17/06/2020

AI-REGULATION.COM

# 17 Key Takeaways From Our Response to the EU White Paper on AI

***Theodore Christakis & Karine Bannelier***

To cite this article:

Theodore Christakis & Karine Bannelier, 17 Key Takeaways From Our Response to the EU White Paper on AI, Chair Legal and Regulatory Implications of Artificial Intelligence, MIAI Grenoble Alpes, June 17th 2020.

AI-Regulation.com

CHAIR LEGAL AND REGULATORY IMPLICATIONS OF ARTIFICIAL INTELLIGENCE

AI Regulation Chair Researchers have prepared a detailed submission to the European Commission's consultation on the AI White Paper. We discuss preferred options and ideas for AI laws in Europe. Read key takeaways and download the full text.

Artificial Intelligence has the potential to make breakthrough advances in several areas, but its growing applications raise complex questions and provoke broad concerns throughout society. In recent years, international organisations and governments around the world have stressed the need to address both the opportunities and societal challenges raised by AI. How can we guarantee that AI is designed and used responsibly? How do we establish ethical and legal rules to protect people, avoid bias and help establish fair and adequate liability schemes? The main mission of the AI Regulation Chair is to research how regulation can support sustainable and ethical innovation. Within this context we prepared a detailed submission for the consultation launched by the European Commission following the publication, in February 2020, of its White Paper on Artificial Intelligence - A European approach to excellence and trust.

In its White Paper the Commission proposed a series of options for AI rules that the EU plans to start releasing early next year and asked for feedback from the academia, public administrations, industry and civil society.

Some of the key findings and suggestions of our submission are the following:

1) The misuse of AI can endanger human rights. The objectives and construction of AI systems should thus be based on existing human rights, ethical principles and solid scientific standards. Measures should also be taken against unauthorized and abusive function creep.
2) Algorithmic decision-making systems are often complex systems that are difficult to understand. Technical solutions could be found to improve explainability but they imply some legal adjustments.
3) The specific characteristics of AI systems, including their complexity, connectivity, opacity, autonomy and vulnerability to cyberattacks, could make it more difficult for victims to present a claim of compensation or establish the causal link between the victim's harm and the defendant's action.
4) The current legal and regulatory framework in Europe already includes an important number of technology-neutral rules and provides satisfactory solutions to several problems. However, improvements to the current regulatory framework are welcome in order to prevent/mitigate/exclude a series of AI-related risks. There are at least four ways to improve the current regulatory framework:
    a) Fragmentation among European countries in the interpretation and application of these rules should be avoided.
    b) European authorities could provide guidance on "technology-neutral" rules.
    c) New rules could be necessary in fields where regulatory gaps exist.
    d) AI applications presenting important risks for human rights should be subject to adequate oversight and control.
5) A risk-based approach to new rules is welcome. However, the Commission's approach to determine "high-risk" might need more work and refinement. We indicate some ways to do so.

6) There are several extremely concerning/high-risk AI applications. Lethal autonomous weapon systems are the most famous and notable example. The use of facial recognition technologies in order to undertake mass surveillance is also scary. The same applies to AI tools that might be used in order to "manipulate" humans or "hack the human mind". The use of data and AI tools for disinformation and influence operations is also a major concern as it could lead to the destabilization of the very democratic foundations of our societies.

7) We agree with the Commission that several mandatory requirements should be considered for a future regulatory framework. We discuss requirements such as the quality of the training data sets (the garbage-in garbage-out principle); the information on the purpose and the nature of AI systems (transparency principle); the robustness of AI systems; solutions for human oversight and the AI control problem; and the need for clear liability and safety rules are essential in order to provide legal certainty to business and to protect the users.

8) The use of Facial Recognition Technologies (FRTs) creates important risks and problems. Some go beyond the subject matter of the only question on FRTs asked by the Commission (their "use in public spaces"). The limited scope of the question does not provide the opportunity for the Commission to receive feedback on some major issues, such as whether the legal framework for the use of FRTs for identification purposes by law enforcement agents in European countries, within the context of criminal investigations, is adequate or not – and what could be the eventual improvements. We recommend than a wider consultation on FRTs should be launched.

9) Concerning the use of FRTs in public spaces, there is already in Europe an important number of rules, starting with the GDPR and the Law Enforcement Directive, permitting to regulate the use of FRTs. However several improvements could be considered:

    a) First, there is a risk of fragmentation in the interpretation and the implementation of the GDPR and the Law Enforcement Directive in this field by EU Member States. The European Data Protection Board could play a very useful role here in providing guidelines for a harmonized interpretation of existing rules.

    b) Second, we believe that there is a lack of adequate information and transparency concerning the existing and intended uses of FRTs in public spaces in the different Member States.

    c) Third, FRTs present an unprecedented surveillance potential, capable of undermining societal choices and raising several important ethical issues. A public and democratic debate should precede the use of FRTs and lead to important political decisions on these matters.

    d) Fourth, this democratic debate should permit to fix from the outset some "red lines" beyond which no use, not even "experimental", of FRTs must be allowed.

    e) Fifth, if and when the use of FRTs is considered as legitimate and proportional, a common understanding of what exactly should be the "appropriate safeguards" would be useful. The question of oversight and control by democratic bodies, regulators and judicial organs is also crucial.

    f) Finally, specific attention should be given to the rules related to relevant databases.

10) Compliance of high-risk applications with the legal and ethical requirements should first be self-assessed by the developers themselves. We could draw here a comparison with the Data Protection Impact Assessment (DPIA) existing in the GDPR under the "privacy by design" principle (Article 35 of the GDPR).

11) More generally, high-risk applications should be tested and certified before they reach the single market and also be submitted to posterior conformity controls. However, these requirements and controls should be subject to several conditions and caveats.

12) The Product Liability Directive, has served since 1985 as an effective tool and contributed to enhancing consumer protection, innovation and product safety. However, some of its technologically neutral rules might need adjustment and update. The following issues should notably be considered:

    a) The Product Liability Directive is based on the principle that the producer is liable for damage caused by the "defect" in a "product" – very much perceived as a material object. However, in complex systems using AI (such as autonomous cars) there might be a constant interaction between traditional products (cars) and associated *services* making a sharp separation between them unfeasible. It is thus necessary to carefully think whether and how the Product Liability Directive might dial with services where there is a significant risk of damage to persons or property.

    b) An issue that needs particular attention is whether software is covered by the legal concept of product or product component and how an updated Directive could deal with this issue without imposing an impossible burden to software developers.

    c) The concept of "defectiveness" also needs careful thinking in relation with sophisticated autonomous systems based on machine-learning.

    d) A better definition of the repartition of liability between the various stakeholders involved, especially between the producer, the operator or what the Voss Report calls the "deployer", should also be considered.

13) National liability legislations offer an important and technology neutral framework in order to cover the challenges arising by the use of new technologies and to enable persons who suffered damage by the operation of such technologies to request reparation. However, EU Member States laws of tort are far from being harmonized. Current national liability rules could thus be updated and harmonized in order to better address some issues discussed in our submission.

14) New regimes of strict liability should be introduced in situations where the operation or the (mis)use of AI-systems involves a significant potential to cause harm to persons. This should be particularly the case when such systems are used in public spaces. In any case, what is important is that persons who suffered damage because of AI systems should be able to seek compensation.

15) An important issue to consider however is whether the same definition of "high risk" should be adopted in relation with the standard regulatory framework analysed in Section 2 of the Commission's survey and in relation with the specific question of liability rules discussed in Section 3. In any case the concept of "high-risk" needs to be clearly defined.

16) The proposal of the Rapporteur of the European Parliament Axel Voss to require from deployers of high-risk AI-systems to hold an adequate liability insurance is important. How exactly such a proposal could be implemented should be studied thoroughly with the participation of the insurance industry.

17) The important issue of hacking should be addressed. Hacking is a serious threat to users of software-based technologies and traditional national tort law rules may often prove insufficient because of the victim's inability to identify the tortfeasor. One solution would be for an eventual future legal regime of strict liability for high risks applications to exclude the exoneration of responsibility of the deployer if the hacker is untraceable or

impecunious. An alternative solution could be, as proposed in the Expert Group's 2019 Report Liability for Artificial Intelligence, to introduce a non-fault compensation scheme, equivalent to that applicable to victims of violent crimes, to ensure that persons who suffered damage from the hacking of high risk AI systems will be able to seek compensation.

The Chair's submission to the European Commission has been drafted by Professors Theodore Christakis and Karine Bannelier with the contribution of the four research fellows of the Chair: Mathias Becuywe, Stephanie Beltran Gautron, Maéva El Bouchikhi and Katia Bouslimani. The authors would like to thank all the members of the MIAI (Grenoble Alpes) for their input and comments during the preparation of this contribution.

You can download the full text of our submission here.



This work has been partially supported by MIAI @ Grenoble Alpes, (ANR-19-P3IA-0003)