Commentary to the Commission's proposal for the "AI Act" – Response to selected issues

Centre for Commercial Law, School of Law, University of Aberdeen

This response is provided by a working group of the Centre for Commercial Law (CCL) at the University of Aberdeen. The working group consists of Dr Péter Cserne, Dr Rossana Ducato, and Dr Patricia Živković, and the response incorporates comments by Prof Abbe Brown, Dr Irène Couzigou, Dr Georgios Leontidis, Prof Nir Oren, Dr Clare Sutherland, Dr Paula Sweeney, Dr Burcu Yüksel Ripley.

The analysis provided in this response contains a preliminary analysis of selected issues.

The views and opinions reported in this response are submitted on behalf of the Centre for Commercial Law and do not necessarily express the position of the School of Psychology, School of Divinity, History, Philosophy & Art History, and the School of Natural and Computing Science.

Table of Contents

1. Market integration, market regulation and fundamental rights 1
2. The scope of application 2
   2.1. Focus on the pre-market stage 2
   2.2. The extra-territorial effect 3
3. The risk-based approach 4
   3.1. Prohibited AI practices 4
   3.2. High risk AI systems 8
4. Measures in support of innovation 11
5. A place for the "AI subject" 13
6. Plain language and beyond 14

We welcome the possibility to contribute to the policy debate on this proposal for a Regulation laying down harmonised rules on Artificial Intelligence and amending certain Union legislative Acts (hereinafter, "Draft AI Act").

The proposal already represents an important step in the EU and at the international level, as it recognises the need to steer the responsible development of innovation. With this proposal, the legal component gains a key role in the EU approach to the creation of trustworthy and reliable AI.

We welcome the horizontal approach of the Commission and the way this proposal is building on the existing legal environment (for instance, the product safety directives and regulations). In principle, the risk-based approach, the emphasis on data governance, the provision on human oversight, the establishment of a public database for AI systems, and the goal of developing human rights-centred AI systems are certainly positive aspects of this proposal.

However, we believe some issues should be addressed or clarified in order to strengthen the proposal and ensure that the human-centred approach is embedded in practice.

1. Market integration, market regulation and fundamental rights

The main justification for the Act (Recital 2) invokes the market integration logic which has provided the underlying motivation for the bulk of EU law, as currently expressed in Article 114 TFEU: in order to remove barriers to market entry, the divergence of national rules should be replaced by uniformity ("maximum harmonisation").

Therefore, the reasons/motivations for the Draft AI Act appear at least semantically to turn around the term "market". Most rules focus on "the placing on the market, putting into service or use" of AI.[1]

As some commentators have noted: "The proposal mixes reduction of trade barriers with broad fundamental rights concerns in a structure unfamiliar to many information lawyers, and with significant consequences on the space for Member State action."[2]

First, many of the high-risk uses of AI regulated in the Draft AI Act take place in non-market settings (e.g., law enforcement) or sectors that only fall under a very broad definition of a market (e.g., medical diagnostic AI). High-risk activities (Title III) that threaten "health, safety and fundamental rights" are regulated even in non-market settings. The tension between market regulation and rights-protection is also expressed in the legal basis of the Draft AI Act.

[1] The Draft AI Act promises to regulate the entire lifecycle of AI, however. Indeed, there are rules both supporting and regulating earlier stages in the lifecycle: innovation and development – we make suggestions for these in sections 2.1. and 4 below.

[2] Veale, Michael, and Frederik Zuiderveen Borgesius. "Demystifying the Draft EU Artificial Intelligence Act." *arXiv preprint arXiv:2107.03721* (2021), p. 3.

Formally, only the use of facial recognition for law enforcement has a different legal basis in the TFEU. The rest is based on Article 114 TFEU (see Recital 23).

There is a danger that the general framing of the Draft AI Act as an instance of market integration obscures the aim of rights-protection and confuses non-expert readers.

To be sure, the market integration logic is broad and allows for substantive limits on market (commercial) activities. In fact, some of the prohibitions in the Draft mirror requirements set by the Unfair Commercial Practices Directive and the Directive would also apply directly to commercial practices that involve the use of AI. Other rules of the Draft AI Act interact with the requirements of the GDPR. Yet the recurrent phrase "the placing on the market, putting into service or use" does not reflect with sufficient clarity that the Draft AI Act also applies to non-market or non-economic activities by private or public providers/users, insofar as they endanger fundamental rights.

Second, there remains uncertainty as to how much competence is left for the Member States to regulate aspects of AI use in a more restrictive way than the Act. This uncertainty arises partly due to the broad definition of the scope of the Draft AI Act, together with the focus of substantive rules on so-called high-risk AI, where the criteria of "high-risk" are problematic. If the Draft AI Act is considered as a measure of maximum harmonisation for all AI systems within the scope of the Draft AI Act – this seems to prevent the Member States from enshrining legitimate policies, goals and values into national rules for non-high-risk AI systems, whenever these go beyond the very light rules envisaged in this Draft for such non-high-risk systems.

2. The scope of application

2.1. Focus on the pre-market stage

Article 3 of the Draft AI Act defines 'artificial intelligence system', 'provider' and 'user'. It also identifies what could be an 'output' of an AI system within the definition of these systems, and it enumerates 'content, predictions, recommendations, or decisions influencing the environments they interact with'. These are all relevant definitions for our feedback below.

Regarding the scope of application, Article 2 states that the Regulation will apply to:

> "(a) providers placing on the market or putting into service AI systems in the Union, irrespective of whether those providers are established within the Union or in a third country;
>
> (b) users of AI systems located within the Union;
>
> […]

This formulation puts emphasis on the final stages of development, such as the "placing on the market or putting into service". It is implicit that providers who plan offering their systems will have to consider the Act's rules from the first stages of development of their technology. However, this preventive approach does not emerge clearly from the Act and its Article 2.

While the Draft AI Act envisages regulating the entire lifecycle of AI, the rules are not equally tight throughout the cycle. Hence, we would suggest extending the scope of application of the Act to developers of AI systems explicitly and establishing specific rules for them to ensure that also the stages of design, research, and experimentation are conducted to protect the safety, health, and fundamental rights of research participants. We will make further comments on the development of AI in the context of "creating a safe space for innovation" below in Section 4.

2.2. The extra-territorial effect

We welcome the approach of the Draft Act providing a broad extra-territorial scope of application while requiring links to the EU (Article 2(1)(a)). Such an approach is necessary due to the specific nature of the subject matter and the globalisation of the market for AI.

Article 2(1)(c) states that the Act will apply to: “providers and users of AI systems that are located in a third country, where the output produced by the system is used in the Union […].”

We would suggest clarifying this ground for the extra-territorial effect and its application. Namely, based on this ground, the AI Act applies to providers and users located in a third country when the output by the system is used in the Union. Recital 11 clarifies that this ground will cover the situation where "an operator established in the Union […] contracts certain services to an operator established outside the Union in relation to an activity to be performed by an AI system that would qualify as high-risk and whose effects impact natural persons located in the Union”. We notice that the formulation of both Article 2(1)(c) and Recital 11 can create some problems of interpretation. First, whereas Article 3 makes the distinction between the users and providers, it would be better to avoid the verb "use" in relation to the outputs only in Article 2(1)(c). A better solution in line with the general principle of legal certainty would be defining these potential “outcome users” in a more precise manner under Article 2. The term "operator" in Recital 11 does not seem to respond to such a challenge. Operator in the AI Act means “the provider, the user, the authorised representative, the importer and the distributor” (Article 3(8) Draft AI Act). Therefore, it does not appear to capture the situation of the actor “using” the AI system’s output in the EU.

In addition, we notice that Recital 11 seems to narrow the actual scope of Article 2(1)(c) when it refers to "high-risk" systems and "effects on natural persons". In line with Articles 2(1)(a)-(b), which do not contain any limitation to the scope of application due to the level of risk of the AI system, it might be useful to consider whether Article 2(1)(c) should encompass all AI systems (including, for instance, the prohibited practices that go beyond the “high risk”).

Moreover, if the rationale of the Draft AI Act is to ensure the protection of the fundamental rights recognised in the EU Charter, the Act should also apply to:

- Providers established in the Union that create and offer AI systems, independently from where they place the system on the market or into service.[3] This extension would avoid the situation where European providers develop systems that would be prohibited or classified as high risk in the EU and make them available in third countries (which do not offer an equivalent level of protection of fundamental rights) without the guarantees established in the Draft AI Act.[4] It will also cover the situations of testing/quality assurance processes that happen in-house, i.e. before the placement of the AI system on the market, as the intended purpose may differ.
- Providers and users of AI systems located in a third country if the outcome of the AI systems is used outside the EU but with an impact on persons located in the EU. This suggestion would be in line with the territorial scope of application of the GDPR, which applies extra-EU, where the controller targets citizens located in the Union (Article 3(2) GDPR). We are in this regard suggesting to centre the territorial application not only around the EU single market but also around EU citizens. In our opinion, connecting the application of the AI Act to EU citizens would lead to an equal footing for providers and users in the EU Member States and third countries. This level playing field would serve as an incentive for innovation for providers established in the EU.

3. The risk-based approach

The Draft AI Act intends to differentiate rules and obligations based on the level of risk entailed by the AI system. It distinguishes between:

- Prohibited AI practices involving a level of risk considered "unacceptable"[5] or contrary to EU values;
- High-risk AI systems, i.e. systems that pose a significant risk to the health, safety, or fundamental rights of natural persons;
- All other cases.

3.1. Prohibited AI practices

Four different practices (listed in Article 5) are prohibited by default because the Commission retained that they pose an unacceptable risk to the health and safety or fundamental rights of natural persons, or they are considered contrary to European values. This attempt to exclude certain systems is commendable, but the formulation is not entirely satisfying. These prohibitions are too narrowly designed and might be difficult to enforce in practice.

[3] In the current Draft, the definitions of "placing on the market" and "putting into service" refer exclusively to the Union market. See, Article 3(9)-(11).

[4] On this risk, see also Veale and Borgesius (n. 2), p. 8.

[5] See point 5.2.2, Explanatory Memorandum, Draft AI Act.

Article 5(1)(a) prohibits systems deploying subliminal manipulative techniques, while Article 5(1)(b) concerns the ban of systems exploiting vulnerable subjects. The Commission provided some examples; however, it will be helpful to have more terminological and conceptual clarification, for instance, with regard to notions such as “manipulation” and “subliminal technique”.

Both in Articles 5(1)(a) and (b), the prohibition seems to require the intention (“in order to”)[6] of the AI practice to materially distort the person’s behaviour. This kind of situation is rightly prohibited, but the current formulation will leave aside those cases where a system is likely to manipulate persons’ behaviour beyond the situations of intended uses (including the foreseeable misuses). A formulation along the lines of the Unfair Commercial Practices Directive’s wording on misleading practices, which focuses on the probability of the occurrence of the unfair effect in practice rather than on the intended purpose of the system (i.e. of its developer and user), would be more helpful here.

The emphasis on individual harm in both provisions is problematic as well. First, we might presume that damage is likely to occur when a manipulative practice entails an arbitrary interference with the rights of self-determination and human dignity. It is, indeed, difficult to imagine a situation where manipulation (which is hidden or covert influence)[7] or non-transparent nudges would be considered acceptable in a democratic society.[8]

Second, the harmful conduct might materialise over time through the cumulation of adverse effects.[9] Finally, manipulative systems could pose collective threats and harms.[10] A wide variety of AI systems – through the adoption of malicious choice architecture, selection, and frame of data - might negatively manipulate our behaviour online and offline, affecting fundamental public interests such as freedom of press, democracy, the rule of law.

Article 5(1)(b) deals specifically with the prohibition of AI systems that exploit the vulnerability of protected groups, such as children or persons with disabilities. This formulation is in line with other EU provisions that tend to protect vulnerable groups (e.g., consumer protection, data protection). Nevertheless, there is already a growing body of literature showing that the protection of traditional vulnerable categories might not be

[6] See also Recital 16 Draft AI Act.

[7] Susser, Daniel and Roessler, Beate and Nissenbaum, Helen F., Online Manipulation: Hidden Influences in a Digital World (December 23, 2018). 4 Georgetown Law Technology Review 1 (2019), Available at SSRN: https://ssrn.com/abstract=3306006 or http://dx.doi.org/10.2139/ssrn.3306006.

[8] On the epistemic distinction between transparent and non-transparent nudging types, see Hansen, Pelle Guldborg, and Andreas Maaløe Jespersen. "Nudge and the manipulation of choice: A framework for the responsible use of the nudge approach to behaviour change in public policy." *European Journal of Risk Regulation* 4.1 (2013): 3-28. See also, Osman, Magda, et al. "Whom do we trust on social policy interventions?." *Basic and Applied Social Psychology* 40.5 (2018): 249-268.

[9] As stressed in Veale and Borgesius (n. 2), p. 5.

[10] Susser, Daniel, Beate Roessler, and Helen Nissenbaum. "Technology, autonomy, and manipulation." *Internet Policy Review* 8.2 (2019).

enough.[11] For example, people in a difficult financial situation can be negatively affected (economically and psychologically) by an AI system that exploits their circumstances to target them with ads of high-interest-rate loans.[12] However, "non-conventional" vulnerabilities (as in the previous example) risk escaping the protection granted by the existing legal framework. We would therefore encourage to embrace a more innovative and forward-looking perspective in the AI Act, shifting from a static conception of vulnerability that looks at some intrinsic conditions of the subject towards an idea of "layered"[13] or "digital"[14] vulnerability that people experience due to the architectural choice of the system.

Article 5(1)(c) prohibits social scoring when it is done by public authorities. This definition is again a very narrow approach, as individuals can suffer similarly significant harms if the system is used by private entities.[15]

Finally, Article 5(1)(d) deals with real-time remote biometric identification for law enforcement purposes. Remote biometric identification is "AI system[s] for the purpose of identifying natural persons at a distance through the comparison of a person's biometric data with the biometric data contained in a reference database, and without prior knowledge of the user of the AI system whether the person will be present and can be identified."[16]. The "real-time" characterisation means that the AI system is capable of instant or short delay identification.[17]

The use of real-time remote biometric identification for law enforcement purposes is prohibited by default unless it is strictly necessary for one of the following purposes:

"(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/JHA and

[11] Wachter, Sandra, and Brent Mittelstadt. "A right to reasonable inferences: re-thinking data protection law in the age of big data and AI." *Colum. Bus. L. Rev.* (2019): 494; Zuiderveen Borgesius, Frederik J. "Strengthening legal protection against discrimination by algorithms and artificial intelligence." *The International Journal of Human Rights* 24.10 (2020): 1572-1593.

[12] Art. 29 Data Protection Working Party, Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679 (2018), p. 10.

[13] As suggested in the data protection framework by Malgieri, Gianclaudio, and Jędrzej Niklas. "Vulnerable data subjects." *Computer Law & Security Review* 37 (2020): 105415.

[14] As suggested in the consumer protection perspective by Helberger, N., et al. "EU Consumer Protection 2.0: Structural Asymmetries in Digital Consumer Markets, A joint report from research conducted under the EUCP2. 0 project." (2021).

[15] On algorithmic governance by governments and corporations, see, Pasquale, Frank. *New Laws of Robotics*. Harvard University Press, 2020, pp. 138 ff.

[16] Article 3(36) Draft AI Act.

[17] Article 3(37) Draft AI Act.

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".

This use of the AI system will have to be authorised by the judicial authority *ex ante*. However, "in a duly justified situation of urgency," the authorisation can be requested later.

In our opinion, there are six main problematic issues with this provision:

1. The distinction between "real-time" and post biometric identification is not entirely convincing. The "significant delay" seems to be the relevant factor to distinguish the two cases, but it is not clearly defined. Furthermore, as stressed by the EDPS and EDPB in their joint opinion on the Draft AI Act, the intrusiveness of biometric recognition systems do not necessarily depend on the "timing" of use: "Post remote biometric identification in the context of a political protest is likely to have a significant chilling effect on the exercise of the fundamental rights and freedoms, such as freedom of assembly and association and more in general the founding principles of democracy"[18];
2. It is not clear from the context and the Preamble why the focus is only on remote real-time biometric identification by law enforcement and not also on such usage by other actors and for other purposes. The activities and the usage by, e.g., private companies when it comes to this type of AI system would be classified as high-risk AI systems. However, the risks of creating a "surveillance society" might be even higher and more challenging to defend against when corporations and other private entities are involved.
3. The conditions legitimising real-time biometric recognition are quite vague (i.e. "duly justified situation of interest") and leave wide space for interpretation, risking affecting the rationale of the prohibition.
4. Article 5 limits the application of the prohibition of 'real-time' remote biometric identification systems to *publicly accessible spaces*. These are defined as "any physical place accessible to the public, regardless of whether certain conditions for access may apply"[19]. While it is clear from Recital 9 that this "notion does not cover places that are private in nature and normally not freely accessible for third parties, including law enforcement authorities, unless those parties have been specifically invited or authorised, such as homes, private clubs, offices, warehouses and factories," it is not evident what the rationale for the exclusion of online spaces is. The same Recital states that "[o]nline spaces are not covered either, as they are not physical spaces;" however, in our opinion, online spaces should be equally covered given the increase of everyday, public activities in the virtual world. The conduct of individuals in that space should be equally treated as the conduct of individuals in physical areas, and

[18] 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), 18 June, 2021, p. 31.

[19] Art. 3(39) Draft AI Act.

they should enjoy the same protection, notwithstanding whether the 'real-time' remote biometric identification systems are used by public authorities or private entities.[20]

5. The prohibition under Article 5 deals exclusively with biometric identification, as defined in Article 3(36) Draft AI Act. However, a similar level of intrusiveness could be reached by an emotion recognition system used in a publicly accessible space, e.g., to scan for signs of emotional agitation in the crowd. What should also be taken into account is the AI system's proficiency (or lack of it) when it comes to the recognition of emotions/affect, which warrants equal caution in use. Therefore, a prohibition on publicly available spaces should extend to emotion recognition systems as well as identity recognition. Furthermore, other emotion/affect recognition AI systems should be independently regulated (see below under 3.2.1).
6. Article 5(1)(d) does not cover biometric categorisation either. In this regard, we concur with the opinion of the EDPB-EDPS that such systems should be prohibited when clustering people based on the grounds that are protected under the EU anti-discrimination framework, where the scientific validity of such systems is not proven, or where their deployment is in contrast with EU values.[21]

3.2. High risk AI systems

3.2.1. Classification of systems as high risk

The Draft AI Act differentiate obligations for operators based on the level of risk that AI systems pose. The declared goal of the Draft is to regulate systems that "have a significant harmful impact on the health, safety and fundamental rights of persons in the Union"[22]. The impact assessment and, therefore, the classification of systems based on their risks is already done by the legislator. Article 6 of the Draft AI Act, which needs to be read in conjunction with Annex II and Annex III, establishes a catalogue of high-risk AI systems and sectors.

This list can be very useful to ensure legal certainty. Nevertheless, catalogues bring a high risk of rapid obsolescence, and the list might need regular updates, subject to technological evolution and the growing awareness about certain AI applications. The power to update the list is centralised and reserved to the Commission.[23]

[20] A similar proposition was made by EDPB and EDPS: "Thus, for consistency reasons, AI systems for large-scale remote identification in online spaces should be prohibited under Article 5 of the Proposal. Taking into account the LED, the EUDPR and GDPR, the EDPS and EDPB cannot discern how this type of practice would be able to meet the necessity and proportionality requirements, and that ultimately derives from what are considered acceptable interferences of fundamental rights by the CJEU and ECtHR." ibid 12.

[21] Ibid., p. 33.

[22] Recital 27 Draft AI Act.

[23] Article 7 Draft AI Act.

This approach risks being too static and might not capture all the systems that can affect fundamental rights based on the interaction between the system, the user, the individual (subject to the AI system), and the environment. Human rights impact assessments are traditionally based on a case-by-case evaluation.[24] The identification of problematic sectors can be an element to be considered in the assessment, but not the only factor.

In addition, we notice that there are already some potential loopholes in the list provided in Annex III.

Point 8 identifies which AI systems in the administration of justice and democratic processes are to be considered high-risk AI systems, and it lists:

> “(a) AI systems intended to assist a judicial authority in researching and interpreting facts and the law and in applying the law to a concrete set of facts.”

The focus on “judicial authority” does not sufficiently take into account the use of AI systems in dispute resolution, as there are many adjudicative dispute resolution mechanisms, such as arbitration, adjudication, and expert determination, which might engage with AI systems in the same manner. This lack of scope would leave the parties to those disputes in a more vulnerable position, and it would undermine the use of dispute resolution methods alternative to litigation. Hence, these types of alternative dispute resolution could be included in Point 8.

Furthermore, Annex III addresses some aspects of emotion recognition, but not in a coherent way. We think that this omission should be remedied by including a separate provision on emotion recognition. This technology is not currently covered by Paragraph 1 of Annex III, which identifies "AI systems intended to be used for the 'real-time' and 'post' remote biometric identification of natural persons" as a high-risk AI system. Emotion recognition does not necessarily aim for the unique identification of a person. Still, its intrusive nature, the fact that unique identification could, in theory, be recovered from the emotion recognition data, and its chilling effects on people’s behaviours warrant a careful approach in regulation.[25]

[24] On the tool of the human rights impact assessment, Mantelero, Alessandro. "AI and Big Data: A blueprint for a human rights, social and ethical impact assessment." *Computer Law & Security Review* 34.4 (2018): 754-772; Mantelero, Alessandro and Esposito, Samantha, An Evidence-Based Methodology for Human Rights Impact Assessment (HRIA) in the Development of AI Data-Intensive Systems (March 22, 2021). Computer Law & Security Review 41 (2021): 105561.

[25] McStay, Andrew, and Lachlan Urquhart. "‘This time with feeling?’ Assessing EU data governance implications of out of home appraisal based emotional AI." *First Monday* 24.10 (2019); Ienca, Marcello, and Gianclaudio Malgieri. “The EU regulates Ai but forgets to protect our mind.” European Law Blog (7 July 2021), https://europeanlawblog.eu/2021/07/07/the-eu-regulates-ai-but-forgets-to-protect-our-mind/. On physiognomy and its dubious use in individual case scenarios, see Foo, Y-Z., Sutherland, C. A. M., et al. “Accuracy in facial trustworthiness impressions: Kernel of truth, or modern physiognomy? A meta-analysis.” Personality and Social Psychology Bulletin, in press; see also Agüera y Arcas, Blaise, Mitchell, Blaise, and Todorov, Alexander. “Physiognomy’s New Clothes.” Medium (7 May 2017), https://medium.com/@blaisea/physiognomys-new-clothes-f2d4b59fdd6a.

There are parts of Annex III that provide for certain usages of emotion recognition to be treated as high-risk AI systems. Point 6 of Annex III lists high-risk AI systems in law enforcement, such as: “(b) AI systems intended to be used by law enforcement authorities as polygraphs and similar tools or to detect the emotional state of a natural person.” Point 7 provides a list of high-risk AI systems in migration, asylum and border control management, and also mentions: “(a) AI systems intended to be used by competent public authorities as polygraphs and similar tools or to detect the emotional state of a natural person.”

Linking the use of emotion recognition systems to a specific type of activity, usually reserved for public authorities, means less scrutiny for private actors in the market. Given the sensitivity of the issue, the possible exploitation in other private areas of one’s life, especially in working places, and the intrusive nature of the technology, the transparency obligations set out in Article 52(2) of the Draft AI Act might not be enough to protect individuals and their fundamental rights. Therefore, we would suggest for consideration a general ban of large-scale emotion recognition systems in publicly accessible places[26], and the inclusion of emotion recognition systems as an additional category of high-risk AI systems in all the other cases.

3.2.2. Requirements for high-risk AI systems and conformity assessment

The core of the Draft AI Act focuses on the systems falling within the high-risk category, which will have to comply with the requirements in Chapter 2.

We welcome, for instance, Article 13, which establishes a transparency obligation towards the users. The latter should receive “concise and clear information”[27] that allows them to “interpret the system’s output and use it appropriately”[28]. In Recital 47, the Draft AI Act recognises the importance to address the opacity of the systems which makes them "incomprehensible or too complex for natural persons”[29]. Therefore, we wonder why an equivalent transparency obligation is not placed onto the user vis-a-vis the individuals who are subject to the AI system. This measure will encompass the technology's whole life cycle, including all the relevant subjects involved.

Before placing high-risk AI systems on the market or putting them into service, providers will have to assess the conformity with the requirements set in Chapter 2. To this end, there are three possible options: the conformity can be presumed if the AI system complies with existing CEN and CENELEC standards (Article 40). Alternatively, where harmonised standards do not exist or are not sufficient, the Commission can adopt commons specifications (Article 41). Also in this case, compliance with the specifications creates a presumption of conformity.

[26] See, above Section 3.1., point 5.

[27] Recital 47 Draft AI Act.

[28] Article 13 Draft AI Act.

[29] Recital 47 Draft AI Act.

Finally, providers can interpret the requirements in Chapter 2 by themselves and follow the conformity assessment procedure described in Annex VII, which includes the involvement of a notified body (Article 43(1)). However, when a high-risk system included in Annex III (points 2-8) is involved, the conformity assessment is based on an internal control procedure (Annex VI), left entirely to the provider.

Our main concerns regard the risk to outsource one of the core goals of this piece of legislation – the creation of human-centred AI - to self-regulation, through a system designed from a very different perspective, less connected with human rights. The Commission seems to recognise such tension, as Recital 64 affirms that professional pre-market certifiers are more experienced in the field of product safety rather than in the area of high-risk AI systems. However, the response to this situation, namely the delegation of the conformity assessment to the provider "at least in an initial phase of application of this Regulation"[30], is problematic.

Considering the emphasis put by the Draft AI Act on the self-assessment procedure (*ex ante* perspective) and the not infinite resources of the enforcement authorities we are seeing in other fields (*ex post* perspectives), we are concerned that the effective compliance with the Regulation will not be adequately controlled.

4. Measures in support of innovation

The idea of regulating the whole lifecycle of AI in the EU in a way that is "innovation-friendly, future-proof and resilient to disruption"[31] is to be welcomed – yet the way these high-level ideas spell out in specific provisions looks rather sketchy and needs to be rethought.

The various stages of development, testing, marketing, and use, require different forms of legal infrastructure/regulation and the Act can be improved with respect to both how these regulatory logics match each other (avoiding loopholes) and in the substantive rules (avoiding hypocrisy and ineffectiveness).

The EU faces the challenge of attracting innovation, in particular R&D activities in the AI sector, to its territory while holding such systems, including their development and testing, against high standards, ultimately justified by fundamental rights and values.

As mentioned in other sections of this response, worryingly, the Act includes some provisions that subtly and perhaps unintentionally allow the development/provision of high-risk and even prohibited technologies in the EU, which would not be allowed to be put to the market here. Yet, they are offered for use abroad, possibly by less scrupulous governments and companies. This concern also possibly applies to the case of Title V, which sets the framework for Member States to provide an innovation-friendly environment in the form of "regulatory sandboxes". The provisions on regulatory sandboxes, in fact, do not seem to prevent the

[30] Recital 64 Draft AI Act.
[31] Section 5.2.5, Explanatory Memorandum, Draft AI Act.

experimentation on AI systems prohibited under Article 5.[32] This limitation under the Draft AI Act is something that should be reconsidered.

Title V also provides regulatory easing and measures of support for SMEs and start-ups. These are to be welcomed.

The provisions on regulatory sandboxes seem to be inspired by the idea of regulatory competition and experimentation: a markedly different concept than the idea of "market integration" via harmonisation/uniform rules that in the main underlies the Act. While innovation support is a legitimate ancillary purpose of the Act, we think that the current provisions on regulatory sandboxes do not set sufficiently clear rules for such research and experimentation.

In regulatory competition, the idea of "race to the top" refers to a mechanism where decentralised "jurisdictions"/units have the incentive to learn from each other and provide regulatory mixes that attract businesses (here, AI developers). The danger of "race to the bottom" is meant to be prevented by common minimum rules and procedural safeguards. Within the EU, Member States can compete in this manner. On a global scale, the EU itself is such a competing unit.

The idea of Member States attracting innovation to their territory by providing "sandboxes" where innovators can experiment with high-risk or even prohibited (not usable) technology may be seen as following this idea of intra-EU competition. But, more likely, it is an instance of decentralisation/subsidiarity pure and simple, with some elements of cooperation and information exchange (Article 53(5) Draft AI Act: "good practices, lessons learnt and recommendations") – without, however, further thoughts on how it may impact the AI market dynamics within the EU: it is not a competition among Member States that is envisaged here.

More importantly, Title V does not seem to reflect systematic thinking on how it may affect the attractiveness of the EU as the locus for AI innovation. Yet these "rules of attraction" are left to be specified elsewhere, perhaps, in "implementing acts" (Article 53(6) Draft AI Act) as well as other rules, measures and policies beyond this Act.

As for its effects, regulatory sandboxes may or may not be taken up by the Member States, and in turn, by AI developers, given their respective costs and benefits. Here various scenarios are possible.

It is possible that the rules for participation in the sandboxes are so restrictive that they push experimentation outside a particular Member State or even the entire EU, risking the capacity of the EU to attract innovation. A key question here is: What makes sandboxes attractive if AI

[32] Recital 16 seems to go into this direction, with the caveat that research for legitimate purposes in relation to AI systems using subliminal techniques or exploiting vulnerabilities of individuals can be conducted only if the "use" of the AI system does not expose individuals to harm and comply with recognised ethical standards for scientific research.

can be developed and tested outside such sandboxes just as well? The current brief and vague rules do not clearly indicate what is likely to happen. This suggests that the rules on regulatory sandboxes are perhaps expressive of some values but ineffective as regulators of behaviour.

Perhaps, it would be sensible to encourage innovation in technologies that are compliant with the Draft AI Act and other EU law, such as the GDPR, by design. This would be a way to credibly demonstrate the EU's aspiration for global leadership in delivering AI systems that promote the values of the EU and respect fundamental rights.

A key provision of this Title is Article 54. It sets up some specific rules permitting the re-use of personal data in sandboxes for the development and testing is of certain "public benefit" AI systems, under few limiting conditions. This provision seems to establish a lawful basis for the further processing of personal data initially collected for a different purpose (Recital 72), but its relationship with Recital 41 ("This Regulation should not be understood as providing for the legal ground for the processing of personal data, including special categories of personal data") should be clarified. More importantly, the rules for the research in the regulatory sandboxes should be better coordinated with the relevant data protection rules.[33]

5. A place for the "AI subject"

Despite the Commission's emphasis on a human-centred approach to AI development, the subject (the individual potentially affected by the AI system) is not considered in the proposal, except for the transparency obligations in Article 52.[34]

In principle, the transparency obligations are welcome, however, they are somewhat limited. The mandatory disclosure about the "mere" existence of the systems encapsulated in Article 52 might not be sufficient information for the "AI subjects" to act upon.

Therefore, we would suggest complementing the Draft AI Act with rights and remedies for the AI subject, such as:

- The right to be informed about the logic involved (not just about the mere existence/interaction with AI system), as provided to users under Article 13 of the Draft;
- The right to access the records and logs of the system - including error rates, (e.g. false alarm rate, miss rate), reliability, bias -, and the information related to the human oversight procedure and outcome;[35]

[33] EDPB-EDPS, Joint Opinion 5/2021, p. 64.

[34] As already mentioned, in our opinion, some systems (i.e. emotion recognition and biometric categorisation) should be prohibited or, at least, considered high risk.

[35] On the relevance of the human operator within an AI system – in particular, a face recognition system – see, White, David, et al. "Evaluating Face Identification Expertise: Turning Theory into Practice. Digested Analysis." *Evaluating Face Identification Expertise: Turning Theory into Practice*. 2020.

- The right to object to the use of the AI system;
- The right to contest the outcome of the AI system;
- The right to be informed about serious incidents that might have a significant impact on their rights and freedoms (similarly to Article 34 GDPR);
- The right to complain within a supervisory authority (similarly to Article 77 GDPR);
- The right to mandate NGOs to exercise the rights that will be recognised under the Act (similarly to Article 80 GDPR).

6. Plain language and beyond

The Draft AI Act addresses a complex issue, and the complexity is somewhat reflected in the text. The proposal contains 82 Articles, 89 Recitals, and 9 Annexes. It is already a very long document, and its reading is becoming much more difficult by the internal references to external acts (see, for instance, the long list in Annex II) and by the provisions that need to be read in conjunction with Annexes (i.e., the definition of AI system). Not only the structure is complex, but also the language used is highly technical. This issue is not peculiar to the Draft AI Act. Still, an accessible communication of its provisions to wider stakeholders (e.g., policymakers, lawyers, IT experts, law enforcement, agencies, businesses, users, academics, citizens) must be ensured.

The use of plain and clear language in a legal text is the first step. However, even if the Regulation will be easy to comprehend, users and providers (especially SMEs and start-ups) might face another set of hurdles, namely the lack of tools and expertise for complying with the rules and for ensuring a high level of protection of fundamental rights. The measures of support mentioned in Article 55 seem to go in this direction; however, they might not be sufficient. Therefore, we would suggest providing more resources and support to a broader number of actors in order to create a level playing field for the development of responsible innovation in the EU.