International Journal of Law and Information Technology, 2022, 30, 88–109
https://doi.org/10.1093/ijlit/eaac007
Advance Access Publication Date: 18 March 2022
Article

Ex machina: preliminary critical assessment of the European Draft Act on artificial intelligence

Vera Lúcia Raposo *

ABSTRACT

This article will unpack the European Draft Act on Artificial Intelligence (AI), the first (both in Europe and in the world) far-reaching regulation in this domain. It categorizes AI systems in three (eventually four) levels of risk and it assigns to each level a particular legal framework, with its own limitations and obligations. The Draft Act was created for laudable purposes, namely to harmonize digital development with fundamental rights and European values. However, similar to many other ambitious projects, it falls short in many ways. As this article will show, several issues are incompletely regulated, and there are doubts about the exact scope and content of the legal solutions outlined in the draft. It also potentially overlaps with several other European norms, introducing the possibility of conflicts. Finally, the law's focus on fundamental rights may come at the price of digital innovation, although critics claim that the Draft Act does not go far enough to protect such rights.

KEYWORDS: artificial intelligence, European law, fundamental rights, innovation, risk assessment

INTRODUCTION

On 21 April 2021, the European Commission (EC) disclosed its proposal for a future European Union (EU) regulation on artificial intelligence (AI).[1] The aim of the legislation is to facilitate and develop the use of AI within the EU to support the establishment of the Digital Single Market[2] and ultimately to achieve digital sovereignty.[3] The

* Faculty of Law, University of Coimbra, Coimbra, Portugal. E-mail: vera@fd.uc.pt

1 A presentation of the Draft Act appears in Yannick Meneceur, 'European Commission's AI Regulation Proposal: Between Too Much and Too Little?' (23 April 2021) <https://www.linkedin.com/pulse/european-commissions-ai-regulation-proposal-between-too-meneceur/> accessed 28 November 2021; Vera Lúcia Raposo, 'May I Have Some Artificial Intelligence with My Human Rights? About the Recent European Commission's Proposal on a Regulation for Artificial Intelligence' (*KSLR EU Law Blog*, 24 May 2021) <https://blogs.kcl.ac.uk/kslreuropeanlawblog/?p=1569> accessed 12 June 2021.

2 EU4DIGITAL, 'EU Digital Single Market' (2021) <https://eufordigital.eu/discover-eu/eu-digital-single-market/> accessed 4 October 2021.

3 Charles Michel, 'Digital Sovereignty Is Central to European Strategic Autonomy (Speech)' (2021) <https://www.consilium.europa.eu/en/press/press-releases/2021/02/03/speech-by-president-charles-michel-at-the-digitaleurope-masters-of-digital-online-event/#:~:text=10%3A25-,Digital%20sovereignty%20is%20central%20to%20European%20strategic%20autonomy%20%2D%20Speech%20by,of%20digital%202021%22%20online%

EC aims to achieve these goals without sacrificing fundamental rights and the EU's values and ethical principles, which may be threatened by AI.[4] As the Draft Act is merely a proposal, it will take some time before it is passed into law as a regulation,[5] possibly with some changes to its content. However, its core concepts are not expected to drastically change.

This article attempts to clarify the content of the Draft Act. Many of its standards remain unclear, with their meaning to be clarified by the EC and subsequently by the entities in charge of the interpretation and application of the future AI regulation and by the European Court of Justice.

ORIGINS OF THE DRAFT ACT

Structurally, the Draft Act is inspired by the General Data Protection Regulation (GDPR).[6] This is evident in its risk-based approach, its extraterritoriality, and the way it envisages penalties for non-compliance. However, they differ on other matters. For instance, while the GDPR allows the Member States some discretion on particular topics,[7] the Draft Act is silent as to possible derogations of its solutions.[8]

Substantially, most of the ideas in the Draft Act can be found in the European Parliament Resolution on the framework of ethical aspects of AI, robotics and related technologies[9] and in the White Paper on AI,[10] both released in 2020. The White Paper, in particular, offers a preview of certain Draft Act solutions. However, the Draft Act distances itself from the White Paper in some nuclear aspects. First, the

20event&text=Digital%20is%20one%20of%20the,had%20heard%20of%20COVID%2D19> accessed 22 October 2021.

4 See C Fernández-Aller and others, 'An Inclusive and Sustainable Artificial Intelligence Strategy for Europe Based on Human Rights' (2021) 40 IEEE Technology and Society Magazine 46, 47. See also Vincent C Muller (ed), *Risks of Artificial Intelligence* (CRC Press 2016).

5 The final draft must be agreed on by the European Parliament and the Council of the European Union, which could take several years. An implementation period of 2 years will follow this before the law becomes binding.

6 reg (EU) 2016/679 of the European Parliament and of the Council, of 27 April 2016, on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Dir 95/46/EC (General Data Protection Regulation).

7 Joanne Vengadesan and Nora Pook, 'United with Differences: Key GDPR Derogations Across Europe' (26 March 2019). <https://www.penningtonslaw.com/news-publications/latest-news/2019/united-with-differences-key-gdpr-derogations-across-europe> accessed 13 October 2021.

8 In terms of a real time, remote, biometric identification system used in publicly accessible spaces for the purpose of law enforcement, 'Member States remain free under this Regulation not to provide for such a possibility at all or to only provide for such a possibility in respect of some of the objectives capable of justifying authorised use identified in this Regulation' (Recital 22). It is unclear whether this description applies to other AI systems and/or only refers to the possibility of imposing more bans and requirements than the ones set forth in the Draft Act or, conversely, whether Member States may exclude some of the requirements and prohibitions of the law. In light of the general spirit of the Draft Act, the latter does not seem possible. See Martin Ebers and others, 'The European Commission's Proposal for an Artificial Intelligence Act—A Critical Assessment by Members of the Robotics and AI Law Society (RAILS)' (2021) 4(4) J-multidisciplinary Scientific Journal 589, 590.

9 European Parliament Resolution of 20 October 2020 with recommendations to the Commission on a framework of ethical aspects of artificial intelligence, robotics and related technologies (2020/2012(INL)) <https://www.europarl.europa.eu/doceo/document/TA-9-2020-0275_EN.html> acessed 13 June 2021.

10 European Commission, 'White Paper on Artificial Intelligence-A European Approach to Excellence and Trust', Brussels, 19.2.2020 COM (2020) 65 final <https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf> accessed 15 May 2021.

Draft Act more carefully details the potential violations of fundamental rights that may arise from the use of AI. Given the severe consequences in terms of liability arising from the violation of the Draft Act rules, this may generate fears concerning placing new forms of AI on the market. Secondly, on a related note, the Draft Act is generally less encouraging of the development of new AI technologies. Whereas the White Paper emphasizes economic and technological development, the Draft Act is more focused on the protection of fundamental rights[11] (although, interestingly, it has also been said that the Draft Act does not go far enough in terms of rights protection),[12] imposing many requirements and limitations on AI providers. The development of trustworthy AI is a clear goal of this proposal[13]; this entails the development of AI systems where 'people can trust that the technology is used in a way that is safe and compliant with the law, including the respect of fundamental rights'.[14]

DEFINITION OF AI

The Draft Act adopts a two-stage model to define 'AI'. First, in Article 3, it presents the relatively vague and generic concept of an 'artificial intelligence system', meaning 'software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations or decisions influencing the environments they interact with.'

Annex I, to which Article 4 refers,[15] further condenses this definition, listing the techniques and approaches in the field of AI covered by the Draft Act as follows:

a. machine learning approaches, including learning supervised, unsupervised and by reinforcement, using a wide variety of methods, including deep learning;
b. approaches based on logic and knowledge, namely knowledge representation, inductive (logic) programming, knowledge bases, inference and deduction engines, reasoning systems (symbolic) and expert systems; and
c. statistical approaches, Bayes estimation, research and optimization methods.

One criticism of this definition is the absence of relevant distinctions. For example, Swedsoft claims that the definition makes no distinction between algorithms

11 Meneceur (n 1).

12 ibid. An earlier version of the Draft Act (which circulated in January 2021) promised more in terms of the protection of rights and the rule of law. Critics point out that the current version has given in to other interests, and probably it could not be in any other way, as there clearly remain several interests and values in this domain to be taken into account.

13 The terms 'trust' and 'trustworthy' are frequently repeated throughout the text of the Explanatory Memorandum and in the proposal itself (European Commission, 'Explanatory Memorandum of the Proposal, Regulation of the European Parliament and of the Council', 21 April 2021 <https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0001.02/DOC_1&format=PDF> accessed 22 April 2021.

14 European Commission ibid 1.

15 When no further specification is added, all quoted norms refer to the Draft Act.

and data on the one hand and models built with combinations of algorithms and data on the other, nor between different types of algorithms and different types of data. This matches the aforementioned general definition of AI.[16]

The overly broad nature of the definition is also a target of criticism.[17] It is broader than most AI definitions, which will cause software not commonly considered to be using AI to be covered by the future regulation.[18] This definition of AI is presented as being broad enough to cover future developments in AI technology, with the aim of extending the life of the law. However, such a broad definition may just as easily become inapplicable to any specific system. In any case, no matter how broad a definition is, it eventually will become outdated, as technology always develops beyond any legal definition.

INSTITUTIONAL FRAMEWORK

The Draft Act sets up specific bodies and entities to operate the regulation. At the national level (Article 59), specific organs are created, namely the notified bodies, ie the entities that will assess the conformity of high-risk AI systems; the notifying authorities, which are the national authorities responsible for the designation, assessment, and monitoring of the notified bodies; and the national supervisory authorities, which may also carry out the tasks of the notifying authorities (Article 59(2) seems to prefer the latter hypothesis).

At the European level, the European Artificial Intelligence Board (EAIB) will be established. This will be a supranational supervisory authority similar to the European Data Protection Board (EDPB), created under the umbrella of the GDPR, and also to the recently proposed European Board for Digital Services, established in the proposal of the Digital Services Act.[19]

RISK ASSESSMENT MODEL

The Draft Act is based on a risk assessment model, under which three categories of AI emerge. First, unacceptable risk AI systems are banned in Title II of the Proposal. This category includes AI that distorts behaviours (dark patterns), explores the vulnerabilities of specific groups (micro-targeting), scores people based on their behaviours (social scoring), and provides real-time biometric identification for law

16 Swedsoft, 'Comments Regarding the European Commission's Proposal for an Artificial Intelligence Act' (24 June 2021), p . 3 <https://www.regeringen.se/49eb04/contentassets/59dff9749d5e4cfa8d51146dd026ff62/swedsoft.pdf> accessed 4 September 2021.

17 This is acknowledged by the Council of the European Union, Proposal for a Regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence. Cf Council of the European Union, '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 - Progress Report' (18 June 2021), 4, <https://data.consilium.europa.eu/doc/document/ST-9674-2021-INIT/en/pdf> accessed 20 July 2021. See also Swedsoft (n 16) 3.

18 'Many associate the term "artificial intelligence" primarily with machine learning, and not with simple automation processes in which pre-programmed rules are executed according to logic-based reasoning' (Ebers and others (n 8) 590).

19 European Commission, 'Proposal for a Regulation of the European Parliament and of the Council on a Single Market for Digital Services (Digital Services Act) and Amending Directive 2000/31/EC' <https://eur-lex.europa.eu/legal-content/en/TXT/?uri=COM:2020:825:FIN> accessed 22 November 2021.

enforcement purposes in public spaces (ie, facial identification). Secondly, high-risk AI systems, as referred to in Title III, are allowed, albeit under particularly strict standards. Thirdly, low-risk AI systems are left more or less unregulated, as established in Title IV. Within this last category, we can distinguish between limited risk (Title IV) and minimal risk (Title IX) systems as follows: AI systems subject to transparency obligations are of limited risk, and of minimal risk are all other AI systems (ie, the majority of those used in the EU), for which suppliers may voluntarily comply with the proposal requirements by creating their own codes of conduct.[20] However, the Draft Act is not clear on this sub-distinction. In addition to these three categories, certain AI systems are not mentioned in the Draft Act and are therefore deemed to be without risk.

The identification of a serious risk is based on the potential threat to health, safety and/or fundamental rights. Notably, the Draft Act is silent on some aspects related to the risk classification, namely who decides the level of risks of a given AI system and when. As for the 'who', the most plausible options are the national supervisory authorities. In terms of the timing, the assessment must not be done too soon; that is, when the AI system is still under development and potential threats remain to be identified. Also, it must not be done too late, just before the AI system reaches the market, as certain requirements based on the level of risk assigned must be met prior to its release. However, the precise moment in the life of the AI system at which the assessment and consequent labelling are to be performed is not defined in the proposal. Moreover, the consequences of erroneous classification of AI systems are not clearly described in the Draft Act. A correction mechanism should be established to amend any misclassification (a reconsideration of the case by the authority that made the first assessment or a kind of 'appeal' to the EAIB); however, the Draft Act remains silent on this issue.

Unacceptable risk systems

Manipulation (dark pattern systems)

AI systems that involve manipulation are described in Articles 5(1)(a) and 5(1)(b) of the Draft Act, the latter describing not only a form of manipulation but also a form of micro-targeting. To clarify the scenarios envisaged by these articles, the EC has presented the following examples: imagine a sound that is inaudible but that affects the behaviour of transport drivers so that they drive longer than is humanly reasonable (allowed) (example of subparagraph (a)); and a toy with an integrated voice assistant that encourages children to adopt dangerous behaviours (example of subparagraph (b)).[21] How real these examples are—that is, how technically feasible they are given the current state of the art—is a question we leave open. However,

20 European Commission, 'New Rules for Artificial Intelligence – Questions and Answers' (21 April 2021) <https://ec.europa.eu/commission/presscorner/detail/en/QANDA_21_1683> accessed 20 July 2021.

21 See Michael Veale and Frederik Zuiderveen Borgesius, 'Demystifying the Draft EU Artificial Intelligence Act—Analysing the Good, the Bad, and the Unclear Elements of the Proposed Approach' (2021) 22(4) Computer Law Review International 97, 100.

there are examples of similar manipulation in real, such as Cambridge Analytica's manipulation of online users for political aims.[22]

The following series of cumulative requirements are used to determine the existence of effective manipulation for the purposes of the prohibition laid out in the Draft Act[23]: The first requirement is the intention to manipulate; subparagraph (a) uses the preposition 'to', whereas subparagraph (b) uses the phrase 'in order to'. Secondly, subparagraph (a) implies the use of subliminal techniques, which must be concealed (specifically, 'that bypass a person's conscience'). Thirdly, subparagraph (b) requires the presence of certain vulnerabilities in the affected people (namely 'age, physical and mental disability'). Fourthly, it is required that the final result of any of these systems is that it 'causes or is likely to cause physical or psychological harm to this or another person'. Note that an effective causal link between the manipulative practice and the harm is not required, as both standards are satisfied by the mere possibility of causing harm.

Some of the manipulative practices covered by this prohibition are already prohibited under national laws and by EU law, such as Article 5 of the Unfair Commercial Practices Directive.[24] Note, however, that this directive applies to a very specific scenario—namely consumer manipulation—whereas Articles 5(1)(a) and 5(1)(b) of the Draft Act refer to any type of manipulation (by public entities).

Article 5 contains a—supposedly absolute—ban on various AI systems considered to be of unacceptable risk. However, a more careful analysis of the rules suggests that the ban contains several loopholes. The ban on manipulation is a paradigmatic example of spurious prohibitions, frequently used throughout the Draft Act. AI systems considered to be manipulative—as defined in Articles 5(1)(a) and 5(1)(b)—are supposedly completely prohibited. However, the way they are defined excludes several scenarios.

First, the norms require manipulative intention on the part of the person or entity that develops, launches in the market or professionally uses these AI systems. However, it is not clear how intent is to be proven if that intention is not declared. Surely few AI providers explicitly state that they use AI for the purpose of manipulating behaviours and emotions. It is not uncommon to find products that are sold for a specific purpose but that everyone, including producers and sellers, knows to be used for other purposes. Furthermore, proving intentions not expressly stated (but only presumed from behaviours) is extremely difficult; thus, it remains to be seen how the authorities will approach this requirement.

22 In this example, not only was manipulation at stake but also profiling, likewise under the purview of the Draft Act. For more details, see Elena Boldyreva, 'Cambridge Analytica: Ethics and Online Manipulation with Decision-Making Process' (2018) 51 The European Proceedings of Social & Behavioural Sciences 91.

23 For more details, see Veale and Borgesius (n 21) 100.

24 Dir 2005/29/EC of the European Parliament and of the Council of 11 May 2005 Concerning Unfair Business-to-Consumer Commercial Practices in the Internal Market and Amending Council Dir 84/450/EEC; Dirs 97/7/EC, 98/2/EC and 2002/65/EC of the European Parliament and of the Council and Regulation (EC) No7 2006/2004 of the European Parliament and of the Council ('Unfair Commercial Practices Directive'). For further information on this directive, see Federico Galli, 'AI and Consumer Manipulation: What Is the Role of EU Fair Marketing Law?' 2020 IV Catolica Law Review 35 <https://doi.org/10.34632/catolicalawreview.2020.9320>.

Moreover, it is not clear whether the harm required by both paragraphs must be caused by a single event capable of producing a manipulative effect by itself or if it can be caused by a sequence of events, none of them individually sufficient to cause such an effect but able to do so when they operate repeatedly over time, as is expected in many cases of manipulative AI systems.[25]

Social scoring

Social scoring systems are prohibited in Article 5(1)(c) of the Draft Act. These are AI systems used by or on behalf of public authorities, intended to generate 'reliability' scores and likely to lead to 'harmful or unfavourable treatment of certain individuals or entire groups of individuals in [...] social contexts' unrelated to those in which the data were originally generated or collected or 'harmful or unfavourable treatment of certain individuals or entire groups [...] unjustified and disproportionate to their social behaviour or the seriousness of it'.

The prohibition does not cover all scoring systems. The norm requires the scoring AI system to operate 'over a certain period of time', thus excluding episodic scoring from the ban. Nonetheless, it is not clear how long 'a certain period of time' is nor why such scoring is more potentially dangerous or presents greater risks to fundamental rights than one-time scoring. Furthermore, Article 5/1/c expressly refers to public authorities (or others acting on their behalf), which excludes scores awarded by non-public authorities as well as their services, such as issuing credit cards or loans (banks), health or professional liability insurance (insurance companies) or compatibility ratings (online dating sites).

Remote biometric identification

Article 5(1)(d) bans 'real-time remote biometric identification systems in publicly accessible spaces for law enforcement purposes', allegedly one of AI's more threatening uses from the perspective of fundamental human rights. Biometric identification involves a wide array of biometric data used to identify a person, such as iris recognition, voice recognition and palm recognition, among others. However, the scenario described in this norm points to a very specific form of biometric identification, namely facial recognition technology (FRT).

The 2020 White Paper did not ban FRT but merely recommended appropriate cautions in its use for remote biometric identification purposes without detailing such recommendations.[26] In contrast, Article 5 of the Draft Act clearly prohibits the use of some forms of FRT.

FRT operates by processing biometric data (ie facial features), a type of processing limited by Article 10 of the Law Enforcement Directive,[27] a directive that is

25 See Veale and Borgesius (n 21) 100.

26 European Commission (n 10) 22. The reactions to this cautionary and vague approach were mixed; however, many supported the use of FRT for remote biometric identification. See European Commission, 'Public Consultation on the AI White Paper Final Report' (November 2020), 11-12, <https://digital-strategy.ec.europa.eu/en/white-paper-artificial-intelligence-public-consultation-towards-european-approach-excellence-and> accessed 22 October 2021.

27 Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the

applicable to the processing of sensitive data in law enforcement scenarios. However, the directive does not forbid FRT but merely imposes certain requirements on its processing, which is required to be strictly necessary, be subject to appropriate safeguards for the rights and freedom of the data subject, have a legal base (prior legal authorization),[28] protect the vital interests of the data subject and of third parties and be restricted to data that are 'manifestly made public by the data subject'. The Draft Act, in contrast, bans certain forms of FRT and imposes additional requirements for the modalities allowed. Thus, it is not a mere repetition of a previous legal solution.[29]

However, critics of this technology may not be appeased by the new solution, as the alleged ban has several exceptions,[30] which is another example of a 'pretended' prohibition. The ban leaves the following loopholes: it does not cover the use of FRT for law enforcement purposes not taking place in real time,[31] that is not carried out in public spaces[32] or that that does not identify people but instead has other purposes (such as identity confirmation, automated recognition of characteristics or the recognition of emotions); it does not cover FRT used by other public entities not related to law enforcement; and it does not cover FRT used by private individuals and companies. Although the criterion for this ban is cited as the potential threat to fundamental rights, such potential threat is also present in these excluded scenarios. Still, most of these other forms of FRT are classified as high-risk AI (Article 6 and Annex III, 1(a)), except for FRT used in emotion recognition systems and biometric categorization systems, which are considered low-risk AI (Article 52(2)).[33]

Moreover, Article 5(1)(d) expressly allows for certain exceptions, arguably justified by potential benefits to crime control.[34] Based on this acknowledgement, the Draft Act permits real-time FRT in public spaces by law enforcement to search for crime victims, including missing children (Article 5(1)(d)(i)); when there is a

purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA ('LED').

28 Note that the Draft Act cannot be invoked as a possible legal basis; as clarified in Recital 23, 'this Regulation is not intended to provide the legal basis for the processing of personal data under Article 8 of Directive 2016/680'.

29 The same conclusion was found by Nikolaos Ioannidis and Olga Gkotsopoulou, 'Pre-Market Requirements, Prior Authorisation and Lex Specialis: Novelties and Logic in the Facial Recognition-Related Provisions of the Draft AI Regulation' (4 April 2021) <https://europeanlawblog.eu/2021/05/04/pre-market-requirements-prior-authorisation-and-lex-specialis-novelties-and-logic-in-the-facial-recognition-related-provisions-of-the-draft-ai-regulation/> accessed 15 November 2021.

30 Criticism of the restricted scope of this ban can be found in the European Data Protection Board and European Data Protection Supervisor, '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)' (18 June 2021), paras 27–35 <https://edpb.europa.eu/our-work-tools/our-documents/edpbedps-joint-opinion/edpb-edps-joint-opinion-52021-proposal_en> accessed 3 September 2021, asking for more bans on AI. This is based on a confusing and repetitive claim that FRT violates human dignity, without identifying which human rights in particular are affected.

31 This is the so-called post remote biometric identification described in art 3(38).

32 As defined in art 3/39.

33 Moreover, both are subject to the requirements of the GDPR and the LED.

34 See Elizabeth McClellan, 'Note, Facial Recognition Technology: Balancing the Benefits and Concerns' (2020) 15 Journal of Business & Technology Law 363, 371–74.

'specific, substantial and imminent threat to the life or physical safety of natural persons, or of a terrorist attack', a vague description potentially applicable to many situations, depending on how strictly or broadly the clause is interpreted (Article 5(1)(d)(ii)); for the detection, localization, identification or prosecution of a perpetrator or suspect of crimes referred to in Article 2(2) of Council Framework Decision 2002/584/JHA (Article 5(1)(d)(iii)).[35] This last exception includes grievous crimes, such as terrorism, trafficking of human beings, murder and rape but also other, less serious crimes such as corruption, fraud and facilitation of unauthorized entry and residency.

As a rule, prior judicial authorization is required for the use of this technology; however, such authorization may be postponed in the case of urgency (Article 5(3)). Therefore, this norm may give rise to abuses only detectable *a posteriori*.

Moreover, it is feared that in the current climate of generalized insecurity, facial recognition cameras will end up being installed in public spaces for preventive purposes.[36] FRT provides many benefits in terms of law enforcement, especially considering that criminals are not faced with legal limitations when looking for ways to accomplish their intents, whose effects can be more harmful than those resulting from the use of this technology. However, in jurisdictions with authoritarian 'mannerisms'—of which some exist in Europe—this can be a dangerous setback for certain fundamental rights.[37]

In their joint opinion, the EDPB and the European Data Protection Supervisor (EDPS) advocated that there should be no exception to the ban on remote biometric identification of individuals in public spaces.[38] However, such a solution would preclude the use of a powerful and useful police investigation mechanism, which we cannot simply disregard. The best solution is to maintain certain bans, limitations, regulations and monitoring but not to impose an absolute ban.

High-risk AI systems

High-risk systems, as described in Article 6 of the Draft Act, are essentially composed of the following two categories of products: products or product components whose regime has been harmonized under EU legislation by the laws listed in Annex II (medical devices, toys, elevators) and independent AI systems that pose special threats to fundamental rights and are listed in Annex III. These lists must be updated according to the latest technological developments.[39] Notably, the label of 'high risk' depends not only on the specific task performed but also on its purpose.

35 Council Framework Decision of 13 June 2002 on the European arrest warrant and the surrender procedures between Member States—Statements made by certain Member States on the adoption of the Framework Decision, 2002/584/JHA <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32002F0584> accessed 22 October 2021.

36 Meneceur (n 1).

37 EDRi, 'Can the EU Make AI "Trustworthy"? No–But They Can Make It Just' (4 June 2020) <https://edri.org/our-work/can-the-eu-make-ai-trustworthy-no-but-they-can-make-it-just/> accessed 15 July 2021.

38 European Data Protection Board and European Data Protection Supervisor (n 30) paras 32 and 33.

39 The European Commission may expand this list, as referred in Article 7 of the Draft Act. Certain relevant but omitted cases have been identified, such as AI systems used for determining insurance premiums, health-related AI systems that are not already covered under Annex II and AI systems deployed for

Two-step procedure

These AI systems are not prohibited; however, their use is subject to various requirements and limitations, in particular a control and monitoring procedure consisting of two stages. The first stage of control is carried out by imposing mandatory requirements on AI systems that must be met before entering the market. For this purpose, Member States are obliged to designate a notifying authority and notified bodies to carry out a third-party conformity assessment,[40] which, if positive, leads to the issuance of a Conformité Européenne (CE) marking of conformity (Article 49). The second control stage consists of an *ex post* monitoring mechanism to be performed by market surveillance authorities, which vary according to the domain in which they act.[41] AI system providers are also required to report serious incidents and malfunctions of their systems.

High-risk AI systems require the CE marking of conformity to enjoy freedom of movement in the EU.[42] This is a logo that can be found on various products circulating on the European market and that attests to compliance with the highest safety, health and environmental protection requirements.

CE marking of conformity requirements

The following series of requirements, largely derived from the Ethics Guidelines formulated by the AI High-Level Expert Group, must be met to obtain the CE marking[43]:

a. Data and data governance: High-risk AI systems should be developed using quality data (ie, the data used to train, validate and test the algorithm). 'Quality' in this context indicates data that are relevant, representative, error-free and complete.
b. Transparency: Providers of high-risk AI systems must disclose adequate information to their users to ensure proper use of the systems, including AI system characteristics, capabilities and limitations; the intended purpose of the systems and care necessary for their maintenance.
c. Human supervision: High-risk AI systems must be designed to be supervised by humans. Supervisory functions include vigilance for automation polarization problems, locating anomalies or signs of malfunctions and deciding whether to override an AI system decision or pull the 'shutdown

housing purposes and predicting policing. See European Data Protection Board and European Data Protection Supervisor (n 30) para 19.

40 European Commission, 'Information from European Union Institutions, Bodies, Offices And Agencies-Commission Notice the "Blue Guide" on the Implementation of EU Products Rules 2016' (2016/C 272/01) <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52016XC0726(02)&from=EN> accessed 10 October 2021.

41 For instance, for AI systems operated by EU institutions, bodies or agencies, the market surveillance authority will be the EDPS (art 63(6)), and for those operated by law enforcement, immigration or asylum authorities, it will be the 'national competent authorities supervising the activities of the law enforcement, immigration or asylum authorities' (art 63(5)), most likely including national data protection authorities.

42 That is, unless the high-risk AI system falls under the exception provided in art 47 of the Draft Act.

43 High-Level Expert Group on AI, 'Ethics Guidelines for Trustworthy AI' (8 April 2019) <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai> accessed 22 May 2021.

switch' if a system poses a threat to safety and security or to the fundamental rights of people.

d. Accuracy, robustness and cybersecurity: The levels of accuracy, robustness and cybersecurity of a high-risk system must be commensurate with the intended purpose of the system. Backup plans should be provided in case of failure to ensure that the system continues to operate and a risks management system must be in place.
e. Traceability and auditability: Suppliers of high-risk AI systems must create and maintain technical documentation containing the information necessary to assess their compliance with the above requirements (see Annex IV of the Proposal). Furthermore, automatic logging of events is mandatory.[44]

If a high-risk AI system on the market is 'substantially modified', it should undergo a new assessment (Article 43(4)). Despite the clarifications provided in Recital 66 on what is a substantial modification, the concept remains unclear.

Self-assessment

As discussed above, high-risk AI systems must obtain the CE marking of conformity to be marketed. However, unlike with other conformity models,[45] the conformity assessment of AI systems can, in most cases, be carried out by the AI providers themselves. That is, except with regard to remote biometric identification systems, which must undergo the more rigorous procedure of third-party assessment, providers may carry out self-assessments.[46] The benefit of this is that the AI provider best knows the product and, consequently, can best certify its conformity. However, there is an undeniable disadvantage to trusting providers to police themselves.

The intervention of notified bodies to perform conformity assessments in high-risk AI systems is insufficient to address this disadvantage. In theory, these bodies will be independent, operating autonomously and transparently (although little is known about how they will perform their functions).[47] However, it is feared that the notified bodies will be limited to private companies, creating what has been called a 'privatized complaint industry'.[48] There is a danger that these entities will only be

44 How to harmonize these obligations with the principle of data minimization, which is set forth in the GDPR, remains to be seen.

45 One paradigmatic case involves pharmaceuticals, whose assessment is carried out by the European Medicines Agency or by national drug authorities.

46 However, third-party assessment can be excluded if the provider of the remote biometric system demonstrates that the system complies with the so-called harmonised standards (art 3(27)) or with the 'common specifications' (art 3(28)), as stated in art 43(1).

47 Jean-Pierre Galland, 'The Difficulties of Regulating Markets and Risks in Europe through Notified Bodies' (2013) 4 European Journal of Risk Regulation 365, 369; Veale and Borgesius (n 21) 110.

48 Daniel Leufer, Digital Policy Analyst in The European Space and Member of the Human Rights Group Access Now, cited in Sebastian Klovig Skelton, 'Europe's Proposed Artificial Intelligence Regulation Falls Short on Protecting Rights' (*Computerweekly.com*, 22–28 June 2021) <https://www.computerweekly.com/feature/Europes-proposed-AI-regulation-falls-short-on-protecting-rights> accessed 5 November 2021, 4–7.

concerned with making the AI systems comply with the formal requirements of the Draft Act, neglecting the defence of fundamental rights.[49]

A double assessment?

Conformity assessments are common for EU products. Indeed, many AI systems are safety components of products subject to this kind of assessment under EU norms. To avoid a double assessment, the Draft Act states that those AI systems will be subject to the same conformity assessment as the product in which they are integrated, the only difference being that such an assessment will consider the new AI requirements in addition to the sectorial legislation in place (Article 63(3)).

However, not all conformity assessments were uniformized by the Draft Act; for example, the assessment of data processing activities was not rendered uniform. The assessment of conformity required by the AI regulation has some similarities to the Data Protection Impact Assessment (DPIA) established in response to various acts on private data (Article 35 GDPR, Article 27 LED). However, the Draft Act does not clarify how they will correlate in practice.[50] The most evident similarity between the two assessments involves the situations in which they are triggered. The DPIA is required for data processing that involves a high risk to fundamental rights, and the assessment of conformity is imposed on high-risk AI systems, which categorization is based, among other criteria, on the risk they pose to fundamental rights (Recitals 27 and 28, Article 7). This represents a substantial overlap of scenarios. As the EDPB and the EDPS state, there is a presumption that a high-risk AI system operating with personal data will also require a DPIA.[51] The conditions that require an assessment may be similar; however, the outcomes may not be. The granting of the CE mark following an assessment of conformity does not imply a positive DPIA result, as the appraisals differ substantially. For consistency, the two assessments should be merged into a single assessment considering all of the relevant criteria. This would avoid conflicting assessments of AI and data protection. Moreover, it would save time and effort and reduce bureaucracy. However, while the assessment of conformity is to be carried out by AI providers, the DPIA is performed by data controllers. These two roles do not always coincide, as many data controllers are AI users, not AI providers.[52]

Low-risk AI systems

Low-risk AI systems receive little attention in the Draft Act. Such systems are not required to undergo the two-step procedure. However, providers can voluntarily do so by creating their own codes of conduct (for minimal risk systems) incorporating some

49 Leufer, quoted in Skelton (n 48) 7.

50 A comparison between the two is presented in Nikolaos Ioannidis and Olga Gkotsopoulou, 'The Palimpsest of Conformity Assessment in the Proposed Artificial Intelligence Act: A Critical Exploration of Related Terminology' (*European Law Blog*, 2 July 2021) <https://europeanlawblog.eu/2021/07/02/the-palimpsest-of-conformity-assessment-in-the-proposed-artificial-intelligence-act-a-critical-exploration-of-related-terminology/> accessed 2 August 2021.

51 European Data Protection Board and European Data Protection Supervisor (n 30) para 21. See also paras 74–76 in the same work.

52 A similar remark appears in European Data Protection Board and European Data Protection Supervisor (n 30) para 21.

of the requirements set forth in the Draft Act. As no system is established to monitor compliance with such codes, it is unlikely that they will have any practical effect.

The few requirements to which low-risk AI systems are subject are mainly limited to transparency obligations imposed to limited-risk systems. This category includes those whose function is to interact with humans, so-called chatbots or simply bots, which must be designed such that people are informed that they are interacting with a machine, unless it is contextually obvious that this is the case. This category also includes AI systems with emotion recognition (the technology that analyses facial expressions from both static images and videos to reveal information about a person's emotional state). AI biometric categorization systems (ie, 'AI systems for the purpose of assigning natural persons to specific categories, such as sex, age, hair colour, eye colour, tattoos, ethnic origin or sexual or political orientation, on the basis of their biometric data') are also in this category. Other systems subject to transparency obligations are deepfakes, ie AI systems that generate or manipulate images or audio or video content that substantially resembles existing people, objects, places or other entities or events and that appear authentic. Providers of these systems are required to disclose the artificial nature of their content. The fight against false news being a major concern, the Draft Act applies to scenarios including both false news and any situation where an alternative reality is created (for educational purposes or marketing).

Problems raised by the risk categorization

It is alleged that the definitions of risk levels focus too much on the respective software, ignoring hardware. For example, AI-based facial recognition software is used via a simple, low-resolution camera, which involves less risk than do high-performance or night-vision cameras.[53] This is an example of how hardware can profoundly shape software performance. However, it is not clear that the Draft Act has taken into account these constraints in its delimitation of risk levels.

Another possible drawback arises from the concept of the 'intended purpose' of AI systems invoked throughout the Draft Act. It is argued that from the consumer's point of view, what matters is not the intended purpose of a system but its foreseeable purpose, which must be evaluated in accordance with the technical and functional characteristics of the AI system, the human behaviour associated to it and its use in conjunction with other products.[54] Therefore, what should guide risk assessments are uses that are reasonably predictable, even if they are not intended uses.

Risk models and fundamental rights

The Draft Act and its risk model strategy received the following criticism based on the argument that fundamental rights cannot be treated and evaluated on an equal footing with 'company interests' (referring to profit interests): 'Our rights, they are non-negotiable and they must be respected regardless of a risk level associated with

53 European Consumer Voice in Standardisation (ANEC), 'ANEC Comments on the European Commission Proposal for an Artificial Intelligence Act', Position Paper, July 2021 <https://www.anec.eu/images/Publications/position-papers/Digital/ANEC-DIGITAL-2021-G-071.pdf> accessed 10 November 2021.

54 ibid.

external factors.'[55] This view, however, fails to acknowledge that AI is not merely a business opportunity and a source of profit but also a potential way to improve our quality of life and even to fulfil fundamental rights.

The so-called risk-based approach is commonly used in regulations. Its presence (or absence) in the GDPR is widely discussed. There are those who claim that the GDPR is based on the protection of rights and not on an analysis of risks.[56] This perspective is allegedly supported by the very wording of the Article 89 Working Party's (A29WP) 'Declaration on the Role of a Risk-based Approach in Legal Data Protection Frameworks'.[57] However, in this document, the A29WP does not reject the risk-based approach in the GDPR but rather clarifies the correct understanding of this approach as follows: 'the risk-based approach is being increasingly and wrongly presented as an alternative to well-established data protection rights and principles, rather than as a scalable and proportionate approach to compliance.'[58] Rights must always be weighed with respect to any legislation. However, 'legislative action related to AI, robotics and related technologies must comply with the principles of necessity and proportionality.'[59] The A29WP explained that the qualification of a technology or procedure as low risk does not diminish the legal protection of the rights involved; it simply mitigates the obligations of those who control that technology.[60] By this, the A29WP was referring to data protection and the obligations of the person or entity responsible for the processing of data; however, the same reasoning holds, *mutatis mutandis*, for the legal solutions of the Draft Act and AI systems.

ACCOUNTABILITY IN THE DRAFT ACT

Who will be held accountable?

A powerful instrument for the protection of fundamental rights is stakeholders' responsibility, which functions through the imposition of duties and rules of conduct and the accountability imposed for non-compliance. The Draft Act imposes tasks and duties on several players (Article 2(1)).[61]

First, the act imposes duties on AI providers, ie those who develop an AI system or have someone develop an AI system for them, even if they are not established in

55 Fanny Hidvegi and others, 'The EU Should Regulate AI on the Basis of Rights, not Risks' (17 February 2021) <https://www.accessnow.org/eu-regulation-ai-risk-based-approach/> accessed 5 May 2021. Note that these authors' comments refer to the White Book, as the text refers to the period before the proposal was public; however, their criticisms also apply to the proposal itself.

56 ibid.

57 art 29 Data Protection Working Party, 'Statement on the Role of a Risk-Based Approach in Data Protection Legal Frameworks' (30 May 2014) <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp218_en.pdf> accessed 4 June 2021.

58 ibid 2.

59 Point 4 of European Parliament resolution of 20 October 2020 with recommendations to the Commission on a framework of ethical aspects of artificial intelligence, robotics and related technologies (2020/2012 (INL)) <https://www.europarl.europa.eu/doceo/document/TA-9-2020-0275_EN.html>, accessed 6 June 2021.

60 art 29 Data Protection Working Party (n 58) 2.

61 Strangely, art 2(4) excludes the application of the Draft Act to international organizations and foreign (non-European) public authorities acting under international agreements for law enforcement and judicial cooperation with the EU or national Member States, which may give rise to impunity.

EU territory. Their obligations are as follows: the obligations to use quality datasets to train, validate and test the AI system; to monitor the system and guarantee human oversight; to design a robust and accurate AI system with an appropriate level of cybersecurity; to carry out a conformity assessment (self-assessment or third-party assessment) in the case of high-risks AI systems; to register the AI system in the AI database, in the case of the AI systems referred to in Annex III; to implement risk management and quality management mechanisms; to maintain detailed technical documentation and automatically generated logs; to comply with transparency obligations and to carry out continuous monitoring on the performance of the AI system.

AI users—ie natural or legal persons,[62] who use AI professionally (thus, excluding personal activities), and who are either established in the EU or not established in the EU but whose outputs of AI systems are used in the EU[63]—also have duties and functions. They are required to use the AI system in accordance with the instruction of the AI provider, monitor the system performance and promptly detect any malfunction, in which case the operation must be suspended, and the incident reported.

In this main list of stakeholders, we must also include those who place AI systems on the EU market—ie importers, distributors, product manufacturers and authorized representatives—if they commercialize the system under their own name or trademark. They are required to confirm that the AI system has an appropriate CE conformity marking and that it is commercialized with proper documentation and instructions. If such players change the intended purpose of the AI system or make substantial modifications to it (Article 28(1)), they are considered to be the AI system's providers and the original provider ceases to be considered as such (Article 28(2)).

In the event of non-compliance, the Draft Act provides sanctions of varying severity, the strictest of which is a 'fine of up to 30 000 000 EUR (thirty million euros) or, if the offender is a company, up to 6% of its total worldwide annual turnover for the preceding financial year, whichever is higher', as referred in Article 71(3) (note the strong similarity with the regime foreseen in the GDPR).

Loopholes in the accountability dimension

The Draft Act is silent on two critical topics related to accountability. First, it does not establish a right to take legal action against suppliers or users of AI systems for non-compliance with its rules[64] (although non-compliance with these rules can be invoked in a civil liability proceeding). This limitation has been repeatedly pointed out both by European authorities (such as the EDPB and the EDPS)[65] and by civil rights groups.[66]

Additional accountability mechanisms should include civil liability for defaulters; however, specific compensation mechanisms are absent from this proposal. Producer

62 'Users' encompasses public authorities and agencies, including EU institutions, bodies or agencies that may be data users or data providers.

63 art 2 refers to the location of AI providers and AI users but not the location of AI systems. It seems that the Draft Act only applies to AI systems located in the EU (the same interpretation appears in Ebers and others (n 8) 591) but not to those that, while developed in the EU, are never used in the EU and whose outputs have no effects in the EU. However, the issue is unclear.

64 Skelton (n 48) 11.

65 European Data Protection Board and European Data Protection Supervisor (n 30) para 18.

66 Skelton (n 48) 11.

liability is currently regulated in Directive No 85/374/EEC[67] on liability arising from defective products, which is currently in the process of being revised to adapt it to new challenges of the digital world.[68] It is expected that suitable solutions for the use of AI in various products will appear in the revisions. Other relevant documents are the 2019 Report on Liability for Artificial Intelligence and Other Emerging Digital Technologies[69] and the European Parliament resolution of 20 October 2020 with recommendations to the Commission on a civil liability regime for AI (2020/2014(INL))[70]; however, these are mere recommendations/guidelines and not enforceable regulations.

There is an additional problem not addressed in the Draft Act that is unrelated to stakeholder liability but instead concerns potential liability of the AI system itself. This has been increasingly discussed in the literature and is based on the premise that AI systems are legal persons, also a widely discussed topic.[71] The fact that the Draft Act remains silent on the issue of AI personhood can be taken as a stand in itself (that AI systems are not legal persons nor can they be held accountable) or simply as the postponement of a resolution of an extremely complex question.

TRANSPARENCY IN THE DRAFT ACT

Transparency obligations

'Transparency' is a keyword of the Draft Act.[72] Stakeholders in the AI domain have the following obligations regarding transparency: to keep technical documentation (Article 11), to maintain records to ensure the traceability of operations (Article 12) and to provide appropriate information to users (Article 13). Such obligations are reinforced in the case of the following AI systems: those that interact with humans, such as chatbots; those used to detect human emotions (eg, in job interviews or in advertising); and those that manipulate content to control human behaviour (namely deepfakes). The obligation to register high-risk system in a public database (Articles 51 and 60 of the Draft Act) is also a reflex of the transparency target.

Such obligations are not exclusive to the Draft Act, rather they are repeatedly invoked by other legal proposals and by courts. For instance, a recent Belgian

67 Dir (EU) 2019/878 of the European Parliament and of the Council of 20 May 2019 amending Dir 2013/36/EU as regards exempted entities, financial holding companies, mixed financial holding companies, remuneration, supervisory measures and powers and capital conservation measures.

68 About the revision of this Directive, see European Commission, 'Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions on Artificial Intelligence for Europe', Brussels, 25.4.2018 COM(2018) 237 final <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2018%3A237%3AFIN> accessed 22 October 2021.

69 Expert Group on Liability and New Technologies, *Liability for Artificial Intelligence and Other Emerging Digital Technologies* (Publications Office of the European Union 2019) <https://op.europa.eu/en/publication-detail/-/publication/1c5e30be-1197-11ea-8c1f-01aa75ed71a1/language-en/format-PDF> accessed 12 May 2021.

70 European Parliament, Resolution of 20 October 2020 with recommendations to the Commission on a civil liability regime for artificial intelligence (2020/2014(INL)) <https://www.europarl.europa.eu/doceo/document/TA-9-2020-0276_EN.html> accessed 4 June 2021.

71 L Floridi and M Taddeo, 'Romans Would Have Denied Robots Legal Personhood' (2018) 557 (7705) Nature 309.

72 Recitals 14, 38, 39, 43, 47, 69 e 70; arts 1(c), 13 and 52.

legislative proposal on AI requires transparency in the use of algorithms by public authorities (law proposal from 6 April 2021),[73] stating that its *modus operandi* should be disclosed online, especially when the algorithms are used for individual decisions. Affected individuals are entitled to receive additional information, such as how the algorithm contributed to the decision, which data were processed and which operations were performed during processing.[74] Likewise, the Italian Supreme Court (Corte di Cassazione No 14381/2)[75] recently concluded that service providers (in this case, reputation-rating services) are required to reveal to their customers the algorithms used. Furthermore, customers not only must consent to the processing of their data to obtain the service but must also have been previously informed about the functioning of the algorithm.

Limitations on transparency

The Draft Act recognizes that in the field of AI, transparency obligations may face limitations arising from the protection of intellectual property rights. Therefore, information obligations are 'limited only to the minimum necessary information for individuals to exercise their right to an effective remedy and to the necessary transparency towards supervision and enforcement authorities, in line with their mandates' (Exploratorium Memorandum).[76]

In addition, under Article 52 of the Draft Act, transparency obligations do not apply to AI systems authorized by law to detect, prevent, investigate or prosecute crimes, unless they are emotion recognition systems, in which case disclosure of their use is always mandatory (for instance, when police forces use facial recognition in undercover operations or when police or courts interrogate suspects).[77]

Moreover, it is difficult to comply with transparency obligations with respect to AI, intrinsic characteristics of which are opacity and extreme complexity (hence its characterization as a 'black box').[78] One might argue that most individuals affected by AI are unable to comprehend information about it. Thus, the requirement of transparency cannot be understood as requiring users of AI systems to have a precise understanding of how they work. Rather, this requirement should be limited to

73 Proposition de Loi - modifi ant la loi relative à la publicité de l'administration du 11 avril 1994 afi n d'introduire une plus grande transparence dans l'usage des algorithmes par les administrations, 6 April 2021 <https://www.lachambre.be/FLWB/PDF/55/1904/55K1904001.pdf>, accessed 11 July 2021.

74 See Joshua Gacutan and Niloufer Selvadurai, 'A Statutory Right to Explanation for Decisions Generated Using Artificial Intelligence' (2020) 28(3) International Journal of Law and Information Technology 193, <https://doi.org/10.1093/ijlit/eaaa016>.

75 Federica Paolucci, 'Consenso, Intelligenza Artificiale e Privacy. Commento a: Corte di Cassazione', sez. I Civ. – 25/05/2021, n. 14381 (16 June 2021) <https://www.medialaws.eu/consenso-intelligenza-artificiale-e-privacy-commento-a-corte-di-cassazione-sez-i-civ-25-05-2021-n-14381/> accessed 7 July 2021.

76 European Commission (n 13) 13.

77 In its joint opinion, the EDPB and the EDPS argue that this exception should not be valid to the detection and prevention of crimes, due to the presumption of innocence (European Data Protection Board and European Data Protection Supervisor (n 30) para 70).

78 Tom Cassauwers, 'Opening the "Black Box" of Artificial Intelligence' *Horizon-The EU Research and Innovation Magazine* (1 December 2020) <https://ec.europa.eu/research-and-innovation/en/horizon-magazine/opening-black-box-artificial-intelligence> accessed 22 October 2021.

providing an understanding of the main limitations of AI systems and identifying their shortcomings.[79]

INNOVATION IN THE DRAFT ACT

Lack of measures to boost innovation

There are only two measures in the Draft Act that promote innovation. One concerns small-scale AI providers and users, and the other is the creation of regulatory sandboxes to encourage innovation, seemingly the most promising initiative. Two measures, establishing no more than three norms, are clearly insufficient to boost innovation.

An additional shortcoming of the Draft Act in terms of promoting innovation is the absence of a specific legal mechanism aimed at the use of AI in research, in parallel terms to Article 89 of the GDPR.[80] That is, the mere development of an AI system for academic purposes, even if not in cooperation with the industry, will be subject to the demanding requirements of this regulation.

Regulatory sandboxes

The word 'sandbox' originally referred to an enclosure filled with sand, in which children play and experiment in a controlled environment. However, the term has acquired new meanings. A regulatory sandbox is a framework set up by a regulator that allows start-ups and other innovators to conduct live experiments in a controlled environment under a regulator's supervision (much like a clinical trial for new pharmaceuticals). Regulatory sandboxes make it possible to test new technologies transparently and contribute to evidence-based law-making.

Regulatory sandboxes carry advantages for AI developers and manufacturers, such as additional regulatory flexibility and an operational legal framework for AI systems that do not fit well within the general legal framework.[81] However, they also pose certain challenges that will most likely be felt in the AI domain. For example, the goal of collecting further data on how these disruptive technologies affect society may be frustrated if such a sandbox is restricted to a few projects (as is usual), as a small sample does not allow for the anticipation of potential legal or ethical hazards. The unregulated nature of the sandbox is a challenge, *per se*, due to its potential derogation of applicable norms. For instance, in the case of AI systems, concerns have been raised as to whether AI systems in a regulatory sandbox would still be bound by data protection laws (notably, the GDPR). In my view, the flexibility of such a sandbox should involve the part of the regulation concerning AI and not the part concerning data protection. This conclusion results from an interpretation of the Draft Act, which specifically clarifies that the purpose of the regulatory sandboxes

79 See Eve Gaumond, 'Artificial Intelligence Act: What Is the European Approach for AI?' (*Lawfare*, 4 June 2021) <https://www.lawfareblog.com/artificial-intelligence-act-what-european-approach-ai> accessed 4 July 2021.

80 The same remark appears in Ebers and others (n 8) 591.

81 Florina Pop and Lukas Adomavicius, 'Sandboxes for Responsible Artificial Intelligence', (*EIPA Briefing*, September 2021) <https://www.eipa.eu/sandboxes-for-responsible-artificial-intelligence/> accessed 2 December 2021.

is to ensure compliance of innovative AI system with the Regulation and other relevant Union and Member States legislation, thus, including data protection legislation[82] (Recital 72). Moreover, Article 53(3) of the Draft Act expressly states, 'The AI regulatory sandboxes shall not affect the supervisory and corrective powers of the competent authorities', thus including the powers of data protection authorities. Therefore, it is not plausible that sandboxes will derogate established mandatory norms.[83]

CONFLICTS BETWEEN NORMS

With several European norms regulating the same matter, it is only natural that conflicts will arise between rules. The Draft Act is part of a much broader project that aims to bring the EU into the digital world. For this purpose, several legal drafts are currently in the pipeline, namely the regulation on digital services,[84] the regulation on digital markets,[85] the regulation on data governance[86] and the Machinery Regulation, released together with the Draft Act.[87] In addition, a lengthy list of existing legislation is equally relevant to the field of AI, including the GDPR, the LED, the EUDPR,[88] the Directive on privacy and electronic communications,[89] the Regulation on medical devices,[90] the Regulation on *in vitro* diagnostic devices[91] and the Defective Products Directive (soon to be revised), just to name a few.

It will be almost impossible to avoid overlaps between all of these legal documents. Even if the norms do not conflict, a challenge remains in that it will be almost impossible to take them all into account when acting or making decisions with respect to AI. As a concrete example, consider that AI systems work with data, most of which are personal data; thus, it will be necessary to articulate the GDPR with the AI

82 Also note that the GDPR does not provide for regulatory sandboxes.

83 Pop and Adomavicius (n 82).

84 Proposal for a Regulation of the European Parliament and of the Council on a Single Market for Digital Services (Digital Services Act) and amending Dir 2000/31/EC <https://eur-lex.europa.eu/legal-content/en/TXT/?uri=COM:2020:825:FIN> accessed 10 November 2021.

85 'Proposal for a Regulation of the European Parliament and of the Council on Contestable and Fair Markets in the Digital Sector (Digital Markets Act)' <https://eur-lex.europa.eu/legal-content/en/TXT/?uri=COM%3A2020%3A842%3AFIN> accessed 10 November 2021.

86 'Proposal for a Regulation of the European Parliament and of the Council on European Data Governance (Data Governance Act)' <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020PC0767> accessed 10 November 2021.

87 'Proposal for a Regulation of the European Parliament and of the Council on Machinery Products', COM/2021/202 final <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0202> accessed 12 November 2021.

88 Regulation (EU) 2018/1725 of the European Parliament and of the Council of 23 October 2018 on the protection of natural persons with regard to the processing of personal data by the Union institutions, bodies, offices and agencies and on the free movement of such data, and repealing Regulation (EC) No 45/2001 and Decision No 1247/2002/EC.

89 Dir 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector.

90 reg (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Dir 2001/83/EC, reg (EC) No 178/2002 and reg (EC) No 1223/2009 and repealing Council Dirs 90/385/EEC and 93/42/EEC.

91 reg (EU) 2017/746 of the European Parliament and of the Council of 5 April 2017 on *in vitro* diagnostic medical devices and repealing Dir 98/79/EC and Commission Decision 2010/227/EU.

regulation, which is no simple task, as both include very detailed rules, namely very detailed assessments (the DPIA under the GDPR and the conformity assessment under the AI Regulation).[92]

CONFLICTS BETWEEN CERTAIN BODIES AND INSTITUTIONS

Likewise, conflicts may arise between jurisdictions. As AI involves administrative, criminal and civil matters, several courts, in theory, will be able to settle disputes under the new regulation. It is expected that some of the 'competence collisions' that currently take place within the scope of the GDPR between various courts, will be repeated in the context of the future AI regulation.[93]

Conflicts may also arise between the bodies and authorities created by the Draft Act and existing bodies and authorities. For example, the Draft Act establishes notified bodies to assess the conformity of certain high-risk AI systems before they are placed on the market. Notified bodies are also present in the two medical devices regulations and perform the same task of conformity assessment. It is not clear whether the two types of notified bodies will perform similar tasks under different legal frameworks or the medical devices notified bodies will also perform AI system assessment. The question is particularly relevant because many AI systems will be also considered medical devices, introducing the possibility of dual conformity assessment; that is, the same device being assessed twice. The use of a single notified body, with the competence and expertise to make both assessments, seems easier, faster and cheaper.[94] However, it is not clear how similar these two assessments will be.

A similar conflict may arise between the EAIB and its twin supervisor, the EDPB,[95] which, until now, has handled certain aspects of AI in its duties concerning data protection.

At the national level, battles may develop between data protection authorities and the national supervisory authority on AI over their respective authority. Some national data protection entities have made it known that they consider themselves to be best suited to act as supervisory entities for AI systems due to their experience in this type of control, as AI systems require data to be trained and therefore frequently raise data protection issues.[96] Moreover, it would be possible to achieve greater

92 In its joint opinion, the EDPB and the EDPS expressed particular concerns about the articulation of the AI regulation and the existing (and vast) data privacy legal framework (European Data Protection Board and European Data Protection Supervisor (n 30) para 15).

93 Meneceur (n 1).

94 For an argument in favour of expanding the competence of the notified bodies acting in the field of medical devices, see Team NB, 'European Artificial Intelligence Regulation, Team NB Position Paper', 6 October 2021 <https://www.team-nb.org/wp-content/uploads/2021/10/Team-NB-PositionPaper-Artificial-Intelligence.pdf> accessed 3 December 2021.

95 Showing some concerns in this regard, the EDPB and EDPS called for a ban in their joint opinion, where it is argued that the regulation 'should equally clarify that the Proposal does not seek to affect the application of existing EU laws governing the processing of personal data, including the tasks and powers of the independent supervisory authorities competent to monitor compliance with those instruments' (European Data Protection Board and European Data Protection Supervisor (n 30) para 15).

96 Commission Nationale de l'Informatique et des Libertés (CNIL), 'Artificial Intelligence: The Opinion of the CNIL and Its Counterparts on the Future European Regulation' (8 July 2021) <https://www.cnil.fr/en/artificial-intelligence-opinion-cnil-and-its-counterparts-future-european-regulation> accessed 4 November 2021.

consistency between the solutions provided by the AI regulation and those provided by the GDPR (as well as the LED Directive, which is also within the purview of the national data protection authorities). It is also argued that establishing a single authority for both domains would avoid the multiplication of supervisory authorities, increase the effectiveness of the regulatory model and reduce legal uncertainty and administrative complexity.

Without contesting the advantages of centralizing the powers and functions of national data protection entities, it nonetheless seems that the reasoning underlying data protection does not entirely coincide with the reasoning necessary to monitor AI systems. That is, AI raises many issues beyond those typical of data protection. However, the recognition of this specificity will inevitably bring about overlaps in authority.

PRELIMINARY CONCERNS ABOUT THE DRAFT ACT

Despite being an important step for the EU in the technological domain, all is not 'rainbows and butterflies' in the AI Draft Act. First, in some regards, it is not enough. It is intended to be a comprehensive regulation on AI; however, it falls short of that goal. A detailed, comprehensive regulation would certainly be impossible to create, as AI is such an extensive and complex domain. However, there are glaring omissions in the act, among them its treatment of liability for damages.

Second, in other respects, it is too much, as certain matters are hyper-regulated by the Draft Act. European AI developers and manufacturers are asked to comply with an overwhelming number of requirements, some of them so demanding that they may be impossible to comply with. For example, Article 10/3 stipulates that 'training, validation and testing datasets must be relevant, representative, error-free and complete'. Experts point out that the idea of a completely error-free dataset is utopian.[97] Another difficulty of complying with this demand lies in its ambiguity.[98] Does the 'error' in question refer to the set of data, its classification, the way the intended behaviour is represented, all of these aspects or other aspects? Furthermore, who will assess data quality and using what criteria?

Finally, the absence of a substantial boost to innovation in the Draft Act is a significant concern. The aim to bring about innovation in AI technologies is restricted to three articles, which, although introducing interesting provisions, fall short of what is required for a truly digital single market. Furthermore, there are only two measures in the Draft Act promoting innovation. AI investment in the EU may be hampered by such an 'innovation hole', which could advantage other leading players.[99] A

This claim is supported by the EDPB and the EDPS (European Data Protection Board and European Data Protection Supervisor (n 30) para 47).

97 Floridi and Taddeo (n 72) 309; Siemens, 'EU's AI Regulation Proposal (21/04/2021) Position & Recommendations' (July 2021), 3 <https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12527-Artificial-intelligence-ethical-and-legal-requirements/F2662941_en?fbclid=IwAR1W6QQXqsZgrCkZM5Z9W22JVFg0RDrh-gc4dFlXvg8z-z-fAhAQi66y39Y> accessed 25 August 2021.

98 Siemens (n 97) 6.

99 This is already a problem, as only 6 of the top 100 AI start-ups worldwide are based in Europe. Cf. CB Insights, 'AI 100: The Artificial Intelligence Startups Redefining Industries' (7 April 2021) <https://www.cbinsights.com/research/report/artificial-intelligence-top-startups/> accessed 10 December 2021.

definition of standards to be adopted by the rest of the world—an expression of the so-called Brussels effects[100]—may not be appropriate for the AI regulation, as most countries would prefer a model that is more balanced between fundamental rights and technological development. Ultimately, if the current proposal indeed becomes the new regulation, Europe may be inescapably relegated to the tail end of the digital revolution.

100 Anu Bradford, *The Brussels Effect. How the European Union Rules the World* (OUP 2020).