Council of the European Union
General Secretariat

PUBLIC

Brussels, 09 September 2021

WK 10581/2021 INIT

LIMITE

TELECOM

WORKING PAPER

This is a paper intended for a specific community of recipients. Handling and further distribution are under the sole responsibility of community members.

WORKING DOCUMENT

From:	General Secretariat of the Council
To:	Working Party on Telecommunications and Information Society
Subject:	Artificial Intelligence Act : PT comments

Delegations will find in annex PT comments on Artificial Intelligence Act.

1st September 2021

PUBLIC

Artificial Intelligence Act

PT comments (COM (2021) 206 final)

PT welcomes the European Commission's legislative proposal for a Regulation laying down harmonized rules on artificial intelligence ("AI Act" or "proposed Regulation").

We consider that it is of paramount importance to regulate Artificial Intelligence (AI). However, it is also imperative to ensure that the proposed Regulation finds the right balance between protecting fundamental rights and stimulating innovation. Otherwise, Europe will be quickly overtaken in the technology field by other continents.

Therefore, we strongly recommend ensuring that AI experiments in controlled environments, for internal research and scientific purposes do not have to comply with the same requirements as AI experiments in other scenarios. Given the relevance of this topic, we propose the introduction in the Regulation proposal of the definition of "controlled environments in AI", reinforcing/clarifying the wording of recitals 71, 72 and article 53.

In addition, it should be taken into account that SMEs and start-ups don't have the same means as bigger techs to deal with the high compliance costs and paperwork resulting from the provisions of the proposed Regulation. Hence, PT considers it is necessary to have support measures for SMEs and start-ups to ensure that the whole European AI innovation system remains competitive and attractive despite the challenges created by this proposal.

- **PT preliminary comments on articles 1 to 4, 8 to 15 and 52 and 69.**

Article 1 -Subject matter

(c)

In order to reduce legal uncertainty and ensure that the scope of the proposed Regulation covers all the current and future artificial intelligence technologies that can be used by natural persons, we suggest ensuring that the examples included in the article are mere examples (**highlighted in bold**):

"(…) (c) harmonised transparency rules for AI systems intended to interact with natural persons, **including** emotion recognition systems and biometric categorisation systems, and AI systems used to generate or manipulate image, audio or video content; (…)"

In addition to the examples already mentioned, we propose a reference to the automatic decision systems/algorithms (credit assignment, social benefits, insurances, etc.), as these are widely available and have a considerable impact in people's lives.

Article 2 - Scope

It came to our attention that AI researchers and software developers regularly upload AI models and other AI related materials to repositories, which have a critical and beneficial role in the software ecosystem. Therefore, given the wording of this article there is a risk that those who upload these materials to software repositories (e.g. open-source), or the operators of these repositories, could be viewed as a regulated entity without "placing on the market" or "putting into service" the system in the EU, which might have an impact on research and open-source software innovation on the EU.

Consequently, we recommend that the terms "placing on the market" and "putting into service" should specifically exclude use of AI systems for internal research and development purposes.

Article 3 - Definitions

First, and as a general comment, we kindly recommend ensuring that the definitions are set in alphabetic order with the aim of facilitating its reading, analysis, and application.

(1) "artificial intelligence system" + Annex 1

We consider that the definition of an artificial intelligence system is too broad including many software technologies applications, that may not always be considered artificial intelligence and that may pose no major concerns around data, opaqueness, safety, and reliability. It is important to note that it is included in the definition as AI techniques "logic-based approaches" (Annex 1, b)), "statistical approaches, Bayesian estimation, search and optimization methods" (Annex 1, c)), which basically covers almost all modern software-based product given the fact that at some level all software is logic-based. If we consider these algorithms as AI we will increase the scope of the proposed Regulation and create legal uncertainty

for companies that use these software's (that are not AI and do not create the same risks the European Commission intends to address through the proposed Regulation), and would have to assess if their software's fall within the legislation scope or not.

Furthermore, the solution expressed in the proposal - which opted for the concretization of the concept of "Artificial Intelligence System" - should adequately translate the concern, expressed, right from the start, in Recital 6. Therefore the definition to be adopted should "(...) be unambiguous to ensure legal certainty while providing sufficient flexibility to adapt to future technological developments" and "(...) be based on the main functional characteristics of software, in particular the ability, in view of a given set of objectives defined by humans, to create outputs such as contents, predictions, recommendations or decisions that influence the interacting environment, either in a physical or digital dimension" .

Yet considering the architecture built based on article 3.1 and its dynamic interaction with Annex I, we are not sure that these premises have been effectively realized.

"Artificial intelligence system" is defined as "a computer program developed with one or more of the techniques and approaches listed in Annex I", leading us to a "circular" specification and failing to fulfil its purpose. This definition has a primordial nature. It also needs to be technically robust to confer a degree of legal certainty compatible with the legal principles and values to be guaranteed by the proposed framework. But for that it will be essential, as we see it, to rethink this concept.

Therefore, we suggest reviewing Annex 1 in order to limit the scope of AI by including only the type of systems/techniques that have some kind of uncertainty on their output given the fact that this uncertainty is the main cause of AI risks. Likewise, we suggest using the AI definition made by the High-Level Expert Group on Artificial Intelligence:
"Artificial intelligence (AI) refers to systems that display intelligent behaviour by analysing their environment and taking actions – with some degree of autonomy – to achieve specific goals. AI-based systems can be purely software-based, acting in the virtual world (e.g. voice assistants, image analysis software, search engines, speech and face recognition systems) or AI can be embedded in hardware devices (e.g. advanced robots, autonomous cars, drones or Internet of Things applications)."

(2) "provider"

This notion appears to consider that all AI systems are developed as a stand-alone product or service and then "placed on the market"/ "put into service". Whereas the AI ecosystem is very diverse and there are many ways AI systems are developed and deployed, and there is almost never a singular entity or person that develops an AI system. AI systems are the result of numerous entities building on top of others' efforts, for example it may start by using open-source repositories created by several contributors and the resulting model might then be shared under an open-source licence for others to build on.

Therefore, we should ask ourselves "who among all the contributors "develops an AI system?" It is of paramount importance to consider the range of developers, researchers, and innovators that make up the open-source community, which has been crucial to advancing the state-of-the-art of AI development.

We need a more nuanced taxonomy to identify the relevant participants in the AI ecosystem and allocate the appropriate responsibilities and obligations to each one rather than a definition of "provider" that risks treating all contributions big and small to the same burdensome regulatory standards irrespective of their nature and role.

(15) "instruction for use"

We propose to clarify the definition "instructions for use" given the fact that we believe that this definition creates legal uncertainty regarding its scope. The "instructions for use" appear to be only applicable to "high risk AI" systems when these are useful and desirable for all AI systems as they are for other types of software.

(23) "substantial modification"

We believe this term needs further clarification.

(29) "training data"

We suggest removing the expression "including the weights of a neural network" because in our opinion it does not add anything to the described concept and creates legal uncertainty concerning its application to other parameters arising from the use of other training techniques.

(36) "remote biometric identification system"

In our opinion, it is also unclear the scope of this definition. The use of this system can pose risks to fundamental rights but can also have positive social benefits, such as monitor health and safety. Consequently, we recommend clarifying certain aspects to enable positive uses of these system.

Further, it is not understandable the meaning of identifying natural persons "at a distance", especially taking into account that high risk uses of remote biometric identification cover, not only "real-time" but also "post" identification, and so it raises the doubt as how can the identification be made "after the fact" in any other way other than "at a distance". It seems that the intention was to cover mass surveillance "where "many people are being screened simultaneously" but the language should be clarified to reflect that intent. Otherwise, commonplace AI systems that identify natural persons at a distance such as smartphones used to identify friends in photos are also regulated under this provision. Moreover, it is also not clear the intention behind the exclusion from the definition "where the "user of the AI system" has "prior knowledge ...whether the person will be present and can be identified." For example, consumers might use their smartphone's AI to find in their photos the faces of family and friends that they trained their device to recognise. In that example, it is unclear who the user of the AI system is. If the consumers are users, they arguably have "prior knowledge" whether the individuals in their contacts or their photo album can be identified by the device. But if the "user of the AI system" is the smartphone or software vendor that designed the AI system for the device, would they have prior knowledge? The language of this article should be clarified in order to not prevent common and beneficial uses of AI to which people would be willing to consent, if given the appropriate opportunity."

(New numbers)

We propose to include a definition of "personal data" in line with the Regulation (EU) 2017/679 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 (General Data Protection Regulation) given the fact that this concept is used several times in the proposed text and that is included in the text a definition of "biometric data".

Therefore, we recommend the inclusion of the following definition:

"(…)"Personal data" means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person; (…)"

Furthermore, it is important to bear in mind that as referred above in the AI ecosystem there are many participants that have an essential role on it and as so it is necessary to ensure that they are all regulated under this proposed Regulation, such as the "person/company that buys the AI product" (who/which is not the end-user), the "technology supplier", the "deployer", the "several contributors to the development of the AI product/system in a co-creation environment", etc.

Article 9 – Risk Management System

As a general comment, we believe that the risks this article intends to address with the recitals, that explicitly state the importance of addressing risks to health, safety, and fundamental rights, should be more harmonized. Furthermore, we consider to be of paramount importance to be very clear on the types of risks we are attempting to address and define clear procedures to help guide providers, developers, etc through the risk assessment process.

Additionally, it is important to stress out the need to define the concept "lifecycle", which is used in this article, as well as in several others. The undefinition of this concept will create legal uncertainty and confusion. Moreover, and as referred above AI products/systems are generally created by several contributors and usually using open-source technologies and as so it is necessary to define how the risks in these cases will be managed, e.g. will the person who used the open-source materials be responsible/liable for the materials used? Or will be the person who created the open-source material?

Finally, considering the principles of the New Legislative Framework it seems that AI system providers will carry most obligations and requirements established in the proposed Regulation.
Nevertheless, it is important to keep in mind that many obligations and requirements can only be managed, in practice, by the user (who controls

the AI system and its use). Even if a provider complies with all its obligations and requirements it cannot foresee all potential uses of the system.

(4)

Moreover, it is established that mitigation should be used until the "overall residual risk of the high-risk AI system is judged acceptable", once again, is not clear what it means "acceptable". Therefore, we consider it is necessary to develop best practices and standards to define these concepts in the proposed regulation.

(6)

It is written that the "testing procedures shall be suitable", however, yet again, it is not explained what it means to be "suitable" and as we referred above, due to the fact there are no best practices or standards, as there are for instance for data protection, the use of these terms will create legal uncertainty and as so we recommend to develop standards and best practices to define these concepts.

(5, 6, 7)

These numbers cover "testing procedures" but it is not clear which type of testing they are referring to. The lack of specification will create misunderstandings, legal uncertainty, and confusion. There are several types of test procedures, such as, unit tests, integration tests, performance tests, operational tests, etc. These tests are of paramount importance and its use will help to mitigate the risks.

Hence, we suggest adding an article to regulate the test phase of AI solutions in order to mitigate the risks of AI. Additionally, we also propose to mention technics such as Explainable Artificial Intelligence (XAI), referred in the EC Communication regarding Artificial Intelligence for Europe[1]. The use of these technics should be encouraged given the fact that these can help debugging and auditing activities.

Finally, this topic could also include a reference to the Responsible AI.

Article 10 – Data and data governance

Although we see the necessity to regulate this subject, we believe some of the requirements set in this article are infeasible and impossible to achieve.

[1] COM (2018) 237

PUBLIC

Data results from capturing real-world events. In many cases, these events were poorly captured or even result from human past decisions that challenge actual values associated with safety and security, fundamental rights, legal certainty, etc. It is normal to find out bias in data.

However, and according to the present proposal, this type of data can represent a threat because it perpetuates past human mistakes. We generally agree with this concern, but there are techniques that can help to identify and reduce a good part of data deviations, but it is difficult to reduce them to zero, as it is impossible for humans to avoid taking bias decisions.

Furthermore, the entire data collection and data management process may include access to old data that did not pass quality criteria such as those required in this proposed article, which does not mean that they have no value at all. So, concepts such as data relevance, representativeness, freedom of errors and completeness should be better defined. We agree that AI providers should share some exploratory data quality pratics with all others operators involved in the value chain, but this should not be mandatory, given the fact we know that data in lab works better for AI systems than data in production.

Additionally, it is important to keep in mind that there are situations where the inclusion of errors in datasets can be advantageous, for instance, in the healthcare context, the dataset used to train an AI model might include innaccuracies or other errors that cannot be cleaned due to their source (e.g. self-reporting of sypmptoms). But training an AI model on data that includes such innaccuracies might be helpful and even preferable where there is reason to expect that such errors might appear in input data once the system is put into service (e.g. in a hospital). Likewise, language models have tremendous potential to help solve critical societal challenges in domains such as healthcare, where there are large amounts of unstructured text from which AI systems could help researchers identify patterns, trends, and connections. But data from these domais often reflect people's assumptions, beliefs, etc which likely will have "errors".

Further, it is important to point out that bias in AI systems does not come exclusively from biased or incomplete data. For example, data could be perfectly accurate and representative but reflect structural or societal biases. Or AI developers could make assumptions during the design of their systems as to how they expect the systems to be used, and as a result the

system may exhibit bias. For example, an AI system that is built to try to predict criminality based on facial features will be biased from the outset due to faulty and biased) assumptions in the design of the system, no matter what data is used. Hence, bias is not just a data issue. Error-filled inputs can still lead to fair outcomes, just as error-free inputs can still lead to unfair outcomes, depending on the overall design and operation of the system. When it comes to bias, we should consider the specific context in which an AI system is used, and whether the system's design, inputs, and outputs are appropriate for that context, rather than focusing exclusively on data sets or any particular technology or methodology.

Besides, we also consider it is not possible to ensure that the training, validation and testing data is complete, given the fact that there is always more data that can be collected. We live in a world ever-changing and the data is increasing on a day to day basis. So paragraph 3 seems to introduce an obligation that is potentially unrealistic or at least very difficult to fulfil.

Moreover, it is important to stress out that in order to stimulate an economy based on the value of data, the EU has been developing several initiatives aimed at increasing the sovereignty of Europe in terms of data, such as, encouraging the reuse of public administration data through the directive on open data and the reuse of public-sector information[2], the development of a framework for the free flow of non-personal data in the European Union[3], or the publication of the European Strategy for Data[4] aimed the developing of an European data market. It is important to understand to what extent this proposed AI Act in the New Legislative Framework (NLF) does not diverge from other European data programs.

Finally, we consider an outcome-based approach to requirements is more likely to achieve our goals. In our view, this approach will promote more the EU values of fairness and non-discrimination by clearly articulating the outcomes that regulated actors should strive to achieve. For instance, the requirements could provide that: high-risk AI systems should provide a similar quality of service for relevant demographic groups impacted by the system; high-risk AI systems that allocate resources or opportunities should do so in a manner that minimizes disparities in outcomes for relevant demographic groups impacted by the system; and high-risk AI systems that describe, depict or otherwise represent people, cultures, or society should

[2] Directive (EU) 2019/1024.
[3] Regulation (EU) 2018/1807.
[4] COM (2020) 66

minimize the potential for stereotyping, demeaning, or erasing relevant demographic groups impacted by the system.

To achieve these goals it should be necessary for the relevant actors to: identify the relevant demographic groups impacted by the system; designing and undertaking an evaluation to assess the extent to which the goal is achieved; reassessing the system design (including the training data, model features, objective functions, etc); re-evaluating the system after incorporating appropriate mitigations and communicating material residual risks to deployers so that appropriate precautions can be taken, including decisions to not use certain systems in particular use cases if they are not fit for purpose. In each case, achievement of the desired outcomes and execution of the procedural steps would need to be assessed by reference to the state of the art and industry best practices, along with a clear-eyed recognition of the highly contextual nature of fairness and the fact that it is never possible to fully “de-bias” an AI system or “guarantee” its fairness. We also suggest that the representativeness of the samples should be guaranteed through appropriate statistical techniques.

In short, we believe that some improvements could be made in the wording of this article. In effect, similarly to other precepts that are part of the proposal, and due to the ambiguity caused by the use of vague concepts, we can also find some loopholes that may work in favour of attributing an inadvisable discretion to the AI system suppliers.

This is particularly the case with the repeated use of the term "appropriate". Indeed, Article 10 requires that " training, validation and test data sets shall be subject to 'appropriate' data governance and management practices" (paragraphs 2 and 6); that the data sets have "appropriate" statistical properties (paragraph 3) and also that to the extent that it is strictly necessary for the purposes of ensuring bias monitoring, detection and correction in relation to the high-risk AI systems, the providers of such systems may process special categories of personal data subject to “appropriate” safeguards for the fundamental rights and freedoms of natural persons (paragraph 5).

Building on this, and while the text can be commended for specifying the minimum considerations that must be taken into account for the data management process to be considered "adequate," it leaves open, for example, the specifics of what, concretely, constitutes an "appropriate "

statistical property: does this require that the data be a representative sample of the entire population, or only of the potential ad hoc groups that may be subject to AI system analysis?

Article 13 - Transparency and provision of information to users

We believe that the proposal article could benefit from some clarification on this point.

In our view, creating an express obligation to the user and the provider of the systems contemplated by this article, to communicate the information listed therein to the persons subject to its use, means also the creation of a corresponding right of those persons to demand this information[5]. And this is because, as we have seen, the mere communication to an individual that he or she has been subjected to the use of a risky AI system, offers little or no protection if that individual is not recognized as having the right to obtain information about that system and is not given access to a specific procedure that he can use it to challenge before an independent authority [6] an adverse result that he believes is the result of such use.

We believe, therefore, that the proposal would benefit from a strengthening of guarantees in the sense of expressly providing for each of these elements. This would mean, moreover, an undoubtedly stronger commitment to empowering citizens to adequately exercise and safeguard their fundamental rights, giving them effective weapons to, on their own initiative, syndicate the concrete use of these systems and seek accountability for any adverse impacts resulting therefrom.

In addition, we would like to draw the attention to the fact that, as in several provisions of Title III, the use in Article 13(1) of expressions such as "sufficiently transparent" or "appropriate type of degree of transparency" seems likely to allow those who make them available, a high degree of discretion in (self)evaluating the level of transparency of their own systems.

[5] Again, we note that the article addresses the provision of information to users, seeming to abstract from the position of the people who are subject to (or affected by) such use...

[6] This proposal could even be articulated with the individual complaint mechanism that we also referred to in point 3.2.

Finally, we would add that, considering the list of information specified in article 13.3, one gets the impression that test-generated results - for example in terms of performance and accuracy - will be sufficient to satisfy the requirements that the standard aims to cover.

However, as experience has proven, these results can be substantially different when the system is tested in a real environment. We would therefore suggest adding to the list a requirement to provide information about the real conditions of normal use of the system and about the parameters topic could also include a reference to the Responsible AI.

Article 14 – Human oversight

We agree that it is extremely important and necessary to have human oversight in some types of AI systems.

Some examples include AI systems that are embedded as other functionality and is not easy for the end-users to detect that an AI system is currently working behind the scene, or as another example, the system is an autonomous solution. For this kind of systems, the human oversight is almost impossible to have during the period of use, and so, robust operational tests are more suitable to safely drive the acceptance of the system than having human oversight. We suggest the revision of the concept of human oversight, or the revision of this article text in order to address different scenarios where the human oversight is necessary. Another example includes the 'automation bias' focused on number 4, paragraph b) as a way to substitute the human workforce, mainly when the solution is a decision support system. If the system is a real high-risk system, then the "conditions of use" may be more important than the "context". Note that these concepts are defined in the "intended purpose" notion (article 3, (12)). The conditions of use may be forced to deployments forms where the decision is taken by the humans and recorded in the system as a first step, and then, as a second step the human decision may be used to confront with the decision of the AI system. This should define an augmented deployment environment, naturally accepted by the end users and reducing the risk of misuse. For instance, if a system produces diagnosis of X-Ray exams, this system could be used for trial and the prioritization of patients in an emergency room, and this is accepted as value-added substitute task even if the system has an accuracy of just 93%. The same system is now used for the final diagnosis that should be signed by the radiologist. In this case, the radiologist decides the diagnosis as first step

and recorded it on the system, as a second step the AI system produce a diagnosis and presents it to the radiologist, he/she verifies if there is a match between the two diagnoses, augmenting the opportunity of correcting a wrong decision. In this scenario the system should never be used as first and unique solution for diagnosis decision. This kind of deployment scenarios should be included in the conditions of use of the intended purpose since, if they are well described some of the concerns about human oversight should be reduced.

Additionally, we highlight that in many cases, it is very difficult, in practice, to ensure that this human guarantee works as intended. This is because the premise underlying the operation of most of these systems is based on algorithms that are beyond the capacity of human perception, meaning that the human being who is responsible for supervising the system will not, in most cases, have the ability to know all the variables inherent in its operation and, therefore, be able to properly perform its supervisory function (except, of course, in cases where human intuition can detect obvious flaws or aberrations).

For the same reason, we believe that the standard as drafted is not suitable to adequately solve the problem of "algorithmic bias". Moreover, despite the remarkable efforts undertaken in the construction of the standard, it remains insufficiently clear whether the human oversight measures positively stated in this article apply to the user or to someone independent of the user, or even whether user refers to the organization using the AI system as a whole or to a specific individual who is responsible for a particular decision[7]. If we see it right, oversight is necessary for all actions related to the development, implementation and use of AI systems, to ensure that fundamental rights are protected in the best possible way at each and every stage. This will include, human oversight of the process, but also regular and independent human oversight of the very people who participate in it and who are ultimately responsible for making the final decision, informed by the outputs produced by the system. It is not enough, therefore, to know whether supervisors are properly aware of the possibility of biases, but it must also be possible to demonstrate, in a transparent and effective way, that the actual decisions were not taken on the basis of excessive confidence in the results produced by the system.

[7] See the text in Recital 38.

Therefore, we recommend defining the concept of “effective human oversight” and the specific results this article intends to seek. In our view, “human oversight” differs depending on the deployment scenario and the nature of the related risks. Consequently, we recommend that the proposed Regulation requires deployers to implement sufficient, qualified human oversight as is appropriate to the deployment scenario at issue.

It is also important to bear in mind that for the “human oversight” to be meaningful and successful it is necessary to ensure that the humans performing the oversight are trained and equipped appropriately in accordance with the instructions of use and other information provided by the supplier. Additionally, the oversight should be tied to the intended use of the AI system and accountability mechanisms should be created to assess the effectiveness of the human overseer.

In this sense, we believe that a third category should be added to Article 14(3) adequately recognizing the need for users to implement organizational measures to ensure robust human oversight, consisting of at least: training for decision-makers, registration requirements, and clear ex-post review processes.

(5)

This paragraph, which imposes a requirement for enhanced supervision when biometric identification systems are used, could benefit from some additional clarification.

According to the wording proposed for this standard, the system user cannot take any action or take any decision based on the identification resulting from the system, unless the result has been verified and confirmed by at least two natural persons.

In our opinion, the requirement of two-person review of biometric identification systems should be limited to specified, uniquely high-risk scenarios. Otherwise, this requirement will be disproportionate and counterproductive. For example, it will imply that for a system used in a restaurant to ensure the staff washes their hands several times during the day it was necessary to have a two-person verification system, which is disproportionate.

On the other hand, and to ensure any useful effect of this standard for uniquely high-risk scenarios, the confirmation carried out by the "two natural persons" should be based on a separate assessment by each of them.

Finally, we would say that reliance on human supervision as a sufficient safeguard should only be considered when it is possible to prove that the use of intrusive systems is necessary and proportionate in a democratic society, preventing it from functioning to legitimize the use of technologies that should not be used in light of their potential to violate fundamental rights. We would therefore reiterate that human supervision cannot act as a panacea for the (very serious) problems that the use of certain systems can give rise to, and consequently cannot be used to validate and – by that way, legitimize – that system or its use in a given context.

Article 15 - Accuracy, robustness and cybersecurity

(1)

According to the proposed Regulation the “high-risk AI systems shall be designed and developed in such a way that the achieve, in the light of their intended purpose, an appropriate level of accuracy, robustness and cybersecurity, and perform consistently in those respects throughout their lifecycle.” However, it is not defined what it means “an appropriate level of accuracy, robustness and cybersecurity”, is 90% accuracy appropriate or 70%? And does the appropriate level changes depending on the context? If it is a critical infrastructure or a system used in a factory? We strongly recommend the development of best practices and standards do define these and other concepts.

(3)

We recommend reviewing the first paragraph of this article, taking into account the definition of “intended purpose” set in article 3, number 12 of the proposed Regulation. Please note that it is not possible to ensure that AI systems are 100% resilient to errors, faults, or inconsistencies. The uncertainty is part of the AI system.

We also suggest adding the importance of quantifying this uncertainty, given its high impact on other topics addressed in this proposal, such as the risk of AI systems.

Article 52 - Transparency obligations for certain AI systems

We would like to signal a precision that is linked to article 5, paragraph 1, a), which refers “the following artificial intelligence practices shall be prohibited: 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”.

It is possible to envisage the existence of systems of this type that do not give rise to these kinds of damages and that are, consequently, exempt from this prohibition. This will be, for example, the case of subliminal manipulations to aid motivation that are consciously chosen by the respective user with their full and informed consent. However, in order to consider the existence, in each case, of a consent of this type, it will be necessary to ensure that the properties of the subliminal manipulative techniques are adequately known by the person being manipulated, with sufficient transparency. With this in mind, we recommend that the subliminal techniques to which the "manipulated" have given their free and informed consent and which do not cause damage or interference with fundamental rights (even excluded from the prohibition of article 5), should be subject to the obligations of transparency included in Title IV (Transparency Obligations for Certain AI Systems), and should be included in the text of this article.

Article 69 - Codes of conduct

We suggest adding a number 5 to this proposed article with the following text or a similar one:
“5 – The Commission and the Board recommend including in these codes of conduct concepts such as: privacy; accountability; safety and security; transparency, explainability and interpretability; fairness and non-discrimination; human control of technology; professional responsibility; promotion of human values.”.

Final remarks:

As referred above, PT welcomes this initiative to create a legislative proposal for a Regulation laying down harmonized rules on artificial intelligence.

However, we would like to highlight some concerns in addition to the ones referred above. According to the reports published by AI Watch, Europe is still not a competitive market, especially considering China and US markets.

Simultaneously, the EU is encouraging the development of the entrepreneurship ecosystem and has demonstrated a constant concern with necessity to stimulate the SME market. We need to ensure that the rules set in the proposed regulation are not too burdensome to guarantee the attraction of the entrepreneurship ecosystem and the SME market. The idea that a regulated market can better define investment limits and act as an attraction to the innovation is interesting and we fully support it. However, if the legislative package does not guarantee the necessary agility for the creation of innovation ecosystems, then only large companies can reap benefits, and the market will certainly be more fragmented, not only because of cultural differences and potential investment capacity between Member States, but also because of the differences in dimension of the actors involved in the development and diffusion of artificial intelligence technology.

We also believe it is necessary to create a test group to validate some of the procedures and obligations that the AI operators will be subject to, in order to “place a product on the market”/ “put into service”. This test group would help us to understand how complex these rules are, and what technology diffusion scenarios can be improved, to ensure that essentially high-risk systems can be approached with less risk, without having to put into practice a process that will significantly delay the widespread adoption of AI and the benefits in the European space for the economic development.

PUBLIC