WilhelmGilliéron
AVOCATS

Auteur: Philippe Gilliéron | Le : 23 February 2024

The EU AI Act – 3 : High-Risk AI Systems

Although GPAIs models have understandably drawn significant recent attention, including in the final rounds of negotiations having led to the adoption of the EU AI Act, as detailed in our second paper of this series, high risk systems lie at the core of the EU AI Act.

The placing on the market, putting into service, respectively use of such systems can only take place in compliance with stringent requirements. While these requirements will be put under scrutiny in our next week(s)' paper, we shall focus this week on the definition of such high risk systems, as set out in Art. 6 of the EU AI Act.

This provision provides for two main categories of AI systems which, by default, are presumed to be high risk:

I. AI systems subject to further compliance requirements

The first category relates to AI systems that are intended to be used as a safety component of a product, *i.e.* a component of a product or system meant to fulfil a safety function, the failure of which endangers the health and safety of persons or property; provided, however, that these AI systems have to be subject to one of the EU regulations listed in Annex II of the EU AI Act and, in accordance with these regulations, to a third party conformity assessment prior to releasing the products on the market.

Annex II lists 19 categories of regulations on a variety of topics: toys, lifts, equipment and protective systems intended for use in potentially explosive atmospheres, radio equipment, pressure equipment, cableway installations, appliances burning gaseous fuels, medical devices and in vitro diagnostic medical devices as well as means of transportations (ranging from civil aviation to two-or-three wheel vehicles, rail or marine equipment).

The fact that the AI system embedded in such a product is considered high risk does not mean that the product embedding such AI system does itself qualify as high risk within the meaning of other regulations to which such product may be subject such as, for instance, Regulation (EU) 2017/745 on medical devices or Regulation (EU) 2017/746 on in vitro diagnostic medical devices, both of which provide for a third-party conformity assessment for medium-risk and high-risk products (as defined by these Regulations).

In these areas, the placing on the market of a product, such as a medical device for instance, is subject to certain governance requirements. While the requirements set out in these existing regulations will not necessarily (and even certainly) match the ones set out in the EU AI Act, so that compliance with the existing regulations does not mean that the stakeholders comply with the requirements set out in the EU AI Act, providers should have flexibility on operational decisions on how to ensure compliance of a product that contains an AI system with all applicable requirements.

Should the system or product covered by one of the 19 categories listed in Annex II however not be subject to a third-party conformity assessment (in particular because it does not match the requirements set out by the applicable Regulation to be subject to such an assessment), the AI system should as a result not be considered as high risk *per se*.

Providers should however then assess whether their AI system may fall under one of the categories listed in Annex III to which I shall now turn:

II. High risk AI systems

Stand-alone AI systems will be considered "high risk" if they pose a serious threat to the health and safety of individuals or their fundamental rights, taking into account both (i) the severity of the possible harm and its probability of occurrence and (ii) the fact that these systems are used in pre-defined categories specified in the Regulation, namely its Annex III.

As a result, the following systems and use cases are considered high-risk:

- *Biometric data* which, by definition, relate to sensitive personal data, trigger specific privacy related risks and may lead in case of technical inaccuracies to discriminatory effects. The EU AI Act qualifies the following use cases of such data as high-risk:

  (i) remote biometric identification systems. These systems have to be distinct on the one side from real-time remote biometric identification, which in general is classified as a prohibited practice as we have seen in our second paper to this series, and on the other side from biometric verification (one-to-one verification), which consists of checking the identity of a person to confirm it for the sole purpose of having access to a service, premises or to unlock a device, bearing in mind that such use will in any case subject to the GDPR or applicable data protection laws.

  (ii) Biometric categorization based upon special categories of data within the meaning of Art. 9 GDPR.

  (iii) Emotion recognition systems.

- *Critical infrastructure*, *i.e.* an asset, facility, equipment, network or a system necessary for the provision of an essential service within the meaning of Art. 2(4) Directive (EU) 202/2557, when the systems are meant to be used as safety components in the management and operation of such infrastructure, road traffic and the supply of water, gas, heating and electricity. Failure of such systems may obviously put at risk the life and health of individuals.

  In both the above use cases, the use of such systems for cybersecurity purposes or personal data protection measures as far as biometric data is concerned are not considered high-risk systems.

- *Education.* While the use of such systems by learners and teachers is welcome, such use may lead perpetuate historical patterns of discrimination and should be classified as high risk when there are intended to be used to:

  (i) determine access or admission;

  (ii) evaluate learning outcomes;

  (iii) assess the appropriate level of education that individuals will access or be able to access;

  (iv) monitor and detect prohibited behavior of students during tests.

  In all these use cases, such systems may affect the individuals' educational and professional course of life, thus their ability to secure their livelihood.

- *Employment.* The EU AI Act retains two use cases in that space that are to be considered high risk:

  (i) for hiring purposes, including to place targeted job advertisements, analyse, filter and evaluate job applications are to be considered high risk.

  (ii) to make decisions affecting terms of the work-related relationships, promotion and termination, or to allocate tasks,

monitor and evaluate performance of individuals.

Such uses may impact future career prospects and worker's rights, potentially perpetuating historical patterns of discrimination.

- *Access and enjoyment of essential private and public services and benefits*. Several uses cases are contemplated by the Commission under this heading, namely:

  (i) to evaluate the eligibility of individuals for essential public assistant benefits and services such as healthcare services, social security or social services. These services target people that are dependent on those services and in a vulnerable position in relation to the authorities; the denial of these services may have a significant impact on those persons' livelihood or fundamental rights.

  (ii) to evaluate the creditworthiness of individuals or establish their credit score, as such systems may there again replicate historical patterns of discrimination and, as a result, discriminate in a biased way individuals' access to financial resources or essential services (such as housing or telecommunication). Systems meant to detect financial fraud and for prudential purposes should however not be considered high-risk.

  (iii) To evaluate and classify emergency calls (police, firefighters, medical aid) as such triage leads to decisions in very critical situations for the life and health of the concerned individuals.

  (iv) To be used for risk assessment and pricing in the case of life and health insurance, as the resulting decisions may have significant impact on persons' livelihood and could infringe their fundamental rights.

- *Law enforcement*. The use of such systems in a legal enforcement context may lead to a significant power imbalance and adverse impacts on fundamental rights, notably if the system is not trained with high quality data, as accuracy, reliability and transparence are particularly important in that field. As a result, the following use cases by law enforcement authorities are considered as high risk:

  (i) to assess the risk of a natural person to become a victim of criminal offenses;

  (ii) as polygraphs (more commonly referred to as lie detector) or similar tools;

  (iii) to evaluate the reliability of evidence in the course of investigation or prosecution of criminal offenses;

  (iv) to assess the risk of a natural person of offending or re-offending not solely based on profiling of natural persons (a phrase which, by the way, we find pretty hard to understand), or to assess personality traits or past criminal behavior of natural persons or groups;

  (v) for profiling natural persons in the course of detection, investigation or prosecution of criminal offenses.

  The preamble however provides that the use of such systems for administrative proceedings by tax or custom authorities as well as financial units in accordance with anti-money laundering legislation to prevent, detect, investigate and prosecute criminal offenses should not be classified as high risk.

- *Migration, asylum and border control management*. The use of such tools in this context may affect people who are in a vulnerable position and who are dependent on the outcome of the actions by authorities. Accuracy, non-discrimination and transparency are therefore key parameters to guarantee the respect of fundamental rights (notably their rights to free movement, private life and international protection). In addition to the use of such systems as polygraphs similarly to such use by law enforcement authorities in general, the following use cases are considered high-risk:

  (i) to assess a security risk, a risk of irregular immigration or a health risk posed by someone willing to enter into the territory of a Member State;

  (ii) to assist authorities to examine applications for asylum, visa and residence permits as well as to examine complaints related to such eligibility, including assessment of the reliability of evidence;

  (iii) to detect, recognize or identify natural persons in the context of migration, asylum and control border management, with the exception of verification of travel documents.

  The implementation of such systems will in any case have to comply with Directive 2013/32/EU and Regulation (EC) 810/2009, and not circumvent international obligations set out under the Convention of 28 July 1951 related to the Status of Refugees as amended by the Protocol of 31 January 1967.

- *Administration of justice and democratic processes.* Taking into account their potential significant impact on democracy,

rule of law and individual freedoms, the use of such systems in the judiciary to research and interpret facts and the law, as well as to apply the law to a concrete set of facts, has to qualify as high-risk to address the risk of biases, errors and opacity. While the preamble retains that the use of such systems can support the decision-making power of judges, it should not replace them, as the final decision-making must remain a human-driven activity and decision.

The same goes with regards to systems meant to influence the outcome of an election or the voting behavior of natural persons in the exercise of their vote; provided, however, that the use of such tools to organize, optimize and structure political campaigns shall not be considered as high-risk.

The fact for a model to be open source does not exempt it from having to comply with the requirements set out in the EU AI Act when it qualifies as a high-risk system.

The list in Annex III may be amended either by adding a new use-case in any of the areas listed in Annex III, or removing some then high risk system that would not pose any significant risk to fundamental rights, health or safety. To qualify as "high risk" or be removed from the list, the Commission shall take into account several factors such as the purpose, whether the system at stake has already been used, the nature and amount of data processed (in particular if special categories are at stake), autonomy of the system, respectively extent of human control, notified harms or related documentation submitted to national authorities, potential impact of such harm (severity and quantity of individuals potentially affected), technical means to rectify a wrongful output, etc.

III. Exemption

In the final round of negotiations, a compromise was found to provide some exemptions to ensure that some AI systems, also potentially covered by Annex III, would not be considered as high risk, bearing in mind that such an exemption does not exist for systems covered by Annex II.

The Commission has considered that there may be specific cases where such systems do not lead to a significant risk of harm to the legal interests protected under those areas, because they do not materially influence the decision-making or do not harm those interests substantially.

Such happens to be the case when:

- the system is intended to perform a narrow procedural task (such as to transform unstructured data into structured ones or classifying documents into categories);
- the task performed by the system is limited to improving the result of a previously completed activity that may be relevant for the purpose of the use case listed in Annex III; in such cases, the AI system only provides an additional layer of human activity (for instance improve the language used in a previously drafted document);
- the system is intended to detect decision-making patterns or deviations from prior patterns, but is not meant to replace or influence the previously completed human assessment without proper human review (such as grading pattern of a teacher to flag potential inconsistencies or anomalies);
- the system is meant to perform a preparatory task to an assessment relevant for use cases listed in Annex III, thus making the potential impact of the output very low in terms of representing a risk for the assessment to follow (for instance smart solutions for file handling that would include various functions such as indexing, searching, etc.).

The Commission may amend these criteria or add new ones, but should always provide concrete and reliable evidence of its *rationale* for doing so, as the overall level of protection of health, safety and fundamental rights should not be compromised.

It will be up to the provider that considers that its system is not high-risk to (i) document its assessment and provide it to the national competent authorities upon request and (ii) still register it in the database as set out in Art. 51(1a).

After having consulted with the AI Board, the Commission shall within 18 months after the entry into force of the EU AI Act, provides guidelines specifying the practical implementation of the Art. 6 by a comprehensive list of practical examples of high risk and non-high risk use cases.

Notwithstanding the above, an AI system that performs profiling of natural persons shall always be considered high-risk.

Having come to an end with regards to the classification of the AI systems under the EU AI Act, we shall turn next week to the requirements these high-risk systems have to comply with under the Regulation.

Source : https://www.wg-avocats.ch/en/actualites/intellectual-property/the-eu-ai-act-3-high-risk-ai-systems/