HIGH-RISK ARTIFICIAL INTELLIGENCE SYSTEMS UNDER THE EUROPEAN UNION'S ARTIFICIAL INTELLIGENCE ACT: SYSTEMIC FLAWS AND PRACTICAL CHALLENGES

Asress Adimi Gikay[*], Pin Lean Lau[§] , Cigdem Sengul[**], Alina Miron[‡] and Ben Malin[†]

ABSTRACT

The European Union's (EU) Artificial Intelligence Act (EU AI Act) has adopted a risk-based approach to artificial intelligence (AI) regulation, where AI systems are subjected to different regulatory standards depending on the seriousness of the risk they pose to public interest. High-risk AI systems, the largest category, are subject to strict regulatory requirements imposed throughout their life cycle, ranging from comprehensive conformity assessment to human rights impact assessment and risk management systems. However, the EU AI Act's high-risk classification system has two systemic fundamental flaws that undermine its ability to strike a fair balance between the risks of various uses of AI technologies and their societal benefits.

First, it defines high-risk AI systems through hyper-technical enumeration, potentially excluding certain AI systems from the high-risk category, even if they pose significant risks to public interest. The Act grants the European Commission the power to revise the high-risk category by adding new AI use cases to the list, if they pose similar or greater risks as the existing ones. But the Commission's power to revise the list does not adequately address the potential loopholes to be created by the restrictive method of defining high-risk AI systems. Second, due to its failure to consider the specific contexts in which AI technologies are used, the EU AI Act could impose disproportionate regulatory burdens on providers and deployers by improperly classifying their AI use cases as high-risk.

By using practical examples based on assessment of several real-world use cases of AI technologies conducted in July 2023 during the St. Gallen University First Grand Challenge on the EU AI Act, this paper argues that the EU AI Act requires revision to adequately regulate AI technologies. The paper proposes a solution to address the EU AI Act's shortcomings, based on the way the law defines high-risk in the context of data protection impact assessment.

KEY WORDS: *EU AI Act, Artificial Intelligence, Machine Learning, High-Risk, General Principle, Low-Risk, Reasonably Foreseeable Misuse*

[*] Senior Lecture in AI, Disruptive Innovation, and Law and Member of Center for AI: Social and Digital Innovation(Brunel University London). Email: asress.gikay@brunel.ac.uk.

[§] Senior Lecturer in Bio Law and General Manager/Member of Center for AI: Social and Digital Innovation (Brunel University London). Email: pinlean.lau@brunel.ac.uk.

[**] Reader, Computer Science Department and Member of Center for AI: Social and Digital Innovation (Brunel University London). Email: cigdem.sengul@brunel.ac.uk.

[‡] Lecturer in the Computer Science Department and Member of Intelligent Data Analysis Research Group (Brunel University London). Email: alina.miron@brunel.ac.uk.

[†] Doctoral Researcher and Member Center for AI: Social and Digital Innovation(Brunel University London). Email: ben.malin@brunel.ac.uk.The authors of this paper took part in First University of St. Gallen Grand Challenge: EU AI Act representing Brunel University London. The Bruenl Team was led by Dr Asress Adimi Gikay. The authors extend special thanks and acknowledgement to Dr. Elena Abrusci, Senior Lecturer in Law and Member of Center for AI: Social and Digital Innovation (Brunel University London). The co-authors of this paper took part in First University of St. Gallen Grand Challenge: EU AI Act representing Brunel University London. The Bruenl Team was lead was led by Dr Asress Adimi Gikay.

CONTENTS

I. INTRODUCTION

The European Union (EU) Artificial Intelligence (AI) Act, often abbreviated as the EU AI Act, is a comprehensive legislative proposal put forth by the European Commission in April 2021.[1] After several months of deliberation, the European Parliament has published a compromise text,[2] pending the conclusion of the trialogue session. The EU AI Act represents a significant step in regulating AI within the EU and carries profound implications for the development and use of AI technologies in the region and around the globe. The core objectives of the EU AI Act are to strike a balance between fostering innovation and ensuring the ethical and responsible use of AI. It establishes a legal framework designed to regulate various aspects of AI, and once passed, it will likely be the first most comprehensive AI law in the world, with major impact on the global AI industry.

The EU AI Act adopts a risk-based framework that formally classifies AI systems into three categories: unacceptable risk, high risk, and non-high risk. AI systems that pose a serious threat to fundamental rights or safety, such as social scoring systems and real-time, remote biometric identification systems, are prohibited under the act. High-risk AI systems, such as those used in employment, healthcare, and law enforcement, must meet strict requirements before they can be marketed or used in the EU. Non-high-risk AI systems, such as those used in chatbots and video games, and permitted emotion recognition systems must comply with certain transparency and accountability requirements. These requirements entail, amongst others, providing natural persons the information that they are interacting with an AI system; securing consent(in case of emotion recognition systems); informing natural persons which functions of the system they are interacting with is AI-enabled; and whether there is a human oversight.[3] These non-high risk AI systems are commonly known as "limited risk AI systems."[4] The EU AI Act is inapplicable to other non-risk AI systems, as they pose minimal-risk or no risk, such as those used in spam filters and calculators.

[1] European Commission, "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" (COM/2021/206 final).

[2] European Parliament, "DRAFT Compromise Amendments on the Draft Report" (2023), available at https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/CJ40/DV/2023/05-11/ConsolidatedCA_IMCOLIBE_AI_ACT_EN.pdf (last accessed 13 September 2023).

[3] Article 52(1)-(3), EU Parliament Compromise Text.

[4] EU Commission, Regulatory framework proposal on artificial intelligence, available at https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai(last accessed 01 November 2023).

The specific classification of non-high-risk AI systems into limited risk and minimal risk is not strictly based on the provisions of the EU AI Act, as this seems to be adopted by commentaries for the sake of convenience (*See* s. II(A)). In addition to its risk-based classification system, the EU AI Act also includes a number of other important provisions, such as requirements for transparency and accountability; prohibition of discriminatory AI systems; and human oversight of AI systems.

Because the primary focus and justification of the EU AI Act revolve around risks, with the intention of establishing proportionate and effective rules, questions naturally arise regarding how effectively it encompasses the protection of the interests of AI system 'end users.'[5] Experts evaluating the EU AI Act also express concerns about the applicability of a top-down allocation of responsibilities, which has traditionally worked well for tangible products, but may face more significant challenges in the context of 'upstream' AI services that can be repurposed in unforeseen ways downstream. Further scrutiny is necessary to address these concerns, including an examination of the transparency and risk management strategies, as well as the categorization of AI systems that could potentially serve diverse purposes.[6]

Beyond these reasons, the authors of this paper identify specifically two systemic fundamental flaws that undermine the ability of the EU AI Act to strike a fair balance between the risks of diverse uses of AI technologies and their societal benefits. These form the basis of the inquiry in this paper and hypothesise the urgency of critical revisions to the EU AI Act before it is adopted and enters into force. First, the Act defines high-risk AI systems through a hyper-technical enumeration method, potentially excluding certain AI systems from the high-risk category, even if they pose significant risks to public interest. Secondly, due to its failure to consider specific context of use of AI technologies, the Act could impose disproportionate regulatory burdens in relation to AI systems that are improperly classified as high-risk AI systems.

[5] Tobias Mahler, 'Between Risk Management and Proportionality: The Risk-Based Approach in the EU's Artificial Intelligence Act Proposal' (30 September 2021) <https://papers.ssrn.com/abstract=4001444> accessed 16 March 2023.

[6] Lilian Edwards, 'Expert Explainer- The EU AI Act: A Summary of Its Signfiicance and Scope' <https://www.adalovelaceinstitute.org/wp-content/uploads/2022/04/Expert-explainer-The-EU-AI-Act-11-April-2022.pdf> accessed 16 March 2023.

To illustrate the claims, the paper is structured as follows. Section II offers a succinct explanation of the high-risk classification system adopted by EU AI Act as well as the categories of such systems. Through an analysis of standalone AI systems and AI systems as a safety component of a product or as a product regulated under harmonised legislation listed in Annex II(A), the paper lays a foundation to analysing the two fundamental flaws. In Section III, the paper highlights the pitfalls of the risk classification system, by using various AI use cases as examples to demonstrate systematic fundamental flaws examined by the authors.

The authors utilise their experience in undertaking conformity assessments of real-world AI systems, products and technologies developed by specific European technology companies, ranging from robotics delivery systems to fertility and reproductive technologies.[7] Whilst the authors are unable to divulge details of the technologies or the companies to whom they rendered the reports for conformity assessments due to confidentiality reasons, this paper uses the lessons learned from the assessment, which exposed inadequacies and lack of clarity of the EU AI Act, specifically in relation to high-risk AI systems. In section IV, the authors offer an alternative method of risk classification, based on the experience from the General Data Protection Regulation (GDPR).[8] Section V concludes the paper by providing an urgent recommendation that could be implemented before the EU AI Act becomes a law. Whilst the recommendation in this paper is made specifically in relation to the EU AI Act, it is equally useful to national jurisdictions crafting their regulation for AI technologies.

II. The High-Risk Classification of AI Systems under the EU AI Act

In this section, we provide a broad understanding of the risk-based approaches that have now become characteristic of the EU AI Act. As will also be demonstrated in the following sub-sections, the risk-based approaches present complications in the closed lists or categorisations of AI systems as high-risk.

A. The Risk-Based Regulation of the EU AI Act

The EU AI Act applies to AI systems defined as "a machine-based system that is designed to operate with varying levels of autonomy and that can, for explicit or implicit objectives,

[7] The authors took part in the first EU AI Act Assessment competition organised by the University of St. Gallen. Details of the team are available at https://www.thegrandchallenge.eu/1st-gc-2023 (Last accessed 01 November 2023).

[8] Regulation (EU) No 2016/679 (OJ 2016 L 119 p.1).

generate outputs such as predictions, recommendations, or decisions, that influence physical or virtual environments."[9] The specific regulatory standard applicable to an AI system depends on which category it falls into, as the Act adopts a risk-based approach. Under the Act, there are three risk categories— (1) unacceptable risks (Title II); (2) high risks (Title III) and (3) low risks (Title IV).

Some scholars distinguish between limited risk and minimal risk and argue that there are therefore four risk categories under the Act.[10] A thorough study of the Act reveals that the act does not necessarily recognise four categories. What scholars refer to as "minimal-risk AI systems" is "non-high-risk AI systems" with respect to which providers can adopt codes of practice similar to the requirements for high-risk AI systems pursuant to Title XI.[11] But this means that the EU AI Act merely suggests that for a non-high-risk AI system, including low-risk AI systems, voluntary compliance with the requirements of the high-risk AI systems is encouraged. There is no distinction in the Act between limited risk and minimal risk. These terms do not appear in any of the Act's versions. Indeed recital 14 of the Act clearly implies that it is concerned with three risk categories as follows:

> *In order to introduce a proportionate and effective set of binding rules for AI systems, a clearly defined risk-based approach should be followed. That approach should tailor the type and content of such rules to the intensity and scope of the risks that AI systems can generate. It is therefore necessary to* ***prohibit certain unacceptable artificial intelligence practices, to lay down requirements for high-risk AI systems and obligations for the relevant operators, and to lay down transparency obligations for certain AI systems***[12]***(Emphasis Added)***

Although a distinction between limited risk and minimal risk could be made in ordinary language, with some difficulty due to the subjectivity of what might amount to limited and minimal risk, the EU AI Act itself does not make such a distinction apparent. Thus, for the purposes of this paper, we operate on the basis of there being three risk categories under the EU AI Act.

[9] EU AI Act Article 3(1).

[10] M. Veale and F. Zuiderveen Borgesius, "Demystifying the Draft EU Artificial Intelligence Act" (2021) 22(4) Computer Law Review International 97 at 98.

[11] Ibid.

[12] EU AI Act Recital 14.

The criterion for assessing risk is the Union's public interest (or public interest for short) including the health and safety of persons,[13] fundamental rights, democracy, the rule or law, and the environment.[14] AI systems are treated differently depending on the degree of risk they pose to these elements of public interests. Those that pose unacceptable risks are prohibited; those posing high risks are strictly regulated and those posing low risks are subject to minimal transparency obligations. The EU AI Act also gives providers and deployers the option to voluntarily apply the strict standards applicable to high-risk AI systems to those systems that do not fall under the high-risk category, i.e., to low-risk AI systems, by way of code of practice.[15]

The EU AI Act lists the prohibited practices in an exhaustive manner.[16] Thus, AI systems that are covered by the prohibition are not allowed to be put on the market, provided that specific requirements for the prohibition are met. These AI systems are not the subject of analysis in this work. This article rather focuses on high-risk AI systems that are subject to stringent regulatory standards and explains why the system of classification is unfit for effective AI regulation, due to its enumerative method that lacks the strength of general principles of application. The delegation of power to the European Commission, to revise the high-risk AI category, is inadequate to address the potential loopholes the current high-risk classification method might create. The proceeding section elaborates on high-risk AI system categorisation in further detail.

B. CATEGORIES OF HIGH-RISK AI SYSTEM

Under the EU AI Act, there are two subcategories of high-risk AI systems. The first one is commonly referred to as stand-alone AI systems: AI systems that exist independently of a product.[17] The second sub-group consists of AI systems that are considered as safety components of products or products themselves that are regulated by the EU harmonisation legislations.[18]

Each sub-category has closed lists and AI use cases that do not fall in the listed areas will not be considered as high-risk, even if they pose serious risks to the public interest. This

[13] EU AI Act Recital 28.
[14] EU AI Act Recital 27.
[15] EU AI Act Art. 69.
[16] EU AI Act Art. 5.
[17] EU AI Act, Annex III.
[18] EU AI Act, Annex II.

allows for such systems to be deployed and put onto market, without meeting the requirements for high-risk AI systems and more critically, without undertaking the strict standards and obligations necessary for high-risk AI systems. Similarly, some use cases may improperly be classified as high-risk AI use cases due to technicality, even if the context in which they are used present low risk to public interest. This is the foundation of the problems that this paper aims to expose.

1. STANDALONE AI SYSTEMS

Standalone AI systems are AI systems that are recognised as high-risk AI systems listed in Annex III of the EU AI Act. In both the EU Commission's proposal and the Parliament's compromised text, the Annex contains eight fixed areas within which these AI systems fall.[19] These are:

1. Biometric and biometrics-based systems;
2. Management and operation of critical infrastructure;
3. Education and vocational training;
4. Employment, workers management and access to self-employment;
5. Access to and enjoyment of essential private services and public services and benefits;
6. Law enforcement;
7. Migration, asylum and border control management; and
8. Administration of justice and democratic processes.

There are three crucial points to remember about stand-alone high-risk AI systems. First, the fact that an AI system is listed in one of the above fields does not necessarily mean that it qualifies as a high-risk AI system, as that only makes it *potentially* [emphasis added] a high-risk AI system. They are classified as high-risk AI systems only if, "in the light of their intended purpose, they pose a significant risk of harm to the health and safety or the fundamental rights of persons and, where the AI system is used as a safety component of a critical infrastructure, to the environment."[20] Such significant risk of harm should be identified by assessing; on the one hand, the effects of such risk with respect to its level of severity, intensity, probability of occurrence and duration combined altogether, and on the other hand, whether the risk can affect an individual, a plurality of persons or a particular group of persons. Such combinations could, for instance, result in a high severity but low probability to affect a

[19] EU AI Act, Annex III.
[20] EU AI Act Parliament Compromised Text Recital 32.

natural person, or a high probability to affect a group of persons with a low intensity over a long period of time, depending on the context. The identification of those systems is based on the same methodology and criteria envisaged also for any future amendments of the list of high-risk AI systems.

Secondly, an AI system that does not fall within the listed categories does not qualify as being a high-risk AI system, regardless of the potential consequences and risks it may pose to public interest. Thirdly, the Commission can amend the list following an appropriate procedure, through a delegated Act.

We argue that this is problematic, because AI systems can be used in a number of ways and areas that are not necessarily on the list. There have already been use cases that were not in the list according to the Commission proposal, but added by the Parliament, such as immigration prediction tools. The Parliament's compromised text itself does not cover AI systems such as those used by financial institutions to tackle money laundering. While these are very obvious cases, there are also more complex use cases that are out of the ambit of the AI systems, despite posing high risk to the protected values in the Act. The Commission's delegated authority is unlikely to lead to an optimal result in closing this loophole (for detailed analysis, see Section III of this paper).

2. AI Systems as Safety Components of Products, or Products Regulated by Harmonisation Law

In this category, an AI system is classified as high-risk if it is a safety component of a product or is itself a product regulated under harmonised legislation listed in Annex II(A) and is subject to third-party conformity assessment under the relevant law.[21]

The EU AI Act addresses two sub-classes of AI systems in this category. First, AI systems are safety components of a product subjected to harmonisation legislation in Annex II(A) and should go through a conformity assessment under the relevant harmonisation legislation which then classifies those as high-risk AI systems. There are many products that are listed in this category including machinery, toys, lifts, aircraft and many others listed in Annex II(A). To qualify as high-risk under this category, the AI system must be a safety component, which means that the failure or malfunction of the AI system should pose a danger to safety and health of persons.

[21] EU AI Act Parliament's Compromise Text Article 6(1).

The determination of whether an AI system is a safety component may be challenging especially because there could be discrepancies between the definitions of safety components in the AI Act and the relevant harmonisation legislation. For instance, the Machinery Regulation requires the independent placing of a system on the market to qualify as a safety component, while the EU AI Act does not explicitly require this.[22] The wording of the relevant EU AI Act provision seems to suggest that the meaning of safety component under the relevant harmonisation legislation prevails. However, we contend that an interpretation of the definition of safety component should also consider the need to enhance responsible use of AI and treat safety components in the same way, even if they are produced as a part of the machinery, rather than being independently supplied. This is in line, for instance, with the Machinery Regulation, which requires third-party conformity assessment for "machinery that has embedded systems with fully or partially self-evolving behaviour using machine learning approaches ensuring safety functions that have not been placed independently on the market, in respect only of those systems."[23]

The problem of clarity as to definitions have also been raised by some stakeholders[24] in the medical devices and healthcare sector in connection with AI systems that are medical devices (MD) or in-vitro diagnostic medical devices (IVD), that are currently regulated under a variety of legislation, including the GDPR, the Medical Devices Regulation (MDR), and the In vitro Diagnostic Medical Devices Regulation (IVDR). For instance, stakeholders have raised the need for definitions to be consistent at both the European and member state levels. Examples raised include the definitions of 'user' (Article 3 (4)) whereby professional users are distinguished from the more general term 'user' to reflect laypersons, including individuals or patients, so that they get the same level of protection afforded by the EU AI Act as professional users, such as healthcare professionals;[25] definition of 'risk' (Article 3 (new)) which lacks a dedicated definition in the European Commission proposal and the Council of the EU General Approach;[26] and definition of 'AI systems' (Article 3 (1)) which is considered to be overly

[22] EU AI Act Art. 3(3). The EU AI Act does not explicitly address independent placing of the product on the market to qualify as a safety component(see EU Act Art. 3(14).

[23] See Regulation (EU) 2023/1230 of the European Parliament and of the Council of 14 June 2023 on machinery and repealing Directive 2006/42/EC of the European Parliament and of the Council and Council Directive 73/361/EEC (the Machinery Regulation) Annex I.

[24] MedTech Europe, 'Stakeholder Joint Statement on Access to Innovative Healthcare under the Artificial Intelligence Act (AI Act)' 2023.

[25] ibid 2.

[26] ibid.

broad, potentially including all medical technologies with software components that are not necessarily considered AI.[27]

Generally speaking, the task of determining whether an AI system is a safety component of another product falling under the harmonisation legislation should not be difficult. For instance, in a robotics system that navigates on the road, the AI system that detects obstacles, pedestrians and facilitates a safe navigation could be considered as a safety component, notwithstanding the fact it may serve other purposes as well. However, clarity and consistency between the AI Act and other legislations would be instrumental in avoiding interpretive inconsistencies and legal uncertainty.

Another sub-category includes AI systems that are themselves regarded as products under the relevant harmonisation law. An example of this type of AI system is software used in healthcare. Regulation (EU) 2017/745 on Medical Devices, or more commonly referred to as the Medical Devices Regulations (MDR) defines such AI systems as medical devices and subjects them to conformity assessment under the MDR, as well as under the EU AI Act.[28] Therefore, software systems used in healthcare are considered as products (medical devices) and are considered high-risk AI systems. This being the case, such products would need to undergo a conformity assessment procedure[29] under the EU AIA. In a nutshell, the EU AI Act regulates AI medical devices that are subject to the MDR. Hence, AI systems which qualify as a medical device and require third-party conformity assessment under the MDR, will automatically be classified as high-risk AI systems under the EU AI Act.[30]

Similarly, to the standalone AI systems, embedded AI systems or AI systems as products are exhaustively listed by the EU AI Act. Any AI system that does not fall under the listed harmonisation legislations may not be considered as a high-risk AI system. Once again, if such systems emerge, it is up to the Commission to amend the list. This can be highly problematic. For example, AI systems that are not regulated as medical devices under the MDR or regarded as low risk under the MDR, are not regulated under the EU AI Act. This effectively excludes the Act from applying to other health and medical AI systems that are not considered to be medical devices. These could be simply fitness or health applications that monitor basic health functions such as heart rate or numbers of steps walked per day, which may generally

[27] ibid.

[28] Article 2(1) MDR; Recital 19, MDR; Annex VIII, Chapter 3, Rule 11 MDR).

[29] Title III, Arts. 6(1)(a)-(b) EU AIA.

[30] Janneke van Oirschot and Gaby Ooms, ‘Interpreting the EU Artificial Intelligence Act for the Health Sector’ (Health Action International 2022) 8.

be considered low risk[31] and not particularly concerning from the perspective of Rule 11 of the MDR. But it does not capture concerns regarding AI systems that might be used in the healthcare sector, but which are excluded from the purview of the EU AI Act – such as hospital systems for patient data management, or AI systems that choose patients in recruitment for clinical trials based on specific medical or biomarker criteria.[32]

These examples, whilst perhaps regarded as medical devices, would not be used in accordance with Rule 11 MDR – and if so, they would be regarded as low risk. The reality, however, is that this artificial segregation may be confusing, and exclude many other types of AI systems from the purview of the EU AI Act, because of their classification under the MDR, even if these applications might have an impact on population health.[33]

3. The Requirements for High-Risk AI Systems: Overview

Under Article 43 of EU AI Act, the provider of a high-risk AI system should undertake a conformity assessment. In this section, we use the example of medical devices to demonstrate the demanding nature of such conformity assessment, and eventual compliance.

As demonstrated above, a medical device is also subject to the requirements of the MDR. Whilst the EU AIA aims to avoid duplication of assessment for AI systems under multiple assessment regimes, we assert that assessment of conformity with the EU AI Act should focus only on requirements of the EU AI Act that are not covered by the MDR (Title III, Article 8; and Article 43, paragraph 3, EU AIA). Therefore, such conformity assessment should include assessment of compliance under Title III, Chapter 2 along with Annex VII, points 4.3 - 4.5 and point 4.6 paragraph (5). Similar to the EU AI Act, the MDR uses a risk-based approach, classifying medical devices into four categories: from the lowest risk (Class I) to medium risk (Classes IIa and IIb), to highest risk (Class III) [Chapter V, Article 51, MDR; and Annex VIII, MDR]. The conformity assessment procedures are listed in Chapter V, Article 52, MDR and manufacturers are to undertake assessment in accordance with Annexes IX to XI, MDR.

[31] ibid 9.

[32] ibid 10.

[33] Janneke van Oirshot, 'Recommendations: Addressing Risks and Harms from Health-Related Artificial Intelligence'<https://haiweb.org/wp-content/uploads/2022/02/Recommendations_Health-related-AI.pdf> accessed 31 October 2023.

Hence, an AI product of software in healthcare can be classified as a software, which is a recognised medical device under Recital 19, MDR; as well as Annex VIII, point 6.3, Rule 11, whereby:

> *software intended to provide information which is used to take decisions with diagnosis or therapeutic purposes is classified as class IIa........except if it is intended for ......variations of parameters ....that could result in immediate danger to the patient, in which case it is classified as class IIb.*"

Regardless, the MDR provides that: *"For devices of classes IIa, IIb and III, an appropriate level of involvement of a notified body is compulsory proportionate to the risk class...*[34]

This further means that a conformity assessment under the MDR will require a third-party conformity assessment by a notified body. Under Annex VII, MDR- the conformity assessment activities cover quality management system auditing (point 4.5.2), product verification (point 4.5.3), pre-clinical evaluation assessment (point 4.5.4), clinical evaluation assessment (point 4.5.5), other specific procedures (point 4.5.6), reporting (point 4.6), final review (point 4.7), decisions and certifications (point 4.8), changes and modifications (point 4.9), surveillance activities and post-certification monitoring (point 4.10), and re-certification (point 4.11).

Within the framework of conformity assessment under the MDR, the notified body will be responsible for assessing the additional requirements of the EU AI Act (i.e., class IIa or above in the MDR). There are also other obligations under Annex II, MDR to provide technical documentation for the medical devices; Annex III, MDR for mechanisms relating to post-market surveillance; and Annex VI regarding registration of devices and economic operators.

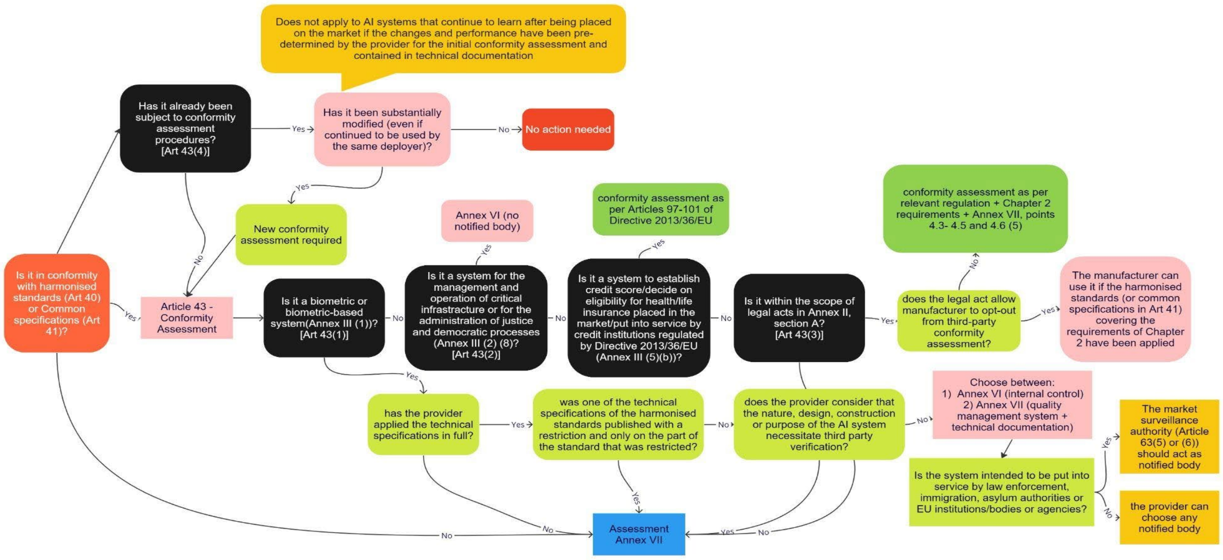
The reality, however, is that there is uncertainty about the EU AI Act's alignment with other sectoral EU-level legislation that would, in fact, prevent duplication of assessments. Because of this, whilst many civil society organisations that champion the human rights of patients in medicine and healthcare may welcome the conformity assessment for medical devices that are also AI systems – other stakeholders such as medical device manufacturers and deployers may find these procedures onerous. One of the key points raised is that the EU AI Act may risk creating two-track systems, one applicable to the AI component of a device, and the other to the MD or IVD component of a device. This could create legal uncertainty and obstacles in delivering the ethical, safe and effective devices that the Act intends to support.[35]

[34] Explanatory Memorandum, EUR-Lex – 52012PC0542- EN- European Union.

[35] MedTech (n 29).

Compliance with the EU AI Act would be costly for companies providing and/deploying AI technologies. According to one study, the cost of compliance including, amongst others, conducting conformity assessment and implementing quality management systems, could be between €193,000 to €330,000.[36] Given the high cost of compliance, especially for small companies or companies that are at an early stage in commercialising their technologies, the criteria to determine whether a given use case is high-risk should be more nuanced, balanced and context-based, to ensure that unnecessary cost of compliance is not borne by businesses. Similarly, the failure of some cases to fit into the high-risk category due to technicality could pose risks to public interest and defeat the purpose of the act. This paper argues that the EU AI Act fails to properly address the concerns raised here.

EU AI Act Assessment Procedure Flow Chart


Does not apply to AI systems that continue to learn after being placed on the market if the changes and performance have been pre-determined by the provider for the initial conformity assessment and contained in technical documentation
Has it already been subject to conformity assessment procedures? [Art 43(4)]
Yes
Has it been substantially modified (even if continued to be used by the same deployer)?
No
No action needed
Yes
New conformity assessment required
No
Is it in conformity with harmonised standards (Art 40) or Common specifications (Art 41)?
Yes
Article 43 - Conformity Assessment
Is it a biometric or biometric-based system(Annex III (1))? [Art 43(1)]
No
Annex VI (no notified body)
Yes
Is it a system for the management and operation of critical infrastructure or for the administration of justice and democratic processes (Annex III (2) (8)? [Art 43(2)]
No
conformity assessment as per Articles 97-101 of Directive 2013/36/EU
Yes
Is it a system to establish credit score/decide on eligibility for health/life insurance placed in the market/put into service by credit institutions regulated by Directive 2013/36/EU (Annex III (5)(b))?
No
Is it within the scope of legal acts in Annex II, section A? [Art 43(3)]
Yes
does the legal act allow manufacturer to opt-out from third-party conformity assessment?
No
conformity assessment as per relevant regulation + Chapter 2 requirements + Annex VII, points 4.3- 4.5 and 4.6 (5)
Yes
The manufacturer can use it if the harmonised standards (or common specifications in Art 41) covering the requirements of Chapter 2 have been applied
Yes
has the provider applied the technical specifications in full?
Yes
was one of the technical specifications of the harmonised standards published with a restriction and only on the part of the standard that was restricted?
No
does the provider consider that the nature, design, construction or purpose of the AI system necessitate third party verification?
No
Choose between:
1) Annex VI (internal control)
2) Annex VII (quality management system + technical documentation)
Is the system intended to be put into service by law enforcement, immigration, asylum authorities or EU institutions/bodies or agencies?
Yes
The market surveillance authority (Article 63(5) or (6)) should act as notified body
No
the provider can choose any notified body
Yes
No
No
Yes
No
Assessment Annex VII

III. The Pitfalls of the EU's High-Risk Classification Systems

A. Problems relating to foresight

Due to the enumerative system of classifying high-risk systems, there are AI use cases that do not fit into the category of high-risk AI systems although their use could pose similarly

[36] CECIMO Paper on Artificial Intelligence at 4, available at https://www.cecimo.eu/wp-content/uploads/2022/10/CECIMO-Paper-on-the-Artificial-Intelligence-Act.pdf)(last accessed 01 November 2023).

significant risks to life, health, safety, rule of law, democracy and the environment. In this subsection, we provide two examples that expose the defects in the system classification of the high-risk category adopted by the EU AI Act.

The Commission's first draft did not include certain AI systems that could have severe implications, generally referred to as predictive tools. One of these tools is a system used to predict migration flow. Whilst migration flow prediction AI systems can be used to predict an influx of migration and deploy the appropriate resource in the relevant hotspot by Non-Governmental Organisations (NGOs), it can also be used to secure borders in a discriminatory manner.

As these tools could rely exclusively on non-personal data such as news reports of war, displacement and natural disaster, rather than using personal data of individuals, data protection law does not extend available protections to individuals who could be adversely impacted by such tools. Privacy law also falls short of protecting potential victims of adverse decisions based on predictive algorithms. As such, a specific provision in the AI Act would be the only appropriate legal safeguard against the potential abuse of immigration prediction algorithms. Such an obvious use case was missed in the initial EU Commission draft, and it has only been addressed in the parliament's compromise text. This shows that the task of determining the AI systems that could be used in an abusive manner is a difficult one, and a closed list of high-risk AI systems is likely to always miss the mark as AI systems and their uses evolve.

Another category of AI system that is not covered either by the EU Commission's draft or the Parliament's compromised text is AI systems used to tackle money laundering by financial institutions. To address the increasingly sophisticated money laundering crime, financial institutions rely on AI systems to identify suspicious transactions. Machine learning algorithms are considered to improve the process of detecting transactions that are suspicious. But using AI systems to tackle money laundering poses significant risks to the rights and wellbeing of individuals. While AI systems analyse trade-offs using abstract parameters (e.g., false positives, false negatives), these trade-offs may have concrete, not abstract, harms on people. These AI systems, if not properly regulated, can lead to discrimination of persons based on race, religion, geographical origin or location.[37] Individuals who could be falsely alerted may also experience undue delay in financial transactions. Overall, the effect of an AI system used to tackle money laundering could be significant on persons, as AI systems used for

[37] Astrid Bertrand 287; Raphaële Xenidis, 'Tuning EU equality law to algorithmic discrimination: Three pathways to resilience' 2020, Vol. 27(6) Maastricht Journal of European and Comparative Law 736–758, 738

instance for consumer credit risk assessment. Nevertheless, the latter is considered as a high-risk AI system subjected to strict regulatory standards, while the former is not.

The late addition of migration prediction tools to the list of high-risk AI systems and the absence of money laundering detection tools from the list even in the parliament's compromise text suggests that attempting to come up with a closed list of high-risk AI systems is a futile exercise. More use cases are likely to be out of the ambit of regulation, more than we could imagine and anticipate. The Commission's delegated act is inadequate to respond to the evolution of the technology and address risks posed by new use cases.

B. Low-Risk AI Systems with Reasonably Foreseeable High-Risk Use

The EU AI Act's risk-based approach leads to certain AI systems being classified as low risk based on their intended purposes. Despite this classification, there's a discernible possibility that these systems could be exploited in ways not originally intended or planned by the provider. This potential for misuse exists because while an AI's design might be inherently benign, its capabilities can be manipulated or repurposed in a foreseeable manner.

An example of such a system could be an AI system that utilises machine learning techniques for monitoring and evaluating production assembly lines in manufacturing environments with the aim of improving efficiency. The main purpose of this system is to monitor and assess production cycles to output various performance metrics for review by a line or factory manager. In addition, the production line assessments can be outputted anonymously to produce training videos showcasing examples of efficient work. Such a system could use cameras mounted either at a workstation or providing a bird-eye view of the assembly line. To avoid processing personal data, the system may use computer vision algorithms to anonymize people in the video stream before further video processing and storage.

The AI system described above does not fall in any of the high-risk categories recognised under the EU AI Act. However, there are clearly several high-risk elements that it has. Firstly, despite using computer vision algorithms to anonymise individuals in the video stream, there is still a risk that the anonymisation may be ineffective. Advanced techniques could reverse the anonymisation, especially if unique identifiers are present (e.g., clothing, jewellery, or even body movements). There is also the risk of the system inadvertently capturing sensitive information present on the production floor. Secondly, the system's assessment of "efficient work" may inadvertently incorporate biases. If the training data for the ML model is not representative of the diversity of workers or working styles, the system might favour certain demographics or working methods over others, leading to unfair assessments. If the provider determines the intended use of this technology, the use of the technology to evaluate worker performance would be considered a misuse. But can an AI system be classified as high-risk based on a potential misuse?

The AI Act is unclear on this issue. The Act defines "reasonably foreseeable misuse" as "the use of an AI system in a way that is not in accordance with its intended purpose *as indicated in instructions for use established by the provider*, but which may result from reasonably foreseeable human behaviour or interaction with other systems*, including other AI systems*."[38] However, several provisions of the act imply that the reasonably foreseeable misuse of the AI system does not lead to its classification from low risk to high-risk. Requirements designed to reduce the risks of AI systems that are marketed or deployed should consider their intended use, the potential for misuse that can reasonably be anticipated, and should align with the risk management framework established by the provider. These requirements should be objective-driven.[39]

The act does not suggest that the knowledge of the reasonably foreseeable misuse could change the classification of the AI system to a high-risk one. On the contrary, "the potential misuse and malicious use of the AI system and of the technology underpinning it" is one of the factors that the Commission should consider when revising the high-risk AI system in Annex II or exercising its delegated power.[40]

Normatively speaking, it is debatable whether the reasonably foreseeable misuse of an AI system should automatically entail its classification into a high-risk AI system. On the one hand, a clear possibility of misuse that the provider can foresee should lead to concerns. If the deployer (user) of the AI system decides to depart from the instruction of use and put the AI system to a misuse, it could pose significant risk that should be tackled in advance. For this reason, it could be argued that classifying the AI system as high-risk based on its potential misuse rather than as low risk based on its intended use is more appropriate. On the other hand, if the provider of the AI system puts in place measures to tackle the potential misuse of the AI system, including legal, technical and organisation measures, then it would be disproportionate to classify the AI system as high-risk and impose the requirements of a high-risk system.

The dilemma is largely due to the enumerative closed list of high-risk high-risk AI systems. If the AI system does not squarely fit into the list, it becomes questionable whether it

[38] Recital 13.

[39] Recital 42. Recital 43 also suggests similar things.

[40] Article 7(2) (e a).

should be placed in the category based on its reasonable misuse. The AI Act's position appears to be that the reasonably foreseeable misuse does not change the classification of the AI system, unless the Commission uses such use as a factor in revising the list of high-risk AI systems. We argue that the solution lies in completely revamping the high-risk classification system under the EU AI Act.

C. Excessive Regulatory Burden on Providers and Deployers of Legally Misclassified AI Systems

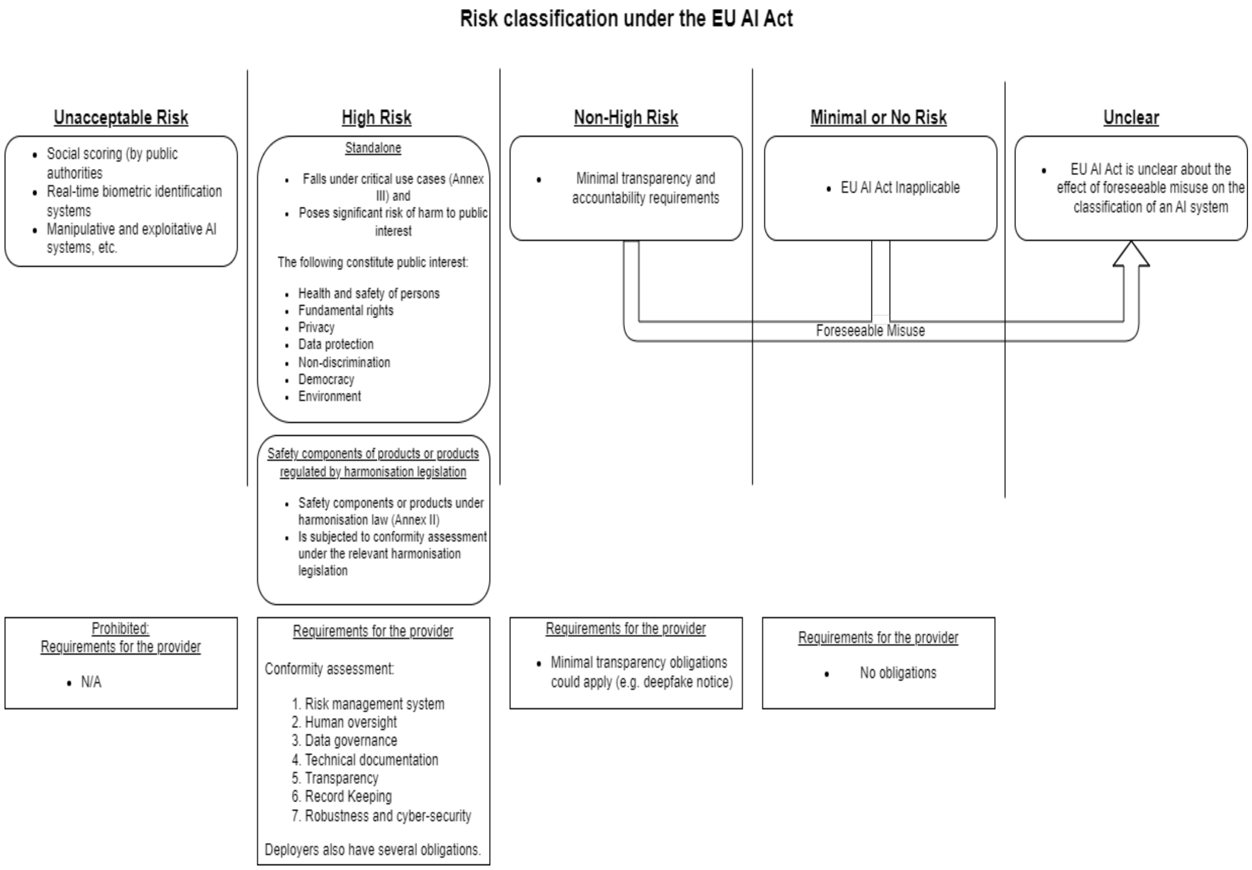
Another example of an AI use case that could demonstrate the flaw in the EU AI Act is an AI-powered delivery robot. Robots designed for deliveries, whether they are ground-based or drones, could transform the methods of transporting and delivering goods. Their use is highly dependent on the context, resulting in categories that range from low risk to high-risk scenarios.

For example, delivery robots operating in controlled environments, like warehouses or large corporate campuses encounter few unpredictable variables. The robots usually navigate pre-defined routes and can avoid obstacles with minimal risk for the safety of people. Nevertheless, the same robots deployed in an urban setting, increases the risk of collisions with pedestrians, vehicles or other obstacles, with heightened levels of potential misuse arising from this setting. The unpredictability of the environment increases the complexity of the algorithms used and can transform the classification from a low risk to a high-risk system.

Facial recognition is another area whereby the context of deployment can dictate the classification of the system, with an implementation in the public domain likely being classified as unacceptable risk and a breach of GDPR. Whereas this technology in a closed, private deployment poses significantly less risk and potential misuse.

Thus, the context in which AI systems are placed is pivotal for their classification. To classify an AI system or use case as high-risk without taking into consideration the specific context of use risks imposing disproportionate burden on providers and deployers that should comply with a number of regulatory requirements, even though their AI use case may pose little or no risk to public interest.

RISK CLASSIFICATION UNDER THE EU AI ACT



Risk classification under the **EU AI Act**

Unacceptable Risk

- Social scoring (by public authorities
- Real-time biometric identification systems
- Manipulative and exploitative AI systems, etc.

Prohibited:
Requirements for the provider

- N/A

High Risk

Standalone

- Falls under critical use cases (Annex III) and
- Poses significant risk of harm to public interest

The following constitute public interest:

- Health and safety of persons
- Fundamental rights
- Privacy
- Data protection
- Non-discrimination
- Democracy
- Environment

Safety components of products or products regulated by harmonisation legislation

- Safety components or products under harmonisation law (Annex II)
- Is subjected to conformity assessment under the relevant harmonisation legislation

Requirements for the provider

Conformity assessment:

1. Risk management system
2. Human oversight
3. Data governance
4. Technical documentation
5. Transparency
6. Record Keeping
7. Robustness and cyber-security

Deployers also have several obligations.

Non-High Risk

- Minimal transparency and accountability requirements

Requirements for the provider

- Minimal transparency obligations could apply (e.g. deepfake notice)

Minimal or No Risk

- EU AI Act Inapplicable

Requirements for the provider

- No obligations

Unclear

- EU AI Act is unclear about the effect of foreseeable misuse on the classification of an AI system

IV. Alternative Method of Risk-Classification

A. General Principle for Qualifying High-Risk AI System

First and foremost, high-risk AI systems should be defined using a general principle under which such systems should be assessed. Whilst the EU AI Act has certain criteria that are taken into consideration in determining whether the AI system is high-risk; namely significant risks to fundamental rights, health, safety and life of person, rule of law, democracy and the environment, these criteria are applied to the AI systems that are already listed as potential high-risk systems in the Annex III. In other words, the criteria are used to potentially exclude AI systems from being qualified as high-risk, even if they are on the list, if they do not pose significant risk to values protected mentioned above. If a system is not on the list, the criteria are inapplicable, as it is automatically considered a non-high-risk AI system.

The proposed general principle requires using the risk criteria to assess the risk posed by AI systems, without enumerating potential candidates for the assessment. This does not exclude the possibility of creating an illustrative list of AI systems that can be considered as high-risk AI systems. An example of this approach can be found in the GDPR provisions governing Data Protection Impact Assessment (DPIA).

Under the relevant provision of the GDPR, data controllers are required to conduct DPIA "where a type of processing in particular using new technologies, and taking into account the nature, scope, context and purposes of the processing, is *likely to result in a high risk to the rights and freedoms of natural persons*."[41] Under the GDPR, whilst not all personal data processing operations require DPIA, the ones that are "likely to result in a high risk to rights and freedoms of natural persons" entails the necessity of DPIA. The GDPR does not provide detailed guidance on "likely to result in a high risk to rights and freedoms of natural persons." However, it does provide an illustrative list of processing operations that meet the criteria. Thus, the DPIA referred to in paragraph 1 shall "*in particular*" be required in the case of:[42]

a. a systematic and extensive evaluation of personal aspects relating to natural persons which is based on automated processing, including profiling, and on which decisions are based that produce legal effects concerning the natural person or similarly significantly affect the natural person;

[41] GDPR Article. 35(1).

[42] Ibid, Article 35(3).

b. processing on a large scale of special categories of data referred to in Article 9(1), or of personal data relating to criminal convictions and offences referred to in Article 10; or

c. a systematic monitoring of a publicly accessible area on a large scale.

Whilst DPIA is mandatory for the above cases, as they are presumed to likely result in high risk to the rights and freedoms of natural persons, the GDPR does not exclude DPIA for other cases, as the above cases are only illustrative.[43] In addition to the illustrative list, the GDPR allows National Supervisory Authorities (NSAs) to draw a list of processing operations that entail DPIA or for which DPIA is not necessary and make it publicly available.[44] The Article 29 Working Party has issued a guideline outlining nine criteria to be used in determining whether DPIA should be conducted.[45]

Learning lessons from the GDPR, a general principle for qualifying AI systems as high-risk AI systems should be introduced under the EU AI Act. Such principles can be supplemented by an illustrative list of high-risk AI systems as well as guidance issued by NSAs. The EU Commission could also be involved in the process of developing such guidance. Applying the general principles, supervisory authorities, individuals, companies, NGOs and other stakeholders should be able to assess whether a given AI use case is high-risk or not depending on the context and the potential adverse effect on public interest. Such a system could potentially lead to disagreements as to whether a given use case qualifies as high-risk. In case of such disagreements, the ultimate determination must be made by a court of law, rather than a supervisory authority or the Commission.[46]

V. Conclusion

Whilst the EU AI Act makes an encouraging effort in addressing the potential risks posed by AI systems, its provisions governing high risk AI systems require revisions to address important systematic challenges. On the one hand, the classification of AI systems as high risk potentially leaves out use cases that could pose significant risk but do not fit into the current high-risk categories.

[43] Article 29 Data Protection Working Party, "Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679" (As last Revised and Adopted on 4 October 2017) at 8. https://ec.europa.eu/newsroom/article29/items/611236, accessed 24/08/2023.

[44] GDPR Articles 35(4) & (5).

[45] Article 9 Data Protection Working Party, pp. 9-10.

[46] For a similar suggestion, see Ada Lovelace Institute, An EU AI Act that works for people and society: Five areas of focus for the trilogue (Policy Brief, 2023), 23.

The Commission's delegated power to revise the list of high-risk AI systems would be inadequate to address the challenge, as this may take time, besides the possibility that the Commission itself may fail to consider the systems as high-risk AI systems. On the other hand, use cases that do not pose significant risk could be classified as high-risk due to the Act's failure to consider specific context of use cases. To address the above challenges, this paper recommends general principles to qualify high-risk AI systems, subject to ultimate judicial review.