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

Ljupcho Grozdanovski,[*] Jérôme De Cooman[**]

Abstract

This article critically examines the inception of the recent European Commission (EC) proposal for a regulation laying down harmonization rules for Artificial Intelligence (AI Act). By establishing a four-level taxonomy of AI-related risks (non-high, limited, high, unacceptable) and corresponding technical standards, this instrument aims at preventing the occurrence of risks caused, in particular, by so-called high-risk AI systems. Though by virtue of its purpose and design, the AI Act follows a risk-based approach to regulation, it displays a specificity when compared to existing risk regulation in the European Union (EU), in such areas as environment and health. This specificity stems from the operative definition of risk the AI Act relies on: the risks covered in this proposal are not scientifically measurable threats of physical harm but threats of human rights violations which are difficult to quantify. In light of this, this article raises the issue of the evidence, if any, the EC gathered in view of designing a proportionate (*i.e.*, reality conform) regulatory framework on AI. Through a critical analysis of the discovery procedures (public consultations) the EC launched for the purpose of drafting the AI Act, this article finds that the Commission disregarded the evidence gathered on several important points, namely the type (horizontal or sectoral) of regulation best suited for AI-related risk prevention. Considering the limited impact of empirical knowledge on the design of the AI Act, this article seeks to determine if, in lieu of being a normative response to factual reality, the instrument under consideration gives specific normative

[*] Permanent Research Associate, Belgian Foundation for Scientific Research (FNRS), University of Liège.
[**] Ph.D Candidate and Research Assistant, University of Liege.

expression to fundamental values and rights, as well as to strategic objectives (trust and excellence) associated with the EU's Digital Single Market. This analysis allows to uncover two key standards having served as referents for the drafting of the AI Act: its axiological congruence with overarching Union values and its consistency with corresponding and already existing EU regulatory instruments. By delving into the origins of the AI Act, this article provides a unique insight into its normative rationale, uncovering the reasons for the relative obsolescence of empirical knowledge in its design and shedding more light on its specific legal nature as both a rights-protecting and risk-regulating instrument.

I. INTRODUCTION

In his 1992 *Risk society*,[1] Beck analysed the modes of production and distribution in a globalized economy, with scientific and technological progress as driving forces of overall social organization. He argued that post-industrial risk society is a concept "based on the importance of bads"[2] and is characterized by "the distribution of bads that flow within various territories and are not confined within the borders of a single society."[3] Artificial Intelligence (hereafter, AI),[4] as the latest offspring of technological innovation, has certainly triggered global debates on series of risks ('bads') that States and regional organizations like the European Union (hereafter, EU) were quick to identify and caution against. Unsurprisingly, regulation of AI became a point of focus of scholars[5] and regulators alike.[6]

Much like the technologies preceding AI,[7] regulators seemed confronted with a familiar 'the old meets the new' scenario, typically experienced when new technologies challenge the scope of application of existing regulatory instruments.[8] Indeed, realizing that current regulation was somewhat ill-adapted[9] for the resolution of -

---

[1] Ulrich Beck, RISK SOCIETY: TOWARDS A NEW MODERNITY (1992).
[2] Ulrich Beck, *Varieties of Second Modernity and the Cosmopolitan Vision*, 33 THEORY CULT. SOC. 257, 258 (2016).
[3] *Id.*
[4] The concept of AI will be defined at the end of this Introduction.
[5] *See* Karen Yeung, *Algorithmic Regulation: A Critical Interrogation*, 12 REGUL. GOV. 505, 505-506 (2018).
[6] Nathalie A. Smuha, *From a 'race to AI' to a 'race to AI regulation': Regulatory Competition for Artificial Intelligence*, 13 L., INNOVATION & TECH. 57, 59 (2021).
[7] *See generally* Lyria Bennett Moses, *The Applicability of Property Law in New Contexts: From Cells to Cyberspaces* 30 SYDNEY L. REV. 639 (2008).
[8] Monroe E. Price, *The Newness of New Technology*, 22 CARDOZO L. REV. 1885, 1889 (2001); Kieran Tranter, *Terrors in the Texts: Technology – Law – Future*, 13 L. & CRITIQUE 75, 76 (2002).
[9] In the field of discrimination, *see* namely Philipp Hacker, *Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies Against Algorithmic Discrimination under EU law*, 55 COMMON MKT. L. REV. 1143, 1145 (2018); *see generally* Sandra Wachter et al., *Why Fairness Cannot be Automated: Bridging the Gap Between EU Non-Discrimination Law and AI*, 41 COMPUT. LAW SEC.

what became - topical risks (e.g. algorithmic biases) associated with the deployment and use of intelligent systems, the need for AI-specific, tailor-made instruments naturally arose.

Unlike previous technologies, AI (as a class of intelligent rather than automated devices) raised never-seen-before challenges, as regulators sought to strike a balance between two competing objectives (and corresponding rationalities): market gains and a rights/values protection. The difficulty in balancing the two is namely due to the diversity and complexity of AI technologies. As stressed in a 2021 Expert Report on AI in Japan "on the one hand, laws and regulations face difficulties in keeping up with the speed and complexity of AI innovation and deployment (...) On the other hand, prescriptive regulation or rule-based regulation can hinder innovation. To address these conflicting problems, it is *necessary to change governance models from conventional rule-based ones to goal-based ones* that can guide entities such as companies to the value to be attained. Because our society shares the *Social Principles of Human-Centric AI, which state the goals for the use of AI, and because principles on AI are slowly but steadily reaching a consensus globally*, it can be said that we are finalizing the building of a foundation for goal-based governance."[10]

The cited Report reveals a key feature of what seems to have become the dominant *method* in AI regulation: since AI is ever-evolving and never fully knowable (in the sense of conclusive evidentiary discovery[11]), the most regulators can aim for are *broadly defined regulatory principles.* These would establish general frameworks within which subsequent and specific regulation could be enacted. While there has been much debate on which principles ought to serve for the setting of a 'golden standard' for AI regulation, one - though not the only[12] - possible candidate for a relatively

---

REV. 105567 (2021).

[10] AI Governance in Japan (1.1.), Report from the Expert Group on How AI Principles Should be Implemented, 20 (2021), www.meti.go.jp (emphasis added).

[11] The conclusiveness of the evidence of AI-related risks will be explored in more detail *infra*, Section III.

[12] *See inter alia* OECD AI Principles Overview, OECD.AI, www.oecd.ai/en/ai-principles (last visited Mar. 31, 2023).

complete list of such principles is that of *Asilomar*,[13] which groups those principles in three main clusters: research, ethics and long-term goals. The 'Research' cluster includes strategies and funding, science-policy links, the development of research cultures and so-called race avoidance.[14] The 'Ethics' cluster includes the principles of safety, failure transparency, judicial transparency, responsibility, value alignment, protection of human values, protection of personal privacy as well as liberty and privacy, the pursuit of shared benefits and prosperity, the preponderance of Human control, non-subversion of social checks schemes and the avoidance of AI Arms Race.[15] Finally, the 'Long-term goals' cluster includes AI capability caution, importance and impact of AI on various types of future global developments, careful identification and managing of risks, recursive self-improvement and the pursuit of the common good.[16]

---

[13] For an outline of these principles, *see* Our Mission, FUTURE OF LIFE INSTITUTE, www.futureoflife.org (last visited Mar. 31, 2023).

[14] This includes research goals and research funding in view of making future AI systems robust, grow prosperity and update legal systems so that they are fairer and more efficient, align AI with sets of ethical values. Implying a constructive and healthy dialogue between researchers and policy makers. Pertaining to culture of cooperation, trust and transparency for AI developers and the avoidance of short-cutting on safety standards.

[15] AI systems should be safe and security throughout their operational lifetime. Meaning that if an AI system causes harm, it should be possible to ascertain why and that any involvement of an autonomous system in judicial decision-making should provide satisfactory explanation auditable by a competent human authority. AI technologies should benefit and empower as many people as possible. The economic prosperity created by AI should be shared broadly, to benefit all humanity. Humans should choose how and whether to delegate decisions to AI systems. The power conferred by control of highly advanced AI systems should respect and improve the social and civic processes on which the health of society depends. An arms race in lethal autonomous weapons should be avoided.

[16] We should avoid strong assumptions regarding upper limits of AI capabilities. Advanced AI could represent a profound change in the history of life on Earth and should be planned and managed with commensurate care and resources. Risks posed by AI systems, especially catastrophic or existential risks, must be subject to planning and mitigation efforts commensurate with their expected impact. AI systems designed to recursively self-improve or self-replicate in a manner that could lead to rapidly increasing quality and quantity must be subject to strict safety and control measures. Superintelligence should only be developed in the

The principles and objectives highlighted in many national AI strategies are variations of the Asilomar principles, the general trend being the emphasis placed on humancentric, ethical values coupled with the fostering of economic efficiency and growth through, say, strategic investments and development of innovation.

This trend of framing ambitious market strategies with strong ethical values is *inter alia* visible in the 2018 German AI Strategy[17] as well as the Swiss Digital Strategy.[18] Alternatively, the 'leadership claim' in the so-called AI race (for markets), is especially strong in the AI strategy of the United States (hereafter, the US). The Executive order from 11 February 2019 [19] expresses the ambition of maintaining American leadership in AI through a concerted effort to promote advancement in technology and innovation "while protecting American technology, economic and national security, civil liberties, privacy, American values and enhancing international and industrial collaboration with foreign partners and allies."[20] To this end, the American strategy is guided by five principles: drive technological breakthroughs through the promotion of scientific discovery, economic competitiveness and national security; the development of appropriate technical standards and reducing the

---

service of widely shared ethical ideals, and for the benefit of all humanity rather than one state or organization.

[17] The Federal Government, Artificial Intelligence Strategy (2018),https://www.ki-strategie-deutschland.de/home.html?file=files/downloads/Nationale_KI-Strategie_engl.pdf; this Strategy outlines twelve fields of action *i.e.,* strengthening research in Germany; innovation competitions and European innovation clusters; transfer to business; fostering the founding of new businesses; shape structural changes in the labor market; using AI for tasks reserved for the state and administrative tasks; make data available and facilitate its use: adapt the regulatory framework; set standards; national and international networking; engage in dialogue with society and continue developing the framework of policy.

[18] Swiss Confederation, DIGITAL SWITZERLAND STRATEGY, www.digitaldialog.swiss (last visited Mar. 8, 2023). This Strategy outlines five key objectives *i.e.,* enabling equal participation for all and strengthening solidarity; guaranteeing security, trust and transparency; continuing to strengthen people's digital empowerment and self-determination; ensure value creation, growth and prosperity and reducing the environmental footprint and energy consumption

[19] Executive Order 13859, 84 Fed. Reg. 31, 3967 (Feb. 11, 2019).

[20] *Id.* at 3967.

barriers to the safe testing and deployment of AI technologies; train current and future generations of US workers with the skills to develop and apply AI technologies; foster public trust and confidence in AI technologies and protect civil liberties, privacy and American values; promote an international environment that supports American AI research and innovation and opens markets for American AI industries, while protecting technological advantage in AI and protecting US critical technologies from acquisition by strategic competitors and adversarial nations.[21]

The EU's regulatory trajectory follows the same general trends, since it also strives to strike the 'right' balance between risk-preventing principles of ethics without, at the same time, hindering the economic gains that AI innovation and use promise to deliver. In the 2020 White Paper on AI, the EC clearly supported a "regulatory and investment-oriented approach with the twin objectives of promoting the uptake of AI and addressing the risk associated."[22] One could argue that the White Paper provided the two-pronged

---

[21] *See id.* at 3867-68; to attain these principles, the Executive Order lists six strategic objectives namely: promote sustained investment in AI R&D in collaboration with industry, academia, international partners and allies and other non-Federal entities to generate technological breakthroughs in AI and related technologies and to rapidly transition those breakthroughs into capabilities that contribute to the US economic and national security; enhance access to high-quality and fully traceable Federal data, models and computing resources to increase the value of such resources for AI R&D, while maintaining safety, security, privacy and confidentiality protections consistent with applicable laws and policies; reduce barriers to the use of AI technologies to promote their innovative application while protecting American technology, economic and national security, civil liberties, privacy and values; ensure that technical standards minimize vulnerability to attacks from malicious actors and reflect Federal priorities for innovation, public trust, and public confidence in systems that use AI technologies and develop international standards to promote and protect those priorities; train the next generation of American AI researchers and users through apprenticeships, skills programs, and education in science, technology, engineering and mathematics, with an emphasis on computer science; develop and implement an action plan to protect the advantage of the US in AI and technology critical to the US economic and national security against strategic competitors and foreign adversaries.

[22] EURO. COMM'N, *White Paper on Artificial Intelligence: A European Approach to Excellence and Trust*, COM (2020) 65 final (Feb. 19, 2020).

regulatory framework within which subsequent EU regulation on AI would take shape. The driving ambition of the European Commission (hereafter, EC) is to realise, on the one hand, *an ecosystem of trust* ensuring "compliance with EU rules, including the rules protecting fundamental rights" and, on the other hand, *an ecosystem of excellence* that supports "the development and uptake of AI across the EU economy and public administration."[23] The latter objective is purely economic and aims at "harnessing the capacity of the EU to invest in next generation technologies and infrastructures."[24] The EC thus sought to increase "Europe's *technological sovereignty* in key enabling technologies and infrastructures for the data economy."[25] This twofold (excellence/trust) ecosystem echoes the Asilomar principles insofar as it places the emphasis on humancentric AI while aiming to foster investments and innovation. Similarly, the European Parliament (hereafter, EP) also acknowledged that AI systems "have the potential to generate opportunities for business and benefits for citizens" while simultaneously waving the need for a regulatory framework "protecting citizens from the potential risks of such technologies."[26] With the Proposal for an Artificial Intelligence Act (hereafter, AI Act), the Commission responded to that call emphasizing that AI "can provide key competitive advantages to companies and support socially and environmentally beneficial outcomes," but also "may generate risks and cause harm to public interests and rights that are protected by Union law."[27]

The interesting and often overlooked question - explored in this study - is the following: which *real-life experiences* (and evidence thereof) can be relied on to design the axiological and regulatory

---

[23] *Id.*

[24] *Id.*

[25] *Id.* (emphasis added).

[26] *European Parliament resolution of 20 October 2020 with recommendations to the Commission on a framework of ethical aspects of artificial intelligence, robotics and related technologies*, 2020/2012(INL), §§ B and H.

[27] EUR. COMM'N, 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, §§ 3 and 4, COM (2021) (Apr. 2, 2021).

shields that regulators should raise to protect their citizens from AI-related risks? Considering the current state and foreseeable development of AI innovation, establishing an *exhaustive* and *definitive list* of AI-related harms is next to impossible.[28] Engineers tell us that any type of - potentially harmful - malfunction of an AI system is due to a form of 'erroneous' (typically biased)[29] data and/or data processing. AI-related harms could thus be assessed under *existing* data protection instruments, like the General Data Protection Regulation (hereafter, the GDPR)[30] in the EU. Being that data is the 'raw material' that algorithms use to function,[31] any specific AI regulation would, in reality, give expression to the safeguards contained in general data protection frameworks which - like the GDPR - may *already* regulate automated decision-making.[32] What then is the added value of specific AI instruments? A plausible answer may be that, in a context of competing claims for global AI leadership, regulators seek to go beyond data protection by establishing a *taxonomy of risks* addressed by new, tailor-made AI regulatory frameworks.[33] For this, it is no longer enough to cast a wide net on instances of data processing gone wrong or to imagine

---

[28] As stated in *AI Governance in Japan (1.1.)*, *supra* note 12 at 29; The AI-related risks cannot be represented in a definitive way, because there are endless possible directions in which AI innovation can evolve entailing endless possible harms that future AI systems may cause.

[29] *See inter alia* Tobias Baer, UNDERSTAND, MANAGE AND PREVENT ALGORITHMIC BIAS. A GUIDE FOR BUSINESS USERS AND DATA SCIENTISTS 70 (2019).

[30] Council Directive 2016/679, art. 22, 2016 (L119) 1 (EU).

[31] *See* Cedric Villani, For a Meaningful Artificial Intelligence. Towards a French and European Strategy 4 (Yann Bonnet *et al.* eds., 2018), https://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf.

[32] *See* Council Directive 2016/679, *supra* note 32; *see also* Melanie Fink and Michèle Finck, *Reasoned A(I)dministration: Explanation Requirements in EU Law and the Automation of Public Administration* 3 *EU L. Rev.* 376, 387 (2022), "given that EU administrative law is applicable across the board, including when AI systems are used, secondary law that establishes specific explanation rights for automated technologies is not absolutely necessary".

[33] *Id.* at 387-88, "secondary law could play an important role in specifying the exact extent to which the public authority's reasoning may rely on an AI system's recommendation and what that means for the degree of explainability required from the system".

all the possible ways in which AI threatens to violate fundamental rights. Regulators would rather need to engage in uncovering *fact-of-the-matter knowledge* for the purpose of designing regulatory responses that would correspond to - because they would accurately apprehend - the *actual* risks in AI programming and use. In short, Asimov's work might be a good reference for making the noble pledge to protect human dignity in the face of AI. However, translating this pledge into a concrete protective scheme against harms that, say, facial recognition systems are likely to cause requires more than futuristic fiction;[34] it requires *evidence.*

The problem of hypothesizing risks rather than seeking to establish them was mentioned, namely, in the Australian Report on the AI Standards Roadmap which cites Australian stakeholders' preference for standards able to "*respond to the real and perceived risks* that might arise in relation to AI."[35] Similarly, a 2020 Memorandum for the Heads of US Executive Departments and Agencies stresses that "regulatory and non-regulatory approaches to AI *should be based on a consistent application of risk assessment and risk management across various agencies and various technologies.*[36] Regarding the extent of the risk assessment, the Memorandum stresses that "*it is not necessary to mitigate every foreseeable risk; in fact, a foundational principle of regulatory policy is that all activities involve tradeoffs.* Instead, *a risk-based approach should be used to determine which risks are acceptable and which risks present the possibility of unacceptable harm, or harm that has expected costs greater than expected benefits.*"[37]

More generally, the requirement for knowledge of facts, derived from evidence, has been increasingly seen as contributing to the

---

[34] *See generally* Jerome De Cooman & Nicolas Petit, *Asimov for Lawmakers*, 18 J. BUS. & TECH. L. 1 (2022).

[35] Standards Australia, *Final Report: An Artificial Intelligence Standards Roadmap: Making Australia's Voice Heard*, 35 (2022) https://www.standards.org.au/getmedia/ede81912-55a2-4d8e-849f-9844993c3b9d/1515-An-Artificial-Intelligence-Standards-Roadmap12-02-2020.pdf.aspx.

[36] Memorandum for the Heads of Executive Departments and Agencies, M-21-06 (Nov. 17, 2020) at 4 (emphasis added).

[37] *Id.* (emphasis added).

transparency - and, by that, the intelligibility and justiciability - of any type of regulation, not only that focused on risks. In the EU, the Better Regulation Agenda stresses that the new guidelines on regulation ought to "re-commit the Commission to use the *best available evidence* and *science* and reinforce the commitment to put in place clear monitoring and implementation plans before measures are adopted."[38] In this context, one might expect that prior to drafting various instruments, the EU regulators might consider undertaking discovery procedures in order to gain a better understanding of the nature and scope of the real-life phenomena they intend to address. However, as will be argued,[39] the knowledge-acquiring procedures followed in view of 'tailoring' empirically adequate (*i.e.*, reality-conform)[40] regulatory responses to AI did generally not include any 'hard' scientific evidence or statistical measurements of AI functionalities. Rather, regulators who undertook discovery (or discovery-akin)[41] enterprises, sought to collect the *views* and *attitudes* of selected stakeholder groups. The choice of this type of evidence is understandable, given that AI-related risks are open-ended, making *conclusive scientific* knowledge thereof difficult to acquire. However, the preference for 'subjective' evidence is also open to criticism given that the risk with such risk-assessments - excuse the pun - is that they may lead to regulatory instruments based on erroneous (typically, exaggerated) perceptions of specific risks.[42]

---

[38] Better regulation for better results - An EU agenda, at 5, COM (May 19, 2015).
[39] *See infra*, Section III at 31.
[40] *See id.* at 32; in the remainder of this study, the expression 'empirically adequate' will be understood as 'reality-conform' *i.e.,* as a representation of a portion of reality (risks in particular) that does is not purely subjective, but that portrays a phenomenon based on opinion- and belief-independent measurements.
[41] *See id.*; the distinction between 'discovery' and 'discovery-akin' alludes to a difference between discovery through the use of scientific measurements and methods (traditionally viewed as discovery proper) and 'mere' stake-holder consultations (which, under the historic *penchant* for unbiased scientific facts, would qualify as discovery *faute de mieux*). However, in the remainder of this study, the expression 'discovery procedure' will be used to designate both scientifically induced and perceptions-induced knowledge.
[42] *See e.g.* Melvin Kranzberg, *Technology and History: 'Kranzberg's Law'*, 15 BULL. SCI. TECH. SOC. 5, 6 (1995).

Considering the regulatory importance of some form of *objective knowledge* of AI-related risks - however open-ended they may be - we will, from the perspective of methodology, follow two approaches. First, and bearing in mind that the AI Act is a risk regulating instrument after all, we will examine the *epistemic validity*[43] of both the discovery procedures launched by the EC and the interpretations of the findings of those procedures. To conduct this analysis, our analytical framework will include *evidence theory*, as transposed into risk theory. This methodological choice is justified by the fact that knowledge of facts in the context of risk-assessment is construed through the application of epistemic principles that originate from, and have been consolidated in the over two-centuries long tradition of 'modern' evidence scholarship.[44] With the development of practice as well as scholarship on risk regulation, a number of those principles (like relevance[45] and the best evidence rule[46]) became important epistemic referents for the unfolding of the two key stages in any risk assessment *i.e.,* risk-*identification* and risk-*characterization*. The former refers to the process(es) followed in identifying the *locus* of uncertainty (*i.e.,* the answering of the '*risk of what?*' question).[47] This is fundamentally an issue of identifying the risks that warrant further discovery

---

[43] The concept of epistemic validity will be further explored; *see infra*, Section III at 23.

[44] *See* Paul Roberts, *Adrian Zuckerman's New Evidence Scholarship*, *in* PRINCIPLES, PROCEDURE, AND JUSTICE: ESSAY IN HONOR OF ADRIAN ZUCKERMAN 59, 70 (Rabeea Assy & Andrew Higgins eds., 2020).

[45] The issue of relevance will be discussed in greater detail, *see infra*, Section III.

[46] The best evidence rule, as applied in risk theory, will be discussed in greater detail *see infra*, Section IV.

[47] This is a 'natural' preliminary stage of chief epistemic importance, since "it is the sake for which we seek truths - the project in which inquiry is necessarily embedded - from which standards of accuracy must be derived." *See* Filip Buekens and Fred Truyen, *The Truth About Accuracy*, *in* EXPERTS AND CONSENSUS IN SOCIAL SCIENCE 212, 221 (Carlo Martini & Marcel Boumans eds., 2014), emphasis added; *see also* Henry Rothstein et al., *The Risks of Risk-Based Regulation: Insights from the Environmental Policy Domain*, 32 ENV'T. INT'L. 1056, 1061 (2006) noting that risk-based approaches "pose a number of epistemic challenges" and that "regulation often deals with issues at the horizons of human knowledge, so there is considerable scope for risk assessments to give false impressions of accuracy or create conditions for regulatory conflict".

through careful selection of both the causes of risk and the relevant scientific data. The latter is essentially a process of sense-making, translating to inferences based on the collected data meant to represent as accurately as possible the causal relationship between a risk and a given harm.

Against the backdrop of evidence theory as our analytical framework, this study will, second, critically assess if in the drafting of the AI Act, the EC: 1. properly identified the AI-related risks ultimately included in the AI Act; 2. properly gathered, interpreted and translated into policy the findings on those risks. More specifically, we will seek to determine how the EC decided on the *relevance* of AI-related risks deserving of further exploration and possible regulatory address. The role and application of the criterion of relevance in the process of risk *identification* will be the point of focus in Section III.

Section IV will examine the conditions that need to be met in order for knowledge of risks to qualify as epistemically *valid* and legitimately warrant a regulatory response. This will allow us to gain a better understanding of the role that evidence usually plays in the shaping of risk-regulation and explain why, in the AI Act, empirical knowledge has a secondary role, this instrument having been chiefly guided by a *desired standard* of protection against fundamental rights violations. Bearing in mind this disregard of evidence, we will raise, in Section V, the issue of whether the AI Act constitutes a policy response based on *adequate* risk representations and is compliant with the principle of proportionality. Section VI will include our main concluding observations.

However, prior to conducting our analysis, it is methodologically necessary to provide definitions of this study's operative concepts,[48] as well as outline, in Section II, the evolution and fine-tuning of the EU's regulatory approach to AI.

Our first operative concept is, naturally, AI. Though it is difficult to provide a one-size-fits-all definition, we shall stress that, in the EU, the EC had - in its Artificial Intelligence for Europe

[48] We call 'operative concepts' those that establish the basic conceptual framework within which our analysis can take shape.

Communication[49] - distinguished between AI as a class of new technologies (or systems) and AI as a field of research. According to the EC, AI *systems* "display intelligent behaviour by analysing their environment and taking actions - with some degree of autonomy - to achieve specific goals."[50] These systems are "purely software-based, acting in the virtual world (e.g. voice assistants, image analysis software, search engines, speech and face recognition systems) or AI can be embedded in hardware devices (e.g. advanced robots, autonomous cars, drones or Internet of Things applications)."[51] AI *as a discipline* includes - according to Annex I of the AI Act - "(a) [m]achine learning approaches, including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning, (b) [l]ogic- and knowledge-based approaches, including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems, [and] (c) [s]tatistical approaches, Bayesian estimation, search and optimization methods."[52] It is interesting to note that the AI Act defines AI *systems* in relation to its definition of AI *as a discipline*. Consequently, those systems are software developed with one or more of the mentioned techniques and approaches that can "for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with."[53] Our references to AI systems in the remainder of this study will be understood within the meaning of the definition provided in the AI Act.

The second operative concept in this study is that of *risk*. Here

---

[49] Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions, Artificial Intelligence for Europe, at 1-2, COM (2018) 237 final (Apr. 25, 2018).

[50] *Id.* at 1.

[51] *Id.*

[52] European Commission, Annexes to the Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonized Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts, at Annex, COM (2021) 206 final (updated Nov. 25, 2022).

[53] *Id.*

again, many plausible definitions can be offered.[54] Möller e.g. gave five meanings: risks as an unwanted event which may or may not occur; the cause of an unwanted event which may or may not occur; the probability of an unwanted event which may or may not occur; the fact that a decision is made under conditions of known probabilities; the statistical expectation value of unwanted events which may or may not occur.[55] Wynne defined risks through the interrelationship between that which is known and that which is not. Following this thread of (un)knowability, risks can be understood as instances in which some parameters of a phenomenon are known, but not the probability distributions.[56] Indeed, depending on the aspects of risks where uncertainty lies, we might know the cause (the generating fact) of a specific risk, but may be able to neither fully measure the probability of that cause actualizing, nor the effects it might have, if and when it actualizes. The specific attributes of risks which distinguish them from the broader class of so-called unknowns will be explored in more detail further in this article.[57] At this stage, we shall stress two key approaches in defining risks. On the one hand, from an objective (*i.e.,* belief- and perception-independent) perspective, risks are synonymous with measurable and quantifiable threats of harm or dangers that upset a pre-established level of safety.[58] On the other hand, from a subjective, perception-based view, risks are specific modes of treatment of certain events capable of happening to a group of individuals. Under this view, risks are not naturally occurring events, but events designated as such by virtue of our understanding.[59]

---

[54] Bob Heyman et al., *The Concept of Risk, in* RISK, SAFETY AND CLINICAL PRACTICES: HEALTH CARE THROUGH THE LENS OF RISK 15, 16 (2009).
[55] Niklas Möller, *The Concepts of Risk and Safety, in* HANDBOOK OF RISK THEORY. EPISTEMOLOGY, DECISION THEORY, ETHICS, AND SOCIAL IMPLICATIONS OF RISK 56, 58 (Sabine Roeser et al. eds., 2012).
[56] Brian Wynne, *Uncertainty and Environmental Learning: Reconceiving Science and Policy in the Preventive Paradigm*, 2 GLOB. ENV'T CHANGE 111, 114 (1992).
[57] *See infra*, Section III.
[58] Möller, *supra* note 57 at 61.
[59] François Ewald, *Insurance and Risk, in* THE FOUCAULT EFFECT 197, 199 (Graham Burchell et al. eds., 1991).

Though these two approaches might seem at odds they are, in truth, complementary in the sense that for any kind of threat to qualify as risk, it ought to be perceived as such, and this perception should - ideally - be corroborated by qualitative and quantitative data produced through commonly accepted discovery procedures and methods. As empirically explorable and measurable phenomena - unlike 'ordinary' uncertainties - risks have always attracted attention because, when they are backed by evidence, they *can shape policy*. This is an important point to keep in mind since, in the remainder of this study, the notion of risks will be primarily understood in relation to *provability* (the question being: how do we go about proving unknowable phenomena like risks?). In this context, we will define risks as actual or perceived threats the veracity[60] of which derives from evidence acquired through established protocols and procedures of discovery. This definition naturally requires that we clarify our understanding of the concept of *evidence*.

Like the previous two concepts, evidence is difficult to concisely define. In its original meaning, it translates to the establishment of factual truths for the purpose of dispute resolution. This primary, adjudicatory function of evidence is *inter alia* visible in Wigmore's work. For him, evidence is any knowable fact or group of facts, considered with a view of being presented - typically before a court - for the purpose of making a claim on the truth of a proposition.[61] This definition contains two key elements that we will refer to in the remainder of this study. First, *the knowability of facts*. This is an important point to keep in mind since the key feature of risks is the total or partial unknowability of some of their properties. Hence, the main evidentiary difficulty is that of defining the conditions under which *actual* (*i.e., empirical* proper) as opposed to *hypothetical* knowledge of risks can be derived. This ties into the second takeaway from Wigmore's definition: the fundamental aspiration of evidence is *truth*. The concept of truth will be understood in a narrow

---

[60] The term veracity here is used in a probabilistic sense (*verisimile*) *i.e.,* as a truth- or accuracy-quality of a statement, derived from reliable but inherently imprecise data; *see* Marianela Garcia Lozano et al., *Veracity Assessment of Online Data*, 129 DECISION SUPPORT SYS. 1, 2 (2020).

[61] John H. Wigmore, *Evidence in Trials at Common Law*, 4th ed., vol. 11 (1961).

sense: not as knowledge of some essential, ontological properties of the metaphysical realm, but as *accurate* (*i.e.*, logically infallible) accounts of tangible, fact-of-the-matter experiences. The criteria that must be met for this type of 'empirical' knowledge to be formed have been the point of focus of the so-called New Evidence Scholarship,[62] characterized by an increased interest in the reasoning models followed by parties and courts when deciding if something is proven and can therefore *be held as true*. This rationalist - more so than procedural - perspective on evidence can be said to have had the greatest epistemic impact on risk theory. After all, "risk research is principally an empirical field of study, but if we are dealing with inadequate conceptualizations, our analysis may be inadequate too."[63] In this context, risk-assessment necessarily includes some type of discovery meant to produce *accurate* knowledge which, like legal evidence, is both *fact-dependent* and *rationally construed*. Hence our use of evidence theory as referent for our analysis of risk identification and characterization in the context of AI regulation.

The fourth - and last - operative concept in this study is that of *regulation*. Despite the rich and varied scholarship, regulation "appears to escape a clear definition."[64] This is due to the relative lack of clarity on "the abstract concept of regulation" and the "opinions about the desirable scope of regulatory powers or desirable regulatory policies."[65] Black, as one of the leading scholars in the field, has explained that "definitional chaos is almost seen as an occupational hazard by those who write about regulation."[66] Linguistic and comparative law approaches do not help as "regulation is not a concept that travels well."[67] Non-English-

---

[62] *See* John D. Jackson, *Analysing the New Evidence Scholarship: Towards a New Conception of the Law of Evidence*, 16 OXFORD J. LEGAL. STUD. 309, 309-311 (1996).

[63] Möller, *supra* note 57 at 62.

[64] Barak Orbach, *What is Regulation,* 30 YALE J. REGUL ONLINE 1, 2 (2012).

[65] *Id.* at 2-3.

[66] Julia Black, *Decentring Regulation: Understanding the Role of Regulation and Self-Regulation in a 'Post-Regulatory' World*, 54 CURRENT LEGAL PROBS. 103, 129 (2014).

[67] Julia Black, *Critical Reflections on Regulation*, 27 AUST. J. LEG. PHILOS. 1, 2

speaking systems often use different terms with different connotations and meanings.[68] French legal scholars use *règlement* and *réglementation* as variations of the generic term 'regulation'. The same goes for Dutch whereby regulation is translated as *verordening* (for instance, an EU Regulation) and as *regeling* or *regelgeving* (with a much broader meaning and closer to the English word 'regulation').

Defining regulation is therefore "by no means a simple matter" due to "the different levels of generality at which the term is used."[69] Yeung has similarly highlighted that "the term 'regulation' has been used to encompass a variety of different conceptions."[70] Jordana and Levi-Faur have, in turn, explained that "the various definitions of regulation reflect specific disciplinary concerns, are oriented towards different research methods, and reflect to a significant extent the unique personal, national and historical experience of the formulator of the definition."[71] This threefold definition shows why regulation is "one of the most controversial topics in law and politics."[72]

One of the reasons why regulation is a confusing polysemic concept is its scope: it includes acts of "controlling, directing, or governing according to a rule, principle or system"[73] and "measures which express such command and control arrangements."[74] Economic regulation is a good example to dive deeper in this large spectrum of meanings of the concept under consideration, as the vast majority of regulatory scholarship equates the target of regulation with economic actors.[75] Legal scholarship usually identifies four regulatory levels, *i.e.*, regulation through governmental agencies or

---

(2002).

[68] *Id.*

[69] Tony Prosser, LAW AND THE REGULATORS 4 (1997).

[70] Karen Yeung, SECURING COMPLIANCE: A PRINCIPLED APPROACH 5 (2004).

[71] Jacint Jordana & David Levi-Faur, *The Politics of Regulation in the Age of Governance*, *in* THE POLITICS OF REGULATION: INSTITUTIONS AND REGULATORY REFORMS FOR THE AGE OF GOVERNANCE 1, 3 (Jacint Jordana & David Levi-Faur eds., 2004).

[72] Orbach, *supra* note 66 at 2.

[73] Prosser, *supra* note 71 at 4.

[74] *Id.*

[75] Anthony Ogus, REGULATION: LEGAL FORM AND ECONOMIC THEORY (1994).

commissions,[76] state intentional act of control,[77] State and non-State intentional act of control,[78] and State and non-State unintentional act of control.[79] In addition, given the prevalence of risk within society,[80] one might go as far as considering that "any governmental interference with market or social processes" in an attempt to control potential adverse consequences can qualify as risk regulation.[81] After all, if one looks hard enough, one can always detect some type of risk explicitly or implicitly[82] impacting the design of various normative frameworks.[83] A too broad definition of regulation, however, "runs into the danger of being so broad that it contributes

---

[76] Gunther Teubner, *After Legal Instrumentalism: Strategic Models of Post-Regulatory Law*, *in* DILEMMAS OF LAW IN THE WELFARE STATE 299 (Gunther Teubner ed., 1986); *see e.g.* Giandomenico Majone, *The Rise of the Regulatory State in Europe*, 17 WEST EUR. POLIT. 77 (1994); Giandomenico Majone, *From the Positive to the Regulatory State: Causes and Consequences of Changes in the Mode of Governance*, 17 J. PUB. POL'Y 139 (1997); Giandomenico Majone, *The Regulatory State and its Legitimacy Problems*, 22 WEST EUR. POLIT. 1 (1999); Black, *Critical Reflections on Regulation*, *supra* note 69 at 30; Jordana & Levi-Faur, *supra* 73 note at 4; REGULATION THROUGH AGENCIES IN THE EU: A NEW PARADIGM OF EUROPEAN GOVERNANCE (Damien Geradin et al. eds., 2006).

[77] *See generally* Robert Baldwin et al., *Introduction*, *in* A READER ON REGULATION 1, 3 (Robert Baldwin et al., 1998); Christopher Hood et al., THE GOVERNMENT OF RISK: UNDERSTANDING RISK REGULATION REGIMES (2nd ed. 2004); *see generally* Julia Black, *Enrolling Actors in Regulatory Systems: Examples from UK financial Services Regulation*, PUB. L. 69 (2003); Jordana & Levi-Faur, *supra* note 73; Roger Brownsword, *Code, Control, and Choice: Why East is East and West is West*, 25 LEG. STUD. 1, 6 (2005); *see generally* Bronwen Morgan & Karen Yeung, AN INTRODUCTION TO LAW AND REGULATION: TEXT AND MATERIALS (2007).

[78] *See generally* Julia Black, *Constitutionalising Self-Regulation*, 59 MOD. L. REV. 24 (1996); John Braithwaite, *The New Regulatory State and the Transformation of Criminology*, 40 BR. J. CRIMINOLOGY 222, 224 (2000).

[79] *See generally* Baldwin et al., *supra* note 79 at 4; Lawrence Lessig, CODE: VERSION 2.0 130 (2006).

[80] Beck, *supra* note 3.

[81] Hood et al., *supra* note 79 at 3; Robert Baldwin et al., UNDERSTANDING REGULATION: THEORY, STRATEGY AND PRACTICE (2011).

[82] *See generally* Lisa Heinzerling & Mark V. Tushnet, THE REGULATORY AND ADMINISTRATIVE STATE (2006).

[83] *See generally* Julia Black, *The Role of Risk in Regulatory Processes*, *in* THE OXFORD HANDBOOK OF REGULATION 304 (Robert Baldwin et al. eds., 2010).

to nothing."[84] A wide-ranging definition, indeed, does not usefully inform us on the criteria against which we can properly qualify the type of 'risk regulation' that is embodied in the AI Act.[85] In light of the above, this article will retain Yeung's definition of regulation as "the sustained and focused attempt by the state to alter behaviour thought to be of value to the community."[86] This definition will be used as it seems to best capture the essence of regulation, *i.e.*, "how to alter behaviour so that people act in a way that they would not otherwise do."[87] This view is confirmed in Orbach's work where regulation is defined as a "government intervention in the private domain or a legal rule that implements such intervention [which is] a binding legal norm created by a state organ that intends to shape the conduct of individuals and firms."[88] This definition is akin to Foucauldian governmentality[89] and offers an understanding of regulation as defining "the conduct of conducts."[90] As such, a cybernetic view of regulation seems to be most helpful to understand what it actually entails. A regulation is a control system cumulatively able to distinguish between preferred and non-preferred states (*standard-setting*), to monitor current and changing states of the system (*information-gathering*), and to change the state of the system to reach the preferred one (*behaviour-modification*).[91]

---

[84] Black, *supra* note 80 at 23.
[85] *See* discussion discussed *infra*, Section IV (The concept of risk-regulation as applied in the AI Act will be discussed).
[86] Yeung, *supra* note 72 at 5.
[87] Morgan & Yeung, *supra* note 79 at 6; *see generally* Black, *supra* note 80 at 69.; Hood et al., *supra* note 79.
[88] Orbach, *supra* note 66 at 6.
[89] Michel Foucault, *Security, Territory, Population* (2004); Mitchell M. Dean, GOVERNMENTALITY: POWER AND RULE IN MODERN SOCIETY (2nd ed., 2010).
[90] Michel Foucault, *Governmentality*, *in* THE FOUCAULT EFFECT: STUDIES IN GOVERNMENTALITY 2, 2 (Graham Burchell et al., eds., 1991), cited in Black, *supra* note 68 at 108.
[91] Hood & Yeung, *supra* note 79 at 22-23.

II. THE EVOLUTION OF THE EU'S REGULATORY APPROACH TO AI

The fine-tuning of the EU's regulatory approach to AI was a gradual, several-year process. The Union institutions gave various incentives - mainly in the form of soft law instruments - that invited the EC to submit a proposal for a regulatory framework on AI (**A**). The culminating point of this progression came with the AI Act proposal in 2021, in view of which the EC had undertaken discovery procedures aimed at gathering evidence on the types of AI-related risks to be included in said proposal. However, an overview of the latter shows a dissonance - explored further in this article - between the risks as evidenced by the discovery procedures and those the EC ultimately chose to give normative translation to (**B**).

A. The Institutional Incentives for AI Regulation

The AI Act can be seen as the result of a several-year process during which the EU institutions tailored the Union's overall regulatory approach to AI. In his 2014 Political Guidelines for the EC, Jean-Claude Juncker stated that "we must take much better use of the great opportunities offered by digital technologies, which know no borders."[92] Juncker's ambition was echoed in the Digital Single Market Strategy for Europe,[93] focused on the creation of "secure and trustworthy infrastructures and content services" that would support digital networks and services.[94] This Strategy viewed cyberthreats as "a borderless problem" and emphasized the need for a reinforcement of security and robustness.

In May 2017, the EC published its mid-term review of the implementation of the Digital Single Market Strategy. In the review,

---

[92] Jean-Claude Juncker, *A New Start for Europe: My Agenda for Jobs, Growth, Fairness and Democratic Change*, EURO. COMM'N, (July 15, 2014), https://ec.europa.eu/commission/presscorner/detail/en/SPEECH_14_546.

[93] *See* Communication from the Commission to the European Parliament, The Council, the European Economic and Social Committee and the Committee of the Regions: A Digital Single Market Strategy for Europe, COM (2015) 192 final.

[94] *Id.* at 3, 12.

the EC argued, first, that a European Data Economy will be achieved through cross-border free flow of personal and non-personal data. Second, regarding online security, the EC stressed that the importance of the 2016 Directive on Security of Network and Information Systems (NIS Directive),[95] as well as of the EU Agency for Network and Information Security (ENISA) which oversees European cybersecurity. Third, considering that digital platforms are "key gatekeepers of the Internet"[96] the EC's aim was to promote those platforms as "responsible players of a fair Internet ecosystem," but also highlighted that they were sometimes engaged in anticompetitive trading practice such as the delisting of their professional users' products or services without noticing and without offering the possibility to contest.[97] The EC also mentioned - albeit briefly - AI technologies, confirming its intention to "continue to monitor the opportunities and challenges brought by artificial intelligence solutions."[98]

In parallel, both the EP and the European Council invited the EC to elaborate and present a European approach towards AI. The European Council stressed the urgency to address issues raised by AI technologies "while at the same time ensuring a high level of data protection, digital rights and ethical standards."[99] Similarly, the EP stated that "the Union could play an essential role in establishing basic ethical principles to be respected in the development, programming and use of robots and AI" and proposed their

---

[95] Directive 2016/1148 of the European Parliament and of the Council of 6 July 2016 concerning measures for a high common level of security of network and information system across the Union, OJ L 194/2016, 1-30.

[96] Communication from the Commission to the European Parliament, the Council, The European Economic and Social Committee and the Committee of the Regions on the Mid-Term Review on the implementation of the Digital Single Market Strategy: A Connected Digital Single Market for All, at 7, COM (2017) 228 final (Oct. 5, 2017).

[97] *Id.*; the Commission added that "there is widespread concern that some platforms may favour their own products or services, otherwise discriminate between different suppliers and sellers and restrict access to, and the use of, personal and non-personal data, including that which is directly generated by a company's activities on the platforms."

[98] *Id.*

[99] European Council Meeting, 19 October 2017, CO EUR 7/CONCL 5.

incorporation "into Union regulations and codes of conduct, with the aim of shaping the technological revolution."[100] The EP expressed a preference for regulation by design[101] and highlighted the importance of the transparency principle as implying the possibility "to supply the rationale behind a decision taken with the aid of AI that can have a substantive impact on one or more persons' lives."[102] The EP also pointed out that "the guiding ethical framework should be based on the principles of beneficence, non-maleficence, autonomy and justice," as well as other European values, enshrined in the Treaty on European Union (hereafter, TEU) and in the EU Charter of Fundamental Rights (hereafter, ECFR).[103]

In response to said recommendations, the EC presented its 2018 *AI for Europe* plan[104] which pursued three objectives: boosting the EU's technological and industrial capacity and AI uptake across the economy, preparing to socio-economic changes driven by AI and ensuring an appropriate ethical and legal framework. On the financial side, the EC aimed at increasing research and innovation investments by supporting AI-focused research excellence centres across Europe and using Digital Innovation Hubs to bring AI systems to small and medium sized enterprises (hereafter, SME). The Commission also drew a parallel between vast amounts of data and significant advances in AI, especially Machine Learning (hereafter, ML) and deep learning (hereafter, DL), recognizing that "access to data is a key ingredient for a competitive AI landscape."[105] On the socioeconomic side, the EC stressed the importance of developing basic digital skills, of workers' replacement due to increased labour automation, and of training AI experts while attracting more talent to the Union. However, across the board, the

---

[100] Eur. Parl., Report with Recommendations to the Commission on Civil Law Rules on Robotics, 2015/2103(INL), §V.

[101] *Id.* at §O (noting that "the developments in robotics and AI can and should be designed in such a way that they preserve the dignity, autonomy and self-determination of the individual").

[102] *Id.* at §12.

[103] *Id.* at §13.

[104] COM (2018) 237 final, *supra* note 70.

[105] *Id.*

ethical issues of AI and their (proper) translation into law seem to have been viewed as the most important.

Recognizing that a thorough analysis of the impact of the development of AI on fundamental rights could not be achieved without the EU Member States (hereafter, MS) joining forces, a declaration of cooperation was signed in April 2018.[106] The signatory MS drafted a Coordinated Plan on AI,[107] which was published in December 2018.[108] In addition to the invitation to the MS to continue their national AI strategies, the EC reassessed the importance of funding research toward public-private partnerships and financing for start-ups and SMEs, of strengthening excellence in AI technology, and of improving both basic and high-level digital skills. A noteworthy point of said Plan is the development of "ethics guidelines with a global perspective and ensuring an innovation-friendly legal framework."[109] This paved the way for the creation of EC's independent high-level expert group on AI (HLEG) comprised of experts in the fields of computer science, law and philosophy.[110] The approach of setting up *ad hoc* advisory expert groups is not without precedent, considering the EC's prior experience in establishing such entities[111] on issues like climate change,[112]

---

[106] Declaration of Cooperation (Apr. 10, 2018) https://wayback.archive-it.org/12090/20180620090945/https://ec.europa.eu/digital-single-market/en/news/eu-member-states-sign-cooperate-artificial-intelligence.

[107] These are Austria, Belgium, Bulgaria, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, Poland, Portugal, Slovakia, Slovenia, Spain, Sweden, the United Kingdom, Norway, Romania, Croatia, Greece and Cyprus.

[108] Communication from the Commission to the European Parliament, the European Council, The Council, the European Economic and Social Committee and the Committee of the Regions: Coordinated plan on Artificial Intelligence, COM (2018) 795 final.

[109] *Id.*

[110] The list of these experts is available at European Commission, *High-Level Expert Group on Artificial Intelligence* https://digital-strategy.ec.europa.eu/en/policies/expert-group-ai (last visited Mar. 31, 2023).

[111] The list of all expert groups is available at: *Register of Commission Expert Groups*, EUR. COMM'N, https://ec.europa.eu/transparency/expert-groups-register/screen/home?do=news.news (last visited Mar. 31, 2023).

[112] *See Commission expert group: Mission Board for adaptation to climate change including societal transformation (E03664)*, EUR. COMM'N,

mobility[113] and cancer research.[114]

The HLEG met for the first time in June 2018[115] and published Draft Ethics Guidelines in December of the same year.[116] It proposed a *human-centric approach* that balanced benefits and risks driven by AI systems. After a public consultation the results of which were published in February 2019,[117] the Ethics Guidelines were published in April 2019[118] and were endorsed by the EC in its plan to build trust in human-centric AI.[119] The Guidelines propose a four-pillar framework of principles which includes: prevention of harm (1), respect of human autonomy (2), fairness (3) and explicability (4). In 2019, the HLEG published their policy and investment recommendations for trustworthy AI[120] followed by sectoral

---

https://ec.europa.eu/transparency/expert-groups-register/screen/expert-groups/consult?lang=en&do=groupDetail.groupDetail&groupID=3664 (last visited Mar. 31, 2023).

[113] *See Group of Experts on the Smart Tachograph (E03663)*, EUR. COMM'N, https://ec.europa.eu/transparency/expert-groups-register/screen/expert-groups/consult?lang=en&do=groupDetail.groupDetail&groupID=3663&news=1(last visited Mar. 31, 2023).

[114] *See Commission expert group: Mission Board for cancer (E03665)*, EUR. COMM'N, https://ec.europa.eu/transparency/expert-groups-register/screen/expert-groups/consult?lang=en&do=groupDetail.groupDetail&groupID=3665 (last visited Mar. 31, 2023).

[115] *High-Level Expert Group on Artificial Intelligence (E03591), Group Details, Meetings*, EUROPEAN COMMISSION, https://ec.europa.eu/transparency/expert-groups-register/screen/expert-groups/consult?lang=en&do=groupDetail.groupDetail&groupID=3591 (last visited Mar. 31, 2023).

[116] *Draft Ethics Guidelines for Trustworthy AI*, EUR. COMM'N (Dec. 18, 2018), https://digital-strategy.ec.europa.eu/en/library/draft-ethics-guidelines-trustworthy-ai.

[117] *Stakeholder Consultation on Guidelines' First Draft,* EUR. COMM'N (Feb. 19, 2019), https://ec.europa.eu/futurium/en/ethics-guidelines-trustworthy-ai/stakeholder-consultation-guidelines-first-draft.

[118] *Ethics Guidelines for Trustworthy AI*, EUR. COMM'N (Apr. 8, 2019), https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai.

[119] European Commission, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: Building Trust in Human-Centric Artificial Intelligence, COM (2019)168 final, 8 April, 2019.

[120] *Policy and Investment Recommendations for Trustworthy Artificial*

considerations on policy and investment recommendations for trustworthy AI.[121] Due to the AI's specificity, blanket recommendations were found to be of limited use. The HLEG therefore focused on three sectors, *i.e.*, the public sector (1), healthcare (2), and manufacturing and Internet of things (3).

The HLEG's Ethics Guidelines were then subject to an evaluation which lasted from June to December 2019.[122] Once this evaluation was completed, the HLEG published a revised (and final) version of the assessment list for trustworthy AI (hereafter, ALTAI) in July 2020.[123] ALTAI particularized the four-pillar framework of principles into seven requirements, *i.e.*, human agency and oversight (1), technical robustness and safety (2), privacy and data governance (3), transparency (4), diversity, non-discrimination and fairness (5), societal and environmental well-being (6), and accountability (7). Meanwhile, the new - and current - President of the EC, Ursula von der Leyen, promised to put forward a legislative proposal for AI.[124] This was followed by the publication, in February 2020, of a White Paper on AI defining a European approach to excellence and trust.[125]

Intelligence, EUROPEAN COMMISSION (June 26, 2019), https://digital-strategy.ec.europa.eu/en/library/policy-and-investment-recommendations-trustworthy-artificial-intelligence.

[121] Jola, Dervishaj, *AI HLEG – Sectoral Considerations on Policy & Investment Recommendations for Trustworthy AI*, EUR. COMM'N (July 23, 2020), https://futurium.ec.europa.eu/en/european-ai-alliance/document/ai-hleg-sectoral-considerations-policy-and-investment-recommendations-trustworthy-ai.

[122] *EU Artificial Intelligence Ethics Checklist Ready for Testing as New Policy Recommendations are Published*, EUR. COMM'N (Jun. 26, 2019) https://ec.europa.eu/digital-single-market/en/news/eu-artificial-intelligence-ethics-checklist-ready-testing-new-policy-recommendations-are; *Pilot the Assessment List of the Ethics Guidelines for Trustworthy AI*, EUR. COMM'N, https://ec.europa.eu/futurium/en/ethics-guidelines-trustworthy-ai/register-piloting-process-0 (last visited Mar. 31, 2023).

[123] *Assessment List for Trustworthy Artificial Intelligence (ALTAI) for Self-Assessment,* EUR. COMM'N, https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment (last visited Mar. 31, 2023).

[124] Ursula von der Leyen, A Union that strives for more: My agenda for Europe (Jul. 16, 2019), https://www.europarl.europa.eu/resources/library/media/20190716RES57231/20190716RES57231.pdf.

[125] COM (2020) 65 final, *supra* note 43.

Capitalizing on the seven requirements developed by the HLEG, the EC stressed in the White Paper that, since AI can cause material[126] or immaterial harms[127] it calls for a regulatory framework that "should concentrate on how to minimize the various risk of potential harm."[128] The risk-based approach in regulating AI was thus announced.

In October 2020, the EP adopted a Resolution with recommendations to the EC on a framework of ethical aspects of AI.[129] In response to the EP's invitation, on 21 April 2021 and following a nearly five-year build-up, the EC finally published a proposal for a regulation of the EP and of the Council laying down harmonised rules on AI and amending certain Union legislative acts (*i.e.,* the AI Act).[130] This Proposal follows the *risk-based approach* to AI regulation, previously outlined in the White Paper on AI. The Proposal was accompanied by a report assessing the proposed regulatory options[131] in which the EC explained that the AI Act is "future-proof and innovation-friendly," designed to "intervene only where this is strictly needed."[132] The MS Coordinated Plan on AI was simultaneously updated to take the Proposal into account.[133]

The progression in the EU's regulatory approach to AI as well as the design of the AI Act as such show that, much like national regulators, the Union institutions (in particular the EC) sought to

---

[126] *Id.* at 10 ("safety and health of individuals, including loss of life, damage to property").

[127] *Id.* ("loss of privacy, limitations to the right of freedom of expression, human dignity, discrimination for instance in access to employment").

[128] *Id.*

[129] 2020 O.J. (404) 63 (Oct. 20, 2020).

[130] *Proposal for a Regulation of the European Parliament and of the Councel Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts*, COM (2021) 206 final.

[131] *Id.*

[132] *Id.*

[133] EUR. COMM'N, *Annexes to the Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: Fostering a European approach to Artificial Intelligence (Coordinated Plan on Artificial Intelligence 2021 Review),* COM (2021) 205 final (Apr. 21, 2021).

strike the right balance between market rationality and rights/values rationality. On the one hand, it was paramount that any new regulation did not stifle innovation by raising disproportionate compliance costs. On the other hand, rights and values seemed to take precedence over other factors (say, economic development) in the EU's political and normative discourse on AI. In this context, the EC stressed the need for effective, risk-preventing mechanisms while reminding that AI developers and deployers were *already* subject to existing EU law provisions on fundamental rights namely, data protection, privacy, non-discrimination, consumer protection, product safety and liability.[134] The dilemma - frequently encountered in AI regulatory discourse - was whether those provisions were apt enough to prevent and sanction violations of fundamental rights or if the novelty and specificity of AI-related risks warranted both an update of the existing legislation and the adoption of new, AI-specific legislation. The EC seemed generally favourable to the update scenario, considering that some AI features - opacity for instance - threatened to render the application and enforcement of existing legislation more difficult.[135]

When it comes to the general method on how AI was to be regulated, the verdict was clear: the EC would follow a risk-based approach carefully tailored against the backdrop of fundamental values and rights protection. But how was the *object* of this approach (*i.e.,* the AI-related risks) defined? What was the level of correspondence between the risks the EC collected evidence on and those included in the AI Act? To answer these questions, we will discuss the discovery procedures launched in view of gathering the evidence on AI-related risks the EC ultimately chose to address in the AI Act.

---

[134] *Id.* At 10.
[135] *Id.*

B. The Divergence Between Established and Regulated AI-Related Risks

In theory, any risk regulating instrument is assumed to adequately regulate risks if it relies on a discovery procedure having yielded correct risk representations. The AI Act is - formally at least - not an exception to this trend. The EC did undertake discovery in the form of two public consultations (**1.**) and, based on the evidence gathered, it induced a four-level taxonomy of risks that the AI Act intends to regulate (**2.**).

1. The Discovery of Evidence of AI-Related Risks

We should mention that the taxonomy of risks now mentioned in the AI Act was originally established in soft-law instruments. The EC's *summa divisio* of risks was initially binary, the distinction being made between high-risk and non-high-risk AI. High-risk systems were further divided into two groups based on two sets of criteria. First, high risks that derive from specific characteristics displayed by AI systems, considering the *sectors* in which they were used, as well as the *ways* in which they were used. Second, the EC admitted - as an exception - that there might be instances where the use of AI applications for *certain purposes* may be considered as high-risk.[136] As a rule of thumb, the EP considered that AI systems can qualify as 'high-risk' when "their development, deployment and use entail a significant risk of causing injury of harm to individuals or society, in breach of fundamental rights and safety rules."[137] To operationalise this definition, the EP proposed an exhaustive list of high-risk sectors and uses or purposes entailing a risk of breach of fundamental rights and safety rules. According to the EP, *high-risk sectors* are employment, education, healthcare, transport, energy, public sector (including asylum, migration, border controls, judiciary and social security services), defence and security, finance, banking, and insurance. High-risk *uses* or *purposes* are recruitment,

---

[136] *Id.* At 18.
[137] *Id.*

grading and assessment of students, allocation of public funds, granting loans, trading, brokering, taxation (and similar activities), medical treatments and procedures, electoral processes and political campaigns, public sector decision that have "a signification and direct impact on the rights and obligation of natural or legal persons," automated driving, traffic management, autonomous military systems, energy production and distribution, waste management and emissions control.[138]

Against the backdrop of this pre-existing taxonomy of risks, the EC claimed that the AI Act was based on "two years of analysis of evidence and involvement of stakeholders."[139] The annexes of the impact assessment accompanying the AI Act listed the evidence used to draft this Proposal. The Commission identified high-risk AI systems[140] through an evaluation of "an external study" that reviewed "available evidence of fundamental rights or safety-related risks created by AI applications" (1), the public consultation on the AI White Paper (2) and on the Inception Impact Assessment (3), five "closed" expert webinars (4),[141] the symposiums organised by the European AI Alliance (5), the participation of Commission representatives to more than fifty online conferences and roundtables (6), the conclusion of the HLEG, the results of the piloting phase of their Ethics Guidelines (7), an "extensive literature review, covering

---

[138] *Id.*

[139] 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, note 237 at 51 SWD(2021) 84 final, Part ½, p. 1.

[140] For the EC's definition of this concept, *see supra*, Section II.

[141] These seminars were online workshop on conformity assessment on 17 July 2020 with 26 participants from the applying industry, civil society and conformity assessment community (1), online workshop on biometrics on 3 September 2020 with 17 external participants from stakeholders such as the Fundamental Rights Agency, the World Economic Forum, the French Commission Nationale de l'Informatique et des Libertés and academia (2), online workshops on standardization on 29 September 2020 with 27 external participants from UNESCO, OECD, Council of Europe, CEN-CENELEC, ETSI, ISO/IEC, IEEE, ITU (3), online workshop on potential requirements on 9 October 2020 with 15 external experts on AI, mainly from academia (4), and online workshop on children's right and AI on 12 November 2020 with external experts (5).

academic books, journals and well as a wide spectrum of policy studies and reports, including by non-governmental organisations" (8),[142] the annex of the European Parliament's Resolution 2020/2012(INL), a list of 132 AI use cases identified by the Final Draft of ISO/IEC TR 24030,[143] and AI Watch analysis.[144]

The difficulty encountered in our analysis was that we neither had access to the external study, nor were we able to attend the closed webinars or access the conclusions the Commission's representatives drew from the European AI Alliance Assembly and other similar conferences. We can therefore only use the public consultation on both the AI White Paper and the Inception Impact Assessment to assess the EC's ambition to draft an evidence-based regulatory proposal.[145] We will qualify both procedures as *discovery procedures* considering that - as stated in the Inception Impact Assessment - the public consultation was, indeed, aimed at "collecting evidence."[146] During the consultation, launched after the

---

[142] Commission Staff Working Document Impact Assessment, Annexes 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 (Apr. 21, 2021).

[143] *ISO/IEC TR 24030:2021*, WEBSTORE, https://webstore.iec.ch/publication/68996#additionalinfo (last visited Mar. 31, 2023). The use cases are clustered by theme, *i.e.*, agriculture, digital marketing, education, energy, fintech, healthcare, home/service robotics, ICT, legal, logistics, maintenance and support, manufacturing, media and entertainment, mobility, public sector, retail, security, social infrastructure, transportation, work and life, and others.

[144] *AI Watch,* EUR. COMM'N https://knowledge4policy.ec.europa.eu/ai-watch_en (last visited Apr. 1, 2023).

[145] Conclusions driven from these two public consultations are similar. As there were only 131 valid feedback instances received during the public consultation on the Inception Impact Assessment - against 1'216 for the public consultation launched after the publication of the White Paper - we choose ton only discuss the latter public consultation, *Artificial Intelligence – ethical and legal requirements*, EUR. COMM'N, https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12527-Artificial-intelligence-ethical-and-legal-requirements/feedback_en?p_id=8242911 (last visited Mar. 31, 2023).

[146] *Inception Impact Assessment*, EUR. COMM'N ARES(2020)3896535 (Jul. 23, 2020) https://ec.europa.eu/info/law/better-regulation/have-your-

publication of the White Paper on AI which ran from 19 February to 14 June 2020, the EC targeted stakeholders with an interest in AI *i.e.,* developers and deployers, companies and business organisations, SMEs, public administrations, civil society organisations, academic, and citizens.[147] There was a total of 1'216 respondents (N=1'216)[148] from 49 countries.[149] The consultation included 60 questions. However, since the present article focuses on uncovering the evidence having warranted the EC's taxonomy of, and regulatory responses to AI-related risks, we will place greater emphasis on the questions that concerned the regulation of (what the EC viewed as being) *high-risk* AI systems. Those questions were either closed (*i.e.,* with *predefined answers*) or open (*i.e.,* allowed free text answers.)[150]

---

say/initiatives/12527-Artificial-intelligence-ethical-and-legal-requirements_en. For the public consultation, *see* European Commission, Public consultation on the AI White Paper: Final Report (Feb. 19, 2020), https://wayback.archive-it.org/12090/20200811222622/https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf.

[147] *Artificial Intelligence – ethical and legal requirements*, *supra* note 146.

[148] *Id.* There are in this regard two difference between the information displayed online and in the European Commission's final report of the public consultation on the AI White Paper. First, the information available online numbered 1'216 respondents. The same goes for the .csv document that contains raw data. However, the final report states that "the public consultation attracted 1'215 contributions"; *id.* This difference has to be found in the number of business associations that answered public consultation (131 according to the online information, 130 in the final report). Second, the consultation lasted from 20 February, 2020 to 14 June, 2020 according to the information available online, but from 19 February, 2020 to 14 June, 2020 according to the final report; *id.* These differences, however, are not enough significant to challenge our conclusions.

[149] Germany, Belgium, France, Spain, United Kingdom, United States, Netherlands, Italy, Austria, Sweden, Finland, Portugal, Denmark, Poland, Romania, Ireland, Switzerland, Greece, Norway, Hungary, Czech Republic, Japan, Malta, Lithuania, India, Bulgaria, Slovenia, Slovakia, Luxembourg, Croatia, Canada, Turkey, Estonia, Cyprus, Serbia, Latvia, China, Vietnam, Syria, Swaziland, South Korea, Mexico, Iraq, Gibraltar, Côte d'Ivoire, Costa Rica, Brazil, Albania, and Afghanistan. Albeit stakeholders from 49 countries responded to the public consultation, it is worth noting that the top-five countries that answered – *i.e.*, Germany (20.64%), Belgium (13.32%), France (9.62%), Spain (8.63%), and United Kingdom (6.25%) – approximately represent sixty percent of the total respondents (58.46%). None of the forty-four other countries display more than 5 percent. Thirty-one countries even represent less than one percent.

[150] Unfortunately, we did not have access to more than 450 position papers

First, the EC looked to assess the *need* for specific regulation with the goal of determining if the risks raised by AI could be addressed by already existing EU legislation. 37.75% of the respondents stressed the need for new regulation, 29.19% considered that existing regulation may apply though it may contain certain gaps and only 2.63% found said regulation adapted to AI.[151] The *consensus* in the respondents' free text answers was that new or gap-filling regulation should be future-proof and *domain-specific* or *sector-based* rather than horizontal.[152] This is an important point to keep in mind for our further analysis of the fact-to-law correspondence[153] (or lack thereof) given that the AI Act, being a *horizontal* regulatory framework, does not seem to reflect the point of unanimity that *sectoral* regulation was seen as preferable.

Second, the EC sought to determine the types of AI systems (and corresponding risks) that should be covered by compulsory regulatory schemes. Interestingly, 37.66% of the respondents considered new compulsory requirements should be limited to *high-risk AI systems* (the closed question defining "high-risk" as the situation "where the possible harm caused by the AI system is particularly high"), 27.22% disagreed and 18.17% disclosed another opinion.[154] The free-text answers also point to a need to further

---

(written answers) attached to some responses. This prevents us from being exhaustive in our analysis. We complete our observations with the European Commission's final report of the public consultation on the AI White Paper (*Artificial Intelligence – ethical and legal requirements*, *supra* note 167) as well as with the Commission Staff Working Document Impact Assessment accompanying the AI Act (Commission Staff Working Document Impact, *supra* note 144).

[151] In public consultation launched after the publication of the Inception Impact Assessment, respondents highlighted that the EC should map legislative gaps before introducing new legislation.

[152] This opinion is shared by those who responded to the public consultation launched after the publication of the Inception Impact Assessment.

[153] *See infra* Section II.

[154] In the public consultation launched after the publication of the Inception Impact Assessment, 90% of the respondents who choose to discuss high-risk AI systems – but we do not know how many amongst the 131 respondents made that choice – consider new mandatory requirements should be limited to high-risk AI systems. *See* Proposal for a Regulation of the European Parliament and of the Council Establishing a Carbon Border Adjustment Mechanism, COM (2021)564

nuance the scale of risks departing from the 'simplistic' binary model which distinguished high- and non-high risk AI systems.

Third, the EC sought to induce an operative definition of the "high-risk" classification. However, the way it went about it was rather curious: in lieu of asking the respondents to give their own understanding of 'high risk,' they were asked to endorse (or not) the definition of this concept provided in the White Paper on AI.[155] In response to the question thus worded, 21.63% of the respondents agreed with said definition, 4.52% disagreed and 63.08% simply gave no answer. However, in their free text answers, respondents observed that both the definition of (high-)risk and the underlying methodology should be further refined.

Against the backdrop of the suggested (as opposed to inquired) definition of 'high-risk' AI, respondents were asked to list the types of AI application and/or use they viewed as most concerning. Their answers included lethal autonomous weapons, remote biometric identification, autonomous vehicles, AI systems dedicated to critical infrastructure (e.g., electricity), health, human resources and employment (*e.g.*, AI-powered recruitment tools), analysing and manipulating human behaviour, predictive policing, mass surveillance, political (dis)information, and law enforcement. Respondents cautioned, however, that this list ought to be regularly reassessed.

It follows that the EC (formally) relied on the gathered views and did indeed provide a list of high-risk AI systems contained in Annex III of the AI Act.[156] This list includes *eight sectors* namely, biometric identification and categorisation of natural persons, management of operation of critical infrastructure such as road traffic or energy supply, education and vocational training, including assessment required for admission to educational institutions; workers recruitment, access to essential public, law enforcement,

---

final (Feb. 7,2021; HLEG on AI , *supra* note 123 at 19-20.

[155] *See Commission White Paper on Artificial Intelligence-A European Approach to Excellence and Trust,* COM (2020) 65 final (July 2, 2020), *supra* note 43.

[156] *Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence and Amending Certain Union Legislative Acts,* COM (2021) 206 final (Sept. 4, 2021); Jordana & Levi-Faur *supra* note 73.

migration, asylum and border control and factual and legal research.[157] This effort of zooming in on sectors where 'high-risk' AI is most likely to be used is of course laudable given AI's inherent complexity, making the inducing of clear-cut taxonomies extremely difficult. However, as will be discussed further,[158] the EC seemed to be selective in its assessments of the probative value of the answers received, given that not all of said answers found their way into the AI Act's provisions. Of course, there is no formal requirement that this be the case; however, if the majority of the gathered views revealed clear preferences on key points (typically, that sectoral regulation was preferred to horizontal regulation), one might assume that those views ought to have *some* impact in shaping the EC's regulatory choices. Moreover, the overall tone of the consultation and the wording of the questions give the impression that the EC had an *already* operative definition of high-risk systems and had - one might argue, pre-emptively - tailored its regulatory response to those systems, the consultation procedure serving the purpose of confirming (or not) the approach that the EC was going to follow anyway. In short - as will be discussed below - some *relevant* facts uncovered by the EC seemed to be largely disregarded as can be induced, *inter alia*, from the taxonomy of risks the AI Act ultimately included.

---

[157] We should stress here that such systems are, in principle, forbidden when used for the purpose of law enforcement. It therefore remains unclear if, when used for purposes other than law enforcement, these systems will fall in the scope of application of the obligations aimed at high-risk AI. 'Recruitment' pertains to the process of screening and filtering applications (4.a) as well as task allocation and performance monitoring (4.b). For example, eligibility for public assistance assessment (5.a), private services, *e.g.*, credit scoring (5.c), including dispatching of first aid services (5.c). 'Law enforcement' encompasses individual risk assessment (6.a), emotion (6.b) or deep fake detection tool (6.c), evaluation of evidence reliability assessment tool (6.d), predictive policing (6.e), profiling (6.f) and data analysis allowing the discovering of hidden patterns and information (6.g). The AI uses referred here are emotion recognition (7.a) risk assessment, verification of authenticity of documents (7.c) and eligibility for asylum and other procedures (7.d). This includes interpretation tools used by judicial authorities (8.a).

[158] *See infra*, Section II.

2. The Regulatory Response to the Evidence Discovered

The EC's stance - expressed before and in the AI Act - is that only high-risk AI systems should be subject to *mandatory requirements* applicable to the training data, robustness, accuracy and human oversight. Alternatively, so-called non-high-risk systems would only be subject to voluntary labelling schemes. Departing from the high-risk/non-high-risk dyad, in the AI Act, the EC included series of obligations for AI systems which it ultimately represented on a four-level risk scale: non-high-risk, limited risks, high risks and unacceptable risks.

Non-high-risk AI systems are defined in opposition to high-risk systems. As high-risks AI systems are exhaustively enumerated, non-high-risks AI systems form a residual (and presumably the largest) category.[159] The regulatory principle for those systems is the absence of a duty to comply with the mandatory requirements which target the high-risks systems (Art. 8). Developers and users of non-high risk AI systems are, however, encouraged to voluntarily apply these requirements through codes of conduct (Art. 69). This voluntary compliance mechanism is reminiscent of the Draft Ethics Guidelines.[160] However, considering the strong criticism received during the open consultation, voluntary compliance as a principled solution was ultimately abandoned in the final report.[161]

Limited risks AI system are, similarly, not subject to mandatory requirements set up in the AI Act (art. 8). However, the AI Act does establish an obligation of transparency for systems which, though formally qualified as non-high risk, interact with natural persons (art.

---

[159] Pauline Bégasse de Dhaem and Denis Philippe, *La digitalisation du secteur bancaire: analyse du projet de règlement européen en matière d'intelligence artificielle, in* ACTUALITÉS EN DROIT ÉCONOMIQUE: L'ENTREPRISE FACE AU NUMÉRIQUE 211, 237 (2021).

[160] HLEG on AI, *supra* note 123 at 2.

[161] *Id.*

52(1)), perform emotion recognition[162] or biometric categorisation[163] (art. 52(2)). Such systems ought to be designed in a way that natural persons know they interact with or are exposed to an AI system. In a similar vein, users of so-called deepfake technology - *i.e.*, hyper-realistic videos using face swaps that leave little trace of manipulation[164] - are required to disclose that the content has been manipulated or artificially generated (art. 52(3)).

High-risk AI systems are - as mentioned - those that pose the most pressing threats to health, safety and fundamental rights. Rather than being altogether prohibited, they are subject to mandatory requirements, chiefly transparency (art. 13) and human oversight (art. 14). Building on the abovementioned criteria (types of use, primary sectors of use, purpose) for qualifying AI systems as 'high-risk,' the AI Act distinguishes between two categories. The first category includes systems intended to be used as safety component of products covered by EU sectorial product legislations listed in Annex II (art. 6(1)(a)) and that are subject to third party *ex-ante* conformity assessment (art. 6(1)(b)), bearing in mind that a safety component is "a component of a product or of a system which fulfils a safety function for that product or system or the failure or malfunctioning of which endangers the health and safety of persons or property" (art. 3(14)). The second category includes stand-alone AI systems with mainly fundamental rights implications that are explicitly listed in Annex III (art. 6(2)).

Finally, AI systems that pose *unacceptable risks* are subject to an *ex officio* ban (art. 5).[165] It should be stressed that military applications are excluded from the scope of the AI Act (art. 2(3)).

---

[162] Art. 3(34) Proposal defines "emotion recognition system" as "an AI system for the purpose of identifying or inferring emotions or intentions of natural persons on the basis of their biometric data."

[163] Art. 3(35) Proposal defines "emotion recognition system" as "an AI system for the purpose of assigning natural persons to specific categories, such as sex, age, hair colour, eye colour, tattoos, ethnic origin or sexual or political orientation, on the basis of their biometric data."

[164] Mika Westerlund, *The Emergence of Deepfake Technology: A Review*, 9 TECH. INNOVATION. MGMT. REV. 40 (2019).

[165] The provisions cited in this Section are articles of the AI Act.

With this exception in mind, AI systems that either use subliminal manipulation of natural person's consciousness (art. 5(1)(a))) or exploit vulnerabilities of a specific group of persons due to their characteristics, *e.g.*, age, physical or psychological disability (art. 5(1)(b)) in order to distort people's behaviour in a way that is likely to cause physical or psychological harm are prohibited.

The ban also extends to AI systems used by public authorities that score natural persons based on their personal and social behaviour, known or predicted (art. 5(1)(c)) as well as those that may lead to detrimental or unfavourable treatment of certain natural persons or groups either "in social contexts which are unrelated to the contexts in which the data was originally generated or collected" (art. 5(1)(c)(i)) or that is "unjustified or disproportionate to their social behaviour or its gravity" (art. 5(1)(c)(ii)).

In a similar vein, the use of real-time remote biometric identification systems[166] in publicly physical spaces accessible to the public for the purpose of law enforcement is a prohibited practice (art. 5(1)(d)). Such systems may be *exceptionally* authorized when they are *strictly necessary* to either search for specific potential victims of crime, including missing children (art. 5(1)(d)(i)), prevent specific, substantial and imminent threat to natural persons' physical safety, including terrorist attacks (art. 5(1)(d)(ii)) or one of the thirty-two criminal offences referred to in article 2(2) of Council Framework Decision and punishable by a custodial sentence or a detention order for a maximum period of at least three years (art. 5(1)(d)(iii)).[167] In case they are exceptionally authorised, law

---

[166] These are AI systems that aim at identifying without significant delay "natural persons at a distance through the comparison of a person's biometric data with the biometric data contained in a reference database, and without prior knowledge of the user of the AI system whether the person will be present and can be identified" (art. 3(36) AI Act).

[167] Article 2(2) of Council Framework Decision 2002/584/JHA, of 13 June 2002, on the European arrest warrant and the surrender procedures between Member States - Statements made by certain Member States on the adoption of the Framework Decision, OJ L190/1 ("the following offences, if they are punishable in the issuing Member State by a custodial sentence or a detention order for a maximum period of at least three years and as they are defined by the law of the issuing Member State, shall, under the terms of this Framework Decision and without verification of the double criminality of the act, give rise to surrender

enforcement authorities are held to carry out a cost-benefit analysis (art. 5(2)(a) and (b)) and will need an *ex ante* authorisation granted by either a judicial or an independent administrative authority (art. 5(3)).

On the surface, the inception in both policy and fact of the AI Act seems beyond reproach. Formally, the EC ticked all the boxes: it complied with its institutional (and political) duty of responding to the invitation to present a legislative proposal on a matter as 'urgent'[168] as AI. The design of this proposal seems to have been guided by the 'right' values resulting in a fact-conform taxonomy of risks and corresponding legal duties. However, the reality-normativity interplay underlying the design of the AI Act does not reveal that evidence played a major role in shaping the latter. In this context, if not reality, what is the AI Act a response to? To address this issue, we must take a closer look into the *epistemic validity*[169] of

---

pursuant to a European arrest warrant: participation in a criminal organisation, terrorism, trafficking in human beings, sexual exploitation of children and child pornography, illicit trafficking in narcotic drugs and psychotropic substances, illicit trafficking in weapons, munitions and explosives, corruption, fraud, including that affecting the financial interests of the European Communities within the meaning of the Convention of 26 July 1995 on the protection of the European Communities' financial interests, laundering of the proceeds of crime, counterfeiting currency, including of the euro, computer-related crime, environmental crime, including illicit trafficking in endangered animal species and in endangered plant species and varieties, facilitation of unauthorised entry and residence, murder, grievous bodily injury, illicit trade in human organs and tissue, kidnapping, illegal restraint and hostage-taking, racism and xenophobia, organised or armed robbery, illicit trafficking in cultural goods, including antiques and works of art, swindling, racketeering and extortion, counterfeiting and piracy of products, forgery of administrative documents and trafficking therein, forgery of means of payment, illicit trafficking in hormonal substances and other growth promoters, illicit trafficking in nuclear or radioactive materials, trafficking in stolen vehicles, rape, arson, crimes within the jurisdiction of the International Criminal Court, unlawful seizure of aircraft/ships, sabotage").

[168] CO EUR 7/CONCL 5, *supra* note 120.

[169] In the context of risk-assessment as explored in this study, epistemic validity will be understood as an attribute that sanctions the acceptability of knowledge of risks, provided that that knowledge is produced through discovery models perceived as reliable. For a critical analysis of this understanding of epistemic validity, *see e.g.,* Rafaela Hillerbrand, *Climate Change As Risk?*, *in* HANDBOOK OF

the evidence-gathering and knowledge-construction performed by the EC.

III. SEEKING KNOWLEDGE OF FACTS FOR THE PURPOSE OF POLICY: THE EPISTEMIC CHALLENGES OF RISK IDENTIFICATION

Let us begin with three postulates. First, the world with all its risks does not represent itself; it must be represented through some means of communication.[170] Second, all inquiries require truth-standards,[171] defined through communal practices of acknowledgement, correction and critique.[172] Third, the risk with risk-assessment is *misrepresentation* due to improper gathering and interpretation of the evidence meant to reveal risks' causes and effects.

Against the backdrop of these observations, a paradox emerges regarding the discovery of risks. As persistent uncertainties, they can never be fully uncovered and yet their discovery should, in principle, yield some form of 'valid' knowledge which ought to warrant a policy response. Hence the importance of selecting well suited procedures and standards which allow to flag generally approved sources of information.[173] However, the launching of such procedures presupposes adequate prior understanding of that which ought to be discovered. In other words, an *inquiry* should be properly formulated[174] which, in the context of risks, occurs in the stage of so-called risk identification. As will be argued in **Sub-Section A.**, identifying risks in the realm of tangible, real-world facts can be challenging because they belong to the broader class of so-called

---

RISK THEORY. EPISTEMOLOGY, DECISION THEORY, ETHICS, AND SOCIAL IMPLICATIONS OF RISK 320, 328 (2012).

[170] Donald Nicolson, *Taking Epistemology Seriously: 'Truth, Reason and Justice' Revisited*, 17 INT'L J. EVIDENCE & PROOF 1, 34 (2013).

[171] Buekens & Truyen, *supra* note 49 at 221: "any sensible inquirer will set for himself standards of accuracy, and only when attainable by the inquirer, will they count as reasonable standards."

[172] Nicolson, *supra* note 172 at 22.

[173] Steven L. Reynolds, KNOWLEDGE AS ACCEPTABLE TESTIMONY 27 (2017).

[174] *See id.* at 2.

unknowns, making it often difficult to circumscribe the portions of reality that call for further exploration. To address this challenge, evidence theory will, as argued in **Sub-Section B.**, provide useful guidance on the application of *relevance* as a criterion of selection of facts more deserving of exploration than others.

A. Identifying the *loci* of Uncertainty

The main challenge with risks is that, try as we might, our knowledge of them will always be incomplete. Since they pertain to possible and/or probable threats, any knowledge we might acquire will necessarily be inconclusive,[175] with lingering uncertainty on the causes and outcomes of the so-called generating facts (*i.e.,* facts which would allow a risk to materialize).[176] Epistemically speaking, this is upsetting because it makes the accomplishment of one important task extremely difficult: that of distinguishing knowledge proper from speculative or hypothetical knowledge.

Risks - scholars tell us - fall in the broader class of 'unknowns'[177] characterized by various degrees of *unknowability*.[178] Though we

---

[175] 'Incomplete knowledge' here is equivalent to Leonelli's 'persistent uncertainty' concept; *see* Giulia C. Leonelli, *Acknowledging the Centrality of the Precautionary Principle in Judicial Review of EU Risk Regulation: Why It Matters*, 57 COMMON MKT. L. REV. 1773, 1779 (2020).

[176] Hauke Riesch, *Levels of Uncertainty*, *in* ESSENTIALS OF RISK THEORY 29, 34 (2013); *see also* EFSA Scientific Committee, Guidance on Uncertainty in EFSA Scientific Assessment: www.efsa.europa.eu (last visited Apr. 1, 2023), at 3: "uncertainty is used in the Guidance as a general term referring to all types of limitations in available knowledge that affect the range and probability of possible answers to an assessment question. Available knowledge refers here to the knowledge (evidence, data, etc.) available to assessors at the time the assessment is conducted and within the time and resources agreed for the assessment."

[177] Rolf Lidskog & Go¨ran Sundqvist, *Sociology of risk*, *in* ESSENTIALS OF RISK THEORY 75, 85 (Sabine Roeser et al. eds., 2013; Leonelli, *supra* note 177 at 1779: "the precautionary principle thus applies when, in the face of persisting uncertainty, a risk may be too high to comply with the EU intended level of protection."

[178] Here we use the term 'unknowability' as a generic term that encompasses several types of unknowns, including risks. However, in the remainder of this paper we will - for simplicity's sake - refer of levels of uncertainty associated with

might be inclined to take a binary stance toward knowing ('we either know or do not know'), unknowability, as a state of relative or total lack of knowledge, is far from being homogenous.[179] Scholars typically suggest four types of unknowns: *risks stricto sensu*, relating to knowledge limited to the odds of an even occurring; *uncertainty*, relating to a lack of knowledge about the odds of a given occurrence; *ignorance*, which is a case of not knowing what we do not know and *indeterminacy*, which characterizes cases where the causal chains or networks between phenomena are open.[180] For the purpose of this article, the concept of ignorance will be excluded because it *prima facie* refers to *total unknowability* of causes, processes and outcomes; it is an 'unknown unknown' whereas risks, uncertainties and indeterminacies are 'known unknowns'[181] that is, events for which the causes (or generating facts) might be known but not their outcomes.

All knowledge-seeking enterprise begins with an inquiry. With risks, the initial inquiry is not necessarily *what* they are but *where* they are: how does one make the decision to explore a portion of reality where risks might reside, if the threats associated with those risks have not yet manifested as tangible, real-life experiences? Properly framing an inquiry - of any kind - is a prerequisite for valid knowledge[182] to be construed. As Haack put it, "an inquirer's business is to discover the true answer to his question; so his obligation is to seek out what evidence he can and assess it as fairly

---

risks.

[179] David J. Spiegelhalter & Hauke Riesch, *Don't Know, Can't Know: Embracing Deeper Uncertainties When Analysing Risks* 369 PHIL. TRANS. R. SOC. A 4730, 4734 (2011).

[180] *Id.*

[181] This is of course a rather simplistic way of distinguishing states of 'ignored ignorance' from states of 'conscious ignorance.' Minimizing, mitigating or excluding ignorance is a complex process which can be represented as a 'progression between determinism and total ignorance and includes statistical uncertainty, scenario uncertainty, recognized uncertainty, and total ignorance; *see* Riesch, *supra* note 178 at 34; *see also* Kenneth Pettersen, *Understanding Uncertainty: Thinking Though in Relation to High-Risk Technologies*, *in* ROUTLEDGE HANDBOOK OF RISK STUDIES 39, 41-42 (Adam Burgess et. al eds., 2016).

[182] The concept of valid knowledge will be defined *infra*, Section IV.

as possible."[183] The risk-ridden real world of course is "a lot messier,"[184] with the congenital unknowability of risks reducing the chances for rock-solid, across-the-board intellectual integrity and scrupulous thoroughness.[185] To determine how inquiries about risks are formulated, we need to look at the *criteria* used to gain awareness of the threats that warrant further exploration and possible regulatory address. Simply put, we need to explore the underlying methodology of risk-identification.

In the process identifying risks, the main problem is the identification of the so-called locations of uncertainty.[186] To better understand the criteria that allow both scientists and regulators to 'know where to look,' scholars have generally raised three questions: 1. what do we want to know about risks?; 2. how are risks perceived and 3. what makes risks relevant?

By addressing the first question, scholarship essentially sought to uncover the *types of uncertainties* likely to attract scientific and regulatory attention. Views vary in this regard. Van Asselt and Rotmans suggested a binary divide of risks in *epistemic* (lack of knowledge proper of a given risk) and *aleatoric* (inherent variability of the phenomenon for which knowledge is sought).[187] Walker *et al.* proposed a pluralist model, suggesting as possible *loci* of uncertainty the context, models, data input, data parameters and final outcomes of future events.[188] Leonelli suggested a source-occurrence-epistemology triad arguing that - in the context of the EU precautionary principle[189] - uncertainties can be hazard-related, risk-

---

[183] Susan Haack, *Epistemology Legalized: Or, Truth, Justice, and the American Way*, 47 AM. J. JURIS. 43, 45 (2004), reprinted in Susan Haack, EVIDENCE MATTERS: SCIENCE, PROOF AND TRUTH IN THE LAW 30 (2014).

[184] *Id.*

[185] *Id.*

[186] Warren E. Walker et al., *Defining uncertainty: a conceptual basis for uncertainty management in model-based decision support*, 4 INTEGR. ASSESS. 5, 9 (2003).

[187] Marjolein B.A. van Asselt & Jan Rotmans, *Uncertainty in Integrated Assessment Modelling: From Positivism to Pluralism*, 4 CLIM. CHANGE 75, 76 (2002).

[188] Walker et al., *supra* note 188.

[189] From the perspective of risk-regulation scholarship, the level of regulatory

related or methodological.[190]

These taxonomies are open to criticism. Asselt's and Rotman's *summa diviso* between epistemic and aleatoric uncertainties can be brought into question. The variability of a phenomenon will necessarily have repercussions on the design of the procedures followed in its discovery. For example, if an environmental hazard is variable *by nature* (*i.e.,* may have multiple causes or produce multiple outcomes[191]), the knowledge-seeking procedures relative to that hazard will likely not be 'solid' in the sense that they will not properly circumscribe the object of the *relevant* knowledge and - by extension - will not employ adequate methodologies to explore it.[192] Similarly, Walker's and Lionelli's views can be questioned, since flawed (in the sense of insufficient, non-probative and/or untrustworthy) information on the *cause* of a hazard will necessarily reflect in the mapping out of the causal chains, the enhancing or inhibiting factors, the possible outcome(s) and the methodologies employed to study the risk and the strategies for its prevention. In other words, an improper circumscription of the *locus* of uncertainty has a *holistic epistemic effect*, insofar as it impacts the protocols and

---

scrutiny over reality can vary depending on whether an instrument applies the principle of *precaution* or that of *prevention*. Prevention in a strict sense - Flückiger argues - aims at preventing harm in cases where the causal link between a generating fact and a harm is supported by evidence. Alternatively, precaution translates to regulatory action in the absence of such evidence. *See* Alexandre Flückiger, *La preuve juridique à l'épreuve du principe de précaution*, 128 REV. EUR. SCI. SOC.107, 112 (2003).

[190] Leonelli, *supra* note 177 at 1776-77.

[191] The inherent variability of hazards is, generally, taken into account in regulation, in particular the EU regulation. Given that all risk assessment is essentially based on presumptive inferences - as will be argued further in this study - regulators must leave open the possibility of rebuttal by subsequent (typically scientific) evidence. This is the reason why most of the EU law instruments that apply the precautionary principle include the possibility for new evidence that could provide more accurate accounts on the risks those instruments address. This point will be discussed *infra*, Section IV.

[192] As Riesch rightly argued: "strictly speaking, all uncertainties are epistemic (...) the boundaries of whether an uncertainty should be considered epistemic or aleatoric seems to be a result of the setup, but the precise boundaries or even existence of the boundary to a large extent depends on our philosophical stances and background assumptions and knowledge." *See* Riesch, *supra* note 178 at 36.

procedures followed and inferences made in the process of discovery.

More importantly however, the cited views focus on locating uncertainties *in the process of risk characterization* but do not allow to flag those uncertainties at the stage of risk *identification*. For example, Van Asselt's and Rotmans' views may be instructive on where uncertainties might lie while assessing the *already identified* AI-related risks but they do not allow us to explain *why* the EU legislature chose to address those risks in the first place.

We may then turn to the second issue mentioned above the gist of which is that *perceptions* of risks point to the types of uncertainties that warrant exploration. The logic here is fairly straightforward: the exploration of a threat associated with a risk is sanctioned by prevailing views on the existence and magnitude (or imminence) of that threat. Möller suggested that perspectives on risks can be scientific, psychological and cultural. From a scientific viewpoint, risks are typically seen as real-world, possible occurrences which can be investigated and measured in systematic ways. The presupposition here is that risks have 'natural,' objective and tangible features which implies that 'veritable' risks are those that can be detected through tried-and-true methods of unbiased scientific discovery. This objectivist view of risks is possibly the oldest and most typical of risk scholarship for essentially two reasons. First, it includes *measurability* as a distinctive feature of risks, against the backdrop of the broader class of 'unknowns.' It is Knight who seminally posited that "'risk' means in some cases a quantity susceptible of measurement (...) and there are far-reaching and crucial differences in the bearings of the phenomena depending on which of the two is really present and operating (...) It will appear that a *measurable uncertainty or 'risk' proper* (...) is so far different from unmeasurable one that it is not in effect an uncertainty at all."[193] The merit of Knight is greater conceptual clarity. Following his reasoning, uncertainties and indeterminacies might be viewed as *descriptive* concepts, in the sense that they merely reveal instances

[193] Frank H. Knight, RISK, UNCERTAINTY AND PROFIT (reprint 1971) cited in Riesch, *supra* note 178 at 31.

where conclusive knowledge lacks. Alternatively, risks can be viewed as *evaluative*, because the level of knowledge - or lack thereof - can be quantified. For example, it is not enough to say that tobacco smoking *likely* shortens people's lifespan. According to the World Health Organization (WHO), there are over 4000 chemicals identified in tobacco smoke for which there is no level of safe exposure, either in first- or second-hand smoking.[194] The statistical data in this context, combined with the fact that in over 8 million deaths world-wide are smoking-related confirms that smoking is, indeed, a risk, not a mere uncertainty.[195] Second, the fact that only risks are measurable unknowns unveils a historic preference for *scientific knowledge* of risks, no doubt in view of avoiding risk-claims based on belief-driven speculations or guess work.[196]

While the cited reasons - revealing a *penchant* for scientific measurability - are epistemically sound, the scope of the *loci* of uncertainty they imply is rather narrow. Following Knight, the *choice* to regulate risks should follow a science-to-policy progression, the assumption being that only 'hard' scientific evidence can legitimately warrant regulatory responses. True as this may be, the AI Act is an exception to the rule. To the best of our knowledge, the EC did not rely on any thorough scientific studies on AI-related risks which would have revealed the high level of risk (of unfair bias) presented by, say, AI systems intended to be used for the purpose of assessing students in educational and vocational training institutions.[197]

---

[194] *See* WHO, Tobacco Fact-Sheet, www.who.int (last visited Sept. 1, 2022).

[195] *Id.*

[196] *See* CJEU, 8 September 2011, *Monsanto SAS et al.*, joined cases C-58/10 to C-68/10, EU:C:2011:553, para. 77: "protective measures adopted (...) cannot validly be based *on a purely hypothetical approach to the risk*, founded on *mere assumptions* which *have not yet been scientifically verified.* On the contrary, such protective measures, notwithstanding their temporary character and even if they are preventive in nature, may be adopted *only if they are based on a risk assessment which is as complete as possible in the particular circumstances of an individual case, which indicate that those measures are necessary*" (emphasis added).

[197] *See* COM (2021) 206 final, Annexes, *supra* note 73, Annex III. That said, we acknowledge that, to qualify AI systems intended to be used for the purpose of assessing students in educational and vocational training institutions, the EC relied

Since science is not everything when it comes to risk-identification, Möller suggested a more subjective, psychological approach which focuses on the ways in which individuals *perceive* risks. The trouble here is that there might be a dissonance between the 'what is actually out there' and 'what we think is out there,' public opinion having often been guilty of exaggerating or undermining the magnitude of the harms associated with various risks.[198] This has certainly been the case with AI. In the face of new technologies with cognitive skills historically attributed to humans, public discourse quickly became polarized with, on the one hand, high hopes in future progress and, on the other hand, apocalyptic scenarios of robots taking over the world.[199] One of the most prominent *topoi* of public debate have been the machines' replacement of human workforce. While laymen's perceptions revealed a fear of total human replacement across the labour market, expert evidence was more sobering, predicting a transition toward a machine-human working model in roughly 40% of all modern-day professions.[200] It follows that the public might inform on the risks to be explored and expect some type of regulation of those risks, but its perceptions might turn out to be extrapolations of reality. The possibility for misinterpretations is exacerbated in the third, so-called cultural approach to risk, the gist being that 'risks' are inherently relativist since cultural preferences shape the views on

---

on its literature review – mentioned above) and highlighted three cases: a discrimination case in Ph.D application in the University of Texas in 2014, a UK grading algorithm discriminating during the covid pandemic in 2020 and a discriminating Irish grading algorithm in 2020. However, there was no evidence akin to scientific evidence (*e.g.* no market surveys, no statistics on how many harms actually occurred in the last few years, which system caused those harms, etc.).

[198] Cass R Sunstein, *Moral heuristics and risk, in* RISK: PHILOSOPHICAL PERSPECTIVES 156, 164-65 (Tim Lewens ed., 2007).

[199] *See* Andrew Berg et al., *Should we fear the robot revolution? (The correct answer is yes), 97* J. MONET. ECON. 117 (2018).

[200] *See, inter alia*, Ljupcho Grozdanovski, *L'automatisation du travail et l'obsolescence du statut de travailleur : réflexions sur l'avenir du travail et de la protection des travailleurs*, 2 REV. FAC. DROIT UNI. LIEGE 21 (2019).

where risks lie and how severe they might be.[201] In extreme cases, 'culturalists' could go as far as contesting the fact-of-the-matter nature of risks, relying on culture to draw a roadmap of what society expects to be protected from by its regulators.[202]

With the AI Act in mind, we might conclude that neither the objectivist nor the subjectivist approaches to risk-identification allow us to determine the reasons why AI-related risks (or any risk for that matter) are viewed as priority candidates for evidence-gathering and policy. On the one hand - as argued - if we followed the 'hard' scientific approach, the scope of the risks addressed would, essentially, be narrowed down to health and environment as areas with already rich scientific research. On the other hand, if we give priority to individual or collective perceptions of risks... they might appear everywhere which is not a workable alternative for the purpose of policy: as much as they might hold their finger on the pulse of society, regulators cannot react to everything that society may qualify as risky. Hence the third question, mentioned above: what makes risks *relevant*? or, to be more exact, which criteria are used to determine whether the knowledge of specific risks is of importance for policy makers? It is on this point of relevance that evidence theory proves to be a useful epistemic referent.

B. Selecting *Relevant* Risks Warranting Further Exploration (and Regulation)

Relevance[203] is essentially a criterion under which an inquirer can exclude certain data sources and prioritize others. Relevance is,

---

[201] Sheila Jasanoff labels the national cultures of risk "civic epistemologies", *see e.g.* Sheila Jasanoff, *Restoring Reason: Causal narratives and political culture*, *in* ORGANIZATIONAL ENCOUNTERS WITH RISKS 209 (Briget Hutter & Michael Power eds., 2005).

[202] Möller, *supra* note 57 at 70.

[203] The basic idea with relevance is that there are propositions to be discussed in the context of an evidentiary debate and propositions to be established through evidence that supports them. In this context, relevance "means being probative directly or indirectly of any material proposition which is an ultimate *probandum*." *See* Jerome Michael & Mortimer J. Adler, *The Trial of an Issue of Fact: II*, 34 COLUM. L. REV, 1462, 1479 (1934).

therefore, paramount in defining the *object* and *scope* of virtually all knowledge-seeking enterprise: save examples of serendipitous scientific discovery, all inquiries (including those on risks) are raised because they present some importance for the scientific community, social institutions (public administrations, courts) and society at large. Both scholarship and regulatory practice - namely in the EU - show that relevance can be induced from two types of sources: it is either facts that reveal which risks warrant regulatory address (**1.**) or it is policy that *prima facie* defines the areas placed under enhanced regulatory scrutiny (**2.**).

1. Relevance Induced from 'Bare' Facts and Shared Perceptions

In law, relevance is most commonly associated with legal evidence. Anglo-American scholarship has viewed it as arguably the most important principle in evidence law[204] for mainly two reasons. First, not all facts are of interest for the law, but only those that enable the resolution of disputes within a given framework of legal remedies. Second, though it can take the shape of a procedural rule, relevance in its generic form appears as a principle of *rational exclusion*. In a trial, the object of a dispute logically dictates the evidence directly or indirectly related to the knowledge needed to resolve that dispute (the 'ultimate' *probandum*).[205] Procedural law is rarely providential to a point where it can micro-regulate the relevant evidence across the board. Rather, rules of procedure establish a basic framework within which the parties can 'compete' in view of persuading a court of their versions[206] of the facts, thus increasing

---

[204] *See* James Bradley Thayer, A PRELIMINARY TREATISE ON EVIDENCE AT THE COMMON LAW (1898).

[205] Michael & Adler, *supra* note 205 at 1279.

[206] Evidentiary accounts are not objective truths but narratives of facts; *see* Russel Brown, *The Possibility of 'Inference Causation': Inferring Cause-in-Fact and the Nature of Legal Fact-Finding*, 55 MCGILL L. J. 1, 20 (2010); ("Legal fact-finders, by contrast, do not replicate events to determine why something happened. Instead, parties offer up various, and typically inconsistent versions of events, all of which are assessed in drawing a conclusion about the event,").

their chances of winning the case. Indeed, procedural law has been called courtroom epistemology[207] because it establishes the 'rules of the game' (as the result of conventional rule-making) while allowing the 'players' the freedom to elaborate strategies on winning (as an exercise in intelligence). History confirms that it was court practice that shaped evidence law as a modern concept, not the other way around, allowing us to assert that distinguishing relevant from irrelevant evidence is a natural stage of any trial, whether it operates under a codified legal framework or not.[208] For this reason, relevance is not so much a *rule* of evidence *per se* "as a *presupposition* involved in the very conception of a rational system of evidence."[209]

The role of knowledge in the formation fact-based policy (like risk regulation) is, of course, different from that in dispute resolution because the regulation-forming processes are not adjudicatory in nature. Regulation does, however, seek to protect (through, say, technical standardisation) which naturally begs the question: 'from what?' More specifically, how can relevance, in the context of evidence theory, allow us to determine *what informs* regulators of the empirical knowledge they ought to gather to uncover a risk worth exploring? Practice, namely in the EU, shows an oscillation between the above-mentioned objectivist and subjectivist perspectives on the sources of risks.

Regarding the influence of the objectivist approach, in an ideal world the facts would 'speak for themselves' (*res ipsa loquitur*)[210] spontaneously revealing the aspects of reality warranting further exploration. The problem with 'bare' facts is that they do not always

---

[207] Kyle McGee, BRUNO LATOUR: THE NORMATIVITY OF NETWORKS 1, 4 (2014).

[208] Evan Bell, *An Introduction to Judicial Fact-Finding*, 39 COMMONWEALTH L. BULL. 519, 521 (2013): "Relevance must, and can only, be judged by reference to the issues which the court is called upon to decide."

[209] Terence Anderson et al., ANALYSIS OF EVIDENCE 289, 291 (2005). In a similar vein, Van Emeren and Grootendorst consider that "relevance or irrelevance always pertains to a certain kind of relation between elements or parts of a discourse or text that is judged (dys)functional to achieving a particular goal or purpose." *See* Frans H. van Eemeren & Rob Grootendorst, A SYSTEMATIC THEORY OF ARGUMENTATION. THE PRAGMA-DIALECTICAL APPROACH 69, 70 (2004).

[210] In evidence theory, this is known as the *res ipsa loquitur* principle, interpreted as meaning that certain facts are self-evident to a point where they can support a legal claim without further evidence required.

provide sufficient grounds for predictive reasoning. This has certainly been true regarding advanced Machine Learning (ML) systems. Though programmers follow strict protocols in the selection of the so-called ground data (*i.e.,* the data used during programming) and in the validation of a given system's real-life performance, it remains that it is often impossible to predict the direction that given system is likely to take, once released in the market.[211] Although the AI Act includes provisions on transparency and explainability,[212] the feasibility of these requirements can be - legitimately - brought into question, considering that programmers are often unable to know in advance 'where to look for the flaw'[213] likely to cause harm down the line.

The arguably only circumstance where bare facts reveal risks is that of... the risks materializing. There have been instances where real-life events in some parts of the world cautioned regulators in other parts of the world against specific risks. For example, Directive 96/82 on the control of major-accident hazards involving dangerous substances[214] was enacted in response[215] to two large-scale industrial disasters in India[216] and Mexico[217] from 1984. Another factor generally seen as revelatory of risks has been the presence of a global call for caution, the gist being that universally shared perceptions of a threat confirm the actuality of that threat. In the case of GMOs e.g.,

---

[211] Weston Kowert, *The Foreseeability of Human-Artificial Intelligence Interactions*, 96 TEX. L. REV. 181, 204 (2017), ("once the artificial intelligence is sent off to the buyer, the programmer no longer has control and the artificial intelligence could be shaped by its new owner in uncountable ways."

[212] *See* our comments on transparency and explainability in the AI Act, *supra,* Section II.

[213] Michael C. Gemignani, *Product Liability and Software*, 8 RUTGERS COMPUT. & TECH. L.J. 173 (1981).

[214] Council Directive 96/82 of 9 December 1996 on the control of major-accident hazards involving dangerous substances, OJ n° *L 10, 14/01/1997, p. 13.*

[215] *Id.*; Preamble, pt 4.

[216] Important quantities of gas escaped an insecticide plant, drifting into densely populated neighborhoods and causing the death of an estimate between 15'000 and 20'000 people.

[217] A gas leakage in a storage and distribution facility for liquified petroleum gas near Mexico City caused around 13 explosions, causing hundreds of deaths and thousands of injured amongst the nearby population.

Regulation n° 1272/2008[218] confirms the EU's intention to contribute to "the *global harmonization of criteria* for classification and labelling, not only at UN level, but also through incorporation of the internationally agreed GHS criteria into [Union] law."[219]

A case of global agitation has been witnessed in the field of AI as the increased business use of certain types of artificial systems began providing examples of harms that those systems might cause. We allude to the topical examples of traffic accidents involving automated vehicles,[220] Amazon's discriminatory recruitment algorithm[221] and trading systems capable of so-called spoofing.[222] We should stress, however, that unlike the industrial hazards mentioned earlier (the magnitude of which rightfully caused concern on the probability of accidents related to 'dangerous' substances), the available data on AI-related harm can be said to illustrate *possible* (*i.e.,* likely) rather than *highly probable* (*i.e.,* quantifiable) harms which begs the question: *are they really risks*?

One of the thorny issues, mainly addressed in doctrines on AI liability,[223] has been that of *causal generalizations*.[224] An evidence scholar might argue that an isolated case of a Tesla car running over a pedestrian is not a strong enough *indicium* to warrant the inference that *all* automated vehicles are likely to cause the same harm.

---

[218] Regulation 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures, amending and repealing Directives 67/548 and 1999/45, and amending Regulation n° 1907/2006, *OJ L 353, 31.12.2008, p. 1.*

[219] Communication from the Commission of 2 February 2000 on the precautionary principle, COM (2000) 1 final; Preamble, pt 6.

[220] *See* Umeda et al. v. Tesla, No 5:20-cv-02926-SVK, 2020 U.S. Dist. LEXIS 175286, at *9 (N.D. Cal. Sept. 23, 2020).

[221] *See* Ljupcho Grozdanovski, *In Search of Effectiveness and Fairness in Proving Algorithmic Discrimination in EU law*, 58 COMMON MKT. L. REV. 99 (2021).

[222] *See* Yavar Bathaee *The Artificial Intelligence Black Box and the Failure of Intent and Causation*, 31 HARV. J. L. & TECH 890 (2018).

[223] *See* Andrea Bertolini & Francesca Episcopo, *The Expert Group's Report on Liability for Artificial Intelligence and Other Emerging Digital Technologies: A Critical Assessment*, 12 EUR. J. RISK REG. 644 (2021).

[224] *See* Ljupcho Grozdanovski, *L'agentivité algorithmique, fiction futuriste ou impératif de justice procédurale ? Réflexions sur l'avenir du régime de responsabilité du fait de produits défectueux dans l'Union européenne*, 232 RÉSEAUX 99, 111 (2022).

Similarly, if a CV-scanning system develops a gender bias, it does not provide sufficient grounds to assume that all such systems are capable of developing the same bias. Some may be perfectly bias-neutral or might develop biases other than the one expected (ethnic background e.g.)... It follows that, with AI-related risks, the delicate question is one of measurability: how many instances of harm should occur for there to be *strong empirical evidence* that AI systems *do present certain risks*?

An overview of fact-finding procedures in the field of AI shows that, when choosing which risks to address, regulators were generally disinterested in statistical evidence on the possibly harmful features of various systems. They seemed more interested in how representative groups felt about certain *preselected* risks. In the United Kingdom (UK) e.g., within the Expl*AI*n project,[225] it was stated that "*understanding public opinion* is vital to develop guidance that is effective at supporting organisations to meet the expectations of individuals when explaining AI decisions."[226] Due to the complexity of AI, those involved in the project consulted citizens' juries[227] who were asked three questions pertaining to four scenarios (medical diagnosis, recruitment, organ transplant, criminal justice) as well as three general questions about AI decisions, focused mainly on explainability. The questions were thus worded so as to test the respondents' views on the importance of explanation of automated decisions and whether they would choose a system which, though accurate, would remain inexplainable. Judging by the

---

[225] *See* THE ALAN TURING INSTITUTE, www.turing.ac.uk (last visited Mar. 8, 2023).

[226] Information Commissioner's Office (ICO) / Alan Turing Institute,*Project ExplAIn, Interim Report*, at 9 , www.ico.org.uk (last visited Jan. 20, 2023).

[227] *See id.*; citizens' juries were created in 1971, under the belief that everyday citizens can provide unique insight into tackling complex issues. In the AI consultation, the ICO formed two such juries with the UK's National Institute for Health Research (NIGH) and Greater Manchester Patient Safety Transnational Research Centre (GM PSTRC). Both juries followed the same design and process. Jurors were made up of a cross-section of the population, representing the demographic breakdown of England as per the 2011 Census. The criteria used for juror selection were gender, age, ethnicity and educational attainment. Each jury comprised of 18 participants (a total of 36).

overall tone of the consultation, the point was to gather information on the citizens juries' attitudes *vis-à-vis* two potentially conflicting objectives *i.e.,* AI performance and privacy protection. However, in this consultation, the risks were, in a way, assumed: the respondents answered the questions bearing in mind the possibility of specific risks materializing; they were not asked *to identify* the risks which, according to them, were most salient or serious in the context of AI.

On the level of the EU - as already argued[228] - the EC's approach was similar. Indeed, the Commission seemed to strongly rely on a pre-existing definition of high-risk, as it designed the discovery procedures to 'nudge' the respondents in the direction of the previously defined risk-based approach.[229] This subtle 'nudging' can be inferred from the wording of the questions. For example, the EC asked how important, in the respondents' views, are the six AI features[230] acting as yardsticks for future regulatory address - and *already proposed* in the White Paper on AI[231] - '1' being not important and '5' being very important. This question essentially asked the respondents to endorse a pre-existing regulatory approach as opposed to suggest one.

Other questions pertained to specific risks. For example, the EC gathered views on the possibility that AI systems may discriminate, or that they may breach fundamental rights. However - here again - raising these issues presupposed that the corresponding risks *had already been identified.* Though the EC gave no specific indication on *how* the risk-identification was performed, it recognized 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 AI as a set of emerging technology."[232] This is reminiscent of what the

---

[228] *See supra*, Section II.

[229] *See supra*, Section II.

[230] These are training data, data and record-keeping, information to be provided, robustness and accuracy, human oversight and specific requirements for certain particular AI applications, such as those used for purposes of remote biometric identification.

[231] EUR. COMM'N, *White Paper On Artificial Intelligence – A European Approach to Excellence and Trust,* COM (2020) 65 final, *supra* note 43.

[232] European Commission, *Commission Staff Working Document Impact Assessment,* SWD(2021) 84 Final, Part 1/2, *supra* note 160 at 51 (emphasis

EC noted in its Inception Impact Assessment *i.e.,* that "AI is a highly dynamic and rapidly evolving industry so that *not a lot of currently valid evidence is available* at this stage."[233]

While collecting 'hard' evidence on the nature and scope of AI-related risks may, indeed, be challenging, the EC's evidence-gathering enterprise can be criticized for various reasons, discussed further in this article.[234] At this stage, may it suffice stressing that regulators seemed largely in favour of the *subjectivist approach* in risk-identification seeking to uncover perceptions and attitudes rather than quantifiable 'bare' facts. As mentioned,[235] the problem with perceptions and attitudes is that they might *amplify* their magnitude, thus triggering a shift from probability-based to caution-based assessments of various risks. In this regard, Lidskog and Sundqvist observe that "*fact finding* and *sense making* are seen as different and discrete spheres of activity, the former populated by technical risk analysts and the latter by various segments of the public."[236] The authors did not discuss *how* the risk (in form of risk events, hazards, or the technical calculation of risk) is constructed, "but only *how it is amplified.* Hence, the bridging ambition also results in a reproduction of the divide between *expert* and *public understandings of risk.* Not only risks are amplified in this approach, but also the divide between risk and understandings of risk."[237]

Lidskog and Sundqvist make an important point: experts and the public can have conflicting views in terms of the *existence* and *intensity* of risks. Indeed, "technical risk analysis cannot provide information about how risks are amplified in society (...) The social amplification approach proposes a division of labour in which technical risk analysis is concerned with investigating the *original signal* whereas social science in general and sociology in particular analyze how this *signal is transformed by society.*"[238]

---

added).

[233] Inception Impact Assessment, *supra* note 148.

[234] *See infra*, Section IV.

[235] *See supra*, Section III.

[236] Lidskog & Sundqvist, *supra* note 179 at 83.

[237] *Id.*

[238] *Id.* at 82. Some scholars are deeply pessimistic regarding the value of public

This *transformative effect* of public opinion can compromise the pursuit of *accuracy* in evidence theory, as applied - *mutatis mutandis* - in the field of fact-based policy. While it is true that reality rarely provides undoubtfully accurate accounts on the existence of risks, the evidence thereof - as imperfect as it may be - should *aspire toward accuracy i.e.,* should be empirically adequate,[239] the danger of failing to do so being errancy on the side of caution.[240]

Considering that neither 'bare' facts nor public opinions allow us to determine the realms of reality where regulators apply the criterion of relevance in choosing to investigate certain risks, another criterion is needed. Surprisingly - or perhaps not - practice in the EU shows that the yardstick for risk identification is seldom reality alone; it is first and foremost *policy*. When regulators commit to achieving specific levels of protection in various *predefined* areas, it is those areas that determine the realms of 'reinforced' regulatory scrutiny over risks. The prior commitment to protect against harms thus acts as Ockham's razor in the exclusion of 'irrelevant' evidence of those harms. In other words, policy seems to act as the ultimate relevance criterion for risk-identification in the EU.

---

opinions; *see e.g.,* Stephen G. Breyer, BREAKING THE VICIOUS CIRCLE: TOWARD EFFECTIVE RISK REGULATION (1993); Cass R. Sunstein, LAWS OF FEAR: BEYOND THE PRECAUTIONARY PRINCIPLE (2005). Others are highly critical regarding the role of the scientific expert. *See e.g.,* Sheila Jasanoff, DESIGNS ON NATURE: SCIENCE AND DEMOCRACY IN EUROPE AND IN THE UNITED STATES (2005).

[239] Accuracy is here considered as an attribute of an account of facts that is - or ought to be - as close as possible to those facts. Indeed, both evidence scholars and courts are aware that evidence is an argumentative reconstruction of facts and is therefore never a perfect mirror of reality. However, for a party to convince a court of their claim, they have to persuade it that their version of the facts presents the highest degree of probability. As Walton put it, "[it is] the initial body of data or 'facts' used as premises, the inferences used to draw conclusions from these premises, and the use of a chain of such inferences to draw out a line of reasoning (chain of inferences) that makes more (or less) probable some proposition that is in doubt, and which needs to be settled." *See* Douglas Walton, LEGAL ARGUMENTATION AND EVIDENCE 106 (2002).

[240] Carol L. Silva & Hank C. Jenkins-Smith, *The Precautionary Principle in Context: U.S. and E.U. Scientists' Prescriptions for Policy in the Face of Uncertainty*, 88 SOC. SCI. Q., 640, 649 (2007).

2. Relevance Inferred from Policy Objectives

It is hardly a surprise that the prioritization of risks for the purpose of regulation is a case of the policy car leading the scientific horse.[241] As most things that capture regulators' attention, risks are also politicized and are subject to selective hearing.[242] This is, no doubt, due to the "multiple interests which bore upon policy-makers, as well as the extent to which policy-makers [are] charged with directly implementing certain decisions as opposed to making guidelines which are implemented by others."[243] In light of this, 'relevance' in the context of risk regulation does not only answer the question '*which evidence of risks is relevant?*' but also '*which risks do policy objectives point to as relevant?*' In the EU, this policy-based prioritization of risks to be regulated is visible in environment and health which, though substantially different from new technologies like AI, are nonetheless useful referents in our analysis of the Union's overall approach to the risk regulation.

Integral to the Internal Market, environment, health safety and consumer protection are listed, in Article 114 of the Treaty on the functioning of the EU (hereafter TFEU), as areas covered by measures on the approximation of laws, ideally warranted by 'new developments' and supported by 'scientific facts.'[244] More

---

[241] Judith A. Curry, *Statement to the Committee on Science, Space and Technology of the United Sates House of Representatives*, Hearing on "The President's UN Climate Pledge" (Apr. 15, 2015) cited in Lucas Bergkamp, *The Reality of Risk Regulation*, 8 EUR. J. RISK REG. 56, 59 (2017). This study will not discuss in detail the issue of whether it is required, or even advised for regulators to have a strong 'bureaucratic grip' on scientific research. Of course, if regulatory decisions lead to inadequate risk prioritization, the science-policy interrelationship is not functional, with a possible effect of hindering scientific research on risks that should be prioritized and therefore, deserve more scientific and regulatory attention..

[242] Patrick R. Brown & Anna Olofsson, *Risk, Uncertainty and Policy: Towards a Social-Dialectical Understanding*, 17 J. RISK RES. 425, 427 (2014).

[243] *Id.* at 428 (*emphasis added*).

[244] Art. 114(3) TFEU: "the Commission, in its proposals envisaged in paragraph 1 concerning health, safety, environmental protection and consumer protection, will take as a base a high level of protection, taking account in particular of any new development based on scientific facts. Within their respective powers, the European Parliament and the Council will also seek to achieve this objective."

specifically in the field of environment, Article 191 TFEU states that the Union's action shall "aim at a high level of protection" and be based on the precautionary principle and on the principles that preventive action.[245] In preparing its environmental policy, said Article further states that the Union shall take into account the available scientific and technical data, environmental conditions in various regions of the EU, the potential benefits of action or lack thereof, the economic and social development of the Union and its regions.[246] This provision can be interpreted as providing a list of possible sources of relevant knowledge about risks. Since it neither prioritizes scientific evidence over political and economic factors, nor does it state that the criteria mentioned are cumulative, it seems to reveal the reticence of the Treaty's drafters to impose rigid constraints on the Union legislature's freedom to decide which evidence to rely on when making regulatory choices. Indeed, the decision to regulate risks ultimately boils down to an assessment of *necessity* which, according to the Court of justice of the EU (hereafter, the CJEU)[247], is performed through the exercise of considerable institutional discretion. In *Enviro Tech*[248] e.g., the Court recognized the breadth of this discretion, "in particular as to the assessment of highly complex scientific and technical facts in order to determine the nature and scope of the measures which [the EU institutions] adopt."[249] In such cases, the scope of the judicial review is limited to the verification of manifest errors of appraisal or misuse of powers, given that by virtue of the Treaties' provisions on the EU's institutional framework, the CJEU "cannot substitute its assessment of scientific and technical facts for that of the institutions on which alone the Treaty has placed that task."[250]

---

[245] Art. 191(2) TFEU.
[246] Art. 191(3) TFEU.
[247] As per Art. 19(1) of the Treaty on the EU (TEU), the CJEU includes the European Court of Justice (ECJ), the General Court and specialized courts. However, for convenience, when referring to the ECJ and the ECJ's caselaw we will use the 'CJEU' reference.
[248] *Enviro Tech (Europe) v. Belgian State*, case C-425/08, CJEU. (2009), EU:C:2009:635.
[249] *Id.*, para. 47.
[250] *Id.*

If the Union's discretion in the identification of risks is broad, policy-driven and virtually immune to judicial review, does scientific evidence matter at all? The short answer is: yes, but it is not everything. A general observation we could make on the method underlying the EU's risk regulation is that it aims at striking the 'right' balance between evidence-based and policy-based considerations, though a state of perfect equilibrium between the two is seldom achieved.

Bearing in mind the issue of relevance, the following dilemma arises: when the existence of a risk is not strongly supported by scientific and/or statistical evidence but animates a public debate, should it still warrant a regulatory response? Ideally, science and public opinion would go hand in hand. In reality, perceptions of risks may pressure regulators into regulating,[251] even if this translates to a relative disregard of (strong) available scientific evidence. This caution-superseding-science trend can certainly be detected in the EU. In the *NFU* case *e.g.*, the CJEU stated that "where there is uncertainty as to the existence or extent of risks (...), the institutions may take protective measures *without having to wait until the reality and seriousness of those risks become fully apparent.*"[252] In a similar vein, the EC's Communication on the precautionary principle states that "what is an 'acceptable' level of risk for society is *an eminently political responsibility.* Decision-makers faced with an unacceptable risk, scientific uncertainty and public concerns have a duty to find answers."[253] Based on this statement, the EC's approach can be described as *in dubio pro libertate* (as opposed to *in dubio pro reo*): persistent doubts do not seem to be resolved by further exploration of reality but by relying on political providence and foresight (as expressed through the exercise of legislative discretion).

In this context, what exactly motivated the EC to submit the AI Act proposal? In the 'Reasons' Section of the latter, the EC stressed, *inter alia*, that the AI Act "*delivers on the political commitment* of

---

[251] *See e.g.*, Emily Hammond, *Nuclear Power, Risk, and Retroactivity*, 48 VAND. J. TRANSNAT. L. 1059 (2015).

[252] Case C-157/96, National Farmers' Union et al., 1998 E.C.R. I-2259.

[253] EUR. COMM'N , *supra* note 221 at 3 (emphasis added).

President von der Leyden"[254] and to explicit requests from the European Parliament and the European Council for a well-functioning internal market for AI systems where both benefits and risks are adequately addressed at the level of the EU.[255] Evidence (in the form of, say, statistical data) is hardly mentioned. The EC stated - as if in passing - that, following the publication of the White Paper on AI,[256] a broad stakeholder consultation confirmed that stakeholders "were largely supportive of regulatory intervention to address the challenges and concerns raised by the increasing use of AI."[257] In the 'Reasons,' greater emphasis is placed on the already accomplished work on AI by the Union institutions or *ad hoc* entities like the HLEG. Indeed, the European Council's 2017 conclusions[258] do stress the "urgency" of addressing 'emerging trends' such as AI and blockchain. While these trends can legitimately be seen as intrinsic to the EU's Digital market, there is - again! - no reference to any specific evidence that might uncover the ways in which AI is actually 'risky.' The series of political and economic reasons that underly the AI Act are embedded in the usual 'value rhetoric' which tends to legitimize legislation by the need to safeguard the Union's fundamental rights and values.[259]

The prevailing market rationale of the AI Act also transpires from the EC's justification of the legal basis of its proposal *i.e.,* Article 114 TFEU.[260] As already mentioned, this provision explicitly

---

[254] EUR. COMM'N, *supra* note 29 at 1.
[255] *Id.*
[256] EUR. COMM'N, *supra* note 135.
[257] EUR. COMM'N, *supra* note 29 at 1.
[258] Regulation (EU) 2016/679 of the European Parliament and the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation), 2016 O.J. L 119/1.
[259] EUR. COMM'N, *supra* note 29 at 1 ("rules for AI available in the Union market or otherwise affecting people in the Union should therefore be human centric, so that people can trust that the technology is used in a way that is safe and compliant with the law, including the respect of fundamental rights,").
[260] In addition to Article 16 TFEU given the AI Act contains specific rules related to the protection of personal data, especially the restricted use of AI systems for "real-time" remote biometric identification in publicly accessible spaces for the

requires that policy in the areas of health, safety, environment and consumer protection have to be based on 'scientific facts.' In justifying its choice of said Article as legal basis for the AI Act, the EC presented two arguments. First, it underlined that "the primary objective of the proposal is to ensure the proper functioning of the internal market by setting harmonized rules."[261] Second, it interpreted the Member States' ambitions to enact horizontal or sectoral AI regulations as a threat of both fragmenting the Internal Market and substantially diminishing legal certainty for both providers and users of AI systems.[262] Though *prima facie* plausible, it would have, no doubt, been useful for the EC to back its observations with some statistical data on, say, the number and types of policy instruments that the Member States are considering.[263]

It follows that neither the reasons stated for the submission of the AI Act proposal, nor those for the selection of its legal basis - both of which include obligations to gather evidence - allow us to discern the process through which the EC selected the risks addressed in said proposal. The relative absence of empirical data among the factors having shaped the AI Act, seems to confirm that policy objectives (such as those associated with the Digital Single Market) more so than fact-of-the-matter evidence were *the* criterion in deciding which AI systems and corresponding risks were regarded as relevant. The rationale of the process of risk-identification for the purpose of the AI Act thus becomes clear: *relevant risks were those the EC said they were.*

However, discretion, in the context of risk-identification, should not be understood as arbitrariness. The CJEU had stated that "the choice of the appropriate legal basis has constitutional significance"[264] and it "must be based on *objective factors which are amenable to judicial review.* Those factors include in particular the

---

purposes of law enforcement.

[261] EUR. COMM'N, *supra* note 29 at 6.

[262] *Id.*

[263] For example, the OECD AI Policy Observatory is a platform which includes a complete list of regulatory strategies, proposals and instruments on AI per each member of this Organization. *See* OCED, www.oecd.ai (last visited Apr. 1, 2023).

[264] Opinion 1/08, 2009 E.C.R. I-11171.

aim and content of the measure."[265] As will be argued further,[266] the CJEU's review of the 'objective factors' underlying Union legislation includes the compliance of the Union's legislature with obligations - like those in Article 114 TFEU - to gather some type of evidence on the risks it chooses to regulate. Indeed, though the policy car may lead the scientific horse, the actuality, magnitude and imminence of risks should have - at least a minimal - rooting in reality. The above-mentioned idea of balance can be reminded here: while policy may often override evidence in identifying relevant risks, it should not altogether set it aside. On the contrary, policy-induced relevance should ideally be *corroborated* with fact-induced relevance. For example, battling against algorithmic biases is not only warranted by policy objectives like data protection, privacy and non-discrimination; it should also be confirmed by *sufficiently probative* (*i.e.,* sufficiently numerous and convincing) examples of algorithms displaying various unfair biases.

In addition to the political importance of not turning a blind eye on reality when identifying risks, an epistemic reason should also be stressed: relevance, in the stage of risk identification, contributes to formulating the basic inquiry (*e.g.* are facial recognition systems a risky branch of AI?) which, in turn, shapes the overall process of discovery (e.g. which harms are facial recognition systems likely to cause?).

Just as risk-identification is - in principle - not a freestyling exercise but a balanced act of policy-meets-science, risk characterization should be an epistemically sound process, conducive to *proper* knowledge - as opposed to speculation - of relevant risks. In other words, risks-characterization should yield *epistemically valid* knowledge, the formation of which will be discussed in the following Section.

---

[265] *See* Case C-268/94, Portugal v. Council, 1996 E.C.R. I-6216.
[266] *See infra*, Section V.

IV. TRANSLATING KNOWLEDGE OF FACTS INTO POLICY: THE IMPACT OF RISK-CHARACTERIZATION ON THE DESIGN OF 'ADEQUATE' REGULATORY FRAMEWORKS

Under the assumption that AI-related risks had been properly identified, their characterization can be performed. Though several reasons can be cited on why characterizing risks is epistemically challenging, we shall focus on one: conclusiveness, understood as *logical infallibility*.[267]

When we assess the conclusiveness of evidence gathered, we tend to look at its probative value.[268] The reasoning here can be represented through a premise-to-conclusion schema: if the 'knowledge' about risks is, in essence, a set of *logical conclusions* derived from evidence gathered, the latter - acting as *logical premises* - should present certain features in order to warrant, what scholars call, *legitimate* inferences.

To better understand which features reinforce the knowledge-inducing value of the facts-as-premises and which inferences can be viewed as 'legitimate,' evidence theory can - again - be a useful point of reference.

In the stage of risk characterization, it is specifically theory of indirect evidence (in particular, presumptions) that provides the most

---

[267] Logical infallibility will be understood as the attribute of evidence and propositions made on the basis of that evidence, when maintaining those propositions as true is more plausible than defeating them with contrary evidence and propositions. The factors that decide in favour of such plausibility are relative to the goals pursued by such propositions, namely, goals pertaining to epistemic cognition (getting to the truth of a matter), goals pertaining to practical cognition (making the best choice in given circumstances), the needs of the agents (finding the solution in time, without using too much cognitive resources, etc.), the kind of interaction in which they are involved, and social values, as well as communal customs and shared practices. *See* Douglas Walton and Giovanni Sartor, *Teleological Justification of Argumentation Schemes*, 27 ARGUMENTATION 111, 112 (2013).

[268] The probative value is, in essence, the aptitude of an item of evidence to support a proposition, by virtue of its properties (trustworthiness of the source, type of evidence at issue) and in relation to the overall information presented before a court in the context of a trial. *See* Richard D. Friedman, *Conditional Probative Value: Neoclassicism without Myth*, 93 MICH. L. REV 439 (1994).

useful analytical framework for examining the premise-to-conclusion passage in the context of risks. The reason to refer to presumption theory is that presumed facts are, like risks, evidence of the 'unknown' or the 'persistently uncertain.' Roughly speaking, to *presume* is to depart from what the available evidence reveals as established in view of drawing probable and/or plausible[269] conclusions of that which we seek to ultimately prove. The parallelism with risk theory is not difficult to grasp. Since *conclusive* knowledge of risks is never possible - because direct experience of risks is lacking - assertions about risks must, like presumed facts, qualify as some form of *valid* knowledge *i.e.,* comply with criteria that allow to distinguish *correct*[270] risk-representations from risk-misrepresentations. If a given risk representation complies with preestablished standards on accuracy, namely defined by the CJEU, it can legitimately warrant a policy response (**A**).

While epistemically valid knowledge of risks is usually paramount for the design of risk regulation, such knowledge seems to play a rather ancillary role in the design of the AI Act. This is no doubt due to the fact that this instrument gives, above all, a tangible normative expression to fundamental values and rights thought as most adequate to protect individuals in particular from 'high-risk' AI systems. True as this may be, the question must be raised of whether the AI Act can be assimilated to 'standard' risk regulation or if, by reason of its focus on fundamental rights, it is a *sui generis* risk-regulating policy that, given its rooting in rights/values as opposed to reality, *does not need* to overly or primarily rely on empirical knowledge (**B**).

---

[269] The interrelationship between probability and plausibility in evidence theory - as transposed into risk-theory - will be discussed in more detail *infra*, Section IV.

[270] 'Correct' risk representations will be understood as empirically adequate (and, in the evidentiary jargon, 'accurate') accounts of the risks concerned.

A. Knowledge of Facts, Paramount in Shaping 'Standard' Risk Regulation

Evidence scholarship allows us to make the preliminary observation that it is *consistency* that warrants the so-called internal validity of knowledge. The likelihood that an assertion be viewed as valid (or true) is often determined by its relationship to assertions already falling under the 'knowledge' category. In the process of inferring conclusions about risks, the 'strength' of the inference is determined by the degree of consistency of its premises, the rule of thumb being that correct (or consistent) premises are more likely to be conducive to correct (relatively infallible) conclusions (**1.**). The nature and features of the premises not only warrant the 'correctness' of a given risk characterization, but they also increase its acceptability: the stronger the evidence establishing a harm's likelihood, the stronger the *reasons to accept* that evidence as 'true' (or accurate) knowledge (**2.**).

1. Probative and Consistent Premises...

In evidence theory, the criterion of consistency of premises (or basic facts) derives from the so-called best evidence rule, the gist of which was summarized by Morgan: "the highest degree of probability must govern [courts'] judgment; and it necessarily follows, that they ought to have before them *the best evidence of which the nature of the case will admit.*"[271]

Seen through the prism of *conclusiveness*, Morgan's principle can be understood as follows: if the knowledge of relevant facts (say, risks) cannot be conclusive, at least the facts from which they are inferred ought to be. From the perspective of *consistency* however, Morgan's principle can be understood as follows: high probability of the knowledge of unknown relevant facts can be achieved if the known facts (*i.e.*, facts for which evidence is available) are consistent in pointing toward that knowledge. But how can the consistency of

---

[271] 1 John Morgan, *The Law of Evidence*, *in* ESSAYS UPON THE LAW OF EVIDENCE, NEW TRIALS, SPECIAL VERDICTS, TRIALS AT BAR AND REPLEADERS 2-3 (1779).

the premises be assessed? In evidence scholarship, this issue has been an object of study (as well as dispute) of various strands focused on presumptions.

Historically, the concept of presumption has been viewed as an evidentiary travesty. To presume is to take something as true, without having the certainty that it actually is.[272] Presuming innocence, good faith or paternity may seem plausible in view of safeguarding social norms, respectively, the equality of arms, the passing of conventions and the stability of family as an institution. However, neither of these facts (innocence, good faith, paternity) are strictly speaking evidence because they do not conclusively prove anything; they are, at best, probabilities considered 'as if'[273] they established some factual truth. Herein lies the great blasphemy of presumptions in the context of general evidence theory: though substantively not evidence, they are fictitiously considered as such and stand as true so long as they are not defeated.[274]

One of the main epistemic issues in presumption theory - as a template for our understanding of risk characterization - has been that of the *conditions* under which a fact can be presumed. In presumptive reasoning, a typical schema includes a distinction between that which is proven and that which proves *i.e.,* the fact for which direct evidence is sought but is unavailable (the *probandum*) and the facts (the *probans*) from which the *probandum* can, *faute de*

---

[272] Presumption can be defined as indirect evidence, the object of which is a known fact which, in a context of normality, is a probable and plausible substitute to a fact for which direct evidence is sought, but is difficult or impossible to produce; Ljupcho Grozdanovski, LA PRÉSOMPTION EN DROIT DE L'UNION EUROPÉENNE 31 (2019).

[273] There is, indeed, an element of fiction in the concept of legal presumption. Though it cannot qualify as evidence, there are various procedural, legal, political and societal reasons why some presumptions (innocence for instance) are considered *as if* they were evidence. The 'as if' expression here is borrowed from Weihinger's seminal study on fictions.

[274] In this study, we focus on rebuttable presumptions. Irrebuttable presumptions, though recognized in many legal systems, have caused much debate on whether they are evidentiary devices or legal fictions. This debate is beyond the scope of this study. May it suffice stressing that, in the context of risk theory, we will consider presumptions and presumptive reasoning as *prima facie* defeasible *i.e.,* open to rebuttal.

mieux, be inferred. As mentioned earlier - and bearing in mind the best evidence rule - the *probandum* will never be certain, but the *probans* should be.[275] .

Against the backdrop of the *probandum/probans* pattern, the soundness of the inference that yields the presumed fact derives from the probative value of the so-called basic facts (*indicia*). The probative value depends on several factors: the quality of the data, the trustworthiness of its sources, the absence of contradictory facts. The CJEU's caselaw on so-called evidentiary presumptions gives hints on specific conditions that should be met for the evidence to facts to act as acceptable premises. Though there are several types of presumptions in EU law,[276] a constant throughout the caselaw has been that the *indicia* from which presumptions derive should be *clear*, *sufficient* and *concordant*.[277]

Structurally speaking, the *probandum/probans* schema in presumption theory can also be discerned in risk characterization: the degree of probability of a given risk is dictated by the nature and probative value of the facts established as premises (the equivalent of *indicia*).[278] Regarding the latter, a parallel can be drawn between the CJEU's triple consistency requirement (clarity, sufficiency,

---

[275] As Grossen put it, the truth of that which is signified (*signifié*) is determined by the certitude of that which 274ignifica (*274ignificant*) ; *see* Jacques-Michel Grossen, LES PRÉSOMPTIONS EN DROIT INTERNATIONAL PUBLIC 161 (1954).

[276] The taxonomy of presumptions in the EU law depends on whether they are formed on the basis of *indicia* presented by the parties (for example tacit collusion) or whether they are established as *prima facie* evidence that do not require prior proof (for example innocence or good faith). For an in-depth study on this point, *see* Grozdanovski, *supra* note 274.

[277] *See* Case T-123/99, JT's Corp. Ltd. V. Comm'n, 2000 E.C.R. II-3272 ¶ 58; these requirements apply *mutatis mutandis* to so-called cognitive presumptions that is, presumptions that epistemically contribute to knowledge-construction. In essence, the classification of presumptions as evidentiary (or practical) or cognitive presumptions depends on the type of deliberation they are used in. Evidentiary presumptions play a role in procedural truth-finding. Cognitive presumptions play a role in epistemic truth finding. *See*, on this point, Petar Bodlovic, *On the Differences Between Practical and Cognitive Presumptions*, 35 *Argumentation* 287 (2021).

[278] *See generally* Gabriele Usberti, *Inference and Epistemic Transparency*, 38 TOPOI 517 (2019), emphasis added.

concordance) in the context of adjudication and the requirements defined by the Court for the evidence used as basis for inferences about risks. In the *Monsanto*[279] case *e.g.*, the French authorities had suspended the licence to place in the market genetically modified maize which had previously been notified under the relevant Union legislation as an 'existing product' in the market. The suspension measure was based on evidence revealing a threat of this product to human health. Monsanto launched an action for annulment before the French courts which, in turn, inquired the CJEU on the scope of the Member States' discretion in the assessing and managing of risks in areas covered by EU legislation.

It is interesting to note that in this case, the CJEU echoed Morgan's high probabilities/best evidence rule in a two-step reasoning. First, it clarified the meaning of, and the evidence required for the establishment of a 'serious risk' - the latter being understood as "a *significant risk which clearly jeopardises human health, animal health or the environment* [and which] must be established on the *basis of new evidence based on reliable scientific*

---

[279] Cases C-58/10 to C-68/10, Monsanto SAS et al., ECLI: EU:C:2011:553 (Sept. 9, 2011); it should be stressed that the Monsanto caselaw is a milestone in the context of EU risk regulation, namely because the CJEU specified the requirements for the epistemic validity of scientific evidence that fall on the EU institutions (in *Monsanto*, the EC). There are, however, cases prior to *Monsanto* in which the Court outlined the evidentiary requirements that also framed the Member States' (discretionary) assessment of risks. In a *Commission vs France* case e.g., infringement proceedings were brought on the grounds that French legislation did not include any provisions on the mutual recognition of certain foodstuff manufactured and marketed in other Member States. The justification was that certain foodstuff were considered to raise a health risk. In considering that France had failed to fulfill its obligations under the Treaty provisions on the free movement of goods, the CJEU specified that a decision to prohibit the marketing of fortified foodstuff can be adopted "if the alleged real risk for public health *appears to be sufficiently established on the basis of the latest scientific data available at the date of the adoption of such decision.*" In this context, "the object of the risk assessment to be carried out by the Member State is to appraise the *degree of probability* of harmful effects on human health from the addition of certain nutrients to foodstuff and the seriousness of those potential risks." This means that "the *risk assessment cannot be based on purely hypothetical considerations*"; *see* C-24/00, Comm'n v. France, 2004 E.C.R I-1306 ¶ 55.

data."[280] Second, the Court clarified the conditions under which public responses to 'serious risks' can be warranted: "protective measures (...) cannot validly be based *on a purely hypothetical approach to the risk, founded on mere assumptions which have not yet been scientifically verified.* On the contrary, such protective measures, notwithstanding their temporary character and even if they are preventive in nature, may be adopted only if they are *based on a risk assessment which is as complete as possible* in the particular circumstances of an individual case, which indicate that those measures are necessary."[281]

Although in *Monstanto*, *completeness* as an evidentiary standard was required from the Member States when acting in an area covered by EU law, other CJEU caselaw shows that the same requirement extends to risk assessments performed by the Union's institutions. For example, in one of the mad cow disease cases,[282] annulment proceedings were brought against a Commission's decision on the export of bovine meat from the UK, upon discovery of new evidence on the transmissibility of the mad cow disease to humans. In favour of the export measures, the Council argued that the Commission had "based its decision on the *best available technical and scientific data*, by means of the obligatory consultation of the Standing Veterinary Committee and also by exercising its option to consult the Scientific Veterinary Committee."[283] Agreeing with the Council's view that the Commission applied the 'best evidence rule' properly, the CJEU ultimately dismissed the action brought by the UK authorities.[284] Similarly, in *Afton Chemical Ltd*,[285] a UK-based producer of metallic additives challenged, before the national courts, the insertion of manganese-based additives in the so-called

---

[280] Cases C-58/10 to C-68/10, Monsanto SAS et al., ECLI: EU:C:2011:553 ¶76 (Sept. 9, 2011).
[281] *Id.* at ¶ 77 (emphasis added).
[282] C 180/96, U.K. v. Comm'n, EU :C :1998 :192 (May 5, 1998).
[283] *Id.* at ¶42.
[284] *See also* Giandomencio Majone, *The evolution of the regulatory state: from the law and policy of antitrust to the politics of precaution, in* ROUTLEDGE HANDBOOK OF RISK STUDIES 216, 222 (Adam Burgess et al. eds., 2016).
[285] C 343/09, Afton Chem. Ltd. EU:C:2010:419 (Jul. 8, 2010).

greenhouse Directive.[286] The UK courts then referred questions for preliminary ruling, inquiring the CJEU on the validity of several of the Directive's provisions considering the claimant's argument that the EC's assessment of the toxicity of said substances was based on arbitrary evaluations of the available evidence.[287] In its assessment of a manifest error in the EC's appraisal of the facts, the CJEU's reasoning was initially 'standard,' in the sense that the Court confirmed the broad discretion that the Union legislature enjoys in attaining desired levels of protection of environment and human health.[288] After this pedagogical reminder however, the Court stressed that its review does is not only limited to the nature and scope of the measures to be taken "but also, to some extent, to the *finding of the basic facts*."[289] The CJEU confirmed that the Union institutions must be able to show if, in the exercise of their discretion, they took into consideration "*all the relevant factors and circumstances of the situation the act was intended to regulate.*"[290] In *Afton Chemical Ltd*, the EC had relied on a report drafted by the European Parliament's Committee on Environment, Public Health and Food Safety which itself relied on various studies including the Canadian Vehicle Manufacturers' Association. Since the Council also relied on data produced by the Council on Clean Transportation, the Court ultimately considered that the evidence gathered was, indeed, sufficient and no manifest error of assessment had been made.

When examined through the lens of presumption theory, the cited cases allow us to posit that the formation of presumptions of risks for the purpose of policy follows similar stages as those followed in dispute resolution contexts. When presuming, both

---

[286] Directive 2009/30 of the European Parliament and of the Council of 23 April 2009 amending Directive 98/70 as regards the specification of petrol, diesel and gas oil and introducing a mechanism to monitor and reduce greenhouse gas emissions and amending Council Directive 1999/32/EC as regards the specification of fuel used by inland waterway vessels and repealing Directive 93/12/EEC, OJ (L 140) 88.

[287] C 343/09, Afton Chem. Ltd. EU:C:2010:419 ¶ 29 (Jul. 8, 2010).

[288] *Id.* at ¶ 32.

[289] *Id.* at ¶ 33.

[290] *Id.* at ¶ 34.

courts and regulators seem to apply criteria - rooted in the best evidence rule - that essentially allow to 'test' if the probative force of the basic facts is such that they warrant specific, legitimate inferences.

However, extensive knowledge of the basic facts is not the knowledge sought: data showing which chemicals are found in cigarettes is not yet knowledge proper of the types of health risks those chemicals entail. What ultimately matters is *how* a fact-finder interprets the facts so as to gain adequate understanding of 'what they are saying.' For example, in *Commission v. CHS*,[291] an appeal was brought against a ruling from the Court of First Instance (hereafter, CFI), by which the latter annulled an EC decision on the withdrawal of marketing authorisations of medicinal products containing certain substances. The reason for this withdrawal was that those substances had been signalled by the Committee for Proprietary Medicinal Products (CPMP) as potentially harmful for human health. The Commission argued that, by not taking under consideration the CPMP's views, the CFI exhibited "a *fundamental misunderstanding* of the procedure leading to the adoption of the contested decision,"[292] and of the assessment of the degree of harmfulness of the substances concerned.[293] In light of these arguments, the President of the CJEU did indeed order that the CFI's ruling be set aside, the takeaway being that fact-finders 1. must consider *all of the relevant indicia* (under the guidance of the 'best evidence rule'); 2. must develop a proper understanding of what those *indicia* actually establish. The issue that then arises is: under which conditions is the understanding of basic facts 'proper'? To address this issue, basic understanding of knowledge *construction* is necessary.

---

[291] *See generally* C 471/00, Comm'n v. CHS, EU:C:2001:218 (Apr. 11, 2001).
[292] *Id.* at ¶ 32.
[293] *Id.* at ¶ 63; s*ee also* Order of the President of the Court of 11 April 2001, C 474/.0.0 P® Comm'n v. Bruno Farmaceuitici SpA and Others EU:C:2001:219 and Order of the President of the Court of 11 April 2001, C 475/00, Comm'n v. Häseler GmbH, EU :C :2001.

2. ... Yielding 'Acceptable' (and Regulation-Worthy) Knowledge of Risks

Although the ambition of this study is not to give a detailed outline of rich and complex knowledge theories, may it suffice stressing that we shall understand the concept of knowledge as a set of *justified beliefs about reality*, which are *accepted as true* so long as they are not defeated by subsequent evidence warranting their revision.[294] Based on this definition, we can posit that knowledge-constituting beliefs are normative while remaining defeasible. This means that what we take to be scientifically established is, in essence, a set of practical or effective *certainties* which are subject to continued (re)evaluation.[295] The possibility of so-called belief revision is a sign that knowledge (of any kind) is in fact tentative,[296] since new evidence can always push us to "accept new beliefs that should rationally be accepted, or to abandon beliefs, because they should rationally be rejected."[297]

Bearing in mind the knowledge-belief interrelationship, a specific issue arises in relation to risks: given that - as already argued - knowledge of risks is by nature presumptive, how do beliefs play into inferring (*i.e.,* presuming) a specific knowledge about risks? An overview of presumption theory, as an epistemic referent for risk theory, revels two sets of properties of 'legitimate' presumptive knowledge: on the one hand, its *objective* infallibility of the conclusions derived from a set of *indicia*, on the other hand, the more subjective *acceptability* of those conclusions.

Knowledge of risks is 'objectively' infallible if it is supported by high probabilities derived through standard models of probabilistic reasoning.[298] In presumption theory, presumed facts

---

[294] *See e.g.,* Steven L. Reynolds, KNOWLEDGE AS ACCEPTABLE TESTIMONY 27 (2017). Defeasibility is, in essence, the capacity of an argument to render void (or unjustified) another argument. *See also* Jacob Hage, STUDIES IN LEGAL LOGIC 9 (2005).

[295] Nicholas Rescher, DIALECTICS - A CONTROVERSY-ORIENTED APPROACH TO THEORY OF KNOWLEDGE 91 (1977).

[296] Hage, *supra* note 296 at 13.

[297] *Id.* at 10.

[298] For an analysis on the evolution of probabilistic reasoning as applied in

can be regarded as evidence if they are both *probable* and *plausible*.[299] Probability essentially pertains to the statistical likelihood of a given fact, the standard 'probabilistic' principle being that repetitive facts are likely to be true (*quod plerumque fit*). The presumption of paternity is a good example here: though not all married men are the biological fathers of their spouses' children, it is a historically confirmed fact that, in the majority of cases, a woman's spouse is likely to be the father of her child. However, while statistical probability allows to quantify the properties (say, the frequency) of a given phenomenon, it does not explain why the knowledge of that phenomenon should be shared in the public sphere. Enter plausibility as the twin-criterion of (statistical) probability.

According to certain strands of evidence and argumentation scholarship,[300] it is plausibility, more so than probability *stricto sensu*, that truly sanctions the validity of presumptive inferences. Rescher[301] seminally underlined the interrelationship between the 'purely' probabilistic and axiological dimensions of presumptions: to say that a proposition is backed by evidence is to claim that it is *prima facie* supported by a standard epistemic source (in a narrow sense), such as sense-perception, memory, testimony, expert-testimony, or common knowledge.[302] By contrast, principles render propositions *plausible* on grounds of simplicity, uniformity, or *normality*, and they come to the fore when presumptive status cannot be assigned on evidential grounds.[303]

Plausibility plays an important role in both presumption theory

---

evidence theory, *see* Lorraine Daston, CLASSICAL PROBABILITY IN THE ENLIGHTENMENT 14 (1988).

[299] *See* Ljupcho Grozdanovski, *Le probable, le plausible et le vrai. Contribution à la théorie générale de la présomption en droit*, 84 REVUE INTERDISCIPLINAIRE D'ETUDES JURIDIQUES 39 (2020).

[300] *See generally* Nicholas Rescher, PRESUMPTION AND THE PRACTICES OF TENTATIVE COGNITION (2006); Nicholas Rescher, PLAUSIBLE REASONING: AN INTRODUCTION TO THE THEORY AND PRACTICE OF PLAUSIBILISTIC INFERENCE (1976).

[301] *See Id.*

[302] *Id.* at 6-7.

[303] *Id.*

and risk theory because it enhances the *acceptability* of the knowledge of facts. As Freeman put it - in the context of informal logic - "our ultimate interest is normative. We are concerned not just with what accepting that p means but with the *conditions under which acceptance is justifiable.*"[304] Freeman thus hinted to the something-more dimension of presumptive inferences: while it remains that statistical probability, as an accurate - or at least, coherent - connection between consistent premises reinforces the logical soundness of certain inferences, other factors influence their legitimacy. A key notion here is that of *context*, understood as an ideological organizing framework comprised of "a *relatively coherent set of assumptions* about the functioning of economic, political and social institutions."[305]

Context informs us on what can be plausibly held as true because, in light of overarching social values and narratives, it points toward not only probable outcomes, but *desired ones*.[306] As Rescher put it, "with probability we ask 'how many alternatives does the thesis engross in its content?'; with plausibility 'how reputable a source or principle speaks for the thesis?' In the former case we orient ourselves towards the content of the thesis, in the other towards its probative credentials."[307]

The *context of normality* - in the sense of Rescher's work - is an important idea to keep in mind. In models of reasoning such as non-monotonic logic,[308] the so-called *normality premise* is often seen as

---

[304] *Id.* at 4 (emphasis added).

[305] Katherine E. Smith, *Beyond 'Evidence-Based Policy' in a 'Post-Truth' World: the Role of Ideas in Public Health Policy*, 33 *in* SOCIAL POLICY REVIEW: ANALYSIS AND DEBATE IN SOCIAL POLICY 151, 154-155 (2017).

[306] Mauro Zamboni, *Legislative Policy and Effectiveness: A (Small) Contribution from Legal Theory*, 9 EUR. J. RISK REG. 416, 423 (2018).

[307] Rescher, PLAUSIBLE REASONING *supra* note 302 at 28.

[308] Non-monotonic logic is defined in opposition to so-called monotonic logic; *see* Hage, *supra* note 296, at 8. (Hage suggested the following definitions: "a system of logic is monotonic, if and only if it is such that if a set of sentences S' is a superset of S, the set of conclusions C' that follow according to this logic from S' is a superset C of conclusions that follow from S. A system of logic is non-monotonic if and only if it is not monotonic.") *See* Hage, *supra* note 296 at 8. Hage suggests that so-called non-monotonic logic systems are a useful model for the analysis of legal reasoning, because the latter is defeasible in a sense that

a factor enhancing the persuasiveness of arguments.[309] Though much can be said about normality as a concept,[310] its key function is to provide - what we might call - basic *background knowledge*[311] (prior beliefs) which we refer to when assessing if new beliefs are based on findings derived from public, objective and rational discovery processes that "combat bias and irrelevancy."[312]

Like in presumption theory, the probability/plausibility interplay is paramount in setting a standard of acceptance of knowledge in risk theory. Möller *e.g.* observed that "what people take to be an acceptable level of risk is relevant for what level we should accept."[313] Indeed, "even *if we take a certain level of safety as a given*, whether that level is sufficiently small to be acceptable or safe

---

cannot be *prima facie* specified. Though 'defeasibility' and 'monotonicity' are different concepts, it can be argued that the latter is assessed with regard to the former.

[309] Hage, *supra* note 296 at 27 ("non-monotonic logic can (...) only justify a conclusion under the presupposition of a normality hypothesis. Without this hypothesis, the argument is unconvincing. Therefore, such a normality hypothesis should be part of the premises. If the normality hypothesis is added to the premises, the argument becomes deductively valid and the conclusion has become unavoidable for those who accept the premises. For instance, the argument that John is a thief, that, barring exceptions, thieves should be punished and that therefore John should therefore be punished, is defeasible, but can be analyzed deductively by adding the premise that in John's case there is no exception to the rule that thieves should be punished,").

[310] Bernard suggested a distinction between two kinds of normality: the one called *descriptive*, reflecting the social consensus and revealing that which is considered as normal by the society, the other called *dogmatic* which does not pertain to what is observable in reality, but to principles, values and collective dogmas, which shape 'normal' conducts in various situations. *See* Elsa Bernard, LA SPÉCIFICITÉ DU STANDARD JURIDIQUE EN DROIT COMMUNAUTAIRE 37 (2010). Rescher viewed normality as a condition for the acceptance of certain beliefs, arguing that there is "a standing presumption in favor of the normal, usual, customary course of things." *See* Rescher, PLAUSIBLE REASONING *supra* note 302 at 57.

[311] Catarina Dutilh Novaes & Herman Veluwenkamp, *Reasoning Biases, Non-Monotonic Logics and Belief Revision*, 83 THEORIA 29, 32 (2017).

[312] E. Allan Lind et al., *Discovery and Presentation of Evidence in Adversary and Nonadversary Proceedings*, 71 MICH. L. REV 1129 (1973); *see also* Hage, *supra* note 296 at 12 ("[a belief is justified] if and only if it is rational to accept this belief if one accepts (all beliefs in) the belief set").

[313] Möller, *supra* note 57 at 71.

is a further normative question about what we have reason to do."[314]

Möller makes an interesting point: risks are disruptions of *given* standards of safety. This suggests that the knowledge of risks we are *likely to accept* is the knowledge that warrants the safety standard we are *likely to expect*, *ceteris paribus*. Just like presumptions are plausible evidence relative to a given context of normality, the acceptance/expectation correspondence in risk theory also relies on a shared conception of normality. To refer to the already mentioned example of smoking:[315] if a world free of tobacco-related ailments is a normative ideal, smoking is clearly a deviation from that ideal. In practice, all we can aspire to is a *workable* standard of health safety which translates, not to a total ban of cigarette production, but to a decrease of tobacco consumption, which the WHO is pushing for through various regulation incentives.[316] In other words, considering the world we live in, we would be more likely to accept knowledge of risks if that knowledge recommended a safety standard that could be practically upheld in the broader context of principles, values and practices we *already* view as constitutive of a 'normal state of affairs.'

However, compared to smoking as a long-standing health problem, AI is a different 'beast:' as a relatively recent phenomenon, we do not have any general, shared idea on what would be a normal or reasonable expectation of safety in the process of, say, AI use. While smoking has been the object of a solid body of 'hard' scientific evidence, the only empirical evidence we have of AI-related risks includes sporadic court cases and public opinions on series of threats AI is likely to raise.[317] In this context, two questions arise: first, how can we know what AI programming or use practice is *normal*? Or, what is our background knowledge of AI that allows us - as laymen - to accept that, say, biometric identification algorithms are, indeed, high-risk?

---

[314] *Id.* at 71-72.
[315] *See supra*, Section III.
[316] *See* WHO, *supra* note 196.
[317] *E.g.*, District Court of the Hague, 6 March 2020, ECLI:NL:RBDHA:2020:865, available (in English) at https://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RBDHA:2020:1878.

Second, are principles such as transparency and explainability the only ones that ought to set the standards of safety we are likely *to expect*? In the 1985 Product Liability Directive it is stressed that "to protect the physical well-being and property of the consumer, the defectiveness of the product should be determined by reference *not to its fitness for use but to the lack of the safety which the public at large is entitled to expect*; whereas the safety is assessed by excluding any misuse of the product not reasonable under the circumstances."[318] The *reasonable expectation* of safety is assessed by taking into account the intended purpose, the objective characteristics and properties of a given product, as well as of the specific requirements of the group of users for whom the product is intended.[319] However, there are various types of AI systems with diverging 'intended purposes,' 'objective characteristics,' and 'properties'... Given this diversity, what warrants the belief that, in the normal world (where black box cases do happen), transparency is a *reasonable expectation* to have from *all AI systems*?...

With these observations in mind, we can raise the issue of the *rationale* (and rationality) of the AI Act: did the Commission make a *valid inference* from the evidence it gathered prior to submitting the AI Act? The short answer is: no. In its evidence assessment, the EC did not seem to comply with the above-mentioned requirements regarding both the nature of the evidence gathered (completeness, best available evidence) and the inferences based on that evidence (adequate understanding, proper application of models of analysis). The downside of the 'fundamental misunderstanding' of the evidence gathered in view of the AI Act is that the latter - as a policy instrument meant to establish safeguards *vis-à-vis* AI-related harms - may rely on *erroneous* characterization of the risks of those harms, which brings up the following question: if not facts, which were the decisive factors in the design of the AI Act? To answer this question,

---

[318] Council Directive 85/374 of 25 July 1985 on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products, OJ L 210, 7.8.1985, p. 29-33.

[319] CJEU, *Boston Scientific Medizintechnik*, joined cases C-503/13 et C-504/13, EU:C:2015, para. 37.

we should have a closer look at the nature of this instrument and the regulatory approach embedded in its design.

B. Knowledge of Facts, Ancillary in Designing the AI Act

Placed in the context of standard epistemic and legal requirements on 'standard' risk regulation in the EU, the AI Act appears to be *sui generis*, fact-neutral yet fact/risk-regulating instrument. This is due to the fact that the AI Act's design does not seem to follow the usual bottom-up approach under which facts are key sources of information on what policy ought to address. In contrast, the AI Act seems to follow a top-down approach as it brings into normative being an axiological meta-narrative on risk-prevention, canonized by the HLEG's Guidelines on Ethics. In other words - as will be argued - the instrument under consideration does not rely on tangible reality *because it does not need to*; it is a legislative embodiment of a pre-established system of values and principles thought to be most protective against AI-related material harm (*i.e.,* safety and health of individuals, including loss of life and damage to property) and immaterial harm (*i.e.,* loss of privacy, limitations to the right of freedom of expression, human dignity, discrimination for instance in access to employment).[320]

In this context, the uncovering of the *true* rationale underlying the design of the AI Act, requires a clarification on its true normative nature. Our analysis on this point, against the backdrop of relevant doctrines on risk regulation, will allow us to argue that designing an evidence-based instrument on AI was never really the EC's intention, or at the very least, its priority: what seemed to supersede the search for fact/law correspondence in regulating AI was the imperative of defining and solidifying a standard of protection *oriented toward fundamental rights protection* (**1.**). In this context, our normative claim is that the decisive factor underlying the method to the madness in drafting the AI Act was not that it be based on empirical knowledge; but that it be *axiologically congruent* with the foundational values that underly the overall EU approach toward regulating new technologies (**2.**)

---

[320] COM (2020) 65, *supra* note 29 at 10.

1. The *ratio legis*: Reasons for the Fact Neutrality of the AI Act

Considering the arguments developed in the previous Sections of this article, it is clear that the AI Act's fact-neutrality makes it stand out from other 'standard' risk regulation instruments. This fact-neutrality can be explained by the specific approach followed in the process of designing the AI Act's regulatory framework (***i***). This approach, as well as the ancillary role of facts, can be justified by a rather original definition of the notion of risk: the latter does not pertain, as is usually the case, to probable threats of physical harm. Within the meaning of the AI Act, risks are defined more broadly as they relate to threats of violations of fundamental rights (***ii***).

i. *'Fact-neutrality' Explained by the Specific Nature of the AI Act as a 'Risk-Regulating' Instrument*

To understand the *normative rationale* of the AI Act, we should not depart from the standard policy-reacts-to-evidence view,[321] but focus on the policy-as-instrument-of prevention view, under which, risk regulation can be defined as an "organised attempt to manage risks or behaviour in order *to achieve a publicly-stated objective or set of objectives*."[322] The fact that policy objectives, rather than facts, can warrant an instrument qualified as risk-regulating allows us to make a subtle distinction between *risk regulation* proper and a *risk-based approach* to regulation. The two are interrelated but not synonymous.[323] Under the former, risks are a *substantive qualifier* of

---

[321] Mark L. Flear, *Regulating New Technologies: EU Internal Market Law, Risk, and Socio-Technical Order*, *in* NEW TECHNOLOGIES AND EU LAW 74, 98 (Marise Cremona ed., 2017).

[322] Julia Black, *Learning from Regulatory Disasters*, 10 POLICY QUART'Y 3, 4 (2014); despite this quite broad definition of risk regulation, Julia Black - still rightly - emphasizes that "not all regulation is about risk, not all regulation is about economics, and not all regulation is about either of those things, but is about ethical issues, or rights, to name but two". *See also* Black, *supra* note 104 at 305.

[323] Julia Black & Robert Baldwin, *Really Responsive Risk-Based Regulation*, 32 LAW & POLICY 181, 184 (2010).

regulation in the sense that they appear as direct objects of regulatory instruments. Under the latter, risks are an *assessment criterion* used by regulatory agencies in prioritising their regulatory actions:[324] State agencies score the risks addressed by a given regulation (*e.g.* environmental protection regulation) and tailor their enforcement actions in reference to a scale of riskiness, prioritizing the 'riskiest' aspects covered in said regulation (*e.g.* addressing water pollution before street litter) and targeting those through specific enforcement strategies.[325]

Risk-based (approach to) regulation allows risk controllers to operate in priority within sectors where intervention is needed.[326] In her analysis of risk-based (approach to) regulation, Black distinguished three types: the tailoring of regulation to fit the particular risks raised by the behaviour of a particular economic operator (type I), the management of the risk that a regulatory agency may not operate properly due, on the one hand, to its own internal organization (type II) or, on the other hand, to the disobedience of regulated economic operators (type III).[327] In each of the three scenarios, risk-based (approach to) regulation is meant to provide systematised decision-making frameworks and procedures that prioritise both regulatory activities and resources deployment - mainly relating to inspection and enforcement - based on a risk assessment of the targeted economic operators' non-compliance with the regulator's objectives.[328]

Although - or perhaps, because - the distinction between risk regulation and risk-based (approach to) regulation is subtle, they are

---

[324] Julia Black, *The Emergence of Risk Based Regulation and the New Public Risk Management in the UK*, PUBLIC LAW 512, 513 (2005) (explaining that risk may serve as a proxy for decision-making).

[325] Baldwin et al., *supra* note 79 at 281-83.

[326] Philip Hampton, REDUCING ADMINISTRATIVE BURDENS: EFFECTIVE ADMINISTRATION AND ENFORCEMENT 4 (2005).

[327] Julia Black, *Constructing and Contesting Legitimacy and Accountability in Polycentric Regulatory Regimes*, 2 REG. GOV. 137, 138 (2008)

[328] Julia Black, *Managing Regulatory Risks and Defining the Parameters of Blame: A Focus on the Australian Prudential Regulation Authority*, 28 LAW & POLICY 1, 3-4 (2006); *see also* Ortwin Renn & Andreas Klinke, *Risk Governance: Concept and Application to Technological Risk*, *in* ROUTLEDGE HANDBOOK OF RISK STUDIES 204 (Adam Burgess et al., eds., 2016).

not mutually exclusive. In recent EU regulation, the GDPR can qualify as a 'hybrid' of sorts. Indeed, the GDPR qualifies as risk regulation because it aims to mitigate risks that arise from the introduction of information and communication technologies (ICTs) into society.[329] It also qualifies as a risk-based (approach to) regulation because supervisory authorities are required to target "compliance action and enforcement activity on areas of greatest risk."[330] More specifically, data protection officers are held to "have due regard to the risk associated with the processing operations" during their activities and therefore concentrate their efforts where they are most needed (art. 39(2) GDPR).[331] It can be argued that this is not risk-based (approach to) regulation strictly speaking, as this type of regulation addresses government agencies and supervisory authorities, which data protection officers - within the meaning of the GDPR - are not. However, Quelle convincingly argued that, in practice, data supervisory authorities (*i.e.,* State agencies) outsource risk assessments to data controllers,[332] which confirms that the GDPR can, after all, fall under the risk-based (approach to) regulation category.

It is true that the EU legislator included scenarios of high-risk data processing directly in the GDPR (Art. 35(3)). However, it also requires that the data controller carry out a data protection impact assessment (DPIA) whenever a specific processing is likely to entail a high risk of violating the rights and freedoms of data subjects (Art. 35(1) GDPR). If the result of the DPIA shows such a risk, the

---

[329] Raphaël Gellert, *Data protection: A Risk Regulation? Between the Risk Management of Everything and the Precautionary Alternative*, 5 INT. DATA PRIV. L. 3, 3 (2015).

[330] Article 29 Data Protection Working Party, "Statement on the role of a risk-based approach in data protection legal framework", WP 2018, 2014, p. 4; *see generally* Christopher Kunner et al., *Risk Management in Data Protection*, 5 INT DATA PRIV. L. 95 (2015).

[331] For an example of economic operators evaluating themselves their own daily and operational risk, *see* Tony Muschara, RISK-BASED THINKING: MANAGING THE UNCERTAINTY OF HUMAN ERROR IN OPERATIONS (2018).

[332] Claudia Quelle, *Enhancing Compliance under the General Data Protection Regulation: The Risky Upshot of the Accountability- and Risk-based Approach*, 9 EUR. J. RISK REG. 502, 511 (2018).

controller is required to consult the supervisory authority (Art. 36 GDPR). Rather than a general prior notification obligation that would have resulted in the data supervisory authority "sifting through endless notifications," they now "sit back and wait until controllers start a prior consultation – as is mandatory for processing operations which the DPIA reveals to be too risky – of their own accord."[333] Controllers are therefore *qua* agents of the data supervisory authority or - to borrow an expression from Brownsword - "surrogate regulators"[334] leading to the emergence of "deregulation,"[335] "meta-regulation,"[336] or "co-regulation."[337] It is interesting to note that the AI Act follows a similar trend as the GDPR. The Explanatory Memorandum makes multiple mentions of the AI Act putting in place "a proportionate regulatory system centred on a well-defined *risk-based regulatory approach.*"[338]

Against the backdrop of the risk regulation/risk-based (approach to) regulation dyad and bearing in mind the 'hybrid' regulatory model embodied in the GDPR, the AI Act can *prima facie* qualify as *risk regulation*, translating the EC's ambition to mitigate risks that arise from the introduction of AI systems into society. However, a comparative analysis between the governance model in the AI Act and that in the GDPR allows to conclude that the AI Act can also qualify as a risk-based (approach to) regulation, since national supervisory authorities will prioritise their enforcement activity based on the exhaustively enumerated high-risk sectors (Annex III AI Act) where, "given the characteristics of the activities typically

---

[333] *Id.*

[334] Brownsword, *supra* note 79 at 5.

[335] *See e.g.,* DEREGULATION OR REREGULATION? REGULATORY REFORM IN EUROPE AND THE UNITED STATES (Giandomenico Majone ed. 1990).

[336] Fiona C. Simon, META-REGULATION IN PRACTICE: BEYOND NORMATIVE VIEWS OF MORALITY AND RATIONALITY (2017); *see also* Reuben Binns, *Data Protection Impact Assessment: A Meta-Regulatory Approach*, 7 INT. DATA PRIV. L. 22 (2017).

[337] Irene Kamara, *Co-regulation in EU Personal Data Protection: the Case of Technical Standards and the Privacy by Design Standardization 'Mandate'*, 8 EUR. J. L. TECH. 1, 6 (2017).

[338] COM (2021) 206 final, *supra* note 48 at 3 (Explanatory Memorandum), emphasis added.

undertaken, significant risks can be expected to occur"[339] – unacceptable risk AI systems having been prohibited and non-high-risk AI systems being subject to voluntary endorsement mechanisms.

The reasons behind this 'hybrid' qualification were stated in the White Paper on AI. First, it ensures that "the regulatory intervention is targeted on the areas where, generally speaking, risks are deemed most likely to occur."[340] Second, although the EC suggested that only high-risk AI systems be subject to prior conformity assessments (Art. 17 AI Act), it stressed that such assessments are "without prejudice to monitoring compliance and *ex post* enforcement" of high-risk and non-high-risks AI systems, "although the high-risk nature of the applications at issue may be reason for the competent national authorities to give particular attention to the former."[341] We find here a clear reference to the combined regulatory approach followed by the EC which consists in 'merging' risk regulation with risk-based (approach to) regulation.

Though the GDPR and the AI Act are comparable in terms of their regulatory design, they differ in the *mise en oeuvre* of the risk-based (approach to) regulation aspects. Indeed, under the AI Act, the discretion left to providers of AI systems is much more limited than that exercised by data controllers. The GDPR does not establish a single way to comply with the requirements it makes mandatory.[342] Rather, it places on data controllers and processors the burden to assess the degree of risk of violating the data subjects' rights and freedoms (risk assessment) and to mitigate the risk thus assessed (risk mitigation).[343] The security of personal data processing under Article 32 GDPR is a case at point as, within the meaning of this

---

[339] COM (2020) 65 final, *supra* note 43 at 17.

[340] *Id.*

[341] *Id.* at 24; it is unclear, however, to what extent this ambition was fully translated in article 63 AI Act organizing market surveillance and control of AI systems in the Union market.

[342] *See generally* Raphaël Gellert, THE RISK-BASED APPROACH TO DATA PROTECTION (2020).

[343] Chris J. Hoofnagle et al., *The European Union General Data Protection Regulation: What it is and What it Means*, 28 INF. COMM. TECHNOL. L. 65 (2019).

provision, the controller and the processor are required to implement appropriate technical and organisational measures to ensure a level of security appropriate to the risk, taking into account the state of the art, the costs of implementation and the nature, scope, context and purposes of processing as well as the risk of varying likelihood and severity for the rights and freedoms of natural persons. The security level of processing is therefore function of an impact assessment carried out by the controller. This collaborative regulatory architecture makes the GDPR a form of co-regulation that differs from traditional command-and-control mechanisms[344] in that it follows a bottom-up approach.[345]

Alternatively, the AI Act is guided by a different logic. While the EU increasingly relies on risk regulation as a normative model for the development of digital technologies, the way the EU legislature designs its regulations is "far from unitary."[346] Contrasting the GDPR's bottom-up approach, the AI Act follows a top-down approach, which translates to very little discretion left to AI producers and importers to assess and mitigate 'high' risks. This is visible in the fact that the categories of risks are directly defined "and set in stone within the law" by the EC.[347] As already mentioned,[348] the AI Act lists the AI systems covered by a ban because assumed to raise unacceptable risks (Art. 5 AI Act). It also lists the AI systems subject to mandatory requirements because they are considered to raise high risks (Art. 6 AI Act). The EC is the only EU institution with the power to amend this list (Art. 7 AI Act).

The compliance and risk-mitigation are also strictly regulated. Providers of high-risk AI systems must ensure that their high-risk AI systems are compliant with the set of mandatory requirements listed in the AI Act (Art. 16 AI Act). They especially have to establish, implement, document and maintain risk management systems meant to eliminate or reduce risks "*as far as possible* through adequate

---

[344] Kamara, *supra* note 339.

[345] Giovanni De Gregorio & Pietro Dunn, *The European Risk-Based Approaches: Constitutional Dots in the Digital Age*, 59 COMMON MKT. L. REV. 473, 477 (2022).

[346] *Id.*

[347] *Id.* at 16.

[348] *See supra*, Section II.

design and development" (Art. 9(5)(a) AI Act). Though the discretion of providers is not altogether eliminated, it is reduced to a trickle.

Based on these arguments, we detect a *specific perception*, more so than evidence of AI-related risks that seemed to dictate the overly cautious framing of those risks, ultimately pushing the EC to establish a rigid normative framework, minutely outlining the types of risks, listing the high-risk sectors and (micro)regulating the compliance and mitigation strategies. But what shaped this perception of risks? It turns out it was policy objectives. As already mentioned, the EC was clear that any "law of AI" will have to *balance an ecosystem of trust* and *an ecosystem of excellence*. The EU institutions (in particular the Commission) are aware that AI is and will be the driver of research, innovation and economic progress. To maximize the various gains that AI development and use promise to deliver, two trends must be ignited and fostered. First, businesses should be encouraged to develop AI-based solutions (this objective was labelled by the EC as an "ecosystem of excellence").[349] This requires an innovation-friendly regulatory framework and financial investments. It is interesting to note that, although only *seven* (out of sixty) questions in the public consultation concerned the ecosystem of excellence,[350] a large majority of the respondents ranked the objectives associated with the 'ecosystem of excellence' as "very important" or "important".[351]

Second, EU citizens must trust AI systems to embrace these solutions, hence the name of the second objective: an "ecosystem of trust."[352] This requires that AI-based solutions are safe and fundamental rights-proof. The objective of the ecosystem of trust is to "give citizens the confidence to take up AI applications and give companies and public organisations the legal certainty to innovate using AI."[353] In blunt terms, the ecosystem of excellence aims at

---

[349] COM (2020) 65 final, *supra* note 43.
[350] EURO. COMM'N, Public consultation on the AI White Paper, *supra* note 148.
[351] *Id.*
[352] COM (2020) 65 final, *supra* note 24 at 3.
[353] *Id.*

ensuring there will be a *supply* of AI-based solutions, the ecosystem of trust that there will be a *demand* for such solutions.

The search for balance between an ecosystem of trust and an ecosystem of excellence gives an initial (though incomplete) explanation on why the AI Act is, at its core, a risk regulation. After all, in its search to strike the right balance between market rationality and fundamental rights protection, it is only logical that risk regulation be heralded as a viable alternative to "over-regulation, legalistic and prescriptive rules, and the high cost of regulation"[354] that reaches "regulatory excellence."[355] The EC made this crystal clear in the White Paper on AI: "the new regulatory framework for AI should be effective to achieve its objectives while not being excessively prescriptive so that it could create a disproportionate burden."[356] Given its ambition to ensure a *proportionate* regulatory intervention that does not stifle innovation,[357] the Commission chose a risk-based approach and emphasised that any mandatory requirements had to be tailored in relation to the gravity of the risk at stake, whether unacceptable, high, limited or non-high. Explained in this way, the EC's choice for a risk regulation framework can be viewed as intuitively the right one. Mandatory requirements as a function of the risk ensure the proportionality of the regulation.[358] The issue of proportionality will be discussed further in this article.[359] At this stage, we will focus on the understanding of this notion under the AI Act and argue that it was the specificity of this notion that ultimately justified the EC's minimal need for, and reliance on evidence to design the mandatory requirements currently enshrined in the instrument under consideration.

---

[354] Milda Macenaite, *The 'Riskification' of European Data Protection Law through a two-fold Shift*, 8 EUR. J. RISK REG. 506, 509 (2017).

[355] Bridget M. Hutter, *A Risk Regulation Perspective on Regulatory Excellence*, *in* ACHIEVING REGULATORY EXCELLENCE 101, 107 (Cary Coglianese ed., 2017).

[356] COM (2020) 65 final, *supra* note 40 at 17.

[357] Nicolas Petit & Jerome De Cooman, *Models of Law and Regulation for AI*, *in* THE ROUTLEDGE SOCIAL SCIENCE HANDBOOK OF AI 199, 212 (Anthony Elliott ed., 2022).

[358] Cary Coglianese, *The Law and Economics of Risk Regulation*, *in* WILEY ENCYCLOPEDIA OF OPERATIONS RESEARCH AND MANAGEMENT SCIENCE (James J. Cochran ed.) (forthcoming).

[359] *See infra*, Section V.

ii. *'Fact-Neutrality' Justified by a Specific Definition of the Notion of 'Risk'*

Considering that the key purpose of risk-regulation and risk-based (approach to) regulation is to mitigate the risk,[360] the EC's risk management,[361] raises doubts on the aptitude of the AI Act to attain that objective.

Typically, risk management encompasses risk identification, measurement, assessment, monitoring and evaluation,[362] and aims at gauging the risk.[363] To adequately (*i.e.,* proportionately) manage risks, both risk regulation[364] and risk-based (approach to) regulation[365] require a scientific risk assessment heralded as a "*Grundnorm*" or, more concretely, "the privileged methodological tool for regulating risk in Europe."[366] Yet, as already mentioned, the EC's risk characterisation was rather 'weak,' from the perspective of the conclusions inferred from the fact-finding procedures the Commission had undertaken.

What is more, though opting for risk regulation of AI seems logical - given the balancing exercise between the ecosystems of trust and of excellence - it is also surprising. A risk-based (approach to) regulation is classically used in an environmental[367] or health and safety contexts[368] notwithstanding the natural (e.g. hurricane,

---

[360] Deborah Lupton, RISK 8-9 (2nd ed., 2013).

[361] Defined as "a range of related activities for coping with risks, including how risks are identified and assessed and how social interventions to deal with risk are monitored and evaluated." *See* Gellert, *supra* note 332 at 26.

[362] Frederick Warner, "Introduction", *in* RISK: ANALYSIS, PERCEPTION AND MANAGEMENT: A REPORT OF A ROYAL SOCIETY STUDY GROUP 5 (The Royal Society ed., 1992).

[363] *See e.g.,* European Union General Data Privacy Regulations, Recitals 75 and 76.

[364] Macenaite *supra* note 357.

[365] Baldwin et al., *supra* note 79 at 281-83.

[366] Alberto Alemanno, *Regulating the European Risk Society*, *in* BETTER BUSINESS REGULATION IN A RISK SOCIETY 37, 41 (A. Alemanno et al. eds., 2013).

[367] *See generally* Clement Tisdell, ENVIRONMENTAL ECONOMICS: POLICIES FOR ENVIRONMENTAL MANAGEMENT AND SUSTAINABLE DEVELOPMENT (1993).

[368] Ellen Vos, INSTITUTIONAL FRAMEWORKS OF COMMUNITY HEALTH AND SAFETY REGULATION: COMMITTEES, AGENCIES AND PRIVATE BODIES 9 (1999).

earthquake and pandemics)[369] or technological (e.g. GMOs and chemical spills) origin[370] of the risk and that the realization of the risk was accidental (e.g. plane crash) or intentional (e.g. plane hijacking).[371] Sanitation, public health, food safety, industrial hygiene, chemicals, nuclear power plant, pharmaceuticals, GMOs, medical devices and even cosmetics provide examples *par excellence* of risk regulation.[372] What they all have in common is that they pertain to risks of *physical* harm. The AI Act – *hic sunt liones!* – defines risks, not *only* in relation to bodily harm, but *also*, if not primarily, in relation to *violation of fundamental rights* (Art. 7(1)(b) *in limine* AI Act). This begs the question of whether regulatory models, applied in the 'historic' areas of health and environment (with their usual focus on physical harm) are good referents for regulation that seeks to manage fundamental rights violations. True, it might not be possible to prevent risks of such violations without at least *some* evidence of those risks. However, as Quelle stressed with regard to data protection law, "we cannot say, for example, that the highest permissible impact on privacy is 2.5 'mg' per data subject per year."[373] Her observation is astute: the "dose-response relationship" as a prevailing model in standard risk regulation, is not a good fit for assessing risks of fundamental rights violations.[374] In light of this, one would be tempted to assume that risk (based) regulation may not be the way to go, when the end-goal of (risk/risk-

---

[369] Shahar Avin et al., *Classifying Global Catastrophic Risks*, 102 FUTURES 20, 20-21 (2018).

[370] Aaron Wildavsky, SEARCHING FOR SAFETY: SOCIAL THEORY AND SOCIAL POLICY 10 (1988, reprinted 2017).

[371] Alberto Alemanno, *Risk and regulation, in* ROUTLEDGE HANDBOOK OF RISK STUDIES 191, 197 (Adam Burgess, Alberto Alemanno & Jens O. Zinn eds., 2016). Alemanno explains (and illustrates) at 194 that "populations face hazards from many sources: from physical forces (for example, sound waves, magnetic fields, radioactivity), chemicals (for example, mercury, ozone, dioxins and pesticides), organisms (for example, viruses and bacteria), as well as human behaviours (for example, smoking, binge drinking, drunk driving, lack of physical exercise)."

[372] *Id.* at 192.

[373] Quelle, *supra* note 335 at 525.

[374] Alemanno, *supra* note 374 at 193 (discussing the prevalence of the dose-response relationship in risk regulation).

based) policy is the protection of fundamental rights.[375]

Risk-based (approach to) regulation is, however, not the only available regulatory model in the law- and policymakers' toolbox.[376] In addition to this approach - which consists in viewing AI as a "techno-economic segment" subject to product compliance schemes - there is an alternative approach that focuses on "how to guarantee fundamental rights and democracy."[377] This is the so-called *rights-based* approach which, according to Linksey, includes two versants:[378] it is either rights-conferring and/or fundamental rights-epitomising. Under the first versant, a regulatory approach is 'rights-based' if it confers rights to individuals. In the second versant - and following the CJEU's formula in *Kükükdeveci* - fundamental rights-epitomising approach translates to giving a "specific expression to" a fundamental right.[379] A regulation is qualified as rights-based if it

---

[375] Fanny Hidvegi et al., *The EU Should Regulate AI On The Basis Of Rights, Not Risks*, ACCESS NOW (Feb 17, 2021 4:10 AM) https://www.accessnow.org/eu-regulation-ai-risk-based-approach/.

[376] Petit and De Cooman, *supra* note 360 at 48.

[377] Bilel Benbouzid et al., *Quatre Nuance de regulation de l'intelligence artificielle : une cartographie des conflits de définitions*, 40 RESEAUX 31, 58 (2022).

[378] Orla Linksey, THE FOUNDATIONS OF EU DATA PROTECTION LAW 35, 36 (2015).

[379] CJEU, 19 January 2010, *Seda Kücükdeveci v Sweden GmbH & Co. KG.*, case C-555/07, EU:C:2010:21, para 21. In this decision, the CJEU held the Council Directive 2000/78/EC of November 2000 establishing a general framework for equal treatment in employment and occupation (OJ L 303, 02 December 2000, 16-22) gives expression to Article 21(1) EUCFR that prohibits any discrimination based on age. *See also*, to some extent, article 52(5) EUCFR, *in limine*, stating that "the provisions of this Charter which contain principles may be implemented by legislative and executive acts taken by institutions, bodies, offices and agencies of the Union, and by acts of Member States when they are implementing Union law, in the exercise of their respective powers". In *Bauer*, the CJEU clarified the usefulness of EU secondary law giving expression to a fundamental right. Even if the fundamental right in question is clear, unconditional and sufficiently precise (*See* Case 26/62, NV Algemene Transport - en Expeditie Onderneming van Gend & Loos v. Netherlands Inland Revenue Administration, 1963 E.C.J. 1.), EU secondary (and national) laws are still relevant as they may be "required to specify (...) certain conditions for the exercise of that right." *See* CJEU, 6 November 2018, *Stadt Wuppertal v Maria Elisabeth Bauer* (C-569/16), *Volker Willmeroth, in his capacity as owner of TWI Technische Wartung und Instandsetzung Volker*

falls in one of these two categories.

A topical example of a rights-conferring regulation is contract law: it recognizes subjective rights to individuals, though these rights are not, strictly speaking, fundamental (Human) rights. Conversely, the rights-epitomizing versant is heralded as the champion of the Human rights-based approach to regulation. The United Nations defined the latter as any activity whose aim "is to contribute directly to the realization of one or several human rights."[380] It is used to enforce, *e.g.*, the right to work,[381] right to access higher education,[382] and right to environmental protection.[383] More specifically, in a (fundamental) rights-based approach, regulation improves the capacity of duty-bearers to meet their obligations and of rights-holders to claim their rights.[384] A (fundamental) rights-based

---

Willmeroth e.K. v Martina Broßonn (C-570/16), joined cases C-569/16 and C-570/16, EU:C:2018:871, para 85. *See also* CJEU, 6 November 2018, *Max-Planck-Gesellschaft zur Förderung der Wissenschaften eV v Tetsuji Shimizu*, case C-684/16, EU:C:2018:874. For more on this, *see* Rüdiger Krause, *Horizontal Effect of the EU Charter of Fundamental Rights: Bauer and Willmeroth, MPG*, 58 COMMON MKT. L. REV., 1173 (2021); *see generally* Nina Poltorak, *The application of the rights and principles of the Charter of Fundamental Rights*, *in* THE PRACTICE OF JUDICIAL INTERACTION IN THE FIELD OF FUNDAMENTAL RIGHTS: THE ADDED VALUE OF THE CHARTER OF FUNDAMENTAL RIGHTS OF THE EU 23 (Federica Casarosa & Madalina Moraru eds., 2022); Eleni Frantziou, *(Most of) the Charter of Fundamental Rights is horizontally applicable*, 15 EUR. CONST. L. REV. 306, 307 (2022).

[380] United Nation Development Group, The Human Rights-Based Approach to Development Cooperation: Towards a Common Understanding Among UN Agencies (Sept. 2003) 1-4, https://unsdg.un.org/sites/default/files/6959-The_Human_Rights_Based_Approach_to_Development_Cooperation_Towards_a_Common_Understanding_among_UN.pdf.

[381] *See generally* Lisette Villacres & Sara Geenen, *Framing Street vending in Guayaquil – Ecuador: From hegemonic discourses to a rights-based approach*, Institute of Development Policy, University of Antwerp, Feb. 2020 at 4, 17.

[382] Jane Kotzmann, THE HUMAN-RIGHTS-BASED APPROACH TO HIGHER EDUCATION: WHY HUMAN RIGHTS NORMS SHOULD GUIDE HIGHER EDUCATION LAW AND POLICY (2018); *see also The Human Rights-Based Approach to STEM Education* (Tanja Tajmel et al., eds., 2021).

[383] Damilola S. Olawuyi, THE HUMAN-RIGHTS-BASED APPROACH TO CARBON FINANCE (2016).

[384] Miao He, A HUMAN RIGHTS-BASED APPROACH TO CONSERVING PROTECTED AREAS IN CHINA: LESSONS FROM EUROPE 354 (2016).

approach therefore entails an obligation on the part of governments to protect the rights concerned, through mechanisms preventing their infringement. Such an approach holistically provides a set of legal standards that serve as the basis for regulation. [385]

In the EU, data protection law provides a telling example of regulation that follows a rights-based approach to regulation since it both confers rights to individuals (rights-conferring) and "gives expression to" the fundamental rights of personal data protection (rights-epitomising).[386] That data protection is a fundamental right is beyond doubt, considering that both Article 8 of the EU Charter of Fundamental Rights (hereafter, ECFR) and Article 16 of TFEU[387] recognize it as such. Since the ambition of EU data protection law is to provide "a minimum and non-negotiable level of protection for all individuals"[388] it is possible to detect, in the EU, a process of "constitutionalisation" of the right to data protection,[389] as can be inferred from both the design of the GDPR and the CJEU's caselaw.

It is also beyond doubt that the GDPR confers rights to individuals. Chapter III of said Regulation - intentionally titled "Rights of the data subject" - guarantees individual rights such as the right to be informed (Art. 13-14), the right to access to personal data (Art. 15), the right to rectification (Art. 16 and 19), the right to erasure or, differently put, to be forgotten (Art. 17 and 19), the right to restriction of processing (Art. 18 and 19), the right to data portability (Art. 20), the right to object (Art. 21), and the right not to be subject to a decision based solely on automated processing, including profiling, which producing legal effects on the data subject

---

[385] *See generally* Asbjorn Eide, *Economic, Social and Cultural Rights as Human Rights*, *in* ECONOMIC, SOCIAL AND CULTURAL RIGHTS: A TEXTBOOK (Asbjorn Eide et al. eds., 2nd ed. 2001).

[386] Linksey *supra* note 381 at 35-36.

[387] For an early review of the impact of data protection law on fundamental rights, *see* Paul de Hert and Serge Gurtwirth, *Data protection in the case law of Strasbourg and Luxembourg*, *in* REINVENTING DATA PROTECTION? 3 (Serge Gutwirth et al. eds., 2009).

[388] Art. 29 Working Party, "Opinion 1/98 Platform for Privacy Preferences (P3P) and the Open Profiling Standards (OPS)", 1998.

[389] De Hert & Gurtwirth, *supra* note 389.

(Art. 22).[390]

In this context, in addition to being a risk-regulating instrument, does the AI Act also confer or protect fundamental rights? The answer seems to be 'yes' since it clearly "gives expression to" some fundamental rights rooted in human dignity and human autonomy which, in the context of AI, translates to the fact that human beings "interacting with AI systems must be able to keep full and effective self-determination over themselves" and that the aforementioned AI systems should not "subordinate, coerce, deceive, manipulate, condition or herd humans."[391]

It is with these values in mind that the HLEG selected, in 2018, the ethical principles of respect for human autonomy, prevention of harm, fairness and explicability. These principles were directly derived from rights enshrined in the ECFR, namely human dignity (Art. 1 to 5 ECFR), freedom of the individual (Art. 6 to 19 ECFR), respect for democracy, justice and the rule of law (Art. 47 to 50 ECFR), equality, non-discrimination (Art. 20 to 26 ECFR) solidarity (Art. 27 to 38 ECFR) and citizens' rights (Art. 39 to 46 ECFR). The choice of these rights, as axiological foundations for the HLEG's principles of ethics, was justified by the goals that these principles were to achieve, when applied in practice. These goals pertain to the effective exercise of... other fundamental rights, namely, the right to liberty and security (Art. 6 ECFR), the respect for private and family life (Art. 7 ECFR), the freedom of thought, conscience and religion (Art. 10 ECFR), of expression and information (Art. 11 ECFR), and of assembly and association (Art. 12 ECFR).

It is important to stress that in creating its four-pillar framework of ethical principles, the HLEG did not make a fact-to-ethics leap; it made a *values-to-rights* leap. This approach can be inferred from the ways in which the HLEG defined each of the four cornerstone principles. Indeed, *prevention of harm* entails "the protection of human dignity [Art. 1 ECFR] as well as mental integrity [Art. 3

---

[390] Hoofnagle et al., *supra* note 345 at 47. There is, however, not a right to explanation of automated decision-making. *See* Sandra Wachter et al., *Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation*, 7 INT. DATA PRIV. L. 76 (2017).

[391] HLEG on AI, *supra* note 123 at 12.

ECFR]."[392] *Fairness* has both a substantive and a procedural dimension. Substantively, it means the absence of unfair bias, discrimination and stigmatisation[393] clearly giving expression to fundamental rights such as equality before the law (Art. 20 ECFR), equality between men and women (Art. 22 ECFR) and non-discrimination (Art. 21 ECFR). In its procedural dimension, fairness translates to the ability of human agents to challenge and seek effective redress against decisions made either by the AI systems or their users.[394] Procedural fairness thus gives expression to the right to an effective remedy and fair trial (Art. 47 ECFR) as well as to the right to good administration (Art. 41 ECFR). Duly contesting algorithmic decisions implies that they have to be explicable; hence the eponym ethical principle which also gives specific expressions to Articles 47 and 41 ECFR.[395]

The value-to-rights leap made by the HLEG in selecting the foundational ethical principles on AI established the basic normative framework within which the AI Act was to be designed. To *operationalize* the standard of effectiveness of the fundamental rights protection, the HLEG - and then the AI Act - outlined a number of key requirements.[396] To uphold and effectively safeguard the ethical principle of respect for human autonomy, human agency and oversight were naturally seen as necessary mandatory requirements for AI providers and users. This meant that human agents must be able to make informed choices and thus exercise their autonomy which presupposed that they receive "appropriate knowledge and tools to comprehend and interact with AI systems to a satisfactory degree and, where possible, be enabled to reasonably self-assess or challenge the system."[397]

Alternatively, the requirement of human oversight seeks to ensure that AI systems do not undermine human autonomy by keeping a human-in-the-loop (at every stage of the AI system), on-

---

[392] *Id.*
[393] *Id.*
[394] *Id.* at 13.
[395] *Id.*
[396] *Id.* at 14.
[397] *Id.* at 16.

the-loop (during the design cycle or system's operation), or in-command (overseeing the overall activity of the AI system).[398] Interestingly, the HLEG stressed that human oversight also implies that "public enforcers have the ability to exercise oversight in line with their mandate" thus alluding to the fundamental right to a good administration.[399] The prevention of harm principle is operationalized through the requirements for technical robustness and safety, as well as privacy and data governance. The former translates to "a preventive approach to risk" that seeks to mitigate or at least minimise the occurrence of harms. The latter implies a control of data quality, integrity and relevance in light of the specific use(s) of an AI system.[400]

The principle of explicability is operationalized through the requirement for transparency. This is tautological: there can be no explicability without transparency, defined as data traceability, explainability, and identification of AI system as such (and not deceivably as human beings).[401]

Finally,[402] the principle of fairness is operationalized through the requirement of diversity and non-discrimination (substantive dimension), and accountability (procedural dimension). Under the substantive dimension, the aim is to avoid the emergence of unfair biases that lead to "unintended (indirect) prejudice and discrimination."[403] Under the procedural dimension, accountability requires mechanisms that ensure adequate redress when unjust adverse consequences occur.[404]

It is interesting to note that, although the EC endorsed these requirements in 2018, it recognised they were "non-binding and as such [did] not create any new legal obligations."[405] Enter the AI Act,

---

398 *Id.*
399 *Id.*
400 *Id.* at 17.
401 *Id.* at 18.
402 This list leaves aside the requirement of societal and environmental well-being (that comes from the principles of fairness and prevention of harm) on purpose as it is closer to a declaration of intent than the basis for a substantive right.
403 HLEG on AI, *supra* note 123 at 18.
404 *Id.* at 20.
405 COM (2019) 168 final, *supra* note 121.

which clearly builds on the HLEG's guidelines, taking over the requirements derived from the HLEG's ethical principles and 'converting them' into binding provisions with mandatory requirements. This can be inferred from a comparative overview of the ethical principles and the corresponding provisions in the AI Act.

First, the explanatory memorandum of the AI Act echoed the HLEG by acknowledging that AI systems can adversely affect a number of fundamental rights enshrined in the ECFR and explicitly stating that "this proposal seeks to ensure a high level of protection for those fundamental rights"[406] by enhancing and promoting the protection of the rights protected by the Charter.[407] The list of fundamental rights potentially infringed by AI is almost endless:[408] the right to human dignity (Art. 1 ECFR), respect for private life (Art. 7 ECFR) and protection of personal data (Art. 8 ECFR), rights to freedom of expression (Art. 11 ECFR) and freedom of assembly (Art. 12 ECFR), non-discrimination (Art. 21 ECFR), equality between women and men (Art. 23 ECFR), rights of the child (Art. 24 ECFR) and the integration of persons with disabilities (Art. 26 ECFR), consumers protection (Art. 28 ECFR), rights to fair and just working conditions (Art. 31 ECFR) , rights to a high level environmental protection and to the improvement of the environmental quality (Art. 37 ECFR), and right to an effective remedy and to a fair trial, the rights of defence and the presumption of innocence (Art. 47 and 48 ECFR).[409]

Second, the Proposal enshrines, in legally binding provisions, the HLEG's ethical requirements. It is the case of human oversight (Art. 14 AI Act), accuracy, robustness and cybersecurity (Art. 15 AI Act), data governance (Art. 10 AI Act), transparency (Art. 13 AI

---

[406] *Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts*, COM (2021) 206 final (Apr. 21, 2021).

[407] *Id.*

[408] Giovanni Sartor, *Human Rights and Information Technologies*, *in* THE OXFORD HANDBOOK ON THE LAW AND REG. OF TECH. 424 (Roger Brownsword *et al.* eds., 2017).

[409] COM (2020) 65 final, *supra* note 24 at 11.

Act), and accountability through risk management system (Art. 9 AI Act), technical documentation (art. 11 AI Act), record-keeping (art. 12 AI Act) and quality management system (Art. 17 AI Act).[410] Figure 1 below represents, in the form of a network, the gradual particularization of ethical principles into first, key normative standards (Ethics guidelines) and second, corresponding requirements for compliance (AI Act).

[410] It is of utmost importance to emphasize the ethical requirements support each other. When endorsing human oversight, the Commission for instance clarified that "the appropriate degree of [...] adaptability, accuracy and explainability of AI-based systems" should be ensured when achieving human oversight, hence referring to the principle of prevention of harm and the requirement of technical robustness and safety; *Id.* The principle of respect for human autonomy gives birth to the requirement of human agency and oversight but is achieved through a combination of the latter and technical robustness and safety. As a result, a single ethical principle might be enshrined in several mandatory requirements within the AI Act. Similarly, a single mandatory requirement of the AI Act might epitomize several ethical principles.

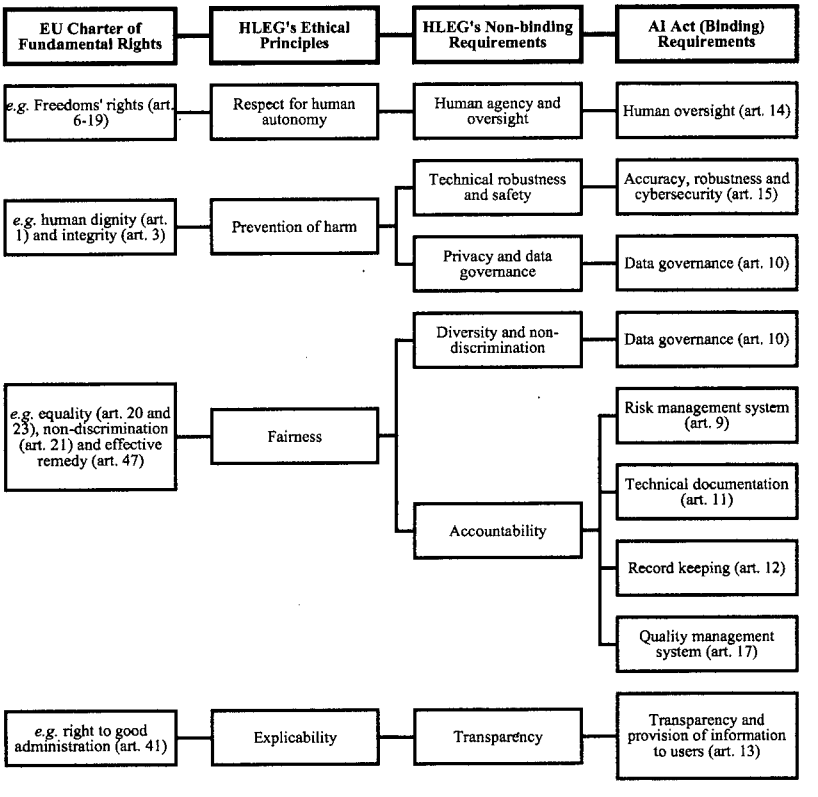

Figure 1. The origins of the AI Act's mandatory requirements

This amphigoric combination of ethical principles, recommendations, fundamental rights and mandatory requirements indicates the AI Act does, indeed, *give expression to fundamental rights*. To exemplify this reasoning, the data governance requirement - enshrined in Article 10 of the AI Act - "gives expression to" the fundamental rights of equality and non-discrimination enshrined in, respectively Art. 20 and Art. 21 ECFR. Since these rights served as referents for the HLEG's ethical principles (fairness) and recommendation (equality and non-discrimination), it is logical that they were 'converted into binding law' in the shape of a mandatory requirement of data governance.

Our reading of the AI Act as giving expression to fundamental rights is compatible with the EC's ambition to build "trust in human-centric Artificial Intelligence."[411] The Commission indeed acknowledged that "the values on which our societies are based need to be fully integrated in the way AI develops" and defined these values as the "respect for human dignity freedom, democracy, equality, the rule of law and respect for human rights."[412]

As a consequence of this three-step regulatory translation (values-rights-duties), the basic regulatory stance in the EU was that AI systems must be developed "in a way that puts people at its centre," meaning that they "should not only be consistent with the law, but also adhere to ethical principles."[413] In light of this premise, the design of the AI Act is hardly surprising:[414] any mandatory requirements contained therein is to allow AI systems to empower citizens and reinforce the respect of their fundamental rights.[415]

Throughout the Explanatory Memorandum and Staff Working Document of the AI Act, as well as in the AI Act itself, the need to protect natural persons from "high risk of harm to the health and safety or the fundamental rights of person" is heralded as the north star.[416] This objective, however, is sought directly through horizontal regulation requiring, in a nutshell, standardisation and certification through *ex ante* safety requirements.

In light of the above, we conclude that the EU's regulation of AI is both a risk/risk-based framework and a fundamental rights-driven framework.[417] To put it more delicately: *structurally* (in terms of its design and governance model) and *functionally* (in terms of the objective of risk-prevention), the AI Act is a risk(-based) regulation. In parallel, *axiologically* (in terms of its foundational values) and *substantively* (in terms of the content of the requirements it contains) it is a rights-based regulation. This *mélange* is surprising, to say the

---

411 COM (2018) 237 final, *supra* note 51.
412 COM (2019) 168 final, *supra* note 121.
413 *Id.*
414 Giovanni Sartor, *Artificial Intelligence and Human Rights: Between Law and Ethics*, 127 MAASTRICHT J. EUR. COMP. 705 (2020).
415 COM (2019) 168 final, *supra* note 121.
416 COM (2021) 206 final, *supra* note 408.
417 De Gregorio & Dunn, *supra* note 347 at 493.

least, as the risk- and rights-based approaches are traditionally considered to be poles apart.[418] Because the rights-based approach is typically rights-conferring, it usually follows a binary logic: an action is either legal or illegal (that is, it either meets a standard of protection, or it does not).[419] A rights-based approach therefore applies the "same rules to everyone irrespective of the level of risk or harm."[420] Risk-regulation, however, does not distinguish between legal and illegal conduct but seeks to tailor an appropriate response to a given risk. The AI Act fits this qualification as compliance with the mandatory requirements is limited to high-risks AI systems (Art. 8 AI Act).

Merging risk- and rights-based regulation is, however, not a premiere. As mentioned earlier, the GDPR is also a risk regulation since it regulates risks that arise from the information and communication technologies; it is a risk-based (approach to) regulation as supervisory authorities focus their efforts on high-risk sectors and finally, it is a rights-based approach to regulation that gives expression to the fundamental rights of data protection and grants several rights to data subjects. Similarly, the AI Act is a risk regulation that mitigates risks of AI systems, a risk-based (approach to) regulation as national supervisory authorities will deploy their enforcement activity in exhaustively enumerated high-risk sectors and - much like the GDPR - it is a rights-based regulation that gives expression to virtually all the rights enshrined in the ECFR. The AI Act is therefore akin to the GDPR in their shared ambition to mitigate "a risk to a right."[421]

This kinship notwithstanding, there is an important difference

---

[418] Raphaël Gellert, *We Have Always Managed Risks in Data Protection Law: Understanding the Similarities and Differences Between the Rights-Based and Risk-Based Approaches to Data Protection*, 2 EUR. DATA PROT. L. REV. 481, 481-83 (2016). However, risk- and rights-based approaches are not mutually exclusive; *see e.g.* Stephen Perry, *Risk, Harm, Interests, and Rights*, *in* RISK: PHILOSOPHICAL PERSPECTIVES 190, 203 (Tim Lewens ed., 2007).

[419] Linksey,, *supra* note 380 at 35-36.

[420] Gellert, *supra* note 331 at 2.

[421] *See generally* Niels van Dijck et al., *A Risk to a Right? Beyond Data Protection Risk Assessments*, 32 COMPUT. L. & SEC. REV. 286 (2016).

between the instruments considered from the perspective of their rights-based dimension. They both give expression to fundamental rights but the GDPR is the only one that actually *confers* rights to individuals. The AI Act 'merely' establishes product safety with technical standardisation and certification schemes meant as safeguards against fundamental rights infringements. With this in mind, scholarship qualified the AI Act as incomplete - some have said, flawed[422] - rights-based regulation. The recent Proposal for a Directive adapting non-contractual civil liability rules to AI may, however, fill this gap (at least partially).[423] This Proposal could be reasonably anticipated, given that *ex ante* safety rules and corresponding *ex post* liability rules are generally not substitutes but complements.[424]

2. The *explanatio legis*: Normative Coherence with Existing EU Law

In lieu of knowledge of facts providing a plausible explanation of the rationale underlying the design of the AI Act, a key argument that can provide such an explanation is the latter's coherence with prior EU law. This coherence can be detected on two levels: on the one hand, the axiological congruence of the AI Act with provisions that recognize or give expression to fundamental rights, on the other hand, consistency with EU Secondary law that falls in the scope of application of the AI Act.

In the context of representational democracy, Bindham Powell Jr. defined ideological congruence as involving "the fit between the

---

[422] *Q&A: How the EU's Flawed Artificial Intelligence Regulation Endangers the Social Safety Net*, HUMAN RIGHTS WATCH, https://www.hrw.org/sites/default/files/media_2021/11/202111hrw_eu_ai_regulation_qa_0.pdf (last visited Apr. 1, 2023).

[423] *Proposal for a Directive of the European Parliament and of the Council on Adapting Non-contractual Civil Liability Rules to Artificial Intelligence (AI Liability Directive),* COM (2022) 496 final (Sept. 28, 2022). On this hypothesis, *see* Martin Ebers *et al.*, *The European Commission's Proposal for an Artificial Intelligence Act – A Critical Assessment by Members of the Robotics and AI Law Society (RAILS)*, 4 MULTIDISCIP. SCIENT. J. 589, 599-600 (2021).

[424] Charles D. Kolstad et al., *Ex Post Liability for Harm vs. Ex Ante Safety Regulation: Substitutes or Complements?*, 80 AM. ECON. REV. 888, 888 (1990).

preferences of the citizens and the committed policy positions of their representatives. Normative theorizing about democracy implies that competitive elections should systematically create a close connection between citizens and their policy makers."[425] In the context of policy, the 'fit' which characterizes congruence is that between a set of values, chosen by policy makers as a foundational axiological matrix for future regulation and the specific regulatory instruments meant to translate those values in series of mandatory requirements.

In the AI Act, axiological congruence can be identified on two levels. The first is a meta-, principle-based level. As already mentioned, the objectives pursued by most regulatory proposals on AI are variations of the Asilomar principles. Like the EU, national and international regulators sought to carefully select amongst those principles in view of reaching a proportionate trade-off between humancentric AI and economic efficiency.[426] Though these selections vary from one system to another (some countries, like Germany, lean more toward human-centrism while others like the US lean more toward market efficiency), they do seem to rely on a common axiological framework of values namely, benevolence, non-malevolence, human autonomy, justice and explicability.[427]

---

[425] G. Binhgam Powell Jr., *Representation in Context: Election Laws and Ideological Congruence Between Citizens and Governments*, 11 PERS. POL. 9, 10 (2013).

[426] *See supra* Section II.

[427] Jaana Leikas et al., *Ethical framework for Designing Autonomous Intelligent Systems*, 5 JOLTMC 18 (2019); *see generally* Luciano Floridi, *Translating Principles into Practices of Digital Ethics: Five Risks of Being Unethical*, 32 PHILOS. TECHNOL. 185 (2019); Luciano Floridi, *A Unified Framework of Five Principles for AI in Society*, 1 HDSR (2019); Thilo Hagendorff, *The Ehics of AI Ethics – An Evaluation of Guidelines*, 30 MINDS MAC. 99 (2020); Anna Jobin et al., *The Global Landscape of AI Ethics Guidelines*, 1 NAT. MACH. INTELL. 389 (2019); Yves Poullet, *About some international documents relating to the ethics of Artificial Intelligence – Some insights*, *in* TIME TO RESHAPE THE DIGITAL SOCIETY 523 (Hervé Jacquemin ed., 2021). For a review of regulatory proposals, *see e.g.* Patricia Gomes Rêgo de Almeida *et al.* , *Artificial Intelligence Regulation: a framework for governance*, 23 ETHICS INF. TECHNOL. 505 (2021); Dominika Harasimiuk & Tomasz Braun, REGULATING ARTIFICIAL INTELLIGENCE: BINARY ETHICS AND THE LAW (2021).

With the exception of the last of these values,[428] the others originate in bioethics[429] and were, first, taken over *talis qualis* in the HLEG's Draft Ethics Guidelines[430] only to, second, be renamed as principles of prevention of harm, respect for human autonomy, fairness, and explicability.[431] As shown in Figure 1 above, it was these principles that provided the foundational axiological framework within which for the mandatory requirements in the AI Act eventually took shape.

In addition to this meta-level value congruence, the AI Act is, second, congruent with the already existing EU regulation that translated those values into subjective rights in the field of data processing: the congruence yardstick here is, of course, the GDPR.

Until the AI Act was drafted, we could have qualified the GDPR as the "law of everything" considering the importance and omnipresence of data processing, automated or not. By addressing advanced-data processing technologies, the AI Act proposal does not reduce the importance of the GDPR; on the contrary, it complements it (hence the congruence argument), the two instruments being, in many ways, two sides of the same axiological coin.[432]

As argued earlier, the GDPR epitomises the fundamental rights of data protection (Art. 8 ECFR) the safeguard of which is framed, in this Regulation, by a number of carefully selected axiological principles namely, lawfulness, fairness and transparency (Art. 5(1)(a) GDPR), proportionality and necessity of the processing (Art.

---

[428] Floridi, *A Unified Framework of Five Principles for AI in Society*, *supra* note 429 at 5.

[429] *See e.g.*, UNESCO Universal Declaration on Bioethics and Human Rights, (Oct.19, 2005); The Convention for the protection of Human Rights and Dignity of the Human Being with regard to the Application of Biology and Medicine: Convention on Human Rights and Biomedicine (ETS No. 164), Council of Europe, 04 April 1997, Oviedo, Spain.

[430] HLEG on AI, *supra* note 123.

[431] *Id.*

[432] Nadezhda Purtova, *The Law of Everything. Broad Concept of Personal Data and Future of EU Data Protection Law*, 10 LAW, INNOV. & TECH. 40, 75 (2018). For more on AI and GDPR interactions, *see also* Sam Wrigley, *Taming Artificial Intelligence: 'Bots,' the GDPR and Regulatory Approaches*, *in* ROBOTICS, AI AND THE FUTURE OF LAW 183 (Marcelo Corrales et al. eds., 2018); Sandra Wachter & Brent Mittelstadt, *A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI*, 2 COLUM. BUS. L. REV. 494 (2019).

5(1)(b), (c) and (e) GDPR) as well as the accuracy of collected data (Art. 5(1)(d) GDPR).[433] The GDPR then particularizes these principles in subjective rights (*e.g.*, right to information and access to personal data (Arts. 13-15 GDPR)) safeguarded through series of mandatory requirements addressed to data controllers and processors (e.g., data protection by design and by default (Art. 25 GDPR)).[434] The safeguard of these rights is, ultimately, guaranteed by a number of mandatory requirements that data processors must observe. Table 1 below summarises the argument.

[433] This analysis leaves aside secure processing (art. 5(1)(f) GDPR) as it is not translated into a subjective right of the data subject, but into an obligation of data controllers and processors, *e.g.*, the implementation of ("appropriate technical and organizational measures to ensure a level of security appropriate to the risk") (art. 32(1) GDPR).

[434] Hoofnagle et al., *supra* note 345.

<table>
<tr><th colspan="2">Ontological Grounds (Principles)</th><th>Subjective Rights</th><th>Mandatory Requirements</th></tr>
<tr><td colspan="2" rowspan="2">Lawfulness, fairness and transparency (Art. 5(1)(a))</td><td>Information and access to personal data (Arts. 13-15)</td><td rowspan="5">• Transparency (Art. 12)
• Appropriate technical and organisational measures demonstrating the lawfulness of the processing (Art. 24(1)), by design and by default (Art. 25)
• Appropriate data protection policies (Art. 24(2))
• In case of automated processing, human intervention (Art. 22(3))[435]</td></tr>
<tr><td>Right not to be subject to a decision based solely on automated processing (Art. 22)</td></tr>
<tr><td rowspan="3">Proportionality and necessity</td><td rowspan="3">• Purpose limitation (Art. 5(1)(b))
• Data minimization (Art. 5(1)(c))
• Storage limitation (Art. 5(1)(e))</td><td>Right to erasure (Art. 17)</td></tr>
<tr><td>Right to restriction (Art. 18)</td></tr>
<tr><td>Right to object (Art. 21)</td></tr>
<tr><td colspan="2">Accuracy (Art. 5(1)(d))</td><td>Right to rectification (Art. 16)</td></tr>
</table>

Table 1. The (non-exhaustive) structure of the GDPR

Our Table allows to comprehensively map out the principles-

[435] The scope of this particular requirement is still heavily debated; *see e.g.* Request for a preliminary ruling, 15 October 2021, Case C-634/21, questioning whether article 22(1) GDPR means ("that the automated establishment of a probability value concerning the ability of a data subject to service a loan in the future already constitutes decision based solely on automated processing, including profiling, which produces legal effects concerning the data subject or similarly significantly affects him or her, where that value, determined by means of personal data of the data subject, is transmitted by the controller to a third-party controller and the latter draws strongly on that value for its decision on the establishment, implementation or termination of a contractual relationship with the data subject?"); *see also* Wachter et al., *supra* note 392; Lilian Edwards & Michael Veale, *Slave to the Algorithm? Why a 'Right to an Explanation' Is Probably Not the Remedy You Are Looking For*, 16 DUKE L. & TECH. REV. 18 (2017); Stefanie Hänold, *Profiling and Automated Decision-Making: Legal Implications and Shortcomings*, *in* ROBOTICS, AI AND THE FUTURE OF LAW 123 (Marcelo Corrales et al., eds. 2018).

rights-obligations structure of the GDPR and, when compared with Figure 1 *supra*, show how this structure is mirrored by the principles-standards-obligations structure of the AI Act. For instance, Art. 9 AI Act requires the elimination or reduction of risks "through adequate design and development" thus echoing Art. 24 ("appropriate technical measures") and Art. 25 ("by design") GDPR. Similarly, Art. 13 AI Act mimics the transparency requirement of the GDPR stating that instructions for use of high-risk AI systems have to be "concise, complete, correct and clear information that is relevant, accessible and comprehensible to users."[436] The right not to be subject to a decision based solely on automated processing (Art. 22(1) GDPR) and, if so, the right to obtain human explanation (Art. 22(3) GDPR) is mirrored in, and completed by, Art. 14 AI Act which makes human oversight mandatory for high-risk AI system.[437]

---

[436] For the record, article 12 GDPR states that information and communication to the data subject have to be provided ("in a concise, transparent, intelligible and easily accessible form, using clear and plain language"). The argument may be extended to the Digital Services Act whose ontological ground is ("transparency, the protection of recipients of the service and the avoidance of unfair or arbitrary outcomes") and to provide ("robust safeguards")that protect ("the right and legitimate interests of all affected parties, in particular their fundamental rights guaranteed by the Charter") (European Commission, Proposal for a Regulation of the European Parliament and of the Council on a Single Market for Digital Services (Digital Services Act) and amending Directive 2000/31/EC, 15 December 2020, COM/2020/825 final, recitals 38 and 41). Article 12(1) DSA hence requires that ("information on any policies, procedures, measures and tools used for the purpose of content moderation, including algorithmic decision-making and human review") to be ("set out in clear and unambiguous language and (...) be publicly available in an easily accessible format.").

[437] Article 14(4) AI Act explains that human oversight implies that a human should be enabled to, e.g., "a) fully understand the capacities and limitations of the high-risk AI system and be able to duly monitor its operation, so that signs of anomalies, dysfunctions and unexpected performance can be detected and addressed as soon as possible; (b) remain aware of the possible tendency of automatically relying or over-relying on the output produced by a high-risk AI system ('automation bias'), in particular for high-risk AI systems used to provide information or recommendations for decisions to be taken by natural persons; (c) be able to correctly interpret the high-risk AI system's output, taking into account in particular the characteristics of the system and the interpretation tools and methods available; (d) be able to decide, in any particular situation, not to use the

Cybersecurity[438] and accuracy[439] of the system similarly constitute a common point... In other words, the GDPR and the AI Act share the same "constitutional dot" and are 'congruent' insofar as they offer different but twin normative expressions (rights, for the GDPR, standards for the AI Act) in operationalizing commonly shared principles and values. In doing so, both instruments participate in the striking of the aspired to balance between an efficient digital single market and the protection of fundamental rights.[440]

In addition to the value-based congruence of the AI Act with fundamental rights recognized constitutionally (in the ECFR) and particularized in EU Secondary Legislation (in the GDPR), another argument that allows us to explain the ancillary role of facts in drafting the AI Act is the latter's consistency with already existing Union legislation which corresponds with one specific normative objective of the instrument under consideration, namely *product safety*.

The AI Act is, indeed, in line with EC's overall digital strategy in its contribution to promoting technology that works for people, one of the three main pillars of the policy orientation and objectives announced in the Communication 'Shaping Europe's digital future.' Its ambition - the EC observed - is to lay down a *coherent, effective* and *proportionate* framework to ensure AI is developed in ways that respect people's rights and earn their trust, making Europe fit for the digital age.

Consistency is essentially derived from the *coherence* between the instruments included in the EU's New Technologies Package. Indeed, the AI Act is viewed as part of a wider cluster of measures

---

high-risk AI system or otherwise disregard, override or reverse the output of the high-risk AI system; (e) be able to intervene on the operation of the high-risk AI".

[438] Compare *e.g.* article 5(1)(f) GDPR requiring personal data to be "processed in a manner that ensures appropriate security" with article 15 AI Act stating that "high-risk AI systems shall be designed and developed in such a way that they achieve (...) an appropriate level of (...) robustness and cybersecurity.

[439] Compare *e.g.* article 5(1)(d) GDPR that requires "accurate" and "kept up to date" personal data with article 15 AI Act stating that "high-risk AI systems shall be designed and developed in such a way that they achieve (...) an appropriate level of accuracy".

[440] De Gregorio & Dunn, *supra* note 347.

that address issues related to the development and use of AI, as outlined in the EC's 2020 White Paper on AI.[441] Consistency is therefore ensured with other ongoing or planned initiatives of the EC that also aim to address digitalization, including the revision of sectoral legislation and initiatives that address liability related to new technologies, including AI systems. Since those systems are commonly (though debatably)[442] perceived as products and safety components of products, the AI Act - the EC argues - is consistent with the sectoral product safety legislation that follows either from the EU New Legislative Framework (NLF) on the one hand, or from the Old Approach, on the other hand.[443]

The NLF includes Regulation n° 765/2008 setting out the requirements for accreditation and the market surveillance of products, Regulation n° 2019/1020 on market surveillance and compliance of products, and Decision 768/2008 on a common framework for the marketing of products. [444] This last document provides reference provisions to be incorporated directly in product legislations whenever they are revised. It was the case for toy safety,

---

[441] COM (2020) 65, *supra* note 24.

[442] The key issue with the commoditization of AI has been that of agency. Unlike 'ordinary' products with relatively fixed and verifiable features, AI is specific in that it displays a level of cognitive autonomy which may make its prediction and decision-making processes unpredictable (the topical example here being that of the black box scenarios). Since advanced ML systems display forms of intelligence akin to human intelligence, scholarship (mainly focused on AI liability) has inquired whether the standard AI commoditization is still tenable. Exploring the topic of AI agency goes beyond the scope of this Article. For a more detailed study on this point, *see* Lawrence B. Solum, *Legal Personhood for Artificial Intelligences*, 70 NCLREV. 1231 (1992).

[443] *New legislative framework,* EURO. COMM'N, https://ec.europa.eu/growth/single-market/goods/new-legislative-framework_fr (last visited Mar. 31, 2023).

[444] Regulation 765/2008 of the European Parliament and of the Council of 9 July 2008 setting out the requirements for accreditation and market surveillance relating to the marketing of products, OJ L 218/2008, 30-47; Regulation 2019/1020 of the European Parliament and of the Council of 20 June 2019 on market surveillance and compliance of products, OJ L 169/2019, 1-44; Decision 768/2008 of the European Parliament and of the Council of 9 July 2008 on a common framework for the marketing of products, OJ L 218/2008, 82-128.

transportable pressure equipment, restriction of hazardous substances in electrical and electronic equipment, construction products, pyrotechnic articles, recreational craft and personal watercraft, civil explosives, simple pressure vessels, electromagnetic compatibility, non-automatic weighing instruments, measuring instruments, lifts, ATEX, radio equipment, low voltage, pressure equipment, marine equipment, cableway installations, personal protective equipment, gas appliances, medical devices, in vitro diagnostic medical devices, and EU fertilising products. [445]

[445] Directive 2009/48 of the European Parliament and of the Council of 18 June 2009 on the safety of toys, OJ L 170/2009, 1-37; Directive 2010/35 of the European Parliament and of the Council of 16 June 2010 on transportable pressure equipment, OJ L 165/2010, 1-18; Directive 2011/65 of the European Parliament and of the Council of 8 June 2011 on the restriction of the use of certain hazardous substances in electrical and electronic equipment, OJ L 174/2011, 88-110; Regulation 305/2011 of the European Parliament and of the Council of 9 March 2011 laying down harmonised conditions for the marketing of construction products, OJ L 88/2011, 543; Directive 2013/29 of the European Parliament and of the Council of 12 June 2013 on the harmonisation of the laws of the Member States relating to the making available on the market of pyrotechnic articles (recast), OJ L 178/2013, 27-65; Directive 2013/53 of the European Parliament and of the Council of 20 November 2013 on recreational craft and personal watercraft, OJ L 354/2013, 90-131; Directive 2014/28 of the European Parliament and of the Council of 26 February 2014 on the harmonisation of the laws of the Member States relating to the making available on the market and supervision of explosives for civil uses (recast), OJ L 96/2014, 1-44; Directive 2014/29 of the European Parliament and of the Council of 26 February 2014 on the harmonisation of the laws of the Member States relating to the making available on the market of simple pressure vessels, OJ L 96/2014, 45-78; Directive 2014/30 of the European Parliament and of the Council of 26 February 2014 on the harmonisation of the laws of the Member States relating to electromagnetic compatibility (recast), OJ L 96/2014, 79-106; Directive 2014/31 of the European Parliament and of the Council of 26 February 2014 on the harmonisation of the laws of the Member States relating to the making available on the market of non-automatic weighing instruments, OJ L 96/2014, 107-148; Directive 2014/32 of the European Parliament and of the Council of 26 February 2014 on the harmonisation of the laws of the Member States relating to the making available on the market of measuring instruments (recast), OJ L 96/2014, 149-250; Directive 2014/33 of the European Parliament and of the Council of 26 February 2014 on the harmonisation of the laws of the Member States relating to lifts and safety components for lifts, OJ L 96/2014, 251-308; Directive 2014/34 of the European Parliament and of the Council of 26 February 2014 on the harmonisation of the laws of the Member States relating to equipment and protective systems intended for use in potentially

AI systems that are safety components of products under the NLF will be considered as high-risk AI system because products regulated under this approach are already submitted to a third party conformity assessment that "already presupposes a risk assessment on the safety risks posed by the products covered by that instruments."[446] The Commission added that "it makes therefore sense to rely on the risk classification of a product under the relevant NLF legislation to define when an AI-driven safety component (of that product) should be considered high-risk."[447] Yet - and perhaps counterintuitively - from these twenty-three EU secondary law instruments, only twelve were selected to be qualified as high-risk.[448]

Regarding AI systems covered by sectoral legislations that do not follow the New Legislative Framework (the so-called old approach), the Commission noted that AI systems that are safety components of products under the old approach legislation will be considered as high-risk AI system because products regulated under

---

explosive atmospheres (recast), OJ L 96/2014, 309-356; Directive 2014/53 of the European Parliament and of the Council of 16 April 2014 on the harmonisation of the laws of the Member States relating to the making available on the market of radio equipment and repealing Directive 1999/5, OJ L 153/2014, 62-106; Directive 2014/35 of the European Parliament and of the Council of 26 February 2014 on the harmonisation of the laws of the Member States relating to the making available on the market of electrical equipment designed for use within certain voltage limits, OJ L 96/2014, 357-374; Directive 2014/68 of the European Parliament and of the Council of 15 May 2014 on the harmonisation of the laws of the Member States relating to the making available on the market of pressure equipment, OJ L 189/2014, 164-259; Regulation 2016/424 of the European Parliament and of the Council of 9 March 2016 on cableway installations, OJ L 81/2016, 1-50; Regulation 2016/425 of the European Parliament and of the Council of 9 March 2016 on personal protective equipment, OJ L 81/2016, 51-98; Regulation 2016/426 of the European Parliament and of the Council of 9 March 2016 on appliances burning gaseous fuels, OJ L 81/2016, 99-147; Regulation 2017/746 of the European Parliament and of the Council of 5 April 2017 on *in vitro* diagnostic medical devices, OJ L 117/2017, 176-332; Regulation 2019/1009 of the European Parliament and of the Council of 5 June 2019 laying down rules on the making available on the market of EU fertilising products, OJ L 170/2019, 1-114.

[446] SWD (2021) 84 Final, Part 2/2, *supra* note 163 at 38 n 33.

[447] *Id.*

[448] *Id.* at 37-39.

this approach "always undergo third party conformity assessments or authorisation procedures in the legislations that will be covered by the new AI initiative."[449] The Commission then labelled as high-risk AI systems that are safety components related to civil aviation,[450] motor vehicles,[451] agricultural and forestry vehicles,[452] two- or three-wheel vehicles and quadricycles,[453] interoperability of railway systems,[454] and marine equipment.[455]

Regarding AI systems provided or used by regulated credit institutions, the EC stressed that the conformity assessment procedures under the EU's financial services legislation and some of the providers' procedural obligations under this proposal are integrated into the procedures under Directive 2013/36[456] on access

---

[449] *Id.* at 39 n. 34.

[450] Regulation 2018/1139 of the European Parliament and of the Council of 4 July 2018 on common rules in the field of civil aviation and establishing a European Union Aviation Safety Agency, and amending Regulations (EC) No 2111/2005, (EC) No 1008/2008, (EU) No 996/2010, (EU) No 376/2014 and Directives 2014/30/EU and 2014/53/EU of the European Parliament and of the Council, and repealing Regulations (EC) No 552/2004 and (EC) No 216/2008 of the European Parliament and of the Council and Council Regulation (EEC) No 3922/91, OJ L 212, 22 August 2018, 1-122.

[451] Commission Regulation 2018/858, of the European Parliament and of the Council of 30 May 2018 on the approval and market surveillance of motor vehicles and their trailers, and of systems, components and separate technical units intended for such vehicles amending Regulations (EC) No 715/2007 and (EC) No 595/2009 and repealing Directive 2007/46/EC, 2018 O.J. (L 151) 1.

[452] Commission Regulation 167/2013, of the European Parliament and of the Council of 5 February 2013 on the approval and market surveillance of agricultural and forestry vehicles, 2013 O.J. (L 60) 1, 51.

[453] Commission Regulation 168/2013, of the European Parliament and of the Council of 15 January 2013 on the approval and market surveillance of two- or three-wheel vehicles and quadricycles, 2013 O.J. (L 60).

[454] Directive 2016/797, of the European Parliament and of the Council of 11 May 2016 on the interoperability of the rail system within the European Union, 2016 O.J. (L 138),44, 101.

[455] Directive 2014/90/EU of the European Parliament and of the Council of 23 July 2014 on marine equipment and repealing Council Directive 96/98/EC, 2014 OJ (L 257), 146-185. As hinted above, this one belongs to the NLF but will be treated like those belonging to the old approach.

[456] Directive 2013/36 of the European Parliament and of the Council, of 26 June 2013, on access to the activity of credit institutions and the prudential supervision of credit institutions and investment firms, 2013 O.J. (L 176), 338–436.

to the activity of credit institutions and the prudential supervision. The AI Act is also said to be consistent with the applicable Union legislation on services, including on intermediary services regulated by the e-Commerce Directive[457] and the Digital Services Act (DSA).[458]

As regards AI systems that are components of large-scale IT systems in the Area of Freedom, Security and Justice managed by the EU's Agency for the Operational Management of Large-Scale IT Systems (EU-LISA), the AI Act will apply to systems that have been placed on the market or put into service after a year has elapsed from the date of its entry into force, unless the replacement or amendment of those legal acts leads to a significant change in the design or intended purpose of the AI system or systems concerned.

The interesting issue, of course, is whether the AI Act's congruence with values and consistency with existing norms are enough for the human-centric risk-based approach to be *effectively* operationalized in practice. The LEADS experts from the University of Birmingham were critical in this regard arguing that "there is a lack of clarity concerning the obligation to ensure that those deploying 'high-risk' systems have quality management systems in place to ensure for respect of fundamental rights extends to mechanisms that guard against rights interferences arising from activities and actions other than those of Member States and EU authorities."[459] In order to ensure that all actors involved - both State and non-State - respect fundamental rights in the process of

---

[457] Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market ('Directive on electronic commerce'), 2000 OJ (L 178) 1,16.

[458] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act) OR *Proposal for a Regulation of the European Parliament and of the Council on a Single Market For Digital Services* (Digital *Services Act) and amending Directive 2000/31/EC*, COM (2020) 825 final (Dec. 15, 2020).

[459] Nathalie A. Smuha et al., *How the EU can achieve Legally Trustworthy AI: A Response to the European Commission's Artificial Intelligence Act*, SSRN (2021), at 40 https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3899991.

deploying, using and market-releasing of AI systems, the experts recommended two series of amendments. First, an amendment aligning the AI Act with the GDPR's obligations on the collection and processing of personal data, applied to all 'data controllers' irrespective of whether they are public or private persons.[460] Second, an amendment through which the obligation to respect fundamental rights would apply to private-sector actors in a direct and unmediated way (*i.e.,* without the mediation of State law or conventional human rights law).[461] This is justified - the experts argue - by the "serious asymmetry between those directly affected by AI systems, and the organisations with the resources and expertise to deploy them, given the capacity of these systems to operate automatically, at scale and in real time."[462]

The criticism voiced by the LEADS group of experts, in combination with our analysis of the fact-neutrality of the AI Act, gives food for thought because of, what we might call, fundamental failures to meet traditional requirements for legal validity.

On the one hand, from a fact-law perspective, it is far from certain that the AI Act gives a normative translation to conclusions drawn from – what we earlier called – consistent premises *i.e.,* reality-based inferences on the risks considered for regulatory address. The AI Act is certainly consistent with overarching regulatory ambitions pertaining to the Digital Single Market, but if one were to perceive this instrument, not as a piece in a larger regulatory puzzle, but as an individual normative choice based on a process that ought to have included fact-based reasoning, we would have little motive to believe that the risk scale this proposal suggests is derived from a form of acceptable (because justified) knowledge.

On the other hand, from a law-policy perspective, and as stressed by the LEADS experts, the AI Act is consistent with the 'risk-based' part of the regulatory approach it follows but fails the consistency test with the 'rights-based' dimension of that approach. This is no doubt due to the EC's omission to fortify the practical means through which the AI Act would effectively operationalize the rights-based

---

[460] *Id.*
[461] *Id.* at 41.
[462] *Id.*

dimension of the regulatory approach it incarnates... Perhaps - here again - a closer look at real-life practices (*i.e.*, evidence!) would have allowed the EC to better detect the data-processing and fundamental rights issues that actually (as opposed to hypothetically) occur in various contexts of AI deployment and use. It is therefore interesting to anticipate if the AI Act, when enacted, will pass the proportionality test the CJEU usually applies to 'standard' risk regulation.

V. A PEAK INTO THE FUTURE: CAN THE AI ACT PASS THE PROPORTIONALITY TEST?

The criteria for the epistemic validity of knowledge (of risks), essentially allow to answer the question 'how do we *rationally* construct knowledge of risks?' Alternatively, the conditions for the validity of the normative translation of that knowledge seek to answer the question 'how does policy impact the construction of knowledge of risks?'

The 'validity' of a body of knowledge *tout court* is usually warranted by compliance with commonly accepted procedures and standards of discovery. However, when knowledge is sought for the purpose of policy, it is not only 'valid' under criteria that frame the *rationality* of the discovery process, but also criteria that frame its *legality*. Fact-based policy (like 'standard' risk regulation) is, no doubt, one of the most telling examples of the *symbiosis* of sorts between epistemic and legal validity. Indeed, to yield valid risk regulation, fact-finding and fact-appraisal should be conducted to the extent of the epistemically justified *and* the legally allowed. This implies that *valid* fact-based regulation should be a response to *correctly* assessed facts; *a contrario*, instruments relying on flawed knowledge of facts should, in principle, be declared invalid. The assessment of the 'normative' validity allows to determine if, when integrated into policy, factual knowledge (ideally issued from the best possible evidence) yielded the best possible standard of protection. In EU law, the adequacy of the fact-to-law translation is assessed under the principle of *proportionality* which regulates the exercise of the powers conferred to the EU.

In modern legal scholarship and practice, proportionality takes the shape of a 'less restrictive alternative' test.[463] However, the justification and legal relevance of this test can be traced back to longstanding doctrines on justice. Roughly since Aristotle,[464] justice has been defined through the concept of proportionality, most saliently expressed in the 'treat like cases alike' principle. Distribution (of say rights and basic goods) and retribution are considered as 'just' or 'fair' if they guarantee desert proportionate to merit.[465] The Aristotelian virtues like justice and moderation[466] were eventually integrated in strands of legal scholarship, namely on points concerning the fact/law correspondence. Montesquieu e.g. posited that a legislator must act in the "spirit of moderation" for their legislation to strike a proper balance between the moral good and the political good.[467] More recently, Schwarze also espoused the proportionality/justice kinship, by distinguishing, on the one hand, proportionate - normative - responses to facts (which he associated with so-called vindicative justice) and, on the other hand, distribution of rights and duties (which derives from the so-called distributive justice).

The proportionality/justice kinship can be interpreted as revealing an important aspect of the 'empirical adequacy' requirement for policy: proportionate regulatory responses to facts are assumed to be most apt to *deliver justice*.[468] The above-mentioned example of tobacco consumption is a good example of this. A total ban of cigarette production would, no doubt, benefit consumers, but it would hardly be seen as fair by cigarette manufacturers who would see their profits plunder, and might

---

[463] Takis Tridimas, *Proportionality in Community law: Searching for the Appropriate Standard of Scrutiny*, *in* THE PRINCIPLE OF PROPORTIONALITY IN THE LAWS OF EUROPE 65, 66 (Evelyn Ellis ed., Hart Publishing 1999).

[464] *See* Aristotle, NICOMACHEAN ETHICS.

[465] For an illustration of the desert approach, *see* Yeung, *supra* note 72 at 72-77.

[466] *See generally* Aristotle, *supra* note 466.

[467] CHARLES DE SECONDAT, BARON DE MONTESQUIEU, SPIRIT OF THE LAWS 67 (Thomas Nugent Trans ed., Botche Books 2001).

[468] We paraphrase Neuwirth who – primarily focusing on adjudicatory contexts – held that law is asked to deliver justice and, to do so, the expectation is that it must be based on truth; *see* Rostam J. Neuwirth, *Law as Mnemonics: The Mind as the Prime Source of Normativity*, 2 EUR. J. LEGAL STUD. 143, 146 (2008).

encounter resistance from governments when substantial portions of States' budgets come from tax schemes on tobacco products. An incitement to *reduce* the consumption of those products seems to strike a workable equilibrium between the need to protect people's health and the need to, nevertheless, ensure the viability of tobacco industries. However, to design this fair distribution of burdens through proper risk management, prior knowledge of risks is paramount. In the US, the Memorandum for the Heads of Executive Departments and Agencies states that "the management of risks created by AI applications should be *appropriate to, and commensurate with, the degree of risk that an agency determines in its assessment.* In general (...) agencies should adopt a tiered approach in which the degree of risk and consequences of both success and failure of the technology determines the regulatory approach, *including the option of not regulating.* Agencies should be aware that there is always likely to be at least some risk, *including that associated with not knowing what is currently unknown.* For AI applications that pose lower risks, agencies can rely on less stringent and burdensome regulatory approaches - or non-regulatory approaches such as requiring information disclosures or consumer education. For higher risk AI applications, agencies should consider, for example, the effect on individuals, the environments in which the applications will be deployed, the necessity or availability of redundant or back-up systems, the system architecture or capability control methods available when an AI application makes an error or fails, and how those errors and failures can be detected and remediated."[469]

On the surface, the AI Act confirms that it is "proportionate and necessary to achieve its objectives, since it follows a risk-based approach and imposes regulatory burdens only when an AI system is likely to pose high risks to fundamental rights and safety."[470] It is interesting to note that the proportionality of the proposal is

---

[469] Memorandum for the Heads of Executive Departments and Agencies, *supra* note 38 at 14.
[470] COM (2021) 206 final, *supra* note 29 at 6-7.

interpreted as *adequacy of the regulatory burdens*[471] in light of risk-taxonomy the AI Act includes. For limited risk AI systems, the EC emphasized that "only very limited transparency obligations are imposed, for example in terms of the provision of information to flag the use of an AI system when interacting with humans."[472] Alternatively, for high-risk AI systems, "the requirements of high quality data, documentation and traceability, transparency, human oversight, accuracy and robustness, are strictly necessary to mitigate the risks to fundamental rights and safety posed by AI and that are not covered by other existing legal frameworks."[473] Moreover, harmonised standards and supporting guidance and compliance tools "will assist providers and users in complying with the requirements laid down by the proposal and minimise their costs. The costs incurred by operators are proportionate to the objectives achieved and the economic and reputational benefits that operators can expect from this proposal."[474]

The EC's view of proportionality seems to rely on a cost-benefit trade-off which is certainly a valid way of evaluating the 'moderation' with which policy ought to respond to facts.[475] However, bearing in mind that the proportionality of risk-regulation is assessed in reference to the evidence of risks upon which it relies, the EC's view of proportionality lacks some empirical justification. The Commission merely asserted that the obligations contained in the AI Act were tailored in reference to the scale of risks in the latter, without providing any explanation on how the evidence gathered justified this scale in the first place. Bearing in mind the *Monsanto*[476] requirement of completeness of the evidence relied on when characterizing risks, the EC's discussion on the principle of proportioanality does not allow to argue that the normative standards in the AI Act truly correspond to (because they would ideally derive

---

[471] *See id.*
[472] *Id.*
[473] *Id.*
[474] *Id.*
[475] *Contra,* to some extent, Carl F. Cranor, *Towards a non-consequentialist approach to acceptable risks, in* RISK: PHILOSOPHICAL PERSPECTIVES 36, 36-53 (Tim Lewens ed., 2007).
[476] *Monsanto SAS et al., supra* note 198.

from) factual reality. As argued before, the impression that the AI Act gives is that, said standards were tailored after a taxonomy of risks suggested in preexisting instruments which seemed to have decisively shaped the EC's views on how to classify AI systems on a scale of risks.[477] Assuming that the proportionality of risk regulation is sanctioned when epistemically valid discovery of risks informs on the standard of protection against those risks, what seems to hold up the AI Act as a 'proportionate' regulatory instrument is, indeed, its *consistency* with prior regulatory incentives and existing EU legislation. But is this enough for the AI Act to stand as a valid (because proportionate) instrument? To answer this question, we need to look at some of the trends in the CJEU's caselaw on the assessment of the proportionality of EU Secondary law. Montesquieu's postulate on legislative moderation' is expressed in Article 5(4) TEU which states that "the content and form of Union action shall *not exceed what is necessary to achieve the objectives of the Treaties*."[478] It follows from the wording of this provision that the metric for proportionality is the *necessity* to achieve the objectives assigned to the Union,[479] the key issue then being how to determine if facts generate a need for a specific type of regulation. An EU scholar might argue that the assessment of this type of 'necessity' is a matter of institutional discretion with no rigid, preestablished criteria on how that discretion ought to be exercised. After all, the CJEU has repeatedly stressed that, in the exercise of their powers, the Union institutions must be allowed broad freedom in areas where their action involves political, economic and social choices, implying complex assessments and evaluations. The only circumstance in which the CJEU would likely sanction 'disproportionate' exercise of discretion is the - statistically rare - case of a manifestly inappropriate measure with regard to the objective which the competent institution is seeking to pursue.[480]

---

[477] *See supra* Section II.

[478] *Monsanto SAS et al., supra* note 198.

[479] *See* Tridimas, *supra* note 465.

[480] See *Inter Alia*, CJEU, 12 July 2001, *Jippes et al.*, case C-189/01, EU:C:2001:420, at 82-83; CJEU, 10 December 2002, *British American Tobacco (Investments) and Imperial Tobacco*, case C-491/01, EU:C:2002:741,

Notwithstanding the breadth of the institutional discretion in deciding on whether facts (as established through evidence) warrant legislative action, both EU positive law and the CJEU's caselaw reveal several requirements that frame the exercise of that discretion. Various EU law provisions include guidelines which, though broad, direct the EU legislature on how the evidence gathered on a specific issue (say, risks) ought to be translated into policy. In EU primary law, Article 2 of Protocol n° 2 annexed to the Treaties clarifies the scope of the EC's fact-finding duty by stating that "before proposing legislative acts, the Commission *shall consult widely*. Such consultations shall, where appropriate, take into account the regional and local dimension of the action envisaged."[481] In EU soft law, the Commission's Communication on the precautionary principle[482] goes even further, by shedding more light on the evidence-proportionality interrelationship. Within the meaning of said Communication, *proportionality* means "tailoring measures to the *chosen level of protection*. Risk can rarely be reduced to zero, but *incomplete risk assessments* may greatly reduce the range of options open to risk managers. A total ban may not be a proportional response to a potential risk in all cases. However, in certain cases, it is the sole possible response to a given risk."[483]

The cited Communication seems to confirm that proportionality is tangibly expressed in a *chosen* level of protection and requires *completeness* of a given risk assessment. Completeness as an evidentiary and an epistemic requirement was discussed earlier.[484] Bearing in mind the *Monsanto* legacy,[485] the CJEU has stated that the scope of its discretion notwithstanding, discovery in view of fact-

---

at 123; CJEU, 12 July 2005, *Alliance for Natural Health et al.*, joined cases C-154/04 and C-155/04, EU:C:2005:449, at 52; CJEU, 7 July 2009, *S.P.C.M. et al.*, case C-558/07, EU:C:2009:430, para. 42; CJEU, 8 June 2010, *Vodafone*, case C-58/08, EU:C:2010:321, at 52.

[481] Protocol n° 2 on the application of the principles of subsidiarity and proportionality, OJ L n° *15, 9/5/2008, p. 206.*

[482] *Communication From the Commission of 2 February 2000 on The Precautionary Principle, supra* note 221.

[483] *Id.* at 3.

[484] *See supra* Section IV.

[485] *See Monsanto SAS et al., supra* note 198.

based policy does not translate to evidentiary free-styling but must be based "on objective criteria."[486] Indeed, "in assessing the burdens associated with various possible measures [the EU legislature] must examine whether objectives pursued by the measure chosen are such as to justify even substantial negative economic consequences for certain operators."[487]

The more interesting question, at this stage, is: how does the EC establish a '*chosen* level of protection'? According to the Communication on the precautionary principle, the yardstick for this is - unsurprisingly - *consistency* with already existing policy. Indeed, though warranted by a body of evidence, new risk regulating measures "*should be of comparable* scope and nature to those already taken in equivalent areas in which all scientific data are available."[488] For example, in the *NFU*[489] case, annulment proceedings were brought against a Commission's decision, adopted on the grounds of two Directives from 1989 and 1990, which set out, *inter alia*, a prohibition on the export of bovine meat to third countries. The applicant argued that this prohibition was disproportionate, as it resulted in a worldwide ban.[490] The Council's reasoning in this case was interesting as it is essentially construed around an *argument of consistency* to justify the extent of the Commission's risk-assessment powers. Indeed, the Council argued that said Directives "form part of a *coherent and exhaustive body of*

---

[486] CJEU, 8 June 2010, *Vodafone*, case C-58/08, EU:C:2010:321, at 52.
[487] COM (2000) 1 final, *supra* note 221.
[488] *Id.* at 4; *see also* Giandomenico Majone, *Foundations of Risk Regulation: Science, Decision-Making, Policy Learning and Institutional Reform*, 1 *E. J. Risk Reg.* 5, 6 (2010) (similarly, Majone argues that the consistency in risk regulation compensates the lack of knowledge about risks. In the context of regulating in certainty (*i.e.*, when the outcomes are determined unambiguously), the key concept is optimization e.g. maximizing an expected return without considering its variance. Alternatively, the key concept in decision-making under uncertainty is consistency if no generally accepted criterion for the correctness of the decision exist, decision-makers are urged to turn to 'procedural rationality', in order for a given system to cope with changing circumstances and the absence of specific solutions).
[489] CJEU, 5 May 1998, *National Farmers' Union et al.*, *supra* note 254.
[490] *Id.* at 11.

law established in order to substitute a set of common rules for unilateral action on the part of each Member State pursuant to Article 36 of the Treaty."[491] In this context, "by *acting in a prudent manner* and pursuing the *safest option for public health*, the Commission has neither made a manifest error in its assessment of the risk to animal or human health nor manifestly exceed the powers conferred on it."[492] The Court followed the Council's views, stating that, under the Common Agricultural Policy, the Union legislature does, indeed, enjoy broad discretion,[493] finding that the 1989 and 1990 Directives were drafted "in very wide terms and empower[ed] the Commission to act 'in all cases' and adopt '*the necessary measures*."[494] Similarly, in a *UK v. Commission* case[495] (on the workers' rest time Directive) the UK government argued, *inter alia*, a violation of the principle of proportionality, by submitting to the Court that the minimum requirements aimed at safeguarding a level of health and safety protection of workers could be attained through less restrictive measures and involve fewer obstacles on competitiveness and the workers' earning capacity. The applicant argued that "neither the Commission's proposal nor the directive *provide any explanation as to why the desired level of protection* could not have been achieved by less restrictive measures, such as, for example, the use of risk assessments if working hours exceed particular norms."[496] By reaffirming the limited judicial review[497] in verifying the EU

---

[491] *Id.* at 12.

[492] *Id.*

[493] *Id.* at 13-14 ("the Council may be prompted to confer on the Commission wide implementing powers, since the Commission alone is able to continually and closely monitor trends on the agricultural markets and to act with urgency if the situation so requires. Such powers are all the more justified when they are to be exercised in accordance with a procedure which allows the Council to reserve its right to intervene.)"

[494] *Id.*

[495] Case C-84/94, *UK v. Commission*, EU:C:1996:431 (Nov. 12, 1996).

[496] *Id.* at 18.

[497] *Id.* at 19 ("as to judicial review of those conditions, however, the Council must be allowed a wide discretion in an area which, as here, involves the legislature in making social policy choices and requires it to carry out complex assessments. Judicial review of the exercise of that discretion must therefore be limited to examining whether it has been vitiated by manifest error or misuse of powers, or

institutions' discretion, the CJEU ruled that "the measures on the organization of working time which form the subject-matter of the directive (...) *contribute directly to the improvement of health and safety protection for workers* (...) and cannot therefore be regarded to the purpose of achieving the objective pursued."[498]

Unsurprisingly, the cited cases reveal the CJEU's *functional* interpretation of the level of consistency that a new measure ought to present with an existing body of law. The Court did not require that an aspect of, say, environment protection be exhaustively harmonized, nor did it specify the type of legislative instruments (regulation, directive) that the 'existing body of law' should include in order to be taken as a referent for consistency. What seems to ultimately matter is whether the 'new' measure is in line with the Treaty objectives which assign, to the Union, the mission to protect namely environment, health and consumers. This functional approach to proportionality and consistency transpires from the AI Act's provisions and may, ultimately, be its saving grace, if the legality of this instrument is challenged on the grounds that it violates the principle of proportionality. Considering the political and societal interests at stake, the CJEU may find that the AI Act's congruence with fundamental values and consistency with existing Union legislation are enough for this instrument to be considered as proportionate. But the spectre of *Monsanto* is still present: though the EC could prove that it formally complied with the obligation to gather evidence (through consultation), it would have a hard time arguing that that evidence met the *Monsanto* requirements of completeness and objectivity as preconditions for the proportionality requirements to be fully met. Perhaps, we will see a two-standard application of this principle in the future. To be proportionate, standard risk-regulation in say environment and health would still need to rely on evidence to pass the CJEU's proportionality test. Alternatively, for risk-regulation, like the AI Act, that addresses risks of fundamental rights violations, consistency with existing

whether the institution concerned has manifestly exceeded the limits of its discretion,").

[498] *Id.* at 59.

policy and congruence with fundamental values may be the decisive factors in assessing conformity with the proportionality principle. At this stage, these are of course speculative observations. We will need to wait and see if the CJEU's caselaw will shed more light on this point in the future...

VI. CONCLUDING REMARKS

This article addresses the issue of whether regulators can or should rely on real-life experience when drafting AI regulation. The EC's ambition was, indeed, to establish an exhaustive list of AI systems labelled as 'high-risk,' the standard assumption being that the nature and intensity of a risk should be backed by some type of empirical evidence. In this context, our article first analysed the epistemic soundness of both the discovery procedures launched by the EC and the validity of the interpretations the Commission gave to the findings of those procedures. Using evidence theory as an analytical framework, we concluded that, while the EC's discovery procedure was *prima facie* aimed at gathering factual knowledge on the types of AI-related risks, that procedure was actually a pro-forma fact-finding enterprise, meant to support a regulatory choice the EC had made prior to launching discovery. Indeed, the opinions gathered by the Commission in its two public consultations are not reflected in the AI Act as shown by our comparative analysis between the results of those consultations and the AI Act's provisions. A closer look into the epistemic validity of the EC's evidence-gathering and the inferences made on the basis of the evidence gathered allows us to assert that the AI Act is not a normative translation of legitimate conclusions, drawn from consistent factual premises. It is rather a response to the policy strategy previously outlined in the 2020 White Paper on AI. In other words, the AI Act is not reflective of facts but of overarching regulatory ambitions pertaining to the Digital Single Market. The upshot of this is twofold. First, choosing a risk-based approach in regulating AI - that defines the level of protection in reference to the intensity of the risk(s) concerned - is coherent with the EC's policy objectives to balance an ecosystem of trust (protection oriented) and an ecosystem of excellence (investment oriented). Second, our analysis of the evidence-to-regulation leap

shows that the chosen risk scale falls short on real-life practices. Considering that the concept of risk is defined in relation to fundamental rights violations, the AI Act is specific in comparison to 'standard' risk-regulating instruments which are generally required to rely on sound empirical and probabilistic evidence of harm. Against the backdrop of these 'standard' requirements, the EC's risk-based approach in regulating AI makes sense (given the goal of preventing AI-related harm) but the EC's definition of risk does not (since 'risks' within the AI Act are not probable occurrences of physical harm but probable violations of fundamental rights).

This study is not a pamphlet against the AI Act; it rather raises a discussion on the rationale of the regulatory model this instrument embodies. In our analysis of this model, we stressed the importance of the so-called axiological congruence of the AI Act with overarching ethical principles and values. It seems that, given the cardinal importance of safeguarding those values and the rights they are expressed through, the empirical knowledge of risks raised by AI systems needed to take the back seat in the factors having decisively shaped the EC's normative choices. Since it seeks to uphold a high level of protection against fundamental rights' violations, the AI Act is both a risk-based and a rights-based regulatory instrument. As such, it does not - surprisingly - grant any individual rights to EU citizens but gives idiosyncratic expression to fundamental rights enshrined in the ECFR, under the assumption that appropriate fundamental rights protection will be achieved through *ex ante* standardization. It remains to be seen if, when the AI Act is enforced, applied and gives way to CJEU caselaw, technical standardization will, indeed, prove to be the appropriate shield against fundamental rights violations the instrument under consideration so ambitiously aims to prevent.