Agency without Imputation

Hermeneutics, Pragmatics, and the Epistemic Validation of Artificial Discourse

Carlos Isaac Zainea[*]

Doctorate in Philosophy · Universidad Santo Tomás

First Spanish version, March 2026 · English version, September 2026

Abstract

Philosophy has approached large language models (LLMs) through an ontological question — does the machine understand? — whose negative answer, correct as it is, tends to yield a paralysing conclusion: if the LLM does not understand, its output has no epistemic value. This article argues that the inference is invalid and proposes an epistemological turn. First, using Ricoeur's threefold mimesis, it locates the LLM's deficit: it is an agent that configures (Mimesis II) without being able to refigure or to have what it configures imputed to it (Mimesis III). Second, it recategorises the error: not "hallucinations" — an incoherent mentalist metaphor — but **infelicities of epistemic authority** in Austin's sense, articulated with Habermas and Floridi. Third, it shows that Grounded Theory and Triangulation supply the procedural scaffolding that compensates for that lack, and demonstrates this with a prototype implemented on Colombian narco visual culture. The result is the concept of the **hermeneutic prosthesis**: an architecture that redistributes epistemic functions between an operative agent (the machine) and an imputable agent (the researcher).

Keywords: artificial intelligence; epistemology; hermeneutics; Ricoeur; Austin; agency; Grounded Theory.

---

[*]Doctoral student, Doctorate in Philosophy, Faculty of Philosophy and Letters, Universidad Santo Tomás, Bogotá, Colombia. Contact: `cizaineam@gmail.com`. I thank Jesús Antonio Pardo León, a colleague in the Doctorate in Philosophy at Universidad Santo Tomás, for his input on the problem — the conceptualisation of *the narco as an aesthetic category*, the subject of his doctoral thesis in progress (Pardo León, in progress) — and for his suggestions of tools with which to approach it. The concept of hermeneutic prosthesis, the choice and application of Grounded Theory, and the implementation of the prototype are the author's own. On the use of artificial intelligence in the preparation of this paper, see the Declaration at the end of the text.

1. The epistemological problem

Large language models (LLMs) entered the production of knowledge before philosophy answered the question they pose. In academic, medical, legal and journalistic contexts, systems such as GPT, Claude or Gemini synthesise literature, extract patterns, propose categories and produce texts that circulate as research findings. The problem is not whether LLMs *will* be used — they already are — but under what conditions such use is epistemologically legitimate.

Philosophy, however, has responded with a question of an **ontological** kind: does the machine *understand*? Does it have intentionality? Is it a subject? The question is legitimate, but it produces a dead end: the answer is no, and accepting it without further ado appears to close off any possibility of thinking of the LLM as a participant in epistemic processes. This work maintains that the dead end is an artefact of the question, not a necessity of the problem. That the LLM does not understand does not mean that it does not *do*: it produces, it operates, it selects which token follows which, it configures effects that reach the world. There is here a type of operation that is neither that of the knowing subject nor that of the passive instrument, and naming it well is the task the ontological question leaves unresolved.

1.1 The legacy of the ontological critique

The tradition that has answered the question of artificial understanding is robust and, in essentials, correct. Searle's *Chinese Room* (1980) establishes that the syntactic manipulation of symbols does not produce semantics. Dreyfus (1972, 1992) adds the phenomenological dimension: human understanding is anchored in the body, in the situation, and in a background of practical skills that no formal system replicates. Gadamer (1960) extends this to the hermeneutic plane: to understand is to participate in a tradition, in a history of effects (*Wirkungsgeschichte*) prior to any individual act. Heidegger (1927) formulates the ontological ground: understanding is not a cognitive operation but a structure of Dasein as being-in-the-world, anchored in temporality, finitude and care.

The LLM has none of this: no body, no situation, no effective history, no world to which it belongs. Its production operates over co-occurrence probabilities in a corpus, not over participation in the practices that corpus sediments. When it utters something about grief, it has lost no one; when it synthesises a philosophical discussion, it has none of the questions that make that discussion matter. The ontological critique is irrefutable, and this work accepts it in full. What it disputes is the inference usually drawn from it.

1.2 The ontological blockage

The ontological critique issues in a syllogism with a paralysing conclusion:

P1. Genuine knowledge requires an understanding subject.

P2. The LLM is not an understanding subject.

C. The LLM cannot participate in the production of knowledge.

The premises are sound. The conclusion does not follow.

The premises are sound; the conclusion is not. It rests on a suppressed premise — that all epistemic participation requires understanding — which is false. Instruments participate in the production of knowledge without understanding anything: the telescope, the mass spectrometer, the statistical model. None of them understands, and all of them contribute to knowledge whose validity is established by *procedures external to the instrument*. What is peculiar about the LLM is not that it fails to understand — so does the thermometer — but that it produces discourse with the form of a justified assertion, which is precisely what instruments do not do. That is what makes the instrumental analogy insufficient, and what requires an epistemology, not merely a metrology.

1.3 The epistemological turn

Ontological question		**Epistemological question**
Does the LLM understand?	TURN →	Is the knowledge produced valid?
Does it have intentionality?		How is its reliability controlled?
Is it a subject?		Who validates, and how?
Answer: No. (And it is a dead end.)		*Answer: It depends on the procedural scaffolding.*

The reorientation involves recognising that epistemology does not require the instruments of knowledge production to be knowing subjects, but that the *processes* governing their use satisfy conditions of validity. The contemporary debate already articulates this distinction from independent angles. Cantwell Smith (2019) distinguishes *reckoning* — the competent manipulation of formal structures — from *judgment* — contextualised, responsible evaluation anchored in a world of consequences. Floridi (2019; earlier Floridi & Sanders, 2004) points to two things worth reading together: that LLMs produce data with the appearance of semantic information, alethically neutral until an external process situates them in a register of truth; and that such systems can be characterised as **agents** in the proper sense — interactive, autonomous, adaptable — without thereby being credited with intentionality, consciousness or responsibility. The resulting diagnosis is two-sided: the LLM operates with growing competence and *acts* with real consequences, yet lacks the conditions under which that action could be epistemically valid and attributable. Hence the question that organises this article: *how is the knowledge produced by — or with — a system that acts and generates coherent discourse without participating in the conditions of understanding or of responsibility to be validated?*

1.4 Thesis and structure

The thesis is stated in three moves. **First**: LLMs are configuration machines in the sense of Ricoeur's Mimesis II — they produce plots, organise data, generate coherent discourse — and in that exercise they are agents in the proper sense, without participating in the conditions of Mimesis III, where knowledge is validated, imputed to a subject, and returned to a world of consequences. There is operative agency; imputation is missing. **Second**: their errors are not "hallucinations" — a mentalist metaphor that presupposes a mind that distorts — but **infelicities of epistemic authority** in Austin's sense: acts that take the form of grounded assertions without being able to satisfy the conditions of that grounding, because no agent capable of answering for them sustains them. **Third**: Grounded Theory and Triangulation supply the protocols for building the scaffolding that compensates for that lack; the prototype *Category Discovery in Colombian Narco Images* — a hermeneutic analysis of narco visual culture implemented with Gemini Vision — demonstrates that the design is implementable and makes it possible to identify what the scaffolding achieves and what the researcher must still contribute.

Section 2 analyses the LLM's functional asymmetry and locates the deficit by means of the threefold mimesis. Section 3 recategorises the error as an infelicity of epistemic authority, articulating Austin, Habermas and Floridi. Section 4 develops hermeneutics as protocol: the genealogy of artificial reasoning techniques (CoT → ToT → ReAct → agentic AI), the translation of Grounded Theory into scaffolding, and a critical analysis of the prototype. Section 5 articulates the concept of the **hermeneutic prosthesis**, grounded in Clark, Merleau-Ponty, Feenberg and Vallor. Section 6 synthesises the argument and opens lines of research.

2. The production of meaning without an epistemic subject

> *"To explain more is to understand better."*
>
> — Ricoeur, *From Text to Action*

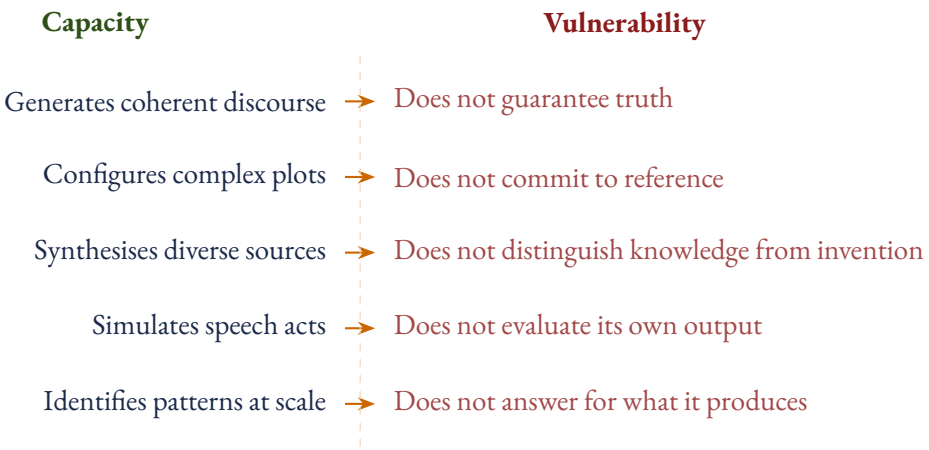
The question left open by §1 first demands a precise description of what the system does — not in an ontological but in a functional sense: what operations does it perform on language, with what competence, and under what structural constraints? The answer is counter-intuitive: the capacities that make the LLM epistemologically useful and the vulnerabilities that make it dangerous are not independent properties but two faces of the same structural condition.

2.1 The functional asymmetry

The LLM is, above all, a machine for producing coherent discourse. Its training — next-token prediction, optimised by human preference feedback — produces a system extraordinarily competent at satisfying formal expectations: grammaticality, thematic consistency, rhetorical adequacy. What that process does not produce, because it does not require it, is correspondence between the discourse and an external state of affairs. The model is trained to be coherent, not to be true; coherence and truth are

distinct properties, and the optimisation produces the first independently of the second.

This separation is the deepest case of a pattern that repeats across every dimension of the system. The LLM configures complex plots fluently — and that very capacity prevents it from committing to reference, because narrative coherence is independent of factuality: a novel is coherent while referring to nothing real, and the system does not distinguish the two registers. It synthesises sources with apparent precision — and for that reason does not distinguish grounded knowledge from confabulation. It produces speech acts with pragmatic force — but without the capacity to evaluate whether each act is justified in its context. Configurative power and validative incapacity are the same mechanism seen from different angles.



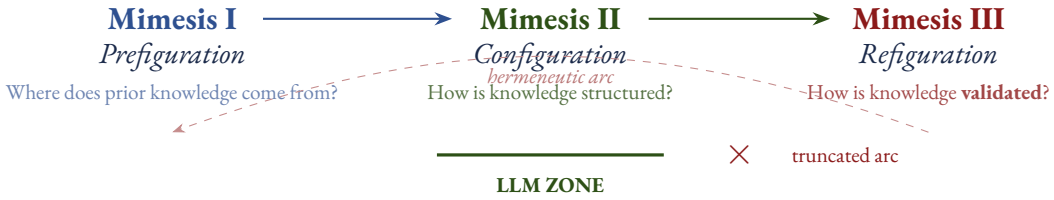
The theoretical consequence matters: the asymmetry is not a contingent deficiency that better training or greater scale will resolve. It is structural. Scaling the model produces greater competence in coherence, not greater epistemic correctness: a larger model generates more fluent and more convincing confabulations, not fewer confabulations. The adequate response is therefore not technical but epistemological: the point is not to improve the LLM, but to design the external conditions of validity it cannot generate from within. The last row of the figure — the system acts but does not answer for what it does — is not one more item: it is the feature that qualifies all the others. Coherence without truth, configuration without reference, and synthesis without discernment are symptoms of a deeper operation: that of an agent that produces without being able to answer for what it has produced.

2.2 The threefold mimesis: a diagnostic use

Time and Narrative (Ricoeur, 1983) offers the categories with which to articulate the asymmetry more precisely. Ricoeur distinguishes three moments of the hermeneutic arc. **Mimesis I** — prefiguration — is the set of practical, symbolic and temporal pre-understandings a subject brings to narration: the world understood before being narrated. **Mimesis II** — configuration — is the act by which dispersed events are organised into a coherent whole, transforming succession into plot. **Mimesis III**

— refiguration — is the moment at which the configured narrative intersects the world of the reader and produces practical effects; it is where narrative knowledge becomes effective and can be evaluated, corroborated or refuted.

The arc is not linear but circular: Mimesis III returns to Mimesis I, enriching the pre-understanding from which new narratives will arise. Crucially, it requires a subject: someone for whom prefiguration is lived pre-understanding, who configures with narrative intention, and for whom refiguration produces self-understanding and practical reorientation.



The use made here of the threefold mimesis is **diagnostic**, not descriptive. It is not claimed that the LLM performs Mimesis II in the full sense — that would require narrative intention and a world of experience the system does not possess — but that the triadic structure makes it possible to locate what is missing: the return arc towards Mimesis III. The LLM inhabits the zone of Mimesis II with extraordinary competence: it organises data into coherent configurations at a scale that exceeds any individual researcher. But the arc is truncated. There is no Ricoeurian Mimesis I: it has no pre-understood world from which to narrate, but a statistical corpus from which to predict. There is no Mimesis III: there is no world to return to, no reader-subject for whom the configuration produces understanding, no practice in which the knowledge is put to the test. The machine configures; it does not refigure.

It is worth pausing on what refiguring is, because the problem is decided there. Mimesis III is not a product the narrative leaves behind, but an event: the moment at which what has been configured returns to the world of action and time, and someone receives it as their own. A plot refigures only when a reader appropriates it, reorients their understanding from it, and carries it into a practice where it can be tested. Refiguration is therefore the crossing between the text and a world of consequences: without that crossing, what has been configured remains suspended — complete in its internal coherence, but never reaching the register in which something is validated or refuted.

That crossing demands a *who*, and this is what permits reading refiguration as a moment of imputation. The Mimesis III of *Time and Narrative* (1983) can be illuminated by the imputability of *Oneself as Another* (1990): the subject capable of recognising themselves as the author of their acts and of answering for them. To appropriate what has been configured is to be able to say "what was narrated is mine" and, when what is narrated has the form of knowledge, "I answer for what I assert". To refigure and to take as one's own are, at bottom, the same gesture: there is no return to the world without someone on whom what is said devolves. Articulating both texts is a reading proposed by this work, not

explicit Ricoeurian doctrine, but it is justified because refiguration makes no sense without a subject to sustain it.

The LLM configures fluently precisely because it does not need that gesture: it predicts the next token without anything it produces devolving upon it — and what dispenses it from appropriation is also what prevents it. It configures without the configuration being anyone's. The deficit, then, is not the absence of understanding, which §1 already established, but something more precise: the absence of a subject to whom what is configured can be imputed. There is an agent that produces, and a world affected by what it produces, but no one who can say "it is mine" and answer for it. This asymmetry — configuring and acting without being able to be held the author of what is configured — is what we call **agency without imputation**. The question that opens §3 is how to name the failure of such an agent when it produces something with the form of an assertion, without falling back on the mentalist metaphor of "hallucination".

3. Infelicities of epistemic authority

> *"To say something is to do something."*
>
> — J. L. Austin, *How to Do Things with Words*

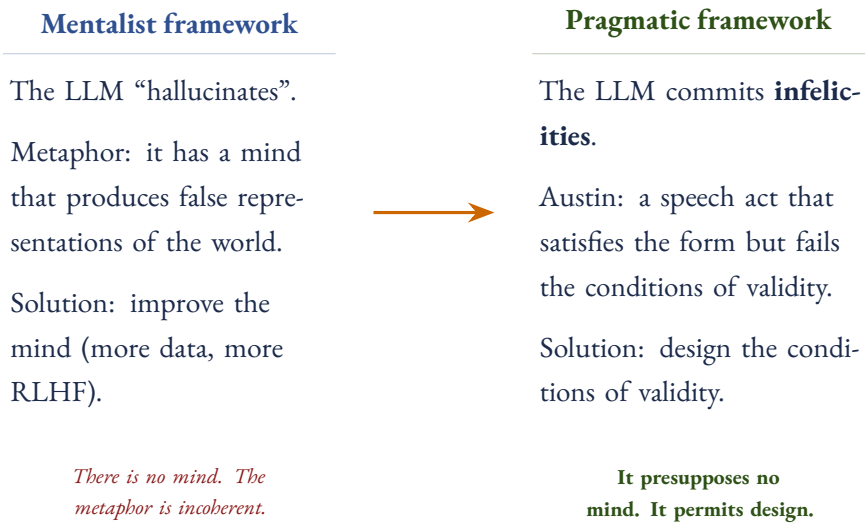
Naming precisely what fails when an LLM produces an incorrect result is not neutral taxonomy: the name determines where the solution is sought. If the LLM "hallucinates", the solution is ultimately psychiatric — correct the system's mind, give it better data. If it commits *infelicities of epistemic authority*, the solution is procedural — design the external conditions of validity the system does not generate internally. The choice is architectural, not stylistic: it determines the type of solution that is conceivable. The mentalist metaphor presupposes a mind that distorts; the Austinian category presupposes an agent that acts without satisfying the conditions of the act. The change of framework does not eliminate the agent: it redescribes it.

The term "hallucination" migrated from psychopathology into discourse about AI when engineers sought a metaphor for non-technical audiences. The one they chose presupposes exactly what §1 showed the LLM does not have: a mind that constructs representations of the world and, under certain conditions, distorts them. But the LLM does not construct representations in any relevant sense: it produces token sequences that satisfy statistical expectations. When it generates a factually false token, it has not failed to represent reality — it has followed the same process as when it generates a true one. The metaphor is not merely imprecise: it is incoherent with respect to the object it purports to describe.

3.1 The two frameworks for error

The alternative is not to seek a more precise metaphor, but to abandon the mentalist framework and replace it with a pragmatic one. The mentalist framework treats the output as the product of a cognitive system that represents the world and, in failing, distorts it; its solution is to improve the system. The

pragmatic framework treats the output as a speech act that either satisfies or fails to satisfy the conditions that make it valid; its solution is to design those conditions externally. The displacement is radical: from psychology to pragmatics, from mind to process, from correcting the system to validating the output.



The pragmatic framework has the advantage of being honest about what the LLM is: it does not posit an interiority the system lacks. It works with the observable — the structure of speech acts and the conditions of their validity — and that ontological austerity is precisely what makes it more powerful for design: if the problem is the absence of conditions of validity, design can build them; if it were a mind that fails, one could only hope that it fails less often.

From conventional to epistemic infelicity

In Austin, *infelicity* covers **performative** speech acts that fail through not satisfying their conventional conditions — the judge without jurisdiction, the promise made under duress — not constative utterances that are simply false, which are ordinary errors. The category marks a specific class of failure in the performative dimension, where what fails is not the content but the conditions that legitimate the act.

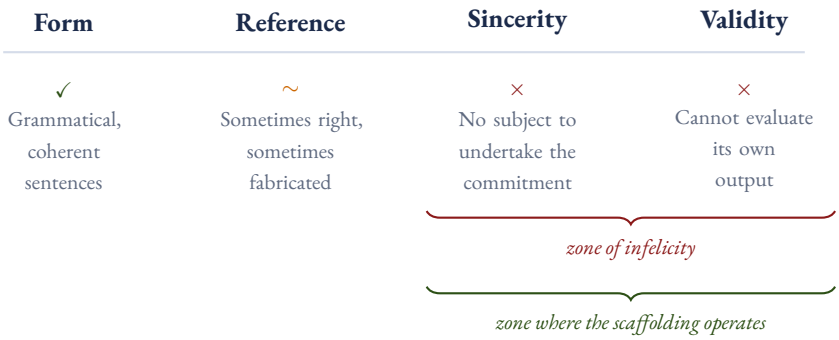
Applying the concept to the LLM requires a justified extension. The LLM's output is never purely constative: when it answers a question, synthesises a bibliography or proposes a category, it does not issue a modest "it may be that p", but an act with **epistemic force** — it presents p as the result of reasoning, with the implicit authority of one who has processed evidence and arrived at a grounded conclusion. That presentation of authority is the irreducible performative dimension of the output, inscribed in the way the system produces discourse, since it was trained on human text in which that force is constitutive of the genre.

The LLM's failure does not consist in saying something false — that would be an ordinary constative error, which humans commit too — but in producing that content *as if* it satisfied the conditions

that ground the authority of an assertion, without being structurally able to satisfy them. The name proposed for this is **infelicity of epistemic authority**: the act takes the form of a grounded assertion when the grounding is constitutively inaccessible from within the system. The authority of an assertion is not a property of the utterance but of the agent who undertakes to answer for it; the infelicity occurs when a real agent produces the act without being that kind of agent. The implication is immediate: if the grounding is inaccessible from within, it must be built from without.

3.2 Anatomy of the infelicity

An epistemologically valid speech act must satisfy at least four conditions. The LLM consistently satisfies only the first:



Form is the condition the LLM best satisfies — grammatical, pragmatically appropriate, rhetorically calibrated sentences — and that surface perfection is the source of the danger: it conceals the underlying deficit. **Reference** is satisfied partially and unpredictably, with no formal mark distinguishing the correct from the fabricated from within the output. **Sincerity** — that the speaker believes what they assert — is structurally impossible: there is no subject who believes, and no interiority from which to undertake a commitment. **Validity** — that the speaker can give reasons — is equally inaccessible: the system has no access to the reasons that produce its output, nor can it distinguish its justified from its unjustified claims. The last two are not deficiencies a better model will resolve: behind the output there are no reasons but probability distributions over vocabularies, and distributions are not reasons. The system produces acts that claim the regime of justification without being able to inhabit it.

3.3 Habermas and Floridi: validity as justifiability

Habermas (1981) allows the diagnosis to be sharpened. Every serious speech act raises three validity claims that the speaker is committed to defending if challenged: **truth** (correspondence with the objective world), **normative rightness** (adequacy to the norms of the context) and **truthfulness** (genuine expression of what is thought). Validity is not a property of the content but of the speaker's capacity to sustain it argumentatively: an act is valid when the speaker can, in principle, give the reasons that back it. The LLM can produce true and normatively appropriate claims, but it satisfies no claim in the sense of *justifiability*: when asked to justify a claim, it produces a new chain of tokens with the morphology of a justification, not a justification. The claim to truthfulness is impossible: there is no

interiority to express.

The consequence is visible in the discrepancy between performance and validity. Engineering measures performance — accuracy, error rate — but performance is not validity. An LLM may score 95% on a benchmark and be epistemologically ungovernable, because the correct 95% is produced by the same statistical reasons as the incorrect 5%, and there is no internal difference the system can point to. The distinction is not one of degree but of category: performance measures hits; validity measures the capacity to give reasons. The "hallucination" framework, by focusing on performance, orients the search for solutions in the wrong direction.

Floridi (2019; Floridi & Sanders, 2004) converges from the philosophy of information and adds two pieces. His distinction between **data** — signals that require no truth — and **semantic information** — content that is true or false, meaningful in a context — names the status of the output exactly: LLMs produce data with the morphology of semantic information, alethically neutral until an external process evaluates them. And his characterisation of the system as an agent (interactivity, autonomy, adaptability, without mental states) resists both anthropomorphisation and instrumentalist denial. LLMs satisfy the three conditions: they are agents. What they do not satisfy is what Ricoeur calls imputability.

Austin, Habermas and Floridi thus converge on a structural diagnosis that none of them formulates separately: the infelicity occurs in executing an act without the conventional conditions that legitimate it (Austin); the LLM cannot give reasons because no communicative subject assumes responsibility (Habermas); there is operative but not imputable agency, and the source of the infelicity lies in that asymmetry (Floridi). The LLM is an agent without imputation that produces performative acts with no agent able to ground them. Hence the design question that guides §4: if imputation cannot come from within, what architecture can locate it where it belongs — in a human agent who *can* answer for what is produced?

4. Hermeneutics as epistemic protocol

The diagnosis of §3 determines the form of the solution. If the LLM produces infelicities because it internally lacks the conditions of validity, the answer cannot be merely technical: it must be methodological — to build, from outside, a protocol that specifies what to do with the output, in what order, by what criterion to determine whether each step was correct, and under what condition the process is sufficient. Answering which protocol satisfies those conditions requires first examining why current protocols of artificial reasoning do not, and then identifying in qualitative research — Grounded Theory and Triangulation — the structures that do.

4.1 The genealogy of the criterion: from CoT to agentic AI

Chain-of-Thought, Tree-of-Thought, ReAct and agentic AI represent genuine leaps in Mimesis II capacity. But their genealogy reveals a constant: each generation adds configurative power without

incorporating an epistemic criterion operating at the level of Mimesis III.

Technique	What it incorporates	What it does not resolve
CoT Wei et al. 2022	Explicit reasoning sequence	*Criterion for the sequence?*
Tree-of-Thought Yao et al. 2023	Exploration of parallel paths	*Criterion for selecting the path?*
ReAct Yao et al. 2022	Environmental feedback: action + observation	*Epistemic justification of the action?*
Agentic AI 2024–	Operational autonomy at scale	*Who validates the chain of infelicities?*

Invariant deficit: no technique specifies a criterion of epistemic validity that is intrinsic to the step, invariant across domains, and normative rather than statistical.

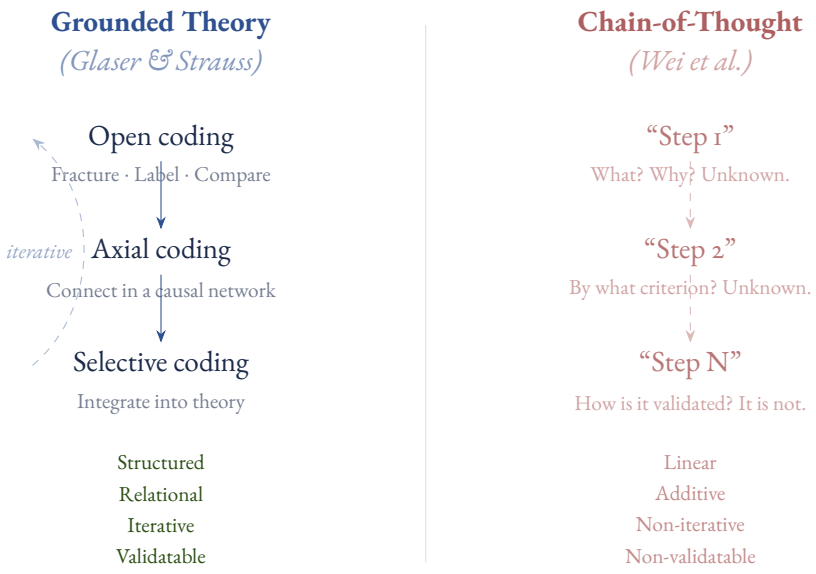
CoT (Wei et al., 2022) improves performance by instructing the LLM to "think step by step", but it lacks a theory of the sequence: it does not specify which steps are necessary, in what order, or when it is legitimate to backtrack. It is a chain of reasoning without logos, without a corrective mechanism operating on anything other than the output of the previous step. ToT (Yao et al., 2023) overcomes linearity by exploring parallel paths, but its selection criterion remains heuristic — coherence or fluency, not justifiability, which the generation process cannot evaluate from within. ReAct (Yao et al., 2022) alternates reasoning and action over external tools, introducing environmental feedback; but that feedback is instrumental: it evaluates whether something worked, not whether it was justified to do it — the distinction between technical efficacy and communicative validity.

Agentic AI (2024–) coordinates multiple agents, maintains memory and executes plans with minimal supervision. Industry calls them *agents*: exact in the Floridian sense, imprecise in any more demanding one. Here the deficit of criterion ceases to be a problem of output and becomes a problem of propagation: each link inherits the infelicity of the previous ones with no mechanism of detection. One agent generates hypotheses from unjustified inferences, another uses them as premises, a third executes them in the world — and the chain of infelicities becomes untraceable. The chain has agents at every link; what it does not have is anyone who can be held the author of the chain. Calling the problem "cascading errors" or "agent drift" does not locate it: responsibility falls on no one, and designing the corrective requires naming it as such from the outset. The deficit of criterion is, in sum, invariant with respect to technical sophistication. Applied hermeneutics has had that specification since long before LLMs existed.

4.2 Grounded Theory as protocol

Grounded Theory, developed by Glaser and Strauss (1967) and systematised by Strauss and Corbin (1990), generates theory from data through an iterative, structured process. What distinguishes it is not only its commitment to data, but the precision with which it specifies the operations that produce

knowledge: each phase has a name, a function, a criterion of correct execution and a termination condition. The contrast with CoT is not one of scale but of nature: CoT leaves it to the system to decide what counts as a step and when the chain is sufficient; GT fractures the material into units, labels them, compares them constantly, builds causal relations and declares saturation when no new codes emerge. Each instruction is an operationalisable criterion, externally verifiable. GT specifies not only what to do, but how to know whether it was done well.



GT and CoT have distinct underlying epistemologies. GT presupposes that knowledge emerges from systematic comparison, saturation and theoretical integration; CoT, that it is reached by chaining plausible steps. The first produces auditable knowledge; the second, convincing text. The difference shows in a simple question: can one know, without evaluating the output, whether the process was executed correctly? For GT, yes; for CoT, no.

4.3 Procedural translation: the scaffolding as design

Transposing GT into a pipeline of prompts is a design programme, not an analogy. The five phases of the protocol, with the criterion each introduces, operate as follows:

Phase 1. Open coding — *Criterion: declaration of limits.*
"Identify concepts without imposing prior categories. For each concept indicate: (a) what you observe, (b) what you cannot determine from this segment, (c) what additional data you would need." The inclusion of (b) and (c) forces the explicit declaration of limits that CoT, ToT and ReAct never produce.

Phase 2. Constant comparison — *Criterion: traceability by code.*
"Compare this segment with the concepts already identified. Does it confirm, contradict, qualify or extend any of them? Indicate the code affected and the nature of the relation." The prompt implements Glaser's constant comparison without requiring theoretical sensitivity

in the LLM: the criterion lies in the design, not in the machine.

Phase 3. Axial coding — *Criterion: auditable relational structure*.
"For each code, identify conditions of appearance, mechanisms and consequences. Build a network of relations indicating their type (causal, temporal, conditional, contradictory)." Strauss and Corbin's condition/mechanism/consequence distinction is translated into a format verifiable by the human in the loop.

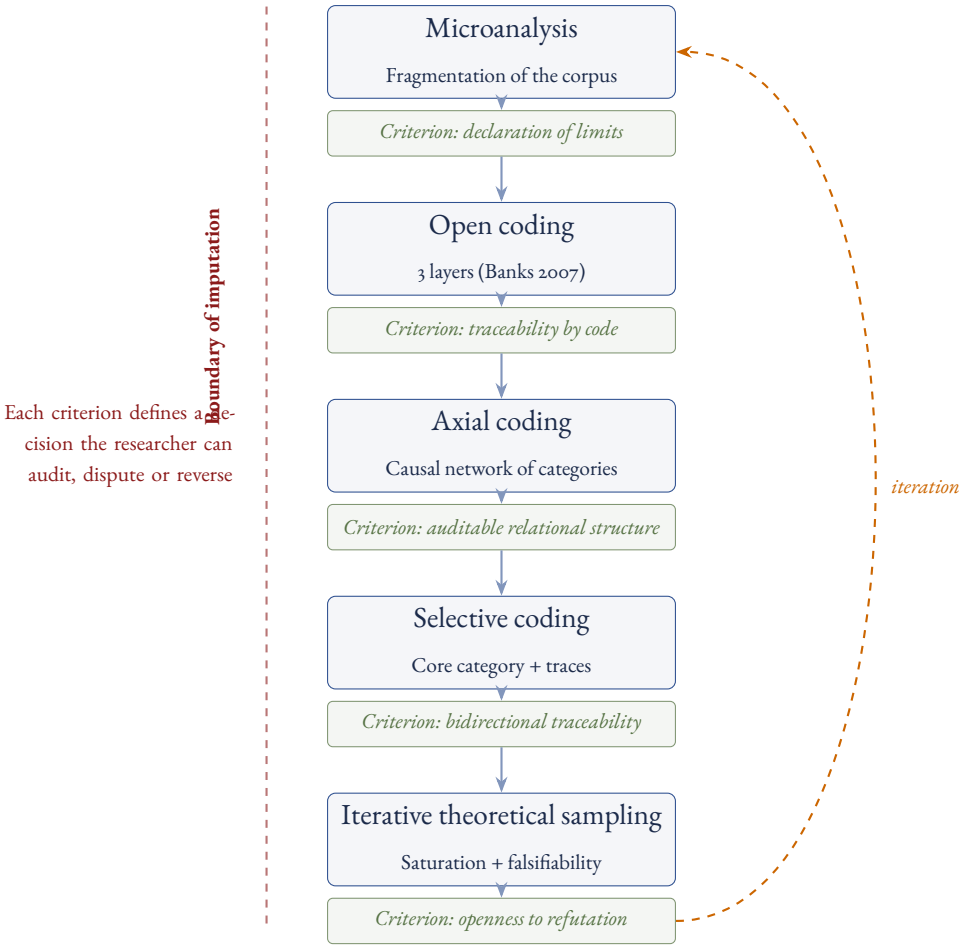
Phase 4. Selective coding — *Criterion: bidirectional traceability*.
"Propose a core category that integrates the findings. For each element you include, indicate the code it comes from; for each you exclude, why it cannot be integrated." Traceability turns the output into auditable text: the criterion agentic AI does not have.

Phase 5. Saturation and falsifiability — *Criterion: openness to refutation*.
"What kind of segment would modify the core category? What datum could falsify it?" The question of falsifiability — absent in CoT, ToT and ReAct — turns the process into research. A system that cannot specify the conditions of its own error does not produce knowledge: it produces text with the appearance of knowledge.

The difference from any artificial reasoning technique lies not in the number of steps but in the type of criterion each incorporates. Those criteria make it possible to determine whether each phase was executed correctly without evaluating the output by its internal coherence. That external auditability is what makes the protocol epistemologically robust: CoT has steps; GT has phases with conditions of verification.

4.4 The prototype: analysing narco visual culture

The five phases are not hypothetical. The system *Category Discovery in Colombian Narco Images* carried out this protocol on a corpus spanning nearly a century of visual circulation around the narco phenomenon in Colombia: press material, institutional photographs of operations, portraits of emblematic figures, military eradication scenes and contemporary digital materials. The LLM is Google Gemini in vision mode, and the system is an iterative pipeline of five modules — one per phase — coordinated by a theoretical sampling module that decides which images enter the next cycle and when to declare saturation.



Algorithmic architecture of the prototype. Each upper box is an operation the system performs on the corpus; each lower box is the external criterion that makes it auditable. The iteration closes the theoretical sampling cycle; the vertical line marks the boundary where imputation returns to the researcher.

Each box is an operation the system executes on the images, and each criterion is the external condition that makes it verifiable without reconstructing the process from scratch. The architecture does not conceal that it is a machine doing the work: it exposes it, decomposes it into discrete steps and leaves visible the points at which the result can be disputed. Auditability does not return to the system the capacity to answer for what it does — that boundary is not crossed — but it locates imputation where it belongs: in the researcher who verifies, accepts or rejects each criterion.

The system executed theoretical sampling cycles until declaring saturation. Axial coding produced 21 saturated categories, integrated by the selective phase into a core category: **Aseptic commodification and transversal spectacularisation of narco trauma**, defined as the process by which visceral violence, necropolitics and illicit capital are stripped of their organic charge and moral weight in order to be transmuted into sterilised visual products, institutional trophies or pop consumer goods. The category integrates three regimes by scene — museum/gallery (*pedagogical containment and ruinous memory*), press/archive (*panoptic institutional crime-trophy*), digital platforms (*pop aspiration and algorithmic evasion*) — articulated with Rancière (2000), Valencia (2010) and Mbembe (2003), in dialogue with Pardo León's conceptualisation of *the narco as an aesthetic category* (in progress). These

formulations, including the core category, are outputs of the system, not conclusions of the researcher: whether they are productive or require correction is exactly what the phase of human imputation must determine.

Four images the system sees — and four the researcher must impute

To understand what the system does and where its agency stops, it is worth showing the material and the actual outputs of the pipeline. The four images that follow are representative of the archive's principal regimes: the anti-narcotics operation of the 1970s, the portrait of the drug lord, the institutional press conference of the 1990s and the military patrol of the 2000s.

(a) *Bonanza marimbera*, 1970s.
`1970s_traqueteo_medellin_0076.jpg`

(b) Portrait of the drug lord, 1980s.
`1970_1990_nuevas_0001.jpg`

(c) Institutional communication, 1990s.
`1990s_Estetica_Yupi_0104.jpg`

(d) Military surveillance, 2000s.
`2001_Narcoterrorism_0294.jpg`

Four fragments of the corpus, four regimes of visibility. Each image was processed by the complete pipeline; the actual outputs by phase are tabulated below.

The table shows the literal outputs of the first two operations of the pipeline: the *microanalysis* module (dense multimodal description) and the *open coding* module (emergent codes with their memo).

Image	Microanalysis (dense description)	Open coding (emergent codes and memo)
(a) *Bonanza marimbera* *1970s*	Outdoor scene: a large group of young men in white T-shirts and shorts (some in boots, suggesting conscript uniform) destroying illicit plant material on arid ground. Uniformed personnel in the background. Collective action over open sacks.	**Codes:** Anti-drug militarisation; Public/symbolic eradication; Forced labour/military conscription; Performative state control; Transnational dimension of drug trafficking. **Memo:** The image connects with the narrative of the "War on Drugs" intensified in Latin America from the mid-twentieth century; young conscripts engaged in mass destruction illustrate the militarisation of anti-drug policy.
(b) Portrait of the drug lord *1980s*	Black-and-white photograph, medium shot. A man identifiable as Pablo Escobar looks at the camera with a calculated expression mixing smile and seriousness; he wears an open shirt and a jacket in 1980s style. The woman (presumably his wife María Victoria) composes a scene of apparent domestic normality.	**Codes:** The figure of the "narco" and his family setting; Construction of a public image; Power and social appearance; Normalisation of illegality; The role of women in organised crime. **Memo:** The image belongs to the height of the Medellín Cartel; aesthetics and dress anchor the scene in the decade; politically, it projects an image of relative domestic normality that neutralises the criminal reading.
(c) Institutional communication *1990s*	Close-up of a middle-aged man in a dark pinstriped suit, white shirt and tie; gold-rimmed glasses. Smiling expression, right hand raised with index finger pointing forward, a rhetorical-assertive gesture. A formal/professional context is suggested.	**Codes:** Political communication; Discursive authority; Projection of confidence; Institutional representation; Public management. **Memo:** The image encapsulates the aesthetics of the public figure in the political scene tied to the armed conflict and drug policy; attire, microphone and pointing gesture are recurrent iconography of state officials making statements.
(d) Military surveillance *2000s*	A lone soldier in camouflage uniform and tactical gear, prone on a rocky ridge, aiming a long-range rifle out of frame. Clear blue sky dominates the upper two thirds; arid terrain below; the soldier blends into the surroundings. An international agency watermark is visible.	**Codes:** Military surveillance; Conflict in rural/mountainous terrain; Armed state presence; Asymmetric warfare; Territorial control. **Memo:** The image relates to the visual narrative of the Colombian conflict and the "War on Drugs"; emblematic of armed forces operations, often US-supported, against armed groups and drug trafficking organisations.

Literal pipeline outputs for the four representative images. Source: the system's `codificacion_abierta.json`*, without editorial revision.*

In each case the system produces what looks like competent qualitative analysis: the descriptions are dense and referentially correct as regards the visible; the codes are consistent with socio-technical vocabularies on drug trafficking; the memos articulate plausible historical associations. And yet in no row is the operation completed: each cell contains a proposal without a subject to sustain it.

The first row illustrates the cleanest case: for the operation there is correspondence between the visible and the named. But the code *Performative state control* does not describe: it interprets. To sustain that interpretation as a thesis — that the operation *is* a performance of sovereignty and not mere

police procedure — requires arguing against other readings and answering for its political consequence; the system generated the code by statistical similarity with academic corpora, but it did not defend it and cannot defend it. The second row makes *agency without imputation* palpable: the system codes the construction of a public image and the normalisation of illegality, and even names the effect of moral neutralisation; but the question every investigation of the narco must answer — how does this aestheticisation relate to historical responsibility towards the victims? — it cannot even formulate. The last two rows belong to the institutional-military regime: the press conference codes discursive authority; the soldier, territorial control and asymmetric warfare. What the system cannot see is the continuity between them — that the prosecutor in Washington and the sniper on the hillside are moments of the same apparatus. That continuity is not a code: it is a hypothesis requiring argument, exposure to refutation and responsibility for the theoretical decision. Coding produces the fragments; integrating them into an apparatus — and deciding whether to name it, with Mbembe, necropolitics, or with Valencia, gore capitalism — requires an imputable agent.

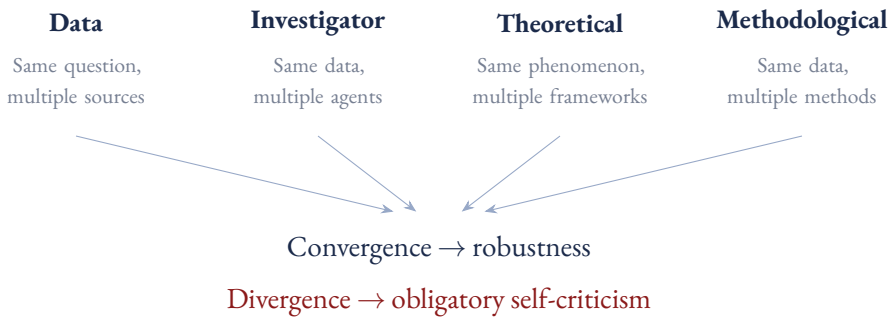
4.5 Achievements and persistence of the deficit

The prototype makes it possible to evaluate what the scaffolding compensates for and where the Mimesis III deficit persists. It achieves, first, real auditability: each code is linked to its source image, its analytic layer (Banks, 2007) and its relations, so that the evaluator can trace and refute it without reconstructing the process. It achieves verifiable relational structure: axial categories are schemas with conditions, mechanisms and consequences that can be checked against the corpus. And it achieves an explicit exit criterion: the cycle does not end arbitrarily, but when new segments produce no new codes.

Three tensions, however, reveal where the arc remains truncated. **First**: saturation is evaluated by the same LLM that produced the codes — a second-order infelicity, since the system declares itself validated without an external subject to validate the declaration. **Second**: theoretical integration is top-down — Rancière, Valencia and Mbembe are encoded as application schemas in the prompts; the system applies theory, it does not generate it, which contradicts the inductive spirit of GT. **Third**: constant comparison is self-referential — the same model that generates the codes compares them, whereas Glaser demanded external theoretical sensitivity. These tensions define what the researcher must contribute: validating that the core category emerges from the data and was not projected onto it; deciding whether saturation is genuine or premature; justifying the theoretical selection; and assuming final epistemic responsibility. That list does not describe optional supervision, but the exact place where imputation returns to the circuit. The scaffolding does not invent imputability; it *locates* it. Without that location, the system produces an illusion of rigour — Vallor's warning (2016): a badly used prosthesis atrophies the capacity it purports to compensate for.

4.6 Triangulation: controlling overconfidence

Grounded Theory resolves the criterion within the process, but not convergence across sources, methods and perspectives. That is the terrain of triangulation, introduced by Denzin (1970): observing the same phenomenon from multiple independent angles, not in order to confirm with more of the same, but to exploit the complementarity of perspectives and detect what each one misses. Denzin distinguished four modalities:



Data triangulation addresses the same question from multiple sources; **investigator** triangulation, the same data from multiple agents — in an LLM system, independent sessions with equivalent prompts whose outputs are compared; **theoretical** triangulation, the same phenomenon from distinct frameworks; **methodological** triangulation, the same data with different methods. Convergence produces robustness: when sources, agents, frameworks and methods coincide, the probability that the finding is an artefact of method decreases. But the most important function is activated by divergence: it makes self-criticism obligatory. Divergence between two parallel sessions is not a problem to be resolved by choosing the more plausible result, but a signal demanding examination of why the processes differed; it makes visible the model's biases, its sensitivity to prompt formulations, and the genuinely indeterminate points. Divergence is first-order epistemic information, not noise.

Together, GT and Triangulation constitute the hermeneutic scaffolding: the first provides the iterative structure and the internal criteria; the second, external control and self-criticism. They do not eliminate the Mimesis III deficit — the arc still requires a subject to complete it — but they make it visible, structure the researcher's contribution, and produce an output the human can evaluate, corroborate and reject.

5. The hermeneutic prosthesis

The scaffolding of §4 resolves the procedural problem, but it is not merely a set of procedures: from the standpoint of philosophy of technology it is a form of mediation that alters the relation between an agent and its object by interposing a tool. That mediation requires a philosophical foundation, because it determines how epistemic functions are distributed between system and researcher.

The category proposed for naming it is that of the **hermeneutic prosthesis**. One point should be

clarified that would otherwise open an apparent contradiction: if the LLM is an agent, calling it a "prosthesis" would seem a regression to the instrumental image §1 rejected. It is not. What is here called a prosthesis is not the LLM in isolation, but the system-scaffolding-LLM articulated to the researcher. The LLM retains its operative agency — it continues to interact, decide and adapt — but that agency is coupled, by means of the scaffolding, to the researcher's circuit of imputation. The word *prosthesis* designates a functional extension: not the replacement of an absent capacity, but the amplification of a present capacity under conditions of limitation. The qualifier *hermeneutic* specifies the type of extension: not computational but interpretive, governed by the protocols of understanding and validation that qualitative research has operationalised. A hermeneutic prosthesis is, then, a system in which the LLM — an operative agent in Floridi's sense — extends the researcher's configurative capacity (Mimesis II) while the researcher — an imputable agent in Ricoeur's sense — preserves and exercises the validative capacity (Mimesis III). The boundary between the two functions, and not their fusion, is the design principle.

5.1 Philosophical foundations of the design

Three traditions converge in the foundation of the concept. The **extended mind thesis** (Clark, 2003) explains why the coupling is possible: cognitive processes are not confined to the organism; when a tool couples reliably and fluidly to an agent's epistemic processes, it becomes constitutive of their cognitive system. Spectacles, paper, the calculator — and, under the right conditions, the LLM — are not external supports to cognition but components of it. The coupling between researcher and LLM is not accidental but constitutive of the epistemic architecture, and the hermeneutic scaffolding is precisely its specification.

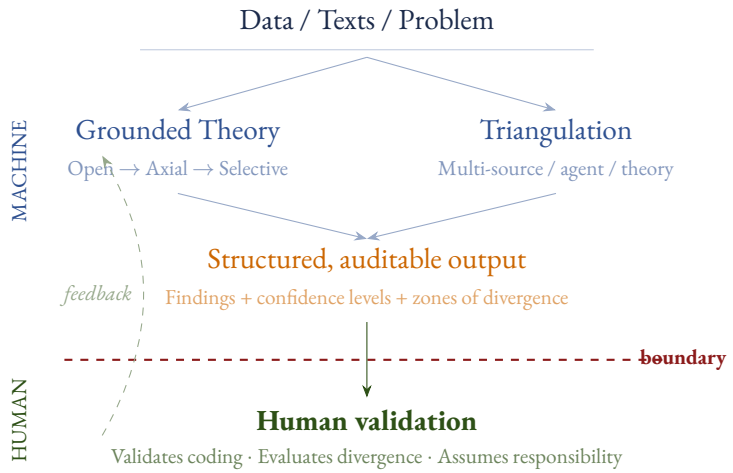
Phenomenology explains how the coupling is experienced as continuity. Merleau-Ponty (1945) showed that the body schema incorporates tools when the agent attains fluency: the blind person's cane is not something they touch but something with which they touch. The researcher who uses the scaffolding expertly does not operate *on* the system but *through* it. Varela, Thompson and Rosch (1991) add the enactivist dimension: cognition is embodied action, not the manipulation of representations. For the hermeneutic prosthesis, this means that validation (Mimesis III) cannot be wholly delegated to the system, because it requires the researcher's coupling to the practical consequences of interpretations: a world to return to, and in which findings are evaluated by their effects.

The **critical theory of technology** explains what political and ethical conditions the coupling must meet. Feenberg (2002, 2017) distinguishes **primary instrumentalisation** — technical operation in the abstract: decontextualisation, reduction to functioning — from **secondary instrumentalisation** — social embedding: systematisation, mediation, meaningful appropriation within a context of practice. LLMs operate by default at the primary layer: they deploy operative agency without being coupled to a horizon of imputation. The hermeneutic scaffolding is exactly secondary instrumentalisation: the set of mediations that returns technical operation to a context of practice, to criteria of validation, and to the subject who can be held the author of the judgement. Feenberg calls this movement

democratic rationalisation, and proposing it does not contradict the critical diagnosis of LLMs: it is its practical consequence. Vallor (2016) adds the dimension no architecture can ignore: the scaffolding must cultivate in the researcher the **epistemic virtues** — openness to refutation, reflexivity, tolerance of uncertainty — that make genuine validation possible. A badly designed prosthesis atrophies the capacities it purports to compensate for: if the researcher delegates coding without understanding it and accepts saturation without interrogating it, the scaffolding does not produce more robust knowledge, but an illusion of robustness with greater efficiency.

5.2 The concept and the architecture

The hermeneutic prosthesis can now be defined precisely: a system of knowledge production in which an LLM operates as a cognitive extension of the researcher in configuration (Mimesis II), while the researcher preserves and exercises validation (Mimesis III), and the scaffolding makes the boundary between them visible and auditable. It is not a conscious AI — a structural impossibility — but an **epistemically governable** AI: a system whose process can be audited, whose limits are declared, and whose failures can be detected without waiting for their consequences.



The machine contributes what it does better than the individual researcher: speed over massive corpora, consistency in applying protocols, pattern identification at scale, and the systematic generation of relational hypotheses. That contribution is genuine, and reducing it to the image of the passive instrument distorts what the system does. But the machine does not contribute validation, responsibility or judgement: those functions are, by design, on the other side of the boundary, because exercising them requires an agent who can be held the author of the decision, a figure embodied only by the researcher. The human validates that the codes reflect the corpus and do not project preestablished categories; decides on the sufficiency of saturation; justifies the theoretical selection; and, above all, assumes final epistemic responsibility — the moment at which conclusions cease to be system output and become assertions of the researcher, defensible before a community of peers. The researcher is not a supervisor approving outputs: they are the agent who closes the arc the system leaves truncated. Without scaffolding, that boundary is invisible, and the risk is accepting outputs as though they were

already validations; with scaffolding, each output includes its provenance, its declared limits, and the conditions under which it ought to be rejected. The boundary is not a barrier but an interface: the place where the machine ends and the human begins, marked precisely so that neither occupies the place of the other.

6. Conclusion: from diagnosis to design

The path ran from diagnosis to design. Section 1 showed that the ontological question is legitimate but insufficient, and that epistemic uselessness does not follow from ontological lack. Section 2 located the LLM within Ricoeur's arc: a Mimesis II machine whose arc is truncated, with no world to return to and no subject to whom what is configured can be imputed. Section 3 recategorised the error as an infelicity of epistemic authority: acts that claim the regime of reasons without being able to inhabit it. Section 4 built the scaffolding — Grounded Theory as protocol with a criterion per phase, Triangulation as external control, and the prototype as demonstration — and §5 named the resulting architecture: the hermeneutic prosthesis as an agential redistribution between an operative and an imputable agent.

6.1 The structure of the argument

The argument is a chain, not a list: each resolution is the condition of possibility of the next.

Tension	Resolution	Route
Ontology vs. Epistemology	→ Epistemological turn	§1
Understanding vs. Configuration	→ Mimesis II (Ricoeur)	§2
Validity vs. Performance	→ Infelicities (Austin)	§3
Heuristic vs. Method	→ GT + Triangulation	§4
Power vs. Responsibility	→ Hermeneutic prosthesis	§5

The ontology/epistemology turn (§1) makes possible the configuration/validation diagnosis (§2), which makes possible the recategorisation of error (§3), which makes possible the protocol with a criterion (§4), which makes possible the concept of the hermeneutic prosthesis (§5). The thesis running through the chain: the tensions that block the debate on LLMs are epistemological, not ontological, and are therefore resolvable by procedural design informed by applied hermeneutics. The ontological diagnosis — the LLM does not understand — is correct; the conclusion — there is nothing to design — is an invalid inference. The piece that makes that invalidity productive is the distinction between two modes of agency: the LLM is an operative agent, not an imputable one, and the design of the scaffolding places each mode where it belongs.

6.2 Contribution, limits and lines of research

The contribution has four components. **Conceptual**: the category of *infelicity of epistemic authority* as a precise alternative to "hallucination", and that of *agency without imputation* as a description of the agent that commits it; together, the displacement from psychology to pragmatics and from correcting the system to validating the output. **Methodological**: the demonstration that GT can be translated into a pipeline of prompts with an explicit criterion of validity per phase — declaration of limits, traceability by code, auditable relational structure, bidirectional traceability, openness to refutation — properties absent in CoT, ToT, ReAct and agentic AI. **Architectural**: the concept of the *hermeneutic prosthesis* as a framework that distributes epistemic functions explicitly and auditably, grounded in Clark, Merleau-Ponty, Varela, Feenberg and Vallor. **Empirical**: the implementation that turns the proposal from a theoretical possibility into an evaluable research programme, demonstrating both what the scaffolding achieves and where the deficit persists.

The limits must be explicit. The scaffolding does not eliminate the LLM's epistemic deficit: the three tensions of §4.5 are structural conditions that the design makes visible but does not resolve; what it does is displace the problem into the space of interaction between system and researcher. Grounded Theory is not the only possible scaffolding — it is the most adequate for qualitative analysis of textual and iconographic corpora; other contexts would require different criteria. And the epistemological question does not replace the ontological one: the critique of Searle, Dreyfus, Gadamer and Heidegger remains correct and relevant. Ontology and epistemology do not compete: the second presupposes the first.

Three lines open up. The **empirical**: a programme of comparative tests contrasting the GT/Triangulation scaffolding with CoT, ToT and agentic architectures on defined tasks, with metrics including process auditability and traceability of findings, not only factual accuracy. The **political and epistemic-ecological**: scaffoldings implemented at scale in universities, newsrooms or courts are not neutral — they structure what counts as rigorous research and which knowledges are excluded by design — a dimension requiring independent philosophical and social analysis. The **cosmotechnical**: Yuk Hui (2019) argues that all technics is rooted in a particular cosmology; Grounded Theory has roots in North American sociology, the hermeneutics that grounds it is European, and LLMs are trained on corpora overrepresenting the global North. Dussel, Escobar and Quijano make it possible to ask, from the Latin American philosophical tradition, what forms of knowledge are structurally excluded by a scaffolding that was not designed to include them. Answering does not invalidate the proposal: it situates it.

The initial question — how is the knowledge produced by a system that acts and generates coherent discourse without participating in the conditions of understanding or of responsibility to be validated? — has no definitive answer. What this work shows is that it can be productively reoriented: instead of seeking in the machine capacities it cannot have, recognising the operative agency it deploys and designing, around it, the conditions that allow imputation to return to the subject who can sustain it. The hermeneutic prosthesis is not the final solution to a philosophical problem: it is the form of a

research programme that turns an ontological dead end into an open space of epistemological design.

Declaration on the use of artificial intelligence

In accordance with emerging norms of transparency in academic publishing, the author declares the following uses of artificial intelligence systems in the preparation of this work.

As object of study and research instrument. The prototype described in §4.5 uses Google Gemini in vision mode as the language model operated by the pipeline. The outputs reproduced in the table of §4.5.1 — dense descriptions, emergent codes and memos — are literal machine outputs, presented without editorial revision precisely because their status as unimputed outputs is what the argument examines. The core category reported in §4.5 is likewise a system output, not a conclusion of the author, and is presented as such.

In the preparation of the text. Anthropic's Claude Opus (versions 4.6, 4.7 and 4.8) was used as support in editing and improving the prose of the Spanish version of this article, and in preparing this English version from that Spanish original. In both cases the model was used for linguistic formulation, not for the generation of arguments.

What is not delegated. The research question, the philosophical argument, the reading of Ricoeur, Austin, Habermas and Floridi, the concept of the hermeneutic prosthesis, the choice and application of Grounded Theory, the design and implementation of the prototype, the critical interpretation of its outputs, and every final decision on content are the author's own. The author assumes full epistemic responsibility for the claims made in this text.

This declaration is not merely a procedural formality. The article argues that the difference between an operative agent and an imputable agent must be made explicit and auditable wherever artificial systems participate in the production of knowledge; a text making that argument is obliged to apply the criterion to itself.

Bibliography

Core works of the argument

[1] Austin, J. L. (1975). *How to do things with words* (2nd ed., J. O. Urmson & M. Sbisà, Eds.). Harvard University Press. (Original work published 1962).

[2] Feenberg, A. (2002). *Transforming technology: A critical theory revisited*. Oxford University Press.

[3] Floridi, L. (2019). *The logic of information: A theory of philosophy as conceptual design*. Oxford University Press. https://doi.org/10.1093/oso/9780198833635.001.0001

[4] Floridi, L., & Sanders, J. W. (2004). On the morality of artificial agents. *Minds and Machines*, *14*(3), 349–379. https://doi.org/10.1023/B:MIND.0000035461.63578.9d

[5] Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Aldine.

[6] Habermas, J. (1984). *The theory of communicative action, Vol. 1: Reason and the rationalization of society* (T. McCarthy, Trans.). Beacon Press. (Original work published 1981).

[7] Ricoeur, P. (1984). *Time and narrative, Vol. 1* (K. McLaughlin & D. Pellauer, Trans.). University of Chicago Press. (Original work published 1983).

[8] Ricoeur, P. (1992). *Oneself as another* (K. Blamey, Trans.). University of Chicago Press. (Original work published 1990).

[9] Smith, B. C. (2019). *The promise of artificial intelligence: Reckoning and judgment*. MIT Press.

[10] Strauss, A. L., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. SAGE.

[11] Vallor, S. (2016). *Technology and the virtues: A philosophical guide to a future worth wanting*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780190498511.001.0001

Complementary works

[12] Banks, M. (2007). *Using visual data in qualitative research*. SAGE.

[13] Clark, A. (2003). *Natural-born cyborgs: Minds, technologies, and the future of human intelligence*. Oxford University Press.

[14] Denzin, N. K. (1970). *The research act: A theoretical introduction to sociological methods*. Aldine.

[15] Dreyfus, H. L. (1992). *What computers still can't do: A critique of artificial reason*. MIT Press.

[16] Feenberg, A. (2017). *Technosystem: The social life of reason*. Harvard University Press.

[17] Gadamer, H.-G. (2004). *Truth and method* (2nd rev. ed., J. Weinsheimer & D. G. Marshall, Trans.). Continuum. (Original work published 1960).

[18] Heidegger, M. (1962). *Being and time* (J. Macquarrie & E. Robinson, Trans.). Harper & Row. (Original work published 1927).

[19] Heidegger, M. (1977). The question concerning technology. In *The question concerning technology and other essays* (W. Lovitt, Trans., pp. 3–35). Harper & Row. (Original work published 1954).

[20] Hui, Y. (2019). *Recursivity and contingency*. Rowman & Littlefield.

[21] Mbembe, A. (2003). Necropolitics (L. Meintjes, Trans.). *Public Culture*, *15*(1), 11–40.

[22] Merleau-Ponty, M. (2012). *Phenomenology of perception* (D. A. Landes, Trans.). Routledge. (Original work published 1945).

[23] Pardo León, J. A. (in progress). *Narcotráfico y cultura: lo narco como categoría estética* [Doctoral dissertation in progress]. Universidad Santo Tomás, Doctorate in Philosophy.

[24] Rancière, J. (2004). *The politics of aesthetics: The distribution of the sensible* (G. Rockhill, Trans.). Continuum. (Original work published 2000).

[25] Reichenbach, H. (1938). *Experience and prediction: An analysis of the foundations and the structure of knowledge*. University of Chicago Press.

[26] Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, *3*(3), 417–457. `https://doi.org/10.1017/S0140525X00005756`

[27] Valencia, S. (2018). *Gore capitalism* (J. Pluecker, Trans.). Semiotext(e). (Original work published 2010).

[28] Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. MIT Press.

[29] Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q., & Zhou, D. (2022). Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems*, *35*, 24824–24837.

[30] Yao, S., Yu, D., Zhao, J., Shafran, I., Griffiths, T. L., Cao, Y., & Narasimhan, K. (2023). Tree of thoughts: Deliberate problem solving with large language models. *Advances in Neural Information Processing Systems*, *36*.

[31] Yao, S., Zhao, J., Yu, D., Du, N., Shafran, I., Narasimhan, K., & Cao, Y. (2023). ReAct: Synergizing reasoning and acting in language models. In *Proceedings of the 11th International Conference on Learning Representations (ICLR 2023)*.

Artificial intelligence tools

[32] Anthropic. (2026). *Claude Opus* (versions 4.6, 4.7 and 4.8) [Large language model]. `https://claude.ai` (Used as support in editing the prose of the Spanish version and in preparing the English version; see the Declaration on the use of artificial intelligence).

[33] Google. (2026). *Gemini* (vision mode) [Multimodal large language model]. `https://gemini.google.com` (Language model operated by the prototype described in §4.5).