Andrzej Madera
Faculty of Social Sciences, University of the National Education Commission in Krakow
ul. Podchorążych 2, 30-084 Kraków
e-mail: amadera@poczta.onet.pl

THE RISK ASSESSMENT PROCESS RELATED TO THE USE OF ARTIFICIAL INTELLIGENCE SYSTEMS IN LIGHT OF THE ARTIFICIAL INTELLIGENCE ACT

Abstract. This article analyses the process of risk assessment related to the use of artificial intelligence systems under the European Union Artificial Intelligence Act. The main purpose of the article is to explain how the classification of an AI system determines the scope of legal obligations imposed on providers and deployers. The article argues that the AI Act does not establish a uniform regulatory model for all AI systems, but instead adopts a graduated, risk-based approach. Under this model, the legal consequences depend primarily on the system's intended purpose, specific use case, context of application, and potential impact on natural persons. The article concludes that risk assessment under the AI Act is not merely a technical or formal compliance exercise, but a functional and contextual legal process. Proper classification requires an active assessment by the entity intending to use the system, including an analysis of its purpose, users, data, decision-making role, and possible effects. Only such an approach allows organisations to determine whether an AI system is prohibited, high-risk, limited-risk, or minimal-risk, and to ensure its lawful, responsible, and safe use

Key words: AI Act; artificial intelligence; risk assessment; high-risk AI systems; transparency obligations.
JEL classification: D80, D81

Introduction

Contemporary artificial intelligence systems increasingly affect decision-making processes, product safety, modes of communication with humans, and the exercise of fundamental rights (Struppek et al., 2023, p. 1017–1023; Truby, 2020, p. 946–950). For this reason, the Artificial In-

telligence Act Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No. 300/2008, (EU) No. 167/2013, (EU) No. 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144, and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828, OJ L 2024/1689 of 12 July 2024, is based on a risk-based regulatory model. Under this model, the scope of legal obligations is not the same for all AI systems but depends primarily on their intended purpose, the context of their use, and their possible impact on health, safety, and individual rights (Stewart, 2024, p.122–130).

The starting point of this article is the assumption that the proper application of the AI Act requires not only knowledge of specific obligations, but above all the ability to correctly classify a given AI system into the appropriate risk category. The aim of the article is to present the risk assessment process related to the use of artificial intelligence systems in light of the AI Act. The analysis focuses on the mechanism for classifying AI systems according to their level of risk, including, in particular, prohibited systems, high-risk systems, and limited-risk systems, which are mainly subject to transparency obligations. The article also seeks to explain the significance of elements such as the intended purpose of the system and the specific case of its use for this classification.

The article primarily applies the doctrinal legal method, consisting in the analysis of the provisions of the AI Act, in particular the rules concerning prohibited practices, the classification of high-risk systems, conformity obligations, the risk management system, and transparency obligations. This method makes it possible to determine the meaning of individual concepts and obligations arising from the regulation and to identify their practical consequences for providers and entities deploying AI systems. Additionally, the analytical and functional methods are used to explain how risk-based regulation is intended to protect health, safety, individual autonomy, privacy, fairness of communication, and fundamental rights.

Legal consequences of an incorrect risk assessment of an AI system in use

Failure to properly assess the risk of an AI system in use may lead to serious legal consequences, since under the model adopted in the AI Act it is precisely the classification of the system that determines the scope of obligations imposed on the provider or deployer. An incorrect classification of a system may result in an organisation using a prohibited system, a high-risk system without fulfilling the required obligations, or a limited-risk system without ensuring the required transparency (Laux et al., 2024, p. 20–32). The most far-reaching consequences will arise where an entity incorrectly assumes that a given system may be lawfully used, even though its application falls within the catalogue of prohibited practices. In such a case, the infringement does not merely consist in the absence of a formal analysis, but in the use of a practice that the AI Act generally prohibits (Gil Gasiola, 2025).

If an incorrect risk assessment results in a system not being classified as a high-risk system, this may lead to the omission of a number of obligations provided for this category. These include, in particular, the obligation to ensure compliance with the requirements of the AI Act, to establish a risk management system, to maintain technical documentation, to ensure appropriate data quality, to keep logs, to ensure transparency of operation, human oversight, robustness, accuracy, and cybersecurity. In practice, this means that the system may be placed on the market or used without the required safeguards, without proper documentation, and without the ability to demonstrate compliance before the competent authorities. Breaches of obligations concerning high-risk systems may result in administrative fines of up to EUR 15 million or up to 3% of the total worldwide annual turnover, whichever is higher.

An incorrect risk assessment may also lead to a breach of transparency obligations. If an entity fails to recognise that an AI system directly interacts with humans, generates synthetic content, creates or manipulates deepfake content, or is used for emotion recognition or biometric catego-

risation, it may fail to provide natural persons with the required information. This creates a risk that users will be misled as to the fact that they are interacting with AI or with artificially generated or manipulated content (Gil Gasiola, 2025). Incorrect classification of the risk of an AI system in use may also have organisational and evidentiary consequences. An entity that has not carried out a proper risk assessment may be unable to demonstrate to the supervisory authority why it adopted a particular classification of the system and why it considered that certain obligations did not apply to it. This is particularly important in the case of systems listed in Annex III to the AI Act, where the provider claims that, despite being formally covered by that annex, the system does not constitute a high-risk system. In such a case, before placing the system on the market or putting it into service, the entity should document its own risk assessment and present it at the request of the competent authority (Kusche, 2024, p. 10-14; Laux et al., 2024, 3-10).

Failure to properly assess risk may also result in the need to take corrective measures, restrict the use of the system, withdraw it, suspend its use, or adapt it to the requirements of the AI Act. In practice, this may mean the need to redesign the system, implement additional safeguards, prepare documentation, carry out testing, ensure human oversight, supplement information provided to users, or change the way in which the system is used within the organisation. The consequences may go beyond administrative liability. Incorrect classification of an AI system may increase the risk of civil liability, especially where the operation of the system causes harm to a natural person, infringes their fundamental rights, leads to discrimination, results in an erroneous decision, or violates privacy or reputation (Boholm & Corvellec, 2011, p. 180-190). It may also give rise to contractual risks if the organisation assured contractors, clients, or users that the system complied with the law, and it subsequently turns out that the system was implemented without the required assessment and without fulfilling the applicable obligations. Consequently, proper risk assessment is not merely a technical or compliance-related stage, but a condition for the lawful use of an AI system. Its absence may

result in the unlawful use of the system, the omission of obligations applicable to a given risk category, financial sanctions, the obligation to suspend or modify the way the system is used, as well as liability towards persons whose rights or interests have been infringed

The risk assessment mechanism as the core of the legal structure of the AI Act

The Artificial Intelligence Act is an EU regulation that establishes harmonised rules for the use of AI systems in the EU, and its legal structure is based on a risk assessment mechanism (Paul, 2024). This concept should be understood as an organised process of identifying, analysing, and classifying threats associated with a specific activity, technology, product, or system in order to determine the level of risk they generate and the legal, organisational, or technical measures that should be applied to them (Renn, 2008, p. 23-50). In the context of the AI Act, the risk assessment mechanism means a method of classifying artificial intelligence systems according to the degree of threat they may pose to health, safety, fundamental rights, human autonomy, privacy, or the fairness of decision-making processes. This mechanism makes it possible to determine whether a given AI system belongs to the category of unacceptable risk, high risk, limited risk, or minimal risk (Struppek et al., 2023, p.1120-1130). The essence of this mechanism is that the scope of legal obligations depends on the level of risk. The greater the risk arising from the intended purpose and manner of use of an AI system, the more far-reaching the obligations imposed on the provider or deployer of the system (Paul, 2024, 1072-1080).

In practice, the risk assessment mechanism requires an examination of what the AI system is intended to be used for, the context in which it will operate, whom it will affect, whether it may influence decisions concerning natural persons, and whether it may cause harm to their rights, freedoms, health, or safety (Szadeczky & Bederna, 2025). It may therefore be assumed that the risk assessment mechanism in the AI Act is a legal and

technical process of assigning an AI system to the appropriate risk category, the purpose of which is to determine the permissibility of its use and the scope of obligations required for its lawful deployment (Japp & Kusche, 2008, p. 76-83; Renn, 2008, p. 23-50).

The risk assessment mechanism implemented in the AI Act is graduated: the greater the risk to health, safety, and fundamental rights, the stricter the obligations (Truby, 2020, p. 950-953). Thus, there is no single package of obligations "for every AI system"; these obligations depend on the purpose for which a given system is used and on the risk layer to which it is assigned. The European Commission describes this as a four-level risk model: prohibited practices, high-risk systems, systems subject to transparency obligations, and systems posing minimal or no risk (European Commission, n.d.)

Unacceptable level of risk of AI systems

The AI Act treats unacceptable risk most strictly. AI systems whose mode of operation creates a particularly high risk of violating the rights, freedoms, or safety of natural persons are prohibited. The prohibition covers placing such systems on the market, putting them into service, and their actual use. Pursuant to Article 5 of the AI Act, AI systems are prohibited if they use subliminal, manipulative, or deceptive techniques, where their purpose or effect is to materially distort the behaviour of a person or a group of persons by impairing their ability to make an informed decision. This concerns situations in which, under the influence of the system, a person makes a decision that they would not otherwise have made, and that decision causes or is likely to cause serious harm to that person, another person, or a group of persons.

AI systems that exploit the particular vulnerability of natural persons or specific groups of persons to manipulation, resulting from their age, disability, or specific social or economic situation, are also prohibited. Prohibited practices also include so-called social scoring, that is, AI systems used to evaluate or classify natural persons or groups of persons

over a certain period of time on the basis of their social behaviour, personal characteristics, or personality traits. The use of AI systems to assess or predict the risk of a specific natural person committing a criminal offence solely on the basis of profiling, personality traits, or characteristics is also prohibited. An exception applies where the AI system merely supports an assessment carried out by a human, and that assessment is based on objective and verifiable facts directly linked to criminal activity. AI systems that create or expand facial recognition databases through untargeted scraping of facial images from the internet or CCTV footage are also prohibited. This prohibition therefore covers the mass and non-selective collection of facial images for the purpose of building databases subsequently used for biometric identification (Gil Gasiola, 2025).The use of AI systems to infer the emotions of a natural person in the workplace or in educational institutions is prohibited. An exception is provided only in cases where such a system is to be deployed or placed on the market for medical or safety reasons (Kasneci et al., 2023). Biometric categorisation systems that individually classify natural persons on the basis of biometric data in order to deduce or infer information concerning their race, political opinions, trade union membership, religious or philosophical beliefs, sex life, or sexual orientation are also prohibited. However, the prohibition does not cover the lawful labelling or filtering of biometric datasets, nor certain cases of biometric data categorisation in the area of law enforcement. As a rule, the use of real-time remote biometric identification systems in publicly accessible spaces for law enforcement purposes is also prohibited. Such use may be permitted only exceptionally, where it is strictly necessary for one of the explicitly specified purposes, such as searching for specific victims of abduction, trafficking in human beings, or sexual exploitation; searching for missing persons; preventing a specific and imminent threat to life or safety; or identifying a person suspected of the most serious criminal offences. In cases where the exceptional use of real-time remote biometric identification is permitted, it may serve only to confirm the identity of a specific person being sought. It must be limited in terms of time, territory, and persons concerned, and

must also be preceded by a fundamental rights impact assessment and, as a rule, by prior authorisation from a competent court or an independent administrative authority. Nor may a decision producing adverse legal effects for a person be made solely on the basis of the output obtained from such a system.

Classification of an AI system as a high-risk system

Classifying a system at the second level, the so-called high-risk AI category, means that the system may be used lawfully, but its use entails the fulfilment of specific operational obligations (Japp & Kusche, 2008, p. 76-79). The classification of an AI system as a high-risk system first requires determining whether the system is linked to a product covered by Union harmonisation legislation. An AI system will be considered high-risk if it is intended to be used as a safety component of a product, or if the AI system itself constitutes such a product, and, at the same time, that product or the AI system as a product is subject to a conformity assessment before being placed on the market or put into service. This therefore concerns situations in which AI performs a function that is significant from the perspective of product safety, and EU law requires an external conformity assessment of that product (Szadeczky & Bederna, 2025).

Irrespective of this basic category, the AI systems listed in Annex III to the Regulation are also considered high-risk systems. These are systems used in particularly sensitive areas in which their operation may have a significant impact on the rights, freedoms, safety, or life situation of natural persons. However, the mere fact that a system falls within the scope of Annex III does not always automatically mean that it will be a high-risk system. An exception is provided for systems that do not pose a significant risk of harm to the health, safety, or fundamental rights of natural persons, in particular because they do not significantly influence the outcome of the decision-making process. A system listed in Annex III may not be considered high-risk if it performs only a limited or auxiliary function. This applies especially where the system performs a narrow proce-

dural task, improves the result of an activity previously carried out by a human, detects decision-making patterns or deviations from previous decisions without replacing or influencing human assessment without appropriate human review, or performs only preparatory tasks as part of a broader assessment relevant to the use cases listed in Annex III (Stewart, 2024, p. 130-135). An important limitation of this exception is the rule that an AI system covered by Annex III will always be considered a high-risk system if it performs profiling of natural persons. In such a case, it is not possible to rely on the argument that the system is merely auxiliary or limited in nature, because the very act of profiling natural persons in these areas has been treated as a circumstance determining high risk. If the provider of an AI system covered by Annex III considers that its system should not be classified as a high-risk system, it must prepare and document its own assessment in this regard before placing the system on the market or putting it into service. The provider must be able to demonstrate why the system does not pose a significant risk of harm to the health, safety, or fundamental rights of natural persons. It is also subject to a registration obligation and, at the request of the competent national authorities, must present the documentation of the assessment carried out.

The classification process is not entirely closed. The European Commission, after consulting the European Artificial Intelligence Board, is to provide guidelines on the practical application of these rules, including examples of AI systems that constitute high-risk cases and those that should not be classified as such. The Commission may also amend the conditions allowing certain systems to be excluded from the high-risk category if specific and credible evidence emerges concerning the actual level of risk posed by such systems. However, such amendments must not lead to a reduction in the overall level of protection of health, safety, and fundamental rights provided for in the Regulation.

High-risk AI systems must meet the specific requirements imposed on them by the AI Act. The assessment of the conformity of these systems with those requirements should take into account their intended purpose, that is, the purpose for which they were designed and are to be used, as

well as the generally acknowledged state of the art in the field of artificial intelligence and related technologies. This means that the provider cannot assess the system in isolation from its actual use or from the current technical standards applicable to a given type of technology. Ensuring the conformity of a high-risk AI system should take place with due regard to the risk management system. The provider must therefore identify, analyse, assess, and mitigate the risks associated with the operation of the AI system, and then use the results of this process to demonstrate that the system meets the requirements laid down in the Regulation. Conformity is therefore not a one-off formality, but should be linked to the ongoing and structured management of risks arising from the functioning of the system (Szadeczky & Bederna, 2025)

Risk management system for high-risk AI systems

The risk management system for high-risk AI systems must be established, implemented, documented, and maintained by the provider. It is not a one-off activity, but a continuous process carried out throughout the entire lifecycle of the AI system. This process should be regularly and systematically reviewed and updated so that it takes into account both risks known at the design stage of the system and risks that emerge later in the course of its use (Renn, 2008, p.44-50). The risk management system primarily includes the identification and analysis of known and reasonably foreseeable risks that a high-risk AI system may pose to the health, safety, or fundamental rights of natural persons when used in accordance with its intended purpose. It is then necessary to estimate and evaluate the risks that may arise both when the system is used properly and under conditions of reasonably foreseeable misuse. This means that the provider should take into account not only the ideal mode of operation of the system, but also realistic scenarios of incorrect, incomplete, or unintended use.

The risk management system should also include the assessment of other risks that may be identified on the basis of data collected as part of

post-market monitoring of the system. The provider must therefore analyse information derived from the system's real-world operation and use it to update the risk assessment. On this basis, the provider should adopt appropriate and targeted risk management measures aimed at addressing the identified threats. This process concerns only those risks that can be appropriately mitigated or eliminated through the proper development or design of the high-risk AI system, or through the provision of adequate technical information. Risk management measures should be designed with regard to the effects and possible interactions between the various requirements laid down for high-risk AI systems. The objective is to effectively reduce risk while maintaining an appropriate balance between the different regulatory obligations (Renn, 2008, p.50-53).

The measures adopted should ensure that the residual risk associated with each hazard, as well as the overall residual risk of the AI system, is assessed as acceptable. In the first place, the provider should seek to eliminate or reduce risk already at the stage of designing and developing the system, to the extent that this is technically feasible. If a given risk cannot be eliminated, appropriate mitigation and control measures should be implemented. In addition, the provider should provide the required information concerning the system and, where necessary, also ensure appropriate training for deployers of the system (Ebers, 2025, 690-695). When mitigating risk, account should be taken of the technical knowledge, experience, education, and training that may be expected of deployers of the system, as well as the intended context of its use. A system intended for specialised professional users should therefore be assessed differently from a system that will be used in a less controlled environment or by persons with limited technical preparation (Szadeczky & Bederna, 2025)

Classification of an AI system as a limited-risk system

The risk associated with maintaining and using AI systems classified as limited-risk systems consists primarily in the fact that, as a rule, such systems are neither prohibited nor subject to requirements as strict as

those applicable to high-risk systems. However, they may significantly affect people's awareness, decisions, and behaviour if the user does not know that they are dealing with artificial intelligence or with artificially generated or manipulated content (Amoozadeh et al., 2024, p. 67-70; Kasneci et al., 2023). Therefore, at this level of the risk pyramid, the main protective mechanism is not a prohibition or a full conformity assessment procedure, but the obligation of transparency. The basic threat is the risk of misleading a person as to the nature of the interaction. This applies to AI systems intended for direct interaction with natural persons, such as chatbots, virtual assistants, or other communication tools. If a person is not informed that they are communicating with an AI system, they may mistakenly assume that they are interacting with a human being. This may affect their trust, the way they formulate statements, the decisions they make, and the scope of information they disclose (Amoozadeh et al., 2024, p. 67-70). This risk is particularly important in situations where the nature of the interaction is not obvious to an average, attentive, and sufficiently informed person.

The second important area of risk is the creation by AI of synthetic content, such as texts, images, audio recordings, or videos. In this case, the threat lies in the fact that the recipient may not be able to independently recognise whether a given piece of content was created by a human or generated or manipulated by an AI system. This may lead to disinformation, manipulation of public opinion, infringement of the reputation of natural persons, and a loss of trust in authentic content. For this reason, providers of such systems should ensure that the outputs of the system are marked in a machine-readable format and are detectable as artificially generated or manipulated (Renn, 2008, p.23-28). A particularly sensitive form of this risk is deepfake content, that is, artificially generated or manipulated images, audio recordings, or videos that may appear authentic. The risk here consists in the possibility of attributing to a specific person statements, behaviour, or an image that did not in fact occur. Therefore, deployers of such systems should disclose that the content has been artificially generated or manipulated. In the case of artistic, satirical, fictional,

or creative content, this obligation may be fulfilled in a manner that does not interfere with the use of the work, but it should still inform the recipient of the artificial nature of the content. Limited risk also includes emotion recognition systems and biometric categorisation systems, insofar as their use is not prohibited or classified as high-risk. In their case, the problem is the possibility of interference with a person's privacy, bodily integrity, and personal data. A person in relation to whom such a system is used should know that their emotions are being analysed or that biometric categorisation is taking place. The absence of such information could lead to covert observation, assessment, or profiling, as well as to a breach of personal data protection principles. Another threat is the publication of texts generated or manipulated by AI for the purpose of informing the public about matters of public interest. In such a case, the risk is that recipients may regard the content as fully human, editorial, or expert in nature, even though it was generated by an AI system. This may affect public debate, democratic processes, and trust in information. Therefore, as a rule, it should be disclosed that the text has been artificially generated or manipulated, unless it has been subject to human verification or editorial control and editorial responsibility is borne by a specific natural or legal person (Ebers, 2025, p.688-690).

The essence of the obligations concerning limited-risk systems is therefore to ensure that a natural person receives clear, explicit, and accessible information, no later than at the time of the first interaction or first use of the system, that they are dealing with AI, an emotion recognition system, biometric categorisation, or artificially generated or manipulated content. The risk of these systems usually does not lie in the fact that their use is inadmissible, but in the fact that, without appropriate transparency, they may operate in a non-transparent, misleading, or difficult-to-recognise manner for humans (Renn, 2008, p. 44-46).

Consequently, the third level of the risk pyramid covers systems that may be used lawfully, provided that information obligations are fulfilled. The purpose of these obligations is to protect human autonomy, trust in information, privacy, personal data protection, and the fairness of com-

munication. Transparency is intended to enable the recipient to make an informed assessment of the situation: whether they are communicating with a human or a machine, whether they are viewing authentic content or a deepfake, whether their emotions or biometric characteristics are being analysed, and whether a text concerning public matters has been generated or modified by AI (Ebers, 2025, p.688-690)

The process of assigning an AI system to a specific risk category

Before starting to use an AI system, in order to assess the risk and assign the system to a specific category, the deployer should determine the answers to the following questions: whether the use of a particular system is prohibited, whether it is a high-risk system, and whether it is subject to the obligations set out in Article 50 of the AI Act. Answering these questions makes it possible to determine the appropriate scope of obligations that must be fulfilled when using the system. In practice, therefore, the commencement of the use of a particular AI system should be preceded by a thorough analysis and its assignment to the appropriate category from the perspective of the risk it may generate. This analysis should concern what the AI system planned for use is intended for, the context in which it is to be used, and the effect it is intended to have on people. Under the AI Act, classification is determined primarily by the intended purpose and the specific use case, rather than by the technology itself, the brand of the model, or its level of "advancement" (Burt et al., 2018).

The preliminary stage should consist in assessing whether the system in question is an AI system within the meaning of the AI Act. The definition contained in Article 3 of the AI Act serves to distinguish ordinary software from systems covered by the AI Act. In practice, this analysis will involve answering the following questions: whether the system operates with a certain level of autonomy, whether, on the basis of input data, it generates predictions, content, recommendations, or decisions, and whether it affects the environment in which it operates.

The normative definition of an AI system is laid down in Article 3(1) of the Artificial Intelligence Act laying down harmonised rules on artificial intelligence and amending certain regulations. According to this definition, an AI system means a machine-based system that is designed to operate with varying levels of autonomy after deployment, that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers from the input it receives how to generate outputs such as predictions, content, recommendations, or decisions that may influence physical or virtual environments (European Commission, n.d.; Kozłowski, 2024, p. 20-25).

The next stage should be the assessment of the "intended purpose" of the specific AI system. This issue comes down to answering the question: what is the AI system to be used for in a given organisation? The same model may have a different classification due to the level of risk it generates depending on the situation in which it will be used. For example, an AI system used as a simple assistant for drafting texts will not generate significant risk, but it may be classified as a high-risk system when used as a tool for pre-selecting job candidates (Burt et al., 2018; Paul, 2024, p. 1073-1078).

Conclusion

In summarising the foregoing considerations, it should be stated that the process of risk assessment related to the use of artificial intelligence systems under the AI Act is of fundamental importance for correctly determining the scope of legal obligations imposed on entities involved in the creation, deployment, and use of AI systems. The Artificial Intelligence Act does not adopt a uniform regulatory model for all AI systems, but is based on a graduated approach in which the intensity of obligations depends on the level of risk generated by a given system, its intended purpose, and the specific context of its use. The analysis shows that the key element in the classification of an AI system is determining its actual or intended use. Classification is not determined by the technology itself,

the brand of the model, its level of advancement, or its general possibility of use in various areas, but primarily by its intended purpose, the specific use case, and its potential impact on natural persons. The same system may therefore be subject to different obligations depending on whether it is used as an auxiliary tool, a system supporting communication, a solution used to make decisions concerning people, or a component of a product relevant from the point of view of safety (Hacker et al., 2023, p.1115-1120).

In response to the main research question posed in the introduction, it should be indicated that the risk assessment process for an AI system should proceed in stages. First, it is necessary to determine whether a given solution constitutes an AI system within the meaning of the AI Act. Next, it should be assessed whether its use falls within the catalogue of prohibited practices, and thus whether it generates unacceptable risk. If not, the next step should be to examine whether the system meets the criteria for being recognised as a high-risk system, in particular because of its connection with a product covered by Union harmonisation legislation or because it falls within the cases listed in Annex III. Only thereafter should it be determined whether the system is subject to transparency obligations applicable to limited-risk systems, especially where it directly interacts with humans, generates synthetic content, enables the creation of deepfakes, or is used for emotion recognition or biometric categorisation (Ebers, 2025, p. 689-695).

The conclusion arising from the analysis of the rules on unacceptable risk is that the AI Act sets a boundary for applications of artificial intelligence that it considers unacceptable from the perspective of human protection. This applies in particular to manipulative systems, systems exploiting the vulnerabilities of natural persons, systems used for social scoring, unlawful criminal profiling, mass scraping of facial images, prohibited biometric categorisation, and real-time remote biometric identification in publicly accessible spaces, except in strictly defined cases. In this area, the EU legislator does not limit itself to imposing precautionary

obligations, but introduces a prohibition of specific practices as contrary to the fundamental values protected by Union law.

With regard to high-risk systems, the basic conclusion is that their use is, as a rule, permissible, but requires the fulfilment of extensive organisational, technical, and documentation obligations. Of particular importance here is the obligation to ensure the system's compliance with the requirements of the AI Act and to establish a risk management system. Risk management cannot be understood as a formal document prepared solely for the purpose of system deployment. It should constitute a continuous, iterative process covering the identification, analysis, assessment, mitigation, and monitoring of risks throughout the entire lifecycle of the AI system. Only such an approach makes it possible to take into account both risks foreseeable at the design stage and those that emerge only during the system's actual use (Ebers, 2025, p. 689-695).

The analysis of limited-risk systems leads to the conclusion that their main regulatory problem is not the mere permissibility of their use, but the risk of a lack of awareness on the part of humans (Christoffersen, 2018, p.1236-1238). These systems may influence the manner of communication, trust in information, the assessment of content authenticity, privacy, and users' decisional autonomy. Therefore, the main protective instrument consists of transparency obligations. Their purpose is to ensure that a natural person knows that they are interacting with AI, that a given piece of content has been artificially generated or manipulated, or that an emotion recognition or biometric categorisation system is being used in relation to them. Transparency does not eliminate risk entirely, but it enables the recipient to make an informed assessment of the situation and reduces the possibility of misleading them.

The considerations presented also show that the proper classification of an AI system requires active conduct on the part of the entity intending to use it. The deployer should not limit itself to the provider's general assessment or to marketing declarations concerning the system. It should independently determine the purpose for which the system will be used in the organisation, what decisions or processes it will support, in relation to

which persons it will be used, what data it will process, and what effects it may produce. Only such an analysis makes it possible to correctly determine whether the system is prohibited, high-risk, limited-risk, or falls within the category of minimal or insignificant risk (Hacker et al., 2023, p. 1115-1122).

Finally, it should be stated that the risk assessment model adopted in the AI Act is functional and contextual in nature (Christoffersen, 2018, p.1232-1237). Its purpose is not to inhibit the development of artificial intelligence, but to organise its use in a manner proportionate to the threats it may generate (Stewart, 2024, p.124-126). The greater the potential impact of an AI system on health, safety, fundamental rights, or the life situation of natural persons, the more far-reaching the obligations imposed by the Regulation. In practice, this means that every organisation planning to use an AI system should implement an internal process of classification and risk assessment, covering the analysis of the definition of an AI system, the system's intended purpose, the catalogue of prohibited practices, the criteria for high-risk systems, and transparency obligations (Renn, 2008, p.28-34). Only such an approach enables the lawful, responsible, and safe use of artificial intelligence under the conditions set out by the AI Act.

References

AMOOZADEH M. DANIELS D. NAM D. KUMAR A. CHEN S. HILTON M. et al., (2024). Trust in Generative AI among students: An exploratory study, *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1*, 67–73.

BOHOLM, Å., CORVELLEC H. (2011). A Relational Theory of Risk. *Journal of Risk Research* 14 (2): 175–190. https://doi.org/10.1080/13669877.2010.515313

BURT A., LEONG B., SHIRRELL S. WANG X.(2018). Beyond Explainability: A Practical Guide to Managing Risk in Machine Learning

Models . *Future of Privacy Forum*. https://fpf.org/wp-content/uploads/2018/06/Beyond-Explainability.pdf

EBERS M. (2025) Truly Risk-based Regulation of Artificial Intelligence How to Implement the EU's AI Act. *European Journal of Risk Regulation*.16 (2), 684-703. doi:10.1017/err.2024.78

GASIOLA G.G. (2025), Rebuilding the pyramid: The AI Act's risk-based approach using a binary decision diagram, *Computer Law & Security Review*, 58, 106189, https://doi.org/10.1016/j.clsr.2025.106189.

HACKER P. ENGEL A. MAUER M. (2023). Regulating ChatGPT and other large generative AI models, *Proceed-ings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*, 1112–1123, https://doi.org/10.48550/arXiv.2302.02337

CHRISTOFFERSEN, M. G. (2018). Risk, Danger, and Trust: Refining the Relational Theory of Risk. *Journal of Risk Research* 21 (10): 1233–1247. https://doi.org/10.1080/13669877.2017.1301538.

JAPP, K. P., KUSCHE I.(2008). "Systems Theory and Risk." In *Social Theories of Risk and Uncertainty. An Introduction*, edited by Jens O. Zinn. 76–103. Malden, MA: Blackwell.

KASNECI E. SEßLER K. KÜCHEMANN S. BANNERT M. DEMENTIEVA D. FISCHER F. et al. (2023). ChatGPT for good? On opportunities and challenges of large language models for education, *Learning and Individual Differences*, 103, Article ID 102274. DOI:10.1016/j.lindif.2023.102274

KUSCHE, I. (2024). Possible harms of artificial intelligence and the EU AI act: fundamental rights and risk. *Journal of Risk Research*, 1–14. https://doi.org/10.1080/13669877.2024.2350720

LAUX, J., WACHTER, S. AND MITTELSTADT, B. (2024). Trustworthy artificial intelligence and the European Union AI act: On the conflation of trustworthiness and acceptability of risk. *Regulation & Governance*, 18, 3-32. https://doi.org/10.1111/rego.12512

PAUL, R. (2024). European artificial intelligence "trusted throughout the world": Risk-based regulation and the fashioning of a competitive

common AI market. *Regulation & Governance*, 18, 1065-1082. https://doi.org/10.1111/rego.12563

RENN O., (2008). Risk governance: Coping with uncertainty in a complex world (2nd ed.), *Earthscan*

STRUPPEK L., HINTERSDORF D., FRIEDRICH F., SCHRAMOWSKI P., KERSTING K., (2023). Exploiting cul-tural biases via homoglyphs in text-to-image synthesis, *Journal of Artificial Intelligence Research*, 78, 1017–1068.

TRUBY J., (2020). Governing artificial intelligence to benefit the UN Sustainable Development Goals, *Sustainable Development*, 28(4), 946–959, https://doi.org/10.1002/sd.2048.

SZADECZKY, T., BEDERNA, Z. (2025). Risk, Risk ,regulation, and governance: evaluating artificial intelligence across diverse application scenarios. Security Journal 38, 35 https://doi.org/10.1057/s41284-025-00495-z

STEWART L.S. (2024). The regulation of AI-based migration technologies under the EU AI Act: (Still) operating in the shadows? *European Law Journal* 30 (1-2),122-135. doi:10.1111/eulj.12516