TPM Vol. 32, No. S3, 2025
ISSN: 1972-6325
https://www.tpmap.org/

WHO DECIDES WHEN THE MACHINE DECIDES? AUTOMATION OF JUDGMENT AND INSTRUMENTAL REASON

VIDAL BASOALTO CAMPOS
DOCTOR OF EDUCATION WITH A MENTION IN CULTURE AND EDUCATION IN LATIN AMERICA;
ASSOCIATE UNIVERSIDAD METROPOLITANA DE CIENCIAS DE LA EDUCACIÓN
EMAIL: vidal.basoalto@umce.cl; OECID ID: HTTPS://ORCID.ORG/0000-0002-7623-6322

Abstract:
This study critically examines the instrumental rationality that underlies the contemporary development of artificial intelligence, articulating the concept of epistemic efficiency within the framework of an epistemology of efficiency. Through a qualitative methodology based on the content analysis of theoretical texts, key tensions are identified between the automation of reason and the ethical principles that should guide emerging technologies. The results show how the hegemony of technical efficiency displaces critical reflection and collective deliberation. The discussion delves into the limits and contradictions of this dominant rationality, while the conclusion proposes a deliberative ethics as a normative horizon and proposes future lines of research aimed at rethinking the epistemological and ethical foundations of technological development.

Keywords: instrumental rationality, epistemology of efficiency, Artificial intelligence deliberative ethics.

INTRODUCTION

EPISTEMIC EFFICIENCY: KNOWLEDGE REDUCED TO CALCULATION

This essay delves into a question of growing urgency in the contemporary intellectual landscape: the way in which artificial intelligence (AI), intrinsically linked to the principles of instrumental reason, is reconfiguring the very foundations of our epistemology. The research challenge lies in unraveling how this technological convergence is shaping the way we conceive, produce and validate knowledge, with the potential emergence of what we call an "epistemology of efficiency". The latter is characterized by an appreciation of knowledge in terms of its ability to optimize processes, displacing other essential dimensions of knowledge, such as depth of understanding, critical reflection and the consideration of ethical and social values.

In this framework, the research question that guides this analysis is the following: in what way is the logic of instrumental reason, embodied in artificial intelligence systems, configuring a new form of epistemology focused on efficiency, and what are its ethical and cognitive implications?

The need to address this problem is based on the growing pervasiveness of artificial intelligence and its ability to radically transform the processes of knowledge production and validation. Understanding the logic that underlies its design and operation requires the adoption of a critical and dialogical approach, which not only technically analyzes its operating principles, but also reveals the normative and epistemological assumptions that sustain it. This approach makes it possible to problematize the way in which AI contributes to shaping new forms of knowledge, guided by criteria of efficiency, prediction and control, to the detriment of other fundamental dimensions such as critical reflection, interpretative complexity and ethical consideration.
The general objective that guides this research is to demonstrate, through an exhaustive bibliographic analysis, that Artificial Intelligence (AI) responds fundamentally to a logic of instrumental reason.

SPECIFIC OBJECTIVES:

1. To theoretically substantiate instrumental reason and analyze the operating principles of AI.
2. Establish the convergence between instrumental reason and AI through bibliographic analysis.
3. Argue the epistemological and ethical implications of instrumental reason in AI.

METHODOLOGY

The methodology adopted is the qualitative analysis of critical bibliographic contents, which allows identifying, organizing and interpreting the relevant theoretical contributions around the concepts of instrumental reason,

technology and knowledge. Through the comparison of classic and contemporary authors, a critical reading of current technological developments linked to AI is articulated.

THEORETICAL FRAMEWORK

This section is structured around three interconnected analytical axes. First, a conceptual cartography is presented aimed at investigating the philosophical and modern sources of instrumental reason and its deployment as the foundation of an epistemology of efficiency. Secondly, it examines efficiency as an expression of an automated reason, and the emptying of the human as an effect of the hegemony of technical rationality in technological contexts dominated by this paradigm. Finally, the question arises about the possibility of an artificial intelligence that operates from a different rationality, capable of incorporating ethical, communicative or non-instrumental dimensions. The development considers both the classical critique of instrumental reason and its contemporary algorithmic reconfiguration, addressing the implications for subjectivity, knowledge, and philosophical tensions that run through this paradigm.

CONCEPTUAL CARTOGRAPHY

This section traces a critical reading of instrumental reason as the foundation of an epistemology of efficiency, from its roots in modernity to its crystallization in contemporary technology.

REASON WITHOUT VALUES: THE ORIGINS OF MODERN EFFICIENCY

Max Weber distinguishes between formal rationality —focused on efficiency and calculation— and substantive rationality —oriented by values— (Weber, 1922). Modern bureaucracy, as a paradigmatic form of formal rationality, transforms human action into predictable procedures, locking the subject in an "iron cage" (Weber, 2009, p. 47). This logic anticipates the operational structure of artificial intelligence, oriented towards optimization without ethical mediation.

CRITIQUE OF INSTRUMENTAL REASON AS A FORM OF DOMINATION

Herbert Marcuse takes up the critique of instrumental reason, understood as "technical rationality emptied of ethical content, centered on the administration of means at the service of given ends" (Marcuse, 1964, p. 8). This technological reason is not neutral: it reproduces forms of control and limits the critical imagination, giving rise to the "one-dimensional man". AI, as an advanced expression of this logic, automates decisions without deliberation, consolidating an epistemology of efficiency.

BEYOND EFFICIENCY: HABERMAS AND THE DEFENSE OF COMMUNICATIVE REASON

Jürgen Habermas contrasts functional rationality with communicative reason, which is based on "intersubjective understanding, dialogue, and consensual validation" (Habermas, 1981, p. 147). The expansion of AI without ethical mediation reinforces the technocratic colonization of the world of life, displacing deliberative spaces (Habermas, 1986). From this perspective, an ethics of artificial intelligence should be inspired by communicative principles that prioritize transparency and participation in the construction of goals.

AUTOMATED REASON: THE TYRANNY OF EFFICIENCY

Artificial intelligence (AI) represents an intensification of instrumental rationality, by automating previously human decision-making processes and reconfiguring the epistemic horizon from technical efficiency. Adela Cortina warns that delegating moral decisions to automated systems can erode autonomy and ethical deliberation, pillars of human judgment. Convergently, Acemoglu points out that autonomous AI could undermine essential human capabilities such as learning and social cooperation.
Zizek (2018) interprets this rationality as an adaptive form focused on precision, while Pérez Palencia introduces the concept of "Epistemolog-AI" to describe the epistemological turn driven by intelligent technologies. Elena Postigo, for her part, claims autonomy, agency and reflexivity as inalienable conditions in the face of a technique that threatens to supplant them.

Recent advances confirm this. The 2024 Nobel Prize in Physics, awarded to Hopfield and Hinton, recognizes the role of neural networks and machine learning in the consolidation of AI capable of optimizing cognitive processes (Hopfield & Hinton, 2024). But this scientific achievement raises fundamental questions: to what extent can judgment be delegated without blurring human agency?

As O'Neil (2016) warns, AI organizes knowledge according to performance parameters, making ethical criteria invisible. Srnicek (2017) argues that this logic subordinates subjective experience to quantifiable metrics, while Harari (2018) highlights that, although these technologies expand our capabilities, they also redefine the very meaning of knowledge, weakening critical autonomy.

This triple function – reproductive, intensifying and transformative – requires a review of the regulatory frameworks that guide the development of AI. Far from being a neutral tool, technology models new forms of rationality that require balancing operational efficiency with ethical principles and humanistic values.

THE EPISTEMOLOGY OF EFFICIENCY IN THE AGE OF ARTIFICIAL INTELLIGENCE

This epistemic displacement deepens the emptying of the human already mentioned: the subjective, the ethical and the deliberative are marginalized in favor of an intelligence without judgment, fast but blind. As Harari (2018) warns, when knowledge is emancipated from human judgment without a critical understanding of the algorithms that generate it, what is at stake is not only truth, but autonomy in the production of meaning.

This section introduces the concept of efficiency epistemology as a critical tool to analyze how artificial intelligence (AI) transforms the modes of knowledge production and validation. The term arises from the convergence between instrumental rationality – focused on utility and performance – and the classic epistemological questions about what is known, how it is known, and who defines valid knowledge (Audi, 2011; Goldman, 2012; Sosa, 2007).

Far from being a simple technical mediator, AI operates as a configuring force of a new rationality, where the knowable is subordinated to optimization criteria. Thus, knowledge is no longer valued only for its truth or justification, but for its operational effectiveness: the ability to be produced, processed, and applied efficiently (Zuboff, 2019; Srnicek, 2017). This epistemic shift shifts understanding towards action, prioritizing agile resolution, automation, and prediction as objectives of knowledge (Pérez Palencia, 2024).

From this logic, efficiency is redefined as an epistemological principle. Instead of validating knowledge for its coherence or grounding, it is legitimized for its practical utility, even if its generation process is opaque or ethically questionable (O'Neil, 2016). This implies a profound mutation: knowing ceases to be a search for understanding and becomes a strategy to operate quickly and obtain quantifiable results.

This displacement carries risks. Harari (2018) warns that by decoupling knowledge from human judgment and entrusting it to unintelligible algorithms, not only truth is compromised, but also critical autonomy. In this context, the epistemology of efficiency marks the transition to a model where the ethical, the subjective and the deliberative are marginalized by a fast but meaningless intelligence.

METHODOLOGY

In this research, a methodological approach of 'qualitative content analysis' (Flick, 2007) was used to decompose and analyze the discourses of the different authors studied. Through 'text analysis' (Denzin & Lincoln, 2015), patterns and categories of analysis were identified, highlighting the points of convergence and divergence between the different perspectives. Main and complementary sources were used, and a rigorous selection was made to ensure coherence with the central theme and to be able to exclude tangential materials. The texts analyzed are summarized in Table 1.

TABLE 1. TEXTS ANALYZED

Author	Text	Year	Category	Editorial
Weber, M.	The Protestant Ethic	2009	Book	www.laeditorialvirtual.com.ar
Weber, M.	Economy and Society	1922	Book	Fondo de Cultura Económica
Marcuse, H.	The one-dimensional man	1964	Book	Ariel
Habermas, J.	Theory of communicative action	1981	Book	Taurus
Habermas, J.	Science and technology as ideology	1986	Book	Tecnos
Cortina, A.	Ethics or ideology of artificial intelligence?	2024	Book	Paidós
Acemoglu, D. & Johnson, S.	Power and progress	2023	Book	Criticism

Zizek, S.	First as a tragedy, then as a farce	2018	Book	Akal
Pérez, M.	Epistemology-IA: the confluence of artificial intelligence and Epistemology in Development Contemporary Scientist	2024	Paper	Ciencia Latina Internacional
Hopfield, J. & Hinton, G.	Could machines rule the world? Nobel laureates have their doubts	2024	Article	El Heraldo de México
O'NEIL, C.	Weapons of Mathematical Destruction. How Big Data Increases Inequality and Threatens Democracy	2017	Paper	Empiria.
Harari, Y.	The 21 Lessons for the 21st Century	2018	Book	Debate
Murillo, Et alt.	Effectiveness, Efficiency, and Value as Modern Approaches to Operations Management	2023	Paper	FIPCAEC.
Covey, S.	The 7 Habits of Highly Effective People	1989	Book	Paidós
Hyatt, M.	Free to Focus: A Total Productivity System to Achieve More by Doing Less	2024	Book	Baker Books
Becerra, P. Lizatte, D.	The measurement of efficiency and productivity by Antonio Álvarez Pinilla (coordinator)	2017	Review	Cuadernos de Economía, National University of Colombia.
Álvarez, A.	Measuring efficiency and productivity	2013	Book	Pyramid
Zuboff, S.	The Age of Surveillance Capitalism: The Struggle for a Human Future in the Face of New Frontiers of Power	2020	Book	Paidós.
Srnicek, N.	Platform Capitalism	2017	Book	Polity Press.
Audi, R..	Epistemology: A Contemporary Introduction to the Theory of Knowledge	2011	Book	(3rd ed.). Routledge
Goldman, A. .	Reliabilism and Contemporary Epistemology: Essays	2012	Book	Oxford University Press
Sosa, E.	A Virtue Epistemology: Apt Belief and Reflective Knowledge	2007	Book	Volume I. Oxford University Press
Kvanvig, J.	The Value of Knowledge and the Pursuit of Understanding	2003	Book	Cambridge University Press.
Pritchard, D.	Epistemic Angst: Radical Skepticism and the Groundlessness of Our Believing	2016	Book	Princeton University Press.
Pérez, C.	Theory of Knowledge: A Contemporary Introduction to Epistemology	2004	Book	Trotta.
Lehrer, K.	Theory of Knowledge	2000	Book	(2nd ed.). Westview Press

Source: Own construction.

The focus of the research is content analysis, focused on identifying and categorizing themes, ideas, and concepts in texts. The aim is to extract relevant information by identifying keywords, repetitive patterns and encoding units of meaning. The analysis focuses on interpreting the layers of meaning, ambiguities, and symbolism in the text. Units of analysis include Units of analysis:

Instrumental rationality. As a critical category that structures modern technical thought (Weber, Frankfurt School, Marcuse, Habermas). As a logic that subordinates means and ends without ethical reflection.

Epistemology of efficiency. As a mode of production, validation and organization of knowledge based on criteria of usefulness, performance and optimization, especially in technological contexts (artificial intelligence).

Automation of reason. Expressed in the delegation of decisions to artificial intelligence systems. Its impact on human subjectivity, agency, and autonomy.

Transformations of artificial intelligence. AI as a force that reproduces, intensifies and transforms instrumental rationality. The emergence of new epistemic configurations (Pérez Palencia, Postigo, Hopfield, Hinton, O'Neil, Srnicek, Harari).

Ethical and communicative dimensions of rationality. Proposed alternatives (Habermas: communicative rationality). The possibility of orienting the development of AI towards horizons that are not merely instrumental.

Each unit of analysis is not isolated, but interrelated with the others (e.g., instrumental rationality runs through the epistemology of efficiency and the configuration of AI). They are based both on fundamental theoretical concepts (such as "instrumental rationality" or "communicative rationality") and on contemporary processes (automation, AI, algorithmic epistemologies).

Frequency and association analysis are used to detect patterns and relationships between concepts, allowing the results to be interpreted and conclusions to be drawn based on the content of the text. Once the units of analysis have been defined and the categories present in the texts worked on have been identified, we codify these categories, quantify the frequency of each of them and examine the relationships between them. We then interpret the possible findings and present results and conclusions.

TABLE 2: QUALITATIVE CONTENT ANALYSIS MATRIX

Unit of analysis	**Category**	**Operational definition**
Instrumental reason	- Means-end relationship - Control and efficiency - Technical mastery	A form of rationality that subordinates knowledge to the achievement of practical, controllable and efficient ends, consolidated in modernity (Weber, Frankfurt School).
Epistemology of efficiency	- Knowledge optimization - Reduced complexity - Instrumentalization of knowledge	Production of knowledge oriented towards efficiency and optimization, leaving critical reflection and meaning in the background.
Automation of the reason	- Human agency and autonomy - Ethics and control	AI transforms human agency and autonomy by making decisions, posing ethical dilemmas about the loss of human control.
Transformations of artificial intelligence.	- Reproduction of instrumental rationality - Transformation of epistemology - New epistemic configurations - Social and human impact of AI	AI amplifies instrumental rationality, transforms the validation of knowledge towards algorithmic logics, redefines epistemic paradigms and reconfigures social organization, affecting subjectivity and human autonomy.
Ethical and communicative dimensions of instrumental rationality	- Communicative rationality - Ethics in technological decision-making - Ethical Alternatives to Instrumental Rationality	Ethical and Communicative Dimensions of Rationality studies how ethical dialogue and critical reflection offer alternatives to the logic of efficiency, seeking to guide technological development towards more humane and just models.

Source: Own construction based on the units of analysis of the texts studied.

RESULTS

THE AGE OF CONTROL: INSTRUMENTAL RATIONALITY AND DIGITAL DOMINATION

The analysis developed confirms that instrumental reason deeply structures modern technical thought, establishing a systematic subordination of human ends to the available means. The means-end relationship is thus marked by a logic where ethical reflection is displaced, privileging exclusively criteria of utility and efficiency. This pattern is intensified in the contemporary context of artificial intelligence, where automated systems reinforce a conception of knowledge oriented towards the control and optimization of results.

In this framework, the category of control and efficiency acquires a central weight: AI not only executes tasks more quickly or accurately, but also organizes the production of knowledge, decisions and social practices under quantifiable performance parameters. Efficiency becomes an end in itself, eroding spaces for criticism and rational deliberation on the objectives pursued.

Finally, the expansion of instrumental rationality in artificial intelligence systems favors new forms of technical domination. The delegation of decisions to algorithmic devices transforms social and political dynamics, affecting the agency and autonomy of individuals. Thus, domination is no longer exercised only by explicit coercion, but through a technical subtlety that shapes behaviors, perceptions, and ways of life.
Taken together, the findings suggest the urgency of building ethical and communicative alternatives that allow us to resist the hegemony of instrumental reason, recovering a critical relationship between knowledge, action and human emancipation.

KNOWLEDGE IN THE AGE OF EFFICIENCY: FROM COMPLEXITY TO OPTIMIZATION

The epistemology of efficiency has profoundly transformed the production and valuation of knowledge, prioritizing the achievement of fast, tangible and measurable results. In this context, knowledge is validated in terms of its practical usefulness and operational effectiveness, configuring a rationality oriented towards constant optimization. Uncertainty and the margin of error are conceived as undesirable elements that must be minimized, even when this implies sacrificing the complexity inherent in the phenomena studied.
This cognitive shift promotes a tendency to simplify reality in order to make it operational, subordinating conceptual depth to immediate applicability. Thus, linear solutions and reductionist models that allow verifiable returns are privileged, to the detriment of holistic approaches capable of capturing the ambiguity and plurality of dimensions present in human experiences and in social reality.
As a result, knowledge is increasingly treated as an instrumental resource, subordinated to objectives of maximization and control. This technical conception reduces knowledge to an instrumental function, displacing forms of knowledge that do not fit into the logic of immediate utility. Ethical, philosophical or critical questions are relegated, while a model is consolidated where knowledge is valuable only to the extent that it can be used to achieve previously defined ends.
Taken together, these findings indicate that the epistemology of efficiency restricts the transformative potential of knowledge and its capacities to nurture critical deliberation, deep understanding, or ethical imagination. Efficiency, converted into a guiding principle, empties knowledge of its reflective dimensions, disarticulating its link with complexity, uncertainty and meaning.

AUTOMATION OF REASON: AUTONOMY AND ETHICS IN THE AGE OF AI

The analysis of the unit "Automation of reason", focused on the categories of human agency and autonomy and ethics and control, reveals profound transformations in the relationship between subjects, technology and power in the context of artificial intelligence (AI). First, the findings indicate that the increasing delegation of decisions to automated systems reconfigures human agency. The outsourcing of deliberative processes to algorithms reduces the ability of individuals to exercise autonomous decisions, diluting the reflexive control over their own actions. This phenomenon affects not only individual autonomy, but also collective forms of self-determination, in an environment increasingly mediated by opaque and self-regulating systems.
Second, the automation of decision-making raises crucial ethical questions. Responsibility for automated decisions becomes fuzzy, especially when these decisions significantly affect social life. Algorithmic opacity makes it difficult to attribute responsibility, increasing the risks of bias, injustice and unintended effects. In this framework, ethical control over AI is weakened, and the absence of transparency in technical processes hinders the exercise of democratic and fair governance.
A third relevant finding refers to the displacement of power structures. The programming, training, and control of AI systems are concentrated in the hands of actors with great technological and economic capacity—such as

corporations and governments—generating an asymmetry in the distribution of cognitive and decision-making power. This centralization transforms the relationship between citizenship and technology, widening the gap between those who design algorithms and those who must live under their consequences.
These findings show that the automation of reason not only introduces functional efficiencies, but also profound challenges for autonomy, ethics, and equity. The need to develop normative frameworks that safeguard fundamental human rights becomes urgent, in a scenario where algorithmic decisions tend to replace – and sometimes supplant – processes of critical human deliberation.

AI: REDEFINING RATIONALITY AND KNOWLEDGE

The analysis of the transformations induced by artificial intelligence (AI) reveals a profound reconfiguration in the logic that guides rationality, knowledge production and contemporary social structures. First, AI reinforces the principles of instrumental rationality by maximizing measurable results and optimizing processes without explicit consideration of their ethical implications. This tendency intensifies the subordination of human ends to technical means, consolidating a logic of performance that dissociates operational efficiency from the values that historically guided reflective human action.

From an epistemological perspective, AI introduces a mutation in the criteria for validating knowledge. Knowledge, traditionally linked to experience, critical deliberation and contextual interpretation, is progressively replaced by algorithmic models based on quantifiable data, statistical correlations and computational patterns. This epistemic mutation displaces human subjects as central agents of knowledge, delegating judgment to automated systems whose internal logic often escapes public scrutiny.
These transformations, however, are not limited to the technical or cognitive realm, but significantly alter social structures and subjective experience. Automating tasks and decisions using AI directly affects human agency, weakening autonomy and the ability to engage thoughtfully in the processes that shape everyday life. Consequently, fundamental ethical dilemmas linked to control, responsibility and justice emerge. The question of who designs, controls and benefits from these systems acquires a particular urgency in contexts where automated decisions affect rights, opportunities and living conditions.

In short, the deployment of artificial intelligence not only redefines rationality in functional terms, but also imposes a new epistemological and political matrix that demands a critical revision. These transformations require approaching AI from a comprehensive ethical perspective that considers both its technical implications and its effects on autonomy, equity, and social justice. Only from a critical approach will it be possible to move towards a responsible integration of these technologies into collective life.

CHALLENGING INSTRUMENTAL RATIONALITY: ETHICS AND COMMUNICATION IN AI

The growing influence of artificial intelligence (AI) in the social, political, and cultural spheres demands an approach that transcends technical efficiency and recognizes the ethical implications of its development. From this perspective, the analysis focused on the communicative and normative dimensions of rationality allows us to identify three critical lines: communicative rationality, ethics in technological decision-making and ethical alternatives to instrumental rationality.

In the first place, the notion of communicative rationality formulated by Habermas (1984) is proposed as a normative way in the face of the dominant instrumental logic, by privileging deliberation, consensus and intersubjective understanding. This perspective challenges technocratic models that reduce decision-making to efficient calculation, and places the inclusion of multiple voices as a condition for socially just artificial intelligence.

Second, ethics in technological decision-making requires a substantive integration of moral values in the design and deployment of AI systems. It is not a question of correcting ethical failures a posteriori, but of preventing inequalities and biases from its algorithmic architecture, interrogating the ideological and distributive structures that sustain these technologies.

Finally, ethical alternatives to instrumental rationality are highlighted, such as the ethics of care (Gilligan, 1982) and Kantian deontologism, which focus on responsibility, attention to vulnerability and the dignity of subjects. These normative orientations propose a link between technology and society not based on mere optimization, but on justice, solidarity and mutual respect.

Taken together, these findings warn about the limits of the technical paradigm and call for a rethinking of the ethical foundations of contemporary technological development. Beyond its functional potential, AI must be

conceived as a social mediation that responds to deliberative and democratic principles aimed at collective well-being.

DISCUSSION

FUNDAMENTAL TENSIONS IN THE EPISTEMOLOGY OF EFFICIENCY

The emergence of an epistemology of efficiency, focused on operational utility and pragmatic optimization, marks a shift that challenges the classical foundations of knowledge. This displacement is not merely technical, but epistemic, affecting the centrality of truth as correspondence (Audi, 2011), which is replaced by the production of useful results, even without causal understanding (Zuboff, 2019; Srnicek, 2017).

Traditional criteria of justification—such as critical evidence (Goldman, 2012; Lehrer, 2000)—as opposed to functional and rapid validation logics. Conceptual understanding is dismissed when operational efficiency is sufficient, as in algorithmic systems that outperform humans without clear explainability.

This logic, moreover, erodes the epistemic agency of the subject, replacing reflective judgment with the passive acceptance of automated decisions (Haack, 2003; O'Neil, 2016). In this context, traditional epistemic values are subordinated to instrumental and productive criteria, in line with the instrumental reason described by Weber and Marcuse, and critically revisited by current authors (Pérez Palencia, 2024).

In short, knowledge ceases to be measured by its truth or justification, and is valued by its effectiveness, configuring a mutation that requires critical revision to preserve the autonomy of thought and the ideals of the Enlightenment.

TABLE 3. CRITICAL TENSIONS: FROM LOGOS TO CALCULATION

EPISTEMIC TENSION	TRADITIONAL LOGIC OF KNOWLEDGE	ALGORITHMIC LOGIC/EFFICIENCY
TRUTH V/S PRAGMATIC UTILITY	Truth as a correspondence with facts (Audi, 2011)	Prioritization of useful outcomes, although without verifiable causal explanation (Zuboff, 2019)
JUSTIFICATION VS. OPERATIONAL SPEED	Justified knowledge requires evidence, coherence, and critical evaluation (Goldman, 2012)	Immediate and functional responses are prioritized over argumentation or deliberation
COMPREHENSION VS. OPTIMISATION	Valorization of conceptual structures, causal relationships and deep understanding	Effective problem-solving without explicit epistemic understanding
CRITICAL REFLECTION VS. AUTOMATION	The subject reflects on conditions of production and validity of knowledge (Haack, 2003)	Uncritical confidence in the results offered by supposedly objective algorithms
EPISTEMIC VALUES VS. UTILITARIAN CRITERIA	Truth, coherence, consistency, and justification prevail as epistemic ends in themselves	Subordination of knowledge to criteria of productivity, efficiency and profitability (Marcuse, Weber; Pérez Palencia, 2024)

Source: Own construction

IS ARTIFICIAL INTELLIGENCE POSSIBLE BEYOND TECHNICAL RATIONALITY?

The current debate around artificial intelligence (AI) raises a crucial question: can its development transcend the frameworks of technical rationality, focused on efficiency, control, and prediction? This question requires exploring alternative models that integrate ethical, deliberative and evaluative criteria, beyond the mere algorithmic logic of optimization.

To address this problem, an interdisciplinary theoretical framework is used that includes Zuboff, Srnicek, Han, Habermas, Crawford, Benjamin and Postigo. These authors offer key conceptual tools to critically analyze the relationship between AI, power, communication, and ethics, and allow us to question whether AI is viable not only for performance, but also for broader human and societal purposes.

SOCIOECONOMIC CRITICISMS OF TECHNICAL RATIONALITY IN AI:

The technical rationality that guides the development of artificial intelligence (AI) is deeply intertwined with the logics of contemporary capitalism. Zuboff (2019) argues that today's AI cannot be understood without the framework of surveillance capitalism, where personal data is transformed into inputs for the prediction and control of behaviors for profit, subordinating ethical considerations to accumulation objectives. This rationality, far from being neutral, reinforces value-extracting mechanisms that reduce individual agency and simplify human complexity.

In a complementary way, Srnicek (2017) warns that AI is structurally integrated into platform capitalism, a model based on the massive collection of data to optimize services, reduce costs, and maximize benefits. This market logic reconfigures the purposes of AI, turning it into an instrument of instrumental efficiency at the service of profitability.

Both perspectives reveal that imagining an ethical and deliberative AI requires not only technical adjustments, but also a structural transformation of the economic models that today condition its development.

THE EROSION OF MEANING AND THE QUANTIFICATION OF EXISTENCE BY ALGORITHMIC LOGIC.

Byung-Chul Han (2017, 2022) offers a broader cultural and philosophical critique of the impact of digitalization on contemporary society. For Han, the digital society has displaced the search for meaning by the accumulation of information, giving rise to a "society of transparency" where truth and understanding are replaced by the proliferation of data. In works such as "Psychopolitics" and "Infocracy," Han argues that the current implementation of AI tends to reduce the richness of human experience to the quantifiable, eliminating essential elements for critical thinking such as negativity, pause, error, and reflection. While Zuboff denounces the capture of AI by economic imperatives, Han problematizes the cultural and existential transformation that algorithmic logic imposes on the contemporary subject, suggesting that an AI outside of technical rationality would imply resisting this reduction of existence to actionable data.

COMMUNICATIVE RATIONALITY AS A PARADIGM FOR A "CRITICAL AI"

In contrast to the predominantly critical perspectives of Zuboff and Han, Jürgen Habermas (1981) offers a normative basis for imagining alternatives. His fundamental distinction between instrumental rationality (oriented towards success and strategic calculation) and communicative rationality (oriented towards mutual understanding through free and non-coercive dialogue) provides a theoretical framework for conceiving forms of social and technological interaction based on intersubjective understanding and agreement, rather than mere efficiency and control. From this perspective, the development of a "critical AI" would imply its insertion in social contexts where the ends are not predefined by the logic of the market or technology, but emerge from deliberative processes oriented towards justice, inclusion and human dignity. Thinking of AI outside of technical rationality, from the Habermasian perspective, involves exploring how it might facilitate non-instrumental communication, ethical debate, and consensus-building, rather than being purely a tool for optimization and one-way influence.

ETHICAL CRITICISMS AND ALGORITHMIC JUSTICE:

Contemporary ethical critiques of artificial intelligence (AI) reveal that its technical rationality, based on ideals of efficiency and neutrality, hides deep power asymmetries, human decisions, and structural biases. Kate Crawford (2021), in Atlas of AI, dismantles the idea of an ethereal and purely informational AI, highlighting its materiality: an infrastructure intensive in natural resources, energy and precarious work. The technical rationality that drives AI, he argues, extends to an exploitative logic that cuts across the entire lifecycle of the system, from its physical base to its algorithmic operations.

Crawford also warns that the data is not neutral or "crude," but rather social constructs laden with prejudice. AI, by looking for efficient patterns, can amplify these distortions, disproportionately affecting historically marginalized groups in critical areas such as criminal justice, employment, or surveillance. Overcoming this technical rationality implies, according to Crawford, moving beyond superficial reforms to a radical critique of the epistemic and political structures that sustain it.

In a convergent vein, Ruha Benjamin (2019), in Race After Technology, denounces how AI can encode new forms of racial discrimination under an appearance of objectivity. With the concept of "New Jim Code," Benjamin points out that algorithmic systems reproduce historical racial hierarchies through mechanisms such

as engineered inequity, default discrimination, and coded exposure. These practices reinforce surveillance and control over racialized populations under the logic of supposedly neutral efficiency.
Faced with this reality, Benjamin proposes a technological abolitionist approach, which is not limited to reforming current systems, but requires actively dismantling the racial power structures inscribed in technology. Her analysis questions the technical rationality of AI from an ethics of racial justice, proposing a profound reconfiguration of the link between technology and equity.
Both authors agree that the technical rationality that structures AI is neither ethical nor neutral, but rather perpetuates material, epistemic and racial inequalities. Imagining a fairer AI therefore requires an incisive critique of its foundational assumptions and a reorientation towards deliberative and emancipatory ends.

THE HUMANIST PERSPECTIVE AND THE QUESTION OF AUTONOMY:

From a bioethical and philosophical perspective, Elena Postigo proposes a critique of the dominant technical rationality in the development of artificial intelligence, underlining the need to preserve the intrinsic dignity of the human person. This dignity, he argues, is based on three essential capacities: autonomy (free and conscious decision), agency (action and influence in the world) and reflexivity (critical evaluation of one's beliefs and motivations).
Postigo warns that AI geared exclusively towards efficiency and prediction can erode these capabilities by reducing humans to actionable data and automated decisions. Against this, he defends an approach to technological development that subordinates algorithmic logic to robust ethical principles. AI, in this view, should be designed to strengthen deliberation, critical judgment, and self-determination, and not to replace them.
His approach is opposed to any functionalist vision that measures human value according to criteria of utility or performance. Instead, it proposes an anthropology that recognizes the complexity and transcendence of the person, orienting AI towards human flourishing and not towards its instrumentalization (Postigo, 2021).

CONVERGENCES AND DIVERGENCES:

The perspectives of Shoshana Zuboff, Nick Srnicek, Byung-Chul Han, Jürgen Habermas, Kate Crawford, Ruha Benjamin, and Elena Postigo converge in a substantial critique of the technical rationality that currently informs the development of artificial intelligence. Despite their different disciplinary approaches and analytical emphases, these authors share a fundamental concern for the primacy of efficiency and optimization, the potential erosion of human autonomy and agency, and the reproduction and amplification of social inequalities inherent in the operational logic of AI.

TABLE 4: COMMENTS ON CONVERGENCES AND DIVERGENCES AMONG THE AUTHORS STUDIED.

Aspect	Authors who agree	Feedback
Convergence 1: Critique of the primacy of efficiency	Zuboff, Srnicek, Han, Crawford, Benjamin, Postigo	The common criticism is that efficiency as a guiding value of AI displaces broader ethical and societal considerations.
Convergence 2: Diminishing Human Autonomy and Agency	Zuboff, Srnicek, Han, Crawford, Benjamin, Postigo	Automation and reliance on efficient systems reduce human self-determination and agency.
Convergence 3: An issue of inequality	Crawford, Benjamin, Zuboff, Srnicek, Han	Inequality is approached from the perspective of biases in AI and the power asymmetries exacerbated by capitalism.

Aspect	Authors who agree	Feedback
Divergence 1: Socioeconomic Analysis vs. Algorithmic Justice	Zuboff, Srnicek vs Crawford, Benjamin	Zuboff and Srnicek address macro economics and its impact on AI, while Crawford and Benjamin focus on AI's internal biases.
Divergence 2: Bioethical and Anthropological Perspective vs. Cultural and Existential Criticism	Shutter vs Han	Postigo focuses on the impacts on the human condition and autonomy from an ethics of care, while Han reflects on the cultural transformation induced by AI.
Divergence 3: Normative Framework vs. Implicit Epistemological Critique	Habermas vs Zuboff, Han	Habermas proposes an alternative normative framework focused on communicative rationality. He argues that

	the solution to the colonization of the world of life by technical systems is not only critical, but also the construction of emancipatory communicative spaces. while Zuboff and Han: offer powerful critiques of the current order, but do not develop a systematic alternative at the epistemological or normative level as Habermas does.

Source: Own construction

CONCLUSIONS

TOWARDS AN ETHICAL AND DELIBERATIVE ARTIFICIAL INTELLIGENCE BEYOND TECHNICAL RATIONALITY.

The convergent analysis of authors such as Shoshana Zuboff, Byung-Chul Han, Jürgen Habermas, Elena Postigo, Nick Srnicek, Kate Crawford and Ruha Benjamin evidences a central concern: the current development of artificial intelligence is strongly anchored in a technical rationality that privileges efficiency, optimization and prediction, to the detriment of fundamental ethical and human considerations.

Zuboff (2019) and Srnicek (2017) denounce how AI has become an instrument of surveillance capitalism, which exploits human behavior as a resource, eroding privacy and reducing individual autonomy. Byung-Chul Han (2012) complements this critique by warning about the tendency towards total quantification that suppresses negativity, reflection and the construction of meaning. In parallel, Crawford (2021) and Benjamin (2019) demonstrate that AI, far from being neutral, reproduces and amplifies pre-existing social inequalities, such as racial biases, under the guise of algorithmic objectivity.

From a bioethical perspective, Postigo (2021) insists on the centrality of the human person—with his or her dignity, autonomy, and reflexivity—in the face of the risk of technological instrumentalization. In turn, Chris Anderson (2008) warns about the reductionism of "data without theory", which ignores the social contexts where AI operates.

Faced with this panorama, an ethical reorientation of technological development becomes urgent. Habermas' (1984) distinction between instrumental and communicative rationality offers a normative horizon for imagining an AI that not only automates, but promotes dialogue, deliberation, and intersubjective understanding. This shift involves moving towards a "critical AI," one that incorporates algorithmic justice as a guiding principle, recognizes power dynamics, and designs systems that empower people, not replace or manipulate them.

Building an ethical and deliberative artificial intelligence requires an interdisciplinary and collective commitment that integrates scientists, humanists, legislators and citizens. It's not just about improving technological efficiency, it's about aligning AI with the values of human flourishing, equity, and shared responsibility.

FUTURE LINES OF RESEARCH

The theoretical framework developed opens several critical lines of inquiry on the development of artificial intelligence (AI) in an ethical-social key. A first line proposes to design AI systems inspired by the communicative rationality of Habermas (1984), integrating principles of dialogue and deliberation beyond mere efficiency. Second, the development of metrics to assess algorithmic fairness is proposed, in line with the warnings of Crawford (2021) and Benjamin (2019) about the structural biases of AI.

From bioethics, Postigo (2021) calls for examining how algorithmic mediation affects the autonomy and agency of subjects. In addition, it is suggested to investigate models of "reflexive opacity" that allow situated forms of social control without requiring total transparency. Zuboff (2019) and Srnicek (2017) inspire a critique of business models based on surveillance and data mining, proposing alternative frameworks.

Likewise, the need to design AI with racial awareness and social justice, facing structural inequalities, is highlighted (Benjamin, 2019). Finally, following Han (2012), he proposes to study how algorithmic logic affects the loss of meaning and the human capacity for symbolization and reflection.

These lines point not only to applied ethics, but also to a profound rethinking of the civilizational assumptions that underpin contemporary artificial intelligence.

ACKNOWLEDGEMENTS AND FUNDING:

This is one of the documents corresponding to the research DIUMCE **" LA NUEVA CARTOGRAFÍA DEL TRABAJO A TRAVÉS DE PLATAFORMAS DIGITALES. UN ESTUDIO DE CASO A TRAVÉS DE LOS DISCURSOS DE ACADÉMICOS/AS DE EDUCACIÓN SUPERIOR ", Code** 33-2024 EFA. Funded by the Academic Vice-Rector's Office and the Directorate of Research and Postgraduate Studies of the Universidad Metropolitana de Ciencias de la Educación.

BIBLIOGRAPHIC REFERENCES

1. Acemoglu, D., & Johnson, S. (2023). Power and Progress: Our Millennial Struggle for Technology and Prosperity (M. Sanz, Trans.). Deusto. Power and Progress: Our Thousand-Year Struggle Over Technology and Prosperity
2. Adorno, T. & Horkheimer, M. (1994). Dialectics of Enlightenment (H. Murillo, Trans.). Trotta.
3. Álvarez, A. (2013). Measuring efficiency and productivity. Pyramid
4. Anderson, C. (2008). The end of theory: The data deluge makes the scientific method obsolete.
5. Audi, R. (2011). Epistemology: A Contemporary Introduction to the Theory of Knowledge (3rd ed.). Routledge.
6. Becerra, P. Lizatte, D. (2017). The measurement of efficiency and productivity by Antonio Álvarez Pinilla (coordinator). Review of Cuadernos de Economía, Universidad Nacional de Colombia
7. Benjamin, R. (2019). Race after technology: Abolitionist tools for the new Jim code. Polity Press.
8. Cortina, A. (2013) What is ethics really for? Paidós
9. Covey, R. (1989). The 7 habits of highly effective people. Paidós
10. Crawford, K. (2021). Atlas of AI: Power, politics, and the planetary costs of artificial intelligence. Yale University Press
11. Denzin, N., & Lincoln, Y. (2015). The SAGE handbook of qualitative research (5th ed.). SAGE Publications
12. Flick, U. (2007). Introduction to Qualitative Research (2nd ed.). Morata.
13. Gilligan, C. (1982). In a different voice: Psychological theory and women's development. Harvard University Press
14. Goldman, A. (2012). Reliabilism and Contemporary Epistemology: Essays. Oxford University Press
15. Haack, S. (2003). Defending science—within reason: Between scientism and cynicism. Prometheus Books
16. Habermas, J. (1981). Theory of communicative action. Taurus
17. Habermas, J. (1986). Science and technology as ideology. Tecnos
18. Han, B.-C. (2012). The Transparency Society. Herder
19. Han, B.-C. (2017). Psychopolitics: Neoliberalism and New Power Techniques. Herder
20. Harari, Y. (2018). The 21 lessons for the 21st century. Debate
21. Hopfield, J. & Hinton, G. (2024) Could machines rule the world? https://heraldodemexico.com.mx/tecnologia/2024/10/9/podrian-las-maquinas-gobernar-el-mundo-los-ganadores-del-nobel-tienen-sus-dudas 644296.html?utm_source=chatgpt.com#google_vignette
22. Hyatt, M. (2024). Free to Focus: A Total Productivity System to Achieve More by Doing Less. Baker Books
23. Kvanvig, J. (2003). The Value of Knowledge and the Pursuit of Understanding. Cambridge University Press
24. Lehrer, K. (2000). Theory of Knowledge (2nd ed.). Westview Press
25. Marcuse, H. (1964). The One-Dimensional Man (A. Tió, Trans.). Ariel
26. Murillo, I., Baldeón, F., Lamilla, I., & Alvarado, P. (2023). Effectiveness, efficiency, and value as modern approaches to operations management. FIPCAEC. Doit: https://doi.org/10.23857/fipcaec.v8i3
27. O'Neil. (2016). Weapons of Mathematical Destruction. How Big Data increases inequality and threatens democracy. https://www.red, alyc.org/journal/2971/297165396011/html/
28. Pérez, C. (2004). Theory of Knowledge: A Contemporary Introduction to Epistemology. Trotta
29. Pérez, M. (2024). Epistemology-AI: the confluence of artificial intelligence and epistemology in contemporary scientific development. International Latin Science. University of Panama. DOIT: https://doi.org/10.37811/cl_rcm.v8i4.13264
30. Pritchard, D. (2016). Epistemic Angst: Radical Skepticism and the Groundlessness of Our Believing. Princeton University Press
31. Postigo, H. (2021). Ethics of artificial intelligence: A view from personalist bioethics. University of Navarra (EUNSA)
32. Sosa, E. (2007). A Virtue Epistemology: Apt Belief and Reflective Knowledge, Volume I. Oxford University Press
33. Srnicek, N. (2017). Platform Capitalism. Polity Press.
34. Weber, M. (2009). The Protestant Ethic and the Spirit of Capitalism (J. A. Pérez, Trans.) Alliance. (Original work published in 1905)

35. Weber, M. (2014). Economy and society. First electronic edition, www.fondodeculturaeconomica.com
36. Weber, M. (1922). Economy and Society: Outline of Comprehensive Sociology. Tübingen: Mohr
37. Wired. https://www.wired.com/2008/06/pb-theory/rationalization of society (T. McCarthy, Trans.). Beacon Press. (Original work published 1981)
38. Zizek, S. (2018). First as a tragedy, then as a farce. Akal
39. Zuboff, S. (2020). The Age of Surveillance Capitalism: The Struggle for a Human Future in the Face of New Frontiers of Power. Paidós
40. Zuboff, S. (2019). The age of surveillance capitalism: The fight for a human future at the new frontier of power. PublicAffairs