MenschenRechtsMagazin
Informationen | Meinungen | Analysen

ISSN 2941-1149 | 30 (2025) 2 | pp. 117–143

Contribution

Responsibility and Accountability for the Use of AI in Law Enforcement in the European Union – Lost in Negotiations?

Steven Kleemann[1] • Milan Tahraoui[2, 3]

[1]University of Potsdam, Law Faculty
[2]Centre Marc Bloch (Berlin)
[3]Paris 1 Pantheon-Sorbonne University

Contents

I. Introduction
II. The Area of Law Enforcement in a Risk-Based Approach
III. Accountability Mechanisms under the AI Act and AI contestability
IV. Conclusion
Vitae

Abstract

This paper examines the EU AI Act's application to law enforcement, highlighting how this sector is incorporated into the risk-based approach and assessing the extent to which such incorporation could weaken safeguards for individuals. It argues that, although the newly created accountability framework is complex, it offers only limited remedies for affected individuals. To ensure genuine protection of fundamental rights, the exceptions ('backdoors') embedded in the framework must be critically examined, contestability mechanisms must be strengthened, and the responsibilities of providers and deployers of high-risk AI must be clarified. Where appropriate, a rights-based approach should be integrated into the risk-based approach to underscore that fundamental rights are non-negotiable. This integration is essential to align the use of AI with the AI Act's twin objectives of protecting fundamental rights and promoting innovation.

Keywords

AI Act, Law Enforcement, Accountability, Contestability, Risk Regulation, Fundamental Rights, Digital Policy

Citation:
Steven Kleemann/Milan Tahraoui, Responsibility and Accountability for the Use of AI in Law Enforcement in the European Union, in: MRM 30 (2025) 2, pp. 117–143. https://doi.org/10.60935/mrm2025.30.2.28.

Received: 2025-08-23
Accepted: 2026-01-09
Published: 2026-02-17

Permissions:
The copyright remains with the authors. Copyright year 2026.

Unless otherwise indicated, this work is licensed under a Creative Commons License Attribution 4.0 International. This does not apply to quoted content and works based on other permissions.

I. Introduction[*]

The European Union Regulation laying down harmonised rules on artificial intelligence (AI Act)[1] entered into force on 1 August 2024. Despite the formal adoption of the AI Act, this contribution will mainly anticipate its future application in practice, as the Act set forth differentiated dates for the entry into force of its various chapters, sections and provisions.

In December 2023, prior to its adoption, the European Parliament, the Council of the European Union and the European Commission reached a compromise during the trilogue negotiations.[2] In this context, the regulation of the use of AI for law enforcement activities was one of the most controversial issues in the negotiations around the AI Act, along with the issue of so-called foundation AI systems, now called general-purpose AI models, which culminated in a three-day marathon trilogue process.[3] Although the dangers posed by AI systems are being addressed more frequently, at least nominally in international policy and legal documents, the risks posed by the increasing use of AI tools for law enforcement purposes have often been narrowly focused on the – albeit undeniably important – aspects of data protection, the legal regulation of data processing and the regulation of facial recognition technologies. However, the use of AI systems by law enforcement agencies raises further questions, particularly concerning the risk of fundamental rights being violated by AI-based decisions or actions. This is especially true for the aspects of the AI Act that have come under strong criticism from the perspective of fundamental rights protection.

Since its final adoption, these criticisms have been somewhat vindicated, as the European Commission has committed itself to a so-called simplification process[4] with the potential to further weaken this

[*] This paper is partially based on the research project “VIKING” (FKZ 13N16242), funded by the German Federal Ministry of Education and Research (BMBF), now the Federal Ministry of Research, Technology and Space (BMFTR). The authors wish to thank Mario Petoshati for his assistance with proofreading and editing this paper. This article is based on a paper presented at the “30th Anniversary Conference Human Rights and Artificial Intelligence Addressing challenges, enabling rights”, 7-8th November 2024 in Potsdam, Germany.

[1] EU Regulation 2024/1689 of 12 July 2024, OJ L, 2024/1689 (AI Act).

[2] For an overview on the notion of trilogue in the European Union, see *Giacomo Rugge*, Trilogues: the democratic secret of European Legislation, 2025.

[3] For an account of the main controversies that took place during the negotiations of the AI Act, see European Digital Rights, EU AI Act Trilogues: Status of Fundamental Rights Recommendations, EDRi of 16 November 2023, available at: https://edri.org/our-work/eu-ai-act-trilogues-status-of-fundamental-rights-recommendations/ (last visited 9 December 2025); *Jeremy Fleming-Jones*, EU AI Act nearing agreement despite three key roadblocks – co-rapporteur, euronews of 23 October 2023, available at: https://www.euronews.com/next/2023/10/23/eu-ai-act-nearing-agreement-despite-three-key-roadblocks-co-rapporteur (last visited 15 September 2025); *Müge Fazlioglu*, Contentious areas in the EU AI Act trilogues, IAPP News of 30 August 2023, available at: https://iapp.org/news/a/contentious-areas-in-the-eu-ai-act-trilogues (last visited 15 September 2025).

[4] See European Commission, Simplification, 2025, available at: https://commission.europa.eu/law/law-making-process/better-regulation/simplification-and-implementation/simplification_en (last visited 21 August 2025); *Sarah Chander/Caterina Rodelli*, One Year On, EU AI Act Collides with New Political Reality, Tech Policy Press of 7 August 2025, available at: https://www.techpolicy.press/one-year-on-eu-ai-act-collides-with-new-political-reality/ (last visited 21 August 2025).

protection.[5] Many of the obligations contained in this regulation are either vaguely worded or, even if specific requirements are well-defined, they contain broad exemptions for law enforcement. Furthermore, the final version of the AI Act incorporates a somewhat limited perspective on contestability, in a broader sense, regarding the impact of AI systems on affected persons, due to the fact that it originally was primarily drafted on the basis of pre-existing EU product safety law.

Moreover, the European Commission and other European institutions are under considerable pressure to alleviate the alleged regulatory burden placed on firms and industry actors,[6] who criticize the regulatory model of the AI Act for impeding innovation and preventing security threats from being addressed.[7] The high-risk AI requirements in Articles 8–27 of the AI Act are central to safeguarding fundamental rights, especially where law enforcement authorities deploy AI systems. The obligations affect providers, distributors and deployers differently. Although Articles 8–15 AI Act do not always specify addressees, they largely concern providers given their focus on system design and development[8]—although this is not the case systematically.[9] This allocation of re-

[5] As this paper was written before the Digital Omnibus Proposal was officially announced by the European Commission, we do not address the many implications of the so-called simplification, which has sparked controversy, particularly with regard to the GDPR and the AI Act. See, European Commission, Digital Omnibus Regulation Proposal of 19 November 2025, available at: https://digital-strategy.ec.europa.eu/en/library/digital-omnibus-regulation-proposal (last visited 1 December 2025); European Commission, Digital Omnibus on AI Regulation Proposal of 19 November 2025, available at: https://digital-strategy.ec.europa.eu/en/library/digital-omnibus-ai-regulation-proposal (last visited 1 December 2025). However, the envisaged amendments by the European Commission could raise several contentious issues, including the weakening of responsibility and accountability frameworks regarding the use of AI for law enforcement purposes. See for instance, European Digital Rights, Press release: Commission’s Digital Omnibus is a major rollback of EU digital protections, EDRi of 19 November 2025, available at: https://edri.org/our-work/commissions-digital-omnibus-is-a-major-rollback-of-eu-digital-protections/ (last visited 1 December 2025).

[6] See for instance, European Digital Rights et al., Open Joint letter against the Delaying and Reopening of the AI Act, EDRi of 9 July 2025, available at: https://edri.org/our-work/open-letter-european-commission-must-champion-the-ai-act-amidst-simplification-pressure/ (last visited 21 August 2025); European Center for Not-for-Profit Law et al., Open Letter to the European Commission on the announced withdrawal of the AI liability, ECNL of 7 April 2025, available at: https://ecnl.org/news/eu-needs-ai-liability-rules (last visited 21 August 2025); *Melissa Heikkilä/Barbara Moens*, EU lawmakers warn against ‘dangerous’ moves to water down AI rules, Financial Times of 25 March 2025, available at: https://www.ft.com/content/9051af42-ce3f-4de1-9e68-4e0c1d1de5b5 (last visited 24 October 2025); *Maria Maggiore/Leïla Miñano/Harald Schumann*, France spearheads member State campaign to dilute Europe AI regulation, Investigate Europe of 22 January 2025, available at: https://www.investigate-europe.eu/posts/france-spearheads-member-state-campaign-dilute-european-artificial-intelligence-regulation (last visited 21 August 2025); *Francesca Palmiotto*, The AI Act Roller Coaster: The Evolution of Fundamental Rights Protection in the Legislative Process and the Future of Regulation, in: European Journal of Risk Regulation 16 (2025), pp. 770-793 (789 f.).

[7] EU Champions AI Initiative, Stop the Clock – Open letter, July 2025, available at: https://aichampions.eu/#stoptheclock (last visited 21 August 2025).

[8] Article 16 lit. a. AI Act states that providers must fulfil the high-risk obligations from Section 2 (Art. 8-15 AI Act), which suggests therefore that they are the main addressees. Art. 15-27 AI Act then differentiate between AI providers, AI distributors, AI importers, AI deployers and other parties involved.

[9] See, for example, Art. 14 AI Act on human oversight, which mainly addresses AI providers but whose obligations can only be fulfilled in cooperation with AI deployers. This is among others indicated in Art. 26(3) AI Act, which deals with

sponsibilities can undermine fundamental rights protection: providers must establish a risk management system for high-risk AI (Art. 9), even though deployers, users or affected persons—particularly in policing contexts—may be better positioned to identify actual rights risks in practice.[10] The AI Act partly compensates for this asymmetry by requiring deployers to conduct a fundamental rights impact assessment (Art. 27), which covers similar concerns, albeit through a different mechanism.

Regarding prohibited AI practices (Art. 5), both deployers and providers carry duties aimed at preventing unlawful interferences with rights. Deployers may neither use prohibited AI systems nor operate systems in ways that amount to a prohibited practice. Providers, in turn, must ensure that their systems cannot function—or be reasonably used—in prohibited ways. They must implement effective, verifiable and proportionate safeguards against foreseeable misuse, include contractual clauses banning unlawful applications, and provide clear guidance on correct use and the need for human oversight.[11] The distribution of responsibilities reflects each actor's control over design, development and deployment, and must be assessed proportionately for each prohibition to ensure that those best placed to prevent rights violating uses actually do so.[12]

By contrast, fewer analyses address the rights granted to individuals under the AI Act to contest AI-driven interferences—such as discriminatory policing tools—or to challenge AI development and deployment projects. Strengthening these avenues of redress will require further measures at the national level. Moreover, rights-based contestation interacts with institutional oversight mechanisms—supervisory authorities, complaint procedures, data protection processes and fundamental rights impact assessments. These mechanisms define obligations but must also be articulated in terms of their underlying protective function: the rights of affected persons, available remedies, and the practical requirements for independent and effective supervision. Many implementation challenges arise precisely at this intersection between institutional responsibilities and the need to secure enforceable fundamental rights protections against AI-based law enforcement practices.

Thus, this paper addresses the following research questions: how are particularly sensitive areas, such as the use of AI for coercive public security purposes like law enforcement, incorporated into the AI Act's regulatory framework? What accountability and responsibility mechanisms are in place for the use of AI in these areas, with regard to the protection of fundamental rights for affected persons? To

the obligations of AI deployers and refers to "the deployer's freedom to organise its own resources and activities for the purpose of implementing the human oversight measures indicated by the provider". See about the distribution of roles between AI providers and AI deployers for the implementation of human oversight obligations, *Johan Laux/Hannah Ruschemeier*, Automation Bias in the AI Act: On the Legal Implications of Attempting to De-Bias Human Oversight of AI, in: European Journal of Risk Regulation 16 (2025), pp. 1519–1534 (1524 ff.).

[10] *Nathalie A. Smuha/Karen Yeung*, The European Union's AI Act: Beyond Motherhood and Apple Pie?, in: Nathalie A. Smuha (ed.), The Cambridge Handbook of the Law, Ethics and Policy of Artificial Intelligence, 2025, pp. 228-258 (241 f.).

[11] European Commission, Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act), C(2025) 884 final of 29 July 2025, para. 40.

[12] Ibid., para. 20.

what extent can these mechanisms sustain avenues of contestability against AI development and deployment in these sensitive areas? We will consider the advantages of formal and informal channels of contestability from a rights-based perspective. In this aim, Section II will elaborate on the area of law enforcement within a risk-based approach, including the definition of risk in the AI Act. Section III introduces some of the main features of responsibility and accountability mechanisms with a focus on contestability in its subsections, while Section IV concludes by providing an overview of the current gaps in the AI Act regarding meaningful fundamental rights protection for the use of AI in law enforcement, and briefly suggesting potential solutions.

II. The Area of Law Enforcement in a Risk-Based Approach

Policing, criminal justice, migration, asylum and border control management are not excluded from the scope of the AI Act, as opposed to the use of AI systems for military, defence and national security purposes. However, due to this integrated approach, special 'backdoors' are used to employ risky AI systems that would be prohibited or restricted if they were used by other state agencies or non-state actors.[13] Many of these 'backdoors' were introduced by the Council of the European Union in its amended version of the AI Act through interventions from representatives of the EU Member States as well as the security and law enforcement communities.[14] The term 'backdoor' refers to a number of different approaches that could be taken to enable the use of risky AI systems. These include special exceptions from generally applicable requirements and enabling conditions for the use of these AI systems, or more indirect legal means. An example of this are the permissive rules contained in the AI Act for regulating the testing of AI systems in real-world conditions for law enforcement purposes (Art. 60 AI Act). Furthermore, an analysis of the AI Act clearly shows that there are even more 'backdoors' for the use of AI systems in the domains of migration, asylum and border management and control.[15] Despite these different legal categorisations, the AI tools used in these two domains of law enforcement are indeed similar, with the major difference being that AI used for migration, asylum and border control management is more permissively regulated for national authorities than their use for law enforcement purposes.

[13] On the issue of 'backdoors' in greater detail, see: *Steven Kleemann/Hartmut Aden*, Die Nutzung Künstlicher Intelligenz durch Strafverfolgungsbehörden – „Hintertüren" der Verordnung der Europäischen Union über Künstliche Intelligenz, in: Wilfried Honekamp/Stefanie Kemme/Jens Struck (ed.), Auswirkungen von KI auf die zukünftige Polizeiarbeit. Technologische Potenziale, rechtliche Rahmenbedingungen, kriminologisch-sozialwissenschaftliche Erkenntnisse, 2025, pp. 3–30.

[14] *Palmiotto* (fn. 6), p. 780, p. 787, pp. 789 ff.; *Ludivine Sarah Stewart*, The regulation of AI-based migration technologies under the EU AI Act: (Still) operating in shadows?, in: European Law Journal 30 (2024), pp. 122-135 (127 f.).

[15] This can be observed by comparing Annex III para. 6 f. of the AI Act. See also *Alberto Rinaldi/Sue Anne Teo*, The Use of Artificial Intelligence Technologies in Border and Migration Control and the Subtle Erosion of Human Rights, in: International and Comparative Law Quarterly 74 (2025), pp. 61-89 (83 f.), arguing that the lines between security and migration have been and continue to be increasingly blurred and that the AI Act "ended up compressing distinct State obligations relating to borders and migration into the same risk bucket".

It can be argued that the inclusion of security agencies within the scope of the AI Act represents an attempt to overcome the opt-out logic that has characterised other EU policy areas, such as data protection. It is also noteworthy that accountability mechanisms in the form of remedies are now included in a separate Section 4 of Chapter IX of the AI Act ("Post-market monitoring, information sharing, and market surveillance"), despite no direct remedial mechanisms for persons affected by AI being foreseen in the European Commission's original proposal. In addition, other obligations, such as the obligation to conduct a fundamental rights impact assessment for high-risk AI systems aim to contribute to accountability (Art. 27 AI Act).[16]

1. Risks as defined in the AI Act

The AI Act defines the term 'risk' in Art. 3 para. 2 as "the combination of the probability of an occurrence of harm and the severity of that harm". Furthermore, the Act distinguishes between different levels of risk intensity. The AI Act essentially differentiates between four risk categories (unacceptable, high, limited and minimal or no risk), which impose varying requirements on such systems. Moreover, during the negotiations of the AI Act, the co-legislators introduced a new category of general-purpose AI models[17] and the notion of 'systemic risk'[18] for those AI models that can be qualified as such under Art. 51. However, this new category is at odds with a truly risk-based approach, and these ambiguities should be taken into account when further examining this topic.

It is first necessary to differentiate between risks and uncertainties. A risk may be defined by Art. 9 AI Act as a 'known known', containing statistical probabilities and quantifiable effects, which is reflected in the general definition of 'risk' under Art. 3 para. 2 AI Act: "the combination of the probability of an occurrence of harm and the severity of that harm". This definitional take can be criticised "as the infringement of a fundamental right does not necessarily require any 'harm' to ensue".[19] An uncertainty is, in comparison, a 'known unknown', that is, a situation or event that cannot be quantified because the effects of a specific technology are not yet known. Furthermore, there are instances of 'unknown unknowns', whereby there is no awareness that certain things or activities may have negative effects, despite the potential for such effects to

[16] *Steven Kleemann/Hartmut Aden*, Die Grundrechte-Folgenabschätzung nach dem Artificial Intelligence Act der Europäischen Union – eine Chance für die polizeiliche KI-Nutzung, in: Sabrina Schönrock/Hartmut Aden (ed.) Breitscheidplatz-Symposium 2024: Zukunftssicherheit: Die Rolle von KI im Kampf gegen den Terrorismus, pp. 47-57 (48 ff.).

[17] See AI Act, Chapter V, Arts. 50-56, pp. 83-87; European Commission, Annex to the Communication to the Commission, Approval of the content of the draft Communication from the Commission – Guidelines on the scope of obligations for general-purpose AI models established by Regulation (EU) 2024/1689 (AI Act), C(2025) 5045 final of 18 July 2025, paras. 12 ff.

[18] Art. 3 para. 65 AI Act defines it as: "a risk that is specific to the high-impact capabilities of general-purpose AI models, having a significant impact on the Union market due to their reach, or due to actual or reasonably foreseeable negative effects on public health, safety, public security, fundamental rights, or the society as a whole, that can be propagated at scale across the value chain."

[19] *Nathalie A. Smuha*, The paramountcy of data protection law in the age of AI (Acts), in: European Data Protection Supervisor (ed.), Two decades of personal data protection. What next?– EDPS 20th Anniversary, 2024, pp. 225-239 (235).

exist.[20] In this regard, it can be argued that the deployment of AI systems for border controls is based on a surveillance logic aiming at discovering ‘unknown unknown’ security risks, by inferring information or creating *sui generis* risk group profiles based on AI technologies.[21] Consequently, it is crucial to achieve consensus on the selection of risks and the severity assigned to them, whether in terms of probability, impact, or both, in order to ensure the success of any risk-based approach.[22] This discussion was particularly pertinent during the genesis of the regulation, when there was an intensive debate about whether specific AI systems or applications should be banned in the EU. There were also fundamental debates regarding the classification of some AI technologies as high-risk, minimal risk, or as the relatively new classification of general-purpose AI models.

Those debates surrounding classifications were of particular importance, as they imply different regulatory consequences. For example, AI systems that are classified as high-risk use-cases, must comply with essential, specific, and procedural precautions. This builds upon various issue areas in which EU law attempts to regulate risks.[23] For the regulation of high-risk activities, the EU has adopted the so-called precautionary principle to regulate risks, notably in environmental policy[24] (Art. 191 para. 2 TFEU)[25]. In accordance with this principle, regulatory measures that restrict economic freedoms and fundamental rights may be implemented at an early stage if an evaluation concludes that a risk is likely to evolve into a danger that could cause serious damage, particularly in relation to human life and health. In the EU context, risk-based AI regulation is therefore closely connected to the *precautionary principle*.[26] This principle allows state authorities to impose restrictions upon technologies or activities, if a technology or behaviour is deemed to be highly risky. Even in the absence of certainty regarding the potential for damage to occur, due to a lack of appropriate knowledge about the full extent of the risks involved, the freedom to develop new technologies and to commercialise them may already be subject to pre-emptive restrictions in the interest of risk prevention. That said, the AI Act in its implementation phase seems to be increasingly under pressure not to excessively constrain market actors in the

[20] *Martin Ebers*, Truly Risk-based Regulation of Artificial Intelligence How to Implement the EU’s AI Act, in: European Journal of Risk Regulation 16 (2025), pp. 684-703.

[21] *Gavin Sullivan/Dimitri Van Den Meerssche*, The Legal Infrastructures of UK Border Control—Cerberus and the *Dispositif* of Speculative Suspicion, in: German Law Journal 25 (2024), pp. 1308–1342 (1310 f.); *Louise Amoore*, The Deep Border, in: Political Geography 109 (2024), pp. 1-9 (1, 5 f.).

[22] Ibid.

[23] *Giovanni De Gregorio/Pietro Dunn*, The European risk-based approaches: Connecting constitutional dots in the digital age, in: Common Market Law Review 59 (2022), pp. 473-500 (496 ff.).

[24] *Jale Tosun*, How the EU Handles Uncertain Risks: Understanding the Role of the Precautionary Principle, in: Journal of European Public Policy 20 (2013), pp. 1517–1528; *David Vogel*, Trading up: Consumer and Environmental Regulation in a Global Economy, 1997; *David Vogel*, Trading up and Governing across: Transnational Governance and Environmental Protection, in: Journal of European Public Policy 4 (1997), pp. 556-571 (557 ff.).

[25] Consolidated versions of the Treaty on European Union and the Treaty on the Functioning of the European Union of 26 October 2012, OJ C 326.

[26] *Smuha/Yeung* (fn. 10), p. 232. See also, *Samantha Besson*, La *due diligence* en droit international, in: Recueil des cours de l’Académie internationale de La Haye 409 (2020), pp. 153-398, (334, 322 ff.).

name of fostering innovation,[27] even if the baseline for assessing what 'excess' means in that context appears to be oftentimes missing.[28]

2. The risk-based approach

In the context of AI regulation, the different concepts of risk-based, principle-based, precautionary principle-based, and rights-based approaches all play a role. This is due to the socio-technical dimensions of AI, where a broad risk-based approach is necessary, considering both individual and societal risks. This approach needs to be complemented by a precautionary principle-based approach, where unacceptable risks need to be defined. These approaches, however, are not uniformly defined and applicable. Different digital regulations follow different concepts. The EU's General Data Protection Regulation (GDPR)[29], for instance, can be described as a bottom-up, risk-based approach, while the AI Act can be considered a top-down approach, and the Digital Services Act (DSA)[30] contains both perspectives.[31] The situation is rendered more complex by the fact that a rights-based approach is also being partially pursued in the AI Act.

Thus, considering fundamental rights violations and measuring threats and impacts on them, based on methods that are imposed by some new and existing legislation, has become a significant dimension in digital regulation. However, despite the explicit call in many regulations to consider fundamental rights impacts, opposing views are fundamentally at odds with this. On the one hand, scholars in business and economics maintain that virtually any phenomenon can be quantified; on the other hand, human rights scholars emphasize the non-discretionary, intrinsic nature of rights such as human dignity, arguing that these values resist measurement altogether.[32] One goal of the new digital legislation attempts mentioned is to bring these views together. This can be achieved by analysing how to measure potential infringements of fundamental rights from an *ex ante* and *ex post* viewpoints. This means, on one side, predicting and quantifying the potential severity of human rights risks before they manifest, through *ex ante* assessments, and also defining the consequences of risk realisation in order to mitigate them (precautionary approach). On the other side, *ex post* assessments are usually realised through courts of law to clarify whether a matter that has already been concluded constitutes a human rights violation. In such cases, it is

[27] For further discussion on this topic, see: *De Gregorio/Dunn* (fn. 23), pp. 477 f.; European Digital Rights et al. (fn. 6).

[28] See for the recent developments in that regard: European Commission (fn. 5). See, for example, the criticisms addressed at the methods of the European Commission and its lack of an evidence-based approach for its proposals: *René Mahieu*, The Ominous Omnibus: Dismantling the Right of Access to Personal Data, Verfassungsblog of 3 December 2025, available at: https://verfassungsblog.de/digital-omnibus-right-of-access-to-personal-data/ (last visited 3 December 2025); *Itxaso Domínguez De Olazábal*, The EU's Digital Omnibus Must Be Rejected by Lawmakers. Here is Why, Tech Policy Press of 3 December 2025, available at: https://www.techpolicy.press/the-eus-digital-omnibus-must-be-rejected-by-lawmakers-here-is-why/ (last visited 3 December 2025).

[29] Regulation (EU) 2016/679 of 27 April 2016, OJ L 119.

[30] Regulation (EU) 2022/2065 of 19 October 2022, OJ L 277.

[31] *De Gregorio/Dunn* (fn. 23), pp. 477 f.

[32] See *Gianclaudio Malgieri/Cristiana Santos*, Assessing the (severity of) impacts on fundamental rights, in: Computer Law & Security Review 56 (2025), pp. 1-18 (1 f.).

usually determined whether the *ex ante* measures, like risk management systems, have been properly in place. Combining *ex ante* (regulatory compliance) and *ex post* (judicial remedies) measures leads to a comprehensive approach safeguarding fundamental rights in AI regulation.[33]

The developments in EU digital legislation are shifting in that regard, seeking so-called 'optimal precaution' in specific contexts are considered more suitable than a pure maximalist precautionary approach, in the sense of minimising risks at all costs through imposing maximum precaution.[34] In addition, rights-based approaches are also integrated and can foster the safeguarding of human rights if the precautionary approach is able to consider fundamental rights as a form of normative uncertainty (which, naturally, imposes limitations).[35] Thus, the different approaches not only coexist in EU digital legislation but are also mutually dependent.

a) Critique of the risk-based approach

A criticism directed to the risk-based approach which involves the determination of the scope or scale of a concrete situation or a perceived threat, contends that it is useful only in technical environments.[36] In such situations, companies evaluate their own operational risk. The AI Act's rules concerning, for example, the notifying bodies foreseen in the Act, seem to be heading in this direction. What are these risks weighed against? The selected approach would have companies evaluate their operational risk against people's fundamental rights. However, from a human rights perspective, we disagree with this interpretation. Human rights, at their core, cannot be weighed against companies' interests and must be guaranteed regardless of a risk category based on external considerations.[37] Regardless of potential business gains, businesses have a responsibility to respect human rights and avoid causing or contributing to human rights abuses.[38]

To fully grasp the complexities of this interaction, it is essential to recognise that the methods used for identifying risks or assessing protected rights differ significantly. Risks to people cannot be easily integrated into corporate risk matrices, as the criteria for prioritization are distinct.[39]

[33] Ibid.

[34] *De Gregorio/Dunn* (fn. 23), p. 478.

[35] *Malgieri/Santos* (fn. 32), p. 5.

[36] *Fanny Hidvegi/Daniel Leufer/Estelle Massé*, The EU Should Regulate AI on the Basis of Rights, Not Risks, Access Now of 17 February 2021, available at: https://www.accessnow.org/eu-regulation-ai-risk-based-approach/ (last visited 21 August 2025).

[37] Ibid. See also for the differences of product safety regulation and human rights protection within the AI Act and the challenges this presents: *Marco Almada/Nicolas Petit*, The EU AI Act: Between the rock of product safety and the hard place of fundamental rights', in: Common Market Law Review 62 (2025), pp. 85–120. Counterpoint: economic interests can be weighed against human rights protection, as in the ECtHR, *López Ostra v. Spain* (16798/90), judgment of 9 December 1994, para. 58, in which the Court considered "that the State did not succeed in striking a fair balance between the interest of the town's economic well-being" and "the applicant's effective enjoyment of her right to respect for her home and her private and family life". Furthermore, private corporations do play an important role in implementing human rights at the international level, taking also their commercial activities into account. We maintain that the logic of fundamental rights risk cannot be incorporated to a commercial risk-based approach, because they do not pursue compatible objectives.

[38] UN Guiding Principles on Business and Human Rights, UN Doc. A/HRC/17/31, Annex, Chapter II.

[39] *Malcolm Rog*, Corporate Human Rights Due Diligence, Harvard Kennedy School, Working Paper No. 81 of December 2022 available at:

Furthermore, as analysed above, a purely risk-based approach rather than a rights-based approach is generally inappropriate to protect fundamental rights, as it fails to address the non-negotiable minimum core of human rights.[40] As already stated, human rights cannot be measured or quantified on a scale from trivial to severe. However, the manner in which the AI Act imposes a fundamental rights impact assessment (Art. 27 AI Act) suggests the opposite. The concept of human rights is, by contrast, based on a binary logic, whereby an act is either legal or illegal. It follows that the AI Act might fail in its own ambition to safeguard fundamental rights at this point. The mutual dependence on risk- and rights-based approaches in regulating AI should be given greater focus in the future.

b) The risk-based approach in the context of law enforcement

In this regard, the area of law enforcement is in a particular state of tension.[41] Law enforcement authorities, decision-makers and society at large can be perceived as needing to achieve greater public safety. One potential method of achieving this might be to enhance the technological capacities of law enforcement agencies, for example by improving their ability to handle large volumes of data.[42] Therefore, some argue that if the use of AI-based systems can improve security, it may be in the interest of the state and society to use such systems.[43] Apart from that, it is one of the areas falling within the scope of the AI Act that poses great ethical and fundamental rights risks, alongside with the highly related field of the use of AI in migration, asylum and border control management. Thus, high standards and consistent rules must be established for the use of AI applications in law enforcement. To some extent this tension between enhancing technological capacities and addressing great challenges for ethics and fundamental rights can be analysed in the final version of the AI Act, by analysing the newly introduced category of ‘sensitive operational data’, which refers to: “operational data related to activities of prevention, detection, investigation or prosecution of criminal offences, the disclosure of which could jeopardise the integrity of criminal proceedings” (Art. 3 para. 38 AI Act).[44] The issue with this definition is that it offers too much interpretational leeway to law enforcement agencies acting as deployers, which may systematically endanger human rights protection by creating a loophole. While

https://www.hks.harvard.edu/sites/default/files/centers/mrcbg/files/CRI_81_AWP_FINAL.pdf (last visited 21 August 2025), pp. 68 ff.

[40] *Ebers* (fn. 20), pp. 689 ff.

[41] For a comprehensive overview of the area of law enforcement in a risk-based approach, see: *Steven Kleemann/Hartmut Aden*, Governing High Risk Artificial Intelligence for Law Enforcement: Strengths and Weaknesses of the European Union’s Risk-Based Approach, in: European Journal of Policing Studies (forthcoming).

[42] See for instance, European Commission, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, Roadmap for lawful and effective access to data for law enforcement, COM(2025) 349 final of 24 June 2025, pp. 12 ff.; *Dean Wilson*, Policing, in: Mareile Kaufmann/Heidi Mork Lomell (ed.), De Gruyter Handbook of Digital Criminology, 2025, pp. 363–370 (368). *Contra*, *Raphaël Challier/Myrtille Picaud/Florent Castagnino*, De la « *safe city* » aux dispositifs numériques de sécurité urbaine, in: Réseaux 251 (2025), pp. 11–43 (25 f., 30).

[43] See: *Yasmine Ezzeddine/Petra Saskia Bayerl/Helen Gibson*, Safety, Privacy, or Both: Evaluating Citizens’ Perspectives around Artificial Intelligence Use by Police Forces, in: Policing & Society 33 (2023), pp. 861-876 (862 f.).

[44] The concept of “sensitive operational data” can be found in various places in the regulation, such as in Art. 5 paras. 4 and 7, Art. 26 paras. 5 and 10, Art. 46 para. 3 AI Act.

other regulations, such as the GDPR for instance, define "special categories of personal data" in Art. 9 GDPR, there appears to be no further specifications required for this newly introduced category of data in law enforcement, which leaves far too much room for interpretation.

With regard to the risk classification, Art. 5 AI Act defines "prohibited artificial intelligence practices" as the highest risk category.[45] As broad exceptions will be allowed, the term is misleading.[46] This applies in particular to the prohibition of biometric facial recognition in public spaces that foresees several exceptions for law enforcement.[47] Furthermore, the use of so-called *'ex post'* remote biometric identification in public spaces is authorised as a non-real-time use of biometric identification. However, this is only permitted for law enforcement under the conditions set out by the AI Act when the AI system is classified as high-risk.[48] These use cases must therefore be documented in police files and made available to the supervisory authorities upon request, as well as being reported annually to those authorities (see Art. 26 para. 10 AI Act). There is a very thin line between so-called 'prohibited' real-time remote identification and its *ex post* forms, which can endanger the protection of fundamental rights if they are not strictly defined and controlled.[49]

During the negotiations, the European Commission's Proposal of the AI Act appears to have bowed to pressure from law enforcement representatives[50] and some EU Member States.[51] In a resolution on AI in criminal law, the European Parliament drew attention to the relevance of this conflict.[52] It also took up this issue in its compromise proposal in favour of a general ban, with no exceptions for law enforcement agencies, on the use of 'real-time' remote biometric identification systems in publicly accessible spaces.[53]

[45] For a detailed overview of prohibited AI practices, see: European Commission (fn. 11).

[46] See: *Dimitrios Linardatos*, Auf dem Weg zu einer Europäischen KI-Verordnung – Ein (kritischer) Blick auf den aktuellen Kommissionsentwurf, in: Zeitschrift für das Privatrecht der Europäischen Union 19 (2022), pp. 58–70 (60); *Michael Veale/Frederik J. Zuiderveen Borgesius*, Demystifying the Draft EU Artificial Intelligence Act — Analysing the Good, the Bad, and the Unclear Elements of the Proposed Approach, in: Computer Law Review International 22 (2021), pp. 97–112 (101); *Andreas Ebert/Indra Spiecker gen. Döhmann*, Der Kommissionsentwurf für eine KI-Verordnung der EU, in: Neue Zeitschrift Für Verwaltungsrecht 6 (2021), pp. 1188–1893 (1189 f.).

[47] European Commission (fn. 11), para. 294.

[48] Ibid., paras. 427 f.

[49] *Daragh Murray*, Police Use of Retrospective Facial Recognition Technology A Step Change in Surveillance Capability Necessitating an Evolution of the Human Rights Law Framework, in: Modern Law Review 87 (2024), pp. 833–863 (837); European Commission (fn. 11), para. 310; *Eric Töpfer/Steven Kleemann*, Polizeiliche Gesichtserkennung – Menschenrechtliche Herausforderungen einer Risikotechnologie, Deutsches Institut für Menschenrechte of August 2025, available at: https://www.institut-fuer-menschenrechte.de/fileadmin/Redaktion/Publikationen/Analyse_Studie/Analyse_Polizeiliche_Gesichtserkennung_01.pdf (last visited 27 October 2025).

[50] *Veale/Borgesius* (fn. 46), p. 98; Access Now, Europe's Approach to Artificial Intelligence: How AI Strategy is Evolving, December 2020, available at: https://perma.cc/X3JM-2M6A (last visited 21 August 2025).

[51] See *Maggiore/Miñano/Schumann* (fn. 6).

[52] European Parliament, Artificial Intelligence in Criminal Law and Its Use by the Police and Judicial Authorities in Criminal Matters, P9_TA(2021)0405 of 6 October 2021, available at: https://www.europarl.europa.eu/doceo/document/TA-9-2021-0405_EN.pdf (last visited 21 August 2025).

[53] European Parliament, DRAFT Compromise Amendments on the Draft Report Proposal for a Regulation of the European Parliament and of the Council on Harmonised Rules on Artificial Intelli-

Several facial recognition applications for law enforcement purposes and the various scenarios in which they could be used are conceivable[54] – 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 Act.[55] Despite protests from civil society organisations against the exceptions, which echoed similar criticisms from the EU’s data protection authorities and the European Parliament, the final agreement does not include a real ban.[56] A step forward safeguarding fundamental rights – though it must be closely monitored –, is the Council of Europe’s so-called AI Framework Convention.[57] Unlike the AI Act, which also seeks to harmonise economic interests, this Framework Convention is primarily concerned with the protection of human rights. However, the extent to which it fulfils its stated objective of protecting fundamental rights remains to be determined, as it has a more limited ambition than the AI Act.[58]

The AI Act generally classifies the use of AI systems by law enforcement agencies as high-risk (Recitals 59-60 AI Act). It remains an open question how the risk categories should be applied to AI systems by law enforcement agencies for purposes not listed in the aforementioned Annex III. To what extent does the AI Act regulate law enforcement per se? Some European legislators have successfully argued that the application of the proposed requirements of the AI Act should be excluded precisely in those contexts where the threats to fundamental rights are the greatest: national security, defence, transnational law enforcement, as well as research and development.[59] This exclusion of the material scope of the AI Act was justified by the argument that national security is generally excluded from the scope of EU law and that, according to Recital 24 AI Act, public international law would be “the more appropriate legal framework for the regulation of AI systems in the context of the use of lethal force and other AI systems in the context of military and defence activities”. In principle, this exclusion depends exclusively “on the purposes of the AI system, not the entities carrying out the activities with that system.” However, such AI systems “must be placed on the market, put into service or used exclusively for military, defence or national security pur-

gence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts, COM(2021)0206 – C9 0146/2021 – 2021/0106(COD) of 9 May 2023.

[54] *Töpfer/Kleemann* (fn. 49).

[55] *Veale/Borgesius* (fn. 46), pp. 101 f.

[56] For a comprehensive overview of the debate on bans, see *Catharina Rudschies/Ingrid Schneider*, The Long and Winding Road to Bans for Artificial Intelligence: From Public Pressure and Regulatory Initiatives to the EU AI Act, in: Digital Society 4 (2025), pp. 1-27.

[57] Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law of 05 September 2024, CETS No. 225.

[58] *Francesco Paolo Levantino/Frederica Paolucci*, Advancing the Protection of Fundamental Rights Through AI Regulation: How the EU and the Council of Europe are Shaping the Future, in: Philip Czech et al. (ed.), European Yearbook on Human Rights, 2024, pp. 3-37 (11 f.); European Data Protection Supervisor, Opinion 20/2022 on the Recommendation for a Council Decision authorising the opening of negotiations on behalf of the European Union for a Council of Europe convention on artificial intelligence, human rights, democracy and the rule of law, 13 October 2022.

[59] *Douwe Korff*, Opinion on the Implications of the Exclusion from New Binding European Instruments on the Use of AI in Military, National Security and Transnational Law Enforcement Contexts, 2022, available at: https://ecnl.org/sites/default/files/2022-10/ECNL%20Opinion%20AI%20national%20security_0.pdf (last visited 27 October 2025), p. 28; *Smuha/Yeung* (fn. 10), p. 235.

poses".[60] However, the applicable criteria can endanger fundamental rights, as the distinction between AI systems that are and are not excluded from the scope of the AI Act does not appear to be consistent in practical terms.[61] The European Commission's Guidelines state that dual-use AI systems—those intended for both civilian and security purposes—are covered by the AI Act.[62] However, this does not limit national security, defence, or military actors from using such systems for those specific purposes, regardless of the entity's nature.[63] Agencies like Europol, Frontex, and national police forces may operate outside the AI Act (and the GDPR) when acting under other legal instruments,[64] and large EU IT systems (Eurodac, SIS, ETIAS) are only subject to the AI Act after 2 August 2030 (Art. 111, Annex X AI Act). Article 2 para. 4 AI Act also excludes third country public authorities or international organisations using AI in EU-linked law enforcement or judicial cooperation, a provision that can be extended to private contractors when they are acting on behalf of those authorities.[65] This raises questions about cases such as EncroChat, where law enforcement, intelligence services, and private firms collaborated to breach encrypted communications.[66] The exemption applies only if the cooperation framework includes adequate safeguards for fundamental rights, overseen by the relevant market surveillance authorities.[67]

Therefore, despite the added clarifications regarding the scope of application of the AI systems that are excluded from the scope of the AI Act, there are concerns that some AI systems might still be used by security or judicial actors without respecting the obligations normally applicable to law enforcement. Only time will tell if the AI Act's scope of application might be more or less protective in comparison with the logic that has led to the adoption of the GDPR and the so-called Law Enforcement Directive (LED).[68]

c) Filtering fundamental rights protection and broadening the scope for avoiding high-risk classification

The Commission's Proposal for the AI Act was logical in that all high-risk applications in Annex III would have to comply with certain obligations. However, due to industry and state interventions an additional 'filter' was integrated. This 'filter'[69] can be found in Art. 6 para. 3 AI Act and states that AI systems that are intended

[60] European Commission (fn. 11), para. 24.

[61] See, *Plixavra Vogiatzoglou*, The AI Act National Security Exception: room for manoeuvres?, Verfassungsblog of 9 December 2024, available at: https://verfassungsblog.de/the-ai-act-national-security-exception/ (last visited 15 August 2025).

[62] Ibid.

[63] Ibid., para. 25.

[64] *Korff* (fn. 59), p. 29.

[65] European Commission (fn. 11), para. 29.

[66] *Jan-Jaap Oerlemans/Dave van Toor*, Legal Aspects of the EncroChat Operation: A Human Rights Perspective, in: European Journal of Crime, Criminal Law and Criminal Justice 30 (2022), pp. 309-328.

[67] European Commission (fn. 11), para. 29; See also Recital 22 AI Act.

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

[69] The so-called filter provision was introduced under the influence of the Council of the EU and the European Parliament during the AI Act negotiations of the AI Act, which concluded with the trilogue, see also: *Palmiotto* (fn. 6) pp. 780 f., 787 f.; *Stewart* (fn. 14).

for a narrow procedural task, such as confirming or improving an accessory factor of a human assessment or performing a preparatory task, can be exempted from categorisation as high-risk systems under certain conditions. Moreover, this new feature was integrated despite a countermobilisation of civil society organisations, a critical letter on this very issue from the UN High Commissioner for Human Rights[70] and a damning negative opinion from the European Parliament's legal service.[71]

The introduction of a structural loophole now seems to broaden the remit of the existing high-risk classification, already open to criticism. Despite the introduction of an *ex post* corrective mechanism in the powers attributed to national market surveillance authorities for controlling and sanctioning that a provider has wrongly classified an AI system as non-high-risk according to Art. 80 AI Act, the self-regulatory powers of AI providers organised under the AI Act can be particularly dangerous in the fields of law enforcement and migration.[72]

d) Drawbacks of the risk-based approach

Lastly, the risk-based approach also contains weaknesses. The current categorisation and the systems listed in Annex III (high-risk systems) need to be differentiated in some respects. The listed risks are classified according to external considerations rather than the legal interests involved, such as law enforcement or critical infrastructure on one side and potentially endangered fundamental rights on the other. The classification itself is based on the idea, that these areas are of particular relevance, which is true for a rough template, but it is not conclusive or sufficient to ensure comprehensive protection of fundamental rights.[73] Thus, the current risk classification, especially the high-risk category, needs to be conceptualised and implemented in a more nuanced way. In this context, risk classification should not be limited to the three broad categories of high-, medium- and low-risk AI systems, in addition to the prohibited AI systems: within these categories, gradations between different sub-levels of risk would facilitate a more differentiated risk assessment. In the field of law enforcement, which is generally and rightly considered in the high-risk category, a gradual distinction from 'low-high-risk' to 'high-high-risk' should be introduced. The dangers posed by such systems for fundamental rights vary. Consequently, there is a need for an interplay of regulatory approaches without creating gaps or unclear risks in the application of AI.[74]

[70] United Nations, Open Letter from the United Nations High Commissioner for Human Rights to European Union institutions on the European Union Artificial Intelligence Act ("AI Act") of 8 November 2023, available at: https://www.ohchr.org/sites/default/files/2024-12/Tu%CC%88rk_open_letter_European_Union_highlights_issues_with_AI_Act_8_11_23.pdf (last visited 23 August 2025).

[71] *Daniel Leufer/Caterina Rodelli/Fanny Hidvegi*, Human Rights Protections... with Exceptions, Access Now of 14 December 2023, available at: https://www.accessnow.org/whats-not-in-the-eu-ai-act-deal/ (last visited 21 August 2025).

[72] *Stewart* (fn. 14), pp. 129 f.

[73] *Hannah Ruschemeier*, AI as a challenge for legal regulation – the scope of application of the artificial intelligence act proposal, in: ERA Forum 23 (2023), pp. 361–376 (373).

[74] For an attempt of a more nuanced risk classification of the area of law enforcement, see: *Kleemann/Aden* (fn. 41).

3. Direct responsibility and accountability mechanisms: late additions to the AI Act

The original proposal of the AI Act did not include a mechanism for individuals who are harmed or otherwise negatively affected by AI systems to file a complaint or seek redress.[75] In the final version, however, a new Section 4 (Remedies) in Chapter VII has been added, comprising Articles 85 and 86 AI Act. It is already foreseeable that these complaints mechanisms will serve as a channel between AI developers, deployers, users and those affected by AI-based decisions. This feature of the institutional architecture for the implementation of the AI Act is of fundamental importance in terms of AI accountability and responsibility, as Art. 85 AI Act provides the main remedy directly available to lay persons affected by AI systems under the scope of the Act, the right to lodge a complaint with the competent national supervisory authority, "[w]ithout prejudice to other administrative or judicial remedies" (Art. 85 para. 2 AI Act). The right to lodge a complaint is widely accessible as it addresses "any natural or legal persons having grounds to consider that there has been an infringement", opening the door to the possibility of initiating collective forms of legal action on the basis of the AI Act.[76]

Furthermore, according to Art. 86 para. 1 AI Act, "[a]ny affected person subject to a decision which is taken by the deployer on the basis of the output from a high-risk AI system [...] and which produces legal effects or similarly significantly affects that person in a way that they consider to have an adverse impact on their health, safety or fundamental rights shall have the right to obtain from the deployer clear and meaningful explanations of the role of the AI system in the decision-making procedure and the main elements of the decision taken". Art. 86 para. 2 AI Act contains some exceptions to this right to an explanation. The extent to which this provision can provide actual protection against a decision taken by a law enforcement authority requires further analysis. Furthermore, the role of national courts and domestic (constitutional) law will be of utmost importance, as fundamental rights violations by national law enforcement agencies will most likely be heard there first.

The analysis of the AI Act's potential for protecting fundamental rights in the context of using AI tools for law enforcement purposes requires examining the institutional architecture that the European Union and its member states are progressively establishing. Accountability and responsibility mechanisms will develop within this complex multi-level institutional environment, which in practice will affect the scope of protection theoretically

[75] *Palmiotto* (fn. 6), pp. 778 f.; European Digital Rights, The EU AI Act and fundamental rights: Updates on the political process, EDRi of 9 March 2022, available at: https://edri.org/our-work/the-eu-ai-act-and-fundamental-rights-updates-on-the-political-process/ (last visited 21 August 2025); European Digital Rights, Civil society calls on the EU to put fundamental rights first in the AI Act, EDRi of 30 November 2021, available at: https://edri.org/our-work/civil-society-calls-on-the-eu-to-put-fundamental-rights-first-in-the-ai-act/ (last visited 21 August 2025); *Veale/Borgesius* (fn. 46), p. 111.

[76] European Center for Not-for-Profit Law, Towards an AI Act that serves people and society: Strategic actions for civil society and funders on the enforcement of the EU AI Act, ECNL of August 2024, available at: https://ecnl.org/sites/default/files/2024-09/AIFUND_ECNL_AI_ACT_Enforcement_2024.pdf, (last visited 21 August 2025), p. 39.

offered to affected persons. Given the nature of the challenges in these sensitive fields of AI application—where the AI Act provides numerous exceptions or 'backdoors' benefiting public security agencies using AI—we assess how AI contestability could legally, formally, or informally, and spontaneously arise in response to these issues.

III. Accountability Mechanisms under the AI Act and AI contestability

Responsibility and accountability under the AI Act require not only an institutional architecture and oversight authorities that can effectively monitor and sanction compliance with its obligations, but also effective substantive rights that at least enable affected persons and laypersons to contest AI-based decisions, particularly those involving the use of AI for law enforcement purposes. In this regard, NGOs have criticised the right to lodge a complaint and Section 5 in Chapter IX on remedies in the AI Act for lacking teeth, stating that "it remains unclear how effectively these [supervisory] authorities will be able to enforce compliance and hold violators accountable".[77] Regarding the future implementation of remedies, the multiplicity of oversight bodies may further weaken the effectiveness of legal means of contestation of AI-based decisions, activities and processes. According to expert consultations organised within the framework of 'Accountability Principles for Artificial Intelligence (AP4AI) in the internal security domain', the "[principle of enforceability] requires that relevant oversight bodies and enforcement authorities have the necessary power and means to respond appropriately to instances of non-compliance with applicable obligations by those deploying AI in a criminal justice context".[78]

This so-called 'many eyes phenomenon'[79] in the context of AI regulation[80] needs to be considered, with respect to the complex institutional architecture set up by the AI Act.[81] This simply means that the effective implementation of responsibility and accountability in relation to AI and its use for law enforcement purposes may be seriously hampered by the multiplicity of European and national authorities that will be responsible for controlling compliance with the requirements of the AI Act. This illustrates the challenges that lie

[77] European Digital Rights and AI coalition partners, EU's AI Act fails to set gold standard for human rights, EDRi of 3 April 2024, pp. 3 f., available at: https://edri.org/our-work/eu-ai-act-fails-to-set-gold-standard-for-human-rights/ (last visited 21 August 2025); Access Now, The EU AI Act: a failure for human rights, a victory for industry and law enforcement, 13 March 2024, available at: https://www.accessnow.org/press-release/ai-act-failure-for-human-rights-victory-for-industry-and-law-enforcement/ (last visited 21 August 2025).

[78] *Babak Akhgar et al.*, Accountability Principles for Artificial Intelligence (AP4AI) in the Internal Security Domain, AP4AI Framework Blueprint. Accountability Principles for Artificial Intelligence (AP4AI), 2022, p. 38.

[79] *Mark Bovens*, Analyzing and Assessing Accountability: A Conceptual Framework, in: European Law Journal 13 (2007), pp. 447-468 (455 ff.).

[80] *Claudio Novelli/Mariarosaria Taddeo/Luciano Floridi*, Accountability in artificial intelligence: what it is and how it works, in: AI & Society 39 (2024), pp. 1871-1882 (1875).

[81] See also for this issue: *Sol Martinez Demarco/Milan Tahraoui/Steven Kleemann*, The siloed logic and the implementation of the Artificial Intelligence Act in the law enforcement context: legal and ethical analysis of the applicability of accountability and responsibility to high-risk AI systems, Routledge Studies in Surveillance (forthcoming).

ahead in creating and embedding concrete mechanisms of accountability and responsibility in the practice of sensitive areas of law enforcement activities.

Firstly, we refer to the two main types of supervisory bodies qualified by the AI Act as the Member States' national competent authorities. We briefly introduce the so-called market surveillance authorities and the notifying authorities, as well as notified bodies for conducting conformity assessments. As we later explain, although the term 'national competent authorities' is generic, it actually adds another layer to the 'many eyes' problem. Indeed, this term is used in different ways, referring either to (i) data protection authorities with additional tasks, competences and means, (ii) newly established bodies responsible for implementing the Act at the national level, or (iii) to several other possible competent independent national public authorities, if needed.[82] Finally, we will frame the interests of considering AI contestability by examining the possible existence of a 'right to contest' in the context of AI, also distinguishing between corrective and non-corrective forms of contestability. The argument is that AI contestability is of critical importance, as the AI Act has significant shortcomings regarding accountability and responsibility for the use of AI for public security purposes.

1. *Insights on the AI Act Complex Governance Architecture and its Role for AI Accountability and Responsibility*

With regard to accountability and responsibility for the use of AI in law enforcement, as well as migration, asylum and border management control purposes, two key institutions established by the AI Act are the market surveillance authorities and the notifying bodies, which will act as national competent authorities (Art. 70 para. 1 AI Act). In particular, they will be responsible for ensuring that the use of AI systems for public security purposes does not compromise the health, safety and fundamental rights,[83] even if the AI Act also confers powers to other institutions for enforcing the AI Act in line with fundamental rights.

Two important powers given to national authorities to ensure that high-risk AI systems in law enforcement comply with the AI Act are notable. First, Article 74 para. 2 AI Act states that these authorities should have full access to the documentation and datasets used for developing such systems, including training, validation, and testing data. This access can be provided through APIs or other secure remote access methods, as long as it is necessary for their tasks. Second, the AI Act stipulates that market surveillance authorities overseeing high-risk AI systems in biometrics—when used for law enforcement, migration, asylum, border control, justice administration, or democratic processes—should have strong investigative and corrective powers (see also Recital 159 AI Act). These include the ability to access all personal data being processed and any other information needed to carry out their duties. One exception to these powers involves the previously mentioned category of 'sensitive operational data'.[84]

The development of a complex institutional architecture for the implementation

[82] See below, for example, the independent public authorities that France has designated on the basis of Art. 77 AI Act.

[83] *Marion Ho-Dac*, La protection des droits fondamentaux dans l'AI Act: Essai de cartographie critique, in: RTD. Euro. 2025, pp. 615-633.

[84] See Section II. 2.

of the AI Act has been informed by debates and criticism raised about serious enforcement issues that affect the implementation of other, already-established regulations, such as the GDPR, also leading to a deficit in terms of responsibility and accountability.[85] One frequent critique concerns the difficulties to enforce the GDPR across the EU, as the EU Member States' national data protection authorities have divergent legal and political preferences.[86] Indeed, significant challenges lie ahead regarding accountability and responsibility for the development and deployment of AI systems in law enforcement contexts. This is particularly pertinent given that EU Member States have demonstrated a propensity to advocate for a reduced set of obligations.[87] Consequently, such preferences may influence the institutional framework of oversight mechanisms within law enforcement domains, which are predominantly shaped by the decisions of EU Member States. 'Excessive accountability'[88] is an interesting concept for criticising this complex architecture from the viewpoint of fundamental rights protection, as it describes the accumulation and network of accountability mechanisms that have produced negative side effects in terms of increasing costs, red tape, and a deterioration of public values such as effectiveness, efficiency, trust, and learning.[89]

2. AI Contestability

As the rapid diffusion of AI in the fields of law enforcement activity and beyond is strongly pushed by states' authorities and dominant firms, without much active role conferred to affected persons and consideration for the broad public,[90] the

[85] In this sense, see for example *Yiran Lin*, More Than an Enforcement Problem: The General Data Protection Regulation, Legal Fragmentation, and Transnational Data Governance, in: Columbia Journal of Transnational Law 62 (2024), pp. 1-39 (20 ff.).

[86] See for instance, *Giulia Gentile/Orla Lynskey*, Deficient by Design? The Transnational Enforcement of the GDPR, in: International and Comparative Law Quarterly 71 (2022), pp. 799–830 (818). *Contra*, a member of the European Parliament involved in the legislative process for this Act, claims that the AI Act will achieve a better level of enforcement due to the quality of the future market surveillance authorities. See, *Laura Caroli*, Will the EU AI Act work? Lessons learned from past legislative initiatives, future challenges, IAPP News of 17 April 2024, available at: https://iapp.org/news/a/will-the-eu-ai-act-work-lessons-learned-from-past-legislative-initiatives-future-challenges (last visited 21 August 2025).

[87] *Palmiotto* (fn. 6); 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 – General approach, 2021/0106(COD), 14336/22 of 11 November 2022, para. 4.2.

[88] *Mark Bovens/Thomas Schillemans*, Meaningful Accountability, in: Mark Bovens/Robert Goodin/Thomas Schillemans (ed.), The Oxford Handbook of Public Accountability, 2014, pp. 673-682 (674).

[89] Remarkably, the proposed Digital Omnibus on AI (European Commission (fn. 5)) aims to "simplify" the supervisory architecture primarily by granting more centralized powers to the AI Office for regulating general-purpose AI models with broad EU-wide impact (pp. 27 f.). Amendments are also envisaged for the supervision of fundamental rights by national authorities and their cooperation, but to a much lesser extent (pp. 29 f.). Essentially, the proposal will further reduce the scope of AI-based activities that can be supervised by reducing applicable obligations for high-risk systems or delaying the applicability of parts of the AI Act (pp. 2, 21 ff.).

[90] See for example, *Marie Petersmann/Dimitri Van Den Meerssche*, On phantom publics, clusters, and collectives: be(com)ing subject in algorithmic times, in: AI & Society 39 (2024), pp. 107–124; *Chris Jones/Romain Lanneau*, Automating Authority: Artificial Intelligence in European Police and Border Regimes, Statewatch of April 2025, available at:

contestability of AI development and deployment is a useful analytical concept for addressing AI accountability and responsibility. Although AI contestability is also partly embedded in international human rights law, including the right to an effective redress, it extends beyond this, encompassing social engagement practices involving AI socio-technical use-cases, even in the absence of anticipated or actual harm. If the risk-based approach had a strong bearing on the negotiation and the adoption of the AI Act, a rights-based approach can command the legal and ethical organisation and reaction to corrective and non-corrective forms of AI contestability. As the 2021 UNESCO Recommendations on the ethics of AI clearly state, it "should be recognized that AI technologies do not necessarily, per se, ensure human and environmental and ecosystem flourishing".[91] This statement reflects a human rights approach that seeks to precondition the use of AI technologies to specific justifications for its use in light of several criteria in which appropriate, proportional and legitimate aims, as well as human rights and rigorous scientific foundations play a key role.[92] As AI can impact societies in which it is deployed far beyond the mere individual situations of persons directly and effectively affected or harmed by a particular use case,[93] it is necessary to reflect on the potential contribution that a right to effective contestation of the use of AI for law enforcement purposes might bring, while contemplating the limits of a perspective arguably excessively focused on individuals[94] and on corrective forms of contestation.

According to the first paradigm, the increasing use of AI-based systems necessitates a stronger emphasis on individuals' right to challenge decisions that affect their lives. This right arises not only from legal frameworks such as the GDPR and parts of the AI Act (Art. 85 ff. AI Act), after their introduction by amendments proposed by the European Parliament to take better into account affected persons,[95] but also from broader human rights principles and procedural guarantees enshrined in democratic legal systems. An important example is the right to an effective remedy under international human rights law[96]

https://www.statewatch.org/automating-authority-artificial-intelligence-in-european-police-and-border-regimes/ (last visited 22 August 2025); *Anaëlle Beignon/Thomas Thibault/Nolwenn Maudet*, Imposing AI: deceptive design patterns against sustainability, Limits '25, 11th Workshop on Computing Within Limits, 26–27 June 2025, available at: https://computingwithinlimits.org/2025/papers/limits2025-beigon-imposing-ai.pdf (last visited 21 August 2025).

[91] UNESCO, Recommendation on the ethics of artificial intelligence, SHS/BIO/REC-AIETHICS/2021 of November 2021, para. 25.

[92] Ibid., para. 26. This paragraph stresses that: "In particular, AI systems should not be used for social scoring or mass surveillance purposes".

[93] *Smuha/Yeung* (fn. 10), p. 258.

[94] *Rinaldi/Teo* (fn. 15), p. 78; European Center for Not-for-Profit Law (fn. 76), p. 39.

[95] European Parliament, Artificial Intelligence Act, Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD)), P9_TA(2023)0236 of 14 June 2023, Amendments 628-630; *Palmiotto* (fn. 6), pp. 783 f.

[96] Art. 8 Universal Declaration of Human Rights of 10 December 1948, UN Doc. A/RES/217 A (III); Art. 2 para. 3 International Covenant on Civil and Political Rights of 16 December 1966, UNTS vol. 999, p. 171; Art. 13 European Convention on Human Rights of 4 November 1950, ETS No. 005; Art. 2 lit. c) Convention on the Elimination of All Forms of Discrimination against Women of 18 December 1979, UNTS vol. 1249, p. 13; Art. 6 International Convention on the Elimination of Racial Discrimination of 7 March 1966, UNTS vol. 660, p. 195.

and emerging international regulatory AI frameworks. In this sense, the contestability of AI-based decisions contributes to preserve fairness, serve justice and autonomy, while at the same time correcting errors, preventing unfair outcomes and improving transparency.

a) A right to contest and the concept of corrective contestability

An effective 'right to contest' certain AI-based decisions can serve as a fundamental mechanism for correcting the asymmetrical power relations created by algorithmic decision-making systems. In this sense, contestability is the ability to appeal or effectively complain about decisions made by systems, which is essential for ensuring agency and fairness in digital environments. This means that contesting a decision is not only a legal or procedural necessity, but a design concept anchored in core principles of AI regulation such as transparency, accessibility and autonomy.[97] When contestability is embedded in the design of decision-making systems, it can serve as a bridge between users, affected parties and systems, offering individuals opportunities to actively engage with and influence decisions that affect their lives.[98] This shows that this area also goes beyond individual considerations of specific legal, ethical or technical aspects and must be viewed holistically within and across individual disciplines. If a right to contest is understood and implemented in this way, it can contribute to serving principles such as fairness and justice to a greater extent and uphold constitutional values by correcting errors, and preventing or changing unfair outcomes retrospectively. This can arguably also lead to more predictable and consistent decisions.[99]

From a European legal perspective, one can refer to the GDPR (which operationalises data protection aspects of the right to privacy enshrined in Article 8 of the Charter of Fundamental Rights of the European Union)[100] to the case law of the European Court of Justice (ECJ) interpreting those provisions, and, where appropriate, to the OECD Council's non-binding recommendations on AI.[101] In its recommendations on AI, the OECD states that, in the context of transparency and explainability, it must also be possible "to provide information that enable those adversely affected by an AI system to challenge its output".[102] Although the contextual definition of the term 'challenge' is not explained in greater detail, the OECD's recommendations have already shaped data protection laws around the world on many occasions in the past, and its recommendations on AI can be influential.[103] However, it is notable that the focus is on *output* and *outcome*,[104] rather than other

[97] *Robert Patrick Collins/Johan Redström/Marco Rozendaal*, The Right to Contestation: Towards Repairing Our Interactions with Algorithmic Decision Systems, in: International Journal of Design 18 (2024), pp. 95-106 (96 ff.).

[98] Ibid., pp. 97 ff.

[99] See: *Margot E. Kaminski/Jennifer M. Urban*, The Right to Contest AI, in: Columbia Law Review 121 (2021), pp. 1957-2048 (1974 f.).

[100] Charter of Fundamental Rights of the European Union of 14 December 2007, OJ C 303/1.

[101] OECD, Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449 of 22 May 2019.

[102] Ibid., p. 9.

[103] *Kaminski/Urban* (fn. 99), p. 1963.

[104] See for the latter, OECD, Advancing accountability in AI: Governing and managing risks throughout the lifecycle for trustworthy AI, OECD Digital Economy Papers No. 349, February 2023, p. 32: "Users of explainable AI systems benefit from being able to understand and challenge or contest an outcome, seek redress, and earn through human-computer interfaces".

possible subjects of contestation.[105] Data protection law stipulates that, pursuant to Art. 22 para. 1 GDPR, individuals must not be subject to decisions based solely on automated processing, and pursuant to Art. 22 para. 3 GDPR, the controller "shall implement suitable measures to safeguard the data subject's rights and freedoms and legitimate interests[,] [...] at least the right [...] to *contest* the decision". In parallel, Art. 11 LED stipulates for the area of criminal law enforcement that "a decision based solely on automated processing, including profiling, [...] [is] prohibited [...]" and that the data subject shall have at least the right to obtain human intervention. Recital 38 LED then states that "in any case, such processing should be subject to suitable safeguards, including [...] the right to obtain human intervention, [...] to obtain an explanation of the decision reached [...] or to *challenge the decision*". Thus, it can be argued that a right to contest can exist at least against certain forms of processing of personal data, namely fully automated data processing. This is likely to apply to a significant proportion of AI-based decisions that affect individuals at an individual level. However, there is some uncertainty regarding the wording "based solely on automated processing" and the requirement that the decision "produces legal effects [...] or similarly *significantly affects*" (Art. 22 para. 1 GDPR). It is not entirely clear what thresholds apply here. Furthermore, if a right to contest is really intended to exist and be enforceable, it is odd that it is 'hidden' in such a place and does not appear elsewhere in the text of the Regulation, nor is it explicitly defined or explained in more detail. Finally, the AI Act has been criticised for focusing excessively on AI providers regarding human oversight (Art. 14 AI Act), rather than on the key role of the deployers in ensuring human intervention, while simply requiring awareness of automation biases but no real obligation to act upon it.[106]

For such a right to be effective, it would have to go beyond a mere right to rectification and at least include an obligation to examine the merits of a complaint and to give reasons for a decision, and that this right requires from the data controller to either make the automated decisions effectively contestable or to discontinue the use of the algorithmic decision-making system altogether.[107] Furthermore, individual rights, procedural rights and transparency rights contained in the GDPR must be taken into account in their entirety to enable an effective right to contest.[108] This means that, at least from the Art. 22 GDPR[109] in conjunction with the Arts. 13, 14, 15 GDPR[110] and Recital 71

[105] See however the slightly broader take on AI contestability in OECD, AI, Data Governance and Privacy: Synergies and Areas of International Co-Operation, OECD Artificial Intelligence Papers No. 22, June 2024, p. 39: "when it comes to helping persons affected by AI systems understand and contest their processes and outputs, or to help users detect algorithmic discrimination, data protection law and AI policy align".

[106] *Laux/Ruschemeier* (fn. 9).

[107] *Emre Bayamlıoğlu*, The right to contest automated decisions under the General Data Protection Regulation: Beyond the so-called "right to explanation", in: Regulation & Governance 16 (2022), pp. 1058-1078 (1063).

[108] *Kaminski/Urban* (fn. 99), p. 1979.

[109] European Data Protection Board, Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models, 18 December 2024, available at: https://www.edpb.europa.eu/our-work-tools/our-documents/opinion-board-art-64/opinion-282024-certain-data-protection-aspects_en (last visited 21 August 2025).

[110] Ibid., para. 63.

GDPR a right to an explanation,[111] or even a right to contest, can be derived.

Moreover, the concept of a right to contest goes beyond the scope of European data protection and privacy regulations and legislative efforts. Requirements such as accountability and transparency are also found in the AI Act, which (in addition to ensuring product safety and compliance with EU law) also aims to protect fundamental rights. As outlined above, the right to contest AI-based decisions is arguably a cornerstone for ensuring fairness, justice and accountability in an increasingly automated world. Based on the GDPR, the AI Act and the broader human rights framework, a right to contest should enable individuals to challenge decisions that affect them and hold AI providers and AI deployers accountable. To be effective, challenge mechanisms must include transparency, human control and clear legal remedies so that individuals can effectively exercise their autonomy. As AI systems influence important decisions, mechanisms for contestation are indispensable tools for upholding the rule of law and addressing the ethical and societal challenges arising from algorithmic decisions. However, the difficulty lies in implementing these rights in practice, as AI systems often lack the necessary transparency to enable individuals to understand, or to challenge, their outcomes. In this regard, the ECJ has issued a remarkable landmark ruling on the transparency of algorithms, confirming the existence of a ‘right to an explanation’ in relation to automated decisions.[112] The ruling also clarified that courts and competent authorities have access to information protected by trade secrets, where necessary, to reconcile this protection with the fundamental rights of the individuals concerned – a finding that will have significant implications, particularly for organisations using high-risk AI systems in decision-making processes affecting individuals.

Thus, a right to contest, object or appeal can be derived from the aforementioned standards of the GDPR, the AI Act, the OECD recommendations and the broader human rights legal framework. In order to enable effective contestation, principles such as transparency, human intervention and explainability must be guaranteed. From a human rights perspective, this is crucial to enable individuals to effectively assert their rights, which is predicated on their awareness of algorithmic decisions or, where applicable, profiling, and their effective ability to challenge the reasons for algorithmic decisions.[113] This means that transparency is particularly important. Transparency should provide context-specific and comprehensible reasons for decisions and allow for case-specific discretion.[114] In order to achieve such transparency, it is also necessary to involve civil society actors, including those

[111] *Bryan Casey/Ashkon Farhangi/Roland Vogl*, Rethinking Explainable Machines: The GDPR’s ’Right to Explanation’ Debate and the Rise of Algorithmic Audits in Enterprise, in: Berkeley Technology Law Journal 34 (2019), pp. 143-188 (155 ff.).

[112] ECJ, judgment of 27 February 2025, Case C-203/22, paras. 57 ff. (Dun & Bradstreet Austria).

[113] See with respect to the awareness of data subjects and their reasonable expectations concerning the processing of their data, European Data Protection Board (fn. 109), paras. 93 ff.

[114] See for instance, *Rita Matulionyte*, Increasing transparency around facial recognition technologies in law enforcement: towards a model framework, in: Information & Communications Technology Law 33 (2024), pp. 66–84; *Evgeni Aizenberg/Jeroen van den Hoven*, Designing for human rights in AI, in: Big Data & Society 7 (2020), pp. 1-14 (7).

who could take on the task of challenging decisions in certain contexts, in order to develop an understanding of what constitutes a comprehensible justification and an effective challenge mechanism.[115] Finally, transparency can also support corrective and non-corrective forms of AI contestability. Regrettably, transparency requirements have been significantly reduced during the negotiations of the AI Act in the case of high-risk AI systems used for public security purposes.[116]

It can be argued that AI contestability is not limited to the right to contest, which can be qualified as corrective AI contestability. In much of the literature dealing with AI contestability in relation to social systems theories, AI contestability is considered a means of potentially improving the development, functioning, deployment, and efficiency of AI systems or models. AI contestability is, therefore, mainly considered from the perspective of the internal logic of AI development and AI deployment to improve its use and mitigate its side effects.[117] One example is the contestability of the acceptability of an AI system's error rate.[118] Furthermore, these strands of literature often focus on the key role of AI developers, conceptualised as an interplay between normative principles and the translation of rules into technical design. Actors outside technical development also tend to be considered as mere recipients or users of AI contestation. A general sensibility for ethical challenges within AI software design can be identified in the perspectives following a corrective AI contestability conception, which is supposed to lead to technology that better protects the rights, interests, and needs of affected communities and persons. Lastly, corrective AI contestability generally intervenes *ex post facto*, which complicates *ex ante* forms of contestation, such as questioning or opposing the development of an AI tool from the outset.[119]

The concept of contestation intervenes in various ways in international and regional AI regulatory frameworks, but most occurrences covered tend to be corrective in nature, as legal mechanisms are anchored in situations where individuals can contest the mere outcomes of AI-based decision-making or the dysfunctionality of AI systems and models.

b) Corrective and non-corrective AI contestability

Against this backdrop, we argue that the provisions of the AI Act read in the broader context of international human rights law also play a role with regard to non-corrective forms of AI contestation, which integrate a second distinct perspective on AI contestability particularly useful to address challenges stemming from the use of AI for law enforcement purposes. By definition, non-corrective AI contestability does not seek to enhance the performance, underlying logic or objectives of an AI system, model or tool. Non-corrective AI contestability can take different forms, including spontaneous contestation as

[115] Ibid.

[116] *Palmiotto* (fn. 6), pp. 790 ff.

[117] *Simon Hirsbrunner/Steven Kleemann/Milan Tahraoui*, Contestation in artificial intelligence as a practice: from a system-centered perspective of contestability towards normative contextualization, situative critique and organizational culture, in: Frontiers in Communication 10 (2025), pp. 1-15 (8).

[118] *Claudia Aragau*, Error, in: Mareile Kaufmann/Heidi Mork Lomell (ed.), De Gruyter Handbook of Digital Criminology, 2025, pp. 215-221 (216, 219).

[119] *Gianclaudio Malgieri/Frank Pasquale*, Licensing high-risk artificial intelligence: Toward ex ante justification for a disruptive technology, in: Computer Law & Security Review 52 (2024), pp. 1-18.

social practice or 'techno-resistance',[120] and can pursue various objectives, such as preventing an AI development or AI deployment project altogether, or targeting more specific deployment contexts or modalities of use.

In the context of the predominant AI regulatory approaches based on risks, this perspective on non-corrective AI contestability is particularly useful and relevant for analysing social practices in light of the relationship between legal imaginaries, traditional modes of law-making and innovation, especially given the fact that most of AI regulatory frameworks focus on the concepts of AI trustworthiness and acceptability. According to the 2019 AI High-Level Expert Group, trustworthiness can be defined as lawful, ethical, and robust AI (technically and socially speaking) throughout the AI lifecycle.[121] Among the requirements suggested by these experts, which influenced the 2021 EU Commission Proposal for an AI Regulation,[122] one refers to accountability and includes the criteria of "auditability, minimisation and reporting of negative impact, trade-offs, and redress".[123] The emphasis placed on trust and trustworthiness in various international and regional AI regulatory frameworks translates into a general objective of inducing people to trust AI, innovation and to use this technology, thereby unlocking its economic and societal potential. However, one fundamental issue is that trust cannot be commanded; it requires fulfilment of the necessary conditions. It has been argued that constant appeals to the concept of trustworthiness can lead to confusion between this concept and acceptability.[124] This problem is exacerbated by the fact that trust is meant as an expert domain under the AI Act and other regulatory frameworks. AI contestability as a social practice is therefore relegated to the background or ignored entirely. Yet one might surmise that contestability is necessary in a democratic context, and inevitable in a non-democratic context. Overreliance on trust and the risks stemming from AI opacity create a basic need for institutions, mechanisms, norms, and cultures that enable effective contestation. Contestability is especially crucial for prohibited AI practices and high-risk AI systems used for security purposes. This is arguably due to the coercive nature of AI tools developed in this context and the power imbalances they generate when deployed, given the involvement of State authorities and the often dominant private corporations.[125] A major advantage of the concept of AI contestability is that it incorporates both 'institutional' and 'formal rules-based reactions' to AI development and deployment, as well as 'spontaneous, informal and cultural interactions and responses'.

[120] See for instance, *Marie Petersmann*, Refusing Algorithmic Recognition, in: European Journal of International Law 35 (2024), pp. 979-989 (986 f.).

[121] High-Level Expert Group on Artificial Intelligence, Ethics Guidelines for Trustworthy AI of 8 April 2019, available at: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai (last visited: 12 December 2025), p. 5.

[122] European Commission, 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 of 21 April 2021.

[123] High-Level Expert Group on Artificial Intelligence (fn. 121), p. 14.

[124] *Johann Laux/Sandra Wachter/Brent Mittelstadt*, Trustworthy artificial intelligence and the European Union AI Act: On the conflation of trustworthiness and acceptability of risk, in: Regulation and Governance 18 (2024), pp. 3-32 (27).

[125] Ibid., see also Art. 7 para. 2 lit. h) and Recitals 59-60 AI Act.

Unlike corrective forms of AI contestability, the concept of non-corrective contestability is not limited to situations where something might go wrong. This enables criticism of the very foundations of an AI development or deployment project to be taken into account, while offering a better focus on systemic risks relating to sensitive intended AI use cases and how they are perceived by legal imaginaries and traditional modes of law-making. Non-corrective AI contestability can pursue various objectives, such as preventing an AI development or deployment project altogether, or targeting more specific deployment contexts or modalities of use. Non-corrective AI contestability is far less anchored in an *ex post* logic, thus offering more critical space for *ex ante* forms of friction, opposition, resistance or refusal.[126] For instance, a key discussion about AI contestability is whether it is possible to effectively contest the very idea of developing or acquiring an AI system or model. A notable example of this is the contestability of AI procurement procedures, given that one of the initial stages of procurement entails identifying the actual needs behind the procurement of such systems or models, along with their requirements.[127] This is particularly difficult in fields of AI development and deployment in the private and public security sectors, as third parties – and even less the broader public – are rarely involved.

First, there are limits to corrective AI contestability, such as empirical or legal ones. For example, Art. 86 AI Act establishes a right to an explanation of individual AI-based decision-making. However, its contestability potential is significantly limited by the fact that this provision only entitles affected persons to obtain clear and meaningful explanations from the deployer regarding the role of the AI system in the decision-making process and the main elements of the decision made. This right to explanation provides a significant legal mechanism that could potentially improve understanding of whether an AI-based decision took place, but also on what basis and according to which rules. There can be, therefore, a potential for contestation. This is, however, only the case indirectly and *ex post facto*, and more generally, without allowing the questioning of the foundations of those AI-based decisions. In fact, there is also a right to lodge a complaint on the basis of Art. 85 AI Act. That said, the AI Act does not actually grant a right to directly contest the development or deployment of an AI system. Rather, this Regulation establishes a right to a mediated formal, legal and institutional contestation, i.e., through an expert or official representation (e.g. oversight bodies). What is then particularly problematic is the fact that the AI Act does not impose on market surveillance authorities either to report on how they handle complaints, or to provide the possibility to 'appeal' their decisions. Indeed, over the course of the negotiations, the European Parliament wished to introduce a "right to an effective judicial remedy against a national supervisory authority",[128] in addition to the right to lodge a complaint. However, the final version of the AI Act

[126] *Petersmann* (fn. 120), pp. 983 f.

[127] Digital Regulation Cooperation Forum (DRC), Transparency in the procurement of algorithmic systems: Findings from our workshops, 2023, available at: https://www.drcf.org.uk/siteassets/drcf/pdf-files/transparency-procurement-algorithmic-systems.pdf?v=381844 (last visited 23 August 2025).

[128] European Parliament, Amendments adopted by the European Parliament on 14 June 2023 on the proposal of the European Parliament and the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), (Ordinary legislative procedure: first reading), OJ C/2024/506 of 23 January 2024, Amendment 629.

has retained a system of remedies that confers more responsibility on the surveillance market authorities, while reducing possibilities for complainants to act upon these authorities' decisions.[129]

Secondly, there are also empirical limits observable in how AI-based technologies transform security politics, affect human rights protection and more generally reshape political and social interactions. For example, the introduction of AI technologies has led to a new surveillance logic in the areas of migration, asylum and border management control, which the AI Act generally classifies as high-risk (Recital 60 AI Act). In this context, there are significant challenges to AI contestability, as the logic of AI-based border surveillance systems aims to discover 'unknown unknowns' not solely on the basis of pre-defined risk concepts and categories, but also based on a "*dispositif* of pre-emptive security or speculative suspicion"[130] or "inferred attributes"[131] that are used to generate fluid categories for sorting persons under surveillance. Similarly, *Rinaldi* and *Teo* argue that "the deployment of AI-driven border and migration management may be challenging the idea of individual empowerment which lies at the core of the human rights protection framework",[132] through datafication,[133] inference and construction[134] as well as algorithmic groupings.[135] In light of this, how can an AI-based socio-technical system that is developed or deployed for public security purposes—and whose underlying logic is difficult to understand—be meaningfully contested? When the system's secrecy is protected either as a matter of public security policy[136] or as a trade secret under public-private partnership agreements, the inherent opacity of such a socio-technical system can sharply shrink the room for corrective interventions.

IV. Conclusion

The EU AI Act marks a pivotal step towards governing AI systems, yet its application to law enforcement contexts remains fraught with ambiguity and competing priorities. While the AI Act's risk-based architecture offers a structured baseline, the introduction of exemptions ('backdoors') for security agencies dilutes that clarity. These carve-outs allow high-risk technologies, including biometric surveillance and predictive policing tools, to bypass safeguards that would otherwise apply to private actors, thereby widening the gap between the Act's stated commitment to fundamental rights protection and its practical enforcement.

Our analysis shows that the current accountability regime — comprising provider-

[129] See in this sense, European Center for Not-for-Profit Law (fn. 76), p. 39; *Griff Ferris/Sofia Lyall*, New Technology, Old Injustice. Data-driven discrimination and criminalisation in police and prisons in Europe, Statewatch of June 2025, available at: https://www.statewatch.org/news/2025/june/police-racism-and-criminalisation-across-europe-increasingly-fuelled-by-digital-prediction-and-profiling-systems/ (last visited 21 August 2025), p. 20.

[130] *Sullivan/Van Den Meerssche* (fn. 21).

[131] *Amoore* (fn. 21).

[132] *Rinaldi/Teo* (fn. 15), p. 78.

[133] Ibid.

[134] Ibid., p. 79.

[135] Ibid., p. 80.

[136] Statewatch, EU's secretive "security AI" plans need critical, democratic scrutiny says new report, Statewatch of 29 April 2025, available at: https://www.statewatch.org/news/2025/april/eu-s-secretive-security-ai-plans-need-critical-democratic-scrutiny-says-new-report/ (last visited 23 August 2025).

centric risk-management obligations, limited post-market monitoring, and a nascent fundamental rights impact assessment — does not fully empower affected individuals to contest AI-driven decisions. The reliance on providers to embed safeguards, coupled with the limited scope of supervisory authorities, creates a systemic asymmetry: law enforcement bodies can deploy powerful AI tools while citizens face procedural hurdles to obtain redress.

Meaningful protection of fundamental rights will depend on several interrelated reforms. First, exemptions should be tightened, and 'backdoor' provisions need to be narrowed or eliminated so that law enforcement applications are subject to the high-risk requirements intended for them. Second, contestability mechanisms must be strengthened by expanding the scope of the fundamental rights impact assessment, guaranteeing transparent documentation, and providing individuals with enforceable rights of review and remediation. Third, the responsibilities of providers and deployers should be clarified, with precise duties assigned to each participant in the AI supply chain, especially where deployers such as police forces possess superior contextual knowledge of the risks. Finally, a rights-based overlay ought to be embedded within the risk-based framework, ensuring that fundamental liberties are treated as non-negotiable rather than merely a variable in a cost-benefit analysis. Only by reconciling the Act's technical risk calculus with a robust, rights centred accountability architecture can the EU ensure that AI enhances public safety without eroding the democratic values it seeks to protect.

Vitae

Steven Kleemann is a doctoral researcher at the Faculty of Law at the University of Potsdam as well as a researcher and policy advisor on digitalisation, AI & human rights at the German Institute for Human Rights. At the time of writing this article, he was a researcher at the Berlin Institute for Safety and Security Research (FÖPS Berlin), working on a project concerning legal aspects of trustworthy AI for police applications. His research focuses on international law, human rights, AI, and security law.

Milan Tahraoui is a doctoral researcher associated with the Centre Marc Bloch, as well as a Ph.D. candidate at both the Paris 1 Pantheon-Sorbonne University and the Free University of Berlin. At the time of writing this article, he was a researcher at the Berlin Institute for Safety and Security Research (FÖPS Berlin), working on a project examining the issues surrounding the use of so-called trustworthy AI applications in law enforcement. His research focuses on international and European law, human rights, digital surveillance, AI and security law.