COMMENTS

On the Regulation laying down harmonised rules on artificial intelligence

(Artificial Intelligence Act)

Dear Madam or Sir,

We, Women in Artificial Intelligence Austria (Frauen in Künstlicher Intelligenz Österreich, "Women in AI Austria"), are gladly taking the opportunity to comment in the public consultation phase on the draft of the Artificial Intelligence Act ("AI Act"):

Preliminary remarks

We welcome the approach taken by European Commission to better protect consumers and their fundamental rights, ensure safe development of new technologies and foster innovation, growth and competitiveness within the single market.

In order to improve transparency and strengthen the rights of users and small businesses alike, an appropriate legal framework is required. We may, thus, draw your attention to the following provisions:

Comments on the proposed AI Act

Art. 3 point (34): 'emotion recognition system' means an AI system for the purpose of identifying or inferring emotions or intentions of natural persons on the basis of their biometric data;

Emotion recognition is based on disproven scientific theories and should therefore not be validated by virtue of being subjected to regulation. We strongly recommend to change this definition to reflect that these systems are at best an attempt to recognise emotion and affect and, if at all valid, require very specific training data sets to perform in very limited contexts due to the abundant variety in the culturally-determined expression of emotion. At the same time, the proliferation in the use of these systems is worrying – precisely because of their shaky scientific grounding – and we do believe that [people] should enjoy heightened protection if their data is processed by such a system.

Furthermore, emotion recognition systems may be developed on data that is not necessarily biometric, for instance based on speech patterns or engagement with content. In our opinion, the definition should therefore be adapted to refer to all types of personal data, citing biometric data as an example.

Finally, the definition of emotion recognition systems provided here would not necessarily include systems employed to detect psychological states, such as depression. The risks to the fundamental rights of persons are equally high, if not greater, if an AI system seeks to define their mental state, which may be done in accordance with a much more solid corpus of scientific theory. We strongly believe that this Regulation should address these risks and encourage the Commission to refine the proposed definition with this background in mind.

Art. 3 point (35) 'biometric categorisation system' means an AI system for the purpose of assigning natural persons to specific categories, such as sex, age, hair colour, eye colour, tattoos, ethnic origin or sexual or political orientation, on the basis of their biometric data;

The proposed definition supposes some relation between biometric data and ethnic origin and sexual or political orientation. We strongly refute the claim that a person's ethnicity, sexual or political orientation may be detected using biometric data, since these are socially constructed categories and are not connected to measurable attributes of a person's physical presence in this world.

Art. 3 point (39): 'publicly accessible space' means any physical place accessible to the public, regardless of whether certain conditions for access may apply

In the course of the extraordinary circumstances of the past nine months, we have all experienced a shift from physical to virtual spaces. An AI system may be employed irrespective of the type of space it engages with because it draws on input data, which can be produced both in physical and in virtual spaces. Given the increasing use of digital space in different contexts – professional, educational, leisurely – and the indifference of AI systems towards different types of spaces, we would recommend to remove the reference to "physical place", instead defining publicly accessible space as "any place accessible to the public, regardless of whether certain conditions for access may apply".

Art. 3 point (44): 'serious incident' means any incident that directly or indirectly leads, might have led or might lead to any of the following:

(a) the death of a person or serious damage to a person's health, to property or the environment,

(b) a serious and irreversible disruption of the management and operation of critical infrastructure.

We strongly support the inclusion of damage to environment in the definition of 'serious incident', in particular since there seem to be very promising applications for AI systems in environmental protection. However, as the definition in Art. 3 para. (44a) stands, the irreversible loss of access or opportunity caused by a misclassification of a person would never have to be reported, despite breaching fundamental rights. The wording of the requirement in Art. 62 para. (1) may be understood as a requirement to report malfunctions constituting a breach of fundamental rights as well as serious incidents: therefore, an explicit reference to fundamental rights in this definition would provide more clarity.

Art 3 (new definition for AI subjects including also companies)

The AI Act sets forth new obligations for operators of AI systems in order to minimise the risk of harm – but one definition conspicuously absent is that of the person(s) who are to be protected. We urge the Commission to include a definition of AI subject in the AI Act, as has been done in the GDPR for data subject. We further believe that the term "AI subject" should be broad enough to include legal persons such as companies as this would be consistent with the broad definition of user and both natural and legal persons can be affected by AI systems. In our view, the AI Act is a better place to introduce such a concept, rather than in other legislation focusing on liability, because it focuses

on minimising harm and will be foundational for subsequent legislation. The benefit of including not only natural, but legal persons in the definition of "AI subject" is that harm arising from situations which have not been covered by the AI Act (e.g. harmful business practices which are enabled by AI systems) could draw on an existing definition to further develop legal debate and theories.

Art. 5 para. (1)

(a) the placing on the market, putting into service or use of an AI system that deploys subliminal techniques beyond a person's consciousness in order to materially distort a person's behaviour in a manner that causes or is likely to cause that person or another person physical or psychological harm;

(b) the placing on the market, putting into service or use of an AI system that exploits any of the vulnerabilities of a specific group of persons due to their age, physical or mental disability, in order to materially distort the behaviour of a person pertaining to that group in a manner that causes or is likely to cause that person or another person physical or psychological harm;

We strongly support measures aiming to protect the fundamental rights and dignity of persons in the European Union. In the proposed provisions, AI systems are prohibited that cause physical or psychological harm. This provision is insufficient to protect persons in the European Union from other serious harms, such as exploitation.

For this reason, we call for the reformulation of these provisions to achieve a better alignment with the European Charter of Fundamental Rights in the following manner:

(a) the placing on the market, putting into service or use of an AI system that deploys subliminal techniques beyond a person's consciousness in order to materially distort a person's behaviour in a manner that **undermines or is likely to undermine the fundamental rights of that person**;

(b) the placing on the market, putting into service or use of an AI system that exploits any of the vulnerabilities of a specific group of persons due to their age, **gender,** physical or mental disability, in order to materially distort the behaviour of a person pertaining to that group in a manner that **undermines or is likely to undermine the fundamental rights of that person**.

d): the use of 'real-time' remote biometric identification systems in publicly accessible spaces for the purpose of law enforcement, unless and in as far as such use is strictly necessary for one of the following objectives:

(i) the targeted search for specific potential victims of crime, including missing children;

(ii) the prevention of a specific, substantial and imminent threat to the life or physical safety of natural persons or of a terrorist attack;

(iii) the detection, localisation, identification or prosecution of a perpetrator or suspect of a criminal offence referred to in Article 2(2) of Council Framework Decision 2002/584/JHA62 and punishable in the Member State concerned by a custodial sentence or a detention order for a maximum period of at least three years, as determined by the law of that Member State.

The prohibition of 'real-time' remote biometric identification systems is, in our opinion, a valuable contribution towards securing freedom from arbitrary surveillance for all Europeans. However, the exemptions provided allow for much leeway for the use of these technologies. Trust is required in order for AI technologies to flourish, but trust will only develop if solid protections provided by law for all Europeans provide good reasons for this trust. In the current proposal, all 'real-time' remote biometric identification systems currently in use for law enforcement purposes would not be regulated, unless updates lead to significant changes in the AI system. These updates, however, are necessary: countless studies and experiments have demonstrated significant biases in these technologies towards non-white and non-male persons. We believe this is unacceptable and regulatory proposals should seek to remedy this issue, not to incentivise law enforcement agencies to purchase systems now and avoid updates as soon as this Regulation comes into force.

Furthermore, the exceptions for the use of 'real-time' biometric identification technologies allow for much leeway in the application of this exemption. There are no requirements for the presentation of objective evidence to the judicial authority regarding the existence and scope of the threats which the use of biometric identification systems aims to counter: the only provision mandating the provision of "objective evidence or clear indications" to judicial authorities regards the necessity and proportionality of the use of 'real-time' remote biometric identification systems in terms of achieving the objectives in Art. 5 para. (1) point (d). Also, there are no requirements for law enforcement authorities to disclose the use of 'real-time' biometric identification systems, even after the threat has passed, or to be transparent about the effectiveness of these systems. As it stands, this provision will lead to the acquisition of AI systems capable of 'real-time' remote biometric identification systems by law enforcement authorities with the aim of using them, when necessary; in turn, less research and funds will be invested in AI technologies which may aid law enforcement in less invasive ways, which do not undermine the fundamental rights of persons in the European Union at a large scale.

Lastly, the formulation of this point allows other public authorities (such as migration and border control agencies) as well as private operators the use of 'real-time' biometric identification systems, since these entities do not fall within the scope of 'law enforcement' as defined in Art. 3 point (41). We believe this is a dangerous loophole which may undermine a number of principles enshrined in

the Charter of Fundamental Rights, especially regarding the dignity and freedoms of persons. Law enforcement and other public authorities operate on the basis of a clear mandate to protect the public interest; private companies are under no such obligation and should not be able to exert control over persons living in the European Union to a greater extent than public authorities. The restrictions accorded to the use of these technologies should not only apply to law enforcement, but to all actors alike.

Art. 7 Amendments to Annex III

Art. 7 empowers the Commission to add new applications to Annex III which fall into one of the previously established high-risk areas and have equal or higher risk than those applications already listed by way of a delegated act. We support the flexibility this provision offers and encourage a regular evaluation of the market for AI systems and the effects of AI systems in certain use cases. However, we believe Art. 7 may not be sufficient to adapt the AI Act to new and emerging AI use cases and offer long-term protection of fundamental rights due to its limited scope. Since only the applications may be updated, but not the use case areas, we believe the definition of high-risk AI systems may not be future-proof and thus not offer long-term safety for people affected by AI systems in a negative way. We therefore support expanding this Article to allow for the amendment of new high-risk areas by the Commission via delegated acts.

Art. 49 CE marking of conformity

In the current version of the AI Act, providers of high-risk AI systems specified in Annex III may conduct self-assessments to ascertain conformity (Art. 43/2 - in case of biometric identification systems, only those applying harmonised standards or common specifications as detailed in Art. 43/1). In effect, this means that providers of high-risk AI systems as defined in Annex III must apply the CE marking called for in Art. 49, yet for users of these systems, there is no possibility to verify whether the provider has acted in due diligence. In our view, there is a high risk that the CE marking may be applied inappropriately and the penalties may not be a strong enough deterrent from non-compliance. We would prefer an obligation for providers of high-risk AI systems which have conducted a self-assessment to nonetheless register with a notified body, which accepts the required documentation and issues a certificate of conformity. This will allow users to verify at minimum that the documentation has been presented to a notified body and is accessible to the market surveillance authority in case of any issues.

Art. 52 Transparency obligations for certain AI systems

In its paragraphs, Article 52 enumerates three use cases of AI systems which shall be subjected to disclosure obligations. The first use concerns automated systems interacting with natural persons (bots for short), the second concerns emotion recognition or biometric categorisation systems and the third prescribes disclosure of the synthetic nature of AI-generated content (or deep fakes). We believe that these provisions may be insufficient in mitigating potential harm caused by the use of these systems. Regarding the interaction with bots in Art. 52(1), affected parties should have the right to interact with the user in a different format, e.g. in the case of a chatbot, affected parties could instead contact the user via a service hotline or an e-mail address. The lack of such alternatives is more severe in Art. 52(2), as it explicitly legitimises the use of emotion recognition systems (which are not scientifically sound). Since these AI systems are not based on valid science, their output cannot be considered reliable, creating greater harm for affected parties who – for whichever reason

– are dependent on its outcomes. Lastly, Art. 52(3) may serve to increase the trust of wider audiences in content presented to them. However, it is an insufficient response to the problem of deep-fake non-consensual pornography. Such images, video and audio can be easily generated, most often targeting women, and easily shared online, harming the reputation, mental well-being or even physical safety of the victims. This is a problem that calls for a multi-layered solution and we encourage the Commission to engage with stakeholders to find an appropriate framework.

Art. 69 Codes of conduct

We welcome the provision in Art. 69 regarding codes of conduct for all operators of AI systems based on the requirements for high-risk AI systems. This will enable start-ups to garner the trust of those companies or public authorities which would like to use AI systems, but hesitate to work with smaller, potentially more innovative but less well-known firms. To make these codes of conduct a reliable instrument, the commitments need to be enforced, as breaches of the codes of conduct which are not penalised can lead to a loss of trust in the codes of conduct themselves.

The codes of conduct regarding topics not covered in the AI Act could also have an important role in furthering the environmental and social sustainability of AI systems – although these could potentially also be covered by standards. Considering the strong gender imbalance in AI-related fields, codes of conduct aiming at diversity of design and development teams are in our view essential to achieve less biased and more balanced AI systems. However, it is important that these codes of conduct are meaningful and have an actual impact on the design, development and deployment process. To avoid that these codes of conduct foster diversity in name but not in fact, we believe it is important to include a variety of stakeholders, in particular civil liberties and fundamental rights organisations.

Annex III: High-risk AI systems referred to in Art. 6(2)

Annex III contains a list of areas where AI applications should be considered high-risk and therefore subjected to more stringent control. We suggest that the list of high-risk applications should be expanded to include:

- Emotion recognition systems – these systems are based on invalid scientific theories, are highly invasive and could cause harm to self-expression. Greater regulatory oversight for these systems (if they are at all to be deployed) is necessary and should be anchored in the AI Act.
- Algorithms deployed in the field of insurance – these algorithms can determine who has access to quality insurance and at what cost, entailing a serious risk for discrimination in access to basic services.
- Algorithms used in health administration – such algorithms have already been found to be discriminatory and again pose a serious risk to fundamental rights (Ziad Obermeyer et al. Dissecting racial bias in an algorithm used to manage the health of populations. In: Science 366.6464 (2019), pp. 447–453. https://science.sciencemag.org/content/366/6464/447)
- Any algorithms that interact with children and youth should be considered high risk, whether or not they form part of a product. As the project Twisted Toys demonstrates, children and youth are exposed to manipulative design and algorithmic systems. The AI Act should seek to prevent any harm to children and youth by way of algorithmic manipulation and actively monitor the development of the market of AI systems targeted to younger

people. Likewise, the impact of algorithms on persons of a higher vulnerability due to age, physical or mental capacity should be further researched, possibly leading to the addition of such a field of application to Annex III.

As mentioned further above, we would furthermore encourage the Commission to regularly evaluate the list of high-risk AI systems and add areas of application if there is a foreseeably high risk – not only, as foreseen in Art. 7, adding AI applications to the areas defined in Annex III. AI systems are emerging technologies and it could prove too restrictive if high-risk areas of application are defined prior to widespread adoption of the technology and cannot be expanded at a later point in time.

Annex IV: Technical documentation

The requirements for technical documentation which have been proposed are a welcome step. We strongly support the points included, in particular the requirement in Annex IV para. (2) point (d) to include a datasheet describing the training methodologies and data sets used, in particular the provenance of the data used. In our opinion, this provision will incentivise the development of high-quality data sets which do not rely on data obtained without consent or contrary to declared licenses (such as Creative Commons licenses).

One of the aspects currently missing is the possibility of adverse effects due to interactions between different AI systems. Given the increased development and application of AI systems, a scenario in which multiple systems form a system complex is likely for use cases as different as home automation and medical treatments. Users should be informed if the provider has tested whether the AI system is interoperable with other types of systems, and if so, which systems were tested.

Exemptions for law enforcement and related activities

We are very concerned with the number of exemptions provided for AI systems used for law enforcement purposes. In a union of constitutional states such as the European Union, it is paramount to ensure balance between the rights and legitimate interests of citizens and the rights and legitimate interests of states. Citizens must have adequate assurances that products or services employed by law enforcement or related agencies (e.g. migration, asylum and border control) are safe, robust and reliable and reflect the values of European society. Yet the exemptions provided in this proposal would at best relax the standards to which AI systems used by these agencies are held; at worst, these provisions could lead to a time-limited free-for-all, incentivising law enforcement and related agencies to obtain AI systems before the Regulation comes into effect, and avoid enhanced compliance costs by not updating the systems previously obtained in a significant manner.

We are particularly concerned about the provisions foreseen in Art. 83 and Art. 43 para. (2) and (6). Considering the relaxation of requirements for AI systems used for law enforcement purposes, the reduction of fines in case these requirements are not complied with (Art. 71 para. (7) and Art. 72 paras. (2-4) for Union agencies, bodies and institutions) are not appropriate. Public authorities have the duty to protect public interest. Should they fail to adhere to legislation protecting the public interest, they must be held responsible for breaking the public's trust – in accordance also with the fundamental right to good administration.

Beyond the concerns expressed above, we are deeply troubled by the exemptions from compliance with the Regulation provided through Art. 2 para. 3, which exempts systems used for military

purposes, and Art. 2 para. 4, which exempts international organisations or public authorities in third countries applying AI in the EU based on agreements for law enforcement or judicial cooperation. As the Regulation stands, the lack of balance between the rights and legitimate interests of citizens and of states is increased by removing military applications from fundamental rights and safety requirements for other applications.

Product safety premise

The proposed Regulation treats AI as a product or as a component of specific types of products, and introduces a specific form of legislation designed to mitigate risks arising from certain uses for this product. While this is certainly a very important aspect of AI technology and users of AI products as well as people who are affected by AI products must have adequate protection from harm, solutions to this aspect of the problem may not be sufficient to handle another aspect of AI technology – namely its function as infrastructure.

We believe that an effective way of observing infrastructure is to observe whether, and if so to which extent, people adapt to these structures. As cars have entered our streets, we have begun to change our streets, outfitting them with traffic lights and pedestrian crossings and speedometers, and shifting expectations of the basic knowledge and skills to be acquired in the course of one's life (e.g. learning to cross the street safely and how to drive a car). AI technologies are developed to deal with an amount of information that humans are not able to process at the necessary speed. People have begun to adapt to this technology, whether by relying on search engines to offer the most relevant information or by actively changing their behaviour online in order to elicit a particular response from a recommendation algorithm.

Cars have had a lasting impact on our society, not only in terms of travel speeds or pollution, but also in the way we build cities. The archetypical suburb is a form of living that is only practical if its residents have cars: suburbs consist only of houses, without grocers or providers of other convenience goods available in easy reach – without a car, that is. Similarly, buildings in city centres may be partially repurposed to offer indoor parking, changing the way pedestrians experience the city. If AI proves to be a technological revolution – and there are strong indicators that this might be the case, considering it is the core technology of some of the most widely used digital products in the world –, it is necessary to think about the changes this technology may bring and develop strategies for how we as a society can act in the face of these changes.

As it has been proposed, the AI Act offers a type of vehicle or building safety norms, but no city planning strategy. We encourage the European Commission to engage with a variety of different stakeholders in order to develop a vision and a concrete plan for the information infrastructure of the future, of which AI might form a part.

Through the product safety perspective, questions of exploitation have unfortunately also been sidestepped. Product safety addresses concerns related to one aspect of the economic lifecycle of production, distribution, consumption and waste: namely what happens when the product is consumed. Environmental damage during production is an issue that the proposed Regulation cannot cope with. Equally, questions of appropriated labour – both in the data collection process and in the contribution of the user in improving the system – cannot be solved through product safety legislation. Questions of competition in an age of AI systems – such as algorithmic pricing

systems – have also remained unanswered by this proposal. Most importantly, the AI Act as it has been drafted does not give affected persons the right to opt out of being affected by an AI system or even the right and means to complain about harmful or defective AI systems.

These issues as well as others are salient and call for well-balanced, long-term strategies. We hope the European Commission will return to deliberate on these questions, together with all interested stakeholders, and develop solutions that will empower humans and ensure healthy and safe environments, be they ecological, social or economic.

***

We commend the work of the European Commission in developing a framework for AI systems. The future of AI development, research and applications will be shaped by this framework and by those frameworks we hope the Commission will next turn its attention to. We hope our feedback contributes to the further refinement of the proposed provisions and will gladly answer any questions about our views on this topic.

On behalf of Women in AI Austria,

Carina Zehetmaier

President of Women in AI Austria

Signiert von: Carina Zehetmaier

Datum: 06.08.2021 15:48:07

Dieses mit einer qualifizierten elektronischen Signatur versehene Dokument hat gemäß Art. 25 Abs. 2 der Verordnung (EU) Nr 910/2014 vom 23. Juli 2014 ("eIDAS-VO") die gleiche Rechtswirkung wie ein handschriftlich unterschriebenes Dokument.

Dieses Dokument ist digital signiert!

Prüfinformation:
Informationen zur Prüfung der elektronischen Signatur finden Sie unter:
www.handy-signatur.at

www.a-trust.at

A-TRUST
einfach sicher