Truly Risk-Based Regulation of Artificial Intelligence

How to Implement the EU's AI Act

Martin Ebers[*]

19.06.2024

Contents

[*] Martin Ebers is President of the Robotics & AI Law Society (RAILS) and Professor of IT Law at the University of Tartu, Estonia. The author gratefully acknowledges the support of the Computer & Communications Industry Association (CCIA Europe) for this research paper, as well as the valuable help and insights provided by the Robotics & AI Law Society (RAILS) and the Munich Institute of Robotics and Machine Intelligence (MIRMI) at the Technical University of Munich (TUM). The views expressed in this paper reflect the author's own expertise and are based on independent academic research. All internet sources referred to were last accessed on June 19, 2024.

Executive Summary

The recently adopted Artificial Intelligence Act (AI Act) of the European Union (EU) claims to be based on a risk-based approach to avoid over-regulation and to respect the principle of legislative proportionality. This paper argues that risk-based regulation is indeed the right approach to AI regulation. At the same time, however, the paper shows that important provisions of the AI Act *do not follow a truly risk-based approach* – contrary to the claims of the European Commission and the co-legislators. Yet, this is nothing that cannot be fixed. The AI Act provides for sufficient tools to support future-proof legislation and to implement it in line with a genuine risk-based approach. Against this background, the paper analyses (i) how the AI Act should be applied and implemented according to its original intention of a risk-based approach, (ii) how the AI Act should be complemented by sector-specific legislation in the future to avoid inconsistencies and over-regulation, and (iii) what lessons legislators around the world can learn from the AI Act in regulating AI.

The following sections are structured as follows:

- *Section 1* shows how risk-based regulation has become the dominant strategy for policymakers to regulate AI – not only in the EU, but globally.
- *Section 2* outlines the key elements of risk-based regulation - discussing the notion of "risk", the distinction between AI risk assessment, impact assessment, and risk management, and the key elements of risk-based regulation.
- *Section 3* criticizes the AI Act, arguing that some of its main provisions are not truly risk-based, leading to over-regulation in some areas and under-regulation in others. In particular, it analyses several problems with the AI Act, such as the lack of a risk-*benefit* analysis, limited reliance on empirical evidence, and lack of case-by-case risk classification.
- *Section 4* examines how the AI Act can be brought into line with a truly risk-based approach. To this end, the paper analyses the relevant instruments to implement the AI Act, such as guidelines, delegated and implementing acts, codes of practice, and harmonized standards.
- *Section 5* analyses how the AI Act should be complemented by sector-specific legislation in the future to avoid inconsistencies and over-regulation.
- *Section 6* draws conclusions on what policymakers outside the EU can learn from the AI Act when regulating AI.

1. Risk-Based Regulation as the Dominant Global Strategy for Regulating AI

1.1. The Risk-Based Approach in the EU's AI Act

The recently adopted Artificial Intelligence Act (AI Act) of the European Union (EU)[1] is the world's first attempt at comprehensively regulating AI. As is well known, the AI Act is claimed to follow a risk-based approach - one that tailors the choice and design of regulatory instruments based on the level of risk, according to the rule: "the higher the risk, the stricter the rules". To this end, the AI Act distinguishes

[1] While the AI Act had not yet been published in the Official Journal of the EU, at the time of writing this paper, the author has taken into account the final version which was approved by the European Parliament (March 13, 2024), by the Council (May 21, 2024) and finally signed into law (June 13, 2024), PE-CONS 24/1/24 REV 1, available at: https://data.consilium.europa.eu/doc/document/PE-24-2024-REV-1/en/pdf.

four risk categories (unacceptable, high, limited, and minimal), defining regulatory requirements based on the risks posed by AI systems.[2]

By adopting this approach, the EU seeks to safeguard fundamental values and rights without unduly hampering the benefits that AI can bring to society.

In this regard, Recital (26) AI Act points out that:

> "In order to introduce a **proportionate** and **effective** set of binding rules for AI systems, a clearly defined **risk-based approach** should be followed. That approach should **tailor the type and content** of such rules to the **intensity** and **scope of the risks** that AI systems can generate. It is therefore necessary to prohibit certain unacceptable AI practices, to lay down requirements for high-risk AI systems and obligations for the relevant operators, and to lay down transparency obligations for certain AI systems."[3]

Hence, the underlying objective of the AI Act's risk-based approach is to strike an optimal (and proportional) balance between innovation and the benefits of AI systems on the one hand, and the protection of fundamental values such as safety, health, and fundamental rights on the other.

Recital (26) of the AI Act refers, particularly, to the principle of (legislative) proportionality enshrined in Art. 5(4) TEU. According to this article, the "content and form of Union action shall not exceed what is necessary to achieve the objectives of the Treaties". Further, the article requires institutions of the Union to "apply the principle of proportionality as laid down in the Protocol on the application of the principles of subsidiarity and proportionality". This protocol, in turn, states that draft legislative acts shall be justified with regard to the principles of subsidiarity and proportionality, taking into account the burden - whether financial or administrative, falling upon the EU, national governments or authorities, economic operators and citizens, "to be minimised and commensurate with the objective to be achieved".[4]

Risk-based regulation can be seen as a legislative technique for promoting a proportionate system of duties and obligations.[5] To this end, as some scholars have pointed out, "risk-based regulation uses risk as a tool to prioritize and target enforcement action in a manner that is proportionate to an actual hazard: in other words, it tends to "calibrate" the enforcement of the law based on concrete risk scores".[6]

As such, the risk-based approach is not new to EU (digital) law. Since the introduction of the Digital Single Market Strategy,[7] the EU has increasingly relied on a risk-based approach towards regulating the digital economy, particularly in the areas of data, online content and platforms, cybersecurity, (digital) products and services, and AI – albeit using different risk-based approaches.[8]

[2] In addition, during the negotiation of the AI Act, the co-legislators added the category of "general purpose AI models". However, as will be shown below (Section 3.4), this new category is inconsistent with a truly risk-based approach.

[3] Emphases added.

[4] Consolidated version of the Treaty on the Functioning of the European Union – Protocol (No 2) on the application of the principles of subsidiarity and proportionality, Art. 5, OJ C 115, 9.5.2008, p. 1.

[5] G De Gregorio and P Dunn, 'The European Risk-based Approaches: Connecting Constitutional Dots in the Digital Age' (2022) 59 Common Market Law Review 473, 499.

[6] *Ibid*, 475. See also Claudia Quelle, 'Enhancing Compliance under the General Data Protection Regulation: The Risky Upshot of the Accountability- and Risk-based Approach' (2018) EJRR 502, 509 et seq.

[7] European Commission, A Digital Single Market Strategy for Europe, COM(2015)192 final.

[8] For a comparison between the different risk-based approaches in the General Data Protection Regulation (GDPR), the Digital Services Act (DSA) and the AI Act cf. De Gregorio and Dunn, (n 5), 475.

1.2. Risk-Based Approaches Outside the EU

The EU's AI Act is not the only piece of legislation that follows a risk-based approach for regulating AI. In fact, in many parts of the world, the risk-based approach has become the dominant strategy for regulating AI systems - both at the international and national levels, and in the work of (international) standard-setting bodies.

- The Blechley Declaration of 1-2 November 2023[9] - signed by 28 countries including the United States, China, and the European Union, at the UK AI Safety Summit, states that countries should take into account the risks associated with AI, and consider, where appropriate, "classifications and categorisations of risk based on national circumstances and applicable legal frameworks."

- The agreement by G7 leaders on International Guiding Principles on Artificial Intelligence (AI) and a voluntary Code of Conduct for AI developers under the Hiroshima AI process,[10] calls to develop, implement and disclose AI governance and risk management policies, in line with a risk-based approach.

- The Council of Europe's "Framework Convention on Artificial Intelligence"[11] combines general principles with a risk-based approach,[12] requiring measures to be taken for "the identification, assessment, prevention, and mitigation of risks and impacts to human rights, democracy, and the rule of law arising from the design, development, use, and decommissioning of artificial intelligence systems" within the scope of the proposed Convention.

- Canada's Directive on Automated Decision-Making[13] requires public bodies to conduct an Algorithmic Impact Assessment - aiming at assessing and reducing risks associated with automated decision systems, by ensuring that "Automated Decision Systems are deployed in a manner that reduces risks to Canadians and federal institutions,." This goal is extended to "any system, tool, or statistical models used to recommend or make an administrative decision about a client." In June 2022, Canada took a step toward updating its legislation by introducing Bill C-27 for a Digital Charter Implementation Act which will enact, in its part 3, the "Artificial

[9] 'The Bletchley Declaration by Countries Attending the AI Safety Summit, 1-2 November 2023' (1 November 2023) <https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023>.

[10] 'Hiroshima Process International Code of Conduct for Advanced AI Systems' (European Commission, 30 October, 2023) <https://digital-strategy.ec.europa.eu/en/library/hiroshima-process-international-code-conduct-advanced-ai-systems>.

[11] Council of Europe, Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law' (No. CM(2024)52-final, 17 May 2024,) <https://rm.coe.int/cai-2023-28-draft-framework-convention/1680ade043>. Cf. also Peggy Valcke and Fien Hendrickx, 'The Council of Europe's road towards an AI Convention: taking stock' (KU Leuven, 9 February 2023) <https://www.law.kuleuven.be/citip/blog/the-council-of-europes-road-towards-an-ai-convention-taking-stock/>.

[12] The EU Council authorized the European Commission to negotiate on behalf of the EU to ensure consistency between the AI Act and the Convention, highlighting that the EU should push the EU should push the Council of Europe towards a risk-based approach that is fully compatible with the AI Act; Recommendation for a Council Decision authorizing the opening of negotiations on behalf of the European Union for a Council of Europe convention on artificial intelligence, human rights, democracy and the rule of law, COM(2022) 414 final (2022), point 11 and 12.

[13] Government of Canada, 'Directive on Automated Decision-Making' (2021) <https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32592>. For more details, see Teresa Scassa, 'Administrative Law and the Governance of Automated Decision-Making: A Critical Look at Canada's Directive on Automated Decision-Making' (2020) <https://ssrn.com/abstract=3722192>.

Intelligence and Data Act" (AIDA).[14] This Act prescribes requirements applicable to "regulated activities" in general, with more restrictive requirements targeting "high-impact" AI systems.

- In the US, President Biden's Executive Order on AI[15] evaluates the risks associated with dual-use foundation models with widely available model weights,[16] and considers potential mechanisms to manage risks and maximize benefits. Also, it mandates agencies to evaluate potential risks related to the use of AI in critical infrastructure sectors,[17] issue public reports on best practices for managing AI-specific cybersecurity risks, and incorporate AI Risk Management Framework into relevant safety and security guidelines for critical infrastructure. Additionally, the order requires actions to understand and mitigate the risk of AI being misused to assist in the development or use of CBRN threats, including evaluating the potential for AI misuse and making recommendations for regulating or overseeing AI models.[18]

- Brazil is currently examining a comprehensive AI bill[19] to establish *a rights-based and risk-based regulatory framework* tailoring the regulatory obligations based on potential AI technology risks.[20]

- The OECD's Framework for the Classification of AI Systems[21] also adopts a risk-based approach in providing guidelines for assessing risks associated with AI systems. It consists of such components as the context for deployment, data governance, algorithm type and characteristics, and performance and outputs. The framework aims to create an environment where legal obligations are tailored to specific risks, ensuring an optimal balance between interests and due diligence. It provides for processes that help refine risk classification criteria based on real-world evidence – such as the development of a risk assessment framework.

- Standards bodies have also developed various risk-based approaches to AI regulation. An example is the Artificial Intelligence Risk Management Framework (AI RMF 1.0)[22] developed by the US National Institute of Standards and Technology (NIST), which provides voluntary guidance for policymakers and companies in organizing their internal AI governance in a risk-based manner. In the same vein, ISO and IEC have developed standards for managing risks,[23]

[14] Canada, Bill C-27, An Act to enact the Consumer Privacy Protection Act, the Personal Information and Data Protection Tribunal Act and the Artificial Intelligence and Data Act and to make consequential and related amendments to other Acts, 1st session, 44th Parliament, 2022, available at: <https://www.parl.ca/legisinfo/en/bill/44-1/c-27>.

[15] Executive Order (No. 14110) on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, The White House (30 October 2023) <https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/> accessed 17 April 2024.

[16] Article 4.6a(i) Executive Order No. 14110.

[17] Article 4.3 Executive Order No. 14110.

[18] Article 4.4 Executive Order No. 14110.

[19] Projeto de Lei n° 2338, de 2023 (Senado Federal, 18 April 2024) <https://www25.senado.leg.br/web/atividade/materias/-/materia/157233>.

[20] Cf. thereto Laura Schertel Mendez and Beatriz Kira, 'The road to regulation of artificial intelligence: the Brazilian experience' (2023) Internet Policy Review <https://policyreview.info/articles/news/road-regulation-artificial-intelligence-brazilian-experience/1737>.

[21] OECD, OECD Framework for the Classification of AI Systems, (OECD Digital Economy Papers, No. 323, February 2022) <https://read.oecd.org/10.1787/cb6d9eca-en?format=pdf>.

[22] NIST (2023, March 30). AI Risk Management Framework. NIST.

[23] ISO/IEC 23894 — Information Technology — Artificial Intelligence — Risk Management, Stage: 30.60, Committee Draft (CD); ISO/IEC AWI 42001 Information Technology — Artificial intelligence — Management system, Stage: 20.00, Preparatory.

and for governance bodies to ensure effective, efficient, and acceptable use of AI within organizations.[24]

1.3. Evaluation

The foregoing overview illustrates that the risk-based approach has become a dominant strategy for regulating and governing AI – not only in the EU, but globally. However, this regulatory technique has influenced legislation in a manner that is far from unitary. In fact, a closer look reveals that there are significant differences in the understanding of what exactly constitutes a risk-based approach.

Accordingly, the next section discusses the key features of risk-based regulation, in particular the notion of "risk", the different types of risk-based regulation and their essential features.

2. Key Elements of Risk-based Regulation

2.1. The Notion of Risk

The global emergence of risk-based approaches to regulating AI systems raises, first and foremost, raises the question as to what policymakers (and the EU in particular) mean when they talk about risk.

Generally speaking, "risk" is the likelihood that a source of hazard will turn into actual harm.[25] Based on this understanding, Art. 3 No. 2 AI Act defines risk as "the combination of the probability of an occurrence of harm and the severity of that harm".

As such, risks are usually distinguished from uncertainties.[26] While risks are "known knowns" with statistical probabilities and quantifiable effects, uncertainties are "known unknowns" that cannot be quantified because we do not know what the effects of a particular technology might be. In addition, there are "unknown unknowns", where we are not even aware that things or activities may have adverse effects at all.[27]

One of the most pressing issues for any risk-based approach is how to reach consensus on which risks to select and how serious they are considered to be (either in terms of probability or impact, or both).[28] There was an active and lively debate during the negotiations on the AI Act about the criteria for determining when AI poses unacceptable risks to society and individuals – to determine when it is banned in the EU, as well as which AI systems should be classified as "high-risk" and thus allowed on the market if certain safeguards are put in place. This illustrates how difficult this endeavour can be; especially if this assessment is not sufficiently based on empirical evidence and a sound methodology.[29]

[24] ISO/IEC DIS 38507 — Information technology — Governance of IT — Governance implications of the use of artificial intelligence by organizations, Stage: 40.20, Enquiry.

[25] Frank Knight, Risk, Uncertainty, and Profit 19–20, 233 (1921); Baruch Fischhoff, Sarah Watson and Chris Hope, Defining risk (1984) **17** *Policy Science* 123–139 <https://doi.org/10.1007/BF00146924>.

[26] Frank Knight, Risk, Uncertainty, and Profit (1921) 19–20, 233.

[27] Julia Black, 'The role of risk in regulatory processes' in Robert Baldwin, Martin Cave and Martin Lodge (eds), *The Oxford Handbook of Regulation* (Oxford University Press 2010) 302-348, 310.

[28] Black (n 27), 311.

[29] On the limited reliance on empirical evidence cf. below, section 3.2.

2.2. The Elements of Risk-based Regulation: Risk Assessment and Categorization, Impact Assessment and Risk Management

Typically, risk-based approaches to AI regulation consist of various elements or phases, namely - risk assessment and categorization, impact assessment and risk management.[30]

At a high level, we can first distinguish between (i) the assessment of risks arising from the use of AI and (ii) the classification of AI systems or applications by risks:[31] The first type assesses the risks posed by the use of AI, which may include risks to safety and health, bias and discrimination, lack of fairness, lack of transparency, invasion of privacy and data protection rights, or other protected interests. In the second type, the assessor looks at the risks associated with the use of AI in order to classify the system or application into a category of risk. This classification process helps stakeholders prioritize their efforts and allocate resources more effectively, contributing to the determination of obligations in line with the extent of risk posed by AI systems.

An impact assessment, on the other hand, goes further than a risk assessment. While a risk assessment is about the identification, analysis and evaluation of AI-related risks, an impact assessment seeks to evaluate the wider impact of AI systems on several stakeholders, including users, society, and the environment, going beyond the mere discovery and analysis of risks. This usually entails a review of governance, performance, communication, threats to safety and security, and other protected interests.

Based on such an impact assessment, risk-based approaches to AI regulation usually also contain requirements for a risk management, which entails the determination, evaluation, and ranking of risks related to AI as well as putting policies in place to reduce, track, and manage the possibility of unforeseen events.

In the AI Act, all of the above-mentioned elements of risk-based regulation are present.[32] The AI Act *assesses* the risks posed by the use of AI and *categorizes* them (mostly on a top-down basis[33]) as unacceptable, high risk, limited risk, and minimal (or no) risk. Moreover, Art. 9 AI Act requires providers of high-risk AI systems to carry out an *impact assessment*, which includes the identification and analysis of foreseeable risks that the high-risk AI system may pose to health, safety, or fundamental rights, and– on this basis – to establish a *risk management system* throughout the entire life cycle in order to take appropriate and targeted risk management measures designed to address the identified risks.

[30] Claudio Novelli, Federico Casolari, Antonino Rotolo *et al.* AI Risk Assessment: A Scenario-Based, Proportional Methodology for the AI Act (2024) 3 *DISO* 13. https://doi.org/10.1007/s44206-024-00095-1; EY, Trilateral Research, "A survey of artificial intelligence risk assessment methodologies: The global state of play and leading practices identified." (2022) https://www.trilateralresearch.com/wp-content/uploads/2022/01/A-survey-of-AI-Risk-Assessment-Methodologies-full-report.pdf.

[31] Cf. EY, Trilateral Research, "A survey of artificial intelligence risk assessment methodologies: The global state of play and leading practices identified." (2022) https://www.trilateralresearch.com/wp-content/uploads/2022/01/A-survey-of-AI-Risk-Assessment-Methodologies-full-report.pdf p. 9 et seq.

[32] Cf. also Tobias Mahler, 'Between Risk Management and Proportionality: The Risk-Based Approach in the EU's Artificial Intelligence Act Proposal' in Luca Colonna and Rolf Greenstein (eds), Nordic Yearbook of Law and Informatics 2020-2021: Law in the Era of Artificial Intelligence (The Swedish Law and Informatics Research Institute 2022) 249 et seq. https://irilaw.files.wordpress.com/2022/02/law-in-the-era-of-artificial-intelligence.pdf.

[33] In the AI Act, risk categorization is mostly top-down, as it is the AI Act itself (and to a certain extent the European Commission, cf. Art. 7 AI Act) that decides into which risk category a particular AI system falls. However, with the so-called "additional layer" provided for in Art. 6(3)-(4) AI Act, providers of systems listed in Annex III have the possibility to demonstrate (and document) that their AI system is not high-risk.

2.3. Essential Features of Risk-based Regulation

Risk-based approaches to AI regulation can take different forms, depending on the ultimate goal of the regulator. As *Coglianese*[34] points out, such a regulation can aim to:

- Eliminate all risk (the zero-risk approach)
- Reduce risk to an acceptable level (the acceptable risk approach)
- Reduce risk until costs become unbearable (the feasibility approach) or
- Strike a balance between risk reduction and costs of regulation (the proportionate or efficiency approach).

As discussed above,[35] the AI Act is based on the latter idea of proportionate regulation - it is aimed at striking an optimal (or proportionate) balance between reducing the risks posed by the use of AI systems on the one hand, and innovation and the benefits of AI systems (or the costs of regulation) on the other.

Accordingly, the question arises as to what key elements a legislator must consider when attempting to adopt such an approach. Arguably, these elements include:

- **Risk-*benefit* analysis:** When assessing the risks of AI, it is necessary to look, beyond the possible harms that AI systems can cause, at their innovative economic and social benefits. After all, “risk” is something we take in the name of benefit; we don’t typically choose to be harmed. Instead, we - as a society, choose to take certain risks in the name of current and potential societal gains.[36] Therefore, in order to assess which risks are acceptable and which risks present the possibility of unacceptable harm, a consistent application of the risk-based approach requires thorough consideration of, not only the negative consequences, but also the positive contributions that AI brings to individuals and society. Such a risk-benefit analysis must include in particular the (opportunity) costs of underuse. As the European Parliament already pointed out in 2020, the underuse of AI can also be considered a “major threat”: missed opportunities for the EU to use AI systems “could mean poor implementation of major programmes, such as the EU Green Deal, losing competitive advantage towards other parts of the world, economic stagnation and poorer possibilities for people.”[37]
- **Technology Neutrality**: A true risk-based approach regulates the risks of applications, not the technology itself. This principle of “technology neutrality” has been recognized by many regulators around the world,[38] including the EU,[39] as an overarching principle for ICT regulation. The main aim of this principle is to ensure equal treatment of technologies with equivalent effects, and to make the law future-proof, i.e., to draft legislation in a way that is flexible enough

[34] Cary Coglianese, 'The Law and Economics of Risk Regulation' (2020) University of Pennsylvania, Institute for Law & Economics Research Paper No. 20-18, 9 https://scholarship.law.upenn.edu/faculty_scholarship/2157/.

[35] See section 1.1.

[36] Margot Kaminski, ‘Regulating the Risks of AI’ (2023) 103 Boston University Law Review 1347, https://ssrn.com/abstract=4195066.

[37] European Parliament, 'Artificial intelligence: threats and opportunities' (23 September 2020, last updated 20 June 2023) https://www.europarl.europa.eu/topics/en/article/20200918STO87404/artificial-intelligence-threats-and-opportunities. Regarding the opportunity costs of the underuse of AI in the health sector, see Ugo Pagallo and others, 'The underuse of AI in the health sector: Opportunity costs, success stories, risks and recommendations' (2024) 14 Health and Technology 1 https://doi.org/10.1007/s12553-023-00806-7.

[38] As to the origins of this principle, see Annika Veerpalu, 'Regulatory challenges to the use of distributed ledger technology: Analysis of the compliance of existing regulation with the principles of technology neutrality and functional equivalence' (PhD thesis, University of Tartu 2021) 30 https://dspace.ut.ee/bitstreams/12ad2896-93f2-4d23-ac81-c28d50c9f25e/download.

[39] See, for example, recital (15) GDPR; recital (10) Digital Content and Services Directive 2019/770.

not to impede future technological development and to avoid the need for constant legislative revision.[40]

- **Evidence-based Risk Assessment and Categorization**: Another important feature of risk-based regulation is that the assessment and classification of risks requires sufficient empirical evidence and a clear methodology. As AI-related risks are being used to justify governmental regulation, there needs to be a common way how to assess and classify these risks. Accordingly, regulators, standard bodies and other stakeholders have been working for many years on risk assessment frameworks and science-based methodologies. Arguably, risk-based regulation works best on quantifiable problems. However, many harms are either not quantifiable at all, or represent a mixture of quantifiable issues with hidden policy choices.[41] In such cases, the question arises as to whether a risk-based approach is appropriate at all or whether other regulatory techniques (such as a rights-based approach) should be adopted instead.[42]
- **Proportionate Regulatory Burden**: Ideally, obligations and other regulatory burdens should be proportionate to the risks posed by AI applications to ensure that regulatory requirements are aligned with the potential harm and impact of AI systems. Risk-based regulation, therefore, seeks to create a legal framework in which legal obligations are tailored to the specific risks posed by the use of a particular AI system for a given purpose, in order to avoid overburdening of the regulated actors.
- **Flexibility and adaptability**: Risk-based regulation must also be flexible enough to adjust retrospectively if it turns out that the original risk assessment or categorization was wrong. As Black puts it, "[r]esponding to risks and attempting to manage them necessarily involves anticipating the future" which by its very nature is unknown.[43] Especially with new technologies such as AI, it is impossible to reliably assess the risks and benefits of AI systems deployed in given areas ex ante. For this reason, a truly risk-based regulation must require the legislator, the regulator and those who manage and mitigate risks to monitor the performance of AI systems throughout their lifetime, periodically re-evaluating risks and implementing the necessary corrections.

Clearly, some caveats are necessary. First, it is important to note that the criteria listed above are by no means exhaustive. Also, it is essential to remember that risk-based approaches to regulation have well-known difficulties – such as how to specify, aggregate and quantify risks, how to reconcile conflicting values and how to set levels of acceptable risk – as well as limitations – such as an overreliance on quantification, how to deal with unquantifiable and, in particular, unknown risks, and how to take into account the risk of harm to individuals.[44] As a result, risk-based regulation usually needs to be complemented by additional set of rules.

These legitimate concerns should not, however, be used as an overall argument against this type of regulation. In fact, risk-based regulation has a number of strengths when properly implemented. First, it rationalizes government intervention by setting clear priorities and objectives. Second, it facilitates the effective use of scarce resources and allows regulators – if implemented according to a true risk-based approach – to focus compliance efforts on products/services and/or systems that pose the

[40] Bert-Jaap Koops, 'Should ICT Regulation be Technology-Neutral' in Bert-Jaap Koops and others (eds), *Starting Points for ICT Regulation: deconstructing prevalent policy one-liners* (2006); Chris Reed, 'Taking Sides on Technology Neutrality' (2007) 4(3) SCRIPTed 263; Brad Greenberg, 'Rethinking Technology Neutrality' (2016) 100 *Minnesota Law Review* 1495 https://scholarship.law.umn.edu/cgi/viewcontent.cgi?article=1206&context=mlr.

[41] Kaminski (n 36), 32.

[42] This issue is particular relevant in the context of fundamental rights, see below, at 3.2.

[43] Black (n 27), 317.

[44] Kaminski (n 36), 32.

greatest risks. Last, but not least, risk-based regulation can be an effective tool for striking the right balance between the benefits and risks of a particular technology.

3. Is the AI Act a Truly Risk-based Regulation?

While risk-based regulation is indeed the right approach to AI regulation, important elements of the AI Act do not follow a truly risk-based approach, especially:

- the choice to protect not only health and safety, but also fundamental rights (3.1.),
- the absence of a risk-*benefit* analysis (3.2.),
- limited reliance on empirical evidence (3.3.),
- abstract risk-categories (3.4.),
- the regulation of GPAI models (3.5.),
- the overly wide AI definition (3.6.),
- double regulatory burdens due to the horizontal approach (3.7.), and
- overlap of enforcement tools (3.8.).

3.1. Protecting Fundamental Rights with a Risk-based Approach?

One fundamental problem is that the AI Act, with its risk-based approach, seeks to protect not only health and safety, but also the fundamental rights of citizens. This is troubling for a number of reasons.

First, the European Union has no general competence to harmonize Member State's laws to protect human rights. As a result, the AI Act "shoehorns"[45] the protection of fundamental rights into the scope of Article 114 TFEU, which gives the EU the competence to remove barriers to trade in the internal market. However, such an approach is very likely to fail, because it does not have the protection of rights as its primary goal, but rather the opening and shaping of markets.[46]

Moreover, the EU's decision to protect human rights in the AI Act primarily through a risk-based approach, rather than of a rights-based approach,[47] is generally ill-suited. Most importantly, such an approach neglects the minimum and non-negotiable nature of human rights. Instead, the AI Act, with its risk-based approach and its fundamental rights impact assessment, implies that fundamental rights violations can be quantified and measured in degrees. This is, however, not the case. As *Yeung* & *Bygrave* point out, while it is possible to speak of different levels of culpability, scale, and magnitude when talking about fundamental rights, "these variations do not imply that fundamental rights

[45] Marco Almada and Anca Radu, 'The Brussels Side-Effect: How the AI Act Can Reduce the Global Reach of EU Policy' (2024) *German Law Journal* 1-18, 3, https://doi.org/10.1017/glj.2023.108.

[46] Hans-W Micklitz and Dennis Patterson, 'From the Nation State to the Market: The Evolution of EU Private Law' (1 June 2012) EUI Working Papers LAW No 2012/15, https://ssrn.com/abstract=2115463; Almada and Radu (n 45).

[47] The AI Act contains only rudimentary individual rights, namely (i) a right to lodge a complaint with a market surveillance authority (Art. 85 AI Act), and (ii) a right to explanation of individual decision-making (Art. 86 AI Act).

violations can be, without problems, ranked on a sliding scale from trivial to serious".[48] Instead, fundamental rights follow a binary logic in that an activity is either legal or illegal.[49]

Finally, risk is typically assessed at the level of the collective/society and not for the individual. Rather than preventing individual harm, risk thinking assesses harm at a social level or a society-wide scale. One of the consequences of this aggregate nature of risk is that individual differences are typically ironed out, as risk analysis often determines acceptable risks by looking at the average citizen.[50] Another consequence is that risk-based regulation often involves society-wide trade-offs (e.g. between fairness and efficiency), with the result that even immense individual harms may be dismissed.[51]

For all these reasons, a risk-based approach is difficult to reconcile with the protection of fundamental rights.

3.2. Missing Risk-*Benefit* Analysis

Another problem with the AI Act is that it lacks a risk-*benefit* analysis - a fundamental component of a truly risk-based approach to regulation.

As outlined above,[52] a risk-*benefit* analysis involves assessing the potential risks and benefits of a particular action or technology to determine whether the benefits outweigh the risks. In addition to determining whether a system has the potential to cause harm and the severity of the likely harm, the analysis must also consider the benefits and likely positive outcomes of using a system, such as advancing scientific discovery.[53] Otherwise, there is no appropriate framework for a proportionate and balanced regulatory regime.

In sharp contrast to this, the AI Act does not consider the potential benefits of AI systems alongside the risks they pose. Instead, the Act focuses primarily on preventing risks and threats to health, safety and fundamental human rights, without considering the potential positive impacts of AI systems.

Besides undermining the idea of truly risk-based approach, this approach ignores the positive contributions of technology and may also result in missed opportunities for societal progress and innovation. In other words, by not considering the positive aspects of the technology in the regulatory framework, the AI Act fails to harness the potential of AI systems to improve the common good.

This absence of a risk-*benefit* analysis in the AI Act also makes it difficult to strike an appropriate balance between potential risks and benefits in dilemma situations. This is illustrated by an example from health care,[54] where it is currently unclear whether AI-based medical devices should have a minimum level of transparency before they can be released to the market. Some data scientists argue that regulators should only allow inherently interpretable algorithmic models while banning AI systems

[48] Karen Yeung and Lee Bygrave, 'Demystifying the modernized European data protection regime: Cross-disciplinary insights from legal and regulatory governance scholarship' (2022) 16 Regulation & Governance 137–155, 146, https://doi.org/10.1111/rego.12401.

[49] Raphael Gellert, 'We Have Always Managed Risks in Data Protection Law: Understanding the Similarities and Differences Between the Rights-Based and the Risk-Based Approaches to Data Protection' (2016) EDPL 481-492, 483.

[50] Kaminski (n 36) 1392.

[51] Kaminski (n 36)

[52] Section 2.2.

[53] London Borough of Waltham Forest, *Risk Assessment & Risk-Benefit Analysis* (London: LBWF Early Years, Childcare & Business Development Service, 2019) 5.

[54] See also Anastasiya Kiseleva, Dimitris Kotzinos and Paul De Hert, 'Transparency of AI in Healthcare as a Multilayered System of Accountabilities: Between Legal Requirements and Technical Limitations' (2022) 5 *Frontiers in Artificial Intelligence* 1, 16 <www.frontiersin.org/articles/10.3389/frai.2022.879603/full>, 11.

with algorithmic opacity that cannot be technically resolved.[55] However, studies show that some opaque AI systems (e.g. deep neural networks) have a much higher degree of accuracy and efficiency than transparent systems (e.g. deductive and rule-based systems).[56] In such a situation, a trade-off between AI's accuracy and transparency must be made.[57]

The Medical Device Regulation (MDR)[58] provides for exactly such a balancing, allowing certain risks to be recognized as acceptable if they are outweighed by the corresponding benefits (Annex I No 4 MDR). Thus, the inherent algorithmic opacity of a medical device may be considered an acceptable risk, if the manufacturer can demonstrate that the benefits of using such a device outweigh the risks.

Whether such a trade-off is also possible under the AI Act – which (in the case of AI based medical devices requiring a third party conformity assessment) applies simultaneously to the MDR – is unclear, given that the AI Act considers only possible risks and their prevention. Therefore, the AI Act does not provide a clear answer to the question of whether a certain degree of algorithmic opacity can be considered an acceptable risk in light of the benefits of using AI. Obviously, the intention of the EU legislator was not to eliminate all risks (the zero-risk approach), but to strike a balance between risk reduction and costs of regulation (the proportionate or efficiency approach), as discussed in sections 1.1. and 2.2. Therefore, a truly risk-based implementation of the AI Act requires striking a balance between algorithmic transparency and efficiency/accuracy.

3.3. Limited Reliance on Empirical Evidence

Another reason why the AI Act is not truly risk-based is its limited reliance on empirical evidence for the design of the different risk categories. As scholars have pointed out,[59] the AI Act does not establish criteria for when AI poses an unacceptable risk to society and individuals. Instead, it merely provides a set list of categories of AI systems that are considered to pose "unacceptable risks" and are therefore banned in the EU.

For high-risk AI systems, Art. 7(2) AI Act sets out the criteria to be taken into account by the European Commission when amending the list of Annex III high-risk AI systems, such as

- The intended purpose and the extent to which an AI system has been used or is likely to be used;
- The extent to which the AI system operates autonomously;
- Whether the use has already caused harm to health and safety or has had an adverse impact on fundamental rights;
- The extent to which affected individuals depend on the output of the AI system;
- whether the output of an AI system is corrigible or reversible;
- and whether EU law is capable of preventing or substantially minimizing those risks.

[55] Cynthia Rudin, 'Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead' (2019) 1(5) *Nature Machine Intelligence* 206 <https://doi.org/10.1038/s42256-019-0048-x>.

[56] Rich Caruana and others, 'Intelligible Models for HealthCare: Predicting Pneumonia Risk and Hospital 30-day Readmission' (2015) *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 1721 <http://people.dbmi.columbia.edu/noemie/papers/15kdd.pdf>. Bernhard Waltl and Roland Vogl, 'Explainable Artificial Intelligence – The New Frontier in Legal Informatics' Jusletter IT (22 February 2018).

[57] Martin Ebers, 'Regulating AI and Robotics' in Martin Ebers and Susana Navas (eds), *Algorithms and Law* (Cambridge 2020) 37-99, 49 et seq.

[58] Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices [2017] OJ L 117/1.

[59] Lilian Edwards, *Regulating AI in Europe: Four Problems and Four Solutions* (Ada Lovelace Institute, March 2022) 11 https://www.adalovelaceinstitute.org/wp-content/uploads/2022/03/Expert-opinion-Lilian-Edwards-Regulating-AI-in-Europe.pdf.

However, neither the recitals of the AI Act nor any accompanying EU document explain and justify how these criteria were applied to identify the areas listed in Annex III in the first place. As *Grozdanovski and De Cooman* conclude, when choosing which risks to address, "regulators were generally disinterested in statistical evidence on the possibly harmful features of various systems."[60]

This applies both to the Commission's original proposal which failed to gather empirical evidence for the design of the AI Act, and to the subsequent legislative process.

Instead of conducting its own practical studies in concrete use cases of AI, the European Commission relied heavily on public consultations[61] for its proposal, despite acknowledging that "robust and representative evidence for harms inflicted by the use of AI is scarce due to the lack of data and mechanisms to monitor Al as a set of emerging technology."[62] Inconsistencies in reported participation numbers as well as the methods employed during the consultation – in particular, the use of closed-ended questions and pre-suggested answers – cast further doubt on the accuracy and representativeness of the data collected.[63] Moreover, there is a discrepancy between the results of the consultation and the final proposal. In certain areas, the proposal deviates from the consensus expressed by the respondents.[64] Given that the consultation process did not really influence the European Commission's decision, the evidence base is further weakened.

The political nature of the definition of risk categories was also evident in the subsequent trilogue negotiations. The Council and the European Parliament proposed new prohibited AI practices and new areas for high-risk AI systems, but with little or no justification as to why these were chosen.[65]

All these lead to the conclusion that the supposedly "risk-based" nature of the Act is neither based on practical evidence nor justified by externally verifiable criteria, but is the result of a political compromise at a particular point in time and is therefore largely arbitrary.[66]

3.4. Pre-defined, Closed Risk Categories

A closely related problem concerns the framing of the risk categories themselves.

At first glance, many of the pre-defined categories do not make sense. For instance, the AI Act does not apply to the most dangerous applications – such as military applications like killer robots (Art. 2(3) AI Act); AI systems developed but not used in the EU to provide support to foreign dictators or hackers (Art. 2(1)(c) AI Act);[67] and autonomous vehicles, drones/airplanes and vessels (Art. 2(2) in conjunction with Annex I.B. AI Act).

[60] Ljupcho Grozdanovski and Jerome De Cooman, 'Forget the Facts, Aim for the Rights! On the Obsolescence of Empirical Knowledge in Defining the Risk/Rights-Based Approach to AI Regulation in the European Union' (2023) 49 Rutgers Computer & Tech LJ 207.

[61] For details see Grozdanovski and De Cooman, 236 et seq.

[62] European Commission, Commission Staff Working Document *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*, SWD(2021) 84 final, Part 1/2.

[63] Grozdanovski and De Cooman (n 60) 239.

[64] Grozdanovski and De Cooman (n 60) 240.

[65] Edwards (n 60) 11.

[66] Edwards (n 60) 11; Grozdanovski and De Cooman (n 60).

[67] However, as I have explained elsewhere, the provision seems justified in light of the fact that the AIA is based on the internal market clause (Art. 114 TFEU), because it is difficult to imagine how the AIA could contribute to the internal market if an AI system is only developed in the EU, but never put into operation there; Martin Ebers and others, The European Commission's Proposal for an Artificial Intelligence Act, Journal "J" (2021) 4, 589–603, 591, https://doi.org/10.3390/j4040043.

On the other hand, many applications qualify as high-risk AI systems under Annex III, simply because they are used in a particular sector, even though they do not pose a serious risk of harm, such as tools to detect duplicates in datasets or tools to improve language.[68] While it is true, that in both of these cases providers have the possibility under Art. 6(3) *sub* 2(a)-(b) AI Act to demonstrate that their systems do not qualify as high-risk, this does not change the fact that these tools are quite often covered by Annex III and only exceptionally exempted, which places the burden of proof (and documentation, Art. 6(4) AI Act) on the provider.

At a more fundamental level, the two examples discussed point to the real problem. The AI Act provides a broad and rather abstract classification of high-risk systems under Annex III. Instead of providing a risk classification on a case-by-case basis, the Act uses a pre-defined, closed list of typical high-risk applications. Whether an AI system used in a specific sector for specific purposes poses a high risk to health, safety and/or fundamental rights, is not assessed for the *concrete* risk, but is pre-defined for typical cases in Annex III. Accordingly, the risk management system required by Art. 9 AI Act is only obligatory in situations that are already classified by the AI Act as high-risk cases. As a result, the risk management obligations of providers under the AI Act consist mainly of risk mitigation rather than risk assessment.[69]

The choice of such a top-down regulation raises several issues. First, this approach leads to over-regulation where, for instance, an AI system falls into one of the eight categories listed in Annex III, but in reality does not pose a significant risk of harm. Second, the list of typical high-risk AI systems (albeit with broad definitions and open to updating) may not be easy for the European Commission to keep up to date in a timely manner, given how rapidly AI technology is evolving.[70] Moreover, the decision to delegate (to the Commission) the power to amend Annex III by adding, modifying and removing high-risk AI systems (Art. 7 AI Act) raises concerns in terms of power allocation.[71]

Finally, the focus on a pre-defined list of high-risk AI systems also creates a sharp rift between this category and other lower-risk categories that are largely unregulated (with the exception of transparency requirements, Art. 50 AI Act). In particular, such a rigid distinction is not justified in cases where an AI system is used in a specific sector (e.g. healthcare sector) but does not qualify as high-risk (e.g. because the system does not qualify as a medical device according to Art. 6(1), Annex I.A.11 AI Act and the MDR),[72] but nevertheless poses numerous risks (e.g. to patients and care recipients due to its direct and indirect effects on the human body and mental health).[73]

[68] Recital (53) of the AI Act lists these cases as examples of cases in which an AI system could be classified as high risk under Annex III of the AI Act, but could be exempted under Art. 6(3) AI Act.

[69] Alessandro Mantelero, *Beyond Data: Human Rights, Ethical and Social Impact Assessment in AI*, (Springer Nature 2022) 169 https://link.springer.com/content/pdf/10.1007/978-94-6265-531-7.pdf

[70] Mantelero (n 69) 170.

[71] Mantelero (n 69) 170. See also Ranjana Achleitner, *Delegierte Rechtssetzung und das Demokratieprinzip: Grenzen und Herausforderungen der exekutiven Rechtssetzung in der EU*, in: Alexander Heger and others (eds.), *Zur Zukunft der Demokratie in der Europäischen Union* (Nomos, Baden-Baden 2023) 101-126.

[72] For example, robots and AI systems used in care for daily communication with the elderly, and applications which provide instructions for workouts, give tips on nutrition, or store the user's weight or pulse.

[73] Martin Ebers, 'AI Robotics in Healthcare between the EU Medical Device Regulation and the Artificial Intelligence Act' (2024) Oslo Law Review, pending for publication.

3.5. Regulation of GPAI Models as a Contradiction to the Risk-based Approach?

The specific legal obligations for providers of so-called “General Purpose AI” (GPAI) models – which were introduced in Art. 51 et seq. AI Act at the last minutes of the trilogue with a “hot needle” – are also inconsistent with a genuine risk-based approach.

By definition, a GPAI model is characterized as such since it “displays significant generality and is capable of competently performing a wide range of distinct tasks regardless of the way the model is placed on the market and that can be integrated into a variety of downstream systems or applications” (Art. 3 No. 63 AI Act). However, the fact that GPAI models can be used for many different purposes in a wide range of areas makes it impossible for their providers to foresee, assess and mitigate their concrete risks.

Clearly, downstream providers integrating GPAI models into their AI systems need “a good understanding of the models and their capabilities, both to enable the integration of such models into their products, and to fulfil their obligations” under the AI Act, as pointed out in recital (101) AI Act. Therefore, it is reasonable for all providers of GPAI to create and regularly update documentation for AI system downstream providers that integrate the GPAI model into their system to help those providers understand the GPAI model’s capabilities and limitations to comply with the AI Act.

However, the provisions of the AI Act that apply to all GPAI model providers go beyond what is necessary:

- Art. 53(1) in conjunction with Annex XI AI Act (which requires a “detailed description”, including information on the “methods to detect identifiable biases”) neglects the fact that bias is not a static phenomenon, but context-specific – as recently demonstrated when Gemini produced ethnically diverse, but historically inaccurate images, such as black Vikings, female popes, and Asian founding fathers.[74]
- Art. 53(1)(b) AI Act (which requires GPAI model providers to disclose sensitive information to any provider that “intends” to integrate the model into its own AI system) not only opens the door to abuse (since “intent” can be easily faked). It is also problematic that such information must be disclosed to *all* providers, even though it is primarily relevant for providers of high-risk AI systems.
- Furthermore, Art. 53(1)(c) AI Act with its obligation to put in place a policy to comply with EU copyright law, is not at all related to the risks that the AI Act seeks to address (safety, health and fundamental rights).

Of even greater concern – from a risk-based regulatory perspective – are the specific obligations that require systemic risk GPAI model providers to conduct model evaluations, assess and mitigate potential systemic risks, monitor and report serious incidents, take corrective action, and ensure an appropriate level of cybersecurity measures (Art. 55 AI Act).

First, by their very nature, “systemic risks” are not limited to specific use cases or applications. Given that GPAI models can be used widely across a variety of industries, it is difficult (*if not even impossible*) to formulate precise standards for classifying risks as systemic. Therefore, the AI Act does not specify which risks are systemic; instead, Art. 3 No. 65 AI Act refers in general terms to negative effects on public health, safety, public security and fundamental rights. Thus, providers have no guidance on what constitutes a systemic risk and how to mitigate it. Arguably, providers can demonstrate compliance

[74] Adi Robertson, ‘Google Apologizes for “missing the Mark” after Gemini Generated Racially Diverse Nazis’ (*The Verge*, 21 February 2024) <https://www.theverge.com/2024/2/21/24079371/google-ai-gemini-generative-inaccurate-historical>.

through codes of practice which will be facilitated by the AI Office until a harmonized standard is published (Art. 55(2)(1) AI Act). However, this does not change the fact that “systemic” risks cannot be quantified and specified in the same way as other risks regulated by the AI Act, because they do not concern the probability of the occurrence of a certain harm (cf. Art. 3 No. 2 AI Act), but rather the impact on the “union market” or the “society as a whole” (cf. Art. 3 No. 65 AI Act).

Such “risk” regulation has nothing in common with the type of risk-based regulation described above in Section 2.

In addition, the AI Act assumes that a particularly high amount of computation used to train GPAI models also increases their risk (Recital 111 AI Act). Therefore, Art. 51(2) AI Act states that a GPAI model is presumed to have high-impact capabilities – such that qualifies it as a model with “systemic risk” under Art. 51(1)(a) AI Act – if the model’s training involves more than 10^25 floating-point operations (FLOPs). However, such a threshold is questionable for at least three reasons:

- First, the (systemic) risk of GPAI models depends, not only on the quantity of computational resources used, but also on a number of other factors - such as the context of the application, the model architecture, and the quality of the training.[75]
- Second, research shows that the 10^25 FLOPs threshold is questionable, since LLMs with fewer FLOPs can be just as risky and even outperform larger models with more parameters.[76]
- And third, the FLOPs threshold was set primarily for political reasons: in order to strengthen the European economy with its two start-ups Mistral (from France) and Aleph Alpha (from Germany), France and Germany in particular successfully lobbied during the negotiations to keep both companies below the threshold[77].

This displays, once again, how arbitrarily the AI Act defines “systemic” risks.
In conclusion, the specific obligations for GPAI are not only inconsistent with the risk-based approach of the AI Act and the principle of technology neutrality, as it means regulating the technology being used rather than the actual risk of the application. Also, the newly introduced category of a “systemic risk” with its FLOP threshold is not based on empirical evidence but rather the result of a political compromise.

3.6. Overly broad AI Definition

Another major setback from the perspective of risk-based regulation is the overly broad definition of AI.

Initially, the European Commission justified its Proposal with the specific characteristics of (unpredictable) software systems based on machine learning, such as (i) opacity (lack of transparency), (ii) complexity, (iii) continuous adaptation and unpredictability, (iv) autonomous behaviour, (v)

[75] Claudio Novelli and others, ‘Generative AI in EU Law: Liability, Privacy, Intellectual Property, and Cybersecurity’ (2024) https://ssrn.com/abstract=4821952, 4.
[76] Cornelia Kutterer, 'Regulating Foundation Models in the AI Act: From “High” to “Systemic” Risk' (AI-Regulation Papers 24-01-1, 12 January 2024) 6 https://ai-regulation.com/wp-content/uploads/2024/01/C-Kutterer-Regulating-Foundation-Models-in-the-AI.pdf; Nicolas Moës and Frank Ryan, 'Heavy is the Head that Wears the Crown: A Risk-Based Tiered Approach to Governing General Purpose AI' (The Future Society, September 2023) https://thefuturesociety.org/wp-content/uploads/2023/09/heavy-is-the-head-that-wears-the-crown.pdf. In contrast, the Executive Order of President Biden (n 15) uses a 10^26 threshold to define (by default) so-called dual-use models, which are then subject to certain reporting requirements (mainly for national security reasons).
[77] However, it appears that France is still unhappy with the current threshold and continues to advocate for a higher threshold (10^26), as companies like Mistral are likely to reach the current threshold in the near future.

functional dependence on data.[78] Accordingly, many of the AI Act's mandatory requirements and obligations for high-risk AI systems – such as testing procedures (Art. 9(6)-(8) AI Act), requirements for training data (Art. 10 AI Act), record keeping (Art. 12 AI Act), transparency (Art. 13 AI Act), human oversight (Art. 14 AI Act) and post-market monitoring systems (Art. 72 AI Act) – attempt to mitigate mainly the risks of AI systems based on machine learning, while such far-reaching obligations are not strictly necessary for other software systems.

From a truly risk-based approach as well as from the principle of technology neutrality, the AI Act should have, therefore, imposed different regulatory burdens on different designs, because predictable AI systems do not pose the same risks as unpredictable systems based on machine learning.[79]

However, this is not the case. According to Art. 3(1) AI Act, an AI system is "a machine-based system designed to operate with *varying levels of autonomy*, that *may* exhibit *adaptiveness after deployment* and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments".[80] This definition - an adaptation of the OECD's updated definition of AI[81] - extends the meaning of AI systems to cover (almost) all software systems,[82] as (i) there is no threshold for the level of autonomy required for a system to be classified as such, and (ii) the use of the word "may" implies that systems do not always have to exhibit adaptiveness after deployment, to be considered AI. Hence, the AI Act applies not only to machine learning, but also to logic- and knowledge-based approaches (recital 12 AI Act).

As a result, even deterministic software systems used in high-risk sectors are subject to the highest requirements. Consequently, as *Schrepel* puts it: "By not discriminating between AI systems based on their functioning, the AI Act indirectly sanctions those that are safer and easier to control"[83] – contrary to the true risk-based approach.

3.7. Double Regulatory Burdens due to the Horizontal Approach

The AI Act does not replace existing EU law, but applies concurrently to it. As a result, companies and individuals will have to observe not only the AI Act, but also other related legislations, such as EU data protection law (Art. 2(7) AI Act), EU copyright law, EU consumer and product safety law (Art. 2(9) AI Act). As many principles and provisions of the AI Act overlap with those of pre-existing legislations, such a *horizontal approach* inevitably leads - in many areas - to legal uncertainty, different interpretations, contradictions and, ultimately, to double regulatory burdens – contrary to the idea of risk-based regulation.

[78] European Commission, *AI Act Proposal*, COM(2021) 206 final, explanatory memorandum, 2; European Commission, *Commission Staff Working Document, Impact Assessment, Accompanying the Proposal for a Regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*, SWD(2021) 84 final, Part 1/2, 28 et seq. See also Martin Ebers, 'Regulating AI and Robotics: Ethical and Legal Challenges' in Martin Ebers and Susana Navas (eds), *Algorithms and Law* (Cambridge University Press 2020) 44 et seq.

[79] Thibault Schrepel, *Decoding the AI Act: A Critical Guide for Competition Experts* (ALTI Working Paper, Amsterdam Law & Technology Institute – Working Paper 3-2023, October 2023) 11 https://ssrn.com/abstract=4609947.

[80] Emphasis added.

[81] OECD, *Recommendation of the Council on Artificial Intelligence*, OECD/LEGAL/0449, amended on 03/05/2024 by the 2024 OECD Ministerial Council Meeting (MCM) https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449. Cf. thereto OECD (2024), "Explanatory memorandum on the updated OECD definition of an AI system", *OECD Artificial Intelligence Papers*, No. 8, OECD Publishing, Paris, https://doi.org/10.1787/623da898-en.

[82] According to Recital (12) AI Act, the only software systems that should not be regarded as AI are "systems that are based on the rules defined solely by natural persons to automatically execute operations".

[83] Schrepel (n 79) 11.

Consider the following three examples from data protection law, medical law and product safety for machinery:

(1) As the AI Act and EU data protection law apply in parallel, both the EDPB and the EDPS have already pointed out during the negotiations, that it is important to clearly avoid any inconsistencies and possible conflicts between the AI Act and data protection law.[84] However, these concerns have largely not been taken into account. For example, both the GDPR and the AI Act impose transparency obligations, but the scope and the requirements are regulated differently in the two laws.[85] Another example is the right to explanation and human intervention/oversight. While the GDPR requires human intervention (Art. 22(3) GDPR) and a right to meaningful information for decisions based solely on automated processing, including profiling (Art. 15(1)(h) GDPR), the AI Act requires human oversight (Art. 14 AI Act) and a right to explanation of individual decision-making (Art. 86 AI Act) for high-risk AI systems – whereby both the content as well as the prerequisites and legal consequences are regulated completely differently.

(2) The relationship between the AI Act and the Medical Devices Regulation (MDR)[86] is also currently unclear.[87] Given that both legislations apply simultaneously, without a formal hierarchy clause in either the AI Act or in the MDR to decide which of the overlapping rules should apply, a number of inconsistencies and contradictions arise. The AI Act is inconsistent with many of the MDR's concepts and definitions. For example, it offers different definitions for certain terminologies - such as "importer", "putting into service", "provider" and "deployer" – from those of the MDR.[88] These differences do not only complicate compliance with both regulations, but could also result in single set of technical documentation[89] defining the same terms differently.[90] In addition, many of the AI Act's mandatory requirements for high-risk AI systems overlap with the MDR's requirements for medical devices in contradictory ways, with no clarity as to which of them takes precedence. For example, both the MDR and the AI Act require the establishment of a risk management system. However, while Annex I of the MDR requires risks to be reduced as far as possible,[91] Art. 9(4) AI Act refers only to "minimising risks more effectively while achieving an appropriate balance in implementing the measures to fulfil those requirements".

[84] EDPB-EDPS, Joint Opinion 5/2021 on the Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), 18 June 2021, para 57.
[85] The GDPR establishes the principle of transparency to facilitate the exercise of data subjects' rights under Art 15-22, including the right to erasure, to rectification and to data portability. In contrast, the AI Act contains transparency obligations only for high-risk AI systems (Art 13 AI Act) and for other certain AI systems (Art 50 AI Act). Moreover, Art 13 AI Act focuses on the interests of the deployer of an AI system rather than on the final user and/or data subject.
[86] Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices [2017] OJ L 117/1.
[87] Cf. Martin Ebers, AI Robotics in Healthcare between the EU Medical Device Regulation and the Artificial Intelligence Act: Gaps and Inconsistencies in the Protection of Patients and Care Recipients, (2024) Oslo Law Review (pending for publication).
[88] For a detailed discussion cf Wimmy Choi, Marlies van Eck and Cécile van der Heijden, 'Theo Hooghiemstra and Erik Vollebregt, Legal analysis: European legislative proposal draft AI act and MDR/IVDR' (January 2022) 16ff, <www.government.nl/binaries/government/documenten/publications/2022/05/25/legal-analysis-european-legislative-proposal-draft-ai-act-and-mdr-ivdr/Report+analysis+AI+act+-+MDR+and+IVDR.pdf>.
[89] According to Article11(2) AI Act, the technical documentation required under the AI Act must be integrated with the technical documentation required under the MDR.
[90] Choi and others (n 88)18.
[91] Cf. in particular Annex I No 2: 'The requirement in this Annex to reduce risks as far as possible means the reduction of risks as far as possible without adversely affecting the benefit-risk ratio'.

(3) A third example is the new Machinery Regulation (MR),[92] which applies alongside the AI Act when an AI system is used in a machine. This also results in a duplication of requirements. For example, both the AI Act and the MR require human oversight, but the two sets of rules differ in detail.[93] Another overlap and contradiction concerns the recoding and retention of decision-making data. While the MR requires "enabled" recording of "data on the safety-related decision-making process" the AI Act requires that high-risk AI systems automatically record logs of "events" throughout the system's lifespan. This discrepancy is likely to create practical challenges for companies to ensure compliance with both regulatory frameworks, resulting in double regulatory burdens.

All three examples – and many others[94] – show that the AI Act will create inefficiencies, regulatory uncertainty and increased compliance costs, due to conflicting or duplicative requirements in both the AI Act and other EU laws.

3.8. Overlap of Enforcement Structures

Since the AI Act applies in addition to other existing EU laws, there is also a risk that the same use of an AI system may be subject to different regulatory authorities in one and the same Member State. Art. 70 AI Act leaves the designation of competent authorities to the Member States.[95] As a result, Member States may choose to entrust the enforcement of the AI Act either to existing bodies (such as national data protection authorities) or to entirely new administrative bodies (such as the Agency for the Supervision of AI in Spain). This will most likely lead to overlapping enforcement structures, duplication of procedures, inconsistencies between these procedures and, in the worst case, double fines for the same set of facts.

In some cases, especially in highly regulated markets, it may even be the case that more than two administrative bodies will compete to enforce their set of rules, as in the example of credit scoring – where Member States may be subject to a financial supervisory authority (responsible for enforcing the Consumer Credit Directive 2023/2225),[96] data protection authorities (responsible for enforcing the GDPR),[97] and market surveillance authorities (responsible for enforcing the AI Act).[98] In all of these cases, coordination between the different regulatory bodies is necessary to avoid double and/or over-

[92] Regulation (EU) 2023/1230 of the European Parliament and of the Council of 14 June 2023 on machinery and repealing Directive 2006/42/EC of the European Parliament and of the Council and Council Directive 73/361/EEC (OJ L 165, 2962023, 1–102.

[93] Tobias Mahler, 'Smart Robotics in the EU Legal Framework: The Role of the Machinery Regulation' (2024) Oslo Law Review (pending for publication).

[94] For more examples cf. Gerald Spindler, 'Algorithms, Credit Scoring, and the New Proposals of the EU for an AI Act and on a Consumer Credit Directive' (2021) 15 Law and Financial Markets Review 239-261: Frictions between the AI Act (proposal), Capital Requirements Directive 2013/36/EU and the (back then: proposed) Consumer Credit Directive 2023/2225.

[95] Cf. also Recital 157 AI Act: "This Regulation is without prejudice to the competences, tasks, powers and independence of relevant national public authorities or bodies which supervise the application of Union law protecting fundamental rights, including equality bodies and data protection authorities."

[96] Cf. Art. 41 Consumer Credit Directive 2023/2225: "Member States shall designate the national competent authorities empowered to ensure the application and enforcement of this Directive and shall ensure that they are granted investigating and enforcement powers and adequate resources necessary for the efficient and effective performance of their duties."

[97] On the application of the GDPR to credit scoring, see in particular the recent Schufa-case; CJEU, case C-634/21 *OQ v Land Hessen* ECLI:EU:C:2023:957.

[98] Cf. Annex III.5.b AI Act (credit scoring as "high-risk AI system").

enforcement, which is contrary to the constitutional principles of *ne bis in idem* (Art. 50 EU Charter of Fundamental Rights) and the principle of proportionality.[99]

However, in the absence of specific provisions on cooperation mechanisms in the AI Act and other EU legislation, the overlap of enforcement structures will be a significant challenge, increasing legal uncertainty and compliance costs – again, contrary to the risk-based approach.

4. How to Implement the AI Act in Accordance with a Truly Risk-based Approach

The foregoing analysis shows that key provisions of the AI Act do not reflect a truly risk-based approach; leading to legal uncertainty and potential over-regulation, as well as unjustified increases in compliance costs. However, this is nothing that cannot be fixed. The AI Act provides for sufficient tools to support future-proof legislation and to implement it in line with a genuine risk-based approach. Accordingly, the following analysis focuses on how to implement the AI Act in accordance with a truly risk-based approach.

To this end, the paper first discusses why the risk-based approach has to be observed as a guiding principle for implementation (4.1.). It will then focus on the various tools the AI Act provides to support future-proof legislation (4.2.), followed by recommendations on how to apply and implement the AI Act in accordance with a truly risk-based approach (4.3.). In particular, this section will assess what the European Commission should consider when issuing and/or adopting guidelines, delegated and implementing acts, the codes of practice and harmonized standards.

4.1. The Risk-based Approach as a Guiding Principle for Implementing the AI Act

The implementation of the AI Act can take several forms. In this paper, the term "implementation" is used to describe the measures taken to ensure compliance with the obligations imposed by the AI Act. Implementation includes:

- the adoption of more specific legal provisions (normative implementation),
- the interpretation, application, and enforcement of the AI Act by public authorities, including guidelines by the European Commission (administrative implementation) and
- its interpretation by the courts (judicial implementation).

Implementation can take place both at the level of the EU (e.g. when the European Commission issues guidelines or adopts implementing/delegated acts) and at Member State level (e.g. when Member States adopt more specific rules or when Member State authorities enforce the AI Act).

When implementing the AI Act, both the European Commission and the Member States must respect the choice of the European legislator to follow a risk-based approach. This follows both from the preparatory work[100] and from the *ratio legis* as laid down in recital (26) AI Act.

As explained above, the overall objective of the AI Act's risk-based approach is to strike an optimal (and proportionate) balance between innovation and the benefits of AI systems on the one hand, and the

[99] According to consistent case law of the ECJ in competition law, if "the possibility of two procedures being conducted separately were to lead to the imposition of consecutive sanctions" for the same acts, a general requirement of natural justice "demands that any previous punitive decision must be taken into account in determining any sanction which is to be imposed"; Case 14/68 *Wilhelm* [1969] ECR 1, para. 11.

[100] Cf. European Commission, 78.

protection of fundamental values such as safety, health, and fundamental rights on the other. To this end, as stated in recital (26) AI Act, "a clearly defined risk-based approach should be followed" which tailors "the type and the content of such rules to the intensity and scope of the risks that AI systems can generate".

The recitals of EU regulations such as the AI Act form an integral part of the legislation, according to Art. 296 TFEU. They are the most important source for determining its objective and meaning. For this reason, the CJEU consistently refers to the recitals in order to interpret EU law. Indeed, it is true that the recitals have "no binding legal force and cannot be validly relied on either as a ground for derogating from the actual provisions of the act in question or for interpreting those provisions in a manner clearly contrary to their wording."[101] However, where there is no (obvious) contradiction between the operative part of an EU legal act (and its recitals) and where EU law needs to be interpreted or concretized, the CJEU regularly refers to the recitals, because:

> "the operative part of a Community act is indissociably linked to the statement of reasons for it, so that, when it has to be interpreted, account must be taken of the reasons which led to its adoption".[102]

Apart from the *ratio legis* of the AI Act itself, an interpretation of the AI Act in the light of EU primary law also supports the view that the implementation of the AI Act should follow a truly risk-based approach. As explained above, the principle of (legislative) proportionality is enshrined in Art. 5 TEU. The CJEU has consistently held that:

> "the principle of proportionality is one of the general principles of Community law. By virtue of that principle, the lawfulness of the prohibition of an economic activity is subject to the condition that the prohibitory measures are appropriate and necessary in order to achieve the objectives legitimately pursued by the legislation in question; when there is a choice between several appropriate measures recourse must be had to the least onerous, and the disadvantages caused must not be disproportionate to the aims pursued."[103]

As a result, both secondary law (*ratio legis* of the AI Act) and primary law (principle of legislative proportionality) favour the implementation of the AI Act on the basis of a truly risk-based approach, provided, of course, that the AI Act possesses sufficient flexible tools to be implemented via a genuine risk-based approach.

4.2. Tools in the AI Act to Support Future-proof Legislation

Although some key provisions of the AI Act are not consistent with a truly risk-based approach, the Regulation provides for sufficient tools to interpret, specify, and even amend it in line with a genuine risk-based approach.

Guidelines of the European Commission

First, many provisions contain broad language that is subject to interpretation. To this end, the Regulation relies, not only on the courts (and ultimately on the CJEU), but also on the European

[101] Case C-134/08 *Hauptzollamt Bremen v J. E. Tyson Parketthandel GmbH hanse j.* [2009] ECR I-2875, para 16. Cf. Case C-162/97 *Gunner Nilsson* [1998] ECR I-7477, para 54; Case C-444/03 Case C-444/03 *Meta Fackler KG v Bundesrepublik Deutschland* [2005] ECR I-3913, para 25; Case C-136/04 *Deutsches Milch-Kontor GmbH v Hauptzollamt Hamburg-Jonas* [2005] ECR I-10095, para 32.

[102] ECJ, 19.11.2009 – Joined Cases C-402/07 *Christopher Sturgeon, Gabriel Sturgeon and Alana Sturgeon v Condor Flugdienst GmbH* and (C-432/07) *Stefan Böck and Cornelia Lepuschitz v Air France SA* [2009] ECR I-10923, 42.

[103] Case C-331/48 *The Queen v Minister of Agriculture, Fisheries and Food and Secretary of State for Health*, ex parte Fedesa et al.

Commission and its AI Office. According to the AI Act, the European Commission shall develop guidelines for the practical implementation of the Regulation, in particular for

- the application of the definition of an AI system (Art. 96(1)(f) AI Act),
- the classification of AI systems as high-risk under Annex III, together with a comprehensive list of practical examples of use cases of AI systems that are high-risk and those that are not (Art. 6(3) AI Act),
- detailed information on the relationship of the Act with the Union harmonization legislation listed in Annex I, as well as with other relevant Union law, including with regard to consistency in their enforcement (Art. 96(1)(e) AI Act), and
- the application of the requirements and obligations for high-risk AI systems set out in Articles 8 to 15 (Art. 96(1)(a) AI Act).

In addition, Art. 66(2)(e) AI Act provides that the AI Board may, at the request of the Commission or on its own initiative, issue recommendations and written opinions on all relevant matters relating to the implementation of this Regulation and to its consistent and effective application, including the Commission's guidelines.

Although guidelines – like recommendations (cf. Art. 288 TFEU) – are non-binding instruments which cannot be regarded as rules of law, they are often used by the CJEU and by national courts as a guide for the interpretation of EU laws. Moreover, guidelines issued by the European Commission constitute "rules of practice from which the administration [here: European Commission; ME] may not depart in an individual case without giving reasons that are compatible with the principle of equal treatment."[104] By adopting such guidelines, the Commission "imposes a limit on the exercise of its discretion and cannot depart from those rules under pain of being found, where appropriate, to be in breach of the general principles of law, such as equal treatment or the protection of legitimate expectations".[105]

Delegated Acts and Implementing Acts of the European Commission

Delegated acts and implementing acts of the European Commission have an even more far-reaching effect.[106] The AI Act confers both types of executive law-making powers on the European Commission. In exercising its delegated powers, the European Commission may not only amend the AI Act, but also "supplement" it (Art. 290 TFEU). Implementing powers, on the other hand, are granted to the Commission "where uniform conditions for implementing legally binding Union acts are needed" (Art. 291(2) TFEU). They authorize the Commission "to adopt all the measures which are necessary or appropriate for the implementation of the basic legislation, provided that they are not contrary to it".[107]

In particular, Art. 97 AI Act gives the European Commission the power to adopt delegated acts in order to

- amend, modify or remove use-cases for high-risk AI systems in Annex III (Art. 7(1) and 7(3) AI Act),

[104] Judgment of the Court (Grand Chamber) of 28 June 2005. CJEU Joined Cases C-189/02 P *Dansk Rørindustri A/S,* C-202/02 P *Isoplus Fernwärmetechnik Vertriebsgesellschaft mbH,* C-205/02 P *KE KELIT Kunststoffwerk GmbH,* C-206/02 P *LR af 1998 A/S,* C-207/02 P *Brugg Rohrsysteme GmbH,* C-208/02 P *LR af 1998 (Deutschland) GmbH, and* C-213/02 P *ABB Asea Brown Boveri Ltd v Commission of the European Communities* ECLI:EU:C:2005:408.

[105] Judgment of the Court, C-189/02 P, para 211.

[106] Cf. European Parliamentary Research Service (EPRS), *Understanding Delegated and Implementing Acts* PE 690.709 – July 2021, Micaela Del Monte and Rafał Mańko https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/690709/EPRS_BRI(2021)690709_EN.pdf.

[107] Judgment of the Court (Grand Chamber) of 1 April 2008. Case C-14/06 and C-295/06 *European Parliament and Denmark v Commission* [2008] ECLI:EU:C:2008:176, para 52.

- modify or add new conditions under which Annex III high-risk AI systems shall not be considered to be high-risk according to Art. 6(3) AI Act (Art. 6(6)-(7) AI Act),
- amend Annex IV-VII regarding technical documentation and the conformity assessment procedures for high-risk AI systems (Art. 11(3), 43(5), 43(6), Art. 47(5) AI Act),
- amend the thresholds for classifying GPAI models as "systemic risk", as well as to supplement the benchmarks and indicators for these thresholds (Art. 51(3) AI Act), including the criteria set out in Annex XIII for the designation of GPAI models with systemic risk (Art. 52(4) AI Act),
- amend Annexes XI-XII regarding the technical documentation for providers of GPAI models (Art. 53(5) - (6) AI Act).

Moreover, several provisions of the AI Act authorize the European Commission to adopt implementing acts, such as

- common specifications (Art. 41(1) AI Act),
- codes of practice for GPAI models (Art. 50(7)(2) and Art. 56(6) AI Act),
- detailed arrangements and the conditions for the evaluation of GPAI models (Art. 92(6) AI Act),
- detailed arrangements and procedural safeguards with regard to fines for GPAI model providers (Art. 101(6) AI Act),
- detailed provisions establishing a template for the post-market monitoring plan and the list of elements to be included in the plan (Art. 101(3)(3) AI Act).

Harmonized Standards and Codes of Practice

Other tools provided by the AI Act are harmonized standards and the aforementioned codes of practice for GPAI.

For high-risk AI systems, the AIA relies mainly on a conformity assessment procedure using harmonized standards, combined with a presumption of conformity - where the provider follows these harmonized standards. As explained in a previous paper,[108] the mandatory requirements for high-risk AI systems (such as quality criteria for training, validation and testing data; provisions for transparency and user information; obligations for human oversight; obligations for accuracy, robustness, and cybersecurity) are worded in a rather broad way. Instead of formulating the requirements for high-risk AI systems itself, the Regulation defines only the essential requirements, leaving the details to standards developed by European Standardization Organizations. To this end, the European Commission has already requested, pursuant to Art. 10(1) Regulation 1025/2012, CEN and CENELEC to develop harmonized standards for the requirements of high-risk AI systems by 30 April 2025.[109] Once CEN/CENELEC have delivered these standards and the European Commission has accepted them and published a reference to such harmonized standards in the Official Journal of the EU, providers of high-risk AI systems will enjoy legal certainty: High-risk AI systems that are in conformity with such

[108] Martin Ebers, 'Standardizing AI – The Case of the European Commission's Proposal for an Artificial Intelligence Act' in Larry A DiMatteo, Cristina Poncibò, Michel Cannarsa (eds.), *The Cambridge Handbook of Artificial Intelligence: Global Perspectives on Law and Ethics*, (Cambridge University Press 2022) 321-344 http://ssrn.com/abstract=3900378.

[109] Commission Implementing Decision of 22 May 2023 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 (Standardization Request), C(2023) 3215 final.

harmonized standards shall be presumed, according to Art. 40(1) AI Act, to be in conformity with the mandatory requirements set out in Chapter III, Section 2 AI Act.

For GPAI models, the AI Act provides a similar mechanism. Here, too, GPAI model providers can rely on the presumption of conformity if they comply with harmonized standards (Art. 40(1) AI Act). However, as technical standardization in this area is still in its infancy, the AI Act provides an additional tool for GPAI model providers to comply with their obligations: the so-called codes of practice (Art. 56 AI Act). As soon as the European Commission decides to approve codes of practice by means of implementing acts, GPAI model providers can rely on it to demonstrate compliance with their obligations (Art. 53(4)(1) and 55(2)(1) AI Act).

4.3. Recommendations on how to Implement the AI Act in Accordance with a Truly Risk-based Approach

All of the instruments discussed above – guidelines, delegated and implementing acts of the European Commission, as well as harmonized standards and codes of practice – are powerful tools to clarify, concretize, amend, supplement, modify or even delete a large number of provisions of the AI Act to "take into account the rapid pace of technological development, as well as the potential changes in the use of AI systems" (recital 52 AI Act). Indeed, such measures are necessary to implement the AI Act in accordance with a truly risk-based approach – in line with its *ratio legis* and the principle of legislative proportionality enshrined in Art. 5 TEU.

For this, the following aspects should be taken into account.

Risk-*Benefit* Analysis and Evidence-based High-risk Categories

As explained above, the AI Act lacks an explicit, independent risk-*benefit* analysis and evidence-based (high-)risk categories. On the other hand, the European Commission has - not only the possibility to issue guidelines for the classification of AI systems as high-risk under Annex III, but also - the power to amend, modify or remove use-cases for high-risk AI systems in Annex III (Art. 7(1) and 7(3) AI Act) and to modify or add new conditions under which Annex III high-risk AI systems shall not be considered to be high-risk according to Art. 6(3) AI Act (Art. 6(6)-(7) AI Act).

Here, the Commission should not only take into account the potential harm that AI systems may cause, but also explicitly consider their economic and social benefits. To a certain extent, the Commission already has to assess the beneficial effects if it wants to amend Annex III by adding or modifying use-cases of high-risk systems, as Art. 7(2)(j) AI Act requires the Commission in these cases to also take into account "the magnitude and likelihood of benefit of the deployment of the AI system for individuals, groups, or society at large".[110] However, the benefits should not be one of many factors to be considered under the umbrella of a risk analysis, but an *independent* criterion to be weighed against the risks.[111]

Moreover, the classification of AI systems as "high-risk" should be based on sufficient empirical evidence. This requirement is also present in the AI Act. In particular, Art. 6(5) AI Act, states that the Commission, when adopting delegated acts to add new conditions under which an AI system referred to in Annex III is not to be considered high-risk, shall follow "concrete and reliable evidence". However, such evidence-based risk analysis should be carried out not only in this context, but generally when

[110] This criterion applies also, as per Art. 7(3)(a) AI Act, when the Commission wants to remove high-risk AI systems from the list in Annex III.

[111] In line with Schrepel's claim to follow a "law + technology" approach; cf. Thibault Schrepel, 'Law + Technology' (Stanford CodeX Working Paper, 19 May 2022) https://www.ssrn.com/abstract=3395293.

defining (or changing) the risk categories, together with a clear methodology, explanation and documentation.

Regulation of GPAI Models

With respect to GPAI models, the European Commission also has flexible tools at its disposal to address at least some of the concerns raised above.

This applies in particular to the thresholds for classifying GPAI models as "systemic risk", which can be amended by delegated acts (Art. 51(3) and Art. 52(4) AI Act). Thus, the Commission can revise the criteria, especially the FLOP threshold which is not based on empirical evidence but rather the result of a political compromise. Art. 51(3) AIA even goes beyond the mere update of the FLOPs threshold, by empowering the Commission to also "supplement benchmarks and indicators in light of evolving technological developments, such as algorithmic improvements or increased hardware efficiency, when necessary, for these thresholds to reflect the state of the art". Accordingly, the Commission has the opportunity to use real-world evidence to set and define the systemic risk threshold by going beyond FLOPs and adding or replacing them with new benchmarks.

In addition, it will be of paramount importance that the codes of practice clearly specify which systemic risks GPAI model providers must assess and mitigate pursuant to Art. 55(1) AI Act. In this respect, the AI Act is very vague and not concrete enough. To be systemic, a risk must have a "significant" impact on the Union market due to the reach of the model, or due to "actual or reasonably foreseeable" negative effects on public health, safety, public security, fundamental rights, or the society as a whole (Art. 3(63) AI Act). These include, according to recital (110) AI Act, major accidents, disruptions of critical sectors and serious consequences for public health and safety, such as chemical, biological, radiological, and nuclear risks; democratic processes, public and economic security; and the dissemination of illegal, false, or discriminatory content. However, there is currently little experience in reliably assessing when these and other systemic risks are "actual or reasonably foreseeable".

Accordingly, there is a need to develop a common methodology to define systemic risks and to establish measures as well as benchmarks for assessing and managing these risks. It is therefore essential that GPAI model providers, standard-setting organizations, national authorities, civil society organizations and other relevant stakeholders work together with the AI Office to develop codes of practice in this area, considering international approaches.[112]

This is likely to be a major challenge as the AI Office and GPAI model providers only have 9 months after the publication of the AI Act in the Official Journal to draft the Codes of Practice - a very short timeframe for drafting a Code based on very vague rules. Against this background, the focus should be primarily on the risk-based approach - addressing only relevant risks and sticking to the letters of the AI Act.

Overly Broad AI Definition and Mandatory Requirements for High-risk AI Systems

Proper risk-based implementation can also mitigate the overly broad definition of AI and the problem that deterministic software systems used in high-risk sectors are subject to the same mandatory (strict) requirements as unpredictable AI systems based on machine learning. While the definition of AI systems in Art. 3 AI Act cannot be changed by the Commission but only be interpreted, many of the requirements in Art. 8-15 AI Act are worded broadly enough to be applied in a manner that takes into account the fact that AI systems pose different risks due to their different levels of autonomy and adaptability:

[112] For example, the Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI Systems https://www.mofa.go.jp/files/100573473.pdf.

- For example, the requirements laid down in Art. 10 AI Act on data and data governance appear to apply only to AI systems "which make use of techniques involving the training of AI models with data", thus excluding other AI systems that are not based on machine learning.
- Moreover, the wording of Art. 13(1) AI Act that high-risk AI systems shall be "sufficiently" transparent to enable deployers to interpret the output of a system and use it "appropriately" is open enough to distinguish between different AI systems with different levels of transparency.
- With respect to human oversight, Art. 14(3) AI Act even explicitly emphasizes that the necessary oversight measures must be "commensurate with the risks, level of autonomy and context of use of the high-risk AI system".
- Furthermore, Art. 15(1) AI Act speaks only of an "appropriate" level of accuracy, robustness and cybersecurity" and Art. 15(4) AI Act even provides specific rules for high-risk AI systems "that continue to learn".

Against this background, it seems sufficient that the European Commission issues guidelines on the application of Articles 8 to 15 (Art. 96(1)(a) AI Act) to clarify how these articles apply with regard to different AI technologies.

Double Regulatory Burdens and Overlap of Enforcement Structures

Guidelines of the European Commission can also help companies to deal with overlaps, inconsistencies and (potential) contradictions between the AI Act and other EU legislation to avoid double regulatory burdens and over-enforcement. Art. 96(1)(e) AI Act requires the Commission to develop guidelines with "detailed information on the relationship of the Act with other relevant Union law, "including with regard to consistency in their enforcement" (Art. 96(1)(e) AI Act).

To this end, the European Commission should conduct an in-depth analysis to identify overlaps and contradictions between the AI Act and other horizontal or sectoral legislation. Based on this research, guidelines could then be rolled out to help clarify the relationship between these laws (i.e. lex specialis, lex generalis, and complementary laws).[113]

Moreover, the European Commission should carry out an assessment of the relationship between the AI Act governance bodies and other governance bodies on EU and Member State level. Guidelines could then clarify whether a particular governance body is the lead authority and how the different bodies should cooperate or relate with each other.[114]

5. Revision of Sector-specific EU Laws

The above-mentioned guidelines on legal and governance overlaps could make an important contribution to clarifying the interplay between the AI Act and other EU laws including the relationship between different enforcement bodies. However, guidelines can only interpret existing laws and not change them. Yet, many issues related to double regulatory burdens and lack of coordination between different enforcement bodies cannot be solved by interpretation alone, as they are rooted in legal inconsistencies, mostly at EU level. Since 2018, the EU has adopted an extensive body of digital legislation, mostly under the Digital Agenda and the Digital Single Market (DSM) initiatives, which has

[113] Similarly, Axel Voss, 'Ten steps to make the AI Act an EU success story' (06 March 2024) <https://www.axel-voss-europa.de/2024/03/06/ten-steps-to-make-the-ai-act-an-eu-success-story/>.

[114] See again Voss (n 113)

produced more than 100 laws/proposals and 66 enforcement agencies and other bodies that are crucial for the digital sector – leading to countless overlaps and contradictions, both at the legislative and the enforcement level.[115]

Therefore, only a revision of sector specific EU laws can deal with these issues. To this end, the legislator should carry out a clear gap analysis and impact assessment in order to avoid over-regulation, taking into account, above all, the following areas.

Standardization of Definitions: To streamline regulatory compliance, definitions used in sectoral regulations and the AI Act should be standardized and aligned. Consistent definitions across different regulatory frameworks can help companies interpret and apply the requirements more effectively, thereby reducing confusion and minimizing the risk of non-compliance.[116]

Clarification of Regulatory Overlap: In order to prevent double regulation, clear guidelines should be established to identify how sectoral regulations apply to AI systems (as for example those embedded in products like medical devices). This clarity will help companies understand which regulations they need to comply with and avoid unnecessary duplication of requirements.

Cross-Referencing and Complementary Application: Where sectoral regulations and the AI Act intersect, there should be provisions for cross-referencing and complementary application. This means that the requirements of the AI Act should complement and enhance the existing sectoral regulations, rather than contradicting or duplicating them. This approach ensures a coherent regulatory framework that addresses relevant risks.

Comprehensive Risk Assessment: Another important reason for revising sector-specific EU laws for AI is to ensure a more comprehensive risk assessment framework. This involves evaluating all measures - both quantitative and qualitative - in the implementation of AI systems in specific sectors. By conducting a thorough risk assessment, regulators can identify potential risks and tailor regulations accordingly.

Sector-Specific Guidelines: When revising EU laws for specific sectors, it is crucial to develop sector-specific guidelines for AI applications within that industry. By this, it is important to avoid adopting a one-size-fits-all regulatory approach for different types of AI systems. Instead, regulators should tailor regulations based on the specific risks associated with different AI applications in various sectors.[117] Sector-specific guidelines should consider the unique risks and requirements of the sector while aligning with the broader risk-based approach advocated by the EU AI Act.

Adaptive Regulatory Framework: In order to avoid over-regulation and ensure a truly risk-based approach, the regulatory framework should be adaptive and responsive to the evolving landscape of AI technology. Regular reviews and updates to sector-specific laws can help ensure that regulations remain effective and proportionate to the risks posed by AI systems.

Clarification of Enforcement Overlap: Further, to avoid double enforcement, sector specific laws should clarify which governance body competent is the lead authority and how the different bodies should cooperate with each other.

Regulatory Guidance and Support: Regulatory authorities should provide guidance and support to businesses operating in sectors where AI is used, such as the healthcare industry. This assistance can

[115] Kai Zenner, J Scott Marcus and Kamil Sekut, 'A dataset on EU legislation for the digital world' (16 November 2023) https://www.bruegel.org/dataset/dataset-eu-legislation-digital-world, last updated https://www.kaizenner.eu/post/digital-factsheet-vol-3.

[116] AppliedAI Initiative, 50.

[117] Kaminski (n 39) 1403

help companies navigate the complex regulatory landscape, understand their obligations, and implement necessary measures to ensure compliance with both sectoral and AI-specific regulations.

These aspects should also be taken into account in any future sector-specific AI-related legislation which should also follow the principles of truly risk-based regulation outlined in section 2 above.

6. Regulating AI Outside the EU: Lessons from the AI Act

Lawmakers around the world are looking at the AI Act to determine whether they should follow the European Union's lead and adopt similar laws to regulate AI systems. This is the so-called "Brussels effect" – a term coined by *Anu Bradford*, to describe Europe's global footprint in terms of triggering emulation in other legal systems. In its original formulation, the Brussels effect was seen mainly as a *de facto* phenomenon where companies voluntarily follow EU rules in standardising a product or service, making their business processes simpler.[118] However, it can also take a *de iure* effect where countries outside the EU adopt EU-like regulations.[119]

Over the past years, scholars have discussed whether the AI Act will unleash a new Brussels effect. While some claim this could be the case,[120] others disagree.[121] EU policymakers strongly believe in the Brussels effect. From the onset, the AI Act was designed with its extraterritorial effects in mind.[122]

In the following sections, this paper argues that it is unlikely that the AI Act will unfold a *de iure* Brussels effect (6.1.). Moreover, in light of the lessons the AI Act can teach lawmakers around the world, such an effect would also be undesirable (6.2.).

6.1. Why the AI Act *Won't* Trigger a Brussels Effect

One major obstacle for foreign legislators to simply copy and paste the AI Act is the complexity of AI as a policy area. Unlike the GDPR – which has indeed served as a model for data protection regulation around the world – Artificial Intelligence does not present a single policy problem (e.g. how to protect the fundamental right to privacy), but rather a set of loosely connected problems – ranging from the protection of health and safety to a variety of fundamental rights. Even when legislators follow the definition of AI – as set out in the AI Act, which is based on the recently updated OECD principles – there is little agreement worldwide on *who* and *what* should be regulated: public administration, law enforcement bodies, the judiciary, alternative dispute resolution mechanisms, and/or the (entire)

[118] Anu Bradford, *The Brussels Effect: How the European Union Rules the World* (Oxford University Press 2020) xiv.

[119] Bradford (n 118), 85.

[120] Fabian Lütz, 'How the 'Brussels effect' could shape the future regulation of algorithmic discrimination' (2021) 1 Duodecim *Astra* 142-63; Charlotte Siegmann and Markus Anderljung, 'The Brussels effect and artificial intelligence: How EU regulation will impact the global AI market' (2022) arXiv preprint arXiv:2208.12645 https://arxiv.org/abs/2208.12645; Gerhard Wagner, 'Liability Rules for the Digital Age: Aiming for the Brussels Effect' (2023) 13(3) Journal of European Tort Law 191-243; Nathalie A Smuha, 'From a 'race to AI' to a 'race to AI regulation': regulatory competition for artificial intelligence' (2021) 13(1) Law, Innovation and Technology 57–84 https://doi.org/10.1080/17579961.2021.1898300 ("regulatory landscape for AI is trending towards at least a basic layer of convergence").

[121] Alex Engler, 'The EU AI Act Will Have Global Impact, but a Limited Brussels Effect' (8 June 2022) https://www.brookings.edu/articles/the-eu-ai-act-will-have-global-impact-but-a-limited-brussels-effect/; M Almada and A Radu, 'The Brussels Side-Effect: How the AI Act Can Reduce the Global Reach of EU Policy' (2024) German Law Journal 1-18 https://doi.org/10.1017/glj.2023.108 accessed 17 April 2024; Ugo Pagallo, 'Why the AI Act Won't Trigger a Brussels Effect' (16 December 2023) in *AI Approaches to the Complexity of Legal Systems* (Springer 2024, forthcoming) https://ssrn.com/abstract=4696148. 1

[122] Gabriele Mazzini and Salvatore Scalzo, 'The Proposal for the Artificial Intelligence Act: Considerations around Some Key Concepts' in Carmelita Camardi (ed), *La Vie Europea per l'intelligenza artificiale* (Cedam 2022) 1.

private sector? Autonomous weapons systems? Self-driving vehicles and other cyber-physical machines? Medical devices and expert systems? Social scoring? Employment? Social welfare? Biometric identification/categorization systems? Credit Scoring? Life and health insurance? Algorithmic recommender systems used by platforms? AI-based contracts?

Moreover, there is little international consensus on the *how* of such a regulation*,* i.e. how to apply fundamental values such as human dignity and autonomy, fairness, transparency etc. in a given context. While global agreements in recent years – such as UNESCO's AI Recommendation and the OECD's AI Principles – recognize such fundamental principles, the ambiguity of these high-level agreements accommodates different political and ethical positions, allowing states to interpret them differently.[123] As *Roberts* and others point out: "Take AI fairness, a principle supported by all G20 member states, as applied to facial recognition technology. The implementation of this principle in the EU context involves the proposed banning of these technologies, while in China, ethnic-recognition technologies are permissible in the name of order and social stability".[124]

Even the adoption of a regulatory technique such as the risk-based approach, which appears to be politically neutral, involves fundamental policy choices: Does a country want to eliminate/ban (certain) risks, reduce them to an acceptable level, or strike a balance between risk reduction and the costs of regulation? If it is the latter: How does it balance risks and benefits in a specific sector and/or use-case?

Another major obstacle for foreign legislators to simply follow the EU's approach, is that the AI Act does not establish a comprehensive legal framework that can be adopted *tel quel*. Instead, it interacts in a very complex way with a rather sophisticated system of existing EU laws. In particular, the AI Act both complements existing product safety legislation[125] and builds, at the same time, on this legislation for the purpose of risk classification.[126] Moreover, the AI Act complements existing EU anti-discrimination law with specific requirements aimed at minimizing the risk of algorithmic discrimination. The AI Act further complements EU data protection law. This means that any processing of personal data by an AI system must comply with EU data protection law (e.g. the GDPR) and the AI Act. Last but not least, the Act provides for a number of future-proof instruments that will complement the Regulation, such as delegated and implemented acts, codes of practice and harmonized standards. As a result, it would not make sense to adopt the AI Act in isolation, as such a piece of legislation would be neither comprehensible nor meaningful without the rich body of existing EU law.

For all these reasons, it is unlikely that the AI Act will become the new global standard.

6.2. Why the AI Act *Shouldn't* Trigger a Brussels Effect

There are also important reasons as to why a *de iure* Brussels effect of the AI Act is not desirable.

First, we have to take into account that the economic, social, legal and political situation of countries are very different and, as a consequence, so are the ways in which countries and citizens are affected by AI.[127] Second, AI-specific regulation is still in its early stages, and it is unclear what the social and economic consequences of the AI Act will be. If it turns out that the AI Act has unforeseen significant

[123] Brent Mittelstadt, 'Principles alone cannot guarantee ethical AI' (2019) 1 Nature Machine Intelligence 501-507 https://doi.org/10.1038/s42256-019-0114-4503, 503.

[124] Huw Roberts, Emmie Hine, Mariarosaria Taddeo and Luciano Floridi, 'Global AI governance: barriers and pathways forward' (2024) 100(3) International Affairs 1275–1286 https://doi.org/10.1093/ia/iiae073, 8.

[125] Insofar as products using AI as a safety component must additionally comply with the specific requirements set out in the AI Act; cf. Art. 2(9) and Art. 6(1) AI Act.

[126] Cf. Art. 6(1)(b) AI Act.

[127] Smuha (n 120) 81.

negative effects, these will be duplicated around the world, with a Brussels effect.[128] Third, the existence of different AI regulations in different countries can – under the right conditions – stimulate experimentation and innovation in regulation through trial and error.[129]

Fourth, from a fundamental rights perspective, scholars rightly point out that the AI Act is the product of constitutional constraints.[130] Since the EU has no general competence to harmonize fundamental rights in the Member States, it relies instead on the competence to promote the internal market (Art. 114 TFEU). As a result, the AI Act does not follow a fundamental rights approach, but instead uses product safety law and risk regulation. However, such an approach is ill-suited to protect human rights. Given that most countries (or even regions) are not subject to the same competence constraints as the EU, they can (and should) use other approaches to address fundamental rights issues in relation to AI.

Finally, legislators around the world should take into account that key provisions of the AI Act do not follow a truly risk-based approach, particularly with respect to proper risk-*benefit* analysis, limited reliance on empirical evidence, pre-defined, closed risk categories, systemic risks of GPAI models, the overly broad definition of AI, double regulatory burdens, and overlapping enforcement structures. It is also *and above all* for these reasons that regulators outside the EU should not blindly follow the EU approach.

[128] Smuha (n 120) 80.
[129] Smuha (n 120) 69.
[130] Almada and Radu (n 45).