Disinformation Between Knowledge and Ignorance An Epistemological Comparison

Antonio Lizzadri

Catholic University of the Sacred Heart of Milan
Largo Gemelli 1, 20123 Milan, Italy
antonio.lizzadri@unicatt.it

Abstract

This paper aims to contribute to the understanding of information disorder from an epistemological perspective, by analysing the internal/cognitive as well as the external/contextual factors that determine knowledge defects. In this regard, I first compare the concept of information with the properly epistemological concept of knowledge by arguing that the concept of knowledge makes explicit the two fundamental prescriptive characteristics that beliefs should have—namely, truth and justification—that, on the contrary, remain implicit in the epistemically neutral notion of information. Therefore, I provide an externalist account of knowledge according to which a belief is true and justified to the extent that it allows for satisfactory adaptation to the environment, in which the ecological, technological and sociological environment itself becomes an integral part of an extended cognitive system. Based on this epistemological exploration of the notion of information through an externalist conception of knowledge, I suggest that disinformation can be understood as the state of an "ignorant" collective cognitive system, that is, a closed system that establishes interactions only within a virtual environment devoid of any semantic relevance and appeal to rational justification. In conclusion, I point out that, although the digital revolution poses risks to the spread of ignorance due to the speed and scale of the dissemination of disinformation, it nevertheless poses challenges that allow epistemological reflection itself to renew itself addressing the crisis of knowledge with extended theoretical resources.

Keywords: Information, disinformation, knowledge, ignorance, extended cognition

1. Introduction

The contemporary debate on information disorder has been decisively shaped by the influential report of Claire Wardle and Hossein Derakhshan (2017), which introduced a now widely adopted tripartite distinction between misinformation, disinformation, and malinformation. According to their framework, *misinformation* refers to false information shared without harmful intent; *disinformation* to false information deliberately created and disseminated to cause harm; and *malinformation* to genuine information shared to cause harm. This taxonomy has become foundational in policy documents, media literacy programs, and academic analyses, especially in the context of digital platforms and the so-called "fake news" crisis.

The strength of Wardle and Derakhshan's proposal lies in its operational clarity. It offers policymakers and researchers a vocabulary capable of distinguishing types of problematic content based on two criteria: falsity and intent to harm. However, precisely because it is primarily oriented toward regulatory and communicative concerns, it rests on a largely implicit epistemology. It treats "information" as a neutral substrate that may acquire normative valence through falsity and malicious intention, but it does not explicitly interrogate the epistemic conditions under which information becomes knowledge, nor the structural features of cognitive systems that enable or disable epistemic reliability.

From an epistemological standpoint, this neutrality is problematic. The concept of *information* employed in the report is descriptive and content-oriented: it concerns the transmission of messages or claims. It presupposes, but does not analyze, the epistemic standards that distinguish well-founded belief from mere opinion or deception. In this sense, the report remains at the level of content classification rather than engaging with the normative dimensions of belief formation. Recent interdisciplinary analyses have emphasized this limitation. Broda and Strömbäck (2024), for example, note that misinformation research often proceeds without a shared epistemology of truth and belief, while Resch et al. (2026) argue that conceptual ambiguity in information disorder studies reflects deeper normative disagreements about epistemic authority and justification.

Classical epistemology has long emphasized that knowledge is not reducible to information. Knowledge entails at least two prescriptive features: truth and justification. A belief must be true and appropriately justified to count as knowledge. In contrast, information, in many contemporary usages—particularly in communication studies and information theory—is epistemically neutral. It may be true or false, justified or unjustified, meaningful or misleading.

Proceedings of the 1st Workshop on Information Disorder (InDor) @ LREC 2026, pages 97–103
12 May 2026.

The Wardle and Derakhshan framework inherits this neutrality. It treats disinformation primarily as false content intentionally produced to harm, but it does not ask how epistemic systems process such content, nor what structural conditions allow false beliefs to stabilize and circulate.

Two epistemological shortcomings emerge from this perspective.

First, the framework is overly *intentionalist*. By distinguishing between disinformation, misinformation, and misinformation based on the presence or absence of harmful intent, it risks overlooking the systemic dimensions of epistemic failure. In digital environments, harmful falsehoods often proliferate without clear malicious agency. Algorithmic amplification, attention economies, and community-based epistemic norms can sustain and reinforce false beliefs independently of identifiable intent (Airoldi, 2021). As Ferrara (2023) shows in the context of generative AI, synthetic disinformation can proliferate through automated systems whose logic is optimization rather than deception. A strictly intentional definition may therefore obscure the distributed and emergent character of epistemic breakdowns.

Second, the framework remains *content-centric* rather than *systemic*. It assumes that the core problem lies in the falsity (misinformation and disinformation) or in the harmful truth (malinformation) of specific pieces of information. Yet epistemic failure may reside not only in isolated false or harmful true claims but in the structural configuration of cognitive environments: echo chambers (Nguyen, 2020), filter bubbles, and platform-induced polarization (Sunstein, 2018; Bail, 2021). In such contexts, even true statements may function within epistemically defective systems, while false statements may appear justified within local epistemic standards.

An epistemologically robust account of information disorder must therefore move beyond the classification of content and intent and consider the conditions under which beliefs are formed, maintained, and revised. It must examine how cognitive agents—individual and collective—interact with their environments, and how these environments shape epistemic reliability. In other words, it must shift from a theory of information to a theory of knowledge (Uscinski et al. 2024). In this regard, the epistemological analysis of information disorder presented in the following sections of this paper will especially focus on disinformation since it includes both negative aspects of information disorder—those concerning content and those concerning intent.

If disinformation is to be understood not simply as harmful false content but as an epistemic pathology, then we must analyze the norms and structures that govern belief formation recognizing that knowledge is not merely a property of isolated propositions but an achievement of cognitive systems embedded in ecological, technological, and sociological contexts. Only then can we adequately address the epistemic crisis often described as the age of fake news.

2. Knowledge Beyond Information

To overcome the epistemic neutrality of the concept of information, we must turn to the concept of knowledge understood in an extended and externalist sense. In fact, a mere internalist conception of knowledge that considers knowledge only as a mental or intrapsychic phenomenon seems not able to face the epistemological challenges posed by the digital revolution. In this regard, the work of Andy Clark (2017) offers crucial theoretical resources (Carter et al. 2018). Clark's extended mind thesis, originally developed with David Chalmers (Clark & Chalmers 1998) and subsequently refined, challenges the internalist assumption that cognition is confined within the biological boundaries of the brain. Instead, cognition is distributed across brain, body, and environment, incorporating technological artifacts and social practices as constitutive components.

Although Clark's primary focus is on cognitive processes, the epistemological implications are profound. If cognitive processes extend into the environment, then so too do the conditions for justification and epistemic reliability. Knowledge cannot be understood solely as a relation between a subject and a proposition; it is an achievement of a coupled system composed of agents and their material, technological, and social surroundings.

According to an externalist account of knowledge, a belief is true and justified to the extent that it allows satisfactory adaptation to the environment. This view resonates with extended and pragmatic strands in epistemology, while also incorporating insights from ecological psychology and cognitive science. Truth and justification are not merely abstract normative standards; they are tied to the functional success of cognitive systems in

navigating their environments. On the other hand, such externalist conception of knowledge might seem to reduce truth to justification and this reduction might seem problematic for objective verification tasks in natural language processing, as a false belief can be highly adaptive within a specific social niche, guaranteeing group cohesion and acceptance. However, it is worth noting that a theory of truth as correspondence and a theory of truth as satisfactory adaptation to the environment are not necessarily incompatible; indeed, they appear to be complementary to the extent that one recognizes that the correspondence between beliefs and reality explains, on an ontological level, why certain beliefs are successful on a pragmatic level and others are not (Putnam, 1978). In other words, satisfactory adaptation has always a semantic relevance, where "semantic" does not mean only lexical or compositional meaning in the narrow sense of formal linguistics but, more broadly, the relation between signs, their referents, and the inferential consequences they carry within practices of verification and action. In this regard, a digitally circulating sign is semantically relevant when it remains answerable to states of affairs that can confirm, disconfirm, or contextually qualify it. Semantic relevance, in other words, is the condition under which information continues to matter for knowing rather than merely for triggering attention or affiliation.

From this perspective, justification is not limited to internal coherence or access to reasons. It concerns the reliability of belief-forming processes within a given ecological niche. A belief is justified insofar as it is produced by processes that reliably track relevant features of the world. Crucially, in a digitally mediated society, these processes include search engines, recommender systems, social media feeds, and collaborative knowledge platforms. The epistemic environment is no longer simply natural or social; it is deeply technological.

An extended conception of knowledge thus requires us to consider at least three intertwined dimensions:

- **Ecological environment**: the physical and biological context in which agents act.
- **Technological environment**: the digital infrastructures and tools that mediate information access and communication.
- **Sociological environment**: the norms, institutions, and communities that shape epistemic practices.

These dimensions jointly constitute the extended cognitive system of the cognitive agents. Within such a system, beliefs emerge from interactions across multiple levels. A social media post, for instance, is not merely a discrete piece of information; it is the product of algorithmic filtering, social endorsement mechanisms (likes, shares), and user expectations shaped by prior interactions. The epistemic status of beliefs formed in this environment depends on the reliability of the entire system, not just on the intentions of individual agents. In this respect, the notion of an extended cognitive system parallels, at a different level of abstraction, Levin's discussion of a cognitive "light cone": what a system can perceive, remember, integrate, and act upon helps determine what it can know and what corrections it can register (Levin 2019; 2022).

The distinction between information and knowledge becomes particularly salient here. Information may circulate abundantly within a digital network, but unless the network supports processes that reliably connect beliefs to relevant environmental states, knowledge does not emerge. Instead, we may witness the stabilization of belief systems that are internally coherent yet environmentally maladaptive.

This externalist and extended view reframes the problem of justification. Justification is not merely a matter of having reasons accessible to consciousness; it is a matter of being embedded in a system that fosters corrective feedback, openness to counterevidence, and responsiveness to environmental constraints. In healthy epistemic systems, errors are detected and revised through interaction with reality and with diverse perspectives. In defective systems, feedback loops become closed, and beliefs are insulated from revision.

Thus, to analyze disinformation, we must ask not only whether specific claims are false or intentionally harmful, but whether the extended cognitive systems in which they circulate are capable of reliable environmental adaptation. If they are not, then the problem is deeper than individual falsehoods; it is structural and systemic.

This systemic perspective sets the stage for a redefinition of disinformation not primarily as false content with malicious intent, but as a state of

epistemic dysfunction within extended cognitive systems.

3. Ignorance Behind Disinformation

Building on the extended conception of knowledge, we can now reconceptualize disinformation through the lens of extended "ignorant cognition" (Arfini 2019, Arfini & Magnani 2022)

Ignorance, in this framework, is not merely the absence of knowledge. It can be an active, structured, and even functional component of cognitive systems. Agents and collectives operate with partial information, selective attention, and bounded rationality. Ignorance may be strategic, unavoidable, or pathological. The crucial question is how ignorance is organized and how it interacts with environmental feedback.

The notion of *ignorant cognition* emphasizes that cognitive systems may become closed in ways that prevent effective interaction with semantically relevant aspects of the environment. Such closure is not simply informational scarcity; it is a structural property of the system. A system is ignorant when it fails to integrate corrective signals, when it selectively filters evidence in a way that reinforces pre-existing beliefs, and when it substitutes internal coherence for environmental responsiveness. Within digital environments and algorithmically curated spaces, the reference to a shared, external world may weaken. Beliefs become calibrated not to environmental constraints but to the dynamics of online communities. In such contexts, justification is internal to the group, and truth becomes secondary if not irrelevant.

Reframed in these terms, disinformation can be understood not merely as the presence of false content but as the state of a collective cognitive system that has become epistemically closed. In such systems:

- Interactions occur primarily within a virtual or self-referential environment.
- Semantic relevance is replaced by engagement metrics or ideological alignment.
- Rational justification is supplanted by emotional reaction.

This redefinition has significant implications.

First, it shifts the focus from *agents who deceive* to *systems that fail to know*. A disinformative environment is one in which the extended cognitive system—comprising users, platforms, algorithms, and social norms—produces and stabilizes ignorance. False beliefs may circulate, but the deeper issue is the structural inability of the system to correct them.

Second, it highlights the role of closure. In ignorant systems, epistemic closure manifests as echo chambers and filter bubbles, where exposure to disconfirming evidence is minimized. However, the problem is not simply exposure; it is the lack of meaningful integration of counterevidence. Even when confronted with opposing views, closed systems may reinterpret them as hostile or irrelevant, thereby reinforcing internal cohesion. In this regard, distinguishing echo chambers from filter bubbles may be appropriate. The two are related but not equivalent. Filter bubbles primarily concern patterned exposure: algorithmic curation, ranking, and personalization narrow the range of materials users are likely to encounter. Echo chambers are stronger social-epistemic formations in which outside voices are not merely absent but actively discredited in advance (Nguyen 2020). For this reason, filter bubbles remain a valid example of structural closure, whereas echo chambers involve a deeper form of immunity to correction.

This distinction also clarifies why "breaking" closure does not automatically improve epistemic conditions. Empirical work associated with Bail shows that exposure to opposing political views can, under some circumstances, intensify polarization rather than reduce it, because hostile content is reinterpreted through identity-protective mechanisms (Bail 2022). I therefore do not treat echo chambers and filter bubbles as interchangeable causes of epistemic failure. The former concern socially reinforced distrust of external sources; the latter concern technologically mediated restrictions of exposure.

On the other hand, the role of technology should also be considered to understand the expansion of ignorant systems (Vaccari & Chadwick 2020). Digital platforms have expanded the scale and speed of information exchange, by also reconfiguring the architecture of epistemic systems. Recommendation algorithms optimize for engagement, not for truth-tracking. Virality may correlate more strongly with emotional arousal than with accuracy. However, the digital revolution did not create ignorance, propaganda,

rumor, or epistemic enclaves ex nihilo. Pre-digital epistemic systems already displayed closure through sectarian communities, partisan presses, rumor networks, and propagandistic state media. What digital platforms changed was not the existence of epistemic pathology as such, but its scale, speed, persistence, and automation. They lowered the friction of publication and circulation, weakened older gatekeeping structures, personalised exposure, and recursively coupled social endorsement with algorithmic amplification (Anderson 2021; Barberá 2020; Vaccari & Chadwick 2020). In this sense, the disruption is real, but it is an intensification and reconfiguration of earlier patterns rather than an absolute historical break.

On the other hand, the disruption seems underlined by the question of responsibility. In an ecosystem of ignorant cognition, responsibility is distributed but asymmetrical. Individual users remain responsible for what they share, for the epistemic virtues they cultivate, and for their willingness to revise beliefs in light of evidence. Yet corporate actors bear a heightened responsibility because they design the infrastructures through which salience, visibility, recommendation, and moderation are organised. Platform companies are not neutral pipes; they structure the effective cognitive environment of publics. Their responsibility is therefore not exhausted by removing illegal content. It also concerns transparency, auditing, ranking design, and the institutional conditions under which public justification remains possible (Gorwa & Garton Ash 2020; Theil 2022).

Crucially, the need of an active responsibility stands out considering again that spreading of ignorance does not necessarily require malicious intent. Even in the absence of deliberate deception, systems may generate and sustain ignorance if their structural incentives undermine epistemic reliability. The Wardle and Derakhshan focus on intent captures an important aspect of strategic manipulation, but it does not fully account for these emergent systemic properties.

Therefore, from the perspective of ignorant cognition, combating disinformation requires more than fact-checking or content moderation. It requires reconfiguring the epistemic architecture of digital environments to reopen systems to environmental feedback and rational justification. This may involve redesigning algorithms, fostering epistemic virtues within communities, and strengthening institutions that mediate between expert knowledge and public discourse.

In short, disinformation is not merely a property of messages; it is a property of systems. It is the condition in which an extended collective cognitive system becomes ignorant—closed, self-referential, and maladaptive.

4. Conclusion

The crisis commonly described as the age of fake news is often approached through technological, political, or sociological lenses. While these perspectives are indispensable, the analysis developed here suggests that epistemology can offer critical perspectives on information disorder studies.

By distinguishing information from knowledge, we expose the limitations of content-based and intentionalist definitions of disinformation. The influential taxonomy proposed by Wardle and Derakhshan provides a valuable starting point, but it does not sufficiently address the normative and systemic dimensions of belief formation. An extended, externalist conception of knowledge reveals that epistemic reliability depends on the interaction between agents and their ecological, technological, and sociological environments.

Within this broader framework, the theory of ignorant cognition allows us to reconceptualize disinformation as the state of an ignorant collective cognitive system. Such systems are characterized by epistemic closure, virtual self-referentiality, and weakened responsiveness to environmental constraints. Disinformation is thus not merely false content with malicious intent, but a structural pathology of extended cognitive systems.

The digital revolution appears, at first glance, to amplify this pathology. The unprecedented scale, speed, and personalization of information flows seem to facilitate the expansion of ignorant systems. Algorithmic curation can intensify closure; attention economies can reward sensationalism over accuracy; online communities can reinforce identity-based epistemic norms. The system of ignorance may appear to easily grow.

Yet digital transformation also presents opportunities for epistemological renewal. The very features that enable rapid disinformation spread—connectivity, data availability, collaborative platforms—also enable new forms of collective intelligence, open science, and participatory verification. The extended nature of cognition, once recognized, becomes a site for deliberate design and intervention. In this regard,

Natural Language Processing (NLP) plays a critical role not just as a set of detection tools but as a means of *epistemic augmentation*. Research in NLP is increasingly oriented toward understanding *perspectivism* — modeling varied human viewpoints in annotation and interpretation — which has direct relevance to epistemic diversity and robustness (Capozzi et. al. 2023, Frenda et. al. 2025). Computational linguistics research also explores model bias and uncertainty in content moderation systems, shedding light on how automated tools may inadvertently reinforce epistemic closure (Konstas et al. 2024, Urbinato et al. 2025).

Epistemology, therefore, must not retreat into abstraction. It must engage with the concrete architectures of digital environments, analyzing how they shape belief formation and justification. At the same time, fake news studies should integrate epistemological reflection, recognizing that the problem is not only political manipulation or technological misuse, but the configuration of collective cognitive systems.

The challenge is to design and cultivate epistemic environments that remain open to correction, responsive to evidence, and capable of adaptive success. This requires interdisciplinary collaboration across philosophy, cognitive science, media studies, computer science, and policy. It also requires a renewed commitment to the normative ideals of truth and justification—not as relics of a pre-digital age, but as guiding principles for the design of extended cognitive systems.

In confronting information disorder, we are not merely managing content; we are shaping the conditions of collective knowledge. The stakes are therefore not only communicative but epistemic. By reframing disinformation as the pathology of ignorant systems within an extended epistemology, we gain both a deeper diagnosis of the crisis and a more robust conceptual foundation for addressing it in the digital age.

5. Bibliographical References

Airoldi, M. (2021). *Machine Habitus: Toward a Sociology of Algorithms*. New York : Wiley.

Anderson, C. W. (2021). Propaganda, misinformation, and histories of media techniques. Harvard Kennedy School Misinformation Review.

Arfini, S. (2017). *Ignorant Cognition: A Philosophical Investigation*. Cham: Springer.

Arfini, S., & Magnani, L. (2022). *Embodied, Extended, Ignorant Minds. New Studies on the Nature of Not-Knowing*. Cham: Springer.

Bail, C. A. (2021). *Breaking the Social Media Prism: How to Make Our Platforms Less Polarizing*. Princeton, NJ: Princeton University Press.

Barberá, P. (2020). Social media, echo chambers, and political polarization. In N. Persily & J. A. Tucker (Eds.), Social Media and Democracy: The State of the Field, Prospects for Reform (pp. 34–55). Cambridge: Cambridge University Press.

Broda, E., & Strömbäck, J. (2024). Misinformation, disinformation, and fake news: Lessons from interdisciplinary research. *Annals of the International Communication Association*, 48(2): 139–156.

Capozzi Lupi, A. T. E., Cignarella, A. T., Frenda, S., Lai, M., & Stranisci, M. A. (2023). Debunker Assistant: A support for detecting online misinformation. In *Proceedings of CLiC-it 2023*, Turin: CEUR, pp. 413–420.

Carter, J. A., Clark, A., Kallestrup, J., Palermos, S. O., Pritchard, D. (eds). (2018). *Extended Epistemology*. Oxford: Oxford Academic.

Clark, A. (2017). *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. Oxford: Oxford University Press.

Clark, A., Chalmers, D. (1998). The Extended Mind. *Analysis*, 58(1): 7–19.

Ferrara, E. (2023). The rise of generative AI and the threat of synthetic misinformation. *Communications of the ACM*, 66(12): 20–23.

Frenda, S., Basile, V., Plank, B., & Stranisci, M. A. (2025). Perspectivist approaches to natural language processing: A survey. *Language Resources and Evaluation*, 59(1), 1–38.

Gorwa, R., & Garton Ash, T. (2020). Democratic transparency in the platform society. In N. Persily & J. A. Tucker (Eds.), Social Media and Democracy: The State of the Field, Prospects for Reform (pp. 286–312). Cambridge: Cambridge University Press.

Konstas, I., Ashraf, S., Ren, A., & Naseem, U. (2024). Large language models for misinformation detection: Challenges and opportunities. *Transactions of the Association for Computational Linguistics*, 12: 845–861.

Levin, M. (2019). The computational boundary of a "self": Developmental bioelectricity drives multicellularity and scale-free cognition. *Frontiers in Psychology*, 10, 2688.

Levin, M. (2022). Technological approach to mind everywhere: An experimentally grounded framework for understanding diverse bodies and minds. *Frontiers in Systems Neuroscience*, 16, 768201.

Nguyen, C. T. (2020). Echo chambers and epistemic bubbles. *Episteme*, 17(2) : 141–161.

Putnam, H. (1978). Reference and Understanding. In Id., *Meaning and the Moral Sciences*. London: Routledge and Kegan Paul, 97–119.

Resch, M. M., et al. (Eds.). (2026). *Trust and Disinformation*. Cham: Springer.

Sunstein, C. R. (2018). *#Republic: Divided Democracy in the Age of Social Media*. Princeton, NJ: Princeton University Press.

Theil, S. (2022). Private censorship and structural dominance: Why social media platforms should have obligations to their users under freedom of expression. Cambridge Law Journal, 81(3), 645–672.

Urbinati, A., Lai, M., Frenda, S., & Stranisci, M. A. (2025). Are you sure? Measuring model bias in content moderation through uncertainty. In *Findings of ACL: EMNLP 2025*. Bangkok: ACL, pp. 980–995.

Uscinski, J., Littrell, S., Klofstad C. (2024). The importance of epistemology for the study of misinformation, *Current Opinion in Psychology*, 57: https://doi.org/10.1016/j.copsyc.2024.101789.

Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation. *Social Media + Society*, 6(1), 1–13.

Wardle, C., & Derakhshan, H. (2017). *Information Disorder: Toward an Interdisciplinary Framework for Research and Policy*. Strasbourg: Council of Europe Publishing.