

Nanopublication – AI opacity has no codex-bearer to disclose it

by Arnaud Quercy

The boundary this source draws for the lens is whether the instituted invariant requires an answerable intending subject. The framework's whole apparatus of disclosure – explicability (P19a), ricochet (P18a), even proto-codex induction (P14a) – is built for practices where someone, however inarticulately, holds the rule. The epistemology-of-AI literature surveyed here raises, without resolving, cases where the rule was never held by anyone: induced from data, legible only (if at all) through mechanistic interpretability that is still immature. The lens has no vocabulary for an invariant instituted by nobody; the source names exactly the condition (opening the black box) that would settle whether such a case is real or merely apparent.

WHAT THE SOURCE SAYS ^[1]

The entry surveys philosophical work on whether humans are epistemically justified in forming beliefs, claiming knowledge, or gaining understanding from AI outputs. It treats AI both as an object of inquiry – asking what 'opacity,' 'interpretability,' and 'explainability' mean, where opacity comes from, and what routes exist to reduce it – and as a tool or quasi-agent whose outputs ground human belief. It surveys accounts of justification (reliabilism, instrument-based knowledge, statistical/inductive justification for machine-learning results) and asks whether black-box models can still deliver scientific understanding, concluding that most of these questions remain open and under active philosophical dispute.

WHERE THE LENS GRIPS

P3a's given/instituted axis and P19a's explicability condition brush directly against Section 2's treatment of opacity. Humphreys's definition – opacity as an agent's lack of knowledge of 'epistemically relevant elements' of a process – reads almost like a formal version of the given/instituted split: what a sensor or user can register at the surface (the given) versus the rule that elects relations in that surface and would need disclosure to be read (the instituted). The article's own worry about whether ML models 'contain

such assumptions' that remain 'unknown until the black box is opened' is structurally the same worry P19a names: an invariant that exists but is mute until articulated in dialogue.

WHERE IT BREAKS

The grip slips at exactly the point the source itself flags as undecided: P19a's explicability condition and P18a's ricochet both presuppose a subject who holds the rule and can, in dialogue, answer 'why these choices.' The entry reports that for many ANNs there may be no such subject even in principle – 'humans do not implement substantive assumptions about their targets in ANNs,' so 'the question of whether such assumptions are appropriately implemented doesn't arise.' That is not an artist withholding or failing to articulate an invariant (P19a's case); it is the possible absence of anyone, human or otherwise, who ever held the rule to begin with. This slips IF one reads 'instituted invariant' as requiring an intending codex-builder answerable in dialogue (the ricochet model, P18a); it holds – does not break – IF one reads P3a's instituted invariant merely as 'the rule that elects relations in the surface,' agent-independent, in which case a trained network's weights simply are an instituted invariant nobody has yet deciphered, no different in kind from P14a's proto-codex, just radically harder to induct from. The source does not adjudicate which reading is correct; it leaves it as exactly this open question.

PROPOSITIONAL GRID ^[2]

Read this grid. "Not evaluated" does **not** mean the source has nothing to say about that proposition. It means **the reading did not instruct it**. That is not a result, and it counts in no denominator. A silence, by contrast, is a finding: the reading touched the proposition without concluding, or noted its absence itself.

Evaluated

proposi- tion		audit status	relational verdict
P3a	Double Structure of the In-variant	audited	pressure-against
P14a	The Proto-Codex	not audited	open-question
P18a	The Ricochet Effect	not audited	open-question
P19a	Explicability as Condition	not audited	pressure-against

What was instructed. These propositions have been judged and carry a current statement: it is **that** statement the reading tested, not the founding one.

▪ **P3a** – judged in Note 05 · current statement : say GIVEN invariant and INSTITUTED invariant, never 'physical' and 'intentional' – the axis is MODE OF EXISTENCE, read as regime of access: the given is what a sensor measures on the surface WITHOUT the codex; the instituted is the rule that elects relations in that surface and that only the disclosed codex deciphers

Not evaluated (33) – P1, P2, P3, P4, P5, P6, P7, P8, P9, P10, P11, P12, P13, P14, P14b, P15, P16, P16a, P17, P18, P19, P20, P21, P22, P23, P24, P25, P26, P27, P28, P29, P30, P31

Overall verdict: Bounded / Open – boundary case – taxonomy of the August 2026 pass, predating the move to relational verdicts. Kept as metadata: the gap between the two forms is data, not residue. This reading’s judgement is in the grid above.

Reading of 2026-10-17 · model claude-sonnet-5 · source 138,094 chars · 1 attempt(s)

REFERENCES

[1] Epistemology of Artificial Intelligence. Stanford Encyclopedia of Philosophy, 2026-10-17. <https://plato.stanford.edu/entries/epistemology-ai/>

[2] Quercy, A. (2026). The 31 Propositions of Ideamorphism. <https://ideamorphism.org/en/publications/2026/03/the-31-propositions-of-ideamorphism.html>

EPISTEMIC PROFILE

Claim type	boundary reading
Voice	third person analysis
Epistemic status	bounded grip
Methodology	Boundary-finding lens applied to one source, model claude-sonnet-5, 1 attempt(s), traceability check passed.

CHECKSUM (SHA-256)

76adcfae251b9b4438b0065075d56ff81c7fb953814d2e257be25f80a738461d

Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)
Author Arnaud Quercy

Asset code	IDS-RNI0335
Identifier	NAN-RDG000288
Version	1
Published	2026-10-17

ISSN: applied for – U.S. ISSN Center, Library of Congress

© 2026 Ideamorphic Studies

Published by Art Quam Anima Publishing New York,
an assumed name of AQA Publishing LLC

c/o Northwest Registered Agent, 418 Broadway Ste N
Albany, NY 12207, USA
+1 917-764-5470

NY DOS ID 7624376 · Assumed Name ID 7658709

publishing.artquamanima.com

Last updated: 2026-10-17

Persistent URI: <https://ideamorphism.org/en/nanopubs/2026/10/IDS-RNI0335-ai-opacity-has-no-codex-bearer-to-disclose-it.pdf>