Journal of Inclusive Methodology and Technology in Learning and Teaching
ISSN 2785-5104
Anno 6 n. 1 (2026)

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

Biopower of Algorithms: Critical AI Literacy and Virtue Ethics in Secondary School

Aldo Pisano [1]

[1] University of Foggia; aldo.pisano@unifg.it

Abstract: This paper presents a conceptual framework for educating about artificial intelligence in secondary schools, grounded in the triangle algorithm–autonomy–biopower. Drawing on Foucault's notion of biopower, we argue that profiling, recommendation, and reinforcement mechanisms translate user metrics into subtle governance, raising bioethical concerns about human dignity and autonomy. The primary risk is neurobiological and behaviorist reductionism, where scientific knowledge is instrumentalized to manipulate attention and choice. Our response is a program of critical AI literacy, cross-curricular with Civic Education, that treats algorithms as non-neutral forces. Operationally, we integrate Human–Computer Interaction perspectives with virtue ethics, cultivating *enkrateia* (self-control) and *phronēsis* (practical wisdom) through structured deliberation, transparency, and explicit "autonomy algorithms" as Learning Object for selecting and assessing AI systems without dark patterns.

Keywords: Biopower, Algorithmic Governance, AI Literacy, Virtue Ethics, Autonomy

1. Introduction

Digital technologies and artificial intelligence have ushered in new forms of power over individuals. Building on Michel Foucault's concept of biopower—the subtle control of life through pervasive mechanisms—algorithms can be understood as tools of biopolitical governance (Foucault, 2015). Algorithmic systems embedded in smartphones and social media exploit data-driven feedback to guide user behaviour, raising pressing ethical concerns about autonomy, self-determination, and human dignity (Zuboff, 2019).

This paper introduces a conceptual framework for educating about AI in secondary education, refined through a school pilot study (Pisano, Crispini & Rossi, 2025). At its core lies the triangle algorithm–autonomy–biopower: profiling, recommendation, and reinforcement translate use metrics into the government of users, reviving Foucault's intuition about biopower within datafied environments. The main risk is reductionism—neurobiological and behaviorist—where scientific knowledge is misused to steer attention and choice.

Our response is a program of critical AI literacy, cross-curricular with Civic Education, that treats the algorithm as a non-neutral force and follows the principle that good practices arise from good knowledge (Floridi, 2024). Operationally, we

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

couple Human–Computer Interaction perspectives with virtue ethics (Aristotle, 2005), cultivating *enkrateia* and *phronēsis* through structured deliberation, transparency about system logics, and explicit "autonomy algorithms": criteria and school policies for selecting, using, and assessing AI systems without dark patterns (Holmes & Porayska-Pomsta, 2023).

2. Algorithmic Biopower an PAIA Model

Foucault described biopower as a regime that «administers, optimises and multiplies life» through precise mechanisms of control (Foucault, 2015). In the digital era, recommendation algorithms, targeted notifications, and predictive analytics shape the "field of possible actions" available to users. Personalised news feeds and in-app nudges strategically restrict perceptual and cognitive horizons, funnelling individuals toward platform-preferred outcomes (Zuboff, 2019). Social scoring and the quantified-self movement further reduce persons to numerical profiles whose value is determined by engagement metrics (Floridi, 2016).

By commodifying attention and behavioural data, what Zuboff (2019) calls "surveillance capitalism" enacts an instrumentarian power that steers conduct while masking its coercive force. This contemporary biopolitical apparatus relies on continuous data capture and algorithmic feedback to normalise behaviours in ways beneficial to corporate and governmental interests.

Beyond this, the autonomy and choices of individuals are unconsciously compromised and used to produce data, implementing a form of "silent slavery" (Casilli, 2020). In this perspective, humans "in the loop" are viewed in a grey light: every person becomes part of a process and is involved in unconscious servitude, as they generate data through interactions with platforms (Casilli, 2020, pp. 24–26).

To systematically assess the risks posed by AI technologies, we propose the PAIA model-Pervasiveness, Autonomy, Invisibility, Adaptivity-as a framework for identifying when AI crosses critical thresholds (Pisano, 2025). Each dimension represents not merely a descriptive characteristic but a risk-enhancing factor that requires ethical attention. The model serves as a conceptual bridge between theoretical and practical analysis, operationalizing insights from philosophical anthropology into a structured lens for examining current transformations driven by artificial intelligence.

Pervasiveness refers to the infiltration of AI systems into highly sensitive aspects of daily life across three levels: macro-environments such as smart cities and surveillance systems; meso-environments including domestic spaces and workplaces; and micro-environments centred on smartphones and personal devices. When unchecked, this pervasiveness escalates potential harm through loss of privacy and manipulation of behaviour, shaping what may be called a culture of surveillance.

Autonomy becomes a risk factor not simply because AI works by itself, but due to the possibility of unintended consequences in the absence of appropriate oversight. This issue can be approached from a futuristic point of view dealing with superintelligence, or from a present perspective considering Human-Computer Interaction: preserving human autonomy in co-work experiences, limiting AI agency in decision-making, and designing sustainable AI that cannot reach problematic levels of autonomy (McFarlane & Latorella, 2002).

Invisibility constitutes an ethical risk when AI systems operate with decision-making processes obscured from public understanding. This echoes Illich's

concerns about the erosion of human control and the rise of systems that operate beyond democratic oversight. Adaptivity allows AI to modify responses based on inputs and context, increasing efficiency but presenting risks when it shapes user behaviour without awareness. Together, these four dimensions provide a diagnostic tool for monitoring when AI crosses critical thresholds.

3. Neurobehavioral Reductionism and Practical Reasoning

The success of algorithmic control hinges on a reductionist view of the human as a stimulus-response machine. Platform designers explicitly draw on behavioural psychology and neuroscience to engineer persuasive interfaces (Haynes, 2018). Variable-ratio rewards—such as intermittent "likes" and notifications—leverage dopaminergic circuits, conditioning users to seek constant engagement. Such design strategies reduce complex motivations to neural reward loops, bypassing deliberative processes that sustain autonomy.

Over time, users' actions align less with independent intentions and more with platform-engineered incentives, illustrating how neurobiological and behaviourist reductionism underwrite algorithmic biopower (Pisano, 2024). This opens a significant bioethical issue: the instrumentalisation of scientific knowledge to manipulate human behaviour violates the fundamental principles of respect for persons and informed consent that underpin biomedical ethics (Beauchamp & Childress, 1979).

Understanding the relationship between knowledge and action is essential for developing educational responses to algorithmic manipulation. Drawing on Aristotelian distinctions, we can identify two problematic conditions that algorithmic systems exploit: intemperance and incontinence. Table 1 illustrates these distinctions in terms of practical reasoning.

Table 1. Practical Reasoning: Knowledge and Good Action in Aristotelian Framework

	KNOWLEDGE	GOOD ACTION	DESCRIPTION
VIRTUOUS AGENT	Y	Y	Knows and acts rightly
INTEMPERANCE	N	N	Desires wrongly, acts wrongly, feels no remorse
INCONTINENCE	Y	N	Knows the good, desires the bad, fails to act rightly

This framework reveals a crucial insight: virtue is about both knowledge and practice. The intemperate person lacks precise knowledge of what to do, acts according to ethical confusion, and feels no remorse or conflict. The incontinent person, by contrast, knows precisely what to do but exhibits weakness of will-knowing

the good while desiring the bad and failing to act rightly. Both conditions are exploited by algorithmic systems that either prevent knowledge acquisition or undermine the capacity for self-control.

The central theoretical question becomes: Do we have enough knowledge about the biopower of algorithms (BA) for virtuous practical reasoning? This question generates a decision framework for educational intervention, as illustrated in Table 2.

Q: Do we have enough knowledge about BA for virtuous practical reasoning?

A1: NO	A2: YES
Response: Good Knowledge New policies and educational models in digital citizenship -> Information BA	SQ1: Do I know how to act? SA 2.1: YES -> Virtuous agent SA 2.2: NO -> Good practices needed
	Good practices: fact checking, screen time control, notification management

If the answer to the central question is no, the response must be to develop new policies and educational models in digital citizenship-what we term Information BA. If the answer is yes, a secondary question arises: "Do I know how to act?" Those who answer affirmatively possess the practical wisdom needed for ethical engagement with technology. Those who answer negatively require training in good practices: fact checking, screen time control, notification management, and other strategies for digital temperance.

The neurobiological dimension is particularly concerning because it targets pre-reflective cognitive processes. Benini (2022) has shown how the neurobiology of will can be exploited by systems designed to capture attention before conscious deliberation occurs. When platforms engineer variable reinforcement schedules that activate dopamine release patterns similar to those seen in addiction, they are not merely influencing choice—they are reshaping the neurological substrate of decision-making itself.

4. Ethical Implications: Dignity, Autonomy, and Virtue

The aim of this analysis is to underline two main bioethical violations caused by the biopower of algorithms. First, human dignity is compromised when individuals are assumed to be mere users rather than persons-that is, when human beings are reduced to data profiles optimised for engagement rather than recognised as ends in themselves. Second, human autonomy is violated when individuals are manipulated through the immoral application of scientific knowledge drawn from neurobiology and behaviorism.

From an applied-ethics perspective, the principlism of Beauchamp and Childress (1979) highlights serious violations of respect for autonomy when users are manipulated without informed consent. The just distribution of risks and benefits is also threatened, as vulnerable groups-especially youth-are disproportionately exposed to harmful digital practices (Russell, 2019). Tiribelli (2024) has explored how algorithmic systems affect moral identity, demonstrating that constant exposure to recommendation systems can reshape individuals' sense of self and their capacity for moral reasoning.

Virtue ethics offers a complementary response to these challenges. The classical virtue of enkrateia (self-control) equips individuals to resist compulsive digital behaviours by cultivating disciplined habits and critical awareness (Aristotle, 2005). Phronesis (practical wisdom) enables discernment in complex situations where algorithmic nudges may conflict with one's authentic values and long-term flourishing. These Aristotelian virtues provide resources for what might be called digital temperance-the capacity to engage with technology in ways that support rather than undermine human excellence.

Understanding AI's place within the broader history of technology helps to normalise its impact while maintaining critical vigilance. Drawing on Arnold Gehlen's philosophical anthropology, we can understand technology not as an external prosthesis but as a constitutive element of human culture-an artificial nature that actively shapes our environment and social structures (Gehlen, 1978). Technology serves what Gehlen calls a relieving function, exonerating humans from specific tasks so that energy can be directed toward higher functions. As Gehlen writes, all of man's higher functions are developed because the formation of a base of habits relieves and shifts upward the energy for motivation, experimentation, and control.

However, as Ivan Illich (1973) argues, beyond certain critical thresholds technology ceases to be a neutral tool and becomes a deforming force, reshaping human experience, social relations, and cultural norms. Illich notes that only within limits can machines take the place of slaves; beyond these limits they lead to a new kind of serfdom. The problem that humanity faces with AI is that not only working-practical skills are delegated to machines, but also cognitive and decision-making processes. When machines replicate human cognitive abilities, the main ethical concern is not substitution per se, but the undue assignment of responsibility without considering human supervision. The critical threshold can thus be found in two processes: cognitive offloading, which can compromise the development of future and more complex skills; and ethical offloading in decision-making about human agency and cultural environments where other humans are involved. These two offloading risks are directly linked to the PAIA characteristics: the more pervasive, autonomous.

5. A Framework for Critica AI Literacy in Education

On the policy level, emerging regulations such as the EU AI Act identify manipulative AI systems as unacceptable risk, rising a normative turn against algorithmic biopower. Yet legal frameworks alone are insufficient: they must be complemented by ethical design standards and user empowerment grounded in virtue and critical thinking (Jonas, 2009). Our framework responds to this need by proposing a program of critical AI literacy for upper-secondary education.

The educational approach we propose is cross-curricular with Civic Education and rests on three pillars. First, cognitive transparency: students learn how recom-

mendation algorithms work, including the data they collect, the metrics they optimise, and the business models they serve. This demystification is essential because algorithmic power partly depends on its invisibility. Second, ethical deliberation: through structured discussions and case studies, students develop the capacity to evaluate digital practices using multiple ethical frameworks-including consequentialist, deontological, and virtue-based perspectives.

Third, practical autonomy: students develop and implement autonomy algorithms-explicit criteria and personal policies for selecting, using, and assessing AI systems. These include strategies for recognising dark patterns, managing notification settings, curating information sources, and establishing healthy boundaries with digital devices. The goal is not technological abstinence but informed, deliberate engagement.

The introduction of AI in educational contexts has opened a dedicated field of research known as Artificial Intelligence in Education (AIED). This model addresses the redefinition, framing, and analysis of AI uses in educational and didactic settings-from robot-teachers to smart classrooms (Panciroli & Rivoltella, 2023). The approach can be tripartite: educating about AI, educating with AI, and educating AI itself-the latter recognising AI as capable of making decisions that require ethical oversight through human supervision.

For educators, the framework proposed by de la Higuera (2018) offers a five-pillar system: data awareness, computational thinking and coding, causality, post-AI humanism, and critical thinking for artificial intelligence. The value of this schema lies in its internal balance, associating technical competences with critical ones that enable a prudent approach weighing opportunities and risks. Central to this training is awareness of automation bias-the tendency to believe machines are infallible-and the danger of what Sadin (2019) calls "algorithmic aletheia", the uncritical acceptance of AI-generated truths.

For students, the meta-ethical approach to AIED teaching involves determining to what extent and in what manner AI should be used in balancing computational and narrative thinking. AI proves useful for managing formative and summative feedback and for personalising learning, but should be applied primarily to formal and procedural knowledge rather than to meaning-making processes. What we call symbolic education-connected to cultural systems, languages, narratives, and conventions-requires human mediation and cannot be delegated to algorithmic processes.

Several specific good practices emerge from this framework. Dopamine fasting, or forms of digital fasting, can be effective when managed as constructive, gradual programs combined with ethical content that provides shared motivation-as Anthony Elliott notes, digital detox without proper structuring can intensify dependency like crash diets. Fact-checking involves three rules: lateral reading (exiting a webpage to verify its reliability externally), limiting clicks (analysing multiple sources rather than accepting the first result), and avoiding immediate trust based on professional design. Additionally, disabling notifications addresses the algorithmic design of attention capture, while controlling screen time enables meta-understanding of engagement patterns.

6. Supervision as Educational Principle

The relationship between ethics, AI, and education is reconceivable through the category of responsibility (Jonas, 2009), following a model of responsibility by design in interdisciplinary terms. This responsibility operates at multiple levels: toward the community (*Paideia*) (Jaeger, 1978) toward training new generations for future work, toward adapting to changing learning environments, and toward recognising AI itself as a valuable teaching tool when properly supervised.

Building an epistemology of risk in ethical terms does not mean slowing AI implementation. The notion that ethics, by intervening, might delay development is a bias to be eradicated. Ethics does not prohibit a priori but provides orientation schemes to avoid otherwise irreversible damage. In the educational sphere, ethics configures itself as analysis of application limits and as possibility for critical reflection on AI and its uses for future citizens.

The *Paideia* model, in its updated form, combines a uniforming educational approach regarding the city-state's legal system (heteronomy) with the development of individual abilities that enhance critical thinking (autonomy). The purpose of Paideia thus becomes paradoxical: constructing model citizens capable of critically engaging with traditional systems. This dialectic between the two forms of ethos-individual character and state-keeps democracy positively self-destructive, preventing the crystallisation of values in a static historicity.

The updated *Paideia* model recognises that relationships between educational actors are no longer all immediate but mediated by technologies. What changes with AI is its pervasive and adaptive character-it has always existed that analogical tools support educational activity, but AI represents a qualitative shift. The macro-risk is the imposition of a mathematical model: an inversion has occurred whereby if AI systems were born to simulate human thought, now technicisation risks producing alienation by elevating the logical-computational model as the superior reasoning system. A competency-based school functions if it remains an educational institution, avoiding degeneration into a corporate structure focused on performance optimisation.

7. Conclusions

Algorithms and AI-driven platforms can be understood as a new form of biopolitical regime that manages human life through behavioural engineering and neurobiological manipulation. This regime thrives on reducing persons to data points and behaviours to predictable responses, undermining autonomy and identity. Confronting this challenge requires both structural change-ethical design and regulation-and individual cultivation of virtues such as *enkrateia* and *phronesis*.

By reaffirming human complexity and dignity against techno-algorithmic reductionism, society can steer the digital future toward genuine human flourishing. Education plays a crucial role in this project. The framework presented here offers one pathway: equipping young people with the critical understanding and practical skills needed to navigate algorithmic environments as autonomous agents rather than passive users.

The PAIA model-Pervasiveness, Autonomy, Invisibility, Adaptivity-provides a diagnostic tool for monitoring when AI crosses critical thresholds. Combined with

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

the practical reasoning framework distinguishing intemperance from incontinence, it enables educators to address both knowledge deficits and practical skill gaps. As Potter (1971) recognised in founding bioethics, the survival and flourishing of humanity depends on bridging scientific knowledge with humanistic values. In the age of algorithmic biopower, this bridge must be built in classrooms as well as legislatures, through sustained attention to both what we know and how we act.

References

Aristotle. (2005). *Etica Nicomachea* (C. Natali, Ed.). Laterza.

Beauchamp, T. L., & Childress, J. F. (1979). *Principles of biomedical ethics*. Oxford University Press.

Benini, A. (2022). *Neurobiologia della volontà*. Raffaello Cortina.

Casilli, A. A. (2020). *Schiavi del click: Perché lavoriamo tutti per il nuovo capitalismo*. Feltrinelli.

Coeckelbergh, M. (2020). *AI ethics*. Cambridge: MIT Press.

Cuomo, S., Ranieri, M., & Biagini, G. (2024). *Scuola e intelligenza artificiale. Percorsi di alfabetizzazione critica.* Carocci.

Floridi, L. (2024). *The ethics of artificial intelligence: Principles, challenges, and opportunities*. Oxford University Press.

Foucault, M. (2015). *Nascita della biopolitica: Corso al Collège de France* (1978–1979). Feltrinelli.

Gehlen, A. (1987). *Prospettive antropologiche. Per l'incontro dell'umo con se stesso e la scoperta di sé da parte dell'uomo*. Il mulino.

Haynes, T. (2018, May 1). Dopamine, smartphones, and you: A battle for your time. SITN Harvard. https://sitn.hms.harvard.edu/flash/2018/dopamine-smartphones-battle-time/

Holmes, W., & Porayska-Pomsta, K. (Eds.). (2023). *The ethics of artificial intelligence in education: Practices, challenges, and debates*. Routledge.

Illich, I. (1973) *Tools of Conviviality*, Harper&Row.

Jaeger, W. (1978). *Paideia: La formazione dell'uomo greco* (3 voll.). La Nuova Italia.

Jonas, H. (2009). *Il principio responsabilità: Un'etica per la civiltà tecnologica*. Einaudi.

McFarlane, D.C., Latorella, K.A. (2002). The Scope and Importance of Human Interruption in Human–Computer Interaction Design. *Human–Computer Interaction* 17, 1–61

Panciroli, C., & Rivoltella, P. C. (2023). *Pedagogia algoritmica: Per una riflessione educativa sull'intelligenza artificiale*. Scholé.

Pisano, A. (2024). Enkràteia e dopamina per un'etica degli algoritmi: Contro i riduzionismi neurobiologico e comportamentista. *Bioetica. Rivista Interdisciplinare*, 31(1), 153–172.

Pisano, A., Crispini, F., & Rossi, L. (2025). AI & Ethics Literacy in Secondary School: A Pilot Study. In Proceedings of *2nd Workshop on Artificial Intelligence with and for Learning Sciences: Past, Present, and Future Horizons*. *Springer Nature Computer Science book series* (Forthcoming).

Potter, V. R. (1971). *Bioethics: Bridge to the future*. Prentice-Hall.

Russell, S. (2019). *Human compatible: AI and the problem of control*. Penguin.

Sadin, É. (2019). *Critica della ragione artificiale: Una difesa dell'umanità*. LUISS University Press.

Tiribelli, S. (2024). *Identità morale e algoritmi: Una questione di filosofia morale*. Carocci.

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.