

Article

Artificial Truth: Algorithmic Power, Epistemic Authority, and the Crisis of Democratic Knowledge

Rosario Palese 

Departamento de Sociología y Comunicación, Facultad de Ciencias Sociales, Universidad de Salamanca,
Campus Miguel de Unamuno, Avda. Francisco Tomás y Valiente s/n, 37007 Salamanca, Spain;
rosario.palese@usal.es

Abstract

This article examines how artificial intelligence and algorithmic systems are reconfiguring truth regimes in digital societies, introducing the concept of “Artificial Truth” to describe an emerging form of epistemic governance where knowledge production and validation become infrastructural functions of sociotechnical systems. The study develops an integrated theoretical framework combining Foucault’s notion of truth regimes, Bourdieu’s theory of symbolic capital and fields, and Actor-Network Theory’s constructivist approach. Through conceptual analysis, the article investigates how algorithmic recommendation systems, generative AI, and automated fact-checking operate as epistemic devices that actively shape what is recognized as credible, authoritative, and true in public discourse. The analysis reveals three fundamental transformations: (1) the restructuring of trust economies, with epistemic authority shifting from institutional expertise to platform-native capital based on engagement metrics and affective proximity; (2) the emergence of generative AI as an epistemic actor producing “synthetic truth” through linguistic fluency rather than propositional understanding; (3) the institutionalization of computational veridiction in algorithmic fact-checking systems that translate situated epistemic judgments into probabilistic classifications presented as neutral. These dynamics configure a regime where truth is evaluated less by correspondence with reality and more by computational plausibility and platform integration. The article’s primary contribution lies in providing a unified theoretical framework for understanding contemporary transformations of epistemic authority, moving beyond disinformation studies to analyze AI as an epistemic actor. By integrating classical sociological perspectives with Science and Technology Studies, it conceptualizes algorithmic systems as epistemic infrastructures that embody specific power relations, restructure symbolic capital economies, and distribute epistemic authority asymmetrically, with profound implications for democratic knowledge, citizen epistemic agency, and public sphere pluralism.



Academic Editors: Felice Addeo and
Gabriella Punziano

Received: 23 February 2026

Revised: 17 March 2026

Accepted: 18 March 2026

Published: 23 March 2026

Copyright: © 2026 by the author.
Licensee MDPI, Basel, Switzerland.
This article is an open access article
distributed under the terms and
conditions of the [Creative Commons
Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: algorithmic epistemic authority; artificial truth; generative AI epistemic agency; computational veridiction; platform epistemic governance

1. Introduction

In the well-known passage from 1984, George Orwell describes a form of power capable of imposing that “two plus two equals five,” intervening not only on bodies but on the very symbolic structures of reality [1]. In an apparently opposite direction, Philip K. Dick, in the essay *How to Build a Universe That Doesn’t Fall Apart Two Days Later*, argued that reality retains its consistency even when the subjective conditions of

belief in it cease, continuing to exist independently of the beliefs projected onto it [2]. In the contemporary information ecosystem, the relationship between persistence and correspondence appears increasingly unstable. What endures in circulation does not necessarily align with empirical accuracy, but with what is reiterated, optimized, rendered visible, and normalized through computational systems of selection and prediction. The contrast between Orwell and Dick can be interpreted as mapping onto the two analytical poles of the framework developed here. Orwell foregrounds the imposition of truth through institutional power; Dick highlights the fragility of reality when collective belief becomes destabilized. The concept of Artificial Truth, as elaborated in this article, is positioned between these poles. It does not denote truth as pure imposition, nor as simple correspondence, but refers to regimes in which truth claims are mediated by computational procedures that combine authority with limited transparency and uneven accountability.

Within this configuration, truth is not merely suppressed or denied, but often simulated, reformulated, and performed in ways that generate reality effects. The widely used term “fake news” requires conceptual clarification in this respect. As Aïmeur et al. [3] demonstrate in their systematic review, the expression conflates analytically distinct phenomena, including misinformation, disinformation, and malinformation, which differ in terms of intentionality, production processes, and social consequences. For this reason, the present analysis moves beyond the folk category of “fake news” toward a framework centered on epistemic authority and regimes of truth.

The central theoretical question guiding this article can be formulated as follows: how can transformations in contemporary truth regimes be conceptualized so as to account simultaneously for the power dimension of computational mediation, the reconfiguration of symbolic capital, and the distribution of epistemic agency across human and non-human actors?

This shift reorients the focus from disinformation as isolated content to epistemic authority as a technologically mediated relation of power. The issue is not limited to whether particular statements are true or false, but concerns how specific claims acquire credibility, visibility, and performative force within digital publics. Rather than diagnosing a generalized “crisis of truth,” the analysis seeks to understand how algorithmic systems participate in instituting modalities through which truth is constituted, legitimized, and contested in ways that differ from, though remain historically connected to, those of institutions such as universities, journalism, and science. The aim is to articulate an integrated theoretical framework for examining this configuration, here conceptualized as Artificial Truth, characterized by the increasing delegation of validation processes to privatized algorithmic infrastructures whose operations are only partially transparent.

The specific objectives are: to integrate three classical theoretical perspectives (Foucault, Bourdieu, Latour) in order to conceptualize algorithmic systems as epistemic devices that operate simultaneously as institutions of power, as fields for the accumulation of symbolic capital, and as heterogeneous networks of human and non-human actors; to document three fundamental transformations: (a) the restructuring of trust economies, marked by the shift from institutional expertise to platform-based epistemic capital; (b) the emergence of generative artificial intelligence as an epistemic actor producing “synthetic truth”; (c) the institutionalization of computational veridiction within algorithmic fact-checking systems; and to analyze the democratic implications of this regime, with particular attention to the privatization of epistemic authority, the erosion of epistemic citizenship, and forms of algorithmic epistemic injustice.

The contribution is theoretical and conceptual in orientation. It does not present an empirical case study of specific platforms or systems; rather, it develops analytical

categories intended to clarify ongoing transformations and to provide a framework for future empirical inquiry.

The distinctiveness of the approach can be summarized in four elements. First, it brings into dialogue classical theoretical perspectives that are often treated separately. By integrating insights on the constitutive power of discourse (Foucault), the dynamics of capital accumulation and symbolic conversion (Bourdieu), and the material agency of technical artifacts (Latour), the framework seeks to conceptualize algorithmic systems as socio-technical assemblages. Second, the article introduces the concept of Artificial Truth as an analytical category that extends beyond the vocabulary of disinformation. Third, the paper treats generative artificial intelligence as an epistemic actor rather than solely as a neutral tool or content amplifier. Fourth, the analysis foregrounds the democratic governance of knowledge, situating algorithmic mediation within broader debates on pluralism, deliberation, and epistemic autonomy.

The article is structured into an introduction, five analytical sections, and a conclusion. Section 2 develops the integrated theoretical framework. Section 3 examines transformations in trust economies. Section 4 analyzes generative artificial intelligence as an epistemic actor. Section 5 discusses algorithmic fact-checking systems as forms of computational veridiction. Section 6 synthesizes these developments under the concept of Artificial Truth and examines their implications for democracy. The concluding section summarizes the theoretical contributions, addresses limitations, and outlines directions for future research.

2. Truth Regimes, Epistemic Authority, and Algorithmic Mediation

The question of the social production of what is recognized as “true” runs through the entire tradition of the sociology of knowledge, from Weberian reflections on rationality to more recent analyses of post-truth. What distinguishes the digital turn is not the mere acceleration of information flows, but a profound transformation of the infrastructures through which truth is constituted, legitimized, and contested. Algorithmic systems do not simply mediate access to information; they actively intervene in the production of “truth effects” [4], reconfiguring criteria of credibility and redistributing epistemic authority through opaque computational logics. Understanding this transformation requires a conceptual apparatus capable of articulating power, symbolic capital, and the material agency of technical artifacts.

2.1. Foucault and Regimes of Veridiction

The Foucauldian notion of “regime of truth” offers a crucial starting point for analyzing epistemic authority in algorithmic societies. For Foucault, a regime of truth does not coincide with a set of true propositions. Rather, it consists of “the ensemble of rules according to which the true and the false are distinguished and specific effects of power are attached to the true” [5]. Truth does not emerge from a simple correspondence between statement and reality. Institutions, verification procedures, and power relations actively define what counts as true within a given historical context.

When applied to digital environments, this concept allows us to examine how algorithmic systems establish specific veridictional logics. Algorithms do not merely process information; they determine which claims become visible, which sources appear credible, and which narratives circulate widely [4]. Facebook’s News Feed algorithm, for example, does not simply mirror users’ preferences or measure informational quality. It selects, ranks, and amplifies content, thereby structuring public reality.

Foucault’s notion of the *dispositif* clarifies the heterogeneous composition of these arrangements. A *dispositif* is “a thoroughly heterogeneous ensemble” that brings together discourses, institutions, architectural forms, regulatory decisions, laws, administrative

measures, and scientific statements [6]. Content moderation systems exemplify this structure. They combine machine learning models, community guidelines, policy frameworks, outsourced labor, user reporting infrastructures, and national regulations. Through these interacting components, platforms actively organize regimes of truth [7].

The concept of algorithmic governmentality extends this analysis to contemporary computational environments. Whereas disciplinary power operated primarily by inducing subjects to internalize norms within bounded institutions, algorithmic governance acts differently. Platforms modulate behavior in real time, profile users predictively, and continuously collect and recombine data [8]. Recommendation systems rarely impose explicit prohibitions. Instead, they orient attention, reward certain forms of expression, and quietly marginalize alternatives, redistributing epistemic agency and co-producing forms of knowledge and experience [9].

2.2. Symbolic Capital and Algorithmic Distinction: A Bourdieusian Perspective

If Foucauldian analysis clarifies how regimes of truth operate, Pierre Bourdieu offers the tools to analyze epistemic authority as relational and institutional power. In Bourdieu's capital theory, authority does not stem from intrinsic qualities. It depends on collective recognition. Actors acquire symbolic capital when others acknowledge their legitimacy, grant them trust, and treat their position as natural [10]. Symbolic capital therefore condenses social esteem into durable authority by converting other forms of capital into publicly recognized legitimacy.

Digital platforms do not simply host these dynamics; they reorganize them. In digital environments, platforms reshape how actors accumulate and convert symbolic capital. Traditional epistemic authority relied on credentials, institutional affiliation, and peer validation. Platforms increasingly reward different forms of capital. They privilege follower counts, engagement metrics, aesthetic appeal, and what Ling et al. [11] describe as "algorithmic meta-capital." Actors who understand platform logics and strategically adapt their content to ranking and visibility systems accumulate this new form of capital more effectively than those who rely solely on institutional credentials.

The concept of field helps clarify how these shifts alter spaces of cultural production. For Bourdieu, a field is a relatively autonomous social space structured by hierarchical positions and struggles over specific forms of capital [12]. Platformization does not merely add a new field. Platforms intervene across fields, mediating between journalism, academia, politics, and cultural production, redefining which forms of capital circulate effectively. A crucial feature of algorithmic capital is its volatility: unlike traditional cultural capital, algorithmic capital depends on mutable technical and regulatory configurations, forcing actors to continuously adapt and reoptimize their practices.

2.3. Distributed Agency and Sociotechnical Assemblages: Actor-Network Theory

Bruno Latour's Actor-Network Theory introduces a third analytical dimension by foregrounding the agency of non-human actors in processes of knowledge production. In ANT, knowledge does not reside in individual subjects, nor does it simply reflect social structures. It emerges from networks of associations that link human and non-human actors. By rejecting a strict ontological divide between subjects and objects, ANT treats the elements of a network symmetrically and examines how they collectively produce effects [13].

When applied to algorithmic systems, this perspective shifts how we interpret computational infrastructures. Algorithms do not merely execute human intentions. They act as mediators. They classify, rank, predict, and recommend, and in doing so they intervene in social relations and decision-making processes [14,15]. Seaver [16,17] describes the divergences between design intentions and situated outputs as "betrayals":

gaps that accumulate as code moves from design to deployment. Crawford [18], in *Atlas of AI*, reconstructs these configurations as global supply chains, showing how corporations render algorithmic authority opaque while presenting outputs as neutral and self-evident. However, ANT's commitment to symmetry also introduces a risk. By treating actors within a network on the same analytical plane, the approach can understate structural power asymmetries. For this reason, ANT must be integrated with a Foucauldian analysis of power. Only by combining distributed agency with structural asymmetry can we account for how specific institutions exercise sustained control over the configuration and governance of algorithmic networks.

2.4. Toward an Integrated Framework: Algorithmic Systems as Epistemic Devices

Table 1 synthesizes the integrated framework developed in this section. It does not simply juxtapose three classical perspectives. It shows how they converge in explaining how algorithmic systems operate as epistemic devices. Each perspective clarifies a different dimension of algorithmic power. A Foucauldian lens reveals how algorithms establish regimes of truth and enact forms of governmentality. A Bourdieusian approach explains how platforms reorganize symbolic capital and recalibrate the hierarchies that structure fields of knowledge production. Actor-Network Theory demonstrates how technical artifacts participate materially in these processes by mediating relations, stabilizing associations, and redistributing agency within sociotechnical networks.

Table 1. Integrated framework for the analysis of algorithmic systems as epistemic devices.

Theoretical Perspective	Key Concept	Algorithmic Operation	Epistemic Effect	Resulting Transformation
Foucault: Regimes of truth and governmentality	Regime of truth as the set of rules that distinguish true/false and attach effects of power to truth	Algorithms institute veridictional logics by determining the visibility, credibility, and authoritativeness of content through heterogeneous dispositifs (machine learning, policy, moderation)	Production of “truth effects”: structuring of public reality through the selection, amplification, and marginalization of narratives	Algorithmic governmentality: real-time behavioral modulation, attention orientation, and optimization of visibility rather than explicit censorship
Bourdieu: Symbolic capital and fields	Symbolic capital as credit and authority derived from collective recognition of legitimacy	Algorithms reconfigure the bases of capital accumulation, shifting from institutional credentials to “algorithmic meta-capital” (capacity to optimize for visibility)	Redefinition of epistemic authority: from formal expertise toward popularity, emotional resonance, perceived sincerity, and affective performativity	Restructuring of fields: platforms operate as meta-fields mediating across different fields, imposing algorithmic metrics and commercial logics; structural volatility of algorithmic capital
Latour: Actor-Network Theory and distributed agency	Knowledge as an emergent product of heterogeneous networks of associations between human and non-human actors	Algorithms act as agentic actants: they inscribe patterns, classify, predict, and recommend, producing systematic “betrayals” along chains of translation	Black-boxing of algorithmic authority: naturalization of computational outputs through the opacification of internal mechanisms	Extended sociotechnical assemblages: global supply chains including data, architectures, interfaces, infrastructures, and policy; power asymmetries in the inscription of interests into technical configurations
Integrated framework (Synthesis)	Algorithms as epistemic devices: sociotechnical assemblages that institutionalize regimes of truth while reordering symbolic capital economies	Simultaneous operation across three registers: (1) non-human epistemic agency; (2) operationalization of power/knowledge regimes; (3) restructuring of symbolic capital	Production of “artificial truth”: truths constituted through technical dispositifs with epistemic logics radically different from those of modern knowledge institutions	Reconfiguration of epistemic infrastructures: criteria of authority shift from objectivity to credibility/sincerity, from institutional distance to affective proximity, and from transparent verifiability to computational opacity

Note. Author's synthesis.

When combined, these approaches allow us to conceptualize algorithms as more than instruments. Platforms deploy them as infrastructures that (1) exercise non-human agency, (2) operationalize specific power/knowledge regimes, and (3) reorder the economies of symbolic capital across fields. This convergence clarifies the transformation currently unfolding: algorithmic infrastructures reorganize the procedures through which societies validate truth claims, redefining how credibility is assigned, how authority is recognized, and how knowledge is stabilized.

3. Algorithmic Trust and Epistemic Capital in Digital Platforms: A Sociotechnical Transformation of Authority

3.1. *Trust, Expertise, and Epistemic Authority: Theoretical Foundations for Digital Analysis*

This section develops the Bourdieusian dimension of the integrated framework introduced in Section 2, examining how digital platforms reorganize the economies of symbolic capital through which epistemic authority is produced, legitimized, and circulated. Trust structures how contemporary societies validate and distribute knowledge. Sociologists have long analyzed trust as a mechanism that reduces complexity [19] and as a condition that sustains expert systems organizing modern life [20]. Luhmann's distinction between personal trust, grounded in direct interaction, and systemic trust, oriented toward abstract institutions and stabilized expectations, remains particularly useful for examining how digital environments reshape epistemic authority.

Before the rise of platforms, institutions anchored credibility. Journalism derived authority from professional norms and editorial routines; science relied on peer review and standardized methodologies [21]. These arrangements did not eliminate the performative dimension of credibility. Scholarship on journalistic authority has long established that credibility was never reducible to accuracy or institutional affiliation: it was produced through stylistic conventions, rhetorical positioning, and audience perceptions of competence and trustworthiness that predated digital mediation by decades [22–24]. Digital platforms do not invent credibility from scratch. They reorganize the material and technical arrangements through which actors produce and circulate it. As Neuberger et al. argue, pre-digital knowledge orders were anchored in stable institutional structures: journalism and science functioned as recognized epistemic authorities, legitimized through professional roles, hierarchical certification procedures, and bounded knowledge contexts that constrained both participation and validation [25]. These mechanisms concentrated authority and constrained participation. They also depended on forms of “epistemic vigilance” [26]: collective gatekeeping practices embedded in limited publication channels and selective certification procedures.

3.2. *The Dual Transformation: Institutional Disintermediation and Algorithmic Re-Intermediation*

Digital platforms reshape epistemic authority through a dual process. First, they weaken the exclusive mediating role of traditional institutions. Second, they insert themselves as new intermediaries governed by proprietary and computational logics [27,28]. Gatekeeping does not disappear. It moves. As some institutional filters lose centrality, platforms establish new systems of selection, ranking, and visibility.

When platforms lower barriers to publication, individuals can circulate content, assemble audiences, and claim credibility without passing through established institutional channels [29,30]. This shift reduces journalism's control over public truth claims [31] and intensifies public challenges to institutional science [32]. In this context, scholars often invoke the notion of “echo chambers” to describe informational enclosure. Yet empirical research complicates this image. Ross Arguedas et al. [33], reviewing cross-national evidence, show that tightly sealed and homogeneous echo chambers rarely materialize in the uniform

way popular discourse suggests. Rather than constructing impermeable informational bubbles, algorithms more frequently skew attention asymmetrically, shaping exposure without fully isolating users from competing claims. At the same time, platforms do more than host or transmit content. They actively re-intermediate communication. Through ranking, recommendation, and curation systems, they define which statements circulate widely and which remain marginal [27].

3.3. New Criteria of Legitimation: From Expertise to Engagement, from Distance to Affective Proximity

Digital platforms alter not only who mediates knowledge, but also how audiences assign credibility. Recent scholarship highlights three interconnected shifts: expertise gives way to popularity and engagement; institutional certification loses ground to platform-native metrics; professional distance yields to affective proximity and perceived sincerity [34,35]. In earlier institutional settings, expertise functioned as the primary currency of authority. Platform ecosystems operate differently. Algorithms amplify content that generates interaction. As a result, highly engaged posts often circulate more widely than institutionally certified knowledge. Engagement becomes a practical indicator of reliability, not because actors normatively redefine truth, but because ranking systems reward measurable interaction [36,37]. At the same time, platforms recalibrate the value of distance. Journalistic detachment and rhetorics of objectivity no longer guarantee trust. Instead, creators cultivate performed authenticity and personal positioning, what Abidin [38] describes as “calibrated amateurism.” Audiences may interpret institutional distance as opacity or elitism, while explicit positionality signals sincerity [21,39]. Platform affordances reinforce this pattern, encouraging personalization, visible metrics of approval, and continuous feedback loops that structure epistemic environments where proximity, responsiveness, and emotional resonance shape what counts as credible.

3.4. Algorithmic Trust as Computational Epistemic Delegation

The most consequential shift concerns what can be described as black-box trust. Users rely on systems whose internal reasoning they cannot inspect [40]. Instead of evaluating arguments or sources directly, they accept outputs generated by opaque computational procedures. In this sense, platforms extend Luhmann’s notion of systemic trust. Complexity is no longer reduced primarily through recognizable institutions. It is reduced when users delegate cognitive labor to algorithms. The concept of algorithmic trust captures this specific form of epistemic delegation: users place confidence in systems they do not understand, accepting rankings, recommendations, and summaries despite opacity, and sometimes because of it. When users cannot directly assess how a system reaches its conclusions, they may default to acceptance rather than scrutiny [40]. This shift reassigns epistemic responsibility. Recommendation systems do not merely transmit information; they actively participate in belief formation. When users cannot evaluate the processes that shape what they see, their capacity for autonomous epistemic judgment weakens accordingly [41].

The consequence is a redistribution of symbolic capital within the epistemic field, of the kind described in Section 2. Recommendation systems do not merely transmit information; they actively structure which claims accrue credibility and which remain marginal. When users cannot evaluate the processes that determine what they encounter, their capacity for autonomous epistemic judgment weakens accordingly [41]. Algorithmic trust thus operates as the mechanism through which platforms exercise symbolic power: distributing epistemic authority unevenly, privileging certain sources, demoting others, and structuring exposure asymmetrically [42]. Over time, repeated interaction

with these systems normalizes their authority. Users habituate themselves to algorithmic mediation, and what once appeared as a contestable technical arrangement becomes an infrastructural background [37]. This habituation prepares the ground for deeper forms of delegation, including the uncritical acceptance of statements generated entirely by artificial systems.

4. Generative Artificial Intelligence as an Epistemic Infrastructure: Agency, Performativity, and Synthetic Truth

4.1. *Generative AI as an Epistemic Actor: Redistribution of Agency and the Performativity of Knowledge*

This section develops the Actor-Network Theory dimension of the integrated framework introduced in Section 2, examining generative AI not as a tool that assists human cognition but as a non-human actant that actively mediates epistemic practices within heterogeneous sociotechnical networks. Science and Technology Studies no longer treat technical artifacts as neutral instruments. Instead, they examine how heterogeneous networks of human and non-human actors produce agency [13,43]. When this perspective is applied to large language models, the interpretive shift becomes clear. LLMs do not simply assist cognition. They intervene in epistemic practices. They generate explanations, summarize evidence, and articulate claims in ways that shape what users recognize as credible, adequate, or true [44,45].

Generative AI does not rupture earlier epistemic arrangements; it deepens the entanglement of human judgment and machine output that posthumanist scholarship had already identified as a structural tendency of information-processing systems [46,47]. Generative systems exert epistemic force through form rather than understanding. They produce statements that resemble expert discourse: specialized vocabulary, technical structure, and argumentative coherence [48,49]. Yet they lack propositional grounding. They predict tokens probabilistically, reproducing linguistic patterns without referential access to the world. Bender et al. [49] characterize such systems as “stochastic parrots.” Despite this absence of semantic grounding, users frequently perceive generative outputs as credible. Under certain conditions, fluency and structural coherence sustain high levels of trust [50,51].

The performativity of generative authority operates through at least three mechanisms. First, fluency functions as a heuristic: users equate coherence with knowledge, even when they know the system relies on computation rather than comprehension [50]. Second, generative systems often adopt assertive tones and limit hedging, amplifying perceptions of competence [52]. Third, conversational interfaces simulate dialogue, activating interpersonal scripts and encouraging users to attribute intentionality or expertise to systems that do not possess them [45,53]. These dynamics reshape epistemic agency: as users adapt to generative infrastructures, machine-generated responses become reference points within everyday reasoning.

4.2. *Synthetic Truth as a Regime of Veridiction: From Correspondence to Statistical Plausibility*

Generative AI alters how societies evaluate truth claims. Rather than retrieving stable bodies of knowledge, large language models assemble heterogeneous signals and generate fluent responses that appear evidentially grounded [44]. In this sense, they operationalize truth, transforming statistical regularities extracted from training data into linguistically plausible statements. As a result, plausibility increasingly competes with correspondence as a criterion of validity.

The concept of synthetic truth captures this shift. It designates an epistemic regime structured by at least four interrelated features. First, generative systems produce state-

ments probabilistically rather than referentially. When such systems generate false claims, the so-called hallucinations do not represent random malfunctions. They follow from architectures optimized for plausibility and coherence rather than strict correspondence [52,54]. Second, textual fluency functions as a marker of authority. Markowitz [51] shows that audiences often rate AI-generated scientific summaries as more credible than human-authored ones, even when experts judge them less cognitively rigorous. Third, generative infrastructures cultivate procedural authority: users increasingly trust outputs because they emerge from technically sophisticated systems rather than from identifiable experts [13,37]. Fourth, generative systems perform knowledge without understanding it. Zhou et al. [55] demonstrate that AI-generated false claims can surpass human-generated ones in persuasiveness when they include descriptive detail, simulated personal tone, and calibrated hedging.

4.3. Automation of Epistemic Practices and the Normalization of Machine Authority

As generative systems become embedded in everyday information environments, they reshape how individuals and institutions produce, evaluate, and circulate knowledge. This shift unfolds along three interconnected dimensions.

First, generative systems automate interpretation and explanation. Large language models now perform tasks that once required expert judgment. While this automation increases efficiency, it also compresses elements central to epistemic rigor. Shanahan [56] observes that these systems frequently produce decontextualized summaries that conceal the probabilistic and fallible character of scientific knowledge.

Second, platforms delegate verification and validation to computational systems. Automated fact-checking tools treat truth as a classificatory output, encoding particular epistemological assumptions directly into training data and system architecture. Although platforms present these tools as neutral solutions, they reshape editorial practice in ways that recalibrate how institutions validate truth claims. The structural implications of this delegation are analyzed in detail in Section 5, which examines automated fact-checking as a form of computational veridiction.

Third, algorithmic intermediation becomes normalized. Generative search engines increasingly mediate access to knowledge by selecting, ranking, and aggregating information [57]. Audits reveal geographic biases, dependencies on commercially dominant sources, and distortions shaped by query framing. As users repeatedly rely on generative outputs, these tools become routine reference points rather than provisional syntheses. Over time, this habituation stabilizes machine authority, and computational logics help structure digitally organized public spheres [58].

4.4. Sociotechnical Conditions of Acceptability: Algorithmic Trust, Automation Bias, and Institutional Legitimation

Synthetic truth does not operate automatically. It becomes effective when specific material, cognitive, and institutional conditions make machine authority appear legitimate. Four mechanisms are particularly salient.

First, interface design shapes how users calibrate trust. Heersmink et al. [45] identify a recurring gap between phenomenological transparency, the fluency and clarity of interaction, and informational transparency, insight into underlying processes. Systems often feel intelligible while remaining opaque. Automation bias reinforces this dynamic: even trained professionals anchor their judgments to AI outputs and continue relying on them despite warning signals [59,60].

Second, repeated reliance on generative systems shifts cognitive practice. Users offload tasks such as synthesis, comparison, and preliminary verification to machines. Cabitza et al. [61] document how, in clinical contexts, prolonged delegation reshapes

professional judgment and attentional habits. This process does more than redistribute tasks. It gradually redefines what counts as adequate reasoning, aligning epistemic standards with the outputs and constraints of generative systems, a dynamic consistent with what Zuboff [62] identifies as a broader structural tendency of digital capitalism to migrate evaluative responsibility toward computational infrastructures.

Third, infrastructures normalize themselves. As Star and Ruhleder [63] note, infrastructures become invisible when they function smoothly. Generative systems operate precisely as the epistemic infrastructures described in Section 2: they filter sources, prioritize certain forms of evidence, and embed normative assumptions into routine interactions without making these operations visible to users. Because users rarely see these filtering processes, they interpret outputs as neutral syntheses rather than as situated constructions. Over time, mediation recedes from view and authority appears natural.

Fourth, power asymmetries shape who benefits and who is disadvantaged. Kay et al. [64] extend the concept of epistemic injustice to generative systems, identifying testimonial, hermeneutical, and access-related harms. Training data reflect historical hierarchies. Procedural opacity makes it difficult to contest automated decisions that affect credibility and participation. Generative systems do not invent these asymmetries, but they can stabilize and amplify them within digital epistemic environments.

5. Generative Artificial Intelligence and Synthetic Truth: The Configuration of a New Epistemic Regime

5.1. Introduction: Algorithmic Verification Systems as Epistemic Infrastructures

This section develops the Foucauldian dimension of the integrated framework introduced in Section 2, examining automated fact-checking systems as contemporary dispositifs of veridiction: arrangements through which the distinction between true and false is operationalized, standardized, and rendered authoritative within digital environments. In continuity with the dynamics outlined in the previous section, algorithmic systems for the detection of disinformation cannot be reduced to mere tools of textual classification, but must be understood as epistemic infrastructures [65] that reorganize the modes through which truth is produced, validated, and governed in digital environments. As already shown, platforms exercise epistemic power through algorithmic dispositifs that define what can be known, by whom, and according to which protocols [27]. Automated fact-checking systems push this logic to its extreme, translating historically situated and contestable epistemic judgments into standardized, scalable computational operations presented as neutral [66]. From this perspective, the term computational veridiction designates a regime of truth production in which factual assessment takes the form of probabilistic outputs generated by machine learning models, evidence is progressively reduced to statistical correlation within training data, and epistemic authority is delegated to neural classifiers optimized on performance metrics [67]. As Shin [66] observes, these systems do not merely identify a pre-existing truth, but actively contribute to constituting it through technical choices that embed controversial epistemological assumptions, cultural biases, and asymmetric power relations.

Automated fact-checking pipelines thus operate as genuine epistemic supply chains [68], decomposing the verification process into modular tasks that include the identification of checkable claims, evidence retrieval, stance classification, and veracity prediction. Each of these stages incorporates decisions concerning what deserves to be verified, which sources count as evidence, how disagreement is treated, and which criteria are used to evaluate performance. Decisions presented as purely technical thereby end up encoding political judgments regarding the distribution of epistemic authority,

the forms of knowledge deemed legitimate, and the communities potentially excluded from the definition of ground truth [69].

5.2. Computational Veridiction: From Argumentative Truth to Probabilistic Calculation

Within this reconfiguration of epistemic authority, the transformation introduced by algorithmic verification manifests first and foremost in the reduction of truth to computational output. Critical scholarship makes it possible to reconstruct how disinformation detection models tend to operationalize veracity through three analytically distinguishable modes: prediction, in the form of continuous credibility scores; classification, via discrete truth labels; and ranking, understood as the relative ordering of the reliability of content or sources [67,70].

Predictive models treat truth as a latent property inferable from observable features, such as linguistic patterns, source characteristics, or diffusion network structures [66]. This approach presupposes that truth leaves systematically detectable traces in data, an assumption that may prove effective for elementary forms of misinformation but becomes fragile in the face of sophisticated disinformation practices capable of mimicking the patterns of truthful content [71]. The confidence scores generated by such models produce an illusion of precision, since the probabilistic value does not express the “actual” truth of a claim, but rather the correlation between its observable features and the labels present in the training set [72].

Classification systems, by contrast, discretize truth into rigid categories, imposing sharp boundaries in contexts characterized by interpretive ambiguity, a plurality of evaluative frames, and legitimate disagreement [72]. As van der Velden et al. [69] show through the concept of factual opinion polarization, many claims incorporate normative judgments or are situated within politically and culturally contested contexts, rendering categorical verdicts inadequate. In such cases, forcing statements into discrete labels can be interpreted as a form of epistemic violence [73], insofar as it marginalizes legitimate perspectives and reduces interpretive plurality in the name of computability and standardization. More generally, the epistemological logic embedded in these systems, here understood as an expression of the computational veridiction regime analyzed throughout this section, tends to reduce uncertainty, ambiguity, and context: dimensions that are central to the exercise of epistemically responsible judgment [72]. Systems predominantly produce point estimates rather than articulated representations of uncertainty; intrinsically ambiguous claims are reduced to univocal categories; and discursive context, speaker intent, and communicative setting are often removed, generating systematic errors in cases where meaning depends on pragmatic and situational elements [74]. This compression of epistemic complexity, structurally produced by optimizing for computational scalability over argumentative adequacy, prepares the ground for a critical reflection on the institutional and political consequences of these dispositifs.

5.3. The Political Construction of Ground Truth

The production of ground truth, understood as the set of training labels that define what a system learns to classify as true or false, constitutes a foundational moment in which epistemic authority is both established and rendered open to potential contestation [67]. As Shin [66] notes, ground truth is not a natural given, but a sociotechnical construction that entails decisions about who is entitled to define truth, which sources are to be considered authoritative, how disagreement should be resolved, and which temporal horizons should be applied.

The sources commonly used in disinformation detection systems include professional fact-checking organizations such as Snopes and PolitiFact, scientific databases,

government statistics, and forms of expert consensus [75]. While these sources enjoy high institutional authority, they reflect specific epistemic cultures and may incorporate biases in topic selection and information framing [69,76,77]. Beyond source selection, additional layers of situated judgment are introduced through annotation practices. In datasets constructed via crowdsourcing, the demographic and cultural characteristics of annotators, often drawn from high-income countries and holding specific political orientations, contribute to shaping the patterns learned by systems [69]. This process also depends on human annotation work that is largely invisibilized. Behind every label and classification lies labor that is often precarious, accelerated, and insufficiently recognized, yet plays a direct role in shaping the datasets from which algorithmic systems learn [78,79]. The distribution of this work is structurally asymmetric: annotation tasks are frequently outsourced to workers located in the Global South, while decisions concerning categories, benchmarks, and definitions of ground truth tend to remain concentrated in institutions and corporations based in the Global North. This asymmetry is not incidental. It instantiates, at the level of technical production, the power relations that the Foucauldian concept of *dispositif* identifies at the level of discourse: who defines the rules of veridiction, and whose knowledge counts as ground truth, are not neutral technical choices but expressions of structured inequality [80]. Couldry and Mejias conceptualize these developments through the notion of “data colonialism,” not merely as a metaphor but as a structural analogy with historical forms of colonial appropriation. When considered from an epistemic perspective, this framework draws attention to how processes of knowledge validation and classification may reproduce global asymmetries. The consequence is not simply unequal representation, but the possibility of epistemic extraction, whereby the cognitive and communicative labor of non-hegemonic communities contributes to the construction of systems that may subsequently operate to their structural disadvantage. When training data disproportionately reflect dominant perspectives, models may classify dissenting or minority narratives as unreliable or false, independently of their epistemic merit [81], generating a structural risk of epistemic injustice for marginalized communities [82].

6. Artificial Truth, Algorithmic Power, and the Crisis of Democratic Epistemology

6.1. Artificial Truth as Epistemic Governance: A Theoretical Framework

The three transformations documented in Sections 3–5 converge in a single systemic configuration. Section 3 showed, through a Bourdieusian lens, how platforms reorganize the economies of symbolic capital that sustain epistemic authority, displacing institutional credentials in favor of engagement-based metrics and affective proximity. Section 4 showed, through an Actor-Network Theory perspective, how generative AI operates as a non-human epistemic actant that produces synthetic authority through fluency and procedural confidence rather than propositional grounding. Section 5 showed, through a Foucauldian analysis, how automated fact-checking systems instantiate computational *dispositifs* of veridiction that translate contested epistemic judgments into probabilistic classifications presented as neutral. Together, these three transformations constitute what this article defines as Artificial Truth. This notion does not refer simply to the production of synthetic content through artificial intelligence. Rather, it designates a systemic configuration of epistemic governance in which the validation of truth claims is increasingly delegated, in structurally asymmetric ways, to privatized algorithmic infrastructures. Table 2 maps this convergence.

Table 2. From analytical framework to Artificial Truth: convergence of three transformations.

Transformation	Theoretical Lens (Section 2)	Mechanism Documented	Mechanism of Artificial Truth	Democratic Implication
Trust economies and epistemic capital	Bourdieu: symbolic capital and fields	Displacement of institutional expertise by engagement-based and affective forms of platform-native capital	Algorithmic trust as epistemic delegation: credibility detaches from certification and attaches to visibility metrics	Erosion of institutional authority; audiences structurally exposed to non-certified epistemic actors
Generative AI as epistemic actant	Latour: Actor-Network Theory and distributed agency	Emergence of generative AI as non-human actant producing synthetic authority through fluency and procedural confidence	Synthetic truth: plausibility competes with correspondence as the dominant criterion of validity	Normalization of machine authority; automation bias; cognitive delegation to opaque infrastructures
Computational veridiction	Foucault: regimes of truth and governmentality	Institutionalization of automated fact-checking as a dispositif translating contested epistemic judgments into probabilistic classifications	Computational veridiction: truth constituted through technical procedures that embed power relations and structural asymmetries	Privatization of epistemic authority; epistemic injustice; exclusion of non-hegemonic epistemologies
Artificial Truth (synthesis)	Integrated framework	Convergence of the three transformations into a systemic configuration of algorithmic epistemic governance	Artificial Truth: validation of knowledge claims delegated to privatized infrastructures optimized for engagement rather than democratic deliberation	Epistocratic drift; erosion of democratic friction; structural compression of epistemic citizenship

Note. Author's synthesis.

This form of governance operates through three interrelated mechanisms. First, algorithms function as gatekeeping devices, structuring access to information and influencing the formation of what has been described as “algorithmic public opinion” [83]. Second, recommendation and ranking systems reconfigure criteria of visibility, shifting emphasis from editorial logics shaped by professional norms toward engagement-oriented metrics often aligned with commercial objectives [35,84]. Third, automated moderation systems allow platforms to regulate public communication at scale, embedding economic and reputational considerations into the operational infrastructure of communication itself [7].

What distinguishes Artificial Truth, in this framework, is not simply the presence of algorithms, but a broader reorganization of the public sphere along all three dimensions identified in Section 2. The shift is visible when measured against the Habermasian model of the public sphere, which presupposed relatively identifiable and normatively constrained gatekeepers [85]: algorithmic mediation substitutes that model with proprietary systems whose decision-making criteria are only partially visible and optimized primarily for engagement rather than deliberative adequacy [86,87]. The contrast is not invoked here as a normative standard but as an analytical reference point that makes the structural novelty of Artificial Truth visible.

It is important to note that “platforms” do not constitute a homogeneous category. Algorithmic architectures differ substantially across systems. The concept of Artificial Truth, as proposed here, should therefore be understood as identifying a regime-level tendency rather than a uniform mechanism. Further empirical research is necessary to specify how this tendency manifests across platform-specific contexts.

6.2. The Privatization of Epistemic Authority and the Platformization of Knowledge

The concentration of epistemic power within platform corporations produces what the framework developed here would identify as an institutional expression of algorithmic governmentality at the level of democratic life. Mendonça et al. [88] describe this tendency as an “epistocratic drift,” in which algorithms operate as institutions that compress spaces

of democratic deliberation and political contestation, partially replacing them with forms of technocratic governance grounded in computational expertise. This drift is not an external addition to the Artificial Truth regime; it is its political form. The privatization of epistemic authority unfolds along three closely interconnected dimensions.

First, infrastructural control. Platforms do not merely distribute knowledge, but govern the entire cycle of information production, circulation, and reception through curation, ranking, and recommendation systems that actively shape what is recognized as relevant knowledge [84]. Second, the privatized validation of truth claims. Fact-checking systems, “authoritative content” labels, and verification tools operate according to criteria defined by corporations, giving rise to a form of “delegated epistemic sovereignty” in which citizens depend on non-democratically accountable actors for definitions of reliability and attention [89,90]. Third, algorithmic mediation assumes quasi-judicial traits: platforms act as “new governors” of online discourse [84], making decisions of amplification, demotion, or removal on the basis of proprietary criteria that are largely optimized for profit rather than for pluralism, accuracy, or equity [7,91].

6.3. *The Transformation of Epistemic Citizenship: From Critical Agency to Passive Consumption*

The platformization of knowledge tends to profoundly reconfigure epistemic citizenship, positioning users increasingly less as citizens endowed with critical capacities and more as epistemic consumers. Algorithmic curation pre-selects information for consumption, reducing the space and the need for autonomous evaluation, source comparison, and critical synthesis [92,93]. The result is a progressive depoliticization of epistemology, in which the cognitive labor of evaluating sources and constructing interpretive orientations is structurally delegated to algorithmic systems [86,88]. This dynamic does not concern individual competences alone, but the material configuration within which public cognition is exercised. Users have limited margins to modify algorithms, understand their functioning, or collectively contest their design choices [94]. When algorithmic systems determine visibility, framing, and the set of available alternatives, citizens’ capacity to form independent judgments is structurally compressed, even in the absence of a conscious renunciation of epistemic agency. This compression is the direct democratic consequence of the three transformations documented in the preceding sections: the displacement of institutional authority by engagement capital, the normalization of synthetic epistemic outputs, and the delegation of veridiction to computational dispositifs. Hyzen et al. [94] capture the material stakes of this configuration through the concept of epistemic welfare, understood as the institutional and infrastructural conditions that make independent epistemic agency possible in the public sphere. When those conditions are structured by platform logics rather than democratic norms, what Öğüç [87] calls epistemic autonomy is not so much violated as rendered structurally unnecessary: algorithmic delegation becomes a *de facto* prerequisite for participation in digital public life, not formally imposed but practically unavoidable [7].

6.4. *Epistemic Injustice, Data Colonialism, and Geopolitical Asymmetries*

Algorithmic governance of knowledge incorporates and reproduces structural inequalities of a social and geopolitical nature. García and Calvo [95] speak of “algorithmic colonization” to describe the use of synthetic content and algorithmic manipulation practices in the formation of public opinion to the advantage of elite interests, particularly in contexts characterized by limited regulatory capacity and reduced informational resilience. Epistemic injustice manifests itself in the systematic marginalization of voices and perspectives, especially non-Western or minority ones, whose communicative practices may not align with the norms embedded in systems designed primarily for Anglophone Western

users [89]. Linguistic hierarchies are particularly evident in systems trained largely on English-language data, producing structural disadvantages for minority languages and the communities that use them [35,96].

The global scale of platforms renders these hierarchies intrinsically transnational. Security policies and priorities defined in the United States and Europe tend to orient visibility and moderation practices on a global scale [96]. At the same time, the political economy of platform power generates winner-take-all dynamics that consolidate the epistemic authority of a small number of dominant actors [92]. In this sense, algorithms operate not as neutral tools, but as institutions that reflect the interests, values, and biases of their designers and of the organizations that deploy them [88,97].

6.5. Democratic Implications: Accountability, Depoliticization, and Pluralism

Algorithmic mediation of the public sphere generates a plurality of democratic deficits. Coeckelbergh [98] observes that technocratic approaches to the governance of artificial intelligence tend to exclude or marginalize citizen participation, proving incapable of adequately incorporating democratic values into system design processes. By framing algorithmic governance as an essentially technical problem, such approaches shift the issue from the terrain of political deliberation to that of functional optimization, reducing the scope for public debate on the purposes and social effects of technologies [88,98].

A crucial effect of this orientation follows directly from the Artificial Truth regime as theorized in this article. When veridiction is delegated to computational dispositifs, when symbolic capital accumulates through engagement rather than deliberation, and when non-human actants stabilize epistemic authority through black-boxed procedures, what Elkin-Koren and Perel [86] call “democratic friction” is structurally eroded. Democracy presupposes friction: the separation of powers, systems of checks and balances, discursive contestation, and the effective possibility of dissent. AI-mediated governance of discourse operates through probabilistic decision-making and recursive optimization that tend to compress precisely these spaces, reducing opportunities for self-government and for the public negotiation of the norms that regulate communication [86]. The erosion is not incidental to the Artificial Truth regime; it is its democratic symptom.

The resulting depoliticization of truth transforms inherently contestable issues into technical problems presented as neutral. When content moderation, disinformation detection, and fact-checking are treated as engineering challenges, the political character of truth-making processes is obscured [90,99]. Zhakin and Mukan [100] describe a shift from truth understood as accuracy toward emotional engagement, and from autonomy as individual freedom toward a form of autonomy mediated by algorithmic systems. The risk that follows is the progressive institutionalization of a platformized “official truth,” not as a monolithic imposition, but as the structural outcome of automated dispositifs applied in the absence of adequate transparency, accountability, and democratic legitimation.

7. Conclusions

This article has developed an integrated theoretical account of how digital information ecosystems reorganize the production, validation, and circulation of knowledge. Generative AI systems and automated verification infrastructures do not simply enhance informational efficiency. Platforms deploy them as epistemic actors. They reshape regimes of truth by structuring visibility, credibility, and authority. The concept of Artificial Truth names this emerging configuration of epistemic governance in algorithmic societies.

The analysis has combined three classical perspectives, Foucauldian regimes of truth, Bourdieusian symbolic capital, and the constructivist insights of Science and Technology Studies, into a unified framework for examining algorithmic mediation. Together, these

approaches show that computational systems do not merely distribute information. They rank, classify, and synthesize. In doing so, they determine which claims circulate widely, which appear credible, and which remain marginal. Artificial Truth designates a regime in which plausibility, performative fluency, and infrastructural positioning increasingly shape what counts as true.

Three lines of inquiry have structured the argument. First, platforms reorganize economies of trust and symbolic capital. Engagement metrics and affective proximity recalibrate authority, often displacing institutional expertise. Second, generative systems act as epistemic agents. They produce synthetic statements that simulate authority through fluency and confidence, even in the absence of propositional understanding. Third, automated fact-checking infrastructures institutionalize computational veridiction. They convert situated and contestable judgments into probabilistic classifications that appear neutral while embedding epistemological assumptions.

Epistemologically, Artificial Truth does not signal a simple “crisis.” It marks a re-orientation. Earlier epistemic orders anchored validation in argumentative procedures, contextual evidence, and publicly contestable standards. Algorithmic systems optimize for performance metrics within opaque infrastructures. As platforms routinize these procedures, computational plausibility competes with correspondence as a dominant evaluative criterion. Politically, this shift concentrates epistemic power within platform corporations. Proprietary systems now validate truth claims at scale, often without meaningful democratic oversight. What emerges is not merely technocracy, but a redistribution of epistemic citizenship: users increasingly consume pre-ranked knowledge rather than participate in the conditions that define it. Democratically, algorithmic mediation alters the material conditions of pluralism. When platforms translate contestable issues into optimization problems, they narrow spaces of disagreement. Training data incorporate historical asymmetries; ranking systems amplify dominant perspectives; marginalized epistemologies struggle for visibility.

This study remains primarily theoretical. It has not provided systematic empirical evidence on how specific users, institutions, or platforms negotiate algorithmic authority in practice. The framework draws largely on Western contexts, limiting its applicability to information ecosystems shaped by different regulatory regimes or epistemic cultures. Moreover, the analysis has treated algorithmic systems in relatively aggregate terms, without differentiating sufficiently between distinct architectures, governance models, and business logics.

Rather than listing generic directions, it is worth specifying how the framework developed here can be operationalized empirically across each of its three analytical dimensions.

The first dimension, the restructuring of trust economies and symbolic capital, calls for comparative platform analyses that measure the relationship between institutional credentials and engagement metrics as competing sources of epistemic authority. Concretely, this means examining how the same claim, produced by an institutional actor and by a high-engagement non-institutional actor, circulates differently across platforms with distinct algorithmic architectures. Network analysis combined with content analysis would allow researchers to trace how visibility translates into perceived credibility across different user groups and regulatory contexts. Cross-platform comparison is essential here, since TikTok, YouTube, and search-based systems embed structurally different logics of capital accumulation.

The second dimension, generative AI as an epistemic actant producing synthetic truth, calls for experimental designs that measure how users calibrate trust toward AI-generated outputs across different disciplinary and cultural contexts. Survey experiments comparing perceived credibility of AI-generated versus human-authored texts, controlling for domain

expertise and interface design, would allow researchers to specify under which conditions fluency displaces correspondence as an evaluative criterion. Qualitative components, such as think-aloud protocols or in-depth interviews with professionals in journalism, science, and law, would complement experimental data by documenting how practitioners negotiate the epistemic status of generative outputs in situated practice.

The third dimension, computational veridiction in automated fact-checking systems, calls for computational audits of specific systems, combined with critical analysis of the training datasets and annotation protocols that define their ground truth. Researchers should examine systematically how disagreement among annotators is resolved, which sources are treated as authoritative, and how geographic and linguistic representation is distributed across training corpora. Connecting audit findings to documented moderation outcomes would allow researchers to specify empirically where epistemic injustice is most likely to be reproduced and amplified.

Normatively, these three lines of empirical inquiry converge on a common question: what institutional and regulatory arrangements are capable of subjecting algorithmic epistemic authority to public oversight and democratic accountability? Addressing this question requires collaboration across computational methods, qualitative sociology, science and technology studies, and democratic theory.

Artificial Truth is not a technological destiny. It is a historically situated configuration of epistemic power. Platforms, designers, regulators, and users actively sustain it through everyday practices. Because it is constructed, it can be contested. The central question is not whether machines tell the truth. It is who designs the procedures that define truth, who controls the infrastructures that stabilize it, and how those procedures align with democratic principles. When algorithmic systems restructure the institutional conditions that make reality publicly recognizable, the risk is not only the circulation of falsehood. It is the erosion of shared criteria through which societies identify and debate truth itself.

Funding: This research received no external funding. The APC was not funded by any external source.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study.

Acknowledgments: During the preparation of this manuscript, the author used DeepL for language refinement and text editing. The author has reviewed and edited the output and takes full responsibility for the content of this publication. The author has read and agreed to the published version of the manuscript.

Conflicts of Interest: The author declares no conflicts of interest.

References

1. Orwell, G. 1984; Arnoldo Mondadori Editore: Milan, Italy, 1989.
2. Dick, P.K. *How to Build a Universe That Doesn't Fall Apart Two Days Later*; ISOLARII: New York, NY, USA, 2024.
3. Aimeur, E.; Amri, S.; Brassard, G. Fake news, disinformation and misinformation in social media: A review. *Soc. Netw. Anal. Min.* **2023**, *13*, 30. [\[CrossRef\]](#)
4. Cvrtila, L. Truth politics and social media. *Politička Misao* **2024**, *61*, 7–30. [\[CrossRef\]](#)
5. Foucault, M. *La Volonté de Savoir*; Gallimard: Paris, France, 1976.
6. Foucault, M. Le jeu de Michel Foucault. In *Dits et Écrits II, 1976–1988*; Gallimard: Paris, France, 1977; pp. 298–329.
7. Cobbe, J. Algorithmic censorship by social platforms: Power and resistance. *Philos. Technol.* **2021**, *34*, 739–766. [\[CrossRef\]](#)
8. Sahakyan, H.; Gevorgyan, A.; Malkjyan, A. From disciplinary societies to algorithmic control: Rethinking Foucault's human subject in the digital age. *Philosophies* **2025**, *10*, 73. [\[CrossRef\]](#)

9. Jarke, J.; Prietl, B.; Egbert, S.; Boeva, Y.; Heuer, H. Knowing in algorithmic regimes: An introduction. In *Algorithmic Regimes: Methods, Interactions, and Politics*; Jarke, J., Prietl, B., Egbert, S., Boeva, Y., Heuer, H., Arnold, M., Eds.; Amsterdam University Press: Amsterdam, The Netherlands, 2024; pp. 7–34. [\[CrossRef\]](#)
10. Bourdieu, P. *Raisons Pratiques: Sur la Théorie de l'Action*; Seuil: Paris, France, 1994.
11. Ling, C.; Shen, Y.; Shanahan, E.A. Algorithmic meta-capital and the narrative policy framework. *Policy Stud. J.* **2025**, *53*, 1108–1122. [\[CrossRef\]](#)
12. Bourdieu, P. *Les Règles de l'Art: Genèse et Structure du Champ Littéraire*; Seuil: Paris, France, 1992.
13. Latour, B. *Reassembling the Social: An Introduction to Actor-Network-Theory*; Oxford University Press: Oxford, UK, 2005.
14. Bucher, T. *If...Then: Algorithmic Power and Politics*; Oxford University Press: Oxford, UK, 2018.
15. Mackenzie, A. *Machine Learners: Archaeology of a Data Practice*; MIT Press: Cambridge, MA, USA, 2017.
16. Seaver, N. Algorithms as culture: Some tactics for the ethnography of algorithmic systems. *Big Data Soc.* **2017**, *4*, 1–12. [\[CrossRef\]](#)
17. Seaver, N. Captivating algorithms: Recommender systems as traps. *J. Mater. Cult.* **2019**, *24*, 421–436. [\[CrossRef\]](#)
18. Crawford, K. *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*; Yale University Press: New Haven, CT, USA, 2021.
19. Luhmann, N. *Trust and Power*; Wiley: Hoboken, NJ, USA, 1979.
20. Giddens, A. *The Consequences of Modernity*; Stanford University Press: Stanford, CA, USA, 1990.
21. Carlson, M. *Journalistic Authority: Legitimizing News in the Digital Era*; Columbia University Press: New York, NY, USA, 2017.
22. Broersma, M. Journalism as performative discourse. In *Journalism and Meaning-Making: Reading the Newspaper*; Rupa, V., Ed.; Hampton Press: Cresskill, NJ, USA, 2010; pp. 15–35.
23. Kioussis, S. Public trust or mistrust? Perceptions of media credibility in the information age. *Mass Commun. Soc.* **2001**, *4*, 381–403. [\[CrossRef\]](#)
24. Dalen, A.V. Journalism, trust, and credibility. In *The Handbook of Journalism Studies*, 2nd ed.; Wahl-Jorgensen, K., Hanitzsch, T., Eds.; Routledge: London, UK, 2019; pp. 356–371. [\[CrossRef\]](#)
25. Neuberger, C.; Bartsch, A.; Fröhlich, R.; Hanitzsch, T.; Reinemann, C.; Schindler, J. The digital transformation of knowledge order: A model for the analysis of the epistemic crisis. *Ann. Int. Commun. Assoc.* **2023**, *47*, 180–201. [\[CrossRef\]](#)
26. Sperber, D.; Clément, F.; Heintz, C.; Mascaro, O.; Mercier, H.; Origg, G.; Wilson, D. Epistemic vigilance. *Mind Lang.* **2010**, *25*, 359–393. [\[CrossRef\]](#)
27. Gillespie, T. *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*; Yale University Press: New Haven, CT, USA, 2018.
28. van Dijck, J.; Poell, T.; de Waal, M. *The Platform Society: Public Values in a Connective World*; Oxford University Press: Oxford, UK, 2018.
29. Carlson, M. Fake news as an informational moral panic: The symbolic deviancy of social media during the 2016 US presidential election. *Inf. Commun. Soc.* **2020**, *23*, 374–388. [\[CrossRef\]](#)
30. Waisbord, S. Truth is what happens to news: On journalism, fake news, and post-truth. *Journal. Stud.* **2018**, *19*, 1866–1878. [\[CrossRef\]](#)
31. Vos, T.P.; Thomas, R.J. The discursive construction of journalistic authority in a post-truth age. *Journal. Stud.* **2018**, *19*, 2001–2010. [\[CrossRef\]](#)
32. Scheufele, D.A.; Krause, N.M. Science audiences, misinformation, and fake news. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 7662–7669. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Ross Arguedas, A.; Robertson, C.; Fletcher, R.; Nielsen, R. *Echo Chambers, Filter Bubbles, and Polarisation: A Literature Review*; Reuters Institute for the Study of Journalism: Oxford, UK, 2022. Available online: <https://ora.ox.ac.uk/objects/uuid:6e357e97-7b16-450a-a827-a92c93729a08> (accessed on 19 February 2026).
34. Dubey, H.V. The use–trust loop: Reel culture, semi-news narratives, and credibility in hyperlocal journalism. *ShodhKosh J. Vis. Perform. Arts* **2023**, *4*, 4719–4725. [\[CrossRef\]](#)
35. Garajamirli, N. Algorithmic gatekeeping and democratic communication: Who decides what the public sees? *Eur. J. Commun. Media Stud.* **2025**, *4*, 54–67. [\[CrossRef\]](#)
36. van Dijck, J. *The Culture of Connectivity: A Critical History of Social Media*; Oxford University Press: Oxford, UK, 2013.
37. Gillespie, T. The relevance of algorithms. In *Media Technologies: Essays on Communication, Distribution, and Difference*; Gillespie, T., Boczkowski, P.J., Foot, K.A., Eds.; MIT Press: Cambridge, MA, USA, 2014; pp. 167–193.
38. Abidin, C. *Internet Celebrity: Understanding Fame Online*; Emerald Publishing: Bingley, UK, 2018.
39. Couldry, N.; Hepp, A. *The Mediated Construction of Reality*; Polity: Cambridge, UK, 2018.
40. Lewis, J.D.; Weigert, A. The social dynamics of trust: Theoretical and empirical research, 1985–2012. *Soc. Forces* **2012**, *91*, 25–31. [\[CrossRef\]](#)
41. Miller, B.; Record, I. Justified belief in a digital age: On the epistemic implications of secret internet technologies. *Episteme* **2013**, *10*, 117–134. [\[CrossRef\]](#)

42. Coady, D.; Chase, J. (Eds.) *The Routledge Handbook of Applied Epistemology*; Routledge: London, UK, 2019.
43. Callon, M.; Law, J. After the individual in society: Lessons on collectivity from science, technology and society. *Can. J. Sociol.* **1997**, *22*, 165–182. [\[CrossRef\]](#)
44. Munn, L.; Magee, L.; Arora, V. Truth machines: Synthesizing veracity in AI language models. *AI Soc.* **2024**, *39*, 2759–2773. [\[CrossRef\]](#)
45. Heersmink, R.; de Rooij, B.; Clavel Vázquez, M.J.; Colombo, M. A phenomenology and epistemology of large language models: Transparency, trust, and trustworthiness. *Ethics Inf. Technol.* **2024**, *26*, 41. [\[CrossRef\]](#)
46. Haraway, D. *Simians, Cyborgs, and Women: The Reinvention of Nature*; Routledge: New York, NY, USA, 1991.
47. Hayles, N.K. *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*; University of Chicago Press: Chicago, IL, USA, 1999.
48. Floridi, L.; Chiriatti, M. GPT-3: Its nature, scope, limits, and consequences. *Minds Mach.* **2020**, *30*, 681–694. [\[CrossRef\]](#)
49. Bender, E.M.; Gebru, T.; McMillan-Major, A.; Shmitchell, S. On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21)*; Association for Computing Machinery: New York, NY, USA, 2021; pp. 610–623. [\[CrossRef\]](#)
50. Jakesch, M.; Hancock, J.T.; Naaman, M. Human heuristics for AI-generated language are flawed. *Proc. Natl. Acad. Sci. USA* **2023**, *120*, e2208839120. [\[CrossRef\]](#) [\[PubMed\]](#)
51. Markowitz, D.M. From complexity to clarity: How AI enhances perceptions of scientists and the public's understanding of science. *PNAS Nexus* **2024**, *3*, pgae387. [\[CrossRef\]](#)
52. Wiesner, F.; Koopman, B.; Gupta, S.; Yang, Y. Hallucinate or memorize? The two sides of probabilistic learning in large language models. *arXiv* **2025**, arXiv:2511.08877. [\[CrossRef\]](#)
53. Verma, R.K. The code of society: Constructing social theory through large language models. *Preprints* **2025**, preprints202505.1792. [\[CrossRef\]](#)
54. Huang, L.; Yu, W.; Ma, W.; Zhong, W.; Feng, Z.; Wang, H.; Chen, Q.; Peng, W.; Feng, X.; Qin, B.; et al. A survey on hallucination in large language models: Principles, taxonomy, challenges, and open questions. *ACM Trans. Inf. Syst.* **2025**, *43*, 42. [\[CrossRef\]](#)
55. Zhou, J.; Zhang, Y.; Luo, Q.; Parker, A.G.; De Choudhury, M. Synthetic lies: Understanding AI-generated misinformation and evaluating algorithmic and human solutions. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*; Association for Computing Machinery: New York, NY, USA, 2023. [\[CrossRef\]](#)
56. Shanahan, M. Talking about large language models. *arXiv* **2023**, arXiv:2212.03551. [\[CrossRef\]](#)
57. Li, A.; Sinnamon, L. Generative AI search engines as arbiters of public knowledge: An audit of bias and authority. *Proc. Assoc. Inf. Sci. Technol.* **2024**, *61*, 205–217. [\[CrossRef\]](#)
58. Napoli, P.M. social media and the public interest: Governance of news platforms in the realm of individual and algorithmic gatekeepers. *Telecommun. Policy* **2019**, *39*, 751–760. [\[CrossRef\]](#)
59. Gaube, S.; Suresh, H.; Raue, M.; Merritt, A.; Berkowitz, S.J.; Lerner, E.; Coughlin, J.F.; Gutttag, J.V.; Colak, E.; Ghassemi, M. Do as AI say: Susceptibility in deployment of clinical decision-aids. *NPJ Digit. Med.* **2021**, *4*, 31. [\[CrossRef\]](#) [\[PubMed\]](#)
60. Vasconcelos, H.; Jörke, M.; Grunde-McLaughlin, M.; Gerstenberg, T.; Bernstein, M.S.; Krishna, R. Explanations can reduce overreliance on AI systems during decision-making. *Proc. ACM Hum.-Comput. Interact.* **2023**, *7*, 129. [\[CrossRef\]](#)
61. Cabitza, F.; Campagner, A.; Balsano, C. Bridging the “last mile” gap between AI implementation and operation: “data awareness” that matters. *Ann. Transl. Med.* **2020**, *8*, 501. [\[CrossRef\]](#)
62. Zuboff, S. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*; PublicAffairs: New York, NY, USA, 2019.
63. Star, S.L.; Ruhleder, K. Steps toward an ecology of infrastructure: Design and access for large information spaces. *Inf. Syst. Res.* **1996**, *7*, 111–134. [\[CrossRef\]](#)
64. Kay, J.; Kasirzadeh, A.; Mohamed, S. Epistemic injustice in generative AI. In *Proceedings of the 2024 AAAI/ACM Conference on AI, Ethics, and Society (AI/ES 2024)*; Association for the Advancement of Artificial Intelligence: Washington, DC, USA, 2024; pp. 684–697.
65. Rieder, B. *Engines of Order: A Mechanology of Algorithmic Techniques*; Amsterdam University Press: Amsterdam, The Netherlands, 2020. [\[CrossRef\]](#)
66. Shin, D. Automating epistemology: How AI reconfigures truth, authority, and verification. *AI Soc.* **2025**, *41*, 1553–1559. [\[CrossRef\]](#)
67. Domínguez Hernández, A.; Owen, R.; Nielsen, D.S.; McConville, R. Ethical, political and epistemic implications of machine learning (mis)information classification: Insights from an interdisciplinary collaboration between social and data scientists. *J. Responsible Innov.* **2023**, *10*, 2222514. [\[CrossRef\]](#)
68. Thorne, J.; Vlachos, A. Automated fact checking: Task formulations, methods and future directions. *arXiv* **2018**, arXiv:1806.07687. [\[CrossRef\]](#)
69. van der Velden, M.A.C.G.; Loecherbach, F.; van Atteveldt, W.; Fokkens, A.; Reuver, M.; Welbers, K. Whose truth is it anyway? An experiment on annotation bias in times of factual opinion polarization. *Commun. Methods Meas.* **2025**, *19*, 332–349. [\[CrossRef\]](#)

70. Qureshi, K.A.; Malick, R.A.S.; Sabih, M. Social media and microblogs credibility: Identification, theory driven framework, and recommendation. *IEEE Access* **2021**, *9*, 137744–137781. [\[CrossRef\]](#)
71. Schuster, T.; Schuster, R.; Shah, D.J.; Barzilay, R. The limitations of stylometry for detecting machine-generated fake news. *Comput. Linguist.* **2020**, *46*, 499–530. [\[CrossRef\]](#)
72. Glockner, M.; Staliūnaitė, I.; Thorne, J.; Vallejo, G.; Vlachos, A.; Gurevych, I. AmbiFC: Fact-checking ambiguous claims with evidence. *Trans. Assoc. Comput. Linguist.* **2024**, *12*, 1–18. [\[CrossRef\]](#)
73. D'Ignazio, C.; Klein, L.F. *Data Feminism*, Open Access ed.; The MIT Press: Cambridge, MA, USA, 2020. [\[CrossRef\]](#)
74. Rubin, V.L. Content verification for social media: From deception detection to automated fact-checking. In *The SAGE Handbook of Social Media Research Methods*; Quan-Haase, A., Sloan, L., Eds.; SAGE Publications: London, UK, 2022; pp. 393–414. [\[CrossRef\]](#)
75. Thibault, C.; Tian, J.-J.; Péloquin-Skulski, G.; Curtis, T.L.; Zhou, J.; Laflamme, F.; Guan, L.Y.; Rabbany, R.; Godbout, J.-F.; Pelrine, K. A guide to misinformation detection data and evaluation. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining*; ACM: New York, NY, USA, 2025; pp. 5801–5809. [\[CrossRef\]](#)
76. Graves, L. *Deciding What's True: The Rise of Political Fact-Checking in American Journalism*; Columbia University Press: New York, NY, USA, 2016.
77. Amazeen, M.A. Journalistic interventions: The structural factors affecting the global emergence of fact-checking. *Journalism* **2020**, *21*, 95–111. [\[CrossRef\]](#)
78. Gebru, T.; Morgenstern, J.; Vecchione, B.; Wortman Vaughan, J.; Wallach, H.; Daumé, H., III; Crawford, K. Datasheets for datasets. *Commun. ACM* **2021**, *64*, 86–92. [\[CrossRef\]](#)
79. Roberts, S.T. *Behind the Screen: Content Moderation in the Shadows of Social Media*; Yale University Press: New Haven, CT, USA, 2019.
80. Couldry, N.; Mejias, U.A. *The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism*; Stanford University Press: Stanford, CA, USA, 2019.
81. Park, J.; Ellezhuthil, R.; Arunachalam, R.; Feldman, L.; Singh, V. Toward fairness in misinformation detection algorithms. In *Proceedings of the Workshop on News Media and Computational Journalism (MEDIATE), 16th International AAAI Conference on Web and Social Media (ICWSM 2022)*; Association for the Advancement of Artificial Intelligence: Washington, DC, USA, 2022. [\[CrossRef\]](#)
82. Neumann, T.; De-Arteaga, M.; Fazelpour, S. Justice in misinformation detection systems: An analysis of algorithms, stakeholders, and potential harms. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22)*; Association for Computing Machinery: New York, NY, USA, 2022; pp. 1504–1515. [\[CrossRef\]](#)
83. Gandini, A.; Gerosa, A.; Gobbo, B.; Keeling, S.; Leonini, L.; Mosca, L.; Orofino, M.; Reviglio, U.; Splendore, S. The algorithmic public opinion: A literature review. *SocArXiv* **2022**. [\[CrossRef\]](#)
84. Schneiders, P.; Stark, B. Ensuring news quality in platformized news ecosystems: Shortcomings and recommendations for an epistemic governance. *Media Commun.* **2025**, *13*, 10042. [\[CrossRef\]](#)
85. Habermas, J. *Storia e Critica dell'Opinione Pubblica*; Laterza: Bari, Italy, 1971.
86. Elkin-Koren, N.; Perel, M. Democratic friction in speech governance by AI. In *Research Handbook on the Law of Artificial Intelligence*; Edward Elgar Publishing: Cheltenham, UK, 2023; pp. 1029–1050. [\[CrossRef\]](#)
87. Nez, E.; Quintana, I.P. Algoritmos e racionalidade pública: Análise da influência dos sistemas automatizados na deliberação democrática à luz da Teoria Habermasiana. *Logeion Filos. Inf.* **2024**, *11*, e7383. [\[CrossRef\]](#)
88. Mendonça, R.F.; Amaral, E.F.L.; Almada, M.P. Algorithms and politics. In *The Oxford Handbook of Deliberative Democracy*; Bächtiger, A., Dryzek, J.S., Mansbridge, J., Warren, M.D., Eds.; Oxford University Press: Oxford, UK, 2023; pp. 123–138. [\[CrossRef\]](#)
89. Tobi, A. Towards an epistemic compass for online content moderation. *Philos. Technol.* **2024**, *37*, 791–815. [\[CrossRef\]](#)
90. Romanishyn, S.; Koval, V.; Petrenko, O. AI-driven disinformation: Policy recommendations for democratic resilience. *Front. Artif. Intell.* **2025**, *8*, 1569115. [\[CrossRef\]](#) [\[PubMed\]](#)
91. Aytac, U. Big Tech, algorithmic power, and democratic control. *J. Polit.* **2024**, *86*, 1431–1445. [\[CrossRef\]](#)
92. Tabarés, R. Plataformización, automatización y aceleración en los medios sociales: Amenazas para la deliberación democrática. *Daimon Rev. Int. Filosof.* **2024**, *93*, 127–142. [\[CrossRef\]](#)
93. Ögüç, Ç. Dead internet hypothesis: AI, censorship, and the decline of human-centered digital discourse. *Imgelem* **2025**, *9*, 89–112. [\[CrossRef\]](#)
94. Hyzen, A.; Loosen, W.; Reimer, J. Epistemic welfare and algorithmic recommender systems: Overcoming the epistemic crisis in the digitalized public sphere. *Commun. Theory* **2025**, *35*, qtaf018. [\[CrossRef\]](#)
95. García, C.S.; Calvo, P. Opinión pública masiva: Colonización algorítmica y sintetificación de la esfera pública. *Rev. CIDOB Afers Int.* **2024**, *138*, 73–98. [\[CrossRef\]](#)
96. Radsch, C.C. On the frontlines of the information wars: How algorithmic gatekeepers and national security impact journalism. In *National Security, Journalism, and Law in an Age of Information Warfare*; Ambinder, M., Henrichsen, J.R., Rosati, C., Eds.; Oxford University Press: Oxford, UK, 2024; pp. 277–300. [\[CrossRef\]](#)
97. Monteiro, R.L.; Almeida, V.A.F.; Doneda, D. Legitimidade democrática na governança algorítmica: Uma análise tridimensional. *Rev. Dir. Fundam. Democr.* **2024**, *29*, 8–32. [\[CrossRef\]](#)

98. Coeckelbergh, M. Artificial intelligence, the common good, and the democratic deficit in AI governance. *AI Ethics* **2024**, *4*, 1089–1098. [[CrossRef](#)]
99. Büscher, B. Artificial intelligence, platform capitalist power, and the impact of the crisis of truth on ethnography. *Annu. Rev. Anthropol.* **2025**, *54*, 253–269. [[CrossRef](#)]
100. Zhakin, S.; Mukan, N. Digital technologies and the reformatting of values in the post-truth era. *Bull. Karaganda Univ. Hist. Philos. Ser.* **2025**, *30*, 241–247. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.