

ARTICLE

HUMAN INTELLECT IN THE FACE OF ARTIFICIAL INTELLIGENCE: HISTORICAL REFLECTIONS THROUGH THE LENS OF THOMAS AQUINAS¹

TEREZINHA OLIVEIRA¹

ORCID : <https://orcid.org/0000-0001-5349-1059>
<toliveira@uem.br>

SOLANGE FRANCI RAIMUNDO YAEGASHI¹

ORCID: <https://orcid.org/0000-0002-7666-7253>
<sfryaegashi@uem.br>

JULIANA CALABRESI VOSS DUARTE¹

ORCID: <https://orcid.org/0000-0003-4088-3752>
<pg54532@uem.br>

¹ Universidade Estadual de Maringá. Maringá, Paraná (PR), Brasil.

ABSTRACT: The aim of this article is to analyze the use of Artificial Intelligence (AI) in human formation, considering the intellectual condition of the human person. To this end, four aspects of the issue are examined. First, it discusses how Thomas Aquinas understands that the human being can only be regarded as truly human insofar as the intellect is developed and exercised. Second, AI is analyzed as a product of capitalism, developed with the purpose of generating profit for its creators. Third, debates that present the use of AI in education as a positive tool with the potential to transform teaching processes are examined. Finally, perspectives that highlight the ethical issues related to the use of AI in schools are discussed. Throughout the analysis, it is shown that AI is a tool designed for human activities but remains dependent on human action for its operation, since it does not produce knowledge autonomously; rather, its outputs result from the compilation and processing of previously existing information. This study neither endorses nor rejects the use of AI in education. Instead, it understands AI, from a historical perspective, as a tool serving the interests of bourgeois society, while recognizing that it is the human intellect that produces, preserves, and disseminates knowledge, thereby shaping social life. The analysis is grounded in the theoretical and methodological principles of Social History, from which two fundamental concepts are derived: memory and the *longue durée*. Memory preserves the understanding of the human person as an intellectual being, whereas the *longue durée* makes it possible to bring theoretical formulations from other historical periods into dialogue with contemporary issues.

Keywords: artificial intelligence, human intellect, human formation, ethics, education.

¹ Article published with funding from the *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior* - CAPES/Brazil) for editing, layout, and XML conversion services.

O INTELLECTO HUMANO FACE AO USO DA INTELIGÊNCIA ARTIFICIAL: REFLEXÕES HISTÓRICAS SOB A LENTE DE TOMÁS DE AQUINO

RESUMO: O objetivo deste artigo é analisar o uso da Inteligência Artificial na formação humana, considerando sua condição intelectual. Para tanto, são examinados quatro aspectos da questão. Em primeiro lugar, discute-se a forma como Tomás de Aquino compreende que o ser humano somente pode ser entendido como tal na medida em que desenvolve e exerce o intelecto. Em segundo lugar, analisa-se a IA como produto do capitalismo, desenvolvida com a finalidade de produzir lucro para seus criadores. Em terceiro lugar, examinam-se debates que apresentam o uso da IA na educação como ferramenta positiva e potencialmente transformadora dos processos de ensino. Por fim, discutem-se perspectivas que destacam os problemas éticos relacionados ao uso da IA na escola. Ao longo da análise, evidencia-se que a IA constitui uma ferramenta destinada às atividades humanas, mas que permanece dependente da ação humana para seu funcionamento, não produzindo conhecimento de forma autônoma, uma vez que seus resultados decorrem da compilação e do processamento de informações previamente existentes. O estudo não adota uma posição de exaltação, nem de rejeição da IA na educação; compreende-a, historicamente, como uma ferramenta a serviço dos interesses da sociedade burguesa, sem perder de vista que é o intelecto humano que produz, conserva e difunde o conhecimento, coordenando a vida social. A análise se fundamenta nos princípios teórico-metodológicos da História Social, dos quais derivam dois conceitos fundamentais: memória e longa duração. A memória preserva a compreensão da pessoa como ser intelectual, enquanto a longa duração possibilita colocar em diálogo formulações teóricas de outros tempos com questões do presente.

Palavras-chave: inteligência artificial, intelecto humano, formação humana, ética, educação.

EL INTELLECTO HUMANO ANTE LA INTELIGENCIA ARTIFICIAL: REFLEXIONES HISTÓRICAS A TRAVÉS DE LA PERSPECTIVA DE TOMÁS DE AQUINO

RESUMEN: El objetivo de este artículo es analizar el uso de la Inteligencia Artificial (IA) en la formación humana, considerando la condición intelectual de la persona. Para ello, se examinan cuatro aspectos de la cuestión. En primer lugar, se analiza cómo Tomás de Aquino comprende que el ser humano solo puede ser entendido como tal en la medida en que desarrolla y ejerce el intelecto. En segundo lugar, se examina la IA como un producto del capitalismo, desarrollada con la finalidad de generar beneficios para sus creadores. En tercer lugar, se analizan los debates que presentan el uso de la IA en la educación como una herramienta positiva y potencialmente transformadora de los procesos de enseñanza. Por último, se abordan las perspectivas que destacan los problemas éticos relacionados con el uso de la IA en la escuela. A lo largo del análisis, se pone de manifiesto que la IA constituye una herramienta destinada a las actividades humanas, pero que sigue dependiendo de la acción humana para su funcionamiento, ya que no produce conocimiento de manera autónoma, sino que sus resultados derivan de la compilación y el procesamiento de información previamente existente. Este estudio no adopta una postura de exaltación ni de rechazo de la IA en la educación; por el contrario, la comprende, desde una perspectiva histórica, como una herramienta al servicio de los intereses de la sociedad burguesa, sin perder de vista que es el intelecto humano el que produce, conserva y difunde el conocimiento, organizando la vida social. El análisis se fundamenta en los principios teórico-metodológicos de la Historia Social, de los cuales derivan dos conceptos fundamentales: la memoria y la *longue durée*. La memoria preserva la comprensión de la persona como un ser intelectual, mientras que la *longue durée* permite poner en diálogo formulaciones teóricas de otros tiempos con las cuestiones del presente.

Palabras clave: inteligencia artificial, intelecto humano, formación humana, ética, educación.

INTRODUCTION

There has been extensive debate regarding Artificial Intelligence (AI) in various fields of human activity, such as science, education, production, the market, and healthcare. Broadly speaking, it is observed that, in practically all actions carried out by human beings, the presence of AI becomes increasingly evident, which has given rise to widely contested positions. In this context, different perspectives can be identified. While some of the literature emphasizes its potential to expand the production of knowledge and optimize processes, other authors highlight the risks associated with its use. There are also perspectives that acknowledge its potentialities but argue that its incorporation should occur in a critical and ethically oriented manner.

The international literature has highlighted that AI has the potential to transform education by favoring the personalization of learning, expanding access to knowledge, supporting teaching work, and promoting innovative pedagogical practices (Holmes et al., 2022). From this perspective, AI is understood as a technology capable of contributing to more flexible and inclusive educational processes. Although it acknowledges these potentialities and discusses aspects related to regulation and the ethical use of AI, part of this body of work devotes less attention to the implications of the growing use of these technologies for the development of human intellect.

In view of this scenario, a reflection on this phenomenon is proposed, considering that the relevance of the intellect for the constitution of the human person and for the knowledge of reality constitutes a recurrent theme in the philosophical tradition. However, the growing incorporation of AI into educational processes and intellectual activities imposes a new inquiry: how, based on the reflections of Thomas Aquinas, are we to understand the implications of the use of Artificial Intelligence for the intellectual condition of the human person and for educational formation? It is this problem that guides the argumentative trajectory of the article.

Thus, in this text, the aim is to analyze the implications of the use of Artificial Intelligence for human formation, considering the intellectual condition of the person in light of the reflections of Thomas Aquinas. From a methodological standpoint, the work is characterized as bibliographic research, of a qualitative and reflective nature, grounded in the critical analysis of contemporary academic productions on AI, education, and ethics, as well as in classical philosophical texts, particularly in the Aristotelian-Thomistic tradition. This approach allows for the establishment of a dialogue between different fields of knowledge, integrating technical, educational, and philosophical perspectives.

In this sense, the text is organized around four axes of discussion. First, reflections of Thomas Aquinas are presented, who developed a philosophical analysis of the meaning of the intellect as a constitutive condition of being a person. Drawing from the Aristotelian-Thomistic tradition, it is argued that the intellectual condition is the foundation of freedom, responsibility, and properly human action, thus offering a consistent theoretical framework for critically evaluating the limits of AI use in human formation. Although elaborated from a medieval author, such reflections remain current and pertinent for understanding the challenges posed by technology in the twenty-first century.

Second, the development and diffusion of AI in the context of capitalist relations of production are analyzed, highlighting how the commodification of knowledge, data concentration, and the opacity of technological processes influence its expansion. This analysis does not aim to reject technology in itself, but to understand that its development does not occur neutrally or detached from the historical, economic, and political conditions that produce it. Based on this understanding, the aim is to make explicit how such conditioning factors interfere in the production, circulation, and use of AI tools, as well as in the way these technologies come to mediate the construction of knowledge and the processes of human formation.

Third, approaches that positively understand the use of AI in education and other fields of knowledge are analyzed, particularly in contexts of higher and professional education. These studies are presented as part of the contemporary debate and allow for an understanding of the arguments that defend AI as a tool to support planning, learning, and decision-making, even though such positions are later problematized in light of the critical reflections developed throughout the text.

Finally, the article dedicates itself to the analysis of perspectives that draw attention to the need for ethical caution in the use of AI, especially with regard to the fallibility of data, the nonexistence of algorithmic neutrality, and the risks related to human autonomy and freedom. This dimension assumes a relevant role in the argumentation, as it allows for the demonstration of how the indiscriminate use of these technologies can negatively impact the intellectual, moral, and political formation of individuals, particularly in the educational field.

By revisiting the reflections of Thomas Aquinas, the article does not intend to offer technical answers to the challenges of AI, but rather to broaden the contemporary debate by bringing back to the forefront a fundamental anthropological question: the intellect as the foundation of the constitution of the human person and of the exercise of freedom.

Evidently, the inherent limitations of the article format prevent the thorough exhaustion of the theme's complexity. The issues discussed here may be further explored in future investigations, especially in view of the rapidity with which AI has been transforming different dimensions of social and educational life.

THE INTELLECT AS THE CONDITION FOR BEING A PERSON

In the first topic, the dialogue established with authors who reflect on AI is presented, having as its axis the consideration of the intellectual condition of the person as constitutive of the human condition. Although the reflections developed therein are grounded in Thomas Aquinas, it is considered relevant to highlight that this understanding of the human being remains present in contemporaneity, including in spaces that are not restricted to the strictly scientific domain.

In this sense, Marcos Silva, PhD in Literature and journalist, defines his understanding of the person, intelligence, and intellect in the following terms:

An intelligent person is one capable of separating the essential from the accidental, the valuable from that which lacks value, that which they need to develop a given activity from that which is irrelevant to it. Intelligence, in a purely etymological sense, refers to

this capacity for discernment. The Latin word *intelligentia* derives from *intelligere*, a term composed of *intus* (between) and *legere*, which means to choose or to read. To be intelligent is, therefore, to know how to choose the best alternative among several, but also to know how to read within things. Definition given by psychology textbooks: it is the capacity and ability to respond in the most adequate possible way to the demands that the world presents. To act with intelligence presupposes a careful and permanent exercise of *observation of the world*. The development of this comprehensive capacity qualifies the content of thought, which is aided by the sensitivity capable of characterizing the human being as a multidimensional unity (Silva, 2013, p. 1, emphasis added).

The language employed by Silva (2013) is direct and does not aim to present a conception of the intelligent person according to philosophical principles. Nevertheless, his words clearly express that which characterizes the person capable of discerning and choosing the best path, of being aware of the place they occupy and of their responsibilities in the world, capabilities that the author designates as observation of reality and, above all, the capacity to understand and act in consonance with the present time.

The lightness of the literary scholar's and journalist's words allows one to enter the thought of Thomas Aquinas with equal lightness, for although more than eight hundred years separate the author from the master of the University of Paris, his writings remain current and relevant to contemporary debate. For the reflections proposed here, two of his texts are drawn upon: Question I of the Part I-II of the *Summa Theologiae*, entitled "The Ultimate End of Man," and the work *On the Unity of the Intellect against the Averroists*. Numerous are the questions developed in the two Summas and in other writings of the Dominican that seek to explain and define the intellectual condition of the person, which is why Thomas Aquinas was chosen as the central reference for the reflections on AI.

A clarification is necessary to understand why the question of the intellect and its use by the person occupies a prominent place in the author's writings. The thirteenth century, the period in which Thomas Aquinas lived, is characterized by the institutionalization of the university, the organization of craft guilds (Le Goff, 2008), the emergence of mendicant orders, the resurgence of struggles for the freedom of communes, among other significant historical transformations. In this context, it became pressing to understand and explain the disputes surrounding the freedom and emancipation of persons based on their own intellectual condition.

The first passage of Thomas Aquinas presented here states that:

[...] Now, the proper operation of man, insofar as he is man, consists in thinking, for it is in this that he differs from animals, and it is for this reason that Aristotle places ultimate happiness in this operation. The principle by which we think is the intellect, as Aristotle says (Thomas Aquinas, 1999, p. 119)².

From the author's point of view, that which characterizes the nature of the human being and distinguishes them from other animals is their capacity to think, which is realized through the use of the intellect. This is a fundamental statement for the reflection on the use of algorithms, since it indicates

² We observe that, in citations from writings prior to the eighteenth century, we will follow, in the body of the text, the norms of the Journal and include the references according to the international standards for citing such works, in footnotes. Thomas Aquinas, *On the Unity of the Intellect*, chap. III, § 77.

that the human being only constitutes themselves as such to the extent that they exercise their intellective capacity. Thus, the question arises: how does this condition take shape in the twenty-first century, when technology comes to guide and, to a certain extent, to define human ‘wills’?

If human actions come to be predominantly conducted by technology, one could affirm that one would be facing a loss of the human condition, as formulated by Thomas Aquinas:

[...] Man differs from irrational creatures in that he has dominion over his acts. For this reason, only those actions over which man has dominion are properly called human. Now, man has dominion over his actions through reason and will. Hence, the *faculty of will and reason* shall be called free will. Thus, actions that proceed from deliberate will are properly called human. If other actions, however, are proper to man, they may be called actions of man, but they are not properly human actions, for they are not of man insofar as he is man (Thomas Aquinas, 2003, p. 32, emphasis added)³.

For Thomas Aquinas, the human being only constitutes themselves as such to the extent that their actions and wills are conducted by the intellect. It is only through the exercise of free will that the animal called man can be defined as a human person, since attitudes devoid of deliberate choice may even be practiced by human individuals, but they are so under the condition of animality, and not according to the specific nature of the human. Thus, the inquiry is resumed: what is the political and philosophical place of the person who is compelled to act not from their own will, but driven by decisions indicated by a machine designed by a concealed subject?

The following passage reinforces the idea that only the person, moved by their will, is capable of intellection: “[...] since the intellect is only moved by the will, as is proved in Book III of *De Anima* [...]” (Thomas Aquinas, 1999, p. 117)⁴. If the will devoid of the intellect does not express the human, then the question arises: what are we when our wills do not derive from the exercise of free will?

In the thirteenth century, this discussion was present and concrete. Precisely for this reason, the master introduced it in his university lectures and in some of his sermons, since it was necessary to explain, through science, a human will that was moving away from feudal values, launching itself into the seas in search of new spaces for the practice of commerce, and daring to question the excessive accumulation of wealth by the Church. It is in this spatial and political context that the Dominican master states:

It is indeed evident that this man in particular thinks, for we would never come to seek to know what the intellect is if we did not think; nor, when we seek to know what the intellect is, do we seek to know from any principle other than that by which we think. Hence Aristotle says: “I call intellect that by which the soul thinks” (Thomas Aquinas, 1999, p. 103)⁵.

When Thomas Aquinas argues that man, in concrete terms, thinks, he grants to the men of the West a power that had not hitherto been attributed to them, insofar as, until the twelfth century, the prevailing conception was that knowledge was a divine gift and that men knew it solely through the divine

³ Thomas Aquinas, *Summa Theologiae*, I-II, q. 1, art. 1, resp.

⁴ Thomas Aquinas. On the Unity of the Intellect, chap. III, § 75.

⁵ Thomas Aquinas. On the Unity of the Intellect, chap. III, § 61.

will. This conception can be found in intellectuals considered *auctoritates* to this day, such as Abelard and Hugh of Saint-Victor. Now, if thinking is proper to man, and if this thinking, as well as his actions, proceeds from his deliberate will as a human being, and not as an animal, it is he, and he alone, who possesses the consciousness capable of defining his life trajectory, and not another or something external to himself. Freedom is thus found in the very intellectual condition of the person.

It is precisely in this sense that Thomas Aquinas states:

[...] the intellect is that which is principal in man and makes use of all the powers of the soul and the members of the body as instruments; it is for this reason that Aristotle says subtly that man is intellect '*or is chiefly it*' (Thomas Aquinas, 1999, p. 129, emphasis in original)⁶.

According to the Dominican master, paraphrasing Aristotle, man is the intellect: the intellectual capacity constitutes the essence and nature of the human being, in the meantime, its first mover.

In light of these Thomistic reflections, it becomes necessary to pay attention to the use of AI when it comes to influencing ideas, choices, and wills. It is, therefore, in the face of this concern that we move to the second topic of this study.

AI AS A COMMODITY OF CAPITAL

The recent popularization of different AI models has intensified the debate regarding their use in the most diverse areas, ranging from relatively simple applications, such as assisted translation (Fantinuoli, 2023), to highly complex processes, such as the identification of patterns aimed at assisting in medical diagnoses (Zeb et al., 2024). This expansion of uses, however, does not occur without criticism. Among them, the following stands out, in the first place: the fact that AI models are predominantly trained on data produced by third parties, without, in general, any form of economic compensation or recognition to the original creators of such content. Secondly, the high expenditure of material and energy resources involved in the development, training, and operation of these systems is pointed out, without necessarily verifying a proportional or socially justified utilitarian return, among other related issues.

For a more adequate understanding of these tools and their applications, it is necessary to reiterate, even though such observation may seem evident, that they do not emerge primarily as social benefits or as disinterested advances in human knowledge. Rather, they are configured as commodities originating from a capitalist mode of production, developed with the central purpose of generating profits or attracting resources for the companies responsible for their creation. Although this logic is common to all tools, digital or otherwise, produced under such a system, a fundamental difference is identified in this case: there is no clear and objective definition of what characterizes a 'good' tool when it comes to AI systems.

In order to illustrate such a distinction, one may take a hammer as an example. If the instrument is purchased and, within a short period of time, presents flaws or breaks, the consumer may

⁶ Thomas Aquinas. *On the Unity of the Intellect*, cap. VI, § 87.

evaluate that the manufacturer did not deliver a quality product and, consequently, choose not to purchase items from the same brand again. However, when resorting to the assistance of an AI, whether for image generation, clarification of doubts, or translation of an excerpt, a different question arises: what kind of guarantee can one have that the result obtained effectively corresponds to what is being sought? It is observed that only those who already have prior mastery of the content in question would be capable of safely evaluating the quality and adequacy of the generated responses. Thus, for example, a medical professional who already possesses a well-founded diagnostic hypothesis may use an AI tool as confirmatory support; on the other hand, a professional without the necessary specialization runs the risk of compromising patient safety by unduly relying on these systems.

In this same sense, the very term “artificial intelligence” tends more to produce mystification than to clarify the technological processes effectively involved. In the field of education, it becomes essential to distinguish the different forms of AI that have emerged over the past decades. Currently, discussions regarding the use of AI in educational contexts are concentrated, above all, on data-based systems, aimed at collecting information about students for analytical and predictive purposes, frequently described by expressions such as learning analytics (Ouyang; Zheng, 2024), as well as on so-called “generative AI” (Ogunleye et al., 2024). The latter includes, for example, robotic tutors and conversational systems that propose to support learning processes through automated dialogues, constructed from the synthesis of internet searches.

These applications, on the one hand, do not exhaust what has been conventionally termed “artificial intelligence”; on the other hand, they end up being obscured when the analysis is limited to the generic use of the term. Analogously to the use of expressions such as “cloud” or “big data,” the recourse to technological jargon tends to detach technology from its material conditions of production, from its historical contexts, and from its concrete modes of operation. In particular, generative AI models are not the fruit of the supposed isolated brilliance of a “garage genius,” but rather result from decades of massive concentration of data, both in terms of access and storage capacity, constituting true latifundia, not only digital, of information (Geiger; Barr; Thomas, 2023).

Unfortunately, the problem arising from this obscuration is not restricted to a merely terminological issue. AI systems are, for the most part, products developed by private companies that hold, in the form of intellectual property, control over algorithmic details, as well as over the processes of collecting and using third-party data. In short, such systems operate as veritable “black boxes” (Casper et al., 2024), which renders them, to a large extent, inexplicable to teachers and students. This opacity directly contributes to the already mentioned difficulty of challenging, supervising, and holding these companies accountable in the face of possible failures, biases, or processes of disinformation.

In view of this scenario, central questions arise: who should be held responsible when a generative model produces false information that ends up being reproduced by students? What methodologies are effectively employed to identify, correct, and mitigate spurious results? In the absence of transparency, it becomes difficult to envision any substantial progress in this direction. The problem is aggravated when one observes numerous cases in which responses generated by AI systems reinforce racial or gender stereotypes (Yang, 2025; Sandoval-Martin; Martínez-Sanzo, 2024; Colón-Vargas, 2025). These occurrences stem from the fact that these models do not possess their own criteria to reliably distinguish adequate sources from inadequate ones, once again transferring to the user the responsibility

of verifying the information through other means – or, ideally, of already having prior mastery of the content in question. In light of this, the inquiry about the real necessity and pertinence of using this technology in such contexts imposes itself.

Another problem, no less relevant, concerns the defense of a restricted interpretation of what is understood by education. Frequently, studies on AI assistance in educational contexts focus on supposed improvements in performance measured by grades, especially in Higher Education (Ouyang; Zheng; Jiao, 2022; Chaudhry, 2023), emphasizing a pedagogy of an instrumental character. Such an approach tends to empty the central purpose of child and youth education, which consists of the integral formation of citizens, understood as autonomous, critical subjects capable of reflecting on their own existence and on the reality that surrounds them; in other words, the intellectual human being presented and defended by Thomas Aquinas.

From this perspective, educational formation should not be conceived as the mere assimilation of a set of disconnected facts, more or less correct, but rather as a process of structuring free thinking and conscious existing. It is, therefore, a matter of understanding education as a privileged space for the development of the human intellect in its fullness, and not merely as a mechanism for optimizing measurable results.

Given this conjuncture, it becomes imperative that such questions be seriously considered within the educational sphere, especially because it is already a reality, at least in distance learning, that students use AI tools as an aid in carrying out assignments and assessments. Should the information provided by these systems be incorrect or, more worryingly, used not as support for learning but as a means of avoiding intellectual effort, the educational system, frequently already overburdened, is confronted with an additional demand for surveillance and mediation (Barbosa; Alves, 2023). The search for solutions in this context tends to involve the participation of third-party agents who, often, have no economic interest in the transparency of the processes of obtaining, validating, and circulating information.

Moreover, excessive dependence on AI systems can lead to a reduction in problem-solving capacity and to the weakening of students' critical thinking (Larson et al., 2024; Essien et al., 2024), contributing to the devaluation of the teaching-learning dynamics in the classroom. This phenomenon occurs insofar as AI favors an instrumentalized pedagogy and a “goal-oriented teaching,” paraphrasing a methodology originating from the field of programming, in which the learning process tends to be reduced to the completion of tasks and the attainment of grades in the most efficient way possible, with the least expenditure of intellectual effort on the part of the student.

It also falls to educators, in a broad sense, to consider the asymmetric impacts arising from the implementation of AI tools in the classroom context (Bulathwela et al., 2024). Among the questions that impose themselves, the following stand out: in what manner will such technologies be incorporated into public education? Who will assume the financial costs of their implementation and maintenance? And, again, how can private companies be held accountable when they would, in theory, be the winners of public bids for the supply of these technological solutions? In the Brazilian context, marked by profound structural inequalities, such inquiries become even more relevant, since many schools, especially in historically marginalized regions, lack adequately equipped computer labs or even regular access to the internet.

An implementation of AI tools devoid of adequate governance and effective control mechanisms tends, therefore, to deepen the socioeconomic and regional divisions already existing in the country. Even though such criticism may be interpreted as excessively cautious or even anachronistic by some sectors, it is worth remembering that approaches guided by the principle of “move fast and break things”, a motto historically associated with technology sector companies, prove to be particularly problematic when what is at risk is not isolated products or systems, but the very social fabric and the educational future of young people enrolled in public education.

To conclude this topic, an additional problem of particular relevance is highlighted. As discussed earlier, the central objective of these tools consists of generating profit for investors and shareholders. In this sense, whenever certain information made available by AI systems is perceived as a threat to these economic interests, there is a concrete possibility that the responsible companies may choose to omit, filter, or alter it, even when it concerns factually correct information. Two recent examples, widely reported by the media, illustrate this phenomenon. The first refers to the Chinese model DeepSeek, which initially refused to answer questions related to the Tiananmen Square Massacre (Martins, 2025). The second concerns the role played by large technology companies in supporting Israel, including the use of AI systems and the commercialization of cloud storage services (Mednick; Burke; Biesecker, 2025). Added to these cases is a set of evidence present in the academic literature, such as practices of content erasure aimed at preventing the circulation of materials considered offensive (Chen, 2024), the generation of distorted or biased images by AI tools (Ungless; Ross; Lauscher, 2023), and the use of automated systems for the exclusion of scientific articles from American databases under the pretext of combating so-called “gender ideology” (Serres; Kowalczyk; Hagen, 2025).

In view of this scenario, accepting forms of censorship imposed by third-party agents and abdicating a potentially transformative didactic experience in exchange for, at best, better performance indicators or a more efficient instrumentalization of the educational process should not be mistaken for educational progress. On the contrary: such a movement tends to reinforce an impoverished conception of education, reduced to the optimization of measurable results and dissociated from its broader formative commitment (Zuin, 2021).

An emblematic example, albeit dating back a few years, is that of the large language model (LLM) T5, developed by Google and released at the end of 2019. For the training of the model from its dataset, a filtering mechanism was employed aimed at excluding terms considered “bad” or “inappropriate” (Gupta, 2024). Such a filter is available in several languages; however, English is taken here as the standard reference (LDNOOBW, [2012]). As can be verified, the exclusion list includes terms classified as “profanities”: “penis,” “vagina,” or “gay sex.” The central issue in this context does not lie in discussing whether such content should or should not be addressed with children of certain age groups, but above all in problematizing which relevant materials may have been excluded from the dataset used in the model’s training. Since there is no access to the raw data, it becomes impossible to know with precision the extent of these exclusions. It is plausible to suppose, for example, that the model was deprived of fundamental articles on the distinction between sex and gender, on sexual health, or even on basic anatomical discussions relating to mammals. In short, one does not know what was effectively removed.

In light of this, the questioning imposes itself regarding the adequacy of such a repository as a basis for the production and mediation of knowledge, as well as its pertinence as a support tool for the educational environment. There was, in this process, no form of democratization of decisions regarding what should be excluded, but rather the adoption of a centralized posture, oriented by the interest in minimizing risks through a strategy considered safer, and described as “overly conservative” by one of the dataset’s own creators (Chen, 2024).

Additionally, considering that generative AI tools depend, to a large extent, on access to materials freely available on the internet, which are predominantly produced in English, one observes the risk of omission of relevant content in other languages or, furthermore, of presenting anglicized and culturally biased versions of certain events, processes, or historical movements. This dynamic contributes to the reproduction of linguistic and epistemological asymmetries already existing in the field of knowledge production. In this sense, as Aceti (2025) points out, the restriction and censorship of image creation, even in embryonic stages (the initial stages of the development of an image concept), are applied based on a decontextualized interpretation of prohibited terms. Moreover, according to the author, such practices of restriction and censorship propagate a generalized conformity with structures that carry imperialist characteristics, hindering the growth of cultural diversity, originality, and alternative modes of thought.

It is observed, therefore, the absence of any form of dialogue between society, especially that situated in countries whose language is not English, and the companies responsible for the development of these tools, with regard to the definition of “prohibited terms” and the filtering criteria employed. In this scenario, it remains for educators to accept such systems as they are made available, without any effective possibility of contestation, contextual adaptation, or democratic participation in the decision-making processes that shape access to knowledge mediated by these technologies.

Despite the critical stance regarding the use of AI tools, understood here both as generative models and as systems for analyzing large volumes of data, the critique developed in this study should not be interpreted as a Luddite or anti-technological position. On the contrary: technology, insofar as it can liberate human beings from certain forms of labor, presents a transformative, inclusive, and disruptive potential. What is problematized, however, is the uncritical acceptance of technological products elevated to the status of contemporary fetish. It is not the mere fact that an innovation emerges surrounded by promises, marketing strategies, and capital investments that automatically qualifies it as technological advancement or progress, especially if it can jeopardize the intellectual capacity of the human being and, consequently, their freedom to arbitrate their ideas, wills, and actions, as discussed in the preceding reflections.

In this sense, a closer articulation on the part of educators with different sectors of society is necessary, especially with legislators and with professionals qualified in areas directly related to AI technologies, such as data engineers, programmers, and data analysts. Likewise, greater transparency is demanded from companies that offer these services in the processes of data collection and processing, as well as more effective forms of legal accountability, including at the international level.

Similar to initiatives aimed at open science, open data, and the development of open-source alternatives, the adoption of standardized audit schemes for these tools becomes imperative. As discussed by Casper et al. (2024), access to “white-box” models, as well as the requirement for updated

documentation that includes code, partial or total, internal testing, and methodological description, constitutes a minimum requirement for the use of these technologies in the classroom to be seriously discussed. Furthermore, it falls to educational managers to delimit, in a clear and unambiguous manner, the scope of use of these tools, especially with regard to applications aimed at the inclusion of persons with disabilities, so as to pay attention to educational spaces, whether formal or informal or social.

AI IN EDUCATION FROM THE PERSPECTIVE OF POSITIVE APPROACHES

As stated at the beginning of the exposition, there are different positions regarding the use of AI in education. Part of the literature highlights its potential to support teaching, learning, and decision-making processes, advocating for its incorporation into educational contexts from different perspectives.

The defense of AI use in education is not limited to the experiences analyzed here. The international literature has recognized its potential to personalize learning, offer individualized feedback, support teaching work, and expand educational opportunities. Holmes and Tuomi (2022) argue that AI can favor more flexible and inclusive pedagogical practices, without dispensing with teacher mediation. In a similar perspective, Luckin (2018) understands AI as a resource capable of enhancing human intelligence and supporting pedagogical decisions. UNESCO (2023) also recognizes that generative AI can expand access to knowledge and promote educational innovation, provided that its use is guided by ethical principles and governance mechanisms. Likewise, the systematic review conducted by Zawacki-Richter et al. (2019) shows that the main applications of AI in Higher Education are concentrated on supporting learning, assessment, and personalization of teaching, demonstrating broad international recognition of its potentialities.

The contributions of these authors show that AI has relevant applications in the educational field and that its use can favor different dimensions of teaching and learning. However, the problem that guides the present study does not consist of denying such potentialities, but of discussing a different question: the implications of the increasing delegation of intellectual activities to automated systems for the formation of the human person. It is precisely this anthropological and philosophical dimension, grounded in the reflections of Thomas Aquinas, that guides the critical analysis developed in the following pages. In this context, three representative studies are examined below, illustrating different ways in which the literature has defended the use of AI in educational and professional contexts.

The first example refers to research that evidences the use of AI in mathematics teaching in an undergraduate teacher training course, aimed at third-year students at a Brazilian federal university. From the perspective of Sant'Anna and Sant'Ana (2023), the technology incorporated into AI proves positive both for lesson planning and for the moment of applying exercises in the classroom. In the aforementioned article, the authors highlight that AI does not present errors.

When requesting that AI elaborate a lesson plan for the development of a class using the aforementioned software, what we obtained as a response did not present divergence from what we normally use. The lesson plans requested for each of the software programs presented the same basic structure, namely: Objective, Necessary

Materials, Target Audience, Lesson Duration, Activities, Assessment, and Conclusion. We noticed that, when proceeding with the request to AI, a crucial point for obtaining a satisfactory response is to have a request and/or question precisely to the extent of what we want, or to carry out an iterative process of refinement thereof, so that the response approaches what we expect; a possible *stopping point* will depend on how specific we are able to be when asking/requesting (Sant'Anna; Sant'Ana, 2023, p. 5, emphasis in original).

For the authors, when the responses provided by AI prove inadequate or present discrepancies in relation to the actual data, in this case previously programmed exercises, the problem lies in the lack of objectivity or clarity in the formulation of the question. No inquiries or concerns are observed on the part of the authors regarding the algorithms used or the databases employed. AI is thus understood as a tool aimed at teaching and learning, being presented as a facilitator for both professors and students in the university context.

The second example, now situated in the field of Law, is authored by a PhD holder in the area, affiliated with a federal university and active in graduate studies, in co-authorship with two doctoral candidates in Law, one of whom is a judge of the Labor Court. The text produced by the three researchers presents a language proper to the legal field and mobilizes, for the validation of its arguments, a significant set of authors and authorities, with the aim of supporting the premise that the use of AI is highly positive for the field of Law..

Caitlan Mulholland states that: “Artificial Intelligence began with the creation of the first algorithm by Ada Lovelace, continued by Alan Turing’s test and by Isaac Asimov’s creation of the Laws of Robotics, and, in the present day, the impressive human imitation capacity of Google Duplex.” The same author also conceptualizes artificial intelligence as “a computational system that simulates the human capacity to reason and solve problems through decision-making based on probabilistic analyses.”

Along the same lines, Alan Turing teaches that “when a machine is capable of behaving in such a way that it is not possible to distinguish it from the human being, it would be demonstrating some kind of artificial intelligence.”

John McCarthy, the father of artificial intelligence, defines it as follows: *the science and engineering of creating intelligent machines, especially intelligent computer programs to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable*. Therefore, artificial intelligence has as one of its objectives the development of machines with intelligent behavior, that is, that can perceive, reason, learn, communicate, and act in complex environments as well as humans can do, or possibly better (Gomes; Vaz; Dias, 2021, p. 4, emphasis in original).

According to the authorities mobilized in the text, AI is presented as a technology similar to human intelligence, capable of performing tasks with perfection, in a manner analogous to the actions carried out by people, and may even surpass them. According to the cited authors, within the scope of Law, AI would have the capacity to overcome and surpass human faculties of learning, reasoning, and acting.

Relying on another authority, the authors affirm that the use of AI in the legal field would constitute a good of public utility.

Isabela Alves recognizes the capacity of artificial intelligence to practice procedural acts necessary for the realization of Law, when she states that “Artificial Intelligence applied

to Law is nothing more than the development of machine performance to carry out tasks that, until then, could only be developed by a legal professional, and which can now be executed through programs and algorithms.” From this perspective, “*artificial intelligence can play a virtuous role, bringing about a significant improvement in the efficiency of the provision of jurisdictional activity*” (Gomes; Vaz; Dias, 2021, p. 5, emphasis in original).

According to the author cited by Gomes, Vaz, and Dias (2021), in the field of Law, AI would enable the overcoming of human performance in the processes of analysis and judgment of legal cases. The author cited by the article’s authors considers that the actions carried out by AI are virtuous, since, due to the evaluative capacity of algorithms, the machine would present greater precision and agility in the conduct of judicial proceedings than the human person. From this perspective, AI would be understood as a virtue for allowing the overcoming of human intelligence. In this regard, it can be inferred from the analysis of these authors that, for the legal field, the use of AI constitutes a social advance, insofar as it surpasses human capacities. Although such discussion is not central at this point in the exposition, the arguments presented indicate the possibility that AI may come to replace the intellectual condition of the human person within the scope of Law.

A third example comes from an article resulting from research conducted by three graduate professors and researchers, two of whom are from the field of Education. For these authors, AI constitutes a relevant resource in the classroom, insofar as it acts as a facilitator of teaching, promoting learning in a rapid manner.

For Santos, Profeta, and Profeta (2023), moreover, when comparing the teaching process conducted by a teacher of a banking and technicist profile, who does not stimulate learning or questioning on the part of students, AI would be more advantageous, since it enables access to a volume of information far superior to what such a human teacher could offer.

[...] if the teacher’s function is limited to answering questions, as the holder of knowledge to be taught (in the sense of simply “*transmitted*”), AI has some advantages from a practical point of view. It is clear, however, that there are limitations: chatbots can “occasionally generate incorrect information, generate harmful instructions or biased content” and have “*limited knowledge of the world*.” All of this is true regarding AI, but let us agree that it is also true regarding flesh-and-blood teachers – if it were not, there would be no need for so many institutional trainings at universities on representativeness, moral harassment, etc. With these caveats, one can reach the conclusion that language models such as ChatGPT are great tools for obtaining quick answers, provided that the user knows how to formulate good questions [...] and has a critical sense to doubt the answers when necessary (Santos; Profeta; Profeta, 2023, p. 11, emphasis in original).

The passage expresses the way in which the authors establish the relationships between teachers and ChatGPT. It is evidently recognized that problems may exist both in teacher training and in classroom practice, and the authors are right to highlight this aspect of pedagogical practice, which is corroborated by national education data presented in the Functional Literacy Indicator – INAF (Ação Educativa; Instituto Paulo Montenegro, 2024). However, the teacher constitutes a human person who interacts physically and intellectually with students. The abandonment of this humanized mediation and its replacement by a technology may prove highly problematic with regard to the intellectual and moral formation of children, adolescents, and young people. In this sense, one must not forget the Aristotelian

maxim according to which human beings learn fundamentally by example, nor Arendt's (2022) formulation that the child is a new being in a world that already exists. Disregarding the relevance of this human contact as a foundation of cognitive development may, therefore, jeopardize the very condition of humanity.

Further on in the text, the authors present another aspect of AI use in the school context that deserves to be considered in this reflection. From the point of view they adopt, the main problems and misunderstandings arising from the use of AI reside in the difficulty of adequately formulating the questions directed to ChatGPT.

Formulating questions for AI to offer the respective answers, or introducing contexts so that AI proposes interactions among the data presented, constitute the most probable potential uses of chatbot-type language models, such as ChatGPT, in school education, and this seems excellent to us. We conclude, however, that it will fall to the contemporary teacher to reflect on the movement that precedes the very act of questioning, which is the promotion of doubt, and on the variability in the quality of the question. It is worth remembering, in the end, something of extreme importance that we have seen regarding the functioning of these language models: the fact that they are stimulated by the incitements of the users themselves is what ultimately leads them to behave differently depending on who is using them (Santos; Profeta; Profeta, 2023, p. 13).

The issue of formulating questions directed to the machine was also observed by Sant'Anna and Sant'Ana (2023) with regard to the use of AI in mathematics teaching. In both studies, it is argued that the inquiry formulated must be clear, objective, and precisely delimit what is requested from the tool. From this perspective, the problem would lie in the subject who formulates the question, and not in possible failures of the technology in the process of data accumulation or processing, since one must not lose sight of the fact that AI operates from algorithms, and not in consonance with the human intellect. Thus, to attribute to the human being the errors or inaccuracies produced by the machine would not imply conferring upon AI a position of superiority in relation to the human and, at the same time, evading the recognition of technological fallibility?

One cannot neglect another observation present in the citation, which bears upon the idea that the teacher must be aware of the doubts that materialize in the inquiries directed to AI. This statement refers, again, to the formative basis of the teacher, since the capacity to doubt and to formulate inquiries is directly related to their level of intellectual formation. Such a finding also raises a new question: if the theoretical-practical formation of the one who questions is a necessary condition for the technology to respond adequately, how will future generations be configured, whose formative processes will be guided, predominantly, by the power of AI, and not by that of the human person? One must not lose sight of the fact that important intellectuals have dedicated themselves to the study of human formation, understanding the development of the intellect as a fundamental condition for the constitution of the person and for the exercise of freedom.

One begins with Aristotle, in the *Nicomachean Ethics*, especially in Book II, in which it is emphasized that every action of the person depends on the habits acquired throughout life through their intellectual and social formation, as well as by the reflections of Thomas Aquinas discussed in the first topic of this article. Or, furthermore, one turns to La Boétie (2025), in the *Discours de la servitude volontaire*,

in which it is argued that freedom depends on knowledge. For this reason, one must not neglect the pursuit of knowledge, for, once lost, servitude to the power of another is installed. Already at the beginning of the nineteenth century, Kant's work *On Pedagogy* (2018) stands out, in which it is affirmed that education constitutes the path by which people insert themselves into the world. It is also affirmed that it represents one of the most arduous and important human tasks, as it is a condition for the preservation of society. No less relevant for the understanding of education as fundamentum of society are the reflections of Arendt (2020, 2022), presented both in *The Human Condition* and in *Between Past and Future*. In these works, the author argues that education constitutes a reference for the preservation of civility in a given historical context and emphasizes that the child is always a new being in an old world, thus depending on the mediation of an adult in their formative process. In this sense, the question is raised about the consequences of a formation mediated predominantly by technologies, in which children, adolescents, and young people may be deprived of contact with a person capable of stimulating them to doubt, reflect, and inquire.

ETHICS IN THE USE OF AI: A NECESSARY APPROACH

After the dialogue with authors who understand AI as a positive and relevant path for teaching and research, a consideration is presented regarding the risks associated with its use when ethical limits are not properly delimited. It is clarified that the understanding of ethics adopted in this work aligns with the presuppositions present in Aristotle's *Nicomachean Ethics*, particularly with the idea that human beings possess two natures of excellence: the moral and the intellectual.

As we have already seen, there are two kinds of excellence: the intellectual and the moral. For the most part, intellectual excellence owes both its birth and its growth to instruction (which is why it requires experience and time); as for moral excellence, it is the product of habit, for which reason its name is derived, with a slight variation, from the word '*habit*.' It is evident, therefore, that none of the various forms of moral excellence is constituted in us by nature, for nothing that exists by nature can be altered by habit (Aristotle, 2009, p. 36, emphasis added)⁷.

Following Aristotle's perspective, it is observed that a person who does not learn to reflect and to inquire will hardly be capable, in cognitive terms, of formulating questions, also presenting great difficulty in understanding and acting in the face of the world that surrounds them, becoming dependent on someone or something to guide them.

Consequently, it is considered fundamental that the use of AI in education, social media, and the market in general be mediated by ethical principles. When it becomes necessary to emphasize the practice of ethical actions, as currently occurs in the academic-scientific sphere, whether in research, in publications, or in the adoption of anti-plagiarism policies, as is observed with the use of AI, this fact indicates the existence of practices that are not guided by good conduct. It is also worth recalling that the long philosophical tradition of ethics, inaugurated by Aristotle, when dealing with the two "excellences,"

⁷ Aristotle, *Nicomachean Ethics*, L. II, § 1103b.

refers exclusively to the behavior and actions practiced by the human being. Thus, to think about the application of ethical principles to a technology created and used by people necessarily leads to the observation that it is human beings who practice ethics or not. Alternatively, it would be necessary to found new ethical bases, should one expect AI to be, by itself, ethical in a human manner. Such a hypothesis would imply removing from human beings the primacy over thinking and acting and attributing to a technology the capacity to evaluate, decide, and answer for its own acts, placing the human being in the position of creator of entities endowed with attributes that, historically, were attributed to the gods in Antiquity or to God in the Latin Christian tradition.

In light of the reflections developed thus far, the formulations of Rossetti and Angeluci (2021) are revisited in order to offer some considerations regarding aspects of AI use that demand ethical practices. The authors list several principles; however, within the limits of this exposition, three are highlighted, as they are considered central to the problematization of the theme: the fallibility of error, the neutrality of algorithms, and freedom and autonomy mediated by algorithmic numbers. According to Zuin (2021), there is even the risk that a directive system, guided by AI, may come to control the mechanisms of punishment and reward in educational processes, compromising the development of ethical consciousness in the relations between teachers and students.

Due to the risks that AI use may entail for human formation, the discussion begins with the question of data fallibility.

Fallibility arises from the data mining process, which occurs inductively by correlating the mined data. Algorithms are created to relate them in a causal manner. However, this causality is not always legitimate. The programmed logic may associate an effect with an undue cause, generating unfair and inconclusive conclusions. Due to the enormous amount of data processed, correlations work well for populations, but when it comes to individuals, the transpositions of the causal relationships found may generate spurious and fallacious results (Rossetti; Angeluci, 2021, p. 8-9).

According to the authors, since algorithms operate from available global information and seek to select it to produce responses according to the inquiry directed to AI, when they receive questions involving singular information, the possibility of errors occurring is high. This is due to the fact that the machine is not trained for individual responses and, because it does not possess human intellect, being a machine, it is not capable of reflecting and selecting, in a contextualized manner, the most adequate response.

In light of this argument, it becomes possible to problematize the positions of the two previously mentioned authors, who attribute the errors produced by AI to the lack of objectivity in the formulation of questions. In reality, because this technology is programmed, as occurs with all machines, and does not possess the neuronal and cognitive plasticity proper to the human being, its responses remain conditioned by the parameters defined at the moment of its creation. Unlike the human, who can think beyond “programming” or, in colloquial terms, “outside the box,” and present variable and unexpected responses, AI operates within previously established limits. It is in this sense that Rossetti and Angeluci (2021) draw attention to the need to analyze the possible errors arising from AI from an ethical perspective, since it is a machine dependent on the programming conceived by its creator, the human being. Thus, one cannot expect from AI an action that has not been previously planned. The

responsibility for its acts, therefore, does not fall upon the technology itself, but upon the person who programmed it, since only the human being possesses the capacity to discern between right and wrong and to make moral choices.

The second point presented by the authors refers to the possibility of AI being neutral, since it does not possess the capacity to act beyond what has been programmed. In this aspect, it is observed that there is no neutrality in technology, since its conception is carried out by a human person, bearer of political positions, worldviews, and even possible prejudices. Such characteristics of the creator may be incorporated into the technology and define the way in which the machine will act in the selection of information and in the responses offered to the user of this product..

In 2015, a user discovered that Google had a “*racist*” algorithm (Salas, 2018): Google Photos’ AI labeled all of his Black friends as gorillas, simply because it could not distinguish the difference between human skin and that of monkeys. The error was corrected, but it demonstrated how algorithms reflect the bias and prejudice of human beings that are present in the databases used by programmers, who also end up mirroring their own prejudices in the codes they develop (Rossetti; Angeluci, 2021, p. 10, emphasis in original).

The example presented in the previous passage proves particularly concerning, insofar as AI may convey data that harms certain social groups and, in the mentioned case, represents a threat to democracy and to societies founded on the equality of rights, due to its segregational character. Since it is a machine, the authors raise the hypothesis that such biases derive from its creator. In this sense, the importance is once again evident that scientists act ethically at the moment of project development and in the use of this technology, so as to avoid the production of prejudiced AI systems, which would be harmful to society.

Thus, it is not the machine that can be held responsible for such practices, but the person who created it, who, being imbued with prejudices and possessing an agent intellect, is capable of discerning between good and evil and, consequently, of understanding the social implications of their actions.

The third point highlighted by Rossetti and Angeluci (2021) regarding ethics in the use of AI refers, centrally, to the question of autonomy and the possible loss of freedom of the human person, constituting, in this sense, the most critical aspect of the indiscriminate use of AI.

The question raised by this reflection is: which ethical principles are violated as a result of the coercion exerted by the invisible force of algorithms in subjects’ decision-making? In this sense, it is a matter of pointing out, for example, the ethical implications of Facebook’s algorithms in the decision-making of its users. The hypothesis is that the ethical principle of freedom is infringed. Algorithms increasingly mediate digital life and decision-making. The aim here is to clarify the ethical importance of this mediation. Making a decision requires an autonomous and free subject capable of choosing their action. However, in the information society, mediatized and computerized, many decisions are made by algorithms (Rossetti; Angeluci, 2021, p. 12).

The reflections presented at this point draw attention to a grave and contemporary issue, since, increasingly, decisions come to be made based on algorithmic guidelines, and not from free thinking. Data from CNN Brasil reports that, in 2025, “More than half (52%) of Brazilians used ChatGPT

or other AI assistants to aid in purchases over the last 12 months, an increase of 52% compared to the previous year. Among these consumers, 18% resorted to the technology for the first time in the last year” (Ramos, 2025). Such a scenario proves concerning, as people come to be guided by machines that do not possess intellectual capacity, even though their creators do.

This observation leads to the reflection that individuals would be, progressively, abdicating free will and deliberately surrendering their destiny to automated systems that, in the last instance, reflect the intentions of their creators. By living, making choices, and consuming through technologies arising from the so-called fourth technological revolution, people end up transferring their freedom and their power of choice to a restricted number of human agents, capable of manipulating information, sustaining segregational practices, and altering the meaning of political discourses through an opaque layer of technological mediation. In this process, the owners and creators of these technologies, such as programmers, designers, manufacturers, designers, and users, remain, often, hidden under the logic of algorithms. Thus, greater transparency and accountability on the part of these actors becomes necessary so that ethics may be effectively practiced in view of the common good.

The concern is not restricted to the reflections developed in this study. Recent international literature has also recognized that, although AI presents potential to support educational processes, its benefits depend on the preservation of human mediation, transparency, accountability, and equity. Rodríguez Tenorio, Llorente Cejudo, and Cabero Almenara (2026) warn, moreover, that excessive delegation of functions to intelligent systems may favor the dehumanization of the educational process, reinforcing the need for ethical principles to guide its use.

In light of these reflections, the need to deepen the ethical discussion regarding the use of AI is reinforced. It is from this perspective that a debate developed by Thomas Aquinas in the thirteenth century is revisited in the sequence, whose reflections on intellect, freedom, and human responsibility offer important contributions for understanding the contemporary challenges arising from the use of these technologies.

FINAL CONSIDERATIONS

By way of conclusion, the initial affirmation is reiterated that the present work does not constitute a crusade against the so-called fourth technological revolution, nor, consequently, against the use of AI in education and in people’s lives in the twenty-first century. It is recognized, on the contrary, that such technologies contribute positively to various everyday activities. However, it is also reaffirmed that one cannot accept or sustain the idea that AI is capable of surpassing or replacing the human intellectual condition.

Throughout the text, we sought to develop this thesis from four articulated axes. In the first place, we discussed the way in which Thomas Aquinas understood and evidenced the importance of the use of the intellect for the person to constitute themselves as a human being. In the second place, it was evidenced that AI is inserted in a context of capitalist production, in which knowledge, data, and the technological processes themselves assume the form of commodities, being traversed by economic interests, concentration of power, and structural opacity. In a third moment, approaches were presented that defend the positive use of AI in fields such as Education and Law, allowing us to understand the

arguments favorable to its adoption and to situate them in the contemporary debate. Finally, the ethical problems related to the use of AI in schools were discussed, especially those associated with data fallibility, the nonexistence of algorithmic neutrality, and the risks that the indiscriminate use of these technologies may pose to human autonomy, freedom, and responsibility.

Thus, what we sought to evidence is that the intellectual condition of the human being grounds their nature as a person and that, whenever they exercise their intelligence in a free and deliberate manner, this nature manifests itself in the defense of the freedom of thinking and acting. Any action, whether practiced by people or mediated by machines, that operates in the opposite direction, that is, that seeks to conduct, replace, or suppress this intellectual condition, may prove profoundly harmful to human formation and to life in society.

It is precisely for this reason that the mediation of ethical principles in the use of AI becomes indispensable, especially in educational processes. More than discussing efficiency or technological innovation, it is a matter of placing at the center of the debate the human person, their capacity to think, to choose, and to act freely, as a fundamental criterion for evaluating the limits and possibilities of AI use in the contemporary world.

In this sense, the reflections of Thomas Aquinas prove particularly fruitful for the contemporary debate on AI, because they shift the discussion from the strictly technological plane to the anthropological and educational plane. Although current approaches emphasize issues related to regulation, algorithmic transparency, biases, and the ethical implications of the use of these technologies, the Thomistic perspective evidences that the deeper question lies in the preservation of the intellectual condition of the human person. Under this focus, the central problem consists not only in determining how to use AI ethically, but in understanding to what extent its growing incorporation into educational processes may favor or limit the development of the human capacity to know, deliberate, judge, and act freely. It is precisely this centrality of the intellect as the foundation of human formation that confers timeliness to the thought of Thomas Aquinas and expands his contribution to the educational discussions of the twenty-first century.

From this perspective, it is concluded that the challenge posed by AI is not restricted to the definition of norms for its use or to the improvement of its systems, but requires preserving education as a privileged space for the development of the human intellect. Technologies may constitute important instruments of support for educational activities, provided they remain subordinate to the conscious action of the person and do not replace their capacity to think, interpret reality, and deliberate upon it. Thus, the main contribution of Thomas Aquinas's reflections to the contemporary debate consists in reaffirming that the human intellect remains the foundation of freedom, responsibility, and the formation of the person, regardless of the technological advances achieved.

REFERENCES

AÇÃO EDUCATIVA; INSTITUTO PAULO MONTENEGRO. *Indicador de Alfabetismo Funcional – Inaf 2024: legado e futuro do alfabetismo funcional*. São Paulo, 2024. Disponível em: <https://www.alfabetismofuncional.org.br>. Acesso em: 15 jan. 2026.

- ACETI, Lanfranco. The necessity of redundancy and the censorship of the AI in the milieu of prostitutional aesthetics. In: LARA-HERNANDEZ, J. A. *et al. The dusk of design: exploring multidisciplinary approaches and evolutionary biology in architecture*. Cham: Springer Nature Switzerland, 2025. p. 259-282.
- ARENDT, Hannah. *Entre o passado e o futuro*. 9. ed. São Paulo: Perspectiva, 2022. *E-book*. Kindle.
- ARENDT, Hannah. O homem: um animal social e a condição humana. In: ARENDT, Hannah. *A condição humana*. Rio de Janeiro: Forense Universitária, 2020. p. 27-234.
- ARISTÓTELES. *Ética a Nicômaco*. Bauru: Edipro, 2009.
- BARBOSA, Renata Peres; ALVES, Natália. A reforma do ensino médio e a plataformização da educação: expansão da privatização e padronização dos processos pedagógicos. *E-Curriculum*, São Paulo, v. 21, 2023. DOI: <http://dx.doi.org/10.23925/1809-3876.2023v21e61619>.
- BULATHWELA, Sahan *et al.* Artificial intelligence alone will not democratise education: on educational inequality, techno-solutionism and inclusive tools. *Sustainability*, [S. l.], v. 16, n. 2, p. 781, 2024. DOI: <https://doi.org/10.3390/su16020781>.
- CASPER, Stephen *et al.* Black-box access is insufficient for rigorous AI audits. In: ACM CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY (FAccT '24), 2024, Rio de Janeiro. *Proceedings...* New York: Association for Computing Machinery, 2024. p. 2254-2272. DOI: <https://doi.org/10.1145/3630106.3659037>.
- CHAUDHRY, Iffat Sabir *et al.* Time to revisit existing student's performance evaluation approach in higher education sector in a new era of ChatGPT: a case study. *Cogent Education*, [S. l.], v. 10, n. 1, p. 2210461, 2023. DOI: <https://doi.org/10.1080/2331186X.2023.2210461>.
- CHEN, Sophia. The lost data: how AI systems censor LGBTQ+ content in the name of safety. *Nature Computational Science*, New York, v. 4, n. 9, p. 629-632, 2024. DOI: <https://doi.org/10.1038/s43588-024-00695-4>.
- COLÓN-VARGAS, Nelson. Exploiting the margin: how capitalism fuels AI at the expense of minoritized groups. *AI and Ethics*, Cham, v. 5, p. 1871-1876, 2025. DOI: <https://doi.org/10.1007/s43681-024-00502-w>.
- ESSIEN, Aniekan *et al.* The influence of AI text generators on critical thinking skills in UK business schools. *Studies in Higher Education*, Abingdon, v. 49, n. 5, 2024, p. 865-882. DOI: <https://doi.org/10.1080/03075079.2024.2316881>.
- FANTINUOLI, Claudio. Towards AI-enhanced computer-assisted interpreting. In: PASTORA, G. C.; DEFRANCQ, B. (orgs.). *Interpreting technologies: current and future trends*. Amsterdam: John Benjamins Publishing Company, 2023. p. 46-71.
- GEIGER, Daniel; BARR, Alistair; THOMAS, Ellen. Em decorrência do boom da IA, centros de dados estão surgindo por toda parte, gerando fortunas, consumindo energia e transformando a América rural. *Business Insider*, São Paulo, 2023. Disponível em: <https://www.businessinsider.com/ai-data-energy-centers-water-energy-land-2023-10>. Acesso em: 18 jan. 2026.

GOMES, Eduardo Biacchi; VAZ, Andréa Arruda; DIAS, Sandra Mara de Oliveira. Limites éticos para o uso da inteligência artificial no sistema de justiça brasileiro, de acordo com a Lei 13.709 de 2018 (LGPD) e Resoluções 331 e 332 do Conselho Nacional de Justiça. *Revista Internacional Consinter de Direito*, ano VII, n. 13, p. 107-124, 2021. DOI: <https://doi.org/10.19135/revista.consinter.00013.04>.

GUPTA, Gagan. Understanding the T5 model: a comprehensive guide. *Medium*, 16 jul. 2024. Disponível em: https://medium.com/@gagangupta_82781/understanding-the-t5-model-a-comprehensive-guide-b4d5c02c234b. Acesso em: 15 jan. 2026.

HOLMES, Wayne *et al.* Ethics of AI in education: towards a community-wide framework. *International Journal of Artificial Intelligence in Education*, v. 32, n. 3, p. 504-526, 2022. DOI: <https://doi.org/10.1007/s40593-021-00239-1>.

HOLMES, Wayne; TUOMI, Ilkka. State of the art and practice in AI in education. *European Journal of Education*, Hoboken, v. 57, n. 4, p. 542-570, 2022. DOI: <https://doi.org/10.1111/ejed.12533>.

LARSON, Barbara *et al.* Critical thinking in the age of generative AI. *Academy of Management Learning & Education*, Briarcliff Manor, v. 23, n. 3, 2024, p. 373-378. DOI: <https://doi.org/10.5465/amle.2024.0338>.

LDNOOBW. *List of dirty naughty obscene and otherwise bad words*. [2012]. Disponível em: <https://github.com/LDNOOBW/List-of-Dirty-Naughty-Obscene-and-Otherwise-Bad-Words/blob/master/en>. Acesso em: 15 jan. 2026.

KANT, Immanuel. *Sobre a pedagogia*. 2. ed. São Paulo: Almedina, 2018.

LA BOÉTIE, Étienne. *Discours de la servitude volontaire*. Paris: Éditions Hatier, 2025.

LE GOFF, Jacques. *Uma longa Idade Média*. Rio de Janeiro: Civilização Brasileira, 2008.

LUCKIN, Rosemary. *Machine learning and human intelligence: the future of education for the 21st century*. London: UCL IOE Press, 2018.

MARTINS, Anna. DeepSeek se autocensura quando perguntado sobre alguns temas políticos da China. *G1*, Rio de Janeiro, 28 jan. 2025. Disponível em: <https://g1.globo.com/tecnologia/noticia/2025/01/28/deepseek-censura-em-alguns-temas-politicos-sobre-a-china.ghtml>. Acesso em: 15 jan. 2026.

MEDNICK, Sam; BURKE, Garance; BIESECKER, Michael. Como gigantes da tecnologia dos EUA forneceram modelos de IA a Israel, levantando questões sobre o papel da tecnologia na guerra. *AP News*, 2025. Disponível em: <https://apnews.com/article/israel-palestinians-ai-weapons-430f6f15aab420806163558732726ad9>. Acesso em: 15 jan. 2026.

OGUNLEYE, Bayode *et al.* A systematic review of generative AI for teaching and learning practice. *Education Sciences*, Basel, v. 14, n. 6, 2024, p. 636. DOI: <https://doi.org/10.3390/educsci14060636>.

OUYANG, Fan; ZHENG, Luyi. AI-driven learning analytics applications and tools in computer-supported collaborative learning: a systematic review. *Educational Research Review*, Amsterdam, v. 44, p. 100616, 2024. DOI: <https://doi.org/10.1016/j.edurev.2024.100616>.

OUYANG, Fan; ZHENG, Luyi; JIAO, Pengcheng. Artificial intelligence in online higher education: a systematic review of empirical research from 2011 to 2020. *Education and Information Technologies*, Dordrecht, v. 27, n. 6, 2022, p. 7893-7925. DOI: <https://doi.org/10.1007/s10639-022-10925-9>.

RAMOS, Marien. Mais da metade dos brasileiros usam IA para fazer compras, diz relatório. *CNN Brasil*, São Paulo, 1 jun. 2025. Disponível em: <https://www.cnnbrasil.com.br/economia/financas/mais-da-metade-dos-brasileiros-usam-ia-para-fazer-compras-diz-relatorio/>. Acesso em: 21 jul. 2026.

RODRÍGUEZ TENORIO, Elvira; LLORENTE CEJUDO, María del Carmen; CABERO ALMENARA, Julio. Retos éticos del uso de la inteligencia artificial en educación: revisión sistemática (2020-2025). *EDMETIC: Revista de Educación Mediática y TIC*, Córdoba, v. 15, n. 2, 2026. DOI: <https://doi.org/10.21071/edmetic.v15i2.18825>.

ROSSETTI, Regina; ANGELUCI, Alan. Ética algorítmica: questões e desafios éticos do avanço tecnológico da sociedade da informação. *Galáxia* (São Paulo), n. 46, p. 1-18, 2021. DOI: <http://dx.doi.org/10.1590/1982-2553202150301>.

SANDOVAL-MARTIN, Teresa; MARTÍNEZ-SANZO, Ester. Perpetuation of gender bias in visual representation of professions in the generative AI tools DALL·E and Bing Image Creator. *Social Sciences*, Basel, v. 13, n. 5, 2024, p. 250. DOI: <https://doi.org/10.3390/socsci13050250>.

SANT'ANA, Irani Parolin.; SANT'ANA, Claudinei Camargo. Uma utilização do ChatGPT no ensino. *Com a Palavra, o Professor*, Vitória da Conquista, v. 8, n. 20, p. 74-86, 2023. DOI: <https://doi.org/10.23864/cpp.v8i20.951>.

SANTOS, Roger; PROFETA, Guilherme Augusto; PROFETA, Rogério Augusto. Por uma (não)reinvenção da educação: a inteligência artificial e o deslocamento do papel tradicionalmente atribuído ao professor. *Inter-Ação*, Goiânia, v. 48, n. 3, 2023, p. 640-657. DOI: <https://doi.org/10.5216/ia.v48i3.75784>.

SERRES, Chris.; KOWALCZYK, Liz.; HAGEN, Neena. “Um caso claro de censura”: os ataques de Trump à Diversidade, Equidade e Inclusão (DEI) abalaram a pesquisa científica. *Boston Globe Media Partners*, Boston, 2025. Disponível em: <https://www.bostonglobe.com/2025/02/10/metro/trump-executive-orders-new-england-boston/>. Acesso em: 15 jan. 2026.

SILVA, Marcos Fabrício Lopes. Sobre a inteligência. *Boletim da UFMG*, Belo Horizonte, ano 39, n. 1810, 2013.

TOMÁS DE AQUINO. O fim último do homem (Ia-IIae, q.1). In: TOMÁS DE AQUINO. *Suma Teológica*. São Paulo: Loyola, 2004. v. 2.

TOMÁS DE AQUINO. *A unidade do intelecto contra os averroístas*. Lisboa: Edições 70, 1999.

UNESCO. *Guidance for generative AI in education and research*. Paris, 2023. Disponível em: <https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research>. Acesso em: 21 jul. 2026.

UNGLESS, Eddie; ROSS, Bjorn; LAUSCHER, Anne. Stereotypes and smut: the (mis)representation of non-cisgender identities by text-to-image models. *arXiv:2305.17072*, 2023. DOI: <https://doi.org/10.48550/arXiv.2305.17072>.

YANG, Yiran. Racial bias in AI-generated images. *AI & Society*, London, v. 40, p. 5425-5437, 2025. DOI: <https://doi.org/10.1007/s00146-025-02282-1>.

ZAWACKI-RICHTER, Olaf. *et al.* Systematic review of research on artificial intelligence applications in higher education— where are the educators? *International Journal of Educational Technology in Higher Education*, Cham, v. 16, n. 39, 2019. DOI: <https://doi.org/10.1186/s41239-019-0171-0>.

ZEB, Shah *et al.* AI in healthcare: revolutionizing diagnosis and therapy. *International Journal of Multidisciplinary Sciences and Arts*, Medan, v. 3, n. 3, 2024, p. 118-128. DOI: <https://doi.org/10.47709/ijmdsa.v3i3.4546>

ZUIN, Antônio A. S. Inteligência artificial e formação danificada: aprendizagem profunda e ética rasa entre professores e alunos. *Educar em Revista*, Curitiba, v. 37, e80158, 2021. DOI: <http://dx.doi.org/10.1590/0104-4060.80158>.

Submetido: 01/24/2026

Preprint: 06/03/2026

Aprovado: 07/29/2026

Editor de seção: Suzana dos Santos Gomes

DECLARATION ON DATA AVAILABILITY

All data underlying the research are included within the manuscript.

AUTHORSHIP DECLARATION

Author 1 – Research coordination; theoretical conceptualization; active participation in data analysis.

Author 2 – Text revision and standardization; data analysis.

Author 3 – Manuscript review; data analysis.

DECLARATION OF CONFLICT OF INTEREST

The authors declare that there is no conflict of interest regarding this article.