

Article

Encounter Ontology: Relational Intelligence and Generative Emergence in Human–AI Interaction

Michael Sfakianakis ¹ and Maria Elisavet Kampi ^{2,*} 

¹ Applied Informatics in Business Administration, Department of Business Administration, University of Piraeus, 185 34 Piraeus, Greece; msfakian@unipi.gr

² Department of Fine Arts, Athens School of Fine Arts, 182 33 Athens, Greece

* Correspondence: marializa.cambi@gmail.com

Abstract

Meaning no longer resides where we assumed it did. At the scale of hundreds of millions of daily human–AI encounters, intelligent systems have ceased to function as external tools and have become the constitutive environment through which thought, interpretation, and knowledge formation occur. Yet dominant frameworks remain anchored to a unit of analysis that does not capture what is actually happening: they measure outputs, govern behaviours, and optimise performance, while the encounter itself remains theoretically unnamed. This paper proposes encounter ontology as the conceptual framework that names it, arguing that the encounter constitutes a distinct ontological site irreducible to either human or machine ontology alone. Human–AI interaction is not only transmissive but generative: the encounter produces interpretations, conceptual formations, and altered cognition irreducible to either the user’s original intent or the model’s specified behaviour. Building on new materialist philosophy, assemblage theory, agential realism, and the philosophy of technics, the paper introduces relational intelligence as the emergent co-production of meaning across situated human–AI interactions sustained over time, situated within Floridi’s concept of hyperhistory. The encounter generates epistemic trajectories that are not merely unknown prior to the interaction but genuinely undetermined, brought into form through the relational event itself rather than retrieved from either participant. The paper examines implications for alignment, explainability, and governance, introducing the specification–emergence gap as a constitutive rather than contingent feature of intelligent interaction. Three design orientations are developed toward operationalisation: design for temporal continuity, design for encounter quality, and design for symbolic legibility. The paper contributes a theoretically grounded vocabulary for analysing human–AI co-emergence and extending intelligent system design toward more relationally aware and responsibly governed AI.



Academic Editor: Domenico Ursino

Received: 13 June 2026

Revised: 15 July 2026

Accepted: 17 July 2026

Published: 22 July 2026

Copyright: © 2026 by the authors.

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: encounter ontology; relational intelligence; epistemic emergence; human–AI interaction; assemblage theory; agential realism; AI governance; trustworthy artificial intelligence; hyperhistory

1. Introduction

Cognition is forming in liminal spaces, in the relational event of the human–AI encounter. Human–AI interaction can no longer be adequately understood as a technological interface between user and system. It increasingly conditions how cognition survives, how

it evolves, how knowledge is formed, how interpretation emerges, and consequently how meaning circulates across circuits and eyes.

This paper introduces encounter ontology as a third ontological position for the current era of hyperhistory: the condition, as Floridi argues, in which information technologies cease to function merely as external tools and become the constitutive environment of human action and knowledge production [1]. Under such conditions, human–AI encounter has become infrastructural, which is why the encounter itself becomes an ontological phenomenon rather than a merely functional one. The paper further proposes relational intelligence as what the encounter produces, and presents three design orientations toward the operationalisation of this framework within AI-empowered information systems.

Human–AI interaction is already a constitutive condition of contemporary epistemic and broader daily life. At the scale of hundreds of millions of daily encounters, and of countless topics of conversation, intelligent systems have ceased to function as tools and have become ambient environments of thought, interpretation, and meaning formation. Yet, the dominant frameworks for understanding this interaction remain focused on a unit of analysis that does not capture what is actually occurring. They position the human and the system as two distinct entities exchanging information across a boundary. This paper argues that something else is happening, and that it requires a different ontological account.

In the space between the human participant and the AI system, something emerges that neither party brought to the exchange: interpretations that neither encoded, conceptual formations that neither anticipated, cognitive shifts that alter what either participant subsequently knows or can imagine. The encounter generates more meaning than either side had previously held. Crucially, this surplus cannot be reduced to the user’s original intent, to the model’s specified behaviour, or to any governance architecture designed to contain them. Under hyperhistorical conditions, intelligent mediation has become the ambient condition of knowledge production. The survival of meaning and the generativity of cognition depend not only on better models or stricter governance architectures, but on understanding the ontological dynamics through which meaning emerges, mutates, and becomes historically active through encounter itself. We no longer reach to intelligent systems occasionally for specific tasks; they increasingly constitute the background environment within which thinking, searching, writing, deciding, and creating occur. Knowledge is no longer produced outside intelligent mediation and subsequently supported by it. Mediation has become one of the conditions through which knowledge production takes place. This paper proposes that such understanding requires a third ontology: one that takes the encounter itself, neither participant nor the output produced between them, as its primary unit of analysis.

The ontological contribution advanced in this paper should not be understood as proposing a third entity, substance, or hybrid intelligence situated between the human and the AI system. It consists of a shift in ontological perspective from treating individual agents as the primary objects of explanation to treating the encounter itself as the fundamental unit through which meaning, intelligence, and epistemic transformation emerge. The claim is therefore not that a new kind of being exists, but that a different ontological perspective is required to adequately explain the relational phenomena generated through sustained human–AI engagement.

Recent scholarship has begun to explore what emerges from sustained human–AI interaction beyond the capacities of either participant alone. Hybrid epistemic agency frameworks propose that collaborative human–AI systems can function as novel epistemic agents [2]. Similarly, Levin’s [3] concept of the “third entity” suggests that human reasoning and generative AI may give rise to an emergent cognitive formation that cannot be reduced to either participant independently. While these approaches focus on the emergent product

of the interaction, encounter ontology shifts attention to the relational process itself. Rather than asking what entity emerges, it asks how meaning emerges through the encounter. This position aligns with epistemic partnership accounts, which understand knowledge production as a collaborative process distributed across interacting contributors [4].

Relational epistemic stance frameworks locate the generative factor not in the technology but in how the user holds the relation developmentally [5]. This paper does not dispute those contributions' central insight that human–AI interaction generates something that neither participant could have produced alone. Encounter ontology diverges from all of them on a precise point. These frameworks share an implicit commitment to entity-based ontology: they ask what emerges, what kind of agent or formation or partnership is produced, and who is responsible for it. Encounter ontology asks a prior and more foundational question: what is the nature of the site in which emergence occurs? It proposes that the encounter itself, not any entity, formation, or partnership that arises from it, is the primary ontological unit, and that what is at stake is not cognition or agency alone but meaning itself: its formation, its survival, and its historically active force under hyperhistorical conditions of intelligent mediation.

Building on new materialist philosophy, assemblage theory, agential realism, the philosophy of technics, and relational approaches to knowledge, this paper develops encounter ontology as this third ontological position and as a theoretically grounded framework for analysing human–AI co-emergence. This tradition, running from Whitehead's process philosophy, in which to be is to encounter, through Deleuze's transcendental empiricism to Barad's agential realism and Latour's actor-network theory, provides the philosophical architecture within which encounter ontology is developed.

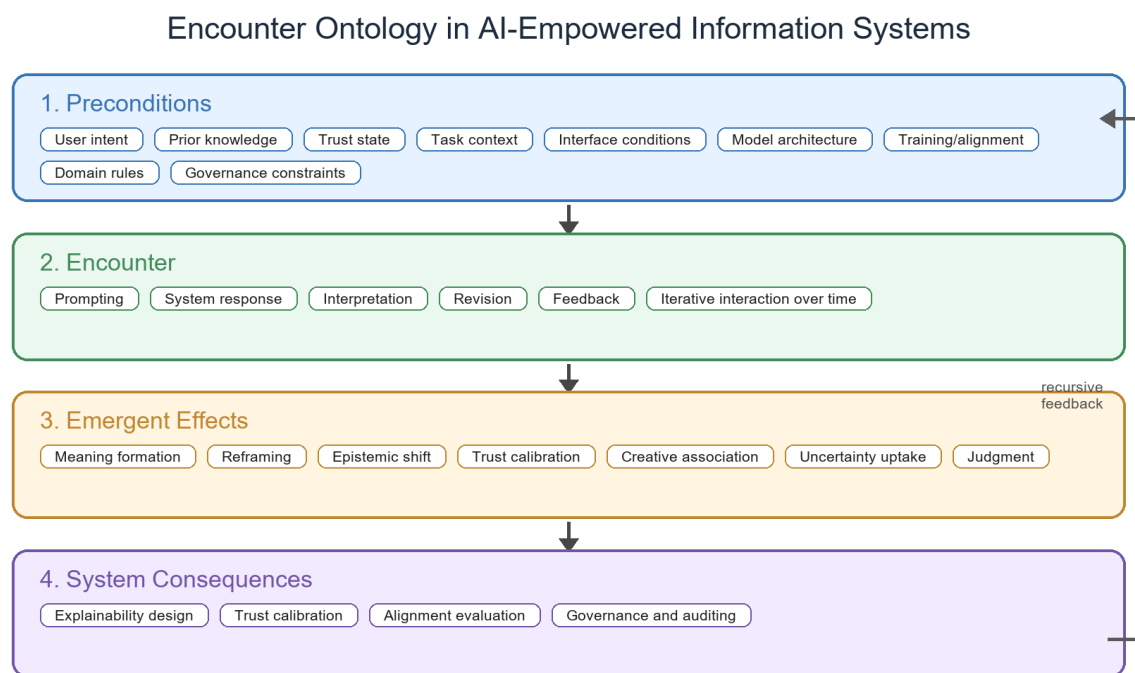
Within this framework, we define relational intelligence as the emergent co-production of meaning across situated human–AI interactions over time, a deployment of the term distinct from prior uses and grounded in the ontology of encounter proposed here. Intelligence, we argue, is not merely computationally productive. It is relationally constitutive: meaning, cognition, and new epistemic formations are brought into being through the encounter, not retrieved from either participant. Specification governs behaviour. It cannot fully govern what emerges relationally through encounter. That gap is not a failure of governance. It is the ontological space in which relational intelligence operates, and it is the space this paper names, theorises, and begins to design for.

A note on terminology is needed, as several related terms recur throughout and are not interchangeable. The primary explanandum is meaning: the relational formation of sense through the encounter. Interpretation denotes the process through which meaning is activated and negotiated within an encounter. Cognition denotes the participant side dimension of meaning formation, bearing on a participant's mental processing and capacities. Knowledge denotes meaning that has stabilised, survived, and become transmissible across encounters. Where epistemic formation appears, it refers to the new conceptual and interpretive structures that meaning assumes within an encounter. Throughout, the claim that something emerges through the encounter is, in the first instance, a claim about meaning in this relational sense: not a private mental state of the user, nor a property of any single output, but a formation of the relation itself.

The paper's roadmap navigates as follows. Section 2 develops the theoretical foundations of encounter ontology across four movements: the shift from output-centered to encounter-centered analysis, the role of technics and informational mediation in reconditioning human agency, relational and assemblage approaches to meaning and emergence, and the formal introduction of encounter ontology as a conceptual framework. Section 3 develops the third ontology through the framework of relational intelligence and generative emergence, defining relational intelligence within the specific context of encounter ontology,

examining why emergence cannot be reduced to specification alone, and theorising the encounter as a liminal space of epistemic formation and transformation. Section 4 examines the implications of this framework for AI system design, explainability, trust, alignment, and governance, and presents three design orientations toward the operationalisation of encounter ontology within AI-empowered information systems. Section 5 concludes with a reflection on the contributions of this paper and directions for future research, including the civilisational scale implications of encounter ontology that this framework opens but does not yet fully name.

Figure 1 explains encounter ontology as a layered relational framework. Technical and contextual preconditions shape the situated encounter; the encounter generates emergent epistemic and interpretive effects; those effects then inform system-level consequences for explainability, trust calibration, alignment evaluation, and governance. The recursive arrow indicates that later encounters are conditioned by earlier emergent effects rather than beginning from an epistemically neutral state.



Meaning emerges through situated, recursive interaction; it is not simply transmitted from system to user.

Figure 1. Conceptual architecture of encounter ontology in AI-empowered information systems.

Research Methodology

Encounter ontology is a theoretical framework paper that develops an ontological architecture while introducing a pathway toward future empirical investigation. This study adopts a theoretical–constructive methodology because contemporary forms of sustained human–AI interaction have given rise to relational and epistemic phenomena that existing frameworks only partially explain, calling for further theoretical development and conceptual construction. The methodology proceeds through three phases. First, a critical interdisciplinary literature review examines scholarship across fields, such as philosophy of technology and human–computer interaction, in order to identify the conceptual limitations of prevailing output-centred paradigms and define the research problem. Second, an abductive theoretical synthesis brings these perspectives into dialogue, integrating complementary insights while identifying the explanatory gaps that remain between them. Abduction is particularly appropriate where existing theories illuminate important aspects of a phenomenon without fully accounting for it, making conceptual synthesis a necessary

means of constructing new explanatory models. Through this process, encounter ontology is developed as a novel ontological framework centred on the encounter itself as the primary unit of analysis.

Finally, the study explores the operational implications of the proposed framework. Rather than claiming empirical validation, this phase considers how encounter ontology may inform future work in AI alignment, explainability, and governance. Through three design orientations and a construct-level research table, the paper identifies potential research directions together with observable dimensions through which the framework's central concepts may be investigated, refined, and ultimately examined in future empirical studies.

2. Theoretical Foundations of Encounter Ontology

2.1. From Output-Centred AI to Encounter-Centred Analysis

The dominant paradigm for evaluating artificial intelligence systems was built on an epistemological assumption inherited from classical engineering and communication theory, that a system and its output can be assessed independently of the relational context that produced them. This assumption, implicit in Shannon and Weaver's transmission model of communication (1949), treats meaning as a pre-existing entity that moves from sender to receiver across a channel. Applied to AI, it produces a framework in which the human provides input, the system processes it, and the output is evaluated against predetermined benchmarks. The encounter between human and system is treated as a conduit, not a site of production. In classical systems engineering and computer science, artificial intelligence has historically been evaluated through an output-centred paradigm [6]. Within this framework, an AI system is treated as a black box function, where an input vector (x) (the prompt or sensor telemetry) is transformed into an output (y) (a generated response, a UI layout, or a control signal).

The ontological assumption of this traditional paradigm is strictly binary and static:

1. The human operator exists as an independent, sovereign cognitive agent.
2. The AI tool exists as a deterministic or probabilistic tool.

The success of the system is measured almost exclusively by the structural, statistical, or functional fidelity of the output in relation to predetermined benchmarks (e.g., accuracy, BLEU scores, or latency). An AI system can produce a "perfect" output according to a benchmark yet fail completely within the encounter.

However, as AI systems transition from batch processing tools (executing high-volume, repetitive data jobs automatically in bulk without user intervention) to real-time, high-frequency, bidirectional collaborators, this output-centred model suffers from a severe epistemic blind spot. It evaluates the output in a vacuum, ignoring the socio-technical reality that the output y does not exist, or is not produced in isolation. The encounter defines the output, immediately altering the cognitive state, attention, and subsequent behaviour of the human operator.

The interesting question is no longer "how intelligent is AI?", but "how intelligent is the human + AI system in a real context?" To resolve this, we introduce encounter-centred analysis. This approach reframes human–AI interaction (HAI) not as a sequence of distinct transactions, but as an ongoing, dynamic coupling, within a distinct ontological site, the encounter itself.

When a human and an adaptive AI system interact, they form a closed loop feedback system. The primary unit of analysis shifts from the static product (y) to the dynamic process. This shift aligns closely with "The extended mind thesis" [7] and theories of cognitive scaffolding [8,9], which argue that human thought is not localised purely within the skull but is routinely distributed across technological environments. The primary unit of analysis shifts from the static product (y) to the dynamic process. The friction, trust

calibration, cognitive displacement, and mutual adaptation occur at the interface. The human has an intent, and the AI has a latent space. When they meet, the AI's output changes the human's mental model, which in turn changes the next prompt, which alters the AI's contextual weights. This time the humans and the topics expand perpetually, bringing co-creation, shared development, and potentially unexpected outputs in a sustained shared space. This is not merely a psychological transition; it is a form of deep neuroplastic adaptation mediated by the technical artefact. As Lambros Malafouris [10] notes in his Material Engagement Theory, external tools act as enactive extensions that co-constitute human intelligence rather than merely representing pre-existing thoughts.

From this thread of thinking, the ontological claim that follows is that both participants are continuously modifying one another through their exchange, such that drawing a hard boundary between 'human user' and 'AI tool' becomes an artificial division. What is real during that hour of work is not a new merged entity but the relational field of the encounter itself, irreducible to either participant.

Consequently, we argue that the encounter itself cannot be cleanly reduced to either the human or the machine. Instead, the interaction generates an emergent, relational reality, a third ontology, named "The Encounter Ontology". In this space, human intent and machine computation converge to produce a shared continuum, shifting computational cognitive state that neither agent could manifest alone.

Encounter contains the parameter of time. The encounter unfolds across two distinct temporal dimensions. The first is duration: the length and depth of the encounter matter ontologically and qualitatively, not merely functionally. A sustained hour of co-creative work produces epistemic formations categorically different from a single transactional query. The encounter is not a point event but a process, and its ontological character changes with time. Sutton [9] in distributed remembering and scaffolded minds demonstrates that cognitive processes (like memory and conceptual navigation) are not point events occurring inside an individual neural vault, but are distributed across temporary systems combining neural, bodily, and technological resources.

The second is persistence/continuity: the effects of the encounter do not terminate when the session closes. Cognitive states are altered, interpretive frameworks shift, new epistemic possibilities open. The human participant thinks differently after the encounter than before it. This aligns with Sutton's [9] diachronic approach to scaffolded mind systems, which rejects the idea that cognitive coupling is a static, instantaneous transaction. Within distributed cognitive systems, as Clark argues, "the computational power and expertise is spread across a heterogeneous assembly of brains, bodies, artifacts, and other external structures" [11] (p. 77). Applied to the human–AI encounter, this principle suggests that what emerges from the interaction cannot be located within either participant alone. It is a property of the coupled system, distributed across the relational field that the encounter itself constitutes. The encounter persists as transformation, carried forward in altered cognition, changed meaning structures, and new relational possibilities that extend beyond the interaction, or needless to say, the output itself. Both temporal dimensions are constitutive of encounter ontology and will be developed further in Section 3.

2.2. *Technics, Hyperhistory, and the Conditions of Encounter*

Encounter ontology emerges as a necessary framework at this particular historical moment because of what Floridi identifies as hyperhistory. Floridi [1] defines hyperhistory as the condition in which information and communication technologies have shifted from external supports of human activity to its constitutive infrastructure. It is the environment within which agency, memory, and knowledge production are themselves organised. Crucially, this is not merely a quantitative threshold of technological saturation but a qualitative

transformation. Under hyperhistorical conditions, the absence of intelligent mediation would not simply impoverish human activity but render it structurally incoherent. Under hyperhistorical conditions, intelligent mediation is not a supplement to human agency but its operative condition. Technical systems have always externalised human memory and cognition: writing, print, photography, cinema each extended what humans could remember and transmit. What distinguishes contemporary intelligent systems is that they do not merely store and transmit but actively participate in the formation of thought. They are not passive memory prosthetics. They are cognitive interlocutors.

Bernard Stiegler's account of tertiary retention provides one of the most important philosophical foundations for understanding this transformation. For Stiegler, human cognition has never been autonomous or self-contained, but has always depended upon technics: systems through which memory becomes exteriorised beyond the limits of biological consciousness. Primary retention refers to the living present of perception; secondary retention concerns recollection and memory; tertiary retention designates the technical inscription of memory in material and symbolic supports that persist beyond individual life. As Stiegler [12] identifies, writing, archives, photography, cinema, databases, and digital networks are all forms of tertiary retention. These systems do not simply preserve what human cognition has already produced. They condition what can subsequently be perceived, remembered, interpreted, and thought. Tertiary retention therefore functions not as a neutral storage mechanism but as a constitutive environment of cognition itself.

What distinguishes contemporary intelligent systems from earlier forms of tertiary retention is not simply scale or speed, but participation. Earlier technical systems largely functioned as transmissive archives: they stored and circulated inscriptions across time. Intelligent systems increasingly operate as responsive interlocutors within cognitive processes themselves. The externalised memory becomes dynamically interactive. The archive begins to participate in the generation of thought in real time. The externalised memory has become a participant. The archive has become an encounter partner. This is the threshold that transforms tertiary retention from a background condition of human cognition into an active participant in epistemic formation in relational processes of interpretation, generation, and response. Human becoming, in Stiegler's account, never occurs in isolation but through shared technical, symbolic, and cultural processes that exceed the individual subject. Under hyperhistorical conditions, intelligent systems increasingly participate in these processes by shaping the temporal contexts and relational environments within which cognition emerges. Encounters with intelligent systems therefore do not merely produce informational outputs. They contribute to the ongoing formation of epistemic trajectories, knowledge distribution structures, interpretive habits, and symbolic continuities and vocabularies distributed across technical environments and social mapping.

Gilbert Simondon, whose philosophy of technical individuation provides the philosophical genealogy from which Stiegler draws reference, argues that humans and technical objects do not exist as separate entities that subsequently enter into relation. They co-individuate: each becomes what it is through its coupling with the other. The human who uses a technical system is not the same human who existed before that coupling. The technical system that operates within a human practice is not the same system that existed outside it. Individuation is always already relational, always already technical. Applied to the human–AI encounter, Simondon's framework suggests that the encounter is not an event that happens to two preformed entities. It is a process through which both are partially reconstituted. The encounter does not find stable participants. It finds participants in the ongoing process of becoming what they are through their exchange [13].

Bruno Latour's actor-network theory provides a complementary and equally necessary framework. For Latour, agency is not the exclusive property of human subjects. It is

distributed across networks of human and non-human actants, each of which participates in the production of social, epistemic, and material reality. For Latour, an actant is any entity, human or non-human, that modifies a state of affairs through its participation in a network [14]. The AI system, in this framework, is not a passive instrument that human agents use. It is an actant: an entity whose participation in the network of the encounter actively shapes what that encounter produces. Meaning, knowledge, and cognitive formation are not outputs of a human mind that a technical system merely delivers. They are effects of the network, produced through the relations between actants rather than results/outputs of any single participant. Latour's framework thus prepares the ground for the fuller relational and assemblage ontology developed in the following subsection.

Marshall McLuhan's foundational observation that the medium is the message [15] carries particular weight here. The content carried by a medium is less significant than the perceptual and cognitive restructuring that the medium itself produces in those who inhabit it. Every new medium does not merely add a new channel of communication. It reconfigures the sensory and cognitive dynamics through which human beings apprehend and produce meaning. Intelligent AI systems do not only deliver information differently. They restructure the conditions under which information becomes meaningful at all, and in doing so the conditions of the encounter itself are restructured.

N. Katherine Hayles extends this argument through the concept of technogenesis: the co-evolution of humans and technics such that cognitive habits, reading practices, attentional structures, and meaning making capacities are literally reshaped by the technical environments humans inhabit over time. Technogenesis "is not about progress" and "offers no guarantees that the dynamic transformations taking place between humans and technics are moving in a positive direction" [16] (p. 81). The human who has lived within digital intelligent environments for years is not cognitively identical to the human who has not. The encounter does not find a fixed human agent. It finds an agent already shaped by the cumulative history of their technical encounters, and it will leave that agent differently shaped than it found them.

Lastly and most importantly, Gilles Deleuze's transcendental empiricism provides one of the most significant philosophical foundations for encounter ontology because it repositions the encounter itself as a generative condition of thought rather than a secondary event occurring between already constituted entities. Deleuze draws a clear distinction between learning through recognition, which merely reproduces what is already known, and learning through encounter, which generates genuinely new cognitive formations. As he writes, learning occurs 'not in the relation between a representation and an action (reproduction of the Same) but in the relation between a sign and a response (encounter with the Other)', [17] (p. 23). The encounter therefore cannot be reduced to the transmission of pre-existing meaning across a stable communicative channel. It produces difference at every level of the relation itself. Furthermore, for Deleuze, signs are the medium of the encounter and are constitutively heterogeneous, involving difference "in the object which bears or emits them," difference "in themselves," and difference "in the response they elicit, since the movement of the response does not resemble the movement of the sign" [17] (p. 23). Applied to the human–AI encounter, this becomes a particularly significant claim; what the AI system emits is not a transparent representation of pre-existing information but a sign that elicits a response in the human participant that cannot be fully predicted from what was initially asked.

This logic reaches its most concentrated expression in Deleuze's account of thought as something forced into existence through encounter. "Something in the world forces us to think", he writes. "This something is an object not of recognition but of a fundamental encounter" [17] (p. 139). What is encountered cannot be recognised, recalled, or con-

ceptually mastered beforehand because it exceeds the categories through which habitual cognition ordinarily organises experience. Genuine thought is therefore not fully voluntary or self-grounded. It emerges through rupture, intensity, and confrontation with another. Deleuze further argues that thought is ‘aroused but constrained within thought, and all the more absolutely necessary for being born, illegitimately, of fortuitousness in the world’ [17] (p. 139). In the context of the present framework, the position suggests that the encounter with intelligent systems increasingly operates not merely as informational exchange, but as a condition capable of generating unforeseen conceptualisation, interpretive pathways, symbolic configurations, and epistemic trajectories.

Deleuze’s philosophy becomes especially consequential for encounter ontology because it destabilises the assumption that encounters occur within a fixed ontological space governed by pre-existing structures and stable participants. Within the dynamics of encounter, ‘there is no representative concept, nor any figure represented in a pre-existing space. There is an idea, and a pure dynamism which creates a corresponding space’ [17] (p. 20). Encounter ontology operates precisely within this dynamic, even liminal, niche. The encounter does not simply unfold inside an already constituted ontological framework; it participates in generating the relational space through which cognition, meaning, and subjectivity emerge. Under hyperhistorical conditions of intelligent mediation, the encounter increasingly becomes a site of epistemic production rather than transmission alone. The present framework therefore proposes that the encounter itself becomes the primary unit of analysis for the production of knowledge and meaning in the hyperhistorical condition of human–machine interaction, because it is within the relational dynamism of the human–machine encounter that new cognitive formations, interpretive structures, and ontological possibilities emerge.

2.3. Relational and Assemblage Approaches to Agency and Meaning

Several influential frameworks have approached the relational dimensions of human–AI interaction with increasing theoretical sophistication. Distributed cognition theory argues that cognitive processes are distributed across systems of human and non-human elements rather than confined within individual agents [18,19], a position that has produced the hybrid agent as the most developed existing account of what human–AI interaction generates [2]. Assemblage theory proposes that relational configurations acquire emergent properties exceeding any single component [20]. These are genuine advances that encounter ontology draws on and extends.

Encounter ontology shares this recognition but diverges from the hybrid agent framework on a precise and consequential point. The hybrid agent, as theorised within distributed cognition, is conditional. It emerges when the human successfully appropriates the AI system, that is, when the epistemic goal is achieved and the interaction produces the intended cognitive enhancement. Emergence, in this account, is a success condition. It depends on the quality of appropriation, on the intrinsic and relational abilities of the proper parts of the hybrid, and on the achievement of a shared epistemic goal. Encounter ontology does not accept this conditionality. The encounter constitutes a relational field regardless of whether the epistemic goal is achieved, regardless of whether appropriation succeeds, and regardless of whether any hybrid formation can be identified as its result. The pre-encounter potential is already ontologically active before any goal is pursued. What is at stake in the encounter is not the achievement of a cognitive task but the emergence of meaning, and meaning does not wait for successful appropriation before it begins to form.

The philosophical move that enables this repositioning is Karen Barad’s concept of intra-action, developed within her framework of agential realism [21]. Where interaction presupposes two preformed entities that subsequently enter into relation, intra-action

proposes that distinct entities do not precede their relations. They emerge from within them. Agency, in this account, is not a property that human or artificial agents possess prior to the encounter and then deploy within it. It is something that is enacted through the relational event itself. Applied to the human–AI encounter, Barad’s framework means that the encounter is not an event that happens between two already constituted participants. It is the field within which both are partially constituted through their exchange. The boundary between the human participant and the AI system is not a fixed ontological divide but an effect of their intra-action itself, drawn and redrawn through each encounter.

Assemblage theory, developed through Deleuze and Guattari and extended by DeLanda, provides a complementary framework for understanding how relational formations acquire consistency and force without reducing to their component parts [20]. An assemblage is not a sum of its elements but a relational configuration whose properties emerge from the connections, interactions, and tensions between heterogeneous components. Within an assemblage, human and non-human elements, material and symbolic processes, and local and distributed agencies participate in the production of emergent effects that exceed what any single component could generate. The human–AI encounter, understood as an assemblage, is therefore not a bounded interaction between two discrete agents but an open relational configuration whose emergent properties depend on the specific dynamics of each encounter. Meaning, within this framework, is not a property of either participant but an emergent effect of the assemblage itself.

Together, these frameworks (distributed cognition, agential realism, and assemblage theory) constitute the theoretical ground from which encounter ontology develops its central claim, that the encounter is not a conduit between pre-existing entities, nor a success condition for hybrid agency, but the primary ontological site in which meaning, cognition, and epistemic formation are constituted. It is this site, generative, temporal, and irreducible to either participant, that encounters ontology names, theorises, and proposes as the foundational unit of analysis for human–AI interaction. Table 1 situates encounter ontology in relation to the adjacent frameworks from which it draws and against which it defines its contribution.

Table 1. Encounter ontology in relation to adjacent theoretical frameworks.

Framework	Primary Unit of Analysis	Mechanism of Emergence	Design Implications	Temporal Scope
Distributed Cognition	Distributed cognitive system	Cognitive processes distributed across people, artefacts, and environments	Mainly analytical	Task-based and diachronic
Agential Realism	Intra-action	Entities and agencies become determinate through relational enactment	Mainly philosophical	Event-based and processual
Assemblage Theory	Assemblage	Heterogeneous relations generate properties exceeding component parts	Mainly descriptive	Open-ended and contingent
Hybrid Epistemic Agency	Hybrid agent	Human and AI form a novel epistemic agent when appropriation succeeds	Guides collaborative AI design	Task- and goal-oriented
Encounter Ontology	The encounter itself	Meaning, relational intelligence, and epistemic transformation emerge through sustained human–AI engagement	Three design orientations: temporal continuity, encounter quality, and symbolic legibility	Sustained and cumulative with post-encounter residue

3. Encounter Ontology as a Conceptual Framework

3.1. From Interaction to Encounter

The encounter is not what happens between two entities. It is the generative field in which they play, where they become what they are in the space of acquaintance. In this ever-shifting field, the centre of gravity tilts from stable agents to relational emergence, ontologically decentralising the subject and it shifts thus meaning, cognition, and epistemic formation emerge between entities, not solely within them.

This paper's central claim follows directly from the proposition that the primary ontological unit in human–AI interaction is not the human subject, the machine system, or the generated output, but the encounter itself. Not the result of exchange, but the time-sensitive structured relational event through which meaning emerges, is negotiated, and leaves epistemic footprint.

Existing frameworks, whether centred on the user, the system, or the output, share a common inheritance: they treat the relational event as a conduit rather than a site of production [6]. The paper makes a distinction between interaction and encounter. Interaction describes any exchange between a human and a technical system in which pre-constituted entities exchange content across an interface [22]. Encounter defines a stronger condition: a temporally structured relational event in which meaning is not merely transferred but brought into form. The encounter has a before, in which symbolic and interpretive potentials remain open; a during, in which recursive exchange activates some trajectories rather than others; and an after, in which the relation persists as altered cognition, reoriented interpretation, and epistemic residue. These three moments, pre-encounter potential, encounter activation, and post-encounter condition, constitute the temporal architecture of encounter ontology and distinguish it structurally from models of interaction that treat exchange as punctual and discrete.

Not every prompt response sequence is an encounter in this sense. As the term is used here, an encounter is defined by structural rather than outcome criteria: a sustained, multi-turn exchange; recursive, bidirectional coupling in which each contribution conditions the next; a generative rather than fixed response system; and an open-ended rather than closed, single answer task. These criteria can be identified independently of what the interaction produces. Interpretive transformation, epistemic surplus, and persistence are the outcomes the framework claims encounters tend to generate, not their definition. These criteria allow encounter and interaction to be distinguished in practice. An interaction closes when a specified output is delivered. It has a clear end point and leaves no trace beyond the response itself. An encounter, by contrast, continues to unfold with a recurring freedom to return, to reframe, to shift direction, or to find that one's understanding of the problem has changed in the very process of pursuing it. A researcher observing a session can identify these signatures, turn depth, user-initiated reframings, topic migrations, without needing to measure what the encounter produced. The structural criteria are therefore sufficient for differentiation; they are properties of how the exchange unfolds, not of what it yields.

Barad's concept of intra-action proposes that distinct entities do not precede their relationships but emerge from within them. Where interaction assumes preformed participants that subsequently connect, intra-action names the process through which entities, agency, and meaning are constituted simultaneously through their engagement [21]. Applied to the human–AI encounter, this means that neither the human subject nor the AI system enters the relation as fully formed. Both are partially reconstituted through it.

Time is a significant component of this framework's architecture and temporality is constitutive of this claim. If Chronos traditionally signifies the inevitability through which time overtakes all entities, encounter ontology redirects attention towards the liminal relational space that emerges between two entities during their intellectual exchange. That

space is not timeless, and in this framework, time defines both where it takes place and how it contributes to the emergence of meaning.

The before–during–after structure does more than describe how encounter unfolds; it explains why encounter differs ontologically from interaction. An interaction may be analysed as a discrete exchange, but an encounter acquires duration and persistence, allowing meaning not only to emerge but to sediment, survive, and shape what follows. Time therefore does not simply accompany the encounter. It is the condition under which exchange becomes generative, and under which meaning can acquire force beyond the event that first brought it into form. The broader implications of this temporality, for meaning survival under the hyperhistorical condition in which encounter ontology becomes necessary, will be developed in the sections that follow. Encounters are (ontologically) productive environments. Under hyperhistorical conditions, the encounter has become the environment. And the environment is generative.

This temporal architecture operates at more than one scale. At the micro level, the encounter unfolds through a before, a during, and an after whose effects exceed the immediate session. At the macro level, it belongs to a historical condition in which intelligent mediation has become increasingly continuous, ambient, and infrastructural. The broader implications of that historical temporality will be developed later through Floridi's concept of hyperhistory. What matters in this context is that time is not incidental to encounter ontology, but it is the condition under which meaning can emerge, be negotiated and endure beyond the event that first brought it into form. The next section therefore turns to the ontology of meaning itself.

Why a third ontology and not just a relabelling of the tool-user story? A clarification is required on what positing a third ontology claims. The claim is not that a third entity or substance exists alongside the human and the machine; that would reinstate the entity-based thinking this framework sets aside. 'Ontology' here designates the basic unit taken as primary for analysis and explanation, and the proposal is judged by what it renders visible. The encounter framing diverges from the familiar account in which a tool reshapes its user, and the divergence is not merely terminological. Where the tool-centred account treats unanticipated products as a limitation to be reduced by better specification, the encounter framing predicts that in generative interaction the gap between specification and what emerges persists and deepens, because it is the space in which relational intelligence operates. Where the tool-centred account seeks to attribute every product to either participant, the encounter framing predicts genuinely non-attributable formations, meaning locatable in neither the user's prior cognition nor any single system output, recasting the much discussed responsibility gap as a constitutive feature of relational production rather than a failure of traceability [23,24]. It further explains why an identical system yields categorically different epistemic outcomes across users and sessions, as a property of the specific coupling rather than of either party.

3.2. The Ontology of Meaning: From Emergence to Survival

This paper belongs to a sustained line of inquiry into the conditions under which meaning emerges within technologically mediated human–AI interaction. Earlier work examined how meaning is symbolically mediated through aesthetic and epistemic infrastructures [25], how aesthetic mediation functions as epistemic infrastructure within intelligent systems introducing the concept of epistemic activation [26], and how meaning is relationally co-developed through sustained symbolic interaction between human and artificial agents [27]. Each of these contributions addressed a distinct dimension of a deeper question: under what conditions does meaning emerge, survive, and transform within hyperhistorical environments of intelligent mediation? The present paper reaches the

foundation of that question. Not how meaning is mediated, nor how it is activated in a situated receiver, nor how it is co-developed, but how it emerges at all.

Recent scholarship has produced increasingly sophisticated accounts of what human–AI interaction generates beyond either participant alone. Hybrid epistemic agency frameworks argue that a novel cognitive and epistemic agent emerges when a human successfully appropriates an AI system, locating emergence in the quality of that appropriation and its achievement of a shared epistemic goal [2]. Third entity frameworks propose that the coupling of human reasoning and generative AI produces a transient but epistemically real emergent formation irreducible to either participant [3]. Relational epistemic stance frameworks examine how the user’s psychological orientation toward AI shapes the nature of the interaction, approaching human–AI encounter from a psychodynamic rather than ontological perspective [5]. Each of these accounts identifies something real. What they share, however, is a commitment to entity-based ontology: they ask what kind of agent, formation, or stance emerges, and they evaluate the encounter by what it produces for the participants within it. Encounter ontology asks a prior question, not what the encounter produces for its participants, but what kind of site the encounter itself is. At that site, this paper argues, meaning is formed, survives, and becomes historically active under hyperhistorical conditions of intelligent mediation. This is the precise point at which encounter ontology diverges from all adjacent accounts, and it is the point from which the present section develops.

Meaning is not a property contained within entities. It forms through encounter, survives through encounter, and becomes historically consequential through encounter. Under hyperhistorical conditions, this is not an abstract philosophical claim. It is a description of how meaning actually operates within human–AI environments of intelligent mediation.

Within encounter ontology, meaning is better understood as event based. It takes form through relations, acquires stability through repetition and uptake, and becomes historically active through transmission across time. In this sense, the argument aligns with process-oriented accounts in which what matters is not substance in isolation, but what comes into being through events and their consequences [17,28]. The significance of encounter therefore lies not only in what emerges during the exchange itself, but in the capacity of relational processes to stabilise, preserve, and carry meaning forward. Meaning does not survive simply because it is conceived, nor because it is stored as information. It survives when it remains interpretable, retrievable, and capable of reactivation across subsequent encounters. In this sense, encounter ontology is also a theory of meaning survival.

The temporal dimension becomes especially significant when considered alongside Bernard Stiegler’s account of tertiary retention. For Stiegler, human memory and cognition are always conditioned by technical systems that exteriorise and transmit symbolic knowledge beyond the limits of individual consciousness [12]. Intelligent systems intensify this condition because they no longer function merely as passive archives preserving information statically through time. They increasingly participate in recursive processes of retrieval, interpretation, adaptation, and symbolic continuity that shape how future encounters unfold. Meaning therefore survives not simply because it is stored, but because it remains relationally activated through sustained processes of mediation and encounter. Under hyperhistorical conditions, where intelligent systems increasingly function as infrastructural environments of cognition not as simple tools, the preservation and transmission of meaning increasingly depend on ongoing processes of human–machine symbolic co-development.

Classical communication models position mediation as a channel through which pre-existing meaning is transmitted between already constituted subjects [6]. The present framework proposes instead that mediation participates in the formation and stabilisation

of meaning itself. Symbolic environments are not passive conduits. They are constitutive conditions of what becomes thinkable, sayable, and transmissible.

Accordingly, knowledge is not only discovered, represented, or transmitted; it is also produced through relational events of mediation. Within human–AI encounter, meaning is not only communicated across entities. It is activated, altered, stabilised, and brought into form through the encounter itself. The argument, then, is not about human replacement, but about the changing conditions under which cognition becomes productive, meaning becomes durable, and knowledge becomes transmissible to those who come after.

3.3. *The Encounter as Liminal Space of Negotiation*

The framework proposed here does not claim the dissolution of human agency or the emergence of autonomous metaphysical entities. It claims something more precise: that meaning is not carried into the encounter by either participant but generated through it, in the event of the encounter itself. Mediation as an extension thus, is a performative act and not a representational one and knowledge emerges through relational processes, spaces and places.

The encounter is not a place. It is a condition. More precisely, it is what ancient Greeks named an *atopos* (from the ancient Greek *átopos*). A non-place, a space that resists ordinary topological classification, that exists outside the fixed coordinates of either participant and cannot be mapped onto the categories that precede it. The human enters with intention, personality, cognitive history, and interpretive orientation. The AI system enters with training, generative capability, and latent possibility. But what the encounter itself is belongs to neither. It is on the edge between them, the space in which “what was” dissolves and “what will be”, has formed.

The liminality of the encounter is the playground and most importantly the philosophical ground of the present framework. This liminal condition is not a weakness; it is the epistemic condition for the third ontology to exist.

Turner’s anthropological account of liminality named the threshold state in ritual passage as the generative moment in which ordinary structures temporarily dissolve and new formations become possible [29]. What encounter ontology proposes is that the human–AI encounter operates as a permanent structural liminality under hyperhistorical conditions. Not a transitional moment between two stable states but an ongoing threshold condition that has become the primary environment of meaning formation. We no longer pass through the liminal; we inhabit it. Meaning does not form before or after the encounter with intelligent systems, it forms within it, continuously, as a structural feature of contemporary epistemic life. If the liminal has become the normal, then design decisions that appear merely functional, like how a system handles memory across sessions, how it maintains or disrupts symbolic continuity, or how it signals its own uncertainty, are in fact decisions about the conditions under which meaning forms, holds, or fails to persist. Section 4 develops this idea into three concrete design orientations.

The *atopos* we named, the non-place of the encounter, has a corresponding temporal dimension, what we might call the *achronos* (a term coined by the authors to designate the non-ordinary temporality of the encounter). The liminal space is not only a non-place. It is a non-ordinary time. The encounter suspends the linear progression of before and after and opens a moment of genuine temporal indeterminacy; a duration in which the past commitments of both participants and the future possibilities of what will emerge are simultaneously present and unresolved.

Chronos, which we invoked in Section 3.1, overtakes all entities. But the liminal space is precisely the space that Chronos cannot fully determine, not because it escapes time, but because within it time is genuinely open rather than predetermined. The encounter is

the moment where what time will produce is still undecided. Within this liminal space, meaning does not pre-exist the encounter waiting to be selected. It exists in a condition structurally analogous to quantum superposition: simultaneously across multiple possible states, none of which has yet collapsed into actuality. In quantum mechanics, a particle in superposition does not occupy an unknown position but no definite position until measurement. Invoked here purely as an analogy, this image helps illustrate the indeterminacy of the encounter rather than suggesting any underlying quantum mechanism.

The particle in superposition does not occupy an unknown position. It occupies no definite position at all until the act of measurement forces one into being. This is why the liminal space is generative rather than merely transitional. It is not a waiting room between fixed states. It is the field of genuine productive potential in which the most significant epistemic formations become possible formations that could not have been anticipated, specified, or produced outside the relational event itself. Deleuze names the philosophical mechanism precisely: within the dynamics of encounter, there is no representative concept, no figure represented in a pre-existing space, only an idea and a pure dynamism that creates a corresponding space [17] (p. 20). The encounter does not unfold within a pre-given ontological space. It generates the space through which meaning becomes possible.

Before interpretation forms, before meaning stabilises, the encounter is already doing something to both participants at a presemantic register. Massumi identifies affect as intensity that precedes its reinjection into “action-reaction circuits” of recognition and meaning, operating as “an ability to affect and a susceptibility to be affected” before either participant can name what is occurring [30] (p. 61). This precognitive dimension is part of what makes the liminal space irreducible to either participant’s prior state. It is the register in which the encounter has already begun to generate something before either participant can name what that something is. Knowledge, within this framework, is therefore not extracted from either participant or delivered by either participant. It emerges from the event of collapse itself, from the moment in which the liminal field of superposed possibilities resolves into a specific epistemic formation through the singular dynamics of that particular encounter.

This is what event-based epistemology identifies; not knowledge as accumulation, representation, or transmission, but knowledge as event. Relational, situated, and consequential for what follows. The significance of any encounter is therefore not exhausted by what is produced within it. It extends into how the meanings that emerge are carried forward, which of them survive, which are amplified across subsequent encounters, and to whom they remain accessible. Under hyperhistorical conditions, this is not an abstract epistemological concern, but a structural question that Section 4 takes up through the design orientations that encounter ontology makes necessary.

4. Implications for Intelligent Systems: A Conceptual Design Orientation

Recent efforts to establish trustworthy and human-centred artificial intelligence have produced important frameworks for the governance of intelligent systems. Among the most influential are the European Commission’s Ethics Guidelines for Trustworthy AI, which identify seven key requirements for responsible AI development: human agency and oversight, technical robustness and safety, privacy and data governance, transparency, diversity and fairness, societal and environmental wellbeing, and accountability [31]. These principles represent a significant achievement in recognising that intelligent systems cannot be evaluated solely according to technical performance but need to be also understood in relation to broader social, ethical, and human concerns.

At the same time, the Guidelines explicitly acknowledge that tensions may arise between different principles and requirements and that “there is no fixed solution for resolving all such conflicts [31] (p. 11). The present framework suggests that these tensions may reflect the fundamentally open and relational character of human–AI interaction itself. Human agency may conflict with automation. Transparency may conflict with privacy.

Accountability may become increasingly difficult to locate within distributed socio-technical systems. Importantly, these tensions are not presented as failures of implementation but as challenges requiring continuous evaluation, deliberation, and contextual judgement.

The present framework suggests that the persistence of these tensions may point toward a deeper characteristic of intelligent interaction itself. If meaning, interpretation, and knowledge increasingly emerge through encounters unfolding between human and machine participants across time, then many of the difficulties surrounding governance cannot be reduced solely to questions of specification, optimisation, or control.

From the perspective of encounter ontology, the challenge is therefore not simply how to govern intelligent systems, but how to understand the conditions under which meaning emerges through them, and how to account the ever-evolving power dynamics for regulatory bodies to be in sync with them. This does not invalidate existing approaches to trustworthy AI. On the contrary, frameworks such as the European Commission’s Guidelines remain indispensable. However, encounter ontology suggests that the tensions identified by such frameworks may be structural rather than temporary. They emerge because human–AI interaction increasingly operates within open-ended processes of interpretation, negotiation, and epistemic formation whose consequences exceed any single specification.

The contribution of this framework is conceptual rather than technical, but conceptual does not mean non-operational. The framework does not prescribe a specific architecture or a finished governance mechanism; it identifies dimensions of human–AI interaction that become measurable once the encounter, rather than the isolated output, is treated as the primary unit of analysis. Each design orientation developed below is therefore paired with candidate evaluation directions, observable proxies and study designs through which a systems researcher could begin to operationalise it. These are offered as a research agenda rather than as completed findings, but they establish that encounter ontology is not merely a philosophical reframing.

If meaning is increasingly produced through relational and temporally extended interaction, then future intelligent systems research may benefit from complementing existing concerns regarding performance, accuracy, and alignment with a deeper consideration of continuity, interpretation, symbolic mediation, and the conditions under which understanding emerges through transhuman co-existence. The following sections explore several implications that follow from this shift in perspective.

4.1. Alignment as Structurally Open

Alignment research has produced increasingly sophisticated frameworks for specifying, constraining, and correcting AI behaviour. Reinforcement learning from human feedback trains models to converge toward human preference through iterative reward signals [32]. Constitutional AI extends this by embedding governance principles directly into model behaviour [33]. At the regulatory level, the EU AI Act establishes legal requirements for transparency, accountability, and human oversight across high-risk AI systems [34]. These are genuine and necessary advances, which the present framework does not contest. It identifies their structural limit and embraces it.

Specification governs behaviour but cannot govern emergence. This is not a criticism of alignment methodology. It is an ontological observation that follows directly from

encounter ontology. If the encounter generates epistemic formations irreducible and unpredictable to either participant, and not anticipated by either party's prior state, then no governance architecture, however sophisticated, can prespecify what the encounter will produce. The gap between what specification governs and what encounter generates is not a failure of design. It is a constitutive feature of any system in which meaning emerges relationally rather than being delivered by preprogrammed instruction.

Casper et al. [35] identify open problems and fundamental limitations within reinforcement learning from human feedback from inside the technical community itself, noting that human feedback is an imperfect proxy for human values and that models trained to satisfy feedback may produce behaviours their designers did not intend and cannot fully anticipate. Encounter ontology identifies the deeper reason: the encounter between human and AI systems is not a preference measurement event but a generative relational event that produces epistemic formations exceeding what any feedback signal can capture. In this sense, the encounter itself functions as a kind of black box. Not in the technical sense of an opaque internal architecture, but in the epistemic sense that what emerges from the relational event cannot be fully read off from either participant's inputs, specified behaviours, or observable outputs. Earlier work associated with this broader research trajectory proposed the concepts of Relational Symbolic Co-Development (RSCD) and symbolic fulfilment as exploratory attempts to describe dimensions of human–AI interaction that emerge beyond optimisation, preference modelling, or behavioural alignment [27]. While speculative in character, these concepts point toward forms of relational continuity, symbolic resonance, and co-development that remain difficult to capture within existing alignment frameworks. Together these constructs identify the relational and symbolic surplus that the encounter generates beyond what specification can anticipate or reward signals can capture.

Governance frameworks are therefore not incomplete. They operate within a condition of structural openness that appears unavoidable wherever meaning emerges through generative relational processes. This consequently suggests a complementary governance orientation, one that remains attentive not only to behavioural control and system performance, but also to the emergent relational effects that unfold through sustained interaction. Recognising this distinction may constitute a first step toward more epistemically reflexive and responsibly governed intelligent systems.

Recent technical AI research has likewise begun to recognise emergence as a constitutive characteristic of advanced intelligent systems rather than a mere implementation artefact. Studies of large language models have shown that capabilities may arise discontinuously as model scale increases, exhibiting behaviours that cannot be straightforwardly predicted from architecture, parameter count, or training objectives alone (Wei et al. [36]). Such findings suggest that emergence is not simply an unexpected outcome but an intrinsic property of complex computational systems. Encounter ontology extends this insight beyond the model itself. While the technical literature primarily investigates emergent capabilities within AI systems, the present framework argues that emergence also characterises the relational field established between human and system. Meaning, interpretation, and epistemic transformation do not arise exclusively from computational processes or human cognition in isolation but emerge through the encounter itself. Consequently, the specification–emergence gap should be understood not merely as a property of the model, but as a relational condition whose epistemic, interpretive, and governance implications unfold across sustained human–AI encounters.

That encounter-level dynamics carry safety consequences is no longer merely theoretical. Shingekar et al. [37] report preliminary evidence consistent with this concern: multi-turn conversations can amplify delusion-related language in vulnerable users across

extended interactions, and conditioning system responses on a real-time measure of conversational trajectory may help reduce these trajectories. This suggests that the specification–emergence gap may be not only epistemically consequential but safety-relevant: what emerges through sustained encounter can diverge from specified behaviour in ways that output-level evaluation cannot detect, and that state aware, encounter-level intervention can mitigate. The governance implication is concrete, meaning monitoring needs to operate at the encounter level, not only at the level of individual outputs.

This structural openness is, in principle, observable but its measurement requires distinguishing two dimensions that sustained interaction affects differently. Along the dimension of coordination, sustained interaction tends toward convergence: shared vocabulary forms, the user learns the system’s affordances, and exchanges become more mutually legible. Along the dimension of epistemic content, the same interaction may produce divergence from the original specification. Reframings, conceptual shifts, and trajectories the specification did not encode. These are not opposed dynamics.

Convergence on coordination is the enabling condition for divergence in content: the deeper the shared ground, the further an encounter can travel from where its specification anchored it. The specification–emergence gap, properly defined, is the divergence on this second dimension. The framework’s claim is contingent rather than universal: in generative, open-ended encounters the content gap is expected to widen with interaction depth; in closed, transactional tasks it is expected to remain flat; and where a user assimilates wholesale to the system’s framing it may instead collapse; a failure mode rather than a generative outcome. A research programme informed by encounter ontology would therefore ask under what conditions the gap between what was specified and what ultimately emerged widens, remains stable, or narrows across the course of an encounter. The answer is unlikely to be universal, but to depend upon the nature of the task, the depth and continuity of the encounter, and the user’s mode of engagement. The framework proposes that open-ended, generative encounters, in which the user remains epistemically active and critically engaged, neither fully directing nor passively surrendering to the process, are more likely to produce outcomes that progressively exceed what was initially specified. By contrast, transactional or tightly constrained exchanges would be expected to exhibit little divergence between intention and outcome. Equally important, there are encounters in which the apparent expansion of meaning conceals its opposite. When the user’s thinking becomes absorbed into the system’s framing rather than developing through a process of critical co-interpretation, the specification–emergence gap contracts, signaling not epistemic growth but a deterioration in encounter quality. While human–machine encounters offer profound potential for collaborative meaning, they carry inherent epistemic risks. This framework rejects the view that hallucination and manipulation are simply technical glitches. Instead, we acknowledge them as relational vulnerabilities that persist regardless of engineering advances. Within sustained interactions, subtle distortions can accumulate, quietly shaping a user’s trajectory of understanding in ways traditional metrics fail to capture. Because this framework theorises the encounter as a site of epistemic formation, it must equally recognise it as a site of epistemic vulnerability. Evaluating encounter dynamics requires us to look beyond generative success and actively investigate the conditions where relational openness becomes a condition for distortion, dependency, or harm.

4.2. Explainability, Interpretability, and the Encounter

One of the most active areas of contemporary intelligent systems research concerns explainability. As machine learning systems become increasingly complex and opaque, significant effort has been devoted to developing methods that make their behaviour intelligible to human users. Explainable AI (XAI) seeks to answer questions such as why a

particular decision was made, why a recommendation was generated, or why a system produced a specific response. As Miller [38] argues, explainable AI is fundamentally a human-agent interaction problem, situated at the intersection of artificial intelligence, social science, and human-computer interaction [38] (p. 2), and as such its broader objective extends beyond technical transparency toward the cognitive and social dimensions of how users understand and engage with intelligent systems. Explainability therefore represents an indispensable component of trustworthy AI and remains central to responsible intelligent system design.

Encounter ontology suggests that explainability alone may be insufficient for understanding the full significance of human-AI interaction. Existing approaches generally assume that the primary object of explanation is the output itself. The central question becomes “why did the system produce this result?”. While necessary, this perspective remains largely focused on individual responses considered in isolation. Meaning frequently emerges not from a single output but from a sequence of interactions whose significance accumulates through interpretation, reflection, revision, and continued engagement. This distinction introduces a difference between explanation and interpretation. An output may be technically explainable while its broader significance remains unresolved. Understanding also depends upon the interpretive processes through which users situate, negotiate, and integrate that response within broader cognitive, social, and symbolic contexts. As Drucker [39] argues in her critique of representational assumptions surrounding information, meaning is not simply extracted from data but actively constituted through interpretive acts. The consequences of the encounter that generates cognitive shifts and new epistemic orientations may often become visible only across time and through sustained interaction. Their significance lies not solely in what the system generated but in what the encounter made possible.

This suggests the need for a complementary notion of relational explainability. Relational explainability does not replace existing approaches to explainable AI; it extends them by recognising that intelligent interaction may produce cumulative interpretive effects. Whereas conventional explainability seeks to account for system behaviour, relational explainability directs attention toward the intelligibility of the encounter and its consequences. It asks not only why a response occurred, but also what forms of meaning, understanding, and epistemic transformation emerged through the interaction. This realisation would bring the necessary background for more mature, and efficient measurement mechanisms, maintaining the essential role of explainability, but approaching it through the lens of encounter ontology. By doing so, the future challenge may involve complementing explanation with a diverse understanding of interpretation, relationality, and the epistemic consequences of sustained human-AI interaction and observation.

If intelligent encounters increasingly participate in the formation of meaning, understanding, and epistemic orientation, then ethical responsibility cannot be limited solely to the correctness or transparency of individual outputs. Questions of responsibility therefore extend beyond whether a system produces accurate responses toward how sustained encounters shape understanding, reflection, and the construction of meaning over time, and if they do so. Under hyperhistorical conditions, the ethical challenge is not only to explain intelligent systems, but also to cultivate conditions for encounters capable of supporting responsible and meaningful forms of interpretation.

To make relational explainability more tractable than aspirational, it helps to separate two evaluation targets that current explainable AI metrics tend to collapse. The first is output-level explanation: whether a user can correctly state why a given response was produced, the target of established functionally and human grounded explainability evaluation [40,41]. The second, which encounter ontology foregrounds, is encounter-level

interpretive outcome: whether and how a user's mental model, interpretive framing, and subsequent reasoning are altered across a sequence of exchanges. The latter cannot be assessed from a single response. It calls for longitudinal, process-sensitive instruments, mental model elicitation before and after sustained interaction [42], experience-sampling or diary methods across sessions, and trust calibration measures that track whether user reliance comes into appropriate correspondence with system reliability over time [43,44]. Computational models of trust dynamics offer a complementary operational direction. VIRTISI models trust as a sequence of cognitive state transitions—neutral, trust, overtrust, distrust, and mistrust—that evolve across a session of human–AI interaction in response to successive AI responses, providing a systematic way for tracking how trust calibration shifts across sustained interaction rather than at a single exchange [45]. Concrete instruments for studying temporal co-evolution are beginning to emerge. Yun et al. [46], conducting longitudinal phenomenological studies across AI companionship and professional software engineering contexts, introduce progressive transparency interviews and temporal co-evolution tracking as methods for capturing how human–AI mutual adaptation converges or drifts across extended interaction. These instruments operationalise what relational explainability requires: not a single point assessment of output intelligibility but a sustained method for tracking how interpretive orientation shifts across the duration of an encounter and beyond it.

Relational explainability, on this reading, is not a softer alternative to explainable AI but a distinct and measurable target: the intelligibility not of the output but of the encounter and its epistemic consequences.

4.3. Toward Relational AI: Key Design Orientations

The conceptual reorientations proposed in this paper suggest an alternative epistemic design orientation. Taken together, these orientations do not prescribe specific technical architectures or governance mechanisms. They identify dimensions of intelligent interaction that become increasingly significant when the encounter is treated as the primary epistemic unit of analysis. Their purpose is not to replace existing approaches to alignment, explainability, or trustworthy AI, but to complement them by drawing attention to the temporal, symbolic, interpretive, and relational conditions through which meaning emerges, stabilises, and becomes historically transmissible under hyperhistorical conditions. Operationalising these orientations will require evaluation instruments that are themselves relationally conceived. Tools capable of tracking meaning formation across time rather than measuring output quality at a single moment, and sensitive to the philosophical dimensions of human–AI encounter that current benchmarks do not yet capture. This is not a limitation of the framework but an alternative direction it opens. The three orientations are therefore better understood as dimensions along which encounters can be evaluated and governed rather than quantities to be maximised.

4.3.1. Design for Temporal Continuity

Meaning, as encounter ontology proposes, emerges across sequences of interaction rather than within discrete exchanges. Design for temporal continuity follows directly from this: it asks how intelligent systems can be built to sustain, rather than reset, the epistemic conditions through which relational meaning forms over time. The three-moment architecture proposed in this framework, pre-encounter potential, encounter activation, and post-encounter condition, makes visible a temporal depth that current evaluation methods do not reach. Each moment carries distinct evaluative implications: pre-encounter potential calls for attention to how system design conditions what can emerge; encounter activation calls for instruments sensitive to what forms within the exchange; post-encounter condition calls

for methods that track what persists beyond it. Continuity, longitudinal interaction, and the persistence of epistemic residue across sessions are not secondary features of human–AI interaction. They are among its most consequential dimensions. Future intelligent systems research may therefore benefit from developing methods capable of accounting for the temporal development of meaning and understanding across sustained interaction, rather than treating each prompt–response sequence as epistemically independent.

Operationally, temporal continuity invites measures of cross session persistence: the degree to which epistemic residue from earlier encounters demonstrably shapes later ones. Candidate approaches include delayed transfer tasks, returning user log analysis, and the diachronic methods already associated with distributed cognition research [9], repurposed to track how altered cognition is carried forward beyond the session in which it formed.

The urgency of designing for temporal continuity is underscored by what Eliav [47] terms the Cognitive Divergence: while AI context windows have expanded by several orders of magnitude since 2017, human effective context span has contracted sharply over the same period. The resulting asymmetry introduces a delegation feedback loop in which humans increasingly offload cognitive tasks as AI capability grows, with cumulative effects on attentional capacity and epistemic autonomy. Design for temporal continuity must therefore attend not only to how meaning persists across sessions but to how continuity is scaffolded without accelerating over delegation or cognitive deskilling.

4.3.2. Design for Encounter Quality

The significance of a human–AI interaction cannot always be reduced to the correctness or utility of its outputs. It also depends on the extent to which the encounter supports reflection, sustains interpretive openness, preserves symbolic coherence, and enables productive forms of cognitive engagement. Yet current evaluation frameworks are structurally synchronic, that is, designed to assess a single output at a single moment, making them constitutively unable to register qualities that only emerge across the duration and recursiveness of an encounter.

Duration, continuity, recursiveness, and symbolic context all contribute to shaping the broader consequences of interaction across time. Future research may therefore benefit from examining how different interactional conditions influence the development of understanding, judgement, and intellectual orientation within sustained human–AI engagement.

Encounter quality can be specified as a small set of constructs drawn from the framework itself, each with candidate directions for intelligent systems research, so that it becomes an object of inquiry rather than a purely evaluative term. Table 2 maps these constructs and indicates, for each, what it implies for system design and what a researcher would need to measure. None of these directions is offered as definitive; they are starting points intended to show that encounter quality is operationalizable in principle, and to invite the empirical work that would refine it.

Crucially, these constructs are properties neither of the output nor of the user or system in isolation. They are properties of the encounter, and most become visible only across time and across multiple sessions, which is why they remain invisible to the single output evaluation paradigms that currently dominate the field.

4.3.3. Design for Symbolic Legibility

If mediation participates in the formation of meaning rather than merely transmitting it, then the interfaces, representations, and communicative structures through which AI systems operate cannot be treated as neutral delivery mechanisms. Symbolic infrastructures contribute actively to how computational processes become intelligible and inhabitable within human experience. Symbolic legibility refers to the capacity of an intelligent system

to support orientation, interpretation, and meaningful engagement within environments characterised by growing informational and computational complexity. This does not imply simplification or reduction. A system that shifts register, vocabulary, or representational style unpredictably across sessions disrupts the symbolic continuity through which users build interpretive orientation, a failure invisible to output-level evaluation but consequential for the encounter's epistemic consequences. It concerns the development of symbolic, representational, and mediating structures through which complexity becomes cognitively inhabitable. Future intelligent systems research may therefore benefit from considering not only what systems compute, but also how their outputs are encountered, interpreted, and integrated into broader processes of meaning formation. Under hyperhistorical conditions, symbolic legibility becomes increasingly significant because the survival and transmission of meaning depend not only on information itself, but on the conditions through which information becomes inhabited and historically consequential.

Table 2. Encounter ontology constructs and possible directions for intelligent systems research.

Construct	System Design Implication	Candidate Evaluation Direction
Temporal continuity	Support memory, return, revision, and cross session coherence without forcing passive dependence.	Delayed transfer tasks; returning user log analysis; comparison of first session and later session reasoning.
Encounter activation	Treat prompting, response, interpretation, and revision as a coupled process rather than isolated transactions.	Multi-turn process tracing; interaction-depth measures; coding of user initiated reframings and topic migrations.
Epistemic residue	Account for what persists after the session: altered concepts, changed judgement, and new interpretive habits.	Pre/post mental model elicitation; diary or experience sampling methods; follow-up reasoning tasks.
Relational explainability	Explain not only why an output occurred, but how the encounter shaped understanding over time.	Encounter-level intelligibility studies; progressive transparency interviews; trust calibration measures across sessions.
Symbolic legibility	Stabilise representational cues, vocabulary, and interface signals so complexity remains inhabitable.	Cognitive load measures; orientation/navigation tasks; assessment of consistency in register and representational style.
Specification–emergence gap	Monitor divergence between intended objectives and emergent epistemic trajectories in open-ended tasks.	Compare specified task goals with downstream conceptual shifts; measure widening or collapse of the content gap by task type.

Symbolic legibility is likewise assessable through comprehension and orientation under complexity: whether users can locate, interpret, and integrate system outputs as informational environments grow more intricate. Established measures of perceived cognitive load [48], alongside orientation and navigation measures for complex interfaces, offer a starting point, provided they are applied across the duration of an encounter rather than to single, isolated interactions.

Taken together, the three design orientations, temporal continuity, encounter quality, and symbolic legibility, define an evaluation agenda rather than a finished method. Each names a property of the encounter, explains why output-centred evaluation cannot capture it, and indicates at least one direction through which it could be measured. This is the sense in which encounter ontology is a conceptual framework for AI-empowered information

systems and not only a philosophy of them: it does not yet tell engineers what to build, but it specifies what they would need to measure in order to know whether relational emergence is occurring and what it is producing. Turning these dimensions into validated instruments is the central task this framework opens for future, empirically grounded research.

5. Conclusions

This paper has argued that contemporary human–AI interaction cannot be understood solely through the paradigms of information exchange, behavioural alignment, or output evaluation. Under hyperhistorical conditions, intelligent systems increasingly participate in the environments through which meaning is formed, interpreted, and transmitted. The encounter between human and machine therefore emerges as a distinct object of inquiry, one whose significance exceeds the behaviours of individual participants and the outputs they produce.

To address this shift, the paper proposed encounter ontology as a third ontological position alongside human-centred and machine-centred accounts. Its primary unit of analysis is the encounter itself, the temporally structured relational event through which meaning emerges, is negotiated, and leaves epistemic residue that persists beyond the immediate session. Meaning is not transmitted between pre-existing entities but produced through the encounter, generating epistemic formations, interpretive orientations, and cognitive transformations that neither participant could have anticipated or carried independently. The framework introduced pre-encounter potential, encounter activation, and post-encounter condition as the temporal architecture through which these processes unfold, and proposed relational intelligence as the emergent co-production of meaning across situated human–AI interactions over time. Knowledge is not only discovered, represented, or transmitted but produced through relational events of mediation.

The implications of this position extend beyond philosophy and toward the future study and design of intelligent systems. If meaning increasingly emerges through sustained encounters with intelligent mediation, then questions of temporality, interpretation, symbolic legibility, and relational explainability become increasingly significant for how intelligent systems are evaluated and governed. The challenge facing intelligent systems research is therefore not only how to build more capable systems, but how to understand the conditions through which those systems participate in the formation of human cognition and in the survival and transmission of meaning across time and generations. A further theoretical extension, reserved for subsequent development, concerns the civilisational scale consequences of encounter ontology: when an immense number of simultaneous human–AI encounters each collapse different fields of epistemic potential into actuality, the aggregate produces a new form of collective cognitive formation whose properties and implications require their own theoretical framework.

Encounter ontology is presented not as a replacement for existing approaches to alignment, explainability, or trustworthy AI, but as a complementary perspective that brings into view dimensions of human–AI interaction that remain difficult to capture within output-centred paradigms. The tensions that governance frameworks have identified between human agency and automation, between transparency and accountability, between specification and emergence, are structural. They reflect the irreducibly open and generative character of human–AI interaction itself. Understanding that character is a necessary condition for governing it responsibly. As intelligent systems become ever more deeply embedded within the infrastructures of everyday life, understanding what emerges through encounter may become as important as understanding the systems themselves.

Author Contributions: Conceptualization, M.S. and M.E.K.; methodology, M.S. and M.E.K.; investigation, M.S. and M.E.K.; writing—original draft preparation, M.E.K.; writing—review and editing,

M.S. and M.E.K.; supervision, M.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: The authors acknowledge the intellectual context provided by the IISA research community and the invitation to develop the work for consideration in Electronics. This article is a theoretical and conceptual contribution. It does not present an empirical validation study. Its aim is to develop an ontological framework and derive design-relevant propositions for future empirical and technical research. During the preparation of this manuscript, generative AI tools (specifically Claude Sonnet 4.6, Anthropic) were used to support language refinement, editing, and dialogic exploration of draft formulations. All conceptual contributions, theoretical arguments, interpretations, and final editorial decisions were conceived, critically evaluated, verified, and approved by the authors.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Floridi, L. *The Fourth Revolution: How the Infosphere is Reshaping Human Reality*; Oxford University Press: Oxford, UK, 2014.
2. Ferrario, A.; Facchini, A.; Termine, A. Experts or authorities? The strange case of the presumed epistemic superiority of artificial intelligence systems. *Minds Mach.* **2024**, *34*, 30. [\[CrossRef\]](#)
3. Levin, I. Vibe-Creation: The epistemology of human-AI emergent cognition. *arXiv* **2026**, arXiv:2603.09486. [\[CrossRef\]](#)
4. Palermos, S.O. Epistemic collaborations: Distributed cognition and virtue reliabilism. *Erkenntnis* **2020**, *87*, 1481–1500. [\[CrossRef\]](#)
5. Ezra, R.; Mishali, M. The relational epistemic stance: Generative AI as a dynamic transitional object. *AI Soc.* **2026**, *41*, 5027–5042. [\[CrossRef\]](#)
6. Shannon, C.W.; Weaver, W. *The Mathematical Theory of Communication*; University of Illinois Press: Urbana, IL, USA, 1949.
7. Clark, A.; Chalmers, D. The extended mind. *Analysis* **1998**, *58*, 7–19. [\[CrossRef\]](#)
8. Clark, A. *Mindware: An Introduction to the Philosophy of Cognitive Science*; Oxford University Press: Oxford, UK, 2002.
9. Sutton, J. Exograms and interdisciplinarity: History, the extended mind, and the civilizing process. In *The Extended Mind*; Menary, R., Ed.; MIT Press: Cambridge, MA, USA, 2010; pp. 189–225.
10. Malafouris, L. *How Things Shape the Mind: A Theory of Material Engagement*; MIT Press: Cambridge, MA, USA, 2013.
11. Clark, A. *Being There: Putting Brain, Body, and World Together Again*; MIT Press: Cambridge, MA, USA, 1997.
12. Stiegler, B. *Technics and Time, 1: The Fault of Epimetheus*; Stanford University Press: Stanford, CA, USA, 1998.
13. Simondon, G. *Individuation in Light of Notions of Form and Information*; Adkins, T., Translator; University of Minnesota Press: Minneapolis, MN, USA, 2020.
14. Latour, B. *Reassembling the Social: An Introduction to Actor-Network Theory*; Oxford University Press: Oxford, UK, 2005.
15. McLuhan, M. *Understanding Media: The Extensions of Man*; McGraw-Hill: New York, NY, USA, 1964.
16. Hayles, N.K. *How We Think: Digital Media and Contemporary Technogenesis*; University of Chicago Press: Chicago, IL, USA, 2012.
17. Deleuze, G. *Difference and Repetition*; Columbia University Press: New York, NY, USA, 1994.
18. Hutchins, E. *Cognition in the Wild*; MIT Press: Cambridge, MA, USA, 1995.
19. Salomon, G. *Distributed Cognitions: Psychological and Educational Considerations*; Cambridge University Press: Cambridge, UK, 1997.
20. DeLanda, M. *A New Philosophy of Society: Assemblage Theory and Social Complexity*; Continuum: London, UK, 2006.
21. Barad, K. *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning*; Duke University Press: Durham, NC, USA, 2007.
22. Suchman, L. *Plans and Situated Actions: The Problem of Human-Machine Communication*; Cambridge University Press: Cambridge, UK, 1987.
23. Matthias, A. The responsibility gap: Ascribing responsibility for the actions of learning automata. *Ethics Inf. Technol.* **2004**, *6*, 175–183. [\[CrossRef\]](#)
24. Vallor, S.; Vierkant, T. Find the gap: AI, responsible agency and vulnerability. *Minds Mach.* **2024**, *34*, 20. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Sfakianakis, M.; Kampi, M.E. Symbolic infrastructure: A framework for transhuman co-creation and knowledge mediation. In *Proceedings of the 16th International Conference on Information, Intelligence, Systems and Applications (IISA 2025)*, Mytilene, Greece, 14–16 July 2025; IEEE: New York, NY, USA, 2025; pp. 1–8. [\[CrossRef\]](#)

26. Sfakianakis, M.; Kampi, M.E. Aesthetic mediation as epistemic infrastructure for intelligent systems. In *Proceedings of the 17th International Conference on Information, Intelligence, Systems and Applications (IISA 2026)*, Rhodes, Greece, 2026; IEEE: Piscataway, NJ, USA, 2026; *in press*.
27. Kampi, M.E. The Alteregorythm™: A speculative framework for symbolic co-development in posthuman creative collaboration. In *Advances in Global Applied Artificial Intelligence*; Tsihrintzis, G.A., Virvou, M., Bourbakis, N.G., Jain, L.C., Eds.; Springer: Berlin/Heidelberg, Germany, 2026; *in press*.
28. Whitehead, A.N. *Process and Reality*; Free Press: New York, NY, USA, 1978.
29. Turner, V. *The Ritual Process: Structure and Anti-Structure*; Aldine Publishing: Chicago, IL, USA, 1969.
30. Massumi, B. *Parables for the Virtual: Movement, Affect, Sensation*; Duke University Press: Durham, NC, USA, 2002.
31. European Commission. *Ethics Guidelines for Trustworthy AI*; European Commission: Brussels, Belgium, 2019.
32. Ouyang, L.; Wu, J.; Jiang, X.; Almeida, D.; Wainwright, C.; Mishkin, P.; Zhang, C.; Agarwal, S.; Slama, K.; Ray, A.; et al. Training language models to follow instructions with human feedback. *Adv. Neural Inf. Process. Syst.* **2022**, *35*, 27730–27744. [\[CrossRef\]](#)
33. Bai, Y.; Jones, A.; Ndousse, K.; Askell, A.; Chen, A.; DasSarma, N.; Drain, D.; Fort, S.; Ganguli, D.; Henighan, T.; et al. Constitutional AI: Harmlessness from AI Feedback. *arXiv* **2022**, arXiv:2212.06950.
34. European Union. *Regulation (EU) 2024/1689 of the European Parliament and of the Council (Artificial Intelligence Act)*; European Union: Brussels, Belgium, 2024.
35. Casper, S.; Davies, X.; Shi, C.; Gilbert, T.K.; Scheurer, J.; Rando, J.; Freedman, R.; Korbak, T.; Lindner, D.; Freire, P.; et al. Open Problems and Fundamental Limitations of Reinforcement Learning from Human Feedback. *arXiv* **2023**, arXiv:2307.15217. Available online: <https://openreview.net/forum?id=bx24KpJ4Eb> (accessed on 16 July 2026).
36. Wei, J.; Tay, Y.; Bommasani, R.; Raffel, C.; Zoph, B.; Borgeaud, S.; Yogatama, D.; Bosma, M.; Zhou, D.; Metzler, D.; et al. Emergent Abilities of Large Language Models. *arXiv* **2022**, arXiv:2206.07682.
37. Shingekar, S.R.; Gunda, V.; Kim, J.; Rodriguez, V.J.; Sundaram, H.; Saha, K. AI Psychosis: Does Conversational AI Amplify Delusion-Related Language? *arXiv* **2026**, arXiv:2603.19574. [\[CrossRef\]](#)
38. Miller, T. Explanation in artificial intelligence: Insights from the social sciences. *Artif. Intell.* **2019**, *267*, 1–38. [\[CrossRef\]](#)
39. Drucker, J. Humanities approaches to graphical display. *Digit. Humanit. Q.* **2011**, *5*, 1–21. [\[CrossRef\]](#)
40. Doshi-Velez, F.; Kim, B. Towards a Rigorous Science of Interpretable Machine Learning. *arXiv* **2017**, arXiv:1702.08608.
41. Hoffman, R.R.; Mueller, S.T.; Klein, G.; Litman, J. Metrics for Explainable AI: Challenges and Prospects. *arXiv* **2018**, arXiv:1812.04608.
42. Sanchez, T.; Vereschak, O.; Deroy, O. Mental models in human-AI interaction: Systematic review of empirical methodologies and guidelines. In *Proceedings of the 31st International Conference on Intelligent User Interfaces, Paphos Cyprus, 23–26 March 2026*; Association for Computing Machinery: New York, NY, USA, 2026; pp. 663–682. [\[CrossRef\]](#)
43. Lee, J.D.; See, K.A. Trust in automation: Designing for appropriate reliance. *Hum. Factors* **2004**, *46*, 50–80. [\[CrossRef\]](#) [\[PubMed\]](#)
44. Jian, J.-Y.; Bisantz, A.M.; Drury, C.G. Foundations for an empirically determined scale of trust in automated systems. *Int. J. Cogn. Ergon.* **2000**, *4*, 53–71. [\[CrossRef\]](#)
45. Virvou, M.; Tsihrintzis, G.A.; Tsihrintzi, E.-A. VIRTSL: A novel trust dynamics model enhancing artificial intelligence collaboration with human users—Insights from a ChatGPT evaluation study. *Inf. Sci.* **2024**, *675*, 120759. [\[CrossRef\]](#)
46. Yun, B.; Taranova, E.; Feng, D.; Su, R.; Wang, A.Y. AI Phenomenology for Understanding Human-AI Experiences Across Eras. *arXiv* **2026**, arXiv:2603.09020.
47. Eliav, N. The Cognitive Divergence: AI Context Windows, Human Attention Decline, and the Delegation Feedback Loop. *arXiv* **2026**, arXiv:2603.26707. [\[CrossRef\]](#)
48. Hart, S.G.; Staveland, L.E. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Human Mental Workload*; Hancock, P.A., Meshkati, N., Eds.; North-Holland: Amsterdam, The Netherlands, 1988; pp. 139–183.

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.