Kempelen
Institute
of Intelligent
Technologies

Stance on

the Proposal for a Regulation of the European Parliament and of the Council Laying down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts

Abbreviations

AI means artificial intelligence.

AIA means proposal for 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.

EDPB European Data Protection Board.

EDPS means European Data Protection Supervisor.

EU means European Union.

GDPR means regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)

KInIT means Kempelen Institute of Intelligent Technologies with its registered seat in Bratislava, Slovakia.

Authors: Matúš Mesarčík, Sára Soľárová, Juraj Podoužek, Mária Bieliková

Whereas KInIT is an independent, non-profit institute dedicated to intelligent technology research bringing together and nurturing experts in artificial intelligence and other areas of computer science, with connections to other disciplines.

Whereas artificial intelligence is at the core of our research.

Whereas our business and academic partners implement artificial intelligence systems in practice.

Whereas we appreciate the value of the public debate on societal impact of artificial intelligence in general.

Whereas European Commission on 21.4.2021 introduced the Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts.

Whereas we have internally discussed and conducted surveys with respect to the proposal and received opinions of our experts.

We are presenting our stance on specific areas of the proposal to contribute to the public debate and present our ideas on the proposed regulation of artificial intelligence.

Contents

Concerns related to...

Executive summary

The definition of AI systems shall be more precise to cover the use of predefined techniques and not only the development. Furthermore, the influence of generated outputs shall be considered and material criteria for techniques added as a binding prerequisite.

The scope of prohibited practice shall be clarified as the current scope may be extensive and additionally contains loopholes for exploitation. More evidence shall be presented as per specific prohibited practices by the legislator.

Several crucial areas of high-risk AI systems are absent in Annex III namely AI systems used in the context of environmental protection, climate change and transport. Furthermore, attention economy (including social networks) and transportation shall be specific areas. From the procedural point of view, due to technological advancement, the EU shall opt for a more dynamic approach in terms of updating Annex III.

AIA shall make direct reference to the HLEG AI work on trustworthy AI. Specifically, acknowledgement of ALTAI in recitals and obligation to conduct ethical assessment shall be part of the binding text of the regulation.

AIA shall contain specific rules for legacy AI systems that are not currently covered by the proposal considering the feasibility of compliance. Furthermore, the notion of significant change shall be provided.

We are of the opinion that transparency obligations shall be covered in a more comprehensive manner. Considering the state-of-the-art, more nuanced approach including delegated acts considering technological development shall be preferred.

Deepfakes require a more specific approach in the regulation and shall be considered as a high-risk AI system considering real-world scenarios severely violating fundamental rights and freedoms. Text-based deepfakes shall be also included in the regulation.

Remote biometry shall be addressed more specifically with precise rules instead of a general ban on the use of remote biometry. Many concerns are already mitigated by existing laws. Further concerns shall be regulated considering fundamental rights at stake.

Stronger EU oversight is needed in case of monitoring of compliance with the AI regulation. KInIT supports the creation of a more active EU supra authority for this purpose.

Public authorities shall not be excluded from imposing administrative penalties. The exception in question shall be interpreted narrowly.

1.

Concerns related to the definition of the Artificial Intelligence

From the point of subject matter, AIA sets forth harmonised rules for the placing on the market, the putting into service and the use of artificial intelligence systems ('AI systems') in the European Union. Therefore, the basic point for triggering the application of the regulation in question is the notion of the artificial intelligence system. AIA provides the legal definition of the artificial intelligence system in Article 3 point 1. The definition is supported by techniques and approaches listed in the first Annex of AIA. These techniques and approaches include. The aim of the legal definition is to be neutral, future proof and easily supplemented in case of dynamic development of new AI technologies[1].

We appreciate the inclusion of the definition into the EU legal order. However, we have remarks towards the contents of the definition.

The definition of AI as provisioned in AIA **is broad.** This fact in practice means that the regulation will cover many systems, many of these may not realistically be appropriate to include. We are **especially aware** of the problematic inclusion of **statistical approaches** under c) in Annex I.

We welcome the functional definition of AI with a list of techniques in Annex I. However, in our opinion AIA shall not apply **only to software "developed with"** one or more techniques but shall be **extended** to AI systems *comprised* of one or more techniques[2] listed in the Annex I. The amendment is necessary because it currently reflects the use of techniques during the development, but the definition does not cover the use of techniques in AI systems in general

Our proposal explicitly requires listed techniques to be a part of the software.

Furthermore, the **effect of the techniques shall be considered as well** to limit the currently proposed definition to situations where AI systems are responsible for generating outputs (e.g. recommendations or decisions).

1 AIA Explanatory Report, 5.2.1.

2 AIA, Annex I:

"(a)Machine learning approaches, including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning;

(b)Logic- and knowledge-based approaches, including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems;

(c)Statistical approaches, Bayesian estimation, search and optimization methods."

From the reasons stated above, we propose the following amended definition of AI:

artificial intelligence system (AI system) means software that comprises ~~is developed with~~ one or more of the techniques and approaches listed in Annex I ~~and can~~ that, for a given set of human-defined objectives, directly or indirectly influence generated outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with.

Additionally, **material criteria characterizing intelligent behaviour as a binding prerequisite for techniques listed in Annex 1 shall be added.** Such an approach would cover situations where AI systems are able to perform in the environment that was not defined or derived during the development. Material criteria shall be determined by the expert community and may *inter alia* include adaptability, human-likeness, sensing, comprehending, or learning.

2. Concerns related to the banned AI systems

AIA differentiates between four levels of AI systems based on risk-approach. Prohibitions are applicable concerning banned technologies. Prohibited practices are enshrined in Article 5 AIA. AIA explicitly prohibits the placing on the market, putting into service or use of an AI systems conducting subliminal techniques beyond a person's consciousness, social scoring by public authorities and exploitation of any of the vulnerabilities of a specific group of persons. AIA further prohibits the use of 'real-time' remote biometric identification systems in publicly accessible spaces for the purpose of law enforcement[3]. However, certain exceptions discussed in the separate part of the document are provided by AIA. The list is composed due to incompatibility with values of the EU and potential violations of fundamental rights and freedoms[4]. Additionally, the selection was supported by specific use cases in the EU as well as in third countries[5].

Evidence for the prohibition is partially disclosed in the Impact Assessment.

In general, we agree with the proposed prohibited practices, in some cases more empirical evidence shall be shown.

3 AIA, Article 5 (1) d)

4 AIA Explanatory Report, 5.2.2.

5 See e.g. AIA Impact Assessment, pp. 46 – 48.

Furthermore, **more clarification** is needed on the application of banned practices especially when it comes to **vague definitions**. For example, deployment of subliminal techniques beyond a person's consciousness to materially distort a person's behaviour in a manner that causes or is likely to cause that person's harm may include marketing algorithms showing specific advertisements that may result in discrimination or chilling effect. Therefore, guidelines on the application of prohibited practices would be welcomed by relevant stakeholders.

3. Concerns related to the classification of high-risk AI systems

Majority of the provisions in AIA are directly applicable to so-called high-risk AI systems. These AI applications are allowed at the EU market but simultaneously are subject to specific requirements including an ex-ante conformity assessment. High-risk AI systems are listed in two different manners. Firstly, high-risk systems are stated in Annex II listing EU safety legislation applicable to specific products. Secondly, the European Commission provides a list of additional high-risk systems in Annex III. The list is composed of the pre-defined areas and specific applications. The European Commission may adopt changes to Annex III considering specific procedure enshrined in the Article 7 AIA.

Our position concerning the classification of high-risk AI systems includes substantial and procedural aspects. **From the point of substantial aspects, several high-risk areas are absent in Annex III.** One of the missing areas is **natural ecosystems and climate** with severe impact on the environment e.g., excessive use of computing power (carbon emissions) in return for just marginal performance improvements. What shall be also included in the high-risk applications is the **area of transport.** The current proposal is limited to the transportation area only in case of being a part of critical infrastructure. The approach seems to be restrictive, and several AI high-risk applications may evade requirements of AIA.

Additionally, we are of the opinion that AI systems used in the **"attention economy"**[6] in the context of spreading disinformation shall be classified as a high-risk AI system as well. If e.g., providers of social media provide space for the spread of disinformation and gain profit out of the attention on such information, the AI systems shall be classified as high-risk. The proposal may be drafted analogically as cases of remote biometry. If remote biometry is not prohibited under Article 5 AIA,

We are of the opinion that AI systems used in the "attention economy" in the context of spreading disinformation shall be classified as a high-risk AI system as well

other biometric applications are classified as high-risk based on Annex III. Similarly, if attention economy models do not fall within Article 5 AIA (especially considering banned use of "subliminal techniques beyond a person's consciousness"), they shall be an integral part of Annex III.

From the point of procedural aspects for potential additions towards the list in Annex III it shall be noted that the European Commission may adopt delegated acts to amend the list in **Annex III.**[7] However, these amendments relate only to the specific applications in the pre-defined areas by AIA. In case of the emergence of a new area with high-impact AI systems, the regulation itself has to be amended within the classic legislative procedure on the EU level. The issue is twofold. Firstly, above mentioned procedural differences may cause the European Commission to artificially expand the list of high-level AI systems in the inaccurate areas limiting the scope and application of AIA. Secondly, the aim of the regulation to be future proof would be hampered by adding areas via traditional (and slow) legislative procedure that may take years allowing the application of risky AI systems threatening fundamental rights and freedoms.

Regulation on electronic identification (commonly referred to as eIDAS Regulation)[8] contains the procedure that might be followed also in case of classification of a standalone high-risks AI system. For example, Commission Implementing Regulation (EU) 2015/1502 of 8 September 2015 on setting out minimum technical specifications and procedures for assurance levels for electronic identification means sets forth details of technical specifications of assurance levels as foreseen by eIDAS Regulation. **Such approach is also proposed in terms of drafting state-of-the-art high risk AI systems** as the procedure for adopting implementing regulations is faster and may reflect dynamics of development of new technologies.

3 See e.g. WU, T.: The Attention Merchants: The Epic Scramble to Get Inside Our Heads. Knopf, 2016.

4 AIA, Article 7 in connection with Article 73.

8 Regulation (EU) No 910/2014 of the European Parliament and of the Council of 23 July 2014 on electronic identification and trust services for electronic transactions in the internal market and repealing Directive 1999/93/EC.

4. Concerns related to the weak ties with ethical framework on trustworthy AI

The ethical and trustworthy side of AI is emphasized by the European Commission and the European Parliament.[9] The European Commission appointed a group of experts to provide advice on its artificial intelligence strategy. The group is referred to as the High-level expert group on artificial intelligence (AI HLEG).[10] During the first terms of the group, experts delivered several documents and guidance in terms of policy, ethics and investments. The most notable works relate to the ethics of AI namely:

- **Ethics Guidelines for Trustworthy AI**[11]
- **The final Assessment List for Trustworthy AI (ALTAI).**[12]

However, the text of AIA mentions ethical implications only in the non-binding form of recitals.[13] The explanatory report only briefly refers to Ethics Guidelines for Trustworthy AI and ALTAI as state-of-the-art minimum requirements towards conformity assessments as foreseen by AIA.[14] As trustworthiness is the key concept provided by Ethics Guidelines for Trustworthy AI, it is of the essence to note that Recital 62 AIA partially acknowledges this concept.

We welcome acknowledgement of the ethical side of AI in the recital and explanatory report of AIA. However, according to us, with the reference to the robust work conducted by AI HLEG, the ties shall be **narrower and assessment of ethical risks shall be considered as a binding part of the conformity assessment.** This position is further supported by conclusions of the European Parliament.[15]

> **It would be of great importance to recognize ALTAI in recitals of AIA and provide an obligation to assess ethical risk within the conformity assessment pursuant to Article 43 AIA.**

9 See e.g. European Parliament resolution of 20 October 2020 with recommendations to the Commission on a framework of ethical aspects of artificial intelligence, robotics and related technologies (2020/2012(INL)).

10 https://digital-strategy.ec.europa.eu/en/policies/expert-group-ai

11 https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai

12 https://ec.europa.eu/digital-single-market/en/news/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment

13 AIA, Recital 5 and 16.

14 AIA Explanatory report, 5.2.3.

15 Point 145.

16 ALLAI: DRAFT AI ACT: ALLAI ANALYSIS AND RECOMMENDATIONS. Available at https://allai.nl/draft-ai-act-allai-analysis-and-recommendations/.

The missing parts of ALTAI have been already the subject of the discussion in other stances.[16]

Based on the findings presented above, we encourage the European Commission to include a more direct and explicit connections to the framework for trustworthy (and ethical) AI. It would be of great importance **to recognize ALTAI in recitals of AIA and provide an obligation to assess ethical risk within the conformity assessment pursuant to Article 43 AIA.** A similar provision is already in the proposed text concerning the need to adopt a strategy for regulatory compliance.[17]

5. Concerns related to double-track effect

AIA explicitly proposes a double-track effect in terms of the applicability of the regulation. This effect is provisioned in the Article 83 AIA considering AI systems already placed on the market or put into service. Firstly, a specific exception is applicable to EU specific information systems e.g. Eurodac, Visa system, Schengen Information System.[18] Other AI applications are covered by AIA in two scenarios (1) AIA is generally applicable to AI systems placed on the market or put into service from the date of application of the regulation; or (2) AIA shall apply to the high-risk AI systems, that have been placed on the market or put into service before the date of application of AIA, only if, from that date, those systems are subject to significant changes in their design or intended purpose.

In practice, AIA will apply to the currently used AI systems only in case those systems will be subject to significant changes in their design or intended purpose.

As the current proposal stands, the legislation will cause a double-track effect due to the application only to new AI systems introduced to the market after the date of application of AIA or in case of significant change of currently used AI systems.

> **We recommend a more balanced approach towards currently deployed AI systems with partial application of requirements provisioned in AIA.**

17 AIA, Article 17 (1) a).
18 See Annex IX AIA for the full list..

Firstly, the regulation will apply differently due to complex obligations required by AIA on new AI systems and current (minimal) requirements on already deployed AI systems. We understand that several obligations as foreseen by AIA are connected to the development phase of AI systems or collection of data and detection of biases. Therefore, it is almost impossible for the currently used AI system to comply with these provisions. **However, not every obligation concerns the development and may be conducted even when the AI system is already deployed.**[19] Furthermore, many requirements related to post-market monitoring including incidents reporting[20] **may be at least partially fulfilled by providers of currently deployed AI systems.**

Secondly, **the notion of significant change is vague** and calls for further guidance and interpretation as the definition is key for triggering the scope of AIA in case of currently used AI systems.

Therefore, we recommend a more balanced approach towards currently deployed AI systems with partial application of requirements provisioned in AIA. The European Commission shall carefully consider if the current proposal will not intervene with the principle of legal certainty for current providers of AI systems.

6. Concerns related to transparency obligations

Transparency of AI is one of the most discussed issues in the political or academic debate as transparency directly increases the trustworthiness of AI.

AIA differs between obligations of transparency related to business users and natural persons. Transparency requirements for business users are covered by Article 13 AIA requiring disclosing of specific information in the technical documentation of a delivered AI system. More importantly, Article 52 contains specific rules in terms of transparency related to natural persons. Similar obligations are set forth considering an emotion recognition system or a biometric categorisation system[21] and deep fakes[22] discussed in the separate section. Specific exceptions apply in the context of law enforcement which are permitted by law to detect, prevent and investigate criminal offences.

> **We believe that transparency is one of the key aspects of trustworthy AI.**

19 For example drawing up technical documentation pursuant to Article 11 AIA.

20 AIA, Article 61.

21 AIA, Article 52 (2).

22 AIA, Article 52 (3).

Furthermore, obligations related to transparency shall be explicitly included in the law. AIA seems to have a very careful approach when it comes to transparency and sufficiency is questionable. For most AI applications only the open communication requirements are set out. **Additionally, provisions are restrictive as they apply only for image, video and audio content, not other forms of content, e.g., text-based content.**

We understand that certain aspects of transparency are already covered by applicable EU legislation namely General Data Protection Regulation (GDPR) and Regulation on fairness and transparency.[23] However, the legislation does not fulfil the goals of transparency in a sufficient way.

Concerning GDPR, Article 22 sets forth rules for the automated decision-making process. Transparency requirements stem directly from Article 22 (3) and related Recital 71[24] and from institutes of information obligations[25] and the right to access.[26] But the scope of Article 22 is even after 3 years of the application still unclear and subject to judicial and academic debate.[27] This in practice means that transparency requirements are not always fulfilled to their potential and many controllers of personal data rely on the non-applicability of the Article 22 GDPR.

Regulation on transparency and fairness has a very limited scope concerning business users of platforms and online search engines. Therefore, obligations on transparency of ranking parameters[28] including an explanation of their selection and application are constructed in a narrow manner.

On the other hand, we are aware of the fact that we are currently not able to fulfil transparency requirements in general. We as researchers understand this situation and contribute to the state of the art in this area, but it should be shown that state of the art is used in particular systems, especially in high-risk systems.

Based on the discussion above we again encourage the European Commission to re-evaluate the approach on the transparency obligations in the AIA proposal. We fully understand the issue of transparency of AI in the technological context and objective problems with feasibility of legal transparency obligations considering state-of-the-art. On the other hand, transparency obligations are significantly important when it comes to trustworthy AI and possibility for means of legal defence against automated decisions.

> **However, the option may lie in drafting of the "general clause" requiring transparency with specifics left to delegated acts that may be adopted after the AIA enters into force to complement state-of-the art in a feasible manner.**

23 Regulation (EU) 2019/1150 of the European Parliament and of the Council of 20 June 2019 on promoting fairness and transparency for business users of online intermediation services. OJ L 186, 11.7.2019, p. 57–79.

24 Recital 71 GDPR: „...In any case, such processing should be subject to suitable safeguards, which should include specific information to the data subject and the right to obtain human intervention, to express his or her point of view, to obtain an explanation of the decision reached after such assessment and to challenge the decision."

25 GDPR, Articles 13–14.

26 GDPR, Article 15.

27 See e.g. TOSONI, L.: The Right To Object to Automated Individual Decisions: Resolving the Ambiguity of Article 22(1) of the General Data Protection Regulation. 11 International Data Privacy Law (2021) (Forthcoming). University of Oslo Faculty of Law Research Paper No. 2021–07. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3845913.

28 Regulation on transparency and fairness, Article 5. See also COMMISSION NOTICE. Guidelines on ranking transparency pursuant to Regulation (EU) 2019/1150 of the European Parliament and of the Council. (2020/C 424/01). Available at https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52020XC1208(01)&from=EN.

7. Concerns related to deep fakes

The proposed regulation categorizes deepfakes in the broad category of low-risk AI systems, which shall be subjected to "minimum transparency obligations."[29] Article 52 (3) of AIA explicitly defines deepfakes including requirement to *"disclose that content has been artificially generated or manipulated."* Despite the definition's soundness, the proposed transparency obligation lacks robustness. This articulation assumes that the mere exposure of users to a disclosure statement would reduce the information asymmetry and thus allow the users (citizens) to combat the effects of deepfakes and still form informed and accurate opinions. The proposed regulation, however, fails to distinguish between pre-and post-engagement with the transparency disclosure.

While deepfakes might not be inherently harmful, they have the ability to have people say or do things they never said or did, thus threatening and potentially disrupting the shared sense of reality and values we as a society hold. The evidence shows the significant impact of deepfakes on individuals (especially women) as well as on society considering cases of revenge porn, political leaders, or harassment of women.[30] In general, people tend to attribute greater validity and believability to information that they recognize from first source, even if the information is fake/falsified and/or has been disputed by independent fact-checkers.[31] Human brain also tends to favour information which can be processed quickly - unconscious bias known as processing fluency. Deepfakes' increasingly sophisticated technical realism and depiction of known political figures can strengthen this bias, hence increase people's believability of the doctored audio-visual content. This means that even when a transparency disclosure statement appears next to deepfakes, people can still favor such content because it can resemble something/someone we recognize and thus human brains evaluate it as valid. **Therefore, the mere focus on transparency is insufficient as it does not take into account the basic psychological biases that prevent people from forming impartial opinions.**

> **The classification of deepfakes under low-risk AI applications leads to an inconsistency with the EC's recognition of other AI systems as high-risk.**

High risks associated with AI systems throughout the regulation are contextualized in potential terms, i.e., their possibility to wrongfully determine individuals' prospects (education, employment, migration) or violate fundamental procedural rights in judicial settings. Deepfakes, similarly to other AI systems, pose both immanent and potential threats regarding "possible hampering of fundamental rights."[32] The European Commission's inability to provide a rationale for deepfakes

29 AIA Impact Assessment, p.3.

30 AIDER, H. – PATRINI, G. et al.: The state of deepfakes: Landscape, threats, and impact. Amsterdam: Deeptrace, 2019. CHESNEY, R. - CITRON, D.: Deep fakes: A looming challenge for privacy, democracy, and national security. Calif. L. Rev., 107, 1753, 2019. SCHICK, N.: Deep Fakes and the Infocalypse: What You Urgently Need To Know. Hachette UK., 2020. ROTH, A.: European MPs targeted by deepfake video calls imitating Russian opposition. The Guardian., 2021

31 PENNYCOOK, G. et al.: Prior exposure increases perceived accuracy of fake news. Journal of Experimental Psychology: General, 147(12), 1865, 2018.

being categorized as low-risk AI applications thus fails to account for violations of the fundamental rights protected by the Charter. **What lacks in the current regulatory proposal is the definition of inappropriate use of deepfakes.** It is critical for society to decide which uses of deepfakes are acceptable and which are not.

Briefly summarized, AIA regulates AI systems for deepfakes detection,[33] prohibits practices that may *inter alia* include deepfakes[34] and leaves other deepfake applications unregulated (with the exception of transparency obligations). This gap **shall be mitigated in the further revisions** of AIA. Additionally, text based deepfakes (also known as auto or machine generated texts or neural fake news) shall be included in the scope of the regulation.[35] We encourage the European Commission to re-evaluate the approach on deepfakes as they represent a significant risk towards fundamental rights as discussed above.

8. Concerns related to remote biometrics

Remote biometric is regulated by AIA in two ways. Firstly, remote biometrics systems used for law enforcement purposes are listed within prohibited practices with certain restrictive exceptions. These exceptions may be further subject to national legislation thus allowing EU Member states to state specific rules for legal use of remote biometrics systems for law enforcement purposes.[36] Secondly, use of biometry for other purposes than law enforcement shall be considered as a high-risk AI system triggering requirements prescribed by AIA.

Although several organizations and EU bodies **propose a general ban** on the use of remote biometrics[37] in public spaces (including online space), we are of the opinion that **such an approach represents the risk of misused opportunity.**

> **Remote biometrics also introduces beneficial opportunities especially in cases of security purposes at the sport events or airports.**

33 AIA, Recital 38.

34 Practices mentioned in the Article 5.

35 See ZELLERS, R. et al.: Defending against neural fake news. In Proceedings of the 33rd International Conference on Neural Information Processing Systems, pp. 9054–9065; HARRAG, F. et al.: Bert Transformer model for Detecting Arabic GPT2 Auto-Generated Tweets. In Proceedings of the Fifth Arabic Natural Language Processing Workshop, pp. 207–214.

36 AIA, Article 5 (2–4).

37 Algorithm Watch: Open Letter to Ban Biometric Surveillance. Available at: https://algorithmwatch.org/en/open-letter-ban-biometric-surveillance/. EDPB-EDPS Joint Opinion 5/2021 on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Available at: https://edpb.europa.eu/our-work-tools/our-documents/edpbedps-joint-opinion/edpb-edps-joint-opinion-52021-proposal_en.

Furthermore, the "seamless" experience provided using remote biometry shall be emphasized as well as far as the purpose of use is clearly defined and the use is technically optimized.[38] According to us, the approach is comparable to the use of identity cards for entering specific premises with logging as an organizational measure.

We **acknowledge and respect concerns of using the technology especially addressing issues of transparency, data processing and potential chilling effect.** However, we are of the opinion that these risks shall be addressed and mitigated by the already existing legal framework (e.g., GDPR) or provisioned in the final wording of AIA. Adhering to the legitimate purpose and defining the retention periods are key requirements towards compliance in case of these systems. Accountability and scalability of the remote biometry shall be carefully assessed and instead of a general ban, biometric applications shall be surrounded by legal requirements specifying obligations for the use as a high-risk AI system. **Therefore, we recommend regulating remote biometrics as high-risk AI systems or restricted AI systems.**

9. Concerns related to supervision and oversight

Supervision and oversight in AIA are proposed as a combination of state oversight and independent notifying bodies. Before AI systems are introduced on the market, they need to undergo so-called conformity assessment. In some cases, the conformity assessment shall be conducted by the notified body accredited by notifying authority.

On top of the obligations referred to above, EU member states shall designate national competent authorities responsible for ensuring the application and implementation of AIA. These national competent authorities may also serve as notifying authorities.

The role of enforcement is crucial when it comes to any regulation. AIA provides a space for accreditation of notified bodies and ex-post oversight of national authorities. Furthermore, the European Artificial Intelligence Board will be founded as the advice body for the EU and national authorities in the field of AI.

We acknowledge the need for mutual international cooperation and state supervision of legal rules. In terms of international cooperation, similar models already exist in the EU law. Pursuant to directive (EU) 2016/1148 of the European Parliament and of the Council of 6 July 2016 concerning

38 See e.g. discussion on the use of biometry for identification and authentication purposes. BYGRAVE, L – TOSONI, L. Article 14 (4). Biometric Data. In KUNER, CH. – BYGRAVE, L. – DOCKSEY, CH.: The EU General Data Protection Regulation (GDPR). A commentary. Oxford University Press, 2020.

measures for a high common level of security of network and information systems across the Union (NIS directive) Cooperation group is established serving *inter alia* also for the purpose of exchanging best practices in the field of cyber-security. Analogically, such cooperation shall be also required by AIA.

> **We are of the opinion that stronger EU oversight is needed.**

This is especially true in cases of large-scale algorithms deployed via international platforms like social media. Furthermore, financial, and mainly personal capacities of some EU member states might not be sufficient for comprehensive and efficient oversight of AI systems. The concept is not novel in the EU law as the European Data Protection Board established by GDPR has a power to decide competence disputes between national data protection authorities.

Based on that, **we support the creation of a more active EU supranational body acting as an appellate authority in cases of large-scale AI systems and guiding national competent authorities.** However, the guidance shall be limited in time.

10.
Concerns related to open clause allowing exclusion of public authorities from administrative penalties

AIA contains the possibility for public authorities to impose severe administrative penalties on providers of AI systems that are not compliant with the regulation. Based on the severity of the violation, AIA sets forth three groups of administrative fines in the Article 71. However, legal systems of several EU Member states do not allow imposing administrative penalties on public authorities. This is explicitly acknowledged in AIA as well providing the room for exception in national laws.[39]

39 AIA, Article 71 (7).

Several public authorities in EU Member states use the AI systems for different purposes. COVID–19 pandemic represents yet another potential for deployment of AI systems for purposes of protection of health and life.[40] This is further supported by the classification of AI systems into the high-risk AI systems, namely in the areas of public services or critical infrastructure.[41] Therefore, public authorities will undoubtedly fall under the requirements of AIA.

A specific regime for imposing administrative penalties apply to public authorities. The approach is not novel as a similar open clause is part of the GDPR. The reasoning for the rule is that some of the EU Member state´s legal systems do not allow sanctioning public authorities by competent bodies. On the other hand, the provision was drafted in a way allowing other EU member states excluding or altering administrative penalties for public authorities. As research shows,[42] 21 EU member states seized the option of the open clause to exclude or alter administrative sanctioning for public authorities. There is a high probability that many EU member states will similarly use the open clause in the context of AIA.

It is argued that the specific position of public authorities as enshrined in the literature[43] does not stand in the area of technology regulation. Firstly, AIA applies to providers of (high-risk) AI systems irrespectively of their private or public nature. Secondly, public authorities provide essential public services to citizens. When AI systems are involved in the process, the potential for violation of fundamental rights and freedoms is increased due to the nature of public services. This is also evident from the point of prohibiting credit scoring systems by public authorities. The third specificity of public authorities concerns funding from public resources and paying potential financial damage from the same source. It is of the essence to note that public authorities may be seen as standards of compliance with norms adopted or recognized by the state. In case that public authority is not compliant with law, taxpayers have the full right to have information about the case and reflect it via mechanisms available in a democracy.

We suggest the re-evaluation of the concept of exclusion of public authorities from administrative penalties

We understand that in some countries it is not possible to sanction public authorities by competent bodies. Nevertheless, the open clause shall be limited to the **narrowest extent possible.**

40 See e.g. ALGORITHM WATCH: AUTOMATING SOCIETY REPORT 2020. Automated Decision-Making Systems in the COVID–19 Pandemic: A European Perspective. 1st September 2020.

41 AIA, Annex III, points 1 and 5

42 See https://www.whitecase.com/publications/article/gdpr-guide-national-implementation#q17.

43 E. g. VAN DAM, C. European Tort Law. 2nd Edition. Oxford: Oxford University Press, 2013

Conclusions

We appreciate the introduction of a comprehensive legal framework for AI in the EU. It is a necessary precondition for mitigation of risks and potential violations of fundamental rights and freedoms. However, the proposal shall undergo several revisions to adhere to its bold goals.

We hope that our stance as a research institution will contribute to the public debate on the development of the final wording of the legislation.

Authors

Matúš Mesarčík

(Ethics and Law Specialist) works as a specialist with a focus on the regulation of new technologies and privacy and data protection. He also conducts research and lectures at the Institute of Information Technology Law and Intellectual Property Law at Comenius University in Bratislava, Slovakia. Apart from his research activities, Matúš worked as a privacy consultant in a law office focused on privacy and new technologies and currently advises public authorities on the topics of data protection and cyber-security.

Sára Soľárová

(Research Intern) just finished her bachelor's degree in Politics, Psychology, Law and Economics at the University of Amsterdam. Throughout, she has developed interest in cyber-policy, which she will pursue in her master's studies at Tel Aviv University. In her thesis, she explored the social impact of deepfakes in the European context, more specifically their conceptualization within the recent EU AI regulation.

Juraj Podroužek

(Senior Researcher) achieved a PhD in philosophy and conducted research in semantics and analytic philosophy at the Slovak Academy of Sciences. In recent years, he has focused on the ethics of ICT, value sensitive design and AI ethics. Together with Miroslav Pikus he founded the informal E-tika group investigating the ethical and societal impact of ICT, which has launched its own podcast. He is also a member of the national committee on the ethics and regulation of AI.

Mária Bieliková

(Expert Researcher & CEO) developed the long-term vision and strategy for the institute. She also conducts research focusing on human-computer interaction analysis, user modelling and personalization. Recently, she has been working in data analysis and modelling of antisocial behavior on the Web. She is active in discussions on trustworthy AI at the national and European levels. For example, she is a former member of the High Level Expert Group on Artificial Intelligence established by the European Commission. Currently she is a chair of the Permanent Committee for Ethics and Regulation of AI established by the Ministry of Investments, Regional Development and Informatization of the Slovak Republic. She was also a member of the European Commission Joint Research Center Board of Governors.

Acknowledgements

The authors would like to thank following people for their valuable comments and insights on the regulation and stance: **Adrián Gavorník**, **Jakub Šimko**, **Ivan Srba**, **Daniela Chudá**, **Michal Kompan**, **Viera Rozinajová**, **Martin Tamajka** and **Marián Šimko**.

KINIT