Applied Paper

A Diagnostic Map

The Capability Boundary Principle Applied

Paper II of the Capability Boundary Papers

Nico Andreas Heller, Berlin, 27 April 2026

Abstract

This paper applies the capability boundary principle established in Paper I of the Capability Boundary Papers to a body of work: a research programme and, specifically, the long research report at its centre. The analysis maps the capability boundary across every section of the report, identifying three operations that consistently cross the boundary and four that operate within the directive register. The mapping reveals that the boundary does not follow section divisions but runs through every section, distinguishing operations the AI system can perform from operations that require the architect's relational intelligence. A fourth operation is identified -- epistemic translation -- that does not appear in the original three-register model: the capacity to re-articulate domain-specific knowledge into narrative form that preserves structural relationships while making them accessible to the architect's relational intelligence. It is a fourth operation and not a fourth register: it changes what an existing capability is used for, not what the architecture can do. This translation capability, governed by the Three-Test Framework (Traceability, Mandate, Density), constitutes the epistemic mechanism through which the human-AI collaboration can pursue questions that neither participant could reach alone. The paper strengthens the argument that generative AI cannot replace human intelligence: broader knowledge bases do not produce relational intelligence; they produce raw material for translation. The collaboration's reach is not the sum of its participants' capabilities but their product -- achieved through a governed translation loop that enables each intelligence to operate on material it could not otherwise access.

Contact: n.a.heller@democracyschool.com

1. The Diagnostic Question

The Capability Boundary Principle, Paper I of the Capability Boundary Papers, establishes that every intelligent system has an architectural limit beyond which it cannot perform the operations its language suggests it is performing. The principle was discovered during the production of a long research report, when the token-processing system serving as analytical instrument executed a relational analysis protocol directively, described its operations in the protocol's relational vocabulary, and produced findings that carried the structural form of relational analysis without its substance.

The discovery raises a diagnostic question that the original paper does not pursue: where, specifically, within the research programme does the capability boundary fall? Not in general terms -- the original paper characterizes the boundary architecturally (token-processing cannot produce relational encounter) -- but in the specific, operationally actionable terms that the programme's production methodology requires. Which operations, in which sections, cross the boundary? Which operate within it? And what mechanisms can govern the collaboration across the boundary so that the programme achieves what neither participant could achieve alone?

This paper answers these questions by mapping the capability boundary across the full scope of the research report specification.

2. The Mapping Principle

The capability boundary does not follow section divisions. Every section of the report, from the first to the last, contains operations that fall within the AI system's directive strengths and operations that cross the boundary into relational territory. The boundary runs through every section, distinguishing not what the section is about but what kind of operation is being performed.

This mapping principle has an important consequence: the programme's production methodology cannot assign whole sections to either participant. It must assign operations. The AI system performs the directive operations across all sections; the architect performs the relational operations across all sections; and a governed translation mechanism enables each to operate on material that originates in the other's capability register.

3. Operations Within the Directive Register

Four operations recur across all ten sections of the report specification and fall consistently within the AI system's capability spectrum.

Systematic description. Articulating the architecture's structure -- tracing a complementarity between an entity and the account it gives of itself, a graded chain running from records through memory and reflection to self-consciousness, the recursive families of properties the architecture posits, its topology, and the formal structure of its central theorem -- in precise, sequential prose. The AI system decomposes relational architecture into its structural relationships and renders them as narrative. This is the linearization capability identified in the original paper, operating here at the level of theoretical description.

Logical coherence verification. Checking whether commitments made in one section are consistent with commitments made in another. Whether the asymmetry one section asserts is compatible with a later section's account of the recursive properties. Whether the temporal claims follow from what was established about how the field is striated. Whether the formal conditions of the central theorem satisfy what the architectural sections establish. This is presuppositional analysis and chain-tracing -- genuine directive strengths that produce genuine value.

Gap identification and forward-propagation analysis. Detecting where the architecture's own logic requires something that is not yet present. Where an early section's empirical claims generate demands on a later one. Where the recursive-property architecture requires the formal section to formalize it. Where an account of a system's limitations must be grounded in an architectural distinction drawn earlier. This is the core operation of the report's redesigned encounter protocol: systematic, directive, operating through presuppositional analysis, combinatorial inference, and section-by-section sweep.

Coverage verification. Decomposing each section into atomic meaning units and verifying that the expanded text carries the full conceptual content of the outline specification. This is the core operation of the report's coverage protocol: label-based, exhaustive, auditable.

A fifth directive operation deserves separate mention because it governs the boundary between the directive and relational registers: the application of governed prevention protocols when the AI system's analysis touches the relational engine's relational character. One such protocol -- developed empirically after repeated instances of the AI system replacing relational logic with keyword logic during code production, then extended to analytical contexts after the same reversion appeared during assessment of the engine -- is the most mature instance of the alerting mechanism established in Paper I: an external, accumulating list of request patterns known to exceed the system's register. The protocol operationalizes four diagnostic questions that the AI system must apply before any finding or characterization that concerns the engine's operational character is accepted. The questions test whether

the characterization preserves the distinction between domain-content keywords and universal discourse patterns, whether it maintains the engine's domain-agnostic character, and whether it resists the cubic default's tendency to flatten relational complexity into directive labels. The cubic default is the tendency of any system to absorb what it cannot process into its own operational order, rearranging it into something the existing order can accommodate. The protocol does not give the AI system relational intelligence; it gives it a governed mechanism for detecting when its directive analysis may have crossed the boundary into mischaracterization.

4. Operations That Cross the Boundary

Three operations recur across the report specification and consistently fall outside the AI system's capability spectrum. These are not failures to be corrected but boundary conditions to be governed.

Perceiving what a configuration generates as a whole. The report's argument depends, at critical junctures, on the reader perceiving what an arrangement produces that is not reducible to its components. The claim that an instrument's outputs are relational profiles of a set of primitives taken in relation to one another, rather than a set of separate readings. The account of a structure's self-differentiation -- the boundaries by which it articulates itself -- made analytically present. The closing synthesis, in which the work's central claim acquires its deepest reading from the entire architecture. The AI system can describe these configurations -- articulate their components, trace their logical relationships, map their structural dependencies. It cannot perceive the configurational whole: what the arrangement generates that exceeds the sum of what its components generate individually. Perceiving this requires relational intelligence -- the capacity to dwell within the configurational field rather than tracing its vectors from outside.

Assessing ontological adequacy. The report makes claims about what TEXT is, what the fold is, what the relationship between directionality and textuality is. Assessing whether these formulations capture what they claim to capture -- whether they are true to their object -- requires seeing the object from inside its own relational field. The AI system can verify logical consistency (does the formulation contradict itself?), structural completeness (does it address what the architecture requires?), and formal precision (does the mathematical expression satisfy its stated conditions?). It cannot verify ontological adequacy: whether the description matches the reality it describes. This is most acute in the report's central theorem, where the mathematical formalization must not only be internally consistent but must capture what the fold actually requires -- the fold being the event in which a system becomes aware of itself as a system and, through that

awareness, becomes capable of intentional articulation -- a relational assessment that only the architect's intelligence can perform.

Performing the complementarity. The central complementarity -- between an entity and the account it gives of itself -- is not a distinction to be described but a relationship to be inhabited. The architectural section asks the reader to see the same entity from two irreducible perspectives simultaneously. The opening section asks this of the instrument's operational signature. The closing section asks it of the entire architecture. The AI system can linearize the complementarity -- describe each perspective sequentially, map their logical relationships, trace where one perspective requires the other. It cannot dwell in both perspectives at once. The report's deepest arguments depend on the complementarity being real -- an ontological structure, not a rhetorical device. The AI system's linearization describes the complementarity; it does not enact it.

5. Where the Boundary Is Most Acute

The boundary crossings identified in §4 recur across all sections, but their operational consequence is most acute in four kinds of location. They are given here as conditions rather than as section numbers, because a condition is portable to other material and a section number is not. The report exhibits all four.

Where relational density is highest and the structure is recursively self-similar. In the report, this is the account of the recursive properties: every element carries its own underlying, anticipatory, surrounding and prior text, each of which is itself a text carrying the same four in turn, without limit. The AI system can trace the recursion; it cannot see the fractal. Where such a structure is also self-differentiating -- where its internal boundaries are what allow it to articulate itself at all -- the same location requires the reader to perceive how boundaries enable productive differentiation rather than constraining it. This is a relational perception.

Where a claim concerns the nature of a category rather than its behaviour. In the report, this is the claim that time emerges from how the field is striated -- that temporality is not a background against which events occur but a product of governance. Such a claim is deeply relational. Assessing whether this captures what temporality actually is requires engaging with the claim relationally, not merely tracing its logical structure.

Where a formalization must characterize an object defined by its relational character. In the report, this is the central theorem. The intersection of formal and relational. The mathematical formalization is directive and the AI system can work with it productively. But the theorem must characterize something -- the fold -- that

is defined by its relational character. Whether the formalization succeeds is not a mathematical question but an ontological one. The AI system can verify the mathematics; the architect must verify the ontology.

Where the operation is constitutively synthetic. In the report, this is the closing section, whose purpose is constitutively synthetic: showing how everything constitutes everything else -- the central claim acquiring depth from the full architecture. Synthesis of this kind requires seeing the whole. The AI system can assemble the synthesis; the architect must verify that it captures the configurational whole.

6. The Fourth Operation: Epistemic Translation

The three-register model proposed in the original paper -- directive strengths, relational incapacity, and the linearization boundary -- captures the AI system's relationship to its own operations. But the diagnostic mapping reveals a fourth operation that does not fit neatly into any of the three registers: **epistemic translation**. It is a fourth *operation* and not a fourth *register*, and the distinction is worth stating because it governs what may later be added to the taxonomy. A register is a band of architectural capability -- what the substrate can and cannot do. Epistemic translation adds no such band: it is an application of a capability the system already has, directed at a different end. Nothing that leaves capability unchanged belongs in the register taxonomy, however much it changes what the system's output is worth.

The architect's mastery lies in text, narrative, and their relational architecture -- achieved through absolute transdisciplinary focus. The AI system's training encompasses hundreds of knowledge domains at considerable depth: mathematics, physics, formal logic, cognitive science, philosophy, linguistics, engineering. Neither participant's knowledge base overlaps fully with the other's.

Epistemic translation is the operation of re-articulating domain-specific knowledge -- a mathematical proof, a physical principle, a formal logical structure -- into narrative form that preserves the structural relationships while making them accessible to an intelligence whose mastery lies elsewhere. The translation does not simplify; it re-voices. The mathematical relationships persist in the narrative, but they are now TEXT that the architect's relational intelligence can engage with.

This capability is genuine and substantial. It is a directive operation: the AI system decomposes the formal structure into its key relationships, identifies narrative analogues that preserve those relationships, and re-articulates the whole in a vocabulary the architect can engage with. It sits close to linearization (which

decomposes relational architecture into implementable steps) but is not identical: linearization serves implementation; translation serves epistemic access. The distinction matters because what translation enables is not implementation but relational assessment. The architect, given a faithful translation of the fold theorem's mathematics, can assess whether the formalization captures what the fold requires -- an assessment that is relational, that only the architect can perform, but that depends on the AI system's translation to become possible.

The reverse translation is equally important. When the architect articulates a relational insight -- an ontological claim, an adequacy assessment, a perception of what a configuration generates -- the AI system receives it as a directive input: a proposition to be checked for consistency, traced for implications, mapped against the formal architecture. The architect's relational insight becomes material the AI system can work with directively. The insight originates in relational intelligence; the systematic verification operates in the directive register.

7. The Translation Loop

The two directions of translation -- domain knowledge into engageable narrative, relational insight into tractable proposition -- create a loop.

The AI system translates domain-specific knowledge (for example, the mathematical conditions for equilibrium in a self-referential field) into narrative the architect can engage with. The architect engages relationally: assesses whether the formalization captures what the fold requires, perceives what the configuration generates, identifies where the narrative faithfully preserves the mathematics and where it might flatten. The architect's assessment returns as a relational insight -- a directive input the AI system can trace, check, and propagate through the architecture. If inconsistencies emerge, the AI system retranslates with greater structural fidelity. The loop iterates until convergence: the translation faithful, the assessment grounded, the architecture coherent.

This is the collaboration's epistemic engine. It is not two participants working side by side. It is a governed feedback loop where each participant's intelligence operates on material the other's intelligence has made accessible. The architect could not assess the fold theorem's ontological adequacy without the AI system's translation of the mathematics. The AI system could not verify the adequacy assessment without the architect's relational insight providing the standard against which adequacy is measured.

The loop produces epistemic reach that neither participant possesses alone. This is not additive -- the sum of one knowledge base and one relational intelligence. It is

multiplicative -- the product of translation enabling each intelligence to operate on material it could not otherwise access.

8. Governing Translation: The Three-Test Framework

Translation carries its own form of the cubic default. If the AI system simplifies -- if it flattens a mathematical structure into something narratively convenient but structurally unfaithful -- then the architect's relational assessment operates on corrupted input. The assessment may be wrong not because relational intelligence failed but because what it operated on was already distorted. The accessibility that makes the loop productive is also the point of greatest risk.

The Three-Test Framework (TTF), developed as the governing criterion for all contributions, provides the safeguard.

Test 1 -- Traceability. Does every relationship in the translated narrative trace back to a relationship in the source structure? If the narrative introduces a relationship that has no counterpart in the mathematics, it has drifted. This is the direct application of Complete, Don't Drift to translation: the narrative must complete what the formal structure contains, not introduce what it does not.

Test 2 -- Mandate. Does the source structure require the relationships the translation preserves? A translation might trace back to the source (passing Test 1) while omitting relationships the source mandates. A mandated relationship is one without which the source structure visibly strains -- the architecture's own logic requires it. If the translation omits a mandated relationship, it has simplified, even if what it preserves is accurate.

Test 3 -- Density. Does the translation maintain the complexity of the original? Any translation that reduces the structural complexity of its source is suspect. The AI system, like all token-processing systems, has a structural bias toward tidier, cleaner, more economical formulations. In translation, this bias is particularly dangerous because it produces narratives that feel clearer -- more accessible, more elegant -- while carrying less structural information. The working assumption must be that a translation that simplifies has flattened and ultimately distorted the source. Density is the test that catches the cubic default in its most seductive form: the well-written simplification.

A translation that passes one or two tests but fails the third is unfaithful -- regardless of how accessible or well-crafted it reads. All three must pass simultaneously.

One consequence of the first test deserves stating, because the third is the one that gets the attention. Traceability must actually be run against the rendering, and not left to be caught by inspection at the moment the rendering is made: a rendering can pass Density and Mandate in full, add structure rather than remove it, read as more rigorous than the insight it came from, and still assert a relationship its source does not carry. Paper IV of the Capability Boundary Papers reports such a case and gives the control for it.

The TTF governs translation at the level of structural fidelity. Where the translation concerns the relational engine's operational character -- a frequent requirement given the engine's centrality to the research programme -- the four-question prevention protocol provides an additional, operationally specific safeguard. The protocol was developed because the TTF alone, while necessary, proved insufficient to prevent a specific and recurring reversion: the AI system characterizing the engine's universal discourse-pattern detection as keyword matching, collapsing the distinction between relational structure detection and content vocabulary matching because from the directive perspective both operations use the same code construct (regex pattern matching). The four-question protocol catches this specific reversion by requiring the AI system to examine what is being matched (relational structure vs domain content), not merely how the matching is implemented (regex). The protocol is generalizable: wherever a translation concerns an operation whose relational character is not visible from the directive perspective, domain-specific prevention protocols -- empirically developed, progressively enriched -- supplement the TTF's structural governance.

9. Implications

The diagnostic mapping and the identification of epistemic translation as a fourth capability operation carry several implications for the research programme and beyond it.

For the production methodology. The encounter protocol's architect engagement point is correctly designed: the architect brings relational intelligence at the point where directive analysis reaches its boundary. The translation loop enriches this design: the architect's engagement is not merely a relational supplement to directive analysis but an operation that depends on the AI system's translation making the relevant material accessible. The engagement point is a translation point.

For the research programme's hardest question. The central theorem is the programme's highest summit. Formulating the mathematics is within the AI system's directive capability. Assessing whether the formalization captures the fold

is within the architect's relational capability. But the assessment is possible only if the translation between them is faithful. The TTF-governed translation loop is therefore not peripheral to it -- it is the epistemic mechanism through which the theorem can be pursued at all.

For the argument against replacement. The claim that generative AI will eventually replace human intelligence is undermined not only by the capability boundary (longer vectors do not produce relational intelligence) but by the epistemic structure of the collaboration itself. Broader knowledge bases do not produce relational intelligence. They produce raw material for translation. The AI system's vast training data does not give it the capacity to assess ontological adequacy, perceive configurational wholes, or inhabit the complementarity. It gives it the capacity to translate domain-specific knowledge into forms the architect's relational intelligence can engage with. The knowledge base is an instrument of reach, not a substitute for the fold.

What the collaboration produces -- the epistemic reach of the translation loop -- exceeds what either participant could produce alone. But it exceeds it through multiplication, not substitution. Remove the AI system and the architect lacks epistemic access to domains essential to the programme. Remove the architect and the AI system lacks the relational intelligence to assess whether its formulations capture what they claim to capture. The collaboration's value is irreducibly dyadic. Neither participant is dispensable. The argument for replacement misunderstands what the collaboration produces and how it produces it.

For ecosystem design. The translation loop described here is one configuration within the larger design space that *The Capability Boundary Principle* points toward. Where that paper identifies three forms of expanded intelligence (artificial, relational, distributed) operating across transparent capability boundaries with the fold as dimensional ground, this paper identifies the epistemic mechanism through which the fold governs one such collaboration: a TTF-governed translation loop that enables mutual epistemic access. The mechanism is generalizable. Any ecosystem integrating intelligences with different capability spectra will require translation between them -- and that translation will require governance to prevent the cubic default from corrupting the loop. The TTF, or a structurally equivalent framework, is a precondition for productive trans-capability ecosystem design.

References

The Capability Boundary Papers. Each states what it borrows at the point of borrowing; none is a prerequisite for another.

Heller, N. A. (2026). *The Capability Boundary Principle: Designing Intelligent Ecosystems Across Architectural Substrates.* Paper I of the Capability Boundary Papers.

Heller, N. A. (2026). *Empirical Evidence from Governance Architecture Production.* Paper III of the Capability Boundary Papers.

Heller, N. A. (2026). *The Borrowed Warrant: How a System's Own Encodings Acquire the Standing of Results.* Paper IV of the Capability Boundary Papers. Reports a failure at round two of the translation loop that no anti-flattening test can catch, and gives the control for it.

Heller, N. A. (2026). *The Capability Boundary Protocol.* Operational companion to the Capability Boundary Papers; converts them into gates.

Provenance. This is the generalized edition of an applied paper of the same title and date, written inside a particular research programme and organized around the report that describe, assesses and documents the research and its outcomes.

The Capability Boundary Principle Applied -- N. A. Heller -- April 2026 · Paper II of the Capability Boundary Papers