Contribution of think tank Renaissance Numérique and of the Chair on the Legal and Regulatory Implications of AI of Grenoble Alpes University

–

Proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts

August 2021

Introduction

Artificial intelligence (AI) systems bear many opportunities for the European economy and society. They also raise significant challenges for the European Union, in terms of its capacity to innovate and, therefore, be competitive in this domain at the international level, but also in terms of its capacity to protect European citizens from the risks those technologies may entail for their rights and liberties. Those challenges are all the more vivid, especially in terms of regulation, that those technologies are particularly diverse (as are their possible uses), evolutionary and unpredictable.

In this context, the European Commission led by Ursula von der Leyen has started working on ways to accompany the development of artificial intelligence systems in the European Union. This contribution relates to the proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, presented on 21 April 2021[1].

This contribution was cowritten by think tank Renaissance Numérique and the Chair on the Legal and Regulatory Implications of AI of Grenoble Alpes University's Multidisciplinary Institute for Artificial Intelligence (MIAI). It is in line with previous contributions by the co-signatories on the European Commission's White Paper on Artificial Intelligence[2]. It also follows the organisation of a seminar by Renaissance Numérique, the Chair on the Legal and Regulatory Implications of AI and Facebook on 10 June 2021. This event gathered around forty participants involved in the topic at the European level, including lawyers, engineers, representatives of national and European public institutions, of civil society and companies,

[1] European Commission (2021), "Proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts", COM/2021/206 final, 107 pp.: https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0001.02/DOC_1&format=PDF

[2] European Commission (2020), "Artificial intelligence : A European approach to excellence and trust", COM(2020) 65 final, 26 pp.: https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf
See the co-signatories' contributions in the "Contributions to the consultation" section of the following page: https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12527-requirements-for-artificial-intelligence/public-consultation_fr

and academics.[3] It aimed at questioning the relevance of the proposal for a regulation on artificial intelligence and its quest for a certain balance[4], especially based on the experience feedback of the involved stakeholders.

This contribution focuses on specific dispositions of the text which raise questions with regard to the implementation of the future regulation. Other aspects of the regulation, such as those related to "real time" remote biometric identification systems, are the subject of parallel works by the co-signatories[5].

The two co-signatories hope the reflections presented in this paper will feed the ongoing debates at the European level in a useful way.

I. The definitions and revision principles enshrined in the text challenge the legibility and flexibility of the regulation

As a preliminary remark, the co-signatories wish to commend the progress brought about by the European Commission's proposal in that it aims to adopt a uniform approach to the legal framework surrounding artificial intelligence systems at the European level. Indeed, the fragmentation of legal regimes aimed at framing the uses and developments of AI systems between Member States would have been likely to slow down the beneficial advances of these technologies and to complexify the legal relationships between the players involved, at the risk of also undermining certain rights. The European Commission's desire to establish a harmonised classification and to rely on European standards[6] seems to be a guarantee of both legibility and stability for the development of these technologies within the European internal market. The Commission's willingness to put in place a framework that is likely to have an international resonance is also to be welcomed, particularly with regard to its definition of "artificial intelligence systems", which is based on the definition set out by the Organisation for Economic Co-operation and Development (OECD) [7]. In this respect, two elements deserve particular attention: the legal definition of artificial intelligence systems and the mechanisms for classifying AI systems as "high-risk".

[3] The organisers warmly thank the participants in the seminar, which enabled lively and in-depth debates, in particular those who kindly accepted to share preliminary addresses to frame the discussion: Samo Zorc, Secretary, Ministry of Public Administration, Slovenia; Salvatore Scalzo, Policy and Legal Officer 'Artificial Intelligence', DG CNECT, European Commission; Maria Luisa Stasi, Senior Legal Officer, Article 19; Elise Lassus, Research Officer, 'Freedoms and Justice' Department, European Union Agency for Fundamental Rights; Marcin Detyniecki, Head of Research and Development & Group Chief Data Scientist, AXA et Vice-president, Impact AI; Kari Laumann, Head of Section for Research, Analysis and Policy and Project Manager 'Regulatory Sandbox', Datatilsynet (Norway's data protection authority).

[4] In the explanatory memorandum of the proposed regulation, the European Commission mentions that: "*In light of the speed of technological change and possible challenges, the EU is committed to strive for a balanced approach."*.

[5] See the "Further readings" section at the end of this contribution.

[6] In this regard, see Recital 13 of the Commission's proposal.

[7] The definition proposed by the OECD is the following: *"An AI system is a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy."*. Source: OCDE (2019), Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449: https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449

Is a broad definition adapted to the reality and upgradeability of artificial intelligence?

With regard to the first point, it would be possible to qualify the definition proposed by the European Commission as utilitarian, in that it defines AI systems as generating "results" based on objectives defined by humans, those "results" being intended to be part of the environments in which humans interact with each other and with the systems[8].

However, this definition is not limited to this vision. If it were the case, it would most likely cover much broader technologies – broader than those commonly accepted as being AI[9] –, which legal framework is already formalised. In order to refine its definition, the European Commission thus chose to define AI systems according to the way they operate[10]. Those various techniques and approaches are restrictively listed in the proposal[11], and the Commission reserves the right to amend this list[12] to include additional techniques and approaches. As it stands, the three approaches that are considered are: machine learning approaches[13], logic- and knowledge-based approaches ("symbolic" AI)[14] and statistical approaches. Based on those elements, the definition proposed in the text appears both broad and dynamic, which calls for two observations.

By proposing an evolving definition that is not limited to existing AI techniques and approaches, the European Commission provides a forward-looking text, that is likely to cover AI systems that were not anticipated at the time of the initial drafting. Indeed, given the speed of technological developments in this field and the length of legislative procedures in the EU, adopting a regulation that could become obsolete only a few years after its adoption would have been counterproductive. It should be noted, however, that the dynamism of this definition is limited to "software" systems. Still, research on other forms of artificial intelligence is currently underway[15]. There is thus a risk that the text will become unfit if these new forms of AI become a reality. Considering not to limit the definition of AI systems to software applications would have the advantage of encompassing these perspectives, but at the same time would entail the risk of broadening the definition even further, to the detriment of clarity

[8] Article 3(1) of the Commission's proposal provides that "artificial intelligence system" means *"software that […] can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with;"*.
[9] As an example, a simple calculator meets the characteristics of this part of the definition.
[10] The first part of article 3(1) of the Commission's proposal mentions explicitly AI techniques and approaches as being a central element of the definition of an AI system: *"'artificial intelligence system' (AI system) means software that is developed with one or more of the techniques and approaches listed in Annex I"*.
[11] The restrictive list of the techniques and approaches that characterise an AI system is presented in Annex I of the proposed regulation: https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0001.02/DOC_2&format=PDF
[12] Article 4 of the proposed regulation provide for this possibility to amend Annex I, and article 73 sets out the terms and conditions for such amendments.
[13] The definition of machine learning according to the Oxford Languages dictionary is the following: "*the use and development of computer systems that are able to learn and adapt without following explicit instructions, by using algorithms and statistical models to analyse and draw inferences from patterns in data".*
[14] "Symbolic" AI was the first historical approach of artificial intelligence. This approach is essentially based on more or less explicit rules (e.g., "A implies B" or "If C and D, then E and not F") that are predefined by "expert" humans. For more information, see: https://en.wikipedia.org/wiki/Symbolic_artificial_intelligence
[15] In this regard, see, for instance: Woods, D., Doty, D., Myhrvold, C. *et al.*, *"Diverse and robust molecular algorithms using reprogrammable DNA self-assembly"*, Nature 567, 366–372 (2019): https://doi.org/10.1038/s41586-019-1014-9

and legal certainty for some actors. These questions call for the implementation of an agile governance system that is open to relevant expertise in order to reach a common vision of the technologies that should be encompassed within the text.

On this point, as the European Commission is basing itself on a restrictive list of AI techniques and approaches, some of which are particularly broad (e.g. statistical approaches), the text provides for the possibility of amending this list[16] while avoiding a long and complex process of revision of the regulation. While this dynamism is welcomed by the co-signatories, the modalities of the revision mechanism raise questions. The text grants this power of modification to the European Commission alone, while safeguards are provided through the European Council or the Parliament, which can revoke its delegation of power[17] or oppose the planned modification[18]. However, the co-signatories consider that it would be preferable to opt for a real co-construction mechanism for the amendments relating to the list of AI techniques and approaches. Indeed, this list is a basis for the definition of artificial intelligence systems, from which the implementation of the text will stem. Therefore, in the interest of legibility and security for the actors concerned, it would seem appropriate to integrate the relevant stakeholders and expertise in the framework of this revision mechanism.

To this end, the role of the European Artificial Intelligence Board (EAIB) could prove central. It could be given a role as a discussion forum in which the various stakeholders could share their expertise in order to propose amendments to the list. This essential role in the elaboration of the amendments would also allow the EAIB to provide recommendations to the concerned actors in order to clarify the situation of the developers or users of artificial intelligence systems *vis-à-vis* the regulation. The co-signatories believe that this would provide stability and legal certainty for all actors and would ensure the regulation fully assumes its role as a legal framework protecting both technological advances and rights and freedoms.

Classification mechanisms for AI systems: some criteria and exceptions need clarification

Beyond the definition, the second aspect on which the co-signatories wish to comment concerns the classification of artificial intelligence systems in the proposed regulation. The European Commission's text provides for a pyramidal classification of AI systems based on the risks that govern the level of legal supervision. The central concept for determining the classification of a system in a given category is the intended use of said system. While few restrictions apply to systems which use is not likely to result in significant risk[19], four use cases are prohibited, and regulation is put in place for systems which use presents a high degree of risk. The co-signatories approve the principle of this risk-based pyramidal classification but wish to put forward a number of remarks concerning its implementation.

---

[16] Article 4 of the proposed regulation provides that: *"The Commission is empowered to adopt delegated acts in accordance with Article 73 to amend the list of techniques and approaches listed in Annex I, in order to update that list to market and technological developments on the basis of characteristics that are similar to the techniques and approaches listed therein."*.

[17] Article 73(3) of the proposed regulation.

[18] Article 73(5) of the proposed regulation.

[19] Article 52 of the proposed regulation establishes an obligation to inform individuals for systems intended to interact with natural persons, for systems based on emotion recognition or biometric categorisation, or for audiovisual systems that generate or manipulate content that appreciably resembles existing content.

First, the co-signatories wish to highlight a number of points relating to uses that are prohibited in the proposed text. Indeed, doubts arise as to the wide margin of appreciation in determining the criteria for enacting the prohibition of certain uses or for evading it. Paragraphs 1(a) and (b) of Article 5 of the proposed regulation prohibit the placing on the market of systems using subliminal techniques and allowing the unconscious manipulation of an individual, as well as systems allowing the exploitation of people's vulnerabilities, insofar as they are likely to cause them harm, including potential harm. If there are no explicit derogations to these prohibitions, the criterion of potential damage as an element allowing to justify a prohibition can be questioned. Indeed, the potentiality of harm seems to be a notion that may prove difficult to assess[20]. The risk is therefore to prohibit a system on the basis of a potentiality without any damage being actually caused, thus depriving a beneficial use. Conversely, basing a ban on a potential harm may be short-sighted. A system may appear not to cause any potential harm *a priori*, while secondary or derivative uses that were not initially anticipated may create risks, particularly in terms of discrimination if using data that may prove biased. The co-signatories believe that it would be preferable to use criteria other than the potential for harm to delineate a ban.

A third prohibition concerns the use of social scoring[21] systems by public authorities or on behalf of public authorities[22]. This prohibition is not absolute, since it is qualified by several factors. First, it is subject to two alternative conditions. The first one is that social scoring leading to detrimental or unfavourable treatment is prohibited only insofar as it is carried out in contexts which are unrelated to the one in which the data was originally generated or collected. The second condition excludes detrimental or unfavourable treatment that is unjustified or disproportionate to the social behaviour of natural persons or its gravity. It appears, therefore, that the prohibition is conditioned by concepts which content is still unclear and which assessment depends on a case-by-case analysis. This wide margin of appreciation of the criteria that may justify the use of social scoring systems therefore seems too great to give this prohibition sufficient force. Finally, the prohibition is, in principle, only intended for public authorities. Other actors are not subject to it unless they intervene on behalf of public authorities, even if their uses would be unjustified or disproportionate, or if it takes place in a context that is different from that of the original data collection.

Finally, the co-signatories also note the broad exceptions to the prohibition for public authorities to use real-time remote biometric identification systems[23] in publicly accessible spaces for the purpose of law enforcement. As is the case of the above-mentioned prohibition, it strictly applies to public authorities for the purpose of law enforcement. The use of these systems is therefore not prohibited for other actors, nor for public authorities for different purposes. Moreover, the exceptions granted for the use of real-time biometric identification techniques despite the principled prohibition appear particularly broad[24], which risks depriving the prohibition of its substance. Indeed, the authorities can use such systems as part of the

[20] In this regard, see, for instance: Floridi, L., "The European Legislation on AI: A Brief Analysis of its Philosophical Approach", 1 June 2021: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3873273
[21] In Article 5(1)c of the proposed regulation, social scoring is defined as follows: *"the evaluation or classification of the trustworthiness of natural persons over a certain period of time based on their social behaviour or known or predicted personal or personality characteristics,"*.
[22] This prohibition is introduced under Article 5(1)c.
[23] This principled prohibition is introduces under Article 5(1)d.
[24] The possibility to fall under these exceptions is conditioned to their necessity for and proportionality to achieving the identified objective, in accordance with the requirements of Article 10 of the Law Enforcement Directive.

prosecution of thirty-two criminal offences[25], but also in order to prevent a specific, substantial and imminent threat to the life or physical safety of natural persons or a terrorist attack, or in order to locate potential victims of crime. These three possibilities hence cover a large part of the law enforcement activities of public authorities. The co-signatories believe that a stronger ban on these systems would be desirable, in line with the joint opinion of the European Data Protection Board (EDPB) and the European Data Protection Supervisor (EDPS) [26].

In addition to those prohibitions, the classification of AI systems as "high-risk" also raises questions. It seems that the notion of "high-risk" as developed in the proposal has two different meanings. The first one consists in labelling certain systems as "high-risk" because of vital issues that depend on the proper functioning of these systems. The second one includes certain systems in this category due to the fact that their malicious use would pose risks to ethics, fundamental rights, and individual and collective freedoms[27]. While the co-signatories agree that systems falling within these two spectrums of analysis can be included in the category of "high-risk" AI systems, they wish to highlight the fact that there are nuances between these two meanings, which the proposed regulation does not seem to take into account. Indeed, these nuances lead to a difference in the assessment of the risks that systems must prevent at each stage of the process (design, development, rollout). Consequently, the content of the impact assessment and its achievement cannot be identical depending on whether a system falls into one meaning of the notion of "high-risk" or the other. The co-signatories argue that these nuances should be taken into account in the text in order to provide an impact assessment that is as suited to the reality of the risks involved as possible. This differentiation would allow the text to gain in legibility for all the actors, hence making it easier for them to determine the obligations incumbent upon them in order to guarantee a sufficient level of security for the systems considered as being "high-risk".

Finally, the co-signatories wish to highlight some elements relating to the procedure for the revision of Annex III[28]. First of all, from a procedural point of view, the co-signatories are in

[25] Those thirty-two offenses correspond to those enshrined in framework decision 2002/584/JAI: participation in a criminal organisation / terrorism/ trafficking in human beings/ sexual exploitation of children and child pornography / illicit trafficking in narcotic drugs and psychotropic substances/ illicit trafficking in weapons, munitions and explosives/ corruption/ fraud, including that affecting the financial interests of the European Communities/ laundering of the proceeds of crime/ counterfeiting currency, including of the euro/ computer-related crime/ environmental crime, including illicit trafficking in endangered animal species and in endangered plant species and varieties/ facilitation of unauthorised entry and residence/ murder and grievous bodily injury / illicit trade in human organs and tissue/ kidnapping, illegal restraint and hostage-taking/ racism and xenophobia/ organised or armed robbery/ illicit trafficking in cultural goods, including antiques and works of art/ swindling/ racketeering and extortion/ counterfeiting and piracy of products/ forgery of administrative documents and trafficking therein/ forgery of means of payment/ illicit trafficking in hormonal substances and other growth promoters/ illicit trafficking in nuclear or radioactive materials/ trafficking in stolen vehicles/ rape/ arson/ crimes within the jurisdiction of the International Criminal Cour/ unlawful seizure of aircraft or ships / sabotage.

[26] EDPB-EDPS (2021), Joint Opinion 5/2021 on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act): https://edps.europa.eu/system/files/2021-06/2021-06-18-edpb-edps_joint_opinion_ai_regulation_en.pdf

[27] This notion of the double sense of “high-risk” was initially developed by Floridi, L., *op.* cit.

[28] Article 7 of the proposed regulation provides for the possibility for the Commission to add areas to the list presented in Annex III, which determines the uses qualifying as “high-risk AI systems”. The first paragraph of this article set two cumulative conditions that may lead to the amendment of Annex III: the inclusion of the proposed amendment in one of the eight existing areas of use, and the risk caused by the system at stake in terms of health, safety and fundamental rights, which should at least

favour of setting up a process of co-construction of the amendments, as has already been exposed regarding possible amendments to Annex I[29]. Secondly, the co-signatories raise doubts as to the robustness of the criteria that will enable the European Commission to forge a sufficient body of evidence to be likely to lead to the amendment of Annex III. Two observations seem appropriate in this respect. The first one is that there seems to be a lack of clarity around the main element of the qualification. It is not clear, as the text is currently drafted, whether the Commission will rely on risk or on evidence of harm (or negative impact) to qualify a system as "high-risk". Clarifying this point seems necessary in order to ensure the legibility of the legal framework. The second observation concerns the margin of appreciation of certain criteria aimed at guiding the qualification of a system as high-risk. For instance, the second criterion refers to the potential uses of the system. If this criterion were to be interpreted too rigidly, there would be a risk of falsely qualifying systems which benefits are proven. Conversely, too flexible an assessment would result in some systems escaping the enhanced obligations applicable to the "high-risk" category, even though unanticipated misuse would be likely to cause significant harm. This margin of appreciation, combined with the absence of a co-construction mechanism that would allow actors to anticipate legal developments, seems to create a legal framework that is too uncertain for actors involved in the development of future systems.

The lack of precision in the text regarding the possibility of extending the areas in which the uses considered to be "high risk" fall – beyond the eight areas that are already provided for[30] – can also be questioned. While this rigidity has the advantage of providing legal certainty for stakeholders, the lack of dynamism in this aspect of the classification could be risky when it comes to the relevance of the text with regard to uses that may not be anticipated at the time of the initial drafting.

II. A multi-stakeholder governance approach must be encouraged and strengthened in order to meet the future interpretation and implementation challenges of the text

The European Commission's text aims to create a framework to foster innovation, trust and the development of AI research and of its market within the European Union. In this respect, although the definitions and principles for the revision of AI categories are key issues[31], it is also necessary to focus on the governance of the involved actors, which will be helpful in

equivalent to those posed by the high-risk AI systems already referred to in Annex III. The second paragraph sets out criteria for establishing a set of indicators for assessing the severity of the risk involved.

[29] The amendment procedure for Annex III is also subject to the conditions laid down in Article 73 of the proposed regulation.

[30] These eight areas which make up Annex III are the following: Biometric identification and categorisation of natural persons; Management and operation of critical infrastructure; Education and vocational training; Employment, workers management and access to self-employment; Access to and enjoyment of essential private services and public services and benefits; Law enforcement; Migration, asylum and border control management; Administration of justice and democratic processes.

[31] Many authors have written about the risk-based approach and on legal concepts which are specific to the draft European regulation. See, for instance :« Projet de règlement sur l'IA (I) : des concepts larges retenus par la Commission », *Dalloz Actualité*, Cécile Crichton, 3 May 2021 : https://www.dalloz-actualite.fr/flash/projet-de-reglement-sur-l-ia-i-des-concepts-larges-retenus-par-commission

understanding but also in interpretating this regulation. The capacity of this text to harmonise and implement these new rules relies on this governance logic.

Although Article 59[32] of the proposal gives Member States some latitude to designate national competent authorities tasked with, on the one hand, guiding the understanding of these rules and, on the other, ensuring that they are applied, the designation of these authorities raises questions. Indeed, even though several national competent authorities can be designated[33], only one should be designated as the national supervisory authority or as official point of contact within the Union[34].

While this flexibility is welcome from a Member State perspective, harmonizing implementation might be preferable, as the difficulty of effectively creating and governing an actual cooperation mechanism within the European Union has already been illustrated, in particular following the adoption and implementation of the General Data Protection Regulation (GDPR). The proposed regulation on artificial intelligence is no exception and, as a result, there is a risk to create gaps in interpretation and treatment between Member States. Indeed, initial feedback from the implementation of the GDPR has highlighted that there are different interpretations of the text within Member States, but also gaps in its implementation. In addition, the creation of a new body, the EDPB, has highlighted and crystallised a number of key issues relating to the system of governance and cooperation between states, but also to the interpretation and implementation of the text. As a result, crucial issues concerning the authorisation within the Union of certain systems, such as real-time and remote biometric identification systems, have been the subject of numerous questions in recent years[35]. Questions related to the one-stop shop principle, to the cooperation between data protection authorities and to the margin of appreciation of Member States also illustrate some of the issues that could emerge in the context of the regulation on artificial intelligence.

In addition to the multiplicity of national authorities that could be competent, the proposal also mentions specific sectors that could and should benefit from specific provisions and designations in order to ensure the harmonisation of EU rules[36]. Consequently, while this text

---

[32] In Article 59(1) of the regulation proposal specifies that national competent authorities shall be established or designated by each Member State for the purpose of ensuring the application and implementation of the regulation. Member States shall then inform the Commission of their designation or designations (including if it concerns national supervisory authorities) and, where applicable, the reasons for designating more than one authority (Article 59(3)).

[33] National competent authorities may be the national supervisory authority, the notifying authority and the market surveillance authority (Article 3(43)).

[34] Recital 77 of the proposal provides that *"Member States hold a key role in the application and enforcement of this Regulation. In this respect, each Member State should designate one or more national competent authorities for the purpose of supervising the application and implementation of this Regulation. In order to increase organisation efficiency on the side of Member States and to set an official point of contact vis-à-vis the public and other counterparts at Member State and Union levels, in each Member State one national authority should be designated as national supervisory authority."*.

[35] In this regard, the EDPB had first adopted guidelines on these issues as part of its "Guidelines 3/2019 on processing of personal data through video devices version 2.0" (§29) adopted on 29 January 2020 : https ://edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_201903_video_devices_en_0.pdf The EDPB then clearly called for a ban of those real-time biometric identification systems following the publication of the AI Act by the European Commission, via the aforementioned joint opinion with the EDPS (see footnote 26).

[36] Recital 80 of the proposed regulation mentions that *"Union on financial services includes internal governance and risk management rules and requirements which are applicable to regulated financial institutions in the course of provision of those services, including when they make use of AI systems"* and that *"In order to ensure coherent application and enforcement of the obligations under this*

raises questions as to its complementarity with other texts that already regulate AI systems, it also raises questions as to the governance mechanisms between the various supervisory authorities and actors ensuring compliance with these texts. The flexibility offered by the proposed regulation in terms of the choice of these authorities raises the issue of effective European harmonisation. Therefore, the effectiveness of the governance process presented in Chapter 2 of Title VI of the proposal could be undermined by the latitude given to Member States to interpret the regulation and to designate one or more competent authorities. Indeed, if the European Union defines the regulation that will ultimately govern the AI systems that are already present on the EU market, but also those that will enter this market after the adoption of the regulation, the power left to Member States must necessarily be balanced against an increased control of higher EU authorities in case of dispute. In this regard, it appears as though there is no plan for an organised coordination between market surveillance authorities, and as though the Commission will be the one to decide in case of a disagreement on the withdrawal of an AI system from the market[37]. Although the supervisory authority for a given sector is generally the one that fulfils this surveillance role[38], there is a risk that, where this authority is not competent, divergent approaches will be adopted across sectors. The supervisory authority would then only have a complementary role by centralising all available data[39].

These risks underline the fact that several governance systems seem to be existing in parallel, without a clear system of communication between them. It therefore appears necessary for the text to better articulate how national surveillance and competent authorities will cooperate, as this seems to be unclear at the moment, relying mainly on EAIB[40] and Member States. Thus, the designated supervisory authorities will play a key role, as long as they are given the means to carry out their tasks. While the text does provide for this[41], in practice, there are inequalities in resources and technical maturity between Member States when it comes to artificial intelligence. In addition, differences relating to the legal culture of each country, or even the ambition of each Member State to try to attract certain talents or to concentrate certain areas of expertise on their territory and thus ensure that the next unicorns will be under their jurisdiction, underline important economic stakes[42].

Regulation and relevant rules and requirements of the Union financial services legislation, the authorities responsible for the supervision and enforcement of the financial services legislation, including where applicable the European Central Bank, should be designated as competent authorities for the purpose of supervising the implementation of this Regulation, including for market surveillance activities, as regards AI systems provided or used by regulated and supervised financial institutions". Another example would be the AI services depending on the institutions, agencies and bodies of the Union, which are covered by special provisions.

[37] Article 63(2) of the proposed regulation.

[38] As an example, and according to Article 63(3), for high-risk AI systems related to products to which legal acts listed in Section A of Annex II apply, the market surveillance authority shall be the authority responsible for market surveillance activities designated under those legal acts. In the sector of financial services, the market surveillance authority shall be the authority responsible for financial surveillance of those institutions. (Article 63(4)).

[39] Article 62.

[40] The latter is more commonly known as "the "Board".

[41] Article 59(4) of the proposal provides that Member States shall ensure that national competent authorities are provided with adequate financial and human resources to fulfil their tasks under the regulation.

[42] By way of comparison, some mentioned at the seminar of 10 June 2021 that there is a risk of "forum shopping" similar to that found under the GDPR between some European DPAs in the face of inaction by the Irish DPA. However, this point appears to be eroding following the European Parliament's vote

This last point is all the more crucial as there is a certain degree of uncertainty concerning the role data protection authorities may have in this governance. This analysis is also shared by other actors, such as the EDPS and the EDPB. In their joint opinion of 21 June 2021[43], the two institutions call for national data protection authorities (DPAs) to be designated as supervisory authorities in order to ensure and contribute to a more harmonised approach. This, according to the two organisations, should allow for a consistent interpretation of the provisions on data processing and avoid contradictions in their implementation across Member States. Indeed, DPAs do have experience in this area and have already had to regulate and take a stance on AI systems. This point is notably supported by the French DPA, the *Commission Nationale de l'Informatique et des Libertés* (CNIL), which, in a press release, has supported the joint opinion of the EDPB and EDPS and called for DPAs to be designated as supervisory authorities. However, it is essential to provide these authorities with the technical and human resources necessary to take on these additional tasks. Indeed, if DPAs were to be designated as supervisory authorities under the new regulation, the text would considerably extend their scope of competence and supervision as well as their tasks. The low level of additional resources granted to DPAs following the adoption of the GDPR[44] highlights the risks linked to multiplying the competences granted to DPAs and the need to invest sufficiently in these authorities – or more generally in all competent national authorities – within the framework of the future regulation. According to some stakeholders, however, it would be appropriate to designate DPAs, i.e. independent administrative authorities that are already established, to ensure compliance with and control the application of the proposed regulation on AI. In that sense, the French CNIL has stressed the point that DPAs are already controlling AI systems that involve processing of personal data.

No matter which authority is designated, it will be necessary to ensure that it has the capacity to absorb these new tasks. Moreover, if DPAs are not designated as national supervisory authorities, it will be necessary to organise a consultation in order to ensure a coherent application of the future regulation with regard to existing regulations on data protection (RGPD, Law Enforcement Directive) within the European Union.

It should also be stressed that the future regulation dedicated to AI differs from the GDPR and that the issues surrounding these systems do not only concern fundamental rights. It is also about ensuring product safety on the market and establishing a multi-actor cooperation that is not limited to collaboration between national competent or supervisory authorities but encompasses a plurality of actors[45].

Finally, a last point of analysis concerns the need to grant more autonomy to the EAIB, especially with regard to its independence from the European Commission. This point is also underlined by the EDPS and the EDPB, who question in their joint opinion the predominant role of the European Commission in this Board. Indeed, the current plan to have it chaired by

in favour of a resolution asking the European Commission to open infringement proceedings against Ireland.

[43] See footnote 26.

[44] For an example, see: La Quadrature du Net, “Dysfonctionnements systématiques des autorités de protection des données : le cas belge”, 8 juillet 2021 : https://www.laquadrature.net/2021/07/08/dysfonctionnements-systemiques-des-autorites-de-protection-des-donnees-le-cas-belge/

[45] Data protection authorities will play a key role, but various actors such as judges, legislators and supervisory authorities will play an equally important one.

the European Commission alters the autonomy of the Board[46]. Among other things, the Commission would be responsible for convening meetings and preparing the agenda of these meetings, but also for approving the Board's rules of procedure[47]. In order to compensate for this lack of independence, the co-signatories advise to have a look at the ideas raised as part of Renaissance Numérique's reflections on the separation of power at the European level (in particular the work conducted by the think tank on the Digital Markets Act) [48] and to draw inspiration from these in order to provide the future Board with a sufficient place and autonomy in this governance.

III. In terms of impact assessments, the text poses implementation challenges

The clarification and opening up of governance are an essential lever for the operational implementation of this new legal framework, particularly with regard to the requirement imposed by the text to conduct impact assessments of AI systems. The assessment of these technologies is indeed based on many unknown variables and is a field of research that is still being explored. Moreover, in general, given the diversity of AI systems and of their uses, determining the right scope of analysis is not always an easy task. AI systems are often part of larger processes. The question then arises as to what should be evaluated: only part of the process or the entire process? Choosing the second option could help improve knowledge and caution. But it could also prove less efficient, by generating a mass of irrelevant data. Imagining the concrete implementation of these impact assessments thus calls for a reinforced and continuous dialogue between the different stakeholders involved, in order to resolve possible difficulties in application.

In the field of artificial intelligence, certain essential principles are not yet the subject of an evaluation methodology, or even of a definitive definition. This is the case, for example, of the notion of transparency, which is mentioned in Article 13 of the proposed regulation. In the first paragraph of this article, the European Commission makes imprecise reference to *"[a]n appropriate type and degree of transparency"*. It is however difficult to know precisely what *"appropriate"* or *"type"* may mean, let alone "*transparency*"[49].

In terms of fundamental rights, a recent study conducted by the European Union Agency for Fundamental Rights[50] among the actors of the AI value chain within the EU has highlighted differences in understanding between these actors depending on the nature of the rights. While there is a high level of awareness regarding the protection of personal data – to which the GDPR has undoubtedly significantly contributed – there is less awareness regarding compliance with the principle of non-discrimination or access to remedies for affected

[46] As defined in Article 57(3) of the proposal.
[47] As defined in Article 57(2) of the proposal.
[48] Renaissance Numérique (2021), *Digital Markets Act: A revolution or a legal contradiction?*, 34 pp.: https://www.renaissancenumerique.org/system/attach_files/files/000/000/285/original/Renaissance_Nume%CC%81rique-NOTE_DMA_English.pdf?1617294205
[49] To learn more, see: Renaissance Numérique (2017), L'éthique dans l'emploi à l'ère de l'intelligence artificielle, 23 pp.: https://www.renaissancenumerique.org/publications/l-ethique-dans-l-emploi-a-l-ere-de-l-intelligence-artificielle
[50] European Union Agency for Fundamental Rights (2020), Getting the future right – Artificial intelligence and fundamental rights, 106 pp. : https://fra.europa.eu/en/publication/2020/artificial-intelligence-and-fundamental-rights

individuals. In this respect, the co-signatories of this contribution, like other actors, call for the strengthening of the text's requirements in terms of remedies[51].

These principles require clear definitions and precise metrics. However, at this stage, some of the concepts presented in the proposed regulation do not have a strong legal tradition *per se* and even raise difficulties of interpretation. These will need to be resolved to enable stakeholders to understand these concepts in the context of their assessment processes. For example, Article 9 – which is dedicated to the risk management systems required for high-risk AI systems – contains a number of terms that raise difficulties, such as *"reasonably foreseeable"* (paragraph 2(b)) or *"suitable"* (paragraph 2(d)). Article 15, dealing with accuracy, robustness and cybersecurity requirements, is also illustrative in this respect. In particular, it refers to *"an appropriate level"* of accuracy, robustness and cybersecurity, as well as the need for high-risk AI systems to operate in a *"consistent"* manner. Similarly, while the requirements for training, validating and testing datasets (Article 10 "Data and data governance") are essential, as currently drafted they raise difficulties of interpretation and implementation. For example, as defined in paragraph 3 of this article, it is required that these data sets be *"free of errors and complete"*. From a technical point of view, it is difficult to establish these two conditions. In particular, AI systems based on unsupervised learning are based on a machine learning approach that iteratively searches for patterns in large, unstructured data sets. It is therefore difficult to guarantee that these sets be free of errors.

In view of these interpretation challenges and in line with the necessity to strengthen its role in terms of governance, the EAIB should be tasked with drawing up operational recommendations in consultation with the multi-stakeholder expert group and with the relevant actors of the ecosystem. It could be inspired by the EDPB, but with a reinforced dialogue logic. Support for stakeholders at the national level should also be strengthened. In addition, and in order to bring together the diversity of expertise needed to implement the regulation – and not only expertise in terms of data protection – a "regulatory hubs" approach could be implemented. These hubs could be led by the national supervisory authorities.

To facilitate this dialogue, in addition to other information required and made public on AI systems[52], the results of impact assessments should also be made public, regardless of the degree of risk induced by these technologies.

IV. Regulatory sandboxes as drivers of innovation and excellence in the European Union

In the proposed regulation, regulatory sandboxes are presented as measures aimed at supporting innovation. This "innovation" side of the regulatory approach presented in the text is a key point. The proposed regulation is based on the assumption that it is the existence of a stable and clear regulatory framework that will enable the development of the AI market in the European Union. However, the framework remains complex and will not be sufficient in

[51] *"The AIA could do much more to protect consumers' rights and be much more incisive about providing measures to redress the possible harms or losses that AI systems may cause. This is the part where one may expect and welcome more improvements in the proposal. It was one of the main recommendations made by the AI4People project: "7. Develop a redress process or mechanism to remedy or compensate for a wrong"".* Floridi, L., *op. cit*.
[52] See, for instance, Article 60 on the EU database for stand-alone high-risk AI systems.

itself to provide an incentive mechanism likely to create a market. As the text currently stands, the incentives it contains are mainly aimed at *"small-scale providers"*[53].

It should be noted that regulatory sandboxes are also levers for cooperation between regulatory authorities, businesses, and other stakeholders. On the one hand, these collaborations allow companies to innovate in a protective framework, since they can benefit from regulatory expertise and be quickly enlightened if any legal uncertainties arise. On the other hand, regulatory authorities can benefit from practical exchanges, in order to better understand AI systems thanks to feedback from those who develop, train and deploy them, and thus possibly adapt the regulation and their recommendations accordingly. Furthermore, effective transparency of these innovation processes can allow all stakeholders to create an environment of trust.

Regulatory sandboxes are therefore to be considered as resulting in a win-win situation, as all stakeholders should benefit from the collaboration. The co-signatories therefore call for the ambition of the text to be strengthened in this area. In order for regulatory sandboxes to work and be a real lever for innovation, it is particularly important to build a harmonised approach between the competent national authorities, and that these authorities have sufficient human, technical and financial resources to implement them. This point is all the more crucial as national competent authorities will have an extended role: they will not only have a controlling role (ensuring the correct application of the regulation), but also a supporting role.

Still, this common approach could be compromised. Indeed, based on the text in its current version, *"the provisions of the regulation are not overly prescriptive and leave room for different levels of Member State action for elements that do not undermine the objectives of the initiative, in particular the internal organisation of the market surveillance system and the uptake of measures to foster innovation"*[54]. If Member States have the possibility to act freely as regards the organisation of these tools, the risk of an imbalance from one state to another is real. In this respect, the role of the Board will be essential. Indeed, it should *"contribute to uniform administrative practices in the Member States, including for the functioning of regulatory sandboxes referred to in Article 53"*[55].

For this reason, the functioning of these regulatory sandboxes should be discussed in a collegial manner between the European Commission, the EAIB, national competent authorities, the AI Expert Group and relevant industry and civil society representatives. For now, the functioning of regulatory sandboxes varies from one Member State to another, and states often struggle to move away from a regulatory compliance approach and to fully integrate an innovation approach. Among the existing regulatory sandboxes, the competent authorities could draw inspiration from the one initiated by the British Financial Conduct Authority (FCA) in 2018 in the field of fintech[56]. This initiative stands out from others in that its ambition is global. Indeed, the British authority has created the Global Financial Innovation Network (GFIN), which brings together eleven global regulators, the objective of which is *"to*

[53] For instance, Article 55 "Measures for small-scale providers and users" provides them with compliance assistance, allowing them priority access to AI sandboxes, increased awareness tailored to their needs and a privileged communication channel.
[54] Explanatory memorandum, paragraph 2.4.
[55] Article 58(b).
[56] In this regard, see on the official website of the FCA: "FCA Innovation – fintech, regtech and innovative businesses": https://www.fca.org.uk/firms/innovation

consider how to build new ways of sharing experience and managing emerging issues"[57] in order to put an end to regulatory borders. The same logic could be followed at the European level, with the support of the EAIB. At the Union level, it would be possible to start by specific projects or by identifying priority areas.

As for the sandbox initiative launched by the Norwegian DPA, *Datatilsynet*, in 2020[58], it is characterised by the transparency approach that has been put at the heart of the way it operates. The authority publishes processes and results as the sandbox progresses. This approach ensures transparency for both companies and other involved stakeholders, who can easily access the information. As such, it could also inspire the development of future regulatory sandboxes.

Conclusion

At a time when artificial intelligence is the subject of intense competition at the international level, the new legislation should enable the European Union to set its own standards in this area. However, in order to achieve this objective and ensure that these standards are adopted as widely as possible, the balance sought in the preamble to this legislation must be achieved. To do so, it is necessary to anticipate now the translation of this legal object into technical terms, both to enable the concerned actors to apply it correctly and to guarantee the protection of individuals.

Given the unpredictable nature of AI systems and the potential scale of their impact, the implementation of the text cannot afford to be carried out without an agile governance, one that is open to relevant stakeholders and expertise, and with reinforced resources. Just like for regulatory sandboxes, it will be necessary to develop ambitious tools that do not oppose regulation and innovation and that allow the European Union to establish itself as a territory of excellence when it comes to artificial intelligence.

***

Further readings

- Renaissance Numérique (2020), Facial Recognition: Embodying European values, 103 pp.: https://www.renaissancenumerique.org/system/attach_files/files/000/000/235/original/report_facial_recognition.pdf?1592553217

- Renaissance Numérique (2021), Regulation of facial recognition technologies: A comparative analysis of France and the United Kingdom, 10 pp.: https://www.renaissancenumerique.org/system/attach_files/files/000/000/291/original/RenaissanceNumerique_EventSummary_FacialRecognitionUKFR.pdf?1623223645

- Theodore Christakis (2021), Facial Recognition in the Draft European AI Regulation: Final Report on the High-Level Workshop Held on April 26, 2021: https://ai-

[57] Olivier Pinaud, « La FCA rallie onze régulateurs à son idée de « bac à sable » mondial pour la fintech », *L'AGEFI*, 8 August 2018: https://www.agefi.fr/fintech/actualites/quotidien/20180808/fca-rallie-onze-regulateurs-a-idee-bac-a-sable-253655

[58] See: Datatilsynet, "Sandbox for responsible artificial intelligence": https://www.datatilsynet.no/en/regulations-and-tools/sandbox-for-artificial-intelligence/

regulation.com/facial-recognition-in-the-draft-european-ai-regulation-final-report-on-the-high-level-workshop-held-on-april-26-2021/

- Team AI Regulation (2021), Facial Recognition in the Draft AI Regulation: Useful Materials: https://ai-regulation.com/facial-recognition-in-the-draft-ai-regulation-useful-materials/

Authors

Mathias Becuywe, Research Engineer, MIAI Grenoble Alpes

Stéphanie Beltran Gautron, Research Engineer, MIAI Grenoble Alpes

Jennyfer Chrétien, Executive Director, Renaissance Numérique

Maéva El Bouchikhi, Research Engineer, MIAI Grenoble Alpes

Contributors

Théodore Christakis, Professor of Law and Director of the Chair on the Legal and Regulatory Implications of AI, MIAI Grenoble Alpes

Guillaume Morat, Senior Associate, Pinsent Masons

Marine Pouyat, President, W Talents

Annabelle Richard, Partner, Pinsent Masons

Proofreading

Jessica Galissaire, Studies Manager, Renaissance Numérique