Contents lists available at ScienceDirect

Computer Law & Security Review: The International Journal of Technology Law and Practice

journal homepage: www.elsevier.com/locate/clsr

Rebuilding the pyramid: The AI Act's risk-based approach using a binary decision diagram

Gustavo Gil Gasiola

Karlsruhe Institute of Technology, Germany

ARTICLE INFO

Keywords:
AI Act
Artificial intelligence
Risk regulation
Risk-based approach
High-risk AI systems
Prohibited AI practices
Minimal-risk AI systems
Binary decision diagram

ABSTRACT

The risk-based approach of the AI Act (AIA) results in a complex normative structure, in which the applicable subset of rules for a specific AI system is determined by the general scope of application and the classification of the system into particular risk levels. A pyramid of risks, a widely accepted explanation of the risk-based approach proposed by the European Commission, fails to provide a comprehensive classification process and does not accurately reflect the risk levels (either directly or indirectly) recognized in the AIA or the relation between classification criteria. This paper proposes a corrective solution to rebuild the pyramid of risks. Given that each AI system must be classified into one risk level and the AIA assigns a specific subset of rules to each risk level, an adaptation of the Commission's risk levels was necessary. Two types of exceptions are included in the list of prohibited AI practices, which significantly impact the classification process. The exception *stricto sensu* (in a strict sense) is the result of a balancing of interests, whereas the exception *lato sensu* (in a broader sense) is due to the absence of excessive regulatory risks. The transparency requirements, identified by the pyramid as a "limited-risk level," operate in parallel with the risk-based approach and do not constitute an independent risk level. Furthermore, as the AIA assigns a specific subset of rules to AI systems used in critical areas that do not pose significant risks, it is necessary to recognize a separate risk level (non-high risk). By analyzing the pyramid of risks, this study suggests representing the classification process as a binary decision diagram. This ensures that the risk-based approach is clearly defined and can help regulators and regulatees classify AI systems in accordance with the AIA.

1. Introduction

The recently enacted EU AI Act (AIA)[1] is based on a "clearly defined risk-based approach."[2] This regulation employs a complex normative structure, comprising three primary components: (1) a general scope of application, (2) classification criteria for AI systems, and (3) subsets of rules applicable to each risk level. The specific subset of rules applied to a given AI system is determined based on a two-step analysis. First, it is necessary to assess whether the system falls within the legal definition of an AI system specified in Art. 3(1) AIA. If the system is not to be subsumed under the legal definition of an AI system, then the regulation does not apply and no further analysis is necessary. Conversely, if it is an AI system, the second step is to classify the AI system or practice[3] into one of the risk levels defined by the AIA. As the risk levels serve as thresholds for a specific subset of rules,[4] the classification of the AI system will determine the applicable rules.

Given that the applicable subset of rules depends on classification of the AI system into a risk level, it is considerably important to ensure clarity regarding the existing risk levels and the respective classification criteria, particularly considering the international influence of the EU

E-mail address: gustavo.gasiola@kit.edu.

[1] Regulation (EU) 2024/1689 of the European Parliament and of the Council of June 13, 2024.

[2] Recital 26 AIA.

[3] The AI Act is not consistent in its distinction between AI systems and practices. For the purposes of this paper, the term "AI system" will be used to refer to both.

[4] Tobias Mahler, 'Between Risk Management and Proportionality: The Risk-Based Approach in the EU's Artificial Intelligence Act Proposal' (2020) Nordic Yearbook of Law and Informatics 247.

https://doi.org/10.1016/j.clsr.2025.106189

Available online 30 August 2025

law.[5] The risk levels are used to tailor the legal requirements to properly address the regulatory risks.[6] Meanwhile, the classification criteria identify the AI systems that present these regulatory risks. Consequently, uncertainty regarding the risk levels or the respective classification criteria can lead to erroneous or unsatisfactory compliance with the AIA.[7] For the regulator, the lack of compliance by the regulatee results in underachievement of the regulatory objectives as the regulatory risks are not properly addressed. For the regulatee, the erroneous classification of their AI systems could lead to either responsibility, including financial consequences,[8] or compliance with an unnecessarily stricter subset of rules.

Despite the relevance of risk levels and classification criteria for the efficient operationalization of the risk-based approach and for the proportionality of the legislative measure, the AIA currently lacks the required clarity.[9] Except for the high-risk AI systems,[10] the regulation does not explicitly define the risk levels. Consequently, there are doubts regarding the risk levels (implicitly) created by the AIA. Furthermore, the classification criteria and respective exceptions suggest a complex dynamics between the different risk levels, which may allow for conflicting interpretations or even loopholes. As a result, the AIA presents a considerable challenge for regulators and regulatees in their task of classifying AI systems and identifying the applicable subset of rules.

A widespread explanation of the risk-based approach of the AIA is provided by the European Commission in the form of a pyramid of criticality[11] or risk (Fig. 1). This visual schema indicates the existence of four risk levels (unacceptable risk, high risk, limited risk, and minimal risk) hierarchically arranged from the top (more regulatory risks) to the bottom (less regulatory risks). The subset of rules for each risk level follows this hierarchy, with the most restrictive rules (prohibition) applied to the highest risk level (unacceptable risk) and the least restrictive rules (no mandatory requirements) applied to the lowest risk level (minimal risk).[12]

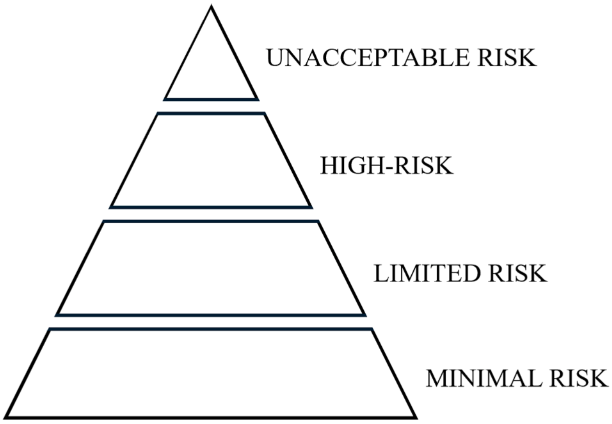

Fig. 1. Pyramid of risks.

Visual tools can contribute to the general understanding of legal frameworks with high levels of complexity,[13] as in the case of the AIA,[14] thereby assisting in compliance.[15] It is therefore not particularly surprising that with only a few critics,[16] the literature[17] has until now relied on this proposed systematization, including the graphical reproduction of the pyramid.[18] The novel nature of the regulation, which is virtually unparalleled elsewhere in the world, and the complexity and length of the legal text have contributed to this scheme becoming a credible explanation of the AIA's risk-based approach.

Nevertheless, the pyramid of risk oversimplifies and distorts the risk-based approach. By assuming that each AI system must be classified at one risk level and this classification determines the applicable subset of rules, the pyramid includes rules that are not actually assigned to a risk level while not mentioning all the risk levels implicit in the AIA. Moreover, the schema fails to represent important nuances and dynamics in the classification criteria, in particular regarding exceptions and declassification. Consequently, there is an urgent need to address the widely reproduced explanation of the risk-based approach and rethink the pyramid in a way that is more in line with the requirements of the AIA.

This study examines the risk-based approach of the AIA and proposes a binary decision diagram that clearly represents the classification assessment and risk levels. By identifying the preconditions for a clearly

[5] Delaram Golpayegani, Harshvardhan J Pandit and Dave Lewis, 'To Be High-Risk, or Not To Be—Semantic Specifications and Implications of the AI Act's High-Risk AI Applications and Harmonised Standards,' *2023 ACM Conference on Fairness, Accountability, and Transparency* (ACM 2023) 905 <https://dl.acm.org/doi/10.1145/3593013.3594050> accessed 4 June 2025.

[6] Mahler (n 4) 248.

[7] Raphaël Gellert, 'The Role of the Risk-Based Approach in the General Data Protection Regulation and in the European Commission's Proposed Artificial Intelligence Act: Business as Usual?' (2021) 3 Journal of Ethics and Legal Technologies 21.

[8] The applicable fines for non-compliance are subject to the risk-based approach. For instance, see Art. 99(3) AIA for the unacceptable risk level and Art. 99(4) AIA for the high-risk level.

[9] Golpayegani, Pandit and Lewis (n 5) 906.

[10] Art. 6 AIA.

[11] Johanna Chamberlain, 'The Risk-Based Approach of the European Union's Proposed Artificial Intelligence Regulation: Some Comments from a Tort Law Perspective' (2023) 14 European Journal of Risk Regulation 1.

[12] Based on European Commission, 'AI Act' <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai> accessed 4 Juni 2025.

[13] Dirk Burkhardt and Kawa Nazemi, 'Visualizing Law - A Norm-Graph Visualization Approach Based on Semantic Legal Data,' *Proceedings of the International Conference of the Virtual and Augmented Reality in Education. The 4th International Conference of the Virtual and Augmented Reality in Education* (VARE2018) 154.

[14] Martin Kretschmer and others, 'The Risks of Risk-Based AI Regulation: Taking Liability Seriously' (2023) <http://arxiv.org/abs/2311.14684> accessed 4 June 2025; Jakob Mökander and others, 'The US Algorithmic Accountability Act of 2022 vs. The EU Artificial Intelligence Act: What Can They Learn from Each Other?' (2022) 32 Minds and Machines 752.

[15] Kevin D Ashley, *Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age* (Cambridge University Press 2017) 59.

[16] Isabelle Hupont and others, 'Use Case Cards: A Use Case Reporting Framework Inspired by the European AI Act' (2024) 26 Ethics and Information Technology 3; Vera Lúcia Raposo, 'Ex Machina: Preliminary Critical Assessment of the European Draft Act on Artificial Intelligence' (2022) 30 International Journal of Law and Information Technology 92.

[17] Golpayegani, Pandit and Lewis (n 5) 906; Raposo (n 16) 91-92; Giovanni De Gregorio and Pietro Dunn, 'The European Risk-Based Approaches: Connecting Constitutional Dots in the Digital Age' (2022) 59 Common Market Law Review 489; Keri Grieman and Joseph Early, 'A Risk-Based Approach to AI Regulation: System Categorisation and Explainable AI Practices' (2023) 20 SCRIPTed 66; Gianclaudio Malgieri and Frank Pasquale, 'Licensing High-Risk Artificial Intelligence: Toward Ex Ante Justification for a Disruptive Technology' (2024) 52 Computer Law & Security Review 7; Jérôme De Cooman, 'Humpty Dumpty and High-Risk AI Systems: The Ratione Materiae Dimension of the Proposal for an EU Artificial Intelligence Act' (2022) 1 Market and Competition Law Review 52.

[18] Kretschmer and others (n 14) 4; Peter Burgstaller, 'The Use of AI and Its Legal Boundaries in the EU,' *2023 1st International Conference on Optimization Techniques for Learning (ICOTL)* (IEEE 2023); Georg Stettinger, Patrick Weissensteiner and Siddartha Khastgir, 'Trustworthiness Assurance Assessment for High-Risk AI-Based Systems' (2024) 12 IEEE Access 22718; Richard Ryan, Saba Al-Rubaye and Graham Braithwaite, 'UTM Regulatory Concerns with Machine Learning and Artificial Intelligence,' *2022 IEEE/AIAA 41st Digital Avionics Systems Conference (DASC)* (IEEE 2022).

defined risk-based approach, this study critically examines the AIA and the explanation proposed by the European Commission. While the definition of an AI system and the specific requirements assigned to each risk level are not the focus of this study, references are provided to facilitate a more thorough analysis of each risk level and the respective classification criteria. Accordingly, the framework proposed in this paper offers a valuable tool for regulators and regulatees in classifying AI systems based on risk levels.

After brief introduction in this first section, the second section analyzes the rationale behind the risk-based approach proposed by the European Commission and its relation to legislative proportionality. In the third section, the risk-based approach is examined from the perspective of its normative structure. Given that the risk levels interact with the scope of application to define the rules applicable to an AI system, it is possible to identify the necessary conditions for well-defined risk levels and classification criteria. In the fourth section, the pyramid of risks is analyzed and a binary decision diagram for the classification of AI systems is developed. Finally, discussion on the proposed diagram and conclusion are presented in the fifth section.

2. Legislative proportionality and risk regulation

The AIA aims to improve the European internal market, including the support for innovation and the protection of fundamental rights regarding the deployment of AI systems.[19] The legal basis for the AIA can be found in Art. 114 of the Treaty on the Functioning of the European Union (TFEU) (the proper functioning of the internal market) and Art. 16(2) TFEU (competence to legislate about the protection of personal data), although the fundamental rights protected by the AIA are not limited[20] by the protection of personal data from Art. 8 Charter of Fundamental Rights of the European Union (Charter). As a legislative measure, the AIA is subject to the principle of proportionality[21] according to Art. 52(1) Charter and Art. 5(4) Treaty on European Union and "may be made only if they are necessary and genuinely meet objectives of general interest recognized by the Union or the need to protect the rights and freedoms of others."

Accordingly, the legislative proportionality underpinned the discussions surrounding the AIA during the drafting process. The European Commission repeatedly expressed its concerns regarding the necessity of establishing an "appropriate legal framework,"[22] considering the specific context in which AI systems are applied and an "impact-based approach."[23] In its White Paper, the Commission advocated a risk-based approach to regulation, emphasizing the need for it to be "effective to achieve its objectives while not being excessively prescriptive so that it could create a disproportionate burden."[24] Finally, the impact assessment accompanying the AIA draft justified the option for "a horizontal EU legislative instrument (…) following a proportionate risk-based approach" in accordance with the principle of proportionality.[25]

In addition to the general requirement for proportional legislative measures, the AIA had to address the challenges posed by the definition of AI systems.[26] The AIA partially[27] adopted the OECD's concept of the AI system, which is quite broad and covers several applications[28] that may, but do not necessarily, pose a threat to health, safety, or fundamental rights. Without a definition that adequately distinguishes the applications that pose regulatory risks, the scope of application based solely on this definition would not ensure legislative proportionality.[29] As argued in the impact assessment, it would be disproportionate to impose burdens, i.e., mandatory requirements, on all systems that fall within the legal definition, regardless of the risk posed by each application.[30]

This risk-based approach[31] is therefore the cornerstone of the AIA as it adjusts and improves the scope of application. The tiered framework, operationalized by risk levels and classification criteria, ensures the necessary legislative proportionality.[32] In particular, the risk-based approach allows the legal framework to be organized according to an abstract[33] risk assessment, with the proportional legislative measure assigned to the AI system in accordance with the risk it poses. This approach thus provides a calibration mechanism[34] for the scope of application, avoiding disproportionate burdens on regulatees.

Risk-based regulation is not entirely new in the EU law, particularly in the implementation of the Digital Single Market Strategy.[35] Previous EU legislative acts have already operationalized risks to structure the application of specific rules. While the New Legislative Framework (NLF)[36] addresses the risks related to product safety,[37] the General Data

[19] Art. 1(1) AIA.

[20] Cf. Art. 1(1) and Recitals 1, 6, 8, 28, 58, 59, 60 and 176 AIA.

[21] As developed by the jurisprudence of the CJEU, the principle of proportionality has the traditional function of protecting liberal values against the excessive exercise of public power. See *Hauer v Rheinland-Pfalz* [1979] ECJ Case 44/79, ECLI:EU:C:1979:290, para 23; and *Nold v Commission* [1974] ECJ Case 4/73, ECLI:EU:C:1974:51, para 14. See also Robert Schütze, 'EU Competences' in Anthony Arnull and Damian Chalmers (eds), *The Oxford Handbook of European Union Law* (Oxford University Press 2015) 96-97.

[22] European Commission, 'Artificial Intelligence for Europe' (2018) COM/2018/237 final <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2018:237:FIN> accessed 4 Juni 2025.

[23] European Commission, 'Building Trust in Human-Centric Artificial Intelligence' (2019) COM/2019/168 final <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52019DC0168> accessed 4 Juni 2025.

[24] European Commission, 'White Paper on Artificial Intelligence' (2020) COM/2020/65 final 268 <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52020DC0065> accessed 4 Juni 2025.

[25] European Commission, 'Impact Assessment Accompanying the 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' (2021) SWD/2021/84 final 48 < https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021SC0084> accessed 4 Juni 2025.

[26] Mahler (n 4) 267; Miriam C Buiten, 'Towards Intelligent Regulation of Artificial Intelligence' (2019) 10 European Journal of Risk Regulation 43-45.

[27] Luca Bertuzzi, 'EU Lawmakers Set to Settle on OECD Definition for Artificial Intelligence' (2023) EURACTIV <https://www.euractiv.com/section/artificial-intelligence/news/eu-lawmakers-set-to-settle-on-oecd-definition-for-artificial-intelligence/> accessed 4 Juni 2025.

[28] Kees Stuurman and Eric Lachaud, 'Regulating AI. A Label to Complete the Proposed Act on Artificial Intelligence' (2022) 44 Computer Law & Security Review 3.

[29] Jonas Schuett, 'Defining the Scope of AI Regulations' (2023) 15 Law, Innovation and Technology 70.

[30] European Commission (n 25) 85.

[31] ibid 85. As described in the Impact Assessment: "a regulatory framework for high-risk AI applications with the possibility for all non-high-risk AI applications to follow a code of conduct".

[32] Mahler (n 4) 249; Kretschmer and others (n 14) 6; Henry Fraser and José-Miguel Bello Y Villarino, 'Acceptable Risks in Europe's Proposed AI Act: Reasonableness and Other Principles for Deciding How Much Risk Management Is Enough' (2023) European Journal of Risk Regulation 4.

[33] Mahler (n 4) 254.

[34] De Gregorio and Dunn (n 17) 475.

[35] Mahler (n 4) 258; De Gregorio and Dunn (n 17) 476.

[36] The NLF consists of a package of measures, including Regulation (EC) No 765/2008 of the European Parliament and of the Council of 9 July 2008 setting out the requirements for accreditation and market surveillance relating to the marketing of products and repealing Regulation (EEC) No 339/93, Decision No 768/2008/EC of the European Parliament and of the Council of 9 July 2008 on a common framework for the marketing of products, and repealing Council Decision 93/465/EEC, and Regulation (EU) 2019/1020 of the European Parliament and of the Council of 20 June 2019 on market surveillance and compliance of products and amending Directive 2004/42/EC and Regulations (EC) No 765/2008 and (EU) No 305/2011.

Protection Regulation[38] and its data protection impact assessment[39] focus on risks to fundamental rights and are based on the principle of accountability.[40] Another example is the Digital Services Act (DSA),[41] where the regulation of platforms is tiered according to different categories of regulatees.[42] Despite the different regulatory concerns and approaches, the concept of risk is operationalized in each of these acts to ensure a proportionate relationship between legislative intervention and regulatory risks.[43]

The AIA integrates the regulatory concerns of product safety[44] and fundamental rights into a "clearly defined risk-based approach."[45] According to the intensity of the regulatory risks, AI systems are classified into predetermined risk levels, which determine the applicable subset of rules. Owing to the convergence of regulatory concerns and the top–-down[46] and centralized[47] categorization of risk levels, the AIA differs significantly from the structures of other legislative acts and therefore requires a systematic analysis to understand its normative structure and functioning.

The literature on AI regulation has highlighted the potential of the risk-based approach to achieve legislative proportionality and strike the difficult balance between the protection of fundamental rights and freedoms and the protection of economic freedoms.[48] However, substantial criticism has cast doubt on the effectiveness of the AI Act's risk-based approach in effectively addressing risks and safeguarding fundamental rights. In particular, it has been claimed that the classification criteria for the unacceptable and high risk levels are too narrow to cover all cases in which an AI system could threaten or even violate fundamental rights.[49] This is in part due to the lack of general prohibitions or objective classification criteria[50] that do not rely on lists of unacceptable AI practices (Art. 5 AIA) or high risk AI systems (in particular, Annexes II and III AIA). Given that minimal risk AI systems are not subject to mandatory requirements, with the exception of a few rules of horizontal application, such as the AI literacy requirement under Art. 4 AIA, the AIA would not establish a minimum or essential level of protection that applies to all AI systems.[51] This is particularly evident due to the lacks substantive rights[52] or an individual complaint mechanism[53] for those affected by AI systems. Moreover, if the AIA is regarded as maximum harmonisation by the EU to regulate AI systems, its pre-emptive effect could hinder national initiatives to fill this alleged regulatory gap.[54]

The next section examines the normative structure of the AIA and analyzes the implications of the risk-based approach. The introduction of a system of risk levels and subsets of rules poses interpretative challenges that could run counter to the intended legislative proportionality. Against this background, we identify conditions for the proper operationalization of the classification process in a binary decision diagram to ensure the consistency of the risk-based approach.

3. Normative structure of the risk-based approach

Legal instruments explicitly or implicitly define their scope of application, i.e., the factual situations (including material, personal, territorial and temporal dimensions) to which their set of rules applies. In turn, each rule can determine legal consequences given one or more rule conditions.[55] The scope of application and the rule conditions reflect the intention of the legislator to regulate a particular situation. The legal consequences, as intended by the legislator, occur only if 1) the concrete factual situation falls within the scope of application, considering all four dimensions, and 2) all the rule conditions are met.

The risk-based approach introduces risk levels as thresholds for the application of rules.[56] The classification of an AI system into one of the risk levels triggers the application of a specific subset of rules. In turn, this subset of rules determines its specific legal consequences given the rule conditions. For instance, only when the AI system is classified as high risk, the rules assigned to this risk level apply. The risk-based approach clusters the applicable rules according to the scope of application, following predefined risk levels. Consequently, the classification of the AI system into a risk level is necessary to determine the relevant rules and the legal consequences.

Consider a provider who has developed a new system and intends to place it on the EU internal market. The desired legal consequence is that the AI system is lawfully placed on the market. Once the system in question can be considered an "AI system,"[57] and the situation can be subsumed under the scope of application ("placing on the market… AI systems… in the Union"[58]), the AIA applies. Consequently, the lawfulness of the placement of the system on the market will be determined by the AIA, without prejudice to other relevant laws. In the second step, the specific subset of rules that will determine the lawfulness of the AI

[38] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC.

[39] Mahler (n 4) 261; De Gregorio and Dunn (n 17) 481.

[40] De Gregorio and Dunn (n 17) 479, 482.

[41] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and amending Directive 2000/31/EC.

[42] De Gregorio and Dunn (n 17) 484.

[43] ibid 482.

[44] Golpayegani, Pandit and Lewis (n 5) 906; Burgstaller (n 18).

[45] Recital 26 AIA.

[46] De Gregorio and Dunn (n 17) 477.

[47] Gellert (n 7) 19.

[48] Mahler (n 4) 248; De Gregorio and Dunn (n 17) 477; Fraser and Bello Y Villarino (n 32) 4; Martin Ebers, 'Truly Risk-Based Regulation of Artificial Intelligence How to Implement the EU's AI Act' (2024) European Journal of Risk Regulation 1.

[49] Chamberlain (n 11) 6; Cooman (n 17) 65; Ebers (n 48) 7; BEUC, 'Regulating AI to Protect the Consumer' (2021) 2; Nathalie A Smuha and others, 'How the EU Can Achieve Legally Trustworthy AI: A Response to the European Commission's Proposal for an Artificial Intelligence Act' (2021) SSRN 17 <https://www.ssrn.com/abstract=3899991> accessed 4 Juni 2025; Lily Ballot Jones, Julia Thornton and Daswin De Silva, 'Limitations of Risk-Based Artificial Intelligence Regulation: A Structuration Theory Approach' (2025) 5 Discover Artificial Intelligence 14.

[50] Smuha and others (n 49) 18.

[51] Gellert (n 7); Ebers (n 48) 7.

[52] Martin Ebers and others, 'The European Commission's Proposal for an Artificial Intelligence Act—A Critical Assessment by Members of the Robotics and AI Law Society (RAILS)' (2021) 4 Multidisciplinary Scientific Journal 593; BEUC (n 49) 2; Smuha and others (n 49) 17.

[53] Smuha and others (n 49) 17; Riikka Koulu and others, 'Artificial Intelligence and the Law: Can We and Should We Regulate AI Systems?' in Bartosz Brożek, Olia Kanevskaia and Przemysław Pałka (eds), *Research Handbook on Law and Technology* (Edward Elgar Publishing 2023) 442.

[54] BEUC (n 49) 2; Koulu and others (n 53) 442; Michael Veale and Frederik Zuiderveen Borgesius, 'Demystifying the Draft EU Artificial Intelligence Act' (2021) 4 Computer Law Review International 108-109.

[55] Daniel Oberle and others, 'Engineering Compliant Software: Advising Developers by Automating Legal Reasoning' (2012) 9 SCRIPTed 291.

[56] Mahler (n 4) 247; Gellert (n 7) 19; Stuurman and Lachaud (n 28) 3.

[57] Art. 3(1) AIA

[58] Art. 2(1)(a) AIA.

system is identified through its classification at a risk level.[59] If the system is classified as high risk,[60] the rules assigned to this risk level set the conditions for the desired legal consequence so in this instance the newly developed AI system may only be lawfully placed on the market if the conditions imposed by the subset of rules applicable to high-risk systems are met.

The scope of application and the risks classification share similarities.[61] While the scope of application defines the general applicability of the AIA to systems or practices, the risk classification identifies the subset of rules from the AIA that applies to the specific system or practice. Both rules are, albeit at a different level, thresholds for the application of other rules. Nevertheless, there is a fundamental difference between them since the scope of application operates on a bivalent logic. The application of the AIA to a given system is either true or false, depending on the criteria established in Art. 2 AIA. In contrast, the risk classification requires that each system or practice that falls within the scope of application be classified into one of the risk levels as defined by the AIA.[62] The question is not whether the AI system could be classified into one of the risk levels, but rather, at which risk level it should be classified.

Significant implications arise from this distinction. Whereas an ill-defined scope of application may include situations that do not present regulatory risks or exclude others with relevant regulatory risks, the classification of risks deals with the assignment of the proper risk level to systems and practices according to the risk they pose. To function consistently, all AI systems (in accordance with the definition from Art. 3(1) AIA) must be classified but this requirement does not preclude the possibility of a residual category (such as the minimal-risk level), which includes all systems and practices that have not been classified at higher levels. Accordingly, the risk-based approach presupposes well-defined risk levels and clear classification criteria.[63] The assignment of a risk level to an AI system implies the applicability of a subset of rules that determine the possible legal consequence. Consequently, different AI systems assigned to the same risk level are all subject to the same legal framework.

The regulatee must be able to assess their own system to identify the appropriate risk level and the corresponding subset of rules that are adjudged applicable. Uncertainty surrounding the classification, similarly to the scope of application, presents a significant challenge in determining the applicable rules. When the risk level of a system is ambiguous or the classification criteria lack a clear definition, the regulation can impose a disproportional burden on the regulatee. Given that the system is subject to the AIA, yet the applicable rules remain undefined, the regulatee confronted with issues related to the classification would have to either comply with the stricter requirements (to avoid the risk of non-compliance), or assume this risk and the corresponding negative consequences related to non-compliance.

The classification criteria for assigning a risk level may rely on a more or less abstract risk assessment.[64] An abstract risk assessment is typically carried out by the legislator or the regulator, who identifies and manages risks considering general aspects of the system. As a generalization, an abstract risk assessment cannot consider all the features and capabilities of a specific system or even the specificities of the context of the application. Conversely, a more concrete analysis could consider these specificities, but at a lower level of abstraction. Whereas the risk analysis carried out by the legislator is typically abstract, also as a matter of equality requirement for issuing laws, the risk analysis required by law and carried out by regulatees is concrete, that is, considering the specific aspects of the regulatees' products or services. The AIA combines these two types of risk analysis for the classification criteria. Although the risk levels are defined explicitly or implicitly throughout the regulation, the classification of AI systems or practices is based on either an abstract risk assessment carried out by the regulator or a concrete risk assessment carried out by the regulatee. In theory, the combination of self- and hetero-regulation—found in the classification criteria for unacceptable- and high-risk levels, as discussed in the next section—guarantees further legislative proportionality, as it could adjust eventual mismatches deriving from the abstract risk assessment. However, it can also increase the complexity of the risk-based approach, making it more opaque (especially for the regulator), thus hampering the achievement of regulatory goals.

These considerations are key to analyzing the pyramid of risks suggested by the Commission, which is explored further in the next section. In particular, it is necessary to investigate each of the risk levels and its corresponding classification criteria according to the normative structure of the risk-based approach to verify both its coherence and the types of risk analysis employed or implied.

4. Deconstructing the pyramid

In October 2019, the German Data Ethics Commission (GDEC) recommended the adoption of a "risk-adapted regulatory approach"[65] to regulate algorithmic systems. This recommendation included a pyramid of risks with five risk levels[66] and respective measures to address each level. According to the recommendation, the classification of an algorithmic system would be assessed by the potential for harm, including the likelihood that harm will occur and the severity of that harm.[67]

The GDEC's approach to the regulation of algorithmic systems may have inspired the Commission to adopt both the AIA's risk-based approach and its representation through a pyramid of risks. Although the Commission does not directly attribute the inspiration, the five-level risk-based system from GDEC is mentioned in its impact assessment.[68] Furthermore, the study prepared for the Commission to support the impact assessment even reproduces the GDEC's pyramid in its entirety.[69]

It is noteworthy that international legal instruments have also developed approaches to AI regulation based on the notion of risk. Art. 16 of the Council of Europe Convention on AI requires its members states to implement a "risk management approach" (Art. 16), which is also

[59] Veale and Borgesius argue that this structure results in an "unusual misalignment between the target of its substantive obligations (primarily high-risk systems) and its material scope (all AI systems)". See Veale and Borgesius (n 54) 109. Although the broad material scope of application is problematic regarding the harmonization and residual competence for member states, the misalignment mentioned can be understood actually as a feature of the risk-based approach to ensure legislative proportionality.

[60] Art. 6 AIA.

[61] Schuett (n 29) 70.

[62] From a technical perspective, Grieman and Early claim that the division of the categories "must be universally applicable to all AI applications" Grieman and Early (n 17) 69.

[63] "[T]here is a pressing need for the categorisation to be as clear as possible as soon as possible, so that businesses can predict whether their systems will be heavily regulated or not regulated at all and adapt their planning for the coming years" Chamberlain (n 11) 5.

[64] About this differentiation, see Mahler (n 4) 253.

[65] German Data Ethics Commission (GDEC), 'Opinion of the Data Ethics Commission' (2018) 20.

[66] The GDEC defines the risk levels as follows: Level 1 – "Applications with zero or negligible potential for harm"; Level 2 – "Applications with some potential for harm"; Level 3 – "Application with regular or significant potential for harm"; Level 4 – "Applications with serious potential for harm"; Level 5 – "Applications with an untenable potential for harm" ibid 19.

[67] ibid 18.

[68] European Commission (n 25) 26.

[69] Andrea Renda and others, 'Study to Support an Impact Assessment of Regulatory Requirements for Artificial Intelligence in Europe: Final Report (D5)' (2021) 85.

mentioned in the OECD AI Principles.[70] Moreover, the United Nations Global Digital Compact[71] and the Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI Systems[72] explicitly call for a risk-based approach to AI regulation. Despite the evident focus on addressing the risks posed by AI systems, which may have similarly informed the Commission, these international legal instruments lack a more concrete implementation of the intended risk-based approach.

The publication of the Commission's pyramid of risk on its website initially served the purpose of explaining the new proposed legal instrument to regulate AI systems and during the legislative process, the proposal underwent substantial modification,[73] while maintaining the structure of the risk-based approach. Following the approval of the final version of the AIA, the Commission continued reproducing the pyramid of risk to explain the regulation.[74] It can, therefore, be reasonably argued that, according to the Commission, the pyramid of risks reproduces the AIA's risk-based approach faithfully.

Against this backdrop, the following section analyzes the Commission's pyramid of risk critically. Each risk level indicated in the pyramid (unacceptable risk, high risk, limited risk, and minimal risk) is assessed according to the prerequisites for a coherent risk-based approach, as outlined in Section 3. In particular, the purpose of this inquiry is to ascertain the extent to which the risk levels are well defined and the classification criteria and exceptions are identified. This critical analysis of the pyramid leads to the proposal of a corrective solution, which is presented in the form of a binary decision diagram. This solution is developed in this section and then discussed further in Section 5.

4.1. Unacceptable-risk level

At the top of the pyramid lies the unacceptable-risk level. AI practices classified at this level present an excessive[75] or unbearable[76] risk to health, security or fundamental rights[77] and, therefore, should be prohibited.[78] Based on an abstract assessment of the risks, these practices are perceived by the AIA as posing a risk that, *a priori*[79], outweighs the potential benefits.[80] From the perspective of the precautionary principle,[81] the identified (or uncertain) risks of these new technologies or practices could justify their prohibition to guarantee fundamental rights and freedoms.[82]

Although Recital 26 AIA mentions "certain unacceptable AI practices," the AIA does not identify this risk level explicitly.[83] Art. 5 AIA simply lists the AI practices[84] that ought to be banned, without mentioning or naming the risk level. Thus, the unacceptable-risk level and the respective classification criteria are intertwined with the rule defining the legal consequence, that is, prohibition. While describing the AI practices assigned to the unacceptable-risk level, the regulation already establishes the prohibition itself.

It is unclear whether Art. 5 AIA truly creates a genuine risk level since the subset of rules assigned to unacceptable AI practices only covers the prohibition of these AI systems. Analyzing Art. 5 AIA within the normative structure of the risk-based approach may appear to be an overcomplication, as the traditional normative structure fits the prohibition rule from Art. 5 AIA perfectly. Nevertheless, the identification of the unacceptable-risk level is justified by the need for the consistency of the risk-based approach. As all AI systems that fall within the scope of application must be classified at a risk level, it is necessary to identify the highest tier that covers cases posing excessive risks, even though the only legal consequence intended is their prohibition. Without the unacceptable-risk level, the risk-based approach loses consistency, as the prohibited AI systems would be excluded from the classification schema.

Given the existence of this risk level, the classification criteria are found in the list of AI systems from Art. 5 AIA. This list describes the more or less abstract characteristics of AI systems that have been identified as posing an excessive risk for health, security, or fundamental rights.

As the list of prohibited AI systems does not specify which technology should be banned, but rather, the different aspects, usages, and purposes[85] that could pose unacceptable risks, the classification criteria can be understood as having a technology-neutral manner. For example, the AIA assigns an unacceptable-risk level to systems that manipulate human behavior harmfully[86] or that exploit the vulnerability of humans.[87] This approach ensures that new AI systems or practices would still be classified as unacceptable AI practices if they fall within the descriptions of at least one[88] of the paragraphs of Art. 5 AIA. However, as the list of prohibitions is exhaustive, i.e., it is not possible to add or remove AI systems without the ordinary legislative procedure, there will still be gaps in the description of AI systems that could pose unacceptable risks[89] from the perspective of the more concrete risk assessment.

While the technologically neutral approach outlined in Art. 5 AIA could be considered partially future-proofed,[90] it also increases the complexity of the classification assessment. Under Art. 5 AIA, it is necessary to ascertain whether the features described, which indicate an excessive risk, are actually present in a specific AI system. This is not a straightforward task, as the description of features and the concepts used

[70] OECD, 'Recommendation of the Council on Artificial Intelligence' (2019).

[71] United Nations, 'Global Digital Compact' (2024) 12.

[72] G7 Japan, 'The Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI Systems' (2023) 1 <https://www.soumu.go.jp/hiroshimaaiprocess/pdf/document05_en.pdf> accessed 4 Juni 2025.

[73] The final version of the AI Act introduced several modifications to the Commission's initial proposal. These modifications included changes to the governance structure, the introduction of new rules for general-purpose AI models, and the revision of the criteria for classifying AI systems as high risk. See Paul Friedl and Gustavo Gil Gasiola, 'Examining the EU's Artificial Intelligence Act' (2024) <https://verfassungsblog.de/examining-the-eus-artificial-intelligence-act/> accessed 4 Juni 2025.

[74] The most recent update to the Commission's website, which presents the pyramid of risk, was published on 3 June 2025. This update was made following the approval of the final version of the AIA. Cf. European Commission (n 12).

[75] Mahler (n 4) 250; Burgstaller (n 18) 1.

[76] Malgieri and Pasquale (n 17) 7.

[77] As Recital 28 AI Act states, these AI practices "contradict Union values of respect for human dignity, freedom, equality, democracy and the rule of law and fundamental rights enshrined in the Charter, including the right to non-discrimination, to data protection and to privacy and the rights of the child."

[78] "The following AI practices shall be prohibited," Art. 5(1) AIA. See Recital 26 AIA.

[79] De Gregorio and Dunn (n 17) 490.

[80] Chamberlain (n 11) 8.

[81] Fraser and Bello Y Villarino (n 32) 11.

[82] Chamberlain (n 11) 5.

[83] Mahler (n 4) 250.

[84] It is interesting to note that the AI Act speaks of AI practices in this category and not of AI systems, as in the case of high-risk AI systems. This may indicate that it is the practice and not the system that should be banned. However, the distinction is unclear. See Mahler (n 4) 266.

[85] Rostam J Neuwirth, 'Prohibited Artificial Intelligence Practices in the Proposed EU Artificial Intelligence Act (AIA)' (2023) 48 Computer Law & Security Review 12.

[86] Art. 5(1)(a) AIA.

[87] Art. 5(1)(b) AIA.

[88] The characteristics described in Art. 5 AIA can overlap, whereby a specific AI system may present several prohibited characteristics. Cf. Neuwirth (n 86) 7, 12.

[89] Chamberlain (n 11) 6.

[90] European Commission (n 12).

have an open texture and are subject to interpretation.[91] For instance, the extent to which the manipulation of human behavior is to be considered harmful[92] or when there is the exploitation of human vulnerabilities[93] are issues that can best be answered on a case-by-case basis. The AIA thus engenders numerous gray areas concerning the classification of unacceptable AI practices and empowers the regulatee to assess the specific AI system according to the list from Art. 5 AIA.[94] Moreover, this indicates that the abstract risk assessment carried out by the regulator must be, in some cases, complemented by the assessment carried out by the regulatee (for example, to identify whether the practice is harmful) to proceed with the classification.

The list of prohibited AI systems comprises some exceptions. The AIA exceptionally permits the deployment of prohibited AI systems under specified conditions. This is the case for the use of real-time remote biometric identification systems in public spaces for law enforcement.[95] Despite the general prohibition of this practice, it can be deployed if it is strictly necessary for the purposes specified in Annex II AIA, such as preventing terrorist attacks. The exceptions introduce an additional step in the classification process. It is necessary to ascertain whether an AI system falls within the prohibited list and the specific use is deemed excepted and, if so, under which conditions.

If an AI system is exempted from the prohibition list (e.g., the use of real-time biometric identification systems in public spaces to prevent terrorist attacks), the system is not classified into the unacceptable-risk level. However, the AIA does not clearly determine how this exempted AI system ought to be classified. There are two possible outcomes due to the exception. First, the exempted AI system could be automatically classified at the risk level directly below the unacceptable-risk level, i.e., the high-risk level. Second, the AI system would not be automatically classified as high risk; rather, it would have to undergo further assessment following the classification process.

These divergent outcomes that may be observed depend on the type of exception created by Art. 5 AIA. When the AIA excludes an AI system from the prohibition list, it does so based on two distinct grounds. The first is an exception *stricto sensu* (or in a strict sense), which results from a balancing of interests and assumes the existence of excessive risks. The second is an exception *lato sensu* (or in a broader sense), which is due to the absence of excessive regulatory risks without denying the existence of any risk. Given the fundamental differences between these two type of exceptions, the distinction between *stricto sensu* and *lato sensu* is a key element in the interpretation of the classification process of the AIA.

The exception *stricto sensu* allows the deployment of AI systems that would otherwise be classified as unacceptable. In this case, the excepted AI system still poses excessive risks to health, security, or fundamental rights, yet it is justified through a balancing of interests. For instance, the general prohibition on the deployment of AI systems to infer emotions of natural persons in the areas of workplace and educational institutions is excepted when applied for medical or safety reasons.[96] Under strict conditions, in the case mentioned, related to the purpose (for medical or safety reasons) of the use, the benefits of the excepted AI system could outweigh the excessive risks posed. The existence of the exception *stricto sensu* does not mean that this AI system does not present excessive risks. Rather, the exception merely indicates that the benefits of this use for medical and safety purposes outweigh the risks identified.

As the exception is due to a balancing of interests, the abstract risk assessment is not affected. The excepted AI practice continues to pose excessive risks, as the harmful properties described remain present, and it is only permitted for the realization of other relevant values. Given that the risks remain excessive, yet justified, it is reasonable to directly assign, that is, without further assessment, the excepted AI systems to the high-risk level. Although the balancing of interests justifies the removal of the prohibition, it does not justify the exception of any further requirements that address the present risks. It is therefore essential to recognize this classification criterion for the high-risk level to ensure the coherence of the risk-based approach and thereby avoid significant gaps in the regulatory framework.

By contrast, the exception *lato sensu* (in a broader sense) allows the deployment of AI systems not due to a balancing of interests but due to a lack of excessive risks. This type of exception is associated with the description of the prohibited practices as both determine the cases of excessive risks. For instance, Art. 5(1)d AIA prohibits an AI system for the prediction of criminal offenses "based solely on the profiling of a natural person or on assessing their personality traits and characteristics." Other AI systems used for the same purpose, but based on different approaches, would not be deemed to pose an unacceptable risk. The exception *lato sensu* identifies the AI systems that do not pose excessive risks. The justification for this exception differs from that of the exception *stricto sensu*: while the latter assumes a balancing of interests, the former determines which AI practices do not pose excessive risks. The exception *lato sensu* affects the abstract risk assessment as the AI systems excepted from the unacceptable-risk level do not, *a priori*, pose excessive risks to health, security, and fundamental rights. Nevertheless, the exception as a delimitation of the description does not allow us to conclude that the AI practice does not pose any significant risks. It is possible that an excepted AI practice could be classified as high risk as well. In contrast to the exception in *stricto sensu*, as the excessive risks cannot be assumed in the exception *lato sensu*, it is necessary to continue the classification assessment to determine whether the AI system should be classified as high risk or minimal risk.

Fig. 2 presents the classification process of an AI system into the unacceptable-risk level in the form of a binary decision diagram. Given an AI system, it is necessary to assess whether the application can be subsumed in the prohibited list outlined in Art. 5 AIA. If an AI system is prescribed in Art. 5 AIA and there is no exception, the AI system is classified in the unacceptable-risk level and is prohibited. Conversely, when an AI system is allowed under specific conditions (i.e., an exception *stricto sensu*), it is classified at the high-risk level. When a specific AI system is not subsumed within the prohibited list, further assessment is required to determine the appropriate risk level.

As the exception *lato sensu* is similar to, and sometimes confused with, the description of the prohibited AI practice, an alternative scheme could imply checking for this type of exception (second question, "excepted AI system?") together with the first question ("prohibited AI practice?"). In this alternative, the third question ("exception *stricto sensu*?") would be unnecessary as the only exception left would be the exception *stricto sensu*. Although less concise, the proposed diagram explicitly asks to assess the nature of the exception and distinguishes between cases of *prima facie* non-subsumption and the exception *lato sensu*. This distinction is not trivial. For example, while some AI systems do not exploit the vulnerabilities of people (not subsumed under Art. 5 (1)(b) AIA), other AI systems exploit them, but not in a way that causes significant harm (exception *lato sensu*). Even if both systems are not considered to pose an unacceptable risk, the assessment is performed at a different level

4.2. *High-risk level*

The subsequent level of the pyramid is the high-risk level. AI systems classified in this level present a significant harmful impact on health,

[91] Veale and Borgesius (n 54) 99-101.

[92] Art. 5(1) AIA.

[93] Art. 5(2) AIA.

[94] De Gregorio and Dunn claim that the list set by the Art. 5 AIA "is independent of any a posteriori risk assessment by providers or users of those systems" De Gregorio and Dunn (n 17) 492. However, the classification of any AI system into the conditions of Art. 5 AIA may also require a risk assessment, as in the case of the harmfulness of the application.

[95] Art. 5(1)(h) AIA.

[96] Art. 5(1)(f) AIA. Other exception *stricto sensu* can be found in Art. 5(1)(d), (g), and (h) AIA.

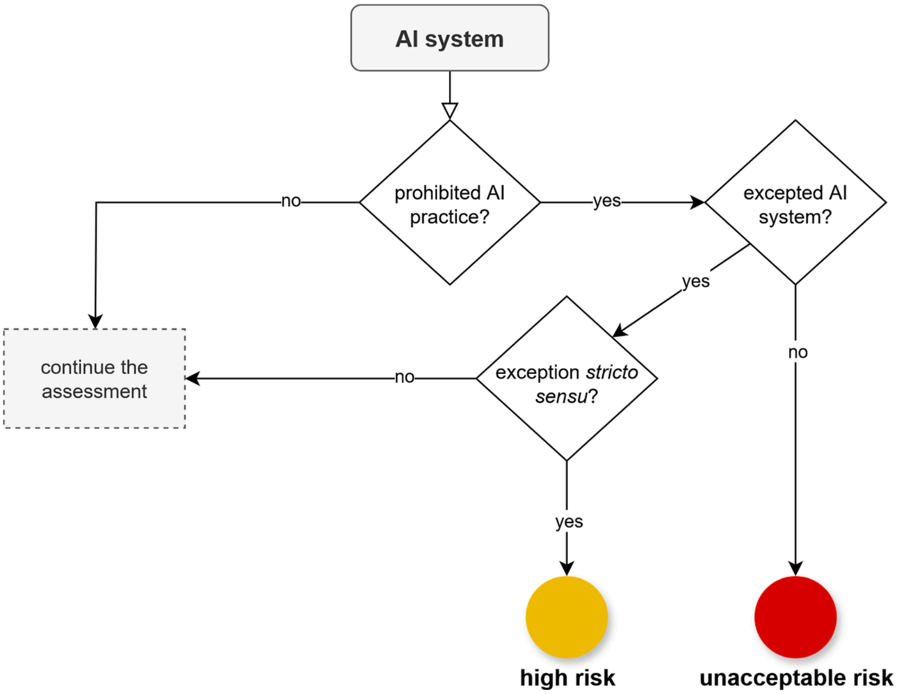

Fig. 2. Classification of AI systems at the unacceptable-risk level.

safety, or fundamental rights.[97] Nevertheless, risks could be managed to an acceptable level through the mandatory requirements and the implementation of a risk management system (Art. 9 AIA).[98] In contrast to the unacceptable-risk level, in which the AI systems are presumed to violate Union values, the AI systems classified in the high-risk level only pose a potential threat to those values.[99]

To classify an AI system at this level, the AIA establishes two criteria, as outlined in Art. 6 AIA. The first criterion relates to the NLF,[100] a set of rules regulating product safety in the EU.[101] According to Art. 6(1) AIA, an AI system is considered high risk if two conditions are met: 1) the AI system is intended to be used as a safety component of a product[102] or the AI system itself is a product covered by Union harmonization legislation, as listed in Annexes I and II ;[103] and 2) this product is subject — by the harmonization legislation — to third-party conformity assessment to be placed on the market or put into service.[104] Both conditions are necessary for the AI system to be classified as high risk based on the first criterion.

The second criterion is the use of an AI system in the critical areas listed in Annex III,[105] which lists eight critical areas and indicates the uses that should be considered as high risk. These critical areas are predefined by the AIA, which identifies domains and specific uses within those domains that could pose risks to health, safety, or fundamental rights. As an example, while point 7 of Annex III of the AIA defines the administration of justice as a critical area, point 7(a) specifies the use of an AI system to assist judicial authorities in interpreting the law. The description of intended use entails, in general, the purpose, AI capability, deployer, and subject.[106] For instance, while employment, management of workers, and access to self-employment are listed as critical domains, the specified uses are the recruitment or selection of natural persons[107] and the decision-making that affects the terms of work-related relationships.[108] Annex III identifies high-risk AI systems from the combination of critical areas and specific use cases within those domains.

This criterion was the subject of considerable debate throughout the legislative process. In particular, the possibility of exempting or declassifying high-risk AI systems based on a concrete risk assessment underwent several revisions before the final version was adopted. The Commission’s original proposal did not include any exception for the classification of AI systems used in critical areas.[109] If an AI system was to be deployed in critical areas following the specified use cases, it would be classified as high risk, without any consideration of the context or the capabilities of the system. The initial proposal assumed that the abstract risk assessment to draft Annex III, which did not consider the concrete application or context, was sufficient to define AI systems as high risk. The regulatee would only need to determine whether its AI system falls within the domains and use the cases identified in Annex III and, if so, the AI system would be classified as high risk and subject to the relevant requirements. This proposal did not permit the regulatee to contest the classification through a concrete risk assessment.

[97] Cf. Recitals 7 and 46 AIA.

[98] See De Gregorio and Dunn (n 17) 491; Schuett (n 29) 4; Fraser and Bello Y Villarino (n 32) 3; Jonas Schuett, ‘Risk Management in the Artificial Intelligence Act’ (2023) European Journal of Risk Regulation 4.

[99] Chamberlain (n 11) 6.

[100] Recitals 9 and 46 AIA.

[101] The NLF encompasses a broad array of products, including machinery (Directive 2006/42/EC), medical devices (Regulation (EU) 2017/745) and motor vehicles (Regulation (EU) 2018/858 and Regulation (EU) 2019/2144).

[102] According to Art. 3(14) AIA, safety component “means a component of a product or of an AI system which fulfils a safety function for that product or AI system, or the failure or malfunctioning of which endangers the health and safety of persons or property.”

[103] Art. 6(1)(a) AIA.

[104] Art. 6(1)(b) AIA.

[105] Art. 6(2) AIA.

[106] Golpayegani, Pandit and Lewis (n 5) 908-909.

[107] Annex III (4)(a) AIA.

[108] Annex III (4)(b) AIA

[109] European Commission, ‘Proposal for a Regulation of the European Parliament and of the Council Laying Down Hermonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts’ (2021) COM/2021/206 final Art. 6(2) and Annex III <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52021PC0206> accessed 4 Juni 2025.

Both the European Parliament (EP)[110] and the Council[111] proposed substantial adjustments to include a concrete risk assessment that would consider the context and specific use of the AI system. According to both proposals, a system would be considered high risk if two conditions are met. First, the AI system is intended to be used in the areas listed in Annex III, and second, the system in question would have to likely lead to significant risk to the health, safety, and fundamental rights.[112] Although the regulator would already assess the risk of use in the critical areas (the first condition), the second condition would consider the context of use and the capabilities of the AI system. In brief, this proposal included a two-step risk assessment to classify an AI system as high risk: the first is an abstract assessment and is carried out by the regulator and the second is a concrete assessment and is performed by the regulatee.

The final version of the AIA differs slightly from the amendments proposed by the EP and the Council; however, it also includes a concrete risk assessment. According to Art. 6(1) AIA, AI systems intended for use in the areas listed in Annex III are initially considered high risk. Nevertheless, Art. 6(3) AIA exempts AI systems that “do not pose a significant risk of harm to the health, safety or fundamental rights of natural persons, including by not significantly influencing the outcome of decision-making.” Furthermore, the article provides examples of situations in which the AI system may not pose significant risks. These include situations where “the AI system is intended to perform a narrow procedural task”[113] or where “the AI system is intended to improve the result of a previously completed human activity.”[114] Thus, as approved, Annex III provides predefined areas,[115] and AI systems should continue to be considered high risk unless the regulatee considers, based on a concrete risk assessment, that the risks are not significant. For classification purposes, there is no obligation for the regulatee to assess whether there are relevant risks of harm to the health, safety or fundamental rights. Given the presumption of significant risks associated with AI systems in critical areas, it is necessary for the regulatee to conduct an assessment if they intend to declassify their system from being high risk.

The introduction of this exemption mechanism based on a concrete risk assessment presents certain challenges for the classification process. If an AI system used in a predefined area does not pose a significant risk, it will not be classified as high risk. However, this exception does create a peculiar situation. Despite not being classified as high risk due to the absence of a significant risk, the AIA does subject this AI system to specific mandatory requirements. For example, the regulatee must document the risk assessment and make the assessment’s documentation available to the competent national authorities upon request.[116] Moreover, the AI system is also subject to registration in the EU database.[117] Due to these mandatory requirements, the AI system that has been declassified from high risk can also not be considered minimal risk (a residual risk level for which no mandatory requirement applies). As a result, the pyramid of risk is unable to properly classify the AI system excepted from the high-risk level.

To achieve coherence in the risk-based approach, all AI systems must be classified according to a specific risk level, which should correspond to a defined subset of rules. Therefore, the AI system that is exempt from the high-risk level constitutes an intermediary risk level, situated between high and minimal risk, as the AIA assigns a specific subset of rules to this level. At this intermediate risk level, the system is subject to some mandatory requirements, in contrast to the minimum risk level, but less strict than those already applied to high-risk AI systems. The identification of this intermediate risk level ensures a coherent risk-based approach, as non-high-risk AI systems are properly identified and duly distinguished from other risk levels.

Fig. 3 presents the binary decision diagram for the classification in the high-risk and non-high-risk levels, under Art. 6 AIA.

When the AI system is not covered by the prohibited AI practices list (Art. 5 AIA), the first criterion for the high-risk level (related to the NLF) must be assessed. The assessment order is not trivial. As the exception from Art. 6(2) AIA (no significant risks) does not apply to the first criterion, the AI system is classified as high risk when the two conditions (harmonization legislation and third-party conformity assessment) are met, without further assessment based on a concrete risk assessment. When the first criterion fails, it is necessary to assess whether the AI system is used in the predefined critical areas from Annex III, taking into account both the domain and the specific use listed. As discussed in the following sections, if the AI system is not used in the predefined areas, it is classified as minimal risk. Otherwise, the AI system is classified into one of two levels: high risk (with significant risks) or non-high risk (without significant risk).

Consider an AI system that determines which applicants are admitted to a university. This practice is not included in the list of prohibited AI practices (Art. 5 AIA). Moreover, this system is not a safety component or a product itself covered by the harmonization legislation. The next step is to check whether the assessment of university applicants falls within the critical areas listed in Annex III. According to point 3 of Annex III AIA, education and vocational training is identified as a critical area, and specifically when the “AI systems intended to be used to determine access or admission or to assign natural persons to educational and vocational training institutions at all levels" (point 3(a)). As the system falls within the critical area, it is then necessary to assess whether this use poses significant risks in order to conclude the classification as high risk or non high risk.

4.3. *Limited-risk level*

The pyramid of risks acknowledges the existence of a limited risk or specific transparency risk level, based on Art. 50 AIA. This risk level is generally recognized in the literature as part of the risk-based approach.[118] Systems classified at this level would pose moderate risks, that is, lower than the unacceptable- and the high-risk levels.[119] The subset of rules assigned to this risk level would cover transparency obligations to ensure that individuals[120] are aware of the applications and thereby strengthen trust in them.[121] Indeed, Art. 50 AIA identifies transparency risks of certain AI systems and addresses them with appropriate measures. For instance, the providers must disclose properly that the audio, image, or video content was generated by AI systems[122] to prevent the uncontrolled dissemination of AI-generated content.

However, in light of the preconditions identified for a coherent risk-

[110] European Parliament, ‘Amendments Adopted by the European Parliament on 14 June 2023 on the Proposal for a Regulation of the European Parliament and of the Council on Laying down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts’ (2023) P9_TA(2023)0236.

[111] Council, ‘Proposal for a Regulation of the European Parliament and of the Council Laying Down Hermonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts’ (2022) 2021/0106 (COD).

[112] European Parliament (n 111); Council (n 112).

[113] Art. 6(3)(a) AIA.

[114] Art. 6(3)(b) AIA.

[115] Recitals 52 and 53 AIA.

[116] Art. 6(4) AIA.

[117] Art. 6(4) and Art. 49(2) AIA.

[118] Golpayegani, Pandit and Lewis (n 5) 906; Raposo (n 16) 100; De Gregorio and Dunn (n 17) 491; Grieman and Early (n 17) 66; Malgieri and Pasquale (n 17) 7; Cooman (n 17) 52.

[119] Mahler (n 4) 251.

[120] Chamberlain (n 11) 8.

[121] ibid 7.

[122] Art. 50(4) AIA.

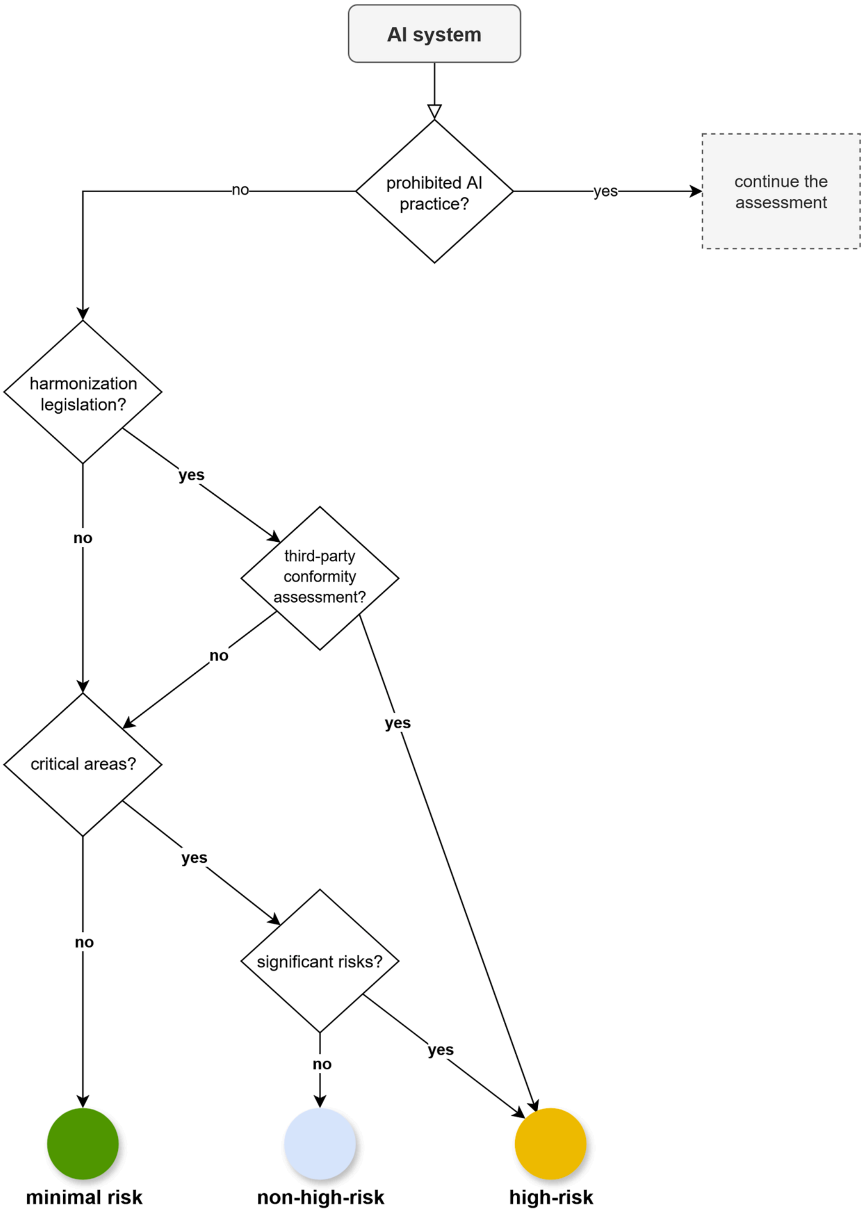

Fig. 3. Classification of AI systems at the high-risk level.

based regulation, Art. 50 AIA does not constitute an independent risk level. The transparency obligations apply not only to the AI systems classified as "limited risk" or a "specific transparency risk"[123] but to all situations described in the legal norm, including high-risk AI systems.[124] Therefore, the legal framework created by the AIA to address certain transparency risks constitutes a parallel regime quite apart from the classification of AI systems at a risk level. In other words, the transparency obligations operate according to the traditional normative structure (rule conditions → legal consequences) and do not employ the risk level to adjust the application of a subset of rules. Art. 50 AIA determines conditions for all AI systems, independently of their classification. For instance, if an AI system interacts directly with natural persons and there is a risk of manipulation, the provider (irrespective of the risk level of the system) must inform the natural person about the interaction with the machine.[125] Moreover, there is no subset of rules assigned to this alleged risk level — each transparency obligation is linked to a specific transparency risk and the regulatee is only subject to those relevant for their AI system. If the AI system interacts directly with a natural person but does not generate synthetic content, only Art. 50(1) AIA applies. As a result, there is no limited-risk level[126] or corresponding classification criteria, as Art. 50 AIA simply determines transparency obligations on all AI systems that meet the conditions set out in the various paragraphs of the article.[127]

[123] European Commission, 'Artificial Intelligence – Questions and Answers' <https://ec.europa.eu/commission/presscorner/detail/en/QANDA_21_1683> accessed 4 June 2025.

[124] Recital 132 AIA.

[125] Art. 50(1) AIA.

[126] Some authors also acknowledge that the transparency obligations are not part of the risk-based approach, see Burgstaller (n 18) 1-2.

[127] Therefore, the transparency obligations under Art. 50 AIA are part of the subset of rules from the AIA that are not affected by the risk classification. Another example of such rules is the obligation for providers and deployers to ensure the AI literacy of their staff and other persons dealing with the operation and use of AI systems (Art. 4 AIA).

The limited-risk level is not part of the risk-based approach, as Art. 50 AIA operates in parallel with the classification process.[128] As a result, the decision diagram does not take the transparency obligations into account.

4.4. *Minimal-risk level*

The next and last tier is the minimal-risk level. The AIA does not mention this level, nor does it provide any criteria for classification. However, the minimal-risk level is conceived of as a residual category[129] and covers all AI systems that are not classified as unacceptable risk, high risk, or non-high risk.[130] As a residual category, the classification criterion is precisely the absence of the conditions already established for the other risk levels and presumably, the majority of AI systems should fall into this category.[131] At this risk level, the AIA considered that the AI system has the potential to offer more benefits than the potential risks,[132] especially compared to the AI systems classified at the other levels,[133] or is simply unproblematic in terms of the scope of the regulation.[134] Accordingly, the AIA does not prescribe mandatory requirements for minimal-risk AI systems and merely encourages the adoption of codes of conduct and voluntary compliance with the requirements for high-risk AI systems.[135]

Although a substantial subset of rules assigned to this risk level is lacking (i.e., no mandatory requirements), the recognition of the minimal-risk level is necessary for the overall coherence of the risk-based approach. Given that all AI systems must be classified at a risk level, the minimal-risk level encompasses all AI systems that do not satisfy the other classification criteria. The minimal-risk level is, therefore, opposed to the unacceptable level, in that both include AI systems that are classified at the higher and lower extremes of risk.

The minimal-risk level is the result of an abstract risk assessment. As no significant risks were identified, the regulator decided not to submit this large group of AI systems to mandatory requirements. However, the classification criteria for a minimal-risk level do not provide any exceptions or corresponding procedures to consider the specific AI application in a more concrete risk assessment. If the classification of an AI system as minimal risk was erroneous (because the application indeed matched the criteria for unacceptable- or high-risk levels), or it was subject to substantial modification[136] (matching the criteria for a different classification), a reclassification is then possible.[137] The situation is less clear when a specific AI system poses significant risks to health, safety, or fundamental rights (identified through a concrete risk assessment), but it does not match the classification criteria for the other risk levels.[138] The classification of this AI system in the minimal-risk level appears misguided, as the AIA does not establish any obligation or mechanism to enable a concrete risk assessment of the minimal-risk AI system, taking into account the context and the AI capabilities.[139] In other words, there is no clear methodology for determining whether a particular minimal-risk AI system should be then classified at a different risk level given the specific context. This is of particular concern when the minimal-risk level is not subject to any minimum set of mandatory requirements, in contradiction to the precautionary principle.[140] In any case, these minimal-risk AI systems could still be subject to different regulatory frameworks.[141]

An AI system is classified as minimal risk if it does not match the classification criteria for the higher risk levels. As a residual category, the representation of the classification for the minimal-risk level corresponds to the complete decision diagram (Fig. 4) for the classification process according to the AIA.

The classification into the minimal-risk category is an outcome obtained via different paths. An AI system included in the prohibited AI practice list from Art. 5 AIA can still be classified as a minimal risk if the application is excepted (exception *lato sensu*) and it does not meet the criteria for the high-risk category. Other AI systems that are not used in the areas covered by the harmonization legislation (or are not subject to third-party conformity checks) and are not deployed in the critical areas according to Annex III are thereby classified as minimal risk.

The AIA does not provide a (re)classification of minimal-risk AI systems that present, in the concrete use case, significant risks. As the classification criteria for unacceptable and high risks are restricted to specific regulatory concerns (for example, any exploitation of vulnerabilities) or predefined domains, a reclassification of minimal-risk AI systems does not seem plausible. This is a significant loophole that should have been addressed by the AIA.

5. Conclusions: Rebuilding the pyramid as a binary decision diagram

Although visual tools can help laypeople and experts in comprehending new and complex regulations, oversimplification can convey inaccurate information. As knowledge of the applicable rules affects compliance, it is crucial to achieve a balance between simplification and schematization on the one hand and precision and depth on the other. In the case of the Commission's pyramid of risk, our analysis suggests that it does not accurately reflect the AIA. To redesign the pyramid accurately and comprehensively, the decision diagram (Fig. 4) presents a clearly defined classification process and a coherent picture of the risk-based approach established by the AIA.

The development of the decision diagram presented interpretative challenges. To ensure legal proportionality and efficiency, the risk-based approach must be coherent and non-ambiguous. If the applicable rules are determined by the classification of an AI system at a risk level, both the existing risk levels and the classification criteria must be unambiguous for the provider/deployer. Consequently, some adjustments are necessary. In particular, the limited-risk level as part of the risk-based approach and the consequences of classification exceptions require further consideration.

Despite the efforts to develop a coherent and non-ambiguous framework, it is noteworthy that the suggested diagram has limitations. The actual classification of AI systems and practices depends on the interpretation of each of the scenarios listed by the AIA to define the prohibited AI practices and the critical areas. The question of whether a particular AI system falls within one of the listed scenarios requires analysis of the specific use case and is not addressed in the diagram. Furthermore, the diagram omits additional steps needed to assess

[128] In addition to the transparency requirements, the rules regarding general-purpose AI models (Chapter V AIA), including the identification of systemic risks (Art. 51 AIA), also operate in parallel to the risk-based approach and the classification process.

[129] Cooman (n 17) 61.

[130] Chamberlain (n 11) 7.

[131] ibid; Cooman (n 17) 61; Initiative for Applied Artificial Intelligence, 'AI Act: Classification of AI Systems from a Practical Perspective' (2023) 24.

[132] Chamberlain (n 11) 8.

[133] De Gregorio and Dunn (n 17) 491.

[134] Mahler (n 4) 268.

[135] In accordance with Art. 95(1) AIA, the AI Office and the Member States shall encourage the drawing up of codes of conduct for voluntary application of specific requirements to "AI systems, other than high-risk AI systems."

[136] Art. 3(23) AIA.

[137] Golpayegani, Pandit and Lewis (n 5) 908.

[138] For example, in domains not considered in the Annex III. Cf. Malgieri and Pasquale (n 17) 7; Cooman (n 17) 65; Veale and Borgesius (n 54) 110; Johann Laux, Sandra Wachter and Brent Mittelstadt, 'Trustworthy Artificial Intelligence and the European Union AI Act: On the Conflation of Trustworthiness and Acceptability of Risk' (2024) 18 Regulation & Governance 26.

[139] Neuwirth (n 86) 12.

[140] Gellert (n 7) 30.

[141] Mahler (n 4) 252; Kretschmer and others (n 14) 5.

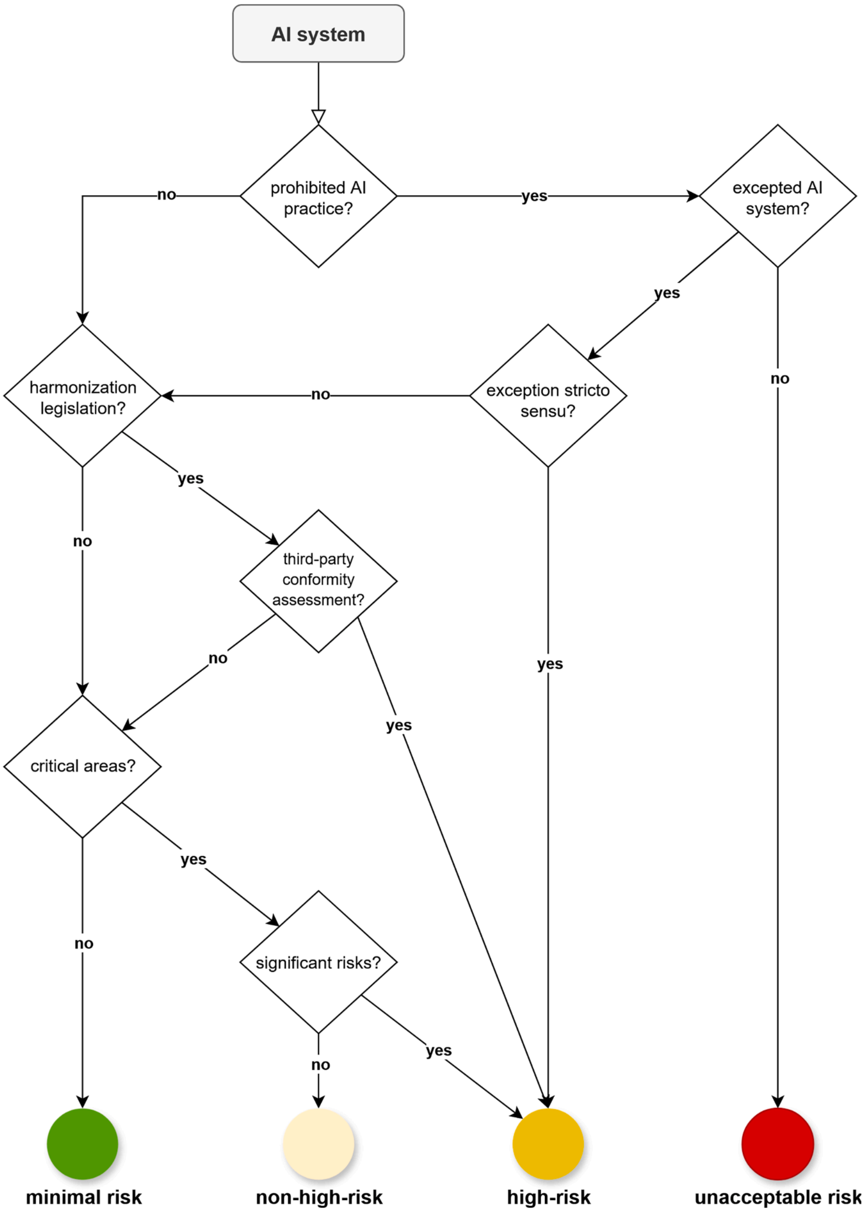

Fig. 4. Classification of AI systems.

significant risks when the AI system is employed within a critical areas. One potential solution to this issue would be the inclusion of the criteria enumerated in Art. 6(3) AIA. Nonetheless, given that the list of criteria is not exhaustive and allows for other justifications, this inclusion would result in a more complex diagram without enhancing its overall completeness.

The limited-risk level, as suggested in the Commission's pyramid of risks, corresponds to *de facto* transparency obligations[142] applicable to all systems that fall within the scope of application of the AIA, regardless of the classification.[143] Furthermore, it is unreasonable to consider the limited-risk level an additional classification of AI system (resulting in classifications such as "high risk + limited risk" or "minimal risk + limited risk"). As discussed in Section 3, the risk-based approach requires an adaptation of the traditional normative structure. The identification of the applicable subset of rules for a given AI system is determined by its classification at a risk level. The risk level is then embedded within the material scope of the application to determine the legal conditions of the desired legal consequence. This adaptation increases the complexity of the legal framework and should be considered only when it helps to clarify the normative structure. As this additional step is not necessary to describe the normative structure from Art. 50 AIA and its paragraphs, its interpretation as a parallel risk level is also inadequate and undermines the coherence of the classification system.

The analysis identified a legal loophole in the AIA regarding the classification exceptions. In fact, the pyramid of risk is not as immutable as it may appear. Although the AIA creates exceptions for the classification of AI systems at the unacceptable- and high-risk levels, the regulation does not clarify the consequences of being excepted from a risk level. This results in two main problems. First, applications that pose an excessive risk but are excepted from the unacceptable-risk level based on a balancing of interests (i.e., an exception *stricto sensu*) could be classified in the minimal-risk level if the classification criteria for the high-risk level are not met. This illustrates the dysfunctionality of the high-risk classification criteria set out in Art. 6(1) and (2) AIA, which

[142] Art. 50 AIA.
[143] Recital 132 AIA.

restrict this risk level to the scope of harmonization legislation and the predefined critical areas and specific uses. If these classification criteria are not met, the AIA would impose no mandatory requirements on the excepted AI system that poses excessive risks (as identified in the AIA). Consequently, the classification criteria lack sufficient flexibility to cover the AI systems that have been excepted from the unacceptable-risk level. Given that the AIA establishes a clearly defined (and coherent) risk-based approach,[144] it seems reasonable to recognize the classification of these excepted AI systems directly into the high-risk level.

Second, the AIA creates a peculiar situation regarding the classification exceptions for the high-risk AI systems that do not pose significant risks.[145] As these applications are not to be classified at the high-risk level, the next risk level (without considering the limited-risk level) would be, according to the pyramid of risks, the minimal-risk level. However, this classification would be inconsistent. Whereas minimal-risk AI systems are not subject to mandatory requirements, the AI systems excepted from the high-risk level are subject to a specific subset of rules (including mandatory requirements). Therefore, it seems necessary to differentiate these excepted AI systems from the others classified in the high- and minimal-risk levels. Moreover, as the risk assessment regarding the significant risks is subject to control by the national authorities, it is of practical importance to classify these AI systems.

The pyramid of risks, which is focused on presenting the risk levels hierarchically, can be rebuilt into a comprehensive decision diagram that explicitly indicates the classification criteria and the paths stemming from exceptions. By concentrating on the interrelation of the classification criteria, the decision diagram becomes more reliable for representing the evaluation carried out to classify AI systems. For regulators and regulatees, the decision diagram offers a useful tool to proceed and control the classification of AI systems according to the AIA.

Declaration of competing interest

The author declares that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

This research was supported by KASTEL (Security Research Labs).

Acknowledgements

The author would like to express his gratitude to Qais Hamarneh, who provided valuable feedback, and to T.-T. Prof. Dr. iur. Frederike Zufall, for her insightful contribution to the review of this article.

Data availability

No data was used for the research described in the article.

[144] Recital 26 AIA.

[145] Art. 6(4) AIA.