adigital

ADIGITAL CONTRIBUTION

European Commission Public Consultation on the White Paper on Artificial Intelligence

On 19 February 2020, the European Commission published its White Paper On Artificial Intelligence – A European approach to excellence and trust. Divided into two sections, Ecosystem of Excellence and Ecosystem of Trust, the White Paper proposes a two-tiered regulatory framework for AI: (i) a mandatory conformity assessment regime for "high-risk" AI; and (ii) a voluntary labelling scheme for all other AI.

Adigital very much welcomes the proposals set out in the Commission's White Paper, and commends the Commission both for taking the initiative, and for its thoughtful approach.

Adigital and its associates are committed to engaging constructively in the multi-stakeholder discussions about AI governance, including on how to create a flexible framework for trustworthy AI in Europe.

The following document presents Adigital suggestions to the European Commission White Paper in the light of the public consultation process. Although this paper contains some general remarks on the proposals for an Ecosystem of Excellence, the core of the document focuses on a series of appreciations and proposals for the Ecosystem of Trust which could set the basis for a European regulatory framework.

Adigital

Gran Vía, 4 Edificio Utopicus
28013 Madrid
+34 (0) 91 598 11 57

adigital.org

Carrer d'Entença, 218
08029 Barcelona
+34 (0) 93 240 40 70

1. General remarks on the Ecosystem of excellence

With a thriving developer community, and world-class universities, Europe is well positioned to play a leading role in AI research and application. For this reason, **Adigital considers that the priorities for the Commission should be**:

- **supporting the research and innovation community**. In particular, the proposal to create a lighthouse centre for AI research, innovation and expertise in Europe is an excellent initiative
- **ensuring the right skills are in place** so all are able to prosper from the AI benefits
- **supporting an ecosystem of European companies and third-country organisations** who live by European values, so that, especially SMEs can take advantage of AI
- promoting the **adoption of AI by the public sector** and assisting the public sector for this purpose

2. General remarks on the Ecosystem of Trust

Adigital agrees **with the need to boost public trust in AI, as in any other technology, but emphasises that this needs to be done in a proportionate way that balances potential harms with benefits**. It is essential that a sense of mistrust on AI is not created as it will probably lead to overregulation, harm AI adoption and divert vital resources from innovation/deployment to compliance. Furthermore, trust cannot be built by regulation alone. We must amplify ongoing efforts to establish industry best practices in the field of responsible AI development, rather than risk cutting them short by prematurely codifying overly specific and inflexible mandates in this fast-changing field.

We strongly believe that the EU regulatory framework already provides strong protections and trust. In fact, the EU's existing legal framework already applies to AI applications, e.g. fundamental rights protections, GDPR, product safety, consumer protection regulations, the P2B regulation on fairness and transparency. However, there may be a need for assessment and for clarification or guidelines, particularly for high risk use cases. However, new legislation should only be considered for specific high-risk cases where it is determined that existing frameworks cannot be adequately clarified or adapted. Any new legislative proposals on AI or in related areas (such as the draft e-Privacy regulation), should be consistent with existing legal frameworks to avoid diverging rules and legal uncertainty, which could have a negative effect

on the uptake of AI within the EU. They should also clearly specify what additional risks any new legislation is seeking to address. Therefore, **Adigital encourages authorities to focus their efforts on clarifying how to observe existing requirements on topics such as unfair discrimination, explainability or interpretability.**

Smart government approaches to regulation will play an important role in building trust and ensuring that AI is used responsibly, while also encouraging innovation. It is important that a proportionate, risk-based approach is taken, balancing potential harms with the social and economic benefits that will be created by AI. Any regulatory framework should be flexible enough to evolve with this dynamic technology space.

2.1. Specific remarks on the Ecosystem of Trust

2.1.1. Positive highlights:

- The **risk-based approach** to regulation, as long as it is accompanied by a proportionate, targeted approach that brings legal certainty. The scope of any regulatory obligations should be a function of the degree of risk and the potential scope and severity of harm. Many AI systems pose extremely low, or even no, risk to individuals or society. Mandatory obligations shouldn't target things like AI used to organize products in a warehouse, or to correct spelling in a document.
- The i**ncremental approach** by limiting the regulation to AI systems or use cases that pose a material risk of individual or societal harm. That includes in particular systems that might harm people or impair their legal or other fundamental rights, such as their ability to access healthcare, or employment opportunities. Many public sector deployments should also sit in this list.
- The **voluntary labelling scheme** for non-high-risk AI. Many companies are working hard to adopt policies and procedures to promote trustworthy AI and such a scheme would act as a benchmark for the development of internal procedures and could become a global reference.

2.1.2. Aspects to be improved

➔ **The definition of AI**

Although the White Paper refers to various earlier definitions of AI, it doesn't expressly endorse any particular definition. This will be an important element of the proposed regulation, and one that will require some further careful thought.

Adigital agrees with the Commission's view expressed in the White Paper that "**the definition of AI will need to be sufficiently flexible to accommodate technical progress while being precise enough to provide the necessary legal certainty**". Given the rapid developments in the field and its wide application, an effective definition should focus on the features that distinguish AI systems from other technologies.

In this regard, it will be important to remember that the term "artificial intelligence" can be, and often is, used to describe a vast array of technologies. These technologies, in turn, can be used in a nearly infinite range of scenarios. Consistent with the Commission's goal of carefully focusing regulatory mandates on high-risk scenarios, **it will be important to ensure that the definition of AI is not so broad** as to sweep in thousands of everyday products and services.

At the same time, it will be important to ensure that the final AI regulation is consistent with the principle of technology neutrality. Although the White Paper does an excellent job of identifying aspects of AI that might give rise to unique risks, it will likely be the case that some, and perhaps many, products or services with an AI component will be no more risky than their non-AI counterparts—and possibly less risky. Where that's the case, EU law should not subject the AI-enabled product to greater regulatory burdens than the non-AI counterpart. Otherwise, the regulation could in fact deter companies from offering AI products and services in Europe. That result would seem contrary to the Commission's broader goals of spurring the uptake of AI.

The definition could be to some extent based on the proposal made by the High-Level Expert Group on Artificial Intelligence. It would need to be adapted in order to have a narrower focus that excludes traditional rule-based systems and focuses just on systems embodying complex machine learning techniques.

➔ **Allocation of responsibilities**

The White Paper points out that the supply chain for AI products can be long and can include many actors, from the point of development of an AI system to its deployment. This point is extraordinarily important when thinking about regulatory obligations. The healthcare sector illustrates well this. There are companies that design and develop AI solutions that are sometimes incorporated later by customers into a medical device, which that customer then sells to a hospital, where it's used by a doctor in providing patient care. There are shared, but different, responsibilities throughout this chain.

Adigital strongly agrees with the Commission that **regulatory obligations should be addressed to the actor in the chain that is best placed to address the potential risk**

involved. An example of this could be the potential bias. Developers of AI tools often will be best placed to address the risk of unfair bias when training an AI algorithm or system. But deployers will often be in the best position to ensure that AI systems aren't used in ways that unfairly discriminate. Similarly, while developers may hold important information about the datasets used to train AI systems, only the deployer may have the ability to monitor the system's operation and impacts once it has been released "in the wild." The same holds true for transparency. Developers are better placed to describe the capabilities and limitations of an AI system, while disclosing the fact of AI use to people likely to be affected by it will typically need to be the responsibility of the deployer.

➔ **Risk-based approach and efficiency of market conformity assessment**

As expressed above, Adigital is in general supportive of a risk-based approach. However, some adjustments are needed to ensure regulation is proportionate, targeted appropriately, and provides legal certainty. Some suggestions in this direction are:

- **Consideration of the opportunity cost of not using AI:** It's vital that any risk assessment takes a holistic view, reflecting not only potential harms but also societal opportunities. The benefits will often outweigh the risks, especially if risks can be mitigated in a thoughtful way with strong safeguards. Regulation must not discourage use in such cases.
- **Proportionality**: Risk assessments must reflect the probability of harm and not just the severity of the harm. They should also take account of the wider operational context when assessing risk, since the same AI application used for the same purpose will pose different risks depending on the way it is integrated into business operations (e.g., extent of human oversight, additional safeguards such as monitoring).
- **Ensure a level-playing field**: The cumulative criteria (combining a list of sectors and clarity over what constitutes high risk use within them) is, broadly, a workable approach. However, the borders between sectors are blurring up, especially in a digital context. Therefore, some activities can be performed by firms belonging to different sectors and, then, subject to different regulatory regimes. In such cases, an activity-based rather than a sector-based approach should be taken in order to prevent regulatory arbitrage. At the same time, focusing only on use cases would results in a highly complex assessment process. Combining both criteria - sector and use cases- could lead to a more targeted and efficient approach.
- **Avoid creating legal uncertainty**. Criteria to identify potential high-risk per se should be clear and objective in order to minimize the discretionality of the classification.

- **Review periods for the risk assessment process need to be proportionate, striking the right balance between stability and a dynamic digital market**. The reclassification as high risk should be based on applications' impact on human lives having changed, or on increased awareness on applications impact, rather than on criteria changes or the assessment process. The five layered risk grading of the German Data Ethics Commission's report would reduce uncertainty by avoiding an all or nothing approach (high -risk or non-high -risk); in the case of a modification in the risk assessment process, applications two levels below the high risk threshold would for example less likely be reclassified as high risk, and thus becoming subject to ex-ante obligations, than applications in the level just below the threshold
- **Remove reference to "immaterial damages" in the risk definition**: This is a vague term, not a legal concept, and the spirit behind it is already covered by other laws (e.g., data protection, non-discrimination, freedom of expression). Far better would be to align it with the phrasing of the Product Liability Directive, which defines damage as death, personal injury or damage to property.

With regards to the mandatory **pre-marketing conformity assessment** to ensure that high-risk AI systems used in the EU meet certain legislative requirements, some observations need to be done. **Some companies have significant experience with these types of assessments in other sectors**, such as cloud computing and computer hardware. Moreover, the EU maintains a range of conformity assessments for products being placed on the market in the EU.

There are lessons to be drawn from these experiences and existing regimes, both with regard to best practices and things to avoid. The European AI strategy should enable a level playing field with other regions of the world for the development of AI applications. Any conformity assessment should be quick and light touch, so consumers in Europe have access to AI innovations at the same time that consumers outside Europe. At the same time, it should not be unnecessarily difficult for organizations to get those innovations into the market.

Cost should also be a consideration. Although a big company can absorb the costs of a conformity assessment, they are likely to be burdensome, and possibly prohibitive, for smaller companies. **Many small or mid-sized companies might not have the resources to cover expensive, and repeated, conformity assessments.**

As a general rule, self-attestation should be the first option. It's most scalable and least likely to unnecessarily extend time to market or unduly burden smaller operators. True third-party,

premarketing conformity assessment should be limited only to the most high-risk scenarios, where the risks to health, safety, or individual rights are at their peak.

The Commission recognized the costs and inflexibility of a prior conformity assessment approach when, in GDPR, consciously abandoned the *ex ante* external "prior checking" system existing under Directive 1995/46/EC. Through GDPR, the Commission instead established the duty to implement accountable data protection programmes that include Data Protection Impact Assessments (DPIAs): ex ante self-assessments for data processing likely to be high risk.

A similar approach to AI would be a more balanced alternative to requiring blanket prior reviews by a regulator of all "high-risk" AI applications as the White Paper recommends.

Any **ex-ante conformity should be carefully designed** to avoid creating a burden for the development of AI applications in Europe versus other regions of the world. Especially if an ex-ante regime would imply regulatory intervention by external entities (be it regulatory or other) during the development process, this would severely hamper the time to market of AI developed in Europe versus other regions in the world.

There should also be clarity about the assessment of existing applications, which are already on the market.

➔ **Mandatory requirements for high risk AI applications:**

On a general note, **Adigital invites the Commission to collaborate with practitioners in drafting specific language to ensure it is workable. As written, some proposals could present serious practical challenges and conflict with existing laws.**

Adigital agrees with the Commission's view expressed in the White Paper that "A risk-based approach is important to help ensure that the regulatory intervention is proportionate", and in particular that "it requires clear criteria to differentiate between the different AI applications, in particular in relation to the question whether or not they are 'high-risk'."

We would suggest that the key factors to consider in assessing risk should be the degree to which human judgment is replaced and the potential negative impact of the application on human lives (severity). In the White Paper**, the Commission's two-criteria approach is complicated** by the proposal that the use of AI for certain purposes would be always considered as high-risk. Examples of such exceptional purposes are given: the use of AI applications for recruitment processes; use in situations impacting workers' rights; and the use of AI for the purposes of remote biometric identification. This is problematic, in particular

because these example exceptions are defined in a very open-ended way, making the scope broad and unpredictable.

It is important that as this risk assessment framework evolves there is strong and practical guidance provided for companies that there can be clarity, consistency and transparency about how applications are deemed to be low-risk or high-risk.

We believe that given the combined criteria of autonomy and severity should suffice to cover all high-risk systems. This approach will remove the need for exceptions.

With regards to specific obligatory requirements made in the White Paper, Adigital makes the following recommendations:

- **Keeping of datasets should not be mandated:**
    - Keeping datasets is **likely to conflict with GDPR** provisions to delete personal data, as well as presenting challenges for copyrighted datasets authorised for only short-term access. The records, documentation and, where relevant, data sets would need to be retained during a limited, reasonable time period to ensure effective enforcement of the relevant legislation. Measures should be taken to ensure that they are made available upon request, in particular for testing or inspection by competent authorities. Where necessary, arrangements should be made to ensure that confidential information, such as trade secrets, is protected".
    - It **would destroy the privacy benefits of on-device processing** because it would effectively force data to be collected and stored centrally - which is precisely what on-device processing seeks to avoid doing.
    - It would **prevent the use of off the shelf, open-source models**, since developers will generally have no access to the data used to train them.
- **Place emphasis on training data quality as well as on testing output:**
    - The **impact of the quality of training data on AI systems will depend on the use case**. For example, in some cases, the focus should be on the quality of the training data (including appropriate diversity, lack of bias etc.), and the outcomes produced by the system, rather than simply the geographical source of training data.

- In some others, **the proposed obligations for developers to "ensure datasets are sufficiently representative" may be impractical**. They could conflict with GDPR under which developers are not meant to have access to sensitive attributes like ethnicity. It is also unclear how to determine what is "sufficient" — especially for providers of multipurpose AI systems.

- Finally, **there might be scenarios where, with enough expertise and care, it's possible to create a high performing model** even using biased, lower quality training data; the reverse is also true.

- Thus, we must aim to models that take into account the need for quality in the training data, the relevance of good testing models using benchmark datasets, are enough to make sure that the outputs are within an acceptable range,

- Pay special **attention to literal translation of the requirements into legislation**:

  - There is **no "one size fits all" approach to ensuring safety, robustness an accuracy of AI systems**. It is important to retain flexibility in the legal interpretation and to collaborate with practitioners to draft rules that are workable from a technical perspective.
  - The White Paper proposes "requirements ensuring that outcomes are reproducible". A too literal interpretation of reproducibility would be impossible to satisfy, as many AI systems have randomness built in, which makes it impossible to guarantee you get the identical output every time even if the input is the same. To be workable, there will need to be scope for alternative definitions of reproducibility that do not require exact matching.

➔ **Safety and liability**

The White Paper also identifies several potential challenges that might arise when we try to apply the existing EU product safety and liability regime to AI. Engaging in this kind of prospective thinking is important. Addressing safety and liability issues early will be key to fostering consumer trust in these new technologies.

At the same time, legislating solutions without being able to point to existing, specific problems is always a challenge. At this stage, we've seen few actual cases where people were harmed by "defective" AI, or where people sought, but couldn't obtain, a remedy. To

the extent we are seeing claims of harm, many of these are arising in the data protection context, where the GDPR already provides a comprehensive regulatory regime.

While we shouldn't be afraid to reform safety or liability rules, **we probably need more data to better pinpoint the problems that need fixing.** The Commission has done a great job in helping identify places where we might look for potential problems. Now we need to assess whether these *potential* problems are *actual* problems.

In general, **Adigital agrees with a "phased approach" for liability** if this means that liability is allocated according to the role of different actors in the life cycle of an AI system. As the White Paper says "while the developers of AI may be best placed to address risks arising from the development phase, their ability to control risks during the use phase may be more limited. In that case, the deployer should be subject to the relevant obligation."

The use of AI systems, and therefore any resulting liability, is context-specific: **the focus of risk should lie on a specific application and the context of its use**. What matters is that consumers have a clear reference point to obtain compensation. There is often a complex chain of various producers and intermediaries involved in any consumer-facing product. We make a difference between, for example, the various producers (software developers, hardware manufacturers, e.g. component manufacturers, or end-product manufactures who embed the software in their products), and back-end operators (who train the AI system) and the front-end operator who is using the system.

This is why having more than a single operator who is liable or introducing joint liability would not be workable, and it would not make sense for anyone involved in making an AI system to be liable for problems they had no awareness of or influence over.

More concretely, in considering liability for a flaw, where compensation is being sought, **there are different approaches**. In the case of a "strict liability" regime (i.e. where no fault needs to be proven by the claimant to obtain compensation), there is no need for explainability/ transparency. However, in case of a "fault-based liability" regime, where the claimant must prove a fault in the product, they would likely need some level of transparency about how the system was developed/trained/used and some level of explanation about how it reached a flawed decision or outcome. In that context, solutions exist to address this. Some companies have created tools to help open up the "AI black box" and provide more transparency to users, deployers and operators.

Given the existing Product Liability Directive and the General Product Safety Directive apply to consumer goods, it would be consistent to apply a new framework or guidance to

consumer goods. Any other criterion would result in uncertainty on the interplay between the existing and new frameworks. In a B2B context, parties can negotiate for a more efficient allocation of risk which consider each specific context and use. Contractual liability is working well and should therefore be maintained.

Only changes that result in the functionality of a product altering in such a way as to impact safety testing or disclosures should be classed as an "important change" requiring a new risk assessment. Generic updates such as security and bug fixes, or simple improvements, should not be included.

No matter how complex the value chain is, the ultimate responsibility for the safety of a product should be with the entity that puts it on the market. This gives the greatest clarity on who people need to seek redress from if there is a problem. Any alternative (e.g.: "joint responsibility") risks making it more complicated and could reduce overall safety, because bad actors will use it as an excuse to duck their responsibilities (on the basis they think they can escape being targeted if something goes wrong).

Strict liability should not be introduced for AI systems because it would mean that anyone involved in making an AI system could be liable for problems, they had no awareness or influence over. Such a regime would have a chilling effect on innovation and undermine the uptake of AI by businesses in Europe.

Software vulnerabilities should not be treated the same as product defects from a legal perspective because of their very different natures. Once launched, physical products are set in stone (unless there is a major recall), however patches to protect software can be rapidly released. There is also greater onus on software users to be responsible in applying the patches offered.

Adigital is opposed to the idea of expanding the definition of "product" in the Product Liability Directive to include software. We believe this changes the scope of the PLD, where standalone software is not covered (at least not in all Member States). Also, it is hard to see in what cases standalone software would result in property damage, bodily injury or death. Usually, this will be caused by hardware or by the use of the software by the human being. However, it would be helpful to have EU guidance to clarify this, and to ensure a consistent approach across Member States.

➔ **Voluntary labelling scheme**

As stated above, **a voluntary labelling system** could act as a benchmark for the development of internal procedures and as a global reference.

However, if finally developed, this labelling system should be oriented to the certification of applications and not firms as a whole and requirements should be proportionate to the actual risk posed by the use case.

➔ **Enforcement and governance**

In general, the ex-ante conformity assessment process seems to be impractical. Adigital makes the following recommendations to address enforcement:

- **Change to ex-post enforcement, supplemented by upfront self-assessment**: There risks being a huge backlog if upfront third-party testing were required prior to launch, and the delays and red tape would deter companies from developing and launching new services in Europe or importing products from third countries. A more balanced approach is to make the expectations clear for risk assessment processes, and allow for self-checking prior to launch, with ex-post investigations carried out only where problems are suspected.

- **Drop the suggestion of retraining AI systems in the EU as a means of mitigating non-compliance**: There is no guarantee that European datasets or training that takes place on European soil will do anything to improve the performance of an AI system. If anything, blocking the use of foundational (non-European) data sets would risk reducing a system's performance, and could even exacerbate the risk of discrimination.
- **Oversight mechanisms and governance structures** should rely on existing agencies as much as possible and be carefully designed to avoid potential fragmentation and overlaps.

About ADIGITAL

Adigital is the Spanish Digital Economy Association. Formed by more than 500 associates, it aims to promote and support the digital economy in Spain through the development of information society services, ecommerce, digital marketing and communication, digital content, mobile applications and other related activities.

Adigital

Gran Vía, 4 Edificio Utopicus
28013 Madrid
+34 (0) 91 598 11 57

adigital.org

Carrer d'Entença, 218
08029 Barcelona
+34 (0) 93 240 40 70