conformity Assessment (CA) is under the proposed EU Artificial Intelligence Act (EU AIA or AIA) and provides a roadmap to execute one, understanding that nothing in this Guide amounts to legal advice. CAs are a key and overarching accountability tool introduced by the AIA for high-risk AI systems and they are expected to play a significant role in the governance of AI in the EU.

This Guide examines the CA as set out under the proposed EU AIA. It does not offer a comparative study with other existing assessment processes required under other European legal acts. The EU AIA includes various documentation obligations. We will refer to those obligations only where necessary, to either highlight their differences from the CA or explain where those other documentation obligations play a role in the performance of the CA.

At the time of writing of this Guide, the EU AIA has still not reached a final agreement. That being said, this Guide lays down the requirements as set out in the official positions of the two co-legislators: the Council of the EU and the European Parliament (for further info, see Section II.A). We hope the Guide serves as an essential resource for those who want to prepare for compliance with the EU AIA. We will update the Guide upon adoption of the final text of the AIA.

In Section II, we provide a brief description of the legislative process that the EU AIA has followed, explaining the current stage and the necessary steps to reach the final adoption of the Regulation. Section II also provides a high-level description of the EU AIA. Section III details the purpose, structure, and function of the CA obligation. Firstly, it identifies the questions that must be answered in order for an actor to assess whether they fall under the obligation to conduct a CA. Section III also explains when and how a CA should be performed and elaborates on all the requirements that need to be met during the CA process. Lastly, Section IV discusses the role of standards and the presumption of compliance with the requirements offered through adherence to harmonized standards.



II. The EU AIA

II.A. The EU AIA is in the final stages of the legislative process

The European Commission is the main European institution that initiates legislation in the EU. In April 2021, it published the legislative proposal for a Regulation laying down harmonized rules on AI, the proposed EU AIA (COM(2021))206). Once adopted as a Regulation, the EU AIA will be of general and direct applicability to all EU Member States and it will be legally binding in its entirety.

The Council of the EU is the EU Body that, inter alia, negotiates and adopts laws. It adopted its General Approach on the EU AIA in December 2022. The European Parliament has the power to adopt and amend legislative proposals. In December 2021, it appointed a Joint Committee to lead the work on the AIA, the Internal Market & Consumer Protection Committee (IMCO), whose Rapporteur is Brando Benifei (S&D Italy), and the Civil Liberties Justice & Home Affairs Committee (LIBE), whose Rapporteur is Dragos Tudorache (Renew Romania). After one and a half years of work, in June 2023, the Parliament adopted its official position on the EU AIA.

At the time of writing this Guide, the two co-legislators, the Council of the EU and the European Parliament, have entered the 'trialogue negotiations' stage, during which they negotiate their respective positions on the Regulation in order to reach a final agreement on the text. The final text is expected to be adopted by the end of 2023 and it is expected to become applicable in late 2025.

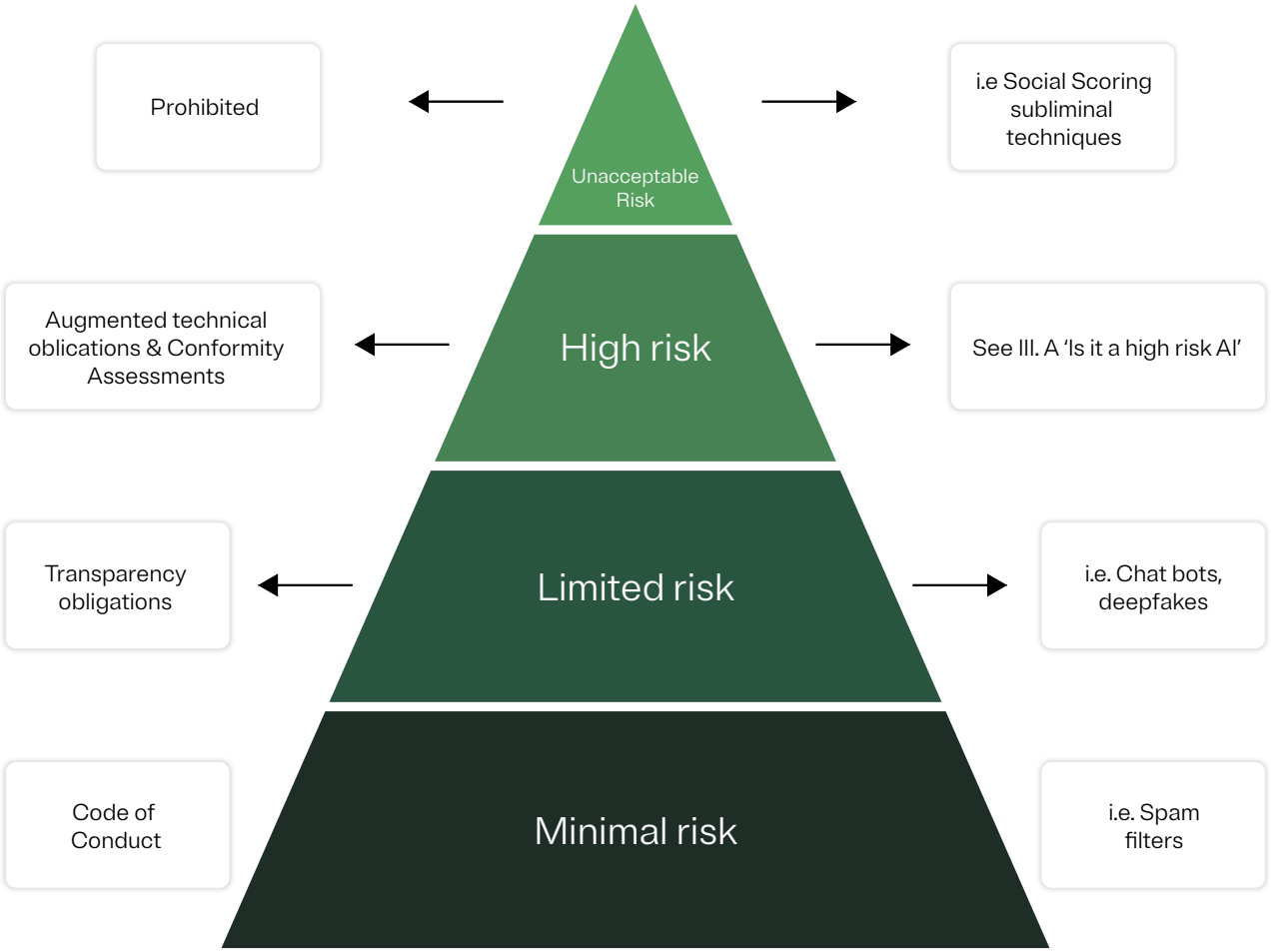
Note: the Guide was created at the time of the trilogue negotiations, with the goal of supporting organizations who have started their road to compliance with the EU AIA.

II. The EU AIA

II.B The EU AIA is a risk-based regulation with enhanced obligations for high-risk AI systems

The EU AIA is structured on the basis of a **precautionary and risk-based approach.**

The EU AIA aims to regulate AI technologies on the basis of the risks that their use is likely to raise to the health, safety, and fundamental rights of individuals. The EU AIA prohibits certain uses of AI systems that raise unacceptable risks (precautionary approach), and it sets rules on the development and deployment of AI systems depending on whether they qualify as high, low, or minimum risk (risk-based approach). **The CA obligation only applies to high-risk AI systems.** The level of risk and whether an AI system qualifies as "high risk" is discussed under subsection Q2: Classification of the AI system as "high-risk."



II. The EU AIA

The EU AIA was initially conceived as a "safety product legislation."

The Regulation aims to align with the processes and requirements found in laws that fall under the New Legislative Framework (NLF) in order to "minimize the burden on operators and avoid any possible duplication" (Recital 63 AIA)[1]. In the EU context, the CA obligation is not new. CAs are also part of several EU laws on product safety, such as the General Product Safety Regulation (GPSR)[2], the Machinery Regulation[3], or the in vitro diagnostic Medical Devices Regulation[4]. **It could thus be the case that an AI system is a safety component of a product that falls under the scope of NLF laws, for which a different CA may have already been performed.** This is relevant for the AI system provider to have in mind when looking at their CA legal obligation under the EU AIA.

The EU AIA will apply without prejudice to other laws.

The application of the EU AIA is intended to be without prejudice to other laws and it is laid down consistently with the Regulations that are explicitly mentioned in the Preamble. One prominent example is the General Data Protection Regulation (GDPR). While the EU AIA and the GDPR have different material scopes, it could very well be the case that both Regulations are applicable to the same processing of personal data that underpins or is the result of an AI system. In that case, legal obligations from both laws must be met. For example, take the data controller's obligation to perform a Data Protection Impact Assessment (DPIA) under Article 35 GDPR. **If the AI system's provider qualifies as a data controller and if the relevant legal conditions are met, the provider will be obliged to perform both a DPIA and a CA.**

[1] An example of this can be found under the Risk Management System requirement, under sub section 4.1.

[2] REGULATION (EU) 2023/988 of the European Parliament and of the Council of 10 May 2023, on general product safety, amending Regulation (EU) No 1025/2012 of the European Parliament and of the Council and Directive (EU) 2020/1828 of the European Parliament and the Council, and repealing Directive 2001/95/EC of the European Parliament and of the Council and Council Directive 87/357/EEC.

[3] REGULATION (EU) 2023/1230 of the European Parliament and of the Council of 14 June 2023, on machinery and repealing Directive 2006/42/EC of the European Parliament and of the Council and Council Directive 73/361/EEC.

[4] REGULATION (EU) 2017/746 of the European Parliament and of the Council of 5 April 2017 on in vitro diagnostic medical devices and repealing Directive 98/79/EC and Commission Decision 2010/227/EU.

III. The Conformity Assessment obligation includes an overarching accountability framework for high-risk AI systems

"Conformity Assessment" is defined under Article 3 AIA as the process of verifying and/or demonstrating[5] that a high-risk AI system complies with the requirements enumerated under Title III, Chapter 2 of the Act. These requirements are:

1. Risk management system;

2. Data governance;

3. Technical documentation;

4. Record-keeping;

5. Transparency and provision of information;

6. Human oversight;

7. Accuracy, robustness, and cybersecurity.

Each of these requirements will be further elaborated under Step 4.

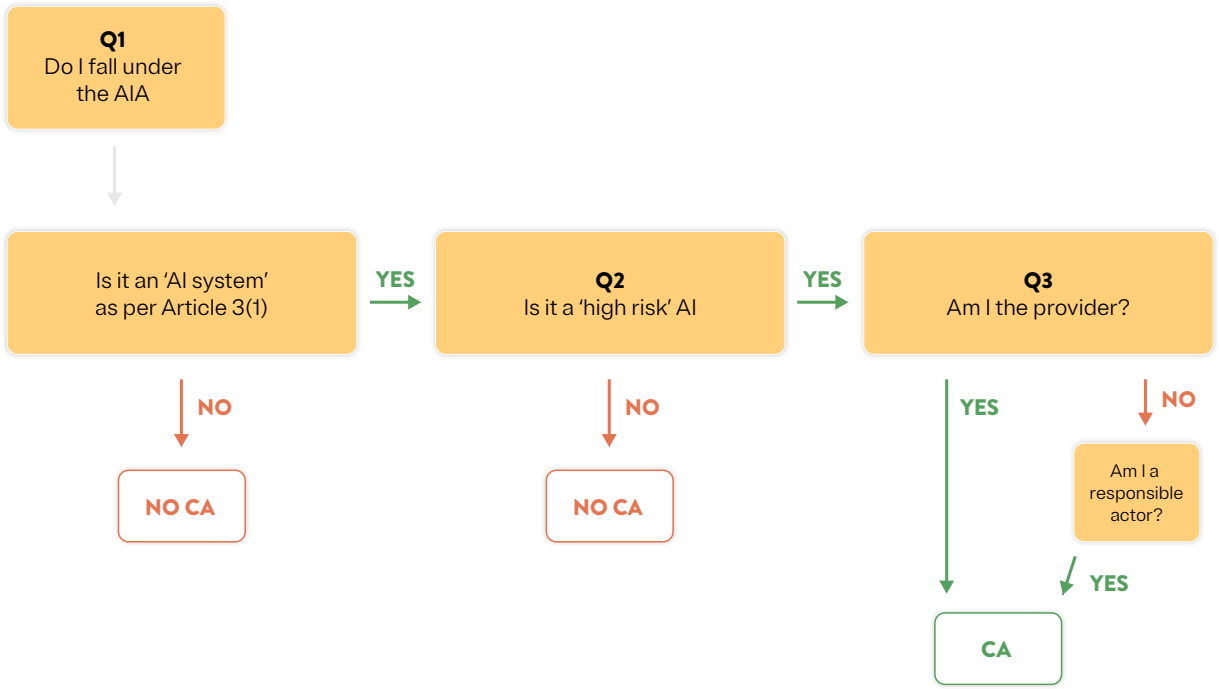
The CA is a process that consists of various assessments, such as the assessment of whether the AI system qualifies as high-risk[6], and the assessment of risks that is part of the risk management system. The CA process additionally consists of requirements that need to be built-in the high-risk AI system (e.g. automatic recording of events, human oversight capacity, transparent operation of the AI system) as well as documentation obligations (e.g. technical documentation). The CA should be understood as a framework of assessments, (technical and non-technical) requirements and documentation obligations.

[5] The Council's version reads "verifying," while the Parliament's version reads "demonstrating."

[6] Whether an AI system qualifies as high-risk and thus a CA needs to be conducted, is an assessment prior to the CA process, albeit necessary in order for a provider to identify whether they are bound by the CA legal obligation.

Step 1: Am I obliged to perform a CA?

The first step in the CA journey is to determine whether an organization falls under the CA legal obligation. Follow the flowchart below, which provides questions that an organization should answer in order to determine whether they need to comply with the CA obligation.



Q1: Do I fall under the AIA?

As a first step, one should identify whether they fall under the material scope of the AIA. The scope is delineated under Article 2(1) AIA. Article 2(4) AIA enumerates the cases that are not covered by the AIA. For a system to fall under the AIA, it should qualify as an "AI system" as defined under Article 3(1) AIA.

AI System definition

The definition of an AI system aims to be as technology-neutral and future-proof as possible, while at the same time aiming to capture all AI systems that are likely to pose risks. It is noteworthy that there is no consensus among the three EU institutions as to a single definition. The Council of the EU and the European Parliament have introduced amendments to better align it with the definition proposed by the OECD[7]. More specifically, their text reads:

Council of the EU: "AI system means a system that is designed to operate with elements of autonomy and that, based on machine and/or human-provided data and inputs, infers how to achieve a given set of objectives using machine learning and/or logic- and knowledge-based approaches, and produces system-generated

[7] It is valuable to note that OECD is currently working on renewing these definitions and it is not entirely clear whether they will be adjusted to the AI Act in time - AI-Principles Overview - OECD.AI

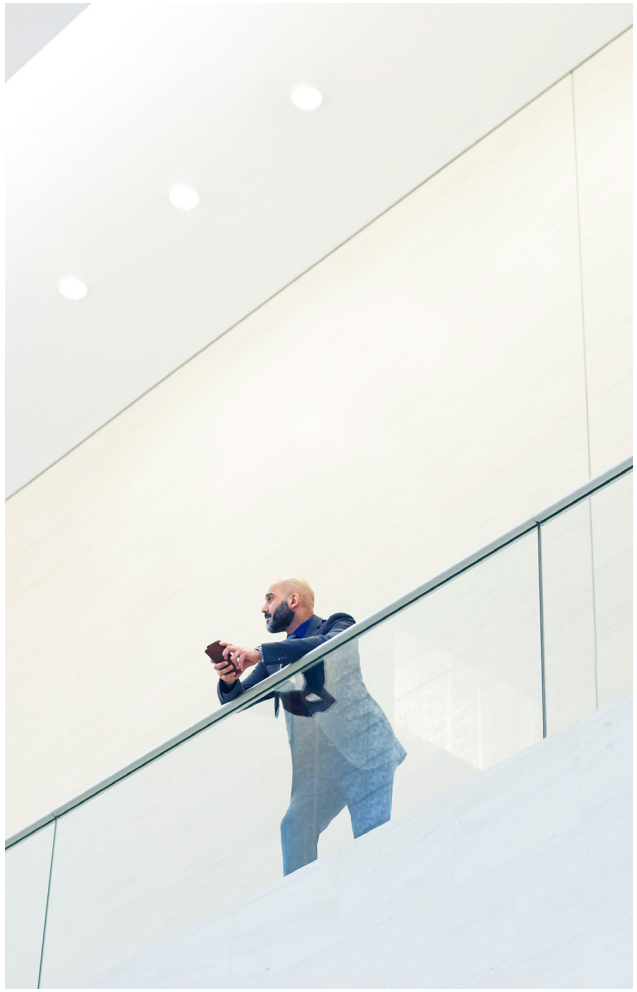
Step 1: Am I obliged to perform a CA?

outputs such as content (generative AI systems), predictions, recommendations, or decisions, influencing the environments with which the AI system interacts" (Article 3(1)).

European Parliament: "AI system means a machine-based system that is designed to operate with varying levels of autonomy and that can, for explicit or implicit objectives, generate outputs such as predictions, recommendations, or decisions, that influence physical or virtual environments" (Article 3(1)). In Recital 6, the EP clarifies that simple software systems or programming approaches that do not present key characteristics of AI such as its learning, reasoning, or modeling capabilities, should not fall under the material scope of the AIA.

Following recent trilogue negotiations, we note that both co-legislators endorse the OECD definition of an AI system. However, the Council's approach is narrower than the one proposed by the European Parliament. The Council differentiates between AI systems from simpler software and programming systems that automatically execute operations. This approach revolves around how outputs are generated ("infers how to achieve a given set of objectives") by machine learning and/or logic-and-knowledge-based approaches.

On the other hand, the European Parliament favors a much wider and more inclusive approach (also broader than the European Commission's proposal), despite carving out simpler forms of software systems. This approach reflects what output is generated. The co-legislators are in ongoing negotiations to iron out this issue.



Step 1: Am I obliged to perform a CA?

Q2: Is it a high-risk AI system?

After having ascertained that the system falls under the AIA, one needs to determine if the AI system qualifies as "high-risk."[8] As already mentioned, the CA obligation only applies to high-risk AI systems. Table 1 presents the classification rules as set under Title III, Chapter 1 AIA:

Table 1 Classification of High-risk AI systems under the AIA
1st Category
AI systems that are safety components of products or are themselves products that fall under Annex II

Safety component of a product covered by the Union harmonization legislation listed in Annex II	**Annex II legislation covers:** - Machinery; - Safety of toys; - Recreational craft and personal watercraft; - Lifts and safety components of lifts; - Equipment and protective systems for use in explosive atmospheres; - Market of radio equipment; - Marker of pressure equipment; - Cableway installations; - Personal protective equipment; - Appliances burning gaseous fuels; - Medical devices (and vitro diagnostic medical devices); - Civil aviation security; - Vehicles; - Marine equipment; - Interoperability of the rail system;
The product is required to undergo a **third-party conformity assessment** pursuant to that legislation * Irrespective of whether an AI system is placed on the market or put into service independently from the product.	
A **product** covered by the Union harmonization legislation listed in Annex II The product is required to undergo a **third-party conformity assessment** pursuant to that legislation * Irrespective of whether an AI system is placed on the market or put into service independently from the product.	

[8] The reader should know that the classification rules is one of the most contentious elements currently corroborated between the co-legislators.

Step 1: Am I obliged to perform a CA?

There are three points that should be made regarding the AI systems that belong to the 2nd Category, i.e. AI Systems that belong to the use cases of Annex III AIA:

1. In its version, the European Parliament (Article 6(2a) AIA) allows the provider to argue that although the AI system falls under one or more Annex III use cases, it does not pose a significant risk to the health, safety, or fundamental rights of natural persons. The provider may argue this through a reasoned notification that should be submitted to the national supervisory authority. If the argument holds true, the system is not high-risk and thus no CA is required.

2. Under Article 7 AIA, the European Commission has the power to amend Annex III list through delegated acts, subject to specific legal conditions.

3. In their respective Article 7(2) AIA, both co-legislators enumerate criteria that should lead the European Commission's assessment of the risks that an AI system poses. Those criteria could also prove useful to providers who wish to argue that their AI system does not pose significant risks to the health, safety, or fundamental rights of natural persons.

<table>
<tr><th colspan="2">2nd Category
AI Systems that belong to the use cases of Annex III</th></tr>
<tr><td>AI system that falls under one or more of the 8 critical areas and use cases referred to in Annex III</td><td rowspan="3">Annex III categories of purposes:
1. Biometrics (and biometrics-based systems);
2. (Management and operations of) Critical infrastructure;
3. Education and vocational training;
4. Employment, workers management, and access to self-employment;
5. Access to and enjoyment of essential private services and public services and benefits;
6. Law enforcement;
7. Migration, asylum, and border control management;
8. Administration of justice and democratic processes</td></tr>
<tr><td>Council

The output of the AI system is not purely accessory[9] in respect of the relevant action or decision to be taken

+

is likely to lead to a significant risk to the health, safety, or fundamental rights</td></tr>
<tr><td>EP

if they pose a significant risk of harm[10] to the health, safety, or fundamental rights of natural persons</td></tr>
</table>

[9] Article 6(3) Council version: "'n order to ensure uniform conditions for the implementation of this Regulation, the Commission shall, no later than one year after the entry into force of this Regulation, adopt implementing acts to specify the circumstances where the output of AI systems referred to in Annex III would be purely accessory (...)"

[10] Article 6(2) EP version: "The Commission shall, six months prior to the entry into force of this Regulation, after consulting the AI Office and relevant stakeholders, provide guidelines clearly specifying the circumstances where the output of AI systems referred to in Annex III would pose a significant risk of harm to the health, safety or fundamental rights of natural persons or cases in which it would not.

Step 1: Am I obliged to perform a CA?

Q3: Am I the responsible actor?

After having classified the AI system as "high-risk," the next question to be answered is whether the actor is the one responsible for performing the CA. The provider is the primary responsible actor for conducting a CA. However, in exceptional circumstances, the obligation might fall on another actor.

The rule is that the actor responsible for performing the CA is the provider of the high-risk AI system (Article 16). The provider is defined as "a natural or legal person, public authority, agency or other body that develops an AI system or that has an AI system developed with a view to placing it on the market or putting it into service under its own name or trademark, whether for payment or free of charge" (Article 3(e))[11]. Even if the provider is not the designer/developer of the system, they still need to make sure that requirements are embedded in the system prior to placing the system on the market or putting it into service.

It might be exceptionally the case that the CA must be performed by the product manufacturer[12], the distributor[13], or the importer[14] of a high-risk AI system, the deployer[15], or a third party.

The exact legal conditions that should be met in order for an actor other than the provider to be obliged to perform the CA are yet to be set. However, the distributor, importer, deployer, or any other third party would, as a rule, be obliged to conduct a CA if they put their name or trademark on a high-risk AI system already placed on the market or put into service, or if they make a substantial modification to a high-risk AI system.

As to the product manufacturer[16], it seems to be straightforward that they will be responsible for a CA if, cumulatively:

- the high-risk AI system relates to products for which the laws in Annex II Section A apply;
- the system is placed on the market or put into service together with the product; AND
- under the name or trademark of the product manufacturer.

[11] In its official position the Councilt reads somewhat differently "develops an AI system or that has an AI system developed and places that system on the market or puts it into service."

[12] "Product manufacturer" is defined only under the Council AIA version. According to Article 3(5a), "product manufacturer" means a "manufacturer within the meaning of any of the Union harmonization legislation listed in Annex II." However, in their respective Recital 55, both the European Parliament and the Council, refer to the "relevant New Legislative Framework legislation" for the definition of "product manufacturer."

[13] According to Article 3(7) AIA, "distributor" means "any natural or legal person in the supply chain, other than the provider or the importer, that makes an AI system available on the Union market without affecting its properties." In its version, the Council deleted the phrase "without affecting its properties."

[14] According to Article 3(6) AIA, "importer" means "any natural or legal person established in the Union that places on the market or puts into service an AI system that bears the name or trademark of a natural or legal person established outside the Union."

[15] It is only the EP that places the CA responsibility on the deployer's shoulders if the relevant legal conditions are met.

[16] Article 23a Council version - Article 24 EP version.

Step 2: When should a CA be conducted?

Once the actor has determined that they are indeed legally obliged to conduct a CA, they need to make sure to do so along the correct timeline.

- **Ex ante:** A CA has to be performed prior to placing an AI system on the EU market, which means prior to making it available (i.e., supplying for distribution or use), or prior to putting an AI system into service, which means prior to its first use in the EU market, either by the system's user or for (the provider's) own use.

- **Ex post:** After the high-risk AI system has been placed on the market or put into service, a new CA will be required in case the AI system is substantially modified. Substantial modification is considered any change that affects a system's compliance with the requirements for high-risk AI systems or results in a modification to the AI system's intended purpose[17].

However, there is no need for a new CA when a high-risk AI system continues to learn after being placed on the market or put into service, as long as these changes are pre-determined at the moment of the initial CA and are described in the initial technical documentation[18].

[17] A final and common approach as to when a change in the high-risk AI system qualifies as "substantial modification" and when it does not, still needs to be carved out. In any case, a "substantial modification" leads to a "new" AI system, for which a new CA has to be conducted.

[18] Inclusion of the possible changes in the high-risk AI system in the technical documentation (see Step 4.3) is a legal requirement for the change to qualify as "non substantial modification". It is an additional obligation of the provider to give, in the context of the transparency requirements (see Step 4.5), information about, inter alia, the "characteristics, capabilities, and limitations of performance of the high-risk AI system." Part of this information shall be any "predetermined changes to the performance of the system."

Step 3: Who should conduct the CA?

There are two ways in which a CA can be conducted – internally, or by a third party. In a nutshell, in the **internal CA** process, the provider (or any other responsible actor[19]) performs the CA. The **third-party CA** is performed by an external "notified body." In Article 43, the AIA mentions explicitly which cases require an internal control CA and which ones should go through the third-party CA process. Table 2 below presents which type of CA process should be followed in each case of a high-risk AI system.

Internal Conformity Assessment (Annex VI AIA)

For a CA procedure based on internal control, both co-legislators agree on the steps that the provider should take, as described in Annex VI of the AIA:

1. The provider verifies that the established quality management system is in compliance with the requirements of Article 17.
2. The provider examines the information contained in the technical documentation in order to assess the compliance of the AI system with the relevant essential requirements set out in Title III, Chapter 2.
3. The provider verifies that the design and development process of the AI system and its post-market monitoring, as referred to in Article 61, is consistent with the technical documentation.

Quality management system (Article 17 AIA)

The provider's obligation to have a Quality Management System (hereafter QMS) in place is found under Article 17 AIA[20]. The QMS shall be documented in a systematic and orderly manner in the form of written policies, procedures, and instructions. Article 17 AIA enumerates the elements that should be included in the QMS, inter alia: a strategy for regulatory compliance, systems, and procedures for data management, the risk management system, the set up, implementation, and maintenance of a post-market monitoring system, policies for communicating with supervisory authorities, means to ensure compliance with the essential requirements, etc.

After having concluded the internal CA, the provider must draw up the so-called "EU declaration of conformity" (Article 48 AIA) (hereafter, the declaration). This declaration shall be kept at the "disposal of the national competent authorities for 10 years after the AI system has been placed on the market or put into service." **Annex V** of the AIA specifies the information that should be included in the declaration, including a statement of compliance with the GDPR if the AI system processes personal data. This is relevant, given that the performance of a DPIA (if the legal conditions are met) is likely to be part of this declaration.

Additionally, pursuant to Article 49 AIA, the provider or other responsible entity has to affix a visible, legible, and indelible CE marking of conformity. Said **CE marking of conformity** would be subject to the general principles set out in Article 30 of Regulation (EC) No 765/2008.

Third-party conformity assessment (Annex VII AIA)

Alternatively, in the case of a third-party CA, it is not the provider but a designated independent notified body that assesses the quality management system and the technical documentation of the high-risk AI

[19] See Q3: Am I the responsible actor?

[20] European Parliament's version, Recital 54 "For providers that have already in place quality management systems based on standards such as ISO 9001 or other relevant standards, no duplicative quality management system in full should be expected but rather an adaptation of their existing systems to certain aspects linked to compliance with specific requirements of this Regulation."

Step 3: Who should conduct the CA?

system, according to the process explained in Annex VII of the AIA.

These 'notified bodies' are conformity assessment bodies, i.e. bodies that perform third-party CA. A conformity assessment body may submit an application for notification to the notifying authority of the Member State in which it is established. Notifying authorities[21] may only notify conformity assessment bodies that satisfy the requirements laid down in Article 33. More information can be found in Title III Chapter 4 of the AIA.

The provider shall submit two applications to the notified body of their choice[22] per Article 43(1) AIA. Annex VII enumerates the information that should be included in each application to the notified body, one for the quality management system and one for the technical documentation. It also lays down the criteria against which the two applications should be assessed. The approved QMS will be surveyed by the notified body in order to make sure that the provider duly fulfills the terms and conditions of the approved QMS.

The notified body shall communicate to the provider the conclusions of the assessment of the documents submitted, as well as the reasoned assessment decision.

(a) **If the notified body finds that the high-risk AI system is in conformity with the requirements of the Act,** it will issue an EU technical documentation assessment certificate (also see Article 44 AIA). The Certificate has limited time validity and can be suspended or withdrawn by the notified body.

The responsibility of the provider to draw up the EU declaration of conformity and affix the CE marking of conformity remains the same as in the internal CA, as described above.

Any change to the AI system that could affect the compliance of the AI system with the requirements or its intended purpose shall be approved by the notified body that issued the EU technical documentation assessment certificate.

Additionally, the notified body shall have the right to make periodic audits of the approved quality management system in order to make sure that the provider duly fulfills the terms and conditions of the approved quality management system (see Annex VII, point 5)

(b) **If the notified body finds that the high-risk AI system is not in conformity with the requirements of the Act,** such a decision shall be communicated to the provider, including the reasons of the refusal.

The provider has the right to appeal against the decision of the notified body (Article 45 AIA).

As a last note, the AIA gives the Commission the power to adopt delegated acts in accordance with Article 73 for the purpose of updating Annex VI ('CA Procedure based on Internal Control') and Annex VII ('Conformity Based on Assessment of Quality Management System and Assessment of Technical Documentation') in light of technical progress (Article 43(5) AIA).

[21] Each Member State shall designate or establish a notifying authority responsible for setting up and carrying out the necessary procedures for the assessment, designation and notification of conformity assessment bodies and for their monitoring, pursuant to Article 30 of the AI Act proposal.

[22] The provider may choose any of the notified bodies unless the system is intended to be put into service by law enforcement, immigration or asylum authorities as well as EU institutions, bodies or agencies. In that case, it is the market surveillance authority referred to in Article 63(5) or (6), that shall act as a notified body.

Step 3: Who should conduct the CA?

Table 2 Internal or Third-party CA according to the high-risk AI system 1st Category AI systems that are safety components of products or are themselves products that fall under Annex II laws.		
	Type of CA	
Safety component of a product / product covered by the Union harmonization legislation listed in Annex II	**Annex II legislation covers:** - Machinery; - Safety of toys; - Recreational craft and personal watercraft; - Lifts and safety components of lifts; - Equipment and protective systems for use in explosive atmospheres; - Market of radio equipment; - Marker of pressure equipment; - Cableway installations; - Personal protective equipment; - Appliances burning gaseous fuels; - Medical devices (and vitro diagnostic medical devices); - Civil aviation security; - Vehicles; - Marine equipment; - Interoperability of the rail system.	The provider shall follow the relevant conformity assessment as required under those legal acts (Article 43(3)).

Step 3: Who should conduct the CA?

Table 2 (cont.) Internal or Third-party CA according to the high-risk AI system **2nd Category** AI Systems that belong to the use cases of Annex III		
Type of CA		
1. Biometrics (and biometrics-based systems)	If harmonized standards or common specifications have been applied	**Internal CA Or Third-Party CA**
	If the provider has not applied harmonized standards/ common specifications or has applied them only in part	**Third-Party CA**
2. Critical infrastructure 3. Education and vocational training 4. Employment, workers management, and access to self-employment 5. Access to and enjoyment of essential private services and public services and benefits 6. Law enforcement 7. Migration, asylum, and border control management 8. Administration of justice and democratic processes		**Internal CA** (The Commission may amend this rule and require third party CA, through delegated acts.)

Step 3: Who should conduct the CA?

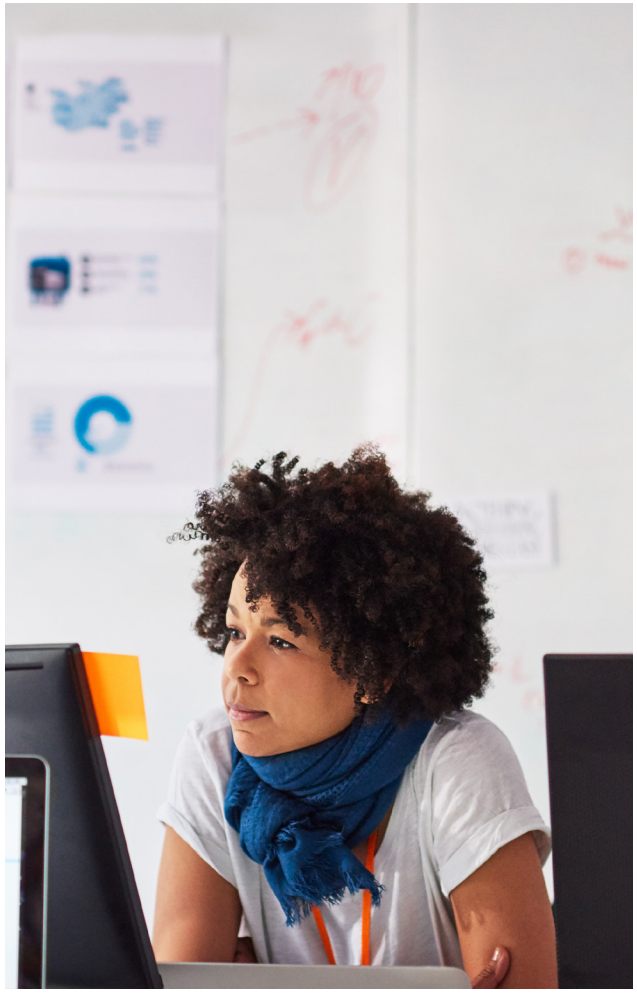
Derogation for exceptional cases (Article 47 AIA)
The proposed AIA introduces exceptional cases under Article 47, in which there can be a derogation from the normal CA process. Only for reasons of public security or the protection of life and health of persons, environmental protection, and the protection of key industrial and infrastructural assets[23] can a high-risk AI system be placed on the market or put into service while the CA has not been concluded. The strict conditions under which such a derogation may take place are yet to be set.

Post-market monitoring system (Article 61 AIA)
A post-market monitoring system is defined under Article 3(25) AIA as "all activities carried out by providers of AI systems to (proactively) collect and review experience gained from the use of AI systems they place on the market or put into service for the purpose of identifying any need to immediately apply any necessary corrective or preventive actions."

The Conformity Assessment is not a one-off exercise. Regardless of whether it is an internal CA or a third-party CA, the provider is required to establish a monitoring system that enables them to verify that the essential requirements are being complied with throughout the lifecycle of the high-risk AI system. For that, Article 61 AIA requires providers to establish a post-marketing monitoring system which will form part of the "quality management system" of Article 17. Since the monitoring takes place after the AI system has entered the market, the user/deployer is also responsible for informing the provider regarding the AI system's performance. The AIA sets the conditions for effective communication and sharing of relevant information between the provider and the user/deployer of the high-risk AI system.

Corrective actions (Article 21 AIA)

In case the provider considers or has reason to consider that the high-risk AI system in use is not in conformity with the AIA, they shall immediately (1) **inform the relevant actors** (e.g. distributors, importers, user/deployer, national competent authority, etc) and (2) **take corrective actions**, as required under Article 21 AIA. Corrective actions might range from bringing the system back to conformity to withdrawing or recalling the system from the market.



[23] The precise grounds on which such a derogation may be allowed, are yet to be set by the co-legislators.

Step 4: Assess conformity with all requirements for high-risk AI systems

All high-risk AI systems must go through the CA process, which aims to verify that all requirements enumerated under Title III, Chapter 2 AIA are complied with. In this Section we will go through these requirements, what they mean, and at what phase of the AI system's life cycle[24] each requirement should be met.

- All requirements should be met before the high-risk AI system enters the market or is put into service - unless otherwise specified.

- The provider is the primarily responsible actor who must ensure that the requirements are met.

- In complying with the requirements, due account shall be taken of the generally acknowledged state-of-the-art, including as reflected in the relevant harmonized standards and common specifications as referred to in Articles 40 and 41, or those already set out in Union harmonization law (Article 8 1(a) Parliament version).

- Compliance with the requirements should be ensured throughout the lifecycle of the system.

- Application of the requirements should account for the intended purpose of the use of the AI system, the reasonably foreseeable misuse of the system, and the risk management system to be established by the provider (Recital 42 AIA).

4.1	Risk Management (Article 9 & Recital 42)
4.2	Data & Data Governance (Article 10 & Recitals 44-45)
4.3	Technical Documentation (Article 11, Recital 46 & Annex IV)
4.4	Record Keeping (Articles 12, 20 & Recital 46)
4.5	Transparency Obligations (Article 13 & Recital 47)
4.6	Human Oversight (Article 14 & Recital 48)
4.7	Accuracy, Robustness & Cybersecurity (Article 15 & Recitals 49-51)

[24] It is only the Council that provides a definition on the 'lifecycle' of an AI system: "'life cycle of an AI system' means the duration of an AI system, from design through retirement. Without prejudice to the powers of the market surveillance authorities, such retirement may happen at any point in time during the postmarket monitoring phase upon the decision of the provider and implies that the system may not be used further. An AI system lifecycle is also ended by a substantial modification to the AI system made by the provider or any other natural or legal person, in which case the substantially modified AI system shall be considered as a new AI system" (Article 3(1a)).

4.1. Is there a risk management system in place (Article 9 AIA)?

Risk Management System
Providers should establish, implement, document, and maintain a Risk Management System throughout the lifecycle of the high-risk AI system.

I. Identification & Assessment of risks (known and reasonably foreseeable);

II. Evaluation of other possibly arising risks (see 'post-market monitoring' and the requirement of 'automatic recording of events');

III. Adoption of Risk Management measures (during the design and development phase of the AI system);

IV. Testing of the high-risk AI system.

Providers of high-risk AI systems should establish, implement, document, and maintain a risk management system (RMS) that runs throughout the entire lifecycle of the high-risk AI system. The provider shall also maintain and monitor the RMS after the AI system has entered the market.

In this sense, the RMS is a continuous and iterative process that should be monitored, reviewed, and updated regularly in order to ensure that it remains relevant and effective. The AIA also requires that the provider and the user/deployer of the high-risk AI system maintain good communication and share information among each other.

A detailed description of the RMS shall be part of the technical documentation under Article 11(1) (see Annex IV point 4) to be attached to the CA, as well as part of the quality management system under Article 17(g). The provider should keep in mind that for AI systems already covered by Union law that require a specific risk management, including credit institutions regulated by Directive 2013/36/EU, the aspects described in Article 9 AIA shall be part of, or combined with, the risk management procedures established by that Union law[25].

Article 9 AIA presents the elements that should be part of a RMS:

1. Identification and assessment of risks

The provider shall identify and evaluate (a) known risks and (b) (reasonably)[26] foreseeable risks that the AI system might pose to health, safety, and fundamental rights of natural persons[27]. The assessment shall be performed on the basis of the intended purpose[28] of the AI system as well as its reasonably foreseeable misuse.

2. Evaluation of other possibly arising risks

The provider shall analyze data that are gathered during the post-marketing monitoring phase and on the basis of this analysis, evaluate risks that might

[25] See Section II.B , point 2 "The EU AIA was initially conceived as a "safety product legislation" , whereby the intention of the European legislature to avoid duplication of processes, is discussed.

[26] The word "reasonably" has been suggested by the European Parliament.

[27] The European Parliament suggests that risks to democracy, the rule of law and the environment should also be identified and assessed.

[28] The Council and the Parliament agree on the definition of "intended purpose": "'intended purpose" means the use for which an AI system is intended by the provider, including the specific context and conditions of use, as specified in the information supplied by the provider in the instructions for use, promotional or sales materials and statements, as well as in the technical documentation" (Article 3(12)).

4.1. Is there a risk management system in place (Article 9 AIA)?

arise during the AI system's use. This part of the RMS is closely related to the requirement of record-keeping (Article 12 AIA). As will be further discussed under subsection 4.4, the automatic recording of events (logs) while the AI system is operating ensures a level of traceability of the AI system's functioning throughout its lifecycle. This enables the monitoring of the AI system, inter alia, for the identification of situations that may result in the AI system presenting a risk. The obligation to monitor the system after it is placed on the market or put into service, and the obligation to evaluate possibly arising risks, justify the requirement of updating an RMS regularly and systematically.

3. Adoption of risk management measures

The AIA does not give examples of potential measures. It does, however, identify the criteria that the provider should take into account when deciding on the most appropriate risk management measures.

- The measures will interact with the other requirements for high-risk AI systems, namely data governance for high-quality datasets, technical documentation, automatic recording of events (logs), transparent operation of the AI systems, measures that enable human oversight and human intervention, and measures that guarantee accuracy, robustness, and cybersecurity. The combined application of the risk management measures, and the measures adopted to satisfy the requirements, shall ensure the **effective mitigation of risks and the appropriate and proportionate implementation of the requirements;**

- any residual risk associated with each identified hazard, as well as the overall residual risks of the AI system as a whole, shall reach **acceptable levels;**

- risk management measures should be **implemented in the design and development phase of the AI system;**

- risk management measures should aim at eliminating or reducing the risks, as far as (technically) feasible (Article 9(4)(a) AIA). Where elimination of risks is not possible, the provider shall implement adequate mitigation and control measures;

- Especially regarding the risks that might appear during the use of the AI system, when adopting the risk management measures, the provider shall take into account:

- (a) the technical knowledge, experience, education, and training to be expected by the **user/deployer.** The provider shall share all relevant information with the user/deployer and, where appropriate, train them;

- (b) the environment in which the system is intended to be used (**context of use**).

4.1. Is there a risk management system in place (Article 9 AIA)?

4.Testing of the high-risk AI system.

Part of the risk management system requires the testing of the high-risk AI system in order to make sure that the AI system performs in a manner that is consistent with its intended purpose, and that the AI system is in compliance with the essential requirements for high-risk AI systems.

- The testing should take place during the development phase of the AI system and, in any case, before placing the AI system on the market.

- Testing shall be made against preliminarily defined metrics and probabilistic thresholds that are appropriate to the intended purpose of the high-risk AI system.

The AIA highlights that while developing the RMS, the provider should pay particular attention to vulnerable groups of people[29] that might interact with or be adversely impacted by the high-risk AI system.

[29] It is only the European Parliament that uses the term"vulnerable groups of people". The Council refers to "persons under the age of 18".

4.2. Are there high-quality datasets used for training, validation, and testing? (Article 10 AIA)

<table>
<tr><td colspan="2">Data Governance
If the high-risk AI system makes use of techniques involving the training of models with data, providers should make sure that they use high-quality datasets for training, validation and testing.</td></tr>
<tr><td>High-Quality Datasets:
- relevant,
- representative,
- appropriately vetted for errors,
- as complete as possible,
- appropriate statistical properties,
- mitigation of bias.</td><td>Data governance practices:
- relevant design choices
- transparency as to the original purpose of data collection
- data collection processes
- data preparation processing operations
- formulation of relevant assumptions
- prior assessment of the availability, quantity and suitability of the data sets that are needed
- examination in view of possible biases
- identification of any possible data gaps or shortcomings.</td></tr>
<tr><td colspan="2">Presumption of conformity:
if the AI system is trained and tested on data reflecting the specific geographical, behavioral, or functional setting within which the AI system is intended to be used.</td></tr>
</table>

High-quality datasets are vital to building safe AI systems that perform as intended and do not lead to discriminatory outputs. High-risk AI systems that make use of techniques involving the training of models with data shall be developed on the basis of training, validation, and testing[30] data sets that meet certain quality criteria. For that, the provider must have in place appropriate data governance on the basis of Article 10 AIA, which shall apply to the development stage of the high-risk AI system.

According to Annex IV(2)(d), detailed information about the datasets used for training, validation, and testing, such as their provenance, scope, and main characteristics, should be part of the technical documentation of Article 11 AIA.

High-quality datasets are datasets that are sufficiently relevant, representative, appropriately vetted for errors, and as complete as possible in view of the intended purpose of the AI system. High-quality datasets should also have the appropriate statistical properties, including regarding the persons or groups of persons for whom the high-risk AI system is intended. Specific attention should be given to the mitigation of possible biases in the datasets, which might create risks to fundamental rights or discriminatory outcomes for the persons affected by the high-risk AI system.

Article 10 AIA enumerates governance practices that providers should adhere to. Those practices include, inter alia, the data collection processes, the assessment of the availability, quantity, and suitability of the data sets needed, the identification of possible biases, etc. Datasets should also take into account the characteristics or elements that are particular to the specific geographical, behavioral, or functional setting within which the high-risk AI system is intended to be used. Article 42 establishes a presumption of conformity with the data governance requirement, where providers have trained and tested their high-risk AI systems on data reflecting the specific

[30] The European Parliament and the Council provide definitions on "training," "validation," and "testing data" under Article 3 (29), (30), and (31) respectively.

4.2. Are there high-quality datasets used for training, validation, and testing? (Article 10 AIA)

geographical, behavioral, or functional setting within which the AI system is intended to be used.

According to Recital 45, providers should be able to access and use high-quality datasets. European common data spaces established by the Commission, as well as the facilitation of data sharing between businesses and with the government in the public interest, will be instrumental in providing trustful, accountable, and nondiscriminatory access to high-quality data for the training, validation, and testing of AI systems (eg. the European health data space).

Article 10 contains a provision that permits the processing of special categories of personal data as defined by the GDPR in Article 9 for the purposes of bias monitoring, detection, and correction. This exceptional processing shall be subject to appropriate safeguards[31], e.g. use of state-of-the-art security and privacy-preserving measures. The European Parliament requires providers who process special categories of personal data, to draw up separate documentation, explaining why that processing was necessary to detect and correct biases.

The Parliament, in its version of the Act, has a specific provision targeted to providers of foundation models. The latter are required to process and incorporate only datasets that are subject to appropriate data governance measures for foundation models, in particular measures to examine the suitability of the data sources and possible biases and appropriate mitigation (Article 28b)

Annex VII on third-party CAs also contains a specific provision referring to data governance practices. It reads that where the AI system does not meet the data governance requirement, and thus the CA is negative, retraining of the AI system will be needed prior to the application for a new CA. In this case, the reasoned assessment decision of the notified body refusing to issue the EU technical documentation assessment certificate shall contain specific considerations on the quality of data used to train the AI system, notably on the reasons for non-compliance.

The European Parliament suggests that the provider who is not able to comply with the data governance requirement because they do not have access to the data and the data is held exclusively by the user/deployer, the latter may, on the basis of a contract, be made responsible for any infringement of Article 10. It remains to be seen whether this provision will make it to the final text of the AIA.

It is expected that the requirements related to the high-quality of datasets will have significant overlap with obligations under the GDPR, whenever the datasets will include personal data. The obligations related to lawfulness, fairness, purpose limitation, and the accuracy principle are among the most relevant ones in this context.

[31] Both the Council and the European Parliament give examples of "appropriate safeguards". However, the European Parliament provides a significantly more elaborate list of conditions that need to be fulfilled in order for the processing of special categories of personal data to take place.

4.3. Has technical documentation been drawn up? (Article 11 AIA)

Technical Documentation
Providers should draw up technical documentation and keep it up-to-date .

1. general description of the AI system;

2. detailed description of the elements of the AI system and of the process for its development;

3. information about the monitoring, functioning, and control of the AI system;

4. detailed description of the risk management system (RMS);

5. description of relevant changes made by the provider to the system through its lifecycle;

6. list of harmonised standards applied in full or in part;

7. copy of the EU declaration of conformity;

8. detailed description of the system in place to evaluate the AI system performance in the post-market phase, including the post-market monitoring plan.

The technical documentation shall be drawn up by the provider before the high-risk AI system is placed on the market or put into service and shall be kept up-to-date. The aim of having technical documentation in place is to demonstrate that the system complies with the essential requirements and to be able to provide national competent authorities and notified bodies with all the necessary information for them to assess compliance with the requirements.

With regard to the content, Annex IV AIA spells out the minimum elements to be included in the technical documentation:

- A **general description of the AI system** (eg. intended purpose, nature of data processed (personal data?), description of hardware and software and the interaction with the AI system, whether the AI system is a component of a product, the system's expected output, scenarios of non-use of the AI system etc);

- **A detailed description of the elements of the AI system** and of the **process for its development;**

- Detailed information about **the monitoring, functioning, and control of the AI system;**

- A detailed description of the **risk management system** in accordance with Article 9;

- A description of relevant changes made by the provider to the system through its lifecycle;

- A list of the **harmonized standards** applied in full or in part. Where no such harmonized standards have been applied, a detailed description of the solutions adopted to meet the requirements set out in Title III, Chapter 2, including a list of other relevant standards and technical specifications applied;

- A copy of the **EU declaration of conformity** (which will be issued after the CA is successfully performed, and hence cannot in fact be part of the technical documentation assessed as part of the CA process);

- A detailed description of the system in place to evaluate the AI system **performance in the post-market phase,** including the postmarket monitoring plan.

4.3. Has technical documentation been drawn up? (Article 11 AIA)

The AIA requires providers to keep the technical documentation at the disposal of the national competent authorities for a period of 10 years after the high-risk AI system has been placed on the market or put into service[32].

It is noteworthy that the European Parliament has included, in its official position, specific obligations for providers of foundation models, under Article 28b, among which is that FM providers have to draw up extensive technical documentation and intelligible instructions for use, in order to enable the downstream providers to comply with their obligations.

Where the national supervisory authority of a Member State finds that the technical documentation is not available, it shall require the provider to act and treat the non-compliance (Article 68).

[32] Article 18 Council version - Article 50(a) Parliament's version.

4.4. Is the automatic recording of events ('logs') possible? (Article 12 AIA)

<table>
<tr><td colspan="2">Record Keeping
Providers should design and develop high-risk AI systems with capabilities that enable the automatic recording of events (logs) while the AI system is operating (AI system traceability).</td></tr>
<tr><td>Aim:
- monitor the system throughout its lifecycle:
- Risks that might arise during use.
- Substantial modifications.
- Facilitate post-market monitoring.

& facilitate the user to comply with their monitoring obligations</td><td>Logs:
e.g. output data, start date, and time
For Remote biometric identification systems (Annex III, para 1, point (a))
- recording of the period of each use of the system,
- the reference database against which input data has been checked by the system,
- the input data for which the search has led to a match,
- the identification of the natural persons involved in the verification of the results.
Logs kept for a period of at least 6 months</td></tr>
<tr><td colspan="2">Conformity with standards/common specification & state-of-the-art</td></tr>
</table>

High-risk AI systems shall be designed and developed with capabilities that enable the automatic recording of events ('logs') while the AI system is operating. Logs include, for instance, output data, start date, and time, and should be kept for a period that is appropriate to enable the responsible actors to fulfill their obligations (Recital 46). Article 20 AIA sets a retention period for logs of at least 6 months. Providers should be mindful that in case these logs qualify as personal data, their retention is a "processing operation" under the GDPR, and should thus comply with all the relevant GDPR provisions.

The AI system's logging capabilities shall conform to recognized standards or common specifications. The European Parliament encourages the adoption of recognized standards or common specifications in order to enable the automatic recording of events ('logs') while the high-risk AI systems are operating (Article 12).

The logging capabilities shall ensure a level of traceability of the AI system's functioning throughout its lifecycle that is appropriate to the intended purpose of the system. The record-keeping obligations are meant to monitor the AI system for the identification of situations that may result in the AI system presenting a risk[33], for any substantial modifications, and, subsequently, to facilitate the post-market monitoring, pursuant to Article 61 AIA, as well as the monitoring of the system's operation by the user/deployer per Article 29(4) AIA.

The Act specifically requires for high-risk AI systems referred to in Paragraph 1, Point (a) of Annex III, that the logging capabilities shall provide, at a minimum, certain information enumerated under Article 12(4).

The provider shall, upon a reasoned request of the national competent authority, give that authority access to the logs to the extent such logs are under the provider's control. In its version, the European Parliament stresses that any information obtained via access to the logs shall be considered a trade secret and should be treated in compliance with the confidentiality obligations set out in Article 70.

[33] See subsection 4.1 on the requirement to have a Risk Management System in place. Part of this RMS is the "Evaluation of the possibly arising risks".

4.5. Is the AI system's operation sufficiently transparent? (Article 13 AIA)

Transparency Providers should design and develop high-risk AI systems to ensure that their operation is sufficiently transparent and that the system's output is interpretable.	
Aim: A. Transparent operation of the AI system - Explainability of the system's decisions (interpretable output)	B. Instructions for use Information provided should be - concise, - complete, - correct, - clear, - relevant, - accessible, - comprehensible to the users. Information should concern - Identity and the contact details of the provider - Characteristics, capabilities, and limitations of performance of the high-risk AI system

Transparency is a concept that is omnipresent in the AIA and it takes various forms. It appears as a general principle applicable to all AI systems (this is so far only under the Parliament's version[34]), as an obligation for providers to register high-risk AI systems in an EU database, as an obligation for providers of limited-risk AI systems (see Recital 70 - Article 1(d) & 52 - Annex IV) and as one of the essential requirements that a high-risk AI system must comply with. Also, transparency becomes particularly important in the context of law enforcement (Recital 38), as well as migration, asylum, and border control management (Recital 39). This section only deals with 'transparency' as an essential requirement for high-risk AI systems.

Article 13 AIA reads that high-risk AI systems should be designed and developed to ensure that their operation is sufficiently transparent and that the system's output is interpretable by the provider and the user/deployer.[35] The Parliament stresses that the user/deployer of the AI system should be able to explain the decisions taken by the AI system in order to, inter alia, be in the position to satisfy the right to explanation of individual decision-making, guaranteed under Article 68 c AIA. An appropriate type and degree of transparency shall be ensured, with a view to achieving compliance with the relevant obligations of the user/deployer and of the provider.

High-risk AI systems shall be accompanied by instructions for use in an appropriate digital format or otherwise. All information provided should support informed decision-making by users/deployers and should be concise, complete, correct, clear, relevant, accessible, and comprehensible to the users/deployers. The information should concern:

1. **Identity and the contact details** of the provider (or authorized representative of the provider);

2. **Characteristics, capabilities, and limitations of performance** of the high-risk AI system (including the intended purpose of the AI system, the level of accuracy, robustness, and cybersecurity,[37] any known

[34] All operators falling under this Regulation shall make their best efforts to develop and use AI systems or foundation models in accordance with the following general principles (...) (d) 'transparency' means that AI systems shall be developed and used in a way that allows appropriate traceability and explainability, while making humans aware that they communicate or interact with an AI system as well as duly informing users of the capabilities and limitations of that AI system and affected persons about their rights" (Article 4(a)).'

[35] Read also Recital 47 AIA.

[37] See also Recital 49 AIA.

4.5. Is the AI system's operation sufficiently transparent? (Article 13 AIA)

or foreseeable circumstance/misuse which may lead to risks to the health and safety, fundamental rights, explainability of the AI system, performance of the system as regards the persons or groups of persons on which the system is intended to be used, specifications for the input data, or any other relevant information in terms of the training, validation and testing data sets use, human oversight measures, predetermined changes to the performance of the system,[38] etc.)

The Parliament adds a provision with transparency obligations specific to providers of foundation models as well as for providers of generative foundation models.[39]

Attention should be given to the fact that when the AI system is processing personal data, the transparency obligations of the GDPR[40] and the ones under AIA should be synced. This is all the more relevant when personal data is processed as part of a solely automated decision-making system that amounts to a high-risk AI system and is subject to enhanced transparency under the GDPR.[41] It is important to note that under the GDPR, transparency is due towards data subjects (an "identified or identifiable natural person", per Article 4(1) GDPR), while transparency under the AIA is due towards the user/deployer. Attention should be paid to Article 68 c, found in the European Parliament version. This Article establishes a "right to explanation of individual decision-making," according to which an affected person shall have the "right to request from the deployer clear and meaningful explanation pursuant to Article 13(1) [AIA] on the role of the AI system in the decision-making procedure, the main parameters of the decision taken and the related input data." If included in the final document of the AIA, Article 68 c AIA will establish another form of transparency, that is transparency of the user/deployer towards the affected person, which will, nonetheless, depend on the system's design and adherence to the transparency requirement of Article 13 AIA. In other words, the right to an explanation of Article 68 c AIA, if adopted as proposed by the European Parliament, will depend on, inter alia, whether the high-risk AI system has been designed and developed on the basis of the transparency requirement[42].

[38] See Step 2 "When to conduct a CA?," bullet point "Ex post".

[39] See Recital 60g and Article 28 b Parliament's version.

[40] Transparency in the GDPR takes the form of a general principle (Article 5(1)(a) GDPR "Personal data shall be processed lawfully, fairly and in a transparent manner in relation to the data subject"), but also the form of legal obligations that fall on the data controller (e.g. Article 12, 13,14) and data subject rights (e.g. Article 15, 22).

[41] Article 22 GDPR reads: "(1) The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her; [...] (3) [...] the data controller shall implement suitable measures to safeguard the data subject's rights and freedoms and legitimate interests, at least the right to obtain human intervention on the part of the controller , to express his or her point of view and to contest the decision." For more information, see the FPF Report, "Automated Decision-Making under the GDPR - A Comprehensive Case-Law Analysis."

[42] See subsection 4.1 on the requirement to have a Risk Management System in place. Part of this RMS is the "Evaluation of the possibly arising risks".

4.6. Is human oversight of the AI system possible? (Article 14 AIA)

<table>
<tr><td colspan="2">Human Oversight
Providers should design and develop high-risk AI systems in a way that they can be effectively overseen by natural persons during the period in which the AI system is in use.</td></tr>
<tr><td>A. effective human oversight

AI systems designed in a way that the user can effectively oversee it and intervene where necessary. (eg. human machine interface tools)</td><td>B. effective human intervention

Built-in operational constraints that cannot be overridden by the system itself and the system is responsive to the human operator. Measures implemented by the provider OR identified by the provider and implemented by the user.</td></tr>
<tr><td colspan="2">- Oversight & intervention while the AI system is in use.
- Natural person that oversees the system:
• Competent,
• trained,
• with the authority to oversee and intervene.
- Provider should inform the user and the user should follow the provider's instructions.</td></tr>
</table>

High-risk AI systems shall be designed and developed in such a way that they can be effectively overseen by natural persons during the period in which the AI system is in use. Human oversight shall aim at preventing or minimizing the risks to health, safety, or fundamental rights that may emerge when a high-risk AI system is used in accordance with its intended purpose or under conditions of reasonably foreseeable misuse.

The requirement of human oversight has two dimensions:

A. The first one refers to the system itself, and the measures taken, which should guarantee that the system is subject to built-in operational constraints that cannot be overridden by the system itself and is responsive to the human operator.[43] Measures that enable human oversight could either be identified and built into the AI system by the provider before it is placed on the market or put into service, when technically feasible, or the measures could be first identified by the provider and then implemented by the user/deployer.

B. The second dimension refers to the natural person responsible for overseeing the system's function. The natural person that has been assigned to oversee the AI system shall have the necessary competence, training, and authority to carry out that role.[44] This dimension raises obligations for both providers and users/deployers. The former shall provide the high-risk AI system to the user/deployer in such a way that the user/deployer is able to:

- understand the capacities and limitations of the system and is able to monitor its operation and be able to detect signs of anomalies, dysfunctions, and unexpected performance;

- remain aware of automation bias;[45]

- correctly interpret the system's output;

[43] Recital 48.

[44] Recital 48.

[45] The Parliament stresses in its version that providers shall "ensure that natural persons to whom human oversight of high-risk AI systems is assigned are specifically made aware of the risk of automation or confirmation bias" (see Article 16 AIA -Parliament's version).

4.6. Is human oversight of the AI system possible? (Article 14 AIA)

- decide not to use the system or to override or reverse the system's output;

- intervene on the operation of the system or interrupt it through a "stop" button or a similar procedure.

Under Article 29, the user/deployer of the high-risk AI system also has an obligation to use the system in accordance with the instructions made available by the provider but also to assign human oversight to a person who is competent, properly qualified, and trained, and has the necessary resources in order to ensure the effective supervision of the AI system.

Detailed information, as well as an assessment of the human oversight measures, including an assessment of the technical measures needed to facilitate the interpretation of the outputs of AI systems by the users, in accordance with Articles 13(3)(d), should be part of the technical documentation.[46]

[46] Annex IV (Technical Documentation), point 2 (e) AIA: the AIA requires the provider to include information on human oversight measures under the "detailed description of the elements of the AI system and of the process for its development" part of the technical documentation.

4.7. Is the AI system accurate and robust? Are there cybersecurity measures in place? (Article 15 AIA).

Accuracy, Robustness, Cybersecurity
Providers should design and develop high-risk AI systems in a way that they achieve, in light of their intended purpose, an appropriate level of accuracy, robustness and cybersecurity.

- Consistent performance of the AI system throughout its lifecycle.
- Accuracy metrics and level of accuracy to be communicated to the user.
- Resilience against errors in the system or interaction with the environment (technical redundancy solutions);
- Resilience against attempts of unauthorized parties, to alter the system's use, behavior, outputs, or performance by exploiting the system vulnerabilities.

Presumption of conformity
where the AI system has been certified under a cybersecurity scheme in so far as the cybersecurity certificate or statement of conformity or parts thereof cover those requirements.

High-risk AI systems shall be designed and developed in such a way that they achieve, in light of their intended purpose, an appropriate level[47] of accuracy, robustness, and cybersecurity,[48] and perform consistently in those respects throughout their lifecycle. In its official position, the Parliament uses the term "security by design and security by default."

The level of accuracy and accuracy metrics should be communicated to the users/deployers and declared in the accompanying instructions of use.[49]

With regard to robustness, high-risk AI systems should be resilient regarding errors, faults, or inconsistencies that may occur within the system or the environment in which the system operates. Providers should adopt technical (and according to the European Parliament, also organizational) measures to achieve resilience. The robustness of high-risk AI systems may be achieved through technical redundancy solutions, which may include backup or failsafe plans. High-risk AI systems that continue to learn after being placed on the market or put into service shall be developed to ensure that possibly biased outputs due to outputs used as an input for future operations ("feedback loops") are duly addressed with appropriate mitigation measures.

Accuracy and robustness are particularly important when there is an interface between the AI system and its user (or any other natural person for this matter).

The AI system should also be resilient against attempts by unauthorized third parties to alter their use, behavior, outputs, or performance by exploiting the system vulnerabilities.[50] The technical solutions aimed at ensuring cybersecurity to address AI-specific vulnerabilities shall include, where appropriate, measures to prevent and control for attacks trying to manipulate the training dataset

[47] The European Parliament suggests, in its version of the AIA, that there will be non-binding guidance that will address the technical aspects of how to measure the "appropriate level" of accuracy and robustness. (Article 15(1a))

[48] In its version of the AIA, the European Parliament has also added the requirement of "safety" next to "accuracy, robustness and cybersecurity" but only in the text and not in the title of the relevant legal provision (Article 15(1)).

[49] Recital 49 AIA, Article 13(3)(b)(ii) AIA, and Section 4.5 of this Guide.

[50] Recital 51 AIA.

4.7. Is the AI system accurate and robust? Are there cybersecurity measures in place? (Article 15 AIA).

("data poisoning"), inputs designed to cause the model to make a mistake ("adversarial examples"), confidentiality attacks, or model flaws.

Providers would benefit from a presumption of compliance with the requirement on cybersecurity where their high-risk AI systems have been certified, or for which a statement of conformity has been issued under a cybersecurity scheme, pursuant to Article 54(3) of Regulation (EU) 2019/881 of the European Parliament and of the Council,[51] as well as the references published in the Official Journal of the European Union (Article 42).

51 Regulation (EU) 2019/881 of the European Parliament and of the Council of 17 April 2019 on ENISA (the European Union Agency for Cybersecurity) and on information and communications technology cybersecurity certification and repealing Regulation (EU) No 526/2013 (Cybersecurity Act).

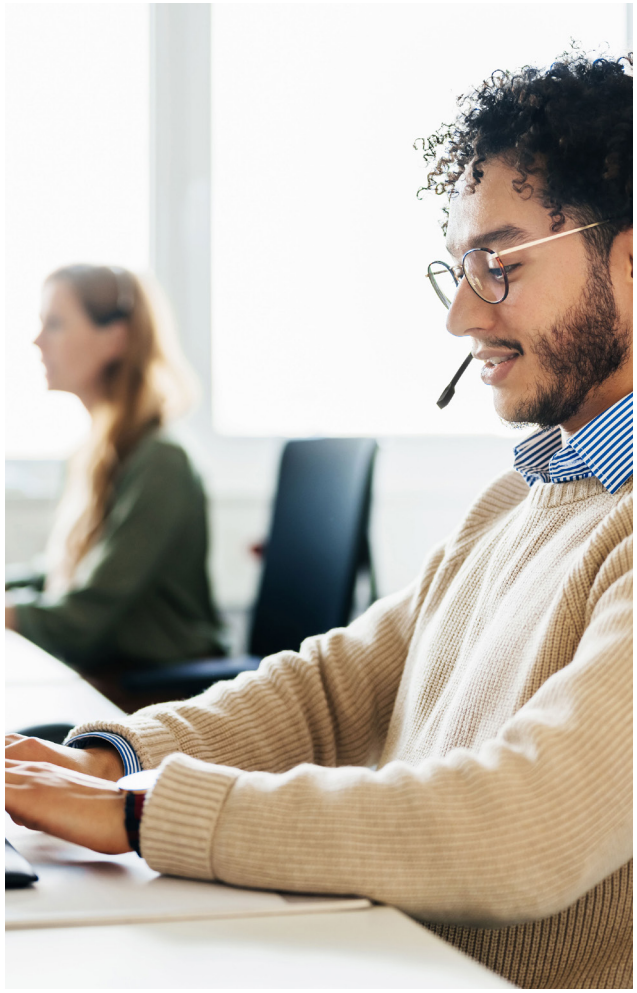
IV. Standards & Presumption of Conformity

Recital 61 highlights that standardization should play a key role in providing technical solutions to providers to ensure compliance with the Regulation. The AIA establishes a presumption of compliance with the requirements for high-risk AI systems, where a high-risk AI system is in conformity with relevant harmonized standards, per Article 40. In case harmonized standards do not exist or are insufficient, the European Commission may adopt common specifications, per Article 41, conformity with which also leads to a presumption of compliance.

The AIA refers to Regulation 1025/2012 on European Standardization for the definition of a harmonized standard. According to Article 2(1)(c) of Regulation 1025/2012, a harmonized standard means a European standard adopted on the basis of a request made by the Commission for the application of Union harmonization legislation.[52]

More specifically, for the adoption of harmonized standards, the European Commission issues standardization requests addressed to the European Standardization Organisations (ESOs, ie. CEN, CENELEC, and ETSI). These requests shall cover all requirements of the AIA, in accordance with Article 10 of Regulation EU (No)1025/2012. The Commission assesses the compliance of the documents drafted by the ESOs with its initial request. Where a harmonized standard satisfies the requirements that it aims to cover and which are set out in the corresponding Union harmonization legislation, the Commission shall publish a reference of such harmonized standard without delay in the Official Journal of the European Union.

In December 2022 (05/12/2022), the European Commission issued a Draft Standardization request to CEN and CENELEC, which are expected to deliver their joint final report by January 2025. The requested standards refer to the essential requirements for high-risk AI systems: to the quality management system (Article 17 AIA) for providers of AI systems, including the post-market monitoring process, and to the



[52] In December 2022, the European Standardization was amended pursuant to the Regulation (EU) 2022/2480 of the European Parliament and of the Council of 14 December 2022 amending Regulation (EU) No 1025/2012 as regards decisions of European standardization organizations concerning European standards and European standardization deliverable.

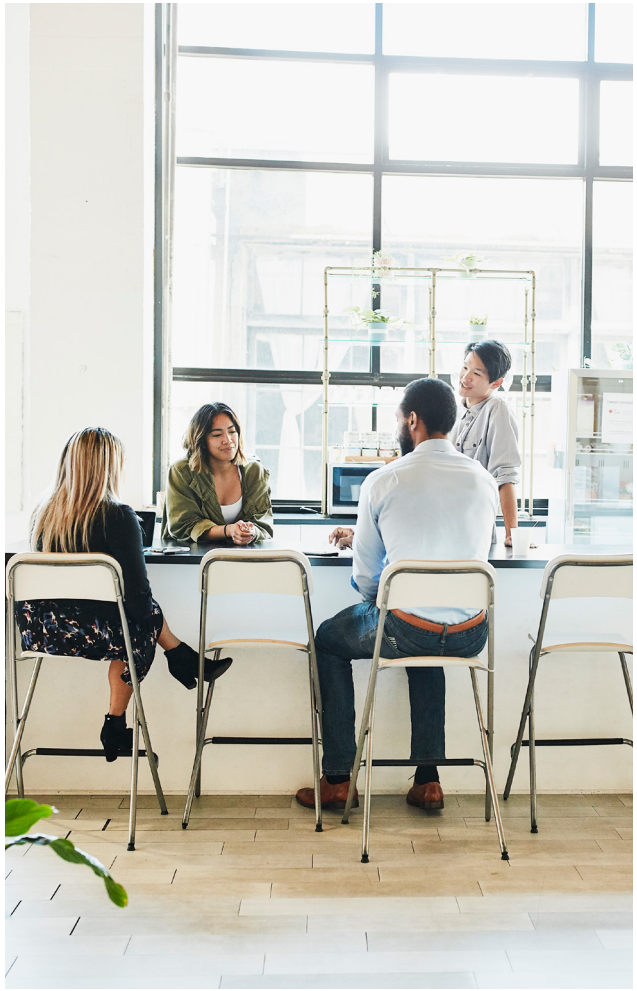
IV. Standards & Presumption of Conformity

conformity assessment procedure for AI systems (see Annex I of the Draft Standardization request).[53]

Specifically, with regard to the CA procedure, the Commission is looking to obtain standards that provide "procedures and processes for conformity assessment activities related to AI systems and quality management systems of AI providers", both in the scenarios of internal CA and third-party CA. Another standardization deliverable is sought to "provide criteria for assessing the competence of persons" tasked with performing a CA.

Such standards will be crucial to developing operational guidance for the implementation of the AIA and are expected to facilitate compliance with the technical obligations prescribed by the Regulation. Given that the AIA is still under negotiation, the Commission clarified that its draft standardization request may be amended when the AIA is finally adopted (Recital 15 Draft Standardization Request).

Pursuant to AIA's measures in support of Innovation, the European Parliament favors, under Article 53(1)(f), a presumption of conformity for those AI systems that are developed within a regulatory sandbox.[54] Prospective providers who develop high-risk AI systems with guidance and supervision on how to fulfill the requirements set out in this Regulation may exit the sandbox being in presumption of conformity with the specific requirements that were assessed within the sandbox.



[53] For more information on Standards-setting for AI systems, see https://artificialintelligenceact.eu/standard-setting/

[54] Regulatory sandboxes are defined in Article 53(1) AIA as follows: "AI regulatory sandboxes established by one or more Member States competent authorities or the European Data Protection Supervisor shall provide a controlled environment that facilitates the development, testing, and validation of innovative AI systems for a limited time before their placement on the market or putting into service pursuant to a specific plan. This shall take place under the direct supervision and guidance of the competent authorities with a view to ensuring compliance with the requirements of this Regulation and, where relevant, other Union and Member States legislation supervised within the sandbox."

Key Takeaways

1. A Conformity Assessment is the process of verifying and/or demonstrating that a "high-risk AI system" complies with the requirements enumerated under Title III, Chapter 2 of the Act. Those requirements consist of: a risk management system; data governance; technical documentation; record-keeping; transparency and provision of information; human oversight; accuracy, robustness, and cybersecurity.

2. The CA should be understood as a framework of assessments, (technical and non-technical) requirements and documentation obligations. The provider should assess whether the AI system qualifies as high-risk and assess known or potential risks as part of the risk management system. The provider should additionally make sure that certain requirements are built-in the high-risk AI system (e.g. automatic recording of events, human oversight capacity, transparent operation of the AI system) as well as whether documentation obligations (e.g. technical documentation) are met.

3. All requirements should be met before the high-risk AI system enters the market or is put into service (unless otherwise specified). Compliance with the requirements should, however, be ensured throughout the lifecycle of the system and until the AI system's withdrawal. For that, all actors involved in an AI system's supply chain should share information among them and should cooperate in a way that ensures compliance with the requirements.

4. Standardization is expected to play a key role in providing technical solutions to providers to ensure compliance with the Regulation. The AIA establishes a presumption of compliance with certain requirements for high-risk AI systems (e.g. cybersecurity requirement, high-quality datasets) as well as in the case where the AI system is developed in the context of a regulatory sandbox.

Sources

https://european-union.europa.eu/institutions-law-budget/institutions-and-bodies/search-all-eu-institutions-and-bodies/european-commission_en

https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence

https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206

https://www.consilium.europa.eu/en/council-eu/

https://www.consilium.europa.eu/en/press/press-releases/2022/12/06/artificial-intelligence-act-council-calls-for-promoting-safe-ai-that-respects-fundamental-rights/

https://www.europarl.europa.eu/about-parliament/en

https://www.europarl.europa.eu/committees/en/imco/home/highlights

https://www.europarl.europa.eu/committees/en/libe/home/highlights

https://www.europarl.europa.eu/news/en/press-room/20230609IPR96212/meps-ready-to-negotiate-first-ever-rules-for-safe-and-transparent-ai

https://single-market-economy.ec.europa.eu/single-market/goods/new-legislative-framework_en

https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32023R0988

https://eur-lex.europa.eu/eli/reg/2023/1230/oj

https://eur-lex.europa.eu/eli/reg/2017/746/oj

https://www.consilium.europa.eu/en/press/press-releases/2022/12/06/artificial-intelligence-act-council-calls-for-promoting-safe-ai-that-respects-fundamental-rights/

https://www.europarl.europa.eu/thinktank/en/document/EPRS_ATA(2023)747926

https://www.consilium.europa.eu/en/press/press-releases/2022/12/06/artificial-intelligence-act-council-calls-for-promoting-safe-ai-that-respects-fundamental-rights/

https://oecd.ai/en/ai-principles

https://www.europarl.europa.eu/thinktank/en/document/EPRS_ATA(2023)747926

Sources

https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022PC0032

https://eur-lex.europa.eu/eli/reg/2019/881/oj

https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32022R2480

https://ec.europa.eu/docsroom/documents/52376

https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32022R2480

https://artificialintelligenceact.eu/standard-setting/

https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206

This page intentionally left blank

FUTURE OF PRIVACY FORUM

The Future of Privacy Forum (FPF) is a non-profit organization that serves as a catalyst for privacy leadership and scholarship, advancing principled data practices in support of emerging technologies. FPF Europe maintains strong partnerships across the EU through its convenings and knowledge-sharing with policymakers and regulators. This transatlantic engagement helps regulators, policymakers, and staff at European Union data protection authorities better understand the technologies at the forefront of data protection law. By building this bridge between European and U.S. data protection cultures, FPF hopes to build a common data protection language. Learn more about FPF Europe by visiting fpf.org/EU.

onetrust

REQUEST A DEMO TODAY AT ONETRUST.COM

As society redefines risk and opportunity, OneTrust empowers tomorrow's leaders to succeed through trust and impact with the Trust Intelligence Platform. The market-defining Trust Intelligence Platform from OneTrust connects privacy, GRC, ethics, and ESG teams, data, and processes, so all companies can collaborate seamlessly and put trust at the center of their operations and culture by unlocking their value and potential to thrive by doing what's good for people and the planet.