# ABB feedback
# Artificial intelligence – ethical and legal requirements

## GENERAL COMMENTS

ABB welcomes the opportunity to comment on the European Commission proposal for a Regulation laying down harmonised rules on Artificial Intelligence ("AI Act"). The AI Act is an important step towards the ambitious goal to promote and facilitate the uptake of Trustworthy AI in Europe.

We believe the debate should equally balance a discussion on potential risks with focus on the significant opportunities linked to the uptake of AI in industrial sectors. Europe's AI regulation can only aspire to become a global blueprint if it can equally address risks and facilitate innovation while promoting the general goal of boosting AI development and application across businesses in Europe.

You will find below our input on specific Articles and our proposals to improve the current text. We remain at your disposal for further information and feedback throughout the legislative process.

## DETAILED COMMENTS

## Article 1, Article 3(1) and Annex I

The proposal is limited to the regulation of high-risk AI systems. This limitation is reasonable and welcome. The definition of high-risk systems in article 3 (1) and Annex I includes industrial systems.

### Issues

The overall limitation is reasonable and should not be changed. Consequently, the definition of high-risk (Article 3 and Annex I, next section) should be re-evaluated. Today, safety in industrial systems is already subjected to continuous observation and is sufficiently regulated and standardized. The simple introduction of AI does not necessarily justify additional regulation.

### Our input

This Regulation shall not apply to AI systems developed for or exclusively used in an industrial environment, unless there is a clear domain-specific gap motivating the need of an additional regulation beyond existing regulations and directives.

## Article 3(1) and Annex I

The list of methods includes machine learning, logic- and knowledge-based approaches, statistical approaches, Bayesian estimation, search and optimization.

### Issues

The definition is too broad, especially in the context of safety components. Functional safety is handled separately from the generic control mechanism. The state of the art has established methods to achieve functional safety in a deterministic and verifiable way. With the broad definition of AI (especially the unspecified mentioning of logic- and knowledge-based approaches) many such methods will fall under the Regulation.

The present directive for safety of machinery and the harmonized standards to implement it already go beyond the intention of the AI regulation by disallowing in practice any non-deterministic, non-verifiable methods as part of functional safety. Hence, data-driven or stochastic methods are excluded from use in functional safety today.

### Our input

We recommend focussing the definition by excluding established “logic” based methods and instead concentrating on non-determinism, non-verifiability, non-predictability, or opaqueness of methods. We would light to highlight a proposal for improved definition:

> **Article 3 (1):**
>
> ‘artificial intelligence system’ (AI system) means a system of which essential functions are controlled by software that is typically developed with one or more

of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with. However, use of these techniques and approaches alone do not necessarily qualify a system as an AI system.

AI systems can produce either deterministic results (i.e. the same output given identical input, regardless of when during the lifetime of the system this is executed), or non-deterministic results (i.e. based on continuous accrual of data or modifications of algorithm during the course of its lifetime).

**Annex I – AI Techniques and approaches referred to in Article 3, point 1 means:**

(a) Machine learning approaches, including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning; potentially based on the approaches listed in (b):

(b) Logic- and knowledge-based approaches, including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems; including statistical approaches, Bayesian estimation, search and optimization methods.

Particularly important is whether the AI method used produces deterministic or non-deterministic output. For applications in safety components, e.g. in safety of machinery, the methods used must lend themselves to reproducible verification and to persistent validation. The results of validation of safety functions must retain validity throughout the lifetime (mission time) of the machinery.

## Article 6

This article defines the scope of “high-risk” AI systems. Principally, these are applications in “safety components” and applications as listed in Annex III, which include uses in various types of social scoring.

### Issues

Safety of machinery has a well-established foundation in the present version of the Machinery Directive 2006/42EC and the derived harmonized standards. Since the Machinery Directive is formulated in a technology-neutral manner, it is applicable even if one were to use novel methods of data analysis and interpretation, including in principle also methods of AI. Adding explicit requirements on uses of AI in safety components and then defining AI very widely to encompass virtually all methods using an algorithm executed on a processor leads to a complex situation that calls into question all legacy technical solutions for the market of safety controllers and safety sensors.

The same applies for the definition of functional safety where IEC 61508 and derived standard like IEC 61511 for process industries.

As a result, AI systems that focus on the non-determinism of the related algorithms should be excluded from use in safety-related control systems. Conversely, the subset of algorithms that deliver deterministic results should not be considered as "high-risk" as their use is well-established today and as their operation is covered by existing harmonisation legislation. AI algorithms of non-deterministic nature or which do not produce verifiable output valid over the lifetime of the associated equipment should not be deployed in safety-related parts of control systems.

Conversely, should it be possible to classify a method used in an AI system as deterministic and verifiable, then such cases should not be burdened with compulsory third-party conformity assessment. Self-declaration should continue to be a possibility, assuming that the machine in question is not listed in Annex IV of the Machinery Directive (or the corresponding future list in the Machinery Regulation).

### Our input

We recommend to focus the scope of paragraph 1 to include only non-deterministic methods used in safety components (even if one might hold the opinion that such things should not exist). Result:

1. Regardless of whether the AI system in question is incorporated into a product or is itself a product, that AI system shall be considered high-risk when all of the following hold:
   a) the AI system is based on non-deterministic algorithms or is data-driven such that it is not possible in principle to deterministically verify that it fulfills the specified requirements for all intended use and foreseeable misuse and such that its validation cannot be guaranteed to hold over the mission time (lifetime) of the application into which it is installed.

b) the AI system is intended to be used as a safety component of a product, or is itself a product, covered by the Union harmonisation legislation listed in Annex II;
c) the product whose safety component is the AI system, or the AI system itself as a product, is required to undergo a third-party conformity assessment with a view to the placing on the market or putting into service of that product pursuant to the Union harmonisation legislation listed in Annex II.

## Article 9

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

### Issues

There is already NLF legislation in place. This leads to double regulation and potential inconsistencies.

### Our input

Risk management system

(1) A risk management system shall be established, implemented, documented and maintained in relation to high-risk AI systems. **It should not be applied to the high-risk AI systems already covered by the NLF legislations.**

Consider a directive-oriented approach rather than a regulation, to allow for national implementations (e.g. HLEG proposed AI check list, April 2019)

## Article 10

Data and data governance

(3) Training, validation and testing data sets shall be relevant, representative, 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.

### Issues

Fundamentally, freedom from errors and completeness of coverage can never be ensured. Therefore, if this is set as a requirement, data-driven methods can never be used. In this case, much of the subsequent regulation becomes void.

Furthermore, the responsibility/accountability for data quality (in testing, experimentation, and operation) needs to be clarified.

Contextual assessment: data quality should be coupled to a certain application/goal/task and only be assessed in that respect. It is preferred to deal with the specific task such as “automated decision making” instead of concrete technologies.

### Our input

Data and data governance

(3) Training, validation and testing data sets shall be relevant and representative., f~~ree 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.

## Article 41

Possibility for the European Commission to adopt common technical specifications" via implementing acts by-passing established standardization process.

### Issues

We believe the established standardisation process is fundamentally superior as it allows the EU's industry to leverage all potential AI solutions including from outside the EU to improve the efficiency of their processes.

### Our input

We recommend maintaining the established standardization process with technical experts (e.g. CEN and CENELEC) in place.

Article 41

Common specifications

1. Where harmonised standards referred to in Article 40 do not exist or where the Commission *together with ESOs and industry representatives* considers that the relevant harmonised standards are *seriously* insufficient or that there is a need to address *pressing and* specific safety or fundamental right concerns, *that cannot be reasonably addressed by development of standardization,* the Commission may, by means of implementing acts, adopt common specifications in respect of the requirements set out in Chapter 2 of this Title. Those implementing acts shall be adopted in accordance with the examination procedure referred to in Article 74(2).

2. The Commission, when preparing the common specifications referred to in paragraph 1, shall gather the views of relevant bodies*, industry, ESOs* or expert groups established under relevant sectorial Union law.

## Article 43

Conformity Assessment

The process to achieve conformity is outlined for cases of different application areas, for domains in which harmonization legislation is already in place and for scenarios in which updates or changes to algorithms and/or data sets are made during the lifetime of the AI system or the application incorporating the AI system.

### Issues

The analogy of the situation that an AI system might either be a component, not yet purpose-bound, or might be part of a larger purposely designed system is reminiscent of the situation with partly completed machinery vs. completed machinery. This article is the place in the proposal where one can take the opportunity to project this situation onto AI systems.

Thus, this is an opportunity to define an analogy to "**partly completed machinery**" as treated in the Machinery Directive (MD).

While subject to the requirements of the MD, **partly completed machinery** may not be CE marked since it is considered a component whose purpose is not yet fully determined (example: industrial robot before being integrated into an application). The **completed machinery**, which may contain one or more instances of partly completed machinery, is then purpose-bound and must conform to the MD and must be CE marked when this is established.

For AI systems that are not in themselves fully purpose-bound, one might have the same concept of "**partly completed AI**", which cannot be CE marked although it must follow the AI Regulation, and "**completed AI system**", which is then a part of a larger, purpose-bound system which must be CE marked according to the AI and possibly also other regulations and directives.

### Our input

For AI systems that are not in themselves purpose-bound and can have multiple applicability, a risk assessment of intended use and foreseeable misuse is not possible. Such AI systems that are components for integration into larger systems shall be considered “**partly completed AI**”. They cannot be CE marked but shall be accompanied by a Declaration of Incorporation for safe integration into a larger system.

When AI systems are incorporated into a complete system, they shall be considered “**completed AI systems**”. The conformity of AI systems as components of a larger system shall be covered by the harmonization legislation applicable to the larger system that is the final, purpose-bound product. Such legislation shall include the relevant requirements of the AI regulation and shall serve as the basis for CE marking the completed system.

## Article 56 and Article 58

A European Artificial Intelligence Board shall be established. One task is to collect and share expertise and best practices among Member States.

### Issue

The formulation of the Article should specify that exchanges should be focussed on regulatory and policy aspects of AI and not technological solutions.

## ANNEX III (4)

Worker management systems are high-risk systems.

### Issues

Worker management is a very broad term. While it is understandable that workers should be protected in fundamental rights, the broad scope does not match with the extensive obligations that follow. The definition should also consider which data about the individual is used within the AI system and worker management and scheduling based on non-person related information (age, gender, etc.) should be excluded from this definition.

### Our input

Replace worker management with “worker management leveraging personal information such as age, gender, etc.”

Alternatively, white-list data that can be used, e.g. employee no., work plan / schedule, work / task history.