Francesca Morganti

AI AND THE CONSTITUTIONAL RIGHT TO HEALTH: EQUALITY AND SUSTAINABILITY IN THE ITALIAN SYSTEM

1. *Introduction. The Constitutional Implications of AI in Healthcare*

The use of artificial intelligence (AI) in the medical and healthcare sectors is already having – and will inevitably continue to have – a significant impact on individuals and their rights. The following analysis approaches these developments from the perspective of Italian constitutional law, relying on the Italian legal system as the main case study through which the broader constitutional implications of AI-driven healthcare can be assessed. In light of this focus, the discussion draws primarily on Italian constitutional scholarship, which offers the doctrinal framework most directly suited to examining the issues addressed here, while also providing a lens through which questions of wider comparative relevance may be considered[1].

[1] On the applications of artificial intelligence in the medical and healthcare sectors, see, among the most recent contributions – from a legal-constitutional perspective and with particular attention to the Italian and European contexts – M. FASAN, *Regulating the Use of Artificial Intelligence in the Doctor-Patient Relationship? A Primer on Supranational and National Legal Frameworks*, in *BioLaw Journal – Rivista di BioDiritto*, No. 1, 2025, pp. 193 ff.; EAD., *Intelligenza artificiale e costituzionalismo contemporaneo. Principi, diritti e modelli in prospettiva comparata*, Naples, 2024, esp. pp. 169 ff.; C. DE MENECH, *Intelligenza artificiale e autodeterminazione in materia sanitaria*, in *BioLaw Journal – Rivista di BioDiritto*, No. 1, 2022, pp. 181 ff.; M. OROFINO, *La questione del sotto-utilizzo dell'intelligenza artificiale in campo sanitario: spunti di rilievo costituzionale*, in *Queste Istituzioni*, No. 4, 2022, pp. 158 ff.; L. SCAFFARDI, *La medici-*

At the outset, it is worth emphasizing that the partial digitalization of healthcare – limited to those segments in which data-driven systems can meaningfully support clinical or organisational processes – may positively affect the right to health, particularly in its dimension as a social right – that is, as an individual entitlement to receive medical care and health-related services[2]. The central research question guid-

na alla prova dell'Intelligenza Artificiale, in *DPCE online*, No. 1, 2022, pp. 349 ff.; C. CASONATO, S. PENASA, *Intelligenza artificiale e medicina del domani*, in G.F. FERRARI (ed.), *Le* smart cities *al tempo della resilienza*, Milan-Udine, 2021, pp. 553 ff.; E.A. FERIOLI, *Digitalizzazione, intelligenza artificiale e robot nella tutela della salute*, in A. D'ALOIA (ed.), *Intelligenza artificiale e diritto. Come regolare un mondo nuovo*, Milan, 2020, pp. 423 ff.; as well as, if one wishes, D. MORANA, T. BALDUZZI, F. MORGANTI, *La salute "intelligente":* eHealth*, consenso informato e principio di non-discriminazione*, in *Federalismi.it*, No. 34, 2022, pp. 127 ff.

Beyond the healthcare context, the potential risks of artificial intelligence – especially those affecting individual rights – have been explored in various studies; see, among others, C. COLAPIETRO, *Digitalizzazione e Costituzione*, in *Italian Papers on Federalism*, No. 1, 2025, pp. 110 ff.; O. POLLICINO, *Regolare l'intelligenza artificiale: la lunga via dei diritti fondamentali*, in G. FINOCCHIARO *et al.* (eds.), *La disciplina dell'intelligenza artificiale*, Milan, 2025, pp. 3 ff.; V. DE SANTIS, *Intelligenza artificiale: identità personale e diritti*, in S. PISANO (ed.), *Intelligenza artificiale. Azienda, lavoro e diritti*, Bari, 2024, pp. 45 ff.; A. SIMONCINI, *Il linguaggio dell'Intelligenza Artificiale e la tutela costituzionale dei diritti*, in *Rivista AIC*, No. 2, 2023, pp. 1 ff.; A. ODDENINO, *Intelligenza artificiale e tutela dei diritti fondamentali: alcune notazioni critiche sulla recente Proposta di Regolamento della UE, con particolare riferimento all'approccio basato sul rischio e al pericolo di discriminazione algoritmica*, in A. PAJNO, F. DONATI, A. PERRUCCI (eds.), *Intelligenza artificiale e diritto: una rivoluzione?*, I, *Diritti fondamentali, dati personali e regolazione*, Bologna, 2022, pp. 165 ff.; C. NARDOCCI, *Intelligenza artificiale e discriminazioni*, in *Rivista del Gruppo di Pisa*, No. 3, 2021, pp. 9 ff.; P. ZUDDAS, *Intelligenza artificiale e discriminazioni*, in *Liber amicorum per Pasquale Costanzo. Diritto costituzionale in trasformazione*, I, *Costituzionalismo, Reti e Intelligenza artificiale*, in *Consulta OnLine*, 2020, pp. 457 ff.; C. CASONATO, *Intelligenza artificiale e diritto costituzionale: prime considerazioni*, in *Diritto pubblico comparato ed europeo*, Special Issue, 2019, pp. 101 ff.

[2] The Italian Constitution contains a specific provision on the protection of health: Article 32, which states that «[t]he Republic shall safeguard health as a fundamental right of the individual and as a social interest and shall guarantee free medical care to the indigent» (para. 1); the following paragraph adds that «[n]o one shall be forced to undergo medical treatment unless provided for by law», and that «[i]n no case shall the law violate the limits imposed by respect for the natural person» (para. 2). As D. MORANA, *La salute come diritto costituzionale*, 5th ed., Turin, 2025, p. 2, has observed,

ing this study is whether, and under what conditions, the integration of AI technologies into healthcare can enhance the enjoyment of this right without simultaneously eroding constitutional guarantees relating to equality, autonomy, and the protection of both present and future communities. Put differently, the inquiry seeks to determine whether the benefits associated with AI in healthcare – benefits that are often emphasized in policy discourse – can be reconciled with the risks, both individual and structural, that such technologies may generate.

The applications of artificial intelligence in the medical field are numerous and heterogeneous. At this stage, however, they are almost all based on machine-learning techniques: systems that are often not easily "readable", yet highly performant, appear capable of replacing the human operator's "slow thinking" with their own "fast perception". The high performance of such programmes comes at a cost – especially in terms of understandability – and entails certain risks for users of the (partially) automated service, but it may also generate considerable benefits.

Notably, deep learning – which is a subset of machine learning – enables systems to recognise and classify images, including medical and diagnostic ones, with remarkable accuracy (consider, for instance, the new and improved version of the so-called *BakeryScan*, which can identify neoplastic urothelial cells[3], or the convolutional neural net-

within this constitutional framework «the single expression "right to health" simultaneously and succinctly encompasses legal positions that differ in structure and content, or are at least not fully overlapping»: on the one hand, a liberty right, which «requires non-interference by others, effectively rejecting any form of intrusion or aggression into the sphere of individual health»; on the other, a right to receive «those healthcare services that, from time to time, prove necessary for the protection of one's health – a claim which, for indigent individuals, is further characterised as a right to the free provision of such services». This "complex" nature of the right to health has been consistently highlighted by constitutional scholarship; see, among others, B. PEZZINI, *Il diritto alla salute: profili costituzionali*, in *Diritto e società*, No. 1, 1983, pp. 21 ff.; and M. LUCIANI, entry *Salute: I) Diritto alla salute – Dir. cost.*, in *Enc. giur. Treccani*, XXVII, Rome, 1991, p. 5. Translations from Italian are the author's own.

[3] The reference is, indeed, to the well-known case of *BakeryScan*, a deep-learning system developed in Japan to "scan" and price various types of baked goods, which has since also been used – under the name *Cyto-AiSCAN* – to identify neoplastic urothelial cells. As detailed by J. SOMERS, *The Pastry A.I. That Learned to Fight Can-*

work that classifies pigmented dermatological lesions as "benign" or "malignant" and assists in melanoma diagnosis)[4]. The convergence of large-scale, high-quality datasets and sophisticated predictive algorithms – which are defining features of machine-learning systems – may allow us, among other things, to detect complex relationships between various and seemingly unrelated individual characteristics, linking such "constellations" of traits with, e.g., the possible onset of a given disease or a more or less favourable response to a certain drug; over time, this could make it possible to: *(i)* predict with high accuracy the occurrence or progression of certain conditions; *(ii)* personalise treatment plans for individual patients; and *(iii)* semi-automate certain "tragic choices", such as decisions about the allocation of scarce intensive-care beds or, more broadly, the rational distribution of limited resources. Predictive models – which typically rely on machine-learning algorithms – also facilitate, and will increasingly do so, the strategic planning of health and social-care services, enabling risk assessment both at the individual and at the collective level (for populations or sub-populations); thus, moving from the person to the system as a whole, these tools allow for: *(a)* the development of targeted, "tailor-made" preventive-medicine programmes; *(b)* a more reasonable allocation of available human and financial resources; and *(c)* an overall containment of healthcare expenditure.

The right to health understood as a right to receive services – one that is increasingly «financially conditioned»[5] – would likely benefit

cer, in *The New Yorker*, web version, March 18, 2021, it was a doctor at the Louis Pasteur Centre for Medical Research in Kyoto who, having come across a television segment about *BakeryScan*, began reflecting on the resemblance between cancerous cells viewed under a microscope and certain pastries. This prompted him to advocate for the adaptation and use of the classification system in the diagnostic field.

[4] On this, see further below, esp. Section 2.

[5] On the problematic framing of the right to health as "financially conditioned" – an understanding that has been accepted and transmitted almost uncritically, and that has found support, among others, in rulings of the Italian Constitutional Court – see, critically discussing some of these positions, B. PEZZINI, *Il diritto alla salute a quarant'anni dall'istituzione del servizio sanitaria nazionale: le criticità strutturali di un diritto sociale*, in *BioLaw Journal – Rivista di BioDiritto*, No. 2, 2019, pp. 126 ff. Pezzini argues that applying such a "financially conditioned" framework to the constitutional right to health is «simplistic and misleading», especially taking into account the

from such developments: the quality of services provided would improve, the organization of the system would become more rational, and territorial disparities – at least in theory – would be reduced. The aim of this contribution is not to deny these potential benefits, but to situate them within a constitutional framework capable of evaluating both the opportunities and the criticalities associated with AI-driven healthcare. This requires a methodology attentive to the coexistence of competing principles and values – effectiveness and efficiency, on the one hand; equality, autonomy, and systemic sustainability, on the other.

It would be misguided, for several reasons, to attribute quasi-magical potential or "capabilities" to machines. Nonetheless, particularly in the clinical context, what is emerging – supported by data – is that the integration of AI systems into diagnostic processes and therapeutic decision-making often yields objectively positive outcomes: when adequately trained, the machine tends to achieve performance levels comparable to, or even exceeding, those of top specialists, and can therefore serve as a valuable aid to average professionals[6].

Of course, these professionals should interact with the machine as an "interlocutor", rather than substitute its judgment for their own. Still, the «practical argumentation» strength of the machine-system often ends up prevailing[7]. It is hardly surprising that a practitioner – es-

«fundamental» nature of this right (see below, esp. n. 10). While it is true that the exercise of the (constitutionally-protected) right to health depends, in practice, on legislative implementation, and that legislators *must* consider financial constraints when exercising their discretion, financial balance – although important, most notably in light of Article 81 of the Italian Constitution – cannot be regarded as an end in itself or an unrelated objective. Rather, it is at most a «tertiary goal», justified only insofar as it supports the «secondary aim» of efficiency, which itself serves the «primary and direct purpose» of guaranteeing the protection of health. Translations from Italian are the author's own.

[6] See further below, esp. Section 2.

[7] Cf. A. SIMONCINI, S. SUWEIS, *Il cambio di paradigma nell'intelligenza artificiale e il suo impatto sul diritto costituzionale*, in *Rivista di filosofia del diritto*, No. 1, 2019, p. 100, who highlight the existence of a genuine «practical argumentative force» that algorithms exert within decision-making processes. According to the Authors, «once an automated evaluation system is introduced into a human decision-making process», that same «automated system tends, over time, to "capture" the decision […] not due

pecially in a high-liability sector like healthcare – may find reassurance in relying on third-party assessments, or even "hiding" behind them, all the more so when the "third party" appears to provide superhuman accuracy and reliability.

Moreover, the machine does not always encourage users to adopt a "critical", dialogical stance toward its automated outputs. Some complex systems are characterized by considerable opacity, making it difficult for users to understand how the system arrived at its outputs – its classifications, predictions, or decisions – based on the input data and information. If the machine's "proposal" is not justified, not argued, not *explained*, the professional will be left only with the option of accepting or rejecting it wholesale, unable to engage in any form of dialogue – however fictional or para-social – with the system. Put differently, the opaque machine inserts itself into the doctor-patient relationship, introducing elements of incomprehension and discontinuity and, in effect, rendering the traditional structure of informed consent inadequate[8]. The attribution of responsibility – or rather, the alloca-

to greater scientific value, predictive accuracy, technical reliability, or indeed evaluative neutrality, but primarily for reasons of practical convenience». Automated evaluations, in fact, relieve «the decision-maker from the burden of justification, from the effort of analysis and motivation», allowing them «to "label" their decision with the aura of "scientificity" or "neutrality" that surrounds algorithmic assessments». Translations from Italian are the author's own.

[8] The expression «informed consent» refers, in fact, to a range of «rights of different nature and structure», and in particular: *(i)* on the one hand, «the right to health as a liberty right, or more precisely, freedom of treatment understood as the *freedom to consent to (or refuse) medical care*»; *(ii)* on the other hand, «the right to health as a right to receive information about medical treatment – namely, the *right to be informed about the treatment* (and more generally about any form of intervention, even if not strictly therapeutic) – that is, to receive the information necessary to form one's consent (or refusal) within the framework of the therapeutic alliance with the physician» (thus D. MORANA, *A proposito del fondamento costituzionale per il «consenso informato» ai trattamenti sanitari: considerazioni a margine della sent. n. 438 del 2008 della Corte costituzionale*, in *Giurisprudenza costituzionale*, No. 6, 2008, p. 4976).

When physicians rely on opaque digital tools that generate diagnostic or therapeutic recommendations whose reliability cannot be verified – because the "reasoning" that led the machine from point A (input data) to point B (classification or prediction) remains unknowable – and the user, even if in disagreement, can only choose to accept or reject the system's output, it becomes materially impossible for those same

tion of responsibilities – for any harm caused to the patient also becomes problematic when the physician operates "in tandem" with the machine.

Furthermore, as will be explored below[9], AI systems – for reasons that will be at least partially examined – can be understood as "pattern-replicators", often reproducing or reinforcing pre-existing discriminatory patterns.

In short, digitalization – particularly through the large-scale deployment of AI systems – carries great promise but also substantial risks, especially when applied to services, such as healthcare, that are instrumental to the enjoyment of constitutionally-protected rights (specifically, the «fundamental» right to health)[10]. Understanding the-

physicians to provide complete and specific information to their patients. The use of such tools introduces a "break" in the therapeutic alliance: the physician is prevented from fully understanding the process behind the proposed course of action, while the patient is (consequently) deprived of the elements needed to make an informed choice. If, in other words, the right to receive information constitutes the logical and legal precondition of the consensual nature of medical treatment, which in turn is a «necessary corollary» of the voluntary nature of the intervention itself (ivi, p. 4973), then the use of artificial intelligence in clinical settings – where it typically involves the deployment of (objectively or intentionally) opaque systems, as is often the case with machine learning – inevitably affects the practical conditions for exercising the freedom to make informed health choices.

On this topic, see also C. DE MENECH, *Intelligenza artificiale e autodeterminazione*, cit., pp. 185 f., who observes that «the opacity» of certain artificial intelligence systems «is clearly ill-suited to the aim of offering patients a clear and reliable diagnostic-therapeutic scenario, one in relation to which they can exercise rational self-determination». Translations from Italian are the author's own.

[9] See further below, esp. Section 2.

[10] As noted above (cf. n. 2), Article 32 of the Italian Constitution states that «[t]he Republic» safeguards health «as a fundamental right of the individual» and a «social interest»; the prevailing interpretation is that this attribute («fundamental») refers to both the *individual right* and the *social* or *collective interest* (see V. CRISAFULLI, *In tema di emotrasfusioni obbligatorie*, in *Diritto e società*, No. 1, 1982, p. 564). D. MORANA, *Il «fondamentale» diritto alla salute nell'emergenza pandemica:* princeps *o tiranno?*, in *Quaderni costituzionali*, No. 4, 2022, pp. 858 ff., identifies the «practical significance» of such "fundamentality" – explicitly recognised with regard to the right to health, and the right to health only – in its function as a «criterion for resolving apparent antinomies within the constitutional system of guarantees: in particular, in cases where a conflict arises between the protection of health and that of other rights of

se risks requires operating on two analytical levels: that of the individual, whose rights and autonomy may be jeopardised by AI-driven processes, and that of the community – both present and future – whose interests are increasingly recognised within the Italian constitutional order[11].

Any legal-constitutional analysis of digitalization and its impact must therefore adopt a dual perspective, continuously moving between an assessment of risks (both actual and potential) and of benefits. The aim is to delineate a constitutional framework capable of accommodating technological innovation without compromising the principles that shape the right to health in the Italian system. The analysis that follows develops this approach, seeking to give due consideration, with regard to risks, not only to those concerning individuals but also to those that may affect present and future communities, even where the latter rest on less settled theoretical grounds.

2. *AI in Clinical Care: Promises, Risks, and Unresolved Questions*

As previously noted, artificial intelligence systems – when adequately trained – can achieve performance levels comparable to, or even exceeding, those of top human specialists. This assertion, however, requires contextualization and clarification.

In the case of highly specialised AI systems, one key limitation concerns the narrow scope of their "knowledge": they excel in depth but lack breadth. Image recognition tools, for instance, may display superhuman abilities – they can "ingest" vast amounts of data and compare each frame to millions of others – yet their "knowledge" remains limited and compartmentalized: «[n]arrow AI for detecting polyps, for example, might "see" a problem polyp that no gastroenterologist

equal constitutional standing, and where such conflict cannot be resolved through the criterion of speciality». See also, on this point, P.F. GROSSI, *Diritti fondamentali e diritti inviolabili nella Costituzione italiana*, in ID., *Il diritto costituzionale tra principi di libertà e istituzioni*, 2nd ed., Padua, 2008, pp. 1 ff., esp. p. 2. Translations from Italian are the author's own.

[11] See further below, esp. n. 45.

would, but it might also be incapable of recognizing other abnormalities that it was not trained to detect»[12].

When it comes to generalist AI systems used in clinical settings, the challenges differ – and can, in some ways, be even harder to pinpoint and address. The issue here is not so much the breadth of the system's "knowledge" as its precision, contextual appropriateness, and the usability of its outputs. Moreover, the performance levels of general-purpose models vary significantly depending on the nature of the task assigned by human users.

A very recent study tested, among other things, the diagnostic potential of several large language models (LLMs), and more precisely their ability to generate well-reasoned diagnostic hypotheses based on detailed descriptions of past clinical cases[13]. The study was inspired by Clinicopathological Conferences (CPCs), which are teaching exercises built around the discussion of complex clinical cases. These conferences are usually retrospective in nature, allowing clinicians to incorporate not only clinical symptoms but also pathological findings, such as autopsy or biopsy results. Typically, an expert physician presents the case, explores diagnostic and therapeutic alternatives, and then proposes a final diagnosis (the clinical hypothesis), which is subsequently compared to the actual pathological outcome (the so-called "pathological truth"). The primary aim, as mentioned, is educational: not simply to arrive at the correct diagnosis, but to trace and explain the reasoning process that has led to it.

The study revealed, among other findings, that current large language models «outperform physicians in several text-based clinical tasks, including generating a differential diagnosis in response to a challenging clinical vignette»[14]. However, their performance «was lower on both literature search and when restricting what information from the curated case presentation is made available to the model. Substantial performance degradation was also observed in tasks re-

[12] F. PASQUALE, *Healing Humans*, in ID., *New Laws of Robotics. Defending Human Expertise in the Age of AI*, Cambridge (MA) - London, 2020, p. 37

[13] See T.A. BUCKLEY *et al.*, *Advancing Medical Artificial Intelligence Using a Century of Cases*, in www.arxiv.org, 15 September 2025, pp. 1 ff.

[14] *Ibidem*.

quiring direct image interpretation, highlighting that clinical image interpretation and multimodal integration are key remaining challenges for generalist clinical AI»[15]. This underscores not only the current limitations of applying general-purpose systems to highly specialized tasks, but also the extent to which the performance of LLMs depends on the structure and completeness of the prompts they are given (that is, the questions, instructions, context, and textual input that guide their responses). More broadly, it highlights the inextricable link between the quality of an AI system's inputs and the reliability of its outputs – a point that remains central in any constitutional assessment of such technologies.

Consider, by way of further example, the deep learning system developed in the United States for the classification of pigmented skin lesions[16] – a tool potentially useful in the diagnosis of melanoma – which nonetheless performed markedly worse when identifying and labelling lesions on darker skin tones as either "benign" or "malignant"[17]. That performance gap was easily traced back to the underrepresentation of certain minority sub-populations in the training data: the algorithm had been trained on incomplete datasets consisting primarily of (data associated with) fair-skinned individuals; as a result, it was naturally more accurate at identifying and classifying lesions on lighter skin during real-world deployment[18].

The data imbalance in favour of socially dominant groups is underpinned by historical, socio-economic, and political factors that defy simple explanation. Medical image repositories are primarily compiled in high-income, technologically advanced countries – such as the United States, Europe, and Australia – whose populations are predominantly white. Clinical trial data similarly reflect long-standing research practices that, for a range of systemic and contextual reasons[19], con-

[15] *Ibidem*.

[16] See A. ESTEVA *et al.*, *Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks*, in *Nature*, Vol. 542, 2017, pp. 115 ff.

[17] As noted by A.S. ADAMSON, A. SMITH, *Machine Learning and Health Care Disparities in Dermatology*, in *JAMA Dermatology*, No. 11, 2018, pp. 1247 f.

[18] *Ibidem*.

[19] In the United States, for instance, African Americans' well-documented lack of trust in the medical establishment – often resulting in lower levels of engagement with

tinue to prioritize the so-called «ideal male subject», typically «young» and «white»[20]. When the incompleteness of datasets stems – at least in

the healthcare system, including limited participation in clinical trials – has frequently been traced back to the «historical disclosure of an unethical and deadly experiment, the Tuskegee Study of Untreated Syphilis in the Negro Male (TSUS)», conducted between 1932 and 1972 by the U.S. Public Health Service itself: «[f]or 40 years [...] the U.S. Public Health Service (PHS) followed hundreds of poor, black men in Tuskegee, Alabama, the majority of whom had syphilis, for the stated purpose of understanding the natural course of the disease. The men were denied highly effective treatment for their condition (most egregiously, penicillin, which became standard of care by the mid-1940s) and were actively discouraged from seeking medical advice from practitioners outside the study. […] News of the Tuskegee Study became public in 1972 in an *exposé* by Jean Heller of the Associated Press, and detailed narratives of the deception and its relationship to the medical establishment were widespread. By that point, the majority of the study's victims were deceased, many from syphilis-related causes» (M. ALSAN, M. WANAMAKER, *Tuskegee and the Health of Black Men*, in *The Quarterly Journal of Economics*, No. 1, 2018, pp. 408 ff.).

On the close correlation between the «low participation of ethno-racial minorities in clinical trials» and the «inability of such individuals – already in a position of subordination – to benefit from technological innovations developed also through data derived from such research activities», see, again with reference to the U.S. context, V. LANDO, *Studi clinici, discriminazioni razziali e intelligenza artificiale: diversity and inclusion nel contesto statunitense*, in *BioLaw Journal – Rivista di BioDiritto*, Special Issue, 2024, pp. 155 ff. Translations from Italian are the author's own.

[20] Cf. Ministero Della Salute, *Il genere come determinante di salute. Lo sviluppo della medicina di genere per garantire equità e appropriatezza della cura*, in *Quaderni del Ministero della Salute*, 2016, p. 74, where it is underlined, with specific reference to clinical drug trials, how pharmaceuticals have traditionally been – and continue to be – generally tested «on an ideal male subject, typically young, white, and weighing around 70 kg». As a result, such trials often «do not guarantee women the same beneficial outcomes as those documented in men». This is hardly surprising: historically, the male body was regarded as the norm, the standard, while the female body was seen as a mere variation, or even a deviation. Clinical research has primarily involved male participants; the specificities of female physiology and lived experience – both biological and social – have long been, and frequently still are, neglected. As early as 2008, the Comitato Nazionale per la Bioetica warned that «in the field of clinical trials, women turn out to be "vulnerable subjects", or in any case not adequately taken into account with regard to their specificities, both in quantitative terms (number of women enrolled compared to men) and in qualitative terms (analysis of data by sex differences)» (Comitato Nazionale Per La Bioetica, *La sperimentazione farmacologica sulle donne*, 28 November 2008, p. 8). Similar observations can be found in Ministero Della

part – from cultural, socio-economic, or political dynamics, overlooking these dimensions can prove harmful both in terms of social justice and technical robustness. As shown in other domains – particularly predictive policing and algorithmic risk assessment for criminal behaviour or recidivism[21] – the failure to integrate explicit and context-

Salute, *Il genere come determinante di salute*, cit., *passim*.

Since then, some progress has been made, including symbolically significant steps – such as the establishment of a *Centro di Riferimento per la Medicina di Genere* at the *Istituto Superiore di Sanità*. These developments are reviewed in Ministero Della Salute, *Piano per l'applicazione e la diffusione della Medicina di Genere*, 6 May 2019, pp. 13 ff., adopted pursuant to Article 3(1) of Law No. 3 of 11 January 2018 («*Delega al Governo in materia di sperimentazione clinica di medicinali nonché disposizioni per il riordino delle professioni sanitarie e per la dirigenza sanitaria del Ministero della salute*»), which, for the first time, explicitly incorporated gender medicine into the *Servizio sanitario nazionale* (cf. M. TOMASI, *Sperimentazioni cliniche e medicina di genere. La ricerca dell'eguaglianza attraverso la valorizzazione delle differenze*, in B. PEZZINI, A. LORENZETTI (eds.), *70 anni dopo tra uguaglianza e differenza. Una riflessione sull'impatto del genere nella Costituzione e nel costituzionalismo*, Turin, 2019, p. 229). See also Regulation (EU) No. 536/2014 of the European Parliament and of the Council of 16 April 2014 «on clinical trials on medicinal products for human use and repealing Directive 2001/20/EC», in force since January 2022, which already in Recital (14) specifies that, «[u]nless otherwise justified in the protocol, the subjects participating in a clinical trial should represent the population groups, for example gender and age groups, that are likely to use the medicinal product investigated in the clinical trial»; according to point 17, letter *y*, of Annex I, the protocol included in the initial application dossier must provide «a justification for the gender and age allocation of subjects and, if a specific gender or age group is excluded from or underrepresented in the clinical trials, an explanation of the reasons and justification for these exclusion criteria».

As further noted in Ministero Della Salute, *Il genere come determinante di salute*, cit., pp. 74–75, the «precautionary approach related to clinical trials on women» is often attributed «to social, environmental, economic and above all biological reasons (hormonal variations such as menstrual cycle, pregnancy, breastfeeding, menopause or use of contraceptives)»; yet all of these "conditions" are, in fact, «constant characteristics of the female population that will use the drug once it is on the market». Translations from Italian are the author's own.

[21] This refers to a widely cited 2016 ProPublica investigation (J. ANGWIN *et al.*, *Machine Bias. There's Software Used Across the Country to Predict Future Criminals. And It's Biased Against Blacks*, in www.propublica.org, 23 May 2016), which revealed significant concerns surrounding the use of the COMPAS algorithm (Correctional Offender Management Profiling for Alternative Sanctions). This system is widely em-

sensitive socio-economic references into automated decision-making processes can lead to the reproduction, and even intensification, of the very "distortions" that algorithms are, at least in theory, meant to mitigate. In the present case, "distortions" refers to both: *(i)* the cognitive biases and stereotype-driven associations inherent in human judgment, which are encoded – often inadvertently – into training data and other

ployed in the United States to assess the risk of recidivism, including in decisions regarding parole eligibility. The report found that, although COMPAS demonstrates comparable overall levels of accuracy across different racial groups, it consistently overestimates the risk of reoffending for African American individuals and underestimates it for white individuals; as a result, the rate of false positives is approximately twice as high for the former, whereas the opposite pattern is observed in relation to false negatives.

Although the questionnaire administered to prospective parolees does not explicitly include questions regarding race, it incorporates a range of variables – including criminal history, educational attainment, employment status, economic conditions, neighbourhood of origin or residence – that are deeply embedded within broader structures of racial inequality. As aptly noted, «[a]s all these variables are structured by racial domination – from job market discrimination to ghettoization – the survey measures the extent to which an individual's life chances have been impacted by racism without ever asking an individual's race» (R. BENJAMIN, *Race After Technology. Abolitionist Tools for the New Jim Code*, Cambridge, 2019, p. 82). In other words, many of the criteria employed to assess social dangerousness are inextricably linked to race – more precisely, to the socio-economic consequences of racial identity. The inclusion, exclusion, or reformulation of such indicators is not a neutral or technical matter, but rather the result of deliberate political choices. Moreover, the overrepresentation of African American individuals in disadvantaged socio-economic conditions and the higher crime rates often observed in predominantly Black neighbourhoods cannot be regarded as coincidental: they must instead be understood as structural outcomes of persistent and systemic forms of racial discrimination. In this light, the purported irrelevance of race within algorithmic systems such as COMPAS appears to be illusory: ostensibly race-neutral models frequently produce outcomes that are as discriminatory – if not more so – than those of explicitly race-conscious systems.

For further discussion of the "COMPAS case" in the Italian legal literature, see: C. CASONATO, *Intelligenza artificiale e diritto costituzionale: prime considerazioni*, cit., pp. 114 ff.; A. CELOTTO, *Come regolare gli algoritmi. Il difficile bilanciamento fra scienza, etica e diritto*, in *Analisi giuridica dell'economia*, No. 1, 2019, pp. 47 ff.; G. RESTA, *Governare l'innovazione tecnologica: decisioni algoritmiche, diritti digitali e principio di uguaglianza*, in *Politica del diritto*, No. 2, 2019, pp. 215 ff.; A. SIMONCINI, S. SUWEIS, *Il cambio di paradigma nell'intelligenza artificiale*, cit., esp. pp. 93 ff.

stages of system design and deployment, thereby contributing to the production of potentially harmful outputs; and *(ii)* structural conditions – such as systemic racism or entrenched gender hierarchies – that lie beyond individual agency yet remain deeply embedded within it. These "upstream" conditions give rise to "downstream" algorithmic outputs that are themselves biased.

Of course, some of these issues can be addressed through regulatory measures – for instance, by requiring those who develop and commercialize AI solutions to rely on high-quality and, crucially, sufficiently representative data. Under the European Artificial Intelligence Act[22], AI systems classified as "high-risk" must be trained, validated, and tested on data that are «relevant, sufficiently representative, and to the best extent possible, free of errors and complete in view of the intended purpose». These datasets must also possess «appropriate statistical properties», including, where applicable, with respect to «the persons or groups of persons in relation to whom the […] system is intended to be used»[23]. Importantly, it is not enough for datasets to be representative in an abstract sense: they must reflect «the characteristics or elements that are particular to the specific geographical, contextual, behavioural or functional setting within which the high-risk AI system is intended to be used»[24]. Given that AI systems employed in the medical domain – for diagnostic or therapeutic purposes – are, by their very nature, deemed "high-risk"[25], they fall squarely within the scope of these requirements.

[22] Cf. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024, «laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act)», hereinafter: AI Act. On the Act itself, see the contributions collected in G. FINOCCHIARO *et al.* (eds.), *La disciplina dell'intelligenza artificiale*, cit., and, in particular, from a constitutional perspective, F. DONATI, *Intelligenza artificiale e diritti fondamentali nel Regolamento sull'intelligenza artificiale*, ivi, pp. 41 ff.; as well as the relevant papers included in Issue No. 3, 2025, of the *BioLaw Journal – Rivista di BioDiritto*, within the *Forum* «AI Act*: un dialogo multidisciplinare*». As for the original Proposal for the Regulation – whose underlying structure is, of course, largely preserved in the final text – see, within Italian constitutional scholarship, A. ODDENINO, *Intelligenza artificiale e tutela dei diritti fondamentali: alcune notazioni critiche sulla recente Proposta di Regolamento della UE, con*

particolare riferimento all'approccio basato sul rischio e al pericolo di discriminazione algoritmica, cit.; A. SIMONCINI, *La proposta di regolazione europea dell'intelligenza artificiale. Prime riflessioni*, in A. ADINOLFI, A. SIMONCINI (eds.), *Protezione dei dati personali e nuove tecnologie. Ricerca interdisciplinare sulle tecniche di profilazione e sulle loro conseguenze giuridiche*, Naples, 2022, pp. 1 ff.; C. CASONATO, B. MARCHETTI, *Prime osservazioni sulla proposta di regolamento dell'Unione europea in materia di intelligenza artificiale*, in *BioLaw Journal – Rivista di BioDiritto*, No. 3, 2021, pp. 415 ff.

[23] Thus, Article 10(3) AI Act. See also Recital (67), which emphasises that: «High-quality data and access to high-quality data plays a vital role in providing structure and in ensuring the performance of many AI systems, especially when techniques involving the training of models are used, with a view to ensure that the high-risk AI system performs as intended and safely and it does not become a source of discrimination prohibited by Union law».

[24] See Article 10(4) AI Act, which aims to prevent so-called contextual biases, *i.e.* those arising from the misalignment between the context in which the artificial intelligence system was designed and trained (reference setting) and the context in which it is subsequently deployed (application context). As an illustrative example, one may consider a programme developed within «high-resource settings: academic medical centers or state-of-the-art hospitals» but deployed in «low-resource settings such as community hospitals, community health centers, or practitioners' offices». The risk lies in the programme exhibiting significantly lower performance and, crucially, not adequately addressing the unique demands and characteristics of the application context (see W. NICHOLSON PRICE II, *Medical AI and Contextual Bias*, in *Harvard Journal of Law & Technology*, No. 1, 2019, pp. 66 ff.).

[25] Article 6(1)(b) of the AI Act defines as "high-risk" «the product whose safety component pursuant to point (a) is the AI system, or the AI system itself as a product» which is «required to undergo a third-party conformity assessment, with a view to the placing on the market or the putting into service of that product pursuant to the Union harmonisation legislation listed in Annex I». Among the legislation listed in Annex I are Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 «on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC», and Regulation (EU) 2017/746 of the European Parliament and of the Council of the same date, «on *in vitro* diagnostic medical devices and repealing Directive 98/79/EC and Commission Decision 2010/227/EU». The Medical Devices Regulation (EU) 2017/745 categorises medical devices – including software, pursuant to Article 2(1) – into risk classes, providing that only class I devices are exempt from conformity assessments by notified bodies. Generally, class I encompasses all "non-invasive" devices (see Rule 1, Annex VIII), but Rule 11 of the same Annex classifies as class IIa or higher any «[s]oftware intended to provide information which is used to take decisions with diagnostic or therapeutic purposes». Consequently, AI systems employed within clinical healthcare settings – as opposed, for instance,

These safeguards have been further reinforced, at the European Union level, by the new Regulation on the European Health Data Space (EHDS), which governs, among other matters, the «secondary use» of electronic health data[26]. Under Articles 53 *et seq.* of the Regulation, health data may be made available – in strictly regulated conditions – for activities such as «training, testing and evaluation of algorithms», provided that robust privacy, security, and fundamental-

to healthcare planning purposes – are considered "high-risk" by virtue of their classification under class IIa or above according to the Medical Devices Regulation and the consequent requirement to undergo the conformity assessments therein prescribed. This assessment, as noted, is a prerequisite for the system-product's designation as "high-risk" under Article 6 of the AI Act.

Paragraph 2 of Article 6 further specifies that AI systems listed in Annex III are also deemed "high-risk". Annex III includes, by way of example, systems related to fields such as «Biometrics», «Critical infrastructure», «Education and vocational training», «Law enforcement», «Administration of justice», among others. AI systems operating in the healthcare domain but outside strictly clinical contexts (*e.g.*, when used for healthcare planning) may reasonably be classified under Sector no. 5 of Annex III, which pertains to «Access to and enjoyment of essential private services and essential public services and benefits». This category includes, notably, «AI systems intended to be used by public authorities or on behalf of public authorities to evaluate the eligibility of natural persons for essential public assistance benefits and services, including healthcare services, as well as to grant, reduce, revoke, or reclaim such benefits and services».

[26] The reference is to Regulation (EU) 2025/327 of the European Parliament and of the Council of 11 February 2025, «on the European Health Data Space and amending Directive 2011/24/EU and Regulation (EU) 2024/2847», hereinafter: EHDS Regulation. With particular regard to the secondary use of electronic health data (Articles 53 *et seq.* of the Regulation), cf., in Italian legal scholarship, A.A. MOLLO, *Prime riflessioni sul Regolamento europeo sullo spazio europeo dei dati sanitari: l'uso secondario e il diritto di esclusione riguardo al trattamento dei dati sanitari elettronici personali*, in *BioLaw Journal – Rivista di BioDiritto*, No. 3, 2025, pp. 11 ff., esp. pp. 18 ff.; M. OROFINO, Digital Health *e diritto alla salute: l'impatto del Regolamento EHDS sui sistemi sanitari nazionali*, in *Italian Papers on Federalism*, No. 1, 2025, pp. 175 ff., esp. pp. 184 ff.; as well as ID., One Digital Health *e circolazione dei dati: tra mercato unico e diritti costituzionali*, in *Corti supreme e salute*, No. 1, 2025, pp. 257 ff., esp. pp. 269 ff.; and, extending beyond a strictly constitutional perspective, the contributions collected in A.M. PINELLI (ed.), *Sanità digitale. Regolamento "EHDS" (UE 2025/327) sullo spazio europeo dei dati sanitari*, I, *Uso dei dati e assetti organizzativi*, Pisa, 2025, and esp. those included in Part III.

rights protections are ensured[27]. The EHDS thus adds an additional layer of guarantees, while also expanding the institutional responsibility for ensuring that the data infrastructures underpinning AI-supported healthcare are equitable, accountable, and constitutionally compatible.

This approach is echoed at the domestic level. The recently approved Enabling Law «*in materia di intelligenza artificiale*»[28] provides – under Article 7, titled «Use of artificial intelligence in the healthcare and disability sectors» – that «artificial intelligence systems used in the healthcare sector and the related data employed» must be «reliable, periodically verified and updated», in order both to «minimize the risk of errors» (paragraph 6) and to ensure that even a partial digitalization of healthcare and social-health services is not carried out according to «discriminatory criteria» (paragraph 2).

While essential, data governance is not in itself a sufficient safeguard against biased processes and outcomes. As will be explored further, the very way a task is framed and presented to an AI system can have a significant impact on its performance, sometimes giving rise – albeit indirectly – to discriminatory effects against members of minority groups.

[27] Article 53 of the EHDS Regulation lists among the purposes «for which electronic health data can be processed for secondary use» those of «scientific research related to health or care sectors that contributes to public health or health technology assessments, or ensures high levels of quality and safety of healthcare, of medicinal products or of medical devices, with the aim of benefiting end-users, such as patients, health professionals and health administrators»; within this category fall activities involving the «training, testing and evaluation of algorithms, including in medical devices, *in vitro* diagnostic medical devices, AI systems and digital health applications». As noted by M. OROFINO, Digital Health *e diritto alla salute: l'impatto del Regolamento EHDS sui sistemi sanitari nazionali*, cit., p. 184, «the textual specification that [the research purpose] includes […] the activity of training, testing and evaluating algorithms» aims, among other things, precisely «to make available *high-quality electronic health datasets* for the training of AI models and systems intended for use in the medical field» (emphasis added). Translations from Italian are the author's own.

[28] The reference is to Law No. 132 of 23 September 2025, entitled «*Disposizioni e deleghe al Governo in materia di intelligenza artificiale*». Translations from Italian are the author's own.

3. *AI, Public Health Policy, and the Reproduction of Intersectional Discriminations*

The "framing of the task" is a critical step in the design of predictive models used for the stratification of patient populations.

In healthcare planning, "population stratification" refers to the process of allocating individuals into graded categories – typically labelled as "risk classes" – according to their expected health risks or, depending on the perspective, their projected healthcare needs. This form of stratification is generally performed through predictive modelling.

Broadly speaking, predictive models process large volumes of data to identify recurring patterns, latent correlations, and underlying trends, which are then leveraged to generate inferences about future events. More specifically, stratification algorithms use observable variables – such as sex, age, and clinical history – to estimate a person's individual health risk. Drawing on patterns extracted from historical data, the algorithm predicts the likelihood of future healthcare needs (understood as the probability of accessing health services) and, on that basis, assigns individuals to a defined risk bracket. These classifications are meant to serve as the foundation for informed clinical and organisational decision-making.

When deployed at an institutional scale[29] – rather than left to the discretion of individual healthcare providers – models of this kind can, as will be further discussed, meaningfully contribute to enhancing the efficiency and appropriateness of healthcare interventions. They allow

[29] As stipulated, *inter alia*, by Ministerial Decree No. 77 of 23 May 2022 («*Regolamento recante la definizione di modelli e standard per lo sviluppo dell'assistenza territoriale nel Servizio sanitario nazionale*») – enacted pursuant to Mission 6 («*Salute*») of the National Recovery and Resilience Plan – only the adoption of a «common stratification model throughout the national territory» is capable of guaranteeing, through «the development of a uniform language», «equity of access and homogeneity of care» across the territory (cf. Annex 1, «*Modelli e standard per lo sviluppo dell'Assistenza Territoriale nel Servizio Sanitario Nazionale*», Section 3, «*Stratificazione della popolazione e delle condizioni demografiche dei territori come strumento di analisi dei bisogni, finalizzata alla programmazione e alla presa in carico*»). Translations from Italian are the author's own.

for the timely identification of individuals at increased risk of acute health events or, more generally, of clinical deterioration, and enable their enrolment in targeted preventive programmes. According to Ministerial Decree No. 77 of 23 May 2022, stratification «uses information related to the individual's clinical, care, and social needs» in order to «identify appropriate, sustainable, and personalised interventions», subsequently articulated in the so-called «Health Project»[30]. The early activation of care pathways – through prompt patient intake, *i.e.* prior to the occurrence of critical events – yields benefits for both the individual, by improving health outcomes, and the system, by enabling a more efficient allocation of limited human and financial resources. In normative terms, the use of stratification algorithms may thus contribute to a fuller realisation of the right to health in its social dimension – namely, as an entitlement to receive actual healthcare services – while also supporting a more appropriate balancing of this right with other legitimate, albeit non-fundamental, policy objectives, such as the rational containment of public healthcare expenditure[31].

It has been aptly observed that «the objectives pursued in developing such [stratification] tools» simultaneously respond «to (at least) three constitutional principles: the *principle of substantive equality*, which entails equitable and uniform access and patient care; the *principle of organisational appropriateness*, which is in turn linked to the constitutional requirement of administrative efficiency, insofar as these tools enhance the efficiency (at equal levels of effectiveness) of service provision; and the *principle of autonomy and differentiation*, as they enable the delivery of public services to be adapted more closely to territorial specificities, in accordance with the organisational autonomy of subnational entities»[32]. Nonetheless, it is precisely in relation to equality that significant concerns may arise.

[30] Cf. Ministerial Decree No. 77 of 23 May 2022, Annex 1, Section 3.

[31] See above, esp. n. 5.

[32] Cf. T. BALDUZZI, *La sanità digitale negli Stati decentrati: stratificazione e modelli predittivi del bisogno di salute nel riparto di competenze tra livelli territoriali*, in D. MORANA (ed.), *La salute tra i diritti e nei territori. Questioni costituzionali nel rapporto Stato-Regioni*, Turin, 2025, p. 186. See also EAD., *Gli indicatori del bisogno di salute: tra appropriatezza, autonomia ed equità nel Servizio sanitario nazionale*, in *Federalismi.it*, No. 3, 2025, esp. pp. 138 ff.

A widely cited study analysing the performance of a predictive model used across numerous U.S. hospitals revealed a troubling pattern: the model systematically underestimated the healthcare needs of patients «who self-identified as Black», often placing them in lower-risk categories than their actual health status would justify – as evidenced by subsequent health trajectories[33]. Importantly, this disparity did not stem from direct racial bias: «Black» individuals were not explicitly penalised based on their racial or ethnic identity. Rather, the issue lay in the selection of variables considered: the model relied heavily on past healthcare expenditure as a proxy for clinical need; however, due to a combination of structural, historical, and socio-economic factors, «Black» individuals – at least in the U.S. context – tend to access healthcare less frequently than warranted by their health conditions, thus generating cost patterns that do not accurately capture their actual health status or requirements[34].

The risk, in short, is that a "framing of the task" lacking adequate methodological reflection and the selection of intrinsically problematic variables – most notably the use of individual healthcare costs as a proxy for health status – may cause stratification algorithms to drift from their intended purpose. Instead of advancing the paradigms of «Population Medicine» and «Proactive Healthcare»[35], they may become amplifiers of inequalities, capable of generating discrimination directly grounded in socio-economic status and indirectly traceable, in the majority of instances, to attributes such as «sex», «race», or the patient's «personal conditions» (to borrow the terminology of Article 3 of the Italian Constitution).

[33] See Z. OBERMEYER *et al.*, *Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations*, in *Science*, Vol. 366, 2019, pp. 447 ff.

[34] As highlighted by Z. OBERMEYER *et al.*, *op. cit.*, p. 450, two main factors help explain this outcome: first, in the U.S. context in particular, «poor patients face substantial barriers to accessing health care, even when enrolled in insurance plans» (structural links between race and socio-economic disadvantage are well established, in the U.S. and beyond); second, «Black patients have reduced trust in the health care system, a fact that some studies trace to the revelations of the Tuskegee study and other adverse experiences». For further details on the so-called Tuskegee study, see above, n. 19.

[35] Cf. Ministerial Decree No. 77 of 23 May 2022, Annex 1, *passim*.

The example of stratification algorithms employed in healthcare planning is especially instructive, even beyond its specific context, as it provides evidence that: *(i)* individual identity markers tend to overlap and interact, generating compound effects (these effects manifest, on the "pathological" side, as instances of multiple or, more commonly, intersectional discrimination)[36]; *(ii)* virtually all legally protected characteristics bear, directly or indirectly, upon an individual's socio-economic positioning, which thereby functions as both a summative representation and a lowest common denominator of marginalisation; *(iii)* omitting explicit consideration of an identity trait "upstream" does not prevent its influence "downstream" (*i.e.*, on subsequent decision-making processes), with a persistent risk of structural or emergent

[36] The simultaneous presence of various forms of disadvantage often acts as a multiplier: rarely can individual grounds of discrimination be isolated and assessed independently; more commonly, «the grounds of discrimination intertwine, creating a unique type of discrimination», known as intersectional discrimination (cf. European Parliament Resolution on «Intersectional discrimination in the EU: the socio-economic situation of women of African, Middle Eastern, Latin American and Asian descent», 6 July 2022, esp. Recital A). When several marginalised identity traits converge in the same individual, their impact is not merely additive: instead, their combined effect tends to exceed the sum of the harms or stigma caused by each factor considered on its own. As Kimberlé Crenshaw – widely recognised as the founding figure of intersectionality – famously observed, particularly with respect to the nexus of race and gender: «the intersectional experience is greater than the sum of racism and sexism» (K. CRENSHAW, *Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics*, in *University of Chicago Legal Forum*, No. 1, 1989, p. 140).

On the theoretical and practical dimensions of intersectionality within the Italian legal framework, see B.G. BELLO, *Diritto e genere visti dal margine: spunti per un dibattito sull'approccio intersezionale al diritto antidiscriminatorio in Italia*, in *Diritto e questioni pubbliche*, No. 2, 2015, pp. 141 ff.; EAD., *Intersezionalità. Teorie e pratiche tra diritto e società*, Milan, 2020; and the interview with Kimberlé Crenshaw by B.G. BELLO and L. MANCINI, *Talking about Intersectionality. Interview with Kimberlé Crenshaw*, in *Sociologia del diritto*, No. 2, 2016, pp. 11 ff.

On the specific risks of intersectional discrimination associated with the use of AI in healthcare – especially with regard to "AI-powered m-Health" solutions – see M.A. WÓJCIK-SUFFIA, *Algorithmic Discrimination in M-Health: Rethinking the US Non-Discrimination Legal Framework Through the Lens of Intersectionality*, in *BioLaw Journal – Rivista di BioDiritto*, No. 1, 2024, pp. 367 ff.

proxy discrimination; and *(iv)* an effective way forward lies in adopting a "conscious" approach – whether race-conscious, gender-conscious, or otherwise – that embraces reparative goals[37], recognises the historical and cultural context, and requires algorithms to integrate contextual awareness, not merely to prevent new forms of discrimination but also, crucially, to rectify or mitigate pre-existing or emerging biases.

4. *Beyond Individual Discrimination: Intergenerational Responsibility as a Constitutional Concern in AI-Driven Healthcare*

Thus far, the analysis has centred on the individual – that is, on what a widespread use of artificial intelligence in the medical and healthcare sectors might mean *for the person*.

Article 32 of the Italian Constitution, however, conceptualizes the right to health as both a «right of the individual» and a «social interest». It is solely by virtue of the latter that individual autonomy in matters of treatment may be subject to exceptional restriction[38]. While the

[37] As noted by J.L. DAVIS, A. WILLIAMS, M.W. YANG, *Algorithmic reparation*, in *Big Data & Society*, No. 2, 2021, p. 1, who advocate for an integration of intersectional theory and reparative practices within the emerging field of algoretics: «By default, technologies reflect and reinforce existing social orders, expressing and materializing hierarchical relations. However, technologies can also be tools of liberation. They can expose, undo, and reshape *status quos*. This latter project necessitates concerted and targeted efforts, underpinned by socially informed perspectives. In service of such efforts, we present algorithmic reparation as a concept and a scaffold for Intersectional approaches to machine learning (ML) systems, displacing *fairness* in favor of *redress*. Beyond improving code, a reparative approach uses computational tools for social intervention, while critically assessing when and where computation does not belong» (emphasis added).

[38] Pursuant to Article 32 of the Italian Constitution, the legitimacy of imposing a «specific» («*determinato*») medical treatment is contingent upon the fulfilment of several cumulative conditions: *(i)* compliance with the statutory reservation (*riserva di legge*) enshrined in the second paragraph of Article 32 (on the nature and scope of which, see, most recently, D. MORANA, *Il vaccino* à la carte*: l'indebolimento della riserva di legge dell'art. 32 Cost. in una pronuncia sugli obblighi vaccinali (indeterminati) nell'ordinamento militare*, in *Giurisprudenza costituzionale*, No. 1, 2023, pp. 269 ff.,

collective dimension of health protection does not fall within the core scope of this inquiry, it is nonetheless noteworthy that the constitutional provision on health embraces a perspective that extends beyond the individual, to encompass both the community and the healthcare system overall.

Although this aspect is largely neglected by current regulatory frameworks[39], artificial intelligence can give rise to an effect «that gen-

analysing C. cost., judgment no. 25 of 12 January 2023, on compulsory vaccination in the armed forces); *(ii)* that the legislatively imposed treatment is, indeed, «specific» («*determinato*») (cf., *inter alia*, P. BARILE, *Diritti dell'uomo e libertà fondamentali*, Bologna, 1984, p. 385; similarly, V. CRISAFULLI, *In tema di emotrasfusioni obbligatorie*, in *Diritto e società*, No. 1, 1982, esp. pp. 561 ff.); *(iii)* that such obligation is «aimed at safeguarding the health not only of the individual subject to the treatment, but also […] of the community» (S.P. PANUNZIO, *Trattamenti sanitari obbligatori e Costituzione*, in *Diritto e società*, No. 4, 1979, p. 903; see also L. CARLASSARE, *L'art. 32 della Costituzione e il suo significato*, in R. ALESSI (ed.), *L'ordinamento sanitario*, I, *L'amministrazione sanitaria*, in AA.VV., *Atti del Congresso celebrativo del centenario delle leggi amministrative di unificazione*, Vicenza, 1967, pp. 103 ff., esp. pp. 110 ff.); *(iv)* that the treatment does not pose a risk of harm to the individual's health (though «temporary and minor adverse effects» may be «tolerable»: C. cost., judgment no. 307 of 14 June 1990, pt. 2 of the *Considerato in diritto*); *(v)* that the «limits imposed by respect for the natural person» are not infringed under any circumstances, as expressly stated in the final clause of Article 32, paragraph 2 (S.P. PANUNZIO, *op. cit.*, p. 903, argues that even the principle whereby «compulsory medical treatments must primarily aim to protect the health of the individual subjected to them» derives from «that fundamental requirement of respect for the human person set forth in the final paragraph of Article 32 of the Constitution»; see also C. MORTATI, *La tutela della salute nella Costituzione italiana*, in *Rivista degli infortuni e delle malattie professionali*, No. 1, 1961, pp. 1 ff., esp. p. 7; and, with particular reference to the intersections between «respect for the natural person» and the freedom of conscience and religion under Article 19 of the Constitution, V. CRISAFULLI, *op. cit.*, esp. p. 562). Translations from Italian are the author's own.

[39] As observed by M. TOMASI, *Intelligenza artificiale, sostenibilità e responsabilità intergenerazionali: nuove sfide per il costituzionalismo?*, in *Rivista AIC*, No. 4, 2024, esp. pp. 53 ff., the final text of the AI Act «eliminates many of the references present in the earlier version [*i.e.*, that adopted by the European Parliament] and, beyond mere declarations of principle, entrusts the concept of sustainability entirely to mechanisms and instruments such as standardisation requests, codes of conduct and information disclosure». References to environmental concerns are confined to the non-binding portion of the Regulation; consider Recital (48), which provides that

erates vulnerabilities on a potentially large scale», impacting «society as a whole»[40]. This occurs for several reasons. First, as already discussed, some of the consequences of AI systems do not concern individuals alone, but rather affect social structures, legal institutions, and the safeguards established by the legal order to prevent insidious forms of domination or subjugation. Second – and in more concrete terms – artificial intelligence entails significant «environmental costs»[41]: it increases global energy demand, contributes to carbon di-

«[t]he fundamental right to a high level of environmental protection enshrined in the Charter and implemented in Union policies should also be considered when assessing the severity of the harm that an AI system can cause, including in relation to the health and safety of persons», or Recital (155), which would oblige providers of "high-risk" AI systems to report to competent authorities «any serious incidents resulting from the use of their AI systems, meaning incident or malfunctioning leading to death or serious damage to health, serious and irreversible disruption of the management and operation of critical infrastructure, infringements of obligations under Union law intended to protect fundamental rights or serious damage to property *or the environment*» (emphasis added) within a requisite «post-market monitoring system». Moving beyond the introductory portion of the Regulation, references to sustainability and environmental themes are consigned, for the most part, to safeguard mechanisms often regarded as minimally effective. One illustrative case is Article 40 of the AI Act, under which the Commission must submit «standardisation requests» to «European standardisation organisations» that call for, among other things, deliverables «on reporting and documentation processes to improve AI systems' resource performance, such as reducing the high-risk AI system's consumption of energy and of other resources during its lifecycle, and on the energy-efficient development of general-purpose AI models». According to M. TOMASI, *op. cit.*, p. 56, «[t]he main shortcomings of this referral-based strategy lie not only in the lengthy approval process required for standard adoption (which may stretch over several years), but also – and more critically – in the fact that standardisation bodies primarily respond to market stakeholders. The involvement of for-profit organizations inevitably complicates efforts to adopt lower-impact solutions that may nevertheless entail higher costs». Furthermore, compliance with standards «is generally to be understood as voluntary and must be demonstrated through self-assessment mechanisms, whose outcomes are not necessarily disclosed» (*ibidem*). Translations from Italian are the author's own.

[40] M. TOMASI, *op. cit.*, pp. 48 f.

[41] See B. MARCHETTI, *I costi ambientali dell'IA*, in *BioLaw Journal – Rivista di BioDiritto*, No. 1, 2025, pp. 525 ff.; and also EAD., *I costi ambientali nascosti dell'intelligenza artificiale*, available at www.diariodidirittopubblico.it, 5 May 2024. On the «ambivalent» nature of digital transformation – which is said to function

oxide emissions, requires large quantities of water (used to cool circuits in data centres), and generates waste, some of which may pose risks to human health[42].

Moreover, these environmental costs are inequitably distributed across various parts of the globe. Consistent with a familiar pattern, the advancement of artificial intelligence is spearheaded and directed – via regulatory instruments and beyond – by technologically advanced nations, but its detrimental effects, particularly in terms of resource extraction and immediate harms, are most acutely felt by the so-called Majority World[43]. In other words, the burdens of artificial intelligence disproportionately impact those who gain the least from its deployment.

Moving again from a global perspective to the domestic context, in accordance with the framework of this study – and attempting in these concluding remarks to synthesize the analysis conducted so far – one must first underscore an *internal* tension between distinct aspects of the right to health. As observed[44], technological instruments can enhance the effective enjoyment of this right (by facilitating access to healthcare services, mitigating territorial disparities, and supporting clinical decision-making processes), yet they simultaneously risk entrenching pre-existing discriminatory frameworks and weakening the

«both as a catalyst for promoting environmental sustainability» and «as a source of concern regarding the protection of the environment» – refer to F. CAMISA, *Ambiente e tecnologia: l'interconnessione tra le 'transizioni gemelle'*, in *Federalismi.it*, No. 14, 2024, pp. 55 ff.; cf. also M. OROFINO, *La tutela dell'ambiente nella costruzione della società digitale europea*, in *Astrid – Rassegna*, No. 4, 2024, pp. 1 ff.

Regarding the concept of *AI for sustainability* – that is, the deployment of AI as a means of supporting sustainable development strategies and advancing positive environmental outcomes – see B.N. ROMANO, *Il delicato ruolo dei poteri pubblici nella gestione dell'IA per il contrasto alle emergenze ambientali. Vantaggi e svantaggi di un'opportunità*, in *Diritto pubblico europeo – Rassegna online*, No. 1, 2025, pp. 6 ff.; A. PIROZZOLI, *Intelligenza artificiale, sviluppo sostenibile e ambiente*, in *Consulta On-Line*, No. 1, 2024, pp. 111 ff. Translations from Italian are the author's own.

[42] Cf. M. TOMASI, *Intelligenza artificiale, sostenibilità e responsabilità intergenerazionali*, cit., pp. 50 ff.; B. MARCHETTI, *I costi ambientali dell'IA*, cit., pp. 525 ff.

[43] M. TOMASI, *Intelligenza artificiale, sostenibilità e responsabilità intergenerazionali*, cit., pp. 62 f.

[44] See above, esp. Sections 2 and 3.

legal institution of informed consent, thereby undermining the logical and normative foundation of treatment voluntariness and freedom of care. Accordingly, the impact of AI-based systems must be addressed through two separate and equally rigorous analyses: one focused on the right to receive healthcare services, the other on the freedom to choose or refuse treatment – both framed by the imperative of ensuring equal access and non-discriminatory treatment across the system of care.

Secondly, an *external* tension must be acknowledged, wherein the right to health – particularly insofar as it is positively influenced by artificial intelligence – may come into conflict with the constitutional imperative to «safeguard the environment, biodiversity and ecosystems, also in the interest of future generations»[45]. This tension is not

[45] Cf. Article 9, paragraph 3, of the Italian Constitution, as amended by Constitutional Law No. 1 of 11 February 2022 («*Modifiche agli articoli 9 e 41 della Costituzione in materia di tutela dell'ambiente*»). According to M. CECCHETTI, *La riforma degli articoli 9 e 41 Cost.: un'occasione mancata per il futuro delle politiche ambientali?*, in *Quaderni costituzionali*, No. 2, 2022, p. 353, the explicit invocation of future generations by Article 9, para. 3, «assumes, in every respect, the rank of a substantive parameter of constitutional legitimacy»: by virtue of it, a «juridical conformative obligation would arise for political choices», rendering them «measurable and reviewable in a judicial scrutiny of reasonableness not limited to non-manifest unreasonableness (or arbitrariness), but extended to the more stringent tests of suitability, necessity and strict proportionality». Translations from Italian are the author's own.

In F. RIMOLI, *Tutela dell'ambiente, ecologismo e interesse delle generazioni future: spunti per una lettura critica*, in *Italian Papers on Federalism*, No. 3, 2024, *passim*, esp. pp. 272 ff., the Author cautions against the potential "instrumentalisation" of concern for future generations, which he deems acceptable only as «an ethical supererogatory act by current generations»: from a legal perspective, and especially from a constitutional standpoint, «it is evident», he argues, «that environmental protection, insofar as it is posited as the foundation of potential limitations upon other liberties (freedom of movement, economic initiative *etc.*) or burdensome impositions relative to individual choices […] must be considered with utmost caution, because, beyond being susceptible to instrumentalization or disguising of economic interests […], it may directly or indirectly generate a significant increase in inequalities already present within the social fabric (in clear contradiction to the intendment of Article 3 of the Constitution)».

On intergenerational solidarity – again in the framework of the reform of Articles 9 and 41 of the Italian Constitution – see also, *inter alia*, F. CIRILLO, *«Anche nell'interesse delle future generazioni»? Un'indagine sul richiamo alla posterità*, in

marginal: it reflects the possibility that technological gains in the enjoyment of the right to health may come at an environmental cost, thereby engaging another constitutionally protected interest. Like any antinomy between rights or interests of comparable constitutional weight, the tension in question necessitates "resolution" through a careful balancing exercise[46].

Taken together, these internal and external tensions allow for a clearer response to the research question posed at the outset. Artificial intelligence can enhance the enjoyment of the right to health, but only within a constitutional framework that safeguards three core requirements: that equality and individual autonomy remain intact; that the organisational guarantees underpinning the collective dimension of health are preserved; and that the environmental and intergenerational costs of technological innovation are kept within constitutionally acceptable limits. Where these conditions are met, AI-driven healthcare can be understood as reinforcing – rather than reshaping or eroding – the constitutional structure of the right to health.

Viewed from this angle, the constitutional meaning of sustainability extends beyond the parameters of clinical or organisational effectiveness: it captures the broader set of conditions – social and environmental alike – on which the protection of health ultimately depends, and should accordingly inform the principles governing the development of AI technologies.

Any serious reflection on what it means for artificial intelligence to

DPCE online, No. 2, 2023, pp. 641 ff.; T. GUARNIER, *La solidarietà intergenerazionale nella prospettiva costituzionale. Prime riflessioni su alcuni nodi da sciogliere*, in *Rivista del Gruppo di Pisa*, No. 3, 2022, pp. 1 ff.; D. PORENA, *«Anche nell'interesse delle generazioni future». Il problema dei rapporti intergenerazionali all'indomani della revisione dell'art. 9 della Costituzione*, in *Federalismi.it*, No. 15, 2022, pp. 121 ff.; and, with particular reference to the climate crisis, R. BIFULCO, *Ambiente e cambiamento climatico nella Costituzione italiana*, in *Rivista AIC*, No. 3, 2023, pp. 132 ff., esp. pp. 136 ff. Translations from Italian are the author's own.

[46] While, as T. GUARNIER cautions (*La solidarietà intergenerazionale nella prospettiva costituzionale*, cit., p. 12), the "balancing technique" may be of limited utility in addressing «the intergenerational problem», as «its grounding in the present makes it unable – at times effectively, at times meaningfully – to engage with future demands». Translations from Italian are the author's own.

be "constitutionally oriented" must therefore address this dimension directly. An approach to AI development that attends (even aspirationally) to the rights and needs of the individual, yet neglects the broader, supra-individual perspective and forgets the «distant»[47] – those living in conditions of exclusion or vulnerability, those beyond the immediate "social gaze", and the generations not yet present – undoubtedly falls short, at this point in time, of the demands of constitutional legitimacy.

***

Abstract[*]

Ita

L'utilizzo di sistemi di intelligenza artificiale in ambito medico-sanitario, tanto nel contesto clinico che a fini di programmazione sanitaria, ha potenzialità molto notevoli ma presenta altrettanto notevoli rischi. Un'indagine giuridico-costituzionale di tali rischi non può fermarsi alla persona e ai suoi diritti, ma deve interessarsi – dati i costi sociali e ambientali dell'intelligenza artificiale medesima – anche alle collettività presenti e future.

Parole chiave: intelligenza artificiale, diritto alla salute, discriminazione intersezionale, responsabilità intergenerazionale

En

The integration of artificial intelligence within healthcare systems, whether in clinical contexts or for strategic planning purposes, presents remarkable opportunities alongside considerable risks. A constitutional and legal inquiry into these risks must move beyond an exclusive focus on individual rights to encompass the broader collective dimension, particularly in light of the social

[47] See A. SPADARO, *L'amore dei lontani: universalità e intergenerazionalità dei diritti fondamentali fra ragionevolezza e globalizzazione*, in R. BIFULCO, A. D'ALOIA (eds.), *Un diritto per il futuro. Tecniche e modelli dello sviluppo sostenibile e della responsabilità intergenerazionale*, Naples, 2008, pp. 71 ff.

[*] Articolo sottoposto a referaggio fra pari a doppio cieco (*double-blind peer review*).

and environmental costs these technologies entail; accordingly, the implications for both present and future communities warrant careful constitutional consideration.

Keywords: artificial intelligence, right to health, intersectional discrimination, intergenerational responsibility