The EU as regulator of Artificial Intelligence – accountability and a new variation of a "Brussels Effect"?

Paper for the European Union Studies Association (EUSA) Conference, May 19-21, 2022, Miami, FL, ***Panel 11E on European Digital Policies,*** Saturday, May 21, 1.45-3.15 pm

Presenting author: **Prof. Dr. Hartmut Aden; co-author** (not presenting) **Milan Tahraoui**[1]

Authors' affiliation and address: Berlin School of Economics and Law (HWR Berlin), Berlin Institute for Safety and Security Research (FÖPS Berlin), Alt-Friedrichsfelde 60, D-10315 Berlin, Germany, email: Hartmut.Aden@hwr-berlin.de

Early draft – not for citation, feedback and critique very welcome!

The paper is based on the research project *FAKE-ID - Videoanalyse mit Hilfe künstlicher Intelligenz zur Detektion von falschen und manipulierten Identitäten* (2021-2024), funded by the German government (BMBF, Grant number 13N15737).

Abstract

In April 2021, the European Commission published a proposal for a Regulation on Artificial Intelligence (COM(2021) 206 final). The draft "Artificial Intelligence Act" combines the objective of reconciling the development of a European market for AI applications with the aim "that AI systems placed on the Union market and used are safe and respect existing law on fundamental rights and Union values". In this paper, we are analyzing the proposal from the perspective of the EU as regulator of accountability settings for the risks related to the use of emerging technologies. The Commission did not opt for loose harmonization with a directive, but for a detailed regulation as an ambitious regulatory approach for artificial intelligence. The paper discusses potential effects of the risk-based approach upon the accountability for the development and use of AI systems and links this to the broader discussion on the effects of EU regulatory policy approaches inside and beyond the EU and a "Brussels Effect" of the EU' regulatory policies.

[1] The authors would like to thank Mr Mario Petoshati for his support.

1. Introduction: The European Commission's ambition to regulate artificial intelligence

In April 2021, the European Commission published a proposal for a Regulation on *artificial intelligence* (AI) (European Commission 2021). This draft ("Artificial Intelligence Act") combines the objective of reconciling the development of a European market for AI applications with the aim "that AI systems placed on the Union market and used are safe and respect existing law on fundamental rights and Union values". This regulatory approach seeks to reconcile a number of conflicting aims in technology regulation, combining the perspective of future economic benefits of new technology with regulatory approaches to minimise potential harm. In the draft AI Regulation, the European Commission opted for a risk-based approach with respect to the protection of fundamental rights.

The Commission lists the conflicting aims in its Explanatory Memorandum for the draft regulation:

> - [to] ensure that AI systems placed on the Union market and used are safe and respect existing law and fundamental rights and Union values
> - [to] ensure legal certainty to facilitate investment and innovation in AI
> - [to] enhance governance and effective enforcement of existing law on fundamental rights and safety requirements applicable to AI systems
> - [to] facilitate the development of a single market for lawful, safe and trustworthy AI applications and prevent market fragmentation. (European Commission 2021, p. 3)

This approach also reacts to concerns with respect to negative impacts of AI. The Conclusions of the Council of the EU of 21 October 2020 mention challenges such as opacity, complexity, bias, a certain degree of unpredictability and partially autonomous behaviour as topics that needed to be addressed in future AI regulation (Council of the European Union 2020, p. 5, para. 5).

The Union's regulatory approach on AI should not be considered in isolation. In a global perspective, several recently established or currently emerging regulatory frameworks for AI systems exist, reflecting various sensitivities towards legal and technological challenges and divergent interests. Established requirements mostly include: "(i) accuracy ii) resiliency vis-à-vis different vulnerabilities, (iii) privacy preservation, (iv) robustness (in different contexts and environments), (v) reliability, (vi) safety, (vii) interoperability, and (viii) explainability of AI systems" (cf. Barros Vale and Matheson 2022, p. 8). This does not necessarily mean that those concepts and related standards are understood conceptually and technically in similar terms. However, it nonetheless indicates the existence of some broad regulatory patterns that can also be found in the emerging regulatory EU legal framework for AI.

This paper specifically looks at the risk-based approach to AI regulation that the Commission opted for and the role that the accountable use of AI plays in this approach. We argue that this kind of regulatory approach can have an effect beyond Europe, building upon what Anu Bradford (2012 and 2020) has called the *Brussels Effect*: the EU's regulatory approaches to emerging technologies that can have far-reaching impacts upon producers and consumers in other parts of the world. The draft EU regulation, as the Commission proposed it, categorically claims global leadership for AI regulation with a qualitative approach of AI regulation explicitly based on an ethical perspective:

> By laying down those rules, this Regulation supports the objective of the Union of being a global leader in the development of secure, trustworthy and ethical artificial intelligence, as stated by the European Council, and it ensures the protection of ethical principles, as specifically requested by the European Parliament. (European Commission 2021, p. 18 = Recital 5).

Against this backdrop, this paper discusses answers to the following questions: (1) What is the role of accountability in the risk-based approach to the regulation of AI? (2) What is the potential impact on the accountability of the use of AI applications – in Europe and beyond?

Methodologically, this paper is based on the analysis of ethical and legal approaches to the regulation of AI and of the draft EU AI Regulation – and on interdisciplinary cooperation with AI analysts and developers in the underlying research project.

2. The definition of AI and its influence on the scope of the regulatory approach

While AI has become a major topic in technology policy and research over the past decades, the ethical, political and legal debate on the regulatory framework for this quickly developing technology shows that there is still some room for discussion on what exactly AI is. The definition of AI is much more than just an academic issue. In the context of AI regulation, the definition directly influences the regulatory scope. Thus, it is not surprising that the definition of AI in the future EU AI Regulation is contested. Typical for EU law, the Commission's Proposal for an AI Regulation suggests a brief definition in the main text, referring to an annex where the scope of cases covered by the regulation. Article 3(1) of the European Commission's Proposal defines 'artificial intelligence system' as:

> software that is developed with one or more of the techniques and approaches listed in Article 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; (European Commission 2021, p. 39 = Art. 3(1)).

In its Compromise text of 29 November 2021, the Slovenian Presidency of the EU Council proposed a substantially modified definition of an AI system that read:

> [AI] system means a system that (i) receives machine and/or human-based data and inputs, (ii) infers how to achieve a given set of human-defined objectives using learning, reasoning or modelling implemented with the techniques and approaches listed in Annex I, and (ii) generates outputs in the form of content (generative AI systems), predictions, recommendations or decisions, which influence the environments it interacts with. (Council Presidency 2021, p. 8= Recital 6)

Compared to the definition in the Commission's Proposal, this amended definition conceived AI as more than just a software. The Slovenian Council Presidency highlighted two objectives behind this revised definition that are both teleological in nature. The new definition would aim to increase legal clarity, while insisting that any AI system should seek "to achieve human defined objectives by learning, reasoning or modelling." An additional aim for this revised definition of any AI system is to ground it "on the key functional characteristics of the software of artificial intelligence distinguishing it from more classic software systems and programming." (Council Presidency 2021, p. 33) In both variations, Annex I will play a crucial role.

The OECD Council provided in its 2019 Recommendation on AI the following general definition of AI systems:

> An AI system is a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy. (OECD Council 2019, p. 7)

In its 2021 report on the right to privacy dedicated to artificial intelligence, the UN Human Rights Council argued that a widely accepted definition of AI did not exist. However, in its own definition of AI, the UNHRC stressed that AI systems are generally aimed at complementing or replacing "specific tasks performed by humans", with the definition also including machine-learning and deep-learning (UN Human Rights Council 2021, p. 4).

Interestingly, the UNESCO in its 2021 *Recommendation on the Ethics of Artificial Intelligence* explicitly refused to establish an abstract general definition of AI system within this document. This would be pointless due to the obsolescence that characterizes AI fields, preferring instead to develop an approach according to which "AI systems […] [are] systems which have the capacity to process data and information in way that resembles intelligent behavior, and typically includes aspects of reasoning, learning, perception, prediction, planning or control." (UNESCO 2021, p. 4). In addition, the UNESCO details three core elements pertaining to its approach of AI systems:

> (a) AI systems are information-processing technologies that integrate models and algorithms that produce a capacity to learn and to perform cognitive tasks leading to outcomes such as prediction and decision-making in material and virtual environments. AI systems are designed to operate with varying degrees of autonomy by means of knowledge modelling and representation and by exploiting data and calculating correlations. […]
>
> (b) Ethical questions regarding AI systems pertain to all stages of the AI system life cycle, understood here to range from research, design and development to deployment and use, including maintenance, operation, trade, financing, monitoring and evaluation, validation, end-of-use, disassembly and termination. In addition, AI actors can be defined as any actor involved in at least one stage of the AI system life cycle, and can refer both to natural and legal persons, such as researchers, programmers, engineers, data scientists, end-users, business enterprises, universities and public and private entities, among others.
>
> (c) AI systems raise new types of ethical issues that include, but are not limited to, their impact on decision-making, employment and labor, social interaction, health care, education, media, access to information, digital divide, personal data and consumer protection, environment, democracy, rule of law, security and policing, dual use, and human rights and fundamental freedoms, including freedom of expression, privacy and non-discrimination. […] In the long term, AI systems could challenge humans' special sense of experience and agency, raising additional concerns about, inter alia, human self-understanding, social, cultural and environmental interaction, autonomy, agency, worth and dignity. (UNESCO 2021, p. 4)

Even if there are differences between artificial intelligence, machine-learning or even deep-learning, this paper uses the broader term of AI or AI systems as indiscriminately encompassing the two other related but distinct concepts. In conclusion, the variety of definitions and descriptions of what AI constitutes underlines the point that AI is still an emerging technology. However, from the perspective of binding regulation, the definition also becomes a contested matter that will have a crucial impact on what exactly will be within the consideration of the emerging binding rules.

3. The risk-based approach of artificial intelligence regulation

The Commission's Proposal for an AI Regulation follows a risk-based approach. Risk-based regulation distinguishes between various levels of risk that require different responses. It usually 'calibrates' the enforcement of the law based on the level of risk that has been identified for the specific context. Thus, a risk-based regulation can be conceived as a tool that helps to prioritize and target enforcement action in a manner that is proportionate to an actual hazard (see De Gregorio and Dunn 2022, pp. 475-476).

In the EU context, risk-based AI regulation can also be seen in close connection with the *precautionary principle* that the EU has established for environmental policy (Article 191 (2) TFEU) and beyond: if a technology or behaviour is highly risky, state authorities are entitled to impose restrictions upon such technologies or activities. Even if it is still uncertain whether damage will occur in a certain situation, due to a lack of appropriate knowledge about the risks,

freedoms on which technology developers and business are able to innovate in their private capacities may already be pre-emptively restricted. This is a fundamentally different regulatory approach, compared to the US. In the US context, laws and public authorities regulate risky technologies more leniently, but companies selling dangerous products must be prepared to pay punitive damages to victims if these products cause harm to consumers.

In the draft EU AI Regulation, differentiated levels of risks include: prohibited AI systems (with exceptions); high-risk AI systems (subdivided between general high-risk AI systems and 'stand-alone' high-risk AI systems, which are defined in the context of their intended purpose, like the usage of face identification tools); and low-level risk AI systems. The applicable obligations for the various stakeholders involved in the development, use or commercialisation of an AI system will vary according to corresponding risk-categories.

In Recital 14, the Commission explains the aim of this risk-based approach as follows:

> In order to introduce a proportionate and effective set of binding rules for AI systems, a clearly defined risk-based approach should be followed. That approach should tailor the type and content of such rules to the intensity and scope of the risks that AI systems can generate. It is therefore necessary to prohibit certain artificial intelligence practices, to lay down risks for high-risk AI systems and obligations for the relevant operators, and to lay down transparency obligations for certain AI systems. (European Commission 2021, p. 21)

The European Commission's Proposal includes the so-called "prohibited artificial intelligence practices" as highest risk-category defined in its Article 5. As the same Article allows broad exceptions, the name is somehow misleading. Several authors have rightly pointed out that these exceptions actually undermine the proclaimed prohibition, due to the extensive scope of those exceptions (cf. Linardatos 2022, p. 60; Veale and Borgesuius 2021, p. 101; Ebert and Spiecker 2021, pp. 1189-1190). This criticism applies in particular to the envisaged prohibition of biometric facial recognition in public spaces that foresees several exceptions for law-enforcement. The Commission's final version of the draft AI Act has apparently already ceded to pressure form law-enforcement representatives (cf. Veale and Borgesius 2021, p. 98; Access Now 2020). The European Parliament in a recent resolution on AI in criminal law (European Parliament 2021, p. 11). As observed by Veale and Borgesius (2021), several facial-recognition applications for law-enforcement purposes are possible under the Commission's Proposal, in addition to the possibility for law-enforcement authorities to use so-called "remote biometric identification systems" if they satisfy the conditions laid out in the proposal.

Major parts of the draft regulation define substantive, and more importantly, procedural requirements for the use of high-risk AI systems (i.e., chapter 2 (Articles 8ff.); European

Commission 2021, p. 46ff.), while only a number of general, mostly procedural rules will apply tom remaining AI systems classified as low risk.

4. The accountability for the development and use of AI systems

Accountability is a frequently used concept at this moment in time. This paper uses *accountability* as a relational concept involving actors carrying responsibility for a specific private or public task and those actors towards whom they are held accountable (Bovens et al. 2014). In the literature and in political debates, the term *accountability* often serves as an umbrella term for concepts such as *transparency*, *efficiency*, *responsiveness*, *responsibility* and *integrity* (cf. Bovens 2007: p. 449f.). The term is sometimes used with normative connotations ('accountability as a virtue'), and sometimes in a more analytical sense ('accountability as a mechanism', see Bovens et al. 2014, pp. 7-9). Accountability can be conceived as the sum of concepts such as oversight, democratic control, responsibility, integrity and transparency.

The Commission's draft EU AI Regulation mentions *accountability* requirements in the context of high-risk AI systems:

> It is therefore appropriate to classify as high-risk a number of AI systems intended to be used in the law enforcement context where accuracy, reliability and transparency is particularly important to avoid adverse impacts, retain public trust and ensure accountability and effective redress. (European Commission 2021, p. 20 (=Recital 38).
>
> For the quality management system to be provided by providers of high-risk AI systems, the proposal asks for "an accountability framework setting out the responsibilities of the management and other staff with regard to all aspects listed in this paragraph. (European Commission 2021, p. 53 (= Article 17 (1)i).

As accurate – ideally original – training data is particularly important for the reliability of AI applications, the Commission suggests "regulatory sandboxes" as a specific tool that could make it possible to use personal data for AI development. These "sandboxes", according to the proposal, will be kept by the European Data Protection Supervisor (EDPS) or by the Member States' Data Protection Authorities (DPA) (Article 52). For accountability purposes, the Commission suggests that

> In the AI regulatory sandbox personal data lawfully collected for other purposes shall be processed for the purposes of developing and testing certain innovative AI systems in the sandbox under the following conditions:
>
> […] the logs of the processing of personal data in the context of the sandbox are kept for the duration of the participation in the sandbox and 1 year after its termination, solely for the purpose of and only as long as necessary for fulfilling accountability and documentation obligations under this Article or other application Union or Member States legislation; (European Commission 2021, p. 71= Art. 54 (1)(h).

In addition to the places where the Commission's proposal explicitly mentions accountability, the basic principles discussed for the regulatory framework of AI systems and covered by the draft EU regulation are closely related to accountability in the broad sense as it is understood here: final human decision-making, transparency and explainability:

4.1 Final human decision-making

AI regulation usually starts from the ethical and normative assumption that humans, rather than computers, should have the final say with regard to decision-making. The heightened importance of human intervention in AI-supported decision-making processes is therefore widely acknowledged, but complex challenges persist with respect to *how* human intervention can be conceived and implemented in various AI applications.

Human intervention must be conceived in a complex technical environment where, since it is difficult to understand how an AI system comes to a concrete outcome, challenging such outcomes may be hard or even impossible to achieve. The European Commission (2020, p. 11) already discussed this in its preparatory whitepaper for the draft regulation. The unpredictability of AI systems is also a generally acknowledged definitional feature of those systems that mark a clear difference with previous computational and algorithmic systems (see for instance, U.S. National Institute of Standards and Technology 2021, p. 1).

This state of affairs is of particular concern in the law-enforcement context, where AI-based decisions and actions can usually be not understood and therefore challenged by affected persons. Against this backdrop, the European Parliament has insisted in its Resolution on AI in criminal law and its use by the police and judicial authorities in criminal matters, on the importance of human intervention for AI systems used in those particularly sensitive contexts:

> [The European Parliament] underlines that in judicial and law enforcement contexts, the decision giving legal or similar effect always needs to be taken by a human, who can be held accountable for the decisions made; considers that those subject to AI-powered systems must have recourse to remedy; recalls that, under EU law, a person has the right not to be subjected to a decision which produces legal effects concerning them or significantly affects them and is based solely on automated data processing; underlines further that automated individual decision-making must not be based on special categories of personal data, unless suitable measures to safeguard the data subject's rights and freedoms and legitimate interests are in place; stresses that EU law prohibits profiling that results in discrimination against natural persons on the basis of special categories of personal data; highlights that decisions in the field of law enforcement are almost always decisions that have a legal effect on the person concerned, owing to the executive nature of law enforcement authorities and their actions; notes that the use of AI may influence human decisions and have an impact on all phases of criminal procedures; takes the view, therefore, that authorities making use of AI systems need to uphold extremely high legal standards and ensure human intervention, especially when analysing data deriving from such systems; requires therefore the sovereign discretion of judges and decision-making on a case-by-case basis to be upheld; calls for a ban on the use of

> AI and related technologies for proposing judicial decisions; (European Parliament 2021, p. 9 (para. 16)).

Requiring human intervention for regulating AI systems constitutes a common theme in most emerging regulatory frameworks, but theoretical approaches to the concrete parameters of human intervention, and how it should transpire in AI systems' functioning and lifecycles vary significantly. The UN Special Rapporteur on the right to freedom of opinion and expression (2018) argued that human intervention as an enabler for human rights scrutiny should at least intervene to limit automation's side effects, given AI systems' adaptability that requires human intervention to monitor and follow-up on their potential evolutions:

> Machine-learning AI systems are adaptable, as the algorithms that power them are able to progressively identify new problems and develop new answers. Depending on the level of supervision, systems may identify patterns and develop conclusions unforeseen by the humans who programmed or tasked them. This lack of predictability holds the true promise of AI as a transformational technology, but it also illuminates its risks: as humans are progressively excluded from defining the objectives and outputs of an AI system, ensuring transparency, accountability and access to effective remedy becomes more challenging, as does foreseeing and mitigating adverse human rights impacts. (UN Special Rapporteur on the right to freedom of opinion and expression 2018, p. 6, para. 8).

Another critical issue identified in this report refers to the risks posed by the AI-based automation of remedies against human rights violations (UN Special Rapporteur on the right to freedom of opinion and expression 2018, p. 15, para. 41).

4.2 Transparency, fairness and explainability

Transparency, explainability and fairness constitute an almost systematic requirement in emerging regulatory frameworks for AI systems. However, due to the complexity of the functioning of AI systems, the implementation of these requirements generates difficult practicalities. For these reasons, the high-level of difficulty to ensure explainability is well illustrated by the fact that an entire field of research ("XAI") is dedicated to this problem (cf. UN Special Rapporteur on the right to freedom of opinion and expression 2018, p. 15, par. 40).

Transparency is a key success factor for the accountability of any data processing (cf. Raab, 2012, p. 24ff.). This idea can also be transferred to AI systems. In this context, *transparency* means that only when citizens are informed of the purposes for which their data will be used – subsequently, they are able to decide if these purposes are acceptable to them. Article 5(1) GDPR mentions transparency as one of the core principles for data protection, along with principles such as the legality and fairness of data processing and purpose limitation. Transparency and fairness are closely connected. Only in situations where the use and

application of a technology is transparent, and therefore understandable for the citizens, may they perceive it as fair and legitimate – and therefore accept it (cf. Aden 2022, p. 125).

In the perspective of accountability, transparency requirements for AI application do not only concern the final human decision-makers, but also those affected by decisions supported by AI systems and other institutional actors such as supervisory authorities. Transparency cannot be a mere abstract and general requirement: it needs to be contextualized against the backdrop of AI systems effects, for instance, on affected persons and communities, in view of offering *intelligible* explanations on some aspects of its functioning.

Additional hindrances for transparency and explainability arise with secrecy and confidentiality obligations which may be imposed by private corporations to protect their business model and property rights-based interests (cf. Independent High-Level Expert Group on Artificial Intelligence 2019). Similarly, public authorities' interests to protect for instance investigative methods of their AI-supported law-enforcement activities can be a limitation to transparency and explainability (cf. for example CENTRIC and Europol Innovation Lab 2022, p. 30-31). This is also reflected in Article 70 of the Commission's Proposal for an AI Regulation stating that

> [n]ational competent authorities and notified bodies involved in the application of this Regulation shall respect the confidentiality of information and data […] to protect, in particular: [] (a) intellectual property rights, and confidential business information or trade secrets or legal person, including source code […][;] (b) the effective implementation of this Regulation, in particular for the purpose of inspections, investigation or audits; public and national security interests; (c) integrity of criminal or administrative proceedings. (European Commission 2021, p. 81).

Indeed, transparency and explainability requirements enable responsibility and accountability, since without them one cannot monitor and oversee how AI systems function and generate real-life impact (cf. Council of Europe Ad Hoc Committee on Artificial Intelligence 2020, p. 33, para. 107).

4.3 Accountability mechanism for the use of AI applications

Due to the complexity of AI systems, accountability cannot be effective without performant accountability mechanisms. The Commission, in its draft AI regulation, does not propose to establish specialized authorities for the supervision of AI systems, but to attribute tasks to existing supervisory bodies, namely to data protection authorities. For its own "assistance", the Commission proposes the establishment of a European Artificial Intelligence Board, involving the EDPS and national supervisory authorities – that must not necessarily be Data Protection Authorities (European Commission 2021, pp. 72-73 (=Article 56-58)). The French Presidency

of the EU Council suggested to expand the role that will be attributed to EU Member States in the future EU AI regulatory approach through the European AI Board. This proposed strengthened role of the European AI Board is framed by the French Presidency of the EU Council as a way to counter-balance the powers of the European Commission, including with respect to common specifications, obligations relating to technical documentation, conformity assessment and to EU declaration of conformity (Conseil de l'Union européenne 2022, pp. 3, 10).

5. Towards a new variation of the "Brussels Effect" – including the accountability for AI systems?

Ambitious regulatory approaches developed by the EU often have an effect beyond the EU, what has been labelled as the "Brussels Effect" (Bradford 2012 , and 2020). As Bradford puts it, "the EU has become a global regulator [… but] does not exert global regulatory power over any policy it desires — market forces successfully globalize some EU regulations but not others, setting limits on the Brussels Effect" (Bradford 2020, p. 25). If the "Brussels Effect" can be originally regarded rather as a non-positive legal concept, Bradford tends to distinguish a "de facto Brussels Effect" from a de jure one, where the former paves the way for the latter (Bradford 2020, p. 78).

The General Data Protection Regulation (EU 2016/680, GDPR) and its impact on the global digital economy have been intensively discussed as a recent example of the "Brussels Effect" (cf. Bradford 2020, pp. 122-169). Bradford's analysis shows that the "Brussels Effect" is legally anchored in mechanisms tending to confer an extraterritorial or extended territorial outreach to its norms (Bradford 2020, pp. 67-68). This transpires in the GDPR broad spatial scope of application and its numerous provisions aiming at regulating international data transfer that interest broadly the European Union (in that sense e.g. Ryngaert and Taylor 2020, pp. 4-9). In this regard, the Commission's Proposal for an AI Regulation confers a broad scope of application, since its Art. 2(1) set forth that:

> This Regulation applies to: […] (a) providers placing on the market or putting into service AI systems in the Union, irrespective of whether those providers are established within the Union or in a third country; [] (b) users of AI systems located within the Union; […] (c) providers and users of AI systems that are located in a third country, where the output produced by the system is used in the Union. (European Commission 2021, pp. 38-39 (=Article 2))

In this section, we will discuss reasons that can make the future EU AI Regulation a serious candidate for a new variation of the "Brussels Effect", despite the fact that the corresponding

EU legislative process is still ongoing. For analyzing what could be the possible "Brussels Effect" in AI fields, it is necessary to first highlight how the EU institutions involved in the emerging Union AI regulatory framework are trying to consolidate the EU regulatory powers, as compared to EU member States' competencies (5.1). We are more specifically looking at accountability as part of what might become internationally recognized standards for AI systems. In this respect, a possible new variation of the "Brussels Effect" must be considered at the global level also in view of competing regulatory approaches with potential international outreach, especially those promoted by the United States and China (5.2).

5.1 The future AI Regulation of the European Union: internal conditions for the "Brussels Effect"

The Commission draft AI Regulation clearly claims global leadership for AI regulation (European Commission 2021, p. 18 = Recital 5). One general reason that makes the Commission's Proposal for an AI Regulation a serious candidate for a new variation of the "Brussels Effect" is the fact that the Union is already perceived having such a global influence in data protection regulatory matters, on which it can capitalize for its emerging AI regulatory framework.

Does the Union possess sufficient regulatory capacity to exercise global regulatory authority? A tentative answer to this question can be provided by zooming in one of the conditions possibly underpinning a "Brussels Effect", namely when the EU can put in "place institutional structures that are capable of adopting and enforcing regulations effectively" (Bradford 2020, p. 25). In this context, we hypothesize that the Commission is trying to maximize its competencies over that of EU Member States in order to maximize the EU's unilateral regulatory effects at the global level. Indeed, the fact that the Commission proposed a regulation – and not a directive – underlines the ambition to go for a high level of harmonization – what is one of the conditions for a potential "Brussels Effect".

As Veale and Borgesius (2021, pp. 108-110) critically observed in their commentary of the Commission's Proposal for an AI Regulation, there seems to be the intention to establish some parts of the future AI Act as a "maximum harmonisation threshold," under which EU member States could not be able to legislate with a higher level of regulatory intervention. This is also relevant in the perspective of accountability, as it limits options for higher standards of fundamental rights protection at member states' level. One particular dimension of AI systems that might be covered by this maximum harmonization approach is transparency (cf. Veale and

Borgesius (2021, p. 108, para. 88). This particular point has been criticized by a broad alliance of mainly Europe-based NGOs in a common statement on the Commission Proposal for an AI Regulation:

> Ensure that harmonisation under the AIA is without prejudice to existing or future national laws relating to transparency, access to information, non-discrimination or other rights, in order to ensure that harmonisation is not misused or extended beyond the specific scope of the AIA. (European Digital Rights et al. 2021, p. 6)

These concerns regarding the balance of competencies between the EU and its member states are still transpiring in the ongoing legislative process at the EU level. They are not limited to the protection of fundamental rights and the rule of law. The European Commission has also been criticized for over-stepping its competencies in relation to the member states (Linardatos 2022, pp. 58-60). This criticism is also reflected in the Second Compromise Text of the French Presidency that would offer to law-enforcement authorities, under so-called exceptional circumstances, the possibility to derogate from the obligation of obtaining an authorisation before putting into service a high-risk AI system (Council of the EU 2022, p. 11).

5.2 The objectives and scope of a potential "Brussels Effect"

In the perspective of accountability, the purposes are crucial, especially regarding AI, human rights, democracy and the rule of law. Indeed, Bradford argues that:

> EU rules do not serve a global template merely because they are detailed and available in multiple languages. Several commentators also emphasize the "normal appeal" of EU rules, which increases their attractiveness as a model for emulation. If this appeal exists, the EU's influence also rests on the quality of its ideas and its normative power of persuasion. Ian Manners developed the concept of "Normative Power Europe" to capture the EU's ability to exert influence through persuasion. Manners argues that the EU is best conceived as a normative power, which vests the EU with "ideational power" and the ability to shape what is normal in international relations. This power can be traced to the goals and values that gave rise to EU integration, and to the EU's commitment to democracy, rule of law, human rights, and fundamental freedoms. The appeal of these principles means the EU sets a "virtuous example," leading to a diffusion of its norms across the world. (Bradford 2020, p. 78-81).

Indeed, the extension of the "Brussels Effect" to AI regulation can already be observed in some respect, for example in a recent Brazilian court case (cf. Barros Vale et al. 2022, p. 7-8).

However, a "Brussels Effect" with respect to high accountability standards conflict with regulatory approaches introduced by other strong global economies, i.e. China. Indeed, some authors have recently coined the Chinese regulatory approach in digital and cyber matters as a "Beijing Effect". By contrast with the Brussels Effect, the Beijing Effect is more focused on the global role of Chinese companies for providing infrastructures, including for digital- and cyber-related activities, as it is significantly supported by Chinese State-driven capitalism

notwithstanding through its flagship global initiative – the Digital Silk Road or Belt and Road Initative:

> While China does not demonstrate Bradford's version of the Beijing Effect, we suggest that there is already a Beijing Effect of a different kind, and one that is likely to grow. We theorize three mechanisms, each a combination of push and pull dynamics, through which China affects transnational data governance: First, foreign governments emulate China's approach to data governance and its promise of data sovereignty, aided by China's promotion of that concept in global Internet governance institutions and other venues. […]
>
> Second, Chinese actors play increasingly important roles in digital technology standard-setting. […] Digital technology standards traverse across borders through adoption in international standard-setting organizations or if multinational companies gravitate towards a common standard to ensure interoperability. It is particularly in this regard that foreign companies maintain an interest in cooperating with China. They may not gravitate towards Chinese law—as the Brussels Effects postulates for E.U. law—but towards common technical standards to maintain interoperability with Chinese technology.
>
> Third, Chinese companies provide digital infrastructures and platforms in host countries along the DSR [Digital Silk Road], thereby shaping the conditions under which these countries transition towards digitally-mediated economies and societies. (Erie and Streinz 2021, p. 21-23).

These authors conclude their proposed definition of a Beijing Effect as resulting from "an endorsement of [China's] underlying data governance principle of governmental and territorial control over data that materializes in different and highly context-dependant domestic data laws." (Erie and Streinz 2021, p. 21-23).

As also US governments may want to regain influence on global technology regulation and standards in the future, accountability standards that the EU establishes in its AI regulation will not automatically set global standards, but conflict with other approaches that may perceive accountability and fundamental rights differently or even as unnecessary restrictions to global business.

6. Conclusion and outlook

This paper has shown that the draft EU Regulation on Artificial Intelligence as the Commission proposed it in April 2021 (European Commission 2021) attempts to reconcile economic interests and a high level of accountability, in particular with respect of the protection of fundamental rights. In this context, ethical standards for the development and use of AI systems may become binding legal standards in a near future. The Commission also explicitly expresses the ambition to use the AI Regulation for setting global standards for AI systems.

However, the accountability standards for the development and use of AI are under pressure in several respect: Public authorities, in particular law enforcement agencies, are interested in

maintaining an unregulated margin of manoeuver for the development and use of AI tools. In many cases, this interest conflicts with the fundamental rights of those affected by AI based decisions or whose data will be used for AI training purposes. Similarly, AI-related business will in many cases be more interested in the unregulated use of personal data and in the development of AI systems with as few restrictions as possible.

In this respect, it is not surprising that many elements of the Commission's draft AI Regulation are the subjects of controversial scholarly and political debates. In the end, the European Parliament and the Council will be under pressure to make compromises. The nature of these compromises will be of crucial importance for the way in which the development and use of AI systems will be held to account in the future – in Europe and beyond.

References

Access Now, (2020). 'Europe's Approach to Artificial Intelligence: How AI Strategy is Evolving' (December 2020) [Online [Available at https://perma.cc/X3JM-2M6A (accessed 3 May 2022)

Aden, Hartmut (2022). 'Privacy and Security: German perspectives, European trends and ethical implications', in: Ron Iphofen & Donal O'Mathuna (eds.), *Ethical Issues in Covert Research and Surveillance*, Bingley (UK): Emerald Publishing 2022, pp. 119-129. https://doi.org/10.1108/S2398-601820210000008009.

Barros Vale, Sebastio, Demetzou Katerina, Matheson Lee (2022). Brussels Privacy Symposium 2021, The Age of AI Regulation: Global Strategic Directions, [Online Available at https://fpf.org/wp-content/uploads/2022/03/FPF_Brussels_Privacy_Symposium-2021.pdf.]

Bovens, Mark (2007). 'Analysing and Assessing Accountability: A Conceptual Framework', *European Law Journal*, 13(4), pp. 447-468.

Bovens, Mark, Schillemans, Thomas and Goodin, Robert E. (2014). 'Public Accountability', in: Bovens, Mark, Schillemans, Thomas and Goodin, Robert E. (eds), *The Oxford Handbook of Public Accountability*, Oxford: Oxford University Press, pp. 1-20.

Bradford, Anu (2012). The Brussels Effect, *Northwestern University Law Review*, 107(1), pp. 1-67.

Bradford, Anu (2020). *The Brussels Effect. How the European Union rules the World*, Oxford. Oxford University Press.

CENTRIC (Sheffield University) and Europol Innovation Lab (2022). Accountability Principles for Artificial Intelligence (AP4AI) – AP4AI Framework Blueprint [Online [Available at https://www.ap4ai.eu/node/14 (accessed on 6 May 2022).

Conseil de l'Union européenne (2022). Proposition de Règlement du Parlement européen et du Conseil établissant des règles harmonisées concernant l'intelligence artificielle (legislation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'Union, Texte de compromis de la présidence – Articles 40-52, 2021/0106(COD), 15 February 2022, [Online [Available at https://aeur.eu/f/i2 (accessed on 27 March 2022).

Council of Europe Ad Hoc Committee on Artificial Intelligence (2020). Feasibility Study, CAHAI (2020)23. Strasbourg: Council of Europe.

Council of the EU (2020). Special meeting of the European Council (1 and 2 October 2020) – Conclusions, EUCO 13/20, 2020.

Council of the EU (2021). Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, Presidency compromise text, 2021/0106(COD), 11481/20, 29 November 2021, [Online [Available at https://data.consilium.europa.eu/doc/document/ST-14278-2021-INIT/en/pdf (accessed on 4 May 2022).

De Gregorio Giovanne; Dunn, Pietro (2022). 'The European Risk-Based Approaches: Connecting Constitutional Dots in the Digital Age', *Common Market Law Review*, 59(2), pp. 473-500.

Ebert, Andreas; Spiecker, Indra (2021). 'Der Kommissionsentwurf für eine KI-Verordnung der EU: Die EU als Trendsetter weltweiter KI-Regulierung', *Neue Zeitschrift für Verwaltungsrecht* 6, pp. 1188-1193.

Erie, Matthew S.; Streinz, Thomas (2021). 'The Beijing Effect: China's Digital Silk Road as Transnational Data Governance', *New York University Journal of International Law and Politics*, pp. 1-92.

European Commission (2020). White paper: On Artificial Intelligence – A European approach to excellence and trust, COM (2020) 65 final.

European Commission (2021). Proposal and Explanatory memorandum for a Regulation of the European Parliament and of the Council Laying down harmonized rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts, COM(2021) 206 final.

European Digital Rights et al. (2021). Civil society calls on the EU to put fundamental rights first in the AI Act [Online [Available at https://edri.org/our-work/civil-society-calls-on-the-eu-to-put-fundamental-rights-first-in-the-ai-act/ (Accessed 4 may 2022)

European Parliament (2020). 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).

European Parliament, (2021). Resolution on Artificial intelligence in criminal law and its use by the police and judicial authorities in criminal matters (2020/2016(INI)), P9_TA(2021)0405, 6. October 2021, [Online [Available at https://www.europarl.europa.eu/doceo/document/TA-9-2021-0405_EN.pdf. (accessed on 4 May 2022)

Independent High-Level Expert Group on Artificial Intelligence, (2019). Ethical Guidelines for Trustworthy AI [Council of Europe]

Linardatos, Dimitrios (2022). 'Auf dem Weg einer europäischen KI-Verordnung – ein (kritischer) Blick auf den aktuellen Kommissionsentwurf', *Zeitschrift für das Privatrecht der Europäischen Union*, 19(2), pp. 58-70.

OECD Council (2019). Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449.

Raab, Charles (2012). The meaning of 'accountability' in the information privacy context. In D. Guagnin, L. Hempel, C. Ilten, I. Kroener, D. Neyland, & H. Postigo (Eds.), *Managing privacy through accountability* Basingstoke: Palgrave Macmillan, p. 15–32).

Ryngaert Cedric; Taylor, Mistal (2020). 'The GDPR as *global* data protection regulation?', *AJIL Unbound*, 114, pp. 4-9.

UNESCO (2021). Recommendation on the ethics of artificial intelligence, SHS/BIO/REC-AIETHICS/2021.

UN Human Rights Council (2021). The right to privacy in the digital age, A/HRC/48/31.

UN Special Rapporteur on the right to freedom of opinion and expression (2018). Promotion and protection of the right to freedom of opinion and expression, United Nations General Assembly, A/73/348.

U.S. National Institute of Standards and Technology (2021). Trust and Artificial Intelligence, NISTIR 8330. Online available at https://www.nist.gov/publications/trust-and-artificial-intelligence (Accessed:4 May 2022).

Veale, Michael; Borgesius, Zuiderveen Frederik (2021). 'Demystifying the Draft EU Artificial Intelligence Act', *Computer Law Review International*, 22(4), pp. 97-112.