EU White Paper

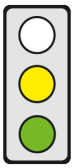
WHITE PAPER ON ARTIFICIAL INTELLIGENCE

cep**PolicyBrief** No. 2020-4

KEY ISSUES

Objective of the White Paper: The Commission proposes binding requirements for high-risk artificial intelligence (AI) applications which must be followed by developers and users of AI across the EU.

Affected parties: AI developers and deployers, companies and individuals who use or are affected by AI.



Pro: (1) The proposed risk-based approach accounts for the fact that the possible consequences for users differ depending on the specific AI application and the sector in which it is deployed.

(2) Mandatory legal requirements for high-risk AI applications and voluntary requirements for low-risk ones respect the principle of proportionality.

Contra: The definition of "high-risk" is not precise enough to create legal certainty. It must be clarified when the risks involved with the use of an AI application have to be considered "significant".

The most important passages in the text are indicated by a line in the margin.

CONTENT

Title

White Paper COM(2020) 65 of 19 February 2020: **On Artificial Intelligence – A European approach to excellence and trust**

Brief Summary

- **General Background**
  - Artificial Intelligence ["AI"] is a set of technologies combining data, algorithms, and computing power [p. 2].
  - The purpose of the White Paper is to "set out policy options" on how to promote the uptake of AI while at the same time addressing the risks associated with AI. To achieve this purpose, the Commission plans to develop in the EU
    - an "ecosystem of excellence" for AI, and
    - an "ecosystem of trust" in AI; to this end the Commission lays out the risks associated with AI, and its plan to address these risks. In this regard, the Commission plans to propose
      - adjustments to existing legislation, and
      - new AI-specific legislation, in particular additional requirements for high-risk AI applications, and a voluntary labelling scheme for low-risk AI applications.

- **Measures to develop an "ecosystem of excellence" for AI**
  - The Commission proposes i.a. the following measures:
    - Extension of the Coordinated Plan on AI [cf. cep**PolicyBriefs** Nos. 2019-10, 2019-12 and 2019-13] by end 2020 [p. 5], in order to ensure i.a. that at least one digital innovation hub per Member State has a high degree of specialisation on AI [p. 7].
    - "Facilitation" of the creation of excellence and testing centres for AI, financed by the EU, Member States, and the private sector, possibly introducing "a new legal instrument" [p. 6].
    - Launch, together with the European Investment Fund, of a pilot fund of €100 million to finance innovative developments in AI [p. 7] [cf. cep**PolicyBrief** No. 2019-10].
    - Presentation of an action plan to facilitate development of, experimentation with, and adoption of AI by public authorities and providers of services of public interest, e.g. utility and transport services. The Commission will first hold a dialogue with stakeholders [p. 8].
    - "Support" for access to and management of data as well as cloud infrastructure, high-performance and quantum computing because the development of AI applications requires both access to data as well as key computing technologies and infrastructures [p. 8].

- **Risks of AI applications**
  - The use of AI applications increases the risks for fundamental rights (e.g. non-discrimination), safety and the effective functioning of the liability regime. This is due i.a. to
    - biases present in an AI application, which could have larger effects than biases in human decision-making because AI lacks the social control mechanisms that govern human behaviour,

Authors: Alessandro Gasparotti, Lukas Harta and Anja Hoffmann| harta@cep.eu
cep | Kaiser-Joseph-Straße 266 | 79098 Freiburg | Germany | Telephone + 49 761 38693-244 | www.cep.eu

- the lack of transparency (opacity) of many AI applications that complicates verifying compliance with legal standards because it is difficult or even impossible to trace back how a decision was taken [p. 11-13]; opacity therefore increases the difficulties of establishing
  - liability if damage is allegedly caused by a defect in an AI-equipped product [p. 13], and
  - a breach of the applicable EU and national safety law [p. 12-13].

► **Possible adjustments to existing legislation applying to AI**

– While EU legislation – like the Product Safety Directive and the Product Liability Directive – in principle also applies to AI, the Commission identifies five shortcomings which should be addressed [p. 14]:
  - Existing EU legislation focuses on safety risks present at the time of placing on the market. However, an AI application can modify its functioning and cause new risks, notably if it relies on machine learning or requires frequent updates.
  - Due to the opacity of AI, it might be necessary to adjust or clarify existing safety and liability legislation in order to ensure its effective application.
  - Existing EU safety legislation is limited in scope because it only applies to products. Therefore,
    - AI-based services are not covered by EU safety legislation, and
    - it is unclear whether stand-alone software constitutes a "product" under EU product safety legislation.
  - Existing EU legislation does not explicitly address other new risks, such as risks that result from cyber threats or loss of connectivity, regardless of whether these risks were present at the time of placing on the market or arise while using the product.
  - The allocation of responsibilities between different economic operators is unclear. EU product liability law only regulates the liability of producers, while national law covers the liability of others in the supply chain. Liability is unclear e.g. when AI is added to a product by a third party, after the product was placed on the market.

► **New AI-specific legislation**

– In addition to updating existing EU law, the Commission is also considering new AI-specific legislation.

– This will require a clear definition of AI, which must be both sufficiently flexible to accommodate technical progress and precise enough to provide the necessary legal certainty [p. 16].

– In order to ensure proportionality, the Commission proposes a risk-based approach: AI applications should be categorized as "high-risk" or "low-risk", subjecting the former to stricter additional requirements [p. 17].

– An AI application is defined as high-risk if it fulfils the following two cumulative criteria [p. 17-18]:
  - it is used in a sector that is included in an exhaustive list of sectors in which significant risks can typically be expected to occur, e.g. healthcare; and
  - it is used in such sector in a manner that is likely to involve significant risks, e.g. if the AI application
    - produces "legal or similarly significant effects" for the rights of an individual or a company, or
    - poses risk of injury, death or significant damage, or
    - produces "effects that cannot reasonably be avoided".

– Irrespective of the sector, the use of AI applications should exceptionally be considered high-risk "as such" if specific risks are involved, e.g. if AI applications are used for recruitment processes, or for remote biometric identification [p. 18].

– The additional requirements for high-risk AI applications should apply to all relevant economic operators providing AI-enabled products or services within the EU, regardless of where they are established [p. 22].

– AI applications involve many actors, including developers, deployers, users, producers and service providers. Obligations under new AI-specific legislation should lie with those actors who are "best placed" to address any potential risks [p. 22].

► **Additional requirements for high-risk AI applications**

– The Commission considers the following additional mandatory legal requirements for high-risk AI applications:
  - Obligations to train AI applications with data sets that are sufficiently broad and representative, or other requirements to ensure [p. 18f.]
    - that AI applications are safe, do not discriminate against people, and
    - that privacy and personal data are adequately protected.
  - Requirements to ensure [p. 20]
    - that AI applications are – over their life cycle – robust, accurate, able to adequately deal with errors or inconsistencies, resilient against attacks, and
    - that their outcomes are reproducible.
  - Different types of human oversight over certain AI applications, i.a. by requiring [p. 21]
    - the output of the AI application to become effective only after human review and validation,
    - possible human review after the output of the AI application has become effective,
    - a human to monitor an AI application while in operation who can intervene in real time and deactivate it, or
    - operational constraints to be imposed on the AI application, e.g. for a driverless car to stop in low visibility.

- Obligations to keep and make available, upon request by authorities, accurate records of the dataset used, documentation on the programming and training methodologies, and – in "certain justified cases" – the datasets themselves, in order to increase the transparency of AI [p. 19].
- Obligations to proactively provide information on the AI application's capabilities and limitations, including its expected level of accuracy, and a general obligation to inform individuals that they are interacting with an AI application, unless this is immediately obvious [p. 20].
- Remote biometric identification systems, e.g. facial recognition systems, may – under the applicable EU data protection law and the Charter of Fundamental Rights – be used only where such use is duly justified, proportionate and subject to adequate safeguards. The Commission wants to launch a broad debate on the circumstances that might justify such use, and on common safeguards [p. 21-22].

– The Commission considers the use of mandatory ex-ante conformity assessments to verify compliance with the requirements, before high-risk AI applications are placed on the internal market and, in particular cases, over their life cycle [p. 23].

► **Voluntary labelling scheme for low-risk AI applications**

– The Commission considers the introduction of a voluntary labelling scheme for AI applications that do not qualify as high-risk. Under this scheme, interested parties could commit to comply with specific requirements and would then be awarded a quality label, signalling the trustworthiness of their AI applications [p. 24].

– These requirements could be the requirements for high-risk AI application set out above, or a similar set of requirements established for the purpose of the labelling scheme [p. 24].

Policy Context

The White Paper on AI is part of the Commission's strategy for shaping Europe's digital future, alongside the European data strategy [cep**PolicyBriefs** to follow] and other key actions set out in the accompanying Communication COM(2020) 67. It builds on the EU's AI strategy of 2018, the Coordinated Plan on AI, and the Communication on Building Trust in Human-Centric AI [cf. cep**PolicyBrief** No. 2019-16]. In its resolution P9 TA-PROV(2020)0032, the European Parliament called on the Commission to develop a risk assessment scheme for AI. In the absence of EU legislation, some Member States have already taken steps toward regulating AI [p. 10].

ASSESSMENT

Economic Assessment

On the whole, the Commission's approach to regulating AI is appropriate for the following reasons:

First: **The proposed risk-based approach accounts for the fact that the possible consequences for users differ depending on** both **the specific AI application and the sector in which it is deployed**. For, setting strict requirements for AI applications may have two negative effects: European start-ups developing AI applications will be harmed if the costs of compliance are disproportionally higher than e.g. in the USA, where start-ups can first scale-up and only have to comply with the stricter EU requirements later when entering the EU market. Moreover, the benefits of deploying AI, e.g. automating long and tedious tasks, would be reduced if companies were to comply with strict requirements, such as always guaranteeing human oversight and redress. **The negative effects of strict regulation**, that might also weaken the development of an ecosystem of excellence, **are reduced if such requirements only apply to high-risk AI applications.**

Second: Given that an AI application can modify its functioning, high-risk AI applications have to be robust, accurate and their outcomes reproducible over their life cycle. This guarantees that e.g. risks of significant damage or injury as the unforeseen consequence of a software update are minimised. Therefore, the Commission rightfully proposes that repeated conformity assessments take place even after the AI applications have been placed in the market.

Third: For high-risk AI applications different forms of human oversight should be envisaged, as suggested by the Commission. As AI applications might lead to different magnitudes of risks even when performing their tasks in the same high-risk sector, the most appropriate form of oversight should be tailored to each specific case. For example, while the output of an AI application which can pose risks of death should not become effective without human approval, for other AI applications ex-post redress might be enough.

Fourth: The introduction of a voluntary labelling scheme for low-risk AI applications respects the freedom to conduct a business. Nevertheless, if the label is highly valued by consumers, companies will be pushed to adhere to the voluntary scheme in order to convince consumers of their "trustworthiness".

Legal Assessment

Legislative Competence of the EU

In exercising its internal market competence and in order to guarantee a high level of consumer protection [Art. 114, Art. 169 (1) TFEU], the EU may also adapt already harmonized law such as the Product Liability Directive and the

Product Safety Directive to new technical developments and risks and clarify or extend their application to AI-based products and services: Not only is free trade with the latter hampered by the legal uncertainty to which extent the current rules apply to AI; there is also a likely risk of legal fragmentation which may distort competition if Member States adopt differing liability and safety rules for AI, e.g. to regulate liability for services which is not covered by the directives.

While the assessment of competence for new AI-specific legislation depends on the exact content of a future proposal, such legislation can, in principle, also be based on Art. 114, TFEU provided that future obstacles to trade are likely to emerge as a result of differing national laws, and the legislation is designed to prevent them [cf. CJEU Case C-376/98, para. 86].

Subsidiarity

Dependent on the actual design of the follow-up measures. The overall aim to create EU-wide trust in AI applications and to ensure the emergence of a dynamic and competitive European AI industry and an EU-wide market uptake of AI cannot be sufficiently achieved by national initiatives.

Compatibility with EU Law in other Respects

Dependent on the actual design of the follow-up measures. Remote biometric identification – in particular facial recognition – technologies pose very high risks to the fundamental rights of those under surveillance and are therefore generally prohibited by the General Data Protection Regulation (GDPR). Their use may be permitted only under strict limitations, e.g. for reasons of "substantial public interest" or national security; it is however unclear when these requirements are fulfilled. As some Member States are already starting to regulate the public use of such technologies, it is reasonable that the Commission is seeking a debate in order to reach a common understanding on the circumstances which might justify such use and the necessary safeguards.

The proposed **mandatory legal requirements for high-risk AI applications and voluntary requirements for low-risk ones respect the principle of proportionality**. Extending the mandatory legal requirements to low-risk applications could be disproportionate.

As the White Paper lacks a comprehensive definition of AI, a legislative act will have to provide one. The White Paper rightly points out that this definition must be both precise enough to provide legal certainty and flexible enough to accommodate technical progress. Likewise, while **the** White Paper's **definition of "high-risk"** is flexible, it **is not precise enough in order to create legal certainty.** Notably, not every use of an AI application in one of the listed sectors should be considered "high-risk" just because it produces a "legal effect", as the Commission appears to say. This should only be the case if the legal effect is sufficiently significant in order to avoid both unnecessary burdens for companies and a discrepancy with the parallel criterion of "posing risk of injury, death or significant damage". **It must be clarified when** such possible legal effect or damage – and thus **the risks involved with the use of an AI application** – **have to be considered "significant",** because the application could then qualify as high-risk. It is also unclear who will decide on the qualification of an AI application as high risk or low risk.

Conclusion

The proposed risk-based approach accounts for the fact that the possible consequences for users differ depending on the specific AI application and the sector in which it is deployed. The negative effects of strict regulation are reduced if such requirements only apply to high-risk AI applications. Mandatory legal requirements for high-risk AI applications and voluntary requirements for low-risk ones respect the principle of proportionality. The definition of "high-risk" is not precise enough in order to create legal certainty. It must be clarified when the risks involved with the use of an AI application have to be considered "significant".