Journal of Media Horizons
ISSN: 2710-4060 2710-4052 Volume 6, Issue 6, 2025

ARTIFICIAL INTELLIGENCE ACTS UNDER THE NATURAL LAW THEORY AND POSITIVISM PERSPECTIVE: A CRITICAL REVIEW

Qazi Naved Ahmad[1], Mudasar Latif Memon[*2], Sardar Ali Shah[3]

[1,3]*Institute of Law, University of Sindh, Jamshoro, Pakistan*
[*2]*Department of Information Technology, University of Modern Sciences, Tando Muhammed Khan, Pakistan*

[*2]mudasar.latif@ums.edu.pk

DOI:https://doi.org/10.5281/zenodo.17774389

Keywords
artificial intelligence regulation; natural law theory; legal positivism; EU AI Act; fundamental rights; algorithmic accountability; AI ethics; risk-based regulation

Article History
Received: 01 October 2025
Accepted: 19 November 2025
Published: 29 November 2025

Corresponding Author: *
Mudasar Latif Memon

Abstract

The collective phenomenon of artificial intelligence (AI) technologies rapidly spreading into the societies of the world has created a desperate need to establish extensive regulatory mechanisms to control the growth of these technologies, their use, and influence. With jurisdictions across the globe struggling to develop effective AI governance structures, there arise underlying issues of what the philosophical basis of such regulatory strategies is. This article thoroughly reviews AI regulation in the two-fold approaches of natural law theory and law positivism and how these two opposing schools of jurisprudence inform AI governance approaches today. The natural law theory assumes the legal justification based on the universal principles of morality, which are available with human reason, which puts the focus on the inherent human dignity, rights, and moral duties. On the other hand, legal positivism upholds that law is a human creation, distinct of morality and has its power originating in institutionalized sources and not metaphysics. We critically examine the expression of these theoretical frameworks in existing AI regulatory efforts with specific attention given to the European Union AI Act, which uses a risk-based approach to classification and compliance. By analyzing the key aspects of the idea of fundamental rights protection, the concept of algorithmic transparency, accountability principles, and ethical concerns, we show how natural law views highlight the role of universal moral imperatives in the governance of AI, whereas positivist views focus on formal legal frameworks and procedural adherence. The review also discusses algorithmic bias and discrimination in both theoretical terms, analyzes the issues in attributing liability and responsibility, and conflicts between innovation needs and rights protection. We single out major weaknesses in current structures such as fragmentation in jurisdictions, enforcing issues, and the necessity to harmonize internationally. Last but not the least, we would suggest balanced AI governance which would incorporate learnings of both schools of jurisprudence but would also deal with the reality of implementing AI to the ground. The provided analysis is part of the existing discussions of the best philosophical underpinnings to the regulation of AI and can be used to provide recommendations and guidance to challenges in AI regulation presented to policymakers, legal experts, and technology developers in the difficult field of AI regulation.

INTRODUCTION

1.1 Artificial Intelligence Regulatory Imperative

Artificial intelligence has been defined as among the most disruptive technologies of the 21st century which has fundamentally transformed the economic framework, social dynamics, and governance systems as well as the nature of human life [1], [2]. Algorithms in autonomous cars, predictive medicine, and robots in criminal justice and finance are just a few examples of how AI is applied in practically every aspect of modern life [3], [4]. The blistering development and popularization of AI technologies, nevertheless, have spawned significant ethical, legal, and societal issues that must be addressed immediately on the regulatory level [5], [6].

The rapid increase in AI capabilities has surpassed the advancement of sufficient regulation systems, which researchers describe as regulatory inertia in the face of increasing technological change [7], [8]. The governance gap takes form in various ways: algorithmic systems that issue consequential decisions about people without enough transparency or accountability, discriminatory AI systems serving to uphold and enhance established biases in society, privacy breaches due to invasive data collection and processing, and the fact that AI systems can pose a threat to the basic human rights and democratic principles [9], [10].

1.2 Jurisprudence Foundations: A Critical Divide

As policymakers, legal scholars, and technology developers look into all these challenges, some very fundamental questions emerge on the philosophical underpinnings on which AI governance models should be shaped [11]. The two schools of jurisprudence represent two different views on the nature of law and the connection between it and morality: natural law theory and legal positivism [12], [13].

The natural law theory, which has an intellectual ancestry stretching back to Aristotle as well as Cicero, argues that the validity of law is based upon compliance with a set of universal moral laws which are available to us because of, and in spite of, the operation of human intellect [14], [15]. This tradition holds that there are some moral truths that are not dependent on human convention, and those laws that violate these basic truths have no real legal power. The natural law focuses on human natural dignity, inalienable rights and the inherent relationship between law and justice [16].

According to legal positivism, as formulated most famously by John Austin, Hans Kelsen, and H.L.A. Hart, the law is a social construct made by the authority of humans and institutional structures. According to positivists, law should be distinctly differentiated with morality and its validity is based on the formal sources of authority instead of the content of the law. Law in this perspective is an order of rules developed by accepted procedures which are not to be confused with ethical or moral judgements [17], [18].

1.3 AI Regulation at the Crossroads: Natural Law vs. Positivism

The conflict between the two schools of thought is all the more relevant when it comes to the regulation of AI. According to natural law views, AI governance should be based on universal moral principles that safeguard human dignity, autonomy and fundamental rights [11], [12]. On this perspective, AI systems, which break fundamental ethical principles, including those that allow mass surveillance, social scoring, or discriminative decision-making, must be banned regardless of their technical complexity or financial advantages [19].

Positivist approaches, in its turn, assert that AI regulation has to be founded on well-set legal standards developed with the help of democratic legislative procedures [13], which are independent of controversial moral decisions. This approach places procedural validity, institutional adherence rules and the extrapolation of precedent cases to new technological situations in the foreground [20]. Positivist logic implies that AI systems need to be governed by explicit laws that define what may and should not be done with them, the ways of monitoring its use, and the principles of liability [13], [21]. Figure 1 shows the Global AI regulations landscape.

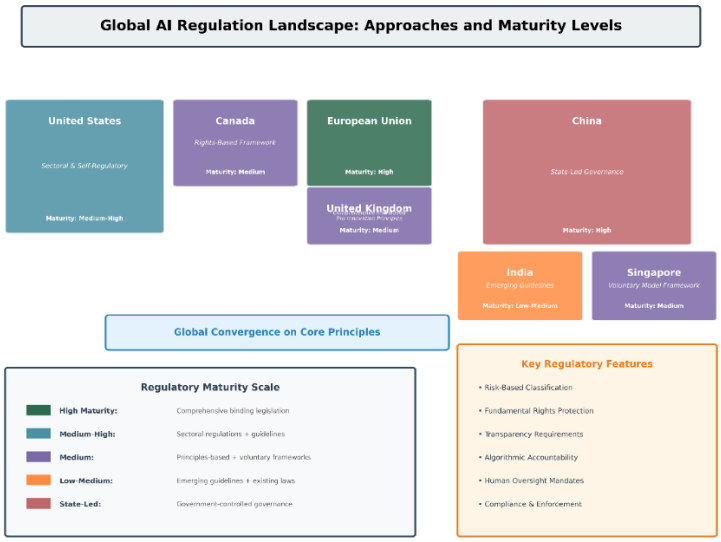

Figure 1 Global AI Regulation Landscape

1.4 The EU AI Act: A Regulatory Paradigm

The Artificial Intelligence Act passed by the European Union in 2024 is the first full-scale legislative initiative on AI regulation in the world, and its analysis is an essential case study of these philosophical contradictions [22], [23]. The AI Act is a risk-based framework, with four levels of AI systems depending on their harm-induced potential: unacceptable risk (prohibited), high risk (strong regulation), limited risk (transparency requirements), and minimal risk (no regulation) [24], [25]. This framework tries to reconcile innovation needs and basic rights protection and entrench transparency, accountability, fairness, and human controls into legal demands [26], [27].

The strategy of the AI Act is indicative of the two forms of jurisprudence. The fact that it puts its focus on fundamental rights affects assessments and bans AI systems that infringe upon human dignity appeals to natural law principles [12], [28]. At the same time, its elaborate scheme of risk classification, compliance requirements as well as enforcement provisions reflect positivist commitments to formal legal systems and procedures [29], [30].

1.5 Objectives and Contributions of This Review

The current comprehensive review paper considers AI regulation in the dual perspective of the natural law theory and legal positivism, and answers the following research questions:

1. What is the relevance of natural law and positivism views on modern AI regulations, especially the EU AI Act?
2. Which aspects of each jurisprudential approach do each of the main challenges in AI governance such as algorithmic bias, transparency, accountability, and basic rights protection imply?
3. What is the best way that policymakers can combine the lessons of these two traditions to create frameworks of AI governance, which are both balanced and effective?

Our analysis makes several contributions to ongoing debates regarding AI regulation:

- **Theoretical Framework:** We offer a procedural analysis of the application of the basic theories of jurisprudence to the new AI regulation issues, explaining the philosophical premises of various regulatory strategies.
- **Comparative Analysis:** We compare the EU AI Act and other regulatory projects in terms of natural law and positivism, finding contradictions and additionalities in these approaches.
- **Practical Implications:** We examine how philosophical commitments translate into concrete regulatory mechanisms, including risk assessment frameworks, transparency requirements, and accountability structures.
- **Critical Evaluation:** We find gaps, constraints and unresolved tensions in existing AI

regulation systems, and suggest solutions to more solid regulatory solutions.

1.6 Organization of the Review

The rest of this paper is structured in the following way. Section 2 gives a background information about natural law theory and legal positivism summarizing the fundamental principles and their historical evolution. Section 3 observes the AI regulation landscape in the world and especially the risk-based framework of AI Act by the EU. Part 4 examines key principles of protecting rights in AI governance in both the jurisprudential viewpoints. The fifth section explores the issue of algorithmic discrimination and bias with the help of natural law and positivism. Section 6 discusses the mechanisms of transparency, explainability, and accountability in AI systems. Section 7 explores the issue of problems in liability attribution and enforcement. In section 8, it addresses the conflicts between innovation and regulation. Section 9 works out the necessity of international collaboration and harmonization. Section 10 offers suggestions of balanced AI governance systems.

2. Theoretical Foundations: Natural Law and Legal Positivism

2.1 Natural Law Theory: Main Principles and the History of Development

The natural law theory is one of the oldest and most powerful philosophical traditions in the jurisprudence, which may have roots in the ancient Greek philosophy and is up to date legal thought with medieval scholasticism [31], [32]. The basic assumption about natural law is that some moral rules exist outside of the human convention and can be obtained by means of rational consideration of human nature and organization of the reality [33].

2.1.1 Classical Natural Law: Aristotle and Cicero

Moral and legal norms are based on the teleological organization of human nature, according to Aristotelian natural law [34]. Aristotle asserted that human beings have a specific nature, which allows humans to be rational and cooperate in society and that laws must help humans achieve happiness [35] . In this context, justice refers to the provision of every individual his or her due based on merit and contribution to the common good [36].

The natural law theory was also advanced by Cicero under the Roman legal tradition where he stated that true law was right reason in consensus with nature [37]; it was a universal law and its nature was immutable and eternal. This idea stressed the universality of natural law principles in various societies and their overruling in relation to positive (manmade) [33].

2.1.2 Thomistic Natural Law: Integration of Faith and Reason

Aristotelian philosophy had been integrated with Christian theology by Thomas Aquinas, who produced the most systematic statement of the theory of natural law during the medieval era [38]. Aquinas identified four laws: eternal law (divine providence), natural law (participation of humans in eternal law through reason), human law (enactments and positive accomplishments of the political authority), and divine law (revealed scriptural commandments) [39]. Aquinas believed that the natural law comprises precepts that are available to human reason, starting with the most basic one of doing good and avoiding evil [40], [41]. Certain moral rules are based on the rational consideration of human instincts that are self-preservation, the reproduction and addressing of descendants, the knowledge of truth and the social collaboration [38], [42]. To have true legal power, human laws need to agree with the natural law; according to Aquinas unjust laws are perversions of the law that do not compel the conscience.

2.1.3 Contemporary Natural Law Theory

Recent natural law philosophers such as John Finnis, Robert George, and Germain Grisez have recently redeveloped the classical natural law arguments without reference to Aristotelian metaphysics or theological assumptions [43], [44]. These theorists also propose "basic goods" basic aspects of human flourishing: life, knowledge, play, aesthetic experience, sociability, practical reasonableness, and religion, which give reasons to act and grounds to think of moral and legal norms [45].

The modern natural law theory highlights a number of principles that are applicable to the regulation of AI [46], [47]:

1. **Universal Moral Principles:** There are moral principles that cut across all across culture and all

historical times which rely on shared aspects of human nature.

2. **Human Dignity:** Constitutive of every Human being is dignity and worth that have to be honored in legal and social provisions.

3. **Rights and Justice:** People have some basic rights that relate to some basic human goods and safeguard the circumstances that humans need to thrive.

4. **Common Good:** Our law ought to be able to support the common good of political societies without any infringement of individual rights and dignity.

5. **Practical Reasonableness:** Legal systems should be guided by rational deliberation about how to coordinate social life in ways that respect and promote human goods.

2.2 Legal Positivism: Separation of Law and Morality

Legal positivism is a more specialised tradition of jurisprudence that developed in the 19 th century; however, its intellectual origins can be traced to the 16 th -17 th -century thinkers like Thomas Hobbes. Positivism opposed the confusion of law and morality in natural law, claiming that law and morality are distinct analyses, and that law is quite valid and moral correct [13].

2.2.1 Austinian Command Theory

John Austin pioneered a powerful early positivism based upon the idea of law as a set of commands supported by sanctions. Austin characterized law as rules given by a sovereign to the subjects and the sovereign is habitually obeyed and not habitually obeyed. On this basis, legal requirements are a result of the risk of punishment of non-adherence and not a result of moral obligations [14], [48].

Austin explicitly dismissed the natural law statement that immoral laws are not truly legal, citing that the nature and existence of law was determinable using social facts (orders of the sovereign) without mentioning of moral judgment. This thesis of separation--the thesis that the notions of law and morality are conceptually distinct--was to be a characteristic of legal positivism [49].

2.2.2 Kelsenian Pure Theory of Law

The more elaborate positivist model was developed by Hans Kelsen in his so called pure theory of law which attempted to explain law by a system of norms without considering any moral, political and sociological factors [13], [50]. Kelsen made a clear distinction between the terms is (descriptive facts) and ought (normative prescriptions), saying that legal science must only be concerned with the latter [50], [51], [52]. The key idea of the theory of Kelsen is the notion of the basic norm (Grundnorm) a supposed foundational norm that gives all norms of a legal system their ultimate source of validity. The foundational norm is not derivable by reference to some higher positive norm but must be assumed in order to sensitize the legal practice [53]. In the case of Kelsen, legal validity is hierarchical and it goes through the fundamental norm, constitutional clauses, and ultimately law statutes and specific legal norms. This hierarchical structure ensures the coherence and self-contained nature of a legal system, independent of external moral justifications [54].

The positivism of Kelsen is neutral about the morality of law. A norm is legally valid when it is formed with regard to procedures spelled out by higher-level norms, whether it is in the morally good or bad. Such formal notion of legal validity divides issues of what the law is, and what it should be [55].

2.2.3 Hartian Legal Positivism: Rules of Recognition

The conceptualisation of law positivism was furthered by H.L.A. Hart in his book The Concept of Law where he responds to criticisms of previous positivist theories. Hart disapproved of the command theory of Austin as being simplistic, developing an alternative viewpoint that legal systems are a union of primary and secondary rules [56], [57]. Primary rules are those rules which create duties or confer powers, whereas secondary rules are those which regulate the making, amendment and adjudication of primary rules.

The rule of recognition is the most significant secondary rule as it outlines rules of identifying legal norms which are valid in a system. Hart has contended that there is a traditional custom of legal officials to accept the rule of recognition, which offers the final justifications of validity of the law [58], [59]. Hart distinguishes the rule of recognition as a social fact, an existent practice of judicial officials, not a

logical assumption, in contrast to the basic norm formulated by Kelsen [59].

The relationship between law and morality is more subtle in how Hart views it than in the way that Austin or Kelsen view it. Hart admitted that legal systems generally involve an element of moral values, and moral factors might be applicable to the interpretation of the law. But he insisted that law need not be at all related to morality--conceptually, immoral rules may be legally valid [60], [61].

2.2.4 Key Principles of Legal Positivism

The legal positivism focuses on a number of principles that are applicable to the regulation of AI [13], [62]:

1. **Social Source Thesis:** Legal rules are based on social facts of institutional practices and not moral truths.
2. **Separation Thesis:** Law and morality are conceptually different; the validity of the law does not require and depend on the morality being correct.
3. **Formal Validity:** The legal norms are valid provided that they are formed in the procedures accepted by the legal system, no matter what they contain.
4. **Rule of Law:** There must be a set of publicly accessible rules that are followed by law enforcers consistently.
5. **Institutional Authority:** Law is formulated and implemented using accepted political setups with a legitimate power.

2.3 Tensions and Complementarities Between Natural Law and Positivism

Natural law has been seen as having a relationship of tension and complementary theory with legal positivism [12], [14], [63]. A number of areas of controversy are:

Nature of Legal Authority: Natural law is the source of legal authority based on the adherence to moral principles and positivism is based on institutional traditions and social recognition [14], [64].

Unjust Laws: The natural law school of thought holds that radically unjust laws are not laws, whereas positivists believe that immoral laws can be legally valid [14], [65].

Role of Morality: Natural law requires an inherent linkage between law and morality, whereas positivism enforces the separation of concept between law and morality [14], [64], [66].

In spite of these disputes, there are scholars who have suggested a combination of the two traditions. An example is the interpretive theory of law by Ronald Dworkin, which holds that legal practice is concerned with moral reasoning between interpretations that best justify the available legal materials in terms of principles of political morality [67], [68]. H.L.A. Hart also recognized the reality that legal systems generally include moral principles and that when rules are indeterminate, judges have sometimes to make moral judgments [18], [69], [70].

The discussion of the natural law versus positivism assumes a new meaning when it comes to the regulation of AI, as many of the fundamental questions about the connection between technology, human values, and legal norms require a keen philosophical approach [12], [71]. Figure 2 shows the natural law theory vs legal positivism.

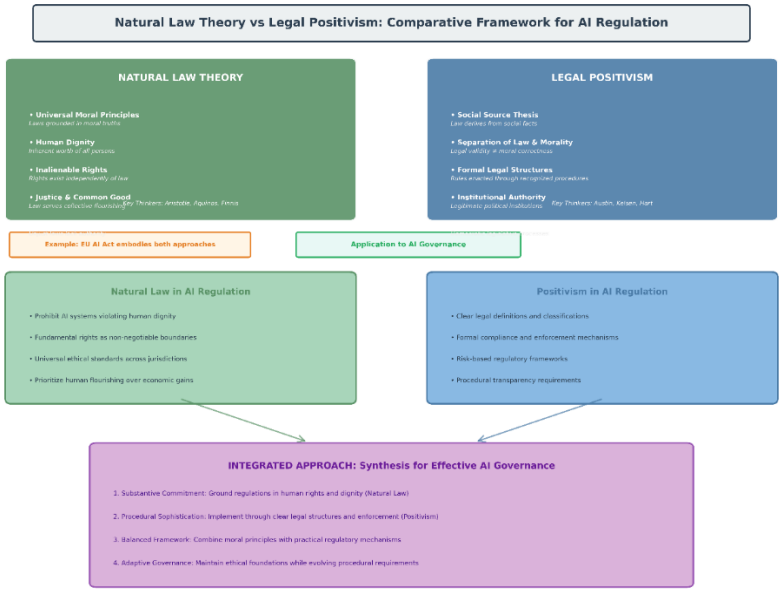

Figure 2: Natural Law Theory vs Legal Positivism

Discussion: The opposing views of the natural law and legal positivism offer different approaches to the conceptualisation of AI governance. The natural law strategies underline that the control of AI should be based on universal moral principles which safeguard human dignity and autonomy and ensure fundamental rights. The AIs that do not align with these principles must be banned no matter how technologically advanced or economically beneficial they may be. Positivist methods focus on formal legal institutions, procedural legitimacy and explicit rules that are implemented via a democratic procedure. The regulation of AI must provide clear set of requirements which are objectively verifiable and their enforcement should be founded on whether formal valid norms are met as opposed to the various moral judgements which are contentious. Awareness of these philosophical underpinnings can help to clarify the assumptions of various regulatory strategies, as well as shedding light on the conflicts that arise within the real-life AI regulation systems.

3. The Landscape of AI Regulation: Global Approaches and the EU AI Act

3.1 Emergence of AI Regulatory Frameworks

The fast changing nature of AI technologies has led to jurisdictions all over the world to come up with regulatory mechanisms to counter the ethical, legal, and societal issues introduced by such systems [72], [73]. Although methods differ greatly in the regions, or legal traditions, political systems and cultural values, there are various common themes that have emerged [74], [75].

Several reasons have informed regulatory efforts: the increase in understanding of AI risks, such as algorithmic bias and discrimination [76], [77]; AI failures or misuse leading to high-profile incidents [76], [78]; pressure on regulatory efforts by civil society organizations seeking greater protection [22], [79]; the realization that failure to regulate AI technologies can undermine societal confidence in the technology [10], [80].

3.2 Comparative Analysis of Global AI Regulatory Approaches

3.2.1 European Union: Risk-Based Comprehensive Regulation

The European Union has chosen the most extensive regulatory response to AI control, that is the Artificial Intelligence Act that came into effect in 2024. The EU AI Act is a horizontal, risk-based framework that would be applicable in all sectors of the economy with the goal of setting up various levels of regulatory requirements on the premise of the level of risk that AI systems may pose [81], [82]. Figure 3 shows the EU AI Act.

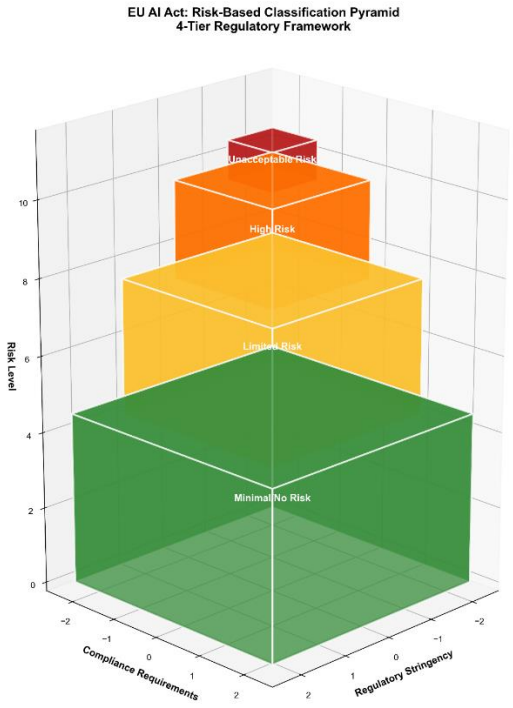

Figure 3 EU AI Act

The AI Act identifies four levels of risk of AI systems:

1. **Unacceptable Risk:** AI systems with evident threat to basic rights are forbidden completely. Social scoring by governments, real-time biometric identification in places of study (with some few exceptions related to law enforcement), manipulative AI systems, which take advantage of vulnerabilities, and systems that allow scraping of facial images on the internet or in CCTV video, are in this category [83].
2. **High Risk:** AI systems with material health, safety, or fundamental rights consequences are eligible to high risk requirements such as conformity tests, risk management systems, data governance, transparency, human oversight, and accuracy. Such high-risk areas as AI in critical infrastructure, education, employment, critical public and private services, law enforcement, migration and border control, and administration of justice are covered [22].
3. **Limited Risk:** AI systems that have certain transparency requirements, including chatbots and deepfakes, need to make users understand that they are engaging with AI-created content [22], [23].
4. **Minimal or No Risk:** AI systems with minimal risk (including AI-based video games or spam filters) do not require regulatory requirements that are specific to them.

The AI Act also creates special conditions concerning general-purpose AI (GPAI) models, especially those with systemic risks, to have the providers evaluate and contain risks, perform adversarial testing, report severe incidents, and have cybersecurity protections [22], [84].

The EU approach has several major characteristics:

Key features of the EU approach include:

- **Fundamental Rights Focus:** The Act gives primacy in the protection of fundamental rights guaranteed by the EU law and fundamental rights impact assessment must be done in high-risk systems [83].
- **Transparency and Accountability:** Providers and deployers must maintain detailed technical documentation, implement risk

management systems, and ensure traceability of AI operations [85].

- **Enforcement Mechanisms:** The national authorities may also audit, fine to the sum of EUR35 million or 7% of the global turnover in case of violations and prohibit non-compliant systems in the market [86].
- **Innovation Balance:** The Act will provide regulatory sandboxes and exemptions to conduct research to prevent suppressing positive AI development [22], [87].

3.2.2 United States: Sectoral and Self-Regulatory Approach

The US has pursued a more disjointed, more sector-centric approach to AI regulation, which is based mainly on legal and agency guidelines instead of legislation [74], [88], [89].

The U.S approach mainly involves the following factors:

- **Algorithmic Accountability Act:** Introduced Bill (undergoing review) would compel the company to implement AI systems in high-stakes situations to prepare impact evaluation assessing the risk of fairness, bias, accuracy, and discrimination [90].
- **Guidelines of the Federal Trade Commission (FTC):** The FTC has cautioned businesses about the employment of AI-based systems that practice unfair or deceptive conduct, imposing current laws on consumer protection to the scenario of AI [91], [92].
- **Blueprint of an AI Bill of Rights:** Published by the White House in 2022, the document is a non-binding framework of five principles of AI harm prevention: safe and effective systems, safeguards against algorithmic discrimination, data privacy, notice and explanation, and human alternatives and fallback options [93], [94].
- **Sectoral Regulation:** The FDA regulates AI applications in the medical field, the EEOC regulates AI applications in the law job sector, the HUD regulates AI applications in the housing sector, etc. [95], [96].
- **Executive Orders:** The federal agencies have been advised on how to use and control AI to ensure safety, security, privacy, equity, and civil rights [88], [97].

The U.S. approach is positivist in its nature that aims at trusting the legal authorities in place and formal institutional mechanisms. Nevertheless, observers believe that this piecemeal solution introduces regulatory loopholes and does not offer holistic safeguards against AI harms [98], [99].

3.2.3 China: State-Led Governance Framework

China has instituted a state-based governance of AI, which entails expanding development assistance and severe content and security regulations [95], [100], [101]. Key features include:

- **Personal Information Protection Law (PIPL):** The law establishes detailed data protection standards, like GDPR, and has special provisions on automated decisions [102], [103].
- **Algorithm Recommendation Regulation:** This regulation sets that the algorithmic recommendation services must offer mechanisms of explanation, prevent the use of information cocoons, and permit users to reject algorithmic recommendations [104], [105].
- **Generative AI Regulations:** Requires generative AI services to uphold social morality, content accuracy, and put in place controls to avoid creation of unlawful content [79], [106].
- **Cyberspace Administration of China (CAC):** Key regulator that has extensive powers to scrutinize, grant, and authorize AI systems [100], [101], [107].

The tactics of China are concerned with national security, social stability and control of Communist Party and economic development which represent the different political and cultural values by the Western democracies [73], [95].

3.2.4 Other Jurisdictions: Emerging Frameworks

United Kingdom: The UK has come up with a pro-innovation strategy since Brexit that is founded on five principles that are implemented in sectors including safety, transparency, fairness, accountability, and contestability. The current regulators apply these principles to their spheres

instead of establishing other institutions related to AI [108], [109].

Canada: A section of Bill C-27, The Artificial Intelligence and Data Act (AIDA) introduces requirements related to high-impact AI systems such as impact assessments and measures to mitigate risk and notification of serious harms [110], [111].

Singapore: The Model AI Governance Framework gives voluntary advice that focuses on transparency, explainability, and human-centricity, and particular sector needs are imposed on legal bases through prior regulations[112], [113].

India: India does not have any specific AI legislation but uses the existing data protection laws and has issued AI principles in the form of guidelines by NITI Aayog focusing on responsible AI [114], [115].

3.3 EU AI Act: In-depth Analysis

As it covers the entire range and has the potential to establish international standards, the EU AI Act is to be scrutinised in the frames of both natural law and legal positivism [116], [117].

3.3.1 Risk-Based Approach: Philosophical Foundations

The risk-based framework in the AI Act is a manifestation of the natural law and the positivist commitments. The ban of AI systems which cause unacceptable risks reflects the natural law principles in that it creates absolute moral boundaries, that is, there are some uses of AI that are prohibited, no matter how beneficial they might indicate, they are illegal and they cannot be allowed to happen due to their violation of the fundamental human dignity and rights [19], [118]. The use of social scoring systems that classify the trustworthiness of citizens according to their behavior, such as those based on social scoring, is forbidden since it infringes on human dignity and autonomy which are two fundamental values of natural law [119].

At the same time, the risk classification system is a good example of positivism methodology, as it defines definite, formal categories with certain legal implications [13]. The providers are able to know what they are required to do using objective criteria except abstract moral reasoning. Positivist adherence to formal legal formulas is evident in the focus of the Act on conformity tests and technical documentation and procedural compliance [13].

3.3.2 Fundamental Rights Impact Assessment

Article 27 of the AI Act mandates deployers of high-risk systems to perform basic rights impact assessments (FRIAs) upon the deployment of AI systems. Figure 4 shows the fundamental rights impact assessment process. FRIAs should determine and assess the possible effects on fundamental rights guaranteed by the EU law [120], [121], which include:

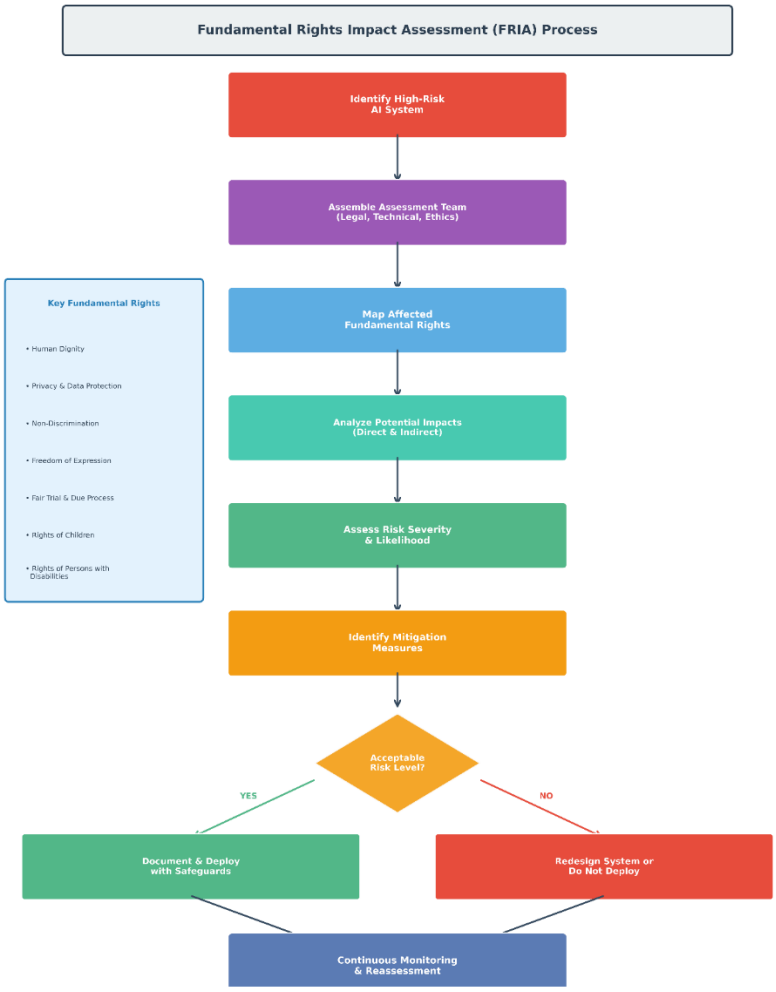

Figure 4 Fundamental Rights Impact Assessment Process

- Right to human dignity;
- Admiration of personal and family existence;
- Protection of personal data;
- Freedom of information and expression;
- Freedom of association and assembly;
- Non-discrimination;
- Child, elderly and person with disability rights;
- The right to an effective remedy and fair trial;
- Right of defence;
- Presumption of innocence.

The FRIA requirement is an expression of the natural law thinking because it mandates a clear [122]consideration of the impact of AI systems on morally fundamental human rights [121], [122]. It puts into effect the natural law principle according to which technology should not violate the inherent human dignity and fundamental freedoms. Nevertheless, the pro forma nature of FRIA requirements, detailing what should be recorded and disclosed, also indicates positivist focus on formal compliance procedures [120], [121].

There is a concern by critics that FRIAs will become box-checking, instead of meaningful rights protections. The views of natural laws would underscore the fact that the protection of real rights would entail a sense of moral dedication and not just formality. Strong FRIAs should be based on substantial and substantial participation with communities who have been affected, human rights knowledge, and readiness to avoid the usage of

systems which jeopardize the fundamental rights despite being formally adherent [123].

3.3.3 Requirement of Transparency and Explainability

The AI Act creates transparency requirements of high-risk and limited-risk systems. The high-risk AI systems should be created in such a manner that there is proper transparency where the user is able to derive meaning in the outputs and use system appropriately [124], [125]. A documentation is a requirement that will require disclosure of system characteristics, capabilities, limitations, and performance measures [126], [127].

In the limited-risk systems, the requirements of transparency are to inform people that they are communicating with AI systems (chatbots), that AI is applied to produce or alter content (deepfakes), or that biometric categorization or emotion recognition systems are in operation [22], [25].

According to the natural law, transparency needs safeguard human autonomy because transparency can allow informed decision-making regarding dealing with AI systems [128]. People have an ethical entitlement to come to know that important decisions are taken by algorithms, and not by the human judgment. Opacity compromises autonomy and dignity by avoiding persons as objects to be manipulated instead of as rational beings with the right to be respected [129], [130].

The positivist analysis concentrates on procedural requirements that are specific to transparency, namely what must be disclosed, to whom it must be disclosed, in what form and under what trade secrets exceptions [131]. This approach of the Act is evidence that the concept of transparency has regulatory purposes (transparency, accountability, user trust), which can be formalized by the formal disclosure requirements [132], [133].

3.3.4 Human Oversight and Algorithmic Decision-Making

The AI Act (article 14) mandates that the high-risk AI systems must be structured in such a manner that facilitates human supervision [134]. The human supervision should make people able to:

- Categorically comprehend the abilities and constraints of the high-risk AI system.
- Be conscious of potential automation bias.
- Interpretation of system output correctly.
- Choose not to utilize the system or ignore it, bypass it or reverse it.
- Stop the functioning of the system by use of a stop button.

These needs indicate natural law obligations to safeguarding human agency and responsibility to consequential decision-making processes [135]. The automated systems are not supposed to fully override the human judgment especially in a situation that touches on fundamental rights [136]. Human dignity and responsibility is safeguarded by the fact that the human is allowed to maintain the ability to override outputs of the algorithm [137], [138].

Under a positivist approach, the human oversight requirement will create clear legal requirements in terms of the system design and operational processes [134], [139]. They define jobs and duties of human operators, developing accountability frameworks based on formal legal obligations, instead of ethical exhortation [139].

3.3.5 Challenges and Criticisms of the EU AI Act

The AI Act has a number of criticisms [22], [25], although it is a rather comprehensive act:

Definitional Vagueness: The critics claim that the primary terminologies such as an AI system and unacceptable risk do not have the required level of specificity, which may cause interpretive conflicts and unstable use [22], [140].

Opt-Out Mechanisms: Article 6(3) permits providers to determine on their own that their high-risk systems create "significant risks," which may be used as an excuse to avoid tough requirements [22].

Enforcement Problems: The implementation is based on the national regulatory authorities having different capacities and resources. There are considerable difficulties with ensuring a degree of uniformity in the implementation of 27 member states [141].

Innovation Concerns: There are worries among industry participants that compliance fees, regulatory onus can make a positive AI innovation hard, especially with smaller firms and startups [142], [143].

Jurisdictional Limits: The Act extends to AI systems located within the EU market or impacting EU

individuals but imposing demands on foreign providers is difficult [22], [144].

Discussion: The fact that various AI regulation frameworks exist around the world is indicative of both the convergence of their fundamental principles (transparency, accountability, and fundamental rights protection) and their differences in terms of implementation strategies. The EU AI Act is by all measures the largest effort to put these ideas into the form of binding legal obligations, reflecting conflicts between natural law imperatives of safeguarding the inherent human rights and positivist insistence on formal legal frameworks. The success of the AI Act would be determined by its successful implementation, sufficient resources in enforcement, collaboration with other countries, and adaptation to the fast-evolving technologies in AI. It is also probable that other jurisdictions will anchor their thinking on the EU model, and tailor solutions to the particular legal traditions and political values of their respective jurisdictions, in creating a complex global picture of AI governance that is coordinated to prevent fragmentation and regulatory arbitrage.

4. Fundamental Rights Protection in AI Governance

4.1 Theoretical Foundations: Rights in Natural Law and Positivist Frameworks

Defense of fundamental rights is one of the core issues AI governance should consider, whereas natural law and legal positivism are based on different versions of the nature and origin of rights [28].

4.1.1 Natural Law Conception of Rights

The concept of natural law of rights may be explained using the economic model of the free market system [44].

In natural law theory, the rights of man are based upon the natural dignity of human beings and needs of human thriving [145]. On this perspective, rights are not legal practices but moral claims, which exist even outside positive legal acknowledgment [145], [146]. Thomas Aquinas held that natural law gives rights commensurate to the most basic human goods, which are life, knowledge, sociability, practical reasonableness and which must be honored to allow persons to thrive [147].

Modern theorists of natural law also underline that the main rights ensure the conditions that safeguard human dignity and agency [146]. According to John Finnis, some of the most fundamental forms of human good lead to correlative rights, life and health, knowledge and aesthetic experience, excellence in work and play, harmony amongst persons, harmony with ultimate sources of meaning, and practical reasonableness in choosing and acting [44], [148].

As applied to AI governance, natural law views are that:

1. **Inherent Dignity:** All human beings have the inherent dignity, which ought to be held by any means in spite of the contingent features of intelligence, social position or economic output [12], [46], [149]. AI systems which view people as objects or instruments are in violation of this primary principle [150], [151], [152].
2. **Universal Rights:** Universal rights, which are inalienable to all people in all cultures at all times, e.g. rights to life and liberty, privacy, non-discrimination, fair process, etc. These rights should be safeguarded under AI governance not because of prudence but as some moral imperative [153], [154].
3. **Playing with Fire:** There are certain applications of AI that cannot be permitted under any circumstances since they are intrinsically deemed to breach human dignity, despite the possible advantages [155], [156]. Such violations include mass surveillance, social scoring, and manipulation of vulnerable people [9].
4. **Common Good:** AI governance must advance the common good without violating individual rights [157], [158]. Technology should serve human values instead of determining them in order to flourish the communities [11], [159].

4.1.2 Positivist Conception of Rights

Legal positivism, in contrast, sees rights as laws or constructions established by the institutions, and not as the rights existing on the basis of moral rights whether known by human beings or not [18]. H.L.A. Hart made a distinction between legal rights

(recognition under positive law) and moral rights (recognition under moral rules), and stated that the presence of legal rights is a matter of social facts about legal systems, not a moral fact [13].

Positivist approaches to rights lay stress on:

1. **Institutional Source:** Rights are based on constitutional text, legislative acts, international treaties, and judicial rulings, which represent the formal legal sources and not the natural law [160].
2. **Contingent Recognition:** Which rights are legally protected depends on the specific commitments of particular legal systems. It does not imply that there are rights that any legal system should acknowledge [161], [162].
3. **Procedural Implementation:** Protecting rights involves clear legal regulations stating rights of those who bear them, the obligations of those who bear rights and those who bear duties, relief mechanisms in the event such rights are violated and methods of enforcement [163].
4. **Democratic Legitimacy:** Democratic processes of law making create the rights or constitutional conventions and obtain their legitimacy based on popular sovereignty and not moral philosophy [163], [164].

When applied to AI governance, positivist views focus on that basic rights protection demands [11], [13]:

1. **Formal Enactment:** Rights should be enshrined into laws, rules or constitutions with express legal force [165].
2. **Specification of Obligations:** Legal frameworks need to indicate what AI developers, deployers, and users should (or should not) do to observe rights [166].
3. **Enforcement Mechanisms:** Rights demand institutional mechanisms on how to monitor compliance, adjudicate on disputes and remedy violation [23], [156].
4. **Procedural Guarantees:** The protection of the rights is not based on abstract moral principles, but instead on procedural requirements, such as impact assessment, transparency requirements, consent requirements [167], [168].

4.2 Fundamental Rights Challenges Posed by AI Systems

They present a unique challenge to the protection of fundamental rights that AI technologies put a strain on both the natural law and the positivist paradigm [169].

4.2.1 Privacy and Data Protection

The use of AI usually involves tremendous amounts of personal information to train and to work, posing significant privacy threats. Machine learning algorithms can make conclusions about personal sensitive issues such as health conditions, sexual orientation, political beliefs, financial status, etc. based on apparently innocent data [170]. Predictive analytics allows unprecedented surveillance opportunities where corporations and governments can monitor people, profile and manipulate large numbers of people at scale [171], [172].

According to the natural law view, privacy breaches lead to diminishing human dignity and independence because it degrades humans to data subjects whose actions can be predicted and manipulated [173]. Right to privacy gives safeguard to the domain of personal life required to genuine self determination. AI-driven systems allowing ubiquitous surveillance or manipulative targeting attack this right [174].

The positivist analysis emphasizes the quality of AI systems to meet the formal data protection requirements [175]. The General Data Protection Regulation (GDPR) of EU and other similar laws have legal requirements on data collection, processing, storage and sharing. The AI systems should comply with the principles of purpose limitation, data minimization, disclosure, and the rights of an individual (access, correction, deletion). Positivists consider the protection of privacy by determining how it adheres to the following procedurally defined requirements in lieu of abstract notion of dignity [176].

4.2.2 Non-Discrimination and Equality

It has been demonstrated that AI systems reinforce and intensify discriminatory trends and generate biased results in the area of employment, lending, criminal justice, health care, and other high-stakes sectors [177]. Algorithms, when discriminated against can take the forms of biased training data,

misbehaving proxy variables, reflecting discriminating objectives, or arising patterns that are not necessarily linked to the characteristics being protected [90], [178].

The natural law theory is of the view that the non-discrimination requirements are a result of the inherent equality and dignity of human beings [44], [137]. All humans have equal moral value irrespective of race, gender, religion among others. This is a fundamental principle of natural justice that AIs systems that are not applied equally to individuals without impeccable reasons. The ethical obligation of avoiding discrimination is not based on whether the bias is deliberate or is a consequence of technical devices [179], [180].

Positivist paradigms deal with discrimination by providing anti-discrimination laws that define the nature of the characteristics of protection (race, sex, religion, national origin, disability, age) and the types of actions which are discouraged (disparate treatment, disparate impact [179], [181]). Legal analysis Pay attention to whether the results of AI systems are considered by the statutes as the discriminating result and whether the given discrimination may be justified under the legal exception (bona fide occupational qualifications, business necessity) [182], [183].

Nevertheless, AI-based discrimination is a problem to both frameworks. Natural law faces the challenge of whether algorithmic clusterings that are not linked to common protected features but are still against the primordial equality principles [179]. Positivist models have challenges when the AI systems are discriminative in relation to new categories that have no protection against discrimination in current laws against discrimination or when the outcomes of discrimination are difficult to attribute causally [177], [184].

4.2.3 Due Process and Fair Trial Rights

Artificial intelligence systems are taking on more and more roles in legal practice or making decisions, including bail, sentencing, parole, legal research and forecasting [185]. Such applications cast serious doubts on the right to due process and fair trial safeguarded under the natural law and under the affirmative legal systems [186].

Natural law underlines the need of justice that involves fair procedures in dispute resolution and legal consequences [185], [187]. People who have been subject to an unfavorable legal ruling are entitled to a notice, hearing rights, fair judgment and interpretation. These basic procedural safeguards are compromised by AI systems that issue opaque prediction of recidivism or future dangerousness without providing information on how they do so [188], [189].

Positivist analysis looks at whether the use of AI in the law has adhered to constitutional due process requirements and the statutory procedural requirements [188], [190]. Fair procedures must be granted by the U.S. Constitution Fifth and Fourteenth Amendments before life, liberty, or property is taken. Article 6 of the European Convention on Human Rights has a right to fair trial which comprises of unbiased tribunals and rational judgements [191]. The AI systems should meet these legal formalities by being transparent, explainable and human controlled [191], [192].

4.2.4 Freedom of Expression and Information

The freedom of expression and information is granted by United Nations [158]. The content moderation, recommendation algorithms, and automated decision making on what content users watch are the ways AI systems define information ecosystems [193]. These systems bring up issues of freedom of expression and access to information- right to fundamental rights under natural law and under positivist systems [193].

Natural law bases the freedom of expression on human rational nature and conditions of social organization that free speech and common deliberation should seek to establish the truth [194]. Individuals possess natural rights of freedom of speech, access to various information and engagement in the societal debate. The manipulation of information flows in artificial ways by AI systems, the formation of filter bubbles, or the act of censoring expression without justifiable reasons is a violation of these basic freedoms [195].

The question that positive analytical approaches address is whether AI content moderation can adhere to constitutional protection of free speech (in a country such as the U.S.) or the compatibility of the rights of expression with other values such as privacy, dignity, and the rule of law (in a country such as the EU). The laws should outline the acceptable reasons

to censor content and the legal safeguards that the aggrieved speakers are entitled to it [196].

4.3 Implementing Fundamental Rights Protection: Convergence and Tensions

Although natural law and positivist approaches agree on a few practical requirements of safeguarding fundamental rights to AI governance, despite philosophical differences:

Prohibition of Core Violations: The two frameworks advocate the banning of AI systems that outrageously violate core human rights. Such prohibitions are considered to be examples of moral imperatives by natural law and of formal legal imperatives by positivism, and are introduced by legitimate processes [166], [197], [198].

Transparency and Explainability: Protection of rights presupposes that a person should be aware of the impact of AI systems on him or her. This is based on natural law on autonomy and dignity, on positivism on procedural due process and accountability [128], [199].

Impact Assessment: Both strategies facilitate requirements to evaluate the possible effects of AI systems on rights during pre-deployment [200], [201]. Natural law considers impact evaluation as an operationalization of moral deliberation on the effects of technology on human flourishing; positivism is an operation aimed at finding and reducing the legal risks [120], [163].

Human Oversight: Consequential decision-making which safeguards moral responsibility (natural law) as well as legal responsibility (positivism) [202].

Tensions remain regarding:

Source of Rights: Natural law claims that there are rights that exist without legal sanction; positivism views rights as constructs of the law [166], [203], [204].

Scope of Protection: Natural law may favor greater protection of human dignity and prosperity than was formally specified in any rights parts, positivism only protects entitlements defined by legal language [163], [205].

Balancing: In cases whereby rights clash or need to be weighed off with other morals (innovation, security), natural law appeals to substantive moral reasoning on human goods, and positivism is based on the procedures and standards provided by the law [14].

Discussion: Both a substantive obligation to a human right (natural law) and explicit procedural arrangements (positivism) are needed to guard the fundamental rights in AI governance. The combination of the prohibition of rights violations, basic rights impact assessment, transparency, and human oversight provisions of the EU AI Act is synthesis of these views. But ultimately, the key to protecting human rights is ensuring that there is a commitment by all concerned parties to ensure that human rights are put first before technical complexity or economic gain is put in consideration, coupled with the ability to enforce the rights and availability of remedies in circumstances where human rights are infringed upon. Both philosophical models do not offer a full set of guidelines, and AI regulation should combine ethical reasoning about human dignity with functional legal frameworks on the protection of rights.

5. Algorithmic Bias and Discrimination: Theoretical Perspectives

5.1 Nature and Sources of Algorithmic Bias

Algorithms bias: The bias generated by AI systems is systematic and unfair and is usually in the form of disparate treatment or disparate impact on people or groups due to some of the protected or sensitive characteristics [206], [207]. To comprehend the problem of algorithmic bias, it is necessary to look at its technical causes as well as the moral and legal consequences of these causes [90], [207].

5.1.1 Technical Sources of Bias

Algorithms may become biased in several points of AI systems development and implementation:

Training Data Bias: Training data are patterns based on which machine learning models are trained. Models replicate and intensify such biases when they are trained using historical discrimination or when they do not adequately represent some groups of people. The facial recognition systems are mostly trained on white male faces, such as, they do not work on women and people of color [208], [209]. Figure 5 shows the sources and propagation of algorithmic bias in AI systems.

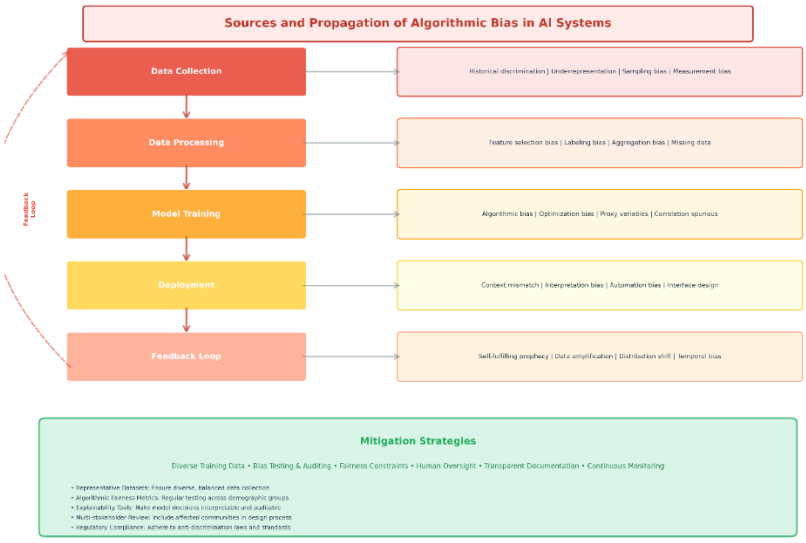

Figure 5 Sources and Propagation of Algorithmic Bias in AI Systems

Proxy Variables: The AI might utilize apparently neutral variables, which are associated with the features to be discriminated, and thus it creates a form of discrimination by proxy. Racial proxy: It is possible that the use of zip economy in lending selections, say by using zip code, is a proxy of race due to residential segregation patterns [210], [211].

Optimization Targets: Machine learning systems optimize on desired goals that can include discriminative priorities. An employment algorithm that optimizes on the basis of characteristics of current employees what has been termed cultural fit can continue to promote homogeneity and reject qualified diverse applicants [212], [213], [214].

Emergent Patterns: AI systems with more complex forms can form discriminatory patterns that are neither programmed nor expected by the developers [215], [216]. The deep learning models can detect and misuse the correlations between sensitive attributes and predictions in a manner that cannot be detected [211], [217].

Feedback Loops: AI systems operating in a dynamic environment are capable of establishing feedback loops that become self-reinforcing. The predictive policing models which focus the enforcement on the already over-policed neighborhoods score more arrests in these neighborhoods, which in turn legitimize the prediction of the model and reinforces the discrimination trends [218], [219].

5.2 Natural Law Perspectives on Algorithmic Discrimination

The natural law theory is a very strong theory of the condemnation of algorithmic discrimination on the basis of certain basic principles of human equality as well as human dignity [178], [220].

5.2.1 Fundamental Equality Principle

According to the natural law, human beings are equal in terms of moral worth and dignity due to their rational nature and moral agency. This inherent equality produces a principle of non-discrimination: individuals are to receive equal treatment without exceptions except in case there are some differences that are of a moral character and which warrant the provision of different treatment [221], [222].

When the AI systems handle individuals differently due to morally irrelevant traits such as race, gender, or ethnicity, algorithmic discrimination breaks the basic equality [130], [178], [223]. Natural law denounces this kind of discrimination not just in that it breaches the positive legal norms but rather in the fact that it opposes the fundamental fact of equality in human beings [178], [220].

5.2.2 Dignity and Objectification

Natural law underscores that human dignity needs a person to treat other persons as ends and not as means to the end of other persons [150]. In that algorithmic

systems categorizing, ranking, and predicting individuals of statistical patterns at group levels objectify people, turning them into an element of data sets instead of acknowledging their agency as individuals and their moral position [174], [178].

AI systems deprive people of opportunities, disadvantage them, and offer unequal treatment opportunities due to being denied by algorithms due to their individual merit and choice, which undermines human dignity [220], [224]. Natural law would accept that person-respect law demands that one look at individual situations, preferences and possibilities instead of placing them in groups out of statistical associations [178], [220].

5.2.3 Justice and the Common Good

The adherence of natural law to justice and the common good gives further reasons to criticize algorithmic discrimination. Justice, which means the rendering of each individual his/her due, must be just distribution of social burdens and social benefits [36], [225]. The fact that AI systems have a systemic disadvantage of already marginalized groups prolongs injustice and is against the common good because it denies these people a full involvement in the social, economic, and political life [226].

In the natural law theory, the common good refers to those conditions that allow everyone in the society to prosper [227]. Such discriminatory algorithms where machines reduce individuals to employees, credit, housing, education, or other goods they need based on biased algorithms are detrimental to the common good because they introduce obstacles to human flourishing to the population segments [220], [228].

5.2.4 Technical Solutions Limitations

According to natural law, there is more to algorithmic bias than technical solutions. Although enhancing information quality, deleting problematic variables, and enacting fairness restrictions are good, they do not provide the essential moral question: Are we treating persons as they deserve to be treated through their dignity and equality [178], [229].

Natural law would highlight that algorithmic discrimination can only be prevented through moral dedication to human equality and not adherence to fairness criteria. Technical procedures to maximize specific mathematical definitions of fairness without particularly reflecting on what morality is can be blind to the underlying moral problems involved [230].

5.3 Positivist Approaches to Algorithmic Discrimination

Legal positivism is considered to use algorithmic discrimination by means of anti-discrimination laws, regulatory systems, and court decisions, which define the forbidden behavior and the enforcement systems [90], [179].

5.3.1 Statutory Anti-Discrimination Frameworks

The first step in the positivist analysis will be to identify the anti-discrimination laws of relevance and whether the algorithmic results are covered by said laws. The most important frame works are [90], [183]:

Civil Rights Act (U.S.): Title VII bars any type of discrimination in employment based on race, color, religion, sex, or national origin. In assessing whether the employment algorithms are in violation of these prohibitions, courts employ disparate treatment and disparate impact theories [183], [231].

Fair Housing Act (U.S.): Bars any discrimination in housing on the basis of the characteristics that are protected. Extends to the algorithmic systems of screening tenants, lending mortgages, and valuation of property [232], [233].

Equal Credit Opportunity Act (U.S.): It prohibits the credit discrimination on basis of race, color, religion, national origin, sex, marital status or age. AI lending algorithms should satisfy the following requirements [234].

EU Anti-Discrimination Directives: Forbid any discrimination on the basis of sex, racial or ethnic origin, religion or belief, disability, age or sexual orientation in the area of employment, goods and services and other respects [211], [235].

GDPR (EU): Article 22 limits the use of automated decision-making with legal or otherwise serious consequences, in which case persons cannot be judged by automated means of action, such as profiling, without consideration [236].

5.3.2 Disparate Treatment vs. Disparate Impact

According to positivist systems of law, there are two types of discrimination:

Disparate Treatment: It is intentional discrimination whereby people are treated differently due to the existence of protected characteristics. When race-

based differences are made by an AI system through an explicit programming action, this amounts to disparate treatment [182], [237].

Disparate Impact: These are policies and practices that seem neutral but which have disproportionate impacts on the protected groups. Even in the absence of discriminatory intent, AI systems with vastly different results on racial or gender lines can be in violation of disparate impact prohibitions [238], [239].

The courts of the US have developed outlines of disparate impact claim analysis: the plaintiffs need to prove that disparate impact is statistically significant, the defendants might provide valid justification of business reasons, and the plaintiffs may demonstrate the less discriminatory alternatives [183], [240]. The application of these frameworks to algorithmic systems brings out complicated issues to do with causation, justification and alternative designs.

5.3.3 Challenges for Positivist Frameworks

The positivist legal approaches have many problems with algorithmic discrimination:

New Classes of protection: AI systems can discriminate against algorithmic groupings that are not related to more traditional, traditional protection characteristics. Assuming that an algorithm produces a category of high risk, which overrepresents people of color but is purportedly determined by other criteria, there is no guarantee that the current anti-discrimination law would be relevant [179], [210].

Causal Attribution: It might be technically challenging to determine that the outcome of the algorithm was due to discrimination and not other justifiable reasons, and this would involve skills in machine learning, statistics, and causal inference [180], [241].

Opacity and Explainability: A large number of AI systems are black boxes, meaning even the developers themselves cannot give any genetic explanations as to why specific predictions have occurred. This obscurity makes it difficult to legalize whether there was any discrimination [178], [242].

Regulatory Gaps: Algorithms are not seen to be covered by current anti-discrimination legislation, which might not be sufficient to address AI-specific issues, such as proxy discrimination, emergent bias, or feedback loops [90], [179].

5.3.4 Procedural and Technical Responses

Positivist methods put procedural and technical procedures to handle algorithmic bias:

Impact Assessments Algorithms: Making developers evaluate the possible discriminatory effects of AI systems prior to deployment, reporting bias testing, and putting in place cures [180].

Fairness Metrics and Auditing: The imposition of mathematical fairness restrictions and audits to ensure that the non-discrimination requirements have been met [92], [243].

Transparency and Documentation: The demand of the technical documentation of data sources, model architecture, testing processes, and performance across demographic groups [92], [244].

Regulatory Oversight: Authorizing agencies such as the FTC, the EEOC, HUD and its peers overseas to research algorithmic discrimination and put it into practice [183], [244].

5.4 Comparative Analysis: Natural Law vs. Positivism on Algorithmic Bias

There are a number of differences between natural law and positivist approaches to algorithmic discrimination:

Normative: Foundation Natural law bases non-discrimination duties on innate human equality and dignity; positivism on statutorily established protected classes [90].

Protection: Natural law may apply protection to more than the formally identified characteristics of protection; positivism restricts protection to those categories that are statutorily defined [14], [163].

Role of Intent Natural law: It is concerned with whether action observes fundamental equality even in the absence of discrimination by intent; positivism draws the distinction between intentional (disparate treatment) and unintentional (disparate impact) discrimination and subjects them to different legal standards [245].

Remedial Approach: Natural law is committed to moral education and moral devotion; positivism is committed to legal observance, sanctions, and remedies to legal breaches [246].

In spite of these distinctions, both frameworks do endorse a number of practical actions:

Ban on Intentional Discrimination: AI systems that are specifically built to discriminate against the

characteristics that are to be protected must be banned [94], [247].

Disparate Impact Analysis: When AI systems have disproportionate negative outcomes on groups that are protected, these results must be justified and alternatives with less discriminatory outcomes should be considered [238].

Transparency and Accountability: Algorithms decision developers and deployers ought to be capable of explaining and justifying algorithmic decisions, i.e. outlier discriminatory patterns can be detected and addressed [248], [249].

Stakeholder Participation: The impacted communities must be included into the AI governance process to make sure that systems do not violate their rights and interests [106].

Discussion: To solve the issue of algorithmic bias, we should combine the knowledge of natural law and positivism. Natural law serves as a reminder, that the basic moral question is the respect of human equality and human dignity--technical responses should not be used instead of this substantive moral response. Positivism also offers legal mechanisms that can be used practically in internationalizing non-discrimination principles by giving clear rules, enforcement procedures and remedies. The AI governance problem is to come up with frameworks that balance moral dedication to equality with procedural protections, meaning that the system of algorithms encourages and does not harm human dignity and justice. This involves not just technical interventions to identify and limit bias but cultural change in the perception of people affected by AI systems within the mindsets of developers, deployers, and policymakers on what they should do.

6. Transparency, Explainability, and Accountability in AI Systems

6.1 Conceptual Foundations: Transparency vs. Explainability

Transparency and explainability are similar, yet different, terms in the context of AI governance, both of which are needed to achieve accountability but need to be clarified.

Transparency means providing insight into the creation, implementation, and use of AI systems, that is, making public the details regarding the sources of data, algorithms, the process of training, performance indicators, and constraints. Transparency is a practice that exposes information to stakeholders such as developers, deployers, users, regulators and people who are going to be affected [250], [251], [252]. Figure 6 shows the multi-level transparency framework for AI governance.

Explainability issues relate to the concept of giving understandable explanations of the particular AI decisions or predictions- to allow users to know why certain things have happened. Explainable AI (XAI) methods provide model behavior that can be comprehended by humans by using algorithms such as feature importance analysis, counterfactual explanations, and attention mechanisms [242].

Transparency and explainability can be used to account for AI systems in the same way that they can be overseen, evaluated, and challenged. They however, treat two different sides of the black box problem transparency sheds light on how systems in general operate, and explainability sheds light on why particular decisions have been made [106], [253].

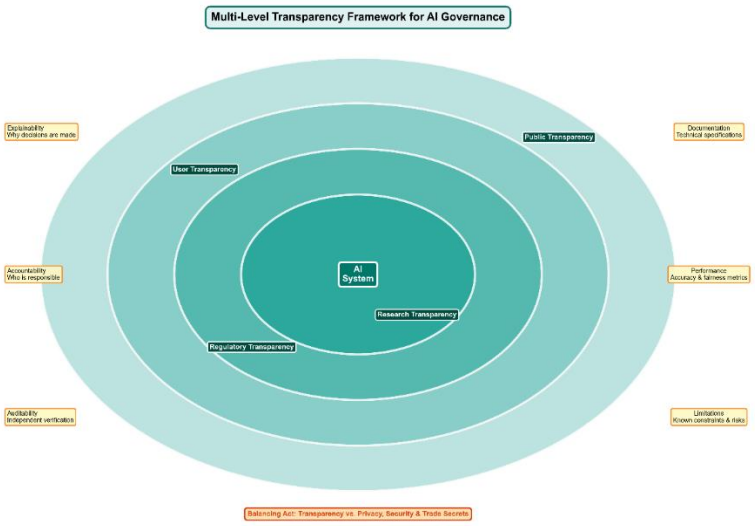

Figure 6 Multi-level Transparency Framework for AI Governance

6.2 Natural Law Perspectives on Transparency and Explainability

In natural law perspective, requirements of transparency and explainability guard the basic human good, such as autonomy, dignity, and the mediums of rational deliberation.

6.2.1 Autonomy and Informed Decision-Making

Natural law lays stress on the fact that human prosperity depends on self-direction, which is governed by practical reason. Individuals should be in a position to comprehend the aspects that determined their lives so that they could make sound decisions that are in line with their objectives and ideals. Transparency in the functioning of AI systems is necessary to protect human autonomy, as significant decisions can be made or influenced by the system without clear information on how it functions, which will not allow people to make an informed assessment and take action in this direction [92], [254], [255].

Explainability is defending the autonomy because the individual is able to conceive how specific results came to pass, evaluate whether the action was reasonable, and decide what to do (appeal, amendment of behavior, legal action). In the absence of such knowledge, individuals have been transformed into receptors of algorithmic decisions as opposed to agents who have control over their lives [256], [257].

6.2.2 Dignity and Respect for Persons

In natural law human dignity is based on the notion that persons are rational beings who are able to perceive and act on reasons. Making consequential judgments about people without explaining why they did not give a person a good reason, AI systems do not exhibit due respect to human rationality [258], [259]. Individuals are entitled to know the grounds of decisions that involve them and whether such decisions can be receptive to explanations that they can express and argue about [259], [260].

Opacity in AI decision-making may be a form of disrespect to act upon persons as objects to be categorized and classified rather than to be entitled to justification [178], [254]. Natural law would consider the need to respect human dignity to mean not only considering the correct predictions, but having understandable reasons, taking into consideration the ability of persons to comprehend and evaluate decisions that may concern them [261].

6.2.3 Justice and Procedural Fairness

The conception of natural law affects the idea of justice in the form of procedures: fair methods of

dispute resolution and of establishing the legal results. Justice mandates that individuals that receive a disadvantaged ruling be notified, given a hearing, and justification that allows them to determine whether they have been treated justly [262].

The AI systems that are transparent and explainable are more appropriate in addressing these requirements of procedural justice [263]. Where the predictions made by algorithms are applicable when deciding bail, sentencing, parole or any other law-related decisions are to be made, the affected individuals have a right to know the conditions behind such predictions and may contest them when necessary. The obfuscated systems, which fail to give rational explanations, are against the fundamental principles of procedural justice [264], [265].

6.2.4 Accountability and Moral Responsibility

According to natural law, moral responsibility is important: agents are to bear responsibility in regard to the outcomes of decisions and behavior. As AI systems are used in making decisions, there are questions concerning responsibility who is responsible, the developers, the deployers, the users or the systems? [266], [267].

Accountability is supported by transparency and explainability as they allow the moral responsibility to be assigned. When developers are not able to describe the reason why their system was able to generate specific results, it would be hard to determine whether they were using due care when designing and testing the system [248], [268]. When deployers implement opaque systems without knowing how they work, an assessment of their culpability in the harm that results would be challenging [269], [270]. Natural law would claim that moral responsibility involves the awareness of actions and their predictable outcomes, which is why opaqueness destroys the awareness and weakens the responsibility [199], [251].

6.3 Positivist Perspectives on Transparency and Explainability

Legal positivism invents transparency and explainability as procedural mandates that are implemented with the help of positive law to accomplish regulatory goals such as oversight, accountability, and user protection.

6.3.1 Regulatory Rationales

Positivist models single out a number of regulatory rationales of transparency and explainability requirements:

Monitoring and Checking Compliance: The regulators require the information concerning AI systems to determine whether it complies with the legal requirements. Technical documentation, performance measures and test outcomes can help authorities to certify that systems comply with statutory requirements [83], [271].

Accountability and Liability: With transparency, responsibility can be attributed in the event of harms by AI systems. Documented development processes, deployment situations, and characteristics of the operations are used to find out whether the developers, deployers or users violated legal responsibilities [272], [273].

User Protection: People who are subject to AI decisions need information to implement legal rights such as access, rectification, deletion of personal data as well as contest automated decisions [249], [251].

Competition in the market: It is possible to foster market competition through opening up the AI capabilities and constraints in a way that leads to informed market choices and to competition among the improved systems [274], [275].

6.3.2 Legal Frameworks for Transparency

Positive law puts up clear transparency requirements based upon jurisdiction and setting:

EU AI Act: The high-risk AI systems should be designed in a way that requires them to provide adequate transparency so that they can be interpreted appropriately by the user and used in a proper way. The providers should provide usage instructions, description of system features and performance, and details of human supervision controls [276].

GDPR: Provides that the logic of automated decision-making, its significance, and its projected consequences be disclosed. Data subjects are entitled to access information regarding processing and substantial explanations of choices [199], [277].

Sector-Specific Requirements in the U.S.: There are several laws that require disclosure of certain AI applications. The Fair credit reporting act dictates that adverse action notices should be issued to explain why credit is denied. The Equal credit opportunity act

demands justifications to credit decisions. Employment and other situations have state laws that progressively require AI to be transparent [232], [278].

6.3.3 Procedural Specifications

Positivist analysis concentrates on proceduralizing of the specifications of transparency: what information shall be revealed, to whom, in what form, being subject to what exceptions.

Technical Documentation: The EU AI Act outlines the technical documentation requirements such as descriptions of AI system components and development process, information of data governance and management, output of risk management system, and metrics to assess accuracy, robustness, and cybersecurity, and system limitations and assumptions [83].

User Information: Systems should present the users with clear and easily available information on system capabilities, correct uses, restrictions and anticipated level of accuracy [279].

Regulatory Access: The authorities need to be able to access technical documentation, testing results, and system operations to perform oversight and investigations [83].

Trade Secret Protections: Legal regulations should have a balance between disclosure and protectiveness of proprietary information to contain procedures of having confidential access to the trade regulation and restricting the disclosure of trade secrets to the general population [280], [281].

6.3.4 Explainability as Legal Requirement

Positivist models develop explainability as a legal requirement, and yet it is difficult to put into practice:

Right to Explanation: GDPR Articles Two states that individuals must not be the subject of decisions made by automated processing, such as profiling, resulting in legal or other similarly significant outcomes, without protections such as the right to human intervention and meaningful information on the logic [282], [283].

Adverse Action Notices: There are several laws that mandate an explanation of adverse outcomes produced by the algorithmic systems like credit rejections, employment rejections or insurance cancellations [284], [285].

Courtroom Affairs: In cases where AI predictions are used during the judgment of a case, due process can demand clarifications that allow the accused to oppose algorithmic evidence [130], [259].

Nevertheless, the realization of explainability requirements has both technical and practical difficulties. Numerous state-of-the-art AI systems (deep neural networks) are not easily explainable even to developers. To obtain "explanations", it can be necessary to make post-hoc approximations which are not intended to be accurate descriptions of model behavior [22], [286]. Legal systems have to deal with these technical constraints and still have significant accountability.

6.4 Challenges and Tensions in Transparency and Explainability

Both natural law and positivist approaches face challenges in realizing transparency and explainability goals:

6.4.1 Technical Complexity

The current AI systems, especially deep learning models, are exceptionally complicated containing millions or billions of parameters. It may not be possible even to experts to know why such systems generate certain outputs because of the fact that such systems exceed the cognitive ability of human beings. Such technical obscurity rises conflict between the principles of transparency and realistic and attainable feasibility [251], [287].

Natural law may state that because systems cannot be made transparent enough and explainable enough, they should not be applied in high stakes areas of application that involve fundamental human goods. Human dignity and autonomy is not a sacrifice to technical convenience or economic gain [178], [288].

Positivism underlines the fact that the law requirements must be pragmatically practicable. Forcing explainability that is impossible will build perverse incentives to fit less complex, but possibly less accurate models, or to produce false explanations that do not describe how the model actually works [289], [290]. The legal frameworks should indicate realistic standards between transparency and technical limitations.

6.4.2 Competing Interests: Transparency vs. Privacy and Security

Transparency may contradict other such values as privacy and security. The detailed disclosure of the work of the AI system may show:

Personal Data: Training information or model parameters that encode sensitive information about people.

Security Vulnerabilities: Weaknesses in the system that enemies might use.

Proprietary Information: Competitive advantages and trade secrets that developers rightfully desire to safeguard.

Where there is a clash between these values, natural law would give precedence to human goods most essential to a prosperous living. Where transparency that safeguards autonomy and averts discrimination demands certain sacrifices of privacy or proprietary rights, natural law may be on the side of disproportionate trade offs. Nevertheless, the unselective disclosure that puts people at risk of harm would be a crime against dignity and objectionable [179], [291].

Positivism aims at striking a balance between conflicting interests using procedural tools. Approaches where a qualified transparency gives regulatory access to confidential information and restricts public disclosure are known as qualified transparency. Tiered systems offer various levels of visibility to various stakeholders depending on their needs and functions [292], [293].

6.4.3 Meaningful vs. Formal Transparency

The fear has been that explainability and transparency requirements can be just a form gone dead. According to the natural law standpoint, when one respects human autonomy and dignity in a meaningful way, relevant transparency is necessary that, in fact, allows one to comprehend and make an informed decision, rather than just complying with disclosure regulations [46], [294]. The legal and regulatory systems should also make sure that the information supplied is understandable, applicable, and practical to people who are being concerned and not technical language or legal disclaimers.

Positivist analysis recognizes the fact that it is difficult to specify meaningful transparency, which is a legal term. It is possible to impose regulations, which require certain information to be disclosed and certain forms to be followed, however, to be sure that disclosures make the user informed and empowered, constant review and changes must be made. Not whether formal requirements are met, but whether transparency practices are going to work, need to be evaluated by enforcement authorities [124], [126].

6.5 Toward Accountable AI: Integrating Transparency and Explainability

A proper AI accountability level will necessitate the incorporation of both transparency and explainability in wider governance systems.

Multi-Stakeholder transparency: Various stakeholders demand various information. Developers require technical information to debug and improve upon it; regulators require access to help in supervision; and those affected should have an understandable description to make decisions. The governance models must state proper transparency of every group of stakeholders [295].

Algorithmic Auditing: Audits of AI systems can be independently conducted to ensure that they meet the requirements of transparency and fairness, and can offer accountability on top of what developers and deployers are willing to self-report [106], [296].

Human Oversight: Transparency and explainability allow human oversight to be an effective measure because it ensures human supervisors can interpret, judge, and intervene in the actions of AI where needed [92], [139].

Remedial Mechanisms: Transparency helps enhance accountability by allowing the identification of the responsible parties and the evaluation of the liability in cases where AI systems result in harms [128], [248].

Discussion: Transparency and explainability are complementary to each other in the context of AI accountability but have different difficulties. A natural law bases its needs on the considerations of human autonomy, human dignity and rational agency-ideals that technology cannot ignore but instead ought to support. Positivism achieves transparency by its specifications of procedures, to balance the requirements of oversight and the technical limitations and competing interests. Good AI governance implies both meaningful commitment to transparency (natural law) and working

mechanisms (positivism). The difference is making sure that requirements of transparency and explainability are not carried out as formal checking exercises but they truly facilitate comprehending, monitoring and responsibility. This requires continuous consultation with the developers of AI, legal academics, ethicists, communities impacted, and policymakers to improve on the requirements as technology and social knowledge advance.

7. Liability, Responsibility, and Enforcement Challenges

7.1 Attribution of Responsibility in AI Systems

Attribution of responsibility in AI systems is an essential component of AI systems, with its prevalent applications evident in smart robots and autonomous defensive driving systems. Attribution of responsibility in AI systems is a crucial element of AI systems, and its most widespread use can be found in smart robots and autonomous defensive driving systems [297], [298]. Implementation of AI in consequential decision-making situations creates complex issues regarding the allocation of legal and ethical accountability in situations where such systems generate detrimental effects [298], [299]. Figure 7 demonstrates AI responsibility attribution network.

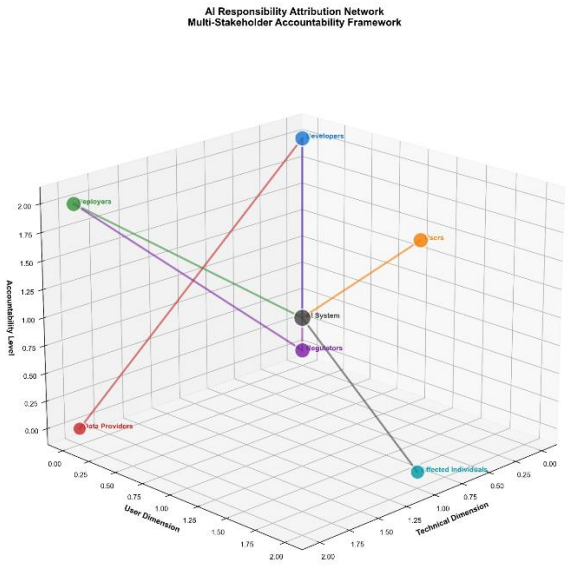

Figure 7 AI Responsibility Attribution Network

7.1.1 The Responsibility Gap

Artificial intelligence systems generate what researchers term a duty of care: The challenge of allocating ethical and legal accountability to choices undertaken by autonomous or semi-autonomous systems [156], [300]. Conventional models presuppose human actors to make choices to which they become responsible. AI makes this harder by spreading the decision-making among many actors (developers, deployers, users) and technical processes (algorithms, data) in a manner that does not allow the attribution of responsibility [301], [302].

When the accident is caused by an autonomous vehicle, who is to blame, the manufacturer, the software developer, the car owner, or the human passenger who did not succeed in intervening? In the case of an algorithm based on employment discrimination, is it the person who developed the model (the algorithm creator) or the company that implemented it (the deployer of the algorithm), the

human resource people that acted upon the recommendation of the algorithm (the actual users of the algorithm) or the data scientists who prepared the training data? [179], [211]. The distributed nature of AI ecosystems, where responsibility is shared across multiple actors, complicates traditional legal frameworks for liability [301].

7.1.2 Natural Law Perspectives on Responsibility

The natural law theory bases moral responsibility upon rational agency and causal contribution to outcomes. There is a variety of principles that guide natural law analysis of AI responsibility:

Human Moral Agency: It is only rational agents who are able to comprehend and make a decision based on reasons that can exercise the true moral responsibility. Being non-conscious, non-intentional, and unaware of morality, AI systems cannot be moral agents. Hence, it should be on human individuals who have participated in AI development, implementation, and application [19], [303].

Causal Contribution: People are causally liable to the consequences of the things that they cause by their decision-making and behaviors. Designers of biased algorithms, deployers who utilize them without proper precautions, and users who misuse the resulting recommendations are all contributors to the resultant harms as they all have a cause-and-effect relationship[301].

Foreseeability and Due Care: A part of moral responsibility is based on foreseeability- whether agents could reasonably foresee the consequences of their actions. The natural law would impose on AI developers reasonable harms that can be caused by insufficient testing, the lack of reduction of known biases, or even the application in the wrong context. Liability involves taking due care that is commensurate to the foreseeable risks [304], [305].

Collective Responsibility: In cases where several parties have led to the negative consequences, the responsibility can be shared or distributed. The focus by natural law on the common good justifies the collective responsibility of institutions and organizations to make sure that the AI systems do not violate human dignity and rights, not the individual liability [201], [302].

7.1.3 Positivist Approaches to Liability

Legal positivism deals with harms that AI brings by implementing existing liability regimes into new technological situations.

Product Liability: There is a strict liability of products due to injuries, which the manufactures can be liable. This framework can probably be extended to AI systems, which malfunction or act in an unsafe manner. The AI Liability Directive proposed by the EU shifts the product liability onto AI systems and introduces rebuttable presumptions of causality to overcome the challenges of evidence [306], [307].

Negligence: The parties have a reasonable careful duty to prevent foreseeable damages. Applicable duties of care have the potential to make developers, deployers and users culpable towards negligence in case they violate the duty and cause injury. The legal issue of how AI systems should be dealt with is a developing one [305], [308].

Vicarious Liability: Employers are liable to the tortious behaviours of employees under employment. Other researchers suggest adding AI systems to the framework, and they are viewed as agents with whom the principals are held accountable. But that demands to answer the question whether AI systems can be agents in a legal sense [309], [310].

Statutory Liability: There are anti-discrimination statutes, consumer protection laws, data protection laws, and industry-specific requirements that put certain responsibilities on AI developers and deployers, and the breach of which carries a penalty [98], [273], [311].

Contractual Liability: AI provider/deployer/user agreements may apportion responsibilities and liability to the extent of restrictions on contractual disavowal of obligatory law safeguards [308], [312].

7.2 Implementation Issues in AI Governance

The proper governance of AI needs strong enforcement systems in place, but it is complicated by many aspects that make it hard to regulate AI and check its compliance.

7.2.1 Technical Expertise Gap

Regulators usually do not have the technical knowledge to assess challenging AI systems. The knowledge of machine learning, statistics, and software engineering required to determine how a

deep learning model satisfies fairness, transparency, or safety is complex and many government agencies lack that kind of expertise [97], [313].

On natural law, it would be stressed that safeguarding fundamental human goods against technological dangers must be seen through empowering the authorities with the knowledge and means to accomplish the protection. In case regulators do not learn about AI systems, they will not be able to make sure that these systems meet human dignity and rights. Technical capacity-building is not an obligation based on practical sense, but on moral obligation [314], [315].

Positivism approaches aim at filling knowledge gaps by using institutional means: employing technical experts into the agencies, hiring external specialists to do audits and evaluations, enforcing standardized testing and certification by a competent third party and creating regulatory sandboxes to conduct controlled experiments [97].

7.2.2 Jurisdictional Fragmentation

The use of AI systems often involves working with more than one jurisdiction with varying legal needs, which presents compliance risks and regulatory arbitrage opportunities. A multinational AI-focused firm can encounter conflicting requirements, including the EU AI Act, the sector-specific regulation of AI in the U.S., the Chinese rules of governing the algorithms, and the national data protection frameworks [73], [95].

The universalism undertakings of natural law underpin the unified global standards safeguarding the basic human rights irrespective of jurisdiction. Dignity of individuals is not any different across geographical boundaries; simple moral rules ought to be used with AI systems wherever they are used [316].

Positivism recognizes the existence of legal duties which are based on laws and legitimate authority of a particular jurisdiction. Harmonization needs to be an international effort and treaty negotiation, rather than extraterritorial application of the rules of certain jurisdictions automatically. Nonetheless, to enhance fragmentation, positivist analysis finds practical advantages to cooperate in regulation [317], [318].

7.2.3 Pace of Technological Change

AI technologies are changing fast and may surpass regulatory adjustment. The regulations that are written to respond to the existing capabilities are likely to get out of date prior to implementation because of the emergence of new techniques, architectures and applications [22], [166].

Natural law highlights the fact that fundamental moral principles do not alter with the changing technology. The human dignity, the preservation of basic rights, the absence of discrimination, these aspects of human rights have no vision after new AI abilities. Nevertheless, the general principles cannot be applied to new situations without the help of practical wisdom and continuous moral reasoning [79].

Positivism emphasizes the need to update legal frameworks by administering formal procedures whenever circumstances evolve. This generates natural delay between development of technology and regulation. Some of the strategies to deal with this challenge include adaptive governance strategies such as principles-based regulation and regulatory sandboxes and periodic review mechanisms [319], [320].

7.2.4 Opacity and Verification Difficulties

Systems opaqueness and the complexity of auditing complex algorithmic processes makes it difficult to verify compliance with AI regulations. Regulators might not have access to training data, architectures of models or operating systems required to determine whether the requirements are satisfied..

The Natural law would contend that developers and deployers bear a moral responsibility to ensure that there is an audit mechanism in place in AI systems, and to be co-operative with the oversight authorities. The actions of denying the transparency to avoid the accountability contravene the obligations of truthfulness and civic responsibility [75], [301].

Positivist models provide legal requirements of documentation, reporting, and regulatory access, violation of which is punishable. Nonetheless, it is difficult to define any practical systems of gaining access to proprietary systems and safeguarding trade secrets [321].

7.3 International Cooperation and Harmonization

The only way to overcome AI governance problems is through international collaboration to come up with common standards, harmonize enforcement efforts, and avoid regulatory discord.

7.3.1 Rationales for International Cooperation

Shared Values: The world is strongly unified on the main principles of AI governance such as transparency, accountability, fairness, safety, and the protection of human rights despite the cultural and political differences. This normative convergence can be turned into working arrangements through international cooperation [158].

Cross-Border Risks: The AI systems that are used internationally expose risks that go beyond the national borders. There is a need to coordinate in order to tackle harms that cannot be adequately regulated through a single jurisdiction [322].

Regulatory Coherence: Disjointed national regulations put compliance costs on AI developers and open regulatory arbitrage. Unified standards impose less load and offer high protections.[323]

Innovation and Trade: Trade in AI technologies and services using interoperable regulatory frameworks with sufficient protection [324], [325].

7.3.2 Existing International Initiatives

OECD AI Principles: Adopted in 2019, the OECD's Principles on Artificial Intelligence represent the first intergovernmental standard on AI. It has five value-based principles, including inclusive growth, sustainability, well-being, values that are human-centered, fairness, transparency, explainability, robustness, security, safety, and accountability [326], [327].

UNESCO Recommendation on AI Ethics: adopted in 2021 by 193 member states, the recommendation of UNESCO offers a broad ethical framework that focuses on the human rights, dignity, life in peace societies, environmental sustainability and cultural diversity [328], [329].

Council of Europe Framework: The Council of Europe has made a framework on human rights, democracy, and rule of law on AI systems, which is being negotiated as a possible binding convention [330], [331].

UN Initiatives: In 2024, the UN General Assembly created two AI cooperation mechanisms, including an Independent International Scientific Panel on AI and a Global Dialogue on AI Governance [332].

Global Partnership on AI (GPAI): Inaugurated in 2020, GPAI is a consortium of 29 countries and the EU to facilitate responsible AI development through the multistakeholder dialogue and projects on areas of priority[156], [333].

7.3.3 International Cooperation Problems

Geopolitical Competition: The great powers believe AI is strategically significant in economic competitiveness, military capacity and world power. This cooperation is made difficult by competition especially between the United States and China [334].

Incompatible Values: Various societies focus on conflicting values when governing AI. Liberal democracies are those that value the rights and privacy of the individual; China is the state that values social stability and security; developing countries value economic growth and access to technology [335], [336].

Sovereignty issues: Nations oppose global structures limiting policy independence in countries in the form of security and law enforcement practices [337], [338].

Capacity Disparities: Developing nations usually do not have sufficient technical knowledge, regulatory frameworks, and resources to engage in the international AI governance entirely, posing threats of wealthy country-dominated standards [158], [339].

Discussion: Accountability of AI-related harms and good governance systems have proven to be a challenging problem that may need to incorporate both natural law and positivism. Natural law serves to remind us that human agents are the ultimate answer to moral responsibility to create, implement, and use AI systems-these individuals and institutions stand under the duty to discharge due care to safeguard the basic human goods that the technology does not weaken. Positivism offers feasible liability schemes and enforcement procedures by which moral responsibilities are implemented as legal ones. The most promising way of coherent international AI governance is through international cooperation based on a mutual commitment to human rights and dignity without interfering with justifiable sovereignty issues. This potential, however, will need political goodwill, sufficient resources and long term commitment by all parties involved in order to ensure

that human flourishing is considered rather than short term competitive advantages.

8. Innovation, Regulation, and Balancing Competing Values

8.1 The Innovation Imperative

The proponents of AI development highlight how it can revolutionize the world to meet the needs of the global population, empower human beings, and bring about economic prosperity [340]. AI applications are predicted to bring healthcare (disease diagnosis, drug discovery), climatic change mitigation (energy optimization, emissions modeling), scientific research (accelerated discovery, hypothesis generation), education (personalized learning, accessibility), and many other fields [341], [342].

Economically, AI is a general-purpose technology similar to electricity or the internet and can make people more productive, and generate new sectors and jobs and enhance the standard of living around the world. The leadership of AIs is considered to be an essential factor in the competitiveness of countries and companies, and there is a push to develop and implement it faster [343].

8.2 Regulatory Concerns: Precaution vs. Permissionless Innovation

AI development has also been moving fast, which has sparked discussions on how regulatory posture should be, and rival frameworks have placed focus on diverse values.

8.2.1 Precautionary Approach

The precautionary principle, which is conspicuous in European regulatory traditions, is that where activities are threatening of serious harm, deficiency of complete scientific certitude should not be utilized to supply justification to avert preventive actions [344]. In the case of AI, it would imply:

- One should limit potentially dangerous AI systems until they are proven safe instead of rolling them out until they are proven harmful [345], [346].
- Applications that are at high-risk should be strongly pre-market approved and monitored [347], [348].
- In the event that the effects of AI systems on the fundamental rights are not fully clear, the regulators must choose to be cautious [26], [349].

The EU AI Act is precautionary in its approach by risk-based, outlawing unacceptable-risk systems and the far-reaching demands that high-risk applications must satisfy before being put on the market [22], [116].

8.2.2 Permissionless Innovation

Instead, permissionless innovation systems point out that the rate of technological advancement is enhanced when engineers are able to innovate with impunity without any prior permission [350], [351]. Some of the major arguments are:

- Excess regulation suppresses good innovation, halting innovations that could deal with large issues [352].
- Markets and social norms are more adaptive mechanisms than strict rules to deal with emerging problems [353].
- The regulatory intervention must be focused on proven harms and not theoretical risks [353], [354].
- Quality improvements and ethical practices are better motivated by competition than regulations [355], [356].

The regulatory philosophy of the U.S. that is based on the current legislation and sector-oriented guidance instead of the broad-based preemptive regulation is more permissive-oriented [22], [95].

8.3 Natural Law Perspectives on the Innovation-Regulation Balance

The natural law theory is used to give a guideline on whether innovation is to be promoted or restricted depending on the connection between innovation and human flourishing. Figure 8 shows the Innovation-regulation balance framework.

8.3.1 Technology as Instrument of Human Good

Natural law is instrumental: technology is the means to the end of the provision of the main human benefits, and not the final goal. Technology that enhances life, health, knowledge, aesthetic life, practical reasonableness, sociability and harmony with ultimate meaning are worth encouraging [357], [358]. The use of AI in medical diagnosis, education, and support of persons with disabilities, or new scientific

knowledge is consistent with the natural law that emphasizes the flourishing of humans.

Nonetheless, technology that endangers basic commodities or instrumentalizes individuals, that is, making them a means to the end of others, should be limited irrespective of the economic or technical advancement [359]. The natural law would justify a ban on AI systems designed to spy on masses, score social, or exploit manipulatively even when this is profitable or a technical wonder [36].

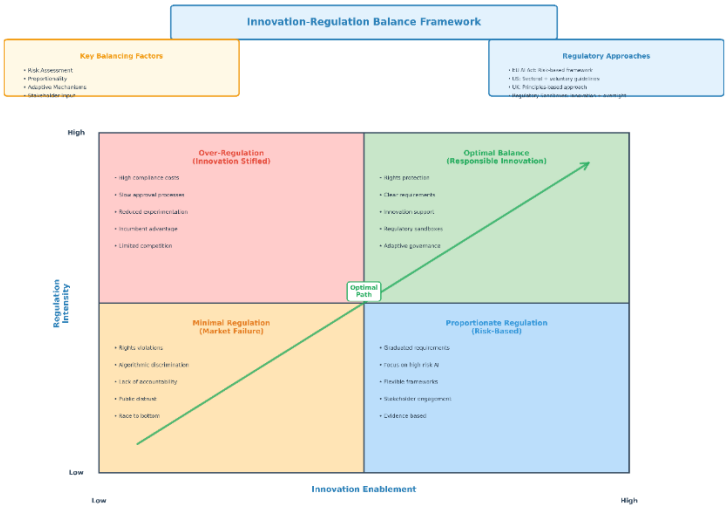

Figure 8 Innovation-Regulation Balance Framework

8.3.2 Principle of Proportionality

The principle of proportionality assumes that the total amount an individual may earn in a job reflects the degree of effort invested, regardless of the job being in a developed or developing country. Principle of Proportionality states that the total amount that an individual can earn in a job should be proportional to the amount of effort he puts in the job irrespective of whether he is working in a developed or developing country [360]. Natural law applies proportionality test in cases where actions are beneficial to some goods in which they pose threats to others. The restrictions on innovations are warranted in the following cases [359].

- **Legitimate Purpose:** Regulations act in the true common goods like defense of basic rights, deterrence of discrimination or security of most people.
- **Increasing Effectiveness:** Restrictions are effective furtherance of their intended purposes.
- **Necessity:** No other less restrictive solutions effectively safeguard pertinent goods.
- **Proportionality Stricto Scilto:** The benefits of regulation are greater than the costs such as forsaken innovation benefits. This framework would suggest that while AI innovation is generally encouraged for its potential to advance human well-being,

- regulatory measures, such as those establishing AI personhood or governing ethical use, are justifiable when they proportionately safeguard fundamental human rights and societal values against potential detriments [361] [362] [12].

Natural law applied to AI governance would underpin risk-based methods of setting the intensity of regulation according to the level of harm that could be caused and letting beneficial innovation be generated. Unconditional bans should apply to those AI applications that pose a risk to the basic human dignity and rights. Where risks to the application are manageable, the requirements must be corresponding-neither too high nor too low to give

reasonable assurance of safety and protection of rights, but not to undue impediment of useful development [296], [363].

8.3.3 Common Good and Distributive Justice

The common good focus of natural law obliges that technological gains be distributed equally instead of being concentrated in the hands of the elite at the expense of the poor groups. There should be AI governance that makes sure:

- The advantages of AI applications cut across all layers of society and not just affluent consumers or advanced nations [158].
- Algorithms do not impose risks and burdens disproportionally within the marginized communities by algorithmic discrimination or exploitation of the latter [152].
- The priorities of development are the real needs of people and not just profitable application [130].
- It is not only the technology giant companies that can be involved in AI innovation, but smaller developers and developing country entities can also be involved [339].

Such a natural law would criticize over-regulation, which entrenches the benefits of the incumbents by increasing entry barriers, as well as under-regulation, which permits the strong actors to internalize the harms upon the vulnerable population.

8.4 Positivist Analysis of Innovation-Regulation Tradeoffs

Legal positivism deals with the questions of innovation-regulation by methods of analysis of institutes competences, procedural legitimacy and empirical evaluation of the regulatory consequences.

8.4.1 Institutional Competence and Regulatory Design

Positivist analysis acknowledges the existence of comparative advantages of various institutional arrangements to govern technological change:

Legislation: democratic legislatures can provide high-level structures based on societal values, but tend to be technologically inept and slow compared to technology [95].

Agency Regulation: Specialist agencies are better informed technically, and can respond more quickly to changes than to legislation, yet they have accountability issues, and may become captives of industries regulated [364].

Judicial Oversight: In the case law approach, the courts may use the established principles of the law to the new technologies by adjudicating on individual cases, creating legal doctrine bit by bit. Nonetheless, the resolution of cases in courts is reactive and it only responds to cases presented before courts [365].

Self-Regulation: The adaptive governance of professionally skilled individuals is made possible by industry standards, professional norms and corporate policies. Nonetheless, self-regulation can fail to secure the interests of the people in situations where the motives of the partners are not aligned [366].

To have a successful AI governance, such institutional mechanisms must be combined: legislation provides general frameworks and banned behavior; agencies create technical requirements and supervising behavior; courts render dispute decisions and interpret the law; industry principles add to formal regulation with best practices.

8.4.2 Regulatory Sandboxes and Adaptive Governance

More and more positivist frameworks use regulatory sandboxes controlled environments in which innovators can experiment with new applications under regulatory oversight with temporary exemption of some conditions [367], [368]. Sandboxes can be used on several occasions:

- **Experimental Learning:** Regulators develop an insight into the risks and benefits of new technologies, based on real-world implementations [369].
- **Adaptive Rule-Making:** Sandbox experiments provide evidence that helps in the formation of proper permanent regulations [367], [368].
- **Lowering Obstacles:** Smaller businesses and startups are able to be innovative without the complete compliance costs of existing requirements possibly being a poor fit with new usage [367], [370].
- **Risk Reduction:** in-Vitro testing allows experimentation and restricts the possible damages to larger groups [367].

The EU AI Act has sandbox provisions that allow high-risk systems to be tested under a regulatory supervision. There are various examples of AI regulatory sandboxes created or suggested by multiple countries as the balancing tool between innovation and protection [22], [371].

8.4.3 Cost-Benefit Analysis and Empirical Assessment

Positivist methods focus on the empirical evaluation of regulatory outcomes. To policymakers, it would be worth considering:

Compliance Costs Direct Costs: Documentation, testing, audits, legal review: direct costs of anonymous compliance Costs Indirect Costs: Delays, less experimentation, other forgone innovation [372], [373].

Enforcement Costs: Resources that regulators need to check compliance, enumerate violations and resolve disputes. This empirical assessment of costs and benefits is crucial for optimizing regulatory frameworks, ensuring they are neither overly burdensome nor insufficiently protective, thereby fostering innovation while safeguarding societal interests [374], [375].

Benefits: The benefits are quantifiable and qualitative including the harms prevented, increased trust, and adoption, minimal risks associated with liabilities, and equal competition playing fields [296].

Distributional Effects: Distribution of costs and benefits among various actors (large vs. small companies, incumbent vs. new entrants, developed vs. developing countries) [376].

Dynamic Effects: Long term functions on the path of innovation, market structure and technological abilities [377].

Empirical research that is systematic can guide judgments as to whether or not certain regulations are appropriate in terms of balance. Nevertheless, positivists recognize that not all the pertinent values can be readily measured, and ultimate policy decisions cannot be made based on the cost-benefit calculations exclusively, but by democratic procedures.

8.5 Synthesis: Responsible Innovation Frameworks

The idea of responsible innovation has been suggested in the literature, and is a set of frameworks whereby ethical development issues are combined with advocacy of positive technological development [378], [379].

Anticipation: Know how AI systems might have a positive and negative impact before implementation [380].

Reflexivity: Attentively scrutinize the presumptions, values, and possible blind spots in the processes of AI development [11].

Inclusion: Add local stakeholders viewpoints and especially those communities that might be compromised by AI systems [381].

Responsiveness: Change AI systems and forms of governance premised on feedback, evidence of effects, and changing values in society [382].

Such principles are consistent with both natural law focuses on practical wisdom and the need to flourish as a human, and positivist stances on adaptive and evidence-based governance [383].

Discussion: This dilemma of encouraging a positive AI innovation and mitigating the risk of potential harm necessitates finer measures that do not lead to stagnation by over-caution and carelessness by under-vigilance. The natural law has substantive guidelines on ethics: AI development should benefit real human goods, should avoid harming fundamental dignity and rights, and should encourage equitably shared prosperity and not concentrate on personal gains. Positivism provides procedural refinement: the integration of a variety of institutional arrangements, the application of adaptive governance instruments, such as regulatory sandboxes, and the systematic evaluation of the regulatory performance. Its ambition should be to facilitate responsible innovation the AI creation with a perspective on human dignity and facilitated by effective governance frameworks that promote but not hinder useful implementation. This will necessitate continuous discussion between technologists, ethicists, policymakers and the communities impacted so as to balance the unavoidable tradeoffs and change strategies as technologies, and social knowledge change.

9. Recommendations for Balanced AI Governance

Considering the discussion in the previous sections, we would offer AI governance frameworks which would integrate the findings of both the natural law and legal positivist point of view and solutions to the practical implementation issues. Figure 9 shows the integrated AI governance ecosystem.

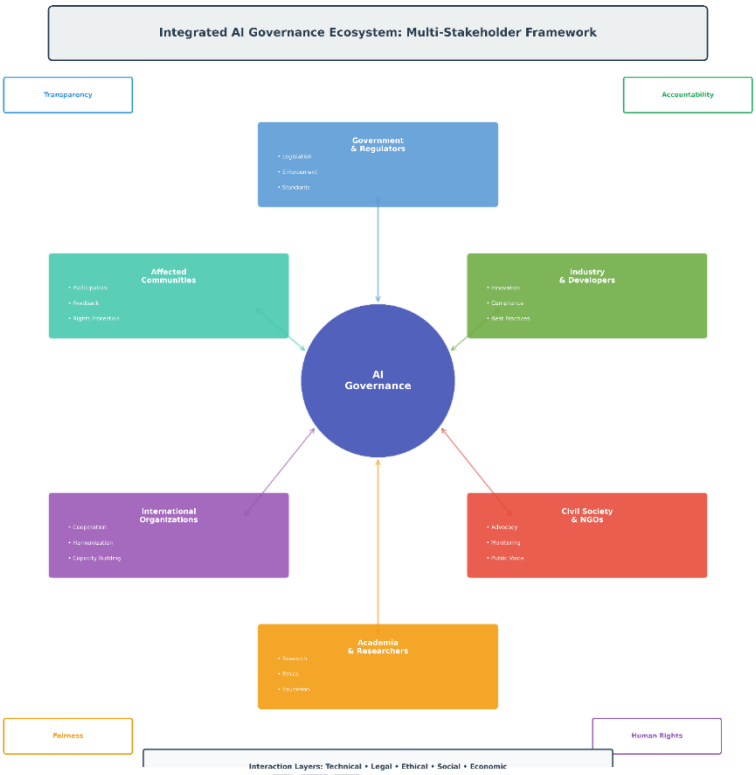

Figure 9 Integrated AI Governance Ecosystem

9.1 Philosophical Foundations: Integrating Natural Law and Positivist Insights

Recommendation 1: Ground AI Governance in Human Rights and Dignity

Regulatory frameworks on AI must explicitly be based on the protection of basic human rights and dignity as it is highlighted on natural law theory. This requires:

- The declaration by the constitution or legal authoritative that AI systems should consider human dignity, basic rights and democracy.
- Outlawing AI applications that constitute the infringement of human decency irrespective of technical detail or economic gain.
- Basic rights have an effect on the assessment since they are compulsory issues of high risk systems.
- Principles of interpretation that would guide the solution of ambiguities in AI rules to favor protection of human rights.

Recommendation 2: Establish Clear Legal Structures and Procedural Mechanisms

Although the AI governance is based on substantive ideas of ethics, it has to be operationalized by using the clarity of law and procedural tools as stressed by legal positivism:

- Specific description of such terms as AI system, high-risk application, algorithmic discrimination, and meaningful human oversight.
- Comprehensive developer, deployer and user obligations specifications with graduated specifications according to the risk levels.
- Open conformity assessments, certification, regulatory and market supervision procedures.
- Easy-to-use tools to enable the complaints, remedies, and challenging algorithmic decisions by the affected individuals.

9.2 Risk-Based Approaches: Calibrating Regulatory Intensity

Recommendation 3: Adopt Graduated Risk-Based Frameworks

Artificial intelligence governance must use risk-focused techniques to scale regulatory intensity to harms with potential while preventing over regulation of low-risk applications and under regulation of risky systems:

- **Banned Systems:** AI Apps that pose an inherent danger to fundamental rights (social scoring, manipulative systems, indiscriminate surveillance) are prohibited in all cases.
- **High-Risk Systems:** Requirement such as risk management systems, data governance, technical documentation, transparency, human oversight, and fundamental rights have an influence on assessments.
- **Medium-Risk Systems:** The obligation to disclose information, a requirement to notify users, and the voluntary inclusion of best practices.
- **Low-Risk Systems:** Light regulation, which permits innovation by having post-market surveillance in case of emerging issues.

Recommendation 4: Establish Dynamic Classification Mechanisms

The classifications of risk should be dynamic because technologies are dynamic and knowledge of effects is dynamic:

- Evidence-based periodic review processes of high-risk categories.
- Systems reclassification mechanisms in case of changes in capabilities, use situations, or risk profiles.
- Sandboxes of regulations that provide controlled testing of new applications to determine the right risk categories.

9.3 Transparency, Explainability, and Accountability

Recommendation 5: Implement Multi-Level Transparency Frameworks

The level of transparency requirements must be adjusted to the needs and functions of different stakeholders:

- **Public Transparency:** Public Transparency: Public knowledge about the existence, purpose, and opportunities of the AI systems available to potentially affected parties.
- **User Transparency:** Comprehensive data to allow users to know what to do with it, what they cannot, and how to interpret the results.
- **Regulatory Transparency:** Full technical documentation, testing and access to the system to its oversight authorities.
- **Research Transparency:** Information and records that allow other researchers to determine fairness, accuracy and risks.

Recommendation 6: Mandate Meaningful Explainability for High-Stakes Decisions

AI systems that generate or cause substantial consequences to individuals by making consequential decisions will have to be meaningfully explanatory:

- Right to explanation where automated decisions have legal or other similarly important effects on individuals.
- Explanations should not only be technical accounts of things, but be understandable to affected persons.
- Human review is a possibility to appeal against algorithmic decisions.
- Onus on AI deployers to prove that systems generate a non-discriminatory result.

Recommendation 7: Establish Algorithmic Auditing Requirements

AI systems must be checked by independent auditing, to ensure their adherence to the legal aspects and ethical standards:

- High-risk system before deployment and periodically after deployment Mandatory third-party auditing.
- Unified procedures of determining fairness, accuracy, robustness, security, and basic rights impacts.
- Transparency (protecting trade secrets) Public report of audit findings to facilitate informed assessment.
- Auditing Regulatory Authority to schedule audits in case of doubt about the compliance with the system.

9.4 Anti-Discrimination and Fairness

It ensures that every employee is treated fairly and without discrimination based on gender, race, ethnicity, religion, or disability (Stewart, 2000, pp.95).

Recommendation 8: Strengthen Anti-Discrimination Protections for Algorithmic Systems

Laws should also be changed to accommodate the problems of AI-related discrimination:

- Clear legislative directives against algorithmic discrimination on the ground of the guarded traits.
- Discrimination based on disparate impact without zero or reasonable justification and demonstrating the absence of less discriminative alternative.
- Guard against discrimination founded on algorithmic classifications that are associated with secured qualities although they need not be founded on them.
- Frameworks in which plaintiffs prove disparate impact and defendants have to prove business necessity and have to seek alternatives.

Recommendation 9: Mandate Bias Testing and Mitigation

The developers and deployers should undertake active actions in order to avoid discriminatory results:

- Bias and discriminatory effects of AI systems on the demographic groups Pre-deployment testing of AI systems.
- Testing procedures, testing results, mitigation measures.
- Continuous process of surveillance on deployed systems relating to the occurrence of emergent bias or discriminatory trends.
- Burden of abandoning or altering systems which show continued discriminatory performances.

9.5 Human Oversight and Meaningful Human Control

Recommendation 10: Ensure Meaningful Human Oversight of High-Risk AI

- The AI systems should be developed in a way that allows and maintains human agency, judgment, and accountability:
- Human-in-the-loop Consequential decisions by AI systems that affect individuals.
- Instructing and empowering human oversight to properly comprehend, assess, and intervene in AI activities.
- Definite human responsibility on AI system deployment and operation.
- The capability to override recommendations made by the algorithm with human operators rejecting or reversing their suggestions.
- Responsibility systems in which humans rather than algorithms are ultimately responsible to make decisions.

9.6 International Cooperation and Harmonization

Recommendation 11: Develop Globally Harmonized Standards

International collaboration must strive to achieve standardized AI governance principles and to be respectful of lawful national differences:

- Regulatory coherence is achieved through common definitions of key terms (core AI system, high-risk application, algorithmic bias).
- Mutual recognition arrangements of conformity assessment and certification.
- Sharing of information between regulatory bodies regarding AI risks and incidents, as well as optimal practices.
- International cooperation in enforcing the basic right against AI systems against all jurisdictions.

Recommendation 12: Establish International AI Governance Institution

The global AI governance work should be organized by a special international institution:

- Multistakeholder governance comprising of governments, industry, civil society, academia and the affected communities.
- Technical advisory groups offering experience in new AI capabilities and dangers.
- International coordination of national regulatory strategies to limit fragmentation and regulatory arbitrage.
- Capacity-building to help developing countries engage effectively in AI governance.
- Tracking and publicizing the trend of AI development, effects, and the efficacy of AI regulation across the world.

9.7 Adaptive Governance and Continuous Improvement

Recommendation 13: Implement Adaptive Regulatory Mechanisms

The AI governance models should be built to support the process of constant learning and change:

- Review and Sunset Costs Assessments of regulations on a periodic basis.
- Sandboxes and pilot programs where regulation can be tested on new applications and forms of governance.
- Requirements on the collection of evidence on analyzing the high effectiveness of regulators and their unintended consequences.
- Rapid response processes to cover risks that are urgent or learn with incident.

Recommendation 14: Foster Multi-Stakeholder Dialogue

To make sure that AI governance is effective, there should be continuous consultations between various stakeholders with different experience and opinions:

- Consultations involving communities, civil society, industry, technical and even ethicists.
- Public involvement in rule-making and allow the citizens to be involved in the development of governance.
- Publicity on the regulatory decisional making and rationale.
- Systems to instill results of frontline deployment experience into systems of governance.

9.8 Resources and Capacity Building

Recommendation 15: Invest in Regulatory Capacity

The effective AI governance demands sufficient resources and technical skills:

- Investment in regulatory agencies to employ AI experts, data scientists, and even ethicists.
- Education and training of regulators, judges, lawyers, and administrators about AI technologies and regulatory governance.
- Technical infrastructure to allow regulatory bodies to audit and evaluate AI systems.
- Global collaboration in the field of capacity-building especially to developing nations.

Discussion: These are recommendations of synthesis of both natural law, substantive ethical commitments and procedural sophistication of positivism. It is important that effective AI governance is based on both: the basis in fundamental human rights and dignity (natural law) and operation through effective legal frameworks, risk-based approaches, transparency conditions, anti-discrimination provisions, human accountability requirements, international collaboration, adaptive systems, stakeholder involvement, and sufficient resources (positivism). AI governance needs no single jurisdiction or institution to handle all those issues; the success can be achieved only through the joint international efforts taking into consideration various values without undermining universal human rights. Most importantly, AI governance should ensure that human flourishing is at the forefront- technology should be used to serve humanity and not to control humanity. This requires continued investment by the entire stakeholder community to focus on ethics and innovation and make sure that the transformative power of AI can be harnessed to the benefit of all humans with the critical protection of inherent dignity and rights.

10. Conclusion and Future Directions

10.1 Summary of Key Findings

This assortment review has scrutinized artificial intelligence regulation by investigating the concept of natural law theory and legal positivism as part of the dual-facet approach to this issue, exploring how these two opposing traditions in jurisprudence shape modern systems of artificial intelligence regulation. Some of the main points that we can conclude are:

Philosophical Foundations: Natural law and legal positivism are two extremely different yet possibly complementary approaches to thinking about AI regulation. Governance is an activity based on natural law principles that are universal moral standards that safeguard innate human dignity, basic rights and conditions of thriving. Legal positivism further states that good regulation is one that has clear legal structures, procedural mechanisms, and formal institutional authority. Each of the two sides has something to offer: natural law includes substantive ethical commitments, which are likely to be used in the development of AI; positivism provides more practical methods of making these commitments operationalizable in legal terms and enforceable.

Risk-Based Regulation: The risk-based framework of the EU AI Act is an endeavor to find the balance between the imperative to innovate and the basic rights protection in the true spirit of both schools of jurisprudence. The ban on AI systems that represent

unacceptable risks is the expression of the understanding of natural law that some applications are intrinsically against human dignity. Examples of positivist methodology of defining definite, formalized categories that have particular legal implications include graduated regulatory requirements given comparative risk levels. Risk-based frameworks rely on the accuracy of their definition, the correct classification processes, the resources necessary to enforce them, and continuous adaptability as technology and knowledge continue to change.

Fundamental Rights Protection: AI governance has to be substantively committed to human rights that have to be protected through dignity and procedures. Natural law states that AI systems should address the rights to privacy, non-discrimination, due process, and autonomy inherent to human nature irrespective of technical ability and economic gain. Positivism offers the means of operationalizing the protection of rights in the form of basic rights impact assessment, transparency needs, human oversight needs, and remedies availability. The protection of rights should be both philosophical: the moral commitment will guarantee the actual respect of rights but not their formal way; the procedural structures will allow observing, enforcing and holding accountable.

Algorithmic Bias and Discrimination: The fact that the AI systems are prone to reproduce and reinforce the discriminatory patterns is a mortal threat to both natural and positivist models. The natural law denounces algorithmic discrimination as being against the basic human equality and dignity. Positivism tackles the issue of discrimination by using anti-discrimination laws, yet the conventional framework cannot handle the AI-specific issues such as proxy discrimination, nascent bias, and new algorithmic groupings. The problem of algorithmic discrimination can be effectively prevented with technical solutions (bias testing, fairness constraints, auditing), legal regulations (disparate impact liability, burden-shifting), and cultural change (educating developers, organizational accountability, participation of affected community).

Transparency and Accountability: Transparency and explainability are vital accountability roles, which have technical and practical constraints. The requirements of transparency based on natural law rely upon respect of agency and autonomy of humans-people need to know about the AI systems which impact them in order to make correct decisions and preserve dignity. Positivism defines transparency by graded disclosure, trade-offs between control requirements and privacy, security, and proprietary issues. They both recognize the fact that meaningful transparency involves more than formal compliance as information should be understandable, relevant and actionable by the target audiences.

Liability and Enforcement: The process of attributing liability to AI-related harms and enforcing governance obligations will be very challenging and will entail the institutional ability, global collaboration, and dynamic processes. Natural law asserts that it is human actors who bear moral responsibility regardless of the sophistication that AI may have because developers, implementers and users are still responsible of the predictable outcomes of their decisions. Positivism offers a liability (product liability, negligence, statutory violations) and enforcement system (regulatory oversight, auditing, penalties). To be better accountable, it is necessary to combat expertise gaps, jurisdictional fragmentation, technological uncertainty, and acceleration by building capacity, collaborating internationally, adopting adaptive governance, and using algorithmic auditing.

Innovation and Regulation Balance: To achieve a compromise between positive innovation of AI and sufficient protection against its negative effects, there must be delicate measures to prevent overprotective and laissez-faire solutions. To promote the actual human good, natural law proposes that technology must be applied in a manner that does not harm the fundamental dignity and rights. Regulation based on proportionality analysis justifies protection of important goods based on effective means without unjustified restriction to important development. Positivism provides institutional sophistication in the form of regulatory sandboxes, adaptive governance, multi-stakeholder engagement and systematic evaluation of regulatory impacts. Responsible innovation models seek to incorporate ethical engagements with viable assistance of advantageous applications.

10.2 Theoretical Contributions

This review contributes to current discussions in the field of AI governance in a number of ways:

Jurisprudential Analysis: We offer a systematic analysis of how the fundamental legal theories, that is, natural law and positivism, might be applied to new governance challenges involving AI and elucidate the philosophical presuppositions behind various methods of regulation and provide insight into conflicts between moral values and procedural needs.

Integrated Framework: We do not only show the incompatibility of natural law and positivism as alternated options; we show how these two can complement each other. It is necessary to have both substantive ethical commitments (natural law) and practical implementation mechanisms (positivism) to good AI governance. None of these frameworks can give full guidance; they are all better integrated to give stronger foundations.

Comparative Institutional Analysis: We explore how various institutional structures (legislation, agency regulation, judicial oversight, self-regulation, international cooperation) are able to handle an AI governance problem, based on either of both philosophical traditions to apply institutional competences and constraints to this matter.

Policy Advice: Our review offers normative advice to policymakers, developers, and other stakeholders who have to navigate the dilemma of AI governance. The recommendations are a synthesis between theory and practice, and they seek to achieve balanced models of protecting core values, but allowing positive innovation to take place.

10.3 Practical Implications

The given analysis has significant practical consequences to the stakeholders of the AI governance:

Policymakers: Should base regulatory frames on the clear statement of values to be upheld (human dignity, fundamental rights, equality, autonomy) and should operationalize these commitments by specifying these commitments via legal mandates, graded requirements, transparency requirements, structures of accountability and enforcement. Risk-based methods have a great potential but must be defined, classified, provide enough resources, and adjusted.

AI Developers: ought to be true ethical responsibilities rather than compliance liabilities. The natural law highlights that developers are responsible towards the foreseeable impacts of their technical decisions in terms of data, algorithms, testing and deployment situations. Legal requirements are the minimum but not all of the ethical demands to respect human dignity and rights.

Deployers and Users: The deployers and users are required to be careful when choosing, implementing, and managing AI systems. The need of human oversight indicates moral (preservation of the human moral agency and accountability) as well as legal (safety and fairness of system operation) requirements. It is a requirement to understand the capabilities, limitations and possible and possible impacts of systems in order to be responsibly deployed.

Civil Society: The civil society plays important role in promoting good governance in AI, compliance, voice amplification of the affected communities and holding regulators and industry to account. A good governance involves continuous questioning and involvement of more than technical professionals and industry players.

International Organizations: This should promote coordination and collaboration to set harmonized standards, exchange information, capacity building and enforcement. The issues of AI governance are not limited either to the national borders or to the fragmentation of approaches; this results in gaps, inconsistencies, and loopholes to evasion. The international institutions need to strike balance between the universal human rights obligations and the need to respect the national legitimate variations

10.4 Future Research Directions

There are numerous directions that should be explored:

Empirical Assessment: It is necessary to undertake systematic empirical research of real impacts of AI regulations. What is the impact that various governance strategies have on innovation levels, market structure, protection of rights, minimization of discrimination and other outcomes? What are the unwanted after effects? The regulatory frameworks should be constantly refined based on evidence.

Technical-Legal Integration: There is need to conduct further research on the conditions of

translating legal requirements to technical specifications and other way round. Which measures of fairness are most appropriate to the principles of legal non-discrimination? What can be done to ensure explainability methods can meet the due process? Which testing procedures are sufficient in evaluating basic rights impacts? It is essential to bridge the area between legal and technical.

Comparative Regulatory Studies: As various jurisdictions adopt diverse AI governance models, comparison can help to discover what works and what faces many problems and what is best adapted to a particular context. Regulatory experiments across the world could be learned to inform best practice and make harmonization easier.

New Technologies: Artificial intelligence is evolving at a fast pace. Generative AI, large language models, artificial general intelligence, and other developments in the future will create new governance issues. It is important to have anticipatory research on what may happen and what regulatory measures should be taken.

Global Justice Dimensions: The effects of AI governance on global justice and development should be discussed more thoroughly. What can structures do to bring benefits to developing nations and disadvantaged groups? What do we need to do in terms of capacity-building in order to have meaningful involvement in global governance? What is the way to treat intellectual property and technology transfer?

Constitutional and Democratic Theory: The connection between the regulation of AI and basic constitutional principles (separation of powers, democratic accountability, judicial review) needs to be investigated. What is the way to roll out AI systems in the government? What are the constitutional limitations? What is the way to have democratic control over more and more self-governing systems?

10.5 Conclusion

Artificial intelligence regulation poses deep problems at the intersection of the law, ethics, governance, and technology. With the ever-growing influence of AI systems on consequential matters in human life, such as employment, credit, healthcare, education, criminal justice, it has become a pressing necessity to make sure the technologies are benefiting human dignity, rights, and thriving instead of eroding them.

This is the case that has been shown in this review in that natural law theory and legal positivism can contribute valuable and complementary information to AI governance. Natural law is a reminder that numerous moral standards need to be applied in technological creation: there should be upholding of human dignity, preservation of basic rights, devotion to equality and justice, that the common good should be encouraged. These principles are not a whim or a tradition but some profound truths concerning the nature of human beings and the conditions of prosperity. Artificial intelligence regulation based on such principles acknowledges that technology should not be telling humans what to value the most, but instead, it should serve the supreme values of humanity.

Legal positivism adds much-needed practical wisdom: moral principles have to be brought to bear on specific legal systems, a system of procedures, enforcement abilities and institutionalization. Good governance involves accurate definitions, progressive duties, openness standards, accountability arrangements, corrective procedures and adaptive procedures. The absence of such workable frames makes moral commitments only aspirational instead of actualised.

The EU AI Act and newly developed governance frameworks across the world are indicative of efforts to unite these points of view, that is to harmonize substantive devotion to human rights with procedural elegance in regulation modeling. It will require long-term efforts on the part of all stakeholders: policymakers should not yield to the pressure of the industry to not properly regulate and activists to not demand highly unrealistic restrictions; developers should take ethical responsibility as obligations rather than as compliance costs; deployers and users should be careful and have meaningful human oversight; civil society should play an accountability role and be able to amplify the voice of affected communities; international organizations should facilitate harmonization and cooperation.

In the most basic sense, AI governance involves keeping in mind what really counts which is the dignity of human persons, their rights and their thriving. Technology has got the colossal potential of improving or degrading human life. The decisions to be taken on AI development and regulation will define the society generations to come. These

decisions should be based on wisdom, the substantive wisdom of natural law to identify basic moral truths, and procedural wisdom of positivism to establish beneficial institutional forms. Such integration is the only way that AI governance can live to its fullest in terms of guaranteeing that these revolutionary technologies can be used to the highest good of humanity, but not to the lowest instincts of its most primordial cravings.

It requires cognitive humility: both philosophers and technologists, regulators and industry, do not have full knowledge of the consequences of AI or ideal solutions to the issue of governance. Product development necessitates continuous communication, trial and error, experience, and the readiness to change with increased knowledge. However, humility does not require paralysis. Basic values or core commitments such as human dignity, basic rights, equality, justice, transparency, and accountability form strong pillars in dealing with uncertainty. Sticking to these values and remaining flexible in the use of means can enable the flourishing of people in the era of intelligent machines through AI governance.

The antecedents could not be greater. AI will either empower human beings, open opportunities, solve global problems, and enrich life or will be concentrated, increase inequalities, weaken autonomy, and jeopardize fundamental rights. The distinction is in decisions that are being made today on the way to create and manage these technologies. This review has attempted to shed light on philosophical grounds, real-world dilemmas and potential solutions towards such choices being made in a wise manner. It is yet to be written, whether AI will be realized or not.

REFERENCES

[1] E. Zaidan and I. A. Ibrahim, "AI Governance in a Complex and Rapidly Changing Regulatory Landscape: A Global Perspective," Humanities and Social Sciences Communications , vol. 11, no. 1, Sep. 2024, doi: 10.1057/s41599-024-03560-x.

[2] O. J. Erdélyi and J. Goldsmith, "Regulating artificial intelligence: Proposal for a global solution," Government Information Quarterly , vol. 39, no. 4, p. 101748, Jul. 2022, doi: 10.1016/j.giq.2022.101748.

[3] J. G. C. Ramírez and Md. M. Islam, "Utilizing Artificial Intelligence in Real-World Applications," Deleted Journal , vol. 2, no. 1, p. 14, Feb. 2024, doi: 10.60087/jaigs.v2i1.p19.

[4] Y. K. Dwivedi et al. , "Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy," International Journal of Information Management , vol. 57, p. 101994, Aug. 2019, doi: 10.1016/j.ijinfomgt.2019.08.002.

[5] S. Raza et al. , "Who is Responsible? The Data, Models, Users or Regulations? Responsible Generative AI for a Sustainable Future," Jan. 2025, doi: 10.36227/techrxiv.173834932.29831105/v1.

[6] F. R. N. Alfiani and F. Santiago, "A Comparative Analysis of Artificial Intelligence Regulatory Law in Asia, Europe, and America," SHS Web of Conferences , vol. 204, p. 7006, Jan. 2024, doi: 10.1051/shsconf/202420407006.

[7] Y. Bengio et al. , "Managing extreme AI risks amid rapid progress," Science , vol. 384, no. 6698, p. 842, May 2024, doi: 10.1126/science.adn0117.

[8] W. Wu and S. Liu, "Dilemma of the Artificial Intelligence Regulatory Landscape," Communications of the ACM , vol. 66, no. 9, p. 28, Aug. 2023, doi: 10.1145/3584665.

[9] M. R. Islam and A. T. Wasi, "Balancing Power and Ethics: A Framework for Addressing Human Rights Concerns in Military AI," in Communications in computer and information science , Springer Science+Business Media, 2025, p. 13. doi: 10.1007/978-3-031-98949-0_2.

[10] V. Maphosa, "The Rise of Artificial Intelligence and Emerging Ethical and Social Concerns," AI Computer Science and Robotics Technology , vol. 3, Dec. 2024, doi: 10.5772/acrt.20240020.

[11] D. S. Allen et al. , "A roadmap for governing AI: technology governance and power-sharing liberalism," AI and Ethics , Jan. 2025, doi: 10.1007/s43681-024-00635-y.

[12] Y.-H. Weng and T. Izumo, "Natural Law and its Implications for AI Governance," Delphi - Interdisciplinary Review of Emerging Technologies , vol. 2, no. 3, p. 122, Jan. 2019, doi: 10.21552/delphi/2019/3/5.

[13] A. Juškevičiūtė-Vilienė, "Legal Positivism, AI, and the Modern Legal Landscape: Challenges in Education, Research, and Practice," Acta Universitatis Lodziensis Folia Iuridica , vol. 109, p. 25, Dec. 2024, doi: 10.18778/0208-6069.109.02.

[14] U. Madagwa, "The Convergence and Conflict Between Legal Positivism and the Moral Principles of Natural Law," SSRN Electronic Journal , Jan. 2024, doi: 10.2139/ssrn.4788603.

[15] J. Finnis, "Natural Law Theory: Its Past and its Present," The American Journal of Jurisprudence , vol. 57, no. 1, p. 81, Jan. 2012, doi: 10.1093/ajj/57.1.81.

[16] P. Soper, "In Defense of Classical Natural Law in Legal Theory: Why Unjust Law is No Law at All," Canadian Journal of Law & Jurisprudence , vol. 20, no. 1, p. 201, Jan. 2007, doi: 10.1017/s0841820900005750.

[17] M. Özcan, "Ethics and Law: Reassessment on the Legal Positivism," Philosophy study , vol. 4, no. 2, Feb. 2014, doi: 10.17265/2159-5313/2014.02.001.

[18] H. L. A. Hart, "Positivism and the Separation of Law and Morals," Harvard Law Review , vol. 71, no. 4, p. 593, Feb. 1958, doi: 10.2307/1338225.

[19] A. Jedličková, "Ethical approaches in designing autonomous and intelligent systems: a comprehensive survey towards responsible development," AI & Society , Aug. 2024, doi: 10.1007/s00146-024-02040-9.

[20] S. Alpers, B.-J. Krings, C. Becker, M. Rill, and M. Weinreuter, "Ethical, legal and social implications (ELSI) of the design of AI systems: Systematization of the consideration through procedure models and guidelines," Zeitschrift für Arbeitswissenschaft , vol. 78, no. 2, p. 160, May 2024, doi: 10.1007/s41449-024-00422-9.

[21] H. Zech, "Liability for AI: public policy considerations," ERA Forum , vol. 22, no. 1, p. 147, Jan. 2021, doi: 10.1007/s12027-020-00648-0.

[22] T. Minssen, H. J. Smith, S. P. Mann, A. M. García, and J. Negrete, "THE EUROPEAN AI ACT : PROMISES, PERILS & TRADE-OFFS IN A COMPLEX REGULATORY ECOSYSTEM," Research Portal Denmark , vol. 1, p. 330, Jan. 2025, Accessed: Jul. 2025. [Online]. Available: https://local.forskningsportal.dk/local/dki-cgi/ws/cris-link?src=ku&id=ku-2fbecc26-4adb-487a-92d7-469ae5a7b774&ti=THE%20EUROPEAN%20AI%20ACT%20%3A%20PROMISES%2C%20PERILS%20%26amp%3B%20TRADE-OFFS%20IN%20A%20COMPLEX%20REGULATORY%20ECOSYSTEM

[23] D. Canavese, A. Ferreira, R. Laborde, and A. Benzekri, "Artificial Intelligence Systems in the European Union: Guidelines and Architectures for Compliance-by-Design," Jan. 2024, doi: 10.2139/ssrn.5037706.

[24] C. Novelli, F. Casolari, A. Rotolo, M. Taddeo, and L. Floridi, "Taking AI risks seriously: a new assessment model for the AI Act," AI & Society , vol. 39, no. 5, p. 2493, Jul. 2023, doi: 10.1007/s00146-023-01723-z.

[25] A. A. Gikay, P. L. Lau, C. Sengul, A. Miron, and B. Malin, "High-Risk Artificial Intelligence Systems under the European Union's Artificial Intelligence Act: Systemic Flaws and Practical Challenges," SSRN Electronic Journal , Jan. 2023, doi: 10.2139/ssrn.4621605.

[26] W. Wu and S. Liu, "A Comprehensive Review and Systematic Analysis of Artificial Intelligence Regulation Policies," arXiv (Cornell University) . Cornell University, Jan. 01, 2023. doi: 10.48550/arXiv.2307.12218.

[27] O. Shumilo and T. Kerikmäe, "The European Approach to Building AI Policy and Governance: A Haven for Bureaucrats or Innovators?," IDP Revista de Internet Derecho y Política , no. 34, p. 1, Dec. 2021, doi: 10.7238/idp.v0i34.387744.

[28] A. Rotolo, B. Ferrigno, J. M. A. G. Godinez, C. Novelli, and G. Sartor, "Foundations for Risk Assessment of AI in Protecting Fundamental Rights," arXiv (Cornell University) , Jul. 2025, doi: 10.48550/arxiv.2507.18290.

[29] T. Keber, R. Schwartmann, and K. Zenner, "The EU AI Act: A Practice-Oriented Interpretation," Computer Law Review International , vol. 25, no. 4, p. 114, Aug. 2024, doi: 10.9785/cri-2024-250403.

[30] L. Lawniczak and C. Benzmüller, "Logical Modalities within the European AI Act: An Analysis," arXiv (Cornell University) , Jan. 2025, doi: 10.48550/arxiv.2501.19112.

[31] S. Acharya, "Natural School; Jurisprudence of Law of Nature," SSRN Electronic Journal , Jan. 2019, doi: 10.2139/ssrn.3458069.

[32] M. L. Stackhouse, "Natural Law Theory and its Relatives," Liberty University law review , vol. 2, no. 3, p. 13, Jan. 2008, Accessed: Jan. 2025. [Online]. Available: https://digitalcommons.liberty.edu/cgi/viewcontent.cgi?article=1045&context=lu_law_review

[33] Y. K. Yonjan, "An Overview on Legal Philosophy of Natural Law: Ancient to Revival Period," SSRN Electronic Journal , Jan. 2023, doi: 10.2139/ssrn.4473274.

[34] M. S. Shellens, "Aristotle on Natural Law," The American Journal of Jurisprudence , vol. 4, no. 1, p. 72, Jan. 1959, doi: 10.1093/ajj/4.1.72.

[35] G. Duke, "Aristotle and Natural Law," The Review of Politics , vol. 82, no. 1, p. 1, Dec. 2019, doi: 10.1017/s0034670519000743.

[36] X. Zhang and M. Yu, "The Labor Theory of Value in the Context of Artificial Intelligence: A Historical and Classical Review," International Critical Thought , p. 1, Nov. 2024, doi: 10.1080/21598282.2024.2431466.

[37] E. Shedeed, "Natural Law Conceptualized in The Republic of Cicero and The Republic of Plato," Jan. 2025, doi: 10.2139/ssrn.5173183.

[38] M. Asla and M. S. Paladino, "From Ethical Naturalism to Aquinas' Notion of Natural Law: A Non-Trivial Convergence?," Religions , vol. 15, no. 12, p. 1560, Dec. 2024, doi: 10.3390/rel15121560.

[39] F. Viola, "Aquinas, Thomas: On Natural Law," 2023, p. 119. doi: 10.1007/978-94-007-6519-1_364.

[40] R. Bradley, "The Relation Between Natural Law and Human Law in Thomas Aquinas," The Catholic lawyer , vol. 21, no. 1, p. 5, Mar. 2017, Accessed: Sep. 2025. [Online]. Available: https://repository.globethics.net/handle/20.500.12424/104113

[41] D. M. Stančienė, "The Ethics of Natural Law according to Thomas Aquinas," Verbum , vol. 6, no. 2, p. 357, Nov. 2004, doi: 10.1556/verb.6.2004.2.6.

[42] E. Alistar, "The Relation Between Law and Morality," SSRN Electronic Journal , Jan. 2019, doi: 10.2139/ssrn.3388103.

[43] J. Finnis, "Natural Law and Legal Reasoning Natural Law," EngagedScholarship @ Cleveland State University (Cleveland State University) , vol. 38, no. 1, p. 1, Jan. 1990, Accessed: Oct. 2025. [Online]. Available: https://engagedscholarship.csuohio.edu/clevstlrev/vol38/iss1/4

[44] R. Tuck and J. Finnis, "Natural Law and Natural Rights.," The Philosophical Quarterly , vol. 31, no. 124, p. 282, Jul. 1981, doi: 10.2307/2219123.

[45] X. Symons and T. J. VanderWeele, "Aristotelian Flourishing and Contemporary Philosophical Theories of Wellbeing," Journal of Happiness Studies , vol. 25, Feb. 2024, doi: 10.1007/s10902-024-00723-0.

[46] S. Veitch, "The Perfect Storm: Artificial Intelligence, Financialisation, and Venture Legalism," Law and Critique , vol. 35, no. 3, p. 609, Nov. 2024, doi: 10.1007/s10978-024-09401-9.

[47] J. Tasioulas, "Artificial Intelligence, Humanistic Ethics," Daedalus , vol. 151, no. 2, p. 232, Jan. 2022, doi: 10.1162/daed_a_01912.

[48] A. T. Kronman, "Hart, Austin, and the Concept of Legal Sanctions," Yale Law School Legal Scholarship Repository , Jan. 1975, Accessed: Nov. 2025. [Online]. Available: https://digitalcommons.law.yale.edu/fss_papers/1077

[49] E. H. Atiq, "Legal Positivism and the Moral Origins of Legal Systems," Canadian Journal of Law & Jurisprudence , vol. 36, no. 1, p. 37, Jul. 2022, doi: 10.1017/cjlj.2022.17.

[50] H. Kelsen, On The Pure Theory of Law , vol. 1, no. 1. Cambridge University Press, 1966, p. 1. doi: 10.1017/s0021223700013595.

[51] N. B. Ladavac, "Sein and Sollen, 'Is' and 'Ought' and the Problem of Normativity in Hans Kelsen," in Law and philosophy library , Springer Nature (Netherlands), 2019, p. 29. doi: 10.1007/978-3-030-18929-7_3.

[52] H. Kelsen, Pure Theory of Law . 1967. doi: 10.1525/9780520312296.

[53] F. A. Samekto, Muh. A. Mahfud, and A. P. Prabandari, "Grundnorm, Pure Theory of Law, and Monism in the Dynamics of International Law," Indonesian Journal of International Law , vol. 22, no. 4, Jun. 2025, doi: 10.17304/ijil.vol22.4.1911.

[54] M. Hildebrandt, "Legal and Technological Normativity," Techné Research in Philosophy and Technology , vol. 12, no. 3, p. 169, Jan. 2008, doi: 10.5840/techne20081232.

[55] F. Forzani, "The Power of Purity: Preliminary Notes for an Archaeology of Modern Jurisprudence," Law and Critique , Jun. 2024, doi: 10.1007/s10978-024-09394-5.

[56] C. E. Pavel, "Is International Law a Hartian Legal System?," Ratio Juris , vol. 31, no. 3, p. 307, Sep. 2018, doi: 10.1111/raju.12218.

[57] H. L. A. Hart, The Concept of Law . 1961. Accessed: Sep. 2025. [Online]. Available: http://all.fh.unair.ac.id/index.php?p=show_detail&id=1023

[58] G. Dajović, "Rule of Recognition and Constitution," 2023, p. 3172. doi: 10.1007/978-94-007-6519-1_378.

[59] N. Varsava, "Derivative Recognition and Intersystemic Interpretation," Legal Theory , vol. 30, no. 3, p. 170, Sep. 2024, doi: 10.1017/s1352325224000120.

[60] O. Gross, "What Both Hart and Fuller Got Wrong," University of Minnesota Law School Scholarship Repository (University of Minnesota) , May 2021, Accessed: Oct. 2025. [Online]. Available: https://scholarship.law.umn.edu/faculty_articles/745

[61] A. E. Afrouzi, "John Robots, Thurgood Martian, and the Syntax Monster: A New Argument Against AI Judges," Canadian Journal of Law & Jurisprudence , p. 1, Nov. 2024, doi: 10.1017/cjlj.2024.17.

[62] B. Townsend et al. , "From Pluralistic Normative Principles to Autonomous-Agent Rules," Minds and Machines , vol. 32, no. 4, p. 683, Oct. 2022, doi: 10.1007/s11023-022-09614-w.

[63] D. Priel, "Legal Positivism and Naturalistic Explanation of Action," Law and Philosophy , vol. 43, no. 1, p. 31, Jul. 2023, doi: 10.1007/s10982-023-09479-9.

[64] "four. Law and Morals: Positivism and Natural Law," in University of Texas Press eBooks , University of Texas Press, 2023, p. 63. doi: 10.7560/326411-005.

[65] B. Flanagan and G. Almeida, "Lawful, but not Really: The Dual Character of the Concept of Law," Law and Philosophy , vol. 43, no. 5, p. 507, Apr. 2024, doi: 10.1007/s10982-024-09501-8.

[66] J. L. Coleman, "On the Relationship Between Law and Morality," Ratio Juris , vol. 2, no. 1, p. 66, Mar. 1989, doi: 10.1111/j.1467-9337.1989.tb00027.x.

[67] R. Dworkin, "The Model of Rules," The University of Chicago Law Review , vol. 35, no. 1, p. 14, Jan. 1967, doi: 10.2307/1598947.
[68] Wisniewski and V. William, "An Application and Defense of Ronald Dworkin's Theory of Adjudication," Jan. 2007, Accessed: Jan. 2025. [Online]. Available: https://scholarworks.umt.edu/cgi/viewcontent.cgi?article=1136&context=etd
[69] L. Alexander, "The Legal Enforcement of Morality." p. 128, Jan. 01, 2005. doi: 10.1002/9780470996621.ch10.
[70] V. Rezadoost, "The Legal Policy of the International Court of Justice: A Conceptual Framework," in Brill | Nijhoff eBooks , Brill, 2024, p. 3. doi: 10.1163/9789004718128_002.
[71] C. Orwat, J. Bareis, A. Folberth, J. Jahnel, and C. Wadephul, "Normative Challenges of Risk Regulation of Artificial Intelligence," NanoEthics , vol. 18, no. 2, Aug. 2024, doi: 10.1007/s11569-024-00454-9.
[72] M. Cole, "International · AI Regulation and Governance on a Global Scale: An Overview of International, Regional and National Instruments," Journal of AI Law and Regulation , vol. 1, no. 1, p. 126, Jan. 2024, doi: 10.21552/aire/2024/1/16.
[73] J. Chun, C. S. de Witt, and K. Elkins, "Comparative Global AI Regulation: Policy Perspectives from the EU, China, and the US," Jan. 2025, doi: 10.2139/ssrn.5104429.
[74] W. Wu and S. Liu, "A Comprehensive Review and Systematic Analysis of Artificial Intelligence Regulation Policies," arXiv (Cornell University) . Cornell University, Jan. 01, 2023. doi: 10.48550/arxiv.2307.12218.
[75] N. K. Corrêa et al. , "Worldwide AI ethics: A review of 200 guidelines and recommendations for AI governance," Patterns , vol. 4, no. 10. Elsevier BV, p. 100857, Oct. 01, 2023. doi: 10.1016/j.patter.2023.100857.
[76] S. Alanoca, S. Gur-Arieh, T. Zick, and K. Klyman, "Comparing Apples to Oranges: A Taxonomy for Navigating the Global Landscape of AI Regulation," p. 914, Jun. 2025, doi: 10.1145/3715275.3732059.
[77] A. Chiappetta, "Navigating the AI frontier: European parliamentary insights on bias and regulation, preceding the AI Act," Internet Policy Review , vol. 12, no. 4, Dec. 2023, doi: 10.14763/2023.4.1733.
[78] M. Anderljung et al. , "Frontier AI Regulation: Managing Emerging Risks to Public Safety," arXiv (Cornell University) , Jul. 2023, doi: 10.48550/arxiv.2307.03718.
[79] M. Constantinides et al. , "Implications of Regulations on the Use of AI and Generative AI for Human-Centered Responsible Artificial Intelligence," p. 1, May 2024, doi: 10.1145/3613905.3643979.
[80] H. Hadan, R. H. Mogavi, L. Zhang-Kennedy, and L. E. Nacke, "Who is Responsible When AI Fails? Mapping Causes, Entities, and Consequences of AI Privacy and Ethical Incidents," International Journal of Human-Computer Interaction , p. 1, Sep. 2025, doi: 10.1080/10447318.2025.2549073.
[81] H. van Kolfschooten and J. van Oirschot, "The EU Artificial Intelligence Act (2024): Implications for healthcare," Health Policy , vol. 149, p. 105152, Sep. 2024, doi: 10.1016/j.healthpol.2024.105152.
[82] A. Sunyaev et al. , "High-Risk Artificial Intelligence," Business & Information Systems Engineering , May 2025, doi: 10.1007/s12599-025-00942-6.
[83] C. W.-L. Ho and K. Caals, "How the EU AI Act Seeks to Establish an Epistemic Environment of Trust," Asian Bioethics Review , vol. 16, no. 3, p. 345, Jun. 2024, doi: 10.1007/s41649-024-00304-6.
[84] F. L. G. Faroldi, "Risk and artificial general intelligence," AI & Society , Jul. 2024, doi: 10.1007/s00146-024-02004-z.
[85] C. de-Manuel-Vicente, D. Fernández-Narro, V. Blanes-Selva, J. M. García-Gómez, and C. Sáez, "A Development Framework for Trustworthy Artificial Intelligence in Health with Example Code Pipelines," medRxiv (Cold Spring Harbor Laboratory) , Jul. 2024, doi: 10.1101/2024.07.17.24310418.

[86] D. Golpayegani et al. , "AI Cards: Towards an Applied Framework for Machine-Readable AI and Risk Documentation Inspired by the EU AI Act," Jun. 2024, doi: 10.31219/osf.io/6dxgt.

[87] A. Lanamäki et al. , "What to Expect from the Upcoming EU AI Act Sandboxes: Panel Report," Deleted Journal , vol. 4, no. 2, May 2025, doi: 10.1007/s44206-025-00201-x.

[88] T. Odelberg, "Understanding the Future of Artificial Intelligence Governance: Comparing the EU AI Act and U.S. Executive Order on Safe AI," Deep Blue (University of Michigan) , Jun. 2024, doi: 10.7302/24392.

[89] A. D. Mustafa, "United States · AI Law Developments: A Late 2024 Update," Journal of AI Law and Regulation , vol. 1, no. 4, p. 479, Jan. 2024, doi: 10.21552/aire/2024/4/13.

[90] X. Wang, Y. C. Wu, X. Ji, and H. Fu, "Algorithmic discrimination: examining its types and regulatory measures with emphasis on US legal practices," Frontiers in Artificial Intelligence , vol. 7, May 2024, doi: 10.3389/frai.2024.1320277.

[91] C. G. Filippi et al. , "Ethical Considerations and Fairness in the Use of Artificial Intelligence for Neuroradiology," American Journal of Neuroradiology , vol. 44, no. 11. American Society of Neuroradiology, p. 1242, Aug. 31, 2023. doi: 10.3174/ajnr.a7963.

[92] A. Nastoska, B. Jancheska, M. Rizinski, and D. Trajanov, "Evaluating Trustworthiness in AI: Risks, Metrics, and Applications Across Industries," Electronics , vol. 14, no. 13, p. 2717, Jul. 2025, doi: 10.3390/electronics14132717.

[93] D. Lage, R. Pruitt, and J. R. Arnold, "Who Followed the Blueprint? Analyzing the Responses of U.S. Federal Agencies to the Blueprint for an AI Bill of Rights," 2024, doi: 10.48550/ARXIV.2404.19076.

[94] A. Oesterling, U. Bhalla, S. Venkatasubramanian, and H. Lakkaraju, "Operationalizing the Blueprint for an AI Bill of Rights: Recommendations for Practitioners, Researchers, and Policy Makers," 2024, doi: 10.48550/ARXIV.2407.08689.

[95] B. C. Larsen, "Governing Artificial Intelligence : Lessons from the United States and China," Research Portal Denmark , p. 300, Jan. 2022, Accessed: Aug. 2025. [Online]. Available: https://local.forskningsportal.dk/local/dki-cgi/ws/cris-link?src=cbs&id=cbs-48598543-6d06-4e04-82ac-07eea8605f9f&ti=Governing%20Artificial%20Intelligence%20%3A%20Lessons%20from%20the%20United%20States%20and%20China

[96] V. Rahimzadeh, "US regulation of medical artificial intelligence and machine learning (AI/ML) research and development," in Edward Elgar Publishing eBooks , Edward Elgar Publishing, 2024, p. 273. doi: 10.4337/9781802205657.ch16.

[97] S. Landau, J. X. Dempsey, E. Kamar, S. M. Bellovin, and R. Pool, "Challenging the Machine: Contestability in Government AI Systems," arXiv (Cornell University) , Jun. 2024, doi: 10.48550/arxiv.2406.10430.

[98] I. Cheong, A. Caliskan, and T. Kohno, "Safeguarding human values: rethinking US law for generative AI's societal impacts," AI and Ethics , vol. 5, no. 2, p. 1433, May 2024, doi: 10.1007/s43681-024-00451-4.

[99] Y. Pi and M. D. Proctor, "Toward empowering AI governance with redress mechanisms," Cambridge Forum on AI Law and Governance , vol. 1, Jan. 2025, doi: 10.1017/cfl.2025.9.

[100] A. H. Zhang, "The Promise and Perils of China's Regulation of Artificial Intelligence," SSRN Electronic Journal , Jan. 2024, doi: 10.2139/ssrn.4708676.

[101] B. Xiao, "Agile and Iterative Governance: China's Regulatory Response to Ai," Jan. 2024, doi: 10.2139/ssrn.4705898.

[102] G. Qiao-Franco and R. Zhu, "China's Artificial Intelligence Ethics: Policy Development in an Emergent Community of Practice," Journal of Contemporary China , vol. 33, no. 146, p. 189, Dec. 2022, doi: 10.1080/10670564.2022.2153016.

[103] X. Ye, Y. Yan, J. Li, and B. Jiang, "Privacy and personal data risk governance for generative artificial intelligence: A Chinese perspective," Telecommunications Policy , vol. 48, no. 10, p. 102851, Sep. 2024, doi: 10.1016/j.telpol.2024.102851.

[104] T. D. Noia, N. Tintarev, P. Fatourou, and M. Schedl, "Recommender systems under European AI regulations," Communications of the ACM , vol. 65, no. 4, p. 69, Mar. 2022, doi: 10.1145/3512728.

[105] B. Goodman and S. Flaxman, "European Union Regulations on Algorithmic Decision Making and a 'Right to Explanation,'" AI Magazine , vol. 38, no. 3, p. 50, Sep. 2017, doi: 10.1609/aimag.v38i3.2741.

[106] R. B. Chellappan, "From Algorithms to Accountability: The Societal and Ethical Need for Explainable AI," Research Square (Research Square) , Oct. 2024, doi: 10.21203/rs.3.rs-5277731/v1.

[107] S. Timcke, "AI Regulation Paradigms and the Struggle over Control Rights," May 2024, doi: 10.21428/bfcb0bff.ae40312a.

[108] R. Owens et al. , "Reimagining AI Governance: a Response by AGENCY to the UK Government's White Paper AI Regulation," SSRN Electronic Journal , Jan. 2023, doi: 10.2139/ssrn.4551757.

[109] P. Li, R. Williams, S. Gilbert, and S. Anderson, "Regulating Artificial Intelligence and Machine Learning-Enabled Medical Devices in Europe and the United Kingdom," Law Technology and Humans , vol. 5, no. 2, p. 94, Nov. 2023, doi: 10.5204/lthj.3073.

[110] D. R. Brown, "Canada's Proposed Artificial Intelligence and Data Act (AIDA): A Critical Review," SSRN Electronic Journal . RELX Group (Netherlands), Jan. 01, 2024. doi: 10.2139/ssrn.4687995.

[111] R. Kim and K. Thomasen, "Submission to The Standing Committee on Industry and Technology on Bill C-27, An Act to enact the Consumer Privacy Protection Act, the Personal Information and Data Protection Tribunal Act and the Artificial Intelligence and Data Act and to make consequential and related amendments to other Acts," SSRN Electronic Journal , Jan. 2023, doi: 10.2139/ssrn.4571389.

[112] J. G. Allen, J. Loo, and J. L. L. Campoverde, "Governing intelligence: Singapore's evolving AI governance framework," vol. 1, Jan. 2025, doi: 10.1017/cfl.2024.12.

[113] J. G. Allen and J. Loo, "Singapore's Evolving AI Governance Framework," in Cambridge University Press eBooks , Cambridge University Press, 2025, p. 151. doi: 10.1017/9781009492553.014.

[114] S. A.V. and N. D. N. -, "Legal Challenges of Artificial Intelligence in India's Cyber Law Framework: Examining Data Privacy and Algorithmic Accountability Via a Comparative Global Perspective," International Journal For Multidisciplinary Research , vol. 6, no. 6, Nov. 2024, doi: 10.36948/ijfmr.2024.v06i06.31347.

[115] R. Bharati, "Navigating the Legal Landscape of Artificial Intelligence: Emerging Challenges and Regulatory Framework in India," SSRN Electronic Journal , Jan. 2024, doi: 10.2139/ssrn.4898536.

[116] N. Rangone and L. Megale, "Risks Without Rights? The EU AI Act's Approach to AI in Law and Rule-Making," European Journal of Risk Regulation , p. 1, Mar. 2025, doi: 10.1017/err.2025.13.

[117] Mag. J. S. Butt, "Analytical Study of the World's First EU Artificial Intelligence (AI) Act, 2024," International Journal of Research Publication and Reviews , vol. 5, no. 3, p. 7343, Mar. 2024, doi: 10.55248/gengpi.5.0324.0914.

[118] M. Anderljung, J. Hazell, and M. von Knebel, "Protecting society from AI misuse: when are restrictions on capabilities warranted?," AI & Society , vol. 40, no. 5, p. 3841, Dec. 2024, doi: 10.1007/s00146-024-02130-8.

[119] M. Chen, S. Engelmann, and J. Großklags, "Social Credit System and Privacy," 2023, p. 227. doi: 10.4324/9781003244677-26.

[120] A. Mantelero, "The Fundamental Rights Impact Assessment (FRIA) in the AI Act: Roots, legal obligations and key elements for a model template," Computer Law & Security Review , vol. 54, p. 106020, Aug. 2024, doi: 10.1016/j.clsr.2024.106020.

[121] T. Rintamaki and H. J. Pandit, "Towards An Automated AI Act FRIA Tool That Can Reuse GDPR's DPIA," 2025, doi: 10.48550/ARXIV.2501.14756.

[122] A. Thomaidou and K. Limniotis, "Navigating Through Human Rights in AI: Exploring the Interplay Between GDPR and Fundamental Rights Impact Assessment," Journal of Cybersecurity and Privacy , vol. 5, no. 1, p. 7, Feb. 2025, doi: 10.3390/jcp5010007.

[123] S. Bertaina et al. , "Fundamental rights and artificial intelligence impact assessment: A new quantitative methodology in the upcoming era of AI Act," Computer Law & Security Review , vol. 56, p. 106101, Jan. 2025, doi: 10.1016/j.clsr.2024.106101.

[124] T. Gils, F. Heymans, W. Ooms, and J. D. Bruyne, "From Policy to Practice: Prototyping The EU AI Act's Transparency Requirements," SSRN Electronic Journal , Jan. 2024, doi: 10.2139/ssrn.4714345.

[125] N. A. Moreira, P. M. Freitas, and P. Nováis, "Towards Transparent AI: How will the AI Act Shape the Future?," in Lecture notes in computer science , Springer Science+Business Media, 2024, p. 296. doi: 10.1007/978-3-031-73497-7_24.

[126] V. Zieglmeier and A. Pretschner, "Trustworthy Transparency by Design," arXiv (Cornell University) , Mar. 2021, doi: 10.48550/arxiv.2103.10769.

[127] F. Königstorfer, "A comprehensive review of techniques for documenting artificial intelligence," Digital Policy Regulation and Governance , vol. 26, no. 5. Emerald Publishing Limited, p. 545, May 29, 2024. doi: 10.1108/dprg-01-2024-0008.

[128] B. C. Cheong, "Transparency and accountability in AI systems: safeguarding wellbeing in the age of algorithmic decision-making," Frontiers in Human Dynamics , vol. 6, Jul. 2024, doi: 10.3389/fhumd.2024.1421273.

[129] B. Karlan, "Authenticity in algorithm-aided decision-making," Synthese , vol. 204, no. 3, Aug. 2024, doi: 10.1007/s11229-024-04716-7.

[130] M. C. Decker, L. Wegner, and C. Leicht-Scholten, "Procedural fairness in algorithmic decision-making: the role of public engagement," Ethics and Information Technology , vol. 27, no. 1, Nov. 2024, doi: 10.1007/s10676-024-09811-4.

[131] A. E. Ali, K. P. Venkatraj, S. Morosoli, L. Naudts, N. Helberger, and P. César, "Transparent AI Disclosure Obligations: Who, What, When, Where, Why, How," p. 1, May 2024, doi: 10.1145/3613905.3650750.

[132] R. Lae and M. P. Smørdal, "Norway's New Transparency Act: An Overview in Light of International Trends," Oslo Law Review , vol. 10, no. 2, p. 1, Apr. 2024, doi: 10.18261/olr.10.2.5.

[133] U.-M. Mylly, "Transparent AI? Navigating Between Rules on Trade Secrets and Access to Information," Deleted Journal , vol. 54, no. 7, p. 1013, May 2023, doi: 10.1007/s40319-023-01328-5.

[134] M. Fink, "Human Oversight under Article 14 of the EU AI Act," Jan. 2025, doi: 10.2139/ssrn.5147196.

[135] B. Wagner, "Liable, but Not in Control? Ensuring Meaningful Human Agency in Automated Decision-Making Systems," Policy & Internet , vol. 11, no. 1, p. 104, Jan. 2019, doi: 10.1002/poi3.198.

[136] F. S. de Sio and J. van den Hoven, "Meaningful Human Control over Autonomous Systems: A Philosophical Account," Frontiers in Robotics and AI , vol. 5, p. 15, Feb. 2018, doi: 10.3389/frobt.2018.00015.

[137] L. M. Haitsma, "Regulating algorithmic discrimination through adjudication: the Court of Justice of the European Union on discrimination in algorithmic profiling based on PNR data," Frontiers in Political Science , vol. 5, Oct. 2023, doi: 10.3389/fpos.2023.1232601.

[138] K. Levy, K. Chasalow, and S. Riley, "Algorithms and Decision-Making in the Public Sector," Annual Review of Law and Social Science , vol. 17, no. 1, p. 309, Jul. 2021, doi: 10.1146/annurev-lawsocsci-041221-023808.

[139] M. Ho-Dac and B. Martinez, "Human Oversight of Artificial Intelligence and Technical Standardisation," Jan. 2025, doi: 10.2139/ssrn.5228774.

[140] N. Silva, "The Artificial Intelligence Act: Critical Overview," Jan. 2024, doi: 10.2139/ssrn.4937150.

[141] B. Ganglmair, J. Krämer, and J. Gambato, "Regulatory Compliance with Limited Enforceability: Evidence from Privacy Policies," SSRN Electronic Journal , Jan. 2023, doi: 10.2139/ssrn.4600876.

[142] L. Tomada, "Start-ups and the proposed EU AI Act : Bridges or Barriers in the path from Invention to Innovation?," Research Portal Denmark , vol. 13, no. 1, p. 53, Jan. 2022, Accessed: Aug. 2025. [Online]. Available: https://local.forskningsportal.dk/local/dki-cgi/ws/cris-link?src=ku&id=ku-e74b43f3-a896-48d8-aa50-0bf1fe755cd9&ti=Start-ups%20and%20the%20proposed%20EU%20AI%20Act%20%3A%20Bridges%20or%20Barriers%20in%20the%20path%20from%20Invention%20to%20Innovation%3F

[143] O. J. Gstrein, N. Haleem, and A. Zwitter, "General-purpose AI regulation and the European Union AI Act," Internet Policy Review , vol. 13, no. 3, Aug. 2024, doi: 10.14763/2024.3.1790.

[144] R. Kilian, L. J. auml ck, and D. Ebel, "European AI Standards - Technical Standardization and Implementation Challenges under the EU AI Act," SSRN Electronic Journal , Jan. 2025, doi: 10.2139/ssrn.5155591.

[145] M. Hussain and N. Akhtar, "Theoretically Justification of Human Rights: A Case of Natural Rights Theory," Zakariya Journal of Social Science , vol. 2, no. 1, p. 42, Jun. 2023, doi: 10.59075/zjss.v2i1.232.

[146] J. C. R. Martí-nez-Villalba, "How to Deduce Human Rights From Natural Law and Other Disciplines," Ius Humani Law Journal , vol. 12, no. 2, p. 27, Oct. 2023, doi: 10.31207/ih.v12i2.327.

[147] A. J. Lisska, "Human Rights Theory Rooted in the Writings of Thomas Aquinas," Diametros , vol. 38, no. 38, p. 134, Dec. 2013, doi: 10.13153/diam.38.2013.542.

[148] R. P. George, "Natural Law, God and Human Dignity," in Cambridge University Press eBooks , Cambridge University Press, 2017, p. 57. doi: 10.1017/9781316341544.003.

[149] J. R. Laracy, V. Kirova, C. S. Ku, and T. J. Marlowe, "Human Dignity and the Ethics of Artificial Intelligence: A Framework for Responsible Design and Use from the Perspective of Catholic Social Teaching," p. 1, Jun. 2025, doi: 10.1109/ethics65148.2025.11098185.

[150] R. Hanna and E. Kazim, "Philosophical foundations for digital ethics and AI Ethics: a dignitarian approach," AI and Ethics , vol. 1, no. 4, p. 405, Feb. 2021, doi: 10.1007/s43681-021-00040-9.

[151] C. Thielscher, "Dignity as a concept for computer ethics," AI and Ethics , vol. 5, no. 4, p. 4061, Mar. 2025, doi: 10.1007/s43681-025-00693-w.

[152] M. Ashok, R. Madan, A. Joha, and U. Sivarajah, "Ethical framework for Artificial Intelligence and Digital technologies," International Journal of Information Management , vol. 62, p. 102433, Oct. 2021, doi: 10.1016/j.ijinfomgt.2021.102433.

[153] I. Gabriel et al. , "The Ethics of Advanced AI Assistants," arXiv (Cornell University) , Apr. 2024, doi: 10.48550/arxiv.2404.16244.

[154] V. Prabhakaran, M. Mitchell, T. Gebru, and I. Gabriel, "A Human Rights-Based Approach to Responsible AI," arXiv (Cornell University) , Jan. 2022, doi: 10.48550/arXiv.2210.02667.

[155] M. K. Cohen, N. Kolt, Y. Bengio, G. K. Hadfield, and S. Russell, "Regulating advanced artificial agents," Science , vol. 384, no. 6691, p. 36, Apr. 2024, doi: 10.1126/science.adl0625.

[156] A. Reis et al. , "ETHICS AND GOVERNANCE OF ARTIFICIAL INTELLIGENCE FOR HEALTH : WHO GUIDANCE," Technical University of Denmark, Jan. 2021. Accessed: Aug. 2025. [Online]. Available: https://local.forskningsportal.dk/local/dki-cgi/ws/cris-link?src=ku&id=ku-e505319d-a404-4d93-bfa7-577735b36fdd&ti=ETHICS%20AND%20GOVERNANCE%20OF%20ARTIFICIAL%20INTELLIGENCE%20FOR%20HEALTH%20%3A%20WHO%20GUIDANCE

[157] M. Coeckelbergh, "Artificial intelligence, the common good, and the democratic deficit in AI governance," AI and Ethics , vol. 5, no. 2, p. 1491, May 2024, doi: 10.1007/s43681-024-00492-9.

[158] U. Nations, Governing AI for Humanity . United Nations, 2024. doi: 10.18356/9789211067873.

[159] B. C. Stahl et al. , "Artificial intelligence for human flourishing – Beyond principles for machine learning," Journal of Business Research , vol. 124, p. 374, Dec. 2020, doi: 10.1016/j.jbusres.2020.11.030.

[160] J. Waldron, "Rights and Human Rights," in Cambridge University Press eBooks , Cambridge University Press, 2020, p. 152. doi: 10.1017/9781316104439.009.

[161] A. Wiener, "The many paths of change in international law," International Affairs , vol. 101, no. 1, p. 346, Jan. 2025, doi: 10.1093/ia/iiae305.

[162] P. R. Dubinsky, T. Higgins, M. Rosenfeld, J. Waldron, and R. Teitel, "What Is a Human Right? Universals and the Challenge of Cultural Relativism," Pace international law review , vol. 11, no. 1, p. 107, Apr. 1999, doi: 10.58948/2331-3536.1248.

[163] S. A. Teo, "Towards Human Rights 2.0? : a meta-theoretical analysis of the disruptions to human rights foundations by artificial intelligence," Research Portal Denmark , p. 429, Jan. 2023, Accessed: Jul. 2025. [Online]. Available: https://local.forskningsportal.dk/local/dki-cgi/ws/cris-link?src=ku&id=ku-5a8c42de-4966-4922-8d3a-58afd08fde80&ti=Towards%20Human%20Rights%202.0%3F%20%3A%20a%20meta-theoretical%20analysis%20of%20the%20disruptions%20to%20human%20rights%20foundations%20by%20artificial%20intelligence

[164] H. Aziz and G. Tom, The Entrenchment of Democracy . Cambridge University Press, 2024. doi: 10.1017/9781009447713.

[165] O. Lobel, "The Law of AI for Good," SSRN Electronic Journal , Jan. 2023, doi: 10.2139/ssrn.4338862.

[166] M. Kiškis, "Legal framework for the coexistence of humans and conscious AI," Frontiers in Artificial Intelligence , vol. 6. Frontiers Media, Sep. 20, 2023. doi: 10.3389/frai.2023.1205465.

[167] A. Akbik and M. Dawson, "From Procedural to Substantive Accountability in EMU Governance," in Cambridge University Press eBooks , Cambridge University Press, 2023, p. 19. doi: 10.1017/9781009228800.003.

[168] K. Buhmann and R. Leal-Arcas, "Towards a Socially Responsible Green Transition : The Role of Investors in Closing Governance Gaps Affecting the Respect for Human Rights Around Raw Materials for a Non-carbon Economy," Research Portal Denmark , p. 157, Jan. 2020, Accessed: Jul. 2025. [Online]. Available: https://local.forskningsportal.dk/local/dki-cgi/ws/cris-link?src=cbs&id=cbs-7921d2c2-e0ad-4dbf-9116-d57b040797d5&ti=Towards%20a%20Socially

%20Responsible%20Green%20Transition%20%3A%20The%20Role%20of%20Investors%20in%20Closing%20Governance%20Gaps%20Affecting%20the%20Respect%20for%20Human%20Rights%20Around%20Raw%20Materials%20for%20a%20Non%2010carbon%20Economy

[169] S. L. Shaelou and Y. Razmetaeva, "Challenges to Fundamental Human Rights in the age of Artificial Intelligence Systems: shaping the digital legal order while upholding Rule of Law principles and European values," ERA Forum , vol. 24, no. 4, p. 567, Dec. 2023, doi: 10.1007/s12027-023-00777-2.

[170] K. D. S. Nonju and A. B. Ihua-Maduenyi, "The Impact of Artificial Intelligence on Privacy Laws," International Journal of Research and Innovation in Social Science , p. 2150, Jan. 2024, doi: 10.47772/ijriss.2024.8090178.

[171] R. Mühlhoff, "Predictive privacy: towards an applied ethics of data analytics," Ethics and Information Technology , vol. 23, no. 4, p. 675, Jul. 2021, doi: 10.1007/s10676-021-09606-x.

[172] R. Mühlhoff, "Predictive privacy: Collective data protection in the context of artificial intelligence and big data," Big Data & Society , vol. 10, no. 1, Jan. 2023, doi: 10.1177/20539517231166886.

[173] L. Floridi, "On Human Dignity as a Foundation for the Right to Privacy," Philosophy & Technology , vol. 29, no. 4, p. 307, Apr. 2016, doi: 10.1007/s13347-016-0220-8.

[174] A. Laitinen and O. Sahlgren, "AI Systems and Respect for Human Autonomy," Frontiers in Artificial Intelligence , vol. 4, Oct. 2021, doi: 10.3389/frai.2021.705164.

[175] R. Schwartmann, T. Keber, K. Zenner, and S. Kurth, "Data Protection Aspects of the Use of Artificial Intelligence — Initial overview of the intersection between GDPR and AI Act," Computer Law Review International , vol. 25, no. 5, p. 145, Oct. 2024, doi: 10.9785/cri-2024-250503.

[176] K. D. Martin and J. Zimmermann, "Artificial intelligence and its implications for data privacy," Current Opinion in Psychology , vol. 58. Elsevier BV, p. 101829, Jun. 22, 2024. doi: 10.1016/j.copsyc.2024.101829.

[177] D. M. Leslie et al. , "AI Fairness in Practice," SSRN Electronic Journal , Jan. 2024, doi: 10.2139/ssrn.4731838.

[178] H. Cossette-Lefebvre and J. Maclure, "AI's fairness problem: understanding wrongful discrimination in the context of automated decision-making," AI and Ethics , vol. 3, no. 4, p. 1255, Nov. 2022, doi: 10.1007/s43681-022-00233-w.

[179] H. Hoch, C. Hertweck, M. Loi, and A. Tamò-Larrieux, "Discrimination for the sake of fairness by design and its legal framework," Computer law & security review , vol. 52, p. 105916, Jan. 2024, doi: 10.1016/j.clsr.2023.105916.

[180] X. Ferrer, T. van Nuenen, J. M. Such, M. Coté, and N. Criado, "Bias and Discrimination in AI: A Cross-Disciplinary Perspective," IEEE Technology and Society Magazine , vol. 40, no. 2, p. 72, Jun. 2021, doi: 10.1109/mts.2021.3056293.

[181] J. Adams-Prassl, R. Binns, and A. Kelly-Lyth, "Directly Discriminatory Algorithms," Modern Law Review , vol. 86, no. 1, p. 144, Aug. 2022, doi: 10.1111/1468-2230.12759.

[182] L. Londoño, J. V. Hurtado, N. Hertz, P. Kellmeyer, S. Vöeneky, and A. Valada, "Fairness and Bias in Robot Learning," Proceedings of the IEEE , vol. 112, no. 4, p. 305, Apr. 2024, doi: 10.1109/jproc.2024.3403898.

[183] E. Black, L. Koepke, P. Kim, S. Barocas, and M. Hsu, "The Legal Duty to Search for Less Discriminatory Algorithms," arXiv (Cornell University) , Jun. 2024, doi: 10.48550/arxiv.2406.06817.

[184] R. Segev, "Artificial Intelligence, Discrimination, Fairness, and Other Moral Concerns," Minds and Machines , vol. 34, no. 4, Nov. 2024, doi: 10.1007/s11023-024-09702-z.

[185] N. Chronowski, K. Kálmán, and B. Szentgáli-Tóth, "Artificial Intelligence, Justice, and Certain Aspects of Right to a Fair Trial," Acta Universitatis Sapientiae Legal Studies , vol. 10, no. 2, p. 169, Dec. 2021, doi: 10.47745/ausleg.2021.10.2.02.

[186] V. Mohod et al. , "Harnessing Predictive Analytics and AI in Judicial Decisions," in Advances in electronic government, digital divide, and regional development book series , IGI Global, 2024, p. 187. doi: 10.4018/979-8-3693-7205-0.ch010.

[187] R. Bharati, "Ethical Implications of AI in Criminal Justice: Balancing Efficiency and Due Process," RESEARCH REVIEW International Journal of Multidisciplinary , vol. 9, no. 7, p. 93, Jul. 2024, doi: 10.31305/rrijm.2024.v09.n07.014.

[188] C. Moore et al. , "Concerning the Responsible Use of AI in the US Criminal Justice System," 2025, doi: 10.48550/ARXIV.2506.00212.

[189] M. Ebers and H. Trasberg, "Due Process, Fair Trial, Transparency, and Explainability," in Data science, machine intelligence, and law , 2023, p. 103. doi: 10.1007/978-3-031-19667-6_6.

[190] I. Cofone and W. Khern-am-nuai, "The Overstated Cost of AI Fairness in Criminal Justice," SSRN Electronic Journal , Jan. 2025, doi: 10.2139/ssrn.5117071.

[191] G. Lopes, "Bias in Adjudication and the Promise of AI: Challenges to Procedural Fairness," Law Technology and Humans , vol. 7, no. 1, p. 47, Apr. 2025, doi: 10.5204/lthj.3812.

[192] J. Collenette, K. Atkinson, and T. Bench-Capon, "Explainable AI tools for legal reasoning about cases: A study on the European Court of Human Rights," Artificial Intelligence , vol. 317, p. 103861, Jan. 2023, doi: 10.1016/j.artint.2023.103861.

[193] G. D. Gregorio and P. M. Dunn, "Artificial Intelligence and Freedom of Expression," in Oxford University Press eBooks , Oxford University Press, 2023. doi: 10.1093/law/9780192882486.003.0005.

[194] "The Study of Right to Freedom of Expression, Digital Media Laws and Demand for Innovation," Law and World , vol. 11, no. 1, p. 114, Mar. 2025, doi: 10.36475/11.1.9.

[195] E. Kavoliūnaitė-Ragauskienė, "Artificial Intelligence in Manipulation: The Significance and Strategies for Prevention," Baltic Journal of Law & Politics , vol. 17, no. 2, p. 116, Dec. 2024, doi: 10.2478/bjlp-2024-00018.

[196] B. Heller and J. van Hoboken, "Freedom of Expression: A Comparative Summary of United States and European Law," SSRN Electronic Journal , Jan. 2023, doi: 10.2139/ssrn.4563882.

[197] O. Bajgar and J. Horenovsky, "Negative Human Rights as a Basis for Long-term AI Safety and Regulation," Journal of Artificial Intelligence Research , vol. 76, p. 1043, Apr. 2023, doi: 10.1613/jair.1.14020.

[198] I. Gabriel, "Artificial Intelligence, Values, and Alignment," Minds and Machines , vol. 30, no. 3, p. 411, Sep. 2020, doi: 10.1007/s11023-020-09539-2.

[199] B. Frész, E. Dubovitskaya, D. Brajovic, M. F. Huber, and C. Horz, "How Should AI Decisions Be Explained? Requirements for Explanations from the Perspective of European Law," Proceedings of the AAAI/ACM Conference on AI Ethics and Society , vol. 7, p. 438, Oct. 2024, doi: 10.1609/aies.v7i1.31648.

[200] A. Rotolo, B. Ferrigno, J. M. A. G. Godinez, C. Novelli, and G. Sartor, "Foundations for Risk Assessment of AI in Protecting Fundamental Rights," 2025, doi: 10.48550/ARXIV.2507.18290.

[201] L. del B. Aranda, C. Custis, O. J. Erdélyi, L. C. C. Galindo, and R. Rozenberg, "Advancing accountability in AI," Feb. 2023. doi: 10.1787/2448f04b-en.

[202] S. Sterz et al. , "On the Quest for Effectiveness in Human Oversight: Interdisciplinary Perspectives," in 2022 ACM Conference on Fairness, Accountability, and Transparency , Jun. 2024, p. 2495. doi: 10.1145/3630106.3659051.

[203] T. Chatinakrob, "Legal Risks and Challenges of Unregulated AI," 2024, p. 268. doi: 10.1163/9789004719934_010.
[204] K. Kyriakou and J. Otterbacher, "In humans, we trust," Discover Artificial Intelligence , vol. 3, no. 1, Dec. 2023, doi: 10.1007/s44163-023-00092-2.
[205] J. Finnis, "Grounding Human Rights in Natural Law," The American Journal of Jurisprudence , vol. 60, no. 2, p. 199, Nov. 2015, doi: 10.1093/ajj/auv013.
[206] E. Ferrara, "Fairness and Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, and Mitigation Strategies," Sci , vol. 6, no. 1, p. 3, Dec. 2023, doi: 10.3390/sci6010003.
[207] J. Simon, P. Wong, and G. Rieder, "Algorithmic bias and the Value Sensitive Design approach," Internet Policy Review , vol. 9, no. 4, Dec. 2020, doi: 10.14763/2020.4.1534.
[208] L. G. Vicente and H. Matute, "Humans inherit artificial intelligence biases," Scientific Reports , vol. 13, no. 1, Oct. 2023, doi: 10.1038/s41598-023-42384-8.
[209] A. Birhane, "The unseen Black faces of AI algorithms," Nature , vol. 610, no. 7932, p. 451, Oct. 2022, doi: 10.1038/d41586-022-03050-7.
[210] L. C. D. Stein, K. Yang, D. Bowen, and M. J. Price, "Measuring and Mitigating Racial Disparities in Large Language Model Mortgage Underwriting," Jan. 2025, doi: 10.15396/eres2025_75.
[211] J. H. Gerards and R. Xenidis, "Algorithmic discrimination in Europe : Challenges and opportunities for gender equality and non-discrimination law," Technical University of Denmark, Jan. 2021. Accessed: Aug. 2025. [Online]. Available: https://local.forskningsportal.dk/local/dki-cgi/ws/cris-link?src=ku&id=ku-7399da8e-ff97-4023-ada0-212038ec4cbd&ti=Algorithmic%20discrimination%20in%20Europe%20%3A%20Challenges%20and%20opportunities%20for%20gender%20equality%20and%20non-discrimination%20law
[212] P. Parasurama and P. Ipeirotis, "Algorithmic Hiring and Diversity: Reducing Human-Algorithm Similarity for Better Outcomes," 2025, doi: 10.48550/ARXIV.2505.14388.
[213] A. Fabris et al. , "Fairness and Bias in Algorithmic Hiring: a Multidisciplinary Survey," 2023, doi: 10.48550/ARXIV.2309.13933.
[214] A. Fabris et al. , "Fairness and Bias in Algorithmic Hiring: A Multidisciplinary Survey," ACM Transactions on Intelligent Systems and Technology , vol. 16, no. 1, p. 1, Sep. 2024, doi: 10.1145/3696457.
[215] K. Wilson, M. Sim, A.-M. Gueorguieva, and A. Caliskan, "No Thoughts Just AI: Biased LLM Hiring Recommendations Alter Human Decision Making and Limit Human Autonomy," 2025, doi: 10.48550/ARXIV.2509.04404.
[216] M. M. A. A. M. Sony, M. B. Amin, A. Ashraf, K. M. A. Islam, N. C. Debnath, and G. C. Debnath, "Bias in AI-driven HRM systems: Investigating discrimination risks embedded in AI recruitment tools and HR analytics," Social Sciences & Humanities Open , vol. 12, p. 102082, Jan. 2025, doi: 10.1016/j.ssaho.2025.102082.
[217] P. Hacker, F. Z. Borgesius, B. Mittelstadt, and S. Wachter, "Generative Discrimination: What Happens When Generative AI Exhibits Bias, and What Can Be Done About It," SSRN Electronic Journal , Jan. 2024, doi: 10.2139/ssrn.4877398.
[218] D. Purves, "Fairness in Algorithmic Policing," Journal of the American Philosophical Association , vol. 8, no. 4, p. 741, Jan. 2022, doi: 10.1017/apa.2021.39.
[219] N. Pagan, J. Baumann, E. Elokda, G. D. Pasquale, S. Bolognani, and A. Hannák, "A Classification of Feedback Loops and Their Relation to Biases in Automated Decision-Making Systems," p. 1, Oct. 2023, doi: 10.1145/3617694.3623227.
[220] C. Orwat, "Algorithmic Discrimination From the Perspective of Human Dignity," Social Inclusion , vol. 12, May 2024, doi: 10.17645/si.7160.

[221] P. S. Shin, "THE SUBSTANTIVE PRINCIPLE OF EQUAL TREATMENT," Legal Theory , vol. 15, no. 2, p. 149, Jun. 2009, doi: 10.1017/s1352325209090090.

[222] T. Loenen, The Conceptualization of Equality and Non-discrimination as Legal Standards . 2024. doi: 10.1163/9789004538368.

[223] H. Machado, S. Silva, and L. Neiva, "Publics' views on ethical challenges of artificial intelligence: a scoping review," AI and Ethics . Springer Nature, Dec. 19, 2023. doi: 10.1007/s43681-023-00387-1.

[224] M. D. Schultz, M. Clegg, R. Hofstetter, and P. Seele, "Algorithms and dehumanization: a definition and avoidance model," AI & Society , vol. 40, no. 4, p. 2191, Nov. 2024, doi: 10.1007/s00146-024-02123-7.

[225] I. Gabriel, "Toward a Theory of Justice for Artificial Intelligence," Daedalus , vol. 151, no. 2, p. 218, Jan. 2022, doi: 10.1162/daed_a_01911.

[226] N. C. Vargas, "Exploiting the margin: How capitalism fuels AI at the expense of minoritized groups," AI and Ethics , vol. 5, no. 2, p. 1871, Jun. 2024, doi: 10.1007/s43681-024-00502-w.

[227] A. Lustig, "Fostering the Global Common Good: The Relevance of Catholic Social Teaching to Public Health Debates," Religions , vol. 14, no. 4, p. 504, Apr. 2023, doi: 10.3390/rel14040504.

[228] R. Fulton, D. Fulton, N. Hayes, and S. Kaplan, "The Transformation Risk-Benefit Model of Artificial Intelligence: Balancing Risks and Benefits Through Practical Solutions and Use Cases," arXiv (Cornell University) , Apr. 2024, doi: 10.5121/ijaia.15201.

[229] P. O. Schenk and C. Kern, "Connecting algorithmic fairness to quality dimensions in machine learning in official statistics and survey production," AStA Wirtschafts- und Sozialstatistisches Archiv , vol. 18, no. 2, p. 131, Jun. 2024, doi: 10.1007/s11943-024-00344-2.

[230] J. Baumann and C. Heitz, "Group Fairness in Prediction-Based Decision Making: From Moral Assessment to Implementation," p. 19, Jun. 2022, doi: 10.1109/sds54800.2022.00011.

[231] D. López-Gómez, "The Quest for Algorithmic Justice in the Workplace: The Equal Employment Opportunity Commission and Other Federal Responses to AI, Technology, and Enhanced Dangers of Employment Discrimination," Seton Hall Journal of Legislation and Public Policy , vol. 49, no. 3, p. 683, Jan. 2025, doi: 10.60095/qyir8109.

[232] J. J.-X. Wu, "Algorithmic Fairness in Consumer Credit Underwriting: Towards a 'Harm-Based' Framework for AI Fair Lending," SSRN Electronic Journal , Jan. 2023, doi: 10.2139/ssrn.4320444.

[233] A. Vidrio, "Tenant Screening Algorithms: Policy Intervention," University of Michigan, Nov. 2023. doi: 10.7302/21915.

[234] I. E. Kumar, K. Hines, and J. P. Dickerson, "Equalizing Credit Opportunity in Algorithms," p. 357, Jul. 2022, doi: 10.1145/3514094.3534154.

[235] S. Wachter, B. Mittelstadt, and C. Russell, "Why fairness cannot be automated: Bridging the gap between EU non-discrimination law and AI," Computer Law & Security Review , vol. 41, p. 105567, Jun. 2021, doi: 10.1016/j.clsr.2021.105567.

[236] G. L. Moratinos and P. D. Hert, "Humans in the GDPR and AIA governance of automated and algorithmic systems. Essential pre-requisites against abdicating responsibilities," Computer Law & Security Review , vol. 50, p. 105833, Jun. 2023, doi: 10.1016/j.clsr.2023.105833.

[237] N. Kallus, X. Mao, and A. Zhou, "Assessing algorithmic fairness with unobserved protected class using data combination," Jan. 2020, doi: 10.1145/3351095.3373154.

[238] E. Black, J. L. Koepke, P. Kim, S. Barocas, and M. Hsu, "Less Discriminatory Algorithms," SSRN Electronic Journal , Jan. 2023, doi: 10.2139/ssrn.4590481.

[239] A. E. R. Prince and D. Schwarcz, "Proxy Discrimination in the Age of Artificial Intelligence and Big Data," Aug. 2019.

[240] F. J. Cavico and B. G. Mujtaba, "Diversity, disparate impact, and discrimination pursuant to Title VII of US civil rights laws," Equality Diversity and Inclusion An International

Journal , vol. 36, no. 7, p. 670, Sep. 2017, doi: 10.1108/edi-04-2017-0091.

[241] C. Chien, A. Zink, and I. Y. Chen, "Bridging Research Gaps Between Academic Research and Legal Investigations of Algorithmic Discrimination," vol. 8, no. 1, p. 591, Oct. 2025, doi: 10.1609/aies.v8i1.36573.

[242] M. A. Mersha, K. N. Lâm, J. Wood, A. K. AlShami, and J. Kalita, "Explainable artificial intelligence: A survey of needs, techniques, applications, and future direction," Neurocomputing , vol. 599, p. 128111, Jun. 2024, doi: 10.1016/j.neucom.2024.128111.

[243] C.-C. R. Yuan and B.-Y. Wang, "Ensuring Fairness with Transparent Auditing of Quantitative Bias in AI Systems," p. 25, Aug. 2024, doi: 10.23919/pnc63053.2024.10697374.

[244] C. O'Neil, H. Sargeant, and J. M. Appel, "Explainable Fairness in Regulatory Algorithmic Auditing," SSRN Electronic Journal , Jan. 2024, doi: 10.2139/ssrn.4756637.

[245] E. Schnapper, "Two Categories of Discriminatory Intent.," Harvard civil rights-civil liberties law review , vol. 17, no. 1, p. 31, Jan. 1982, Accessed: Oct. 2025. [Online]. Available: https://digitalcommons.law.uw.edu/cgi/viewcontent.cgi?article=1309&context=faculty-articles

[246] Muh. Sjaiful and W. Zahrowati, "The Integration For Morality Values In The Concept of Law," Journal of Law Policy and Globalization , Nov. 2019, doi: 10.7176/jlpg/91-18.

[247] G. Bangura, "The European Union Artificial Intelligence Act: Mitigating Discrimination In Artificial Intelligence Systems," SSRN Electronic Journal , Jan. 2025, doi: 10.2139/ssrn.5087870.

[248] J. R. R. Kumar, A. Kalnawat, A. M. Pawar, V. D. Jadhav, P. Srilatha, and V. Khetani, "Transparency in Algorithmic Decision-making: Interpretable Models for Ethical Accountability," E3S Web of Conferences , vol. 491, p. 2041, Jan. 2024, doi: 10.1051/e3sconf/202449102041.

[249] J. Matthews, "Patterns and Antipatterns, Principles, and Pitfalls: Accountability and Transparency in Artificial Intelligence," AI Magazine , vol. 41, no. 1, p. 82, Mar. 2020, doi: 10.1609/aimag.v41i1.5204.

[250] Q. V. Liao and J. W. Vaughan, "AI Transparency in the Age of LLMs: A Human-Centered Research Roadmap," Harvard Data Science Review , Feb. 2024, doi: 10.1162/99608f92.8036d03b.

[251] H. Felzmann, E. Fosch-Villaronga, C. Lutz, and A. Tamò-Larrieux, "Towards Transparency by Design for Artificial Intelligence," Science and Engineering Ethics , vol. 26, no. 6. Springer Science+Business Media, p. 3333, Nov. 16, 2020. doi: 10.1007/s11948-020-00276-4.

[252] B. Li et al. , "Trustworthy AI: From Principles to Practices," arXiv (Cornell University) , Jan. 2021, doi: 10.48550/arXiv.2110.01167.

[253] M. Almada, "Governing the Black Box of Artificial Intelligence," SSRN Electronic Journal , Jan. 2023, doi: 10.2139/ssrn.4587609.

[254] B. Vaassen, "AI, Opacity, and Personal Autonomy," Philosophy & Technology , vol. 35, no. 4, Sep. 2022, doi: 10.1007/s13347-022-00577-5.

[255] A. Mei, M. Saxon, S. Chang, Z. C. Lipton, and W. Y. Wang, "Users are the North Star for AI Transparency," arXiv (Cornell University) , Jan. 2023, doi: 10.48550/arxiv.2303.05500.

[256] G. Wang and R. Pea, "Algorithmic Autonomy in Data-Driven AI," arXiv (Cornell University) , Nov. 2024, doi: 10.48550/arxiv.2411.05210.

[257] W. Lu, "Inevitable challenges of autonomy: ethical concerns in personalized algorithmic decision-making," Humanities and Social Sciences Communications , vol. 11, no. 1, Oct. 2024, doi: 10.1057/s41599-024-03864-y.

[258] W. Seymour, M. V. Kleek, R. Binns, and D. Murray-Rust, "Respect as a Lens for the Design of AI Systems," p. 641, Jul. 2022, doi: 10.1145/3514094.3534186.

[259] A. Kesari, D. Sele, E. Ash, and S. Bechtold, "A Legal Framework for eXplainable Artificial Intelligence," Jan. 2024, doi: 10.2139/ssrn.4972085.

[260] R. M. J. Byrne, "Good Explanations in Explainable Artificial Intelligence (XAI): Evidence from Human Explanatory Reasoning," p. 6536, Aug. 2023, doi: 10.24963/ijcai.2023/733.
[261] R. Deinhammer, "Can Natural Law Ethics be Tenable Today? Towards a Critical Natural Law Theory," The Heythrop Journal , vol. 62, no. 3, p. 511, Jul. 2016, doi: 10.1111/heyj.12345.
[262] D. Nicolitsas, "Fairness, expectations and life satisfaction: evidence from Europe," Empirica , vol. 51, no. 2, p. 313, Jan. 2024, doi: 10.1007/s10663-023-09602-y.
[263] R. Warner and R. H. Sloan, "Making Artificial Intelligence Transparent: Fairness and the Problem of Proxy Variables," Criminal Justice Ethics , vol. 40, no. 1, p. 23, Jan. 2021, doi: 10.1080/0731129x.2021.1893932.
[264] L. Edwards and M. Veale, "Enslaving the Algorithm: From a 'Right to an Explanation' to a 'Right to Better Decisions'?," IEEE Security & Privacy , vol. 16, no. 3, p. 46, May 2018, doi: 10.1109/msp.2018.2701152.
[265] L. Metikoš and I. van Domselaar, "Procedural Justice and Judicial AI; Substantiating Explainability Rights with the Values of Contestation," Jan. 2025, doi: 10.2139/ssrn.5242905.
[266] Z. Porter et al. , "Unravelling Responsibility for Ai," Jan. 2024, doi: 10.2139/ssrn.4871675.
[267] H. Altehenger, L. Menges, and P. Schulte, "How AI Systems Can Be Blameworthy," Philosophia , vol. 52, no. 4, p. 1083, Sep. 2024, doi: 10.1007/s11406-024-00779-5.
[268] I. Taylor, "Is explainable AI responsible AI?," AI & Society , vol. 40, no. 3, p. 1695, Apr. 2024, doi: 10.1007/s00146-024-01939-7.
[269] A. Kierans, K. Rittichier, U. Sonsayar, and A. Ghosh, "Catastrophic Liability: Managing Systemic Risks in Frontier AI Development," 2025, doi: 10.48550/ARXIV.2505.00616.
[270] H. Smith, "Clinical AI: opacity, accountability, responsibility and liability," AI & Society , vol. 36, no. 2, p. 535, Jul. 2020, doi: 10.1007/s00146-020-01019-6.
[271] K. Schweighofer et al. , "Safe and Certifiable AI Systems: Concepts, Challenges, and Lessons Learned," 2025, doi: 10.48550/ARXIV.2509.08852.
[272] B. B. Arcila, "AI Liability Along the Value Chain," SSRN Electronic Journal , Jan. 2025, doi: 10.2139/ssrn.5209735.
[273] A. T. da Fonseca, E. V. de Sequeira, and L. B. Xavier, "Liability for AI Driven Systems," in Law, governance and technology series , Springer International Publishing, 2023, p. 299. doi: 10.1007/978-3-031-41264-6_16.
[274] A. G. Yaşar et al. , "AI and the EU Digital Markets Act: Addressing the Risks of Bigness in Generative AI," arXiv (Cornell University) , Jul. 2023, doi: 10.48550/arxiv.2308.02033.
[275] N. de Marcellis-Warin, F. Marty, E. Thelisson, and T. Warin, "Artificial Intelligence and Market Manipulations: Ex-ante Evaluation in the Regulator's Arsenal," SSRN Electronic Journal , Jan. 2020, doi: 10.2139/ssrn.3744497.
[276] D. Liga, "The Interplay Between Lawfulness and Explainability in the Automated Decision-Making of EU Administration," in Oxford University Press eBooks , Oxford University Press, 2024, p. 239. doi: 10.1093/9780198919575.003.0009.
[277] B. Goodman and S. Flaxman, "EU regulations on algorithmic decision-making and a 'right to explanation'.," arXiv (Cornell University) , Jun. 2016, doi: 10.48550/arXiv.1606.08813.
[278] Y. Wang, "Ethical considerations of AI in financial decision," Computing and artificial intelligence. , vol. 2, no. 1, p. 1290, Jun. 2024, doi: 10.59400/cai.v2i1.1290.
[279] C. González-Gonzalo et al. , "Trustworthy AI: Closing the gap between development and integration of AI systems in ophthalmic practice," Progress in Retinal and Eye Research , vol. 90. Elsevier BV, p. 101034, Dec. 10, 2021. doi: 10.1016/j.preteyeres.2021.101034.
[280] P. S. Menell et al. , "Trade Secret Case Management Judicial Guide," SSRN Electronic Journal , Jan. 2023, doi: 10.2139/ssrn.4360102.

[281] J. Chally, "The Law of Trade Secrets: Toward a More Efficient Approach," Vanderbilt law review , vol. 57, no. 4, p. 1269, May 2004, Accessed: Aug. 2025. [Online]. Available: https://scholarship.law.vanderbilt.edu/cgi/viewcontent.cgi?article=1682&context=vlr

[282] R. Binns and M. Veale, "Is that your final decision? Multi-stage profiling, selective effects, and Article 22 of the GDPR," International Data Privacy Law , vol. 11, no. 4, p. 319, Sep. 2021, doi: 10.1093/idpl/ipab020.

[283] G. Malgieri and G. Comandé, "Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation," International Data Privacy Law , vol. 7, no. 4, p. 243, Oct. 2017, doi: 10.1093/idpl/ipx019.

[284] M. Veale and L. Edwards, "Clarity, surprises, and further questions in the Article 29 Working Party draft guidance on automated decision-making and profiling," Computer Law & Security Review , vol. 34, no. 2, p. 398, Jan. 2018, doi: 10.1016/j.clsr.2017.12.002.

[285] Z. Bednarz and L. Przhedetsky, "AI Opacity in the Financial Industry and How to Break It," in Cambridge University Press eBooks , Cambridge University Press, 2023, p. 70. doi: 10.1017/9781009334297.007.

[286] D. Martens et al. , "Beware of 'Explanations' of AI," 2025, doi: 10.48550/ARXIV.2504.06791.

[287] W. J. von Eschenbach, "Transparency and the Black Box Problem: Why We Do Not Trust AI," Philosophy & Technology , vol. 34, no. 4, p. 1607, Sep. 2021, doi: 10.1007/s13347-021-00477-0.

[288] D. G. Grant, J. Behrends, and J. Basl, "What we owe to decision-subjects: beyond transparency and explanation in automated decision-making," Philosophical Studies , Aug. 2023, doi: 10.1007/s11098-023-02013-6.

[289] S. Bordt, M. Finck, É. Raidl, and U. von Luxburg, "Post-Hoc Explanations Fail to Achieve their Purpose in Adversarial Contexts," in 2022 ACM Conference on Fairness, Accountability, and Transparency , Jun. 2022, p. 891. doi: 10.1145/3531146.3533153.

[290] M. Ebers, "Truly Risk-based Regulation of Artificial Intelligence How to Implement the EU's AI Act," European Journal of Risk Regulation , p. 1, Nov. 2024, doi: 10.1017/err.2024.78.

[291] A. Vedder and D. Spajić, "Moral autonomy of patients and legal barriers to a possible duty of health related data sharing," Ethics and Information Technology , vol. 25, no. 1, Mar. 2023, doi: 10.1007/s10676-023-09697-8.

[292] A. E. Holm, "Corrective Corporate Responses to Accusations of Misconduct on Societal Issues," Research Portal Denmark , p. 273, Jan. 2022, Accessed: Aug. 2025. [Online]. Available: https://local.forskningsportal.dk/local/dki-cgi/ws/cris-link?src=cbs&id=cbs-36ed4f2a-f46c-4d0a-94ad-d3249c4b28b3&ti=Corrective%20Corporate%20Responses%20to%20Accusations%20of%20Misconduct%20on%20Societal%20Issues

[293] D. Baum, K. Baum, S. Zamani, C. Bennoit, and D. Werth, "Transparent Transparency Developing a Scheme for Understanding Transparency Requirements," in Lecture notes in computer science , Springer Science+Business Media, 2024, p. 55. doi: 10.1007/978-3-031-75434-0_5.

[294] H. Felzmann, E. Fosch-Villaronga, C. Lutz, and A. Tamò-Larrieux, "Robots and Transparency: The Multiple Dimensions of Transparency in the Context of Robot Technologies," IEEE Robotics & Automation Magazine , vol. 26, no. 2, p. 71, Jun. 2019, doi: 10.1109/mra.2019.2904644.

[295] A. Bell, O. Nov, and J. Stoyanovich, "Think about the stakeholders first! Toward an algorithmic transparency playbook for regulatory compliance," Data & Policy , vol. 5, Jan. 2023, doi: 10.1017/dap.2023.8.

[296] G. Falco et al. , "Governing AI safety through independent audits," Nature Machine Intelligence , vol. 3, no. 7, p. 566, Jul. 2021, doi: 10.1038/s42256-021-00370-7.

[297] Y. Qi, B. Schölkopf, and Z. Jin, "Causal Responsibility Attribution for Human-AI Collaboration," arXiv (Cornell University) , Nov. 2024, doi: 10.48550/arxiv.2411.03275.
[298] M. Dastani and V. Yazdanpanah, "Responsibility of AI Systems," AI & Society , vol. 38, no. 2, p. 843, Jun. 2022, doi: 10.1007/s00146-022-01481-4.
[299] G. Dodig-Crnković, G. Basti, and T. Holstein, "Delegating Responsibilities to Intelligent Autonomous Systems: Challenges and Benefits," Journal of Bioethical Inquiry , vol. 22, no. 3, p. 507, May 2025, doi: 10.1007/s11673-025-10428-5.
[300] S. Vallor and T. Vierkant, "Find the Gap: AI, Responsible Agency and Vulnerability," Minds and Machines , vol. 34, no. 3, Jun. 2024, doi: 10.1007/s11023-024-09674-0.
[301] A. Ibitoye, M. Nkwo, and R. Orji, "Rethinking responsible AI from ethical pillars to sociotechnical practice," AI and Ethics , Aug. 2025, doi: 10.1007/s43681-025-00809-2.
[302] I. Taylor, "Collective Responsibility and Artificial Intelligence," Philosophy & Technology , vol. 37, no. 1, Feb. 2024, doi: 10.1007/s13347-024-00718-y.
[303] G. Lima and M. Cha, "Responsible AI and Its Stakeholders," arXiv (Cornell University) , Jan. 2020, doi: 10.48550/arxiv.2004.11434.
[304] H. Fraser and N. Suzor, "Locating Fault for AI Harms: A Systems Theory of Foreseeability, Reasonable Care and Causal Responsibility in the AI Value Chain," SSRN Electronic Journal , Jan. 2025, doi: 10.2139/ssrn.5190797.
[305] H. Fraser and N. Suzor, "Locating fault for AI harms: a systems theory of foreseeability, reasonable care and causal responsibility in the AI value chain," Law Innovation and Technology , p. 1, Mar. 2025, doi: 10.1080/17579961.2025.2469345.
[306] J. D. Bruyne, O. Dheu, and C. Ducuing, "The European Commission's approach to extra-contractual liability and AI – An evaluation of the AI liability directive and the revised product liability directive," Computer Law & Security Review , vol. 51, p. 105894, Oct. 2023, doi: 10.1016/j.clsr.2023.105894.
[307] N. Th. Nikolinakos, "Reforming the EU Civil Liability Framework Applicable to Artificial Intelligence and Other Emerging Digital Technologies: The Proposed AI Liability Directive," in Law, governance and technology series , Springer International Publishing, 2024, p. 377. doi: 10.1007/978-3-031-67969-8_8.
[308] M. J. Okuno and H. G. Okuno, "Legal Frameworks for AI Service Business Participants: A Comparative Analysis of Liability Protection Across Jurisdictions," Jan. 2025, doi: 10.2139/ssrn.5170869.
[309] M. Herbosch, "Liability for AI Agents," SSRN Electronic Journal , Jan. 2025, doi: 10.2139/ssrn.5236649.
[310] C. Novelli, G. Bongiovanni, and G. Sartor, "A conceptual framework for legal personality and its application to AI," Jurisprudence , vol. 13, no. 2, p. 194, Dec. 2021, doi: 10.1080/20403313.2021.2010936.
[311] B. Custers, H. Lahmann, and B. I. Scott, "From liability gaps to liability overlaps: shared responsibilities and fiduciary duties in AI and other complex technologies," AI & Society , vol. 40, no. 5, p. 4035, Jan. 2025, doi: 10.1007/s00146-024-02137-1.
[312] A. Aguirre, G. Dempsey, H. Surden, and P. B. Reiner, "AI Loyalty: A New Paradigm for Aligning Stakeholder Interests," IEEE Transactions on Technology and Society , vol. 1, no. 3, p. 128, Aug. 2020, doi: 10.1109/tts.2020.3013490.
[313] B. Judge, M. Nitzberg, and S. D. Russell, "When code isn't law: rethinking regulation for artificial intelligence," Policy and Society , May 2024, doi: 10.1093/polsoc/puae020.
[314] A. Reuel, L. Soder, B. Bucknall, and T. A. Undheim, "Position Paper: Technical Research and Talent is Needed for Effective AI Governance," arXiv (Cornell University) , Jun. 2024, doi: 10.48550/arxiv.2406.06987.
[315] M. Moloney et al. , "Ethical Artificial Intelligence, Fintech and Data Protection: A Path Forward for Training in Europe," Jan. 2024, doi: 10.2139/ssrn.4885037.

[316] E. Gorian and N. D. Osman, "DIGITAL ETHICS OF ARTIFICIAL INTELLIGENCE (AI) IN SAUDI ARABIA AND UNITED ARAB EMIRATES," Malaysian Journal of Syariah and Law , vol. 12, no. 3, p. 583, Dec. 2024, doi: 10.33102/mjsl.vol12no3.798.

[317] E. G. C. Bernal, "Harmonization of cross-border contract law: Legal solutions in a globalized world," vol. 2, no. 2, p. 2273, Dec. 2024, doi: 10.59400/jps2273.

[318] M. Lehmann, "Legal Fragmentation, Extraterritoriality and Uncertainty in Global Financial Regulation," Oxford Journal of Legal Studies , vol. 37, no. 2, p. 406, Oct. 2016, doi: 10.1093/ojls/gqw028.

[319] D. Ahern, "The New Anticipatory Governance Culture for Innovation: Regulatory Foresight, Regulatory Experimentation and Regulatory Learning," European Business Organization Law Review , Apr. 2025, doi: 10.1007/s40804-025-00348-7.

[320] S. Oh, C. Choi, and J.-H. Lee, "Regulatory Reform for Fostering an Innovation Ecosystem," SSRN Electronic Journal , Jan. 2020, doi: 10.2139/ssrn.4054295.

[321] A. R. Gatt, P. V. Bonanno, and R. Zammit, "Ethical considerations in the regulation and use of herbal medicines in the European Union," Frontiers in Medical Technology , vol. 6, Jun. 2024, doi: 10.3389/fmedt.2024.1358956.

[322] S. Cha, "Towards an international regulatory framework for AI safety: lessons from the IAEA's nuclear safety regulations," Humanities and Social Sciences Communications , vol. 11, no. 1, Apr. 2024, doi: 10.1057/s41599-024-03017-1.

[323] R. Trager et al. , "International Governance of Civilian AI: A Jurisdictional Certification Approach," arXiv (Cornell University) , Jan. 2023, doi: 10.48550/arxiv.2308.15514.

[324] J. Vujicic, "Global AI regulation and its impact on technology business: A comparative legal framework analysis," World Journal of Advanced Research and Reviews , vol. 24, no. 3, p. 3457, Dec. 2024, doi: 10.30574/wjarr.2024.24.3.3911.

[325] S. Park, "Bridging the Global Divide in AI Regulation: A Proposal for a Contextual, Coherent, and Commensurable Framework," SSRN Electronic Journal , Jan. 2024, doi: 10.2139/ssrn.4950781.

[326] A. Nguyen, H. N. Ngo, Y. Hong, B. Dang, and B.-P. T. Nguyen, "Ethical principles for artificial intelligence in education," Education and Information Technologies , vol. 28, no. 4, p. 4221, Oct. 2022, doi: 10.1007/s10639-022-11316-w.

[327] N. Tomašev et al. , "AI for social good: unlocking the opportunity for positive impact," Nature Communications , vol. 11, no. 1. Nature Portfolio, May 18, 2020. doi: 10.1038/s41467-020-15871-z.

[328] N. AllahRakha, "UNESCO's AI Ethics Principles: Challenges and Opportunities," International Journal of Law and Policy , vol. 2, no. 9, p. 24, Sep. 2024, doi: 10.59022/ijlp.225.

[329] F. Morandín-Ahuerma, "Ten UNESCO Recommendations on the Ethics of Artificial Intelligence," Sep. 2023, doi: 10.31219/osf.io/csyux.

[330] E. H. Morawska, "Council of Europe standards and activities related to AI: towards a Framework Convention on AI and human rights?," in Edward Elgar Publishing eBooks , Edward Elgar Publishing, 2024, p. 25. doi: 10.4337/9781035337934.00009.

[331] D. Leslie, C. Burr, M. Aitken, M. Katell, M. Briggs, and C. Rincón, "Human Rights, Democracy, and the Rule of Law Assurance Framework for AI Systems: A Proposal," SSRN Electronic Journal , Jan. 2021, doi: 10.2139/ssrn.4027875.

[332] S. Kiden et al. , "Responsible AI Governance: A Response to UN Interim Report on Governing AI for Humanity," Nov. 2024, doi: 10.48550/arxiv.2412.12108.

[333] Y. Bengio, "The First International AI Safety Report," vol. 2, no. 2, Jun. 2025, doi: 10.70777/si.v2i2.14755.

[334] C. Colther, J. P. Doussoulin, and G. Tontini, "Artificial Intelligence and Global Power Dynamics: Geopolitical Competition, Strategic Alliances, and the Future of Ai Governance," Jan. 2025, doi: 10.2139/ssrn.5251303.

[335] C. Sun and X. Chen, "Destined for Balance? Centralized and Decentralized Approaches to AI Governance," Politics and Governance , vol. 13, Oct. 2025, doi: 10.17645/pag.10197.
[336] J. Ding, "Dueling Perspectives in AI and U.S.–China Relations," in Oxford University Press eBooks , Oxford University Press, 2022, p. 881. doi: 10.1093/oxfordhb/9780197579329.013.53.
[337] H. Roberts, E. Hine, and L. Floridi, "Digital Sovereignty, Digital Expansionism, and the Prospects for Global AI Governance," in Philosophical studies series , Springer International Publishing, 2023, p. 51. doi: 10.1007/978-3-031-41566-1_4.
[338] N. Mishra and K. Kugler, "INTERNATIONAL COMMUNITY IN THE GLOBAL DIGITAL ECONOMY: A CASE STUDY ON THE AFRICAN DIGITAL TRADE FRAMEWORK," International and Comparative Law Quarterly , p. 1, Nov. 2024, doi: 10.1017/s0020589324000290.
[339] Technology and Innovation Report 2025 . United Nations, 2025. doi: 10.18356/9789211068016.
[340] P. Carvlão, "The dual imperative: innovation and regulation in the AI era," International Journal of Technology Policy and Law , vol. 3, no. 3, p. 236, Jan. 2024, doi: 10.1504/ijtpl.2024.10068130.
[341] Q. Mao and M. Xu, "Navigating the Ethical Landscape of AI Innovation: Challenges and Opportunities," Deleted Journal , Jun. 2025, doi: 10.1515/ijdlg-2025-0005.
[342] B. Anderson and E. Sutherland, "Collective action for responsible AI in health," Jan. 2024. doi: 10.1787/f2050177-en.
[343] I. Bose, A. K. Kar, G. M. Lee, A. C. M. Leung, and D. J. Kim, "Artificial Intelligence and the Bright Internet: Transforming the Internet for a Better Future," Information Systems Frontiers , Aug. 2025, doi: 10.1007/s10796-025-10632-z.
[344] A. Donati, "The Precautionary Principle Under EU Law : The Knots and the Links of its Network," SSRN Electronic Journal , Jan. 2022, doi: 10.2139/ssrn.4026140.
[345] Y. Bengio, "Government Interventions to Avert Future Catastrophic AI Risks," Harvard Data Science Review , Apr. 2024, doi: 10.1162/99608f92.d949f941.
[346] J. S. Gans, "How Learning About Harms Impacts the Optimal Rate of Artificial Intelligence Adoption," Feb. 2024. doi: 10.3386/w32105.
[347] K. A. Kilian, "Beyond Accidents and Misuse: Decoding the Structural Risk Dynamics of Artificial Intelligence," arXiv (Cornell University) , Jun. 2024, doi: 10.48550/arxiv.2406.14873.
[348] A. R. Wasil, J. Clymer, D. Krueger, E. Dardaman, S. Campos, and E. R. Murphy, "Affirmative safety: An approach to risk management for high-risk AI," arXiv (Cornell University) , Apr. 2024, doi: 10.48550/arxiv.2406.15371.
[349] G. Malgieri and F. Pasquale, "Licensing high-risk artificial intelligence: Toward ex ante justification for a disruptive technology," Computer law & security review , vol. 52, p. 105899, Nov. 2023, doi: 10.1016/j.clsr.2023.105899.
[350] A. D. Thierer, "Getting AI Innovation Culture Right," SSRN Electronic Journal , Jan. 2023, doi: 10.2139/ssrn.4404402.
[351] A. D. Thierer, "Permissionless Innovation and Public Policy: A 10-Point Blueprint," SSRN Electronic Journal , Jan. 2016, doi: 10.2139/ssrn.2761139.
[352] M. Park, S. Wu, and R. J. Funk, "Regulation and innovation revisited: How restrictive environments can promote destabilizing new technologies," Organization Science , Aug. 2024, doi: 10.1287/orsc.2022.16770.
[353] D. Acemoğlu and T. Lensman, "Regulating Transformative Technologies," Jul. 2023. doi: 10.3386/w31461.
[354] A. Bradford, "The False Choice Between Digital Regulation and Innovation," SSRN Electronic Journal , Jan. 2024, doi: 10.2139/ssrn.4753107.
[355] V. Kulothungan and D. Gupta, "Towards Adaptive AI Governance: Comparative Insights from the U.S., EU, and Asia," 2025, doi: 10.48550/ARXIV.2504.00652.

[356] G. K. Hadfield and J. A. Clark, "Regulatory Markets: The Future of AI Governance," arXiv (Cornell University) , Jan. 2023, doi: 10.48550/arxiv.2304.04914.
[357] P. Brey, "The strategic role of technology in a good society," Technology in Society , vol. 52, p. 39, Feb. 2017, doi: 10.1016/j.techsoc.2017.02.002.
[358] D. Crouch, "Transhumanism within the Natural Law: Transforming Creation with Nature as Guide," Religions , vol. 15, no. 8, p. 949, Aug. 2024, doi: 10.3390/rel15080949.
[359] P. Woodward, "Technological Innovation and Natural Law," Philosophia Reformata , vol. 85, no. 2, p. 138, Nov. 2020, doi: 10.1163/23528230-8502a001.
[360] G. Hermerén, "The principle of proportionality revisited: interpretations and applications," Medicine Health Care and Philosophy , vol. 15, no. 4, p. 373, Nov. 2011, doi: 10.1007/s11019-011-9360-x.
[361] S. A and D. Prasad M, "AI as an Inventor Debate under the Patent Law." 2025.
[362] "Current ethical challenges of artificial intelligence in the context of human work," Scientific Papers of Silesian University of Technology Organization and Management Series , vol. 2024, no. 211, Jan. 2024, doi: 10.29119/1641-3466.2024.211.14.
[363] H. Fraser and J.-M. B. y Villarino, "Acceptable risks in Europe's proposed AI Act: Reasonableness and other principles for deciding how much risk management is enough," arXiv (Cornell University) , Jan. 2023, doi: 10.48550/arxiv.2308.02047.
[364] J. Clark and G. K. Hadfield, "Regulatory Markets for AI Safety," arXiv (Cornell University) , Jan. 2020, doi: 10.48550/arXiv.2001.00078.
[365] O. Akpobome, "The Impact of Emerging Technologies on Legal Frameworks: A Model for Adaptive Regulation," International Journal of Research Publication and Reviews , vol. 5, no. 10, p. 5046, Oct. 2024, doi: 10.55248/gengpi.5.1024.3012.
[366] J. Braithwaite, "The Healthcare Regulatory Ecosystem," in SpringerBriefs in applied sciences and technology , Springer Nature, 2024, p. 33. doi: 10.1007/978-3-031-49570-0_4.
[367] OECD, Regulatory sandbox toolkit . 2025. doi: 10.1787/de36fa62-en.
[368] A. Noya, "Regulatory sandboxes in artificial intelligence," Jul. 2023. doi: 10.1787/8f80a0e6-en.
[369] C. Crampes and A. Estache, "Efficiency vs. equity concerns in regulatory sandboxes," Research Square (Research Square) , Dec. 2023, doi: 10.21203/rs.3.rs-3706419/v1.
[370] E. Leckenby, D. Dawoud, J. C. Bouvy, and P. Jónsson, "The Sandbox Approach and its Potential for Use in Health Technology Assessment: A Literature Review," Applied Health Economics and Health Policy , vol. 19, no. 6. Adis, Springer Healthcare, p. 857, Jul. 13, 2021. doi: 10.1007/s40258-021-00665-1.
[371] R. Plato-Shinar and A. Godwin, "Regulatory Cooperation in AI Sandboxes: Insights from Fintech," SSRN Electronic Journal , Jan. 2025, doi: 10.2139/ssrn.5199887.
[372] W. Wu and S. Liu, "Compliance Costs of AI Technology Commercialization: A Field Deployment Perspective," arXiv (Cornell University) , Jan. 2023, doi: 10.48550/arxiv.2301.13454.
[373] F. Trebbi and M. B. Zhang, "The Cost of Regulatory Compliance in the United States," Nov. 2022. doi: 10.3386/w30691.
[374] C. M. Pinckaers, "Toward a Holistic Understanding of AI and ESG: Exploring Their Interplay and Risks." 2024.
[375] L. Pöhler, K. Diepold, and W. Wallach, "A Practical Multilevel Governance Framework for Autonomous and Intelligent Systems," arXiv (Cornell University) , Apr. 2024, doi: 10.48550/arxiv.2404.13719.
[376] E. Brynjolfsson et al. , "The Digital Welfare of Nations: New Measures of Welfare Gains and Inequality," SSRN Electronic Journal , Jan. 2023, doi: 10.2139/ssrn.4567676.

[377] OECD and W. T. Organization, Economic Implications of Data Regulation . 2024. doi: 10.1787/aa285504-en.
[378] H. Herrmann, "What's next for responsible artificial intelligence: a way forward through responsible innovation," Heliyon , vol. 9, no. 3, Mar. 2023, doi: 10.1016/j.heliyon.2023.e14379.
[379] R. Frodeman, "International handbook on responsible innovation. A global resource," Journal of Responsible Innovation , vol. 6, no. 2, p. 255, May 2019, doi: 10.1080/23299460.2019.1613612.
[380] T. A. Bach, M. Kaarstad, E. Solberg, and A. Babic, "Insights into suggested Responsible AI (RAI) practices in real-world settings: a systematic literature review," AI and Ethics , Jan. 2025, doi: 10.1007/s43681-024-00648-7.
[381] C. J. van Rees and B. Müller, "All that glitters is not gold: trustworthy and ethical AI principles," AI and Ethics , vol. 3, no. 4, p. 1241, Nov. 2022, doi: 10.1007/s43681-022-00232-x.
[382] S. Dick, W. H. K. Chun, and M. Canute, "How AI Sets Limits for Human Aspiration," Issues in Science and Technology , vol. 40, no. 2, p. 87, Dec. 2023, doi: 10.58875/zclm4010.
[383] A. Reuel and T. A. Undheim, "Generative AI Needs Adaptive Governance," arXiv (Cornell University) , Jun. 2024, doi: 10.48550/arxiv.2406.04554.