ISSN-e [2675-5602]

Global Academic Nursing Journal

REFLECTION ARTICLE

Algorithms versus autonomy: the erosion of bioethical foundations in the digital health landscape

Algoritmos versus autonomía: la erosión de los fundamentos bioéticos en el panorama de la salud digital

Algoritmos versus autonomia: a corrosão dos fundamentos bioéticos no cenário digital da saúde

Victor Augusto Fontenelle Ramos Monteiro[1*]
ORCID: 0009-0006-0214-0068
Lísia Andrade Probo[2]
ORCID: 0000-0003-3993-5615
Tânia Rodrigues Furtado[1]
ORCID: 0000-0002-8171-1140
Ícaro Soares de Carvalho Pinheiro[1]
ORCID: 0000-0002-3051-223X
Raquel Pereira da Cruz Silva[3]
ORCID: 0000-0003-1034-1143
Elyssandra Keila da Costa Veloso[4]
ORCID: 0000-0002-2955-2402
Andrienny Santana da Silva[5]
ORCID: 0009-0004-4929-4836
Fabiano Fagundes Moser da Silva[6]
ORCID: 0000-0003-3866-696X
Marla Dias de Carvalho[7]
ORCID: 0009-0004-6949-3939
Vyrna Rebeca de Carvalho Alves[2]
ORCID: 0000-0002-1641-908X

[1]Universidade Estadual do Piauí. Piauí, Brazil.
[2]Universidade Federal do Piauí. Piauí, Brazil.
[3]Centro Universitário Adventista de Ensino do Nordeste. Bahia, Brazil.
[4]Hospital Universitário da Universidade Federal do Piauí. Piauí, Brazil.
[5]Faculdade do Piauí. Piauí, Brazil.
[6]Universidade de Brasília. Distrito Federal, Brazil.
[7]Centro Universitário Santa Terezinha. Maranhão, Brazil.

How to cite this article:
Monteiro VAFR, Probo LA, Furtado TR, Pinheiro ISC, Silva RPC, Veloso EKC, Silva AS, Silva FFM, Carvalho MD, Alves VRC. Algorithms versus autonomy: the erosion of bioethical foundations in the digital health landscape. Glob Acad Nurs. 2025;6(4):e527. https://dx.doi.org/10.5935/2675-5602.20200527

***Corresponding author:**
pesquisavictoraugusto@gmail.com

Submission: 10-21-2025
Approval: 11-30-2025

Abstract

The digitalization of healthcare, while promising greater efficiency and personalization, introduces profound ethical tensions by operating under algorithmic logics that can subvert traditional bioethical principles. This theoretical reflection study aimed to analyze, considering critical bioethics and the philosophy of technology, how the architectures of digital platforms erode the principle of autonomy and challenge the care relationship in healthcare. Through an integrative critical literature review of the Medline, SciELO, and Google Scholar databases, it is argued that this erosion occurs in three interconnected dimensions: the redefinition of autonomy as algorithmic compatibility, the transformation of the therapeutic relationship by engagement metrics unrelated to care, and the colonization of the clinical space by the rationality of dataism. It concludes that this is a structural conflict (mismatch) between the relational ethics of clinical care and the instrumental ethics of platforms, requiring, as a response, a resilient ethical praxis in the field of health, based on critical literacy, advocacy for transparency, and the development of alternative technologies centered on the therapeutic values of the bond and patient safety.

Descriptors: Bioethics; Personal Autonomy; Digital Technology; Patient Safety; Professional-Patient Relations.

Resumén

La digitalización de la atención médica, si bien promete mayor eficiencia y personalización, introduce profundas tensiones éticas al operar bajo lógicas algorítmicas que pueden subvertir los principios bioéticos tradicionales. Este estudio de reflexión teórica tuvo como objetivo analizar, a la luz de la bioética crítica y la filosofía de la tecnología, cómo las arquitecturas de las plataformas digitales erosionan el principio de autonomía y desafían la relación de cuidado en la atención médica. Mediante una revisión crítica e integradora de la literatura en las bases de datos Medline, SciELO y Google Scholar, se argumenta que esta erosión se produce en tres dimensiones interconectadas: la redefinición de la autonomía como compatibilidad algorítmica, la transformación de la relación terapéutica por métricas de compromiso no relacionadas con la atención, y la colonización del espacio clínico por la racionalidad del dataísmo. Se concluye que se trata de un conflicto estructural (desajuste) entre la ética relacional de la atención clínica y la ética instrumental de las plataformas, que requiere como respuesta una praxis ética resiliente en el campo de la salud, basada en la alfabetización crítica, la defensa de la transparencia y el desarrollo de tecnologías alternativas centradas en los valores terapéuticos del vínculo y la seguridad del paciente.

Descriptores: Bioética; Autonomía Personal; Tecnología Digital; Seguridad del Paciente; Relaciones Profesional-Paciente.

Resumo

A digitalização da saúde, embora prometa maior eficiência e personalização, introduz tensões éticas profundas ao operar sob lógicas algorítmicas que podem subverter os princípios bioéticos tradicionais. Este estudo de reflexão teórica teve como objetivo analisar, à luz da bioética crítica e da filosofia da tecnologia, como as arquiteturas das plataformas digitais corroem o princípio da autonomia e desafiam a relação de cuidado em saúde. Por meio de uma revisão crítica integrativa de literatura das bases Medline, SciELO e *Google Scholar*, argumenta-se que a corrosão ocorre em três dimensões interligadas: a redefinição da autonomia como compatibilidade algorítmica, a transformação da relação terapêutica por métricas de engajamento alheias ao cuidado e a colonização do espaço clínico pela racionalidade do dataísmo. Conclui-se que este é um conflito estrutural (*mismatch*) entre a ética relacional do cuidado clínico e a ética instrumental das plataformas, exigindo como resposta uma práxis ética resistente no campo da saúde, baseada em alfabetização crítica, advocacia por transparência e desenvolvimento de tecnologias alternativas centradas nos valores terapêuticos do vínculo e da segurança do paciente.

Descritores: Bioética; Autonomia Pessoal; Tecnologia Digital; Segurança do Paciente; Relações Profissional-Paciente.

Introduction

Contemporary healthcare is inexorably mediated by digital technologies. The promise of more democratic, precise, and personalized healthcare, driven by big data, artificial intelligence, and connectivity platforms, coexists with an ethical paradox of profound consequences[1,2]. While these tools broaden access to information and new modes of care, their underlying infrastructure, governed by opaque algorithms and economic logics of engagement and data extraction, can systematically undermine the autonomy of patients and professionals, eroding the relational foundations upon which bioethics and the practice of care have been historically built[3,4].

In the Brazilian context, marked by profound social, political, and ideological inequalities, the penetration of smartphones and social media has transformed them into primary tools for seeking health information, self-managing chronic conditions, and even for professionals offering services[5,6]. However, this immersion occurs in a predominantly commercial digital environment, where global platforms dictate the rules of interaction. Studies demonstrate how recommendation algorithms, designed to maximize screen time, frequently prioritize emotionally charged and polarizing content, including misinformation about vaccines and treatments, creating informational environments ("bubbles" or "echo chambers") that limit, rather than expand, exposure to diverse and evidence-based perspectives[7,8]. This phenomenon directly attacks autonomy, a pillar of principlist bioethics, which presupposes the capacity to make free and informed decisions[9].

The problem, therefore, goes beyond the mere quality of the available information; it involves a structural erosion of bioethical foundations. The logic of dataism, being the belief in objectivity and the decisional primacy of quantified data, tends to reduce the person to a set of data points, neglecting their narrative, existential context, and values[10,11]. Simultaneously, healthcare professionals see their practical autonomy restricted by algorithmic and non-transparent clinical decision support systems, and their relationship with patients is reconfigured by communication dynamics that prioritize performance and immediacy over attentive listening and bonding[12,13]. The research question that arises is: "In what ways do the technical and economic architectures of digital platforms reconfigure or subvert classic bioethical principles, especially autonomy, and what implications does this have for the effectiveness and integrity of healthcare?”.

The hypothesis explored here is that there is a fundamental incompatibility, termed a mismatch, between the relational ethics of care and the instrumental ethics of dominant digital platforms. While the former is centered on the person, trust, responsibility, and uniqueness, the latter is centered on data, engagement, predictability, and scale[4,14]. The relevance of this theoretical reflection is twofold: first, it offers an urgent critical diagnosis, articulating debates from the philosophy of technology and bioethics with the concrete challenges of digital health. Second, by denaturalizing the neutrality of digital tools, it seeks to provide a conceptual framework so that professionals, managers, and policymakers can critically evaluate, adopt, and demand technologies that truly serve an ethical and emancipatory care practice.

Given the above, this study aims to analyze, from the perspective of critical bioethics and the philosophy of technology, how the algorithmic architectures and operational logics of digital platforms erode the principle of autonomy and challenge the foundations of the healthcare relationship, proposing frameworks for a resilient ethical praxis.

Methodology

This is a theoretical reflection study, qualitative and analytical-critical in nature, whose objective is to examine and problematize assumptions, conceptual relationships, and ethical implications within the scope of digital health. This design is suitable for investigations that seek to produce knowledge through the logical and argumentative analysis of complex ideas, theories, and phenomena, without generating primary empirical data[15]. The investigation was conducted through an integrative critical literature review, a strategy that allows for the synthesis and analysis of theoretical perspectives and evidence from diverse sources to develop a new interpretation or conceptual model on a given topic[16].

The search was conducted between March and May 2025 in the indexing databases of the US National Library of Medicine (Medline), Scientific Electronic Library Online (SciELO), and Google Scholar. The search combined controlled descriptors (MeSH/DeCS) and free terms in Portuguese, English, and Spanish, articulated by the Boolean operators "AND" and "OR". The following search strategies were developed and subsequently integrated: for the Technology/Platforms axis ("Algorithms" OR "Artificial Intelligence" OR "Social Media" OR "Digital Platforms") AND ("Ethics" OR "Bias" OR "Transparency"); for the Bioethics/Principles axis: ("Bioethics" OR "Personal Autonomy" OR "Principlism" OR "Beneficence" OR "Non-maleficence" OR "Justice") AND ("Digital Health" OR "Health Communication"); and for the Philosophical Critique axis: ("Philosophy" OR "Critical Theory" OR "Surveillance Capitalism" OR "Dataism") AND ("Health" OR "Medical Care").

The screening, eligibility, and selection process followed clear inclusion and exclusion criteria. Original and review scientific articles, books, book chapters, and technical reports from international organizations, published between 2015 and 2025, whose central focus was the analysis of the intersection between ethics, digital technologies, and healthcare practices, were included. Priority was given to material that adopted a critical perspective, examined socio-technical impacts, or discussed philosophical foundations for the problem in question. On the other hand, studies of an exclusively technical or instrumental nature lacking ethical or social reflection, opinion editorials without systematic argumentative support, and so-called grey literature not submitted to peer review were excluded. An exception to this last criterion was made for normative and position papers issued by regulatory agencies of recognized relevance in the field, considered essential primary sources for the regulatory context.

The analysis of the selected material followed a conceptual and thematic content analysis approach, adapted for theoretical studies[17]. The process was conducted in three interrelated stages: 1) exploratory reading and organization, in which the texts were read in full and organized into a summary table containing key concepts, central arguments, and theoretical frameworks; 2) analysis and problematization, a stage in which the concepts and arguments were compared and contrasted, with the aim of identifying tensions (e.g., autonomy vs. algorithmic determinism), paradoxes, and divergent lines of argumentation in the literature; and 3) synthesis and argumentative construction, the moment in which, based on the identified tensions, the study's own critical line of reasoning was constructed, articulating the frameworks of critical bioethics and the philosophy of technology to answer the proposed objective.

Results and Discussion

The reflection undertaken reveals that the tension between algorithms and autonomy is not merely an undesirable side effect of the digitalization of healthcare, but rather a symptom of a deeper ontological conflict. The elucidation of this conflict shows that the algorithmic infrastructure that underpins the digital healthcare landscape operates according to an instrumental rationality fundamentally distinct from the practical rationality that guides ethical care. This dissonance generates a process of corrosion that acts on three interconnected fronts: the reconfiguration of moral agency itself, the transformation of the therapeutic relationship, and the colonization of the ethical space by a logic of prediction and control.

The first and most subtle form of erosion concerns the redefinition of autonomy as algorithmic compatibility. The principle of autonomy, as formulated in principlist bioethics, is based on the capacity for self-determination of a reflective subject, situated within a context of values and relationships[9]. However, in the platform ecosystem, autonomy is progressively reinterpreted as a system's ability to predict and manage preferences. Personalization, presented as an unequivocal good, often functions as a nudging mechanism, being a choice architecture that subliminally guides behavior[3]. When a social media or health app algorithm prioritizes content that confirms existing biases or suggests pre-defined paths based on population data correlations, it is not expanding the freedom of informed choice. In fact, it is restricting the horizon of possibilities available to the individual, replacing moral deliberation with automated suggestion. Autonomy is thus emptied of its substantive content; critical reflection and authorship over one's own life are reduced to a series of binary choices (like/dislike, accept/reject) within a technologically determined menu. The paradox lies in the fact that the individual feels increasingly empowered by the customization of their digital experience, while their real agency over the criteria that shape that experience drastically diminishes. This process is aggravated by the phenomenon of the "filter bubble," in which recommendation algorithms, in seeking to maximize engagement, create homogeneous informational environments that reinforce pre-existing views and isolate the user from dissonant and challenging perspectives[18]. In a healthcare context, this can mean selective exposure to information that confirms erroneous beliefs about treatments or the perpetuation of stigmas in mental health, severely hindering the formation of a truly informed and autonomous opinion for decision-making.

This reconfiguration of individual agency has direct and profound consequences for the care relationship, which constitutes the core of healthcare practice. The analysis demonstrates that digital platforms frequently operate disruptive mediation in this relationship, transforming it from a dialogical encounter between people (professional-patient) into a triadic interaction in which the algorithm assumes an active and non-transparent role. On the one hand, clinical decision support systems based on machine learning can introduce a new type of authority, algorithmic authority, which challenges professional judgment and can lead to unjustified distrust or, conversely, to an uncritical delegation of clinical responsibility[11]. The "black box" of many of these systems prevents professionals from understanding the reasoning behind a recommendation, hindering their critical evaluation and ultimate accountability for the clinical act. On the other hand, communication mediated by commercial platforms between professional and patient is shaped by engagement metrics (likes, views, screen time) that are unrelated to therapeutic objectives. The pressure for instant responses and constant accessibility performance can erode necessary professional boundaries and transform the therapeutic relationship into a successful customer service[12]. The relationship, once based on trust and attentive presence, risks becoming a transaction governed by the logic of immediate return and consumer satisfaction. Care, understood in the Heideggerian sense, is an active and present concern for the being of the other, involving patience, listening, and singular responsibility[14], it is thus threatened by a logic of interaction that prioritizes speed, scale, and quantifiable metrics. The very notion of confidentiality, the cornerstone of the therapeutic relationship, is called into question by the data economy of the platforms, which transforms the intimacy of the consultation into a commodity for targeted advertising[19].

The third axis of corrosion is systemic and ideological, manifesting itself in the rise of dataism as a new rationality for health[10]. This logic, which understands reality primarily as a flow of data and believes that optimal decision-making emerges from the algorithmic analysis of these flows, promotes a dangerous equivalence between what is quantifiable and what is relevant. Health conditions, experiences of suffering, and life projects are translated into data points, traceable and optimizable. In this process, personal narrative, ambiguity, uncertainty, and subjective values, which are inextricable elements of the human condition and clinical practice, are systematically devalued or excluded because they are not fully operationalizable. The promise of an objective and unbiased medicine, guided by data, ignores the fact that the data sets themselves are historical and social products, often carrying and reproducing structural inequalities of race, gender, and class[11,20]. The belief in the neutrality of the algorithm thus serves as a camouflage for the crystallization of prejudices, now cloaked in the seemingly unquestionable authority of data science. Justice, another bioethical pillar, is doubly compromised: first, when already marginalized populations are further harmed by systems trained on data that do not represent them, generating less accurate diagnoses or prognoses[20]; and secondly, when access to the best digital health resources is determined by the ability to generate data that actually has value for analysis, decision-making, and improving quality of life, for example, and not by need, new digital social determinants of health are created. Furthermore, the logic of dataism, in its incessant search for predictive correlations, tends to pathologize normal variations and promote a medicalization of daily life, where moods, eating habits, and sleep patterns become constant targets of monitoring and optimizing intervention, under the pretext of well-being[21].

The initial hypothesis of a fundamental mismatch is therefore dramatically confirmed. The relational ethics of care, which is contextual, narrative, and focused on individual well-being, clashes head-on with the instrumental ethics of platforms, which is abstract, correlational, and focused on efficiency and engagement. Eroded autonomy is not an accident, but a predictable result of this asymmetrical encounter. This does not imply, however, a Luddite rejection of technology. On the contrary, criticism is a necessary step towards a more reflective and sovereign appropriation. The urgency of developing a resilient ethical practice is highlighted, involving: a) critical literacy in algorithms for professionals and patients, demystifying their functioning and potential biases; b) advocacy for transparency, auditability, and regulation of systems used in healthcare, demanding interpretable explanations and safeguards against algorithmic discrimination; and c) the creation and valorization of alternative socio-technical technologies designed from relational values of care, privacy by design, and public interest, and not from data extraction and surveillance capital. The future of digital health will depend not on uncritical submission to the logic of platforms, but on our collective capacity to subordinate them to an ethical project of care that reaffirms, in the digital space, the primacy of the person, their narrative, and their reflective autonomy.

Conclusion

It was found that the erosion of autonomy in the digital health landscape is not a passing anomaly, but rather a structural outcome of the colonization of healthcare by the instrumental logic of platforms. The research revealed that the reconfiguration of the capacity for moral action in terms of compatibility with algorithms, the transformation of the therapeutic relationship into a triadic interaction mediated by engagement metrics, and the rise of data fixation as hegemonic rationality act synergistically to erode the relational foundations of clinical bioethics. This process undermines not only the capacity for free and informed decision-making of patients and professionals, but also clinical integrity itself, representing a tangible risk to patient safety, treatment effectiveness, and health equity. The central hypothesis of a fundamental and irreconcilable incompatibility between the ethics of clinical care and platform ethics is confirmed, pointing to the inadequacy of merely normative or technical responses in the field of health. The outlined path, therefore, demands active resistance grounded in clinical practice: in the critical literacy of professionals and users of health services, in the demand for transparency, the possibility of auditing and specific regulation for artificial intelligence systems in health, and in the promotion of digital technologies designed with privacy from the outset and centered on the therapeutic values of bonding and comprehensiveness. True innovation in digital health will thus reside in the capacity of the health community to subordinate algorithmic architectures to an ethical-clinical project that reaffirms, on new technical bases, the non-negotiable primacy of the person, their narrative of illness, and their reflective autonomy as pillars of the safety and quality of care.

References

1. Topol EJ. Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again. New York: Basic Books; 2019.
2. Mathews SC, McShea MJ, Hanley CL, Ravitz A, Labrique AB, Cohen AB. Digital health: a path to validation. NPJ Digit Med. 2019;2:38. doi: 10.1038/s41746-019-0111-3.
3. Zuboff S. The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. New York: PublicAffairs; 2019.
4. Van Dijck J, Poell T, de Waal M. The Platform Society: Public Values in a Connective World. New York: Oxford University Press; 2018.
5. Machado KRS, Pereira MEBC, Carvalho DQ, Rocha RSB. Fontes de informação em saúde utilizadas pela população brasileira: uma revisão integrativa. Cienc Saude Colet. 2021;26(11):5745-5758. doi: 10.1590/1413-812320212611.32102020.
6. Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação (CETIC.br). TIC Domicílios 2022: Pesquisa sobre o uso das tecnologias de informação e comunicação nos domicílios brasileiros [Internet]. São Paulo: Comitê Gestor da Internet no Brasil; 2023 [citado 2024 Nov 15]. Disponível em: https://cetic.br/pt/publicacao/pesquisa-sobre-o-uso-das-tecnologias-de-informacao-e-comunicacao-nos-domicilios-brasileiros-2022/
7. Gisondi MA, Barber R, Faust JS, Raja A, Strehlow MC, Westafer LM, et al. A Deadly Infodemic: Social Media and the Power of COVID-19 Misinformation. J Med Internet Res. 2022;24(2):e35552. doi: 10.2196/35552.
8. Vosoughi S, Roy D, Aral S. The spread of true and false news online. Science. 2018;359(6380):1146-1151. doi: 10.1126/science.aap9559.
9. Beauchamp TL, Childress JF. Principles of Biomedical Ethics. 8th ed. New York: Oxford University Press; 2019.

10. Prainsack B. The political economy of digital data: introduction to the special issue. Policy Stud. 2020;41(5):439-446. doi: 10.1080/01442872.2020.1723517.
11. Mittelstadt BD, Allo P, Taddeo M, Wachter S, Floridi L. The ethics of algorithms: Mapping the debate. Big Data Soc. 2016;3(2):1-21. doi: 10.1177/2053951716679679.
12. O'Connor S. Using social media in nursing education and practice: a bibliometric analysis. Nurse Educ Today. 2022;118:105523. doi: 10.1016/j.nedt.2022.105523.
13. Sharon T. When digital health becomes digital medicine, and why it matters. Lancet Digit Health. 2018;1(3):e101-e102. doi: 10.1016/S2589-7500(18)30121-0.
14. Figueiredo LC. Heidegger e a questão da ética: o cuidado (Sorge) como ethos originário. Nat Hum. 2003;5(1):155-180. doi: 10.1590/S1517-24302003000100008.
15. Crossan MM, Apaydin M. A multi-dimensional framework of organizational innovation: A systematic review of the literature. J Manag Stud. 2010;47(6):1154-1191. doi: 10.1111/j.1467-6486.2009.00880.x.
16. Whittemore R, Knafl K. The integrative review: updated methodology. J Adv Nurs. 2005;52(5):546–553. doi: 10.1111/j.1365-2648.2005.03621.x.
17. Bowen GA. Document analysis as a qualitative research method. Qual Res J. 2009;9(2):27–40. doi: 10.3316/QRJ0902027.
18. Pariser E. The Filter Bubble: What the Internet Is Hiding from You. New York: Penguin Press; 2011.
19. Cohen IG, Mello MM. HIPAA and Protecting Health Information in the 21st Century. JAMA. 2018;320(3):231–232. doi:10.1001/jama.2018.5630.
20. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science. 2019;366(6464):447-453. doi: 10.1126/science.aax2342.
21. Lupton D. The quantified self: A sociology of self-tracking. Cambridge: Polity Press; 2016.