Acta Univ. Sapientiae, Legal Studies, **11**, 2 (2022) 19–34

DOI: 10.47745/AUSLEG.2022.11.2.02

The European Approach to Artificial Intelligence. Ethical and Regulatory Implications[1]

Kitti MEZEI
PhD, Research Fellow
Centre for Social Sciences (Budapest, Hungary),
Institute for Legal Studies
Assistant Professor
Budapest University of Technology and Economics (Hungary),
Faculty of Economics and Social Sciences, Business Law Department
Researcher
University of Public Service (Budapest, Hungary),
Eötvös József Research Centre, Cybersecurity Research Institute
e-mail: mezei.kitti@tk.hu

Anikó TRÄGER
PhD, Junior Research Fellow
Centre for Social Sciences (Budapest, Hungary),
Institute for Legal Studies
Assistant Lecturer
Budapest University of Technology and Economics (Hungary),
Faculty of Economics and Social Sciences, Business Law Department
e-mail: trager.aniko@gtk.bme.hu

Abstract. The development of artificial intelligence (AI) must ensure human-centred and ethical operations, transparency and respect for fundamental rights. In addition to its obvious benefits, AI also entails a number of risks such as opaque decision-making. The aim of this paper is to present and analyse in detail the legal environment for AI in the European Union, with a particular focus on the principles and directives, as well as the current and possible future legal framework, the draft EU AI Act. The article discusses the concept and framework of the EU AI Act on artificial intelligence. A separate chapter reviews the risk-based approach at the heart of the regulation. It provides a

1 The publication was supported by the European Union project RRF-2.3.1-21-2022-00004 within the framework of the Artificial Intelligence National Laboratory and by the Ministry of Innovation and Technology NRDI Office within the framework of the FK_21 Young Researcher Excellence Programme (138965).

detailed analysis of the systems categorzed by risk, their requirements, and the regulatory solutions developed by the draft.

Keywords: artificial intelligence, EU AI Act, trustworthiness, risk-based approach, high-risk AI

1. Introduction

The application of Artificial Intelligence (AI) is expanding into ever more areas of life (e.g. it can improve healthcare, help law enforcement authorities fight crime more effectively, make transport safer, or even help detect fraud and cybersecurity threats, etc.). It is therefore undoubtedly one of the biggest challenges of our time, both from an economic and a regulatory perspective. This is illustrated by the publication by the EU Commission in 2020 of a White Paper on Artificial Intelligence, which forms the basis for specific regulation of AI development and applications at the EU level.[2] It sets out that AI can significantly impact our society, that it is necessary to build trust and confidence in it, and that the AI sector must be based on fundamental rights and values such as human dignity and privacy. Human-centred AI assumes technology that people trust because it aligns with the values underpinning human societies. Non-binding soft law solutions, such as ethics guidelines for AI, play a crucial role in establishing trustworthy AI, assessing risks, and managing the technology in a regulatory context.[3] These are embedded in regulation and can mitigate risks during the legislative process. Ethics by design is an approach to ensure that ethical requirements are appropriately considered in developing an AI system or technique. It aims to address ethical issues at the earliest stages of development rather than as an afterthought. In addition, this trend can have a positive cultural impact, particularly in the technology industry, where market leaders seek to get ahead of regulation rather than being left behind, designing their products and services to comply with legislation still in draft form.[4]

In the general design of AI regulation, four main ethical directions should be highlighted, as set out in the ethical guidelines of the High-Level Expert Group on AI: respect for human autonomy: do not control/manipulate humans, do not compromise democratic processes; prevention of harm: including resistance to unintended external influences that may result in harm; fairness: the development and use of AI systems should be fair; explainability: means transparency of

2 European Commission 2020a.

3 See the criticism of ethical principles in Héder (2020. 57) and Hagendorff (2020. 99).

4 An example of this is the Netherlands, which has already started to apply rules similar to the draft regulation, even though the regulation is not expected to enter into force until a later date. Bertuzzi 2022.

operation[5] (trusted AI systems can be traced and their decisions explained, in particular users should be informed that they have interacted with an AI system and also how the AI system works, what its capabilities are, in what ways and how reliably it uses the datasets provided to it). Other requirements include: human empowerment and human oversight; technical stability and security; data protection and management; diversity, non-discrimination, and equity; social and environmental well-being; accountability.[6]

Technology can be the target of regulation and a tool, even embedded in technology as a command. This encourages developers to address regulation by design at the early stages of development. For example, the White Paper states that AI systems are expected to have built-in safety and security mechanisms to ensure that any operation carried out by the system is demonstrably safe for the physical and mental well-being of the individuals involved. The European Union's regulation[7] points in this direction in several digital regulatory areas (e.g. data protection[8] and algorithmic trading[9]). Traditionally, technology vendors have tested their products ex-post after the risk has materialized. They should have taken measures to correct their processes and compensate for any damages if and when liability was found. This reactive model, which has always struggled to keep pace with technological developments, is becoming obsolete. Instead, legislators are encouraging companies to set up compliance teams around 'product advisors' and to take account of the harm and risks posed by a product at an early stage, to carry out an ethical and regulatory risk assessment. However, such regulation is flexible, requiring standards such as 'appropriate technical and organizational measures' that can be adapted to the company or product/service. Privacy by design and privacy by default are key concepts and are now the bases of

5 For example, in this context, a new draft transparency standard, IEEE P7001, is now available, one of the P70XX series of 'human standards' that are emerging from the IEEE Standards Association's global initiative on the ethics of autonomous and intelligent systems. P7001 aims to create a standard that has 'measurable, testable levels of transparency, so that autonomous systems can be objectively assessed, and levels of compliance determined'. P7001 is also generic in nature; it aims to be applicable to all autonomous systems, including robots (autonomous vehicles, assistive robots, drones, robotic toys, etc.) as well as software-only AI systems such as AI used in medical diagnostics, chatbots, facial recognition systems, etc. P7001 defines five different groups of stakeholders, and AI systems must be transparent to each group in different ways and for different reasons. Winfield et al. 2021.

6 European Commission High-Level Expert Group on Artificial Intelligence (AI HLEG) 2019. 14.

7 See Hanani 2022. 137; Codagnone et al. 2022; Mökander et al. 2022. For a comparison with Chinese AI regulation, see also Roberts et al. 2021. 3659–3677.

8 With privacy by design, privacy safeguards must be built into products and services from the earliest stages of development. In other words, companies need to think about security measures at the design stage of data management processes before they start processing data. For example, pseudonymization, or encryption of personal data, is one way of ensuring compliance with *built-in* data protection.

9 Directive 2014/65/EU requires Member States to ensure that algorithmic trading systems do not create or contribute to disorderly trading conditions in the market and to address disorderly trading conditions that such algorithmic trading systems do create.

digital regulation. Engineers and developers need to address legal and regulatory requirements from the very beginning of the design of their digital products.[10]

Furthermore, in 2020, the European Parliament issued a report to the Commission with recommendations on the civil liability regime for AI.[11] In response, in September 2022, the Commission took the initiative to modernize the rules on the objective liability of manufacturers for defective products (from smart technology to pharmaceuticals).[12] The revised rules aim to create legal certainty for businesses, making investing in new and innovative products easier. They will ensure fair compensation in the event of damage caused by a defective product, including a digital or refurbished product. On the other hand, the Commission has proposed a targeted harmonization of national liability rules for AI for the first time.[13] A single set of rules would make it easier for victims of damage caused by AI to get compensation.[14]

The most important step forward in the comprehensive regulation of AI is the publication in April 2021 of the Commission's proposal for a draft Artificial Intelligence Act (hereinafter as AI Act), which contains important restrictions on AI systems used in or in connection with the EU.[15] The use of AI with specific characteristics, such as opacity due to the black box effect, complexity, dependence on data, and autonomous behaviour, may adversely affect several fundamental rights enshrined in the Charter of Fundamental Rights of the European Union. Because of these characteristics, both public authorities and the individuals concerned may lack adequate means to verify how a given algorithmic decision was made and whether the relevant rules have been respected. Therefore, the proposal aims to ensure a high level of protection of these fundamental rights and to address the different sources of risk through a clearly defined risk-based approach. This paper analyses the AI Act in detail.

2. The European Union's Draft AI Act

2.1. Scope of the AI Act and Definition of the AI System

The draft EU AI Act aims to implement a minimum set of horizontal rules applicable to all AI systems placed on the market or used in the EU. The new regulation would apply primarily to service providers established in the EU or third countries placing

10 82 Clarke 2022.
11 European Parliament 2020.
12 European Commission 2022a.
13 European Commission 2022b.
14 European Commission 2022c.
15 It should be noted that the AI Act should be read in conjunction with other major legislative packages, such as the Digital Services Regulation (DSA), the Digital Markets Regulation (DMA), and the Digital Governance Regulation (DGA), the first two of which primarily regulate large commercial online platform providers such as Google, Amazon, Facebook, and Apple (GAFA).

AI systems on the EU market or installing them in the EU and to users of AI systems located in the EU.[16] To prevent circumvention of the regulation, the new rules would also apply to providers and users of AI systems located in third countries if the output produced by these systems is used in the EU. However, the draft regulation would not apply to AI-based systems developed or used exclusively for military purposes, to authorities in third countries, international organizations, or to authorities using AI-based systems in the framework of international agreements on law enforcement and judicial cooperation. Another exemption has been added for people using AI for non-professional purposes, which would fall outside the scope of the AI regulation, except for the transparency obligations.[17]

The Commission proposes a technology-neutral definition of AI in Article 3(1) of the draft, which states that an AI system is an 'artificial intelligence system' meaning 'software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with'. Accordingly, the term AI system would refer to software-based technologies that include machine learning, logic and knowledge-based systems, and statistical approaches. This broad definition includes AI systems that can be used independently or as part of a product. An AI system can be designed to operate with varying degrees of autonomy. It can be used standalone or as part of a product, whether the system is physically integrated into the product (embedded) or serves the functions of the product without being integrated (non-embedded). The AI Act aims to be future-proof and cover current and future AI technology developments. To this end, the Commission would – using delegated acts (Article 4) – add new approaches and techniques for AI regulation to the list in Annex I as needed. Furthermore, Article 3 contains a long list of definitions, including the concepts of 'provider' and 'user' of AI systems, covering both public and private entities, as well as 'importer', 'distributor' and 'operator', 'sentiment recognition', and 'biometric categorization'.

2.2. The Risk-Based Approach

The use of AI, with its specific characteristics, can adversely affect several fundamental rights and the security of users. To address this, the AI Act adopts a risk-based approach, whereby AI applications are classified into risk classes, and

16 See Article 2. The AI Act would also apply to EU institutions, offices, bodies, and agencies acting as providers or users of AI systems.

17 Some members of the Council and the European Parliament would extend this by excluding from the scope of the regulation AI systems where national security issues are at stake. This would allow (autocratic) governments to use biometric mass surveillance or social scoring in the name and under the guise of 'national security' even if these are prohibited by the regulation. Bertuzzi 2022.

legal action is tailored to the specific risk level.[18] To this end, a distinction is made between unacceptable, high-risk, moderate-risk, and minimal-risk AI systems. Under this approach, AI applications would be regulated only to the extent strictly necessary to address specific risk levels.

2.2.1. AI Systems Falling into the Prohibited Category

With this in mind, the AI Act distinguishes a completely prohibited category (Title II), which includes the prohibition of facial recognition[19] (with exceptions)[20] in public places, subliminal manipulation, mass surveillance, or the unlawfulness of the social scoring system[21] (similar to the one used in China). All AI systems that clearly threaten people's safety, livelihoods, and rights are banned, from social scoring by governments to voice assistant games that encourage dangerous behaviour. Of these, subliminal manipulation has been the most criticized because the draft does not provide a precise definition of what should be understood by this or what cases might fall into this category. According to the literature, it generally refers to sensory stimuli that consumers cannot consciously perceive; for example, visual stimuli that last less than 50 milliseconds. However, most applications of AI will not be subliminal, as users will perceive it consciously. Thus, the AI Act in its current form still allows for many forms of AI-based manipulation.[22]

2.2.2. Moderate-Risk AI Systems

In addition, it identifies high-risk AI applications (Title III), for which it establishes binding rules, and other applications that are less risky (Title IV) but still deserve some attention, and it addresses the risks associated with these applications by supporting them with provisions to enhance transparency. These rules are contained in Article 52, which requires the AI to inform the person at all times that s/he is facing an AI. Systems capable of detecting emotions must inform the persons concerned, deepfake videos must be labelled, and it must be known that they are machine-forged moving images. These categories are neither prohibited nor high-risk in themselves. Interestingly, the AI Act classifies tools used by law

18 For more on this, see Mahler 2021.

19 For more on the regulation of facial recognition programmes in the EU, see Madiega–Mildebrath 2021.

20 The use of AI systems for the 'real-time' remote biometric identification of natural persons in publicly accessible locations for law enforcement purposes necessarily involves the processing of biometric data. The rules of the AI Act, based on Article 16 TFEU, which prohibit such use, subject to certain exceptions, should be applied as *lex specialis* to the rules on the processing of biometric data contained in Article 10 of Directive (EU) 2016/680, and therefore exhaustively regulate such use and the processing of the biometric data concerned.

21 See AlgorithmWatch 2022.

22 See Franklin et al. 2022. 35; Vergnolle 2021; Hacker 2021.

enforcement agencies to detect deepfake as high-risk, while deepfake content in general falls into the low-risk category. This is a curious discrepancy, which appears to be based on the assumption that deepfake technologies (which are mainly used in the private sector for the time being) constitute a lower risk than deepfake detection AI systems in the hands of state actors. However, under the AI Act, this labelling obligation does not apply to law enforcement. This means that when some law enforcement authorities use deepfake, they do not have to label it as such [Article 52(3)].[23] Biometric categorization systems – systems that biometrically group individuals according to categories such as 'gender, age, hair colour, eye colour, tattoos, ethnic origin, sexual or political orientation, based on their biometric data' – or emotion recognition systems, which are used in the context of Article 3(34), are not prohibited and are not included in the list of high-risk AI systems. Consequently, they fall into the category of AI systems of limited risk and are therefore covered by the provisions of Article 52(2) for both public and private actors, with the exception of law enforcement authorities. Finally, the draft leaves AI applications not falling into either category to regulation by codes of conduct, i.e. self-regulation. So, the AI Act does not contain specific rules for the use of AI, which is neither prohibited nor high-risk in itself, beyond the basic requirements, but would refer it to the so-called codes of conduct.[24] The use of codes of conduct is not new among the European Union's regulatory solutions. It is currently used, or more precisely required, in the field of media regulation, but the Digital Service Act[25] will also further strengthen its role in media regulation. In the case of the already cited DSA, the Regulation already makes it clear in the recitals[26] that the application of codes of conduct is to be made in the context of self- and co-regulation, similar to the current media regulation.

Although the self-regulation and co-regulation models have very similar elements, it is necessary to briefly distinguish between the two regulatory models. Self-regulation in principle does not have a mandatory nature. The actors in the field or some joint organization of such actors develop some professional-ethical standards. Those who consider it essential will voluntarily join this agreement and accept these standards as binding on themselves, possibly jointly handling complaints about them. Although there is no classical binding force of the state behind the regulation, if the participants mean what they say in their code of conduct, it can

23 Georgieva et al. 2022. 14.

24 Article 69(1) of the Proposal for a Regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts (AI Act). See European Commission 2021.

25 European Commission 2020b.

26 Recital (68) of Proposal for a Regulation of the European Parliament and of the Council on a Single Market for Digital Services (Digital Services Act) and amending Directive 2000/31/EC. European Commission 2020b.

significantly help achieve a high level of compliance with professional, ethical principles.[27]

Co-regulation generally builds on the self-regulation of organizations described above. However, in this case, the state is already involved in the regulation. The responsibility for compliance with and enforcement of the principles of the codes of conduct is shared between the state and the professional organizations involved in the regulation.[28] The classic legislative model is that the state (or, in this case, the European Union) creates a binding rule, and then the public authority enforces it. The stakeholders are only involved in the preparation and impact assessment of the legislation. By contrast, the central actor in the process of co-regulation, and thus in the drafting and applying of codes of conduct, remains the one to whom their content will otherwise be binding. Self- and co-regulation also provide a more flexible regulatory mechanism that considers regional specificities.

Going back to the text of the AI Act, its recitals[29] and the provisions of Article 69 only provide for the encouragement and support of the adoption of codes of conduct in general. This leads to the conclusion that the aim of EU legislation on regulating AI is essentially to promote self-regulation and that there is currently no thought of developing a co-regulatory model.

The AI Act provides in Article 69(1) that:

> The Commission and the Member States shall encourage and facilitate the development of codes of conduct aimed at promoting the voluntary application of the requirements set out in Chapter 2 of Title III to AI systems other than high-risk AI systems based on technical specifications and solutions which, in the light of the intended use of the systems, constitute an appropriate means of ensuring compliance with those requirements.

The quoted text of the regulation refers to the mandatory requirements for using high-risk AI. On this basis, the draft would aim to ensure that they are also applied to the highest possible degree in the case of lower-risk AI.

In this context, the Commission and the Council would therefore encourage AI actors and their organizations to adopt and implement codes of conduct setting out requirements in the areas of the risk management system, data, data governance, technical documentation, record keeping, transparency and provision of information to users, human oversight, accuracy, robustness and cybersecurity, in line with the AI Act.[30]

27 Examples of how self-regulation works can be found in the field of media regulation. On this, see Tófalvy 2013. 85–87.

28 Hegedűs 2015, Csink–Mayer 2012.

29 Recital (81) of the AI Act.

30 Title III, Chapter 2 of the AI Act.

In addition to the general support for self-regulation, the draft regulation also sets out, by way of example, a list of areas where it considers particularly important for non-high-risk schemes to develop a code of conduct in which they accept to be bound by more stringent provisions. Examples of such areas include the promotion of the voluntary application of requirements relating to environmental sustainability, accessibility for persons with disabilities, stakeholder participation in the design and development of AI schemes, and diversity of development teams to AI schemes based on clear objectives and key performance indicators to measure the achievement of those objectives.[31] As to who is entitled to adopt codes of conduct, the AI Act – similarly to the current models of media regulation – designates the regulated parties themselves, i.e. the individual providers of the AI systems or the organizations representing them, or both, included through the involvement of users and stakeholders and their representative organizations.[32] The draft regulation also states that 'codes of conduct may cover one or more AI systems, taking into account the similarity of the intended purpose of the relevant systems.'

Regarding codes of conduct, the draft also briefly states that the Commission and the Council will consider the specific interests and needs of small service providers and start-ups in encouraging and facilitating their development.[33]

The motivation behind the EU's move towards self-regulation is partly the time factor: AI, like the media, is a fast-moving field with many different areas. Classical legislative instruments are slow at the Member State level, but even more so in the EU. We should think here of the AI Act itself, which has been years in the making and is still only a draft. Furthermore, if market players can be involved in developing regulation based on their existing self-regulation and ethical principles, this will allow for significantly faster adaptation. Developers and professional organizations involved in self-regulation have the expertise and knowledge to develop and, where appropriate, monitor the principles. Greater acceptance and cooperation can be expected if they are involved in regulation. A further advantage could be that if AI developers move to codes of conduct under the Regulation, this could lead to more effective, detailed regulation than the current codes of ethics, which may or may not have any substance to them.[34]

In addition to the expected positive aspects, it is also necessary to briefly discuss the disadvantages of self-regulation. One of the main disadvantages of this model is that participation in it and adherence to codes of conduct is entirely voluntary. As a result, it ultimately lacks the classic binding force and enforceability. In line with this, it is also apparent that the AI Act – applying the risk-based approach here, too – has not opted for this regulatory model for higher-risk schemes but

31 Article 69(2) of the AI Act.
32 Article 69(3) of the AI Act.
33 Article 69(4) of the AI Act.
34 For more details on the criticisms of these, see: Ződi 2020; Larsson 2020. 437–451.

for 'classic' mandatory regulation, with the threat of heavy fines in the event of non-compliance.[35]

The question remains, then, how interested will AI developers – who are not required to apply the stricter rules of the Regulation – be to even adopt the much more stringent rules that are mandatory for high-risk AI. In the light of this development, it is also questionable whether the European Union will leave this area to self-regulation at all or whether it will move towards co-regulation, as in the case of the media, or whether it will return to the classic centralized regulatory solution but with less stringent rules for lower-risk areas.

2.2.3. High-Risk AI Systems

Proposed rules for high-risk AI are of interest because most of the provisions of the new regulation are built around this risk category. An AI system is considered high-risk either because it is a security component of an already tightly regulated product group (listed in Annex II, from toys through craft to medical instruments) or because it is used in an area where human rights are particularly at risk. The latter list includes two dozen specific applications in eight areas such as AIs for biometric identification of natural persons, AIs for the control of critical infrastructures (transport, gas, water, electricity), and some other AIs 'active' in various areas (such as recruitment, university admissions, credit assessment, and advice to judges). Indeed, the AI Act states that AI systems used in employment, management of workers and access to self-employment, in particular recruitment and selection of persons, decisions on promotion and dismissal, and the allocation of tasks to persons with a contractual relationship to work, as well as the monitoring or evaluation of such persons, should be considered as high-risk, as they may have a significant impact on the future career prospects and livelihood of these persons. A significant power imbalance characterizes law enforcement authorities' actions involving specific AI systems. They may lead to the surveillance, arrest, or privation of liberty of a natural person and other adverse effects on fundamental rights. In particular, if an AI system is not trained with good-quality data, does not meet adequate standards of accuracy or stability, or is not properly designed and tested before being placed on the market or otherwise put into service, it may select people in a discriminatory or otherwise unfair or unjust manner. It may also hinder the enforcement of important fundamental procedural rights such as the right to an effective remedy and a fair trial, as well as the rights to defence and the presumption of innocence, in particular, if such AI systems are not sufficiently transparent, explained, and documented. The AI systems used in migration management, asylum, and border management[36] affect people who are often in

35 Article 71 of the AI Act.

36 See Dumbrava 2021.

a particularly vulnerable situation and whose lives are affected by the outcome of the actions of the competent authorities. The accuracy, non-discriminatory nature, and transparency of the AI systems used in this context are therefore of particular importance in ensuring respect for the fundamental rights of the persons concerned, namely their rights to free movement, non-discrimination, privacy and protection of personal data, international protection, and due process. In previous compromises, the EU Council already moved towards curbing significant leeway for law enforcement. The new text extends the exemption to the four-eye principle, which requires at least two persons to verify a decision of a high-risk system. Moreover, public authorities using high-risk systems in law enforcement, migration, asylum and border control, and critical infrastructure have been exempted from registering on the EU database.[37]

Specific AI systems designed to administer justice and democratic processes should be considered high-risk, considering their significant impact on democracy, the rule of law, individual freedoms, and the right to an effective remedy and a fair trial. In particular, to address the risk of possible distortions, errors, and opacity, AI systems that aim to assist judicial authorities in researching and interpreting factual and legal elements and in applying the law to specific facts should be considered high-risk. However, this classification should not cover AI systems intended for purely ancillary administrative activities that do not affect the actual administration of justice in individual cases such as anonymization or pseudonymization of court decisions, documents or data, staff communications, etc., administrative tasks, or the allocation of resources. For the use of high-risk systems in this area, Member States might decide to appoint police forces or judicial authorities as market surveillance authorities. The text now specifies that such market surveillance activities should not affect the independence of the courts. Systems for pollution control have been removed from the list of high-risk use cases, while systems to calculate risks and pricing for insurance have been added, except if the provider is an SME.

The requirements for high-risk AIs in the regulation (Chapter 2) provide that risk assessment systems must always be established, implemented, documented, and maintained (Article 9). They must be operated in conjunction with appropriate data governance systems, and the data used for teaching the AI, validation, and testing must be 'clean' (Article 10). High-risk AIs must be accompanied by detailed documentation, and event logging systems must be associated (articles 11–12). Systems of this type must operate transparently and always retain human oversight and intervention (articles 13–14). They must also meet the requirements of accuracy, robustness, and cybersecurity (Article 15). Most of these requirements must be incorporated into the design of the high-risk AI system. In addition to the technical documentation to be prepared by the service provider, the other requirements

37 Bertuzzi 2022.

should be taken into account at the earliest stages of the design and development of the AIs. A new transparency obligation has been added, requesting the providers of systems susceptible of causing significant harm to include the expected output in the instructions for use when appropriate. For the quality management systems that high-risk AI providers will have to implement, new wording was introduced to align them with similar systems mandated under sectorial legislation.

Finally, it is worth addressing one of the fundamental rights most at risk when using AI systems, especially in the case of algorithmic decision-making: the right to equal opportunities and non-discrimination. The main cause of this is the incompleteness or error of the dataset used by the AI or used to train the AI or the inherent bias in the system. The bias in algorithmic decision-making that the problems mentioned above in the dataset may cause can lead to infringement without any intentionality or human awareness behind it. AI in decision-making can also produce discriminatory results if the system learns from biased training data and the AI Act imposes strict training data requirements.[38] Comprehensive and well-chosen teaching data (the examples used to train the AI) are key here. The role of the code producer also changes from being responsible for the programming (its being error-free) to be primarily responsible for the quality of the data and the correct choice of examples (see Article 10).

3. Concluding Thoughts

The AI Act is forward-looking, detailing the general requirements for high-risk AI systems (the so-called 'essential requirements'). In contrast, the detailed technical requirements will be defined mainly by European standards developed in the framework of European standardization. Although detailed technical standards have already played an important role in Chapter 5 of Title III, they are still largely missing. Their development will be crucial for the effective implementation and enforcement of the proposed AI Act. This observation can be made more generally concerning the implementation of the conformity assessment mechanism of the proposal. Conformity assessment of AI systems will be carried out according to technical rules defined entirely by notified bodies, i.e. private bodies that are supposed to be remunerated for their activities. Therefore, it is of the utmost importance to ensure that national authorities are given as much power as possible to democratically control how these organizations carry out their activities and how they implement the proposal's standards in concrete terms.

The mandatory requirements for high-risk AI systems are broadly based on the 'requirements for trustworthy AI' listed in the ethical guidelines of the High-Level Expert Group on Artificial Intelligence. They must be met before a system

38 Zuiderveen Borgesius 2018. 6.

can be placed on the market or put into service. These relate to data quality and management, documentation and record keeping, transparency and user information, human supervision, robustness, accuracy, and security. Introducing such mandatory requirements is a significant step forward in protecting against the harmful effects of AI systems. However, the proposal still needs to be significantly revised in terms of how high-risk systems are defined, and the requirements, which are currently based on a list, and the provisions are prescriptive.

By granting the notified body the right to have full access to teaching, validation, and testing data and to request access to source codes, the draft creates a tension between the need to regulate the activities of organizations responsible for the development of high-risk systems and the protection of the intellectual property of these organizations, in line with the freedom to conduct a business and the right to the protection of intellectual property, both of which are protected by the Charter of Fundamental Rights of the European Union. It is necessary to ensure that the know-how of undertakings is adequately protected, with appropriate confidentiality requirements, and that access requests are targeted and proportionate to the specific task.

A possible criticism is that it is difficult to predict the future use of AI systems and that it is too early to establish a definitive list of prohibited AI practices. The prohibition of subliminal manipulation under the AI Act provides a low level of protection. It only applies to a limited range of abuses and remains open to other non-subliminal but manipulative AI practices.[39]

References

ALGORITHM WATCH 2019. *Personal Scoring in the EU: Not Quite Black Mirror Yet, at Least if You're Rich*. https://bit.ly/3MAQBvM (accessed on: 14.11.2022).

BERTUZZI, L. 2022a. *EU Council Nears Common Position on AI Act in Semi-final Text*. https://bit.ly/3Hxyd6t (accessed on: 14.11.2022).

2022b. Once Bitten, Netherlands Wants to Move Early on Algorithm Supervision. *Euroactiv*. https://www.euractiv.com/section/digital/news/once-bitten-netherlands-wants-to-move-early-on-algorithm-supervision/ (accessed on: 19.11.2022).

CLARKE, O. 2022. *Legislators Worldwide Move to Adopt Regulation by Design*. https://bit.ly/3Tm4H6f (accessed on: 10.10.2022).

CODAGNONE, C. et al. 2022. *Identification and Assessment of Existing and Draft EU Legislation in the Digital Field. Study for the Special Committee on Artificial Intelligence in a Digital Age (AIDA)*. Luxembourg.

39 For a detailed analysis of the Artificial Intelligence Act and suggestions for amendments, see: Smuha et al. 2021; Ebers et al. 2021. 589.

CSINK, L.–MAYER, A. 2012. *Variációk a szabályozásra* [Variations on Regulation]. Budapest.

EBERS, M. et al. 2021. The European Commission's Proposal for an Artificial Intelligence Act – A Critical Assessment by Members of the Robotics and AI Law Society (RAILS). *Multidisciplinary Scientific Journal* 4. https://www.mdpi.com/2571-8800/4/4/43 (accessed on: 19.11.2022).

EUROPEAN COMMISSION. 2020a. *White Paper on Artificial Intelligence: A European Approach to Excellence and Trust*. https://commission.europa.eu/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en (accessed on: 19.11.2022).

2020b. *Proposal for a Regulation of the European Parliament and of the Council on a Single Market for Digital Services (Digital Services Act) and Amending Directive 2000/31/EC COM/2020/825 Final*. https://eur-lex.europa.eu/legal-content/en/TXT/?uri=COM:2020:825:FIN (accessed on: 19.11.2022).

2021. *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*. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52021PC0206 (accessed on: 19.11.2022).

2022a. *Proposal for a Revision of the Product Liability Directive. Brussels, 28.9.2022 COM(2022) 495 Final 2022/0302 (COD).* https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52022PC0495 (accessed on: 19.11.2022).

2022b. *Proposal for a Directive Adapting the Rules on Non-contractual Civil Liability to Artificial Intelligence. Brussels, 28.9.2022 COM(2022) 496 Final 2022/0303 (COD)*. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52022PC0496 (accessed on: 19.11.2022).

2022c. *New Liability Rules for Products and AI to Protect Consumers and Promote Innovation*. https://ec.europa.eu/commission/presscorner/detail/en/IP_22_5807 (accessed on: 19.11.2022).

EUROPEAN COMMISSION HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE (AI HLEG) 2019. *Ethics Guidelines for Trustworthy AI*. https://www.aepd.es/sites/default/files/2019-12/ai-ethics-guidelines.pdf (accessed on: 19.11.2022).

EUROPEAN PARLIAMENT. 2020. *Resolution of 20 October 2020 with Recommendations to the Commission on a Civil Liability Regime for Artificial Intelligence (2020/2014(INL))*. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52020IP0276 (accessed on: 19.11.2022).

FRANKLIN, M. et al. 2022. Missing Mechanisms of Manipulation in the EU AI Act. *FLAIRS*: https://journals.flvc.org/FLAIRS/article/view/130723 (accessed on: 19.11.2022).

FUTURE OF LIFE INSTITUTE.

2022. *Manipulation and the AI Act.* https://futureoflife.org/wp-content/uploads/2022/01/FLI-Manipulation_AI_Act.pdf (accessed on: 19.11.2022).

GEORGIEVA, I. et al. 2022. *Regulatory Divergences in the Draft AI Act – Differences in Public and Private Sector Obligations*. https://www.europarl.europa.eu/thinktank/en/document/EPRS_STU(2022)729507 (accessed on: 19.11.2022).

HACKER, P. 2021. Manipulation by Algorithms. Exploring the Triangle of Unfair Commercial Practice, Data Protection and Privacy Law. *European Law Journal* (forthcoming): https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3835259 (accessed on: 19.11.2022).

HAGENDORFF, T. 2020. The Ethics of AI Ethics. An Evaluation of Guidelines. *Minds and Machines* 30. https://link.springer.com/article/10.1007/s11023-020-09517-8 (accessed on: 19.11.2022).

HANANI, R. J. 2022. The Politics of Artificial Intelligence Regulation and Governance Reform in the European Union. *Policy Sciences* 55. https://link.springer.com/article/10.1007/s11077-022-09452-8 (accessed on: 19.11.2022).

HÉDER, M. 2020. A Criticism of AI Ethics Guidelines. *Információs Társadalom* 4. https://inftars.infonia.hu/pub/inftars.XX.2020.4.5.pdf (accessed on: 19.11.2022).

HEGEDŰS, L. 2015. *Az ön- és társszabályozás vizsgálata egyes európai államok médiaigazgatásában* [Examining Co-regulation in the Light of Media Governance in Hungary and Abroad]. https://blszk.sze.hu/images/Dokumentumok/diskurzus/2014/2/heged%C5%B1s.pdf (accessed on: 19.11.2022).

LARSSON, S. 2020. On the Governance of Artificial Intelligence through Ethics Guidelines. *Asian Journal of Law and Society* 7. https://doi.org/10.1017/als.2020.19 (accessed on: 19.11.2022).

MADIEGA, T.–MILDEBRATH, H. 2021. *Regulating Facial Recognition in the EU*. https://www.europarl.europa.eu/RegData/etudes/IDAN/2021/698021/EPRS_IDA(2021)698021_EN.pdf (accessed on: 19.11.2022).

MAHLER, T. 2021. Between Risk Management and Proportionality: The Risk-Based Approach in the EU's Artificial Intelligence Act Proposal. *Nordic Yearbook of Law and Informatics*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4001444 (accessed on: 19.11.2022).

MÖKANDER, J. et al. 2022. The US Algorithmic Accountability Act of 2022 vs. The EU Artificial Intelligence Act: What Can They Learn from Each Other? *Minds and Machines* 32. https://link.springer.com/article/10.1007/s11023-022-09612-y (accessed on: 19.11.2022).

ROBERTS, H. et al. 2021. The Chinese Approach to Artificial Intelligence: An Analysis of Policy, Ethics, and Regulation. *AI & Society* 36. https://link.springer.com/article/10.1007/s00146-020-00992-2 (accessed on: 19.11.2022).

SMUHA, N. et al. 2021. *How the EU Can Achieve Legally Trustworthy AI: A Response to the European Commission's Proposal for an Artificial Intelligence*

Act. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3899991 (accessed on: 19.11.2022).

TÓFALVY, T. 2013. Média a törvényen túl? [Media beyond the Law?]. *Médiaelmélet*. https://www.mediakutato.hu/cikk/2013_04_tel/06_media_onszabalyozas.pdf (accessed on: 19.11.2022).

VERGNOLLE, S. 2021. Identifying Harm in Manipulative Artificial Intelligence Practices. *Internet Policy Review*. https://policyreview.info/articles/news/identifying-harm-manipulative-artificial-intelligence-practices/1608 (accessed on: 19.11.2022).

WINFIELD, A. F. T. et al. 2021. A Proposed Standard on Transparency. *Frontiers in Robotics and AI*. https://doi.org/10.3389/frobt.2021.665729 (accessed on: 19.11.2022).

ZŐDI, Zs. 2020. *On the Futility of Codes of Ethics in Regulating Artificial Intelligence*. https://bit.ly/3W6ynGc (accessed on: 01.12.2022).

ZUIDERVEEN BORGESIUS, F. Z. 2018 *Discrimination, Artificial Intelligence, and Algorithmic Decision-Making*. https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73 (accessed on: 01.12.2022).