Ordine internazionale e diritti umani
International Legal Order and Human Rights
Ordenamiento Jurídico Internacional y Derechos Humanos
Ordre Juridique International et Droits de l'Homme
Diretta da Claudio Zanghì, Lina Panella, Carlo Curti Gialdino
EDITORIALE SCIENTIFICA

ALESSIA DEL PIZZO*

SHAPING FAIR AI: OVERCOMING ALGORITHMIC *BIAS* AND FOSTERING INCLUSION

SUMMARY: 1. Artificial intelligence as *locus* of conflict between efficiency and justice. – 2. *Bias* and vulnerability: automated discrimination in critical domains. – 3. The role of law: from formal equality to algorithmic inclusion. – 4. Conclusions.

1. *Artificial intelligence as locus of conflict between efficiency and justice*

At the heart of contemporary constitutional democracies lurks a profound tension between the ideal of efficiency – exalted by algorithmic logic[1] – and the principle of substantive justice, the cornerstone of the modern legal system[2]. Artificial intelligence (AI), in its progressive penetration into vital sectors such as healthcare, criminal justice and education, does not merely act as a technical aid, but tends to transform itself into a silent regulatory actor, capable of modulating rights and redefining the relationships between individuals and institutions.

The emergence of AI as an enabling technology marks, in fact, a dislocation of the decision-making function: from the human subject – historically invested with legal responsibility, moral awareness and deliberative capacity – to automated systems that operate according to statistical-probabilistic logic. The concept of "*enabling technology*" should be

* PH.D in Law and Cognitive Neuroscience, Niccolò Cusano University, Rome, Italy.

[1] Algorithmic logic is a set of formal and computational rules that govern the transformation of *input* into *output* by means of a finite sequence of deterministic or probabilistic operations, encoded in an algorithm, and oriented towards the maximisation of efficiency, internal consistency and repeatability of decisions. It replaces or complements human forms of deliberation with automatic mechanisms based on data, models and objective functions. D. E. KNUTH, *The Art of Computer Programming*, Boston, 1997.

[2] The tension between the ideal of efficiency and substantive justice is a central issue in the contemporary context, in which legal and social institutions are called upon to reconcile economic objectives with ethical values. Efficiency, often understood as the optimisation of time and costs, sometimes runs the risk of clashing with the demands of fairness and justice. While it pursues the quick and pragmatic resolution of conflicts, justice requires a deeper analysis of the circumstances and consequences of decisions. For a more in-depth look at the topic, see also F. MEDICO, *Il doppio custode*, in *Sociologia del diritto* (on-line), 2023, p. 1 ff., available at: https://doi.org/10.30682/sg328.

understood in the proper sense: AI does not merely support or speed up already existing activities, but creates the conditions for the structural transformation of decision-making, legal and administrative processes. It enables – i.e. makes possible – new forms of "*governance*", categorisation and intervention, often removing them from the explicit control of the human being. This shift, only apparently neutral, raises numerous axiological and constitutional questions.

If, on the one hand, AI is celebrated for its ability to process large volumes of data in a very short time, offering seemingly optimised solutions to complex problems, on the other, it introduces a new rationality into public decision-making processes, based on performance and predictivity criteria, which risks circumventing or compressing the principles of equality, proportionality and inclusion.

In the absence of conscious regulatory structuring, the promise of efficiency may turn into a silent multiplier of inequalities, capable of consolidating systemic forms of exclusion under the veil of technical objectivity.

In the framework of the most advanced studies on the interaction between law and technology, AI emerges not as a mere tool, but as a true "*epistemic technology*", capable of reshaping the very construction of social reality[3] . Mireille Hildebrandt has clarified how such systems do not merely manage data, but define in advance what is relevant, which subjects are recognised for legal purposes, and which attributes count in the assessment of reality[4] . In this perspective, the algorithm takes on the function of an implicit norm: it already operates upstream of decisions, defining the criteria for selecting, categorising and weighing data, and selects what can enter the space of the considerable and the judiciable.

In other words, the algorithm's decision-making power does not only manifest itself in the executive phase of the judgement, but acts upstream in the prefiguration of the conditions that make the decision itself possible. Each decision-making model, in fact, does not merely provide computational outcomes, but organises the field of the relevant, defining semantic priorities, activating selective filters and delimiting what can be considered visible and relevant for the deliberative outcome[5]. In this function, it does not assume a prescriptive guise in the classical sense of the norm, but embodies a structural normativity: silent, opaque, but capable of affecting the trajectories of recognition and access to rights.

It has been observed how this normativity presents itself under the appearance of technical neutrality, concealing, in reality, value choices embedded in the design of the algorithm[6]. The selection of data, the construction of categories, and the definition of operational thresholds are operations that reflect cultural and political decisions, and which, if not adequately scrutinised, tend to stabilise existing inequalities, removing them from critical scrutiny and the possibility of opposition. For this reason, the algorithm cannot be considered a neutral entity: it operates as a discriminating device, which requires legal instruments capable of revealing its logic and guaranteeing its democratic controllability.

The operational dimension of algorithms directly affects the distribution of opportunities and forms of social recognition: they include or exclude, they attribute visibility or determine marginalisation, shaping the cognitive and value space within which meanings

---

[3] M. HILDEBRANDT, *Smart Technologies and the End(s) of Law*, Cheltenham, 2015.

[4] M. HILDEBRANDT, *Law as Information in the Era of Data-Driven Agency*, in *Modern Law Review*, 2016.

[5] K. MARTIN, *Ethical Implications and Accountability of Algorithms*, in *Journal of Business Ethics*, 2018.

[6] On this point see: S. VENKATASUBRAMANIAN, *Structural disconnects between algorithmic decision-making and the law*, in *International Review of the Red Cross Blog*, 2019; M. KRANZBERG, *Technology and History: 'Kranzberg's Laws'*, in *Technology and Culture*, 1986, p. 544 ff.

are attributed and subjective positions are constructed. In this sense, each algorithmic system is configured as an interpretative grid of an axiological nature, capable of orienting institutional and intersubjective dynamics well before a normative elaboration or formalised decision is made.

This shift of the cognitive-decisional axis, operated by algorithmic architectures, highlights how the power exercised by AI is not exhausted in punctual or circumscribed outcomes, but extends to the systemic configuration of the spaces of the possible. It is precisely in the capacity to orientate the parameters of evaluation in advance – and thus to delimit what can become the object of consideration, attribution or exclusion – that its ordering power manifests itself. In this context, the theme of algorithmic normativity can no longer be confined to the sphere of justice or public administration alone, but transversally invests all the sectors in which automated decisions contribute to structuring the conditions of access to rights, resources and social recognition. Hence the need to question the rationality underlying these systems not only in functional terms, but in their axiological, normative and performative dimensions.

In this sense, the normativity that emanates from algorithmic systems does not take the prescriptive form typical of classical sources of law – such as laws or regulations – but manifests itself as a structural and diffuse normativity, analogous to that which Michel Foucault recognised in knowledge-power devices: it does not directly impose behaviour, but conditions the very premises of the thinkable and the practicable[7]. Although clothed with the authority of mathematical objectivity, such normativity is actually the outcome of design, technical and cultural choices embedded in the structure of the code, which are often neither visible nor contestable by users or stakeholders. Precisely for this reason, the algorithm tends to naturalise pre-existing inequalities, inscribing them in computational processes that guarantee their systematic and silent reproduction[8].

This awareness forces law to abandon a merely reactive or regulatory posture. The algorithm should not be seen as a neutral object to be "*disciplined*" but as a *"form of informational power"*, i.e. a mechanism that redistributes knowledge and opportunities, with direct implications on legal subjectivity and individual autonomy[9]. Law must therefore become a critical instrument of unveiling and direction, capable of questioning the epistemic assumptions of technology and opposing the apparent neutrality of efficiency with the criterion of public justifiability.

At European level, Regulation (EU) 2024/1689, known as the AI Act[10], marks a fundamental step in the construction of a unified regulatory framework on artificial intelligence, inspired by the principles of accountability, transparency and protection of fundamental rights. The regulation introduces a risk-based approach, distinguishing between low, medium and high-risk systems, and imposing specific obligations for the latter such as rights impact assessment, transparent technical documentation and traceability of the entire life cycle of the system. However, the binary classification between "*high risk*" and "*not high*

[7] M. FOUCAULT, *Sorvegliare e punire*, Turin, 1976.

[8] J. MONTEIRO, M. MARRAFON, *Legitimidade democrática na governança algorítmica*, in *Revista Direitos Fundamentais & Democracia*, 2024.

[9] J. E. COHEN, *Between Truth and Power: The Legal Constructions of Informational Capitalism*, 2019.

[10] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828

risk" may prove insufficient to capture the pervasive and transversal nature of algorithmic inequalities, which often manifest themselves even in contexts not formally labelled as critical.

A comparative look allows one to grasp the parallel evolution of AI law in other legal contexts. In the United States, the regulatory debate is focusing on the introduction of genuine "*algorithmic accountability*", which imposes transparency and verifiability obligations on all automated systems employed in sensitive sectors[11] . In Canada, the *Artificial Intelligence and Data Act* (AIDA)[12], originally part of Bill C-27 (2022), was supposed to establish a risk classification system and governance requirements for the development and responsible use of AI. However, the draft legislation lapsed in January 2025 with the early end of the federal legislature and the resignation of Prime Minister Trudeau[13]. Brazil also saw significant progress with *Projeto de Lei no. 2338/2023*, approved by the Senate on 10 December 2024 and currently under consideration by the Chamber of Deputies[14].

These developments, although heterogeneous, converge on a shared vision of the need to make artificial intelligence compatible with the principles of the rule of law: transparency, accountability, inclusion and effective protection of fundamental rights. Within this framework, the comparison not only highlights normative convergences between legal systems, but also offers critical criteria for enhancing the capacity of law to act on an increasingly hybrid terrain, where the algorithm tends to silently replace formal normativity. It is precisely from this comparative analysis that a common thread emerges: law cannot confine itself to intervening once algorithmic decisions have already produced their effects,

---

[11] In 2023, the bill known as the *Algorithmic Accountability Act* (S.2892), sponsored by Senators Ron Wyden and Cory Booker, was introduced in the US Senate. The text included significant obligations for companies that develop or use automated decision-making systems, requiring algorithmic impact assessments, audits on the risk of *bias*, and the obligation to make transparent the criteria used for user profiling in sectors considered sensitive (such as healthcare, criminal justice, education and employment). The bill also envisaged the creation of a public register of the systems used, under the supervision of the Federal Trade Commission, and the possibility for citizens to access the automated decision-making logic that affects them.
However, the proposal did not make it through the committees and was not passed by the end of the 118th Congress, thus lapsing. Nevertheless, it sets a relevant precedent for the legal debate on *algorithmic accountability*, as it represents the first organic attempt in the United States to regulate the transparency and accountability of artificial intelligence systems in a structured manner. The text has also helped lay the foundations for future legislative interventions, both at federal and state level, and has fostered a broader reflection on the balance between technological innovation and the protection of fundamental rights in the digital society. Algorithmic Accountability Act of 2023, S.2892, 118th Congress. Available at: https://*www.congress.gov/bill/118th-congress/senate-bill/2892*.

[12] Government of Canada, Artificial Intelligence and Data Act (AIDA) - Companion Document, https://ised-isde.canada.ca/site/innovation-better-canada/en/artificial-intelligence-and-data-act-aida-companion-document.

[13] Currently, the Canadian approach is articulated in local initiatives, such as Ontario's Bill 194, and non-binding instruments, including a voluntary Code of Conduct published by the federal government to promote ethical and transparent practices in the AI lifecycle.

[14] Brazil now ranks among the most active countries in Latin America in the construction of a legal framework for AI, inspired by the European AI Act but enriched by specific elements aimed at strengthening the protection of fundamental rights and the responsibility of providers. PL 2338/2023 is distinguished by the introduction of guaranteed digital rights - such as the right to know when interacting with an AI system, the right to an explanation of automated decisions, to challenge and to non-discrimination - and by the adoption of a risk-based classification of systems, which distinguishes between excessive risk (prohibited, as in the case of social scoring) and high risk, subject to transparency, impact assessment, data management and human oversight obligations. Senado Federal do Brasil, *Projeto de Lei no. 2338/2023*, https://www25.senado.leg.br/web/atividade/materias/-/materia/157233.

but must help shape the initial conditions within which those decisions take form. The asymmetries that AI tends to amplify indeed require legal involvement in the very configuration of such systems.

2. *Bias and vulnerability: automated discrimination in critical areas*

The contemporary legal debate on the use of algorithms in decision-making processes focuses in particular on two problematic nodes: the determination of legal responsibilities and the structural risk of *bias* inherent in machine learning models. In a decision-making ecosystem increasingly mediated by automated systems, it becomes essential to precisely identify legally clear criteria for the allocation of liability, especially in the presence of erroneous or discriminatory outcomes. The increasing adoption of artificial intelligence in strategic areas – such as the labour market, the provision of public services or the assessment of individual rights – has in fact raised serious questions in terms of accountability and transparency, exposing asymmetries between the decision-making power of algorithms and the ability of those involved to exercise effective control over such processes[15]. In the context of automated decision-making, the issue of accountability is one of the central challenges for the rule of law in the digital ecosystem. Traditionally, accountability in technology has been traced back to the individual developers or technical teams responsible for the design and implementation of systems[16]. However, this approach today proves inadequate with respect to the structural and organisational complexity of artificial intelligence systems, the effects of which are distributed along an articulated chain of actors, stages and application contexts. In this perspective, the AI act marks an important evolution, proposing a model of shared and distributed responsibility, which involves not only suppliers, but also users, importers, distributors and – for high-risk systems – even third parties in charge of conformity assessment. This configuration reflects a pluralistic conception of liability, based on transparency, traceability, monitoring and risk management obligations, which must be fulfilled throughout the system's life cycle, from design to post-marketing[17].

This evolution in the legal configuration of liability makes it all the more urgent to reflect systematically on the other major critical issue related to the use of artificial intelligence: that of algorithmic *bias*. If, on the one hand, the challenge is to ensure that the systems are controllable and legally accountable, on the other hand, it is necessary to question the nature of the data and logics that feed these systems, since it is precisely in these that the most insidious distortions lurk, capable of compromising the impartiality of automated decisions and profoundly affecting the effective protection of fundamental rights[18].

The concept of algorithmic *bias*, if from a technical-scientific point of view it can be considered a statistical distortion or an accidental dysfunction of learning models, requires –

---

[15] M. GÜNTHER, A. KASIRZADEH, *Algorithmic and human decision making: for a double standard of transparency*, in *AI & Society*, 2021, p. 375 ff.

[16] G. COMANDÉ, *Intelligenza artificiale e responsabilità tra «liability» e «accountability». Il carattere trasformativo dell'IA e il problema della responsabilità,* in *Analisi Giuridica dell'Economia*, 2019, 18(1), p. 169 ff.

[17] S. KLEANTHOUS, M. KASINIDOU, P. BARLAS, J. OTTERBACHER, *Perception of fairness in algorithmic decisions: future developers' perspective*, in *Patterns*, 2022.

[18] A.E. WALDMAN, *Algorithmic Legitimacy*, in W. BARFIELD (ed.), *The Cambridge Handbook of the Law of Algorithms*, Cambridge Law Handbooks, Cambridge University Press, Cambridge, 2020.

from a legal and philosophical-normative perspective – a much deeper and structural analysis. Reducing *bias* to a technical anomaly is tantamount to defusing its political and legal significance: it is not the unforeseen effect of an error, but the consistent outcome of systems that learn from data steeped in historical stratifications, social hierarchies and systemic inequalities. *Bias*, in this sense, does not emerge when the algorithm fails, but precisely when it functions as designed: it takes the form of automated discrimination, which translates and reproduces pre-existing social asymmetries into computational form. Training datasets are never neutral: they reflect worldviews, cultural representations and power relations[19].

From the point of view of fundamental rights, the algorithmic *bias* represents an indirect but concrete violation of the principle of non-discrimination, which is an inalienable pillar in any democratic system.

This principle is recognised and guaranteed both at the domestic constitutional level and at the supranational and international level. In modern constitutional democracies, it takes the form not only of a prohibition of arbitrary differential treatment, but also of a positive obligation for states to prevent and counteract all forms of systemic inequality, even when it manifests itself through technologically mediated mechanisms.

At European Union level, the principle is formally enshrined in Article 21 of the Charter of Fundamental Rights[20], which prohibits any form of discrimination based in particular on sex, race, colour, ethnic or social origin, genetic features, language, religion or belief, political or other opinion, membership of a national minority, property, birth, disability, age or sexual orientation. This precept, of direct applicability, imposes on legislators and administrations a reinforced obligation to scrutinise the indirect effects produced by the adoption of digital technologies, even when such effects cannot be traced back to a discriminatory intent. Article 21 of the Charter, interpreted in light of the case law of the Court of Justice, in fact requires a substantive understanding of the principle of equality: the assessment must not stop at the formal neutrality of measures, but must extend to the actual consequences they produce within society[21]. This approach, which emphasises the material dimension of equality, implies that even apparently neutral technological tools may be subject to scrutiny where, in their concrete application, they generate a differential impact on vulnerable groups.

Scholarly work has highlighted how the interaction between algorithmic systems and European anti-discrimination law reveals a structural risk: harmful effects often emerge not as explicit violations of the prohibition of discrimination, but as collateral outcomes of complex decision-making architectures, which are difficult to detect without an analysis directed at the results they produce. For this reason, the prohibition set out in Article 21 requires, according to a significant strand of contemporary studies, that public authorities adopt assessment tools capable of capturing forms of inequality that do not appear in the formulation of the rule, but in its social operation[22]. In this sense, the obligation to scrutinise

[19] As the US mathematician and data scientist C. O'Neil, "*algorithms are opinions embedded in code*". C. O'NEIL, *Weapons of Math Destruction*, New York, 2016.

[20] European Union, Charter of Fundamental Rights of the European Union, in OJEU C 326, 26 October 2012.

[21] On this point see: Court of Justice of the European Union, judgment of 16 July 2015, Case C-83/14, *CHEZ Razpredelenie Bulgaria AD*, ECLI:EU:C:2015:480; Court of Justice of the European Union, judgment of 11 April 2019, Case C-397/18, *DW v Nobel Plastiques Ibérica SA*, ECLI:EU:C:2019:309.

[22] Several authors have argued that a structural incompatibility exists between EU anti-discrimination law and the statistical fairness metrics developed in machine learning. While EU law requires a contextual and

does not amount to a merely procedural control, but to a form of substantive oversight requiring institutions to examine the mechanisms through which technologies may structure or amplify pre-existing inequalities.

This evolving interpretation of Article 21, which links the principle of non-discrimination to a positive duty to prevent indirect harmful effects, provides the logical foundation for the subsequent analysis of the European Convention framework. It allows one to understand how anti-discrimination protection within the European multilevel legal order operates through the control of the concrete outcomes generated by practices, procedures and ostensibly neutral criteria, regardless of whether they originate from public authorities or private actors. This focus on effects, rather than on subjective intent, finds a coherent counterpart in Article 14 of the European Convention on Human Rights (ECHR) and in the case law of the European Court of Human Rights on the relevance of indirect discrimination.

From this perspective, Article 14 of the ECHR prohibits any form of discrimination «in the enjoyment of the rights and freedoms set forth» in the Convention and, although ancillary in nature, requiring a connection with another Convention right[23], it has progressively acquired substantive weight in the Court's jurisprudence. The Court has clarified that the prohibition does not concern only intentional disparities, but also requires the examination of the effects produced by rules, practices or ostensibly neutral mechanisms where they systematically or foreseeably affect specific social groups. Within this interpretative framework, the Court has recognised the full legal relevance of indirect discrimination, qualifying as potentially incompatible with Article 14 all situations in which a disadvantage resulting from the uniform application of a measure translates, in practice, into a differential impact lacking an objective and reasonable justification[24].

Numerous empirical studies confirm the urgency of addressing the legal implications of algorithmic *bias* in areas with a high impact on fundamental rights. In healthcare, extensive research by Ziad Obermeyer et al. published in 2019 in *The New England Journal of Medicine* found that an algorithm widely employed by a major US insurance company to identify

substantive assessment of the differential effects produced by a given measure, technical metrics operate through standardised and abstract criteria that are unable to capture the complexity of indirect discrimination as recognised by the Court of Justice. In algorithmic systems, the social cues typically associated with human discrimination (such as stereotypes, prejudices or observable behaviours) are absent, which makes the legal identification of disparate treatment considerably more complex. On this basis, it has been proposed to adopt a baseline statistical measure, *conditional demographic disparity* (CDD), which may serve as a minimal reference point for detecting automated disparities while preserving the interpretative margin required by the European model of non-discrimination. S. WACHTER, B. D. MITTELSTADT, C. RUSSELL, *Why Fairness Cannot Be Automated: Bridging the Gap Between EU Non-Discrimination Law and AI*, in *arXiv* (online), 2020, arXiv:2005.05906.

[23] European Court of Human Rights, *, Guida all'articolo 14 della Convenzione*, updated edition 31 December 2020. Available:
www.giustizia.it/cmsresources/cms/documents/guida_cedu_articolo14_e_articolo1_protocollo12.pdf.

[24] A paradigmatic case is represented by the Grand Chamber judgment *D.H. and others v. Czech Republic*, in which the Court condemned the practice of systematically placing Roma children in special schools on the basis of psycho-aptitude tests that were formally neutral, but which were substantially discriminatory against that community (ECHR, Grand Chamber, *D.H. and others v. Czech Republic*, Appl. No. 57325/00, judgment of 13 November 2007, ECLI:CE:ECHR:2007:1113JUD005732500). In *Opuz v. Turkey*, the Court stated that the inaction of public authorities - in that case the absence of adequate protective measures towards a woman victim of domestic violence - may also constitute discrimination under Art. 14 ECHR, if it produces asymmetrical consequences to the detriment of vulnerable groups, such as women, even in the absence of a direct discriminatory intention (ECHR, *Opuz v. Turkey*, Rec. no. 33401/02, judgment of 9 June 2009, ECLI:CE:ECHR:2009:0609JUD003340102).

patients in need of complex care systematically assigned lower risk scores to African-American patients than to white patients with identical medical conditions[25]. The system, in fact, used past health expenditure as a proxy variable, believing it to be indicative of clinical need. However, this parameter proved to be severely distorting, as it reflected systemic inequalities in access to care by African American communities, historically excluded from adequate and continuous health care pathways. This distortion was not an accidental technical error, but the predictable outcome of a decision-making architecture that, by selecting an economic metric as the criterion of relevance, reproduced – rather than corrected – the racial asymmetries already present in the healthcare system.

Other recent studies have highlighted how deep learning models applied to radiological diagnostics tend to perform significantly poorer for ethnic minority patients[26]. In particular, several automatic classification algorithms used on radiographic images were analysed and found that their accuracy was systematically higher for white patients than for black, Hispanic or Asian patients. This discrepancy is not accidental: it depends on unbalanced training datasets, in which images of white patients are over-represented, and on metrics optimised on majority populations. This leads to a systemic underdiagnosis affecting specific groups, thereby compromising equitable access to care and potentially infringing on the right to health[27].

The use of proxies that are not validated from a social equity perspective shows how the selection of algorithmic indicators is not a neutral operation, but an axiological and normative choice that directly affects the structure of rights[28].

In the context of the criminal justice process, the introduction of artificial intelligence systems to support judicial decisions – e.g. in assessing the risk of recidivism, the allocation of pre-trial detention or the determination of alternative measures – has raised growing questions about the respect of constitutional and conventional guarantees, such as the principle of due process, the right of defence and the prohibition of discrimination. A case in point is the use of predictive algorithms such as COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*), used by several US courts to estimate the probability of a defendant committing a new crime. Independent studies have revealed a systematic tendency for the algorithm to classify as 'high risk' African-American subjects who, in fact, would not have committed crimes after the assessment. In contrast, white defendants with a similar profile were more frequently classified as 'low risk' even in the presence of significant prior history or comparable past behaviour[29]. This predictive *bias* was not due to technical error, but reflected a decision-making structure based on historical data contaminated by systemic biases – such as selective policing policies and disparities in access to alternative measures –

---

[25] Z. OBERMEYER, B. POWELL, C. AUFMAN, S. VASAN, *Dissecting racial bias in an algorithm used to manage the health of populations*, in *New England Journal of Medicine*, 2019, vol. 381, p. 2471 ff.

[26] L. SEYYED-KALANTARI, H. ZHANG, M.B.A. MCDERMOTT, I.Y. CHEN, M. GHASSEMI, *Underdiagnosis bias of artificial intelligence algorithms applied to chest radiographs in under-served patient populations*, in *Nature Medicine*, 27, 2021, p. 2176 ff.

[27] *Ibid.*

[28] Health data are never neutral or purely numerical elements: they always implicitly reflect the social structures and value hierarchies that define the meaning and treatment of human lives. Every piece of clinical data collected, every indicator used in healthcare, in fact incorporates political and cultural choices that influence what is measured, for whom it is measured and for what purpose. In this sense, the very construction of the data is already a normative act, reflecting institutional priorities and prior power relations. V. EUBANKS, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*, New York, 2018.

[29] J. ANGO WINNER, J. LARSON, S. KIRCHNER, *The Machine Bias*, in *ProPublica (*online journal), 2016.

which the algorithm replicated, codified and reiterated in its output[30]. Such biased predictions, although formally neutral, resulted in asymmetrical judicial treatment, influencing the granting of probation, bail or alternative sentences. The proprietary nature of the software, protected by industrial secrecy, has long prevented defendants from accessing the algorithm's operating criteria, making effective judicial review of the underlying decision-making mechanisms impossible[31].

The problem, however, is not limited to the quality of the data[32] or statistical errors. In a broader perspective, it is necessary to consider the active role of artificial intelligence in the construction of social reality: algorithmic technologies do not merely describe the world, but help to shape it on the basis of narrative, cultural and power structures embedded in data and models[33]. The algorithm operates as a performative device, capable of selecting what is to be visible, categorised or measured, thus contributing to the definition of standards of relevance, deservingness and legal recognition. This constitutive character has been highlighted by authoritative studies that have highlighted how algorithmic systems, far from being neutral instruments, interact with social structures and contribute to their reproduction[34]. The technical infrastructure that supports AI incorporates worldviews, values and political aims that are often neither transparent nor accessible to those involved in automated decision-making processes. In doing so, the algorithm can reinforce pre-existing social hierarchies, making them part of a codified symbolic order that is difficult to challenge.

This performative capacity of the algorithm highlights a further criticality: the emergence of forms of automatic classification that tend to crystallise individuals into opaque and unchangeable groups, with no possibility of reconfiguration or exit. This is the phenomenon described by Sandra Wachter as *algorithmic immutability*, according to which the sets constructed by AI systems – although not based on intrinsic characteristics such as race or gender – end up behaving, in decision-making practice, as if they were fixed and unchallengeable attributes[35].

This algorithmic crystallisation produces an effect similar to that of discrimination based on protected categories: it impedes social mobility, reduces the possibility of rectification or explanation and restricts individual autonomy, effectively excluding individuals from fundamental opportunities such as work, credit, education or health care.

---

[30] C. THOMAS, A. NUNEZ, *Automating Judicial Discretion: How Algorithmic Risk Assessments in Pretrial Adjudications Violate Equal Protection Rights on the Basis of Race*, in *Law & Inequality*, 2022, p. 371 ff.

[31] *Ibid.*

[32] Data quality is a central aspect in the construction of algorithmic systems, as it directly influences the reliability and fairness of automated decisions. A quality dataset is not simply an ordered set of information, but a product constructed through methodological, epistemological and often also political choices. The main requirements that guarantee its suitability include completeness (i.e. the ability to adequately represent all relevant subjects with respect to the application domain), accuracy (the correspondence of the data to empirical reality), timeliness (the use of up-to-date information), representativeness (the balanced presence of the different social categories), and the absence of systemic biases (preventing stereotypes or historical prejudices from being encoded in the data). P. CHEN, L. WU, L. WANG, *AI Fairness in Data Management and Analytics: A Review on Challenges, Methodologies and Applications*, in *Applied Sciences*, 2023, 13(18), p. 10258 ff.

[33] K. CRAWFORD, *Atlas of AI. Power, Politics, and the Planetary Costs of Artificial Intelligence,* New Haven, 2021.

[34] N. SEAVER, *Algorithms as culture: Some tactics for the ethnography of algorithmic systems*, in *Big Data & Society* (online journal), 2017, p. 1 ss; C. O'NEIL, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, New York, 2016; J. NIDA-RÜMELIN, N. WEIDENFELD, *Digitaler Humanismus. Eine Ethik für das Zeitalter der Künstlichen Intelligenz*, München, 2018.

[35] S. WACHTER, *The Theory of Artificial Immutability: Protecting Algorithmic Groups under Anti-Discrimination Law*, in *SSRN Electronic Journal* (online journal), 2022.

Classification is thus transformed into a technical predestination, which operates silently but systemically, consolidating new forms of inequality not recognised by the current anti-discrimination legal architecture.

3. *Towards an inclusive algorithmic normativity: law between substantive justice and computational design*

The growing affirmation of algorithmic technologies in decision-making processes poses numerous questions for legal theory: is law still able to exercise an effective regulatory function in an ecosystem in which decisions are increasingly entrusted to opaque, dynamic and self-learning models?

Computational architectures, in fact, do not merely perform support functions, but are configured as regulatory devices in their own right, capable of affecting social relations, redistributing power and determining access, exclusions and recognitions. In this perspective, law is called upon to overcome a reactive and formalistic conception of legality in order to assume a design posture, capable of intervening *ex ante* in the definition of algorithmic logics, according to a paradigm of legal protection integrated in the design phase[36].

This approach implies an epistemological reversal: law can no longer be read as a separate system with respect to technology, but must become an integral part of its infrastructure. The aim is to prevent the executive force of the code from taking the place of the normative force of the law, making automated decisions transparent, contestable and respectful of fundamental rights[37].

A legally pregnant analysis of algorithmic performativity requires abandoning a technicist and static vision of the code, to embrace a relational and dynamic paradigm, in which algorithms are understood as "*socio-technical assemblages"* whose normativity stratifies over time. It is in this perspective that the concept of the "*biography of algorithms"*, elaborated within the framework of organisation theory to describe the evolutionary trajectory of algorithms as socio-technical artefacts, fits in. According to this approach, every algorithmic system is structured through a progressive stratification of "*programmes of action"* – worldviews, operational metaphors and performative expectations – that guide its functions and condition its effects, often on the basis of ideological assumptions or implicit structural interests[38] . This, from a critical perspective, raises essential questions about rule of law in digital environments. If, in fact, the algorithm is the bearer of embedded logics and values, and if these logics influence decisions that affect fundamental rights, then the very architecture of the code must be subject to legal scrutiny. The apparent neutrality of the machine – the result of an iterative process of ideological and technical "*layering*" – risks concealing the concentration of power in the hands of technical or commercial actors who

[36] M. HILDEBRANDT, *Smart Technologies and the End(s) of Law,* cit., p. 34.

[37] M. HILDEBRANDT, *Legal Protection by Design: Objections and Refutations*, in *Legisprudence*, vol. 5, no. 2, 2011.

[38] During its lifecycle, an algorithm does not merely perform predetermined functions, but incorporates implicit normative representations: models of the intended user, visions of organisational efficiency, selection logics of relevant information. These - often invisible - inscriptions guide the translation of abstract goals (e.g. security, productivity, merit) into technical computational mechanisms that, in turn, rewrite the criteria of access, evaluation and recognition. It is in this sense that algorithmic design is configured as an eminently legal moment, in which the spaces of individual freedom, procedural transparency and substantive equality are 'programmed'. S. GLASER, A. BIJL, P. STORM, *The Biographies of Algorithms: Understanding the Performative Dynamics of Algorithmic Systems in Organisations,* in *Organisation Theory*, vol. 2, no. 3, 2021, pp. 1-21.

are not democratically accountable. Hence the urgency of a legal gaze that is both epistemological and structural, capable of interrogating the design process as the *locus* of contemporary normativity.

The effective inclusion of fundamental legal principles in computational mechanisms is an unavoidable condition to ensure that artificial intelligence develops and operates in compliance with the canons of constitutional legality and fundamental human rights. In a historical phase in which normativity tends to migrate from legislative text to source code, law can no longer limit itself to an *ex post* control function, but must claim a constitutive role in the definition of digital architectures. Algorithms do not merely support decision-making; they prefigure, structure and, in many cases, determine decisions in a prescriptive manner. This occurs in strategic domains such as personnel selection, creditworthiness assessment or the distribution of information on digital platforms, where code becomes a genuine normative infrastructure[39]. In automated CV-screening processes, for example, the algorithm determines which candidates progress to the next stage: the system does not simply assist HR, but effectively prescribes the ordering of opportunities. Empirical studies confirm that candidates filtered by AI systems are subsequently hired at significantly higher rates than those selected through traditional methods, demonstrating the performative power of algorithmic mechanisms over occupational mobility[40]. Similarly, in credit scoring models, the automated attribution of a score determines access to, or exclusion from, essential services such as mortgages, loans or credit lines. Such decisions appear neutral and data-driven, yet they embed and amplify pre-existing social vulnerabilities, particularly for individuals without established financial histories or from migrant backgrounds[41]. On digital platforms, moreover, algorithmic ranking exercises an even more radical normative function: it determines which content becomes visible and which remains invisible, thereby shaping the distribution of economic opportunities (monetisation), access to information and even the formation of public opinion[42]. As extensively documented in digital sociology and information science, these systems do not simply reflect user preferences, but prescribe what becomes salient, redefining hierarchies of value and visibility[43].

These examples reveal a structural truth: algorithmic normativity operates upstream, in the phase of study and design, and it is precisely at this stage that law must intervene in a constitutive manner, no longer as a mechanism of post-hoc correction. If code determines who sees what, who gains access to which opportunities, and who is deemed "*reliable*" or "*deserving*", then the algorithmic architecture becomes the new terrain on which the conditions of possibility for social action are articulated. To remain faithful to its guarantee function, law must therefore act within the algorithmic environment, shaping its evaluative premises and ensuring that conditions of access and recognition remain consistent with the principles of the rule of law.

---

[39] L. LESSIG, *Code: And Other Laws of Cyberspace*, Basic Books, New York, 1999.

[40] M. RAGHAVAN, S. BAROCAS, J. KLEINBERG, K. LEVY, *Mitigating bias in algorithmic hiring: evaluating claims and practices,* in «Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT '20)», Association for Computing Machinery, New York, 2020, pp. 469-481.

[41] M. HURLEY, J. ADEBAYO, *Credit Scoring in the Era of Big Data*, in *Yale Journal of Law and Technology*, vol. 18, 2017, pp. 5 ss.

[42] T. GILLESPIE, *Custodians of the Internet. Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*, New Haven–London, Yale University Press, 2018.

[43] On this point see: R.K. NIELSEN, S.A. GANTER, *The Power of Platforms. Shaping Media and Society*, New York, Oxford University Press, 2022; T. GILLESPIE, *Custodians of the Internet. Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*, New Haven–London, Yale University Press, 2018, ISBN 9780300173130.

In this perspective, the shift from a merely prescriptive conception of law to a co-design paradigm is not a rhetorical formula, but a structural requirement clearly articulated in the Ethics Guidelines for Trustworthy AI of the High-Level Expert Group on AI. The Guidelines identify three inseparable dimensions of trustworthy artificial intelligence: lawfulness, ethical soundness and technical robustness[44]. These requirements do not operate as independent layers, but as components that must be integrated into system design from the earliest stages. The configuration environment of decision-making structures can therefore no longer be conceived as an exclusively technical space, but as one in which legal knowledge contributes to defining limits, orienting criteria and structural conditions of operation.

Within this framework, the contribution of the jurist does not consist in imposing an already-formed regulatory framework from the outside, but in translating fundamental legal categories, such as proportionality, non-discrimination, dignity and transparency, into constraints and parameters that shape design choices. This occurs, for example, in the selection of relevant variables, the definition of optimisation metrics, the calibration of algorithmic sensitivity thresholds or the establishment of effective mechanisms for contestability. In this sense, legal intervention takes the form of a conceptual mediation that brings into the computational environment the principles that sustain constitutional legality.

System design must be oriented towards ensuring contestability, transparency and meaningful avenues for review. In the European Guidelines, these elements are not ethical accessories but essential conditions for the legitimacy of any automated system. This entails building technologies that allow users to understand decision-making logics, verify the proportionality of interventions and influence the ways in which their data are processed and their positions determined. In this way, legal protection does not operate solely a posteriori, correcting errors, but acts upon the very conditions of decision-making, shaping the technological environment so that it reflects the fundamental principles of the rule of law.

A particularly illuminating reflection in this regard comes from the European research project *COHUBICOL – Counting as a Human Being in the Era of Computational Law*, directed by Mireille Hildebrandt and funded by the European Research Council. The investigation conducted within COHUBICOL has shown how the transposition of law into computational form – in both *data-driven* and *code-driven* models – profoundly transforms the very nature of legal normativity. Law, from being a system of rules open to interpretation, dissent and contestation (and thus conforming to the Rule of Law), runs the risk of becoming a closed, opaque and self-executing system, where the rule is inscribed in automatic procedures that admit neither exceptions nor interpretative pluralism[45].

---

[44] The High-Level Expert Group on AI of the European Commission articulated the notion of "*trustworthy AI*" around three core requirements: lawfulness, understood as conformity of the system with the legal order; ethics, conceived as respect for fundamental European values such as dignity, autonomy, justice and non-discrimination; and technical and societal robustness, which includes resilience, safety, risk management and human oversight. The Guidelines further emphasise the *by design* character of these requirements: they are not to be verified *ex post*, but must be embedded in the design choices concerning data, models, metrics and mechanisms of contestability. Particular importance is also attributed to accountability, traceability and explainability of outputs, regarded as necessary instruments for ensuring that legal principles operate as structural constraints rather than as mere ethical aspirations.
HIGH-LEVEL EXPERT GROUP ON AI, *Ethics Guidelines for Trustworthy AI*, European Commission, 2019.

[45] The concept of *law as computation*, which is at the heart of the project, takes two main forms: on the one hand, *data-driven law*, based on the use of machine learning techniques to predict or suggest the application of rules;

In this context, the principle of contestability – i.e. the right of every individual to understand, discuss and challenge decisions that affect him or her – appears threatened by the opaque automation of decision-making processes. As pointed out by part of the doctrine, the computational codification of law tends to flatten the deliberative and dialectical dimension of law itself, replacing *practical reason* with *technical rationality* and reducing legal subjectivity to an algorithmic function[46]. The effect is structurally political: the legitimacy of the decision no longer lies in its public justifiability, but in its procedural efficiency.

In a decision-making ecosystem increasingly permeated by computational opacity, formal equality proves to be a legally inadequate instrument. Article 3(2) of the Italian Constitution in fact imposes a positive obligation on the state: it is not enough to avoid discriminatory treatment, but action must be taken to actively remove economic and social obstacles that effectively limit the freedom and participation of citizens. Applied to the context of algorithmic technologies, this principle implies the need to critically investigate the operating logics of automated systems, making them transparent, accessible and subject to democratic control.

Algorithmic rationality, precisely because it acts *ex ante* by selecting, categorising and attributing relevance, must be evaluated not only in terms of technical efficiency, but especially with respect to its ability to promote or hinder inclusion[47]. A constitution that recognises equality as a dynamic and substantive principle also requires technological artefacts to be judged according to their impact on distributive justice and effective access to rights.

European Union law already offers a first regulatory basis for containing the risks associated with automated decision-making. Article 22 of Regulation (EU) 2016/679 (*General Data Protection Regulation – GDPR*)[48] grants the data subject the right not to be subjected to decisions based solely on automated processing, when such decisions produce legal effects or significantly affect the individual. If interpreted substantively, this provision does not merely prohibit automated decision-making, but also imposes an obligation to ensure

---

on the other hand, *code-driven law*, which takes the form of the automation of legal rules in the form of executable code - think, for instance, of *smart contracts* on *blockchain* platforms. Both approaches, if not properly governed, risk turning law into a closed, deterministic and opaque device, unable to guarantee transparency, contestability and interpretability.

Hildebrandt points out the dangers associated with such a transformation, emphasising how the computation of law can compress the individual's ability to challenge decisions that affect him, reduce the interpretative flexibility that characterises legal argumentation and, above all, shift the centre of gravity of normative power from public and democratically legitimised institutions towards private actors, such as technology companies and code programmers. Faced with this scenario, the COHUBICOL project proposes the development of a *new computational legal hermeneutics*, capable of integrating tools of classical legal theory with technical computer skills. The aim is to preserve the fundamental principles of legality - publicity, predictability, accountability - even within automated digital ecosystems. In this sense, law must not passively adapt to technology, but must know how to 'write' it according to its own principles: this is the profound sense of the notion of *legal protection by design*. EUROPEAN COMMISSION, *Cordis - Cohubicol project fact sheet counting as a human being in the era of computational law*, available at: https://cordis.europa.eu/project/id/788734.

[46] P. HACKER, *Explainable AI under Contract and Tort Law: Legal Incentives and Technical Challenges*, in *Artificial Intelligence and Law*, 29, 2021.

[47] L. VIOLINI, *Costituzione, tecnologie e inclusione: verso una cittadinanza digitale sostanziale,* in *Costituzione e Inclusione Digitale,* Giappichelli, 2022.

[48] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of individuals with regard to the processing of personal data and on the free movement of such data and repealing Directive 95/46/EC (General Data Protection Regulation), published in OJEU. L 119 of 4 May 2016

genuine, effective and informed human control, which is not merely a formal ratification of the algorithmic output[49].

However, the scope of Article 22 is both circumscribed and insufficient. The rule protects the individual only in the presence of direct and relevant consequences, potentially leaving a vast grey area uncovered: that in which algorithms do not produce immediate legal effects, but nevertheless contribute, right from the design and profiling stage, to generating structural asymmetries and systemic discrimination. In this perspective, the provision of the GDPR proves incapable of addressing the forms of exclusion that are rooted in the very design of automated decision-making systems, making it necessary to integrate it with substantive principles of algorithmic justice and inclusion.

It is precisely in this direction that a deeper reflection on the role of law is required: the construction of a genuinely inclusive algorithmic normativity cannot stop at the mere enunciation of principles of non-discrimination nor be satisfied by an ethical supervision entrusted to soft law instruments or to the deontological responsibility of developers[50]. Such an approach – although important on a cultural level – has proved insufficient to guarantee effective protection of fundamental rights in digital contexts, where inequalities no longer manifest themselves only in the blatant violation of norms, but creep into the technical assumptions that structure automated decisions. The challenge facing contemporary legal culture is therefore more radical: it is necessary to go beyond the traditional horizon of law as a discipline of behaviour, in order to recognise its role as an enabling – or limiting – architecture of the conditions of possibility of social action.

It is in this perspective that regulatory intervention must address not only the explicit contents of algorithms, but their very design, training datasets, optimisation metrics, and verification and validation methods. The Italian case of the automated system for the management of the Citizenship Income[51] – which generated discriminatory exclusions on

[49] This reading has also been transposed by national case law: consider, for example, ruling No. 8472/2019 of the Council of State, which deemed illegitimate the use of an algorithm for the automatic allocation of school seats, in the absence of a real margin of discretionary assessment by the administration49 . The administrative judge clarified that entrusting the decision entirely to the algorithm empties the procedure of its human dimension, violating the principle of good conduct and impartiality (art. 97 of the Constitution), as well as the participatory rights of the recipient of the measure. Sent. Cons. Stato, Sec. VI, 13 December 2019, no. 8472.

[50] European governance of artificial intelligence was initially based on soft law instruments and non-binding ethical principles without direct legal effect. The Ethics Guidelines for Trustworthy AI published in 2019 by the High-Level Expert Group on AI promote values such as transparency, justice, accountability and non-discrimination, but do not provide for coercive obligations or enforcement mechanisms.
In parallel, even voluntary international instruments - such as the Toronto Declaration (2018) or the Hiroshima Code of Conduct (2023, promoted by the G7) - while offering important value references, do not by themselves guarantee effective protection of fundamental rights, especially with respect to the systemic biases embedded in algorithmic systems.

[51] The management system for the Citizenship Income, used in Italy as of 2019, was also based on automated mechanisms for assessing the 'adequacy' of applications, including algorithms for profiling beneficiaries and *ex ante* controls. Some critical studies have denounced the use of opaque criteria not accessible to citizens and not verifiable in court, which would have excluded individuals in a socially vulnerable condition without the possibility of cross-examination. The case raised important questions in terms of transparency, accountability and human control over automated decisions.
M. DE ANGELIS, *Vulnerabilità e misure di sostegno al reddito in Italia: il caso del reddito di cittadinanza,* in *Economia e società regionale*, no. 1/2023, pp. 116-136.

the basis of opaque and unchallengeable criteria – and the AMS algorithm[52] in Austria, which introduced penalising forms of profiling for categories already vulnerable in the labour market, represent two paradigmatic examples of the systemic risk of exclusion that is generated not by discriminatory intentionality, but by the naturalisation of pre-existing structures of inequality. In such cases, the algorithm functions as a device for consolidating the existing social order, through a computational rationality that interprets reality through historically stratified statistical regularities, without questioning its ethical or constitutional implications.

Far from being attributable to a programming error, the discriminatory risk then derives from an epistemic and axiological deficit that invests the entire life cycle of the system: from the engineering of the model to the absence of public control. For this reason, inclusion cannot be considered a final attribute to be verified a posteriori, but must become a structural parameter of the algorithmic design, the object of strict legal regulation and democratic accountability.

The AI Act represents an important attempt at a response, introducing a binding framework based on risk assessment and shared responsibility throughout the systems' life cycle. The provisions on high-risk systems, in particular Articles 29 and 64, provide for tools such as fundamental rights impact assessments and independent audits, which aim to ensure a form of *prior* and participatory *accountability*. However, the binary high/high-risk categorisation is insufficient to intercept the structural and ubiquitous nature of computational inequalities.

Particular attention also deserves the role of sectoral codes of conduct provided for in Articles 29 and 95, which – although being soft law instruments – can be an effective lever for algorithmic inclusion, provided they are co-designed with the active involvement of impacted communities. In this sense, algorithmic fairness is measured not only in the statistical correctness of results (*fairness by metrics*), but in the democratic quality of decision-making processes and in the possibility for citizens to understand, challenge and influence the logics that govern their daily lives.

In this framework, the idea of justice cannot be reduced to the procedural fairness of the algorithm: it must also include the possibility for citizens to understand, contest and influence the decisions that affect them. As Aristotle recalled in the *Nicomachean Ethics*, equity (*epieikeia*) is what enables law to adapt to the variety of concrete situations, where the general rule proves insufficient[53]. Translated into the context of artificial intelligence, this means constructing a *computational legality* that does not abdicate the technicalities of algorithmic rationality, but critically traverses them, orienting them towards constitutionally compatible

[52] Austria's AMS (Arbeitsmarktservice) system, introduced in 2018, employed a predictive algorithm to classify the unemployed into three 'employability' categories, based on variables such as age, gender, nationality, level of education and work history. This classification resulted in differential access to public employment services, particularly penalising women, migrants and people with disabilities. The project drew strong criticism from civil society, the academic community and the Austrian Data Protection Authority, which considered the system to be non-compliant with the GDPR due to violation of the principle of fairness and lack of transparency. In 2020, its implementation was suspended. D. ALLHUTTER, F. CECH, F. FISCHER, G. GRILL, A. MAGER, *Algorithmic Profiling of Job Seekers in Austria: How Austerity Politics Are Made Effective*, in *Frontiers in Big Data*, vol. 3, 2020.

[53] Aristotle defines equity as the corrective to universal law when the latter, due to its generality, is inadequate with respect to the complexity of the concrete case. Equity, as superior and flexible justice, assumes a central role in complex regulatory contexts, such as today's dominated by automated decision-making systems. ARISTOTELE, *Etica Nicomachea*, transl., intr. and notes by C. NATALI, Laterza, Roma-Bari, 1999.

ends. It is in this perspective that the jurist is called upon not only to regulate AI, but to co-design it, translating the principles of substantive equality into normative design criteria capable of preventing injustice before it manifests itself.

4. *Conclusions*

The growing centrality of artificial intelligence in public and private decision-making processes calls for a profound reconfiguration of the role of law. In a context in which normativity tends to migrate from the text to the computational structure, it is no longer sufficient to preside *ex post* over the effects of automated decisions: it is necessary to preside *ex ante* over the architectures that generate them. In other words, the jurist can no longer limit himself to controlling the legality of algorithmic decisions: he must participate in their semantic and axiological design.

This change requires a new legal paradigm: law no longer as a reactive system, but as an embedded normative infrastructure, capable of directly affecting the design of AI systems. This is achieved concretely through at least three strategic directions:

1. Constitutionalisation of algorithmic design: fundamental principles – equality, transparency, due process – must become computable requirements, i.e. technical specifications embedded in datasets, optimisation metrics and variable selection criteria. Inclusion can no longer be a criterion assessed downstream, but must become a structural design constraint.
2. Structural and cyclical accountability: there is a need for a legal model of distributed accountability, which accompanies the system throughout its life cycle, imposing audit, documentation and explainability obligations not only at the time of placing on the market, but also at the time of updating, reconfiguration or migration of systems. Law must act as an active memory of the algorithmic cycle, not as a footnote to technology.
3. Regulatory co-production between law and technology: the jurist must participate in development processes, adopting collaborative methodologies (co-design) with computer scientists, engineers, sociologists and impacted communities. In this perspective, tools such as *ex ante* fundamental rights impact assessments, in-house ethical-legal committees and user-accessible dispute platforms become new-generation legal devices, capable of anticipating conflict, not just mitigating its effects.

It is therefore not a question of "*humanising*" the algorithm with declarations of principle, but of legalising its operational logic. The law is called upon to exercise a new form of vigilance: no longer just on the legitimacy of the act, but on the grammar of the code. From this perspective, legality is no longer a constraint external to the technical system, but a design criterion. Algorithmic fairness, therefore, is not the fortuitous outcome of a good technical operation, but the result of a legal-political choice, embedded in the very form of the automatic decision.

Only in this way can law continue to exercise its historical function: not that of chasing the transformations of technology, but of orienting them towards the horizon of constitutional justice.

At its core, the task of law in the algorithmic age is not merely to police the boundaries of technology, but to preserve the space of subjectivity and democratic deliberation. Justice is not an emergent property of computation, but a political project that requires constant vigilance, even when decisions are filtered or reorganised by computational systems. Otherwise, there is a risk that logics of automatic efficiency will replace forms of public evaluation grounded in contestation, justification and responsibility. Law thus remains the forum in which these transformations must be interpreted, debated and rendered compatible with human dignity and with the substantive equality that democratic constitutions commit to guaranteeing.