Conceptual Reflection Article

Received: 19 June 2026
Accepted: 24 October 2026
Published: 07 November 2026

Cognitive Security and Philosophical Dialogue: A Historical-Philosophical Analysis of Ideological Manipulation in Information Warfare

Liudmyla Krymets
Department of Social Sciences, Kruty Heroes Military Institute of Telecommunication and Information Technology, 45/1 Knyaziv Ostrozkyh St., Kyiv, 01011, Ukraine; liudmylakrymets@gmail.com

Vitalii Chornyi
Department of Social Sciences, Kruty Heroes Military Institute of Telecommunication and Information Technology, 45/1 Knyaziv Ostrozkyh St., Kyiv, 01011, Ukraine

Natalia Logvinenko
Department of Social Sciences, Kruty Heroes Military Institute of Telecommunication and Information Technology, 45/1 Knyaziv Ostrozkyh St., Kyiv, 01011, Ukraine

Serhii Khovrych
Department of Social Sciences, Kruty Heroes Military Institute of Telecommunication and Information Technology, 45/1 Knyaziv Ostrozkyh St., Kyiv, 01011, Ukraine

Maksym Udod
Department of Social Sciences, Kruty Heroes Military Institute of Telecommunication and Information Technology, 45/1 Knyaziv Ostrozkyh St., Kyiv, 01011, Ukraine

Cite this article:
L. Krymets et al. (2026). Cognitive Security and Philosophical Dialogue: A Historical-Philosophical Analysis of Ideological Manipulation in Information Warfare. Revista Cultura Científica, (24), pp. 2477–2490.
https://doi.org/10.70517/revcc2624199

2026 © The Author(s). Published by Revista Cultura Científica. This article is published as open access under the Creative Commons Attribution 4.0 International (CC BY 4.0) license (http://creativecommons.org/licenses/by/4.0/).

Abstract

The intensification of information warfare in algorithmically mediated environments has shifted manipulation from the content of beliefs to the conditions under which perception, interpretation and judgement are formed. The objective of this study is the conceptualisation of cognitive security through a historical-philosophical reconstruction of the relationships among cognition, truth, power and algorithmically mediated perception. The study takes the form of a conceptual philosophical reflection article and combines historical-philosophical analysis, conceptual genealogy, comparative philosophical interpretation and conceptual synthesis. The positions of Plato, Aristotle, Al-Farabi, Machiavelli, Bacon, Kant, Heidegger, Arendt and Foucault are correlated with contemporary scholarship on algorithmic governance, surveillance capitalism, information warfare and epistemic injustice. The analysis develops the argument that the classical problem of illusion has moved from the representational to the infrastructural level: ranking, filtering and personalisation pre-structure the field in which truth and falsity can be distinguished before judgement begins. Information warfare is accordingly interpreted as a struggle over the conditions of perception rather than over particular narratives. On this basis, cognitive security is defined as a philosophical category denoting the preservation of reflective judgement, interpretive autonomy and epistemic plurality, distinct from its technical meaning in cybersecurity. Philosophical dialogue is identified as a practice of epistemic resistance and epistemic repair that introduces reflective friction into accelerated communication, restores the relational conditions of shared intelligibility and underpins educational formation. The proposed framework offers conceptual grounds for digital governance, media literacy and educational policy, while its empirical operationalisation remains a task for further research.

Keywords: cognitive security, information warfare, algorithmic epistemology, philosophical dialogue, power–knowledge, interpretive autonomy

1. INTRODUCTION

Information warfare has become a structural feature of contemporary public life rather than an exceptional condition of armed conflict. Twentieth-century mass communication had already demonstrated that public perception can be organised on a large scale: the propaganda model proposed by Herman and Chomsky [1] described how the institutional structure of the mass media filters what reaches the public, and the state propaganda systems of the 1930s showed that control over perception could be pursued deliberately as a political project. The digital transformation has changed both the scale and the mechanism of this control. The Cambridge Analytica case during the 2016 electoral campaigns and the activities attributed to the Internet Research Agency drew public and scholarly attention to the use of personal data for micro-targeted political messaging and to sustained computational influence campaigns across social media platforms [2, 3]. These cases illustrate the expansion of ideological persuasion to include algorithmically mediated behavioural steering, in which the visibility of information is ordered before reflection begins.

In networked media environments, information flows are shaped by asymmetric media ecosystems and recursive feedback loops that blur the distinction between exposure and persuasion [2]. Memetic forms of conflict use content not as neutral information but as a strategic means of organising attention, understanding and perception in the digital sphere [4]. The decisive change therefore concerns not isolated instances of deception but the continuous modulation of attention and interpretation. When the direction of perception, the allocation of attention and the range of available interpretations are set in advance, the independence of judgement is weakened, and the trust on which judgement relies becomes dependent on opaque algorithmic institutions [5]. For Arendt, factual truth is the precondition of a common world, and the erosion of stable facts leads citizens not to disagree about a shared reality but to inhabit realities that no longer coincide [6, 7]. Contemporary social epistemology describes the same process as the fragmentation of shared reality under conditions of polarisation, in which plurality ceases to be a diversity of perspectives on one world and turns into a set of mutually incompatible informational worlds [8].

The scholarly literature offers several strong diagnoses of this situation. The theory of surveillance capitalism explains how human experience is converted into behavioural data used to predict and modify behaviour [9]; the theory of the network society describes the redistribution of knowledge across interconnected informational infrastructures [10]; studies of the social power of algorithms show that computational systems order what becomes visible, relevant and knowable [11]; and recent Foucauldian analyses interpret the transition from disciplinary to algorithmic societies as a transformation in the construction, validation and dissemination of knowledge [12]. Research on algorithmic authority demonstrates that the ranking, filtering and prediction of information shift epistemic authority from human interpretive judgement towards computational selection [13], while the theory of epistemic injustice clarifies how credibility and interpretive resources are distributed unequally among knowers [14]. Security-oriented studies examine the use of artificial intelligence in information warfare and the military framing of information manoeuvre [3, 15], the methods of countering disinformation [16] and the spread of fake news [17].

Despite their diagnostic strength, these accounts address different aspects of the problem and require closer integration for the specific philosophical inquiry pursued here. Sociotechnical analyses describe how platforms organise visibility, while the present study examines how the classical problems of truth, appearance, perception and judgement are reconfigured when cognition operates within algorithmically structured epistemic spaces. Security studies treat cognition as a target of influence operations, while the notion of security applied to cognition is usually borrowed from the technical vocabulary of cybersecurity and artificial intelligence safety, where it concerns the protection of systems and the prevention of harmful failures or malicious uses [18, 19]. A research gap therefore persists between the historical-philosophical tradition, which has long examined the vulnerability of knowledge to illusion, rhetoric and ideological structuring, and contemporary research on algorithmic manipulation, which can benefit from a more explicit philosophical framework connecting cognition, power and truth across historical and contemporary contexts.

The conceptual problem of the study follows from this gap. The central issue is no longer whether individuals are deceived by particular messages but whether the conditions of independent judgement persist within technologically prestructured epistemic environments, and which philosophical resources allow these conditions to be described and protected. The study addresses this problem by placing digital epistemology within a long philosophical tradition that extends from classical theories of perception to contemporary analyses of algorithmic governance. Its theoretical contribution is threefold: a genealogical reconstruction of the relationship among cognition, truth and power that links classical philosophy with the algorithmic information environment; a philosophical definition of cognitive security that distinguishes it from its technical meaning; and a justification of philosophical dialogue as a practice of epistemic resistance and epistemic repair. The objective of this study is the conceptualisation of cognitive security through a historical-philosophical reconstruction of the relationships among cognition, truth, power and algorithmically mediated perception. Section 2 describes the methodology, Section 3 presents the findings of the conceptual analysis, Section 4 discusses them in relation to existing scholarship and sets out the limitations of the study, and Section 5 formulates the conclusions.

2. METHODOLOGY

The study is a conceptual philosophical reflection article in the field of social philosophy and the philosophy of information. It does not collect or analyse empirical data and does not follow the protocol of a systematic or scoping review. The research design combines four interrelated approaches. Historical-philosophical analysis reconstructs how classical and modern thinkers formulated the problems of truth, appearance, cognition and power. Conceptual genealogy, developed in the Foucauldian tradition [12, 20], traces how the relationship between truth and power has been reconfigured from classical philosophy to the age of algorithmic governance and treats knowledge as a set of historically layered formations of truth regimes rather than as a stable faculty. Comparative philosophical interpretation sets the positions of different thinkers against one another and against contemporary scholarship in order to identify continuities and ruptures. Conceptual synthesis integrates the results into the concept of cognitive security and into an account of philosophical dialogue as a form of epistemic resistance.

The study adopts a dialogical model of interpretation characteristic of the hermeneutic tradition [21]. Classical and contemporary philosophers are treated not as isolated theoretical positions but as participants in a continuous inquiry into cognition as a condition of the disclosure of the world. The aim of this model is not a reductive synthesis but the identification of a stratified ontology of cognition, in which cognition is distributed across shifting structures of power, discourse and technical mediation.

The analytical corpus consists of two groups of texts that perform different functions in the argument. The first group comprises classical and modern philosophical works that supply the conceptual categories of the analysis: Plato's allegory of the cave and his critique of illusion in the Republic, read with the assistance of established commentary [22]; Aristotle's account of habituation and the formation of judgement [23]; Al-Farabi's conception of the virtuous city [24]; Machiavelli's analysis of political appearance [25]; Bacon's doctrine of the idols [26]; Kant's critical theory of the conditions of experience [27]; Hegel's dialectic [28]; Heidegger's analysis of technology as enframing [29]; Arendt's conception of factual truth and the common world [6, 7]; and Foucault's analysis of power and knowledge [20], alongside its contemporary interpretation [12]. The theories of communicative action [30], philosophical hermeneutics [21], epistemic injustice and epistemic resistance [14, 31] and the epistemologies of the South [32] serve as conceptual resources for the analysis of dialogue.

The second group comprises contemporary scholarship that describes the algorithmic information environment. It includes studies of surveillance capitalism, the network society and the social power of algorithms [9–11]; the analyses of psychopolitics by Han [33] and of media culture by Postman [34]; research on algorithmic authority, the phenomenology of algorithms and the aesthetics of algorithmic disinformation [13, 35, 36]; studies of truth as social practice, trust in algorithmic institutions and artificial intelligence, and the fragility of truth under polarisation [5, 8, 37, 38]; and research on information warfare, disinformation and the malicious use of artificial intelligence [2–4, 15–19, 39]. Works on attention, echo chambers and the phenomenology of experience [40–42] connect the two groups. The groups are related asymmetrically: the philosophical texts provide the categories through which the problem is interpreted, whereas the contemporary literature provides the description of the environment in relation to which these categories are interpreted. The corpus is purposive and interpretive and does not claim exhaustive coverage of the literature on any of the topics involved.

The analysis proceeds in four steps. First, the categories of truth, appearance, cognition, power, judgement and autonomy are identified in the philosophical texts, together with the specific epistemic problem that each thinker associates with them, such as illusion in Plato, habituation in Aristotle, the management of appearances in Machiavelli, structural distortion in Bacon and the regime of truth in Foucault. Second, each category is correlated with the operations of the algorithmic information environment described in the contemporary literature, namely ranking, filtering, personalisation, recommendation, classification and prediction. Third, the correlations are evaluated in terms of continuity and transformation: the analysis establishes whether a classical problem persists, intensifies or changes its mode of operation, for example by moving from the representational to the infrastructural level. Fourth, the results are synthesised into the concept of cognitive security and into an account of philosophical dialogue as a practice of epistemic resistance and epistemic repair, and the outcome of the comparison is summarised in Table 1.

The credibility and coherence of the interpretation are assessed through the consistency of the conceptual framework, the explicit correlation of each classical category with its contemporary analogue and the use of established scholarship for the description of the algorithmic environment. As the study is based exclusively on published texts and involves no human participants, personal data or experimental procedures, ethics committee approval was not required.

3. FINDINGS

This section presents the findings of the conceptual analysis. In accordance with the reflective design of the study, the findings are interpretive propositions derived from the correlation of classical philosophical categories with the operations

of the algorithmic information environment rather than empirical results.

3.1. FROM EPISTEMIC ILLUSION TO ALGORITHMIC MEDIATION

The first finding concerns the nature of the digital transformation itself. Digital technologies do not merely change the means of communication; they transform the conditions under which knowledge is acquired, justified, disseminated and contested. Sahakyan et al. [12] show that the transition from disciplinary to algorithmic societies rests on a change in how knowledge is constructed, validated and distributed under technological governance. In this environment, digital systems do not simply transmit knowledge but frame what counts as knowledge, truth and credibility. Epistemology must therefore address the infrastructural conditions of belief formation alongside its reflective concern with justification, because cognition increasingly operates through continuously reorganised and externally mediated processes. The problem is therefore phenomenological as well as epistemological, since it concerns the structures through which experience is organised before it becomes knowledge [42].

This shift reactivates the classical problem of the relationship between appearance, truth and perception. In the allegory of the cave, Plato described human beings who take mediated appearances for reality itself [22]. In the digital environment, this image acquires renewed interpretive significance: algorithmic infrastructures can be understood by analogy as a contemporary cave in which visibility is curated, attention is organised and perception is pre-shaped before conscious reflection. Phenomenological analyses of how social media algorithms appear to their users show that these systems are encountered not as transparent tools but as opaque conditions of experience [35]. In this interpretation, the epistemic illusion represented by Plato's allegory can acquire an additional basis in computational infrastructures.

Aristotle's theory of habituation offers a second point of correlation. For Aristotle, intellectual and moral dispositions develop through repeated practice, sustained action and deliberation, and perception and judgement are formed gradually rather than given immediately [23]. In digital ecosystems, habituation remains a process of engagement with the world, but it is also increasingly shaped by algorithmically structured patterns of interaction that mediate perception and attention. Habits of attention are formed through repetition designed for engagement, and repetition itself operates as a means of persuasion [37]. The question of the conditions under which judgement is formed therefore acquires a technological dimension.

Kant reconceived cognition as organised by the conditions of possible experience: perception is never purely passive but always structured by the categories of the understanding [27]. In algorithmic settings, this insight acquires a new form, as external systems of computational classification increasingly shape the material available to judgement. This is an analogy concerning the mediation of experience, rather than a claim that algorithms replace Kant's transcendental categories. Heidegger developed a related line of thought in his analysis of modern technology as enframing (Gestell), a mode of revealing in which beings appear primarily as resources to be ordered and optimised [29]. Algorithmic systems intensify enframing by converting human attention and cognition into quantifiable data streams, so that what appears to be information processing turns into an ontological restructuring of the way in which reality is disclosed.

Critical theory adds the dimension of power to this transformation. Truth cannot be separated from the structures of power that govern what can be seen and said, and in the digital environment these structures take the form of algorithms that determine what can be known, filtered, ranked and predicted [11]. Zuboff [9] describes surveillance capitalism as an economic order that converts human experience into behavioural data used to predict and influence behaviour, including cognition, attention and decisions. Han's analysis of psychopolitics [33] complements this account by showing that the dominant form of power operates less through direct coercion than through the imperatives of self-optimisation and self-discipline, which produce a paradoxical sense of freedom in which individuals govern their own behaviour. In media theory, Postman [34] argued that television established a communication framework that prioritised entertainment over sustained rational discourse, and the digital environment intensifies this tendency. Within the network society, these developments shape social relations, identity formation and cognitive structures through interconnected digital infrastructures, and knowledge is no longer controlled by a single source or institution but distributed across communication networks [10].

Contemporary interdisciplinary research extends these observations. Analyses of algorithmic authority show that platforms automate the ranking, filtering and visibility of information and shift epistemic authority from human interpretive judgement towards computational systems of selection and prediction [13]. The resulting governance of visibility is often algorithmically organised without corresponding opportunities for transparent public scrutiny or democratic negotiation. In attention economies, engagement metrics increasingly replace epistemic justification as the criterion of relevance, so that visibility and circulation can become more commercially significant than the truth of a claim, and disinformation operates through aesthetic and attentional frameworks rather than through argument [36]. The first finding can therefore be formulated as follows: the digital transformation moves the classical problem of illusion from the content of knowledge to the conditions of perception, and power is exercised through the organisation of these conditions.

3.2. COGNITION AND THE GENEALOGY OF TRUTH AND POWER

The second finding concerns the historical depth of the problem. The manipulation of perception precedes digital technologies, and the genealogy of cognition, truth and power reveals a continuous line of reflection on the vulnerability of knowledge. Since Machiavelli, political power has been conceived in terms of the management of appearances rather than the possession of truth. In The Prince, the ruler is advised to manage the appearance of virtue while retaining the capacity to depart from conventional virtues when political circumstances demand it [25]; politics becomes a theatre of perception, and judgement becomes the exercise of a spectator. Applied to digital communication, this logic implies that visibility is the principal political resource and that this resource is now governed by algorithmic ranking.

Bacon's doctrine of the idols in the New Organon [26] formulates a parallel epistemological warning. The idols are systemic distortions of human reason that arise from habits of perception, the use of language and the organisation of society; error is therefore not incidental but embedded in the way human beings interpret reality. Digital systems can be interpreted as a new type of idol, since they introduce technologically produced cognitive bias that shapes perception before the subject becomes aware of it.

For Plato, cognition is inseparable from the ethical orientation of the soul towards truth, and knowledge is transformational rather than merely representational, since it requires the reorientation of perception itself [22]. The prisoners of the cave are not merely deceived by particular images; they live in a condition of structural illusion. Read as a problem for contemporary communication, the Platonic model suggests that algorithmic infrastructures form a comparable structural illusion, with the difference that the regulation of perception now serves the purposes of control and power within a data-driven economy [9].

Al-Farabi's political philosophy [24] embeds cognition in the normative order of the virtuous city, where truth is not only epistemic but also civic: the stability of political order depends on the integrity of knowledge and its proper circulation within the community. Epistemic corruption is therefore not merely an individual error but a systemic condition that destabilises collective rationality. Translated into the conditions of an algorithmic information environment, this framework implies that the structures governing the circulation of information participate directly in the formation or deformation of collective rationality and civic coherence.

Arendt shifts the focus to the political consequences of epistemic instability. Factual truth is the fragile foundation of the common world and the condition of possibility of political judgement; when factual truth dissolves, the common world disintegrates into non-communicating realities [6, 7]. Digital systems intensify this vulnerability by producing epistemically isolated informational spheres in which individuals no longer share a horizon of appearance, so that political judgement loses its grounding in a common world.

Foucault examines the historically situated procedures and power relations through which statements acquire authority as truth and become recognised as knowledge [20]. Contemporary interpretations extend this account to digital governance [12]. In algorithmic environments, this logic becomes operational: the visibility, credibility and relevance of information are not neutral epistemic properties but outcomes of computational selection. Beer [11] shows that algorithmic systems function as forms of social and epistemic ordering by shaping what becomes visible, relevant and knowable. Cognition is accordingly not only internally generated but externally scaffolded by infrastructures that organise attention and mediate perception before reflection, so that mediation precedes interpretation. Cognition becomes infrastructural, embedded in networks of selection, amplification and suppression, and is restructured into a distributed process within technological systems of knowledge production rather than abolished.

Taken together, these positions can be assembled into a comparative conceptual genealogy. Plato formulates the ontological distortion of perception, Aristotle its dependence on habituation, Machiavelli and Bacon its political and structural sources, Al-Farabi its civic embeddedness, Arendt its vulnerability within a shared reality and Foucault its entanglement with power. Within this genealogy, cognition appears simultaneously as the object, the medium and the effect of regimes of truth. The genealogy supports the interpretation that the so-called new information crisis is an extreme case of a long-standing problem, namely the preservation of truth in a world where appearance is readily mistaken for reality, dialogue is liable to disintegrate and the common world is under threat.

3.3. INFORMATION WARFARE AND THE CONDITIONS OF PERCEPTION

The third finding concerns the transformation of information warfare. Its historical precedents are well known: the media controls of the Soviet Union in the 1930s and the activity of the Reich Ministry of Public Enlightenment and Propaganda show that the control of perception has long been a deliberate political practice. Contemporary information warfare differs from these precedents not in its intention but in its mode of operation. Research on misinformation raises concerns that algorithmically reinforced information ecosystems amplify interpretive divergence and erode the common factual basis required for deliberation [17], while studies of trust show that the digital transformation changes not only what people believe but also how trust is generated and maintained [5, 38]. The instability of truth is thus correlated with

the instability of trust.

Security research positions cognition as a domain of strategic vulnerability. Studies of the use of artificial intelligence by major powers in information warfare and influence operations show that perception is treated as a functional target rather than as a passive receptacle of data [3], and military discourse increasingly conceptualises information manoeuvre as a domain of operations in its own right [15]. The methods proposed for countering disinformation, accordingly, extend beyond the correction of individual messages [16]. Information warfare is therefore not limited to state-centred propaganda but is embedded in everyday digital life, where memetic content organises attention and interpretation [4]. Manipulation can therefore operate within the ordinary conditions of digitally mediated experience, rather than appearing only as an external interruption of otherwise independent cognition.

In Foucauldian terms, this change marks the emergence of a regime of visibility in which power operates less through suppression or censorship than through the organisation of the conditions of knowledge [12]. Visibility hierarchies, recommendation systems and data classification determine what may appear as knowledge in the first place; truth is not abolished but reallocated within technologically controlled informational architectures. The Platonic problem of the instability of appearance and reality is thereby intensified: illusion is no longer only rhetorical or cognitive but infrastructural, and interface design, algorithmic ranking and attention optimisation precondition perception. Research on artificial intelligence safety and on the malicious use of artificial intelligence identifies vulnerabilities, possible failures and opportunities for scaled misuse, which can make some forms of epistemic manipulation difficult to detect and counter in real time [18, 19].

These systems do not merely condition particular decisions; they can shape the range of decisions presented as available and thereby alter the conditions under which rational agency can be exercised. The techniques of behavioural influence developed in military computing, artificial intelligence governance and strategic communication are becoming interdependent, and manipulation can no longer be reduced to propaganda or misinformation in the traditional sense. The finding of this subsection is that information warfare in the digital age is a struggle over the possibility of perception, thought and understanding, conducted through the pre-structuring of the interpretive space rather than through the control of particular narratives. This conclusion moves the problem beyond technical and policy responses and calls for a philosophical account of truth, power and cognition.

3.4. COGNITIVE SECURITY AS A PHILOSOPHICAL CATEGORY

The fourth finding is the definition of cognitive security as a philosophical category. Under conditions in which cognition is continuously constructed, disrupted and restructured by algorithmic, social and technological systems, cognitive security can be defined as the sustained capacity to maintain reflective autonomy, interpretive independence and epistemic integrity. This meaning differs from the technical use of security vocabulary in cybersecurity and artificial intelligence safety, where it concerns the protection of digital infrastructure and the prevention of harmful system failures and attacks [18]. In the philosophical sense, cognitive security concerns the protection of judgement as such and is therefore a meta-epistemological issue: it addresses not only what is known but also how knowing is possible in environments saturated with mediated information. In digital ecosystems, the subject is not simply placed within flows of information but positioned within epistemically pre-ordered environments that format perception before reflection.

The concept has a distributive dimension, which the theory of epistemic injustice makes explicit. Fricker [14] shows that individuals can be wronged specifically in their capacity as knowers, through deficits of credibility and gaps in shared interpretive resources. Digital systems can reproduce such inequalities of interpretive authority and credibility when algorithmic visibility amplifies dominant voices and structurally marginalises alternative epistemologies. This is an application of Fricker's framework, rather than a finding reported in her book about contemporary platforms. Medina's epistemology of resistance [31] adds that epistemic agency depends on the capacity to recognise and resist such distortions. Cognitive security therefore includes not only the protection of individual reasoning but also the preservation of a fair distribution of epistemic agency.

From the Platonic perspective, cognitive security corresponds to the problem of the liberation of the soul from the cave, with the difference that the chains of perception are now located in computational mechanisms rather than imposed by rhetoric alone [22]. From the perspective of Al-Farabi [24], cognitive security has a civic dimension: if the stability of political order depends on the integrity of knowledge and its correspondence to truth, the algorithmic fragmentation of knowledge into individualised informational systems undermines shared rationality and collective intelligibility. Cognitive security is thus required to maintain not only individual reasoning but also the coherence of political life.

Arendt's analysis indicates that political judgement is endangered when the shared factual basis of public discussion disintegrates [6, 7]. Personalisation systems can contribute to epistemic isolation in which people no longer share a common world of facts but inhabit fragmented interpretive worlds; cognitive security is therefore a precondition for preserving a shared world in which judgement and responsibility remain meaningful. The Foucauldian concept of power–knowledge adds a structural aspect to this analysis. If digital regimes of truth are operationalised through algorithmic visibility,

classification and prediction, power works through epistemic configuration rather than through direct repression [12], and cognitive security must be understood as resistance to the structural preformation of thought.

Developments in artificial intelligence complicate this picture further. Philosophical accounts of trust distinguish interpersonal trust, reliance and the grounds on which AI systems may be considered trustworthy [38]. These distinctions matter where decision-making is distributed across human and computational components, because transparency alone does not settle the question of justified trust. The growing role of autonomous agents in shaping informational environments extends cognitive influence even where no human intention directs it, and the locus of cognition shifts towards a co-constructed socio-technical process. In this situation, cognitive security converges with the issues of cognitive liberty and the integrity of attentional and behavioural processes that are vulnerable to external modulation. Cognitive security is accordingly not a merely defensive stance but the active construction of enabling conditions under which depth of reflection, stability of attention and interpretive diversity can be preserved under information overload. In its most general form, it redefines security as the protection of thinking as a condition of human freedom.

3.5. PHILOSOPHICAL DIALOGUE AS EPISTEMIC RESISTANCE

The fifth finding identifies philosophical dialogue as a practice that responds to the conditions described above. Contemporary algorithmic communication is structured around immediacy, affective reinforcement and the optimisation of engagement rather than reflective understanding, which contrasts with the orientation towards mutual understanding that Habermas [30] placed at the centre of communicative action. Under these conditions, communication tends to rely on speed, repetition and confirmation loops that stabilise existing beliefs instead of exposing them to critical examination [41]. Philosophical dialogue introduces a different temporality of thought: it decelerates thinking and creates time for questioning and conceptual elaboration, thereby counteracting the industrial capture of attention that Stiegler [40] identified as a central threat to the formation of younger generations.

In the Platonic tradition, dialogue is connected with dialectic, understood not as an exchange of opinions but as an organised movement of thought towards truth, in which constant questioning gradually distinguishes appearance from reality and opinion from knowledge [22]. Dialogue in this sense is not a means of transferring information but a transformation of cognition, in which truth is sought through intellectual activity rather than received. Philosophical hermeneutics supports this view by treating dialogue as the medium in which understanding occurs across difference and in which the horizons of the interlocutors are transformed [21]. Decolonial scholarship adds that dialogue among different knowledge traditions can destabilise dominant epistemic hierarchies and create space for marginalised forms of rationality [32]. Dialogue is thus not only communicative but structurally transformative: it reconstitutes the conditions of intelligibility through ongoing interpretation and resists the stabilisation of knowledge within fixed informational structures.

The central function of dialogue in the algorithmic environment is the production of epistemic friction. Whereas engagement-optimised systems deepen repetition and limit contact with alternative perspectives, dialogue confronts beliefs with other beliefs and with reasons, which Medina [31] identifies as a condition of epistemic self-correction. This friction is constructive rather than obstructive, since it allows thought to detach itself from immediacy and to enter a space in which assumptions can be questioned, refined and repositioned. On a broader level, this relationship expresses the Hegelian dialectical movement, in which understanding emerges through the sustained examination and working through of contradiction, so that meaning is not given at the outset but developed through the mediation of differences [28]. Meaning in dialogue accordingly remains provisional and open to reconstruction.

Dialogue also has an ethical and political dimension. Where some voices are marginalised or structurally silenced, dialogue becomes a form of epistemic resistance that restores agency to excluded knowers and counteracts asymmetries of epistemic power [14, 31]. Understood as an inclusive practice of philosophical dialogue, it acquires a pluralist character as a space in which multiple rationalities coexist without being reduced to a single dominant epistemic regime, and epistemic diversity is treated as a philosophical necessity rather than as a problem to be eliminated. The finding of this subsection is that philosophical dialogue functions as a counter-power to algorithmic acceleration and the automation of epistemology: it resists the closure of meaning, interrupts the recursive reinforcement of belief and maintains the conditions under which thinking remains open, self-critical and directed towards truth rather than towards mere information.

3.6. EDUCATION AND COGNITIVE FORMATION

The sixth finding concerns education as the institutional site in which cognitive security is formed. In an environment of algorithmically mediated communication, fragmented attention and continuous informational saturation, the conception of education as the transfer of information becomes inadequate. Education acquires the function of forming cognitive resilience, interpretive autonomy and critical judgement, and the emphasis shifts from what is known to the ways in which knowing is structured and maintained. From a philosophical perspective, education is the formation of the interpretive subject within a complex epistemic milieu, in the sense of Bildung elaborated in philosophical hermeneutics [21]. Since

cognition is always situated within historically, culturally and technologically structured conditions of intelligibility, education teaches individuals not only to acquire knowledge but also to inquire into the conditions under which knowledge appears valid, meaningful or authoritative. This capacity can be described as epistemic sensitivity, that is, the ability to perceive how perception, interpretation and judgement are themselves shaped by external systems of influence. The transformation of information into knowledge, and of knowledge into understanding, depends on regulated practices of attention, interpretive reflection and meta-cognitive awareness; cognitive resilience is therefore a cultivated disposition rather than a spontaneous effect of exposure to information.

In the allegory of the cave, education is not the acquisition of new content but the reorientation of perception, in which the soul is turned away from shadows towards the conditions of truth [22]. This Platonic conception of paideia acquires renewed significance when algorithms structure the reality within which reflexive consciousness must be attained. Aristotle's account of habituation complements it by showing that judgement is formed through sustained practice [23]; under digital conditions, this implies that reflective habits of attention must be deliberately cultivated as a counterweight to habits formed by engagement-driven design. Stiegler's analysis of the capture of attention in younger generations explains the urgency of this task, since the formation of attention is a condition of the formation of thought itself [40].

Cognitive formation is also relational rather than purely individual. Learning environments are not passive backdrops but co-constitute the movement of thought, and meta-cognitive awareness, including the recognition of one's own interpretive biases and presuppositions, is a condition of intellectual responsibility. Dialogical forms of learning that bring different perspectives into contact can bring the epistemic friction described in Subsection 3.5 into educational practice. Collaborative learning that brings together participants with different cognitive and cultural backgrounds deepens analytical interaction and turns the formation of knowledge into a distributed social process rather than an achievement of isolated individuals. Cognitive resilience is therefore best conceptualised not only as an individual capacity but also as a property of educational structures and pedagogical frameworks that give priority to reflective and critical thinking, and higher education institutions bear particular responsibility for developing interpretive autonomy and protecting the conditions of critical reasoning. The finding of this subsection is that education is a principal institutional practice through which cognitive security is formed, as it trains the subject not only to know but also to reflect on knowing.

3.7. PHILOSOPHICAL DIALOGUE AS EPISTEMIC REPAIR

The seventh finding concerns the possibility of repair. Information warfare in digital environments produces a deep epistemic disturbance in which shared reality is divided into interpretive domains that are often incommensurable, and the shared world of facts, meanings and judgements is gradually replaced by individualised informational spaces curated by algorithms [8, 41]. Under these conditions, the problem is not only the weakening of consensus but also the instability of the horizon of intelligibility within which agreement and disagreement become possible. When truth is no longer anchored in common experience or shared rational standards, communication tends to devolve into parallel monologues instead of dialogue.

Epistemic repair cannot consist in the restoration of uniformity or enforced consensus, since uniformity would eliminate the plurality on which genuine understanding depends. Repair must instead be understood as the reconstruction of the relational conditions within which differences become intelligible to one another again. Philosophical dialogue performs this function as a reparative epistemic practice: it does not eliminate divergence but reorganises it within a space of shared questioning, justification and interpretive openness, so that disagreement becomes productive tension rather than disintegration. The movement through which understanding emerges from the encounter of different horizons [21] provides the model for this reconstruction. Epistemic repair is therefore not a return to homogeneity but the rebuilding of a relational architecture of understanding in which meaning is produced across differences rather than predetermined by narrow informational enclosures. The restoration of the epistemic world is accordingly a philosophical rather than a merely technical task, as it requires the reinstatement of dialogue as a form of collective thinking that can connect divided realities without erasing their differences.

3.8. SYNTHESIS OF THE CONCEPTUAL FINDINGS

The findings presented can be summarised in three propositions. First, the classical philosophical problems of truth, appearance, cognition and power persist in the algorithmically structured environment, but their mode of operation changes. Illusion moves from the representational to the infrastructural level, habituation is shaped by engagement-optimised design, the management of appearances is automated through ranking, and regimes of truth are operationalised through classification and prediction. The field in which truth and falsity can be distinguished is thereby pre-structured before judgement begins (Table 1).

Table 1. *Classical Epistemic Problems and Their Contemporary Algorithmic Analogues*

Thinker	Core Epistemic Problem	Contemporary Algorithmic Analogue	Relevance to Cognitive Security
Plato [22]	Illusion: mediated appearances taken for reality	Curated visibility and pre-shaped perception within algorithmic infrastructures [35]	Liberation of perception from infrastructural illusion
Aristotle [23]	Formation of perception and judgement through habituation	Habits of attention formed by engagement-optimised patterns of interaction [37]	Deliberate cultivation of reflective habits of attention
Al-Farabi [24]	Integrity and proper circulation of knowledge as a condition of civic order	Algorithmic fragmentation of collective rationality	Protection of shared rationality and civic coherence
Machiavelli [25]	Power as the management of appearances	Visibility as a political resource governed by algorithmic ranking	Critical awareness of how visibility is produced
Bacon [26]	Idols as structural distortions of reason	Technologically produced cognitive bias operating before awareness	Recognition of error as structural rather than incidental
Kant [27]	Categories organising the conditions of experience	External computational classification shaping the material available to judgement	Preservation of the subject's own capacity for judgement
Heidegger [29]	Enframing: beings revealed as resources	Attention and cognition converted into quantifiable data streams [9]	Resistance to the reduction of thought to data
Arendt [6, 7]	Factual truth as the foundation of a common world	Personalised informational realities that no longer coincide [8]	Preservation of a shared world as the condition of judgement
Foucault [12, 20]	Power–knowledge and regimes of truth	Visibility hierarchies, recommendation and classification [11]	Resistance to the structural preformation of thought

The comparison in Table 1 shows that the classical thinkers do not supply direct answers to the problems of the digital environment; rather, they provide categories that make it possible to identify which dimension of cognition is affected by each algorithmic operation. The comparison reveals a consistent pattern: in each case, a problem that classical philosophy located in rhetoric, habit, political performance or the structure of reason is now also located in technical infrastructure, which operates continuously and largely below the threshold of conscious deliberation.

Second, the category of cognitive security adds to existing approaches a normative concept that connects their diagnoses. Whereas the theories of surveillance, networks and algorithmic power describe how the environment is organised, and the theory of epistemic injustice describes how epistemic agency is distributed, cognitive security names what is at stake in all of them, namely the preservation of reflective judgement, interpretive autonomy and epistemic plurality as conditions of human freedom. Its distinctive emphasis is on judgement and the construction of enabling conditions, while technical security and AI safety remain relevant complementary concerns rather than wholly separate domains.

Third, philosophical dialogue occupies a dual position within this framework. As epistemic resistance, it introduces reflective friction into accelerated communication, interrupts the recursive reinforcement of belief and restores agency to marginalised knowers. As epistemic repair, it rebuilds the relational conditions under which divided interpretive worlds can become mutually intelligible without being forced into uniformity. Education is the institutional practice through which both functions are transmitted and sustained. Together, these propositions constitute the conceptual result of the study: cognitive security is the normative horizon, the genealogy of truth and power is its theoretical foundation, and philosophical dialogue and education are its principal practices.

4. DISCUSSION

The findings support and extend several established lines of research. They are consistent with Arendt's account that factual truth is the precondition of a common world [6, 7] and specify the mechanism through which this precondition is eroded in the digital environment: through the personalisation of informational worlds that may increasingly diverge, alongside the continuing effects of organised lying. In this respect, the findings agree with social-epistemological accounts of the fragility of truth under polarisation [8] and with analyses of echo chambers in social media [41], while adding a genealogical dimension that links this erosion to the classical problem of appearance.

The findings also extend the Foucauldian line of interpretation. Sahakyan et al. [12] describe the transition from disciplinary to algorithmic control, and Beer [11] demonstrates the social power of algorithms as a form of ordering. The present analysis shows that this ordering affects not only the conduct of the subject but also the pre-reflective conditions of perception, and it connects the Foucauldian concept of regimes of truth with the Platonic problem of illusion and the Baconian doctrine of structural distortion. This connection explains why algorithmic power is difficult to perceive: it operates at the level at which classical philosophy located the sources of error rather than at the level of explicit claims.

With respect to the theories of surveillance capitalism [9], psychopolitics [33] and the network society [10], the findings occupy a complementary position. Zuboff, Han and Castells explain the economic, political and structural organisation of

the digital environment, whereas the concept of cognitive security articulates the philosophical stake of this organisation, namely the autonomy of judgement. The analysis thus moves from the description of the techno-economic system to the question of which conditions of thinking must be preserved within it.

The findings extend the application of the theory of epistemic injustice by proposing that algorithmic visibility systems can produce credibility deficits and interpretive marginalisation through ranking and amplification alone, without requiring an overtly prejudiced act by a particular individual [14]. Medina's concept of epistemic resistance [31] acquires a specific application in this context, since philosophical dialogue appears as an institutionalisable form of epistemic friction that counters both individual and structural epistemic distortions.

The account of dialogue developed in Subsections 3.5 and 3.7 also clarifies its relationship to established theories of communication. It shares with Habermas [30] the conviction that communicative rationality is oriented towards understanding rather than strategic success, and it interprets engagement-optimised communication as a systematic distortion of this orientation. At the same time, the findings depart from the Habermasian emphasis on rational consensus: epistemic repair, as conceptualised here, aims not at agreement but at the restoration of mutual intelligibility among positions that may remain different. In this respect, the account is closer to Gadamer's [21] understanding of dialogue as a transformation of horizons and to the ecology of knowledges proposed by Santos [32], both of which treat plurality as a condition of understanding rather than as an obstacle to it. The educational dimension of the findings, in turn, gives concrete form to Stiegler's [40] thesis that the capture of attention threatens the transmission of the capacity to think across generations, since it identifies dialogical education as a practice through which attention can be formed anew.

At the same time, the empirical literature on disinformation calls for caution regarding the scale of algorithmic influence. Guess et al. [39] found that the sharing of fake news on Facebook during the 2016 campaign was relatively rare and concentrated among particular groups of users, and Benkler et al. [2] attribute the dynamics of disinformation in the American media ecosystem primarily to political and institutional asymmetries rather than to technology as such. These results do not refute the argument of the present study, which concerns the conditions of perception rather than the measurable effects of individual messages, but they indicate that cognitive security cannot be grounded in technological determinism. Human interpretation, institutional contexts and practices of civic resistance [4] remain active factors in the formation of judgement.

The theoretical novelty of the study lies in three results. First, the study develops a comparative conceptual genealogy of the relationship among cognition, truth and power from Plato, Aristotle, Al-Farabi, Machiavelli and Bacon through Kant, Heidegger, Arendt and Foucault to contemporary analyses of algorithmic governance and argues that the algorithmic environment transforms classical epistemic problems by shifting them from the representational to the infrastructural level. Second, it conceptualises cognitive security as a philosophical category, distinct from the technical vocabulary of cybersecurity [18, 19], which denotes the preservation of reflective judgement, interpretive autonomy and epistemic plurality and integrates individual, civic and distributive dimensions. Third, it justifies philosophical dialogue as a practice that performs the dual function of epistemic resistance and epistemic repair and links this practice with education as the institutional site of cognitive formation.

The proposed concept of cognitive security has definite boundaries. It may overemphasise the role of algorithms in shaping thought and underestimate the capacity of individuals and communities to interpret, appropriate and resist algorithmic mediation; perception is not wholly determined by infrastructure, and shifts in thought also depend on context and individual dispositions. The concept also relies on normative notions of truth and autonomy that remain contested, particularly from post-structuralist positions that question the possibility of a non-mediated standard of truth. For this reason, cognitive security should be understood not as a final state but as a regulative framework of critique that requires continuous refinement, including the clarification of how cognitive independence can be safeguarded without recourse to an idealised model of individual rationality and of which institutional structures can support plurality within distributed cognition.

The study has several limitations. First, it offers a theoretical genealogy of power and cognition but does not provide an empirical framework for measuring cognitive manipulation or epistemic vulnerability, which limits its immediate applicability for decision-makers and applied researchers. Second, the concept of cognitive security remains at an early stage of development and requires operational refinement before it can be applied in digital policy, cognitive science and interdisciplinary research. Third, the genealogy, despite its breadth, does not cover all relevant epistemological and critical traditions and draws predominantly on the Western philosophical canon, with Al-Farabi as the principal exception.

Future research may proceed in three directions. The first is the operationalisation of cognitive security through the methods of cognitive science, artificial intelligence ethics and computational social science, including the development of indicators of epistemic vulnerability and interpretive autonomy. The second is comparative research that integrates non-Western traditions of knowledge [32] into the genealogy of cognition and power. The third is the examination of dialogical educational practices as instruments for developing cognitive resilience, which would allow the normative claims of the present study to be tested in educational settings.

5. CONCLUSION

The study has conceptualised cognitive security through a historical-philosophical reconstruction of the relationships among cognition, truth, power and algorithmically mediated perception. Its main result is a reasoned conceptual argument that information warfare in the algorithmic environment is a struggle over the conditions of perception as well as over particular narratives: ranking, filtering and personalisation pre-structure the field in which truth and falsity can be distinguished before judgement begins, and the classical problem of illusion accordingly moves from the representational to the infrastructural level.

In response to the objective of the study, cognitive security is defined as a philosophical category denoting the preservation of reflective judgement, interpretive autonomy and epistemic plurality under conditions of algorithmic mediation. The theoretical contribution of the study consists in the genealogical connection of classical philosophical problems with the operations of the algorithmic environment, in the distinction between the philosophical and the technical meanings of cognitive security, and in the justification of philosophical dialogue as a practice of epistemic resistance and epistemic repair. The theoretical significance of these results lies in a normative concept that integrates the diagnoses offered by theories of surveillance, networks, algorithmic power and epistemic injustice around the question of which conditions of thinking must be preserved.

The practical implications concern several domains. For digital governance and policy, the framework provides grounds for treating algorithmic transparency and the preservation of informational diversity as conditions of public judgement rather than as technical options. For education and media literacy, it supports the inclusion of dialogical and reflective practices aimed at interpretive autonomy and epistemic sensitivity among explicit learning outcomes. For security policy, it indicates that responses to information warfare must address the conditions of perception and trust and not only the content of individual messages.

The principal limitation of the study is its conceptual character, as the framework has not yet been operationalised or empirically tested. Further research should develop indicators of cognitive security and examine dialogical educational practices as means of building cognitive resilience.

Conflicts of Interest

The authors declare no conflicts of interest.

Funding

This research received no external funding.

Data Availability

Data sharing is not applicable to this article, as no new data were created or analysed in this study.

Author Contributions

Liudmyla Krymets: Conceptualization, Methodology, Formal analysis, Writing – original draft, Writing – review & editing. Vitalii Chornyi: Conceptualization, Validation, Supervision, Writing – review & editing. Natalia Logvinenko: Investigation, Writing – original draft (education and cognitive formation), Writing – review & editing. Serhii Khovrych: Investigation (historical context of information warfare), Formal analysis, Writing – review & editing. Maksym Udod: Investigation (cognitive and psychological aspects of manipulation), Resources, Writing – review & editing. All authors have read and agreed to the submitted version of the manuscript.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this manuscript, the authors used Claude (Anthropic) and ChatGPT (OpenAI) for language editing, reference checking and manuscript formatting. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

REFERENCES

[1] Herman, E. S., & Chomsky, N. 1988. Manufacturing Consent: The Political Economy of the Mass Media. New York: Pantheon Books.

[2] Benkler, Y., Faris, R., & Roberts, H. 2018. Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics. New York: Oxford University Press.

[3] Hunter, L. Y., Albert, C. D., Rutland, J., Topping, K., & Hennigan, C. 2024. "Artificial intelligence and information warfare in major power states: how the US, China, and Russia are using artificial intelligence in their information warfare and influence operations." Defense & Security Analysis, 40(2), 235–269. https://doi.org/10.1080/14751798.2024.2321736

[4] Munk, T. 2025. "Digital defiance: Memetic warfare and civic resistance." European Journal on Criminal Policy and Research, 31(3), 501–528. https://doi.org/10.1007/s10610-025-09613-4

[5] Dowding, K., & Taylor, B. R. 2024. "Algorithmic decision-making, agency costs, and institution-based trust." Philosophy & Technology, 37(2), Article 68. https://doi.org/10.1007/s13347-024-00757-5

[6] Villa, D. R. 1996. Arendt and Heidegger: The Fate of the Political. Princeton: Princeton University Press.

[7] Arendt, H. 1967. "Truth and politics." The New Yorker, 25 February. https://www.newyorker.com/magazine/1967/02/25/truth-and-politics

[8] Kirmayer, L. J. 2024. "The fragility of truth: Social epistemology in a time of polarization and pandemic." Transcultural Psychiatry, 61(5), 701–713. https://doi.org/10.1177/13634615241299556

[9] Zuboff, S. 2019. The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. New York: PublicAffairs.

[10] Castells, M. 2009. The Rise of the Network Society: The Information Age: Economy, Society, and Culture. Vol. I (2nd ed.). Chichester: Wiley-Blackwell. https://doi.org/10.1002/9781444319514

[11] Beer, D. 2017. "The social power of algorithms." Information, Communication & Society, 20(1), 1–13. https://doi.org/10.1080/1369118X.2016.1216147

[12] Sahakyan, H., Gevorgyan, A., & Malkjyan, A. 2025. "From disciplinary societies to algorithmic control: Rethinking Foucault's human subject in the digital age." Philosophies, 10(4), 73. https://doi.org/10.3390/philosophies10040073

[13] Coşgun Solak, M. 2025. "The societal epistemology of artificial intelligence: The rise of algorithmic authority in the production of knowledge in the social sciences." İnsan ve Toplum Bilimleri Araştırmaları Dergisi, 14(4), 144–160. https://doi.org/10.15869/itobiad.1730658

[14] Fricker, M. 2007. Epistemic Injustice: Power and the Ethics of Knowing. New York: Oxford University Press.

[15] Pijpers, P. B. M. J., & Ducheine, P. A. L. 2023. "If You Have a Hammer...: Shaping the Armed Forces' Discourse on Information Maneuver." International Journal of Intelligence and CounterIntelligence, 36(4), 1164–1183. https://doi.org/10.1080/08850607.2023.2197560

[16] Dowse, A., & Bachmann, S. D. 2022. "Information warfare: methods to counter disinformation." Defense & Security Analysis, 38(4), 453–469. https://doi.org/10.1080/14751798.2022.2117285

[17] Lazer, D. M. J., Baum, M. A., Benkler, Y., et al. 2018. "The science of fake news." Science, 359(6380), 1094–1096. https://doi.org/10.1126/science.aao2998

[18] Yampolskiy, R. V., & Spellchecker, M. S. 2016. Artificial Intelligence Safety and Cybersecurity: A Timeline of AI Failures. arXiv preprint arXiv:1610.07997. https://doi.org/10.48550/arXiv.1610.07997

[19] Brundage, M., Avin, S., Clark, J., et al. 2018. The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation. arXiv preprint arXiv:1802.07228. https://doi.org/10.48550/arXiv.1802.07228

[20] Foucault, M. 1980. Power/Knowledge: Selected Interviews and Other Writings, 1972–1977. Ed. C. Gordon; trans. C. Gordon, L. Marshall, J. Mepham & K. Soper. New York: Pantheon Books.

[21] Gadamer, H.-G. 2004. Truth and Method (2nd rev. ed.). Trans. J. Weinsheimer & D. G. Marshall. New York: Continuum.

[22] Ferrari, G. R. F. (Ed.). 2007. The Cambridge Companion to Plato's Republic. Cambridge: Cambridge University Press.

[23] Aristotle. 2009. The Nicomachean Ethics. Trans. D. Ross, rev. L. Brown. Oxford: Oxford University Press.

[24] Al-Farabi. 1985. Al-Farabi on the Perfect State. Ed. and trans. R. Walzer. Oxford: Clarendon Press.

[25] Machiavelli, N. 1988. The Prince. Ed. Q. Skinner & R. Price; trans. R. Price. Cambridge: Cambridge University Press.

[26] Bacon, F. 2000. The New Organon. Ed. L. Jardine & M. Silverthorne; trans. M. Silverthorne. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139164030

[27] Kant, I. 1998. Critique of Pure Reason. Ed. and trans. P. Guyer & A. W. Wood. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511804649

[28] Hegel, G. W. F. 2018. The Phenomenology of Spirit. Trans. T. Pinkard. Cambridge: Cambridge University Press.

[29] Heidegger, M. 1977. "The question concerning technology." In: The Question Concerning Technology and Other Essays. Trans. W. Lovitt. New York: Harper & Row, 3–35.

[30] Habermas, J. 1984. The Theory of Communicative Action. Vol. 1: Reason and the Rationalization of Society. Trans. T. McCarthy. Boston: Beacon Press.

[31] Medina, J. 2013. The Epistemology of Resistance: Gender and Racial Oppression, Epistemic Injustice, and the Social Imagination. New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199929023.001.0001

[32] Santos, B. de S. 2014. Epistemologies of the South: Justice Against Epistemicide. Boulder: Paradigm Publishers.

[33] Han, B.-C. 2017. Psychopolitics: Neoliberalism and New Technologies of Power. Trans. E. Butler. London: Verso.

[34] Postman, N. 2005. Amusing Ourselves to Death: Public Discourse in the Age of Show Business. 20th anniversary ed. New York: Penguin Books.

[35] Longo, A. 2025. "How do social media algorithms appear? A phenomenological response to the black box metaphor." Minds and Machines, 35(2), Article 15. https://doi.org/10.1007/s11023-025-09716-1

[36] Ferreira, G. B. 2025. "The aesthetics of algorithmic disinformation: Dewey, critical theory, and the crisis of public experience." Journalism and Media, 6(4), 168. https://doi.org/10.3390/journalmedia6040168

[37] Foster, C. L. E. 2023. "Truth as social practice in a digital era: iteration as persuasion." AI & Society, 38(5), 2009–2023. https://doi.org/10.1007/s00146-021-01306-w

[38] Durán, J. M., & Pozzi, G. 2025. "Trust and trustworthiness in AI." Philosophy & Technology, 38(1), Article 16. https://doi.org/10.1007/s13347-025-00843-2

[39] Guess, A., Nagler, J., & Tucker, J. 2019. "Less than you think: Prevalence and predictors of fake news dissemination on Facebook." Science Advances, 5(1), eaau4586. https://doi.org/10.1126/sciadv.aau4586

[40] Stiegler, B. 2010. Taking Care of Youth and the Generations. Trans. S. Barker. Stanford: Stanford University Press.

[41] Sunstein, C. R. 2017. #Republic: Divided Democracy in the Age of Social Media. Princeton: Princeton University Press.

[42] Briegel, H. J., & Müller, T. 2025. "Phenomenology." In: Projective Simulation in Action. Synthese Library. Cham: Springer Nature Switzerland, 333–374. https://doi.org/10.1007/978-3-031-98119-7_6