Council of the
European Union

Brussels, 13 January 2022
(OR. en)

5293/22

Interinstitutional File:
2021/0106(COD)

LIMITE

TELECOM 9
JAI 43
COPEN 13
CYBER 12
DATAPROTECT 5
EJUSTICE 2
COSI 11
IXIM 12
ENFOPOL 13
FREMP 8
RELEX 37
MI 26
COMPET 18
CODEC 37

NOTE

From:	Presidency
To:	Delegations
No. Cion doc.:	8115/21
Subject:	Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts - Presidency compromise text - Articles 8-15 and Annex IV

I. INTRODUCTION

1. The Commission adopted the proposal for a Regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act, AIA) on 21 April 2021.

2. On 12 October 2021, the SI Presidency requested the delegations in the WP TELECOM to provide written comments and drafting suggestions on **Articles 1-29 and Annexes I-III** of the proposed AIA by 26 October, with a view to start working on the first, partial compromise text of the proposal. Based on the written submissions from the delegations, as well as taking into account the input from the discussions held within the Council, the SI Presidency drafted the first, partial compromise proposal, which covers Articles 1-7 and Annexes I-III of the proposed AIA. This partial compromise proposal was presented to WP TELECOM on 30 November 2021 by the SI Presidency and it was discussed in detail during the meeting of WP TELECOM on 11 January 2022 under the FR Presidency.

3. The FR Presidency has resumed the drafting work where the SI Presidency left off, and it has drafted **the next part of the first compromise proposal, covering Articles 8-15 and Annex IV**, which is contained in the Annex of this document.

4. **The FR Presidency invites the delegations to discuss the proposed changes to Articles 8-15 and Annex IV during the WP TELECOM meeting on 18 January 2022**.

5. The changes in the document compared with the Commission's proposal are underlined: additions are marked with **bold**, deletions with ~~strikethrough~~.

II. MAIN CHANGES

1. **Article 8 - Compliance with the requirements**

a) The text added in the introductory **Article 8(1)** has been moved from **Article 9(3)**, because the reference to 'state of the art' applies to all the requirements specified in the subsequent Articles 9-15.

2. **Article 9 - Risk management system**

a) The modifications in **Article 9(2)** are meant to clarify which types of risks are being addressed with the provisions concerning the requirements for high-risk AI systems. The text in **paragraph 2(b) of Article 9** referring to the risks related to the intended purpose of a high-risk AI system has been deleted because this concept is now covered by the revised **paragraph 2(a)**.

b) The changes in **Article 9(3)** have been made to acknowledge that some of the requirements could be at odds with one another, which may necessitate trade-offs during the implementation, e.g. as regards accuracy vs. robustness, privacy (data minimization) vs. fairness etc.

c) The second part of **Article 9(6)** has been deleted because it is not needed from the legal perspective.

3. **Article 10 - Data and data governance**

a) The modification in **Article 10(3)** is meant to acknowledge the fact that training, validation and testing data sets can never be completely free of errors and to clarify that the requirement is to ensure that they are free of errors to the best extent possible.

b) The text in **Article 10(6)** has been rewritten to indicate that for the development of high-risk AI systems not using techniques involving the training of models, the requirements specified in Article 10 should apply only to the testing data sets (and not to training and validation data sets).

c) A new **paragraph 6a** has been added in Article 10 to clarify that the data minimization principle as laid down in Regulation (EU) 2016/679 should be applied with consideration for the full life cycle of the AI system.

4. Article 11- Technical documentation

a) The changes made in **Article 11(1)** concerning the equivalent documentation that could be drawn up to comply with the requirements of this Article have been made in order to provide more flexibility for SMEs and start-ups.

5. Article 12 - Record-keeping

a) The last sentence in **Article 12(1)** has been deleted to reflect the fact that standards are not obligatory.

b) The modifications in the structure of **Article 12(2)** are meant to make it more readable, and the textual changes have been introduced in order to better explain for what purposes data should be kept.

6. Article 13 - Transparency and provision of information to users

a) The changes in **Article 13(1)** have been introduced in order to simplify the text and to indicate that transparency requirements should help users with interpretability of the output of high-risk AI systems.

b) The modifications in **Article 13(3)** clarify or add further elements which should be included in the instructions of use for high-risk AI systems.

7. Article 14 - Human oversight[1]

a) Changes in **Article 14(4)** are meant to clarify that the technical requirements concerning human oversight fall on the provider and to ensure that they are not too excessive, depending on the circumstances.

[1] Some comments related to human oversight regarding the user have been addressed by the Presidency in the new version of Article 29, which will be presented in WP TELECOM at a later date.

b) In **Article 14(5)** the word 'separately' has been added to ensure the effectiveness of the provisions concerning human oversight in relation to biometric identification systems.

8. **<u>Article 15 - Accuracy, robustness and cybersecurity</u>**

a) The modification in **Article 15(3)** is meant to clarify that the possibility of biased output due to 'feedback loops' and their indirect effects should be prevented with appropriate mitigation measures.

9. **<u>Annex IV - TECHNICAL DOCUMENTATION referred to in Article 11(1)</u>**

a) The additions in **point 1(c)** provide examples of the different forms in which the AI system can be placed on the market or put into service.

b) The changes in **points 2(d) and 5** provide further clarifications as regards the information that needs to be included in the detailed description of the elements of the AI system to be included in technical documentation.

Annex

Proposal for a

REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL

LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS

Chapter 2 (Articles 8-15) and Annex IV

CHAPTER 2

REQUIREMENTS FOR HIGH-RISK AI SYSTEMS

Article 8
Compliance with the requirements

1. High-risk AI systems shall comply with the requirements established in this Chapter**, taking into account the generally acknowledged state of the art, including as reflected in relevant harmonised standards or common specifications**.

2. The intended purpose of the high-risk AI system and the risk management system referred to in Article 9 shall be taken into account when ensuring compliance with those requirements.

Article 9
Risk management system

1. A risk management system shall be established, implemented, documented and maintained in relation to high-risk AI systems.

2. The risk management system shall consist of a continuous iterative process run throughout the entire lifecycle of a high-risk AI system, requiring regular systematic updating. It shall comprise the following steps:

   (a) identification and analysis of the known and foreseeable risks **most likely to occur to health, safety and fundamental rights in view of the intended purpose of the high-risk AI system** ~~associated with each high-risk AI system~~;

~~(b) estimation and evaluation of the risks that may emerge when the high-risk AI system is used in accordance with its intended purpose and under conditions of reasonably foreseeable misuse;~~

(c) evaluation of other possibly arising risks based on the analysis of data gathered from the post-market monitoring system referred to in Article 61;

(d) adoption of suitable risk management measures in accordance with the provisions of the following paragraphs.

The risks referred to in this paragraph shall concern only those which may be reasonably mitigated or eliminated through the development or design of the high-risk AI system, or the provision of adequate technical information.

3. The risk management measures referred to in paragraph 2, point (d) shall give due consideration to the effects and possible interaction resulting from the combined application of the requirements set out in this Chapter 2**, with a view to minimising risks more effectively while achieving an appropriate balance in implementing the measures to fulfil those requirements**. ~~They shall take into account the generally acknowledged state of the art, including as reflected in relevant harmonised standards or common specifications.~~

4. The risk management measures referred to in paragraph 2, point (d) shall be such that any residual risk associated with each hazard as well as the overall residual risk of the high-risk AI systems is judged acceptable, provided that the high-risk AI system is used in accordance with its intended purpose or under conditions of reasonably foreseeable misuse. Those residual risks shall be communicated to the user.

In identifying the most appropriate risk management measures, the following shall be ensured:

(a) elimination or reduction of **identified and evaluated** risks as far as possible through adequate design and development **of the high risk AI system**;

(b) where appropriate, implementation of adequate mitigation and control measures in relation to risks that cannot be eliminated;

(c) provision of adequate information pursuant to Article 13, in particular as regards the risks referred to in paragraph 2, point (b) of this Article, and, where appropriate, training to users.

In eliminating or reducing risks related to the use of the high-risk AI system, due consideration shall be given to the technical knowledge, experience, education, training to be expected by the user and the environment in which the system is intended to be used.

5. High-risk AI systems shall be tested for the purposes of identifying the most appropriate risk management measures. Testing shall ensure that high-risk AI systems perform consistently for their intended purpose and they are in compliance with the requirements set out in this Chapter.

6. Testing procedures shall be suitable to achieve the intended purpose of the AI system ~~and do not need to go beyond what is necessary to achieve that purpose~~.

7. The testing of the high-risk AI systems shall be performed, as appropriate, at any point in time throughout the development process, and, in any event, prior to the placing on the market or the putting into service. Testing shall be made against preliminarily defined metrics and probabilistic thresholds that are appropriate to the intended purpose of the high-risk AI system.

8. ~~When implementing t~~**T**he risk management system described in paragraphs 1 to 7 **shall give** specific consideration **to** ~~shall be given to~~ whether the high-risk AI system is likely to be accessed by or have an impact on **persons under the age of 18** ~~children~~.

9. For credit institutions regulated by Directive 2013/36/EU, the aspects described in paragraphs 1 to 8 shall be part of the risk management procedures established by those institutions pursuant to Article 74 of that Directive.

Article 10
Data and data governance

1. High-risk AI systems which make use of techniques involving the training of models with data shall be developed on the basis of training, validation and testing data sets that meet the quality criteria referred to in paragraphs 2 to 5.

2. Training, validation and testing data sets shall be subject to appropriate data governance and management practices. Those practices shall concern in particular,

(a) the relevant design choices;

(b) data collection **processes**;

(c) relevant data preparation processing operations, such as annotation, labelling, cleaning, enrichment and aggregation;

(d) the formulation of relevant assumptions, notably with respect to the information that the data are supposed to measure and represent;

(e) a prior assessment of the availability, quantity and suitability of the data sets that are needed;

(f) examination in view of possible biases **that are likely to affect health and safety of persons or lead to discrimination prohibited by Union law**;

(g) the identification of any possible data gaps or shortcomings, and how those gaps and shortcomings can be addressed.

3. Training, validation and testing data sets shall be relevant, representative, **and to the best extent possible,** free of errors and complete. They shall have the appropriate statistical properties, including, where applicable, as regards the persons or groups of persons on which the high-risk AI system is intended to be used. These characteristics of the data sets may be met at the level of individual data sets or a combination thereof.

4. Training, validation and testing data sets shall take into account, to the extent required by the intended purpose, the characteristics or elements that are particular to the specific geographical, behavioural or functional setting within which the high-risk AI system is intended to be used.

5. To the extent that it is strictly necessary for the purposes of ensuring bias monitoring, detection and correction in relation to the high-risk AI systems, the providers of such systems may process special categories of personal data referred to in Article 9(1) of Regulation (EU) 2016/679, Article 10 of Directive (EU) 2016/680 and Article 10(1) of Regulation (EU) 2018/1725, subject to appropriate safeguards for the fundamental rights and freedoms of natural persons, including technical limitations on the re-use and use of state-of-the-art security and privacy-preserving measures, such as pseudonymisation, or encryption where anonymisation may significantly affect the purpose pursued.

6. **For the development of high-risk AI systems not using techniques involving the training of models, paragraphs 2 to 5 shall apply only to the testing data sets.**

~~Appropriate data governance and management practices shall apply for the development of high risk AI systems other than those which make use of techniques involving the training of models in order to ensure that those high risk AI systems comply with paragraph 2.~~

6a. In order to comply with the requirements laid out in this Article, the data minimisation principle referred to in Article 5 paragraph 1c of Regulation (EU) 2016/679 shall be applied with consideration for the full life cycle of the system.

Article 11
Technical documentation

1. The technical documentation of a high-risk AI system shall be drawn up before that system is placed on the market or put into service and shall be kept up-to date.

The technical documentation shall be drawn up in such a way to demonstrate that the high-risk AI system complies with the requirements set out in this Chapter and provide national competent authorities and notified bodies with all the necessary information to assess the compliance of the AI system with those requirements. It shall contain, at a minimum, the elements set out in Annex IV **or, in the case of SMEs and start-ups, any equivalent documentation meeting the same objectives, subject to approval of the competent authority**.

2. Where a high-risk AI system related to a product, to which the legal acts listed in Annex II, section A apply, is placed on the market or put into service one single technical documentation shall be drawn up containing all the information set out in Annex IV as well as the information required under those legal acts.

3. The Commission is empowered to adopt delegated acts in accordance with Article 73 to amend Annex IV where necessary to ensure that, in the light of technical progress, the technical documentation provides all the necessary information to assess the compliance of the system with the requirements set out in this Chapter.

Article 12
Record-keeping

1. High-risk AI systems shall ~~be designed and developed with capabilities enabling~~ **technically allow for** the automatic recording of events (‘logs’) **over the duration of the life cycle of the system** ~~while the high-risk AI systems is operating~~. ~~Those logging capabilities shall conform to recognised standards or common specifications.~~

2. ~~The logging capabilities shall ensure~~ **In order to ensure** a level of traceability of the AI system’s functioning ~~throughout its lifecycle~~ that is appropriate to the intended purpose of the system**,**~~.~~ ~~3. In particular,~~ logging capabilities shall enable the **recording of events relevant for** ~~monitoring of the operation of the high-risk AI system with respect to the occurrence of~~

   (i) identification of situations that may result in the AI system presenting a risk within the meaning of Article 65(1) or ~~lead to~~ **in** a substantial modification**;**~~, and~~

   (ii) ~~facilitate~~ **facilitation of** the post-market monitoring referred to in Article 61~~.~~**; and**

   (iii) monitoring of the operation of high-risk AI systems referred to in Article 29(4).

4. For high-risk AI systems referred to in paragraph 1, point (a) of Annex III, the logging capabilities shall provide, at a minimum:

   (a) recording of the period of each use of the system (start date and time and end date and time of each use);

   (b) the reference database against which input data has been checked by the system;

   (c) the input data for which the search has led to a match;

   (d) the identification of the natural persons involved in the verification of the results, as referred to in Article 14 (5).

Article 13
Transparency and provision of information to users

1. High-risk AI systems shall be designed and developed in such a way to ensure that their operation is sufficiently transparent ~~to enable users to interpret the system’s output and use it appropriately. An appropriate type and degree of transparency shall be ensured,~~ with a view to achieving compliance with the relevant obligations of the user and of the provider set out in Chapter 3 of this Title **and enabling users to understand and use the system appropriately**.

2. High-risk AI systems shall be accompanied by instructions for use in an appropriate digital format or otherwise that include concise, complete, correct and clear information that is relevant, accessible and comprehensible to users.

3. The information referred to in paragraph 2 shall specify:

(a) the identity and the contact details of the provider and, where applicable, of its authorised representative;

(b) the characteristics, capabilities and limitations of performance of the high-risk AI system, including:

(i) its intended purpose**, inclusive of the specific geographical, behavioural or functional setting within which the high-risk AI system is intended to be used**;

(ii) the level of accuracy, **including its metrics,** robustness and cybersecurity referred to in Article 15 against which the high-risk AI system has been tested and validated and which can be expected, and any known and foreseeable circumstances that may have an impact on that expected level of accuracy, robustness and cybersecurity;

(iii) any known or foreseeable circumstance, related to the use of the high-risk AI system in accordance with its intended purpose or under conditions of reasonably foreseeable misuse, which may lead to risks to the health and safety or fundamental rights;

(iv) **when appropriate,** its ~~performance~~ **behaviour regarding specific** ~~as regards the~~ persons or groups of persons on which the system is intended to be used;

(v) when appropriate, specifications for the input data, or any other relevant information in terms of the training, validation and testing data sets used, taking into account the intended purpose of the AI system.

(c) the changes to the high-risk AI system and its performance which have been pre-determined by the provider at the moment of the initial conformity assessment, if any;

(d) the human oversight measures referred to in Article 14, including the technical measures put in place to facilitate the interpretation of the outputs of AI systems by the users;

(e) **the computational and hardware resources needed,** the expected lifetime of the high-risk AI system and any necessary maintenance and care measures to ensure the proper functioning of that AI system, including as regards software updates~~.~~**;**

(f) a description of the mechanism included within the AI system that allows users to properly collect, store and interpret the logs, where relevant.

Article 14
Human oversight

1. High-risk AI systems shall be designed and developed in such a way, including with appropriate human-machine interface tools, that they can be effectively overseen by natural persons during the period in which the AI system is in use.

2. Human oversight shall aim at preventing or minimising 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, in particular when such risks persist notwithstanding the application of other requirements set out in this Chapter.

3. Human oversight shall be ensured through either one or all of the following **types of** measures:

   (a) **measures** identified and built, when technically feasible, into the high-risk AI system by the provider before it is placed on the market or put into service;

   (b) **measures** identified by the provider before placing the high-risk AI system on the market or putting it into service and that are appropriate to be implemented by the user.

4. ~~The measures referred to in paragraph 3 shall enable the individuals~~ **For the purpose of implementing paragraphs 1 to 3, the high-risk AI system shall be provided to the user in such a way that natural persons** to whom human oversight is assigned **are enabled,** ~~to do the following~~, as appropriate **and proportionate** to the circumstances:

   (a) ~~fully~~ **to** understand the capacities and limitations of the high-risk AI system and be able to duly monitor its operation~~, so that signs of anomalies, dysfunctions and unexpected performance can be detected and addressed as soon as possible~~;

   (b) **to** remain aware of the possible tendency of automatically relying or over-relying on the output produced by a high-risk AI system (‘automation bias’)~~, in particular for high-risk AI systems used to provide information or recommendations for decisions to be taken by natural persons~~;

   (c) ~~be able~~ to correctly interpret the high-risk AI system’s output, taking into account **for example** ~~in particular~~ ~~the characteristics of the system and~~ the interpretation tools and methods available;

   (d) ~~be able~~ to decide, in any particular situation, not to use the high-risk AI system or otherwise disregard, override or reverse the output of the high-risk AI system;

   (e) ~~be able~~ to intervene on the operation of the high-risk AI system or interrupt the system through a “stop” button or a similar procedure.

5. For high-risk AI systems referred to in point 1(a) of Annex III, the measures referred to in paragraph 3 shall be such as to ensure that, in addition, no action or decision is taken by the user on the basis of the identification resulting from the system unless this has been **separately** verified and confirmed by at least two natural persons.

Article 15
Accuracy, robustness and cybersecurity

1. High-risk AI systems shall be designed and developed in such a way that they achieve, in the light of their intended purpose, an appropriate level of accuracy, robustness and cybersecurity, and perform consistently in those respects throughout their lifecycle.

2. The levels of accuracy and the relevant accuracy metrics of high-risk AI systems shall be declared in the accompanying instructions of use.

3. High-risk AI systems shall be resilient as regards errors, faults or inconsistencies that may occur within the system or the environment in which the system operates, in particular due to their interaction with natural persons or other systems.

   The robustness of high-risk AI systems may be achieved through technical redundancy solutions, which may include backup or fail-safe plans.

   High-risk AI systems that continue to learn after being placed on the market or put into service shall be developed in such a way to ensure that possibly biased outputs ~~due to outputs used as~~ **influencing** ~~an~~ input for future operations (‘feedback loops’) are duly addressed with appropriate mitigation measures.

4. High-risk AI systems shall be resilient as regards attempts by unauthorised third parties to alter their use or performance by exploiting the system vulnerabilities.

   The technical solutions aimed at ensuring the cybersecurity of high-risk AI systems shall be appropriate to the relevant circumstances and the risks.

   The technical solutions to address AI specific vulnerabilities shall include, where appropriate, measures to prevent and control for attacks trying to manipulate the training dataset (‘data poisoning’), inputs designed to cause the model to make a mistake (‘adversarial examples’), or model flaws.

***

ANNEX IV
TECHNICAL DOCUMENTATION referred to in Article 11(1)

The technical documentation referred to in Article 11(1) shall contain at least the following information, as applicable to the relevant AI system:

1. A general description of the AI system including:

   (a) its intended purpose, the person/s developing the system the date and the version of the system;

(b) how the AI system interacts or can be used to interact with hardware or software that is not part of the AI system itself, where applicable;

(c) the versions of relevant software or firmware and any requirement related to version update;

(d) the description of all forms in which the AI system is placed on the market or put into service **(e.g. software package embedded into hardware, downloadable, API etc.)**;

(e) the description of hardware on which the AI system is intended to run;

(f) where the AI system is a component of products, photographs or illustrations showing external features, marking and internal layout of those products;

(g) instructions of use for the user and, where applicable installation instructions;

2. A detailed description of the elements of the AI system and of the process for its development, including:

(h) the methods and steps performed for the development of the AI system, including, where relevant, recourse to pre-trained systems or tools provided by third parties and how these have been used, integrated or modified by the provider;

(i) the design specifications of the system, namely the general logic of the AI system and of the algorithms; the key design choices including the rationale and assumptions made, also with regard to persons or groups of persons on which the system is intended to be used; the main classification choices; what the system is designed to optimise for and the relevance of the different parameters; the decisions about any possible trade-off made regarding the technical solutions adopted to comply with the requirements set out in Title III, Chapter 2;

(j) the description of the system architecture explaining how software components build on or feed into each other and integrate into the overall processing; the computational resources used to develop, train, test and validate the AI system;

(k) where relevant, the data requirements in terms of datasheets describing the training methodologies and techniques and the training data sets used, **including a general description of these data sets**, ~~including~~ information about ~~the~~ **their** provenance ~~of those data sets~~, ~~their~~ scope and main characteristics; how the data was obtained and selected; labelling procedures (e.g. for supervised learning), data cleaning methodologies (e.g. outliers detection);

(l) assessment of the human oversight measures needed in accordance with Article 14, 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);

(m) where applicable, a detailed description of pre-determined changes to the AI system and its performance, together with all the relevant information related to the technical solutions adopted to ensure continuous compliance of the AI system with the relevant requirements set out in Title III, Chapter 2;

(n) the validation and testing procedures used, including information about the validation and testing data used and their main characteristics; metrics used to measure accuracy, robustness, cybersecurity and compliance with other relevant requirements set out in Title III, Chapter 2 as well as potentially discriminatory impacts; test logs and all test reports dated and signed by the responsible persons, including with regard to pre-determined changes as referred to under point (f).

3. Detailed information about the monitoring, functioning and control of the AI system, in particular with regard to: its capabilities and limitations in performance, including the degrees of accuracy for specific persons or groups of persons on which the system is intended to be used and the overall expected level of accuracy in relation to its intended purpose; the foreseeable unintended outcomes and sources of risks to health and safety, fundamental rights and discrimination in view of the intended purpose of the AI system; the human oversight measures needed in accordance with Article 14, including the technical measures put in place to facilitate the interpretation of the outputs of AI systems by the users; specifications on input data, as appropriate;

4. A detailed description of the risk management system in accordance with Article 9;

5. A description of ~~any~~ **relevant** change**s** **made by the provider** to the system through its lifecycle;

6. A list of the harmonised standards applied in full or in part the references of which have been published in the Official Journal of the European Union; where no such harmonised 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;

7. A copy of the EU declaration of conformity;

A detailed description of the system in place to evaluate the AI system performance in the post-market phase in accordance with Article 61, including the post-market monitoring plan referred to in Article 61(3).