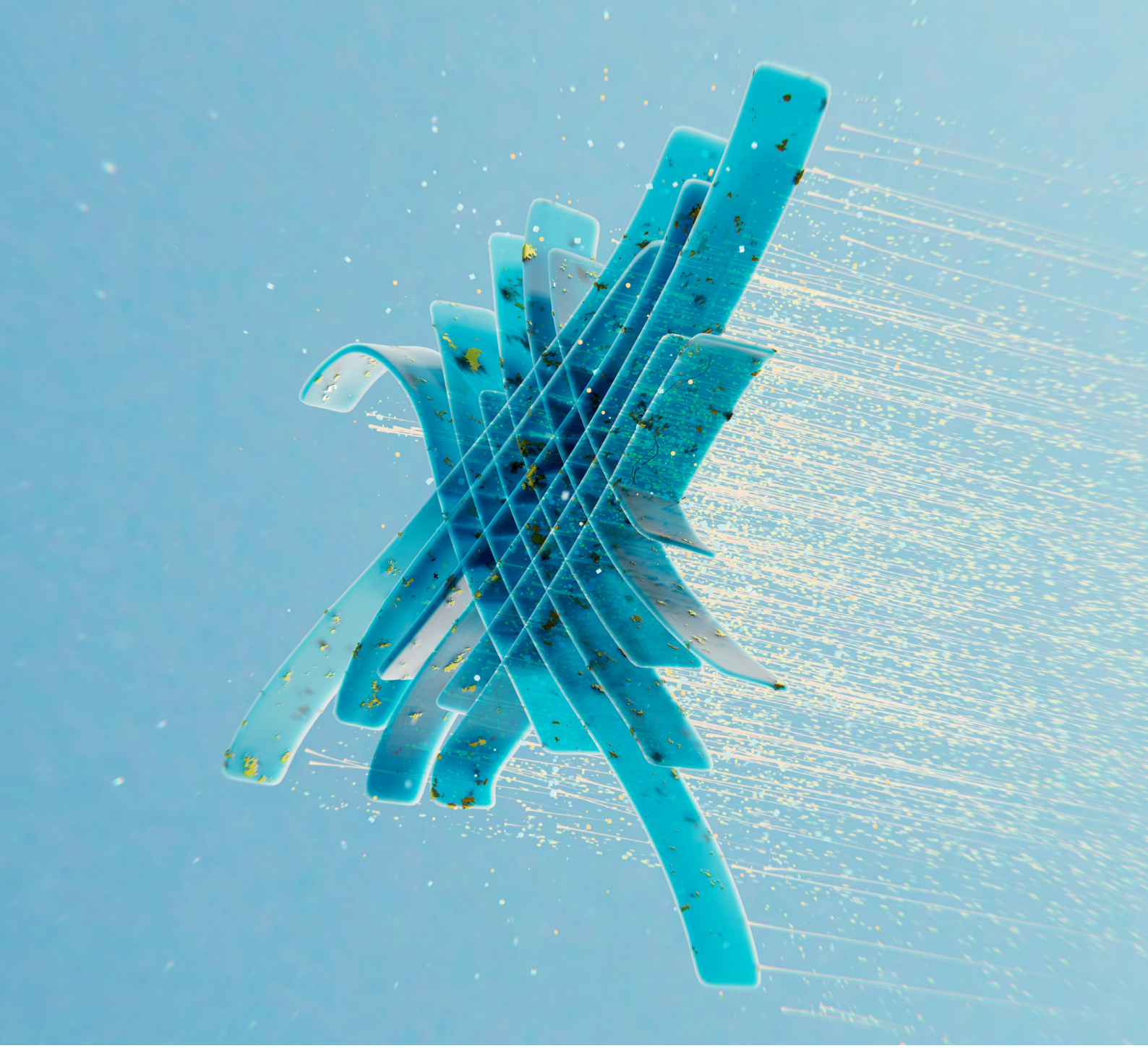


Assessing High-Risk AI Systems Under the EU AI Act

A Practical Four-Step Guide for General Counsels

By Ong Johnson, Lo Khai Yi
and Jerrine Gan

October 2024

HHQ
ADVOCATES & SOLICITORS

Identifying High-Risk AI Systems: A Step-by-Step Guide for General Counsels

In our previous article, *"8 Prohibited AI Practices Under the EU AI Act You Must Know: Are Your AI Systems at Risk?"*, we provided an in-depth discussion on the types of AI practices that will be prohibited under the EU AI Act, which is set to be implemented on 2 February 2025. We were also privileged to be invited by BFM to further explore some of these prohibited AI practices, and if you are interested, you can listen to the full discussion here: https://www.bfm.my/podcast/enterprise/enterprise-biz-bytes/prohibited-ai-practices

Building on our previous discussion of prohibited AI practices, this article will now explore another critical element of the EU AI Act, which is high-risk AI systems. The EU AI Act adopts a risk-based approach to regulating AI, classifying AI systems into several tiers: prohibited AI practices, high-risk AI systems, and General-Purpose AI Models, with lower or minimal risk. Unlike prohibited AI systems, which are outright banned, high-risk AI systems are even more relevant to most organizations, as they are often embedded within various AI applications.

Given the complexity and extensive regulatory requirements associated with high-risk AI systems, this article aims to provide general counsels with a practical guide on how to conduct an internal assessment of whether your AI systems fall under the high-risk category. Once classified as high-risk, there are specific obligations that organizations need to fulfill, depending on whether they act as providers, deployers, importers, or distributors of these AI systems.

Step 1: Is the AI System Covered by the Union Harmonization Legislation?

The first step in assessing whether an AI system is considered high-risk is to verify if it falls within the scope of the Union harmonisation legislation listed in Annex I of the EU AI Act. This legislation covers specific product categories, and AI systems that are either (i) one of these products or (ii) intended to be used as a safety component in these products which are subject to stricter regulatory oversight.

1. Machinery
2. Toys
3. Watercraft
4. Lifts
5. Explosives
6. Radio equipment
7. Pressure equipment
8. Cableways
9. Personal protective equipment
10. Gas appliances
11. Medical devices
12. Civil aviation
13. Vehicles
14. Marine equipment
15. Rail systems

Step 2:
Assess the Need for Third-Party Conformity Assessment

If your AI system is either a product or a safety component in one of the products listed above, the next key question is whether that product is required to undergo a third-party conformity assessment before being placed on the market, as mandated by EU harmonisation legislation.

If the answer is yes, then the AI system is automatically classified as a high-risk AI system.

If the AI system is not a product listed in Annex I or does not serve as a safety component for such products, or if the product does not require a third-party conformity assessment, your assessment does not end there, and you should move to step 3.

Step 3:
Review Annex III for Additional High-Risk AI Systems

Even if the AI system does not meet the criteria from Steps 1 and 2, it may still be categorized as high-risk if it falls under one of the AI systems listed in Annex III of the EU AI Act.

Annex III outlines 8 types of high-risk AI systems:

1. **Biometrics AI Systems** – These include biometric AI systems that are not prohibited. It encompasses: (i) remote biometric identification systems (excluding biometric AI systems used solely for biometric verification and authentication purposes, such as unlocking devices or granting access to premises); (ii) AI systems used for biometric categorisation according to sensitive attributes or characteristics; or (iii) emotion recognition AI systems.

2. **Critical Infrastructure AI Systems** – These are AI systems intended for use as safety components in the management and operation of critical digital infrastructure, road traffic, or in the supply of water, gas, heating, or electricity. Critical infrastructure AI systems are considered high-risk because their failure or malfunction may endanger the life and health of individuals on a large scale and lead to significant disruptions in the normal conduct of social and economic activities. Examples of safety components of such critical infrastructure may include systems for monitoring water pressure or fire alarm control systems in cloud computing centres.

3. **Education and Vocational Training AI Systems** – These AI systems are intended to: (i) determine access or admission to education or vocational training institutions; (ii) evaluate learning outcomes; (iii) assess the level of education an individual will receive or be able to access; or (iv) monitor and detect prohibited behaviour of students during tests. These are categorised as high-risk AI systems because they may influence the educational and professional trajectory of a person's life, affecting their ability to secure a livelihood. When improperly designed or used, such systems can be particularly intrusive and may violate the right to education and training.

4. **Employment AI Systems** – This includes AI systems used: (i) in the recruitment process, such as targeted job advertisements, job application filtering, and candidate evaluation; or (ii) to make work-related evaluations that affect employment relationships, such as promotion, termination, task allocation, or performance assessment. These systems are considered high-risk because they can significantly impact future career prospects, livelihoods, and workers' rights.

5. **Private and Public Service and Benefit Access AI Systems** – This includes AI systems intended to: (i) evaluate the eligibility of individuals for public assistance benefits and services; (ii) assess a person's creditworthiness; (iii) perform risk assessment and pricing for life and health insurance; or (iv) evaluate and classify emergency calls or establish priorities in dispatching emergency first-response services. These systems are classified as high-risk because they are often used by individuals in vulnerable positions, dependent on these benefits and services.

6. **Law Enforcement AI Systems** – These AI systems are intended for: (i) assessing the risk of a person becoming a victim of criminal offenses; (ii) being used as polygraph tools; (iii) evaluating the reliability of evidence during investigations or prosecutions; (iv) assessing the risk of a person committing or re-committing a crime, based on factors other than personal profiling, such as personality traits or past criminal behaviour; or (v) profiling individuals in the detection, investigation, or prosecution of criminal offenses. These systems are high-risk because they may unjustly single out individuals in a discriminatory or otherwise incorrect manner. Furthermore, their use could undermine important procedural rights, such as the right to an effective remedy, a fair trial, the right of defence, and the presumption of innocence.

7. **Immigration AI Systems** – These AI systems are used in migration, asylum, and border control management for: (i) polygraphs or similar tools; (ii) assessing risks, such as security risks, irregular migration risks, or health risks of individuals; (iii) assisting authorities in examining applications for asylum, visas, or residence permits, including eligibility assessments and evaluating the reliability of evidence; or (iv) detecting, recognising, or identifying individuals in migration or border management processes, excluding travel document verification. These systems are classified as high-risk because they affect individuals in particularly vulnerable situations who depend on the decisions of competent public authorities. The accuracy, non-discriminatory nature, and transparency of these AI systems are especially crucial to ensure respect for the fundamental rights of affected persons, including free movement, non-discrimination, privacy, international protection, and good administration.

8. **Administration of Justice and Democratic AI Systems –** This includes AI systems intended for: (i) use by or on behalf of judicial authorities to research and interpret facts and law, and apply the law to those facts, or for similar use in alternative dispute resolution; or (ii) influencing the outcome of elections or referenda, or the voting behaviour of individuals exercising their right to vote. These systems are high-risk because of their potential impact on the fundamental processes of justice and democracy.

If your AI system falls into one of these categories, it is automatically classified as high-risk. However, this does not conclude the assessment. You should proceed to step 4, where it is essential to evaluate whether the rebuttal assumption is applicable.

Step 4:
Assessing the Impact on Health, Safety, and Fundamental Rights

The 8 types of AI systems referred to in Annex III are generally categorized as high-risk. However, this presumption can be rebutted if the AI systems do not pose a significant risk of harm to the health, safety, or fundamental rights of individuals, including not materially influencing decision-making outcomes. The following four exceptions provide a basis for rebuttal:

1. The AI system is intended to perform a narrow procedural task, such as transforming unstructured data into structured data, classifying incoming documents into categories, or detecting duplicates among a large number of applications

2. The AI system is intended to improve the result of a previously completed human activity, such as AI systems intended to improve the language used in previously drafted documents, for example, improving professional tone, academic style, or aligning text with certain brand messaging.

3. The AI system is intended to detect decision-making patterns from prior decision-making instances and is not meant to replace or influence the previously completed human assessment, without proper human review.

4. The AI system is intended to perform preparatory tasks for assessments relevant to the use cases listed in Annex III, such as smart solutions for file handling, which may include various functions like indexing, searching, text and speech processing, or linking data to other data sources.

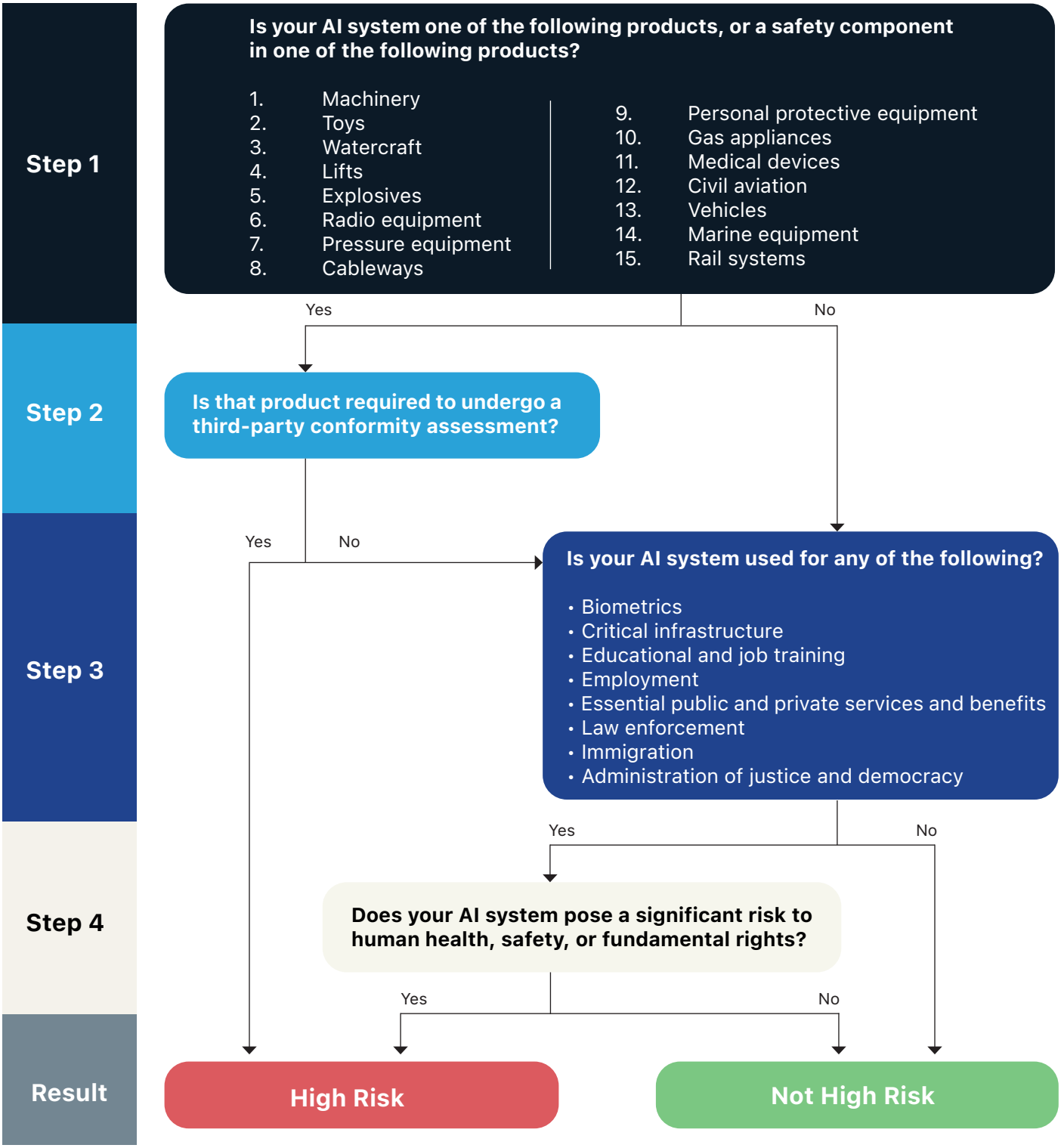
If the AI systems referenced in Annex III can demonstrate that they do not pose a significant risk of harm to health, safety, or fundamental rights—specifically, that they do not materially influence decision-making outcomes—by satisfying the aforementioned exceptions, they will not be classified as high-risk AI systems.

Conversely, if these AI systems fail to meet the criteria for rebuttal, they will be considered high-risk.

Flowchart for Simplifying the Internal Assessment Process

We understand that determining whether an AI system is classified as high-risk under the EU AI Act is a complex undertaking that necessitates a structured and methodical approach. Therefore, to aid general counsels in conducting a preliminary internal self-assessment of their AI systems, we have developed a visual flowchart below.

This flowchart below serves as a practical internal guide for evaluating whether an AI system falls into the high-risk category – while this tool is beneficial for initial assessments, it is important to note that it may not encompass the full depth of a comprehensive legal audit.



Core Requirements for High-Risk AI Systems

Once an AI system is classified as "high-risk" under the EU AI Act, such AI systems must comply with seven core requirements specified in the legislation. These requirements encompass critical areas such as:

(i) risk management systems
(ii) data governance
(iii) technical documentation
(iv) record keeping
(v) transparency
(vi) human oversight
(vii) accuracy, robustness, and cybersecurity

Given the length of this article, we will delve into each of these seven core requirements and the corresponding legal obligations for various stakeholders within the AI value chain—whether you are a provider, deployer, distributor, manufacturer, or importer of a high-risk AI system—in subsequent articles. Understanding these distinct compliance measures is essential for ensuring that your organization meets its obligations under the EU AI Act.

Implementation Timeline and Penalties

The EU AI Act officially came into force on 1 August 2024; however, its implementation will occur in stages. As outlined in our previous article, the enforcement of prohibited AI practices will commence on 2 February 2025, while the requirements for high-risk AI systems will take effect on 2 August 2026. This extended timeline for high-risk AI systems reflects the significant obligations that organizations will need to meet.

It is important to recognize that non-compliance with these high-risk AI systems can result in substantial penalties, including fines of up to €15 million or 3% of global revenue. Therefore, organizations must prepare adequately to ensure compliance and mitigate risks associated with these new regulations.

The Technology Practice Group at Halim Hong & Quek is well-versed in technology law, including the EU AI Act, and we are currently providing training to multinational corporations in Malaysia on this subject. Should you require assistance or wish to schedule a more detailed discussion to ensure compliance, please let us know.

Key Authors

Ong Johnson
Partner | Head of Technology Practice Group
johnson.ong@hhq.com.my

Lo Khai Yi
Partner | Co-Head of Technology Practice Group
ky.lo@hhq.com.my

Jerrine Gan
Pupil-in-Chambers | Technology Practice Group