ROBERT
SCHUMAN
CENTRE

RSC 2023/59
Robert Schuman Centre for Advanced Studies

WORKING PAPER

The EU AI Act: A Medley of Product Safety and Fundamental Rights?

Marco Almada, Nicolas Petit

European University Institute
Robert Schuman Centre for Advanced Studies

The EU AI Act: A Medley of Product Safety and Fundamental Rights?

Marco Almada, Nicolas Petit

RSC Working Paper 2023/59

ISSN 1028-3625

Published in October 2023 by the European University Institute.
Badia Fiesolana, via dei Roccettini 9
I – 50014 San Domenico di Fiesole (FI)
Italy
www.eui.eu

With the support of the
Erasmus+ Programme
of the European Union

The European Commission supports the EUI through the European Union budget. This publication reflects the views only of the author(s), and the Commission cannot be held responsible for any use which may be made of the information contained therein.

Robert Schuman Centre for Advanced Studies

The Robert Schuman Centre for Advanced Studies, created in 1992 and currently directed by Professor Erik Jones, aims to develop inter-disciplinary and comparative research on the major issues facing the process of European integration, European societies and Europe's place in 21st century global politics.

The Centre is home to a large post-doctoral programme and hosts major research programmes, projects and data sets, in addition to a range of working groups and ad hoc initiatives. The research agenda is organised around a set of core themes and is continuously evolving, reflecting the changing agenda of European integration, the expanding membership of the European Union, developments in Europe's neighbourhood and the wider world.

For more information: http://eui.eu/rscas

The EUI and the RSC are not responsible for the opinion expressed by the author(s)

Abstract

The European Union ("EU") draft for an Artificial Intelligence Act (the AI Act) is a legal medley. Under the banner of risk-based regulation, the AI Act combines two repertoires of European Union (EU) law, namely product safety and fundamental rights protection. However, the proposed medley can fail if it does not account for the structural differences between the two legal repertoires. This paper maps how three classical issues of law and technology—the pacing problem, a mismatch between means and ends, and institu-tional path dependence—manifest themselves in the AI Act. After this diagnosis, it propos-es some adjustments to the text and spirit of the AI Act.

Keywords

risk regulation; product safety; fundamental rights; harmonized standards; legitimacy

Acknowledgements

The authors thank Anca Radu, Albert Sanchez-Graells, Anupam Chander, Deirdre Ahern, Andrea Bertolini, and Mireille Hildebrandt for their comments and suggestions.

Previous versions of this paper were presented at *Governing Artificial Intelligence: Designing Legal and Regulatory Responses* (Trinity College Dublin, 3 June 2022) and The Gift of the Evil D-AI-ty. Regulating Data and AI in Europe (Scuola Sant'Anna Pisa, 1–2 December 2022), as well as the Lisbon Digital Rights and Freedoms's *Workshop on Ongoing Work* (Lisbon, 2 February 2023).

Marco Almada's work on this paper was partially funded by doctoral grants from the ASPIRE programme and Fundación Carolina.

1. Introduction

Artificial intelligence ("AI") dominates regulatory debates around the world. Keeping with its reputation as a global rule maker in digital matters,[1] the European Union (EU) is developing a pioneering legislation known as the AI Act. A draft text was issued on 21 April 2021,[2] followed by the Council's general approach on 6 December 2022[3] and the European Parliament's negotiating position on 14 June 2023.[4] Based on their institutional positions, all three institutions are now engaging in informal discussions to agree on the final shape of the text,[5] with a view to adopting a final version by the end of 2023.[6]

While the informal negotiations may still lead to changes in the final form of the AI Act, the positions adopted by the European Commission, the Council, and the Parliament already suggest considerable agreement regarding crucial elements of the regulatory approach. At this stage, it is a safe bet that the AI Act will be oriented towards a specific regulatory target: laying down rules for "secure", "trustworthy", and "ethical" AI in the EU.[7] These ambitions, however, might be undermined by the choice of regulatory technique underpinning the AI Act.

The main issue is this: the AI Act is conceived as a product safety instrument. There are valid pragmatic reasons for the EU to follow a product safety approach, not least to leverage its broad regulatory competencies in this sector.[8] However, in at least three fundamental ways, the legal principles governing EU product safety regulation risk misdirecting regulators towards sub-optimal solutions regarding AI systems. First, a product safety framework invites regulators to think that AI systems have well-defined purposes, mischaracterizing the challenges and opportunities posed by a class of AI technology called general-purpose systems. Second, a product safety framework might tempt regulators to address issues in terms of safety risks, a framing that is ill-suited to deal with the fundamental rights issues that justify an AI Act in the first place. Third, the regulatory structure contemplated in the AI Act might be tempted to piggyback on existing product safety institutions and underestimate the need to develop new capabilities required for the "residual" protection of fundamental rights in AI-related contexts.

The AI Act is a difficult read, even for lawyers. Its technical formulation has allowed diverse claims about what it supposedly does—or fails to do. Activists, techno-solutionists, and sceptics have all railed against the Act, advancing high-level arguments that sometimes pay little heed to the fine print of the instrument. Contrastingly, this paper engages with the AI Act on its own terms, using tools

[1] See, *inter alia*, Anu Bradford, *The Brussels Effect: How the European Union Rules the World* (Oxford University Press 2020).

[2] *Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative acts* [2021]. Any references to articles and recitals of the "AI Act" throughout the paper refer to the Commission proposal, unless specified otherwise, and any references to page numbers refer to its explanatory memorandum and other accompanying materials.

[3] 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' (25 November 2022), adopted on 6 December 2022. References to articles and recitals introduced or changed by the Council will be labelled in this paper as "Article/Recital x AI Act (Council general approach)".

[4] *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)) (Ordinary legislative procedure: first reading)* [2023]. References to articles and recitals introduced or changed by the Parliament will be labelled in this paper as "Article/Recital x AI Act (Parliament amendments)".

[5] For an overview of the EU legislative procedure and the role played by informal trilogues, see Tiago Sérgio Cabral, 'A Short Guide to the Legislative Procedure in the European Union' (2020) 6 UNIO – EU Law Journal 1, 161.

[6] Luca Bertuzzi, 'AI Act Enters Final Phase of EU Legislative Process', *EURACTIV* (14 June 2023) <https://www.euractiv.com/section/artificial-intelligence/news/ai-act-enters-final-phase-of-eu-legislative-process/> accessed 22 June 2023.

[7] See European Commission, 'Artificial Intelligence for Europe' (2018).

[8] Yet, this approach is not unanimous in EU legal scholarship. See, e.g., Pieter Van Cleynenbreugel, 'EU By-Design Regulation in the Algorithmic Society: A Promising Way Forward or Constitutional Nightmare in the Making?' in Hans-W Micklitz and others (eds), *Constitutional Challenges in the Algorithmic Society* (Cambridge University Press 2021) 211; Lilian Edwards, 'Regulating AI in Europe: Four Problems and Four Solutions' (Expert Opinion, March 2022).

from law and technology scholarship to cast a concrete light on some of its systemic implications.[9] After mapping the issues outlined above, the paper suggests actionable directions for law and policy changes to the AI Act.

2. Leitmotiv: the AI Act as Risk-based Regulation

The AI Act describes itself as a risk-based instrument.[10] It does not lay down uniform rules for all AI systems. Instead, the AI Act delineates three distinct legal regimes for AI systems.[11] Each regime lays down legal obligations that correlate with anticipated risks to public interests and values protected by EU law.[12] In every case, the determination of the applicable regime requires examining the context, intent, or technique underpinning an AI system.

2.1 Regulatory Regimes in the AI Act

The first regime covers AI systems raising an unacceptable risk to EU values such as the respect for human dignity, democracy, or the protection of fundamental rights.[13] The original proposal by the European Commission lists four "artificial intelligence practices" deemed unacceptable and forbids the placing on the market, putting into service or use of an AI system for any such purpose.[14] For example, the AI Act prohibits AI techniques that distort a person's behaviour in a way that is likely to cause physical or psychological harm,[15] either through subliminal manipulation of individual behaviour[16] or the exploitation of the vulnerability of children or older people.[17] Through these prohibitions, the AI Act adopts a precautionary approach, denying market access whenever the risks stemming from their use are deemed too high for risk-mitigating interventions.[18]

The second regime concerns AI systems that pose a high risk to health, safety, and fundamental rights.[19] Under the AI Act, such systems can only access the EU market if they conform to various legal requirements.[20] Some of these requirements are technical, specifying properties that any high-risk AI system must have,[21] while others require the creation of organizational mechanisms to monitor risks related to that system.[22] To a large extent, these requirements are directed at the provider of

[9] The paper thus relies on "law and technology" as a repository of knowledge about previous technical challenges faced by the law, along the "menu" approach discussed by Michael Guihot, 'Coherence in Technology Law' (2019) 11 Law, Innovation and Technology 311.

[10] Recital 14 AI Act.

[11] Articles 5 and 6 AI Act. Some sources, and the Commission itself, refer to *four* risk tiers, as Article 52 AI Act establishes transparency obligations for certain AI systems. However, these obligations—as well as the Parliament's proposed rules for foundation models under Article 28b AI Act (Parliament amendments)—also apply to high-risk systems that meet the conditions. They are, as such, better understood as additional rules rather than a separate regulatory regime.

[12] Recital 4 AI Act. The AI Act relies on an *ex ante* definition of risk categories, instead of a contextual assessment as in the GDPR, leading to a *top-down* approach to risk: Giovanni De Gregorio and Pietro Dunn, 'The European risk-based approaches: Connecting constitutional dots in the digital age' (2022) 59 Common Mkt L Rev 473, s 4.

[13] Recital 15 AI Act

[14] Article 5(1) AI Act. Both the Parliament and the Council have proposed changes to the list of forbidden AI applications, and civil society organizations are pushing for further alterations. See, e.g., 'EU Lawmakers Must Regulate the Harmful Use of Tech by Law Enforcement in the AI Act' (*European Digital Rights (EDRi)*, 20 September 2023) <https://edri.org/our-work/civil-society-statement-regulate-police-tech-ai-act/> accessed 25 September 2023. Yet, the overall approach to systems posing an unacceptable risk remains unchanged.

[15] Recital 16 AI Act.

[16] Article 5(1)(a) AI Act. On the AI Act's formulation of subliminal manipulation, see Rostam J Neuwirth, *The EU Artificial Intelligence Act: Regulating Subliminal AI Systems* (Routledge 2022).

[17] On vulnerability in the context of AI, see, *inter alia*, Sofia Ranchordas and Luisa Scarcella, 'Automated Government for Vulnerable Citizens: Intermediating Rights' (2022) 30 Wm & Mary Bill Rts J 373.

[18] On the role of the precautionary principle in EU regulation, see Kenisha Garnett and David J Parsons, 'Multi-Case Review of the Application of the Precautionary Principle in European Union Law and Case Law' (2017) 37 Risk Analysis 502.

[19] Article 6 AI Act specifies that a system is classified as high-risk if it meets one of two criteria. A system is a high-risk system if it is a product that falls into the purview of existing EU product safety law, or a safety component of such a product. Alternatively, any system destined for the applications listed in Annex III AI Act is a high-risk system.

[20] Article 8(1) AI Act.

[21] Articles 10–15 AI Act.

[22] Article 9 AI Act.

the system,[23] who is expected to comply with these rules before the system is placed on the market, put into service, or used. Compliance with the rules, therefore, requires foresight of risks connected with each system and its intended applications.[24] In addition, this regime establishes a complex post-market monitoring procedure. The procedure seeks to address risks not eliminated upfront, laying down obligations for providers of AI systems[25] and other actors like national market supervisory authorities[26] and the users of said systems.[27]

According to the Commission's estimates, the high-risk classification should cover between 5% and 15% of the AI systems in use in the EU.[28] However, their regulatory regime constitutes the main part of the AI Act. This is because the Commission's proposal wants to establish harmonized rules for high-risk systems.[29] To do so, the AI Act must substantially constrain the possibility of national and subnational rulemaking directed at high-risk AI systems. For example, Member States cannot stipulate that certain applications are deemed high-risk for the purposes of national law,[30] or eliminate requirements established at the EU level.[31] All high-risk AI systems in the EU are subject to the same regulatory regime, and the role of Member States is mainly restricted to enforcing the applicable rules through their national authorities.

The third regime is a baseline that covers all other AI systems.[32] As proposed by the Commission, the AI Act does not introduce general rules directed at systems not covered by the previous two regimes,[33] and even the Parliament text only goes as far as establishing a set of guiding principles applicable to all AI systems.[34] Systems caught in this residual regime are to be governed chiefly by other legal instruments at the EU and national level. For example, AI systems placed on the market as consumer goods, or components thereof, are covered by the General Product Safety Regulation (GPSR[35]) unless they are subject to a specific consumer protection norm. Furthermore, Member States retain the competence to adopt AI-specific rules in any domains not covered by EU-wide

[23] Article 3(2) AI Act defines "provider" as the natural or legal person, public authority, agency or other body that develops an AI system or has it developed with a view to commercializing it or putting the system into use under its own name or trademark. Article 29 AI Act contains a set of obligations directed at the deployers of AI systems, which the Parliament mandate seeks to extend. Nonetheless, the bulk of the obligations stemming from the AI Act remains directed at providers of high-risk AI systems.

[24] Article 16 AI Act ascribes the bulk of these responsibilities to the provider, though Articles 24–29 AI Act present some situations in which other actors—such as distributors or the users that deploy the AI system in a given context—are also subject to said obligations.

[25] Article 61 *ss.* AI Act.

[26] See Article 63 AI Act.

[27] Article 29 AI Act.

[28] *Impact Assessment Accompanying 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* [2021] 69. There is, however, some controversy surrounding these estimates: Meeri Haataja and Joanna J Bryson, 'What Costs Should We Expect from the EU's AI Act?' (20 August 2021); Andreas Liebl and Till Klein, 'AI Act: Risk Classification of AI Systems from a Practical Perspective' (March 2023).

[29] With the goal of ensuring "a uniform level of protection across the Member States and facilitate the creation of a single market for lawful, safe, and trustworthy AI": COM (2021) 206, s. 2.4.

[30] See the lack of Member State discretion in Article 6 AI Act and the restricted power that the Commission (but not the Member States) has to add new high-risk applications under Article 7 AI Act.

[31] One of the corollaries of the primacy of EU law is that national law cannot override provisions of EU legal instruments; see, e.g., Case 34/73 *Fratelli Variola S.pA v Amministrazione italiana delle Finanze* [1973] ECLI:EU:C:1973:101. This means, for example, Member States cannot weaken the data governance requirements laid down in Article 10 AI Act.

[32] According to European Commission estimates, the rules on high-risk AI systems should apply to between 5 and 15% of the AI systems used in the European Union: *AI Act Impact Assessment* (n 28) 71.

[33] Article 52 AI Act establishes transparency requirements for AI systems designed for some applications, which apply regardless of whether a system is classified as high-risk.

[34] Article 4a AI Act, as proposed by the Parliament, lists six guiding principles: human agency and oversight; technical robustness and safety; privacy and data governance; transparency; diversity, non-discrimination and fairness; and social and environmental well-being. These principles apply without prejudice to obligations set up by existing EU and national law, and compliance with them 'can be achieved, as applicable, through the provisions of Article 28, Article 52, or the application of harmonised standards, technical specifications, and codes of conduct as referred to in Article 69, without creating new obligations under this Regulation.' (Article 4a(2) AI Act).

[35] Regulation (EU) 2023/988 of the European Parliament and of the Council of 10 May 2023 on general product safety, amending Regulation (EU) No 1025/2012 of the European Parliament and of the Council and Directive (EU) 2020/1828 of the European Parliament and the Council, and repealing Directive 2001/95/EC of the European Parliament and of the Council and Council Directive 87/357/EEC (Text with EEA relevance) [2023] OJ L 135/1. The GPSR is now in force, and shall apply from 13 December 2024, replacing the General Product Safety Directive (Directive 2001/95/EC of the European Parliament and of the Council of 3 December 2001 on general product safety (Text with EEA relevance) [2001] OJ L 11/4) referenced by Recital 82 AI Act.

harmonization instruments.[36] Ultimately, providers (and users) of AI systems with lower risk levels might be subject to standard liability rules if the perceived risks lead to actual harm.

Table 1: Regulatory regimes in the EU AI Act

Regulatory regime	Unacceptable risk	High risk	Baseline
Definition in the AI Act	Article 5	Article 6	Residual
Regulatory approach	Prohibition	Harmonized rules at the EU level	General principles + sector-specific law
Legal approach	Precautionary principle	New Legislative Framework	General product safety framework
Other normative sources	Member States cannot add or remove applications from the prohibited list.	EU provides additional norms. Private and national standards have non-binding value.	EU and Member States can provide further norms. Private and national standards have non-binding value.

Table 1 above synthesizes the key features of the three regulatory regimes. Each risk regime strikes a different balance between the benefits that AI might bring and the potential hazards associated with their use. In all cases, the measures must reconcile core EU values, such as fundamental rights, with economic goals, such as market access.[37] This balancing act is, in its essence, similar to the risk-based approach followed in other EU digital regulation instruments,[38] like the General Data Protection Regulation (GDPR),[39] the Digital Services Act (DSA),[40] and the Digital Markets Act (DMA).[41] Therefore, the AI Act's reliance on risk models shares many of the virtues and vices of risk regulation and the EU approach to risk in digital technologies.[42]

[36] Though their exercise of said competence is constrained by other EU norms on technical regulation: Martin Ebers, 'Standardizing AI - The Case of the European Commission's Proposal for an Artificial Intelligence Act' in Larry A DiMatteo and others (eds), *The Cambridge Handbook of Artificial Intelligence: Global Perspectives on Law and Ethics* (Cambridge University Press 2022) 339–40.

[37] European Commission, 'Fostering a European Approach to Artificial Intelligence' (2021) 6.

[38] Gregorio and Dunn (n 12).

[39] 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 (General Data Protection Regulation) [2016] OJ L 119/1.

[40] 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) (Text with EEA relevance) [2022] OJ L 277/1.

[41] 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) (Text with EEA relevance) [2022] OJ L 265/1.

[42] See, *inter alia*, Fiona Haines, 'Regulation and Risk' in Peter Drahos (ed), *Regulatory Theory* (1st edn, ANU Press 2017); Margot E Kaminski, 'Regulating the Risks of AI' (WeRobot, Seattle, 15 September 2022) <https://papers.ssrn.com/abstract=4195066> accessed 22 August 2022.

2.2 What Regulatory Technique in the AI Act?

Risk-based regulation is a broad concept. Several regulatory techniques can be deployed under a risk-based approach, some of which can lead to different framings of risk than the ones reflected in the AI Act.[43] The Act follows a prescriptive risk-based technique, which assumes legislators can identify classes of AI practices and systems that impose unacceptable or high risk levels. From that premise, the regulation specifies exhaustive lists of AI practices and systems falling into the unacceptable[44] and high risk[45] categories.[46]

The *ex ante* specification of risk levels has been criticized as being both too vague and too strict. Any prescriptive approach cannot overcome a degree of generality, abstraction, and vagueness.[47] These features introduce some degree of legal uncertainty, and industry-connected groups have claimed that the current criteria for high-risk AI are unclear regarding the classification of many AI applications currently in use.[48] Besides, both the Parliament and the Council have sought to water down the specification contained in Article 6 AI Act by providing "escape clauses". Under the Council's general approach, classification under Article 6 AI Act would involve an additional step: a system is not considered to be high-risk if its output is purely accessory regarding the relevant action or decision to be taken.[49] Similarly, the Parliament position allows providers to notify the national supervisory authority[50] that their system does not pose 'a significant risk of harm to the health, safety or fundamental rights of natural persons'[51] and should, therefore, be exempted from the high-risk classification. These proposals have been criticized by scholars[52] and civil society organizations[53] as creating "loopholes" for circumventing regulation. Still, the cross-institutional support for them suggests that the final version of the AI Act is likely to weaken the *ex ante* approach to risk assessment.

Another distinctive approach of the AI Act's risk-based approach concerns how it frames the risks it is expected to address. Other digital regulation instruments display a preference to address risks related to specific fundamental rights—the protection of personal data for the GDPR[54] or free speech for the DSA[55]—over the safeguarding of market access. By contrast, the AI Act leans more explicitly towards internal market goals.[56] While they are endlessly mentioned in the Act's recitals and explanatory documents, fundamental rights appear sparingly in the main text,[57] and are always

[43] On the broad tent of risk-based regulation, see Kaminski (n 42) pt II.

[44] Article 5 AI Act.

[45] Article 6 AI Act defines high-risk applications, further specified in the Act's Annexes II and III.

[46] In doing so, the Commission attempts to mitigate legal uncertainty by providing more explicit guidance about what is the applicable regime for any given AI system, addressing concerns that were frequently raised throughout the public consultation process preceding the AI Act: Gloria Golmohammadi, 'Realizing the Principle of Participatory Democracy in the EU: The Role of Law-Making Consultation' (Doctor of Laws in Legal Science, University of Stockholm 2023) 234.

[47] On the open character of the technical measures required by the AI Act, see, e.g., Madalina Busuioc and others, 'Reclaiming transparency: contesting the logics of secrecy within the AI Act' (2023) 2 Eur L Open 79.

[48] See, e.g., Liebl and Klein (n 28) 27–42.

[49] Article 6(3) AI Act (Council general approach).

[50] Under Article 6(2a) AI Act (Parliament amendments).

[51] Article 6(2) AI Act (Parliament amendments), which also stipulates that environmental risk is to be taken into account for systems covered by point 2 of Annex III AI Act.

[52] Meeri Haataja and Joanna J Bryson, 'The European Parliament's AI Regulation: Should We Call It Progress?' (2023) 4 Amicus Curiae 3, 707, 712–13.

[53] 'Joint Statement: EU Legislators Must Close Dangerous Loophole in AI Act' (*Access Now*, 9 July 2023) <https://www.accessnow.org/press-release/joint-statement-eu-legislators-must-close-dangerous-loophole-in-ai-act/> accessed 25 September 2023.

[54] Historically, EU data protection law has focused on the rights of the individuals whose data is processed and the GDPR accordingly sets up measures aimed at mitigating the risks to said rights. See, *inter alia*, Przemysław Pałka, 'Data Management Law for the 2020s: The Lost Origins and the New Needs' (2020) 68 Buff L Rev 559; Thomas Streinz, 'The Evolution of European Data Law' in Paul Craig and Gráinne de Búrca (eds), *The Evolution of EU Law* (3rd edn, Oxford University Press 2021).

[55] The DSA focuses on structural risks that online platforms may create to the exercise of fundamental rights, laying down transparency and redress requirements that bind the exercise of power by said platforms. Giovanni De Gregorio, *Digital Constitutionalism in Europe: Reframing Rights and Powers in the Algorithmic Society* (Cambridge Studies in European Law and Policy, Cambridge University Press 2022) 211–14.

[56] Indeed, the absence of individual rights and redress mechanisms has been subject to critique by European civil society organizations: Letter from EDRi and others, 'An EU Artificial Intelligence Act for Fundamental Rights. A Civil Society Statement.' (30 November 2021).

[57] See, e.g., Article 14 AI Act, in which the protection of fundamental rights is presented as a safety concern.

accompanied by the health and safety imperatives that are the traditional object of EU product safety law.[58] As a result, the AI Act can sometimes seem closer to market-oriented instruments such as DMA than to instruments couched in fundamental rights terms, such as the GDPR.[59]

The Act's bias towards market access is also visible in the choice of the (AI) system as its main regulatory target.[60] The GDPR and the DSA include provisions establishing technical requirements for digital systems,[61] which resemble the technical requirements for high-risk AI in the AI Act.[62] These technical requirements are, just like the ones in the AI Act, expected to play a central role in enforcing legal provisions in the contexts in which AI systems are used.[63] However, the AI Act's focus on market access introduces a major difference: its technical requirements are directed at procedural objectives, setting standards that allow the AI system to be placed on the market (or continue in it). By contrast, the technical obligations in the GDPR and the DSA aim to ensure particular substantive values.[64] As the following sections show, this focus on the object placed on the market, put into service, or used has considerable implications for the AI Act's protective framework.

This paper uses "medley" as a running metaphor for the AI Act. More specifically, the choice of that metaphor is intended to denote the mix between fundamental rights and product safety in the Act's provisions. The Commission proposal created such a mix, and the Parliament and Council appear unlikely to deviate from it, even if, as the following sections show, they add their own twists to the song. But that particular medley needed not be the only way to regulate AI systems. In fact, the High-Level Expert Group on AI convened by the Commission to guide its work towards the AI Act never really envisioned product safety as a possible legal framework. It recommended, instead, a rights-based approach to AI regulation, to be operationalized through risk mechanisms.[65] The question then is, why did the Commission decide to adopt a product safety approach?

The best possible response is this: a product safety approach reflects the EU's limitations and strengths as an AI regulator. Because the AI Act is intended as a horizontal instrument, setting general rules for AI systems,[66] it must have a broad scope. However, the legislative competencies of the EU are, for the most part, sector-specific.[67] To address this shortcoming, most provisions of the AI Act are grounded on the EU's legislative competence for harmonizing the single market.[68] As a result, these provisions must be framed as market regulation instruments, lest they fall foul of the constitutional limits of their legal basis.[69]

[58] The AI Act uses the formula "risks to the health, safety, and fundamental rights" throughout the text in a way that mirrors the use of "risks to the health and safety" in instruments such as the GPSR.

[59] See Section 4 below.

[60] Article 1 AI Act.

[61] See, e.g., Article 25 GDPR and Article 27 DSA.

[62] Articles 8–15 AI Act.

[63] See, in the case of the GDPR,

[64] See, e.g., how Article 25 GDPR targets the data processing operation, binding controllers to adopt measures to protect a broad range of principles, rather than focusing on the computer systems processing personal data: Marco Almada and others, 'Article 25. Data Protection by Design and by Default' in Indra Spiecker gen. Döhmann and others (eds), *General Data Protection Regulation: Article-by-article commentary* (Beck; Nomos; Hart Publishing 2023) para 3.

[65] See, e.g., AI HLEG, 'Policy and Investment Recommendations for Trustworthy AI' (Independent High-Level Expert Group on Artificial Intelligence, 29 June 2019).

[66] *COM(2021) 206 final* (n 2) 3.

[67] See, e.g., Kieran Bradley, 'Legislating in the European Union' in Steve Peers and Catherine Barnard (eds), *European Union Law* (4th edn, Oxford University Press 2023) 108–15. In particular, it is important to highlight that, under Article 51(2) CFR, the EU has no stand-alone competence to legislate on fundamental rights. Instead, the EU's obligations to protect fundamental rights provide side constraints that must be observed as the EU exercise the competencies conferred to it by the Member States.

[68] Article 114 TFEU, as highlighted in *COM(2021) 206 final* (n 2) 6. See, however, the discussion on the role of Article 16 TFEU in Section 4 below.

[69] On the breadth and limits of Article 114 TFEU, see, *inter alia*, Robert Schütze, *European Constitutional Law* (Oxford University Press 2021) 233–35.

Yet, constitutional constraints only tell part of the story. The DSA and the DMA are also grounded on the same legal basis, but they do not follow product safety approaches. Patterning the AI Act after product safety legislation is a deliberate choice by the Commission and the European legislator, not a constitutional requirement.

By relying on a product safety framework, the EU plays to its regulatory strengths. EU product safety law has been in force for decades,[70] influencing regulation in other jurisdictions.[71] In addition, recent product safety instruments have increasingly dealt with risks relating to software systems.[72] Using a regulatory framework based on product safety thus ensures the AI Act can benefit from decades of accumulated knowledge and scholarship on product safety law. This approach also allows the EU and its Member States to rely on existing institutional capabilities for enforcing the Act.[73] The choice of framing the AI Act as a product safety instrument[74] allows the EU to draw from a tried-and-true approach instead of creating new norms from the ground up.

This is not to say that regulating AI as a product has no downsides. In fact, the very effectiveness of the product safety framework may distract regulators from significant differences between AI systems and products. In the following sections, we consider three ways in which reducing AI risk to product risk can be tempting and argue that each reduction makes the AI Act less adequate for the goals it intends to pursue.

3. Off-Beat Specification: AI Functionalities and The Pacing Problem

The AI Act considers that an AI system's risk level generally depends on "the function performed … but also on the specific purpose and modalities for which that system is used".[75] As the previous section shows, all three regulatory subsystems in the Act are defined in terms of how AI is used: Article 5 defines some practices as unacceptable, and the distinction between high-risk and minimal-risk systems is also made by distinguishing between specific application areas. Much care is taken to describe in advance the specific "function" or "use case" that gives rise to the risks that the AI Act wants to eliminate or regulate.[76]

Specification and categorization are hallmarks of the continental legal tradition. They give subjects of the law maximum predictability.[77] But specification and categorization produce over or under-inclusive legal rules. When deployed towards a dynamic technology, the problems of over and under-inclusion are not conjectures. They are certainties. Given the existing state of AI technology, the AI Act's choice to submit AI systems to risk regimes by specifying narrow functionalities walks straight into problems of legal loopholes and overreach.

[70] Geraint Howells and others, *Rethinking EU Consumer Law* (1st edn, Routledge 2017) 262–64.

[71] Charlotte Siegmann and Markus Anderljung, 'The Brussels Effect and Artificial Intelligence: How EU Regulation Will Impact the Global AI Market' (August 2022) 78–80.

[72] Two European Commission officials have stressed that a key source of inspiration for the AI Act has been the regulatory framework developed for medical devices, which often rely on software, and has "been put to the test successfully in the course of the last 15 years": Gabriele Mazzini and Salvatore Scalzo, 'The Proposal for the Artificial Intelligence Act: Considerations around Some Key Concepts' in Carmelita Camardi (ed), *La via europea per l'Intelligenza artificiale* (Cedam 2023).

[73] See, e.g., Recital 64 AI Act and its acknowledgement of "the more extensive experience of professional pre-market certifiers in the field of product safety".

[74] Even if, as discussed below, one with considerable departures from existing practices in product safety regulation.

[75] *COM(2021) 206 final* (n 2) 13.

[76] The lists and categories work like absolute presumptions. Risk levels are not studied, they are assumed. The significance of the risks incurred cannot be rebutted by a showing of facts.

[77] Patricia Popelier, 'Five Paradoxes on Legal Certainty and the Lawmaker' (2008) 2 Legisprudence 47.

3.1 Misunderstanding Technological Evolution? Towards General Purpose AI

The EU lawmakers assume they can specify the particular "function" or "use case" of AI systems (hereafter, "functionalities"). Such an assumption does not appear unreasonable. Many current applications of AI are highly specialized, developed and validated for operation in a narrow context and scope of activities, such as content moderation, document review, or credit scoring.[78]

Nevertheless, the assumption above can be problematic in the context of technological change. If the technology is unstable, the law must catch up to adapt to the specified functionalities. This point is particularly salient regarding computing technologies. The history of computing technologies is one of instability, owing to the relentless process of recombining hardware and software components towards new purposes.[79] To cope with this instability, regulators can use strategies to future-proof law, such as proactive updates of legislation and experimental legislation.[80] In fact, the AI Act already includes some mechanisms to that effect.[81]

Now, the problem is not so much one of speculation but one of feasibility. The assumption that functionalities can be specified in the law is untenable regarding a technology that evolves towards general-purpose attributes.

A recent development in machine learning involves the construction of large-scale models that can perform some tasks of a relatively general character. As of 2023, the most salient example of this approach is provided by large language models (LLMs), such as GPT4 or LLaMA, trained on huge amounts of text data.[82] This training allows such systems to perform various tasks, such as generating coherent text, which are themselves valuable. Yet, the main value of these systems comes from the possibility of fine-tuning them for specific tasks through processes that are much less complex and data-intensive than the creation of the original large language models.[83] AI systems such as GPT-4 are therefore presented as general-purpose tools which lack a specific function or use case.

How does the general-purpose evolution of AI technology impact the AI Act framework? With a general-purpose AI system, it is tough to specify in advance functionalities or use cases that raise safety concerns.

The problem is not to predict whether new functionalities or use cases will raise safety concerns. Some will, and some are already a reality.[84] The difficulty is to predict which functionalities and what safety concerns will arise. The problem is compounded by the fact that it is reasonable to assume that the functionalities and safety issues arising from general-purpose AI could be very different from those that concern us today.

[78] Some progress has been made towards the development of AI systems that can perform more than one task—see, e.g., Scott Reed and others, 'A Generalist Agent' (5 December 2022)—but even the most general systems that exist today are still competent only in very narrow domains.

[79] See, e.g., Sara Hooker, 'The Hardware Lottery' (2021) 64 Commun ACM 58.

[80] On future-proofing, see Sofia Ranchordás and Mattis van 't van't Schip, 'Future-Proofing Legislation for the Digital Age' in Sofia Ranchordás and Yaniv Roznai (eds), *Time, Law, and Change: An Interdisciplinary Study* (Hart Publishing 2020).

[81] See, in particular, the regulatory sandboxes enabled by Article 53 AI Act and the mechanisms for updating Annex III AI Act laid down in Article 7.

[82] For overviews of the current state of the art, aimed at non-technical audiences, see OECD, 'AI Language Models: Technological, Socio-Economic and Policy Considerations' (Organisation for Economic Co-operation and Development 13 April 2023); Samuel R Bowman, 'Eight Things to Know about Large Language Models' (*arXiv.org*, 4 February 2023) <https://arxiv.org/abs/2304.00612v1> accessed 14 June 2023; Johanna Okerlund and others, 'What's in the Chatterbox? Large Language Models, Why They Matter, and What We Should Do About Them' (Report, Technology Assessment Project, 2022); Matthias Gallé, 'Foundation Models in AI: What Impact for Policies and Law?' (Memo, Frontier Talks, 30 May 2022).

[83] See, e.g., OECD (n 82) 32.

[84] See, *inter alia*, Laura Weidinger and others, 'Taxonomy of Risks Posed by Language Models' in (FAccT '22, ACM 2022) FAccT '22 214; OECD (n 82); Noëlle Gaumann and Michael Veale, 'AI Providers as Criminal Essay Mills? Large Language Models Meet Contract Cheating Law' (SocArXiv 9 June 2023); Josh A Goldstein and others, 'Generative Language Models and Automated Influence Operations: Emerging Threats and Potential Mitigations' (2023).

Regulators are thus faced with a difficult choice. Suppose they act at the early stages of technological development. In that case, they might miss important information about general-purpose AI technologies and their impact on society, leading to inadequate solutions to legal problems. But, if they take too long, regulation will only be adopted once adverse effects of these technologies are already manifest, with potentially catastrophic consequences.[85] Law and technology scholars refer to this quandary as the pacing problem.[86]

The deep uncertainty surrounding general-purpose AI technologies makes the pacing problem particularly hard. First, AI-generated texts might allow cheaper production of science, evidence, and expertise. Are there safety risks arising from delegating knowledge production to AI systems? In what contexts?[87] Second, AI voice applications can generate chatbots from dead people. Are there safety risks arising from talking with the dead? In what contexts?[88] In questions such as these, the stakes associated with the pacing problem become particularly high, as both a premature and ill-conceived response and a precise but delayed one might cause considerable societal damage.

3.2 Dealing with Legal Loopholes and Overreach? Deep Uncertainty and Hard Tradeoffs

The AI Act, however, offers few answers to questions like these. By focusing on known functionalities and safety risks, the Act directs attention towards problems that have already been detected or anticipated. Doing so might foster confidence that a general-purpose AI system is safe if it meets all the criteria specified in the law.[89] Given the mixed track record of forecasts of technological development,[90] this confidence might be unwarranted.

When it comes to risks that have been effectively foreseen, the specification of functionalities or use cases in the law might lead to overcorrection by developers of general-purpose AI. Like lawmakers, developers cannot predict the breadth and depth of applications of general-purpose AIs. Yet, it is reasonable to anticipate that general-purpose AI systems can be deployed towards a narrow functionality or safety concern covered in the AI Act. With this background, searching for an optimal allocation of the compliance burden between developers and deployers involves a hard tradeoff. The tradeoff consists in ensuring (i) an appropriate level of incentives towards the development of general-purpose AI; and (ii) targeting the agent with the best technical capabilities to ensure compliance with safety concerns. It is hard to have both at the same time. The developer is often the agent whose incentives to develop general-purpose AI must be protected. But the developer is often the agent with the best technical capabilities to mitigate safety concerns.

When the Commission originally introduced its proposal, general-purpose AI was not a salient topic in public discourse. By contrast, the Council and the Parliament produced their negotiating positions in a context marked by intense debates about the capabilities and risks of large language models and other systems claiming general capabilities. The result is that the Council's general approach and the Parliament's amendments both include provisions—albeit different ones—on certain classes

[85] On the different kinds of impact AI might have in society, see Nicolas Petit and Jerome De Cooman, 'Models of Law and Regulation for AI' in *The Routledge Social Science Handbook of AI* (Routledge 2021) 208.

[86] See, e.g., Gary E Marchant, 'Addressing the Pacing Problem' in Gary E Marchant and others (eds), *The Growing Gap Between Emerging Technologies and Legal-Ethical Oversight: The Pacing Problem* (The International Library of Ethics, Law and Technology, Springer Netherlands 2011).

[87] See, e.g., Abeba Birhane and others, 'Science in the Age of Large Language Models' (2023) 5 Nat Rev Phys 5, 277; Joshua Habgood-Coote, 'Deepfakes and the Epistemic Apocalypse' (2023) 201 Synthese 103.

[88] See, e.g., Edina Harbinja and others, 'Governing Ghostbots' (2023) 48 Computer Law & Security Review 105791.

[89] Indeed, trustworthiness has been closely connected to compliance with the law in the AI Act: Charly Derave and others, 'The Risks of Trustworthy Artificial Intelligence: The Case of the European Travel Information and Authorisation System' (2022) 13 Eur J Risk Regul 389. The Parliament amendments makes such a connection explicit in Article 4a(2), which specifies that the principles laid down in Article 4a(1) are guidelines for implementing existing legal obligations rather than a source of those in themselves.

[90] See, e.g., Frank Bannister and Regina Connolly, 'The Future Ain't What It Used to Be: Forecasting the Impact of ICT on the Public Sphere' (2020) 37 Government Information Quarterly; Neil Pollock and Robin Williams, 'The Venues of High Tech Prediction: Presenting the Future at Industry Analyst Conferences' (2015) 25 Information and Organization 115.

of AI systems perceived as general-purpose. Neither of the proposed rules abandons the idea that developers should be at the forefront of risk mitigation. Still, they both provide additional requirements that a provider must observe when creating general-purpose systems.

For the Council, a 'general purpose AI system' is intended to carry out generally applicable functions,[91] which can be used in a plurality of contexts and integrated into other systems. Under Title IA of the Council general approach, general-purpose AI systems that may be used as high-risk AI systems or as components thereof are subject to the technical requirements applicable to high-risk AI,[92] and their providers are subject to informational requirements aimed at governing the potential high-risk uses.[93]

At first glance, the Council's approach creates a burdensome regulatory regime. Any system that can be used for high-risk purposes would be subject to all requirements for high-risk AI. This would amount to classifying general-purpose AI systems as inherently high-risk, as the value of these systems comes precisely from the possibility of using them in many contexts. However, the Council text provides a simple mechanism to avoid this classification. If a provider of a general-purpose system specifies to the deployers that their system is not to be used for high-risk purposes, the system is not subject to the high-risk rules.[94] As a result, the stricter regime is likely to apply only for those general-purpose AI systems that are marketed for high-risk applications.[95]

The Parliament proposal does not single out general-purpose AI.[96] Instead, it introduces new rules directed at a different legal figure: the foundation model, defined as 'an AI system model that is trained on broad data at scale, is designed for generality of output, and can be adapted to a wide range of distinctive tasks'.[97] These models are not defined as AI systems in themselves but rather as components that can be used to power general-purpose AI systems or systems with an intended purpose.[98] Accordingly, the obligations targeted at foundation model providers are meant to ensure that their models do not introduce issues into the downstream systems that rely on them.

Two categories of obligations fall into providers of foundation models. Some of these measures impose substantive requirements: providers must adopt measures to mitigate reasonably foreseeable risks,[99] ensure the quality of the datasets used to train the model,[100] and ensure the technical quality of model outputs.[101] Others establish transparency requirements, such as supplying documentation and instructions of use for downstream providers[102] or registering the model in an EU database.[103] Unlike the Council's approach, the Parliament does not make the providers of general-purpose AI tools directly responsible for high-risk use. However, it obliges those providers to adopt measures to mitigate risks and ensure that foundation models do not preclude downstream providers from discharging their own duties regarding high-risk AI systems built upon these models.

[91] Article 4(1b) AI Act (Council general approach): '…such as image and speech recognition, audio and video generation, pattern detection, question answering, translation and others'.

[92] Article 4b(1) AI Act (Council general approach).

[93] Article 4b(2–6) AI Act (Council general approach).

[94] Article 4c(1) AI Act (Council general approach).

[95] Under Article 4c(2) AI Act (Council general approach), a provider cannot invoke this carve-out in bad faith or if they have have sufficient reason to believe the product will be misused.

[96] The term is defined in Article 3(1d) AI Act (Parliament amendments), but it is only used once in the binding text, within Article 28(1)(ba) AI Act (Parliament amendments).

[97] Article 3(1c) AI Act (Parliament amendments).

[98] Recital 60e AI Act (Parliament amendments).

[99] Article 28b(2)(a) (Parliament amendments).

[100] Article 28b(2)(b) (Parliament amendments).

[101] Article 28b(2)(c-d) (Parliament amendments).

[102] Article 2b(2)(e) (Parliament amendments).

[103] Article 2b(2)(g) (Parliament amendments).

The Parliament compromise offers a reasonable answer to the current regulatory challenges of general-purpose AI. It allows for the accountability of actors in the AI supply chain,[104] tapping into their technical knowledge but using the informational transparency requirements to control any risks that might require further attention. However, the use of 'foundation model' as the regulatory target raises two concerns. The first one concerns the term itself, which reflects a specific paradigm of general-purpose AI[105] and might bind regulation to one conception of AI. Second, the term 'model' remains undefined in the AI Act, leading to uncertainty about how this concept relates to the general definition of 'system' that guides the rest of the AI Act. For these reasons, we believe the trilogues would do well to follow the Parliament's approach, but the final text of the AI Act should use the term instead of 'foundation models' throughout the text.

4. Different Drums: The Mismatch Between Product Safety Means and Fundamental Rights Ends

Fundamental rights are mentioned throughout the AI Act as an overriding public interest that warrants legislative protection.[106] In particular, Article 65(1) AI Act extends the definition of product risks to include risks to fundamental rights.[107] The result is a product safety instrument heavily couched in fundamental rights language.

The AI Act is not the first product safety instrument to cover fundamental rights.[108] The EU regulation that lays down harmonized rules for medical devices explicitly refers to the protection of fundamental rights in general[109] and personal data more specifically[110] while including extra safeguards to two specific freedoms: freedom of expression and freedom of the press.[111] More generally, the EU is constitutionally required to protect fundamental rights as it exercises its powers, including in product safety.[112]

However, the AI Act displays a higher level of engagement with fundamental rights than other EU product safety instruments.[113] This can be seen in the practical requirements imposed on AI systems. The segmentation of AI systems into various risk tiers, outlined in Section 2, puts risks to fundamental rights on an equal footing with the risks to health and safety that are the bread and butter of product safety law.[114] Various essential requirements laid down for high-risk AI systems are formulated in terms of fundamental rights, such as the need to indicate circumstances in which the use of the AI system may impose risks[115] or to design suitable mechanisms for human oversight of the AI system.[116] Finally, conformity with essential requirements must be assessed, considering how well an AI system minimizes or eliminates risks to fundamental rights.[117] Fundamental rights are not an

[104] More broadly, see Jennifer Cobbe and others, 'Understanding Accountability in Algorithmic Supply Chains' in 2023 ACM Conference on Fairness, Accountability, and Transparency 1186.

[105] Closely associated with research developed by Stanford University: Rishi Bommasani and others, 'On the Opportunities and Risks of Foundation Models' (arXiv 7 December 2022).

[106] COM (2021) 206, p. 11. Both the explanatory memorandum and the impact assessment for the AI Act discuss the potential impact of AI technologies on the various individual and collective rights protected by the CFR. These discussions guide the interpretation of the AI Act and are reflected in provisions that specifically mention human rights.

[107] See, e.g., how Article 65(1) AI Act extends the definition of product risks to include risks to fundamental rights.

[108] Consumer protection is itself a fundamental right in EU law, enshrined in Article 38 CFR.

[109] Recital 89 of the Medical Devices Regulation: Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC (Text with EEA relevance) [2017] OJ L 117/1.

[110] Recital 69 Medical Devices Regulation.

[111] Article 1(16) Medical Devices Regulation.

[112] Article 51(1) CFR.

[113] See Article 114(3) TFEU.

[114] These protections are not a radical departure from the EU product safety framework. Instead, they reflect the constitutional value of the CFR and the role of approximating law in protecting the health and safety of the EU public. See, e.g., the use of risks of adverse impact to fundamental right as one of the criteria for updating the list of high-risk AI systems: Article 7 AI Act.

[115] Article 13(3)(b)(iii) AI Act.

[116] Article 14(2) AI Act.

[117] Article 43(6) AI Act. Article 41(1) AI Act empowers the Commission to adopt "common specifications" for certain technologies if stan-

"afterthought"[118] to the AI Act but its backbone.

This centrality notwithstanding, the role of fundamental rights in the AI Act is one of its most criticized aspects.[119] In response to these critiques, the Parliament proposes three major alterations to the AI Act. First, its amendments include a variety of rights-based provisions, such as a right to an explanation for those affected by decisions based on high-risk AI systems.[120] Because these new provisions exceed the purview of Article 114 TFEU, a second alteration changes the legal basis of the Act. Under the Parliament amendments, the AI Act is entirely grounded both on Article 114 TFEU and the EU's competence to regulate personal data protection under Article 16 TFEU.[121] Finally, the Parliament text further expands the "health, safety, and fundamental rights" formula by including new values to be protected: democracy, the rule of law, and the environment.[122] The Parliament amendments thus respond to a perceived gap by underscoring the centrality of fundamental rights to the AI regulatory regime.[123]

Such measures miss the forest for the trees. Critiques of the AI Act are right to suggest that it will not ensure an adequate by protection of fundamental rights. But they are wrong to suggest that this is due to insufficient attention to rights-based demands—and that adding more fundamental rights provisions in the Act is the best solution to the problem.

Inadequate protection of fundamental rights arises instead from a mismatch between means and ends in the AI Act. The Act aims to achieve a fundamental rights protection end by recourse to product safety means. At an abstract level, the approach works. Product safety and fundamental rights protection both purport to mitigate or prevent risks. Yet, product safety and fundamental rights protection differ in substantial methodological ways. The instruments used for protecting individuals from product risks rely on a formalized logic of evaluation, which differs from the logic of proportionality that guides fundamental rights assessments. Furthermore, product safety instruments aim to keep risk below a level deemed satisfactory rather than minimize it. Attempts to address fundamental rights concerns through a product safety lens are, therefore, likely to overlook important concerns. Table 2 below summarizes the main differences between the approaches, which are discussed in the remainder of the section.

dard setting processes are insufficient to address fundamental rights concerns in a given context.

[118] See, e.g., Céline Castets-Renard and Philippe Besse, 'Ex Ante Accountability of the AI Act: Between Certification and Standardization, in Pursuit of Fundamental Rights in the Country of Compliance' in Céline Castets-Renard and Jessica Eynard (eds), *Artificial Intelligence Law Between Sectoral Rules and Comprehensive Regime Comparative Law* (Bruylant 2023) 624

[119] Edwards (n 8) 13–15; EDRi and others (n 56).

[120] Article 68(b) AI Act (Parliament amendments).

[121] Recital 2a AI Act (Parliament amendments).

[122] See, e.g., Recital 1 AI Act (Parliament amendments): 'The purpose of this Regulation is to promote the uptake of human centric and trustworthy artificial intelligence and to ensure a high level of protection of health, safety, fundamental rights, democracy and rule of law and the environment from harmful effects of artificial intelligence systems in the Union while supporting innovation and improving the functioning of the internal market.'

[123] As proposed, the extension also clashes with established EU constitutional practice. As the Parliament amendments provide no grounds to distinguish between provisions based on Article 16 TFEU and those stemming from Article 114 TFEU, they arguably fail to provide a clear legal basis for the act, as required by CJEU case law: see, e.g., Elise Muir, *An Introduction to the EU Legal Order* (Cambridge University Press 2023) 171–73. This issue seems to be addressable, but the constitutional discussion exceeds the scope of this paper.

Table 2: Comparison between rationales in product safety law and constitutional reasoning

Product safety	Fundamental rights
Actuarial risks predominate	Actuarial, sociopolitical, and cultural risks
Risks stem from the technical object	Risks stem from the sociotechnical context
Small world: known and consistent problems	Multidimensional harm and wicked problems
Satisficing technical baselines	Constrained maximization of principles

4.1 Dissonant Conceptions of Risk

The AI Act defines risk as "…the combination of the probability of an occurrence of a hazard causing harm and the degree of severity of that harm".[124] Under this conventional definition, a risk is a combination of its likelihood and magnitude. And risk regulation focuses on changing the value of one or both terms until an acceptable level.[125]

The above definition is characteristic of a product safety conception of risk, in which harm can be measured quantitatively.[126] Four properties inform such a conception. First, risks are actuarial: the undesirable outcome is harm of measurable severity to an individual, group, or the environment.[127] Second, risks are probabilistic: a reliable estimate of how likely a given event is can be produced.[128] Third, risks are discrete: each risk correlates with a specific undesirable event that might come to occur.[129] Fourth, risks are unidimensional: all risks can be described using the same harm metric, which means two risks are equivalent if the combination of their likelihood and severity leads to the same result.[130]

These four properties do not hold in contexts in which AI systems have the potential of adversely impacting fundamental rights. The hypothesis of actuarial risks does not capture contexts in which the severity of harm cannot be adequately quantified.[131] Probabilistic risk is unappealing in complex systems, in which the complete set of events that might occur is hard to describe a priori (especially when those involve human-machine interactions).[132] The premise of discrete risks can also be inadequate in many AI use cases, as undesirable outcomes might arise from cumulative

[124] Article 3(1a) AI Act (Parliament amendments). Such a definition is not present in the Commission version, which nonetheless refers to it in Article 65(1), which refers to the definition of risk in Article 3(18) Regulation (EU) 2019/1020 of the European Parliament and of the Council of 20 June 2019 on market surveillance and compliance of products and amending Directive 2004/42/EC and Regulations (EC) No 765/2008 and (EU) No 305/2011 (Text with EEA relevance) [2019] OJ L 169/1.

[125] *OECD Regulatory Policy Outlook 2021* (OECD Publishing 2021) 186.

[126] EU product law is complemented by harmonized rules for product liability, which often entail financial compensation to harmed consumers. On the current EU rules on product liability, see Howells and others (n 70) 258–89; Joasia Luzak, 'A Broken Notion: Impact of Modern Technologies on Product Liability' (2020) 11 European Journal of Risk Regulation 630.

[127] On the connection between actuarial risk and other forms of risk, see Haines (n 42) 185.

[128] The uncertainty of such measurements is usually addressed through precautionary heuristics in EU product legislation. See, *inter alia*, Garnett and Parsons (n 18); Roberto Baldoli and Claudio M Radaelli, 'Foundations of Regulatory Choice : Precaution, Innovation and … Nonviolence?' (2021) 17 Journal of contemporary European research 186; Brice Laurent, 'Regulatory Precaution' in *European Objects: The Troubled Dreams of Harmonization* (The MIT Press 2022).

[129] OECD (n 125) 185.

[130] ibid 186.

[131] In this sense, actuarial calculations provide a utilitarian reasoning for regulation, which is at odds with the deontological framing of fundamental rights as imposing limits (or side constraints) to the exercise of power in a society. See, *inter alia*, András Sajó and Renáta Uitz, *The Constitution of Freedom: An Introduction to Legal Constitutionalism* (Oxford University Press 2017) 378–96.

[132] Ortwin Renn, *Risk Governance: Coping with Uncertainty in a Complex World* (Earthscan risk in society series, Earthscan 2008) 14. To further complicate things, experts might disagree with one another—and with political stakeholders—on how to properly assess risks, a point not examined in this paper.

events, not specific ones.[133] Finally, the assumption of unidimensional risks struggles with value incommensurability, that is, in situations where competing values cannot be reduced to manifestations of a more fundamental value.[134]

Some examples can illustrate these limits. Harm from AI systems to human personality cannot be expressed in a numerical estimate.[135] For example, how can one account for the loss of dignity felt by lawyers who realize that LLMs just destroyed 50% of their comparative advantage in brief writing? At best, an ordinal measure of risk is possible.

Similarly, AI systems may not harm a specific individual or group but instead affect social interactions between these groups.[136] For example, the diffusion of deep fakes may undermine social trust if individuals assume that everything—or, even worse, everything they dislike—is artificially generated.[137] And the Parliament's inclusion of democracy and the rule of law in the AI Act suggests that concerns of this kind are increasingly relevant for AI governance.

Lastly, harm from AI systems can also generate an "outside context problem",[138] that is, a situation so far away from the usual expectations of a society that it cannot be predicted until it occurs.[139] One example is Nick Bostrom's "treacherous turn", in which an AI system reaching super intelligence would enslave humans to optimize its performance concerning its designated preference function.[140]

By focusing on product risks, the product safety framework leaves out various types of risk to fundamental rights that may be caused or amplified by the adoption of AI technologies.

4.2 Different Ambitions Towards Risks

In product safety, risk is addressed through a satisficing logic. As long as an AI system meets specified safety requirements, it does not matter if the system barely meets the applicable standard or surpasses it by far.[141] To a certain extent, fundamental rights protection in the EU legal system adopts a similar binary approach. Each right has an "essence" that must be upheld.[142] However, beyond that essence, the protection of fundamental rights follows a logic of optimization.[143] Fundamental rights must be protected and promoted to the maximum extent. Only measures that are the least restrictive of fundamental rights can be tolerated under the principle of proportionality.

[133] Burkhard Schafer, 'Death by a Thousand Cuts: Cumulative Data Effects and the Corbyn Affair' (2021) 45 Datenschutz und Datensicherheit - DuD 385.

[134] Some of the most widely accepted theories on fundamental rights rely on balancing as a means to resolve—perceived or real—conflicts between said rights: Aharon Barak, 'Proportionality Stricto Sensu (Balancing)' in *Proportionality: Constitutional Rights and their Limitations* (Cambridge Studies in Constitutional Law, Cambridge University Press 2012); Jan-R Sieckmann, 'To Balance or Not to Balance: The Quest for the Essence of Rights' in Jan-R Sieckmann (ed), *Proportionality, Balancing, and Rights: Robert Alexy's Theory of Constitutional Rights* (Law and Philosophy Library, Springer International Publishing 2021).

[135] This concern becomes particularly salient when one considers that some aspects of human personality might be inherently irreducible to the kind of metrics used to quantify the harms generated by a product: Mireille Hildebrandt, 'Privacy as Protection of the Incomputable Self: From Agnostic to Agonistic Machine Learning' (2019) 20 Theoretical Inquiries L 83.

[136] Haines (n 42) 184.

[137] On the general effects of *deep fakes*, see, *inter alia*, Bobby Chesney and Danielle Citron, 'Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security' (2019) 107 Calif L Rev 1753; Catherine Kerner and Mathias Risse, 'Beyond Porn and Discreditation: Epistemic Promises and Perils of Deepfake Technology in Digital Lifeworlds' (2021) 8 Moral Philosophy and Politics 81.

[138] Iain M Banks, *Excession* (The Culture 4, Ebook, Orbit Books 2008) 82.

[139] For example, there is a considerable body of literature on existential risks that might be connected to AI, in line with the idea that "An Outside Context Problem was the sort of thing most civilisations encountered just once, and which they tended to encounter rather in the same way a sentence encountered a full stop.": ibid. See, *inter alia*, Seth D Baum, 'Risk and Resilience for Unknown, Unquantifiable, Systemic, and Unlikely/Catastrophic Threats' (2015) 35 Environ Syst Decis 229; Matthijs M Maas and others, 'Reconfiguring Resilience for Existential Risk. Submission of Evidence to the Cabinet Office on the New UK National Resilience Strategy' (27 September 2021); Petit and De Cooman (n 85).

[140] Nick Bostrom, *Superintelligence: Paths, Dangers, Strategies* (Oxford University Press 2014) 114–15.

[141] See the conformity assessment procedure in Article 43 AI Act.

[142] Maja Brkan, 'The Concept of Essence of Fundamental Rights in the EU Legal Order: Peeling the Onion to Its Core' (2018) 14 European Constitutional Law Review 332.

[143] See, *inter alia*, Robert Alexy, *A Theory of Constitutional Rights* (Julian Rivers tr, Oxford University Press 2002) 47; Giovanni Sartor, 'The Logic of Proportionality: Reasoning with Non-Numerical Magnitudes' (2013) 14 German LJ 1419.

Achieving fundamental rights protection through product safety regulation entails a shift from a satisficing logic to an optimizing logic. The merits of this regulatory arrangement are not easy to visualize. On the one hand, a satisficing logic is a pragmatic approach in a context of technological uncertainty,[144] as in the case of AI systems.[145] In the absence of knowledge about what AI can do in a context, the problem of specifying red lines that any AI system cannot cross is simpler than the problem of evaluating whether an AI system provides the best realization of the principles at stake. On the other hand, the AI Act might lead to prioritizing the protection of those fundamental rights that can be more easily described in terms of satisficing, such as economic rights, to the detriment of less measurable rights, such as those connected with political participation.

One way out of this conundrum consists of imposing on product safety processes a logic of optimization when it comes to fundamental rights protection. The problem, this time, is practical. Frequent and ongoing updates of technical standards would be required under a norm of optimization of fundamental rights protection. Such updates, in turn, would deprive the product safety regime of the stability associated with standardization processes. Plus, changing a technical standard is not just a matter of revising a document. Drastic consequences might follow when existing compliant products are no longer compatible with new requirements. In particular, updating a standard is of little use if the actual systems are not updated to match the new specifications. Yet, changes to the system might not be feasible or, even if they are, might take a long time.[146] In that case, technology would lag behind the demands of the law,[147] and even the best standards would be insufficient to ensure the expected level of optimization. The protection of fundamental rights through product safety means is thus unlikely to satisfice the aims of either branch of the law, let alone optimize their fulfilment.

5. Playing the Same Old Song: The Issue of Regulatory Inheritance

The AI Act is built on solid institutional foundations. A major feature of the AI Act is its appeal to conventional product safety regulation processes and institutions. The pre-market controls adopted for high-risk AI systems—standard-setting,[148] conformity assessment,[149] certification,[150] and labelling[151]—correspond to long-established models in other sectors,[152] and European standardization bodies are already working on specifying AI-related technical standards.[153] Likewise, the post-market controls for AI systems placed on the market or put into service are modelled upon the EU surveillance

[144] Florian M Artinger and others, 'Satisficing: Integrating Two Traditions' (2022) 60 Journal of Economic Literature 598.

[145] Laurin B Weissinger, 'AI, Complexity, and Regulation' in Justin Bullock and others (eds), *The Oxford Handbook of AI Governance* (Oxford University Press 2022); Sue Anne Teo, 'How Artificial Intelligence Systems Challenge the Conceptual Foundations of the Human Rights Legal Framework' (2022) 0 Nordic Journal of Human Rights 1.

[146] On the persistence of old technologies after they cease to be in the state of the art, see David Edgerton, *The Shock of the Old: Technology and Global History since 1900* (Profile Books 2019); Mar Hicks, 'Built to Last' [2020] (11) Logic Magazine; Marianne Bellotti, *Kill It with Fire: Manage Aging Computer Systems* (No Starch Press 2021).

[147] Lyria Bennett Moses and Monika Zalnieriute, 'Law and Technology in the Dimension of Time' in Sofia Ranchordás and Yaniv Roznai (eds), *Time, Law, and Change: An Interdisciplinary Study* (Hart Publishing 2020); Marco Almada, 'Regulation by Design and the Governance of Technological Futures' [2023] Eur J Risk Reg FirstView.

[148] These standards are produced by private (or quasi-private) standardization bodies, but some of them acquire a distinctive legal character. When the European Commission confers a mandate to a standardization body for producing a harmonized standard, compliance with said standard gains a mandatory character: Marta Cantero Gamito, 'Europeanization through Standardization: ICT and Telecommunications' (2018) 37 YB Eur L 395, 400–02. In the AI Act, this mandatory character is sustained by Article 40 AI Act, which establishes that compliance with harmonized standards generates a presumption of conformity with the AI provisions covered by said standards. Divergence from the standard is theoretically possible, but, as Veale and Zuiderveen Borgesius point out, these standards might be unavoidable in practice: Michael Veale and Frederik Zuiderveen Borgesius, 'Demystifying the Draft EU Artificial Intelligence Act — Analysing the good, the bad, and the unclear elements of the proposed approach' (2021) 22 Comput L Rev Int'l 97, 105.

[149] Article 43 AI Act.

[150] Article 44 AI Act.

[151] Article 49 AI Act.

[152] See, *inter alia*, Kira JM Matus and Michael Veale, 'Certification Systems for Machine Learning: Lessons from Sustainability' (2021) 16 Regulation & Governance 177, 7.

[153] Commission Implementing Decision on a standardisation request to the European Committee for Standardisation and the European Committee for Electrotechnical Standardisation in support of Union policy on artificial intelligence [2023] C(2023)3215.

processes applicable to certain categories of products.[154] Substantial regulatory know-how is readily available to guide the application of the Act, promising swift application[155] and limited litigation.[156]

That said, the AI Act specificities will require regulatory processes and institutions to develop capabilities beyond their usual remit.[157] This section highlights some of the challenges already visible in the current versions of the Act.

5.1 Capability Challenges

Somewhat ironically, the first challenge relates to the product safety framework itself. Applying the AI Act and its open-ended provisions requires agents involved in product safety regulation processes to develop new capabilities. Four sets of agents are concerned:

- Standard Setting Organizations (SSOs) that define technical standards to further specify the high-level requirements imposed by the Act;
- Providers and users of AI systems that assess their systems for compliance with said requirements;
- Certification bodies that provide third-party assessments; and
- Market surveillance authorities (MSAs) asked to assess risks in particular contexts.

The first type of capability concerned is technical. Agents subject to the AI Act need to develop technological literacy. Put simply, they need to understand the technologies they are regulating and/or using. The challenge concerns MSAs, SSOs, and users of AI systems. Technical capabilities are concentrated in a small population of market organizations that pay salaries beyond the reach of most, if not all, regulators.[158] In fact, the concentration of AI expertise, coupled with the fact that large AI systems are resource-intensive technologies,[159] means that most users of AI technologies rely on outside expertise through mechanisms such as outsourcing or using “AI as a service” technologies.[160]

The second type of capability concerned is legal. The AI Act also requires agents involved in product safety regulation to address risks to fundamental rights protection. As shown before, assessing fundamental rights involves methods and objectives that differ from those used to evaluate product safety concerns. The challenge concerns SSOs and MSAs,[161] who must use external expertise or allocate resources to train their staff.[162] Either option will be difficult for resource-constrained public sector institutions at the European and Member State levels.[163] The problem of resources is compounded by the addition of legal instruments on digital matters that the EU is adopting in parallel

[154] The current rules governing market surveillance of products subject to EU-wide harmonization are laid down by Regulation (EU) 2019/1020, OJ 2019 L 169/1.

[155] Yet, various factors raise concerns about whether the promise of swift implementation will come to pass Elsewhere, it has been described as “dangerously optimistic”: Veale and Borgesius (n 148) 111.

[156] Thus avoiding the enforcement challenges facing other pieces of EU digital regulation, such as the GDPR: Giulia Gentile and Orla Lynskey, ‘Deficient by Design? The Transnational Enforcement of the GDPR’ (2022) 71 International & Comparative Law Quarterly 799.

[157] A comprehensive analysis of these risks is not feasible at the current stage of the AI Act legislative procedure. In fact, the EU regulators’ ambition of making the Act a “future-proof” regulation suggests that many concerns can only be detected in the long run.

[158] This is reflected, for example, in the concentration of research work on AI in established market actors, instead of state research centres or academic institutions: Abeba Birhane and others, ‘The Values Encoded in Machine Learning Research’ in *2022 ACM Conference on Fairness, Accountability, and Transparency* (FAccT ’22, New York, NY, USA, Association for Computing Machinery 21 June 2022); Nestor Maslej and others, ‘The AI Index 2023 Annual Report’ (2023).

[159] See, e.g., Kate Crawford, *The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence* (Yale University Press 2021).

[160] Jennifer Cobbe and Jatinder Singh, ‘Artificial Intelligence as a Service: Legal Responsibilities, Liabilities, and Policy Challenges’ (2021) 42 Comput L & Sec Rev.

[161] As well as the actors who are expected to comply with AI Act requirements.

[162] And face the risk of losing these professionals shortly after they are trained, as the private sector offers substantially better remuneration and resources for work.

[163] On the resource issues being faced by EU digital enforcement authorities, see, e.g., Letter from Andrea Jelinek and Wojciech Wiewiórowski, ‘Open Letter on EDPB Budget Proposal for 2023’ (9 December 2022).

to the AI Act, such as the DSA and the DMA.[164]

The AI Act anticipates capability challenges by shifting some enforcement work towards market agents. The regulatory models used for high-risk and minimal-risk systems are meta-regulatory models, in which Member States and the EU direct the governance efforts of private actors.[165] For high-risk systems, state institutions play a stronger role, but one that is mostly cast in terms of market surveillance, overseeing whether the regulated actors carry out their legal obligations in a satisfactory manner.[166] To a large extent, though, these actors are free to choose how to carry out their obligations. This freedom is particularly salient in minimal-risk systems, in which the duties of the regulated actors are defined in very lax terms.[167]

Reliance on market agents has led to concerns about the ineffectiveness of public oversight mechanisms and related issues of lack of transparency[168] or regulatory capture.[169] More to the point, perhaps, it is conceivable that outsourcing to market agents can limit the technical literacy challenge. But it is much less clear that outsourcing to market agents can ever mitigate the fundamental rights expertise challenge.

5.2 Legitimacy Challenges

Another challenge of the AI Act concerns an issue of interest to EU constitutional and administrative lawyers: legitimacy. EU law strictly conditions any delegation of regulatory power.[170] A usual argument made in legal scholarship is that widespread recourse to standardization in the EU constitutes an unlawful delegation of regulatory power.[171]

The AI Act does not escape the predicament. Instead, the text expands the scope of standardization to fundamental rights and other public interests. It is thus open to question whether standardization holds the legitimacy required to deal with the diverse regulatory concerns that motivate the AI Act. If technical standards are already seen with suspicion under EU law,[172] additional concerns are warranted in the AI Act.

Legal scholarship draws a line between three aspects of legitimacy. Input legitimacy refers to the stakeholders' participation in the production of standards that are relevant to them.[173] Throughput legitimacy refers to procedural controls—such as transparency and accountability—applying to the production of standards.[174] Finally, output legitimacy relates to the benefits arising from standardization.[175]

[164] For an overview of some of these instruments, see Gregorio and Dunn (n 12).

[165] Peter Grabosky, 'Meta-Regulation' in Peter Drahos (ed), *Regulatory Theory: Foundations and applications* (ANU Press 2017).

[166] Veale and Borgesius (n 148) s VI.

[167] See, e.g., the GPSR, which establishes a general duty of only placing safe products on the market and allows providers considerable discretion when it comes to proving safety.

[168] Busuioc and others (n 47) s 5.

[169] On the various dimensions of regulatory capture, see, *inter alia,* Barry M Mitnick, 'Capturing "Capture": Definition and Mechanisms' in David Levi-Faur (ed), *Handbook on the Politics of Regulation* (Edward Elgar Publishing 2011); Andrea Saltelli and others, 'Science, the Endless Frontier of Regulatory Capture' (2022) 135 Futures 102860; Meredith Whittaker, 'The Steep Cost of Capture' (2021) 28 interactions 50.

[170] The landmark CJEU case here is *Meroni:* Case 9-56 *Meroni & Co, Industrie Metallurgiche, SpA v High Authority of the European Coal and Steel Community* [1958] ECLI:EU:C:1958:7.

[171] For an entry point into these debates, see Mariolina Eliantonio and Megi Medzmariashvili, 'Hybridity Under Scrutiny: How European Standardization Shakes the Foundations of EU Constitutional and Internal Market Law' (2017) 44 Legal Issues of Economic Integration 323.

[172] This is not to say that traditional technical standards are somehow apolitical. See, *inter alia,* Cantero Gamito (n 148); Corinne Cath, 'The Technology We Choose to Create: Human Rights Advocacy in the Internet Engineering Task Force' (2021) 45 Telecomm Pol'y 102144; Brice Laurent, *European Objects: The Troubled Dreams of Harmonization* (The MIT Press 2022).

[173] Mariolina Eliantonio and Caroline Cauffman, 'The Legitimacy of Standardisation as a Regulatory Technique in the EU – A Cross-Disciplinary and Multi-Level Analysis: An Introduction' in Mariolina Eliantonio and Caroline Cauffman (eds), *The Legitimacy of Standardisation as a Regulatory Technique* (Edward Elgar Publishing 2020) 8.

[174] ibid 13.

[175] ibid 11.

Each of these three dimensions is questionable in the AI Act. Regarding output legitimacy, issues of method and purpose might prevent SSOs from addressing all the relevant risks stemming from AI systems. The output legitimacy of AI standards dealing with fundamental rights might thus be weaker than the legitimacy of standards perceived as strictly technical.[176]

By contrast, the issue of input legitimacy is not specific to standards in the AI Act. Instead, it builds upon a broader critique of the EU approach to product safety. Product safety regulation has been accused of depoliticizing regulatory choices. Debates on norms, values, and morality would be reduced to discussions over technical facts.[177] As Section 4 shows, a product safety framing risks suppressing relevant political dimensions of fundamental rights.

In turn, the latter issue is tightly coupled with a throughput legitimacy one. Debates posed in highly technical terms are restricted to experts. There is little room for public participation.[178] But fundamental rights are political matters par excellence, as their protection lies at the heart of liberal constitutionalism and participative democracy.[179] And attempts of product safety institutions to present themselves as apolitical bodies have not fueled confidence in their legitimacy in fundamental rights issues.[180]

The AI Act compounds issues of input and throughput legitimacy in several ways. First, the AI Act is a maximum harmonization instrument.[181] As such, Member States are not allowed to legislate on matters already covered by it or extend the existing requirements with national-level provisions meant to increase input or throughput legitimacy.[182] They can still legislate on matters not covered by the AI Act,[183] but any additional legislation must not create obstacles to the EU single market. Reliance on a standardization framework thus empowers market actors with questionable legitimacy for laying down rules on fundamental rights while pre-empting the national legislative institutions that would be expected to pursue such rules at a national level.[184]

[176] Even among standard-setters themselves: Cath (n 172).

[177] See, e.g., Mark L Flear, ‘Regulating New Technologies: EU Internal Market Law, Risk, and Socio-Technical Order’ in Marise Cremona (ed), *New Technologies and EU Law*, vol 1 (Oxford University Press 2017) 75. More specifically about standardization as an inherently political practice, see Niels ten Oever and Stefania Milan, ‘The Making of International Communication Standards: Towards a Theory of Power in Standardization’ (2022) 1 Journal of Standardisation.

[178] Some standardization organizations take pride in being open to the public: the IETF (Internet Engineering Task Force) opens participation to “individuals willing to contribute technical expertise to help make the Internet work better.”: ‘Participate in the IETF’ (*IETF*, no date) <https://www.ietf.org/about/participate/> accessed 28 July 2022. But, for the most part, debates on standardization remain closed to small groups of experts, all too often male and hailing from a few countries and businesses: JoAnne Yates and Craig Murphy, *Engineering Rules: Global Standard Setting since 1880* (Johns Hopkins University Press 2019).

[179] On the specific context of fundamental rights in the EU legal order, see Koen Lenaerts and José Antonio Gutiérrez-Fons, ‘The Place of the Charter in the European Legal Space’ in Steve Peers and others (eds), *The EU Charter of Fundamental Rights: A Commentary* (Hart Publishing 2021).

[180] Nathalie Smuha and others, ‘A Response to the European Commission’s Proposal for an Artificial Intelligent Act’ (8 May 2021) 10. On the challenges of ensuring accountability and legal oversight over expertise-driven organizations, see, *inter alia*, Deirdre Curtin, ‘“Accountable Independence” of the European Central Bank: Seeing the Logics of Transparency’ (2017) 23 European Law Journal 28; Olia Kanevskaia, ‘Governance of ICT Standardization: Due Process in Technocratic Decision-Making’ (2019) 45 NC J Int’l L 549.

[181] Veale and Borgesius (n 148) 109.

[182] ibid 108–10.

[183] It has been argued that the AI Act precludes further AI regulation by Member States, because its Article 1 defines its scope as covering all AI systems, including those systems covered by the baseline we mention above: ibid 108–10. However, Protocol 25 to the Treaties specifies that the exercise of a competence by the EU is not enough to pre-empt the Member States from legislating in the entire area covered by the competence; instead, pre-emption is limited to the elements ‘governed by the Union act in question’. Furthermore, the AI Act explicitly relies on general product safety law as a fall-back regulatory regime for baseline systems (Recital 82), and that regime leaves room for Member State requirements in the absence of an EU-wide standard: see, e.g., Article 7(1)(b) GPSR. Hence, the AI Act does not pre-empt in and of itself national requirements for baseline systems, provided that such norms are compatible with the existing EU framework for that particular category of product, including the relevant harmonized technical standards.

[184] The EU, however, would retain its ability to create further legislation on AI matters.

To mitigate the legitimacy challenges to SSOs, the Commission has restricted the AI Act's standardization mandates to CEN and CENELEC,[185] two SSOs governed primarily by national standardization bodies from EU member states. As such, the production of harmonized standards for AI does not involve the ETSI,[186] in which private organizations and non-EU actors are seen to have a stronger voice.[187] This decision builds input and throughput legitimacy in two ways. First, it mitigates the risk of conflicts of interest involved in having industry organizations determine the content of the standards that they will need to follow.[188] Second, it addresses concerns with EU digital sovereignty[189] by ensuring that non-EU organizations are not involved in elaborating standards that affect sensitive matters such as fundamental rights, democracy, and the rule of law. Such results, however, come at the expense of access to a reduced pool of expertise, thus potentially constraining output legitimacy.

Another threat to the AI Act's legitimacy is the dearth of mechanisms for natural or legal persons to file complaints against risks stemming from AI systems.[190] These mechanisms were completely absent from the Commission proposal, meaning that natural and legal persons would need to rely on courts to settle disputes over the provisions of the AI Act that are directly applicable to them. The Parliament amendments seek to mitigate this issue by providing three remedies to people affected by the use of high-risk AI systems: the right to lodge a complaint with national supervisory authorities,[191] the right to an effective judicial remedy against said authorities,[192] and the right to an explanation of individual decision-making.[193] Such measures would potentially boost the AI Act's legitimacy by ensuring individuals have a voice regarding the use of AI (input) and avoiding or redressing AI-related harms (output).

Finally, the general problem of throughput legitimacy discussed above is compounded by compromises to apply a product safety framework to AI systems not covered by product safety law. These tradeoffs are particularly salient in Article 70 of the Act, which imposes various confidentiality duties on MSAs tasked with enforcing the AI Act. In particular, MSAs are subject to specific duties of secrecy when exercising their powers concerning high-risk systems deployed in law enforcement or migration, asylum, and border control management applications.[194] The security interest protected here is easy enough to understand.[195] The resulting framework, however, further mangles the legitimacy of MSAs by limiting their transparency and restricting scrutiny of the AI systems concerned to a few MSA staff members who enjoy the necessary clearances.[196]

[185] Commission Implementing Decision on a standardisation request to the European Committee for Standardisation and the European Committee for Electrotechnical Standardisation in support of Union policy on artificial intelligence.

[186] A telecommunications standardization organization that, with CEN and CENELEC, comprises the European Standardization Organizations.

[187] Cantero Gamito (n 148) 414.

[188] See, e.g., Justus Baron and Olia Kanevskaia, 'Wearing Multiple Hats—The Role of Working Group Chairs' Affiliation in Standards Development' (2023) 52 Research Policy 104822.

[189] On the debates surrounding sovereignty, see, *inter alia,* Tomáš Gábriš and Ondrej Hamuľák, 'Digital Sovereignty or Sovereignty with Digital Elements?' in David Ramiro Troitiño and others (eds), *Digital Development of the European Union: An Interdisciplinary Perspective* (Springer International Publishing 2023); Nathalie A Smuha, 'Digital Sovereignty in the European Union: Five Challenges from a Normative Perspective' (SSRN Scholarly Paper, 7 January 2023); Andrea Calderaro and Stella Blumfelde, 'Artificial Intelligence and EU Security: The False Promise of Digital Sovereignty?' (2022) 31 European Security 415.

[190] The AI Act as proposed by the Commission does not include any such mechanism, but some product safety instruments include them: Veale and Borgesius (n 148) 111. Both the Council and the Parliament have proposed allowing individual complaints to market surveillance authorities, so it is likely—though by no means certain—that such a mechanism will appear in the final version of the Act. If that is the case, input legitimacy might be slightly increased.

[191] Article 68a AI Act (Parliament amendments).

[192] Article 68b AI Act (Parliament amendments).

[193] Article 68c AI Act (Parliament amendments).

[194] Article 70(2) AI Act.

[195] Such as the principle of originator control, which empowers security and law enforcement agencies to govern the circulation of information they put into interoperable databases: Deirdre Curtin, 'Second Order Secrecy and Europe's Legality Mosaics' (2018) 41 West European Politics 846, 853.

[196] On this point, see sections 4.C and 5 of Busuioc and others (n 47).

6. Coda: how to avoid a regulatory cacophony?

The AI Act raises classic issues of law and technology. The abovementioned pacing problem, the mismatch between means and ends, and the problems of regulatory path dependence towards product safety are well-known themes in the law and technology literature. From the standpoint of the literature on law and technology, the AI Act provides new clothing for old problems rather than entirely new challenges.

As ever, the hardest task consists in finding solutions. This paper submits that SSOs and MSAs should not underestimate the challenges of creating new capabilities. And there should be an awareness that even the best standards will still leave considerable space for interpretation and necessitate context-sensitive evaluations.

In addition to this, some structural recommendations for regulation are possible. While product safety requirements and standardization can play an important role in AI governance, they should not be the only game in town. Instead, their shortcomings should be addressed by ad hoc instruments, dealing with the risks to fundamental rights created by specific applications. Alternatively, the AI Act can be construed as establishing two sets of requirements: one concerning risks to health and safety, framed in product safety terms; and a parallel system dealing with risks to fundamental rights and other general values, which frames risk—and the required measures—in terms that address the shortcomings identified in Section 4 above.[197] Of course, either approach leads to increased regulatory fragmentation compared to a centralized, horizontal instrument. But it is the price to ensure suitable protection of fundamental rights.

Alternatively, the fundamental rights protection goals of the AI Act might be better served by narrowing the Act down and focusing its provisions on the technical aspects of AI systems. The protection of fundamental rights could then be supplied by international treaties, such as the proposed Council of Europe treaty on AI and human rights, democracy, and the rule of law.[198] In this variant, regulatory fragmentation can ensure that fundamental rights are protected in their full depth and breadth instead of being reduced to a simplification sketched in the language of product safety. Otherwise, one might end up sacrificing both fundamental rights and product safety concerns on the altar of legal uniformity.

[197] The authors thank Mireille Hildebrandt for raising this alternative to a product-centric approach.

[198] On the interplay between the AI Act and the proposed treaty, see Marco Almada and Anca Radu, ‘The Brussels Side-Effect: How the AI Act Can Reduce the Global Reach of EU Policy’ [forthcoming] German LJ.

Authors

Marco Almada

Department of Law, European University Institute

Marco.Almada@eui.eu

Nicolas Petit

Department of Law and Robert Schuman Centre for Advanced Studies, European University Institute
Invited Professor, College of Europe, Bruges

Nicolas.Petit@eui.eu