European Journal of Risk Regulation (2025), **16**, 770–793
doi:10.1017/err.2024.97

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

The AI Act Roller Coaster: The Evolution of Fundamental Rights Protection in the Legislative Process and the Future of the Regulation

Francesca Palmiotto

Centre for Fundamental Rights, The Hertie School, Berlin, Germany and IE Law School, Madrid, Spain
Email: f.palmiotto@hertie-school.org

Abstract

This paper traces the legislative process of the EU Artificial Intelligence Act (AI Act) to provide an empirical and critical account of the choices made in its formation. It specifically focuses on the dynamics that led to increasing or lowering fundamental rights protection in the final text and their implications for fundamental rights. Adopting process-tracing methods, the paper sheds light on the institutional differences and agreements behind this landmark legislation. It then analyses the implications of political compromise for fundamental rights protection. The core message it aims to convey is to read the AI Act with its institutional setting and political context in mind. As this paper shows, the different policy aims and mandates of the three EU institutions, compounded by the unprecedented level of redrafting and the short time needed to reach a political agreement, influenced the formulation of the AI Act. Looking forward, the paper points to the role of implementation, enforcement and judicial interpretation in enhancing the protection of fundamental rights in the age of AI.

Keywords: AI Act; EU Law; fundamental rights; Artificial Intelligence; process-tracing; law-making

I. Introduction

Artificial Intelligence (AI) is one of the most pressing challenges to the protection of fundamental rights in our modern society.[1] AI may perpetuate or create inequalities, exercise new forms of power, and erode the core tenets of democracy and the rule of law.[2] The risks that AI presents have prompted different governments and international

[1] The EU Agency for Fundamental Rights considers AI as one of the key priorities in their Strategic Plan "FRA Strategic Plan 2023–2028" (*European Union Agency for Fundamental Rights*, 17 October 2023) <http://fra.europa.eu/en/publication/2023/strategic-plan-2023-2028> accessed 29 February 2024; See also their report "Getting the Future Right – Artificial Intelligence and Fundamental Rights" (European Union Agency for Fundamental Rights 2020) <https://fra.europa.eu/en/publication/2020/artificial-intelligence-and-fundamental-rights> accessed 2 January 2022.

[2] See among many others Virginia Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor* (First Edition, St Martin's Press 2017); Mireille Hildebrandt, "Algorithmic Regulation and the Rule of Law" (2018) 376 Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 20170355; Mark Coeckelbergh, *Why AI Undermines Democracy and What To Do About It* (John Wiley & Sons 2024); Paul De Hert, "The Future of Privacy. Addressing Singularities to Identify Bright-Line Rules That Speak to Us Foreword" (2016) 2 European Data Protection Law Review 461; Jeroen Temperman and Alberto Quintavalla (eds), *Artificial Intelligence and Human Rights* (Oxford University Press 2023).

organisations worldwide to adopt ethical guidelines, soft-law instruments, international conventions and legislative measures.[3] From a global perspective, the European Union (EU) is portrayed as a leader in advancing human rights in the digital age through regulations.[4] Legal scholars refer to "digital constitutionalism" to describe the EU's normative commitment to a digital society founded on fundamental rights and constitutional values, reflected in its secondary legislation.[5] Recent legislative interventions include the Digital Services Act package,[6] the Data Act,[7] the proposed Artificial Intelligence Liability Directive[8] and the landmark Regulation on Artificial Intelligence (hereafter AI Act).[9] The AI Act, officially published in July 2024 and entered into force in August 2024, is the first comprehensive framework on AI worldwide. It ambitiously aims to foster trustworthy AI in Europe by ensuring that the development and deployment of AI systems respect fundamental rights, safety, democracy and the rule of law while supporting innovation.[10] Protecting fundamental rights from the harmful effects of AI is a core policy objective of the Regulation that seeks to achieve with a proportional risk-based approach. If AI systems pose unacceptable risks to fundamental rights, the use of such systems is prohibited. On the contrary, if an AI system poses a high risk to fundamental rights, their use is permissible, subject to requirements and safeguards. By adopting the AI Act, the EU proudly regards itself as a guarantor of fundamental rights and values in the digital age.[11]

With its recent entry into force, the AI Act will be the object of much doctrinal inquiry for the years to come.[12] This paper, however, adopts a different approach and analyses the

[3] For a mapping of policy and legislative initiatives globally see "Global AI Law and Policy Tracker" <https://iapp.org/resources/article/global-ai-legislation-tracker/> accessed 29 February 2024; "Artificial Intelligence and Democratic Values" (Center for AI and Digital Policy 2024) <https://www.caidp.org/reports/aidv-2023/> accessed 17 April 2024.

[4] Anu Bradford, *Digital Empires: The Global Battle to Regulate Technology* (Oxford University Press 2023).

[5] Giovanni De Gregorio, *Digital Constitutionalism in Europe: Reframing Rights and Powers in the Algorithmic Society* (Cambridge University Press 2022); Anu Bradford, "Europe's Digital Constitution" (2023) 64 Va. J. Int'l L. 1; Francisco de Abreu Duarte, Giovanni De Gregorio and Angelo Golia, "Perspectives on Digital Constitutionalism" in Bartosz Brożek, Olia Kanevskaia and Palka Przemyskaw (eds), *Research Handbook on Law and Technology* (Edward Elgar 2023); Angelo Jr Golia, "Critique of Digital Constitutionalism: Deconstruction and Reconstruction from a Societal Perspective" [2023] Global Constitutionalism 1.

[6] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act), OJ L 227/1 and Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act), OJ L 265/1.

[7] Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on harmonised rules on fair access to and use of data and amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 (Data Act), OJ L 2023/2854.

[8] Proposal for a Directive of the European Parliament and of The Council on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive) (COM/2022/496 final), 28 September 2022.

[9] Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), OJ L 2024/1689.

[10] Art 1(1) of the AI Act.

[11] See for instance the press release from the European Parliament "Artificial Intelligence Act: MEPs Adopt Landmark Law" (13 March 2024) <https://www.europarl.europa.eu/news/en/press-room/20240308IPR19015/artificial-intelligence-act-meps-adopt-landmark-law> accessed 26 November 2024 [quoting the MEP Dragos Tudorache "The EU has delivered. We have linked the concept of artificial intelligence to the fundamental values that form the basis of our societies"].

[12] See for instance Jonas Schuett, "Risk Management in the Artificial Intelligence Act" [2023] European Journal of Risk Regulation 1; Oriol Mir, "The AI Act from the Perspective of Administrative Law: Much Ado About Nothing?" [2024] European Journal of Risk Regulation 1; Irina Carnat, "Addressing the Risks of Generative AI for

AI Act backwards.[13] Its core aim is to understand how the AI Act evolved through the legislative process, focusing on the dynamics that led to increasing or lowering fundamental rights protection in the final text. In the first part, the paper sheds light on institutional differences and political compromises behind the adoption of landmark legislation. The second part analyses how political agreements affect fundamental rights protection and the formulation of the AI Act.

The paper is grounded on two premises. The first is that law and politics are intertwined. In this sense, this article is not a mere description of the AI Act's legislative process but an empirical and critical account of the choices made in its formation. By looking at its political and institutional context, the paper aims to provide a deepened understanding of the AI Act and support a contextualised interpretation of its core provisions.

The second is the need to consider the EU in light of its peculiar constitutional features as a supra-national order,[14] and specific policy aims.[15] Since it lacks direct competence in fundamental rights policies, the EU has often used peculiar legislative instruments, such as internal market legislation, to promote fundamental rights.[16] Additionally, while the EU has direct competence in data protection, legislative interventions may collide with Member States national prerogatives, especially when crossing other policy areas such as migration, asylum and law enforcement.[17] Finally, regulating AI systems also presents peculiar regulatory challenges,[18] such as protecting fundamental rights without hindering innovation and balancing individual rights protection with other public interests, such as national security.[19] As the paper will show, the AI Act is the result of such a balance

the Judiciary: The Accountability Framework(s) under the EU AI Act" (2024) 55 Computer Law & Security Review 106067; Emilija Leinarte, "The Classification of High-Risk AI Systems Under the EU Artificial Intelligence Act" (2024) 1 Journal of AI Law and Regulation 262.

[13] For a comprehensive account of the history of the AI Act see Nikos Th Nikolinakos, "The Proposed Artificial Intelligence Act and Subsequent 'Compromise' Proposals: Commission, Council, Parliament" in Nikos Th Nikolinakos (ed), *EU Policy and Legal Framework for Artificial Intelligence, Robotics and Related Technologies* (Springer International Publishing 2023) <https://doi.org/10.1007/978-3-031-27953-9_8> accessed 9 November 2023; For an analysis of the Regulation's overall rationale and policy aims see Tatjana Evas, "The EU Artificial Intelligence Act" (2024) 1 Journal of AI Law and Regulation 98.

[14] See generally Mark Dawson, *The Governance of EU Fundamental Rights* (Cambridge University Press 2017); Some scholars have been sceptic towards the EU in fundamental rights policies. See among others Andrew Williams, "Human Rights in the EU" in Damian Chalmers and Anthony Arnull (eds), *The Oxford Handbook of European Union Law* (Oxford University Press 2015); Dimitry Vladimirovich Kochenov, "False Accountability, Elusive Rule of Law" [2018] Verfassungsblog <https://verfassungsblog.de/false-accountability-elusive-rule-of-law/> accessed 29 February 2024; Philip Alston and JHH Weiler, "An 'Ever Closer Union' in Need of a Human Rights Policy" (1998) 9 European Journal of International Law 658.

[15] Daniel Mügge, "EU AI Sovereignty: For Whom, to What End, and to Whose Benefit?" [2024] Journal of European Public Policy 2200 [arguing that "the EU AI strategy de facto embraces a jurisdictional conception of sovereignty, meant to boost Europe's position in a global AI competition, with benefits mostly tailored to stakeholders in the EU"].

[16] The use of internal market legislation in promoting fundamental rights has been criticised in the literature. See for instance Jason Coppel and Aidan O'Neill, "European Court of Justice: Taking Rights Seriously, The" (1992) 29 Common Market Law Review 669; Alexander Somek and Alexander Somek, *Engineering Equality: An Essay on European Anti-Discrimination Law* (Oxford University Press 2011); On the contrary, see Bruno De Witte, "A Competence to Protect: The Pursuit of Non-Market Aims through Internal Market Legislation" in Philip Syrpis (ed), *The Judiciary, the Legislature and the EU Internal Market* (Cambridge University Press 2012); Vasiliki Kosta, *Fundamental Rights in EU Internal Market Legislation* (Hart 2015).

[17] See for instance in the field of data protection Teresa Quintel, *Data Protection, Migration and Border Control: The GDPR, the Law Enforcement Directive and Beyond* (Hart 2022).

[18] See among many others Araz Taeihagh, "Governance of Artificial Intelligence" (2021) 40 Policy and Society 137; Mark Coeckelbergh, "Artificial Intelligence: Some Ethical Issues and Regulatory Challenges" [2019] Technology and Regulation 31.

[19] See also Mügge (n 15).

between policy objectives and compromise among the contrasting visions of the three core institutions involved in law-making: the European Commission, the Council of the EU and the European Parliament. The title of the paper, the "AI Act Roller Coaster," metaphorically evokes the divergences among political actors on how fundamental rights ought to be protected and the way in which compromise was achieved.

Methodologically, the paper adopts process tracing, a qualitative research method used to observe causal processes and interactions.[20] Taking the opinions on the AI Act by the European Data Protection Supervisor (EDPS) and Board (EDPB) as a benchmark,[21] the paper quantifies and assesses the increase or decrease of fundamental rights protection throughout the legislative process. After a note on methods, the first part unravels each institution's distinct visions before the interinstitutional agreement, known as "trilogues." The paper then focuses on the implications of political compromises for the final text. Quantitively, it assesses whether fundamental rights standards were increased or lowered after the political negotiation between the Council, the European Parliament and the Commission. Qualitatively, it analyses the implications of such political choices for fundamental rights protection. In the conclusive remarks, the paper provides a set of recommendations to better enforce and align the AI Act with fundamental rights.

II. Tracing the AI Act

Process tracing methodology is a qualitative research method used to observe processes and interactions and draw inferences on their dynamics.[22] In his book "The Governance of EU Fundamental Rights," Mark Dawson applied process tracing methodology to the EU legislative process to illustrate how institutional interaction increases the level of rights protection in the EU.[23] Inspired by his research, I apply process tracing methodology to the legislative process of the AI Act. In this paper, I consider the legislative procedure as the process, the institutional interactions as different mechanisms within the process, the Commission, Council and Parliament as actors, and the final version of the AI Act as the outcome.

1. The process and actors

The AI Act legislative process was initiated in April 2021 when the Commission published its proposal.[24] Both co-legislators, the Council and the Parliament, discussed the text in parallel.

[20] See generally Andrew Bennett and Jeffrey T Checkel (eds), *Process Tracing: From Metaphor to Analytic Tool* (Cambridge University Press 2014); David Collier, "Understanding Process Tracing" (2011) 44 PS: Political Science & Politics 823; Derek Beach, "It's All about Mechanisms – What Process-Tracing Case Studies Should Be Tracing" (2016) 21 New Political Economy 463.

[21] EDPS-EDPB Joint Opinion 5/21 (18 June 2021) (hereafter EDPB-EDPS Opinion) <https://www.edpb.europa.eu/our-work-tools/our-documents/edpbedps-joint-opinion/edpb-edps-joint-opinion-52021-proposal_en> accessed 6 December 2024 and EDPS Final Opinion 44/23 (23 October 2023) (hereafter EDPS Final Opinion) < https://www.edps.europa.eu/data-protection/our-work/publications/opinions/2023-10-23-edps-opinion-442023-artificial-intelligence-act-light-legislative-developments_en> accessed 6 December 2024.

[22] Collier (n 20); See also Beach (n 20); Bennett and Checkel (n 20). For an application of process tracing to study EU policies see Benedetta Voltolini and Rainer Eising, "Framing Processes and Lobbying in EU Foreign Policy: Case Study and Process-Tracing Methods" (2017) 16 European Political Science 354.

[23] Dawson (n 14).

[24] Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts (COM/2021/206 final), 21.4.2021 (hereafter Commission Proposal).

The process leading the Council to form a general approach was led by the expert working groups TELECOM under three different presidencies of the Council. Between June and December 2022, Member States (hereafter MS) were invited to engage in a discussion paper with crucial policy priorities and then send final remarks before the final agreement was reached. The documents produced during these meetings are vital for understanding political interests and normative justifications underpinning their approach. Unlike the Parliament, however, the Council benefits from broader secrecy protection during their meetings. Generally, while negotiations are ongoing, the Council does not make its internal documents public until the end of the legislative process but allows access to document requests. Following my request, the Council granted me access to 90 per cent of the documents during the research phase of this article, which now, after the adoption of the AI Act, are publicly available.[25]

The Council published the "general approach" on 6 December 2022.[26] This document gives the Parliament an idea of the Council's position and aims to speed up the legislative procedure. Meanwhile, the Parliament discussed the proposal and adopted a negotiating position on 14 June 2023.[27] The general approach and the Parliament negotiating position formed the basis for the negotiations in the so-called "trilogue." The AI Act was formed mainly through trilogue negotiations, which took place over seven months, finally resulting in a provisional agreement on 8 December 2023.[28]

The trilogues are informal interinstitutional meetings which aim to reach an agreement between the three institutions. If agreement is reached, the resulting text has to be approved by the co-legislator according to the rules of procedures of each institution. Presently, as shown by Brandsma and others, "around 99% of new European laws are fast-tracked, with political compromises mostly found behind closed doors" in the trilogues.[29] Trilogues are held *in camera* and secluded from public scrutiny, raising issues of legitimacy and transparency in EU law-making.[30] Nonetheless, even without documentation of negotiations, process tracing allows us to open the "black box" of trilogues.[31] For this purpose, the paper divides the process into two temporal segments: before and after the trilogue.

Before, both co-legislators could amend, revise and propose their version of the AI Act. Sections III.1, III.2 and III.3 focuses on the pre-trilogue positions, illustrating how EU institutions conceptualise AI regulation and fundamental rights protections. After that, the resulting text represents a compromise between the different institutional visions. Section III.4 analyses how a political agreement was reached and its impact on increasing or lowering fundamental rights standards.

[25] The documents can be found in the register of Council documents <https://www.consilium.europa.eu/en/documents-publications/public-register/public-register-search/> accessed 6 December 2024.

[26] Council of the European Union General Approach on the Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, 25 November 2022 (hereafter Council Approach).

[27] 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, 14 June 2023 (hereafter Parliament Mandate).

[28] "Commission Welcomes Political Agreement on AI Act" (*European Commission*, 12 September 2023) <https://ec.europa.eu/commission/presscorner/detail/en/ip_23_6473> accessed 6 December 2024.

[29] Gijs Jan Brandsma and others, "Inside the Black Box of Trilogues: Introduction to the Special Issue" (2021) 28 Journal of European Public Policy 1.

[30] Deirdre Curtin and Päivi Leino, "In Search of Transparency for EU Law-Making: Trilogues on the Cusp of Dawn" (2017) 54 Common Market Law Review 1673.

[31] Brandsma and others (n 29) 5.

2. The Standards

Tracing the evolution of fundamental rights protection in the AI Act requires a benchmark to assess whether the legislative process increased such protection.[32] This paper uses the EDPB-EDPS Opinion[33] as a comparison baseline for two reasons.

First, both institutions are independent authorities. The EDPS, established by Regulation 2018/1725,[34] is an independent supervisory authority overseeing the processing of personal data by EU institutions and bodies. It ensures compliance with data protection laws, advises on relevant legislation and policies, and collaborates with other authorities to maintain consistency in data protection across the EU. The EDPB, established under the General Data Protection Regulation (GDPR),[35] is also an independent authority, comprising the heads of each member state's supervisory authority and the EDPS. Its primary function is to ensure the uniform application of data protection laws and provide guidance to EU institutions.

Second, both authorities have a specific mandate to ensure the respect of fundamental rights and freedoms when personal data are processed. Article 42 of Regulation 2021/1725 grants the EDPS a legislative consultation role, particularly when proposed legislation impacts individuals' data protection rights. Additionally, when a legislative proposal is "of particular importance for the protection of individuals' rights and freedoms with regard to the processing of personal data,"[36] the consultation can be coordinated between the EDPS and the EDPB issuing a joint opinion.

In their joint opinion on the AI Act, published right after the Commission's proposal, the EDPS and the EDPB highlighted their role in safeguarding fundamental rights, emphasising the importance of privacy and data protection as prerequisites for upholding other fundamental rights.[37] In their view, the AI Act supplements the GDPR in protecting "basic human rights,"[38] including the right to human dignity, non-discrimination and privacy, which are potentially affected when AI processes personal data. While the opinion generally welcomes the Commission's proposal, it also provides twenty-two recommendations to improve the protection of fundamental rights in the AI Act. The recommendations are summarised in the Table below.

	EDPB-EDPS Opinion
R1	Compliance with EUDPL in Article 1 and corresponding Recital
R2	Article 16 TFEU as the main legal basis
R3	Include international law enforcement cooperation in scope in Article 2(4)
R4	Remove the exemption from the scope in Article 83

(*Continued*)

[32] This paper took large methodological inspiration by the work of Dawson (n 14).

[33] EDPB-EDPS Opinion (n 21).

[34] Regulation (EU) 2018/1725 of the European Parliament and of the Council of 23 October 2018 on the protection of natural persons with regard to the processing of personal data by the Union institutions, bodies, offices and agencies and on the free movement of such data, and repealing Regulation (EC) No 45/2001 and Decision No 1247/2002/EC, OJ L 2018/295 (hereafter Reg 2018/1725).

[35] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC, OJ L 2016/119 (hereafter GDPR).

[36] Art 42 of Regulation 2021/1725.

[37] EDPB-EDPS Opinion (n 21) 7.

[38] EDPB-EDPS Opinion (n 21) 22.

(*Continued*)

	EDPB-EDPS Opinion
R5	Include compliance with GDPR and EUDPR in Chapter 2 "Requirements"
R6	Prohibit any type of social scoring
R7	Prohibit any automated biometric recognition in publicly available spaces
R8	Prohibit profiling in law enforcement
R9	Prohibit polygraphs in law enforcement
R10	Prohibit emotion detection
R11	Prohibit AI systems categorizing individuals from biometrics
R12	Clearly establish the independence and roles of supervisory authorities
R13	Give more autonomy to the EAIB without political influence of the Commission
R14	Clarify the scope and objectives of sandboxes & compliance with EUDPR
R15	Include a clear relation to EUDP law in the certification system
R16	Stronger link to EDPR and effective implementation of DP principles
R17	Address the rights and remedies available to individuals
R18	Initial risk assessment by provider and subsequent DPIA by user
R19	Restrict the exception to transparency obligations for law enforcement in the public database
R20	Add FRA as observer to the Board in Article 57
R21	Designate data protection authorities as national supervisory authorities in Article 59
R22	CAP with ex ante third party assessment

Using their recommendations as a baseline, Section III analyses each institutional approach to the AI Act before the trilogue and identifies whether and who followed their recommendations. Subsequently, it investigates whether their recommendations were implemented in the final text resulting from the trilogue negotiations. In this way, it is possible not only to assess the overall level of rights protection in the final text but also to tease out which actor was responsible for lowering or increasing the standards and political justifications.

III. The evolution of fundamental rights protection in the legislative process

Since the start of the legislative process, EU institutions had wildly divergent visions on the nature of the issues that AI raises for fundamental rights and the suitable regulatory framework to address them. They disagreed on the role of fundamental rights in the regulation, how they should be protected, and, most importantly, if exceptions to such protection should exist. By looking at the Commission's Proposal, the Council's general agreement and the Parliament's negotiating mandate, three different visions of AI emerge: (1) the AI Market of values, (2) the Trade-off AI and (3) the Human-centric AI.

1. The commission: The AI "Market of Values"

The proposal by the Commission can be framed as internal market legislation with injected public values. The core aim is to achieve an AI market that complies with Union values and

public interests, including protecting fundamental rights.[39] Despite the primary objective of the proposal is to improve the functioning of the internal market, fundamental rights play a crucial role in shaping the regulation, giving rise to a peculiar "medley" of product safety legislation and fundamental rights protection, as Almada and Petit aptly name it.[40]

On the one hand, the proposal is loyal to the traditional repertoire of product safety legislation in considering AI as a harmful product, which, therefore, needs to comply with specific requirements and certifications before entering or being put into service in the EU internal market.[41] On the other hand, however, the proposal recognises that the harm produced by AI systems is not comparable to a defective dishwasher. AI systems can discriminate, perpetuate or create social inequalities, harvest personal data and monitor individuals. When AI is used in decision-making processes, individuals are not consumers of AI products but are subject to them. AI is not a dishwasher but a technology that poses risks to fundamental rights and democratic values, which transcend market regulation and consumer law. The concept of risk is the core of the proposal, which classifies AI systems according to their risk level: unacceptable risk, high risk and low risk. Fundamental rights are crucial in drawing the line between the three categories.

First, fundamental rights are used as the normative justification to prohibit specific uses of AI. Title II establishes a list of prohibited uses that compromise AI systems whose use is considered unacceptable as contravening Union values, for instance by violating fundamental rights. An example of this is social scoring by public authorities, ie, the use of AI to surveil, profile and rank citizens (akin to the Chinese system). The prohibition is justified in light of the right to non-discrimination and the inviolable right to human dignity.[42]

Second, fundamental rights constitute the metrics distinguishing between low-risk and high-risk AI. High-risk AI systems are the core object of the regulation: if a system fulfils the classification rules in Article 6, then the AI system provider must comply with the requirements set in Chapter II, the conformity assessment procedure, pre- and post-market monitoring and reporting obligations. On the contrary, if a system poses only low or minimal risk, the provider is not obliged to comply with the regulation but can voluntarily follow the requirements as a code of conduct. Together with the definition of an AI system, the classification rules for high-risk systems represent the pivot of the AI Act as they determine which company is subject to the regulation.

In the proposal, the classification rules for high-risk AI systems aim to ensure legal certainty and foreseeability.[43] Instead of defining when an AI system "poses a significant harmful impact on health, safety and fundamental rights,"[44] the Commission provides a list of areas and systems which automatically classify as high risk. Annex III lists, among others, AI systems used in law enforcement, migration and asylum, education, administration of justice and employment. High-risk AI systems include biometric identification, polygraphs, risk assessment, or monitoring of workers. In the Commission's view, this approach ensures legal certainty for providers, as they will not have to interpret and assess their AI system's impact but only to check if their system is on the list. The underlying idea of this "automatic" approach to risk classification is that the Commission

[39] Recital 1 and Art 1 of the Commission Proposal.

[40] Marco Almada and Nicolas Petit, "The EU AI Act: A Medley of Product Safety and Fundamental Rights?" [2023] SSRN Electronic Journal <https://papers.ssrn.com/abstract=4308072> accessed 9 November 2023 (forthcoming in Common Market Law Review with the title "The EU AI Act: Between the rock of product safety and the hard place of fundamental rights").

[41] For a detailed analysis see ibid.

[42] Recital 17 of the Commission Proposal.

[43] Legal certainty is also widely referenced in the text and highlighted as a key aim of the proposal.

[44] Recital 27 of the Commission Proposal.

interprets the concept of fundamental rights (as well as health and safety) and assesses when an AI system adversely impacts them.

What happens if, in the future, a new AI system is developed in high-risk areas or used in other areas that are not listed and pose a threat to fundamental rights? In other words, how does the proposal ensure a future-proof classification of high-risk AI? In the proposal, the list in Annex III can be amended by the Commission via delegated acts.[45] When updating the list, the Commission must follow specific criteria provided by Article 7(2) to assess whether an AI system "poses a risk of harm to the health and safety or a risk of adverse impact on fundamental rights." Also, in this case, the Commission actively interprets and assesses the impact on fundamental rights. As the following two sections will show, the Council's and the Parliament's approaches to risk classification radically differ as they delegate the power to interpret fundamental rights to the provider.

Fundamental rights represent the benchmark to distinguish between AI systems and related regimes under the AI Act. When AI is inherently incompatible with fundamental rights, its use is prohibited; when AI poses a "risk of adverse impact" to fundamental rights,[46] the AI Act aims to provide protection to prevent or minimise such risk.[47] Fundamental rights protection is implemented in the Commission's proposal as a process to follow ex-ante before an AI system enters the market or is put into service. In this process, the provider[48] is the crucial addressee of the regulation.

Providers of high-risk AI must comply with the requirements set in Chapter 2, which include provisions on quality of training data and bias prevention (a well-known cause of algorithmic discrimination), ex-ante testing, risk management and human oversight. The user of the system, for example, an employer using an AI system to shortlist candidates, also plays a role. According to Article 29 of the Proposal, the user must use the AI system in line with the instructions to avoid function creeps, monitor and record the system logs when in use, and interpret and use the system appropriately.[49] The providers must inform users about the systems' functionalities, limitations and level of accuracy. In the Commission's view, obligations for the provider and the user will facilitate the respect of fundamental rights "by minimising the risk of erroneous or biased AI-assisted decisions in critical areas."[50] In case infringements of fundamental rights still occur, the requirements of traceability, documentation keeping and (limited) transparency would ensure, according to the Commission, effective redress for affected persons.[51] Despite the references to "affected persons" in different Recitals, individuals whose rights can be violated by AI systems have almost no role in the proposal. Only two provisions in the proposal address them: Article 52 on transparency requirements and Article 60 on the public database of high-risk AI systems. These provisions aim to increase transparency towards individuals interacting with AI systems (in the first case) and the public by setting a public registry of high-risk AI systems. Individuals also have no remedies if a violation of the AI Act occurs. Under the proposal, the enforcement of fundamental rights protection is

[45] Art 7 of the Commission Proposal.

[46] Ibid.

[47] For a critical take on the conflation between risk and trust see Johann Laux, Sandra Wachter and Brent Mittelstadt, "Trustworthy Artificial Intelligence and the European Union AI Act: On the Conflation of Trustworthiness and Acceptability of Risk" n/a Regulation & Governance <https://onlinelibrary.wiley.com/doi/abs/10.1111/rego.12512> accessed 7 February 2023.

[48] Defined in Art 3(2) of the Commission Proposal as the "natural or legal person, public authority, agency or other body that develops an AI system or that has an AI system developed with a view to placing it on the market or putting it into service under its own name or trademark, whether for payment or free of charge."

[49] Art 13(1) of the Commision Proposal.

[50] Explanations of the Commission Proposal, p 11.

[51] Ibid.

achieved through conformity assessment procedures for high-risk systems, post-market monitoring, and penalties for non-compliance.[52]

The lack of access to information, rights and remedies for individuals was harshly criticised by legal scholars, civil society organisations and the EDPB and EDBS. While the choice of excluding individuals from the AI Act can be debated from many perspectives, the Commission's strategy seems to be coherent with the role of the proposal: the AI Act is an internal market legislation which *complements* EU primary and secondary law (notably data protection) and national laws on fundamental rights. In the Commission's view, the AI Act minimises fundamental rights violations by AI and makes remedies for violations easier to achieve.

Moreover, a vital objective of the proposal is to foster the development and use of AI, attracting companies in the EU market with clear rules and proportionate regulatory burdens while also guaranteeing fundamental rights and values. The balancing effort between fundamental rights protection and companies' interests emerges when considering transparency obligations, which are limited to the "minimum necessary information for individuals to exercise their right to an effective remedy and the necessary transparency towards supervision and enforcement authorities."[53] In this sense, fundamental rights protection is balanced against the right to intellectual property protection.

A second class of exceptions to transparency and fundamental rights protection is for AI systems used in law enforcement and migration management.[54] The provision of information to individuals interacting with AI is limited when the system is used for law enforcement,[55] and exceptions to information provided in the database apply to law enforcement and migration management.[56] Real-time biometric identification, such as facial recognition, is generally prohibited but exceptionally allowed in three specific cases, including preventing a terrorist attack.[57] Moreover, AI systems that are part of EU databases, such as Eurodac and the Schengen Information System, are excluded from the scope of the AI Act if operational one year before the entry into force of the regulation.[58] EU databases are used mainly in migration management, border control, and asylum and raise critical issues for the protection of fundamental rights, which are widely studied in the literature.[59] Therefore, their exclusion from the scope of the AI Act risks significantly impairing legal protection for asylum seekers, migrants, and refugees.

In their Join Opinion,[60] the EDPS and EDPB criticised the proposal's exclusionary approach and recommended restricting the "broad exceptions" in law enforcement and including EU databases in the scope of the AI Act.

[52] Art 71 of the Commission Proposal.

[53] Explanations of the Commission Proposal, p 11.

[54] On transparency in the Commission proposal see Madalina Busuioc, Deirdre Curtin and Marco Almada, "Reclaiming Transparency: Contesting the Logics of Secrecy within the AI Act" (2023) 2 European Law Open 79.

[55] Art 52 of the Commission Proposal.

[56] Art 60 and Annex VIII of the Commission Proposal.

[57] Art 5(1)d of the Commission Proposal.

[58] Art 83 of the Commission Proposal.

[59] Simona Demkova, *Automated Decision-Making and Effective Remedies: The New Dynamics in the Protection of Eu Fundamental Rights in the Area of Freedom, Security and Justice* (Edward Elgar Publishing 2023); Niovi Vavoula, 'The "Puzzle" of EU Large-Scale Information Systems for Third-Country Nationals: Surveillance of Movement and Its Challenges for Privacy and Personal Data Protection' (2020) 45 E.L. Rev. 348; Niovi Vavoula, *Immigration and Privacy in the Law of the European Union: The Case of Information Systems* (Brill Nijhoff 2022).

[60] EDPS–EDPB Opinion (n 21).

2. The council: A trade-off approach

Compared to the amendments proposed by the Parliament, the changes in the Council's version were minimal yet remarkable. The Council kept the main structure of the AI Act intact – including Article 114 TFEU as a legal basis – while carving out several exceptions for security and law enforcement purposes.

The most prominent example is in Article 2(3), where the Council extended the exclusion from the scope of the regulation for AI systems in the military sector to any activity concerning defence or national security.[61] Additionally, exemptions to protection standards are disseminated throughout the Council Mandate. Real-time biometric identification – allowed only in three exceptional cases under the EC Proposal – is more widely permissible;[62] several use cases in law enforcement are deleted from the list of high-risk AI systems, as well as the use of document verification in migration management;[63] AI systems deployed in law enforcement, border control, migration and asylum management or for the operation of critical infrastructures will not be registered in the public database.[64] A recurring justification by the Council is the need to preserve the ability of law enforcement and migration authorities to carry out their activities, use information systems, and identify people who could be involved in crimes or unwilling to disclose their identities.[65] In other words, the Council views AI as a threat to fundamental rights but also as an attractive opportunity for law enforcement and security-related activities.

The list of prohibited AI practices in Article 5 has been the most debated of the AI Act, particularly regarding real-time biometric identification. Member States in the Council had very different views. While some supported a total ban,[66] others agreed on an exception in law enforcement supported by stronger guarantees.[67] According to some, requiring a prior judicial authorisation would provide a solid safeguard for individuals while also obtaining "a better position for the EP negotiations."[68]

The Council also conceives the AI Act as a regulatory burden that should not be borne by authorities pursuing the public interest of preserving national security, preventing and prosecuting crimes, and controlling migration and borders. As Austria commented on the proposal to delete document verification technology from Annex III, "the associated administrative burden [of classifying the system as high risk] would eliminate the added value gained from the system."[69] The AI Act is a regulatory burden that providers should not bear unless strictly necessary.

In order to ease the burden on providers, a key amendment was introduced in Article 6(3) of the AI Act. When classifying a system as high-risk, the Council proposes to refer to the list in Annex III "unless the output of the systems *is purely accessory* in respect of the relevant action or decision to be taken." The Council justifies this amendment as follows:

> "A number of Member States expressed some doubts as regards the classification of AI systems as high risk based on the broad terms of the proposal, leading to concerns that such an approach may also capture AI systems that are not likely to cause serious

[61] Recital 12a of the Council Mandate, as these are "sole responsibilities of Member States."
[62] Recital 18 and 5(1)d of the Council Mandate.
[63] Annex III of the Council Mandate.
[64] Art 51 of the Council Mandate.
[65] Recital 18 of the Council Mandate.
[66] See for example Germany's submission "Paper on Separate Regulation of AI Systems for Public Administration" (WK 12308/2022 INIT, 20.9.2022).
[67] See for example Spain's comment on Art 5(1)(d) in the comments on 1st part of 3rd compromise proposal (WK 13372/2022 INIT, 5.10.2022).
[68] Ibid.
[69] See Austria's comments on Arts 1–29 (WK 13191/2021 INIT, 3.11.2021), 14.

> fundamental rights violations or other significant risks. The Czech Presidency has analysed the feedback received in response to the options proposed in the policy paper, and it has proposed to modify the regime by introducing another horizontal layer on top of the high-risk classification made in Annex III. More specifically, Article 6(3) has been extended, and it now contains new provisions inspired by ideas from the High-level expert group on AI and from the OECD classification framework of AI systems, according to which the significance of the output of the AI system in relation to the decision or action taken by a human, as well as the immediacy of the effect, should also be taken into account when classifying AI systems as high risk".[70]

The Council's proposed approach to classification depends on a self-assessment by the provider, which, therefore, determines whether it is subject to the regulation or not.[71] Although intending to ease the burden of regulation on providers, the Council introduced a dangerous loophole, capable of "jeopardising the safeguards applicable to AI systems."[72]

3. The Parliament: A human-centric AI

When OpenAI released Chat-GPT in March 2023, the Parliament was discussing amendments to the AI Act. Undoubtedly, the vivid discussions about future threats and present concerns about the social harm of generative AI influenced the Parliament, which shifted the focus of the AI Act to reinforce individual rights.

First and foremost, the Parliament changed the legal basis (from Article 114 TFEU to Article 16 TFEU) and, therefore, the core aims of the AI Act. In the Parliament's view, AI is a threat not only to fundamental rights – including the right to a high level of environmental protection - but also to broader values of democracy and the rule of law.[73] Hence, regulation has the primary aim of promoting the uptake of human-centric AI and ensuring a "high level of protection of health, safety, fundamental rights, democracy and the rule of law, and the environment from harmful effects of artificial intelligence systems in the Union *while* supporting innovation" (emphasis added).[74] The Parliament's version of the AI Act builds on the Commission's core ideas of a risk-based approach but radically changes its market-based nature by anchoring the regulation to protect fundamental rights and democratic values. The result is a hybrid instrument between internal market legislation – with design requirements, conformity assessment procedures and post-market monitoring – and fundamental rights legislation – riddled with general principles,[75] rights and remedies for individuals.

The Parliament uses fundamental rights as normative arguments to prohibit specific uses of AI, going much further than the Commission's proposal. Four new prohibitions are

[70] Second Presidency Compromise text (11124/22, 15.7.2022), 4 <https://artificialintelligenceact.eu/wp-content/uploads/2022/07/AIA-CZ-1st-Proposal-15-July.pdf> accessed 17 April 2024.

[71] This policy option was also justified as "likely easier for providers" in the Policy Options prepared by the Czech Presidency in view of the discussion in WP Telecom on 5 July 2022 (WK 8862/2022 INIT, 17.6.2022) <https://artificialintelligenceact.eu/wp-content/uploads/2022/07/AIA-CZ-Options-Paper-17-June-2022.pdf> accessed 17 April 2024.

[72] EDPS Final Opinion (n 21).

[73] Recital 1, 13, 27 and new Recital 28a of the Parliament Mandate.

[74] Art 1 of the Parliament Mandate.

[75] See the proposed Art 4a of the Parliament Mandate establishing general principles for any AI system.

added to Article 5 of the AI Act[76]– including profiling by law enforcement authorities and indiscriminate scraping of data for facial recognition databases and emotion detection systems – based on the unacceptable risk they pose to fundamental rights. The Parliament widely refers to the Charter of Fundamental Rights (CFR), in particular, the right to non-discrimination, privacy, and human dignity,[77] but also to technical limitations of systems, such as emotion detection, and their limited reliability.[78] Most importantly, the Parliament bans any type of real-time biometric identification by public or private parties without exceptions.[79] The Parliament justifies the total ban in light of the core rule of law principles, as real-time biometric identification evokes a "feeling of constant surveillance" and gives parties deploying such systems "a position of uncontrollable power."[80]

In the original proposal, prohibiting AI systems meant that such systems could not be allowed in the EU market, regardless of where the provider is established or used in the EU.[81] However, in the Parliament's view, this was insufficient "in order for the Union to be true to its fundamental values."[82] If specific AI systems are deemed unacceptable under the regulation, then providers based in the EU should also not be allowed to export such systems to third countries.[83] In this way, the Parliament significantly extend the scope of the regulation and its extra-territorial reach.

As in the Commission's proposal, fundamental rights play a crucial role in the classification rules for high-risk AI systems, although with noteworthy differences. For the Parliament, AI systems should be classified as high risk not only in the cases listed in Annex III (what I called an "automatic" approach in the Commission's proposal) but only *if* they pose "a significant risk of harm to the health, safety or fundamental rights of natural persons."[84] In other words, instead of having a risk assessment carried out upstream, the Parliament delegates this role to two subjects: the provider and the user.

Firstly, the provider has to assess whether their AI system is a high risk by 1) checking if it falls in the critical areas under Annex III *and* (if so) 2) performing a risk assessment[85] following the guidelines provided by the Commission.[86] Similar to the Council's approach, the Parliament gives the provider the task of assessing whether their systems fall within the regulation but also provides verification mechanisms to avoid misclassifications. Under Article 6(2)a of the Parliament Mandate, if the provider considers their system risk-free, they shall submit a reasoned notification to the competent authorities,[87] who shall review and reply. If the provider has misclassified the AI system, they can be fined.

Secondly, the deployer (or "user" in the terminology of the Commission's proposal) has to perform a fundamental rights impact assessment (FRIA)[88] before using the system.[89] According to the Parliament, deployers are best placed to understand how the high-risk AI

[76] Art 5(1)ba; Art 5(1)da, Art 5(1)db, Art 5(1)dd of the Parliament Mandate.
[77] Recital 16a, Recital 26a, b, c of the Parliament Mandate.
[78] Recital 26c of the Parliament Mandate.
[79] In contrast with the Commission's and Council's proposals to allow it in three exceptional cases.
[80] Recital 18 of the Parliament Mandate.
[81] Art 2 of the Commission Proposal.
[82] Recital 10 of the Parliament Mandate.
[83] Recital 10 and Art 2ca of the Parliament Mandate.
[84] Art 6(2) of the Parliament Mandate.
[85] Whereby risk is defined as "the combination of the probability of an occurrence of harm and the severity of that harm" in Art 3(1)1a of the Parliament Mandate.
[86] Art 6(2) of the Parliament Mandate.
[87] National supervisory authority or the AI office.
[88] For more see Alessandro Mantelero, "The Fundamental Rights Impact Assessment (FRIA) in the AI Act: Roots, Legal Obligations and Key Elements for a Model Template" (2024) 54 Computer Law & Security Review 106020.
[89] New Art 29a of the Parliament Mandate.

system will be used concretely and can identify potential significant risks that were not foreseen in the development phase. The choice of obliging deployers to perform an FRIA was also highly influenced by the debates on the so-called "general purpose" AI systems, which emerged vividly after the release of Chat-GPT. One of the core problems of the original Commission's proposal was that it did not consider AI systems, which can be used for very different purposes that are determined by the user (and not by the provider). Consider Chat-GPT, a widely known AI chatbot. The system can be applied to many different contexts and for different purposes. A student can ask Chat-GPT: "Can you suggest a structure for my policy brief assignment?" Nevertheless, judges can also use it in deciding cases and writing their judgments.[90] The difference is that in the second case, the AI system does pose potential risks to fundamental rights.[91] Considering this issue, the Parliament introduced several provisions for general purpose AI systems and the FRIA for the deployers.

As to fundamental rights protection, the Parliament develops guiding principles that shall be followed by any provider of AI systems,[92] including fairness, transparency, non-discrimination, human agency, technical robustness and safety.[93] In the case of high-risk AI systems, the general principles are "translated" into the requirements set out in Chapter 2 of the regulation. In this sense, the Parliament attempts to bypass the rigid regulatory scheme of the Commission, developing an overarching framework for AI "in line with the Charter as well as the values on which the Union is founded."[94] Unlike the Council's approach, fundamental rights protection cannot be subject to broad exceptions, especially in sensitive areas of law enforcement and migration management. The Parliament deletes the exception to registering high-risk AI systems and their deployers for law enforcement and migration management.[95] It even extends the information to be published in the public database when the provider or deployer is a public authority.[96] Regarding the exclusion of EU databases from the scope of the AI Act, the Parliament narrows down the exception in Article 83 of the Proposal by excluding only systems implemented before the entry into force of the regulation (and not until one year after, as in the Commission's and Council's proposals). In these cases, however, the operators of such systems "must take all necessary steps to comply with the AI Act."[97]

After levelling the playfield, the Parliament started to build on the Commission Proposal with new rights and remedies for individuals. The role of "affected"[98] persons in the AI Act is undoubtedly the most prominent change in the Parliament's approach. In line with the GDPR, individuals have new rights, such as the right to be informed of an AI-supported decision and the right to request a "clear and meaningful explanation on the role of the AI system in the decision-making procedure, the main parameters of the decision taken and the related input data."[99] The new informational rights for affected persons are instrumental in exercising remedies for violations of the regulation. In particular, the Parliament built on the Council's idea to allow any natural or legal persons

[90] See for instance Luke Taylor, "Colombian Judge Says He Used ChatGPT in Ruling" *The Guardian* (3 February 2023) <https://www.theguardian.com/technology/2023/feb/03/colombia-judge-chatgpt-ruling> accessed 17 April 2024.

[91] See more in Lilian Edwards, "Regulating AI in Europe: Four Problems and Four Solutions" (Ada Lovelace Institute 2022) <https://www.adalovelaceinstitute.org/report/regulating-ai-in-europe/> accessed 22 April 2022.

[92] Defined in Art 3(1) of the Parliament Mandate.

[93] New Art 4a of the Parliament Mandate.

[94] Ibid.

[95] Art 51 of the Parliament Mandate.

[96] Annex VIII of the Parliament Mandate.

[97] Art 83 of the Parliament Mandate.

[98] Defined in Art 2(1)8a of the Parliament Mandate.

[99] New Art 68c of the Parliament Mandate.

to submit a complaint to market surveillance authorities[100] by introducing two rights. First, a right to lodge a complaint with national supervisory authorities for infringements of the AI Act. Second, a right to an effective judicial remedy against a decision of the supervisory authority.[101]

Overall, the objectives, language and solemn enunciation of principles and rights in the Parliament's version give the AI Act a new look: from an internal market tool to a comprehensive framework for human-centric AI.

4. The trilogue: Reaching political compromise

The trilogue represented the critical moment in the legislative process to reach a political agreement on the most debated issues. It was hard to imagine how compromise could be achieved since the three institutional positions diverged immensely. Before entering the (closed) doors of the trilogue room, the Parliament and the Council held opposite views on sixteen out of twenty total recommendations from the EDPB–EDPS. While the Parliament had implemented most of them, the Council only did so in four cases.

Parliament		Council	
not implemented	2	not implemented	17
implemented	15	implemented	4
partly implemented	2	partly implemented	1
Opposite positions		Convergent positions	
R1, R2, R3. R4. R5, R7, R8, R10, R11, R12, R15, R17, R18, R19, R20, R22		R6, R9, R13, R14, R16, R21	
Total opposite: 16		Total convergent: 6	

Before the trilogue, fundamental rights protection was at a crossroads. The stakes were high, causing mobilisation from civil society organisations and scholars, who submitted letters pleading to the Parliament to safeguard individual rights and ban facial recognition.[102] The trilogue negotiations also prompted the EDPS to submit his final opinion on his initiative in October 2023.[103] In his opinion, the EDPS reiterated most of his previous recommendations while adding new specific concerns raised from the Council general mandate and the Parliament-amended version of the AI Act. With the hope of influencing the EU institutions during the trilogues, the opinion aimed to ensure that "persons impacted by the use of AI systems enjoy both an appropriate level of protection and legal certainty."[104]

After a three-day "marathon" at 1:00 AM on Friday, 8 December 2024, the press conference started with an announcement: *Habemus* the AI Act.[105] The AI Act was born, and

[100] Art 63(7d) of the Council Mandate.

[101] New Arts 68a and 68b of the Parliament Mandate.

[102] See, among others, the coalition "Protect Not Surveil" which includes more than 300 supporters and is led by AccessNow, EDRi, PICUM and the Refugee Law Lab.

[103] EDPS Final Opinion (n 21).

[104] EDPS Final Opinion (n 21), 3.

[105] The press conference was the only moment open to public, live streamed on the website of the Parliament, and where journalists could intervene after long and secluded discussions.

all three institutions' representatives acknowledged it was a historic moment. It took two more months for the public to obtain more information about the results of the trilogue. As the graphic below shows, the trilogue resulted in a compromise between the different positions, with ten recommendations implemented, seven partly implemented, and five not implemented. A visual representation of the "AI Act Roller Coaster" is provided in the Figure 1 below.

Where the positions were broadly divergent, political compromise was reached in most cases, ending in the partial implementation of the EDPB–EDPS recommendation. The Parliament obtained agreement on their position in six cases, whereas the Council only in three.

Compromised version	Parliament version prevailed	Council version prevailed
R1, R2, R4, R8, R10, R11, R22	R3, R12, R15, R17, R18, R20	R5, R7, R19
Total 7	Total 6	Total 3

Remarkably, the co-legislators agreed to fully implement some important recommendations for fundamental rights protection. Among others, the full prohibition of any type of social scoring,[106] the inclusion of a clear link to data protection law in the certification system,[107] the obligation for the deployer to perform a Fundamental Rights Risk Assessment (FRIA),[108] and the addition of the Fundamental Rights Agency as a permanent member of the Advisory Forum.[109] More importantly, the final agreement included new rights and remedies for individuals as suggested by the EDPS–EDPB and the Parliament. These include a novel right to an explanation for AI-driven decision-making that will complement the protection provided by Article 22 of the GDPR for solely automated decisions.[110] Additionally, individuals affected by AI systems will also have the possibility to lodge a complaint for violations of the AI act to marker surveillance authorities.[111]

However, by examining the individual issues, the process tracing reveals that the Council was most successful in specific areas: the use of AI for law enforcement, migration control, and national security. These findings may be unsurprising for many scholars working on EU migration policies. Research shows that the EU intervention in migration policy has faced strong resistance from national governments, especially when negotiating in the Council.[112] At the EU level, while the Parliament is generally more progressive, the Council remains protectionist.[113]

[106] Art 5(c) of the AI Act.

[107] Art 48(5) and Annex V point 5 of the AI Act.

[108] Art 27 of the AI Act.

[109] Art 65(5) of the AI Act.

[110] On the new right to explanation see Simona Demkova, "The AI Act's Right to Explanation: A Plea for an Integrated Remedy" (*MediaLaws*, 31 October 2024) <https://www.medialaws.eu/the-ai-acts-right-to-explanation-a-plea-for-an-integrated-remedy/> accessed 1 November 2024; Georgios Pavlidis, "Unlocking the Black Box: Analysing the EU Artificial Intelligence Act's Framework for Explainability in AI" (2024) 16 Law, Innovation and Technology 293; On the limitations of Article 22 GDPR see Francesca Palmiotto, "When Is a Decision Automated? A Taxonomy for a Fundamental Rights Analysis" (2024) 25 German Law Journal 210.

[111] Art 85 of the AI Act.

[112] Terri Givens and Adam Luedtke, "The Politics of European Union Immigration Policy: Institutions, Salience, and Harmonization" (2004) 32 Policy Studies Journal 145; Gallya Lahav and Anthony M Messina, "The Limits of a European Immigration Policy: Elite Opinion and Agendas within the European Parliament" (2005) 43 JCMS: Journal of Common Market Studies 851.

[113] Givens and Luedtke (n 112); Adam Luedtke, "Uncovering European Union Immigration Legislation: Policy Dynamics and Outcomes" (2011) 49 International Migration 1.

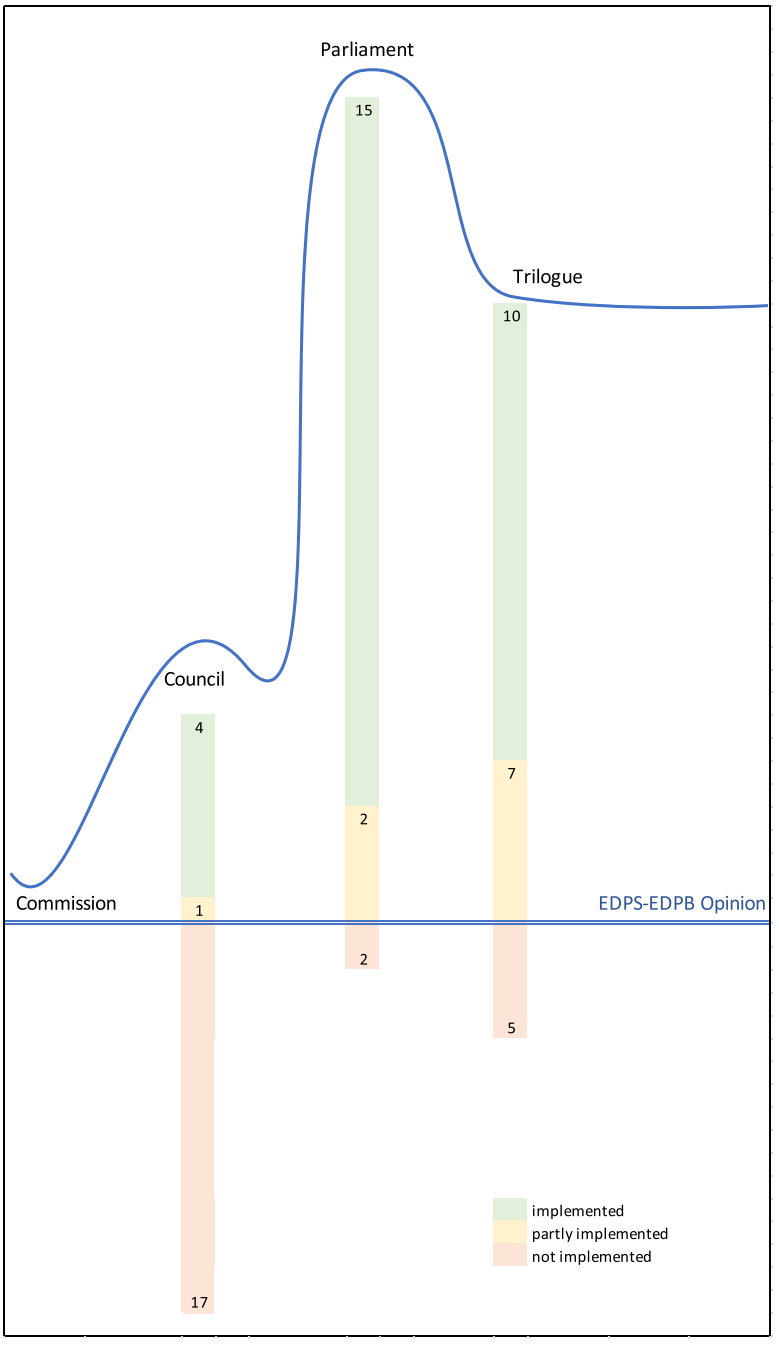

Figure 1. The AI Act Roller Coaster.
Source: author's elaboration.

Recommendation	Outcome
R3: Include international law enforcement cooperation in scope in Art 2(4)	Compromise with Council victory for excluding national security from the scope
R7: Prohibit any automated biometric recognition in publicly available spaces	Council position prevailed
R8: Prohibit profiling in law enforcement	Compromise achieved with a broad exception adopted
R10: Prohibit emotion detection	Compromise with Council victory for exclusion in law enforcement
R11: Prohibit AI systems categorizing individuals from biometrics	Compromise with Council victory for exclusion in law enforcement
R19: Restrict the broad exception to transparency obligations for LE in the public database	Council position prevailed

This analysis provides an alarming picture of the AI Act resulting from the trilogue: while most of the recommendations were implemented, therefore increasing fundamental rights protection, different standards will apply in the fields of law enforcement and migration. The next session will reflect on the implications of this choice in more detail.

Finally, a critical recommendation by the EDPS regarding the scope of the application was not implemented in the final version. The newborn Article 6 deviates from the Commission's automatic approach to risk classification, resulting in a hybrid version between the Parliament's proposal and the Council's approach, requiring providers to self-assess whether the AI systems pose a risk to fundamental rights, health or safety or not.[114]

Conclusively, although co-legislators achieved significant progress, particularly in establishing rights and remedies for individuals, dangerous compromises were made. The next Section will reflect on the implications of such political choices and their significance for fundamental rights protection. More specifically, it will focus attention on two critical aspects resulting from political compromise: (1) the uncertain scope of protection in Articles 2 and 6 and (2) the double standards of protection for AI systems in law enforcement, migration and asylum.

IV. Implications for fundamental rights protection

1. Uncertain scope of protection

In addition to the prohibited use of AI, the AI Act's protective function rests on Article 6, the core pillar that defines classification rules for AI systems.

Classifying AI systems as high-risk triggers the requirements for AI systems, as well as the obligations and duties for providers and deployers. Conversely, if the system is not classified as high-risk, the provider is not subject to legal obligations but can voluntarily follow codes of conduct. Deployers are not obliged to perform the FRIA, and individuals affected by AI decision-making cannot exercise the right to an explanation in Article 86. Therefore, Article 6 – in combination with Articles 2 and 3(1) of the AI Act – has a cornerstone role for the AI Act's applicability. In the final version of the AI Act, an AI system is classified as high-risk if two cumulative requirements are fulfilled: (1) the system is listed in Annex III, and (2) it poses a significant risk of harm to the health, safety or

[114] Art 6(3) of the AI Act.

fundamental rights of natural persons.[115] This would not be the case, according to Article 6(3), where the system does not "materially influence the outcome of decision-making," for instance, when performing a narrow procedural or preparatory task. However, a system that performs profiling of natural persons shall always be considered high risk.[116] After the assessment, providers must keep the documentation and register the system in the database if they consider that a derogation applies. Market surveillance authorities can, upon request, access the documentation and eventually order compliance with the regulation in case of misclassification.[117]

Undoubtedly, this provision raises several interpretative questions that legal scholars and, eventually, the Court of Justice of the EU will address in the future. In particular, it will be crucial to set tangible standards for vague concepts such as "materially influencing the decision" or "narrow procedural tasks." For the purpose of this paper, however, I want to focus on the rationale behind this provision, the underpinning political justification and the (perhaps unintended) consequences for fundamental rights.

As Section 3 showed, the Commission's proposal's original version of Article 6 aimed to ensure legal certainty for providers by offering a non-rebuttable list of high-risk AI systems. Both the Parliament and the Council proposed an additional layer of classification rules, delegating the risk assessment to providers. The reason behind these proposals was to add a horizontal level of granularity and flexibility to ensure that the Regulation applies only to cases with significant risk.[118] How such a horizontal level had to be designed was, however, subject to contrasting views. Several Member States were concerned about the insufficient level of legal certainty of the exemptions, especially when linked to the role of AI systems in decision-making. In these cases, it will be impractical for the provider to determine *a priori* how the systems will be concretely used by human decision-makers.[119] In the end, a compromise was reached by adding more detailed conditions for the derogation and the introduction of a monitoring duty in cases of misclassification. Rather than reflecting a clear intention from the co-legislators, the final version seems to be the result of bargaining between different stakeholders' interests and policy goals.

The aim of having a stricter proportionality approach to risk, however, resulted in a deeply uncertain formulation of Article 6 and a worrisome delegation of powers to providers. In fact, Article 6 empowers providers to decide whether their system falls within the scope of the AI Act or not. This is a dangerous delegation of fundamental rights protection duties to providers. This delegation requires the provider to determine what fundamental rights are and whether their system adversely impacts them. Based on their assessment, the safeguards for fundamental rights will or will not apply.[120] Interpreting fundamental rights and assessing when they are at risk is not only a task that requires expertise but also legitimacy. The delegation of these tasks, traditionally a remit of the judicial and legislative branches, raises important questions of oversight and governance of fundamental rights.

The risk is to ground the Regulation in the hope of effective ex-post enforcement through market surveillance. Worryingly enough, when misclassified by the providers, the system would still enter the market, potentially harming individuals' fundamental rights, until the market surveillance authority takes action.[121] In other words, the AI Act accepts the risk of fundamental rights being violated until competent authorities decide to act.

[115] For a detailed analysis see Leinarte (n 12).
[116] Art 6(3)(d) of the AI Act.
[117] Art 80 of the AI Act.
[118] See Policy Options Paper by the Czech Presidency (n 71).
[119] See among others the submissions by Belgium, the Netherlands and Austria.
[120] However, national competent authorities shall monitor cases of misclassification according to Art 80.
[121] Arts 80 and 79(5) and (9) of the AI Act.

A second concern regarding the scope of protection arises from Article 2(3). While the original proposal excluded only AI systems in the military, the final version broadens the derogation to defence and national security purposes. The exclusion of national security was uncontested in the Council.[122] Formally, the exclusion was justified based on Article 4(2) TFEU, as national security is the sole responsibility of Member States. Therefore, the EU is not competent to regulate AI in this field.[123] The exclusion of national security also clearly emerges as a political priority of Member States, which should be left free to organise the use of AI in public administration and public services.[124] For the effective protection of fundamental rights, it is critical that the notion of "national security" is strictly interpreted to delimit the discretion of law enforcement authorities and instrumentalisation.[125]

2. *Double standards of protection*

The AI Act embeds double standards for individuals affected by AI systems, with lower protection for individuals suspected or accused of having committed a crime, migrants, asylum seekers and refugees. From a legal and ethical perspective, this is a critical weakness of the Regulation.[126] But from a political perspective, the Council considered the trilogue a victory by introducing several exceptions for law enforcement and migration authorities, in line with their mandate or "general approach."[127]

When justifying exceptions for AI systems, Member States argued for the need to avoid "unnecessary administrative burdens" on public authorities,[128] which would affect the effectiveness of their activities. For instance, commenting on Article 83 of the Proposal, Austria proposed keeping the EU IT systems entirely out of the regulatory scope. Their argument was rooted in concerns about implementation difficulties and the potential hindrance to the European Entry–Exit System (EES). This stance highlights a broader concern among some Member States about the impact of regulation on operational flexibility and efficiency of migration authorities. In the final version, large IT systems are not entirely outside of the scope of the AI but benefit from an extended compliance deadline in 2030.[129] As Vavoula argues, this provision would allow a three-year grace

[122] See the Policy Option by the Czech Presidency (n 71), where they state that: "the vast majority in the Council strongly supports an explicit exclusion of national security from the scope" (p 6).

[123] See, for instance, Denmark's comment supporting the exclusion as "outside of EU Law" (DK Comments on the AI Act, WK 11769/2022 INIT, 9.9.2022) or Austria's support for the exclusion as justified by the specificities of MS and common Union defence policy subject to public international law, which is, therefore, the most appropriate framework (AT Comments on the AI Act, WK 11910/2022 INIT, 13.9.2022).

[124] Or "protect the situation of the Austrian Armed Forces," in the words of Austria (AT Comments on the AI Act, WK 11910/2022 INIT, 13.9.2022).

[125] On the risks of national security as an "exemption card" see Valsamis Mitsilegas and others, "Data Retention and the Future of Large-Scale Surveillance: The Evolution and Contestation of Judicial Benchmarks" (2023) 29 European Law Journal 176.

[126] See also Sandra Wachter, "Limitations and Loopholes in the EU AI Act and AI Liability Directives: What This Means for the European Union, the United States, and Beyond" (2024) 26 Yale Journal of Law and Technology 672; Niovi Vavoula, "Unpacking the EU Proposal for an AI Act: Implications for AI Systems Used in the Context of Migration, Asylum and Border Control Management" (*Turkish Policy Quarterly*) <http://turkishpolicy.com/article/1100/unpacking-the-eu-proposal-for-an-ai-act-implications-for-ai-systems-used-in-the-context-of-migration-asylum-and-border-control-management> accessed 5 May 2022; Dimitri van den Meerssche and Rebecca Mignot-Mahdavi, "Failing Where It Matters Most? – The Digital Constitutionalist" (22 December 2022) <https://digi-con.org/failing-where-it-matters-most/> accessed 2 January 2023.

[127] Council's comment on the Interinstitutional Agreement 5662/24 (26 January 2024), 4.

[128] See, for instance, Finland's comment on Art 51 of the Commission Proposal (FI Comments on the AI Act, WK 13236/2022 INIT, 4.10.2022).

[129] Art 111(1) AI Act.

period where IT systems can operate without complying with the requirements and safeguards of the AI Act.[130]

Framing AI as a tool for the protection of national security also allowed the Council to propose similar exceptions in law enforcement and migration control, with Germany advocating for a separate regulation specifically tailored to AI use in public administration. In their submission, "Separate Regulation of AI Systems for Public Administration," they argue that such a regulation should address the unique needs and challenges faced by security, migration, and asylum authorities, as well as tax and customs administration.[131] Germany emphasised the importance of enabling government functions through AI while ensuring the protection of fundamental rights. For this purpose, it was crucial to strike a better balance between transparency and protecting confidential information and narrow the list of high-risk use cases in law enforcement and migration. In the end, controversial systems such as deepfake detectors, crime analytics and document verification systems were deleted from the original list of high-risk AI systems in Annex III.

A further political narrative, which gained strong support in the Council, was to portray transparency requirements as harmful to law enforcement and migration. Regarding the registration duty in the public database, several Member States argued for exceptions due to security concerns. In their view, the database could pose a security risk and affect the capabilities of the authorities.[132] Moreover, it would expose the investigative methods of law enforcement to criminals and "hostile states."[133] As a result, transparency obligations – what I consider the most innovative and protective tools for fostering procedural rights and remedies – were watered down. The registration of high-risk AI systems in the public database, a revolutionary and much-needed provision, contains exceptions for AI systems used in the area of law enforcement, migration and asylum. According to Article 71, such systems will be contained in a non-public section of the database, thus perpetuating the existing legal barriers that suspects, migrants and asylum seekers face when exercising their rights and remedies against AI-driven decisions.[134]

Additionally, specific disclosure duties enshrined in Article 50 will not apply to AI systems authorised by law to detect, prevent, investigate or prosecute criminal offences. If subject to emotion recognition systems or biometric categorisations, two extremely invasive and harmful AI systems, suspects or defendants will not be aware of it. Additionally, the watermarking obligation for deepfake does not apply if the use is authorised for law enforcement purposes. This exception is particularly worrisome, as it fails to minimise the risks of wrongful convictions based on deepfake evidence. Fortunately, the new "right to an explanation" proposed by the Parliament remained unaffected by the exclusionary approach for law enforcement and migration management.

[130] Niovi Vavoula, "Tr-AI-nsforming Migration, Asylum and Border Management in the EU: The Roles of the AI Act, Interoperable Large-scale IT Systems and EU Migration Agencies" in N. Vavoula & A. Karaiskou (eds) *Towards Autonomous Borders? Artificial Intelligence, Human Rights and Rule of Law Challenges in Contemporary Border and Asylum Governance,* Computer Law & Security Review (forthcoming 2025).

[131] "Separate Regulation of AI Systems for Public Administration" (DE Comments on the AI Act, WK 12308/2022 INIT, 20.9.2022).

[132] See also Germany's comment on Art 51 of the Commission Proposal (DE Comments on the AI Act, WK 13423/2022 INIT, 6.10.2022).

[133] See Finland's comment on Art 60 of the Commission Proposal (FI Comments on the AI Act, WK 13236/2022 INIT, 4.10.2022).

[134] See, among others, Francesca Palmiotto and Derya Ozkul, "Climbing a Wall: Strategic Litigation Against Automated Systems in Migration and Asylum," German Law Journal (forthcoming 2025).

Overall, framing the exceptions was not about balancing security versus fundamental rights protection[135] but about the practical implications of regulation on the efficiency and effectiveness of public authorities. For Member States, AI gives an unprecedented advantage to public authorities, which the regulations risk annulling. However, the consequence of framing the AI Act as a "burden" is a regulatory framework with double standards for fundamental rights protection. By carving out exceptions and prioritising flexibility, the Council created gaps in the regulatory framework that left fundamental rights inconsistently protected in the AI Act. The table below summarises the core exceptions in the final text of the Regulation.

Provision		General Rule	Exception
Art 2(3)	Scope of application	(original proposal) Exclusion for military purposes	(final version) Exclusion of military, defence and national security
Art 5(f)	Prohibition of emotion detection	Prohibited in education institutions and workplace	Allowed in law enforcement, migration and asylum
Art 5(g)	Prohibition of biometric categorisation	Prohibition when the categorization aims to deduce or infer protected characteristics	Exception for the labelling or filtering of lawfully acquired biometric datasets or categorizing of biometric data in law enforcement
Art 50	Transparency for end-users	Users must be informed of the fact that: 1. They are interacting with an AI system 2. The content is AI-generated or manipulated 3. They are exposed to emotion recognition systems or biometric categorization systems	Exception for AI systems authorized by law to detect, prevent, investigate or prosecute criminal offences
Arts 71 and 49(4)	Public Database for high-Risk AI systems	Registered in a publicly available and accessible section	AI systems in law enforcement, migration and asylum are registered in a non-public section
Arts 111 and 113	Compliance deadline for high-Risk AI systems	Deadline for High-Risk AI systems is 2 August 2027	Deadline for AI systems that are component of large-scale IT systems is postponed to 2030

V. Conclusions

The legislative process of the Artificial Intelligence Act has been a roller-coaster for fundamental rights protection. Proposed in 2021 by the European Commission, the AI Act was an internal market legislation imposing requirements for certain types of AI systems which pose high risks to fundamental rights. The Council then proposed significant exceptions to crucial requirements of transparency and the scope of application of the AI Act in the areas of law enforcement and migration management. Civil society organisations and scholars advocated for improving fundamental rights protection in the AI Act with several initiatives targeting the European Parliament. After lengthy

[135] Gwendolyn Sasse, "Securitization or Securing Rights? Exploring the Conceptual Foundations of Policies towards Minorities and Migrants in Europe*" (2005) 43 JCMS: Journal of Common Market Studies 673.

discussions in Parliament, the AI Act had a new look: new rights for individuals and different objectives to foster a human-centric approach to AI.

Before the trilogue negotiations, significant disparities in fundamental rights protection existed among the institutions involved in the AI regulation process. Notably, the Parliament incorporated fifteen recommendations of the EPBS-EDPB opinion into its position, while the Council adopted only four. Consequently, the trilogue phase emerged as a pivotal moment in the regulatory process. Analysis of the final text resulting from political negotiations revealed that while compromises were reached in many cases to reconcile divergent views, the Parliament played a crucial role in increasing standards of protection for fundamental rights. However, the Council wielded the most decisive influence in law enforcement, migration, and national security matters.

A deep dive into the documents submitted by Member States within the Council shed light on the underlying reasons for this dynamic. Member States strongly resisted ceding competencies to the EU, particularly in areas deemed national interests, such as migration, asylum, and criminal law enforcement.[136] Throughout the discussions, a prevailing political narrative framed AI as a significant asset for bolstering security, thus positioning regulation as an unnecessary administrative burden. Notwithstanding the vital role played by the EDPS and EDPB, their recommendations gained less traction in areas encroaching upon Member States' prerogatives.[137] Despite important achievements in the trilogue, critical choices made present challenges to the effective protection of fundamental rights in the AI Act.

What does the future hold for the protection of fundamental rights in the AI Act after its entry into force? Law-making is only the first step of law. The implementation, enforcement and judicial interpretation of law hold crucial importance for the protection of fundamental rights in the age of AI.

Looking forward, a critical role can be played by the Commission, which has the power to expand the list of high-risk AI systems in Annex III through delegated acts[138] and adopt guidelines for the application of Article 6 with a comprehensive list of practical examples that are high-risk or not high-risk. This is an important power that can reduce the uncertainty of Article 6 and keep the Regulation updated in the long term. Researchers, the Fundamental Rights Agency and civil society organisations can support this effort by monitoring and providing evidence of emerging or overlooked risks that AI systems pose to fundamental rights.

Undoubtedly, market surveillance authorities are essential for effectively enforcing the protective scope of the AI Act. For this purpose, Member States must equip authorities with sufficient economic and human resources to act quickly and systematically whenever a provider claims to be exonerated from the AI Act (through the derogation mechanism in Article 6). In a similar vein, monitoring will be critical to avoid the instrumentalisation of Article 2(3) as a "national security exemption card" for AI systems in law enforcement and migration management.

Additionally, it is important to recall the role of the EU Court of Justice in interpreting the AI Act in light of EU primary law, most notably the Charter of Fundamental Rights, and litigation more broadly. Individual and strategic litigation against AI tools can support regulators and enforcement agencies in assessing risks to fundamental rights and keeping

[136] See similarly in the field of EU data protection Quintel (n 17) [where she argues that data protection rules in law enforcement are interpreted broadly, including the field of border control and migration management, lowers standards of protection].

[137] See also Ronit Justo-Hanani, "The Politics of Artificial Intelligence Regulation and Governance Reform in the European Union" (2022) 55 Policy Sciences 137 [explaining policy preferences at the domestic level as a key driver for EU policymaking and AI governance].

[138] As also argued by Wachter (n 126) 716.

them alert on emerging ones. For this purpose, the AI Act, combined with the GDPR, offers litigants a new set of legal tools to challenge AI-driven decisions.

Finally, the commitment of the EU and its Member States to fundamental rights protection in the age of AI does not end with the adoption of the AI Act. Next to duty of States under international human rights law, the recent "Council of Europe Framework Convention on Artificial Intelligence," requires signatory States to adopt or maintain legislative measures to ensure that activities within the lifecycle of artificial intelligence systems are fully consistent with human rights, democracy and the rule of law. To fulfil their obligations, EU Member States should introduce specific rules and safeguards for AI systems used by public authorities, particularly in asylum procedures and criminal proceedings.[139] In the words of the German representatives, "the AI Act must not be a regulatory ceiling for specific requirements imposed by Member States,"[140] especially in areas that remain emblematic of national sovereignty. Indeed, as Mir highlights, the AI Act does not prevent Member States from adding safeguards.[141] In the field of criminal justice, the AI Act should complement *ad hoc* regulation, setting procedural rules tailored to the specific national criminal legal system.[142] These would include, as I have argued elsewhere, rules on the admissibility of AI as evidence at trial, a right to examine AI systems at trial, procedural safeguards for the use of AI as an investigative tool, and strict requirements for judicial uses of AI.[143]

Urgent attention should be given to transparency, a corollary of privacy and procedural rights, to ensure that *all* affected data subjects – *particularly* migrants, asylum seekers, suspects and defendants – are aware of the use of AI systems and are put in a position to challenge their use. Last but not least, Member States still retain the power *not to* authorise the use of specific AI systems, such as live facial recognition, if deemed incompatible with their values and constitutional rights.

Acknowledgments. This research is part of the Algorithmic Fairness for Asylum Seekers and Refugees (AFAR) Project, funded by the Volkswagen Foundation under its Challenges for Europe Programme. I would like to thank the AFAR team and all the members of the Centre for Fundamental Rights and the Centre for Digital Governance at the Hertie School and the Working Group "The Digital Public Sphere" at the EUI for their support and feedback throughout the research. A special thanks goes to Mark Dawson and Marco Almada for the fruitful discussions and comments.

[139] On the need for ad hoc regulation see Sachoulidou A, "Harnessing AI for Law Enforcement: Solutions and Boundaries from the Forthcoming AI Act" (2024) 15 New Journal of European Criminal Law 117 and Palmiotto F, "Artificial Intelligence and the Transformation of Criminal Trials : Preserving Fairness in Europe" (Thesis, European University Institute 2023) <https://cadmus.eui.eu/handle/1814/75243> accessed 30 May 2023.

[140] Germany's comment on Art 2(7) of the Commission Proposal (DE Comments on the AI Act, WK 13423/2022 INIT, 6.10.2022).

[141] Mir (n 12) 13.

[142] Athina Sachoulidou, "Harnessing AI for Law Enforcement: Solutions and Boundaries from the Forthcoming AI Act" (2024) 15 New Journal of European Criminal Law 117; Francesca Palmiotto, "Artificial Intelligence and the Transformation of Criminal Trials : Preserving Fairness in Europe" (Thesis, European University Institute 2023) <https://cadmus.eui.eu/handle/1814/75243> accessed 30 May 2023; See also in public administrative law Mir (n 12).

[143] Palmiotto (n 142).

Cite this article: F Palmiotto (2025). "The AI Act Roller Coaster: The Evolution of Fundamental Rights Protection in the Legislative Process and the Future of the Regulation". *European Journal of Risk Regulation* **16**, 770–793. https://doi.org/10.1017/err.2024.97