Rozgonyi, Krisztina, Mari-Liisa Parder and Rodrigo Conde Jiménez. "Ethics in Regulating Artificial Intelligence: An Overview of the Recent Legislation in the European Union." *Genealogy+Critique* 12, no. 1 (2026): 1–36. DOI: https://doi.org/10.16995/gc.25624

Open Library of Humanities

Ethics in Regulating Artificial Intelligence: An Overview of the Recent Legislation in the European Union

Krisztina Rozgonyi, Austrian Institute of Technology, Krisztina.Rozgonyi@ait.ac.at
Mari-Liisa Parder, University of Tartu Center of Ethics
Rodrigo Conde Jiménez, Austrian Institute of Technology

This article examines how ethical values and principles are embedded and operationalised within the European Union's artificial intelligence (AI) legal framework. Combining a systematic interpretive mapping of ethical principles with legal doctrinal and ethics-informed analysis, the study analyses six core EU instruments relevant to AI governance: the AI Act, GDPR, Digital Services Act, Data Governance Act, Data Act, and DSM Directive. It investigates which ethical values are reflected in these laws, how they are prioritised, and to what extent they are legally enforceable. The analysis shows that non-maleficence, fairness, and privacy form the enforceable ethical core of EU AI regulation, primarily implemented through risk-based safeguards and data protection obligations. Explicability plays a prominent but largely supportive role, mainly operationalised through transparency requirements. By contrast, values such as democracy, solidarity, and sustainability remain weakly embedded and appear predominantly in non-binding provisions. This distribution reveals a friction between the EU's ambition to position itself as a global leader in ethics-based, human-centric AI governance and the selective legal translation of ethical commitments.

Genealogy+Critique is a peer-reviewed, open-access journal published by the Open Library of Humanities.
© 2026 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See http://creativecommons.org/licenses/by/4.0/.
OPEN ACCESS

1. Introduction

Artificial Intelligence (AI) represents a multifaceted construct within societal and political imaginaries. On the one hand, it is perceived as a driving force for economic growth, technological innovation, and global competitiveness with the capacity to significantly enhance productivity, reshape labour markets, and influence long-term economic trajectories (Filippucci et al. 2024). Governments worldwide have been investing heavily in AI development, endorsed by major international organisations, envisioning progress through its wide-scale deployment. On the other hand, AI has increasingly become associated with societal fears and dystopian projections of the future dominated by uncontrollable and dehumanised systems that diminish human agency and threaten individual well-being (e.g., Cave et al. 2018). These competing narratives not only coexist but frequently move in antagonistic directions, often without effective mechanisms for reconciliation or mitigation. In times of such technological and normative uncertainty, many actors seek guidance through rules, norms, standards, and principles intended to govern change and ensure a smooth transition into an uncertain future. Among these normative instruments, the *law* plays a distinctive role due to its binding character and formal authority in structuring governance and societal change. However, AI, as an evolving and disruptive technological phenomenon, presents a regulatory challenge.

Historically, during the phases of digitisation and digitalisation, regulatory responses often relied on self-regulation and co-regulation to manage rapid technological changes, engaging the tech industry as an assumed key stakeholder with intimate knowledge of the technologies involved (Marsden 2011). These models, in several instances, successfully fostered legal clarity and supported innovation (Brownsword 2008). In such cases, shared interests between governments and technology actors facilitated cooperative regulatory frameworks. However, in the absence of aligned policy agendas and considering the increasing dominance of oligopolistic and powerful tech companies, this cooperation frequently gave way to unilateral, corporate-driven 'solo-regulation' (Milosavljević & Micova 2016). Efforts to regulate AI have mirrored many aspects of the earlier digital technology regulatory pathways (i.e. risk-based approach). Yet, the uncertainties surrounding AI's societal implications have triggered a somewhat distinct regulatory trajectory. Rather than pursuing wide-scale collaborative regulatory models, where minimal consensus on basic normative standards might have provided a foundation, 'ethics' was introduced as a concrete policy option. This move was accompanied by the promise that "striving for ethics and ethical decision-making [...] will make technologies better" (Wagner 2018, 84). However, ethics was also critiqued as a soft and potentially evasive substitute for regulation, offering a veneer of

responsibility in contexts where resistance to binding legal norms prevailed: "the rise of the ethical technology debate ran in parallel to the increasing resistance to any regulation at all" (ibid., 85).

Ethics and law are closely connected yet distinct domains: ethics offers philosophical guidance on what is right or wrong, while law sets enforceable societal rules. Ethical theories, such as deontology, utilitarianism, virtue ethics, and care ethics, among others, inform moral judgment, whereas laws are shaped through political and social processes and may reflect, but not fully encompass, ethical values. While law establishes minimum standards and legal responsibility within specific jurisdictions, ethics can go beyond legal compliance, can be applied universally and more flexibly. This distinction becomes particularly important when addressing emerging and unknown challenges – such as AI – where ethical reasoning may fill gaps left by slower-moving legal systems. Hence, across jurisdictions, a wide array of ethical values, such as fairness, transparency, accountability, and human-centricity, are reflected in AI policy guidelines and strategies (Hagendorff 2020) underlying legal frameworks.

Within this context, the presence of ethics in the body of AI policy and legal instruments has been a feature since their conception, often due to these instruments' interlinkage with the concept of 'ethics' as a (desirable) guideline for the development of AI as a technology. In this respect, authors such as Hagendorf (2020), Corrêa et. al (2023), or G'Sell (2024) have mapped the presence of ethics within this corpus: either as a collection of principles providing general guidance on AI development (Hagendorff 2020, 100–101), as a growing body of global policy guidelines that use ethical values as a guiding precept (Corrêa et al. 2023, 3), or as an implicit feature in some existing governance options for generative AI, albeit with a focus on 'weaker' forms of governance, such as self-regulation (G'sell 2024, 20). This last topic has raised the most criticism among several authors, who often perceive this ethical basis as an excuse towards more deregulatory efforts, and thus, as an avenue that can be exploited by private individuals seeking to opt out of more strict forms of AI regulation (Hagendorff 2020, 99–100; G'sell 2024, 21).

However, there remains a limited understanding of how, and to what extent, ethical values, principles and ethics-based policies have been embedded in binding legal acts that are directly or indirectly relevant to AI. This gap constitutes the starting point for our review article, which analyses current legislation in the European Union (EU) – an actor that aspires to function as a global ethics-based standard setter and to exercise normative regulatory power in relation to emerging technologies such as AI. To this end, the article first outlines the ethics-informed principles relevant to AI law and regulation, drawing on a survey of policy documents that have shaped EU legislation. Next,

we identify the legal acts subject to closer examination, whereby our corpus comprises seven EU legal acts (483 pages in total) selected to provide a comprehensive legal context for AI governance. We introduce these legal texts and explain their relevance after first tracing the origins and political dynamics behind AI regulation in the EU, particularly the tensions between safeguarding European values and ensuring global competitiveness. Through interpretative legal analysis, we assess the extent to which these principles are reflected in the legal texts, highlighting those that are prioritised and those that are marginalised. In particular, we examine whether ethics-based principles are incorporated into the enforceable (binding) or interpretative (non-binding) provisions of the legislation. We then focus on the most prevalent principles and contrast them with their modes of legal incorporation through doctrinal legal analysis (Byrne and Olsen 2024). Finally, we reflect on the EU's ambition to realise an ethics-informed regulatory approach to AI.

2. Defining Artificial Intelligence for Legal Inquiry

Artificial Intelligence systems have existed since the rise of computing and automation. However, definitions of what constitutes "Artificial Intelligence" have varied throughout the last decades, especially as systems evolve, becoming more complex, varied, and autonomous. The earliest mentions of AI stem from the birth of computing and automation in the 1950s. Here, the Dartmouth Proposal, written by McCarthy et al. (1955), defines AI as the possibility for a machine to simulate any aspect of learning or any other feature of intelligence as precisely as a human can (McCarthy et al. 2018). This definition was later expanded upon to include "the science and engineering of making intelligent machines, especially intelligent computer programs [...] that can solve problems and achieve goals in the world as well as humans" (McCarthy 2007, 2).

Other definitions for AI have underscored what capabilities a system would need to be considered intelligent. For example, Russel and Norvig (2022) provide a definition that is based on the Turing Test, that is, a system that can achieve natural language processing, knowledge representation, automated reasoning, and machine learning (Russell and Norvig 2022). However, since the advent of the most recent AI boom in the 2010s and 2020s, later definitions for AI focused on the technology's capabilities and societal implications have emerged from policy actors. For example, the EU High-Level Expert Group on Artificial Intelligence defines AI as a system "that displays intelligent behaviour by analysing its environment and taking actions – with some degree of autonomy – to achieve specific goals" (High-Level Expert Group on Artificial Intelli-

gence 2019, 24). Most notably, the Organization for Economic Cooperation and Development (OECD) defined AI as "a machine-based system that infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments" that can also act autonomously and adapt to their environment after their deployment (OECD 2024, 6–7).

The definition of an AI system included in the EU AI Act[1] is borrowed practically unchanged from the OECD's definition (EU AI Act, Art. 3), a widely accepted description eligible for legal interpretation. Therefore, for the purpose of this paper, we also define AI as a machine-based system that is designed to operate with varying levels of autonomy, that receives inputs and generates outputs with specific goals behind that can influence physical or virtual environments, and that may exhibit adaptiveness to its environment after its deployment.

3. Regulating AI in the European Union (EU)

Technological regulation in the EU has evolved rapidly, accelerating since 2022 with the adoption of a risk-based approach aimed at mitigating digital harms, exemplified by the Digital Services Act (DSA)[2] and the Digital Markets Act (DMA)[3]. Together with the General Data Protection Regulation (GDPR)[4], these instruments have reinforced the EU's normative power (Manners 2002), consolidating its reputation as a global standard-setter through norms (Wessel 2021) and underscoring its ambition for leadership in technology regulation as a form of geopolitical influence (Broeders, Cristiano, and Kaminska 2023). Within this broader context – particularly given the rapid rise of AI as a frontier technology – the notion of a global 'race' for AI regulation has gained prominence in policy debates. Scholars have observed the emergence of a competitive

[1] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance). PE/24/2024/REV/1; *OJ L, 2024/1689, 12.7.2024*.

[2] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act) (Text with EEA relevance) PE/30/2022/REV/1. *OJ L 277, 27.10.2022, p. 1–102*.

[3] Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act) (Text with EEA relevance) PE/17/2022/REV/1. *OJ L 265, 12.10.2022, p. 1–66*.

[4] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance). *OJ L 119, 4.5.2016, p. 1–88*.

dynamic in global AI governance. Smuha (2021) and Koniakou (2023), for example, describe an intensifying race for regulatory hegemony, reflecting both the pace of technological development and the strategic importance attributed to AI. Within this landscape, ethics-informed, 'trustworthy', and human-centric AI regulation has become strongly associated with the European policy sphere (European Commission 2025). This approach has attracted attention for its emphasis on ethical principles rather than alternative regulatory paradigms (Auernhammer 2020). At the same time, it has been subject to sustained critique, and some have questioned the adequacy of ethics-based regulation in the context of global competitiveness and innovation (Schneider 2024), while others highlight concerns about vagueness and limited enforceability, which may undermine regulatory certainty (Csernatoni 2025; Paul 2024). Such critiques have also been echoed by actors within EU institutions themselves (European Commission, European Political Strategy Centre 2025).

Against this background, understanding how and why ethical principles became central to the EU's approach to AI regulation is crucial. Ulnicane (2022) frames this development as both a continuation of the EU's foundational values and part of a longer debate on the legal and moral status of robots and autonomous systems – technologies historically shaped by ethical considerations. Similarly, Müller and Kettemann (2024) link ethics-based AI policy to established European legal traditions, notably respect for human rights and the Charter of Fundamental Rights of the European Union. In parallel, the EU's articulation of digital sovereignty – closely tied to human-centredness, trustworthiness, and ethical AI – has become a key element of its geopolitical positioning (Broeders, Cristiano, and Kaminska 2023) and an explicit political ambition (Delponte 2018; Christou, Meyer, and Fanni 2025), particularly in a global environment where ethical considerations are often marginalised in AI development (European Parliament 2020; Daly et al. 2021; Franke 2021).

Taken together, these dynamics suggest that an ethics-focused AI regulation constitutes a European geopolitical ambition that both reaffirms the EU's legal and policy traditions and positions it as a normative counterweight to less ethically grounded AI development globally (Sajduk and Dziwisz 2024; Franke 2021). At the same time, it may also serve as a means for the EU to assert global relevance in AI governance, primarily as a regulator and norm diffuser rather than as a technological frontrunner. This ambition must therefore be understood within the broader legislative ecosystem shaping AI in Europe, with the AI Act as a central pillar alongside the GDPR, DSA, DMA, and related frameworks governing data and intellectual property.

4. Laying Out an Ethics-Based Analytical Frame for Legal Analysis

Ethics, as a branch of philosophy, is concerned with questions of what is right, what is wrong, and how individuals ought to behave. Deontological theories, such as those deriving from Kant's categorical imperative, assess actions based on their intrinsic nature, namely, whether the act itself is right or wrong. In contrast, teleological theories, such as utilitarianism, focus on outcomes and judge actions by their consequences, for example, by aiming to maximise overall well-being. Other significant ethical theories include virtue ethics, which traces back to Ancient Greece and Aristotle and centres on the moral character of the actor, and care ethics, which has its roots in nursing ethics and prioritises relationships, empathy, and context-specific moral considerations.

More specifically, political philosophy is concerned with the relationship between ethics and law. Burg (2011) describes ethics and law as twin disciplines, noting the difficulty of clearly delineating the boundaries between legal norms and moral values, given the existence of numerous exceptions (Shklar 1964, as cited in Burg 2011). While the law usually reflects society's minimum expectations and codifies and enforces ethically desirable behaviour, ethics extends beyond these minimum standards (Trevino and Nelson 2021). Lennerfors (2019) argues that laws are socially and politically constructed and that the relationship between law and ethics is not inherent; laws may emerge from political negotiations that prioritise certain ethical values while neglecting others, and they may reflect unequal distributions of power within society. Nevertheless, it is also argued that unjust laws can and should be challenged through ethical critique and public discourse. Furthermore, actions may be considered ethical but not legal, and vice versa (Kotsios, Lennerfors, and Laaksoharju 2025; Chance et al. 2025).

A further dimension concerns legal compliance. Adhering to the law merely as a minimum standard, without considering broader ethical implications, can lead to a form of moral complacency. Taebi (2021) argues that ethical guidance can help clarify what constitutes the "right" action, particularly in situations where the law remains silent or insufficient. In the context of AI, it is important to note that law often lags behind technological developments, whereas ethics can be more responsive and adaptive. This has been demonstrated by Poel and Royakkers (2023), for example, in their analysis of the evolving concept of responsibility in relation to AI. However, the risk of "ethics washing" has become increasingly salient. Wagner (2018) warns that while policymakers broadly agree on the need for ethical oversight of AI, there is far less consensus on what this should entail. The invocation of vague ethical principles without concrete enforcement mechanisms risks delaying substantive regulation and the pre-

vention of societal harms, and ethics "toothless" (Rességuier and Rodrigues 2020) and "useless" (Munn 2023).

The EU's ethics-focused approach to AI regulation draws on a longstanding tradition of embedding fundamental rights, human dignity, and democratic values within its legal and policy architecture. While this vision is normatively compelling, it raises important questions regarding how abstract ethical principles are selected, operationalised, and translated into binding legal norms. To critically examine EU legal instruments relevant to AI, it is therefore necessary to systematically explore how ethical reasoning is embedded in the law, to what legal effect (binding or non-binding), and how key ethical concepts are interpreted in legal terms. Accordingly, our analysis first outlines the ethics-informed principles relevant to AI law, drawing on a survey of policy documents that have shaped EU legislation.

A wide range of ethical values – such as fairness, transparency, accountability, and human-centricity – feature prominently in AI strategies and guidelines across jurisdictions, including those of the EU (Hagendorff 2020). An expanding body of scholarship addresses which values should guide AI development and which ought to be embedded in AI systems themselves. Floridi and Cowls (2019), for instance, propose five core principles for ethical AI, drawing on the bioethical principles of beneficence, non-maleficence, autonomy, and justice, while adding explicability as a fifth principle. Hagendorff (2020, 103) further observes that certain values – such as accountability, explainability, privacy, justice, and robustness – are more readily operationalised in technical systems and therefore more likely to be implemented in AI design. At the same time, Sutrop (2020, 57) raises normative concerns regarding value alignment, questioning whether shared understandings of desirable outcomes are feasible given the plurality of human goals and preferences.

This plurality of values is also reflected in major international AI ethics instruments that have informed the EU's legislative work, including the Ethics Guidelines for Trustworthy AI (European Commission 2019)[5], the OECD AI Principles (OECD 2019)[6], and the UNESCO Recommendation on the Ethics of Artificial Intelligence (UNESCO 2021)[7]. These documents constitute forms of applied ethics and articu-

[5] European Commission. 2019. *Ethics Guidelines for Trustworthy AI*. Brussels: European Commission, High-Level Expert Group on Artificial Intelligence. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai.

[6] OECD. 2019. *OECD Principles on Artificial Intelligence*. Paris: Organisation for Economic Co-operation and Development. https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449.

[7] UNESCO. 2021. *Recommendation on the Ethics of Artificial Intelligence*. Paris: United Nations Educational, Scientific and Cultural Organization. https://unesdoc.unesco.org/ark:/48223/pf0000381137.

late key themes such as non-maleficence, autonomy, and justice, while seeking to operationalise them within policy frameworks. Methodologically, the ethical values and principles articulated in these instruments provide analytical access for examining EU AI legislation by tracing how such principles are embedded, prioritised, or diluted within the legal texts. By grounding our analytical categories in these authoritative documents, we aim to capture both the internal logic of EU AI policymaking and its external projection of regulatory power. Accordingly, we combine the values and principles from these instruments into a single interpretative framework (see Appendix 1) for analysing EU laws relevant to AI, with particular attention to the distribution of ethical values across binding and non-binding provisions, in order to assess the normative weight accorded to ethics within the EU's AI regulatory framework (Floridi 2018).

5. Legal Interpretive Analysis of AI-Relevant Laws and Regulations of the EU

Our empirical work followed a two-step approach. First, we identified a set of key legal instruments that directly or indirectly shape the regulatory landscape for AI within the European Union. These instruments were selected because they represent the EU's most comprehensive and strategic efforts to govern different dimensions of AI development, deployment, and use. While the AI Act constitutes the central legislative instrument specifically dedicated to AI, the remaining texts reflect the broader regulatory ecosystem within which AI operates. These include foundational legislation on data protection (GDPR), platform governance (DSA), data access and sharing (DGA[8]; Data Act[9]), and intellectual property rights (DSM Directive[10]) (see **Table 1** for an overview of the most relevant EU legal acts with direct or indirect relevance to AI).

[8] Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act) (Text with EEA relevance) PE/85/2021/REV/1. *OJ L 152, 3.6.2022, p. 1–44*.

[9] Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on harmonised rules on fair access to and use of data and amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 (Data Act) (Text with EEA relevance) PE/49/2023/REV/1. *OJ L, 2023/2854, 22.12.2023*.

[10] Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC (Text with EEA relevance.) PE/51/2019/REV/1. *OJ L 130, 17.5.2019, p. 92–125*.

Policy	EU Act/Regulation	Summary
Artificial intelligence policy	Artificial Intelligence Act (EU AI Act) (2024)	The AI Act uses a risk-based approach through four risk categories, prohibiting the use of "unacceptable risk" AI systems and setting more strict requirements for "high-risk" AI systems. It lays out the governance of AI at the EU and the MS level and provides for regulatory experimentation and flexibility in the implementation.
Regulation on data storage and privacy	General Data Protection Regulation (GDPR) (2016)	GDPR and AI Act are complementary in ensuring lawful, fair, and transparent processing of personal data in AI systems.
Regulation on digital platforms & services	Digital Services Act (DSA) (2022)	Introduction of the risk-based approach to regulation. Specific obligations on very large online platforms (VLOPs) and very large online search engines (VLOSEs), which are also key actors in AI deployment.
Regulation on data sharing	Data Governance Act (DGA) (2022)	The Data Governance Act creates policy for data sharing in sectors such as health, environmental, mobility, agricultural, and public administration. The objective is to improve data availability, promote the reuse of data, ensure regulated data intermediaries, and help data sharing across sectors and borders.
Regulation on data ownership, storage and transactions	Data Act (2023)	The Data Act, complementing the Data Governance Act, is a horizontal set of rules on data access designed to help users of connected devices to gain access to data from those devices, create protections for consumers from certain data sharing contracts, and enable users to switch cloud providers to access private sector data when necessary.
Copyright in the Digital Single Market	Directive on copyright and related rights in the Digital Single Market (DSM 2019)	The DSM Directive is relevant to IP protection for AI technology, regulation of information and data used as inputs for AI, specifically about the text and data mining exceptions under the DSM-IP protectability of AI's output, Digital Rights Management (DRM) and IP enforcement through AI.

Table 1: Overview of the most relevant EU legal acts with direct or indirect relevance to AI.

Analysing this corpus of legislation enables us to trace how ethics-based values and principles, as identified in the three AI policy guidance documents, are translated into binding and non-binding legal norms across different domains of EU law relevant to

AI governance. This analysis was guided by three central research questions: (1) which ethical values and principles are embedded in EU AI-relevant legislation; (2) which of these values are prioritised or marginalised; and (3) to what extent they are legally enforceable.

In the second step, we draw on the results concerning the most prevalent ethical values to conduct an in-depth legal doctrinal and ethics-informed analysis, focusing on the alignment – or lack thereof – between ethical principles and their legal operationalisation. Finally, we reflect on the EU's ethics-informed regulatory ambitions in light of our empirical findings.

The empirical analysis in this study is grounded in a broadly applied legal doctrinal methodology (Hutchinson and Duncan 2012) involving a systematic and interpretive engagement with legal texts in order to identify, explain, and reconstruct the normative logic underpinning legal rules and principles. As applied here, this approach enables a close reading of how legal instruments express and embed ethical values and is particularly well suited to examining the internal coherence, justification, and normative content of EU AI-relevant legislation. Our interpretative reading adopts an internal perspective on law, treating legal norms as part of a normative order shaped by shared meanings and legal reasoning, entering into a dialogue with the legal texts to reconstruct the reasoning structures (Byrne and Olsen 2024).

We applied such legal interpretation to the selected set of EU legal instruments that either directly regulate AI (such as the AI Act) or indirectly shape the governance environment in which AI systems are developed and deployed (including the GDPR, DSA, DGA, Data Act, and DSM Directive). These instruments were analysed for references to the predefined set of eleven ethical values and principles derived from the three international guidelines (Appendix 1). Particular attention was paid to the legal placement of ethical values, specifically whether they appear in binding provisions (such as Articles and Annexes) or in non-binding elements (such as Recitals). This distinction is central to assessing the legal enforceability of ethical values, as binding provisions impose enforceable obligations, while Recitals primarily guide interpretation and clarify legislative intent.

The EU AI Act

Adopted in 2024, the AI Act establishes harmonised rules for the placing on the market, use, and governance of artificial intelligence systems within the European Union. Its cornerstone is a risk-based regulatory concept, which prohibits certain "unacceptable risk" AI practices (such as social scoring and manipulative techniques) and imposes stringent obligations on "high-risk" AI systems, including requirements for

conformity assessments, human oversight, and post-market monitoring. In addition, the AI Act incorporates innovation-supportive instruments, such as regulatory sandboxes, and establishes a multi-level governance framework involving both EU-level and national authorities.

General Data Protection Regulation (GDPR)

Enacted in 2018, the GDPR establishes data protection as a fundamental right, emphasising lawful and transparent processing and individual control over personal data. In the AI context, developers and deployers may qualify as data controllers or processors and must comply with GDPR obligations throughout the lifecycle of personal data processing. The GDPR therefore operates as a horizontal safeguard in AI governance, complementing the AI Act's sector-specific requirements. As a technology-neutral framework, its relevance evolves alongside technological developments, including AI (Nemitz 2018).

Digital Services Act (DSA)

The DSA (2022) establishes the EU's framework for regulating digital platforms, imposing due diligence obligations on online intermediaries, with stricter requirements for Very Large Online Platforms and Search Engines. It introduces systemic risk assessments, crisis response mechanisms, and enhanced transparency obligations for content moderation and recommender systems. The DSA and the AI Act share policy objectives relating to public health, public security, and fundamental rights. However, concerns have been raised about potential tensions – particularly regarding the dissemination of illegal content and civic or electoral discourse – and about the risk that parallel application of the two regimes may unduly restrict freedom of expression if not carefully implemented (Calvet-Bademunt and Barata Mir 2024).

Data Governance Act (DGA)

The DGA (2022) aims to foster trust in voluntary data sharing across the EU by facilitating the reuse of protected public-sector data and establishing safeguards for its responsible use. It introduces a regulatory framework for data intermediation services and data altruism mechanisms, designed to enhance transparency, neutrality, and accountability in data sharing practices. By promoting data availability, public interest objectives and fundamental rights, the DGA supports the development and deployment of data-driven technologies, including AI, within a regulated governance framework.

Data Act

Adopted in 2023, the Data Act complements the DGA by establishing horizontal rules for fair access to and sharing of data generated by connected devices and related services. It seeks to address contractual imbalances, promote interoperability, and facilitate switching between cloud service providers. The Data Act primarily advances 'fairness' through more equitable value distribution among data economy actors. It also supports solidarity by enabling public sector access to private-sector data in situations of exceptional need, and sustainability through provisions that facilitate repair, maintenance, and reuse markets.

DSM Directive

The DSM Directive (2019) modernises EU copyright law for the digital environment, introducing mandatory exceptions for text and data mining, new rights for press publishers, and obligations for online content-sharing platforms to license or remove infringing material. The text and data mining provisions are particularly relevant for AI developers and rightsholders alike. However, the Directive predates the rapid rise of generative AI and the resulting intensification of copyright conflicts. Its balancing mechanisms, therefore, face pressure to accommodate AI development without undermining innovation or rights protection, raising concerns about interpretive uncertainty and regulatory gaps (Guadamuz 2024; Peukert 2024).

6. Ethical Values and Principles in the EU Laws

Drawing on the eleven ethical values and principles defined as analytical categories (Appendix 1), we interpreted the selected EU legal instruments. For each principle, key terms, indicative expressions, and definitional markers were extracted from relevant policy guidelines and translated into the formal legal language of EU law. For example, *fairness* encompassed references to bias mitigation, equality of treatment, and equitable data governance; *explicability* included traceability, auditability, and duties to inform affected persons; and *non-maleficence* was linked to harm prevention and risk-based safeguards. Coding focused on the proximity between ethical categories and their legal expression, allowing assessment of substantive presence, normative priority, and enforceability. This approach moved beyond keyword searches towards a context-sensitive interpretive analysis. Results are summarised in Appendix 2.

Embedded Ethical Values and Principles

Across the AI Act, GDPR, DSA, DGA, Data Act, and DSM Directive, four ethical principles dominate: non-maleficence, fairness, privacy, and explicability. Non-maleficence –

understood as harm prevention and risk avoidance – is most prominently embedded in the AI Act and the DSA. The AI Act operationalises this principle through prohibitions of certain AI practices and stringent safeguards for high-risk systems, while the DSA adopts a harm-prevention logic via systemic risk assessments addressing threats to public health, safety, and fundamental rights.

Fairness appears in two main forms. In AI governance, it is associated with non-discrimination and procedural equity; in data governance, it is reframed as market fairness, focusing on equitable value distribution and access to data. This latter interpretation is central to the DGA and Data Act, while the DSM Directive adds an emphasis on fair remuneration for rightsholders.

Privacy is the foundational value of the GDPR, providing a comprehensive rights-based framework for personal data protection. In other instruments, privacy functions more as a supporting safeguard, often framed as a prerequisite for trust in AI or data sharing. Meanwhile, conceptual debates persist over whether privacy protection should prioritise private life or informational self-determination, understood as individuals' control over data disclosure and processing (Flayyih et al. 2025; Lundgren 2020).

Explicability is present across the AI Act, DSA, GDPR, and – more marginally – data governance legislation. It is operationalised through transparency and traceability requirements, including documentation and disclosure obligations. While explicability has broad cross-instrument presence, it often plays a supporting role, enabling accountability or risk management rather than functioning as an independent policy driver. Its frequent placement in non-binding sections limits enforceability.

Other values – accountability, democracy, sustainability, and solidarity – are less systematically embedded. Accountability is reinforced through auditing and oversight mechanisms; democracy appears mainly in preambular references to civic discourse and rights protection; sustainability is addressed sporadically; and solidarity emerges in niche provisions such as data altruism and public-interest data access.

Prioritisation and Enforceability of Values

Non-maleficence and fairness emerge as the most prioritised values, benefiting from detailed operative provisions and risk-based regulatory logic. Privacy remains a structural right, though often framed instrumentally outside the GDPR. Explicability and accountability are moderately enforceable, while sustainability, solidarity, and democracy are largely confined to non-binding provisions. This reliance on aspirational language risks creating a gap between rhetorical commitment and enforceable norms. However, the limited explicit codification of values such as democracy and solidarity

may reflect assumptions – similar to those identified by Rendtorff (2015) in bioethics – that such values are already embedded elsewhere within the legal order.

7. Discussion of the Ethical and Legal Conceptualisation of the Core Principles

Non-Maleficence

The legal doctrine of non-maleficence, rooted in the ethical principle "first, do no harm" (primum non nocere), has a prominent place in medical ethics (Frischhut and Werner-Felmayer 2020). It refers to the obligation not to inflict harm and to take reasonable steps to prevent foreseeable harm. This principle is often considered a prima facie obligation that can be weighed against other duties and principles, such as patient autonomy around complex end-of-life decisions, as legal cases adjudicated by the European Court of Human Rights (ECtHR) have reaffirmed.[11] In the case of Artificial Intelligence, specifically, it is a principle embedded in UNESCO's Recommendation on the Ethics of AI, the first multilateral agreement for the governance of this technology. Meanwhile, non-maleficence in the surveyed EU laws is implicitly central yet normatively different: while ethical frameworks construe it as a positive duty to prevent and minimise harm, the AI-related legal instruments – especially the DSA, the GDPR and the AI Act (Kusche 2024) – operationalise it mainly through *risk* and *fundamental rights* proxies rather than an explicit "do no harm" norm (see Appendix 2 – Non-maleficence).

The AI Act institutionalises a risk-based regulatory concept built around high-risk uses, conformity assessment and post-market monitoring, rather than a substantive prohibition of harmful AI except in narrowly defined scenarios (such as those defined by Article 5 AI Act as manipulative or exploitative, which by their inherent nature violate fundamental rights and EU values). The AI Act also constructs a risk management framework for the high-risk systems, "understood as a continuous iterative process planned and run throughout the entire lifecycle of a high-risk AI system, requiring regular systematic review and updating" (Article 9 AI Act). The GDPR refers to harm regarding the "risk to the rights and freedoms of natural persons", including physical, material and non-material harms such as discrimination and reputational damage (Recital 75 GDPR). The non-maleficence principle was also referred upon by national courts, as illustrated by the Dutch SyRI litigation[12], while applying human rights law to assessing the function of algorithmic systems. The case cast doubts, whether interna-

[11] CASE OF LAMBERT AND OTHERS v. FRANCE, Judgement of *25 June 2015; Application no. 46043/14*.

[12] District Court of the Hague, 6 March 2020, ECLI:NL:RBDHA:2020:865, available in English at: uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RBDHA: oncastto2020:1878 (accessed on 1.1.2026).

tional human rights law and data protection were well-suited to address digital harms (Rachovitsa and Johann 2022). The DSA extends the risk-management idiom to platforms, requiring VLOPSEs to "diligently identify, analyse and assess any systemic risks in the Union stemming from the design or functioning of their service and its related systems, including algorithmic systems, or from the use made of their services" (Article 34 DSA) and mitigate them in a "reasonable, proportionate and effective" manner (Article 35 DSA), yet leaving "harm" contestable and governance heavily proceduralised. The interpretation of non-maleficence as harm "management" rather than a strict obligation of harm avoidance was specifically apparent in the first reports on the implementation of the DSA published recently.[13]

We highlight the ethical-legal incoherences in this instance. Non-maleficence in an ethical sense would demand symmetrical protection against comparable harms regardless of institutional context and would treat structural, cumulative harms (e.g. surveillance-driven chilling effects or discriminatory profiling) as intrinsically undesirable and up to the person potentially affected by the harm to be able to assess what kind and to what level harm they are willing to tolerate. EU digital legislation and case law, frequently instrumentalise harm within balancing exercises (e.g. against innovation, trade secrets or migration control), resulting in a hierarchy of acceptable harms and exposes the conceptual gap between non-maleficence as an ethical principle and damage limitation as a legal technique. A telling example of this tension is the regulation of the use of AI for intrusive biometric and emotion-recognition systems, which are banned in the workplace and education (Article 5(1)(f) AI Act) but tolerated in migration and law-enforcement contexts (Annexe III: High-Risk AI Systems Referred to in Article 6(2) AI Act). This has been harshly criticised by civil society actors during the legislative process (James 2024), for example, for effectively normalising disproportionate harms against already marginalised groups. We can argue that this inconsistency is recurrent and systemic: while the ethical interpretation of non-maleficence emphasises avoidance of unjustified or unnecessary harms, the legal approach to digital and AI-related harms focuses on procedural harm management. However, given this contradiction, here it is important to notice that this approach is nonetheless consistent with UNESCO's interpretation of non-maleficence in its most literal sense, especially

[13] See the "First report of the European Board for Digital Services in cooperation with the Commission pursuant to Article 35(2) DSA on the most prominent and recurrent systemic risks as well as mitigation measures" on 18 November 2025; available at: Press statement of the European Board for Digital Services following its 16th meeting | Shaping Europe's digital future (accessed on 1.1.2026).

given the instrument's vague definitions of what "legitimate aims and objectives" and "appropriate context" shall mean.[14]

Fairness

Fairness is a broad and complex principle, running through all aspects of life; it implies some level of impartiality existing in the context (Murphy 2011). As a plural ethical concept, fairness spans distributive justice (who gets benefits/burdens), procedural justice (how decisions are made), and interactional justice (how people are treated and informed). In AI ethics, it is often operationalised through the avoidance of unjustified disparate impacts and the design of decision procedures that are contestable and intelligible – yet "fairness" metrics in machine learning (e.g., statistical parity, equalised odds) can be mutually incompatible, making fairness an essentially *normative* and context-sensitive choice rather than a purely technical property (Barocas et al. 2023). Fairness in an ethical and in a technical sense is often at odds. Ethical accounts treat fairness as a requirement to avoid unjustified differential treatment and to remedy structural inequalities, including group-based disadvantages reproduced by data-driven systems (Panarese et al. 2025). Technical AI-fairness literature reframes this as the pursuit of mathematically specified parity criteria at the model or dataset level, which can diverge sharply from legal understandings of equality and non-discrimination (Deck et al. 2024). Fairness in the analysed AI-related regulations operates at the intersection of (constitutional) equality, data-protection "fairness", anti-discrimination law and economic "fairness" (see Appendix 2 – Fairness), yet it remains conceptually and doctrinally fractured, which complicates any coherent translation of ethical fairness into positive law (Grozdanovski 2025). In the specific laws, these conflicts are even more apparent.

The AI Act embeds fairness mainly through ex ante design-level requirements for high-risk systems, particularly data-governance obligations to detect and mitigate bias in training, validation and testing datasets and through references to equality and non-discrimination as protected fundamental rights (i.e. Article 10 (2 f and g) AI Act). However, it largely fails to address deep ethical concerns related to the impact of AI on distributed justice. Furthermore, following the OECD's and the High Level Expert Group on AI's (AI HLEG) approach (OECD, 2021, 6–7; High-Level Expert Group on Artificial Intelligence, 2019), the ethics-informed principles incorporated in the AI Act embed fairness in the concept of trustworthy AI, attached to equal access, gender equality and

[14] See "UNESCO Recommendation on the Ethics of Artificial Intelligence", p. 20.

cultural diversity, and the avoidance of unfair biases (Recital 27 AI Act), which adds even more confusion to the legal interpretation with conflicting results.

Under the GDPR, fairness appears as an overarching principle governing data processing (Article 5(1)(a) GDPR requires that personal data be processed "lawfully, fairly and in a transparent manner"), supplemented by specific safeguards for automated decision-making (i.e. Articles 15 (1 h) and 22 (1) GDPR). Meanwhile, scholars argue that the GDPR's fairness principle is under-specified and has been treated as largely *procedural* (compliance with notice, lawful basis, purpose limits), risking a "fairness-by-formality" that does not reliably address power asymmetries, manipulation, or discriminatory outcomes in AI-mediated profiling (Häuselmann and Custers 2024, 3). The corresponding litigation by the Court of Justice of the EU (CJEU) on automated decision-making illustrates the fault-line between ethical and legal fairness: in the *SCHUFA* case[15], the Court treated credit scoring – where third parties rely heavily on a score – as falling within GDPR Article 22's automated decision-making safeguards, thereby strengthening procedural protections around consequential automated judgement, reiterated also by the *D&B* decision[16] recently. None of the cases recalled for deeper, substantive or structurally unequal ethical impacts.

The DSA addresses fairness mostly on the policy level as part of the DSA's strategy for enhancing user autonomy and ensuring a fairer and safer online environment. The regulation of fairness is indirect and could be best explained through the obligations for systemic-risk governance for recommender systems to provide explanation and user choice (Article 27 DSA). Yet, many argue (Hakkarainen and Savolainen 2025) that framing fairness as an architectural risk reduces the ethical claim to a procedural legal technique, to risk assessments, transparency and mitigation, rather than a right to "fair ranking" or "fair feeds".

Finally, the Data Act and Data Governance Act – both of which aim for data fairness as their main policy objective – introduce different fairness concepts, such as the FAIR data principles, data altruism for users' trust-building for data sharing and the fair allocation of value from data among actors in the data economy (see Appendix 2 – Fairness: DGA and Data Act references). Importantly, none of them addresses bias as a matter of fairness. Arguably, fairness is primarily economic/structural (balancing bargaining power, addressing unfair contractual terms), referring to competition law

[15] Judgment of the Court (First Chamber) of 7 December 2023, *SCHUFA Holding (Scoring)*, Case C-634/21 ECLI:EU:C:2023:957.

[16] Judgment of the Court (First Chamber) of 27 February 2025, *CK v Dun & Bradstreet Austria GmbH and Magistrat der Stadt Wien*, Case C-203/22, ECLI:EU:C:2025:117.

to curb "unfairness" of data markets and related "unfair practices" (Kianzad 2025), rather than e.g. anti-discrimination in AI, potentially pulling the concept in a direction that sits uneasily with ethical fairness as equal respect and non-subordination.

The ethical-legal tensions regarding fairness are apparent. It is questionable to what level EU laws entail robust protection against individual or group-based disadvantage, nor do they address power asymmetries from the ethical justice perspective. The legally inscribed fairness in data protection, platform regulation and data (re) distribution relies heavily on procedural matters of mitigating unfairness, and at best aims at balancing interests in AI innovation.

Privacy

Privacy in the European AI legal context is formally entrenched through the GDPR (see Appendix 2 – Privacy: GDPR) and fundamental rights (Articles 7 and 8 of the Charter of Fundamental Rights of the EU). Meanwhile, the ethical concept of privacy frames it as a condition for autonomy, freedom (Sax 2018), intimacy and democratic participation, not merely as control over personal data. This discrepancy between the ethical and the legal can be well observed in our sample, too.

Under the GDPR, privacy is largely operationalised as data protection (Brkan 2019) through principles of lawfulness, fairness, transparency, and data-centric duties such as purpose limitation, data minimisation, integrity/confidentiality, and "data protection by design and by default" (Article 25 GDPR), complemented by rights against certain forms of automated decision-making (Article 22 GDPR). Case-law analysis (Vale and Zanfir-Fortuna 2022) shows that supervisory authorities and courts rely heavily on Article 5 principles, transparency duties and impact assessments, while Article 22's "right not to be subject" to automated decisions is interpreted variably as a prohibition or a waivable entitlement, revealing doctrinal uncertainty at the core of digital privacy protection – notwithstanding the lack of ethical references of privacy. This produces a predictable inconsistency with ethical privacy regarding many salient AI harms, such as behavioural manipulation, chilling effects, sensitive inferences from "non-sensitive" data, or harms arising from non-personal but re-identifiable datasets.

The AI Act–GDPR interplay did not help to clarify nor to ethically better embed privacy protection (De Luca and Federico 2025). The example of the differing legal interpretations of algorithmic discrimination reveals already uncertainties to be addressed through legislative reform or further guidance. The claim for "trustworthy" AI called for clearer and deeper ethical guidance to solve the conflicts on how to mitigate discrimination and bias through large-scale AI training and monitoring practices with GDPR data minimisation and purpose limitation. The AI Act's risk-based approach to

privacy protection – especially in the areas of migration and public benefits – raised several concerns about the possibility of an ethically ingrained implementation (Fundamental Rights Agency 2025).

The DSA, the Data Act and the Data Governance Act entail ethics-unrelated approaches to privacy. The DSA – other than the online protection of minors (Article 28 DSA) – addresses privacy only indirectly and through systemic-risk assessments for very large platforms (Article 34 DSA), encompassing risks to fundamental rights (including privacy), but without any specific, ethics-informed obligations. The Data Act and Data Governance Act are rather concerned with unlocking industrial and public-sector data, and their approach to privacy is largely one of coordination (see Appendix 2 – Privacy: DGA and Data Governance Act).

The inconsistencies between the ethical and the legal understanding of privacy are alarming regarding AI. While ethically privacy would require meaningful limits, especially on surveillance infrastructures or strong restraints on high-risk inferences, the legal means treat privacy primarily as compliant data processing and as a parameter in risk management for AI, tolerating even extensive data extraction where formal GDPR conditions are met. These conflicts entrench a procedural, individualised, and often innovation-accommodating understanding of privacy in the EU laws that sits uneasily with richer ethical accounts of privacy inevitable to AI.

Explicability

Explicability has been proposed as one of the core principles for ethical AI, entangling two sides of it in "the epistemological sense of *intelligibility* (as an answer to the question "how does it work?") and in the ethical sense of *accountability* (as an answer to the question: "who is responsible for the way it works?")" (Floridi and Cowls 2019, 2). Hence, the concept of Explainable Artificial Intelligence (XAI) refers to a set of processes and methods that allow human users to comprehend the decisions or predictions made by AI models and systems. Explicability also occupies a central position between ethics and law, functioning both as an enabling condition for other ethical principles (i.e. fairness, accountability, autonomy) and as a normative demand in its own right, also instrumental to the EU's concept on "trustworthy AI" (High-Level Expert Group on AI 2019). Despite the central position in AI policy, the legal provisions related to explicability in the surveyed laws were scattered and non-comprehensive, fragmenting explicability across procedural duties, rather than recognising a freestanding "right to explanation" (see Appendix 2 – Explicability).

The AI Act embeds explicability connected to transparency and, to some extent, accountability. There are distinct provisions on obligations on transparency (Article 50

AI Act in general) and information for high-risk AI systems (i.e. Article 13 and Article 11 (with Annexe IV) AI Act) to allow deployers (and indirectly individual users) to interpret system outputs and understand limitations, which embeds explainability in documentation and transparency duties. In parallel, human oversight requirements in case of high-risk AI (i.e. Article 14 AI Act) are to enable human interpretation of outputs, often through built-in explainability features. The stipulation of the right to explanation of individual decisions (Article 86 AI Act) under the remedies regarding high-risk AI is perhaps the closest to the ethical demands for individuals to be entitled to clear and meaningful explanations, in case their health, safety or fundamental rights are involved, of how the AI system contributed to the decision. While the ethical-legal connections in the case of the AI Act are plausible, it is also apparent that legal conceptualisation of explicability is limited, risk-calibrated and audience-specific (authorities and deployers first), and framed around regulatory compliance rather than contestability or justice for individuals (Veale and Borgesius 2021).

Regarding EU data protection law and regulation, explicability is implicit in duties to provide "concise, transparent, intelligible and easily accessible" information about processing and profiling, as well as "meaningful information about the logic involved" in automated decision-making (European Data Protection Supervisor 2023), 9). Articles 13–15 GDPR require the provision of "meaningful information about the logic involved" regarding personal data collection to the data subject in automated processing (Article 13 4 (f) GDPR), while Article 22 GDPR establishes safeguards against certain automated decisions with legal or similarly significant effects, connecting explicability both to transparency and accountability. Notably, these requirements focus on the procedural safeguards of explainability rather than the individual's contextualised needs for understanding AI.

The DSA approaches explicability through platform architecture and recommender systems, including the obligation for VLOPSEs to "explain the design, the logic, the functioning and the testing of their algorithmic systems, including their recommender system" to the regulators (Article 40 (3) DSA). Similarly, systemic transparency reporting obligations (i.e. Articles 15, 24, 27, 39 and 42 DSA) aim to ensure the flow of information about the most crucial elements of platforms' operation, assuming to enable explainability. Explainability is the DSA serves mostly regulatory purposes, providing for systemic reporting obligations without additional contextualisation, left largely to researchers seeking further information via Article 40 DSA access opportunities for vetted inquiries.

Overall, there are significant divergences between the ethical and legal accounts of the notion of explicability. Ethical notions of explicability demand genuinely under-

standable, context-appropriate accounts of how AI systems shape opportunities, risks and distributions of advantage, especially for structurally vulnerable groups. EU digital legislation instead operationalises explicability through transparency, documentation and reporting requirements adjusted to risk categories, tailored to regulatory oversight rather than human control, producing a procedural, compliance-oriented conception that only partially supports the ethical ideal of explanation as a vehicle for individual empowerment, contestation and human agency. This divergence risks incoherence: systems may satisfy formal transparency requirements while remaining effectively unintelligible to those most affected by their decisions.

8. Conclusion and Looking Ahead

The EU's AI-relevant legislative framework reflects a multi-value ethical architecture, with non-maleficence, fairness and privacy forming the enforceable core, and explicability playing a significant supporting role. Sustainability, solidarity and democracy remain underdeveloped in binding terms, largely confined to aspirational statements. The reliance on non-binding language for key ethical commitments risks leaving them under-realised, particularly as generative AI and other emerging technologies expose new normative challenges. We are not naïve: our finding goes in line with the political agreements at the international level on this matter, where non-binding recommendations have taken the forefront in international AI regulation – at the cost of risking a high level of protection of several fundamental rights.

Our aim was to draw attention to values that underpin European democracies and to ask whether legislation on AI reflects these values adequately, and whether the EU's proclaimed policy of trustworthy and human-centric AI is truly embedded in law. The findings reveal shortcomings: values tied to democratic processes are treated more as an "afterthought" than a priority, while human-centricity lacks binding and enforceable guarantees of meaningful human oversight. At the same time, the vagueness of fairness across instruments and its differing interpretations highlight the instability of this principle as a regulatory foundation, also due to its multi-faceted and highly context-dependent nature.

A limitation of our paper is that it does not explore in depth the political and legislative dynamics behind these outcomes, nor the division of responsibilities between EU institutions and member states in governing AI. Due to the complexity of a study of that nature, It rather takes a pragmatic stance in the geopolitical realities into which these instruments were conceived: a global environment where few global players concentrate AI development, and an incipient yet growing global regulatory response

that is rich and consistent in terms of principles and standardised practices but vague and reliant on non-binding instruments to ensure their functioning. Nevertheless, the analysis demonstrates how this global landscape results in ethics-based commitments that are then selectively translated into law and identifies the gaps in leaving central democratic and societal values outside the enforceable core. Future research should investigate not only the effectiveness of ethics-based regulation, but also how ethical values can be more robustly enacted in legal frameworks to support both technological development and the legitimacy of AI governance in Europe.

Appendix 1: AI Values and Principles

Values-principles/ documents	EU ethics guidelines for trustworthy AI (2019)	OECD AI principles (2019)	UNESCO AI recommendations (2021)
Human rights and dignity	Respect for human autonomy	Human rights and democratic values, including fairness and privacy	Human rights and human dignity. Respect, protection and promotion of human rights and fundamental freedoms and human dignity
Non-maleficence	Prevention of harm		Proportionality and Do no Harm
Fairness	Fairness	Human rights and democratic values, including fairness and privacy	Fairness and non-discrimination
Explicability	Explicability	Transparency and explainability?	Transparency and explainability
	Human agency and oversight		Human Oversight and determination
Robustness and safety	Technical robustness and safety	Robustness, security and saftety	Safety and Security
Privacy	Privacy and governance	Human rights and democratic values, including fairness and privacy	Right to privacy and data protection
Transparency	Transparency	Transparency and explainability	Transparency and explainability
Diversity and non-discrimination	Diverity, non-discrimination and fairness		Ensuring diversity and inclusiveness; Fairness and non-discrimination
Sustainability	Societal and environmental well-being	Inclusive gorwth, sustainable development and well-being	Sustainability; Living in peaceful, just and interconnected societies/Environment and ecosystem flourishing
Accountability	Accountability	Accountability	Responsibility and accountability

Appendix 2: Ethical values and principles in the EU AI-relevant legislative instruments

Values-principles/EU Laws	AI Act		GDPR		DSA		DGA		Data Act		DSM	
	Recitals	Articles	Recitals	Articles	Recitals	Articles	Recitals	Articles	Recitals	Articles	Recitals	Articles
Human rights and human dignity (fundamental rights)	1; 3; 5; 6; 8; 9; 10; 17; 20; 22; 27; 28; 31; 32; 34; 43; 46, 48; 52; 53; 57; 58; 59; 60; 65; 66; 67; 70; 72; 75; 77; 80, 91; 92; 93; 96; 118; 121; 139; 140; 155; 157; 171; 176	1; 2/4; 3/49a; 5/2b; 6/3; 6/6; 6/7; 6/8; 7/1b; 7/2e,i; 7/3a,b; 9/2a; 10/2f; 0/5; 13/3b (iii); 14/2;27/1,2,4; 28/7; 36/7e; 36/8a; 36/9c; 40/3; 41/a(iii); 43/6; 57/6; 57/11; 58/2i; 58/4; 66/h; 70/3; 77/1,3; 79/1,2> 82/1; 86/1; 112/10; Annex IV/3; Annex VIII/C/4	1; 2; 3; 4; 10; 16; 47; 51; 52; 53; 69; 73; 102; 104; 109; 111; 113; 114; 153; 166; 173	1/2; 4/24; 6/1f; 9/2b, g, j; 23/1; 45/2a; 50/b; 51/1; 88/2	3; 9; 22; 36; 39; 40; 41; 47; 51; 52; 54; 63; 79; 81; 86; 109; 116; 140; 153; 155	1/1; 14/4; 34/1b; 35/1, 3; 36/1, 3a, 8c; 48/4e, 5	1; 2; 3; 6; 19; 22; 46; 62	No binding rules	7; 8; 56; 80; 101	No binding rules	70; 84; 85	17/9, 10
Non-maleficence (no harm, harmless, malicious, risk avoidance)	1; 5; 20; 26; 27; 28; 29; 32; 33; 46; 48; 52; 53; 75; 76; 96; 110; 115 + risk avoidance – Risk-based approach – 342 non-binding reference – Main aim of the AI Act	1; 3/2; 3/4 a, d; 3/56; 3/61 a,b; 5/1 a,b; 5/2 ; 6/3; 6/6; 7/1b; 7/2 d,e,f,g,h; 27/1d 99/7j + risk avoidance – Risk-based approach – 434 binding reference – Main aim of the AI Act	9; 15; 28; 38; 39; 49; 51; 65; 71; 74; 75; 76; 77; 80; 81; 83; 84; 85; 89; 91; 94; 96; 98; 116; 122; 144	4/24; 23/2g; 24/1; 25/1; 27/2a; 30/5; 32/1, 2; 33/1; 34/1, 3b, 4; 35/1, 7c, d, 11; 36/1, 2; 39/2; 49/1a; 57/1b; 70/1h	63; 69; 79; 96; 114; 116; 137; 140 + risk avoidance – Risk-based approach – 73 non-binding reference – Main aim of the DSA	42/5; 52/2e, 3b; 82/1 + risk avoidance – Risk-based approach – 40 binding reference – Main aim of the DSA	15; 24	5/13; 30/d	No mention of harms at all! Risks: 8; 18; 31; 42; 112	Harm: 11/2b; Risks: 2/29; 17/2g; 25/2a(iii), f	16; 17; 24; 60;	No binding rules

(Contd.)

Values-principles/EU Laws	AI Act		GDPR		DSA		DGA		Data Act		DSM	
	Recitals	**Articles**	**Recitals**	**Articles**	**Recitals**	**Articles**	**Recitals**	**Articles**	**Recitals**	**Articles**	**Recitals**	**Articles**
Fairness (fair(ness), bias, equal(ity), altruism)	27; 28; 31; 32; 48; 54; 59; 61; 67; 70; 74; 75; 80; 94; 96; 110; 142; 154; 156; 157; 173	7/2k; 10/2 f,g; 10/5a, e, f,; 14/4b; 15/4; 58/2b; 70/1; 95/2e; Annex XI/1/2c	4; 39; 42; 45; 60; 71; 129; 155	5/1a; 6/2; 6/3b; 13/2; 14/2; 40/2a;	58; 59; 67; 68; 94; 95; 134; 152; 153	21/3f; 46/2; 47/1	2; 20 Key policy objective is about FAIR data principles and data altruism – 51 non-binding references. No mention of bias though!	12/f; 26/2; 30/h(ii); Key policy objective is about FAIR data principles and data altruism – **139 binding references.** No mention of bias though!	5; 6; 19; 26; 28; 38; 40; 42; 44; 52; 56; 58; 59; 60; 61; **62;** 111; 119 Main policy objective: fairness in the allocation of value from data among actors in the data economy: 40 non-binding references to (un)fairness.	8/1, 2; 10/1, 5a,b; **13/1, 2, 3, 4, 5, 7**; 20/2; 30/2; 41; 49/1a (altruism!), g Main policy objective: fairness in the allocation of value from data among actors in the data economy: 20 binding references to (un)fairness.	3; 6; 24; 48; 61	5/4; 8/1b; 12/3b; **16**; **18**; 20/1 Main policy objective: fair balance of rights and interests.

(Contd.)

Values-principles/EU Laws	AI Act		GDPR		DSA		DGA		Data Act		DSM	
	Recitals	Articles	Recitals	Articles	Recitals	Articles	Recitals	Articles	Recitals	Articles	Recitals	Articles
Explicability (trust, explainability, traceability, auditability, communicate, inform, consult)	3; 27; 53; 59; 61; 65; 71; 73; 74; 92; 93; 116; 121; 131; 133; 143; 161	3/15; 12/2; 13/3b(iv); 17/1h; 20/1, 2; 22/4; 23/4; 24/2, 4; 26/5, 7, 8, 11; 33/1; 36/3, 4, 5, 7d, 9; 37/4; 40/2; 41/1, 2, 6; 45/1, 2; 46/3; 50/3; 54/5; 57/11, 15; 58/2g, I; 62/1c, 3c; 70/2; 71/1; 73/7; 74/13b; 75/2; 76/5; 77/1; 78/1b; 79/2,3, 7; 80/3; 81/1; 82/5; 90/2; 97/4; 101/2, 4; Annex IV/2c; Annex V/1; Annex VII/4.6, 4.7, 5.3; Annex VIII/A/4, B/4; Annex XI/2/3	7; 30; 71; 66; 86; 87; 107; 113; 127; 141; 158	11/2; 12/3, 4; **13–15**; 17/2; 19; **22**; 28/2, 3a,h; 33/3b; 34/1, 4; 35/4, 5; 36/2; 37/7; 39/1a,b; 41/4; 47/1j; 49; 58/1b; 56/3; 57/1f; 58/2e; 60/3, 7, 8, 9, 10; 61/5; 64/1, 4, 5, 7; 65/1c, 5, 6; 66/1; 68/5; 77/2; 78/2	3; 9; 12; 31; 32; 33; 34; 40; 45; 51; 52; 54; 56; 63; 68; 72; 73; 74; 90; 92; 93; 112; 118; 123; 126; 133; 139; 145; 152; +Trusted flaggers (specific policy instrument):17 non-binding reference. +Specific independent auditing requirements: 37 non-binding references.	9/1, 2ii, 5; 10/1, 2iv, 5; 11/1, 2, 3; 12/1; 14/2, 3; 15/1a; 18/1, 2; 20/5; 21/4c; 22/3, 4, 6, 8; 24/3; 27/2; 28/4; **30;** 32/1; 33/3, 4; 34/3; 36/3, 4c, 7; 37/4f; 39/3; 40/3, 6, 8, 10, 11; 41/4, 13; 42/4e; 44/1; 47/2b; 49/3; 55/1, 2b; 57/1, 3; 58/5; 59/2, 3; 60/2, 3, 4; 62/6; 68/2; 69/4, 6; 73/2; 74/3; 75/23; 87/4; 91/7; +Trusted flaggers (specific policy instrument): 16 binding reference +Specific independent auditing requirements: 61 binding reference)	3; 5; 15; 23; 24; 31; 32; 33; 38; 43; 45; 46; 47; 52; 54; 58; 61	No binding rules on trustworthiness! 2/15; 5/5, 9; 11/9; 12/k; 17/2; 21/1, 5; 22/1c; 24/5a; 25/1; 27/2; 31/5; 32/4	2; 14; 15; 17; 20; 21, 22; **24**; 25; 34; 36; 45; 52; 64; 66; 69; 71; 73; 76; 80; 82; 85; 92; **95**; 98; 100; 101; 102; 107; 108; 109; 113; 118	2/1, 7; 3/2, 3e; 4/5; 5/4; 8/3 5; 9/7; 10/10, 11; 11/2c, 3; 16/1; 17/1c, 2f; 19/1c; 20/2; 25/2a(iii), b; **26**; 28/1; 29/4, 5, 6; 30/1,2; 31/3; 32/5; 33/6, 7, 10; 35/7; 36/1c; 38; 49/1a	15; 37; 41; 48; 50; 71 Main objective: countering informational unbalances between users and rights-holder's: 46 non-binding reference to information obligations.	11; 12/3d, 5; 17/9, 10 Main objective: countering informational unbalances between users and rights-holder's: 29 binding reference to information obligations.

(Contd.)

Values-principles/EU Laws	AI Act		GDPR		DSA		DGA		Data Act		DSM	
	Recitals	**Articles**	**Recitals**	**Articles**	**Recitals**	**Articles**	**Recitals**	**Articles**	**Recitals**	**Articles**	**Recitals**	**Articles**
Human oversight (autonomy, self-determination, agency, human (individuals') intervention, human control, consent, objection, authorisation)	12; 27; 29; 42; 61; 66; 72; 73; 91; 96; 110; 111; 134 +individuals' consent – 7x non-binding	3/1; 6/3 b, c; 7/2 d; 13/3 d; 14/3; 26/2, 3; 27/1 e; 50/4; Annex 2 e, 3; Annex XIII/e +individuals' consent – 9x binding	71; 87; 124; Individuals' consent: 42 non-binding reference – Main aim of the GDPR: personal data protection	4/24; 22/3; 25/2; 60/4, 5; 65/1a; Individuals' consent: 31 binding reference – Main aim of the GDPR: personal data protection	40; 42; 43; 44; 45; 58; 64; 67; 68; 103; 109; 111; 114; 124; 136; 141	14/1; 42/2a; 50/1; 51/1c; 68/1	5; 22; 26; 30; 32; 33; 46 + Control and the provision of consent are among the main policy objectives: 21 non-binding references to consent.	31/3b + Control and the provision of consent are the main policy objectives: 27 binding references to consent.	34; 36; 38; 40; 80; 96; 104; 107	4/4; 6/2a; 10/12; 37/9	Copyright main policy: granting exclusive control rights to the authors of copyrighted works: 28 non-binding references to authorisation.	Copyright main policy: granting exclusive control rights to the authors of copyrighted works: 13 binding references to authorisation.
Robustness and safety (technical resilience, security, fallback/ fail-safe plans, accuracy, reliability, reproducibility)	27; 44; 47; 58; 59; 60; 66; 74; 75; 76; 77; 94; 103; 110; 115; 122; 133; 138; 166	5/57; 10/5b; 13/3b(ii); 14/4e; 15; 50/2; 58/2(i); 59/1a (iv); 74/12; 78/2; Annex III 7/b; Annex IV 2g, 3; Annex VII/4.3	39; 42; 49; 50; 56; 62; 71; 75; 78; 81; 83; 90; 94; 102; 107; 108; 109; 110; 122; 156	4/4, 12; 5/1d,e,f; 6/4e; 18/1a; 22/3, 4; 23/2f; 25/1; **32**; 35/7d,9; 40/2h; 45/2a; 47/2d	3; 9; 12; 26; 28; 29; 40; 52; 61; 62; 71; 72; 73; 74; 79; 96; 104; 108; 109; 148; 155	1/1; 15/1e; 28/1; 39/1; 42/2c; 47/2a; **Chapter III;** 85/1	2; 3; 7; 8; 15; 19; 20; 21; 23; 46; + Data safety and security are among the main policy objectives: 16 non-binding references to data security.	5/5, 9; 20/1; 30/h(iii); + Data safety and security are among the main policy objectives: 26 binding references to data security.	20; 24; 30; 31; 39; 42; 43; 64; 74; 81; 82; 83; 92; 94; 97; 100; 102; 115	1/1e; 2/29, 39; 3/1; 4/1, 2, 5; 5/1, 4; 6/2f; 17/1g; 19/1b; 25/2a(iv); 30/6; 35/1d; 36/1a, b	15; 16; 22	3/2, 3; 5/1a, 3

(Contd.)

Values-principles/EU Laws	AI Act		GDPR		DSA		DGA		Data Act		DSM	
	Recitals	Articles	Recitals	Articles	Recitals	Articles	Recitals	Articles	Recitals	Articles	Recitals	Articles
Privacy (data protection, quality and integrity of data, data access, data protocols)	9; 27; 28; 43; 54; 57; 67; 68; 69; 80; 110; 141	10/5b; 26/9; 27/4; 59/1e; 70/3; Annex VII/C/5	Data protection – Main aim of the GDPR: 157 non-binding references.	Main aim of the GDPR: 166 binding references.	10; 21; 28; 52; 68; 69; 71; 72; 77; 81; 84; 94; 96; 97; 98; 103; 107; 122; 130; 134; 139; 141; 143; 148	5/1d; 22/3; 24/3, 5; 26/3; 28/1, 2, 3; 34/1b 2e; 37/2; 39/1; 46/2; **40**; 44/1g; 45/1; 46/2; 69/1d, 5	7; 10; 15; 24; 26; 30; 62 Data protection and integrity are among the main policy objectives: 78 non-binding references to data protection.	20/2c, 20; 5/4; 7/4c; Data protection and integrity are among the main policy objectives: 28 binding references to data protection.	2; 3; 6; **7**; 8; 10; 12; 15; 18; 19; **20**; **21**; 22; 31; 36; 57; 72; 82; 100; 101; 102; 107; Data access regulation – main policy objective: 121 non-binding references to data access provisions.	1/5; 2/39; 4/6; 5/9, 13; 6/1; **11/1,2,4**; 17/1g; 19/1b; 21/5; 30/6; 35/1d; 37/3; 49/1a; Data access regulation – main policy objective: 72 binding references to data access provisions.	No references.	No binding rules.
Transparency (identifiability)	9; 26; 27; 52; 53; 59; 60; 65; 66; 67; 72; 73; 74; 94; 95; 96; 101; 102; 104; 107; 114; 131; 132; 133; 134; 135; 137; 157; 173; 174	1/2d; 9/2a, d; 9/5a; 9/5a; 10/2h; 12/2a; 13; 14/5; **50**; 56/2c; 79/6; 82/3; 96/1d; 99/4g; 112/2b, 11c; Annex V/1; Annex VII/4.6; Annex VIII/A4; Annex XI/1.2c, 2.1; **Annex XII**	13; 30; 33; 39; 58; 60; 71; 77; 78; 100	5/1a; 11; **12**; 26/1; 36/2; 40/2a; 41/2c; 42/3; 43/2d; 53/1; 88/2	40; 45; 48; 49; 50; 52; 64; 65; 66; 68; 72; 94; 100; 107; 111; Major regulatory instrument of the DSA: transparency obligation: 18 non-binding reference.	**Chapter III; 15; 24; 27;** 37/2; **39; 42;** 44/1g; 46/1, 2; 50/1 Major regulatory instrument of the DSA: transparency obligation: 12 binding reference.	5; 11; 13; 15; 43; 46; 47; 52	4/3, 5; 5/2, 4; 11/9; 12/f, m; 17/2; **20 (strong transparency requirements)**; 22/1a	5; 21; 25; 29; 34; 42; 61; 69	4/6, 7; 5/9, 10; 8/1; 10/1, 5b; 19/3; 32/3	Main objective: countering informational unbalances between users and rights-holder's: 22 non-binding reference to transparency obligations.	Main objective: countering informational unbalances between users and rights-holder's: 5 binding references to transparency obligations.

(Contd.)

Values-principles/EU Laws	AI Act		GDPR		DSA		DGA		Data Act		DSM	
	Recitals	Articles	Recitals	Articles	Recitals	Articles	Recitals	Articles	Recitals	Articles	Recitals	Articles
Diversity and non-discrimination (inclusion, equal treatment, accessibility, balanced stakeholder participation)	7; 21; 27; 28; 31; 32; 44; 45; 48; 54; 56; 57; 58; 59; 60; 67; 68; 70; 72; 80; 92; 95; 96; 110; 116; 121; 139; 142; 143; 149; 150; 165; 173	10/2f; 10/3; 16/l; 26/4, 7; 40/3; 50/5; 56/3, 5, 6; 58/1a, b, c; 56/2b; 60/4h; 61/1b; 62/1c; 68/2; d; 71/1b, 6; 77/1; 95/2d, e; 95/3; Annex IV/2g, 3	4; 71; 75; 85; 99; 155	50/c; 88/1	3; 26; 29; 47; 52; 58; 69; 81; 94; 95; 103; 104; 107; 108; 134; 137; 152	20/2; 34/1b; 45/2, 3; 47/1, 2; 48/4e; 63/1e	2; 6;10; 11; 15; 16; 24; 26; 27; 38; 58	1/2; 4/2; 5/2, 7/4b; 13; 12/f, m;	Main policy objective: reasonable and non-discriminatory terms and conditions: 10 non-binding references to (non-)discrimination obligations.	Main policy objective: reasonable and non-discriminatory terms and conditions: 10 binding references to (non-)discrimination obligations.	2; 6; 47; 48; 52; 60; 75	8/1b; 12/3b; 18/2
Sustainability (societal and environmental well-being, ecological responsibility, environmental friendliness, social impact, democracy)	1; 2; 4; 6; 8; 27; 28; 31; 48; 55; 61; 62; 130; 142; 155; 165; 176 democracy – 16 non-binding reference	1; 5/1b; 3/49d; 7/2h; 46/1; 59/1a(ii), (iii); 95/2b; 112/7 democracy only 3 binding reference	19; 50; 56; 73; 153 – only 5 non-binding references to democracy	6/4; 23/1; – only 2 binding references to democracy	12; 69; 76; 79; 82; 98; 104; 111; 128; 145 Democracy: only 2 non-binding reference!	No binding obligation	11; 24; 53	No binding obligation	1; 7; 15; 64;	No binding obligation	54; 55	No binding obligation
Accountability (responsibility, obligations, enablement of the assessment, minimisation and reporting of negative impacts, trade-off, redress, evaluation and review, monitoring)	10; 22; 27; 34; 59; 68; 79; 85; 89; 91; 96; 101; 114; 125; 134; 138; 141; 161; 177	7/2k(i); 17/1m; 23/4; 24/3; **25**; 31/3; 33/2; 36/3, 8b, 9b; 37/1; 47/4; 50/4; 60/4h; 66/e(vi); 71/5; 91/5; 99/7g; 100/1b; Annex V/3	74; 79; 80; 82; 85; 104; 108; 146; 148;	5/2; 9/3; **24**; 30/1; 42/4; 43/4; 45/2a, b; 47/2e;62/4; 82/5; 83/2d;	3; 27; 39; 40; 49; 54; 59; 64; 65; 90; 92; 93; 94; 99; 112; 121	9/2v, 5; 10/2v, 5; 14/5; 16/5; 17/3f, 4; 20/5; 32/1c, 2; **34**; 35/2a; 40/4; 41/2, 5; 42/4a; 49/2; 50/3; 51/2; 57/1; 58/5; 59/3; 64/2, 3; 67/4	7; 15; 21	5/12c; 7/4e; 9/2; 26/3, 4; **35**	7; 31; 34; 45; 56; 74; 85; 94; 108	4/3, 9, 12; 24	38; 48; 66; 70; 75; **77**; 79;	3/4; 17/5, 8, 9, 10; **19**; 21

AI Use Statement

The authors used an AI-based language model (ChatGPT, version 5.2) solely for language-related assistance, including improving clarity, coherence, consistency of terminology, and proofreading of the manuscript. The AI tool did not contribute to the development of the research design, the formulation of research questions, the analysis or interpretation of data, or the development of substantive arguments. All intellectual content, analytical decisions, and conclusions presented in the article are the sole responsibility of the authors.

Funding Information

This work was partially supported by DIACOMET ("Fostering capacity building for civic resilience and participation: Dialogic communication ethics and accountability" project, funded from the European Union's Horizon Europe research and innovation grant agreement No 01094816).

Competing Interests

The authors declare that they have no competing interests.

Author Contributions

Krisztina Rozgonyi coordinated the project behind this research. Krisztina Rozgonyi and Mari-Liisa Parder conceptualized the study and designed the research methodology. Krisztina Rozgonyi, Mari-Liisa Parder, and Rodrigo Conde Jiménez contributed to the theoretical framework. Krisztina Rozgonyi carried out the formal analysis and data collection for this article. Krisztina Rozgonyi drafted sections 1, 5, and 6 of the manuscript, and contributed to sections 2, 7, and 8. Mari-Liisa Parder drafted sections 4 and contributed to sections 7 and 8. Rodrigo Conde Jiménez drafted section 3 and contributed to sections 2, 7, and 8. All authors contributed to reviewing and editing the manuscript and approved the final version for publication.

References

Auernhammer, Jan. 2020. 'Human-Centred AI: The Role of Human-Centred Design Research in the Development of AI'. *Paper presented at DRS2020: Synergy*. August 11. https://doi.org/10.21606/drs.2020.282.

Barocas, Solon, Moritz Hardt, and Arvind Narayanan. 2023. *Fairness and Machine Learning: Limitations and Opportunities*. MIT Press. https://mitpress.mit.edu/9780262048613/fairness-and-machine-learning/.

Brkan, Maja. 2019. 'The Essence of the Fundamental Rights to Privacy and Data Protection: Finding the Way Through the Maze of the CJEU's Constitutional Reasoning'. *German Law Journal* 20 (6): 864–83. https://doi.org/10.1017/glj.2019.66.

Broeders, Dennis, Fabio Cristiano, and Monica Kaminska. 2023. 'In Search of Digital Sovereignty and Strategic Autonomy: Normative Power Europe to the Test of Its Geopolitical Ambitions'. *JCMS: Journal of Common Market Studies* 61 (5): 1261–80. https://doi.org/10.1111/jcms.13462.

Brownsword, Roger. 2008. *Rights, Regulation, and the Technological Revolution*. 1. publ. Oxford University Press.

Burg, Wibren van der. 2011. *Law and Ethics. The Twin Disciplines*. https://repub.eur.nl/pub/77446.

Byrne, William Hamilton, and Henrik Palmer Olsen. 2024. 'Doctrinal Legal Science: A Science of Its Own?' *The Canadian Journal of Law and Jurisprudence* (New York, USA) 37 (2): 343–67.

Calvet-Bademunt, Jordi, and Joan Barata Mir. 2024. '*The Digital Services Act Meets the AI Act: Bridging Platform and AI Governance | TechPolicy. Press*'. Tech Policy Press, May 29. https://techpolicy.press/the-digital-services-act-meets-the-ai-act-bridging-platform-and-ai-governance.

Cave, Stephen, Kanta Dihal, and Sarah Dillon. 2020. *AI Narratives: A History of Imaginative Thinking about Intelligent Machines*. 1st ed. Oxford University Press.

Chance, Shannon, Tom Børsen, Diana Adela Martin, Roland Tormey, Thomas Taro Lennerfors, and Gunter Bombaerts, eds. 2025. *The Routledge International Handbook of Engineering Ethics Education*. Routledge. https://doi.org/10.4324/9781003464259.

Christou, George, Trisha Meyer, and Rosanna Fanni. 2025. 'The European Union: Assessing Global Leadership through Actorness in Artificial Intelligence'. *Journal of European Integration* 47 (3): 383–401. https://doi.org/10.1080/07036337.2024.2377200.

Corrêa, Nicholas Kluge, Camila Galvão, James William Santos, et al. 2023. 'Worldwide AI Ethics: A Review of 200 Guidelines and Recommendations for AI Governance'. *Patterns* 4 (10): 100857. https://doi.org/10.1016/j.patter.2023.100857.

Csernatoni, Raluca. 2025. 'The EU's AI Power Play: Between Deregulation and Innovation'. *Carnegie Endowment for International Peace*, May 20. https://carnegieendowment.org/research/2025/05/the-eus-ai-power-play-between-deregulation-and-innovation?lang=en.

De Luca, Stefano, and Marina Federico. 2025. Algorithmic Discrimination under the AI Act and the GDPR. *European Parliamentary Research Service*. https://www.europarl.europa.eu/RegData/etudes/ATAG/2025/769509/EPRS_ATA(2025)769509_EN.pdf.

Deck, Luca, Jan-Laurin Müller, Conradin Braun, Domenique Zipperling, and Niklas Kühl. 2024. '*Implications of the AI Act for Non-Discrimination Law and Algorithmic Fairness*'. arXiv:2403.20089. Preprint, arXiv, June 26. https://doi.org/10.48550/arXiv.2403.20089.

Delponte, Laura. 2018. 'European Artificial Intelligence (AI) Leadership, the Path for an Integrated Vision'. *European Parliament: Policy Department for Economic, Scientific and Quality of Life Policies*. https://www.europarl.europa.eu/RegData/etudes/STUD/2018/626074/IPOL_STU(2018)626074_EN.pdf.

European Commission. 2025. *'International Outreach for Human-Centric Artificial Intelligence Initiative | Shaping Europe's Digital Future'*. February 18. https://digital-strategy.ec.europa.eu/en/policies/international-outreach-ai.

European Commission. European Political Strategy Centre. 2025. The Future of European Competitiveness. *Part A, A Competitiveness Strategy for Europe*. Publications Office. https://data.europa.eu/doi/10.2872/9356120.

European Data Protection Supervisor. 2023. *TechDispatch #2/2023 – Explainable Artificial Intelligence*. https://www.edps.europa.eu/data-protection/our-work/publications/techdispatch/2023-11-16-techdispatch-22023-explainable-artificial-intelligence.

European Parliament. Directorate General for Parliamentary Research Services. 2020. The Ethics of Artificial Intelligence: Issues and Initiatives. *Publications Office*. https://data.europa.eu/doi/10.2861/6644.

Flayyih, Najlaa, Mohammed Hasson Ali, Ahmad Fadii, and Khaled Aljasmi. 2025. "The Right to Informational Self-Determination between Legislation and Implementation." *AJEE* 10 (3). https://ajee-journal.com/the-right-to-informational-self-determination-between-legislation-and-implementation.

Floridi, Luciano. 2018. 'Soft Ethics, the Governance of the Digital and the General Data Protection Regulation'. *Philosophical Transactions: Mathematical, Physical and Engineering Sciences* 376 (2133): 1–11.

Floridi, Luciano, and Josh Cowls. 2019. 'A Unified Framework of Five Principles for AI in Society'. *Harvard Data Science Review* 1 (1). https://doi.org/10.1162/99608f92.8cd550d1.

Franke, Ulrike Esther. 2021. Artificial Divide: How Europe and America Could Clash Over AI. *European Council on Foreign Relations*. https://www.jstor.org/stable/resrep29123.

Frischhut, Markus, and Gabriele Werner-Felmayer. 2020. 'A European Perspective on Medical Ethics'. *Medicine* 48 (10): 634–36. https://doi.org/10.1016/j.mpmed.2020.07.001.

Fundamental Rights Agency. 2025. Assessing High-Risk Artificial Intelligence – Fundamental Rights Risks. *European Union Agency for Fundamental Rights*. https://fra.europa.eu/sites/default/files/fra_uploads/fra-2025-assessing-high-risk-ai-fundamental-rights-risks_en.pdf.

Grozdanovski, Ljupcho. 2025. 'Non-Discrimination Law, the GDPR, the AI Act and the – Now Withdrawn – AI Liability Directive Proposal Offering Gateways to Pre-Trial Knowledge of Algorithmic Discrimination'. *AI and Ethics* 5 (5): 5039–62. https://doi.org/10.1007/s43681-025-00754-0.

G'sell, Florence. 2024. Regulating under Uncertainty: Governance Options for Generative AI. *Stanford Cyber Policy Centre*.

Guadamuz, Andres. 2024. 'The EU's Artificial Intelligence Act and Copyright'. *The Journal of World Intellectual Property* n/a (n/a): 1–7. https://doi.org/10.1111/jwip.12330.

Hagendorff, Thilo. 2020. 'The Ethics of AI Ethics: An Evaluation of Guidelines'. *Minds and Machines* 30 (1): 99–120. https://doi.org/10.1007/s11023-020-09517-8.

Hakkarainen, Jenni, and Laura Savolainen. 2025. 'Individual Choice, Collective Effects: Recommender Systems, Law by Design, and the DSA's Double Choice Architecture'. *Information, Communication & Society* 0 (0): 1–18. https://doi.org/10.1080/1369118X.2025.2595663.

Häuselmann, Andreas, and Bart Custers. 2024. 'Substantive Fairness in the GDPR: Fairness Elements for Article 5.1a GDPR'. *Computer Law & Security Review* 52 (April): 105942. https://doi.org/10.1016/j.clsr.2024.105942.

High-Level Expert Group on Artificial Intelligence. 2019. Ethics Guidelines for Trustworthy AI. *European Commission*. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai.

Hutchinson, Terry, and Nigel Duncan. 2012. 'Defining and Describing What We Do: Doctrinal Legal Research'. *Deakin Law Review* 17 (1): 83–119. https://doi.org/10.21153/dlr2012vol17no1art70.

James, Ness. 2024. 'EU's AI Act Fails to Set Gold Standard for Human Rights'. Artificial Intelligence. *European Disability Forum*, April 3. https://www.edf-feph.org/publications/eus-ai-act-fails-to-set-gold-standard-for-human-rights/.

Kianzad, Behrang. 2025. 'Fairness, Digital Markets and Competition Law – Reconciling Fairness Norms in Digital Markets Act, Data Act and AI Act with Competition Law'. *Journal of Law, Market & Innovation* 4 (1): 133–60. https://doi.org/10.13135/2785-7867/11807.

Koniakou, Vasiliki. 2023. 'From the "Rush to Ethics" to the "Race for Governance" in Artificial Intelligence'. *Information Systems Frontiers* 25 (1): 71–102. https://doi.org/10.1007/s10796-022-10300-6.

Kotsios, Andreas, Thomas Taro Lennerfors, and Mikael Laaksoharju. 2025. 'Law in Engineering Ethics Education: An Exploration'. In *The Routledge International Handbook of Engineering Ethics Education*. Routledge.

Kusche, Isabel. 2024. 'Possible Harms of Artificial Intelligence and the EU AI Act: Fundamental Rights and Risk'. *Journal of Risk Research*, May 11, 1–14. https://doi.org/10.1080/13669877.2024.2350720.

Lennerfors, Thomas Taro. 2019. *Ethics in Engineering*. Studentlitteratur Ab.

Lundgren, Björn. 2020. 'Beyond the Concept of Anonymity: What Is Really at Stake?' In *Big Data and Democracy*, edited by Jai Galliott and Kevin Macnish. Edinburgh University Press. https://www.cambridge.org/core/books/big-data-and-democracy/beyond-the-concept-of-anonymity-what-is-really-at-stake/D32C88A08C843F1D0B769E98C4C2AA6A.

Manners, Ian. 2002. 'Normative Power Europe: A Contradiction in Terms?' *JCMS: Journal of Common Market Studies* 40 (2): 235–58. https://doi.org/10.1111/1468-5965.00353.

Marsden, Christopher T. 2011. *Internet Co-Regulation : European Law, Regulatory Governance and Legitimacy in Cyberspace*. Cambridge University Press.

McCarthy, John. 2007. "What Is Artificial Intelligence?" November 7. https://www-formal.stanford.edu/jmc/whatisai.pdf.

McCarthy, John, Marvin L. Minsky, Nathaniel Rochester, and Claude E. Shannon. 2018. "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence (1955)." Reprint of 1955 proposal. https://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html.

Milosavljević, Marko, and Sally Broughton Micova. 2016. "Banning, Blocking and Boosting: Twitter's Solo-Regulation of Expression." *Medijske studije / Media Studies* 7 (13): 43–58. https://doi.org/10.20901/ms.7.13.3.

Müller, Martin, and Matthias C. Kettemann. 2024. 'European Approaches to the Regulation of Digital Technologies'. *Introduction to Digital Humanism*, 623–37.

Munn, Luke. 2023. 'The Uselessness of AI Ethics'. *AI and Ethics* 3 (3): 869–77. https://doi.org/10.1007/s43681-022-00209-w.

Murphy, S. P. 2011. "Fairness." In *Encyclopedia of Global Justice*, edited by D. K. Chatterjee. Springer. https://doi.org/10.1007/978-1-4020-9160-5_257.

Nemitz, Paul. 2018. 'Constitutional Democracy and Technology in the Age of Artificial Intelligence'. *Philosophical Transactions: Mathematical, Physical and Engineering Sciences* 376 (2133): 1–14.

OECD. 2019. *OECD Recommendation of the Council on Artificial Intelligence*. OECD/LEGAL/0449. OECD. https://www.oecd.org/en/topics/ai-principles.html.

OECD. 2024. *'Explanatory Memorandum on the Updated OECD Definition of an AI System'*. OECD, March. https://www.oecd.org/content/dam/oecd/en/publications/reports/2024/03/explanatory-memorandum-on-the-updated-oecd-definition-of-an-ai-system_3c815e51/623da898-en.pdf.

Panarese, Paola, Marta Margherita Grasso, and Claudia Solinas. 2025. 'Algorithmic Bias, Fairness, and Inclusivity: A Multilevel Framework for Justice-Oriented AI'. *AI & SOCIETY*, *ahead of print*, July 15. https://doi.org/10.1007/s00146-025-02451-2.

Paul, Regine. 2024. 'European artificial intelligence "Trusted throughout the World": Risk-based Regulation and the Fashioning of a Competitive Common AI Market'. *Regulation & Governance* 18 (4): 1065–82. https://doi.org/10.1111/rego.12563.

Peukert, Alexander. 2024. 'Copyright in the Artificial Intelligence Act – A Primer'. *GRUR International* 73 (6): 497–509. https://doi.org/10.1093/grurint/ikae057.

Poel, Ibo van de, and Lamber Royakkers. 2023. *Ethics, Technology, and Engineering: An Introduction*. John Wiley & Sons.

Rachovitsa, Adamantia, and Niclas Johann. 2022. 'The Human Rights Implications of the Use of AI in the Digital Welfare State: Lessons Learned from the Dutch SyRI Case'. *Human Rights Law Review* 22 (2): ngac010. https://doi.org/10.1093/hrlr/ngac010.

Rendtorff, Jacob Dahl. 2015. 'Update of European Bioethics: Basic Ethical Principles in European Bioethics and Biolaw'. *Bioethics Update* 1 (2): 113–29. https://doi.org/10.1016/j.bioet.2015.12.004.

Rességuier, Anaïs, and Rowena Rodrigues. 2020. 'AI Ethics Should Not Remain Toothless! A Call to Bring Back the Teeth of Ethics'. *Big Data & Society* 7 (2): 2053951720942541. https://doi.org/10.1177/2053951720942541.

Russell, Stuart J., and Peter Norvig. 2022. Artificial Intelligence: A Modern Approach. Fourth edition, Global edition. With Ming-wei Chang, Jacob Devlin, Anca Dragan, et al. *Prentice Hall Series in Artificial Intelligence*. Pearson.

Sajduk, Błażej, and Dominika Dziwisz. 2024. *Comparative Analysis of AI Development Strategies: A Study of China's Ambitions and the EU's Regulatory Framework*. Beyond the Horizon ISSG. https://doi.org/10.31175/eh4s.2014.12.

Sax, Marijn. 2018. 'Privacy from an Ethical Perspective'. In *Handbook of Privacy Studies: An Interdisciplinary Introduction*, edited by Aviva de Groot and Bart van der Sloot. Amsterdam University Press. https://doi.org/10.1017/9789048540136.006.

Schneider, Henrique. 2024. 'Europe's AI Regulation Will Stifle Innovation'. *TECHNOLOGY*. GIS Reports, October 10. https://www.gisreportsonline.com/r/ai-act-eu-regulation-innovation/.

Smuha, Nathalie A. 2021. 'From a "Race to AI" to a "Race to AI Regulation": Regulatory Competition for Artificial Intelligence'. *Law, Innovation and Technology* 13 (1): 57–84. https://doi.org/10.1080/17579961.2021.1898300.

Sutrop, Margit. 2020. 'Challenges of Aligning Artificial Intelligence with Human Values'. *Acta Baltica Historiae Et Philosophiae Scientiarum* 8 (2): 54–72. https://doi.org/10.11590/abhps.2020.2.04.

Taebi, Behnam. 2021. Ethics and Engineering: An Introduction. *Cambridge Applied Ethics*. Cambridge University Press. https://doi.org/10.1017/9781316822784.

Trevino, Linda K., and Katherine A. Nelson. 2021. *Managing Business Ethics: Straight Talk about How to Do It Right*. John Wiley & Sons.

Ulnicane, Inga. 2022. 'Artificial Intelligence in the European Union'. In *The Routledge Handbook of European Integrations*, 1st ed., by Thomas Hoerber, Gabriel Weber, and Ignazio Cabras. Routledge. https://doi.org/10.4324/9780429262081-19.

UNESCO. 2021. *Recommendation on the Ethics of Artificial Intelligence*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000381137.

Vale, Sebastião Barros, and Gabriela Zanfir-Fortuna. 2022. FPF Report: Automated Decision-Making Under the GDPR – A Comprehensive Case-Law Analysis – Future of Privacy Forum. *Future of Privacy Forum*. https://fpf.org/blog/fpf-report-automated-decision-making-under-the-gdpr-a-comprehensive-case-law-analysis/.

Veale, Michael, and Frederik Zuiderveen Borgesius. 2021. 'Demystifying the Draft EU Artificial Intelligence Act – Analysing the Good, the Bad, and the Unclear Elements of the Proposed Approach'. *Computer Law Review International* 22 (4): 97–112. https://doi.org/10.9785/cri-2021-220402.

Wagner, Ben. 2018. 'Ethics As An Escape From Regulation. From "Ethics-Washing" To Ethics-Shopping?' In *BEING PROFILED: COGITAS ERGO SUM: COGITAS ERGO SUM: 10 Years of Profiling the European Citizen*, edited by Emre Bayamlioglu, Irina Baraliuc, Liisa Albertha Wilhelmina Janssens, and Mireille Hildebrandt. Amsterdam University Press. https://www.degruyterbrill.com/document/doi/10.1515/9789048550180-016/html.

Wessel, Ramses A. 2021. 'Normative Transformations in EU External Relations: The Phenomenon of "Soft" International Agreements'. *West European Politics* 44 (1): 72–92. https://doi.org/10.1080/01402382.2020.1738094.