Obutchénie
Revista de Didática e Psicologia Pedagógica

DOI: http://doi.org/10.14393/OBv9.e2025-21

# Artificial Intelligence in a Historical-Cultural perspective: a theoretical essay[1]

## Inteligência Artificial e a perspectiva Histórico-Cultural: um ensaio teórico

*Mariana Matulovic da Silva Rodrigueiro*[2]
*Franciele Santos Teixeira*[3]
*Maria Teresa Zampieri*[4]
*Sueli Liberatti Javaroni*[5]

**ABSTRACT**
This text addresses perceptions of Artificial Intelligence, providing a brief history of its emergence and implications, articulated with the premises of historical-cultural theory. We argue for the necessity of a philosophical and epistemological discussion regarding the use and potential directions of this object, including its role in the interaction process. Throughout the text, we discuss concepts such as sign and meaning and ways in which this interpretation can be applied to Artificial Intelligence.

**Keywords:** Artificial Intelligence. Digital Technologies. Mathematics Education.

**RESUMO**
Neste texto abordamos algumas percepções sobre a Inteligência Artificial (IA), trazendo um breve histórico de seu surgimento e implicações, bem como uma possível articulação com pressupostos da Teoria Histórico-Cultural. Defendemos que é necessário fazer uma discussão filosófica e epistemológica da IA, apontando possíveis caminhos para seu uso, inclusive, no processo de interação entre suas áreas de conhecimento, a Teoria Histórico-Cultural e a relação da apropriação cultural. Ao longo do texto abordamos conceitos como signo e significado e maneiras que essa leitura pode ser feita com a Inteligência Artificial.

**Palavras-chave**: Inteligência Artificial. Tecnologias Digitais. Educação Matemática.

[1] English version by Ivanir Liberatti Almeida Prado. E-mail: ivanirlaprado@gmail.com.
[2] PhD in Philosophy. Graduate Program in Mathematics Education, Brazil. Orcid: https://orcid.org/0000-0001-6626-4621. E-mail: mariana.matulovic@unesp.br.
[3] Master in Mathematics Education. Graduate Program in Mathematics Education, Brazil. Orcid: https://orcid.org/0000-0002-7244-5506. E-mail: fs.teixeira@unesp.br.
[4] PhD in Mathematics Education. Graduate Program in Mathematics Education, Brazil. Orcid: https://orcid.org/0000-0002-6656-2538. E-mail: mt.zampieri@unesp.br.
[5] PhD in Mathematics Education. Graduate Program in Mathematics Education, Brazil. Orcid: https://orcid.org/0000-0002-1948-4346. E-mail: sueli.javaroni@unesp.br.

DOI: http://doi.org/10.14393/OBv9.e2025-21

## 1 Introduction

Tikhomirov (1981, p. 264, our translation) advocated a more widespread view of Artificial Intelligence (AI), emphasizing that: "Artificial intelligence is the science whose objective is to develop methods that allow machines to solve problems that would require intelligence if solved by humans". For this author, this discussion could be expanded to aspects that address the supplementation or replacement of human capacity in the development of activities, highlighting a critique of possible reductionist views on human activity and computerization[6].

Considering discussing this vision, which at first glance does not encompass, in particular, all the potentials of AI, we aim in this text possible approaches to the conception of AI through the historical-cultural perspective and for this, we start from the question: "Is AI and Historical-Cultural Theory possible?". In this sense, we will discuss aspects related to learning and human development and possible relations between humans and machines, that is, we seek to articulate authors from the historical-cultural perspective, mainly in the line of Vygotsky, with discussions about AI.

Although AI does not have "needs"[7] or "motives" in the human sense, since it does not act to satisfy its own biological or social need[8], as it performs tasks based on objectives programmed by humans, we can think of its activity as a set of automated operations and actions aimed at achieving a pre-defined end by its creators. However, the way AI is developed and used can generate phenomena that Leontiev and other thinkers who share the Historical-Cultural Theory help us to highlight.

In view of what has been presented so far, in the following section, we discuss briefly how AI has been constituted in seeking to weave considerations in this regard and the Historical-Cultural Theory.

[6] We retained the term "informatization", present in Tikhomirov's original text (1981), to preserve the historicity of the concept, even though other terms like "digitalization" are more common today
[7] The quotation marks used throughout the text, which appear without reference, serve only to highlight certain terms and do not refer to quotations.
[8]AI, in almost all its applications, acts and operates due to programming, thus reflecting the programmer's needs.

DOI: http://doi.org/10.14393/OBv9.e2025-21

## 2 Theoretical Framework

### 2.1 A brief presentation on Artificial Intelligence

If we had to answer the question: "What is AI?" we would have a complex task without a unique answer, such that it was a well-defined answer, as we need theoretical support from various fields of knowledge, such as Human Sciences, mainly Philosophy and Psychology, Biological Sciences with Genetics and Biology, and Exact Sciences, with Logic, Mathematics, and Computing (COPPIN, 2004).

But before delving into this universe of AI, we need to understand the first attempts to formalize the concepts of intelligence and artificial since the works of Alan Turing (in the 1940s and 1950s) until 1956, when a group of researchers — including John McCarthy, Marvin Minsky, Nathaniel Rochester, and Claude Shannon — used the term for the first time at the Dartmouth Summer Research Project on Artificial Intelligence workshop (RUSSELL; NORVIG, 2020).

In 1950, Turing writes: "Can machines think?" In the quest to try to understand the depth of this concern and the complexity of the meaning of the words "machine" and "think", he proposes to replace the original question with another more concrete and testable one, incorporated into the "Imitation Game", later known as the Turing Test (TURING, 1950): "Can machines perform tasks at such a high level that we cannot discern whether it was executed by a machine or by a human?".

To illustrate, he describes a game with three participants: an interrogator (C), who can be of any gender, and two competitors, a man (A) and a woman (B). All players involved remain in isolated rooms, and communication is strictly textual, attempting to eliminate any vocal, visual, or tactile clues. The primary objective of the interrogator is to discern which of the two competitors is the man and which is the woman, identifying them only by the labels X and Y. The man's (A) task, however, is actively to deceive the interrogator, trying to make him

DOI: http://doi.org/10.14393/OBv9.e2025-21

believe that he is the woman. On the other hand, the woman (B) seeks to assist the interrogator in the correct identification, generally providing truthful answers (TURING, 1950).

Turing's turnround lies in his proposition to modify this game. He raises the fundamental question: What will happen when a machine assumes the role of the man (A) in this game? If the machine, by trying to pass as a human – and here, gender distinction is not the focus, but rather the indistinguishability of human behavior, succeeds in deceiving the interrogator as frequently as the man would, leading the interrogator to an incorrect identification, then, for the purposes of the Turing Test, this machine would demonstrate behavior indistinguishable from human behavior (TURING, 1950; FEIGENBAUM, 2003).

Turing was concerned with what machines could do, instead of what intelligence is. For him, intelligence could be measured by the ability of a machine to perform tasks that we normally associate with human intelligence, such as reasoning, learning, and problem-solving (GUGERTY, 2006). Thus his approach emphasized practical results instead of theoretical definitions. Turing adopts the term thinking machines (TURING, p. 436, 1950) to designate thinking machines, which is a prelude to the use of Artificial Intelligence, which emerged in 1956.

According to Teixeira (1998, p. 33):

> In essence, the Turing test establishes the following criterion for deciding whether a machine thinks: if a machine's behavior is indistinguishable from that exhibited by a human being, there is no reason not to attribute to this machine the ability to think.

It is important to point out that in 1936, Turing was restless in seeking a rigorous and formal definition of what was understood by an effective procedure or mechanical method. Hilbert, by specifying the Entscheidungsproblem (Decision Problem) had postulated, in essence, whether there was an algorithm — a finite series of well-defined steps that could be executed mechanically — capable of

DOI: http://doi.org/10.14393/OBv9.e2025-21

determining, for any given mathematical proposition, whether it was true or false, that is, if it was demonstrable from a set of axioms. To answer these questions and fill this conceptual gap, Turing conceived a theoretical machine that would become the central concept of the entire theory of Computation: the Turing Machine (TEIXEIRA, 1998).

In 1955, Newell, Simon, and Shaw developed the first program, the Logic Theorist, which successfully proved theorems from Whitehead and Russell's Principia Mathematica in a manner very similar to the human algorithmic process, conceiving "thought as processing (i.e., transformation) of symbols in short and long-term memories. These symbols were abstract and amodal, that is, not connected to sensory information" (GUGERTY, 2006, p. 881, our translation, author's emphasis). Thus, from this perspective, thought can be seen as the manipulation of symbols governed by formal rules, in the bias of classical logic, within memory structures, reducing it to simple algorithmic operations.

In 1956, John McCarthy, Marvin Minsky, Nathaniel Rochester, and Claude Shannon organized the Dartmouth Summer Research Project on Artificial Intelligence (MACCARTY et al., 2006, p. 02, our translation) with the aim of strengthening Artificial Intelligence as a research area. In the words of the idealizers themselves:

> We propose that a two-month study, with ten men, on artificial intelligence to be conducted during the summer of 1956 at Dartmouth College, in Hanover, New Hampshire. The study should proceed based on the conjecture that every aspect of learning or any other characteristic of intelligence can, in principle, be described with such precision that a machine can be made to simulate it.

McCarthy and his collaborators believed that some aspects of human intelligence, mainly those related to symbolic processing, could be described with enough precision to allow the construction of machines capable of simulating such capacities. They maintained that by working interdisciplinarily, researchers from

DOI: http://doi.org/10.14393/OBv9.e2025-21

different fields of knowledge could understand the mechanisms by which machines would use language, from abstractions and concepts, to solve problems traditionally attributed to humans. Furthermore, they believed that these machines would have the ability to progressively improve over time (NORVIG; RUSSELL, 2013; MCCARTHY et al., 2006).

Initially, the machine must be able to use language, that is, it would be able to understand and generate natural language. Natural language is a system of signs culturally and socially structured, used by humans to express thoughts, feelings, desires, instructions, or narratives, with complex and ambiguous rules that vary according to context, culture, and communicative intention. Then, the ability to form abstractions and concepts is highlighted, allowing the system to create symbolic representations that capture the essence of objects and ideas, enabling complex reasoning. Moreover, the machine must be able to solve problems traditionally attributed to human intelligence, applying strategies to deal with challenging situations in various domains (MACCARTY et al., 2006).

Another crucial aspect is the ability to learn from experience, allowing the system to improve its performance from previous data and interactions, incorporating learning mechanisms. Logical reasoning, in turn, is fundamental for the machine to make inferences and make informed decisions. The proposal also contemplates the ability to perceive the environment, interpreting sensory stimuli such as images and sounds, enhancing the machine's interaction with the mundane world. Finally, the machine must be able to control physical devices, executing autonomous actions and interacting directly with its surroundings (MACCARTY et al., 2006).

From 1956 onwards, we have an intense development of this new knowledge area, AI, with all its complexity, intensity, and interdisciplinarity of concepts. For this reason, giving a precise characterization of what AI is or what is studied in AI is difficult and requires care.

Russell and Norvig (2004) propose a taxonomy that organizes AI definitions according to two fundamental dimensions: the type of criterion adopted to evaluate

DOI: http://doi.org/10.14393/OBv9.e2025-21

intelligence (thought or behavior) and the reference model used (similarity to the human being or ideal rationality). The combination of these dimensions results in four major approaches for the study and development of AI: thinking like humans, acting like humans, thinking rationally, and acting rationally (RUSSELL; NORVIG, 2004).

In the work, Russell and Norvig (2004) characterize each of the four perspectives, with examples and developments that occurred in AI. Thus, in the category of thinking like a human, AI involves using computational models to replicate the internal cognitive processes of the human mind, drawing inspiration from psychology and neuroscience to simulate reasoning. In the third category, thinking rationally, the authors already bring to the title the rationality of the Laws of Thought of Aristotelian Syllogisms. Thus, in these systems, we have the prioritization of formal logic, aiming to build systems that deduce valid conclusions and operate according to the laws of logical thought.

In terms of behavior, we have the categories of acting like humans and acting rationally. Systems that act like humans focus on external behavior, evaluating AI by its ability to mimic human performance indistinguishably, with the Turing Test as its exponent. An interesting fact is that Russell and Norvig (2004), when formulating the test, Turing opens six major areas that compose AI: natural language processing to allow the machine to communicate clearly in a natural language; knowledge representation to store acquired and sent information; automated reasoning, which uses stored data to answer questions and infer new conclusions; machine learning, which allows adjusting to new situations and identifying, besides predicting, to manipulate items and recognize environments (RUSSELL; NORVIG, 2004).

Finally, the approach of systems that act rationally focuses on developing intelligent agents that maximize their objectives in a given environment, regardless of whether their internal processes resemble humans or strictly follow formal logic, thus encompassing perception, learning, and optimized action (RUSSELL; NORVIG, 2004).

DOI: http://doi.org/10.14393/OBv9.e2025-21

Since its official emergence in 1956, AI has gone through periods of heightened enthusiasm and accelerated progress (AI Boom), followed by phases of disillusionment and reduced investments, denoting these periods as AI winters (AI Bust). It is evident, therefore, that the trajectory of AI is marked by distinct phases of optimism, challenges, and paradigmatic advances (HAIGH, 2024; RUSSELL; NORVIG, 2004). Thus, we understand that the history of AI is a complex tapestry of advances and setbacks, shaped by cycles of optimism and “winters”[9].

Since its beginning (1952-1969), AI was driven by notable enthusiasm, with the development of programs such as the General Problem Solver (GPS) and the concept of the Advice Taker, solidifying the belief in the potential of machines to imitate human reasoning (RUSSELL; NORVIG, 2004). However, this initial phase faced a dose of reality (1966-1973), according to Russell and Norvig. Predictions by pioneers like Herbert Simon did not materialize, such as the one that in 10 years, a computer would be the world chess champion and would prove an important and unprecedented mathematical theorem — exposing severe limitations: the absence of domain knowledge in tasks such as automatic translation, the insurmountable combinatorial explosion in complex problems, and inherent limitations of learning models such as Perceptrons, culminating in funding cuts and questions about the field's viability (RUSSELL; NORVIG, 2004).

In response to these limitations, the period from 1969 to 1979 marked a strategic shift to knowledge-based systems. Programs like DENDRAL and MYCIN demonstrated the effectiveness of incorporating specialized domain knowledge and dealing with uncertainties, moving away from the “weak methods” of generic search. This approach also extended to Natural Language Processing, where robust understanding revealed the need for vast world knowledge. The 1980s were

[9] Since its inception in 1956, AI has gone through periods of great optimism, where society envisions and anticipates actions and developments that, at the time, are beyond AI's capabilities. These periods are known as AI "springs" or "booms." However, when conceptual problems are scientifically verified and evidenced, hindering all the optimistic projections of AI—such as the logical XOR problem in Perceptron modeling—AI enters a period of slow developments and even stagnation in research and resource allocation. These times are referred to as AI "winters" or "busts" (STRICKLAND, 2021).

DOI: http://doi.org/10.14393/OBv9.e2025-21

marked by a leap in the application of AI in industry, for example, the development of the expert system R1 that, from the quest to replicate a certain capacity of human reasoning, allowed devices with different configurations to be programmed with logical rules, making them compatible, attracting eminent global investments. However, the excess of promises inflated a “bubble” that burst, resulting in a prolonged period of little attention and investment in AI research, the famous “AI winters” (RUSSELL; NORVIG, 2004).

Despite this contraction, the field restructured itself. The return of neural networks in the second half of the 1980s, driven by the rediscovery of backpropagation, and the adoption of a rigorous scientific methodology were crucial. AI abandoned isolationism, integrating with disciplines like Statistics and Information Theory, and demonstrating robustness in applications such as speech recognition and automatic translation, using data-driven models and mathematical theories. The incorporation of Probability and Decision Theory, notably with Bayesian networks, formalized uncertain reasoning and solidified AI's transition into an interdisciplinary science with practical impact (RUSSELL; NORVIG, 2004).

From 1995 onwards, AI consolidated itself, focusing on the concept of intelligent agents and benefiting from the explosion of large data sets, mainly from the development and implementation of communication networks. This integration allowed reasoning and planning systems to deal with uncertainty and brought AI closer to other fields like economics and control theory, resulting in advances like autonomous vehicles and other IoT elements (Internet of Things). A paradigm shift occurred with the perception that massive amounts of data could, in many cases, outperform algorithmic complexity in problem-solving, circumventing the challenge of how to express all the knowledge a system needs and catalyzing the development of practical and robust applications (RUSSELL; NORVIG, 2004).

DOI: http://doi.org/10.14393/OBv9.e2025-21

In this evolutionary context, Generative Artificial Intelligence (GAI) emerges as an important and promising point within AI itself. GAI represents an effective response to AI's historical limitations, especially in overcoming the “combinatorial explosion” and excessive dependence on manually coded explicit knowledge. By inferring complex patterns from vast data implicitly, generative models produce coherent and contextualized content, a capability that overcame the failures of primary automatic translation and the difficulty of scaling expert systems[10].

By offering a natural interface and unprecedented creative capabilities, GAI not only enhances human-machine interaction but also revives the debate about Artificial General Intelligence (AGI), approaching the aspiration for thinking and creating machines. GAI drives the current AI “boom”, redefining or setting expectations and generating new ethical and social challenges regarding its use.

## 2.2 Generative Artificial Intelligence

The development of GAI represents one of the most significant technological achievements of recent decades, being a direct result of accumulated advances in various AI sub-areas. This AI strand is characterized by the ability to autonomously produce new content, such as texts, images, sounds, videos, and programming code, with quality levels that often resemble the results produced by humans (SOLANKI; KHUBLANI, 2024).

In the 1960s to the 2000s, in the AI field, the potential of systems like ELIZA[11], was envisioned, which although rudimentary, demonstrated the potential of human-machine interaction and natural language processing.

[10] The combinatorial explosion refers to the vast universe of possible states or situations to be analyzed when designing an expert system, such as a system for playing chess. To program a device capable of playing chess, one must analyze all possible move positions, with each game element in every possible position on the board. This results in an explosion of combinations and possibilities.

[11] *ELIZA* was one of the first natural language processing programs, developed by Joseph Weizenbaum in 1966. It simulated a conversation with a Rogerian psychotherapist through simple substitution rules and pattern recognition, without real understanding of the content, but capable of creating the illusion of meaningful dialogue with the user (Berry, 2023).

DOI: http://doi.org/10.14393/OBv9.e2025-21

Despite the early recognition of the neural networks' potential, their large-scale application was historically restricted by the computational limitations of the time. Only with the improvement in processing power and the increasing availability of data at the turn of the millennium (BIG DATA), deep neural networks (deep neural) could effectively handle massive volumes of information and model complex relationships between input and output variables (SOLANKI; KHUBLANI, 2024).

Machine Learning is an AI field whose goal focuses on developing algorithms capable of learning patterns and optimizing a performance criterion from example data or previous experience. Unlike traditional programming, machine learning allows computers to learn autonomously, without the need to be explicitly programmed for each specific task (ALPAYDIN, 2020). In this paradigm, data is often input into an algorithm in a structured and relevant format or undergoes a manual and intensive treatment process known as feature engineering. This means human experts need to identify and extract the most important characteristics of the data before the algorithm can learn from them (ALPAYDIN, 2020).

Deep Learning represents a significant advancement within Machine Learning. Characterized by the use of artificial neural networks with multiple hidden layers, it offers a significant capacity: the ability to autonomously extract and learn data representations. Instead of relying on human preprocessing to identify features like shapes in an image or pattern words in a text, Deep Learning models can infer these features on their own from large volumes of data (ALPAYDIN, 2020).

The true GAI revolution was catalyzed by the introduction of Generative Adversarial Networks (GANs)[12] by Ian Goodfellow in 2014. This model, based on a duel dynamic between neural networks — where two networks compete in the

[12] Generative Adversarial Networks (GANs), proposed by Ian Goodfellow in 2014, operate through an adversarial model between two neural networks: the generator, responsible for creating synthetic data, and the discriminator, tasked with distinguishing between real and generated data. The generator aims to deceive the discriminator, while the discriminator attempts to identify fakes. This competitive process allows the generator to learn to produce increasingly realistic data (Creswell et al., 2018).

DOI: http://doi.org/10.14393/OBv9.e2025-21

sense of the generator trying to deceive and the other seeking to find counterfeits — enabled a new era in image generation, allowing for the creation of increasingly realistic and complex content.

After the 2010s, a rapid expansion was consolidated, with the diversification of generative models, including convolutional (CNN) and recurrent (RNN) neural networks for text and video, as well as variational autoencoders and diffusion models. In this context, the development of Large Language Models (LLM) like GPT-1 marked a qualitative leap in text generation, demonstrating a high capacity for text synthesis and comprehension.

Up to this point, we have discussed the characterization of AI, and to give direction to this work's proposal, we bring elements of Historical-Cultural Theory that allow us to approximate AI concepts with elements of this theory.

### 2.3 Um diálogo com elementos fundamentais da Teoria Histórico-Cultural

We understand that the constitution of the human is not a mechanical and linear process but a dynamic process mediated by another human immersed in culture, because "[...] the subject's relationship with the world is only possible through the mediation of another subject" (WERNER, 2015, p. 34). We perceive then that the constitution of the subject is not a mechanical or linear process, but also a process of cultural immersion. Thus, we highlight that, as pointed out in Vygotsky's works, the central idea lies in the development of higher psychological functions, which is based on a sociocultural process, that is, mediated by culture and social interactions.

Interactions are mediated by cultural and psychological instruments, where the first consists of the thing, and the second is considered development with the other through language and speech.

> Typically, human thought is constituted by language because it is from the moment language comes into play, in the course of development, that thought becomes verbal and speech rational. The emergence of verbal thought does not happen mechanically. The internalization of language and the development of verbal thought

DOI: http://doi.org/10.14393/OBv9.e2025-21

> occur through a long process of changes, which, in turn, alter the way the subject operates reality (WERNER, 2015, p. 34).

Thus, we understand that cultural instruments, also called tools, are external objects that humans use to act upon the world and transform it. They are socially created and transmitted from generation to generation, incorporating humanity's accumulated experience and knowledge, according to Vygotsky (2001). For this author, the use of cultural instruments is not only a way to facilitate action but also to transform the human mind itself. By using them, the individual internalizes the logic and structure of this tool, which modifies their cognitive processes. For example, we can interpret that using a computer not only speeds up writing but also changes how we think and organize the content we write. An important aspect within this theory is that language is the main symbolic instrument for human thought, being a fundamental mediator when considering human development (VYGOTSKY, 2001).

In this sense, the sign is seen as a symbolic mediator in organizing behavior and thought (VYGOTSKY, 2001). Thus, we can envision three concepts related to the sign: it being a thing, which when linked to language, we can think of the material integrated into human activity; sign as something real and in the common consciousness of being; and interpretations about the structure and functions of the sign. We will consider the first interpretation to reflect on its role in human activity.

Meaning, without considering language, does not exist since, according to Vygotsky (2001), it is the central aspect of the linguistic sign, that is, meaning consists of connecting the word (the linguistic sign) to reality and thought. Thus, the meaning of the sign is not just a simple copy of real connections and relations, but as Leontiev (1989, p. 246) states:

> Objectively, the sign confronts the subject as a real sign with everything behind it, including those functional characteristics that come from the characteristics of the activity in which that sign is included. Subjectively, however, it is perceived as a psychological formation in which the social real content of that

Obutchénie
Revista de Didática e Psicologia Pedagógica
DOI: http://doi.org/10.14393/OBv9.e2025-21

> sign is incorporated and transformed. The subject's consciousness in this case remains as a contemplative consciousness; and, from his/her point of view, the sign emerges as an image of the sign, and the meaning emerges as an image in which he/she fixes and experiences his/her social personal experience without worrying about the problems of its real core and its real nature (LEONTIEV, 1989, p. 246).

Furthermore, these aspects are connected to the concepts of sign and meaning, so that we can investigate factors involved in the origin of meaning and how it functions in human social activity (LEONTIEV, 1989). Thus, as Werner (2015, p. 34) highlights:

> Typically, human thought is constituted by language because it is from the moment language comes into play, in the course of development, that thought becomes verbal and speech rational. The emergence of verbal thought does not happen mechanically. The internalization of language and the development of verbal thought occur through a long process of changes, which, in turn, alter the way the subject operates reality.

As human beings do not exist in isolation, it is through communication that they get developed because “knowledge is not in the subject nor in the object, but in the interaction between both” (MELLO; TEIXEIRA, 2012, [no page]). Therefore, we consider, corroborating Silva (2023), that the basis of knowledge requires interaction with the external environment so that the individual, having contact with the object of his reality, develops driven by his needs, motivations, and objectives (SILVA, 2023). Thus, we seek to address the subject's development by considering AI in this interaction process. To make this discussion, we rely on the concepts of pseudo-concreteness, fetishization, and reification.

Fetishization, or more specifically, the commodity fetishism, is the process by which complex social relations between people — in work, production, exchange — are disguised and appear as natural and intrinsic relations between things (commodities). The commodity, a product of human labor, seems to acquire a life

DOI: http://doi.org/10.14393/OBv9.e2025-21

of its own, magical value and power that are, in fact, attributes of the social interactions that produced it.

Both pseudo-concreteness and fetishization dialogue with the concept of alienation, as by having a superficial perception of reality and having complex social relations disguised, the subject alienates themselves in the sense that there is the creation of obstacles to the appropriation of culture that should be constituted as heritage by all beings. It is still possible to affirm that "concerning psychological processes, the rupture between the sense and meaning of human actions has as one of its consequences the restriction of the process of human personality development" (DUARTE, 2004, p. 59).

Thus, thinking about technological development, as brought about by AI, can transform the dynamics of the society we live in or even alter communication processes. Mello and Teixeira (2012, [no page]) highlight how technological development enables increasingly complex levels of interaction to happen and that still:

> It is possible to affirm that the advent of high-speed internet subjects interaction itself to a virtualization process, enhancing the communicational capacity of subjects and enabling a hypermedia process of interaction with others and the environment. In the same sense, recognizing the importance of interactive processes for human development and, especially for the learning process, it seems urgent to discuss the potential of these technologies that are symbiotically embedded in people's lives and have gained space in the educational environment.

We understand that discussing this technological development in the educational environment is complex and requires a multifaceted analysis, in addition to a specific direction, as we are doing in this article by focusing on AI.

By including information and communication technologies as mediators of thought, Tikhomirov (1981) maintains conceptual fidelity to Vygotsky's ideas, expanding this theoretical framework to the computerization scenario, highlighting that we should pay attention to how human mental activity is transformed through this "new" mediation, contrasting substitution theory and supplementation theory.

DOI: http://doi.org/10.14393/OBv9.e2025-21

By refuting the substitution theory, Tikhomirov (1981) argues that the computer does not replace human thought since human heuristics are more complex, flexible, and qualitative than those implemented in a computer. On the other hand, the computer (we can expand to DT) offers a complementary and qualitatively different cognitive mediation (Tikhomirov, 1981), not replicating human thought but reorganizing it.

To talk about supplementation theory, Tikhomirov (1981) initially contrasts information theory with the psychological theory of thought, saying that while the former reduces thought to mental processes involving data transmission, encoding, and storage, the latter, based on Vygotsky's concepts, understands thought as a symbolic, socially and historically constructed and mediated activity, which, as previously discussed, takes into account contexts, meanings, and subject transformations.

Therefore, by refuting these two theories, Tikhomirov (1981) argues that the computer, or Digital Technologies in general, should not be understood as a problem solver or as a memory increment but rather as a new agent of cultural mediation, reorganizing human thought. Thus, AI will be debated here in this article within this perspective.

## 3. Methodology

This article fits the methodological modality of a theoretical essay because

> Instead of the general objective, specific objectives, justification, theoretical foundation, methodology that defines data collection and analysis criteria, and conclusion, in the essay, the orientation is not given by the search for answers and true statements, but by questions that guide subjects to deeper reflections (MENEGHETTI, 2011, p. 321).

Thus, we do not propose an empirical investigation but a construction of conceptual approaches between different frameworks that harmonize with each other, seeking to discuss the question: “Is AI and Historical-Cultural Theory possible?”. The theoretical essay is characterized by critical elaboration

DOI: http://doi.org/10.14393/OBv9.e2025-21

and the proposition of new articulations so that there is greater freedom for authors to "[...] defend a particular position without having to rely on the rigorous apparatus of empirical or bibliographic documentation [...]" (SEVERINO, 2013, [no page]).

In a theoretical essay, "truth does not materialize through dogmatic affirmation resulting from the simple use of models or systems. Truth is reduced to means, not confusing truth with the principle of identity" (MENEGHETTI, 2011, p. 323), thus we do not seek to defend a single and absolute truth, but to foster discussion and reflection about AI elements and their use in the perspective of Historical-Cultural Theory, that is, we argue in such a way that "the argumentative act needs to be constructed throughout the process of involvement with the object. Arguments are the result of the interaction of subject and object, where logic and truths are means and not ends in themselves" (MENEGHETTI, 2011, p. 324).

Therefore, we seek, through reflections and interpretations, to tension AI theory, based on its logical-formal premises since Turing, with Historical-Cultural Theory, especially in the contributions of Vygotsky, Leontiev, Kosik, and Tikhomirov, based on concepts such as sign, meaning, pseudo-concreteness, fetishization, and thought reorganization, making approximations through the questioning: "Is AI and Historical-Cultural Theory possible?".

## 3. Artificial Intelligence and Historical-Cultural Theory

We must emphasize that in this article we make an articulation between ideas from distinct scientific fields, but that dialogue with each other, to weave a philosophical and epistemological discussion, pointing possible paths for the use of AI.

According to the concepts of fetishization and pseudo-concreteness proposed by Kosik (1965), caution is needed with the alienated use of AI, that is, in the sense of having a superficial perception of reality and the complex social relations embedded in it, creating obstacles to the appropriation of culture.

DOI: http://doi.org/10.14393/OBv9.e2025-21

To do this, it is necessary to go beyond the superficial, seeking depth in the roots of AI history, as discussed earlier. From a longitudinal study of this history, in the field of Logic and the foundations of Computer Science, we discuss its beginning, from the works of Alan Turing, who, in general lines, was concerned with what machines could do, instead of what intelligence is (GUGERTY, 2006).

Next, we brought the concern of transforming AI into a field of knowledge or research area, from the contributions of Maccarty et al., (1955), who believed that some aspects of human intelligence, mainly those related to symbolic processing, could be described with enough precision to allow the construction of machines capable of simulating such capacities. We emphasize the advancement of natural language processing to allow the machine to communicate clearly in a natural language (RUSSELL; NORVIG, 2004).

We also discussed the transition of AI to an interdisciplinary science with practical impact (RUSSELL; NORVIG, 2004), about AI "winter" and the recent advancement of generative AIs.

In the field of psychology, we dialogue with concepts from Historical-Cultural Theory, particularly with theorists like Vygotsky and his followers. We highlight Tikhomirov's argument (1981), which argues that the computer should not be understood as a problem solver (replacing the subject), nor as a memory increment, but as a new agent of cultural mediation, like reorganizing human thought.

Extending this reasoning to AI, we can understand it as a cultural instrument, performing a complementary and qualitatively different cognitive mediation of human activity, not in the sense of replicating thought but reorganizing it. And this is intrinsic to epistemology, particularly cultural appropriation. But from this, what would be the underlying questions when dealing with AI use in the most diverse fields of knowledge?

Reflecting on this question, philosophically and epistemologically, under the light of the references discussed, we understand that in AI we constantly deal with representations of the world. Language models (LLM), neural networks, and computer vision systems do not operate directly with reality itself, but with data

DOI: http://doi.org/10.14393/OBv9.e2025-21

that are representations of aspects of that reality. The meaning that an AI model assigns to a word or an image is not intrinsic, but emergent from the vast amounts of data and training with which it interacted.

The data that feeds AI models (texts, images, audios, videos) are, in essence, records of a systematic reality. When an LLM is trained on billions of pages of text, it is ingesting the connections and relations that humans have established in the world. For example, the frequency with which "coffee" appears near "wake up" or "work" reflects a systematicity of human activity – people drink coffee to wake up and work. Or still, the way "cat" and "meow" are correlated reflects a systematicity in the animal world. More than that, these data are the product of systematized human activity. The way we write, speak, label images, organize information – all this reflects the systematicity of our actions, mediated by cultural and psychological instruments (WERNER, 2015).

Just as Leontiev (1989) points out that the introduction of sign systems can generate distortion or misrepresentations in human reflection, in AI, we have the challenges of data biases, model interpretability, and ensuring that a system's internal representation corresponds in a useful and ethical way to the real world. The material objects (the model's internal representations) can contain and propagate these distortions if not well managed. With this, we can think that the meaning of a word does not come only from the object it represents but from the way we interact with that object and others through language, within a social context, that is, not being a simple reproduction of social connections and relationships (Leontiev, 1989).

Making an analogy with AI, a photo of a cat is a pixel matrix. Each pixel has a numerical value (materiality). However, what the model "reads" in this matrix are the converted relationships that form the "cat" pattern – the edges, textures, colors that together represent the concept of cat to the model. The materiality of the pixels (the raw numbers) is the base, but what matters is the representational value that emerges from the organization of these pixels, just as the value of money is not in the paper, but in the economic relations it represents, considering its logical-historical movement, in the sense defended by Leontiev (1989).

DOI: http://doi.org/10.14393/OBv9.e2025-21

According to the field of formal logic, AI models, particularly LLMs, are specialized in learning and replicating this systematicity. They do not receive explicit grammatical rules; they infer these rules (the "formal systematicity") from observing how humans use language in vast datasets. An LLM does not have a pre-programmed rule that “after ‘the’, usually comes a masculine noun”; instead, it sees this regularity billions of times in the data ("the car," "the book," “the dog”) and thus, this “rule” emerges in its architecture statistically. This formal regularity is a reflection of objective relations and human activities (“the car” as a transport object).

Transformers are the architecture that revolutionized LLMs, allowing them to process data sequences (like text) in a way that captures long-distance relationships between elements. Transformers are trained on a large volume of human-generated text – which is essentially a massive record of human social activity. Leontiev (1989) argues that meaning arises from the connections and relations of reality, shaped by social activity.

Thus, when processing this text, the model is not just “reading words”; it is inferring patterns of how these words connect, how concepts relate, how arguments are constructed, how emotions are expressed. It learns that “doctor” is related to “hospital”, “disease” to “cure”, and that these relations are part of human activity of caring for health. These are the social content connections that form the objective content of the sign (Leontiev, 1989).

Furthermore, Leontiev (1989) warned us that when introducing sign systems, we can have “distortions and misrepresentations”. This is mirrored in AI in challenges such as biases. If the training data of a GPT (OpenAI) contains social biases, the model will “convert” and replicate these biases in its responses because its formal systematicity reflects the (problematic) systematicity present in human activity data. This prevents a critical understanding of reality because what is presented as “truth” is an algorithmic construction.

The hallucination of LLMs can also be seen as a form of generating texts that appear to be grammatically correct but were actually generated just to give a response, regardless of its truthfulness or reality. The model produces an

DOI: http://doi.org/10.14393/OBv9.e2025-21

"expression" that seems meaningful but has no correspondence with objective reality because the "connections and relations" it internalized were distorted or are insufficient for the task (LEMOS, 2024).

Thus, we understand that AI can "objectify" human relations by automating complex interactions, transforming people into "data" or "resources", dehumanizing relations and reducing subjectivity to quantifiable metrics. Beyond that, with it, especially in its use and perceptions in everyday life, it can be a powerful producer or reinforcer of pseudo-concreteness (Kosik, 1965). It operates on a level of data and patterns that, if not critically understood, can lead us to superficial conclusions about the world.

Recommendation systems (movies, products, news) give us the impression that AI knows what we want or that the user experience is optimized for our benefit. This hides the praxis of algorithmic manipulation, where AI does not serve our autonomy but commercial interests, data collection, and the maintenance of consumption patterns (MARTIN, 2019). The apparent convenience prevents us from seeing the power relationship and behavior engineering behind it.

AI is trained on historical data that reflects the biases and ideologies of our society. By processing and reproducing these patterns, it projects these common and ideological representations back into the user universe, often amplifying them. A situation exemplifying this is when a facial recognition AI consistently fails with darker skin tones, or a language model associates certain professions with specific genders; these are "common representations" (social biases) being reproduced and naturalized by technology. The statistical "truth" that AI presents ("this is how the data shows") hides the historical process of discrimination and inequality that generated these data and representations. AI does not question these biases; it reproduces them and makes them more "evident", reinforcing ideological pseudo-concreteness, as defended by Kosik (1965).

AI is becoming increasingly ubiquitous, embedded in our phones, cars, homes, healthcare systems, education, etc. It becomes perceived as a "natural" and inevitable condition of modern existence, a "fixed object". This naturalization prevents us from

DOI: http://doi.org/10.14393/OBv9.e2025-21

recognizing AI as the result of intense social activities, political choices, massive investments, and power struggles. The existence of autonomous cars or AI diagnostics, the Internet of Things (IoT), seems like a “natural advancement”. But this pseudo-concreteness and fetishization (Kosik, 1965) ignore the decisions about who develops, who finances, who regulates, who benefits, and who is marginalized by these technologies, which are, in fact, results of complex human activities.

## 3 Final Considerations

In this text, we presented a theoretical essay in which we sought approximations and possibilities to discuss the use of AI from the perspective of Historical-Cultural Theory. Throughout the text, we aimed to deepen different conceptions and views of AI based on its historicity, seeking connections between its links. By presenting its development, we emphasized the importance of interpreting biases when using AI, as its database corresponds to the interpretations of those who create the algorithm. AI makes connections according to the database it receives, and for this reason, it cannot and should not be considered neutral or entirely responsible for creating meanings for the signs it interprets.

We also addressed the fact that believing AI possesses this neutrality or can develop human characteristics that arise from social interaction with others falls into the belief that we can be replaced.

Therefore, throughout the text, we start from the roots of AI, based on the field of Philosophy and Cognitive Science, and engage with sociological and epistemological implications to bring a conception of AI to support research in the most diverse areas of knowledge that have it as an object of study, based on Historical-Cultural theory. With this, we aim to offer subsidies that can dialogue with the epistemology of these fields, contributing to the deepening of discussions.

DOI: http://doi.org/10.14393/OBv9.e2025-21

# Inteligencia Artificial y la Perspectiva Histórico-Cultural: un Ensayo Teórico

**RESUMEN**
En este texto, abordamos algunas percepciones sobre la Inteligencia Artificial (IA), presentando una breve historia de su surgimiento e implicaciones, así como una posible conexión con los supuestos de la Teoría Histórico-Cultural. Argumentamos que es necesario abordar una discusión filosófica y epistemológica sobre la IA, destacando posibles vías para su uso, incluyendo la interacción entre sus campos de conocimiento, la Teoría Histórico-Cultural y la relación con la apropiación cultural. A lo largo del texto, abordamos conceptos como signo y significado, y las maneras en que esta interpretación puede lograrse mediante la Inteligencia Artificial.

**Palabras clave:** Inteligencia Artificial. Tecnologías Digitales. Educación Matemática.

## 4 Referências

ALPAYDIN, E. *Introduction to machine learning*. MIT press, 2020.

BERRY, D. M. *The limits of computation: Joseph Weizenbaum and the ELIZA chatbot*. *Weizenbaum Journal of the Digital Society*, v. 3, n. 3, 2023. Disponível em: https://doi.org/10.34669/WI.WJDS/3.3.2. Acesso em: 30 jun. 2025.

COPPIN, *B. Artificial intelligence illuminated*. Jones & Bartlett Learning, 2004.

CRESWELL, A. *et al. Generative adversarial networks: An overview*. IEEE Signal Processing Magazine, v. 35, n. 1, p. 53-65, 2018.

DUARTE, N. *Formação do indivíduo, consciência e alienação: o ser humano na psicologia de A. N. Leontiev*. Cadernos CEDES, São Paulo, v. 24, n. 62, p. 44–63, abr. 2004. DOI: https://doi.org/10.1590/S0101-32622004000100004.

FEIGENBAUM, E. A. *Some challenges and grand challenges for computational intelligence*. Journal of the ACM (JACM), v. 50, n. 1, p. 32-40, 2003.

GUGERTY, L. *Newell and Simon's logic theorist: Historical background and impact on cognitive modeling*. In: PROCEEDINGS OF THE HUMAN FACTORS AND ERGONOMICS SOCIETY ANNUAL MEETING. Los Angeles, CA: SAGE Publications, 2006. p. 880-884.

HAIGH, Thomas. *How the AI boom went bust*. Communications of the ACM, v. 67, n. 2, p. 22-26, 2024.

KOSIK, K.; NEVES, C.; TORÍBIO, A. Dialética do concreto. Milão: Bompiani, 1965.

LEMOS, A. L. M. Erros, falhas e perturbações digitais em alucinações das IA generativas: Tipologia, premissas e epistemologia da comunicação. MATRIZes, v. 18, n. 1, p. 75-91, 2024.

DOI: http://doi.org/10.14393/OBv9.e2025-21

LEONTIEV, A. A. *Sign and activity*. In: WERTSCH, J. V. (ed.). The concept of activity in Soviet psychology. Armonk, NY: M. E. Sharpe, 1981. p. 241-255.

MARTIN, K. *Ethical implications and accountability of algorithms*. Journal of business ethics, v. 160, n. 4, p. 835-850, 2019.

MCCARTHY, J.; MINSKY, M. L.; ROCHESTER, N.; SHANNON, C. E. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955. *AI Magazine*, v. 27, n. 4, p. 12, 2006.

MELLO, E. F. F.; TEIXEIRA, A. C. *A interação social descrita por Vigotski e a sua possível ligação com a aprendizagem colaborativa através das tecnologias de rede*. In: SEMINÁRIO DE PESQUISA EM EDUCAÇÃO DA REGIÃO SUL – ANPEd-Sul, 9., 2012, Caxias do Sul. Anais… Caxias do Sul: Universidade de Caxias do Sul, 2012. p. 1–15. Disponível em: https://www.ucs.br/etc/conferencias/index.php/anpedsul/9anpedsul/paper/viewFile/6/871. Acesso em: 20 ago. 2025.

MENEGHETTI, F. K. O que é um ensaio-teórico? *Revista de Administração Contemporânea*, Curitiba, v. 15, n. 2, p. 320–332, mar./abr. 2011. DOI: https://doi.org/10.1590/S1415-65552011000200010.

RUSSELL, S.; NORVIG, P. *Inteligência artificial: uma abordagem moderna*. 1. ed. São Paulo: Campus, 2004.

SEVERINO, A. J. *Metodologia do trabalho científico*. São Paulo: Cortez, 2013.

SILVA, E. C. *Desenvolvimento do pensamento computacional em uma dinâmica pedagógica baseada na perspectiva histórico-cultural*: possibilidades para a formação das ações mentais de estudantes e do conceito polígono regular a partir da produção de um pensamento geométrico. 351 p. Tese (Doutorado em Educação Matemática) – Instituto de Geociências e Ciências Exatas, Universidade Estadual Paulista "Júlio de Mesquita Filho", Rio Claro, 2023.

SOLANKI, S. R.; KHUBLANI, D. K. *Generative artificial intelligence: exploring the power and potential of generative AI*. Berkeley, CA: Apress, 2024.

STRICKLAND, E. *The Turbulent Past and Uncertain Future of AI*: Is there a way out of AI's boom-and-bust cycle?. IEEE Spectrum, v. 58, n. 10, p. 26-31, 2021.

TIKHOMIROV, O. K. The psychological consequences of the computerization. In J. Werstch (Ed.), The concept of activity in soviet psychology (pp. 256 - 278). New York, United States: Sharp, 1981.

TEIXEIRA, J. F. *Mentes e máquinas: uma introdução à ciência cognitiva*. Porto Alegre: Artes Médicas, 1998.

DOI: http://doi.org/10.14393/OBv9.e2025-21

THUMALA, S. R.; MADATHALA, H.; MANE, V. M. *Azure Versus AWS: A Deep Dive into Cloud Innovation and Strategy*. In: 2025 International Conference on Electronics and Renewable Systems (ICEARS). IEEE, 2025. p. 1047-1054.

TURING, A. *Computing machinery and intelligence. Mind,* v. 59, n. 236, p. 433-460, 1950.

VYGOTSKY, Lev Semenovich. Pensamento e Linguagem. Edição eletrônica Ridendo Castigat Mores. E- book. Rio de Janeiro: eBooks Brasil / Jahr.org, 2001.

WERNER, J. A relação linguagem, pensamento e ação na microgênese das funções psíquicas superiores. *Fractal: Revista de Psicologia*, v. 27, p. 33-38, 2015.

Received in June 2025
Approved in September 2025