Thomas Burri and Fredrik von Bothmer, The New EU Legislation on Artificial Intelligence: A Primer, available on SSRN: https://ssrn.com/abstract=3831424 – version of 21 April 2021 (we may update the text).

The New EU Legislation on Artificial Intelligence: A Primer

Thomas Burri and Fredrik von Bothmer[1]

On the following pages, the Commission's proposal of 21 April 2021 of new Union legislation to regulate artificial intelligence (AI)[2] *is explained and discussed. The proposal for a Union regulation only marks the beginning of the legislative process. The Council and the European Parliament will most likely modify the proposal substantially.*

Overall, the Commission's proposal is clever, well written and balanced. Its scope is not overly broad, for it focuses on the most problematic uses of AI and contains only minimal obligations with regard to other AI. The proposal undoubtedly marks a defining moment in the history of AI, since it will ultimately lead to the first comprehensive legislative measure globally containing binding rules on AI. The proposal controversially bans certain uses of AI, but it also contains a whole host of new duties for those who put into circulation what the proposal designates as "high-risk AI". The proposal carefully attempts to avoid posing an undue administrative burden on market actors. Obviously, some of the lessons from the General Data Protection Regulation have been learned. Nevertheless, the proposal also marks the advent of the regulatory state (and the regulatory Union) in AI and the end of unbridled freedom in all things AI.

The proposed regulation distinguishes three categories of AI, namely certain uses of AI that it bans, high-risk AI which it regulates in detail, and low-risk AI which it addresses to a limited extent. Since the first two categories are relatively narrowly circumscribed, the vast majority of existing intelligent algorithms falls into the third category where the proposed regulation essentially requires AI to be flagged.

---

[1] Thomas Burri, Professor of International Law and European Law at the University of St. Gallen (Switzerland), Dr. iur. (Zurich), LL.M. (College of Europe, Bruges), admitted to the bar in Switzerland (corresponding author – contact: thomas.burri@unisg.ch). Fredrik von Bothmer, Dr. iur. (St. Gallen), LL.M. (Fletcher), Manager Human Rights in Social Compliance at Daimler AG (contact: fbothmer@icloud.com); this text was written in a purely personal capacity and does not represent the view of Daimler AG or engage it in any way.

[2] 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, 21 April 2021, available at: https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-european-approach-artificial-intelligence.

The proposed regulation prohibits certain uses of AI. It bans the use of AI: a) to materially distort a person's behaviour; b) to exploit the vulnerabilities of a specific group of persons; c) public social scoring and d) for real time remote biometric identification in public places.[3] The ban is limited to the use of AI ("AI practice"). It does not preclude the development of AI which could potentially be used in cases covered by the ban. Given the general-purpose character of AI, this limitation is essential.

The material distortion of behaviour (a) and the exploitation of vulnerabilities (b) are only banned, if they causally affect a human person's physical or psychological harm, or if they are likely to do so.[4] This condition qualifies the ban and makes it narrower. But the ban is still rather vague and broad with regard to these two categories. The terms used in (a) and (b) are not standard legal terms. No case law exists that could indicate how they are to be understood. Some ill intent, maybe fraudulent or deceptive intent, seems implicitly required. Potentially, the prohibition would cover widely used tools, such as social media bots masking as regular users with a view to influencing voters, especially if the psychological harm caused were to be construed widely. It is doubtful whether such tools, whilst of ill repute, need necessarily be banned.

Importantly, the ban of biometric recognition is not absolute. Derogations from it are possible broadly speaking in case of public security concerns. Article 5(1)(d) specifically lists the search for individual potential victims of crimes; the concrete looming threat to life or physical safety of human persons or a terrorism threat; and the finding or prosecution of suspects or criminals in case of certain crimes as grounds for legitimate derogation. *E contrario,* grounds other than those mentioned cannot justify derogations from the ban. This is noteworthy, because derogations are often framed more broadly. Restrictions of fundamental and human rights, for instance, can regularly be justified on grounds other than the protection of public security. The ban of biometric recognition is furthermore subject to an assessment of the circumstances of the situation, the potential harm as well as necessity and proportionality (including temporal, geographic and personal limitations) according to article 5(2), and an authorization to be granted by an independent national authority.[5] Given this framing, the derogation from the ban of biometric recognition appears appropriately narrow (if an exception from the ban is considered desirable in the first place). Yet it is not entirely clear whether the Union has the

[3] See article 5(1)(a)-(d) proposed regulation.

[4] "[…] in a manner that causes or is likely to cause that person […] physical or psychological harm" (article 5(1)(a) and (b)). The difference between letters (a) and (b) seems to be that under letter (a) a "subliminal technique beyond a person's consciousness" is deployed, while for persons who are vulnerable due to age or disability no such subliminal technique is required. For the latter a material distortion of behaviour likely causing damage is sufficient.

[5] For a slightly broader law enforcement derogation according to national law, see article 5(4).

power to ban biometric recognition (see *infra*). But if it does, it should also have the power to regulate the details of the exception to the ban, even if it is on the ground of public security in a broad sense.

The most detailed rules of the proposed regulation apply to high-risk AI, a notion which the regulation construes rather broadly. Firstly, AI within the scope of the Union harmonization legislation listed in annex II (as a product itself or part of another product) and subject to a conformity assessment by a third party according to that legislation is considered to be high-risk AI. This *inter alia* includes toys, lifts, appliances burning gaseous fuel, and medical devices.[6]

Secondly, annex III of the proposed regulation lists other intelligent algorithms that represent a high risk within the sense of the regulation, namely AI intended to be used in: remote biometric identification systems (assuming a derogation from the ban in article 5), critical infrastructure, educational institutions (including access to and performance within such institutions), employment (including recruitment and promotion), the award of assistance (essential private services and public services and benefits, including eligibility, creditworthiness, and emergency first response services), law enforcement and criminal law (including certain aspects of sanctioning such as prediction of recidivism, predictive policing, detection of deep fakes, determination of reliability of evidence, profiling, and crime analytics), management of migration, asylum, and border control, as well as assistance of judicial authorities. For all AI listed in annex III, the use intention is the factor that determines the qualification as high-risk. This list of high-risk AI will likely be debated controversially and may be changed by the Council and the Parliament.

The category of high-risk AI is not only comprehensive from a substantive perspective, since it includes most intelligent algorithms that proved controversial in recent times, article 6 including annexes II and III of the proposed regulation is also to be considered legally exhaustive. The only possibility to expand the list is for the Commission to exercise the power it is delegated by the regulation to add other high-risk AI falling broadly within the domains listed in annex III.[7] This power, which the Commission exercises jointly with the Council and the Parliament, should exclude any possibility for the Court of Justice of European Union to apply the

[6] See article 6 proposed regulation. Further Union legislation is included, namely that applicable to machinery (as revised in parallel), watercraft, equipment to be used in potentially explosive atmosphere, radio and pressure equipment, and in vitro medical devices; for the specific legislation covered, see Annex II.

[7] See articles 7 and 73 proposed regulation. According to article 7 proposed regulation, adding further high-risk AI depends, broadly speaking, on the harm an AI could cause.

provisions concerning high-risk AI to other algorithms. Even so, the Court has handled some annexes flexibly in the past.[8]

The qualification of an AI as high-risk gives rise to a number of obligations, which essentially must be met by what the proposed regulation calls the "provider"[9] when the AI is brought into circulation.[10] According to the regulation, *inter alia* a comprehensive risk management system must be established;[11] the quality of data (for training, validation and testing) and data governance must be ensured;[12] technical documentation and logs must be kept and transparency established;[13] human oversight is to be guaranteed through design or by enabling users;[14] accuracy, robustness, and cybersecurity are required;[15] and high-risk AI, if it is a "stand-alone" AI, i.e. not an AI integrated into other regulated products, must be registered in a newly established public database[16]. Users – a term which excludes those who make use of AI in a personal, non-professional capacity[17] – are subject to obligations too. Namely, they are bound by the provider's instructions and must monitor the high-risk AI they use.[18] Third parties that modify a high-risk AI automatically assume the obligations incumbent upon the provider.[19] A conformity assessment, which aims to ensure that the obligations contained in the proposed regulation are complied with, must be completed by the provider itself. Only in certain cases in which biometric recognition is used lawfully, an external assessment by a third body according to the proposed regulation is required.[20]

While the obligations on providers and others are quite onerous, they only arise with high-risk AI. Although this category includes a number of algorithms, the vast majority of AI falls into

[8] Notably in the coordination of social security, i.e. the application of the Regulation on the coordination of social security systems, European Parliament and Council, Regulation (EC) No. 883/2004 (O.J. L 166 of 30 April 2004, p. 1-123, repeatedly amended), 29 April 2004 (and the preceding regulations).

[9] Article 16 proposed regulation; a provider is defined in article 3(2) as follows: "a natural or legal person, public authority, agency or other body that develops an AI system or that has an AI system developed with a view to placing it on the market or putting it into service under its own name or trademark, whether for payment or free of charge".

[10] Some of the obligations incumbent on providers are shifted to importers and distributors in certain constellations pursuant to articles 26 and 27 proposed regulation as well as other third parties according to article 28.

[11] Article 9 proposed regulation.

[12] Article 10 proposed regulation.

[13] Articles 11-13 proposed regulation.

[14] Article 14 proposed regulation which includes important requirements, such as the possibilities to avoid automation bias, intervene in the operation, and correctly interpret output.

[15] Article 12 proposed regulation.

[16] Articles 51 and 62 proposed regulation.

[17] See article 3(4) proposed regulation.

[18] Article 29 proposed regulation.

[19] Article 28(1) proposed regulation; this obligation of third parties may have ramifications for AI based on open source code.

[20] Article 43 proposed regulation.

the category of low-risk algorithms ("certain AI systems"[21]), with regard to which the proposed regulation imposes only limited transparency and disclosure requirements. Notably, natural persons must be informed, when an AI is part of an interaction they are having; emotion recognition and biometric categorization systems must be flagged as such vis-à-vis natural persons; and users who create "deep fakes" must disclose the nature of the output, unless it is used for purposes of law enforcement or, for instance, in satire or comedy.[22]

Apart from the categorization into banned uses of AI, high-risk and low-risk AI, which specifies the substantive purview of the legislation, the scope of the proposed regulation is both wide in certain respects, while being narrow in others. The regulation adopts the expansive approach to the geographical scope which has been characteristic of the General Data Protection Regulation[23]. Specifically, the proposed regulation applies when providers established outside of the Union bring AI into circulation in the Union. It even applies when output produced by an AI is used in the Union, even if all other relevant facts (notably providers and users) are limited to third countries.[24] In addition, providers established outside of the Union must designate an authorized representative in the Union for purposes of compliance with the regulation.[25]

The substantive scope of the proposed regulation is limited by the well-balanced definition of AI which is contained in annex I (and the reach of the categories of AI the proposed regulation addresses). The definition does not simply cover all algorithms.[26] It may well be the first legally binding definition of AI (if the proposal becomes law). But the substantive scope is also otherwise limited. The proposed regulation does not on the whole apply, for instance, to civil aviation or cars that incorporate safety components using AI.[27] AI that is used exclusively for military purposes, including in autonomous weapons systems, is also excluded from the scope of the proposed regulation.[28] This latter exclusion is necessary, as the Union lacks the power to regulate the military matters. Moreover, the exclusion is not overly broad, since the proposed

[21] Article 52 proposed regulation

[22] Ibid.

[23] Regulation on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), European Parliament and Council, Regulation (EU) 2016/679 (O.J. L 119 of 4 May April 2016, p. 1-88), 27 April 2016.

[24] Article 2(1)(a) and (c) proposed regulation.

[25] See article 25 proposed regulation; this does not apply when there is an importer.

[26] See article 3(1); Annex I further defines "artificial intelligence techniques and approaches" as follows: "a (a) Machine learning approaches, including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning; (b) Logic- and knowledge-based approaches, including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems; (c) Statistical approaches, Bayesian estimation, search and optimization methods." The Commission has the power according to article 4 to amend this definition.

[27] See article 2(2) proposed regulation.

[28] Article 2(3) proposed regulation.

regulation *e contrario* does apply to dual use AI (which the Union has the power to regulate). Even if justified and well circumscribed, the exclusion of weapons systems, in particular those with autonomy, has the consequence that (lethal) autonomous weapons systems remain unregulated. This, in turn, reflects the stalemate in the discussions within the forum of the Certain Conventional Weapons Convention.[29]

The proposed regulation carefully attempts to foster innovation and not stifle research and development of AI. Regulatory sandboxes are permitted under the proposed regulation.[30] Further measures are designed to alleviate the regulatory burden on small and medium-sized enterprises.[31] It is unclear at this moment whether these measures will be effective and actually of much use in practice. Moreover, while the proposed regulation basically does not cover AI which has not yet been placed on the market and hence refrains from directly interfering with research and development, it is possible that the mere existence of the legislation and the need to comply with it as soon as a product becomes ready for circulation, has a chilling effect on innovation and development. However, regardless of such an effect, with the proposed regulation the heavy hand of the regulatory state has undoubtedly arrived in AI. Under the new regulation, there will be some role for the Union (or the Member States) in almost all things related to AI. Hence, the proposed regulation likely marks the end of unbridled freedom in AI. Some may see this as welcome progress, since the development of AI in the wild is brought to an end; others will deplore the loss of the greatest possible freedom.

Finally, some further aspects of the proposed regulation should be highlighted. i) A careful analysis is necessary to determine whether the Union has the power to adopt the regulation with the content proposed. It is not clear whether articles 16 of the Treaty on the Functioning of the European Union on data protection and article 114 TFEU, which requires a link to the establishment and the functioning of the internal market, together can serve as foundation for the whole regulation.[32] While the power to impose penalties remains with the Member States, article 72 of the proposed regulation determines that certain behaviour, including contravention of the ban of certain uses of AI within the sense of article 5 (which is rather vague in some respects), shall be subject to fines totalling no more than 30 Mio. Euro or 6% of worldwide annual turnover. Depending on the size of the market actor, this may of course amount to a

[29] Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons which may be deemed to be Excessively Injurious or to have Indiscriminate Effects (with Protocols I, II and III), 1342 UNTS 163 (engl.), 10 October 1980; but see the consensus on certain basic principles with regard to autonomous weapons systems: Group of Governmental Experts (under CCW), *Eleven Guiding Principles on Lethal Autonomous Weapons Systems*, 2020, available on the internet at: <https://multilateralism.org/wp-content/uploads/2020/04/declaration-on-lethal-autonomous-weapons-systems-laws.pdf>.

[30] Articles 53 and 54 proposed regulation.

[31] Article 55 proposed regulation.

[32] See the beginning of the preambulatory clause.

significant, if not eyewatering sum. Furthermore, some provisions touch upon procedural criminal law which remains largely within the power of the Member States. It is, for instance, not clear whether the provisions concerning recidivism and predictive policing are covered by Union powers. A question of power also arises with regard to AI that assists judicial authorities other than the courts of the Union.[33]

ii) The proposed regulation does not per se preclude AI that is learning live and on the go while being in use (adaptive AI), but imposes some additional obligations beyond those already applicable to high-risk AI in which the learning is frozen when it is put into circulation.[34] iii) AI that was trained exclusively with data stemming from the Union benefits from a presumption of compliance with the regional specificity requirement applicable to training data.[35] It is not strictly necessary, though, to retrain AI which was trained abroad in the Union with data stemming from the Union. iv) The requirement to ensure human oversight of high-risk AI in article 14 of the proposed regulation begs the question as to its association with the notion of meaningful human control, which has been discussed for some time within CCW. v) Compliance with harmonised standards, which have been recognized as authoritative by way of publication in the Union Official Journal, is rewarded with a presumption of conformity.[36] This presumption may be important for bodies like the IEEE which are currently developing standards of AI.

***

[33] See the provision as to law enforcement and judicial authorities in annex III(6) and (8).

[34] See articles 15(3) and 43(4) proposed regulation.

[35] See article 42 in conjunction with article 10(4) proposed regulation.

[36] Article 40 proposed regulation; see also the definition of "harmonised standard" in article 3(27) proposed regulation.