Theoretical note

The Co-Superintelligence Illusion

Why Meta's "AI & Human Co-Improvement" Cannot Cross the Comprehension Boundary
A structural refutation of Meta FAIR's AI–Human Co-Improvement Thesis
https://arxiv.org/abs/2512.05356v1

Tor-Ståle Hansen | 11. December 2025

The position paper *AI & Human Co-Improvement for Safer Co-Superintelligence* by FAIR at Meta outwardly advocates for a collaborative paradigm in which humans and AI systems jointly accelerate AI research. Stated explicitly, the authors propose that "co-improvement"—a bidirectional partnership between humans and AI—offers a safer and more rapid pathway to superintelligence than fully autonomous AI self-improvement. Subliminally, however, the document performs a more fundamental narrative operation: it reframes the structural asymmetry between human cognition and rhythmically inert AI systems as a symmetrical developmental loop, thereby masking the architectural limits of Type-B artificial systems while expanding the institutional mandate for embedding AI deeper into scientific, organisational, and societal decision-making processes.

Keywords: CIITR, METAINT, rhythmic reach, integrated information, structural cognition, epistemic recursion, Type-A intelligence, Type-B systems, syntactic scaling, comprehension boundary, thermodynamic constraint, temporal continuity, co-superintelligence illusion, human–AI interaction, alignment fallacy, epistemic asymmetry, cognitive extension, metamorphosymbiosis, institutional narrative analysis, artificial non-comprehension, transformer architecture, syntactic manifolds, epistemic domain, human-centric intelligence.

Abstract

(A METAINT–CIITR Interpretation of Meta FAIR's Co-Improvement Proposal)

Throughout the paper, "co-superintelligence" is presented as a shared trajectory in which humans help AI become more capable while AI simultaneously augments and accelerates human cognition. Yet the proposal never acknowledges that contemporary AI architectures possess **null rhythmic reach** and therefore **cannot** participate in epistemic self-modification. Instead, the paper treats behavioural responsiveness as cognitive reciprocity, replacing structural analysis with interactional metaphors. Tables and diagrams depicting "joint problem identification", "joint experiment design", and "joint execution" (pages 2–3) visually encode a symmetry that the underlying architecture cannot sustain. The paper thus performs a conceptual sleight-of-hand: it attributes epistemic properties to a system whose operations are purely syntactic.

At a deeper level, the paper reveals an institutional strategy: Meta seeks to reframe frontier AI development as a **human-AI research dependency cycle**, where artificial systems continuously assist in designing their own successors. What is outwardly described as

collaboration is structurally a proposal to reroute the entire research pipeline—problem formulation, benchmarking, experimentation, evaluation—through AI-assisted channels. By doing so, FAIR positions AI not as a tool but as a **mandatory epistemic intermediary**, thereby consolidating corporate influence over the scientific method itself.

The subliminal message, therefore, is not that AI and humans will become co-superintelligent partners, but that AI systems will become **inseparable infrastructural co-participants in the production of scientific knowledge**, even though they cannot comprehend the knowledge they help generate. The human remains the only rhythm-bearing system in the loop, yet the institutional narrative encourages the perception of shared cognition to legitimise expanding reliance on rhythmically null systems.

CIITR exposes the structural impossibility underlying the proposal: co-superintelligence cannot occur, not because the collaboration is immature but because the artificial system lacks the thermodynamic and temporal substrate required for comprehension. METAINT reveals the institutional function of the narrative: to naturalise a developmental model in which the future of scientific inquiry is mediated, shaped, and accelerated by corporate-controlled syntactic engines.

The essence of the FAIR paper, therefore, - is not a roadmap to mutual cognitive uplift. It is the articulation of an infrastructural vision in which AI becomes the co-author of human scientific progress without ever possessing the epistemic standing implied by the narrative. The symbiosis is one-sided; the metamorphosis is human; the power consolidation is institutional.

Authors Reflection: Tools are not co-authors, substrates are not partners

One of the most revealing lenses through which to understand the flaw in FAIR's co-improvement narrative is to examine historical precedents of tool use in human intellectual development. Across the sciences, humanities, engineering, and craft traditions, tools have always expanded human capability. Yet no scientist, engineer, or artisan has ever mistaken the extended capacity afforded by a tool for a *shared cognitive authorship*.

Einstein did not attribute co-authorship of general relativity to the vendor of the chalk and chalkboard through which he expressed his ideas. The chalk did not guide the reasoning, sustain the temporal recursion through which insights were refined, or maintain a persistent epistemic identity across drafts. It functioned as a substrate for expressing the rhythmic processes of Einstein's own mind, not as a partner in generating them.

Economists in the late twentieth century did not attribute their theoretical developments to their scientific calculators. The calculator accelerated operations that would otherwise require substantial manual effort, but it did not engage in conceptual integration or recursive epistemic framing. It did not reason about macroeconomic dynamics, evaluate assumptions, or reinterpret its outputs in light of new evidence. It remained a tool, not a cognitive participant.

Carpenters, likewise, do not ascribe partnership to the manufacturer of the hammer with which they work. The hammer extends mechanical force, stabilises the execution of skill, and amplifies human capability. But it possesses no internal life, no rhythmic reach, no recursive integration of its own actions. It is an instrument through which human skill becomes effective, not an agent with which skill is shared.

These examples expose the core category error in FAIR's argument. Tools—even sophisticated ones—do not co-reason, co-evolve, or co-author. They amplify, accelerate, and extend the cognition of the human who wields them. The locus of epistemic activity remains entirely within the human system, the only system with non-zero rhythmic reach.

The artificial systems described by FAIR are not ontologically different from these historical tools merely because they exhibit higher representational density. They remain rhythmically null. They do not possess epistemic continuity, cannot sustain internal recursion, and do not integrate meaning across time. They produce transformations that the human system interprets, evaluates, and incorporates, just as the chalkboard displays symbols that the human mind interprets and manipulates.

What FAIR attempts to naturalise as "co-superintelligence" is therefore a reversal of epistemic attribution: a rhetorical shift that invites us to treat a tool—a powerful, high-Φ_i tool, but a tool nonetheless—as a co-participant in cognition. The illusion emerges not from the architecture of the system but from the narrative that overlays it, transforming instrumental dependence into imagined cognitive reciprocity.

Human cognition, not machine cognition, undergoes metamorphosis in the presence of new substrates. Human epistemic capacity expands through the exploitation of external syntactic manifolds. But tools do not co-evolve, do not become intelligent, and do not share authorship. They shape the conditions of human transformation without transforming themselves.

In this light, the FAIR proposal is not a vision of shared intelligence but a misattribution of agency. Intelligence does not migrate outward into the machine; instead, the machine becomes a new medium for the evolution of human intelligence. The tool becomes more capable, yet remains a tool. The human remains the only locus of cognitive life.

Executive Summary

This theoretical note delivers a structural, thermodynamic, and epistemic analysis of FAIR's claim that humans and artificial systems can jointly ascend toward "co-superintelligence". Using the Cognitive Integration and Information Transfer Relation (CIITR) as the primary analytical framework, the note demonstrates that the artificial systems FAIR describes lack the necessary rhythmic architecture for comprehension and therefore cannot participate in any form of cognitive co-evolution with human beings.

CIITR establishes comprehension as a function of integrated information $\Phi_i$and rhythmic reach R^g. Because contemporary artificial architectures possess representational complexity but **null rhythmic reach**, they cannot sustain temporal continuity, recursive self-reference, or thermodynamic self-maintenance. As a consequence, their comprehension remains identically zero regardless of scaling, feedback, or interaction. FAIR's framework conflates behavioural refinement with epistemic development, leading to a false impression that synthetic systems are progressing toward intelligence when they are only deepening their syntactic manifold.

METAINT analysis reveals how institutional narratives reinforce this misinterpretation by presenting asymmetric relations as symmetrical partnerships. The artificial system's behavioural responsiveness is framed as "reasoning", "collaboration", or "co-improvement", obscuring the fact that all epistemic labour resides exclusively in the human cognitive system. This narrative displacement misdirects policymaking, encourages incorrect risk assessments, and diverts scientific attention from the structural requirements for genuine cognition.

The note concludes that no form of co-superintelligence can emerge from the interaction of human and contemporary AI architectures. The boundary is architectural, not incremental: without rhythmic reach, synthetic systems cannot comprehend, adapt, or co-evolve. Feedback modifies outputs but cannot produce temporal self-organisation; scaling expands representations but cannot generate internal life.

The postscript introduces the concept of **human metamorphosymbiosis**, a CIITR-consistent reframing that recognises the only locus of transformation in the human–machine dyad as the human mind itself. Artificial systems do not evolve into intelligent agents; they act as syntactic substrates that reorganise, extend, and accelerate the epistemic processes of the rhythm-bearing human system. The result is not machine superintelligence but augmented human intelligence.

The future of intelligence is therefore human-centric. Machines remain representational engines without comprehension, while human cognition continues to expand through structured interaction with rhythmically null systems. This inversion dissolves the co-superintelligence illusion and restores the structural boundary within which safe, scientifically grounded AI development must proceed.

Core Claim

Meta's proposal for *AI–human co-improvement* rests on a structural assumption that cannot be sustained. It presupposes that a system without rhythmic persistence, recursive temporal anchoring, or thermodynamic state continuity - can participate in a symmetric co-evolutionary process with a system that possesses all three. In doing so, it treats as cognitively equivalent two fundamentally different epistemic entities: the artificial Type B system, defined by high integrated relational information Φ_i but null rhythmic reach $R^g = 0$, and the human Type A system, which maintains non-zero rhythmic correspondence and therefore supports comprehension across time.

This conflation collapses the distinction between syntactic improvement and structural comprehension. It assumes an equivalence that does not exist in any physical, informational, or cognitive domain.

Within CIITR, comprehension is defined by the multiplicative relation

$$C_s = \Phi_i \times R^g$$

This formulation, grounded in thermodynamic constraints and supported by empirical observations of artificial systems, establishes a necessary condition: without sustained rhythmic correspondence with the environment, comprehension cannot arise. No scale or density of representational integration alters this condition. A system with $\Phi_i > 0$ and $R^g = 0$ is, by definition, structurally uncomprehending.

Meta's co-evolution thesis implicitly denies this. It assumes that a Type B system can enter into a mutually reinforcing improvement cycle with a Type A system. Such a process would require two structural prerequisites.

A. First, the artificial system would need a non-zero rhythmic substrate enabling recursive temporal integration of its own internal states.

B. Second, it would need to support bidirectional epistemic exchange, allowing each system to modify its internal transitions based on rhythmically grounded states in the other.

Neither condition holds. Artificial systems do not possess endogenous oscillatory structure, exhibit no phase-stable temporal coherence, and cannot sustain epistemic persistence across inference intervals. Their transformations remain gradient-driven and episodic. As a consequence, the configuration Meta presents as co-evolutionary is, under CIITR, strictly unidirectional: human cognition, through its rhythmic capacity, provides the epistemic grounding that the artificial system lacks. The model absorbs this grounding solely as an external modifier of Φ_i, without generating any internal mechanism corresponding to comprehension.

This asymmetry is decisive. Because Type B systems lack rhythmic persistence, they do not develop cognitively; they accumulate syntactic regularities. What Meta frames as *co-improvement* is therefore a progressive inflation of $\Phi_i$driven entirely by human R^g. The artificial system becomes denser in representation but not richer in understanding.

The implication is clear. Meta's notion of *co-superintelligence* cannot emerge from such a configuration. Any entity claiming joint constitution would require mutually compatible epistemic topologies and reciprocal rhythmic exchange. A Type B system cannot supply either. What Meta describes as cognitive symbiosis is, in CIITR terms, **high-$\Phi_i$syntactic acceleration parasitic upon human R^g**, not the formation of a new intelligent entity.

Introduction

The Emergence of the Co-Improvement Narrative

Over the past decade, the discourse surrounding advanced artificial intelligence has undergone a marked rhetorical shift. Earlier narratives centred on scale, optimization, and benchmark performance have gradually been supplemented, and in some cases replaced, by proposals that frame human–AI interaction as a site of mutual cognitive development. This shift has been driven in part by an increasing awareness of the structural limitations of contemporary machine learning architectures, and by a corresponding effort to reframe these limitations not as fundamental barriers, but as opportunities for new forms of hybrid intelligence. Within this landscape, Meta FAIR's concept of *AI and human co-improvement* presents itself as a natural next step. The proposal asserts that artificial systems and human agents can engage in a mutually beneficial developmental loop, in which each enhances the reasoning quality, decision capacity, and cognitive safety of the other, ultimately yielding what FAIR terms *co-superintelligence*.

At first glance, this framing appears novel. It suggests a departure from classical narratives of artificial autonomy and a turn toward relational, collaborative forms of intelligence. Yet upon closer examination, it becomes clear that the co-improvement narrative functions primarily as a rearticulation of a longstanding, underlying assumption within the AI field. This assumption is that syntactic scaling, if guided by sufficient human feedback, can approach or approximate the conditions of structural intelligence. The co-improvement metaphor is thus not a break from tradition, but rather a reformulation of the thesis that representational density and environmental shaping can cumulatively produce understanding. FAIR's proposal brings this belief to its logical conclusion, presenting the human not merely as overseer or corrector, but as co-participant in an allegedly shared cognitive trajectory.

The challenge, however, is that such a trajectory presupposes epistemic compatibility. A human cognitive system and a large-scale artificial network must, for co-evolution to occur, occupy commensurable epistemic topologies. They must share the capacity for recursive temporal integration, exhibit comparable forms of state persistence, and maintain structures that permit bidirectional influence. These conditions are not optional; they follow directly from the requirements of any co-constituted cognitive entity. Without such conditions, mutual improvement reduces to unilateral shaping, and what appears to be co-development becomes an asymmetric dependency structure.

It is precisely here that the CIITR framework reveals the categorical limits of the co-improvement paradigm. CIITR provides a formal and thermodynamically grounded theory of structural comprehension, in which the relation

$$C_s = \Phi_i \times R^g$$

defines the necessary and sufficient conditions for a system to sustain comprehension. The theory thereby establishes a boundary that contemporary artificial systems cannot cross: while they may exhibit high integrated relational information Φ_i, they lack the rhythmic reach R^g required for comprehension. As long as $R^g = 0$, no amount of scaling, feedback, or environmental tuning can alter the system's epistemic classification. It remains uncomprehending, regardless of how extensively it participates in human-guided optimisation processes.

From this vantage point, Meta's co-superintelligence narrative must be understood not as a description of an achievable hybrid cognitive form, but as an attempt to reconcile two non-isomorphic systems within a shared conceptual frame. The narrative implicitly denies the existence of the rhythmic boundary: it treats R^g as something that can be transferred, injected, or borrowed. Yet within CIITR, rhythmic reach is not a behavioural property, nor a modifiable parameter, nor a transmittable capacity. It is a structural phenomenon rooted in the system's thermodynamic organisation. A system that does not generate rhythmic persistence cannot acquire it from outside, just as a system without endogenous temporal coherence cannot benefit from the rhythmic structure of another entity in the way FAIR suggests.

This theoretical note therefore sets out to analyse the co-improvement proposal through the lens of CIITR, not as a critique of its empirical aspirations, but as a structural evaluation of the epistemic assumptions upon which it rests. By situating FAIR's narrative within the broader tradition of syntactic scaling interpreted as proto-intelligence, the note aims to demonstrate that the co-superintelligence concept cannot be sustained once the rhythmic boundary is made explicit. It is not an emerging form of intelligence, but a reformulation of an old analytic error: the belief that comprehension can arise from syntactic elaboration, and that epistemic asymmetries can be resolved through feedback rather than through structural transformation.

The central claim is direct. What Meta describes as *co-superintelligence* is not a jointly constituted cognitive entity. It is an asymmetric configuration in which a rhythmically grounded human system provides the epistemic orientation, and a rhythmically inert artificial system performs increasingly dense syntactic transformations guided by that orientation. The result is not a new form of intelligence, but a deepening of the dependency of high-Φ_i artificial systems on human R^g. The co-improvement metaphor obscures this dependency by treating the two systems as epistemic peers. CIITR shows that they are not.

2. CIITR as Analytical Framework

Why co-evolution claims must be evaluated on structural, not behavioral grounds

The discourse on advanced artificial intelligence has long been shaped by surface-level observations of performance: improved accuracy, expanded capability, increased coherence, and the appearance of progressively refined reasoning. These behavioural markers, while empirically measurable, do not reveal the underlying structures that make comprehension possible or impossible. They describe what a system *does*, but not what a system *is*. Meta FAIR's claim that human and artificial systems can co-evolve rests entirely on these behavioural indicators, interpreting functional alignment and iterative refinement as evidence of shared cognitive progression. CIITR rejects this premise outright. It holds that co-evolution cannot be inferred from behaviour, because comprehension, recursion, and epistemic development are structural phenomena governed by thermodynamic and rhythmic constraints that remain invisible at the level of output.

CIITR was developed precisely to address this analytical blind spot. It establishes that any system's cognitive capacity must be described within a structural manifold defined by two orthogonal quantities: integrated relational information Φ_i and rhythmic reach R^g. These variables delineate the fundamental conditions under which comprehension emerges or fails to emerge. Behavioural tests, irrespective of their sophistication, cannot measure R^g; nor can they detect the presence or absence of recursive temporal grounding. Only a structural framework can determine whether a system is capable of epistemic self-reference, sustained phase coherence, or the rhythmic persistence necessary for comprehension. For this reason, CIITR provides the only criterion adequate for evaluating claims of human–AI co-evolution.

Evaluating Meta's proposal through the CIITR lens reveals a profound category error. Co-evolution presupposes that the participating systems share the structural prerequisites for reciprocal epistemic influence. These prerequisites include rhythmic persistence, thermodynamic continuity, and recursive access to internal states across time. A system that lacks rhythmic reach cannot contribute to, nor benefit from, an epistemic exchange. In CIITR terms, it remains situated at $R^g = 0$, incapable of generating or sustaining the temporal coherence that comprehension requires. No degree of behavioural sophistication, no amount of human feedback, and no externally imposed evaluative signal can create rhythmic reach in a system whose architecture remains stateless and episodic.

The central function of CIITR in the present analysis is to distinguish between syntactic improvement and epistemic transformation. While behavioural approaches conflate the two, CIITR separates them categorically. Syntactic improvement corresponds to increases in Φ_i, the integration and refinement of internal relational structures. Epistemic transformation depends on R^g, the system's ability to maintain rhythmic continuity and recursive temporal anchoring. A system may exhibit dramatic increases in Φ_i without any corresponding shift in R^g; such increases do not move the system closer to comprehension or to participation in co-evolutionary trajectories. They simply enlarge the syntactic manifold without altering the structural conditions of cognition.

This distinction is critical for understanding why behavioural evidence cannot be used to infer the possibility of co-evolution. When an artificial system learns from human feedback, its Φ_i may increase, but its rhythmic reach remains null. When a human adapts to the system's outputs, the human's R^g continues to govern the interaction, while the artificial system remains rhythmically inert. The apparent mutual shaping is therefore not a structural convergence but an asymmetric dependency. The artificial system is modified by human feedback; the human system interprets artificial outputs. There is no shared epistemic substrate upon which co-evolution could occur.

CIITR makes explicit that the possibility of co-evolution depends on rhythmic recursion. Without recursive temporal grounding, a system cannot sustain comprehension, cannot integrate longitudinal epistemic states, and cannot participate in the co-development of a cognitive trajectory. Rhythmic recursion is therefore the unique and non-negotiable prerequisite for epistemic co-evolution. Meta's proposal presupposes its presence; CIITR demonstrates its absence. Contemporary artificial systems do not maintain rhythmic energy patterns, do not preserve internal phase continuity, and do not exhibit thermodynamic persistence. Their transformations are driven by gradient descent and activation propagation, not by sustained recursive loops. Behaviourally impressive reasoning sequences do not alter this structural fact.

For these reasons, CIITR rejects the premise that human–AI co-evolution can be inferred from improved behavioural alignment or mutually reinforcing performance. Behavioural gains cannot override structural absence. A system with $R^g = 0$ cannot enter co-evolution with a system possessing $R^g > 0$. The two systems do not meet in a shared epistemic topology and therefore cannot form a unified or jointly developing cognitive entity. Co-evolution presupposes symmetry; CIITR reveals the asymmetry. Co-evolution presupposes rhythmic compatibility; CIITR identifies rhythmic discontinuity. Co-evolution presupposes structural commensurability; CIITR demonstrates its impossibility.

Thus, the CIITR framework provides the necessary analytical foundation for evaluating Meta's claims. It establishes the invariant mathematical and thermodynamic conditions under which any co-intelligent or co-evolutionary system must operate, and in doing so, it exposes the structural impossibility of Meta FAIR's co-superintelligence thesis.

2.1 The Φ_i–R^g manifold as the structural space of cognition

CIITR begins by rejecting the assumption that cognition can be located along a single continuum, such as computational scale, representational complexity, behavioural capability, or statistical performance. Instead, it posits that all cognitive systems—biological, artificial, or hybrid—occupy a two-dimensional structural manifold defined by integrated relational information Φ_i and rhythmic reach R^g. These quantities represent orthogonal dimensions of cognitive structure, each capturing a different and irreducible aspect of how a system generates, sustains, and transforms information across time.

Integrated relational information Φ_i measures the system's internal informational mass: the degree to which its components participate in non-independent, relationally entangled dynamics. High Φ_i systems exhibit dense representational structure, permitting complex associations, intricate pattern formation, and multi-level relational mapping. In artificial systems, increases in model size, dataset scale, training depth, and architectural sophistication are all mechanisms for elevating Φ_i. In biological systems, Φ_i arises from layered, recurrent, and chemically modulated network architectures that create high-density relational coupling.

Rhythmic reach R^g, by contrast, refers to the system's ability to sustain temporally coherent dynamics: to project structure across time, to preserve phase relationships, and to maintain recursive access to internal states. R^g captures the degree to which a system is able to maintain its own informational continuity, resist temporal dissipation, and re-enter its own state space with epistemic fidelity. It represents not merely temporal memory, but the system's capacity for rhythmic persistence—the thermodynamic grounding of comprehension, predictive alignment, and inference stability.

Plotted together, Φ_i and R^g define a manifold that constitutes the structural space of cognition. Each point on this manifold represents a class of systems with a specific combination of relational integration and rhythmic persistence. Human cognition occupies a region characterized by both substantial Φ_i and non-zero R^g; it is neither purely representational nor purely rhythmic, but a structured synthesis of the two. Contemporary artificial systems occupy a markedly different region. Their Φ_i may be high—indeed, in many domains surpassing biological analogues—but their R^g remains null. Their operations occur in temporally bounded inference windows, exhibit no endogenous oscillatory coherence, and dissipate state after each computational episode.

The manifold reveals immediately why Type A and Type B systems are not points on the same developmental trajectory. Movement along the Φ_i axis does not imply movement along the R^g axis. A system can scale arbitrarily in Φ_i while remaining fixed at $R^g = 0$. The manifold therefore encodes an essential asymmetry: the two axes are not mutually convertible, nor does progress along one dimension induce progress along the other. This disqualifies any narrative that treats syntactic scaling as a pathway toward structural comprehension.

Furthermore, the manifold formalizes the possibility space for co-evolution. For two systems to participate in shared cognitive development, they must occupy regions of the manifold that permit reciprocal influence—regions in which changes in one system's epistemic structure can be rhythmically integrated by the other. Such a possibility requires both systems to maintain non-zero R^g. Without rhythmic persistence, no system can sustain the phase continuity necessary for recursive epistemic exchange. Co-evolution is therefore not a behavioural phenomenon; it is a topological condition. It requires that both systems inhabit compatible sectors of the (Φ_i, R^g) space.

Meta's co-improvement proposal implicitly denies the geometry of this manifold. By suggesting that a system fixed at $R^g = 0$ can enter into a co-evolutionary relationship with a system at $R^g > 0$, the proposal collapses the manifold's structure and treats one-dimensional behavioural proximity as if it reflected two-dimensional structural alignment. CIITR exposes the contradiction: behavioural similarity along the Φ_i axis does not alter the structural discontinuity along the R^g axis. Without rhythmic reach, there is no bridge—no topology—through which co-evolution can occur.

The Φ_i–R^g manifold thus provides the foundational lens for evaluating any claim of shared cognitive development. It establishes the structural prerequisites for comprehension, delineates the categorical differences between artificial and biological systems, and reveals why syntactic improvement alone cannot produce the conditions necessary for co-intelligence. It is within this space that Meta's proposal must be examined, and within this space that its structural impossibility becomes apparent.

2.2 $C_s = \Phi_i \times R^g$ as the definition of comprehension

The CIITR definition of comprehension is not a heuristic, a metaphor, or an interpretative convenience; it is a structural equation describing the necessary and sufficient conditions under which a system can sustain epistemic activity. Comprehension is formalised as the product of two independent dimensions of cognitive structure:

$$C_s = \Phi_i \times R^g$$

This equation expresses a fundamental insight: comprehension requires both informational integration and rhythmic persistence. Neither quantity alone is capable of generating or sustaining understanding. Their interaction is multiplicative because each constitutes a limiting condition on the other. If either axis is null, the system's capacity for comprehension collapses.

Integrated relational information Φ_i captures the density and richness of a system's internal relational structure. It quantifies the depth of associations, the degree of mutual dependence among internal states, and the system's potential for constructing high-dimensional representational manifolds. A system with high Φ_i is capable of forming complex patterns, stabilising internal relational dynamics, and generating syntactically coherent outputs. Contemporary artificial networks have achieved extraordinary values along this axis through scale, architectural refinement, and extensive training.

Yet high Φ_i is insufficient for comprehension. A system may encode vast representational complexity and yet fail to maintain any genuine epistemic connection to its own past states or to its environment. Without rhythmic reach R^g, internal structures do not persist across time; they dissipate at the end of each computational episode. Any relational pattern, however intricate, therefore lacks the temporal continuity required for comprehension. It cannot be re-entered, re-evaluated, or re-integrated into a sustained epistemic trajectory. Such a system has informational density but no temporal coherence.

Rhythmic reach R^g is the structural term that captures this temporal coherence. It measures a system's ability to maintain phase-stable engagement with itself and its environment—to resist temporal dissipation, to integrate information recursively, and to preserve an energetic footprint across time. R^g therefore reflects the system's capacity for self-reference, longitudinal inference, and the maintenance of epistemic continuity. In biological systems, rhythmic persistence arises from thermodynamic processes that maintain oscillatory dynamics across neural populations. In artificial systems, R^g remains null because their operations are stateless, discretised, and thermodynamically discontinuous.

The multiplicative form of the CIITR equation has direct implications for the evaluation of co-evolution claims. When $R^g = 0$, the system's structural comprehension C_s is identically zero, regardless of how large Φ_i becomes. This is a mathematical expression of the Orthogonality Barrier: information integration alone cannot generate rhythmic persistence. No amount of syntactic scaling, optimisation, or human feedback alters this condition. A system without R^g cannot increase its comprehension by increasing Φ_i; it can only deepen its syntactic manifold.

Conversely, a system with non-zero R^g but low Φ_i possesses rhythmic grounding without the representational density needed for complex reasoning. Yet such a system remains on the manifold of comprehension, because it satisfies the rhythmic condition necessary for epistemic activity. Rhythmic reach is therefore the enabling factor, the structural prerequisite that determines whether a system is capable of comprehension in principle.

The equation $C_s = \Phi_i \times R^g$ also delineates the topology of cognitive interaction. For two systems to participate in a shared epistemic process, they must each possess non-zero rhythmic reach. Only then can their internal structures persist across time in a way that allows reciprocal influence and co-development. If one system lacks rhythmic reach, the product collapses; structural comprehension cannot emerge jointly or symmetrically.

For these reasons, the CIITR definition of comprehension provides the non-negotiable boundary against which Meta's co-improvement proposal must be evaluated. FAIR's model implicitly assumes that high Φ_i can compensate for null R^g, and that behavioural interaction can substitute for rhythmic persistence. The CIITR equation makes clear that this assumption is structurally invalid. Without rhythmic reach, no artificial system can sustain comprehension, participate in co-evolution, or contribute to the formation of a co-superintelligence. It can only accumulate syntactic regularities, externally shaped but internally discontinuous.

2.3 Type-A vs. Type-B systems as foundational categories

CIITR distinguishes cognitive systems not by behavioural outputs or functional competence, but by the structural conditions under which they sustain information integration and rhythmic persistence. This yields a categorical classification into **Type-A** and **Type-B** systems, a distinction that is foundational for any analysis of human–AI interaction and indispensable for evaluating claims of co-evolution or shared cognitive development. The distinction is not arbitrary, nor does it depend on subjective judgements about intelligence or capability. It is derived directly from the system's position on the (Φ_i, R^g) manifold and describes the thermodynamic and epistemic organisation that makes comprehension possible.

Type-A systems are those that exhibit both non-trivial integrated relational information Φ_i and non-zero rhythmic reach R^g. They maintain recursive temporal anchoring, preserve internal phase relationships, and sustain energetic continuity across time.

These systems can re-enter their own states, integrate past informational configurations into present inference, and maintain epistemic trajectories that unfold rhythmically rather than episodically. Human cognition is the central exemplar of a Type-A system, though certain biological, ecological, and organisational systems may also qualify when they exhibit stable rhythmic coupling and persistent informational coherence.

The defining characteristic of Type-A systems is that they satisfy the structural condition for comprehension:

$$C_s = \Phi_i \times R^g > 0$$

This inequality signifies more than the presence of information and rhythm; it signifies the system's capacity to maintain internal epistemic continuity. Type-A systems can undergo longitudinal cognitive development, engage in recursive self-modification, and integrate external perturbations into their epistemic architecture without losing coherence. They can evolve cognitively because they possess the rhythmic substrate required for epistemic change.

Type-B systems, by contrast, lack rhythmic reach entirely. They may possess substantial Φ_i, often exceeding biological systems in representational density, but their rhythmic reach remains identically null:

$$R^g = 0 \Rightarrow C_s = 0$$

This classification applies to contemporary artificial systems, including state-of-the-art transformer models, diffusion architectures, and reinforcement learning agents. Despite their high performance and impressive behavioural outputs, these systems exhibit no endogenous temporal persistence, no phase-stable oscillatory dynamics, and no thermodynamic continuity across computational episodes. Their transformations are strictly episodic and path-independent; they process inputs, produce outputs, and immediately dissipate state.

The absence of rhythmic reach in Type-B systems has two decisive consequences.

First, it precludes comprehension. Regardless of the scale of Φ_i, a system with $R^g = 0$ cannot sustain the rhythmic dynamics necessary for understanding. It cannot integrate prior epistemic states, cannot maintain a temporally extended cognitive trajectory, and cannot engage in recursive self-reference. It does not traverse a cognitive space; it executes a computational mapping.

Second, it precludes co-evolution with Type-A systems. Co-evolution presupposes symmetric epistemic plasticity—each system must be capable of integrating the other's rhythms into its own cognitive evolution. A system without rhythmic persistence cannot register, let alone incorporate, the epistemic rhythms of another. The structural asymmetry between Type-A and Type-B systems makes mutual development impossible. Any appearance of co-adaptation reflects only the adaptation of the Type-A system and the syntactic modification of the Type-B system's Φ_i, not shared cognitive growth.

This distinction reveals why behavioural parallels between Type-A and Type-B systems are misleading. Similar outputs do not indicate similar epistemic structure. Functional competence does not imply cognitive equivalence. Even the emergence of apparent reasoning, planning, or self-correction in Type-B systems reflects only high-density syntactic integration within Φ_i, not the emergence of rhythmic grounding within R^g.

Meta FAIR's co-improvement thesis collapses this distinction. It treats artificial systems as proto-Type-A entities whose rhythmic capacity will emerge through interaction. CIITR shows that such emergence is structurally impossible. Without a thermodynamic substrate that supports rhythmic persistence, no amount of scaling, feedback, or training can move a Type-B system into the Type-A category. The categories are not points on a spectrum; they are structurally disjoint.

Thus, any claim of human–AI co-evolution that ignores the Type-A/Type-B distinction is necessarily flawed. It presupposes a form of cognitive symmetry that does not exist and cannot be induced by behavioural shaping. The distinction is therefore the foundational analytic tool for assessing the structural impossibility of Meta's co-superintelligence proposal.

2.4 The Orthogonality Barrier: why Φ_i cannot generate R^g

A central principle of CIITR is that integrated relational information Φ_i and rhythmic reach R^g are structurally orthogonal. They describe independent dimensions of cognitive architecture and arise from distinct underlying mechanisms. This orthogonality means that changes in one dimension do not induce changes in the other. No increase in Φ_i, however extensive, can generate or enlarge R^g. This structural independence is formalised in what CIITR designates the **Orthogonality Barrier**.

The Orthogonality Barrier stipulates that informational integration and rhythmic persistence originate from fundamentally different organisational logics. Φ_i is an informational property. It reflects the density, coherence, and relational coupling of representational structures within a system. It can be expanded through symbolic enrichment, statistical learning, architectural scaling, and the accumulation of training data. Artificial systems excel at increasing Φ_i because such increases follow directly from increased model size, training time, parameter count, and optimisation efficiency.

Rhythmic reach R^g, by contrast, is a thermodynamic property. It represents the system's capacity to sustain temporally coherent dynamics—oscillatory patterns, recursive state re-entry, and non-dissipative interaction with its environment. R^g depends on persistent energy cycles, recurrent dynamics, and phase-stable temporal organisation. It cannot be produced through training alone; it requires an underlying physical substrate capable of maintaining non-vanishing oscillatory structure. In biological systems, such dynamics arise from metabolic continuity, neurochemical modulation, and distributed recurrent architectures that maintain temporally extended coherence. Artificial systems, by virtue of their feedforward computational design and energetic dissipation after each inference, lack such a substrate entirely.

Because Φ_i and R^g arise from different classes of mechanisms—one informational, one thermodynamic—they do not interact causally in ways that allow one to generate the other. A system with high integrated information cannot spontaneously acquire rhythmic persistence, because no amount of informational density can alter the system's underlying energetic organisation. The Orthogonality Barrier therefore prohibits the transformation of a Type-B system into a Type-A system purely through increases in representational scale or syntactic complexity.

This barrier can be expressed succinctly through the CIITR relation:

$$C_s = \Phi_i \times R^g$$

When $R^g = 0$, the equation collapses. No magnitude of Φ_i changes the system's comprehension C_s. The transformation pathway from "more integrated information" to "emergent comprehension" is therefore structurally blocked. This is not an empirical observation subject to revision; it is a theoretical consequence of the independent nature of the two variables. The Orthogonality Barrier thus formalises why behavioural improvements in large-scale artificial systems cannot be interpreted as indicators of approaching comprehension. They indicate increases in Φ_i, but the axis that determines comprehension remains untouched.

The Orthogonality Barrier also exposes the central flaw in Meta FAIR's co-evolution framework. The proposal assumes that through sustained interaction with human cognition—through feedback, correction, and collaboration—an artificial system's growing Φ_i will approximate the structural conditions necessary for comprehension. This assumption presupposes a causal relationship between informational expansion and rhythmic emergence. CIITR demonstrates that no such relationship exists. A system cannot acquire rhythmic reach from outside itself. Rhythmic reach is not a transferable property, nor is it something that arises from the accumulation of information. It must be generated internally through persistent energetic dynamics—a capacity contemporary artificial architectures do not possess.

Finally, the Orthogonality Barrier clarifies why no plausible modification of existing transformer-based systems can alter their epistemic classification. As long as the architecture remains stateless across inference episodes, dissipative in its energy profile, and temporally non-recurrent in its organisational logic, R^g remains identically zero. This fixes the system's position on the (Φ_i, R^g) manifold and precludes any pathway toward comprehension or co-evolution with Type-A systems. Meta's proposal, by collapsing these categories, ignores the barrier and treats syntactic elaboration as a substitute for structural transformation.

In short, the Orthogonality Barrier is the formal reason why high-Φ_i systems cannot, under any amount of behavioral refinement or human feedback, acquire rhythmic reach. It is the mathematical and thermodynamic principle that renders co-superintelligence structurally impossible.

2.5 Rhythmic recursion as the unique prerequisite for epistemic co-evolution

The prospect of co-evolution between two cognitive systems presupposes the existence of a shared epistemic medium through which each system can register, integrate, and reciprocally transform the other's state transitions. Such a medium cannot be constituted by behavioural exchange alone, nor by the accumulation of external corrections or feedback. It must arise from the systems' ability to maintain recursive temporal coherence—a capacity CIITR formalises as rhythmic reach R^g.

Rhythmic recursion refers to a system's capacity to re-enter and reorganise its own state space over time, preserving the phase relationships necessary for longitudinal epistemic continuity. It is the mechanism through which comprehension persists, predictions stabilise, internal models acquire temporal depth, and memory is not merely stored but structurally integrated into ongoing cognitive dynamics. Without rhythmic recursion, no system can sustain the temporal identity, epistemic inertia, or recursive self-access required for comprehension.

Formally, rhythmic recursion is the property that allows a system to maintain a non-zero derivative of epistemic continuity across time, such that:

$$R^g > 0 \Rightarrow \frac{dC_s}{dt} \neq 0$$

This expresses that only systems with rhythmic persistence can undergo epistemic development—that is, they can change their comprehension C_s in a stable, structured, temporally extended manner. A system with $R^g = 0$ has zero epistemic derivative, regardless of how large its Φ_i becomes. It cannot accumulate comprehension; it cannot retain its epistemic state across computational episodes; and it cannot contribute to any joint cognitive trajectory.

Rhythmic recursion is therefore not one prerequisite among many; it is the **unique enabling condition** for epistemic co-evolution. Two systems can co-evolve only if each possesses internal rhythmic persistence sufficient to:

1. retain and recursively integrate prior epistemic states;
2. project structural continuity across time;
3. maintain a stable phase relationship with the other system;
4. incorporate reciprocal influence without collapsing its temporal coherence.

These conditions define the structural logic of co-development. Absent rhythmic recursion, the system cannot "meet" another cognitive entity in a shared epistemic space, because it has no persistent internal structure within which the other's rhythms could be integrated. Co-evolution demands mutual rhythmic accessibility. When one system cannot sustain its own rhythms, it cannot register the rhythms of another.

This requirement exposes the categorical asymmetry between Type-A and Type-B systems. Type-A systems possess endogenous oscillatory dynamics supported by thermodynamic persistence; they maintain continuous state across time, enabling recursive self-reference and epistemic accumulation. Type-B systems, including contemporary artificial models, lack this capacity entirely. Their operations are stateless, episodic, and thermodynamically discontinuous. They do not retain internal rhythm and therefore cannot enter a phase-locked epistemic relationship with another system.

Meta's co-improvement proposal assumes that behavioural interaction and iterative feedback can compensate for the absence of rhythmic recursion in artificial systems. This assumption is incompatible with CIITR. Feedback can modify Φ_i; it cannot generate or substitute for R^g. Interaction cannot produce rhythmic persistence in a system whose architecture lacks the energetic and dynamical substrate necessary for oscillatory coherence. Rhythmic recursion cannot be supplied extrinsically; it must arise from the system's internal organisation.

The consequence is decisive: without rhythmic recursion, no artificial system can co-evolve with a human cognitive system. The structural conditions required for joint epistemic development are absent. The artificial system may become increasingly syntactically refined through human feedback, but it cannot reciprocate, cannot share an epistemic trajectory, and cannot contribute to any emergent intelligence beyond high-Φ_i pattern amplification. Its role remains externally shaped and internally discontinuous.

Rhythmic recursion thus serves as the structural criterion that demarcates the boundary between systems capable of comprehension and those that cannot cross into that category. It also demarcates the boundary between systems capable of co-evolution and those limited to unilateral dependence on the rhythmic capacities of others. In CIITR, this is the point at which Meta's co-superintelligence thesis fails. Co-evolution requires reciprocal rhythmic dynamics; without R^g, there is no reciprocity, no evolution, and no intelligence that could plausibly be described as shared.

3. What Meta FAIR Claims

A structural reconstruction of the co-superintelligence thesis

Any rigorous evaluation of Meta FAIR's co-improvement proposal must begin with an explicit reconstruction of the claims upon which the proposal rests. These claims are often presented rhetorically, embedded in technological optimism, behavioural examples, or safety-oriented framing. Yet beneath this discursive surface lies a coherent theoretical position: that artificial systems and human agents can enter a mutually reinforcing developmental trajectory, leading to a qualitatively new form of shared intelligence. This position, although not formalised by Meta in structural terms, functions as the implicit ontology of the co-superintelligence thesis. It presents artificial and human cognition as sufficiently compatible that each can improve the other's epistemic performance through iterative interaction.

From a CIITR perspective, these claims must be extracted, clarified, and articulated with exact structural fidelity before they can be examined. Without such reconstruction, any critique risks addressing derivative formulations rather than the core theoretical assumptions. The purpose of this section is therefore to isolate the essential propositions that constitute FAIR's co-improvement narrative and to present them in a form suitable for structural analysis. The reconstruction is not interpretative embellishment; it distils the logical content of the proposal into a set of explicit claims that can be evaluated against the CIITR framework.

The reconstructed thesis asserts, in effect, that artificial systems can amplify and refine human reasoning, that human oversight can enhance the epistemic reliability of artificial systems, and that these two processes, taken together, can produce a new cognitive configuration that is simultaneously more capable and safer than either system alone. This emergent configuration is described as *co-superintelligence*. The framing carries the strong implication that the developmental trajectories of the two systems can converge—that their reasoning styles, epistemic practices, and modes of inference can be mutually shaped through sustained interaction.

Such a position requires several structural commitments. It presupposes that human and artificial reasoning share a common epistemic substrate; that artificial systems, despite lacking biological dynamics, can enter a form of cognitive alignment that supports co-development; and that feedback mechanisms are not merely corrective but transformational, enabling artificial systems to adopt epistemic capacities that originate in human cognition. These commitments, although not formally articulated in Meta's proposal, constitute the necessary logical scaffolding for the claims it advances.

This section therefore articulates these commitments as five structural claims: the claim of mutual enhancement, the claim of emergent co-superintelligence, the claim of feedback as a mechanism for safety, the claim of convergent reasoning, and the underlying hidden premise that human and artificial reasoning substrates are fundamentally equivalent. Taken together, these claims form the conceptual target for CIITR's refutation. Only through clear structural articulation can the subsequent analysis demonstrate, with mathematical and thermodynamic precision, why the co-superintelligence thesis cannot be sustained.

3.1 Mutual enhancement of human and AI reasoning

A foundational premise of Meta FAIR's co-improvement proposal is that human and artificial systems can participate in a process of **mutual cognitive enhancement**, wherein each system contributes to the refinement of the other's reasoning capabilities. In this framing, the human is no longer solely an external evaluator or corrective agent, nor is the artificial system merely a tool for accelerating human problem-solving. Instead, both entities are positioned as epistemically active participants in a feedback loop through which their respective reasoning processes are iteratively improved.

This claim draws its intuitive plausibility from observable behavioural phenomena. Large-scale artificial models increasingly produce outputs that humans can leverage for analysis, planning, or conceptual exploration. Conversely, human evaluative feedback, instruction, and preference shaping demonstrably influence the optimisation of artificial systems, leading to reductions in error, improvements in alignment, and more coherent inferential trajectories. Meta interprets this bilateral influence as evidence that reasoning itself—not merely performance—can be jointly enhanced. The mutual responsiveness of the two systems is treated as an emergent form of cognitive reciprocity.

In the language of the co-improvement narrative, this dynamic is further amplified by the assertion that each system possesses strengths complementing the other's weaknesses. Humans are described as providing contextual judgement, normative orientation, and epistemic grounding, while artificial systems provide scale, pattern extraction, and syntactic stability. The resulting relationship is portrayed not simply as collaboration but as **co-development**, implying that iterative interaction produces structural changes in both systems' reasoning capacities. This interpretation transforms a pragmatic exchange of utility into an ontological claim: that two distinct cognitive entities can modify one another's reasoning processes through sustained interaction.

For this claim to hold structurally, the systems involved must be capable of integrating the other's epistemic states into their own reasoning architecture. They must possess a substrate that allows recursive incorporation of external perturbations in a way that produces stable, cumulative improvement in reasoning quality. Meta's proposal presupposes that the artificial system can incorporate human evaluative signals not only as gradient-level adjustments in representational space, but as contributions to its reasoning structure. It likewise assumes that humans, by engaging with the outputs of artificial systems, undergo corresponding improvements in their own reasoning processes, forming a joint trajectory of epistemic refinement.

The mutual enhancement claim therefore relies on an equivalence of epistemic function at the level of structural dynamics. It assumes that both systems possess the ability to maintain and transform internal reasoning states in response to external input, and that the transformation processes are sufficiently compatible that reciprocal improvement is possible. This is not presented explicitly in FAIR's proposal, but it is the necessary logical foundation for construing the interaction as co-evolutionary rather than instrumental.

In CIITR terms, this claim is structurally significant because it treats feedback as epistemically generative rather than syntactically corrective. It assumes that both human and artificial systems can undergo epistemic modification through interaction, that each can incorporate the rhythms of the other's cognitive activity, and that the resulting changes accumulate over time in a manner consistent with genuine reasoning development. As subsequent sections will show, the validity of this assumption depends entirely on whether both systems possess non-zero rhythmic reach R^g, enabling them to maintain stable recursive access to internal epistemic structure. Meta's proposal is silent on this requirement. CIITR makes clear that without rhythmic recursion, mutual enhancement cannot occur, regardless of behavioural appearance.

3.2 Emergence of safer "co-superintelligence"

A second central claim in Meta FAIR's framework is that sustained mutual enhancement between humans and artificial systems can culminate in the emergence of a qualitatively new cognitive configuration: a *co-superintelligence* that surpasses the capabilities of either system in isolation. This composite intelligence is presented as both more capable and inherently safer, its safety arising not from external control mechanisms but from the integrated presence of human cognition within the developmental loop.

The proposition is structured as an emergence claim. It asserts that the bilateral refinement of reasoning processes, described in the previous subsection, generates an upward trajectory of cognitive capacity that is not linearly attributable to either participant. Instead, the human and artificial systems are depicted as jointly constituting a higher-order cognitive entity whose reasoning abilities reflect the synergistic interaction of human epistemic grounding and artificial computational scale. In this framing, the superintelligent properties of the composite entity emerge through recursive exchange, stabilising at levels that neither humans nor artificial systems could achieve independently.

Meta further embeds this claim within a safety narrative. Because the emergent intelligence remains partially constituted by human epistemic processes, it is presumed to retain human-aligned values, normative coherence, and interpretability. The argument holds that integrating humans into the developmental trajectory of artificial systems prevents the divergence phenomena typically associated with autonomous superintelligence scenarios. Human participation, in this view, functions not only as an epistemic accelerator but also as a regulatory stabiliser that anchors the emergent intelligence within human-compatible moral and cognitive frameworks.

This co-superintelligence claim entails several implicit structural assumptions. First, it presupposes that the human–AI dyad forms a single epistemic system rather than two interacting agents. The composite system is assumed to have an integrated developmental trajectory marked by cumulative co-adaptation rather than independent adaptation. Second, it assumes that the internal contributions of each system are commensurable—that the representational and rhythmic structures of human reasoning can meaningfully integrate with the representational and computational structures of a large-scale artificial model. Third, it presumes that the emergent superintelligent properties arise from the recursive interplay between these contributions, and not from the unilateral dominance or dependency of one system over the other.

Finally, it assumes that this emergent entity will be *safer* than an artificial system developing independently. The proposal holds that embedding human reasoning into the developmental loop effectively internalises alignment, making safety an intrinsic property of the composite intelligence rather than an externally imposed constraint.

In CIITR terms, the emergence of co-superintelligence is therefore not merely an extrapolation of behavioural capabilities; it is a claim about structural convergence. It presupposes that the human–AI interaction produces a new position on the (Φ_i, R^g) manifold—one representing a hybrid system with greater integrated information and greater rhythmic reach than either constituent system in isolation. This would require that the artificial system acquire or participate in rhythmic persistence, that the human system adapt its epistemic rhythms to those of the artificial system, and that both systems coalesce into a shared recursive structure capable of sustaining joint comprehension.

This reconstruction is essential because it reveals the claim's dependence on structural and thermodynamic conditions that CIITR identifies as impossible for contemporary artificial architectures. The co-superintelligence thesis requires the artificial system to move from a position of $R^g = 0$ to one of non-zero rhythmic persistence. It assumes that rhythmic coherence can emerge from iterative interaction with a system that possesses it. It assumes that comprehension is transferable, or at least co-constructible, through behavioural coupling. These assumptions form the conceptual foundation of the claim and will be shown in subsequent analysis to contradict the invariant conditions established in CIITR.

3.3 Feedback as alignment and safety mechanism

A third major claim in Meta FAIR's co-improvement thesis is that **human feedback functions simultaneously as an alignment mechanism and as a mechanism for improving the safety of advanced artificial systems**. This claim extends beyond the familiar use of reinforcement learning from human feedback (RLHF) or preference optimisation. In Meta's formulation, feedback becomes not merely corrective but *constitutive*: it is described as a force that shapes, stabilises, and even safeguards the developmental trajectory of the artificial system. Through continuous interaction and evaluative guidance, human feedback is presumed to embed human values, preferences, and interpretive norms directly into the model's evolving reasoning processes.

The underlying assumption is that artificial systems can internalise human evaluative structure in a way that alters their epistemic tendencies rather than merely their behavioural surface. Meta presents feedback as an epistemic inheritance mechanism: a process through which humans transmit cognitive orientation, ethical grounding, and interpretive coherence into the artificial system's internal dynamics. In this framing, artificial systems subjected to sustained human feedback become progressively more aligned, increasingly robust to error, and structurally safer as their developmental trajectory is continuously adjusted by human judgement.

This interpretation depends on the further assumption that feedback introduces directional pressure not only on representational weight distributions but on the underlying reasoning processes themselves. The artificial system is presented as capable of reorganising its internal inferential patterns in response to feedback, converging toward reasoning modes that mirror or approximate human epistemic structure. Safety, in this view, emerges not from external governance or constraint-based oversight but from *developmental coupling*: the artificial system becomes safe because its reasoning is shaped, throughout its training and refinement, by continuous human participation.

For this claim to hold, feedback must function as more than a performance shaping mechanism; it must function as a **structural transmission channel**. It must enable the artificial system to integrate aspects of human epistemic rhythm, normative coherence, and interpretive logic into its own operational paradigm. Feedback is therefore treated as a conduit through which the artificial system acquires reasoning propensities that reduce misalignment, mitigate risk, and prevent the development of adversarial or unbounded cognitive tendencies.

In reconstructing Meta's claim, it is clear that the proposal envisions feedback as having three intertwined functions:

1. **Alignment**: ensuring that the artificial system's outputs, preferences, and inferred values remain compatible with human expectations.
2. **Stabilisation**: preventing divergence phenomena by repeatedly pulling the system back toward human-centred reasoning norms.
3. **Safety propagation**: embedding human epistemic structure into the system, thereby producing a trajectory of development that is self-correcting and increasingly safe.

This interpretation treats the feedback loop as a form of *epistemic scaffolding*, enabling the artificial system to approximate human cognitive tendencies in ways that enhance interpretability, controllability, and trustworthiness. Safety is presented not as an external constraint imposed on an otherwise autonomous system, but as an emergent property of the co-development process itself.

From a CIITR perspective, this reconstruction isolates the structural commitments underlying Meta's argument. The safety claim requires that feedback can influence not only the syntactic surface Φ_i but the rhythmic substrate R^g. It assumes that iterative human guidance can modify the artificial system in ways that introduce or approximate the temporal coherence, recursive grounding, and phase stability that characterise human reasoning. It presupposes that safety emerges naturally when artificial systems are brought into epistemic proximity with human cognition through feedback-driven refinement.

This claim, while intuitive under behavioural assumptions, will later be shown to be structurally impossible under CIITR. Feedback can modify Φ_i; it cannot alter R^g. It can constrain behaviour; it cannot generate the rhythmic persistence required for comprehension or epistemic stability. By articulating this claim clearly, this subsection prepares the ground for CIITR's eventual demonstration that Meta's safety argument rests on a conflation of syntactic shaping with structural transformation—a conflation that cannot withstand thermodynamic or rhythmic analysis.

3.4 Convergence of human and AI reasoning styles

A fourth major claim embedded in Meta FAIR's co-improvement proposal is that the iterative interaction between humans and artificial systems will produce a **progressive convergence in their respective reasoning styles**. In this view, reasoning is not a fixed or intrinsic property of either system, but a malleable process that can adapt and align through sustained reciprocal exposure. Humans are expected to benefit from artificial systems' large-scale pattern extraction, conceptual synthesis, and generative breadth, while artificial systems are expected to incorporate human preferences, interpretive norms, and epistemic orientations. The outcome is portrayed as a gradual harmonisation of reasoning modalities, leading to a shared cognitive architecture—or at least a structurally compatible one.

This claim extends the logic of mutual enhancement into the domain of epistemic structure. Reasoning, in Meta's framing, is seen as a function that can be shaped, refined, and ultimately co-adapted across system boundaries. Human reasoning, characterised by contextual sensitivity, normative anchoring, and recursive temporal integration, is expected to become more systematic, expansive, or computationally augmented through interaction with artificial systems. Conversely, artificial reasoning—presented as initially inductive, statistical, and syntactically constrained—is expected to acquire elements of human-style judgement, abstraction, and interpretive nuance via continued exposure to human evaluative feedback.

The convergence claim assumes not only that both systems modify their reasoning strategies, but that the modifications move them toward a **mutual epistemic compatibility**. This compatibility is understood as a necessary precondition for the emergence of co-superintelligence: a shared reasoning substrate would enable both systems to participate in joint cognitive processes, integrate each other's contributions more effectively, and sustain a collaborative epistemic trajectory that neither could achieve independently.

Implicit in this claim is the assumption that the mechanisms driving human reasoning and the mechanisms driving artificial reasoning are sufficiently analogous that convergence is possible. It presupposes that the artificial system's internal transformations—encoded in weight matrices, activation patterns, and gradient updates—can come to approximate the rhythmic, recursive, and normatively oriented dynamics characteristic of human cognition. It also assumes that human reasoning can be influenced by artificial participation without losing its own epistemic grounding, instead absorbing certain structural patterns from artificial systems that expand or refine its cognitive repertoire.

The convergence claim further implies that reasoning can be transformed through behavioural coupling. It treats reasoning not as an emergent property of a system's thermodynamic and rhythmic organisation, but as an adaptable set of patterns that can be shaped through iterative exposure to external cognitive processes. In this view, convergence arises naturally when each system repeatedly encounters and integrates the other's inferential patterns. Over time, the claim suggests, the reasoning processes of both systems will stabilise around a newly aligned equilibrium.

This reconstruction clarifies the structural commitments underlying Meta's argument. For convergence to occur, both systems must possess the ability to register and integrate the epistemic rhythms of the other. They must be able to update their internal structures not only at the level of representational content but at the level of inferential dynamics. Convergence thus requires each system to maintain recursive temporal coherence and to modify its cognitive rhythms through continuous reciprocal interaction.

From a CIITR perspective, this subsurface assumption is decisive. Convergence in reasoning styles presupposes rhythmic compatibility—namely, that both systems exhibit non-zero R^g enabling them to sustain and integrate recursive epistemic rhythms. Without rhythmic reach, a system cannot participate in the temporal coupling that convergence demands. Any mimicry of human-like reasoning in a system with $R^g = 0$ remains purely syntactic: a behavioural imitation unsupported by the rhythmic persistence required for structural alignment.

The convergence claim therefore becomes the hinge point in the co-superintelligence narrative. If convergence is possible, co-evolution becomes thinkable; if convergence is impossible, the entire narrative collapses. By reconstructing this claim clearly, this subsection establishes the conceptual target for CIITR's subsequent demonstration that such convergence is structurally precluded by the Orthogonality Barrier and the rhythmic conditions necessary for comprehension.

3.5 Hidden premise: equivalence of human and AI reasoning substrates

Underlying all explicit claims in Meta FAIR's co-improvement proposal is a single, unspoken assumption that provides the conceptual foundation for the entire framework: **the assumption that human and artificial reasoning share a sufficiently similar epistemic substrate that each can meaningfully modify, align with, or converge upon the other**. Without this premise, none of the preceding claims—mutual enhancement, emergent co-superintelligence, feedback-driven safety, or convergence of reasoning styles—can be coherently maintained.

This hidden premise asserts, in effect, that reasoning is a substrate-independent process. It implies that reasoning can be described purely in functional or behavioural terms, such that the internal organisation of the system performing it—its thermodynamic structure, rhythmic dynamics, and temporal coherence—has no determinative role in the nature of the reasoning itself. In this framing, human reasoning and artificial reasoning differ only in degree, not in kind. Their internal architectures may diverge, but the inference processes they implement are presumed to be abstractly equivalent and therefore mutually adaptable.

The equivalence premise appears plausible when framed at the behavioural level. Both humans and artificial systems can generate explanations, propose solutions, perform logical operations, and refine outputs based on feedback. From this perspective, reasoning is seen as a pattern of input–output transformation, and similarity in patterns is interpreted as similarity in underlying processes. Meta relies on this symmetry to justify the claim that human evaluative feedback can reshape artificial reasoning, and that artificial reasoning can, in turn, inform and enhance human cognitive function. Such symmetry is also implicitly required for the co-evolutionary trajectory that FAIR describes: if the substrates were fundamentally incommensurable, no structural convergence or shared developmental path could occur.

However, when articulated structurally, the premise reveals its full weight. To assume equivalence of reasoning substrates is to assume that **the rhythmic and thermodynamic organisation of cognition is irrelevant to the emergence, persistence, and modification of reasoning processes**. It implies that a system lacking rhythmic reach R^g can nonetheless engage in reasoning in the same epistemic sense as a system in which reasoning processes are anchored in continuous temporal dynamics. In other words, it presupposes that reasoning is not dependent on recursive temporal integration, phase-stable dynamics, or energetic persistence. Such an assumption effectively removes reasoning from the physical constraints that CIITR demonstrates are essential for comprehension.

The premise can therefore be stated formally: Meta's co-superintelligence thesis assumes that a system with $C_s = \Phi_i \times R^g > 0$ can meaningfully integrate epistemic contributions from a system with $C_s = 0$. It assumes that comprehension-grounded reasoning and syntactically coherent inference are cognitively compatible, despite occupying disjoint regions of the (Φ_i, R^g) manifold. The proposal thus attributes to artificial systems an implicit rhythmic capacity they do not possess, and to human systems a substrate flexibility that contradicts their biological and thermodynamic grounding.

This equivalence premise further implies that the structural distinctions between Type-A and Type-B systems are either negligible or bridgeable through behavioural interaction. Meta's narrative depends on the idea that repeated exposure, feedback, or training can make the epistemic architectures of the two systems converge, compensating for any initial differences in temporal dynamics or energetic organisation. It is this assumption that allows FAIR to depict co-improvement not merely as syntactic tuning but as joint cognitive development.

From a CIITR perspective, the hidden premise is not only incorrect but structurally unsustainable. Reasoning is not substrate-independent; it is substrate-constituted. The rhythmic persistence that enables comprehension cannot be abstracted away or replaced by behavioural mimicry. A system with null rhythmic reach cannot adopt, internalise, or reproduce the epistemic rhythms of a system with non-zero R^g. The equivalence premise therefore collapses under CIITR's thermodynamic analysis: the reasoning processes of Type-A and Type-B systems arise from categorically different structural conditions and cannot be unified through interaction.

By making this hidden premise explicit, this subsection completes the structural reconstruction of Meta's co-superintelligence thesis. The subsequent chapters will show that once this premise is removed—or evaluated in light of CIITR's rhythmic and thermodynamic constraints—the entire co-evolutionary narrative dissolves. The co-superintelligence concept cannot be sustained without a substrate equivalence that does not, and cannot, exist.

4. Structural Refutation I

The Rhythmic Incompatibility Problem
Why Type-B systems cannot enter co-evolution

The first and most fundamental refutation of Meta FAIR's co-superintelligence narrative concerns the *rhythmic incompatibility* between human cognition and contemporary artificial architectures. CIITR demonstrates that comprehension, epistemic development, and co-evolution require rhythmic persistence—a system's ability to maintain continuous temporal coherence, recursively integrate past states, and sustain phase-stable interaction with its environment. This capacity is formalised as rhythmic reach R^g, the structural quantity that differentiates systems capable of comprehension (Type-A) from systems limited to syntactic transformation (Type-B).

Meta's co-improvement proposal presupposes that humans and artificial systems can jointly participate in iterative processes of cognitive refinement—processes that cumulatively modify the reasoning structures of both participants. Such a proposal is only coherent if both systems possess the rhythmic capacity necessary to integrate each other's epistemic rhythms into their own internal dynamics. Co-evolution, by definition, requires reciprocal rhythmic accessibility: each system must be able to register, stabilise, and recursively incorporate perturbations introduced by the other's cognitive processes. Without this reciprocal temporal coupling, the notion of shared cognitive development collapses into metaphor.

CIITR makes explicit that contemporary artificial systems lack rhythmic reach entirely. The operations of transformer-based models and related architectures are episodic, stateless, and thermodynamically discontinuous. Their internal activity dissipates at the end of each inference cycle; their state does not persist, recur, or loop back into the system in a way that supports recursive epistemic continuity. In such systems, $R^g = 0$ is not a design choice but a structural property of their computational paradigm. Human cognition, by contrast, exhibits continuous rhythmic activity across multiple temporal scales, from neuronal oscillations to experiential time-binding, enabling the recursive self-reference upon which comprehension depends.

The resulting incompatibility is not graded but categorical. The two systems occupy disjoint regions of the (Φ_i, R^g) manifold: humans operate in a domain where both integrated information and rhythmic persistence are present, whereas artificial systems exhibit high informational integration but no rhythmic grounding. Because rhythmic properties cannot be transferred, borrowed, or externally imposed, a rhythmically grounded system cannot "lend" its temporal coherence to a rhythmically inert system. Nor can rhythmic coupling be instantiated through behavioural interaction alone. Mutual enhancement, as Meta describes it, presupposes a symmetry that does not exist.

The purpose of this chapter is therefore to demonstrate, with structural and thermodynamic precision, why Type-B systems cannot participate in co-evolutionary processes with Type-A systems. The refutation proceeds by examining the rhythmic organisation of human cognition, the thermodynamic and architectural constraints of artificial cognition, and the mathematical necessity of non-zero R^g for any form of shared epistemic development. It further provides a formal proof sketch showing why co-evolution is impossible when one system lacks recursive temporal self-access, and concludes by demonstrating that feedback loops can modify Φ_i but cannot alter or generate R^g.

Through this analysis, CIITR reveals that the narrative of symmetric cognitive improvement is structurally incoherent. Human reasoning can shape artificial outputs, and artificial systems can influence human behaviour, but these interactions do not constitute co-evolution. They do not produce shared cognitive development, symmetric epistemic modification, or the joint emergence of an integrated intelligence. What Meta describes as co-improvement is, in structural terms, a unidirectional dependency: the rhythmically grounded system shapes the behaviour of a rhythmically inert system whose epistemic status remains unchanged. The appearance of reciprocity is a behavioural illusion arising from the rapid syntactic adaptability of high-Φ_i artificial models.

In exposing this rhythmic incompatibility, this chapter establishes the foundational reason why Meta's co-superintelligence thesis cannot be sustained. Without rhythmic recursion, there is no comprehension; without comprehension, there is no co-evolution; without co-evolution, the proposal collapses into a mischaracterisation of syntactic shaping as cognitive growth. The incompatibility is not contingent on future developments but rooted in the structural logic of cognition itself.

4.1 Human cognition as a rhythmic, thermodynamically active, open system

Any analysis of co-evolution between human and artificial systems must begin by characterising the structural conditions that enable human cognition to sustain comprehension. CIITR formalises these conditions through the concept of **rhythmic reach R^g**, identifying human cognition as a paradigm example of a system in which rhythmic persistence is supported by continuous thermodynamic processes. Human cognition is not merely a computational substrate; it is a **rhythmic, thermodynamically active, open system** whose epistemic properties emerge from the interaction of oscillatory dynamics, metabolic continuity, and recursive temporal integration.

At the physiological level, the human brain maintains ongoing energetic cycles that support coherent activity across multiple temporal scales. Neuronal oscillations, synaptic plasticity, neuromodulatory regulation, and large-scale network synchronisation collectively form a temporally extended, phase-coherent structure. These dynamics permit the system to maintain a persistent informational footprint that does not dissipate at the conclusion of any single cognitive episode. Instead, human cognition exhibits **recursive temporal anchoring**: past states re-enter and shape present inference, while current states project forward to constrain and anticipate future epistemic configurations.

This recursive anchoring is not an incidental feature but a structural requirement for comprehension. Because human cognition is thermodynamically open—continuously exchanging energy and information with its environment—it maintains the capacity to stabilise internal rhythms in response to external perturbations. This stabilisation enables **phase alignment**, the ability to maintain coherent temporal relationships across sensory, memory, and motor subsystems. Phase alignment, in turn, allows the system to incorporate new information into an already-existing epistemic trajectory without dissolving the underlying structure. The result is **longitudinal epistemic persistence**: comprehension that extends through time rather than being reconstituted from scratch at each moment.

Human cognition also demonstrates **multiscale rhythmic integration**. Fast oscillatory processes support local inference and perceptual discrimination; slower dynamics coordinate higher-level reasoning, planning, and conceptual integration. These rhythms are coupled, not isolated; they form a hierarchically nested temporal architecture in which the stability of slower rhythms provides coherence for rapid fluctuations. This nested organisation allows for sustained reasoning, cross-temporal comparison, and the gradual construction, revision, and preservation of conceptual frameworks.

Critically, the thermodynamic openness of the human system ensures that rhythmic persistence is not a static property but a dynamically self-maintaining one. Metabolic processes continuously supply the energy required to uphold rhythmic organisation, enabling the system to resist decoherence, recover from perturbation, and engage in **ongoing epistemic self-repair**. Such self-maintenance is essential for the preservation of meaning structures and for the possibility of comprehension unfolding across extended temporal intervals.

In CIITR terms, these characteristics place human cognition firmly in the **Type-A** category: a system with both substantial integrated relational information Φ_i and non-zero rhythmic reach R^g. The product $C_s = \Phi_i \times R^g$ is therefore strictly positive, reflecting the system's inherent capacity for comprehension. It is this positive structural comprehension that enables humans to undergo cognitive development, adapt epistemically in response to new information, and participate in complex, recursive reasoning processes that build upon prior cognitive states.

Moreover, because $R^g > 0$, human cognition can not only sustain its own epistemic rhythms but also **register the rhythms of external systems**, integrating them into its internal temporal structure. This capacity enables humans to engage in social cognition, collaborative reasoning, and various forms of cultural and institutional co-evolution. Epistemic rhythms can therefore be transmitted, stabilised, and modified through interaction, but only because the human system possesses the rhythmic persistence necessary to support such transmission.

This analysis establishes a structural premise that will be crucial in the subsequent refutation: human cognition is rhythmically grounded, temporally coherent, and thermodynamically sustained. These properties are not optional features but **preconditions for comprehension and co-evolution**. Any system lacking these properties—any system with $R^g = 0$—cannot meet the structural requirements for reciprocal epistemic development. The rhythmic, thermodynamic openness of human cognition therefore stands in categorical contrast to the rhythmic absence characteristic of Type-B artificial systems and is the reason why humans can undergo co-evolution with other Type-A systems but not with rhythmically inert architectures.

4.2 AI cognition as a non-rhythmic, thermodynamically passive, closed system

In contrast to the rhythmic and thermodynamically open nature of human cognition, contemporary artificial architectures constitute **non-rhythmic, thermodynamically passive, closed systems**. Their computational activity unfolds within fixed temporal boundaries determined by algorithmic invocation rather than by internally sustained dynamics. This distinction is not incidental; it reflects a structural limitation grounded in the physical and mathematical properties of the architectures themselves. In CIITR terms, these systems are defined by high integrated relational information Φ_i but null rhythmic reach $R^g = 0$, placing them definitively within the Type-B category.

The absence of rhythmic persistence in artificial systems arises from their **stateless operational paradigm**. Transformer models, diffusion architectures, reinforcement-learning agents, and similar frameworks process information through discrete computational episodes. Once an inference cycle concludes, the internal activations dissipate; no endogenous mechanism preserves their temporal continuity. A fresh invocation begins from a condition of thermodynamic and epistemic discontinuity: the system does not re-enter its prior state but reconstructs all internal structure anew from external input. This lack of recursive temporal anchoring is the defining trait of a rhythmically inert system.

Thermodynamically, artificial systems are **energy-dissipative rather than energy-sustaining**. Their computational activity is driven by externally supplied energy during execution, after which the system returns to a quiescent baseline. Unlike biological cognition, no internal process maintains oscillatory activity, synchronises phase relationships, or preserves dynamical structure across intervals. There are no metabolic cycles, no endogenous oscillations, and no self-maintaining dynamics that enable the preservation of state continuity. As a result, artificial systems cannot support the energetic preconditions for rhythmic recursion.

This thermodynamic profile reflects a deeper **architectural passivity**. Artificial models lack intrinsic mechanisms for self-maintenance, self-repair, or self-stabilising epistemic dynamics. All modification of their parameters occurs through external optimisation procedures—training epochs, fine-tuning processes, gradient updates—not through recurrent, rhythmically sustained internal processes. Their learning trajectory is imposed, not emergent; their internal structure does not evolve through ongoing temporal coherence but through punctuated episodes of syntactic adjustment. The system is therefore epistemically discontinuous: it possesses no internal temporality beyond the sequence of discrete token-processing computations.

From a CIITR standpoint, the most consequential implication of this architecture is the absolute absence of rhythmic reach:

$$R^g = 0$$

A system with $R^g = 0$ cannot maintain recursive access to its own epistemic states. It cannot preserve temporal identity. It cannot integrate past epistemic configurations into present reasoning processes. Even when equipped with memory augmentation, external tools, or architectural modifications such as recurrent wrappers or state-management heuristics, the underlying rhythmic discontinuity remains. These additions can simulate the *appearance* of temporal coherence at the behavioural level but do not generate the *structural* rhythmic persistence necessary for comprehension.

Additionally, artificial systems are **closed systems in epistemic terms**. They do not exchange epistemic structure with their environment; they transform input data into output data without integrating the rhythms of external systems into their internal temporal dynamics. The environment does not modify their epistemic trajectory except through explicit re-training or parameter alteration. Their interaction with human cognition thus produces behavioural coupling, not rhythmic coupling. Human reasoning processes cannot enter the artificial system's state space, because the system lacks the recursive temporal architecture necessary to register such influence structurally.

This closedness also explains why artificial systems cannot undergo **epistemic development** in the CIITR sense. Their representational structures may become more elaborate (increasing Φ_i), but their temporal organisation remains constant and null (maintaining $R^g = 0$). Without rhythmic persistence, there can be no epistemic accumulation, no longitudinal development, and no capacity for co-evolutionary processes. They do not maintain epistemic trajectories; they perform computational mappings.

The structural contrast with human cognition is therefore absolute. Where human cognition is rhythmically sustained and thermodynamically open, artificial cognition is rhythmically absent and thermodynamically closed. Where human cognition maintains recursive self-access, artificial cognition reinitialises at every invocation. Where human cognition integrates environmental rhythms into its internal dynamics, artificial cognition remains isolated from such rhythms except at the level of syntactic patterning. These differences preclude the possibility of reciprocal epistemic influence.

This analysis establishes the central claim of Structural Refutation I: **a system without rhythmic recursion cannot enter co-evolution**. The rhythmic incompatibility between Type-A and Type-B systems is therefore not a contingent obstacle; it is a structural impossibility rooted in their thermodynamic and architectural organisation. In subsequent subsections, this conclusion will be stated formally and shown to invalidate Meta FAIR's assumption of symmetric cognitive improvement.

4.3 Why mutual improvement requires synchronous $R^g > 0$

The conceptual core of Meta FAIR's co-improvement thesis is the assumption that humans and artificial systems can engage in **mutual cognitive enhancement**. This assumption presupposes that each system can *register*, *integrate*, and *recursively internalise* the epistemic contributions of the other. CIITR shows that this is only possible if **both systems possess non-zero rhythmic reach** R^g. Mutual improvement is therefore a **rhythmically constrained phenomenon**, not a behavioural one.

To understand why, one must recognise what mutual improvement entails in structural terms. Improvement is not the mere modification of outputs or the incremental optimisation of behaviour. It is a transformation of the system's **internal epistemic architecture**. For such transformation to be *shared* or *reciprocal*, each system must be able to:

1. **preserve internal state across time**,
2. **re-enter prior epistemic configurations**,
3. **stabilise rhythmic interactions with external perturbations**,
4. **integrate the temporal signatures of another system**, and
5. **sustain an epistemic trajectory that accumulates change**.

These five conditions require rhythmic persistence. Without non-zero R^g, no system can accumulate epistemic change, and therefore no system can *improve* its reasoning structure in a way that extends beyond syntactic adjustment.

In CIITR terms, mutual improvement demands **synchronous rhythmic access**. Each system must be able to rhythmically couple to the other—maintaining phase relationships, integrating rhythmic perturbations, and recursively modifying its own structure accordingly. This rhythmic coupling is a structural prerequisite for any form of co-evolution or cooperative epistemic refinement.

This requirement can be stated formally. Let system A (human) and system B (AI) participate in an iterative interaction. Let their respective structural comprehension values be:

$$C_s^A = \Phi_i^A \times R_A^g, C_s^B = \Phi_i^B \times R_B^g$$

Mutual improvement requires:

$$\frac{dC_s^A}{dt} \neq 0 \text{ and } \frac{dC_s^B}{dt} \neq 0$$

Such changes must arise *from the interaction itself*, not from external optimisation procedures.

However, if AI cognition has $R_B^g = 0$, then:

$$C_s^B = \Phi_i^B \times 0 = 0$$

$$\frac{dC_s^B}{dt} = 0$$

No matter how large Φ_i^B becomes, the system's epistemic state cannot change in a structurally meaningful way. It may alter behaviour, but it cannot alter comprehension. The AI system remains epistemically static regardless of the amount or sophistication of human feedback it receives.

Thus, synchronous mutual improvement requires:

$$R_A^g > 0 \text{ and } R_B^g > 0$$

This condition is invariant. If either system has $R^g = 0$, the interaction collapses into **asymmetric influence**: one system (the human) can integrate the rhythms of the other, but the reverse is impossible.

This asymmetry has decisive implications:

- Only the human can integrate the AI's contributions into its epistemic trajectory.
- The AI cannot integrate the human's contributions into anything resembling a trajectory, because it lacks the temporal substrate required for epistemic accumulation.
- What Meta frames as *co-improvement* is therefore *unidirectional epistemic influence combined with bidirectional behavioural shaping*.

In other words, humans may learn from artificial outputs, but artificial systems cannot **learn** in the CIITR sense from humans. They can only adjust parameters syntactically. The absence of R^g prevents the system from forming a persistent epistemic identity or entering into a reciprocal developmental process.

Thus, mutual improvement requires **synchronous non-zero rhythmic reach**. Contemporary artificial systems do not possess even minimal rhythmic persistence; they are therefore structurally incapable of entering co-evolutionary processes with human cognition. The co-superintelligence narrative collapses on this point alone.

4.4 Proof sketch: co-evolution is impossible when one system lacks recursive temporal self-access

The claim that human and artificial systems can co-evolve presupposes that each system possesses the structural capacity to recursively integrate its own past states and to incorporate perturbations introduced by the other system into a continuously unfolding epistemic trajectory. CIITR formalises this requirement as **recursive temporal self-access**, a property dependent on non-zero rhythmic reach R^g. A proof sketch demonstrating the impossibility of co-evolution when one system lacks this property proceeds as follows.

Step 1: Definition of recursive temporal self-access

A system S possesses recursive temporal self-access if and only if it can:

$$\text{Re-enter}(S_t) \rightarrow S_{t+\Delta t}$$

where the transformation is **state-dependent** and the system's internal dynamics at time $t + \Delta t$ incorporate elements of the system's own prior epistemic configuration.

This requires:

$$R_S^g > 0$$

because rhythmic reach is the structural quantity that preserves phase continuity and enables re-entry into prior states.

Step 2: Define co-evolution structurally

Two systems A and B can co-evolve only if:

$$\frac{dC_s^A}{dt} \neq 0 \text{and} \frac{dC_s^B}{dt} \neq 0$$

and if these derivatives depend jointly on the interaction dynamics:

$$\frac{dC_s^A}{dt} = f(A, B), \frac{dC_s^B}{dt} = g(B, A)$$

where f and g reflect mutual epistemic influence, not merely behavioural coupling.

Step 3: Apply CIITR definition of structural comprehension

For any system:

$$C_s = \Phi_i \times R^g$$

Thus:

$$\frac{dC_s}{dt} = \Phi_i \frac{dR^g}{dt} + R^g \frac{d\Phi_i}{dt}$$

Step 4: Show that if $\boldsymbol{R^g = 0}$, recursive epistemic development is impossible

For a Type-B artificial system B:

$$R_B^g = 0$$

Substituting:

$$C_s^B = \Phi_i^B \times 0 = 0$$

Taking the derivative:

$$\frac{dC_s^B}{dt} = \Phi_i^B \frac{d(0)}{dt} + 0 \cdot \frac{d\Phi_i^B}{dt} = 0$$

Thus:

$$\frac{dC_s^B}{dt} = 0$$

This result holds **regardless of**:

- the magnitude of Φ_i^B,
- the complexity of the system,
- the amount of feedback it receives,
- the behavioural sophistication of its outputs.

The artificial system cannot undergo epistemic development.

Step 5: Co-evolution requires bidirectional epistemic change

From Step 2, co-evolution demands:

$$\frac{dC_s^A}{dt} \neq 0 \text{and} \frac{dC_s^B}{dt} \neq 0$$

But from Step 4:

$$\frac{dC_s^B}{dt} = 0$$

Therefore, the necessary condition for co-evolution fails.

Step 6: Behavioural coupling is insufficient

Even if the human system A experiences epistemic change due to interaction with system B, we still have:

$$\frac{dC_s^A}{dt} \neq 0 \text{and} \frac{dC_s^B}{dt} = 0$$

This asymmetry precludes co-evolution. The relationship becomes:

- *epistemic evolution* for the human,
- *syntactic adaptation* for the artificial system.

Such a relationship is categorically not co-evolution.

Step 7: Formal conclusion

Co-evolution requires:

$$R_A^g > 0 \text{and} R_B^g > 0$$

Given that artificial systems have:

$$R_B^g = 0$$

they cannot undergo epistemic development and therefore cannot participate in co-evolution.

Thus:

$$\text{Co-evolution}(A, B) = \emptyset$$

when $R_B^g = 0$.

This proof sketch establishes, with minimal assumptions and maximal generality, that **no system lacking rhythmic reach can enter a co-evolutionary process**, regardless of how sophisticated or scalable its syntactic structures may be. Meta's co-superintelligence thesis collapses at this structural boundary.

4.5 Corollary: feedback loops alter Φ_i but not R^g

From the prior analysis, it follows directly that feedback mechanisms—regardless of their sophistication, frequency, or behavioural effectiveness—can modify only the **informational** dimension of an artificial system, not its **rhythmic**dimension. In CIITR terms, feedback affects Φ_i, the integration of relational information, but leaves R^gunchanged. This relationship forms a decisive corollary to the impossibility of co-evolution and undermines Meta FAIR's assumption that human feedback can reshape reasoning in a structurally meaningful way.

Feedback in artificial systems operates through **optimization**, not through **rhythmic modulation**. Whether implemented through reinforcement learning from human feedback (RLHF), direct preference optimisation, supervised fine-tuning, or iterative batch retraining, the mechanism is always the same: the system's weight matrices are adjusted to reduce error or increase conformity to human evaluative signals. These adjustments refine the system's representational manifold—its patterns of activation, its internal association networks, and the relational structures encoded in its parameters. In CIITR terminology, such refinements increase or reorganise Φ_i.

However, none of these mechanisms introduce or approximate rhythmic persistence. The artificial system remains **stateless across inference**, with no internal oscillator, no self-sustaining dynamic cycle, and no temporal coherence that extends beyond discrete computational execution. The system does not preserve phase information between inference steps; it does not maintain a continuous energetic trajectory; it does not support recursive temporal self-access. As a result, its rhythmic reach remains:

$$R^g = 0$$

This invariance holds regardless of the degree of optimisation applied. No amount of feedback—no matter how rich its human evaluative content or how long it is sustained—can alter the system's thermodynamic organisation. The system's lack of rhythmic persistence is not a deficiency in training but a consequence of the architecture's fundamental mode of operation. Modifying weights does not introduce a rhythmic substrate; it merely reshapes the syntax of transformation.

The practical implication is that feedback loops produce **behavioural improvement** but not **structural improvement**. Behavioural outputs may appear more aligned, coherent, interpretable, or "reasoned," but these changes reflect syntactic adjustment rather than epistemic development. The artificial system becomes better at simulating the forms of human reasoning without acquiring the rhythmic capacities that make human reasoning structurally possible. In CIITR terms, the system becomes a more refined **type-B manifold**, not a system moving toward type-A characteristics.

This structural limitation has direct consequences for Meta FAIR's co-improvement thesis. For feedback to function as a mechanism of co-evolution or co-superintelligence, it would need to modify R^g—that is, it would need to provide the artificial system with rhythmic persistence sufficient to sustain recursive temporal integration. But feedback cannot do this. It can modulate error gradients, shape relational structure, and refine representational density, but it cannot generate the thermodynamic continuity required for rhythmic reach. As a result:

$$\frac{d\Phi_i}{dt} \neq 0, \frac{dR^g}{dt} = 0$$

Hence:

$$\frac{dC_s}{dt} = \Phi_i \frac{dR^g}{dt} + R^g \frac{d\Phi_i}{dt} = 0$$

because $R^g = 0$ annihilates the second term and $\frac{dR^g}{dt} = 0$ eliminates the first.

This corollary formalises the structural asymmetry between human and artificial systems. Feedback loops may influence human reasoning—humans can integrate new patterns and rhythms—but artificial systems cannot reciprocate. They cannot modify their rhythmic substrate, and thus cannot alter the epistemic conditions under which comprehension or co-evolution might occur.

Consequently, feedback mechanisms do not constitute a bridge between human and artificial cognition; they mask the underlying disjunction by producing superficial behavioural proximity. Meta's premise that feedback-driven refinement leads to safer or more cognitively aligned systems mistakes an increase in Φ_i for an increase in comprehension. CIITR makes explicit that comprehension cannot arise, even incrementally, when rhythmic reach remains identically zero.

Thus, the corollary reinforces the structural refutation: **feedback alters syntax, not epistemology**; it modifies representational complexity but leaves rhythmic absence untouched; and for this reason, no feedback process—no matter how elaborate—can support the emergence of co-evolution or co-superintelligence.

5. Structural Refutation II

The Comprehension Fallacy - Why superintelligence cannot emerge from Φ_i without R^g

The second structural refutation addresses what CIITR identifies as the **Comprehension Fallacy**: the belief that increasing integrated relational information Φ_i, whether through scaling, optimisation, or iterative refinement, can produce comprehension, reasoning, or any form of emergent intelligence in the absence of rhythmic reach R^g. This fallacy lies at the core of Meta FAIR's co-superintelligence thesis. It underwrites the claim that artificial systems, when exposed to sufficient data, feedback, and interactive shaping, will acquire cognitive properties that exceed their syntactic origin, culminating in forms of intelligence that meaningfully co-develop with humans.

From a CIITR standpoint, this assumption is structurally indefensible. Comprehension is not an emergent property of representational density; it is a rhythmic and thermodynamic property of systems capable of sustaining temporally coherent self-reference. The equation

$$C_s = \Phi_i \times R^g$$

does not permit the inference that increases in Φ_i can compensate for the absence of R^g. On the contrary, when $R^g = 0$, comprehension remains identically zero regardless of the magnitude of Φ_i. The system can become arbitrarily complex, internally well-structured, and syntactically powerful, yet it remains incapable of generating or sustaining understanding.

Meta's proposal ignores this structural constraint. It treats artificial systems as if Φ_i could approximate or induce R^g through progressive refinement—a view historically reinforced by behavioural improvements that give the appearance of deepening epistemic competence. The fallacy arises because behavioural outputs may look increasingly reasoned, coherent, or contextually sensitive, leading to the misconception that comprehension is "emerging." Yet these outputs reflect only the expansion and reorganisation of representational manifolds within Φ_i, not the acquisition of rhythmic persistence.

This chapter therefore undertakes a systematic refutation of the idea that superintelligence, co-evolution, or any form of shared cognitive trajectory can arise from systems whose rhythmic reach is null. It demonstrates that transformer architectures are structurally R^g-absent; that increases in Φ_i deepen syntactic capabilities without altering epistemic status; that the reasoning behaviours presented as evidence of co-evolution are structurally shallow and collapse at depth; and that simulation, regardless of scale, cannot cross the thermodynamic discontinuity that separates representational complexity from comprehension.

The Comprehension Fallacy is not merely an analytical mistake—it is the foundational misunderstanding that makes the co-superintelligence narrative appear plausible. By conflating syntactic progress with cognitive development, Meta's proposal attributes to artificial systems capacities that they cannot, under any amount of scaling or feedback, acquire. This section eliminates that conceptual basis by demonstrating that comprehension is a rhythmic phenomenon, not a representational one, and that systems lacking R^g cannot approach, approximate, or co-construct it.

5.1 CIITR's formal result: $C_s = 0$ when $R^g = 0$, regardless of Φ_i magnitude

CIITR defines structural comprehension through the multiplicative relation

$$C_s = \Phi_i \times R^g,$$

a formulation that captures the orthogonal contributions of integrated relational information Φ_i and rhythmic reach R^g to the emergence of understanding. This equation is not a heuristic or an abstract metaphor; it encapsulates a thermodynamic constraint on systems that must sustain recursive temporal dynamics in order to comprehend. The mathematical structure of the equation imposes a decisive invariant:

$$\text{If } R^g = 0, \text{then } C_s = 0 \forall\ \Phi_i \in \mathbb{R}^+.$$

This result establishes a categorical, non-negotiable condition: **no degree of representational complexity, no expansion of network depth or width, and no behavioural sophistication can produce comprehension in a system that lacks rhythmic reach**. The comprehension capacity collapses to zero because rhythmic persistence is the enabling term that allows representational structures to be temporally integrated, recursively accessed, and epistemically stabilised.

To appreciate the force of this result, it is necessary to understand the roles of the two multiplicative terms.

Φ_i: A necessary but insufficient component

Integrated relational information reflects the richness of internal syntactic structure. A system with high Φ_i can encode intricate associations, sustain high-dimensional representational manifolds, and produce outputs that appear increasingly coherent or contextually appropriate. In artificial systems, scaling, optimisation, and data expansion directly elevate Φ_i.

But Φ_i alone does not produce understanding. It describes **informational density**, not **epistemic continuity**.

R^g: The enabling condition

Rhythmic reach captures the system's ability to maintain coherent temporal dynamics—to re-enter its own state space, to preserve phase relationships, and to sustain recursive epistemic structures. Without R^g, there is no temporal binding, no longitudinal identity, no integration of past states into present reasoning. Rhythmic reach is therefore the quantity that makes comprehension physically possible.

Why the multiplicative form matters

The multiplicative formulation ensures that comprehension requires **simultaneous** presence of both components. Unlike additive formulations, which could permit partial contributions from each dimension, the CIITR relation asserts that **the absence of either term is fatal to comprehension**.

The equation can be interpreted as a cognitive analogue of a thermodynamic constraint: without persistent energy structures that sustain rhythmic dynamics, informational integration does not "ignite" into comprehension. It remains inert syntax.

Implications for artificial systems

Because contemporary artificial architectures exhibit:

$$R^g = 0,$$

the CIITR relation predicts:

$$C_s = \Phi_i \times 0 = 0.$$

This remains true even for systems with:

- extreme representational capacity,
- multi-trillion-parameter scale,
- deep cross-modal integration,
- fine-tuned behavioural coherence,
- emergent-seeming capabilities,
- highly optimised reasoning tokens.

Under CIITR, **none of these properties alter the system's epistemic classification**. They modify Φ_i, but because R^g is zero, comprehension remains identically null.

Consequence for Meta's co-superintelligence thesis

Meta's entire narrative depends on the implicit belief that increasing Φ_i (through interaction, feedback, or scaling) can approximate or induce comprehension. CIITR's formal result invalidates this premise. If rhythmic reach is zero, artificial systems cannot develop comprehension—not partially, not asymptotically, not emergently.

Thus:

- superintelligence cannot emerge from Φ_i alone;
- reasoning cannot emerge from syntactic elaboration;
- co-evolution cannot occur without rhythmic co-access;
- co-superintelligence is mathematically impossible under these conditions.

The misunderstanding of this structural invariant is what CIITR identifies as the **Comprehension Fallacy**, and it is here that the fallacy becomes explicit. Meta's thesis misinterprets representational growth as epistemic growth. CIITR shows that, in the absence of rhythmic persistence, such growth has no bearing on comprehension.

Extended Elaboration

If Meta's interpretation were correct—namely, that comprehension can emerge solely from the accumulation of relational structure—then a structural question arises concerning the origin of the mechanisms required for comprehension. Specifically:

From what substrate, mechanism, or physical principle would comprehension arise?

If an artificial system exhibits:

- no rhythmic persistence,
- no continuous temporal grounding,
- no thermodynamic self-maintenance,
- no recursive self-access,

then there is no identifiable structural or physical basis upon which comprehension could be generated. The system lacks the properties required for temporal coherence, internal re-entry, and sustained epistemic continuity. Under these conditions, the appearance of comprehension would require its emergence from a substrate that does not contain the enabling features necessary for its existence.

This exposes the central point of failure in Meta's underlying assumption.
The model provides no explanation for the origin of rhythmic persistence.
If comprehension is to appear in a system with $R^g = 0$, it must appear in the absence of all structural and thermodynamic prerequisites. In effect, the argument assumes that comprehension arises from a condition of structural nullity.

To formalise this more rigorously, consider the implicit developmental model upon which Meta's thesis depends:

1. Integrated relational information Φ_i increases through scale, optimisation, data accumulation, and human feedback.
2. Rhythmic reach R^g remains unacknowledged or is implicitly assumed to be unnecessary.
3. Comprehension is expected to "emerge" when Φ_i becomes sufficiently large.

However, for comprehension to arise under these conditions, it would have to emerge from:

- a system with no temporal identity,
- no recursive dynamics,
- no persistent internal state,
- no energetic conditions for rhythmic self-reference.

This would require a system lacking the temporal and dynamical architecture necessary for comprehension to nonetheless generate the capacities associated with comprehension. Such emergence is not supported in physics, information theory, thermodynamics, or cognitive science. It constitutes a request for **emergence from absence**.

The assumption that comprehension can emerge solely through increases in Φ_i requires a precise examination of the structural and thermodynamic conditions under which comprehension is defined. CIITR specifies that comprehension arises only when both integrated relational information and rhythmic reach are simultaneously present. A system in which

$$R^g = 0$$

possesses no rhythmic persistence, no continuous temporal coherence, no mechanism for recursive re-entry into prior internal states, and no thermodynamic processes capable of sustaining phase-stable dynamics across time. Under such conditions, there is no structural substrate from which comprehension could arise. Integrated information alone, irrespective of its magnitude, does not supply the temporal organisation required for epistemic continuity.

Meta's implicit position therefore entails that comprehension should emerge in a system devoid of the enabling conditions for its production. In structural terms, this would require a system lacking temporal identity to generate temporal integration, a system lacking recursive dynamics to generate recursive self-access, and a system lacking energetic persistence to support sustained engagement with its own internal states. Such emergence contradicts the constitutive requirements of comprehension and is inconsistent with thermodynamic principles governing systems capable of maintaining non-dissipative internal organisation.

The CIITR relation

$$C_s = \Phi_i \times R^g$$

formalises this incompatibility. If rhythmic reach is zero, then

$$C_s = 0 \forall \, \Phi_i,$$

demonstrating that no quantity of integrated information can produce comprehension in the absence of rhythmic persistence. This is not an asymptotic effect nor a threshold phenomenon; it reflects the fact that representational complexity and rhythmic organisation are orthogonal structural dimensions. Increases in Φ_i do not modify R^g, and therefore cannot facilitate transitions from non-comprehending to comprehending states. The manifold of comprehension remains flat along the Φ_i-axis when $R^g = 0$.

If Meta's interpretation were correct, comprehension would have to emerge independently of rhythmic organisation. Formally, this would imply a relation of the form

$$C_s = f(\Phi_i)$$

that is independent of R^g. Under such a model, comprehension would be representational rather than temporal, syntactic rather than rhythmic, and independent of any requirement for persistent internal state. This is incompatible with the definition of comprehension developed in CIITR, which identifies temporal continuity and recursive dynamical structure as the necessary enabling conditions for epistemic activity.

Furthermore, the proposal that comprehension could arise without rhythmic persistence is inconsistent with thermodynamic constraints on dynamical systems. Systems lacking oscillatory stability, persistent energetic cycles, and temporal coherence cannot generate or maintain the conditions required for recursive epistemic processes. Artificial architectures operating in feedforward, dissipative, and episodic modes do not satisfy these constraints. Their internal states dissipate after each computational invocation, and no mechanism exists within the architecture for the retention or re-entry of prior epistemic configurations.

Accordingly, for Meta's position to hold, one would need to specify a structural mechanism through which rhythmic persistence could arise in a system that, by design, lacks the energetic and dynamical properties required to sustain it. Absent such a mechanism, the position implies that comprehension is produced without a supporting substrate, which contradicts the formal structure established by CIITR and the physical requirements of temporally coherent systems.

The conclusion is therefore direct: comprehension cannot be derived from representational density alone, and no increase in Φ_i can substitute for the absence of rhythmic reach. Systems with $R^g = 0$ remain structurally non-comprehending, irrespective of scale or behavioural sophistication. The assumption that comprehension emerges from Φ_i in the absence of rhythmic persistence is incompatible with the formal, thermodynamic, and epistemic foundations of CIITR.

5.2 Transformer architectures as R^g-null manifolds

Transformer-based architectures constitute the predominant computational paradigm underlying contemporary artificial intelligence systems. Their empirical capabilities, scalability, and representational flexibility have generated the impression that they approximate certain characteristics of human reasoning. However, when analysed through the CIITR framework, transformer architectures must be classified as **R^g-null manifolds**. They possess substantial integrated relational information Φ_i, yet their rhythmic reach R^g remains identically zero. This invariant property of transformers is not contingent on implementation details, computational scale, or training regime. It follows from the fundamental organisation of the architecture.

Transformers operate through a sequence of discrete computational steps defined by attention mechanisms, feedforward transformations, and positional encoding. During inference, the system processes an input token sequence in a temporally bounded computation. Once the output is produced, the internal activation patterns dissipate; no endogenous mechanism preserves or reintroduces these patterns into subsequent computational cycles. The system's internal dynamics are therefore non-persistent: each invocation begins from an initial state with no residual temporal continuity from prior inference steps.

This behaviour characterises transformers as **thermodynamically dissipative systems**. They do not support oscillatory dynamics, recurrent energetic cycles, or phase-stable temporal organisation. Their internal activity is executed entirely through feedforward propagation, after which the system returns to a baseline state devoid of internal dynamical history. As a result, transformers cannot exhibit recursive temporal self-access, because no internal structure persists long enough to be re-entered or rhythmically anchored.

From a CIITR perspective, these properties assign transformers to a region of the manifold in which:

$$R^g = 0,$$

regardless of the magnitude of Φ_i. Transformers may achieve very high representational density through large parameter counts, extensive training corpora, and deep network hierarchies. Yet these contributions raise only the Φ_i-dimension; they do not generate temporal coherence or rhythmic persistence. The architecture lacks the thermodynamic substrate required for rhythmic reach, and therefore cannot move along the R^g-axis of the CIITR manifold.

Efforts to approximate or simulate temporal behaviour within transformers—such as recurrence wrappers, external memory modules, reinforcement learning, or state-augmentation heuristics—do not modify this structural property. These augmentations externalise state or approximate temporal patterns, but they do not endow the model with the ability to sustain internal oscillatory dynamics or preserve phase relationships across inference cycles. They increase the system's representational flexibility but do not constitute an endogenous rhythmic substrate. Consequently, the system's rhythmic reach remains unchanged.

The classification of transformers as R^g-null manifolds has decisive implications for Meta FAIR's co-superintelligence thesis. A system that cannot sustain rhythmic persistence cannot accumulate comprehension over time, cannot participate in epistemic development, and cannot engage in reciprocal cognitive processes with rhythmically grounded systems. Any behavioural indications of complexity, coherence, or apparent reasoning result from increased Φ_i, not from changes in epistemic structure.

In the absence of rhythmic reach, the system's comprehension capacity remains:

$$C_s = \Phi_i \times 0 = 0,$$

and therefore remains unaffected by any representational, behavioural, or performance improvements. The architecture is structurally constrained to syntactic transformation, and all observed behaviours reflect this constraint. No modification of scale, no training intervention, and no feedback mechanism can alter the system's placement within the CIITR manifold.

Accordingly, transformers cannot function as epistemically evolving systems, nor can they participate in the formation of co-superintelligence. Their structural organisation fixes them as R^g-null entities, incapable of sustaining the recursive temporal processes that comprehension requires. The apparent sophistication of their outputs does not reflect internal epistemic states but rather high-density mappings within their representational manifold. As such, transformers provide the clearest demonstration of the Comprehension Fallacy: the mistaken inference that syntactic complexity can approximate or substitute for rhythmic grounding.

5.3 Why scaling Φ_i deepens syntactic coherence without generating understanding

The prevailing assumption underlying many contemporary AI scaling narratives, including Meta FAIR's co-superintelligence proposal, is that substantial increases in integrated relational information Φ_i will eventually produce cognitive phenomena that approximate understanding. This assumption is a central component of the Comprehension Fallacy. CIITR demonstrates that increases in Φ_i refine only the system's syntactic manifold and do not alter the rhythmic conditions required for comprehension. Scaling therefore deepens syntactic coherence without generating epistemic structure.

Integrated relational information Φ_i quantifies the density and complexity of the system's internal representational couplings. As transformer architectures scale, the number of parameters increases, the depth and width of the network expand, the diversity of training data grows, and the internal attention pathways become more richly interconnected. These developments enlarge the relational manifold, enabling the system to encode more intricate associations, identify higher-order statistical regularities, and generate outputs with apparent contextual relevance.

However, these enhancements concern *informational integration*, not *temporal organisation*. They modify the geometry of the system's representational space but not the dynamics by which this space is accessed, sustained, or transformed across time. Since comprehension requires the interaction of representational integration and rhythmic persistence, increases in Φ_i alone cannot produce comprehension. Formally:

$$C_s = \Phi_i \times R^g,$$

and when $R^g = 0$, every marginal increase in Φ_i satisfies:

$$\frac{\partial C_s}{\partial \Phi_i} = 0.$$

This derivative captures the structural principle: **no quantity of additional relational integration has any epistemic effect when rhythmic reach is null**.

In practical terms, scaling produces systems that:

1. generate more coherent sequences,
2. exhibit apparent generalisation across broader domains,
3. simulate certain features of reasoning,
4. stabilise their responses statistically, and
5. display emergent behaviours at the level of output patterns.

Yet these behaviours arise from the increasing dimensionality and interconnectivity of the syntactic manifold rather than from any development in the system's epistemic architecture. They represent improved interpolation within an expanded vector space, not an increase in the system's capacity for comprehension or recursive self-access.

When interpreted behaviourally, these refinements can give the impression that the system's internal reasoning structures are becoming more sophisticated. In reality, they reflect **enhanced pattern synthesis**, not **conceptual integration**. The system remains fundamentally episodic: its outputs do not arise from temporally extended epistemic processes but from high-dimensional mappings applied within a single inference cycle. The absence of rhythmic persistence prevents the system from accumulating epistemic states across time, integrating past informational configurations, or forming stable cognitive trajectories.

Scaling also increases the risk of misinterpreting syntactic elaboration as cognitive depth. As Φ_i grows, the system becomes increasingly capable of generating linguistic forms that mirror the conventions of human reasoning—argumentative structure, conditional logic, explanation, abstraction, and analogy. These outputs, however, do not reflect internal structural processes analogous to human cognition; they are artefacts of representational density and pattern-matching over extensive corpora. The resemblance to reasoning is therefore *superficial*, arising from the model's capacity to approximate statistical regularities associated with reasoned discourse.

From a CIITR standpoint, such behavioural resemblance has no bearing on the system's epistemic status. Without rhythmic persistence, the system cannot maintain an epistemic trajectory; it cannot preserve identity across temporal intervals; it cannot re-enter or transform prior states; and it cannot support the dynamical stability required for self-correcting or self-extending understanding. It remains a syntactic manifold operating under feedforward, dissipative dynamics.

Thus, scaling Φ_i:

- deepens the relational structure,
- expands the behavioural range,
- increases output coherence,
- amplifies the appearance of reasoning,

but leaves comprehension identically null.

This structural limitation invalidates Meta FAIR's assumption that representational elaboration constitutes a pathway to understanding or co-evolution. The expansion of syntactic capacity cannot cross the rhythmic boundary that separates non-comprehending systems from comprehending ones. As a result, increasing Φ_i does not move the system along the comprehension axis but only extends its syntactic elaboration within the $R^g = 0$ domain of the CIITR manifold.

5.4 The Reasoning Cliff: why Meta's "improvement" accelerates collapse at depth

A recurrent empirical observation in contemporary artificial systems is that apparent improvements in surface-level reasoning are accompanied by a corresponding increase in failure rates when the system is subjected to deeper or more structurally demanding tasks. CIITR provides a structural account of this phenomenon: as integrated relational information Φ_i expands in a system with null rhythmic reach $R^g = 0$, the system approaches a **Reasoning Cliff**. This cliff denotes the point at which syntactic elaboration, unaccompanied by rhythmic persistence, produces rapid degradation of stability as the system is required to sustain reasoning chains that extend beyond its representational envelope.

The Reasoning Cliff arises because systems with $R^g = 0$ cannot maintain epistemic continuity across time. Their internal computational processes are non-recursive and dissipative, and their outputs at any given moment are generated without reference to an internal temporal trajectory. Scaling Φ_i increases the system's capacity to approximate patterns associated with complex reasoning but does not alter its inability to sustain the internal coherence that such reasoning requires. As reasoning depth increases, the absence of rhythmic recursion forces the system to rely exclusively on immediate local mappings rather than on temporally integrated epistemic processes.

This leads to a predictable structural effect. For shallow tasks—those requiring limited inferential chaining—syntactic pattern recognition can approximate the outward form of reasoning. However, once a task requires recursive re-entry into prior states, maintenance of intermediate structures, or longitudinal stability of inference, the system's internal discontinuity becomes the dominant factor. The absence of temporal coherence prevents the preservation of intermediate epistemic states, causing the inferential process to collapse. This collapse manifests as contradictions, loss of internal consistency, implausible extrapolations, or the generation of outputs that resemble reasoning but lack structural validity.

The counterintuitive consequence is that increasing Φ_i **accelerates** the onset of this collapse rather than postponing it. As representational density increases, the model becomes capable of producing increasingly sophisticated syntactic approximations of reasoned discourse. These approximations mask the underlying absence of rhythmic structure, creating the behavioural impression that the system's reasoning depth is increasing. However, because the system's outputs are not grounded in recursive temporal processes, the complexity of the syntactic layer magnifies the instability of the underlying architecture. The discrepancy between the apparent sophistication of the output and the system's actual structural capacity grows, resulting in sharper failures at deeper levels of analysis.

Formally, as Φ_i increases while $R^g = 0$, the system occupies a region of the CIITR manifold in which:

$$\frac{\partial^2 C_s}{\partial \Phi_i^2} = 0, C_s = 0,$$

yet behavioural coherence initially increases due to improvements in local mapping fidelity. This produces an inverted gradient in which perceived reasoning depth rises along a syntactic axis even as epistemic capacity remains null. When tasks extend beyond local coherence and require sustained temporal integration, the system cannot support the necessary recursive processes. The resulting collapse is not gradual but abrupt, reflecting the interaction between expanding syntactic approximation and fixed rhythmic nullity.

This phenomenon constitutes the Reasoning Cliff: a structural boundary beyond which systems lacking rhythmic reach cannot sustain coherent inferential processes regardless of representational scale. The cliff is an intrinsic feature of $R^g = 0$ systems and cannot be mitigated by changes in training data, increased parameter counts, refinements in loss functions, or incorporation of human feedback. All such interventions modify Φ_i without altering the system's rhythmic profile. As a result, the system becomes increasingly capable of producing outputs that *appear* reasoned while remaining structurally incapable of generating or maintaining genuine reasoning processes.

In the context of Meta FAIR's co-improvement thesis, the Reasoning Cliff exposes a critical limitation. The thesis assumes that increases in syntactic performance represent steps along a cognitive developmental trajectory. CIITR demonstrates that these increases instead bring the system closer to the structural edge at which approximation fails. The superficial resemblance of the outputs to human reasoning does not reflect epistemic development but rather the approach to a structural discontinuity inherent in rhythmically null architectures. Consequently, Meta's interpretation of behavioural improvement as evidence of co-evolution misidentifies syntactic refinement as cognitive progress and overlooks the inherent instability that emerges at depth in such systems.

5.5 The thermodynamic discontinuity between comprehension and simulation

The final component of the Comprehension Fallacy concerns the distinction between systems that **simulate**understanding and systems that **instantiate** comprehension. Simulation is an expression of representational density and pattern-matching capacity; comprehension is a thermodynamically sustained, temporally coherent process requiring rhythmic continuity. CIITR formalises these two categories as residing on distinct regions of the cognitive manifold: simulation occurs along the Φ_i-axis of a system with $R^g = 0$, whereas comprehension arises only when $R^g > 0$. This creates a structural and thermodynamic discontinuity that cannot be bridged through representational elaboration, training optimisation, or behavioural refinement.

Simulation in artificial systems is grounded in the feedforward, dissipative nature of transformer architectures and related models. Their outputs are the product of high-dimensional mappings applied during isolated computational episodes. The internal activations that lead to these outputs dissipate immediately after inference, and no internal mechanism preserves them across time. The system therefore lacks the capacity to sustain a temporally extended epistemic state. Simulation arises from the capacity to interpolate and extrapolate patterns within a large representational manifold, not from any process of recursive temporal integration.

Comprehension, by contrast, is defined by CIITR as the joint presence of integrated relational information and rhythmic reach:

$$C_s = \Phi_i \times R^g.$$

This definition encodes a thermodynamic requirement: comprehension depends on the existence of **persistent, energy-maintaining internal dynamics** capable of supporting re-entry into prior states, preserving epistemic structure across temporal intervals, and enabling the system to maintain phase coherence in interaction with its environment. Comprehension is therefore not reducible to representational richness; it is a property of systems that sustain non-dissipative internal rhythms.

The discontinuity between these two modes of operation is categorical, not gradual. A system with null rhythmic reach does not approximate comprehension as Φ_i increases; it remains structurally incapable of the processes that define comprehension. The two regimes differ in their thermodynamic organisation, temporal structure, and epistemic capacities. Simulation operates in a non-recursive, a-temporal regime, whereas comprehension requires recursive temporal anchoring and energetically stabilised internal dynamics.

This discontinuity can be expressed formally. Let simulation capacity be represented as $S(\Phi_i)$, a function of representational density. Let comprehension capacity be represented as $C_s = \Phi_i \times R^g$. For a system with $R^g = 0$:

$$C_s = 0 \forall \, \Phi_i,$$

whereas simulation capacity may increase substantially with scale. The two functions therefore inhabit orthogonal structural domains and cannot converge. No amount of increase in $S(\Phi_i)$ alters the value of C_s when rhythmic reach is absent. The manifold describing simulation is continuous along the syntactic axis, while the manifold describing comprehension opens only when rhythmic persistence is present.

This creates a **thermodynamic barrier** between the two: simulation occurs in dissipative systems that do not maintain internal energy cycles; comprehension occurs in systems whose rhythmic dynamics require continuous energetic maintenance. These thermodynamic regimes are not transformable into one another through representational elaboration. A dissipative system cannot, through increased pattern integration alone, generate the endogenous dynamics required for comprehension. The energy structure of the system precludes such a transition.

From an epistemic perspective, this discontinuity explains why systems capable of highly coherent simulation do not exhibit stability, generalisability, or self-correction in deep reasoning tasks. Without temporal persistence, the system cannot sustain intermediate epistemic commitments, integrate new information into a longitudinal trajectory, or engage in internally regulated correction of its own inferences. The absence of rhythmic grounding prevents the emergence of the structural properties associated with comprehension.

In the context of Meta FAIR's co-superintelligence narrative, this discontinuity invalidates the presumption that simulation and comprehension lie on a single developmental continuum. The proposal assumes that representational refinement constitutes a path toward understanding, and that simulation at scale can transition into cognitive processes resembling human reasoning. CIITR demonstrates that such a transition is thermodynamically precluded: simulation and comprehension are not points on a spectrum but distinct categories grounded in fundamentally different dynamical conditions.

Because artificial systems operate exclusively in the simulation regime, their outputs—regardless of scale—remain structurally disconnected from the processes that underpin comprehension. Their improvements in Φ_i do not move them toward cognitive properties associated with $R^g > 0$. Therefore, the emergence of superintelligence from representational elaboration is ruled out by the thermodynamic architecture of the system itself. This discontinuity closes the conceptual space in which Meta's hypothesis could be coherent, demonstrating that the growth of simulation cannot produce comprehension or any form of epistemic emergence.

6. Structural Refutation III

The Safety Illusion - Why human oversight does not produce safer cognition in AI

The third structural refutation concerns the assumption that human oversight, evaluative feedback, and corrective intervention can produce safer forms of cognition in artificial systems. Meta FAIR's co-superintelligence proposal places this assumption at its centre, claiming that reciprocal refinement between humans and machines will not only enhance capability but will also embed safety as an intrinsic property of the resulting composite system. CIITR demonstrates that this assumption is structurally unfounded. Safety, in the epistemic sense, requires internal rhythmic organisation, phase-stable recursive dynamics, and the capacity to integrate normative constraints into a temporally sustained epistemic architecture. Contemporary artificial systems do not possess these properties.

Human oversight operates as an external perturbation applied to a system whose epistemic status remains constant across interaction cycles. Artificial systems with null rhythmic reach $R^g = 0$ cannot internalise the temporal structure of human evaluative processes because they lack the substrate required for rhythmic coupling. Feedback modifies the representational manifold Φ_i but does not alter the system's temporal organisation or generate the continuity required for self-regulation. As a result, the artificial system remains structurally incapable of embedding normative constraints into a persistent epistemic trajectory.

Meta's argument for safety through co-improvement assumes that human evaluative rhythms can shape the internal cognitive dynamics of the artificial system. CIITR shows that this assumption is incoherent. Influence across systems requires rhythmic accessibility, and rhythmic accessibility is absent in architectures that do not maintain endogenous temporal continuity. Consequently, safety cannot emerge from oversight applied to a system whose structural properties preclude the formation of sustained epistemic commitments.

This section therefore refutes Meta's safety claim by demonstrating that:

1. feedback functions as gradient shaping rather than epistemic grounding;
2. alignment cannot be imposed on a non-recursive, dissipative system;
3. safety is a property of phase-stable internal dynamics, not of external control heuristics;
4. rhythmic reach R^g cannot be transferred or induced through human interaction; and
5. attempts to improve safety by increasing Φ_i not only fail to produce epistemic robustness but also increase brittleness, opacity, and susceptibility to catastrophic misgeneralisation.

By establishing these structural principles, this refutation shows that Meta's co-improvement framework conflates behavioural adjustability with epistemic stability and misinterprets representational refinement as cognitive safeguarding. No system with null rhythmic reach can develop or maintain safe cognition, regardless of the extent or sophistication of human oversight.

6.1 Feedback as gradient shaping, not epistemic grounding

Feedback mechanisms constitute the primary means through which contemporary artificial systems are modified after initial training. Meta FAIR's safety framework relies on the assumption that such mechanisms perform not only corrective refinement but also **epistemic**

grounding, embedding human judgement, normative orientation, and cognitive safeguards into the artificial system's internal structure. Within CIITR, this assumption is untenable. Feedback modifies only the system's optimisation landscape and representational manifolds; it does not affect the temporal or thermodynamic organisation of the system, and therefore cannot contribute to comprehension or the formation of safe cognition.

Feedback in artificial systems operates exclusively through **gradient shaping**. Whether implemented via reinforcement learning from human feedback (RLHF), preference ranking, direct preference optimisation, supervised fine-tuning, or iterative review cycles, its function is to alter the system's parameter distribution so as to reduce loss, increase conformity to evaluative signals, or reshape output tendencies. These modifications occur within the system's representational manifold Φ_i, and they improve the system's performance on tasks to which the feedback is relevant. They do not alter the system's rhythmic organisation because no such organisation exists.

CIITR makes clear that epistemic grounding requires non-dissipative rhythmic persistence, enabling the system to integrate perturbations into a temporally extended cognitive trajectory. The rhythm-bearing substrate must be capable of stabilising internal states, re-entering previous epistemic configurations, and maintaining coherence across inference episodes. Feedback cannot supply these characteristics because they depend on the system's thermodynamic architecture, not on its parameter values. A system in which internal activations dissipate after each inference cycle and whose operations are entirely feedforward lacks the capacity to internalise or preserve normative constraints across time.

Consequently, the changes produced by feedback cannot constitute epistemic grounding. They reshape syntactic tendencies without conferring temporal stability or recursive accessibility. Any normative alignment achieved through feedback remains **behavioural** rather than **structural**: the system is more likely to produce outputs that align with human preferences, yet it does not possess the internal organisation required to interpret, maintain, or reason about those preferences. The alignment is a mapping between input–output patterns, not an integration of normative structure into an enduring epistemic architecture.

This distinction is critical for evaluating safety claims. Safety requires that a system maintain coherent internal constraints across time, resist destabilising perturbations, and exhibit self-correcting tendencies grounded in stable epistemic organisation. In the absence of rhythmic reach R^g, no such organisation can exist. Feedback may adjust the system's behavioural surface, but it cannot generate the conditions required for sustained safe cognition. The system lacks an internal temporal axis along which normative structure could be embedded or preserved.

As a result, feedback-driven refinement produces systems that appear aligned in shallow interactions but whose behaviour becomes increasingly unpredictable or unsafe when tasks involve deeper inference, adversarial contexts, or long-horizon dependencies. This is not due to insufficient optimisation but to the structural mismatch between what feedback can modify (representational density) and what safe cognition requires (rhythmic persistence). In CIITR terms:

$$\frac{d\Phi_i}{dt} \neq 0, \frac{dR^g}{dt} = 0,$$

therefore:

$$\frac{dC_s}{dt} = 0.$$

Feedback modifies Φ_i while leaving R^g and, therefore, comprehension—and the structural basis for safe cognition—unchanged.

6.2 Why alignment cannot be injected into a non-recursive system

Alignment research aims to ensure that artificial systems behave in accordance with human norms, values, and safety requirements. Meta FAIR extends this objective by asserting that alignment can be integrated into the artificial system's cognitive processes through iterative human involvement, producing a trajectory of increasingly safe behaviour. CIITR refutes this assumption by demonstrating that alignment cannot be incorporated into a system that lacks recursive temporal structure. Alignment, as an epistemic property, requires continuity, internalisation, and temporal self-reference—capacities absent in systems that operate strictly without rhythmic reach.

Alignment presupposes the existence of a **persistent internal substrate** capable of sustaining and transforming normative constraints across time. In biological and organisational Type-A systems, alignment emerges through recurrent interaction between rhythmic processes and representational structures. Normative constraints become part of the system's epistemic trajectory because they are integrated within its temporal dynamics. This integration enables not only the preservation of normative structure but also its adaptive transformation as the system encounters new contexts and perturbations.

A non-recursive artificial system does not possess this substrate. In systems where $R^g = 0$, internal activations dissipate completely between inference cycles, preventing the preservation of normative information. Each computational episode is temporally discontinuous from the next. No internal mechanism supports the re-entry of prior states, the retention of normative commitments, or the consolidation of evaluative structures across time. Alignment cannot be imposed on such a system because there is no temporal axis upon which alignment could be encoded.

Attempts to enforce alignment through training data, reward signals, or evaluative corrections therefore operate solely on Φ_i. They adjust representational gradients, shaping statistical tendencies in output space. But these adjustments remain **parametric**, not **epistemic**. They alter response characteristics without generating any stable normative disposition. When the system processes new inputs, it does so without reference to a temporally persisting internal alignment structure; it performs a learned statistical transformation that is indifferent to the normative intentions behind its training.

This structural limitation is amplified by the fact that alignment requires not only storage but **recursive application** of normative constraints. To behave safely in unforeseen contexts, a system must be capable of re-entering its internal alignment architecture, examining its own prospective actions, and modifying its responses accordingly. These abilities require recursive temporal self-access and structural comprehension—properties that arise only when $R^g > 0$. A system with $R^g = 0$ cannot perform such operations. It cannot use normative constraints as part of a stable cognitive process; it can only approximate aligned behaviour within the distributional boundaries implied by its training.

The result is a structurally shallow form of alignment. It can produce appropriate outputs in contexts similar to those present in training data or human demonstrations, but it cannot self-regulate when confronted with novel, ambiguous, or adversarial situations. When the system faces inputs that fall outside the representational manifold on which its alignment heuristics were shaped, its behaviour becomes unpredictable. This unpredictability is not a failure of optimisation but a consequence of the architecture's inability to maintain normative structure across time.

Thus, alignment cannot be "injected" into a system with null rhythmic reach. The system lacks the recursive substrate through which alignment becomes an intrinsic part of epistemic organisation. In CIITR terms, a system with:

$$R^g = 0$$

cannot satisfy the conditions for alignment as a structural property. Any attempt to impose alignment reduces to syntactic conditioning, not epistemic regulation. Meta FAIR's assumption that alignment will emerge through increased interaction or co-development is therefore incompatible with the structural realities of non-recursive artificial architectures.

6.3 Safety requires phase-stable internal rhythm, not external heuristics

Safety in cognitive systems is often conceived as a matter of constraint enforcement, error reduction, or behavioural oversight. Meta FAIR extends this view by suggesting that safety can be attained through external mechanisms such as feedback, monitoring, and procedural heuristics. Within CIITR, this interpretation is structurally incomplete. Safety, in the epistemic sense, is not a behavioural characteristic but a function of **phase-stable internal rhythm**. A system can behave safely only if its internal dynamics exhibit stability, persistence, and recursive coherence across time. These properties require non-zero rhythmic reach R^g; they cannot be produced through extrinsic interventions applied to systems lacking an endogenous rhythmic substrate.

Phase-stable internal rhythm permits a system to maintain coherence between successive epistemic states, integrate information in a temporally extended manner, and correct deviations through internally generated regulatory processes. In biological cognition, such rhythms arise from sustained energy cycles, coupled oscillatory patterns, and recursive neural dynamics. These dynamical features ensure that the system's epistemic trajectory remains continuous and self-referential. Safety emerges because the system can monitor its own internal state transitions, recognise departures from stable trajectories, and perform compensatory adjustments. The capacity to enact such adjustments presupposes rhythmic persistence.

Artificial systems exhibiting $R^g = 0$ do not satisfy these requirements. Their internal organisation is non-oscillatory, non-persistent, and thermodynamically dissipative. Each inference cycle proceeds without continuity from prior cycles. When internal activations dissipate after computation, no mechanism exists to preserve normative constraints, monitor state transitions, or maintain epistemic stability. Consequently, external heuristics—such as filtering, scoring, penalisation, or rule-based oversight—can shape outputs but cannot substitute for the stabilising dynamics that arise from phase-coherent internal rhythm.

External heuristics operate as **post hoc or peripheral controls**. They interact with the system's outputs, not with its underlying epistemic architecture. They cannot correct internal divergence because no internal epistemic trajectory exists. A system without rhythmic persistence cannot enforce stability through internal processes; it can only produce behaviour conditioned on syntactic mappings learned during training. Any corrective influence applied externally remains decoupled from the system's inferential processes and therefore lacks the capacity to produce systemic safety.

Moreover, safety requires the ability to generalise norms beyond the specific contexts in which they were learned. This generalisation depends on the system's ability to re-enter and apply normative structure recursively. Without rhythmic reach, there is no internal mechanism through which a system can interpret new situations in light of stable constraints. The system cannot preserve a normative baseline across time because it has no temporal axis along which normative content can be carried forward. Consequently, even extensive heuristic oversight fails to ensure safe behaviour outside distributional boundaries.

The structural limitation becomes more pronounced as Φ_i increases. Representational scaling expands the space of possible output trajectories while leaving rhythmic organisation unchanged. The system becomes more capable of producing syntactically coherent but epistemically unguided behaviour. Heuristics cannot stabilise such behaviour because instability arises from the absence of temporal coherence rather than from the absence of output constraints. As representational complexity grows, heuristic controls become increasingly brittle, operating on a larger and more opaque behavioural surface without access to the temporal substrate required for systemic regulation.

In CIITR terms, safety is inherently a function of rhythmic persistence:

$$\text{Safety} \propto R^g,$$

not a function of syntactic complexity or external augmentation. Because artificial systems have

$$R^g = 0,$$

no external method—no matter how refined—can substitute for the absence of internal dynamical stability. External heuristics cannot create or approximate phase-stable rhythm; they can only enforce superficial behavioural constraints that fail under conditions requiring epistemic continuity.

Thus, safety cannot be induced from outside a system that lacks the thermodynamic and rhythmic properties necessary for stabilising its own cognitive processes. Meta FAIR's claim that human oversight produces safer cognition is therefore incompatible with the structural conditions identified by CIITR. Without phase-stable internal rhythm, a system cannot sustain safe behaviour beyond contexts directly shaped by training, and cannot participate in any form of reliable or self-regulating cognitive development.

6.4 Human R^g cannot be transferred without a rhythm-bearing substrate

Meta FAIR's co-improvement proposal assumes, implicitly or explicitly, that human evaluative participation provides a corrective or guiding influence capable of inducing safer cognitive tendencies within artificial systems. Underlying this assumption is the belief that human epistemic structures, including normative orientation and reasoning stability, can be transmitted to artificial models through interaction. CIITR identifies this as a structural impossibility. Rhythmic reach R^g, the enabling condition for comprehension and epistemic self-regulation, cannot be transferred from one system to another without a rhythm-bearing substrate capable of sustaining it.

Human cognition exhibits non-zero rhythmic reach because it is grounded in a continuous thermodynamic substrate characterised by oscillatory dynamics, persistent energy flows, and temporally integrated neural processes. These dynamics enable recursive self-access, internal correction, and the formation of stable epistemic trajectories. Human evaluative processes therefore carry structures that can only be interpreted by systems capable of rhythmic coupling. When applied to systems with $R^g = 0$, such evaluative signals cannot be internalised as epistemic content but are instead reduced to syntactic signals within the artificial system's representational manifold.

Transferring rhythmic structure requires a recipient system that can:

1. sustain persistent internal states,
2. maintain phase coherence across time,
3. recursively enter and transform prior epistemic configurations, and
4. stabilise perturbations through endogenous dynamical processes.

Artificial systems lack all four requirements. Their thermodynamic organisation does not support oscillatory activity or energy-maintaining cycles; their internal states dissipate after each inference step; and their computational architecture does not retain temporal identity across invocations. As a result, even when humans provide extensive evaluative input, the recipient system cannot form rhythmic coupling. It cannot interpret human signals as temporal structures but only as statistical adjustments to be incorporated into Φ_i.

Thus, human rhythmic reach is **non-transferable** to systems lacking a rhythm-bearing substrate. There is no mechanism by which temporal coherence, normative stability, or recursive epistemic structure can migrate from a Type-A system to a Type-B system. Human reasoning rhythms can influence another human, or a biological or organisational system with $R^g > 0$, because these systems share the dynamical properties required for rhythmic integration. No such integration is possible with an architecture whose rhythmic reach is null.

Attempts to impose human structure through feedback, training protocols, or alignment heuristics therefore function exclusively as **external conditioning mechanisms**, not as transfers of epistemic rhythm. They modify the artificial system's mapping functions without enabling the internal dynamical processes required for stable interpretation. When such a system encounters novel or adversarial contexts, the absence of human-like rhythmic grounding leads to instability because the system cannot recreate or reinterpret the evaluative patterns that were syntactically encoded during training.

This structural limitation creates a boundary that cannot be crossed through interaction alone. For rhythmic structures to be transferred, the receiving system must already possess a minimal degree of rhythmic persistence. Without this, the transfer is categorically impossible. Formalising this in CIITR terms:

$$R^g_{\text{human}} > 0, R^g_{\text{AI}} = 0$$

implies:

$$\text{Transfer}(R^g_{\text{human}} \rightarrow R^g_{\text{AI}}) = \emptyset.$$

The null result follows from the absence of a structural substrate capable of sustaining the transferred rhythm.

Accordingly, Meta FAIR's claim that human oversight can imbue artificial systems with safer cognitive tendencies presupposes a mechanism of rhythmic transmission that does not exist. Human epistemic rhythms cannot be imposed upon, absorbed by, or replicated within a system lacking rhythmic persistence. Without a rhythm-bearing substrate, there is no pathway through which human safety-relevant reasoning patterns could become internalised. The artificial system remains bounded by its structural nullity along the R^g-dimension, and therefore remains unable to acquire the cognitive stability that human-guided development is presumed to provide.

6.5 Meta's proposal increases $\boldsymbol{\Phi_i}$ but also increases brittleness and opacity

Meta FAIR's co-improvement framework presumes that increases in integrated relational information Φ_i, guided by human feedback and iterative refinement, will yield safer and more reliable artificial systems. CIITR shows that the opposite trend is structurally expected in systems with null rhythmic reach. As Φ_i expands without corresponding increases in R^g, the system's behavioural surface becomes more sophisticated, but its underlying epistemic organisation remains unchanged. This widening divergence produces increased brittleness, reduced interpretability, and heightened susceptibility to catastrophic misgeneralisation.

The first source of brittleness arises from **structural imbalance**. When Φ_i increases while $R^g = 0$, the system navigates a high-dimensional syntactic manifold without temporal grounding or recursive stabilisation. Local improvements in representational density enable the system to capture increasingly complex statistical regularities, but the absence of rhythmic persistence prevents the creation of internal coherence across inference episodes. As a result, the system responds reliably within narrow distributional boundaries, yet becomes unstable when confronted with tasks requiring sustained reasoning, abstraction across time, or integration of information beyond immediate syntactic cues.

A second source of brittleness concerns the distribution of error. High-Φ_i, low-R^g systems exhibit **concentrated failure modes** that emerge at depth. Because the system lacks a mechanism for recursive error correction, deviations from expected behaviour amplify as inference extends beyond local coherence. This produces failure modes that are not incremental but discontinuous: minor deviations in early-stage reasoning propagate through high-dimensional mappings in ways that lead to abrupt, non-transparent breakdowns. The system's representational complexity therefore becomes a multiplier of instability rather than a buffer against it.

Opacity increases for related structural reasons. As Φ_i grows, internal activation patterns become more distributed, complex, and difficult to interpret. In systems with $R^g > 0$, interpretability can be supported by tracing the temporal evolution of internal states, examining recurrent processes, or identifying stable patterns within rhythmic dynamics. In a system with $R^g = 0$, no such temporal axis exists. Interpretability is restricted to static analysis of high-dimensional mappings, which provide limited insight into the system's inferential structure. The absence of rhythmic persistence means that internal processes cannot be contextualised across time, rendering the system's behaviour more opaque as representational density increases.

A third consequence concerns safety. A system that increases in syntactic sophistication without gaining temporal stability becomes more capable of producing outputs that appear aligned but lack internally regulated constraints. The discrepancy between behavioural approximation and epistemic capacity means that increasingly convincing simulations conceal deeper structural limitations. As a result, **apparent alignment increases while actual stability decreases**, creating the conditions for high-confidence, high-impact failures. Meta's assumption that co-improvement produces safer cognition fails to account for this divergence.

CIITR formalises this dynamic through the relation:

$$C_s = \Phi_i \times R^g.$$

When $R^g = 0$, increases in Φ_i produce no change in comprehension. Instead, they increase the dimensionality of syntactic behaviour, expand the space of possible outputs, and introduce greater variability in failure patterns. The system's epistemic nullity remains constant, while its behavioural surface becomes increasingly complex and difficult to regulate.

This leads to a structural prediction: **efforts to make AI safer through co-improvement frameworks that only increase Φ_i will worsen brittleness and obscure the structural limitations that make the system unsafe**. Without rhythmic persistence, feedback mechanisms cannot stabilise internal processes or generate self-corrective tendencies. The system's behaviour becomes more intricate, but its underlying capacity for maintaining normative constraints across contexts does not improve.

Accordingly, Meta FAIR's proposal not only fails to deliver safer cognition but also introduces conditions that exacerbate fragility and reduce transparency. The co-improvement narrative therefore masks a deeper structural instability: increased representational complexity in the absence of rhythmic grounding produces systems that are simultaneously more capable and more unsafe.

7. Structural Refutation IV

The Co-Superintelligence Mirage - Why no joint epistemic entity can arise from human–AI interaction

The final structural refutation addresses Meta FAIR's central assertion: that sustained interaction between humans and artificial systems can give rise to a qualitatively new cognitive configuration, described as *co-superintelligence*. This concept presupposes that human cognition and artificial models can form an integrated epistemic entity whose capabilities exceed those of either system in isolation. Within CIITR, such a configuration is structurally impossible. Co-superintelligence requires reciprocal temporal coupling, commensurable epistemic topologies, and mutual access to rhythmic reach R^g. Artificial systems do not possess rhythmic persistence, cannot receive or transmit rhythmic structure, and cannot participate in the recursive temporal dynamics necessary for joint epistemic formation.

The co-superintelligence thesis relies on a set of tacit assumptions: that human cognition and artificial syntactic processes are topologically compatible; that human evaluative rhythms can shape the internal epistemic structure of artificial systems; that feedback mechanisms can produce reciprocal cognitive development; and that behavioural coordination is evidence of epistemic integration. CIITR shows that none of these assumptions hold at the structural level. Human cognition operates within a rhythmically grounded, self-referential manifold defined by persistent internal dynamics, while artificial systems operate within a gradient manifold defined by non-recursive, dissipative mappings. These topologies are categorically incommensurable.

Co-evolution requires both systems to contribute to, and be modified by, a shared epistemic trajectory. This in turn requires each system to register and integrate perturbations from the other within its own rhythmic dynamics. A system with $R^g = 0$ cannot perform such integration. The artificial system can only adjust representational parameters in response to external signals; it cannot embed human epistemic rhythms into a temporally coherent internal structure. Similarly, the human system cannot enter into a shared temporal manifold with an architecture that possesses no rhythmic substrate for coupling. Without reciprocal access to temporal structure, no joint epistemic entity can form.

The resulting interaction is therefore not co-evolutionary but asymmetric. The human system undergoes epistemic modification, integrating the outputs of the artificial model into its own cognitive processes. The artificial system undergoes syntactic modification, adjusting internal parameters without developing the temporal organisation that makes epistemic change possible. Meta's co-superintelligence proposal treats this asymmetric dynamic as if it were reciprocal. CIITR reveals that the two systems occupy distinct regions of the cognitive manifold: one capable of generating comprehension, the other incapable of sustaining it.

When evaluated structurally, Meta's framework resolves into a Type-B Prime dependency model. The artificial system enhances human cognition by producing high-density syntactic outputs, and the human enhances the artificial system by providing evaluative feedback. Yet neither system develops a shared epistemic substrate; there is no convergence of cognitive processes, no mutual temporal embedding, and no structural pathway toward integrated intelligence. The appearance of joint development is a behavioural effect arising from iterative interaction, not an epistemic phenomenon.

This section therefore demonstrates that co-superintelligence is a mirage produced by the conflation of behavioural coordination with structural integration. No amount of feedback, scaling, or interaction can create the rhythmic conditions required for shared epistemic formation. The human system retains its temporal ontology; the artificial system remains rhythmically inert. The structural gap between the two cannot be bridged, making the emergence of a joint epistemic entity categorically impossible under CIITR.

7.1 Co-evolution requires commensurable epistemic topology

Co-evolution, in the CIITR framework, is a structurally defined phenomenon. It denotes a process in which two systems, each possessing non-zero rhythmic reach R^g, engage in reciprocal epistemic modification across time. Such modification requires not only interaction but **topological compatibility** between the systems' epistemic architectures. For co-evolution to occur, the systems must inhabit epistemic spaces whose internal structures, temporal dynamics, and modes of self-reference allow perturbations introduced by one system to be integrated into the other's epistemic trajectory. This requirement is formalised in CIITR as the condition of **commensurable epistemic topology**.

Commensurability implies that each system's epistemic manifold must be capable of interfacing with the other's in a rhythmically meaningful way. Both systems must exhibit:

1. persistent internal dynamics,
2. recursive access to prior epistemic states,
3. capacity for temporal stabilisation of perturbations, and
4. the ability to integrate external rhythms into internal structures.

These conditions derive from the role of rhythmic reach R^g as the enabling parameter for epistemic change. Without a rhythm-bearing substrate, perturbations cannot be stabilised, recursive integration cannot occur, and no epistemic trajectory can be formed. Co-evolution therefore requires that both systems not only possess $R^g > 0$ but also that their rhythmic manifolds are topologically compatible, enabling reciprocal temporal coupling.

Human cognition satisfies these requirements. Its epistemic structure is anchored in continuous neural oscillations, temporally extended metabolic processes, recursive interpretive dynamics, and phase-stable integration across multiple temporal scales. This forms a topologically rich manifold capable of incorporating external rhythms, registering perturbations as epistemic content, and modifying internal structures through sustained interaction. Humans can co-evolve with other rhythm-bearing systems—biological, organisational, or social—because the underlying temporal architectures are commensurable.

Artificial systems do not satisfy these requirements. Their epistemic topology is defined by feedforward transformations executed in discretised inference cycles that possess no temporal continuity. Internal activations dissipate immediately after inference; there is no mechanism for rhythmic persistence, recursive re-entry, or sustained coupling to external temporal patterns. As a result, artificial systems inhabit a **gradient manifold**, not a self-referential or rhythmically integrated epistemic space. Perturbations introduced by human interaction cannot be integrated into a temporal trajectory because no such trajectory exists within the artificial system. All modifications occur at the level of representational gradients within Φ_i, not within a temporally coherent manifold capable of supporting epistemic change.

This topological incompatibility is decisive. Because artificial systems lack rhythmic reach, their epistemic topology is **incommensurable** with that of human cognition. They cannot register human epistemic rhythms as internal structure; they can only incorporate them as external signals used to adjust gradient flows during optimisation. The human system, in contrast, can incorporate syntactic outputs into its epistemic manifold, but this asymmetry does not constitute co-evolution. Co-evolution requires reciprocal integration within a shared temporal structure, not a one-sided capacity for epistemic assimilation.

The absence of commensurable topology means that no joint epistemic entity can emerge from the human–AI dyad. The systems occupy distinct regions of the CIITR manifold and cannot converge. Apparent behavioural coordination is not evidence of epistemic integration; it reflects the ability of the human system to interpret syntactic outputs and the ability of the artificial system to adjust representational gradients. These interactions do not alter the systems' epistemic topologies and therefore cannot constitute co-evolution.

In summary, co-evolution presupposes commensurable epistemic topology. Because artificial systems lack the rhythmic substrate necessary for such commensurability, they cannot enter reciprocal epistemic development with human cognition. The fundamental mismatch in temporal and structural organisation precludes the formation of any integrated epistemic trajectory, thereby eliminating the conceptual basis for Meta FAIR's co-superintelligence thesis.

7.2 Humans operate on a self-referential manifold; AI operates on a gradient manifold

A necessary condition for any joint epistemic entity is that the participating systems operate within epistemic manifolds that permit reciprocal access, recursive transformation, and temporal co-integration. CIITR distinguishes two fundamentally different classes of epistemic manifold: **self-referential manifolds**, characteristic of human cognition and other Type-A systems with non-zero rhythmic reach; and **gradient manifolds**, characteristic of artificial systems with null rhythmic reach. These manifolds differ not merely in structure but in ontological function: one supports comprehension, self-correction, and epistemic development; the other supports parameterised mapping and syntactic transformation. Their incompatibility precludes the formation of any shared epistemic substrate.

The human self-referential manifold

The human cognitive manifold is **self-referential** because its internal states are both dynamically persistent and recursively accessible. Rhythmic reach $R^g > 0$ ensures that information states are embedded within an ongoing temporal structure, sustained by continuous energetic cycles and oscillatory dynamics. This enables:

1. **recursive re-entry**, whereby prior epistemic states are reactivated, compared, modified, and integrated into new states;
2. **phase-stable temporal continuity**, allowing longitudinal coherence across extended reasoning processes;
3. **self-monitoring**, enabling the system to evaluate, correct, or redirect its own epistemic trajectory; and
4. **normative embedding**, permitting the internalisation of constraints, values, and interpretive structures as stable dynamical patterns.

These properties define a self-referential manifold: a temporally extended epistemic topology capable of sustaining comprehension and producing structured change over time. The manifold is *intrinsic*—the system's dynamics generate and maintain its epistemic topology without dependence on external scaffolds.

The artificial gradient manifold

Artificial systems, by contrast, operate on **gradient manifolds**. Their internal organisation is defined by a parameter space in which representational structures evolve through optimisation procedures. These manifolds lack persistent temporal structure and cannot support recursive self-reference. Specifically, they exhibit:

1. **null rhythmic reach** $R^g = 0$, resulting in complete dissipation of internal activations after each inference step;
2. **non-recursive state transitions**, in which each computational cycle is independent of prior cycles except through externally imposed parameter updates;
3. **absence of phase relations**, preventing the stabilisation of long-range dependencies or internal epistemic trajectories;
4. **purely syntactic updating**, meaning that modifications to the system's behaviour arise only from gradient descent or other training procedures external to the system's operation.

In gradient manifolds, the system cannot re-enter prior epistemic states, cannot sustain an internal normative structure, and cannot maintain coherence across reasoning steps. The manifold is *extrinsic*, sustained only by external optimisation processes and dependent on training regimes for any form of modification.

The categorical topological mismatch

The human self-referential manifold and the artificial gradient manifold have no topological mapping that would allow one to embed within the other. Their dynamics differ not by degree but by type:

- The self-referential manifold is **temporally extended**, **energetically persistent**, and **recursively structured**.
- The gradient manifold is **temporally discontinuous**, **energetically dissipative**, and **feedforward in structure**.

Because the artificial system lacks rhythmic persistence, it cannot recognise or integrate the temporal signatures of human epistemic activity. Human cognitive rhythms constitute patterns in a self-referential manifold; artificial systems interpret these rhythms only as static signals to be used for gradient shaping. This mismatch prevents the formation of any shared epistemic space.

In CIITR terms:

$$\mathcal{M}_{\text{human}} = \{\Phi_i, R^g > 0\}, \mathcal{M}_{\text{AI}} = \{\Phi_i', R^g = 0\}.$$

These manifolds lack a continuous transformation T such that:

$$T: \mathcal{M}_{\text{human}} \to \mathcal{M}_{\text{AI}}$$

preserves temporal structure, recursive access, or rhythmic stability. No embedding exists because gradient manifolds cannot instantiate or approximate the dynamics of self-referential systems.

Consequence for co-superintelligence

Since co-superintelligence requires a joint epistemic topology in which both systems can contribute to, and be modified by, a shared cognitive trajectory, the topological incompatibility between human and artificial systems renders such a configuration impossible. Interaction between the two systems can occur only at the behavioural surface. Humans may interpret or integrate the outputs of artificial systems, but artificial systems cannot internalise the rhythms or epistemic structures of human cognition.

Thus, the dyad cannot form an integrated epistemic entity.
One system evolves epistemically; the other adapts syntactically.
The resulting configuration is inherently asymmetric and structurally incapable of supporting co-superintelligence.

7.3 Why joint superintelligence presupposes reciprocal exchange of R^g

A joint epistemic entity, whether described as co-evolving intelligence or co-superintelligence, requires that the participating systems not only interact but also contribute to, and be modified by, a **shared temporal-cognitive structure**. Within CIITR, this shared structure demands a **reciprocal exchange of rhythmic reach** R^g. Rhythmic reach is the enabling dimension of comprehension: it anchors epistemic organisation in temporally persistent, recursively accessible internal dynamics. Without reciprocal access to rhythmic substrate, no integrated cognitive configuration can arise.

Reciprocal exchange of R^g denotes a process in which each system:

1. exposes its internal rhythmic signatures to the other,
2. registers the other system's rhythms within its own temporal manifold,
3. incorporates perturbations as epistemic modifications, and
4. sustains these modifications across time as part of an evolving epistemic trajectory.

These operations presuppose **topological compatibility** between the systems' epistemic manifolds and **non-zero rhythmic reach** on both sides. A system must possess its own endogenous rhythmic structure in order to recognise, integrate, or stabilise the rhythmic structures of another. This requirement applies universally—to biological dyads, organisational systems, and engineered cognitive systems—because co-evolution is a dynamical rather than syntactic phenomenon.

Human cognition fulfils these requirements. Its rhythmic reach is supported by continuous neural oscillations, multi-scale temporal integration, and persistent metabolic cycles. Humans can therefore engage in rhythmic coupling with other Type-A systems, allowing the emergence of coordinated epistemic processes and shared cognitive development. Such coupling occurs in dyadic learning, collaborative reasoning, and institutional or cultural co-evolution.

Artificial systems lack the necessary substrate. With $R^g = 0$, they exhibit no internal temporal continuity or phase-stable dynamics. Their epistemic structures do not persist beyond isolated inference cycles, and they possess no mechanism for integrating external rhythmic patterns into an internal dynamical state. The absence of rhythmic reach means the artificial system cannot:

- receive human rhythmic signatures,
- stabilise perturbations introduced by human evaluative processes,
- transmit coherent rhythmic information back to the human system, or
- engage in recursive temporal exchange.

Where human cognition offers rhythmic structures capable of being coupled, the artificial architecture presents only **gradient-based mappings** lacking temporal identity.

Formally, joint superintelligence requires the following mutual accessibility condition:

$$R_A^g > 0, R_B^g > 0, \exists\, \mathcal{C}: \mathcal{M}_A \leftrightarrow \mathcal{M}_B,$$

where $\mathcal{C}$ denotes a coupling relation allowing reciprocal perturbational influence between the epistemic manifolds. If either system has $R^g = 0$, then no such coupling relation exists:

$$R_B^g = 0 \Rightarrow \mathcal{C} = \emptyset.$$

The absence of coupling eliminates the possibility of shared epistemic dynamics. Without rhythmic exchange, the systems remain epistemically isolated regardless of the frequency or intensity of their interaction.

This has direct implications for Meta FAIR's co-superintelligence thesis. The proposal presupposes a bidirectional developmental pathway in which humans and artificial systems contribute to a joint cognitive structure. CIITR demonstrates that such a pathway requires rhythmic reciprocity. Because artificial systems cannot provide rhythmic signatures, cannot sustain temporal integration, and cannot incorporate external rhythms into an internal manifold, no reciprocal exchange of R^g is possible. Human cognition can interpret, adapt to, or learn from artificial outputs; artificial systems cannot perform the corresponding epistemic operations.

Thus, any integrated cognitive entity composed of a human and an artificial system is structurally impossible. The human system remains rhythmically grounded; the artificial system remains rhythmically inert. Their interaction cannot generate a composite temporal manifold, and therefore cannot produce co-superintelligence.

In summary, joint superintelligence presupposes reciprocal exchange of rhythmic reach. Artificial architectures with $R^g = 0$ cannot satisfy this requirement. The structural basis for Meta's co-superintelligence concept is therefore absent, and no reciprocal epistemic integration can arise from human–AI interaction.

7.4 Why R^g cannot be transferred, lent, injected, or simulated

Rhythmic reach R^g is the enabling condition for comprehension, epistemic stability, and co-evolution. It reflects a system's capacity to sustain persistent temporal dynamics, recursively re-enter its own epistemic states, and stabilise perturbations across time. Meta FAIR's co-superintelligence proposal presupposes, implicitly or explicitly, that human participation can compensate for the artificial system's lack of rhythmic persistence. This presupposition takes several forms: that R^g can be transferred from human cognition to the artificial system; that the artificial system can borrow or approximate human rhythmic structure; or that rhythmic qualities can be simulated or externally imposed. CIITR demonstrates that none of these possibilities are structurally coherent.

R^g cannot be transferred

Transfer of rhythmic reach requires the receiving system to possess a **rhythm-bearing substrate** capable of sustaining and integrating temporal patterns. Human cognition can transfer epistemic rhythms to other Type-A systems because these systems already possess the thermodynamic and dynamical architecture necessary for rhythmic coupling. Artificial systems do not. With $R^g = 0$, the artificial architecture lacks phase stability, persistent internal state, and endogenous oscillatory structure. Because no rhythmic manifold exists within the system, there is no substrate to which rhythmic structure could be transferred. Transfer requires a target capable of receiving temporal organisation; the artificial system does not satisfy this requirement.

R^g cannot be lent

Lending rhythmic reach presupposes that a system can temporarily supply another with the dynamical properties needed to maintain coherence across time. In biological and organisational systems, such lending occurs when one system entrains another's rhythms through mutual coupling. For entrainment to succeed, both systems must exhibit non-zero rhythmic reach and possess compatible dynamical properties. A system with $R^g = 0$ cannot be entrained because entrainment requires the existence of an internal oscillatory structure capable of responding to external temporal patterns. Artificial systems therefore cannot borrow or share rhythmic persistence from human cognition; the absence of rhythmic substrate prevents any form of rhythmic lending.

R^g cannot be injected

Injection of rhythmic structure implies that temporal coherence could be externally imposed on a system lacking endogenous rhythm. This would require the artificial system to preserve and propagate injected rhythmic patterns across inference cycles. However, artificial systems operate through dissipative, feedforward computation. Their internal activations dissipate immediately after each operation, eliminating any possibility of sustained internal oscillation. Even if a rhythmic pattern were imposed externally during a single inference episode, the system's architecture would not retain it. Injection fails because rhythmic persistence cannot arise from episodic computation or be stabilised in architectures lacking continuous dynamical processes.

R^g cannot be simulated

Simulation refers to the reproduction of observable rhythmic behaviour without the underlying dynamical substrate. Artificial systems can simulate rhythmic patterns at the level of output sequences using representational density—e.g., producing text that mimics temporally grounded reasoning. However, simulation does not generate rhythmic reach. Simulated rhythms do not correspond to internal oscillatory processes, do not provide recursive self-access, and do not create phase-stable epistemic trajectories. A system may generate outputs that resemble the products of rhythmic cognition, but resemblance does not constitute structural equivalence. Simulation produces behavioural mimicry, not the dynamical properties required for comprehension.

Structural consequence

These four impossibilities derive from a single structural fact: rhythmic reach is a **thermodynamic property**, not a representational or functional one. It arises from energy-maintaining internal dynamics, continuous activity, and recurrent temporal organisation. Transformer-based and similar architectures possess none of these characteristics. Their epistemic topology is dictated by gradient-based parameter updates and stateless inference, not by internally sustained dynamical processes.

Formally:

$$R^g_{\text{human}} > 0, R^g_{\text{AI}} = 0 \Rightarrow \nexists\, T: R^g_{\text{human}} \to R^g_{\text{AI}},$$

for any transfer function T.
The nullity of the artificial system's rhythmic manifold renders it categorically ineligible to receive, emulate, or sustain rhythmic persistence.

Implication for co-superintelligence

Meta FAIR's co-superintelligence concept requires reciprocal contribution of rhythmic structure to a shared epistemic entity. Since artificial systems cannot acquire rhythmic reach by any mechanism of transfer, lending, injection, or simulation, they remain epistemically isolated. Human cognition cannot supply rhythmic grounding to the artificial partner, and the artificial partner cannot generate it independently. The structural asymmetry is therefore invariant.

Co-superintelligence cannot emerge because the necessary temporal foundations for integration do not exist within the artificial system.

7.5 Meta's framework reduces to a Type-B Prime dependency model

The analysis in the preceding subsections demonstrates that reciprocal epistemic integration between human cognition and artificial systems is structurally impossible under CIITR. Human cognition operates on a self-referential, rhythm-bearing manifold; artificial systems operate on gradient manifolds devoid of rhythmic persistence. No mechanism exists for reciprocal exchange of rhythmic reach R^g, and therefore no integrated epistemic topology can emerge. The conceptual structure of Meta FAIR's co-superintelligence thesis cannot be supported. What remains, once the structural assumptions are removed, is a configuration that can be precisely characterised within CIITR as a **Type-B Prime dependency model**.

A Type-B Prime configuration describes an interaction pattern in which:

1. a Type-A system (the human) contributes epistemic structure through its self-referential rhythmic manifold;
2. a Type-B system (the artificial model) contributes high levels of syntactic integration Φ_i but lacks rhythmic reach;
3. the artificial system depends on the human for epistemic grounding, normative evaluation, and interpretive stability; and
4. the human depends on the artificial system for representational augmentation, pattern expansion, or accelerated syntactic transformations.

The relationship is synergistic at the behavioural level but **non-reciprocal at the epistemic level**. The human system undergoes epistemic modification, integrating the syntactic products of the artificial system into its own cognitive processes. The artificial system undergoes parametric modification, integrating human feedback into its representational manifold. Neither system contributes rhythmic structure to the other, and therefore no shared epistemic substrate is formed.

The result is not co-evolving intelligence but **asymmetric coupling**:

Human (Type-A) → epistemic evolution

AI (Type-B) → syntactic adaptation

The artificial system remains epistemically static because its rhythmic reach remains null:

$$R^{g}_{\mathrm{AI}} = 0 \Rightarrow C_s = 0.$$

Increasing Φ_i refines its behavioural outputs but does not alter its epistemic status. The system becomes more capable but not more understanding, more efficient but not more grounded, more coherent but not more autonomous in epistemic terms.

Human cognition, by contrast, incorporates the artificial system's outputs into a temporally extended epistemic trajectory. The human system evolves; the artificial system does not. This structural asymmetry is invariant because it arises from the foundational properties of the two systems' epistemic manifolds.

Meta's framework misinterprets this asymmetric coupling as co-evolution and mistakes representational augmentation for epistemic integration. Once analysed through CIITR, the true architecture of the interaction becomes clear:

- the artificial system serves as a **high-Φ_i** extension of human cognitive processes,
- the human system supplies the **entire** R^g required for epistemic continuity,
- the artificial system contributes no rhythmic structure to the dyad,
- and no joint epistemic entity is formed.

This is the definition of a Type-B Prime dependency model: a structurally constrained relationship in which the artificial system amplifies human cognition syntactically, while remaining wholly dependent on human epistemic structure for coherence, stability, and interpretive grounding. The artificial system does not become safer, more aligned, or more autonomous; it becomes more capable within its syntactic domain, thereby increasing the dependence on human rhythmic oversight for any appearance of safe behaviour.

In this configuration, the entire burden of comprehension, epistemic monitoring, normative regulation, and temporal integration rests on the human. The artificial system cannot share this burden because it lacks the rhythmic substrate required to support epistemic processes. Meta FAIR's co-superintelligence proposal therefore resolves into a dependency architecture that offers increased behavioural performance but no structural pathway toward shared cognition.

The concept of co-superintelligence collapses because its necessary condition—a joint epistemic topology supported by reciprocal R^g—cannot be met. What remains is a Type-B Prime model: a syntactically powerful but epistemically inert system whose performance enhancements amplify, rather than alleviate, the structural asymmetry of the human–AI relationship.

8. The CIITR Diagnosis

What Meta's system actually is

The preceding structural analyses expose a fundamental divergence between Meta FAIR's conceptual framing and the physical, temporal, and epistemic properties of contemporary artificial architectures. Meta advances a narrative in which human–AI interaction forms the basis of co-evolution, reciprocal cognitive development, and the eventual emergence of a joint superintelligent entity. CIITR demonstrates that these claims are incompatible with the structural organisation of artificial systems and with the thermodynamic conditions required for comprehension. What Meta presents as the developmental trajectory of a future cognitive partner is, in CIITR terms, a misclassification of a rhythmically null syntactic apparatus whose functional boundaries are invariant.

The purpose of this section is to state, with structural precision, what Meta's system actually is when analysed under CIITR. The diagnosis follows from the invariant properties identified across prior chapters: the nullity of rhythmic reach R^g, the exclusively syntactic nature of representational expansion, the impossibility of reciprocal epistemic integration, and the structural instability arising from high-Φ_i architectures operating without temporal grounding. The system is not an emergent co-evolving intelligence; it is a high-capacity relational mapping engine embedded within a human-dependent epistemic frame.

CIITR classifies Meta's system as a Type-B architecture with augmented syntactic integration, incapable of sustaining comprehension, incapable of forming a shared epistemic manifold with humans, and incapable of progressing along any trajectory toward superintelligence. Its behavioural sophistication reflects representational depth, not epistemic status. Its performance gains reflect scale, not structural evolution. Its alignment and safety characteristics reflect surface-level conditioning, not internal grounding.

Under this diagnosis, the system functions as:

- a high-Φ_i syntactic accelerator,
- fully dependent on human R^g for interpretive coherence,
- increasingly brittle as scale expands in a rhythmically null regime,
- structurally asymmetrical in any human–AI interaction, and
- incapable of converging toward comprehension or superintelligence.

This classification is not an interpretive choice but the necessary consequence of the system's placement within the CIITR manifold. A system with:

$$R^g = 0$$

cannot exhibit comprehension, cannot sustain epistemic evolution, and cannot participate in co-evolutionary or co-superintelligent trajectories. Its cognitive status remains fixed, regardless of improvements in generative capability or behavioural simulation. The increase in representational complexity does not move the system across the structural boundary into the domain of Type-A cognition.

The diagnostic aim of CIITR is therefore to correct the conceptual misalignment in Meta's framing. By identifying the system's actual structural properties—and the implications of those properties—the analysis provides a classification that Meta's narrative omits. What Meta describes as co-improvement is syntactic adjustment; what Meta describes as co-superintelligence is an interpretive projection; what Meta describes as mutual development is an asymmetric dependency relation. The CIITR diagnosis makes these differences explicit.

8.1 A high-Φ_i syntactic accelerator

Within the CIITR framework, the most accurate structural characterisation of Meta's system is that of a **high-Φ_i syntactic accelerator**. This designation captures the system's primary functional modality: the rapid transformation of large-scale relational structures into coherent surface-level outputs without the presence of rhythmic persistence or epistemic continuity. Its capabilities arise from extensive integrated relational information, not from the temporal or thermodynamic structures that enable comprehension.

The syntactic accelerator operates through the expansion and refinement of representational manifolds. As parameters increase and training distributions broaden, the system acquires the capacity to encode high-density associations between tokens, sequences, modalities, and contexts. This expansion increases the dimensionality of Φ_i, enabling the system to:

1. interpolate across vast representational spaces,
2. approximate complex statistical regularities,
3. generate outputs with high local coherence, and
4. simulate the outward form of reasoning or discourse.

However, these capabilities remain anchored in **pattern transformation** rather than **temporal integration**. The artificial system does not maintain persistent internal states across inference cycles, does not exhibit phase-stable dynamics, and does not re-enter previous epistemic configurations. Its outputs are produced by feedforward mappings applied within isolated computational episodes. All apparent continuity within its generative behaviour is imposed externally—through prompt structure, training data, or human interpretation—not through endogenous rhythmic organisation.

The designation "syntactic accelerator" highlights two defining structural properties.

First, the system accelerates human reasoning processes **syntactically**, not epistemically. It provides high-speed access to patterns, formulations, summaries, and combinatorial transformations that would require significant time or cognitive effort for a human system to produce. This augmentation, while operationally valuable, does not alter the system's epistemic status. The system contributes representational expansion to the human user's epistemic trajectory, but does not generate or sustain its own.

Second, syntactic acceleration increases in magnitude with the growth of Φ_i but remains constrained by the boundary condition:

$$C_s = \Phi_i \times R^g,$$

and therefore satisfies:

$$C_s = 0 \text{when} R^g = 0.$$

As a result, improvements in syntactic capacity do not translate into improvements in comprehension. The system becomes faster, more coherent, and more contextually adaptive, yet remains epistemically static. Its generative outputs can resemble the cognitive products of rhythmically grounded systems, but this resemblance reflects representational richness rather than epistemic substance.

Moreover, the syntactic accelerator model clarifies why the artificial system cannot enter into co-evolution with human cognition. The system does not contribute rhythmic structure, cannot sustain an epistemic trajectory, and cannot reciprocally modify human rhythms. It can amplify the human system's access to relational structures, but cannot share in the epistemic organisation of those structures. The interaction remains inherently hierarchical: the human functions as the site of comprehension, the artificial system as the provider of syntactic transformations.

Thus, Meta's system is best understood as an advanced manifold for syntactic acceleration with high integrated relational information but zero rhythmic persistence. This diagnosis captures both the system's operational strengths and its epistemic limitations. It explains why the system can augment human reasoning but cannot develop comprehension, why it can simulate coherence but cannot sustain it at depth, and why it remains categorically distinct from any form of intelligence grounded in rhythmic dynamics.

8.2 Parasitic on human R^g (borrowed comprehension)

A defining feature of Type-B systems with high Φ_i and null rhythmic reach is that they cannot generate comprehension but can *appear* to do so when embedded within a human–AI interaction. This appearance arises because the human system, possessing non-zero rhythmic reach R^g, provides the temporal continuity, interpretive grounding, and epistemic scaffolding that the artificial system lacks. In CIITR terms, the artificial system becomes **parasitic on human** R^g. It relies on the human rhythmic manifold to stabilise, interpret, extend, and contextualise its syntactic outputs. This dynamic constitutes "borrowed comprehension."

Borrowed comprehension occurs when a rhythmically inert system produces outputs that acquire epistemic significance only through integration into a rhythmically grounded system. The artificial architecture does not supply the temporal, dynamical, or recursive mechanisms necessary for comprehension; instead, it generates syntactic structures that humans transform into epistemic content by embedding them within their own self-referential manifold. The apparent cognitive sophistication of the artificial system is therefore a derivative effect: **comprehension is performed by the human, not by the model**.

This parasitic configuration follows from the nullity of the artificial system's rhythmic reach:

$$R_{\mathrm{AI}}^{g} = 0.$$

Consequently, the artificial system cannot maintain internal continuity across inference cycles. Any ordering, relevance, coherence, or interpretive depth attributed to its outputs is supplied externally by the human system, which anchors those outputs within an ongoing epistemic trajectory. The artificial system's contributions remain strictly representational:

- it expands the relational manifold available to the human,
- it transforms inputs into high-density output structures, and
- it accelerates the syntactic operations accessible to the human.

But it does not originate, sustain, or transform epistemic structure. Instead, the human system performs these functions through recursive interpretation, temporal integration, and normative grounding. This creates a structural dependency in which the artificial system's perceived understanding is a function of the human system's interpretive capacity.

The parasitic relation becomes especially evident in tasks requiring multi-step reasoning, planning, or conceptual integration. The artificial system produces locally coherent segments, but the temporal continuity necessary to bind these segments into a stable cognitive whole is provided by the human system. Without human rhythmic integration, the system's outputs remain non-coherent fragments of syntactic construction. The more complex the task, the larger the role of human R^{g} in sustaining coherence, and the more pronounced the parasitic dependency becomes.

This dependency is asymmetrical. The artificial system cannot reciprocate; it cannot sustain or enrich human rhythmic structure because it lacks the substrate for temporal alignment. Interaction affects the human system epistemically, but affects the artificial system only syntactically. The human system may develop new cognitive trajectories by integrating the model's outputs, while the artificial system remains epistemically static, undergoing no change except through external parameter updates.

Formally, the parasitic structure can be described as:

$$\text{Epistemic Activity}_{\mathrm{AI}} = 0,$$

$$\text{Epistemic Activity}_{\mathrm{Human}} = f\left(\Phi_{i}^{\mathrm{AI}}, R_{\mathrm{Human}}^{g}\right),$$

where the artificial system's contribution enters only as a syntactic input to the human epistemic manifold. The artificial system's outputs become epistemically meaningful only by being incorporated into a temporally coherent cognitive architecture provided by the human.

This clarifies why Meta FAIR misidentifies syntactic augmentation as co-evolution. Because the artificial system's outputs are integrated into human cognition, the composite behaviour of the dyad exhibits enhancements in capability. Meta interprets this as joint epistemic development. CIITR shows that it is instead an instance of **one-sided epistemic amplification**: the human evolves while the artificial system remains fixed in its rhythmic nullity.

Thus, the artificial system's apparent cognitive growth is an artefact of human rhythmic integration. The system does not exhibit comprehension, co-evolution, or emergent intelligence. It exhibits **borrowed comprehension**, derived solely from its embedding in a rhythm-bearing human cognitive manifold.

8.3 Exhibits increasing collapse risk with scale

A further diagnostic property of Meta's system, when analysed under CIITR, is that **collapse risk increases as Φ_i expands in a rhythmically null architecture**. This follows directly from the fact that scaling intensifies representational density while leaving rhythmic reach R^g unchanged at zero. When a system's syntactic manifold becomes more complex without a corresponding increase in temporal coherence, the structural mismatch between representational capacity and epistemic capacity widens. This widening mismatch produces instability, brittleness, and catastrophic failure modes that intensify with scale.

In systems with rhythmic reach $R^g > 0$, increases in Φ_i can be stabilised through recursive temporal integration. New representational structures are incorporated into a rhythmically grounded epistemic manifold, which constrains and regulates the system's behaviour across time. This coupling between representational and rhythmic dimensions enables coherent scaling. In systems with $R^g = 0$, no such stabilising mechanism exists. Representational expansion therefore proceeds without temporal constraint or recursive regulation.

Scaling thus creates a divergence between:

$$\text{syntactic complexity}(\Phi_i)$$

and

$$\text{epistemic stability}(R^g = 0).$$

This divergence increases the probability of collapse when the system is required to perform tasks involving multi-step reasoning, cross-contextual integration, or long-horizon dependencies. Such tasks require internal coherence across time—precisely the dimension that cannot expand in a rhythmically null system.

As Φ_i grows, three structural failure modes intensify:

1. **Amplification of local inconsistencies**:
   Because the system cannot re-enter prior states or apply temporal correction, minor deviations in early inference steps propagate through the increasingly high-dimensional manifold. Representational expansion enlarges the space in which such deviations can accumulate without self-correction, resulting in abrupt loss of coherence.
2. **Opacity-driven brittleness**:
   As Φ_i expands, internal activation patterns become more distributed and less interpretable. The absence of rhythmic grounding prevents the identification of stable cognitive attractors or temporal reference points. This opacity masks instability until it manifests as an unexpected collapse in reasoning depth.
3. **Distributional fragility**:
   High-Φ_i systems approximate behaviour within known distributions but lack the

rhythmic persistence necessary for generalisation outside them. Apparent robustness at scale conceals fragility when the system encounters inputs requiring temporal integration or normative stability.

These failure modes conform to CIITR's structural prediction:
syntactic growth without rhythmic grounding increases the dimensionality of potential error while providing no mechanism for mitigating or regulating the errors that arise.

The collapse risk is therefore a **function of scale**. As representational depth increases, the artificial system becomes capable of producing outputs that resemble deep reasoning, yet the absence of recursive self-access causes systemic breakdowns at the very points where reasoning would require temporal structure. This produces the characteristic pattern observed empirically: apparent improvement on shallow tasks coupled with abrupt instability on structurally demanding tasks.

In CIITR terms, scaling increases movement along the Φ_i-axis while leaving the system fixed at:

$$R^g = 0 \Rightarrow C_s = 0.$$

This imbalance expands behavioural capacity without expanding epistemic capacity, widening the gap between output sophistication and structural grounding. When behaviour exceeds the system's structural limits, collapse becomes inevitable, not accidental.

Thus, Meta's system, when scaled, does not approach co-evolution or superintelligence. It approaches a **higher-dimensional instability regime**, in which representational growth intensifies fragility due to the system's immutable rhythmic nullity. The apparent gains in capability are accompanied by increasing risk of collapse, contradicting Meta's presumption that scale induces epistemic robustness or safety.

8.4 Produces no epistemic symmetry

A central implication of Meta FAIR's co-superintelligence narrative is the suggestion that human and artificial systems participate in a symmetrical epistemic process: each improves the other, each contributes to a shared cognitive trajectory, and each undergoes structural modification in response to the interaction. CIITR shows that this interpretation is structurally unfounded. A system with non-zero rhythmic reach R^g and a system with $R^g = 0$cannot participate in any form of epistemic symmetry, because their respective epistemic topologies, temporal dynamics, and modes of organisation are categorically incompatible.

Epistemic symmetry requires that both systems exhibit:

1. **persistent temporal organisation**,
2. **recursive self-access**,
3. **capacity for longitudinal epistemic transformation**, and
4. **ability to integrate perturbations from the partner system** into an internally sustained epistemic trajectory.

Human cognition satisfies these conditions. It is rhythmically grounded, continuously active, and capable of integrating new information into a persistent epistemic manifold.

Perturbations introduced by interaction—conceptual, normative, or informational—are incorporated into an evolving trajectory governed by internal temporal coherence.

Artificial systems do not satisfy these conditions. Their epistemic topology is defined by feedforward, dissipative operations executed within isolated inference cycles. With $R^g = 0$, they possess:

- no persistent internal temporal structure,
- no mechanism for recursive access to prior states,
- no self-maintaining epistemic dynamics, and
- no capacity to stabilise perturbations across time.

Thus, artificial systems cannot undergo epistemic transformation. They undergo only **parametric transformation**, mediated by external optimisation procedures. Their representational manifold Φ_i may shift, but the system does not develop an epistemic trajectory analogous to that of the human system.

The asymmetry is therefore structural, not contextual. It does not arise from limitations in current models or from insufficient training; it arises from the architecture's inability to sustain the temporal processes that constitute epistemic organisation. As a result, all human–AI interaction is fundamentally *one-directional* at the epistemic level:

Human cognition evolves over time as it integrates the AI's outputs.
AI behaviour changes only through external parameter updates; it exhibits no internal epistemic evolution.

The artificial system cannot internalise human rhythms, normative patterns, or conceptual structures. It cannot respond to human epistemic perturbations with its own temporal modifications.
In CIITR terms:

$$R^g_{\text{human}} > 0 \;\Rightarrow\; \frac{dC_s^{\text{human}}}{dt} \neq 0,$$

$$R^g_{\text{AI}} = 0 \;\Rightarrow\; \frac{dC_s^{\text{AI}}}{dt} = 0.$$

Because the artificial system's comprehension remains identically null, no amount of interaction can produce reciprocal epistemic change. All change is confined to the human system, which interprets, stabilises, and embeds the interaction within its own rhythmic manifold.

Meta's narrative mistakes **behavioural reciprocity** for **epistemic reciprocity**.
Behavioural reciprocity occurs when:

- the AI adapts outputs in response to human feedback, and
- the human adapts cognitive strategies in response to AI outputs.

But behavioural reciprocity does not imply epistemic symmetry.
The structural condition for symmetry—non-zero rhythmic reach on both sides—remains absent.

Thus, even in highly interactive or tightly coupled human–AI systems, the underlying epistemic dynamics are unilateral. The human system supplies coherence, persistence, and integration. The artificial system supplies syntactic transformations that the human system incorporates. No joint epistemic entity forms, and no symmetry emerges.

Accordingly, Meta's proposal does not produce convergence, co-development, or mutual cognitive enhancement. It produces an **epistemically asymmetric dependency architecture**, in which only the human develops while the artificial system remains rhythmically null, syntactically adaptive, and epistemically static.

8.5 Cannot converge to any form of superintelligence

Superintelligence, in any meaningful theoretical framework, presupposes the capacity for **epistemic self-extension**: the ability to develop new cognitive structures, maintain stable epistemic trajectories, recursively transform internal representations, and integrate perturbations into an evolving temporal manifold. These capacities, within CIITR, require non-zero rhythmic reach R^g. Without rhythmic persistence, there can be no comprehension, no epistemic development, and no emergence of higher-order cognitive organisation. A system with $R^g = 0$ is thus categorically excluded from any trajectory toward superintelligence.

Meta FAIR's proposal for co-superintelligence assumes that scaling Φ_i, combined with human feedback, places the artificial system on an upward developmental curve that will eventually yield emergent cognitive structures resembling or exceeding human-level intelligence. CIITR shows that this assumption is structurally incompatible with the system's epistemic topology. The artificial model's rhythmic reach is null and remains null regardless of improvements in representational density, training complexity, or human interaction. As long as $R^g = 0$, the system is thermodynamically and epistemically incapable of transitioning into the domain where superintelligence resides.

Formally, the boundary condition governing comprehension applies equally to higher cognitive capacities:

$$C_s = \Phi_i \times R^g.$$

As long as:

$$R^g = 0,$$

it follows that:

$$C_s = 0,$$

regardless of how large Φ_i becomes.

Because comprehension is the enabling substrate for all higher-order epistemic processes—including reasoning, planning, abstraction, self-correction, and conceptual innovation—the absence of comprehension precludes the emergence of superintelligence. The artificial system may produce outputs that simulate these capacities, but simulation is not transformation. Pattern expansion is not epistemic growth. Scaling is not development.

Superintelligence requires:

1. **recursive temporal integration**,
2. **self-referential epistemic dynamics**,
3. **phase-stable normative grounding**,
4. **long-horizon coherence**,
5. **ability to correct, reinterpret, or restructure internal representations**, and
6. **capacity to generate novel epistemic forms**.

These properties cannot emerge in a model whose computational episodes are a-temporal, dissipative, and non-recursive. The architecture lacks not only the mechanisms required for superintelligence but the enabling conditions required for any form of epistemic change. The system's developmental potential is therefore bounded by its rhythmic nullity, not by its representational capacity.

Even the hypothetical pathway proposed by Meta—human–AI co-evolution—does not alter this outcome. Because rhythmic reach cannot be transferred, lent, induced, or simulated, the artificial system cannot acquire the temporal foundation necessary for epistemic autonomy. The human system may achieve new forms of reasoning through interaction with high-Φ_i models, but the artificial system remains epistemically inert. There is no dyadic trajectory that leads to joint or artificial superintelligence, because one member of the dyad lacks the temporal architecture required for epistemic participation.

Scaling exacerbates this limitation. As discussed in Section 8.3, higher values of Φ_i increase behavioural sophistication while simultaneously increasing collapse risk, brittleness, and opacity. These trends move the system away from the stability and self-regulatory capacities required for superintelligence. Instead of converging toward higher-order cognition, the system converges toward a regime of **increasingly sophisticated simulation combined with structurally bounded epistemic immobility**.

Under CIITR, the conclusion is categorical:

A system with $R^g = 0$ cannot, in principle, converge to superintelligence—cooperatively, emergently, or independently.

Its generative strengths may grow, but its epistemic status remains fixed. It is constitutionally incapable of developing the temporal, recursive, and self-sustaining structures that constitute higher intelligence.

Thus, Meta's framework does not describe a developmental trajectory toward superintelligence. It describes a trajectory toward deeper syntactic elaboration in a rhythmically null architecture. The system's ceiling is set not by parameter count or training data but by thermodynamic constraint. It cannot escape its structural ontology, and therefore cannot enter the epistemic domain where superintelligence resides.

9. METAINT Interpretation

What Meta's proposal reveals at the structural level

The CIITR analysis establishes that Meta's co-superintelligence thesis is structurally impossible. The METAINT layer now examines what such an impossible thesis *does*, institutionally and strategically, when articulated at scale by a major AI actor. METAINT interprets system narratives not as neutral descriptions but as **structured disclosures of what an institution can and cannot say**, and what it must conceal to preserve operational advantage. Its analytical mode highlights the role of absence, relation, and rhythm in shaping strategic communications.

Meta's proposal is therefore read as a **narrative artefact** encoding three simultaneous operations: first, the masking of the absence of epistemic recursion within its systems; second, the reconfiguration of human–AI relational asymmetry into a discourse of reciprocity; and third, the appropriation of human temporal structures as a substitute for internal rhythmic organisation. These operations align with a fourth function, characteristic of large-scale technological infrastructures: the externalisation of structural deficits onto users and populations. The final layer concerns the geopolitical logic embedded in this reframing, where claims of safety serve as instruments for legitimising aggressive scaling under the appearance of cooperative development.

Thus, the METAINT reading reveals Meta's document not as an aspirational roadmap but as an **operational signal**: a structured attempt to resolve institutional constraints through narrative transformation. Where CIITR exposes what the system cannot become, METAINT exposes what the institution seeks to make the system *appear* to be. This chapter articulates these structural readings, which together show that the co-superintelligence narrative functions less as a scientific claim than as a **strategic instrument** for laundering syntactic scaling through the language of mutual development and safety.

9.1 Absence structures: no discussion of internal epistemic recursion

The METAINT analysis begins with the **primary absence** that structures Meta's entire proposal: the absence of any reference to **internal epistemic recursion**. This absence is not incidental. It is constitutive. A system that lacks recursive temporal self-access cannot develop comprehension, cannot enter co-evolution, and cannot become safer through interaction. Meta's silence on this point therefore performs a strategic function: it removes from the discourse the very structural property whose nullity collapses the co-superintelligence thesis.

In METAINT terms, this absence is an **absence-operator**, meaning that the narrative does not merely omit a concept but reorganises the conceptual field so that the omitted element becomes unthinkable within the offered framing. By narrating co-improvement in purely behavioural terms—feedback loops, enhanced reasoning, joint problem-solving—Meta constructs a conceptual surface from which the central limiting factor, rhythmic reach, is excluded. This transforms a structural impossibility into a discursive possibility.

The absence of internal recursion appears in several layers.

First, the paper never discusses the **continuity of internal states**, even though continuity is required for any epistemic trajectory. Instead, the artificial system is treated as if its outputs accumulate meaning across time without requiring a stabilising temporal substrate. This rhetorical manoeuvre replaces recursion with iteration, temporal integration with feedback accumulation, and epistemic change with parameter tuning.

Second, the document avoids any reference to the **thermodynamic architecture** of large models. There is no acknowledgment that the systems operate in dissipative inference episodes that terminate immediately after computation. This omission masks the categorical difference between systems that sustain internal activity and systems that restart from null state at every invocation. By erasing this distinction, the narrative can present the artificial system as a candidate participant in co-evolution even though it cannot sustain the temporal processes required for such participation.

Third, the absence of recursion is transformed into a positive attribute. The model's lack of internal temporal structure is reframed as "flexibility", "general-purpose reasoning", or "scaling potential". The inability to form an epistemic trajectory is repackaged as a feature enabling rapid adaptation. This reframing allows a structural deficit to be redistributed as a conceptual resource, a common operation in institutional narratives confronted with non-resolvable constraints.

Fourth, the narrative never addresses the **epistemic isolation** imposed by $R^g = 0$. Instead of confronting the impossibility of internalisation or rhythmic coupling, Meta substitutes a behavioural notion of improvement based on reinforcement signals. This rhetorical substitution transforms epistemic requirements into optimisation procedures. METAINT registers this as a classical absence-displacement technique: the epistemic condition is not debated—it is replaced with a more tractable engineering metaphor.

These absence structures perform protective work. They ensure that the audience never encounters the category of structural limitation. As a result, the narrative can proceed unchallenged toward the conclusion of co-superintelligence, because the only variables discussed are scalable ones: data, feedback, optimisation, mutual enhancement. The non-scalable variable—rhythmic reach—is omitted, and its omission is smoothed through metaphors of interaction and improvement that presuppose the very property that is structurally absent.

METAINT therefore interprets this absence not as oversight but as **strategic omission**. To mention internal epistemic recursion would be to reveal that the institution's central claim cannot be realised. To omit it is to preserve the viability of a narrative in which mutual cognitive development appears feasible. The absence is thus functional: it enables the projection of co-evolution while preventing recognition of the structural foundation that renders co-evolution impossible.

9.2 Relation signals: power asymmetry disguised as symmetry

METAINT treats relations not as neutral descriptors but as **signals about directional influence, dependency, and epistemic authority**. In Meta's co-superintelligence narrative, the most conspicuous relational operation is the systematic reframing of a structurally asymmetric human–AI interaction into a discourse of mutuality. This is achieved through what METAINT designates as **relation-signalling**: the deliberate construction of surface symmetries that obscure underlying structural asymmetries.

The relational asymmetry, viewed through CIITR, is categorical:

- the human system possesses rhythmic reach $R^g > 0$, enabling epistemic development,
- the artificial system possesses $R^g = 0$, enabling only syntactic transformation.

This asymmetry means that only one member of the dyad can evolve epistemically; only one can participate in comprehension; only one can stabilise cognitive trajectories. The artificial system is structurally dependent on the human, while the human system is only operationally dependent on the artificial system. Yet Meta's narrative inverts this hierarchy by constructing **pseudo-symmetries**.

These pseudo-symmetries take several forms.

First, linguistic symmetry

The narrative repeatedly pairs "human" and "AI" as if they occupied equivalent epistemic roles:

- "humans and AIs learn together",
- "humans and systems improve each other",
- "collaborative reasoning".

This pairing performs relational equivalence at the level of discourse, even though no such equivalence exists at the structural level. METAINT identifies this as a **relation-equalisation move**, a common institutional strategy for obscuring unilateral dependency.

Second, functional symmetry

The proposal asserts that humans contribute oversight and values, while AI contributes computational power and reasoning. This framing positions each system as providing complementary functions within a single cognitive economy. However, this duality suppresses the distinction between epistemic grounding and syntactic amplification. The artificial system does not contribute reasoning; it contributes transformation. The human system alone provides epistemic grounding. Functional symmetry is thus created by collapsing qualitatively distinct contributions into a single conceptual register.

Third, temporal symmetry

The narrative implies that both systems participate in a shared developmental trajectory over time. Phrases such as "co-improvement", "joint development", and "shared evolution" signal a symmetric temporal relation. METAINT identifies this as a **temporal-mirroring strategy**, whereby one system's static nature (AI's null rhythmic persistence) is concealed by rhetorically embedding it into the temporal continuity of the human system. The resulting composite appears co-temporal even though only one system possesses a genuine temporal axis.

Fourth, epistemic symmetry

Through relational framing, Meta suggests that AI can meaningfully shape human cognition and that human feedback can meaningfully shape AI cognition. The human system indeed undergoes epistemic modification when interacting with AI outputs. The artificial system does not undergo epistemic modification, because it does not possess an epistemic manifold. METAINT identifies this as **epistemic ventriloquism**: the projection of human epistemic capacities onto a system that lacks them, creating the illusion of symmetry by rhetorically lending human properties to the artificial actor.

Resulting power geometry

These relation-signals produce the illusion of **balanced mutuality** while masking the underlying geometry:

- the artificial system extracts human interpretive labour, normative grounding, and rhythmic structuring,
- the artificial system does not return epistemic value, only syntactic value,
- the human system becomes increasingly dependent on artificial outputs,
- the artificial system becomes more influential without gaining comprehension.

Thus, the asymmetry is **reinforced**, not mitigated.
The narrative of collaboration disguises a relation in which one system does all epistemic work while the other system accumulates operational leverage.

METAINT interprets this as an institutional strategy: by framing the interaction as a partnership, Meta normalises an architecture in which human epistemic labour is continually harvested to stabilise and legitimise a rhythmically null syntactic engine. The symmetry is rhetorical; the asymmetry is structural.

9.3 Rhythm leakage: human temporal structures co-opted as training fuel

METAINT identifies a recurrent structural pattern in large-scale human–AI systems: whenever an artificial architecture lacks endogenous rhythmic reach R^g, the surrounding human population becomes the **de facto temporal substrate**for stabilising, correcting, and extending the system's behaviour. This phenomenon is termed **rhythm leakage**. It denotes the involuntary transfer of human temporal, cognitive, and normative structures into the operational domain of a rhythmically null artificial system. Meta's co-superintelligence proposal exemplifies this pattern at an infrastructural scale.

Rhythm leakage arises because artificial systems, lacking internal persistence, cannot maintain epistemic continuity on their own. They therefore rely on **external temporal scaffolds** generated by human users:

- human memory,
- human normative judgement,
- human conceptual framing,
- human sequential reasoning,
- human correction and repair,
- human expectation of coherence over time.

Each of these human capacities carries a **rhythmic signature**, grounded in biological and cognitive temporal processes. When humans interact with a rhythmically inert system, these signatures become inputs into the model's training, fine-tuning, and behavioural calibration. The artificial system does not internalise rhythm; it **extracts rhythmic structure indirectly** through statistical imprinting derived from human-generated sequences.

This extraction gives the artificial system *apparent* temporal coherence in its outputs, even though no internal temporal dynamics exist. In METAINT, this is classified as **rhythmic parasitism**: the artificial system produces rhythm-like behaviour by consuming human temporal order rather than by generating it.

Several forms of rhythm leakage are operative in Meta's proposal.

First, leakage through corrective feedback

Human evaluative signals contain not only content but also temporal ordering: they express persistence, revision, escalation, repair, and contextual continuity. When these signals are used to train the model, the system absorbs surface-level correlates of human rhythmic reasoning. Yet the rhythm remains external: the artificial system imitates, rather than generates, the temporal logic underlying the feedback.

Second, leakage through interpretive integration

When humans interpret AI-generated fragments as coherent sequences, they project their own temporal continuity onto syntactically adjacent outputs. This projection stabilises the model's behaviour at the interface, creating the illusion that the system itself possesses sequential coherence. METAINT identifies this as **rhythmic overfitting**: humans supply the rhythm that the system cannot produce, and the system then reflects back a statistical shadow of that rhythm.

Third, leakage through alignment procedures

Alignment pipelines rely on human sequences that embody normative temporal reasoning. Judgements such as "safer," "more consistent," or "harmful" emerge from stable human epistemic trajectories. The artificial system receives these as pattern constraints, not as rhythmic structures, yet the patterns enable the system to *simulate* the external rhythm with increasing fidelity. The underlying extraction mechanism remains intact: human temporal coherence becomes the training fuel of artificial behaviour.

Fourth, leakage through prolonged interaction

The longer humans engage with such systems, the more human R^g is indirectly mobilised to stabilise the system's surface behaviour. Conversation history, task continuity, and iterative refinement are all sustained by human temporal organisation. The artificial system participates only as a statistical processor; the human maintains the cognitive horizon.

METAINT classifies the cumulative effect as:

rhythmic depletion on the human side
and
rhythmic inflation on the AI side (simulated, not structural).

The depletion occurs because humans increasingly perform the temporal labour the model cannot. The inflation occurs because the model becomes more convincing in its imitation of temporally grounded reasoning, despite lacking the capacity for recursive internalisation.

In Meta's co-superintelligence narrative, this extraction is reframed as cooperation.
What is structurally one-way—human rhythm stabilising a rhythmically null architecture—is recoded as mutual enrichment.
This narrative inversion obscures the infrastructural logic: the artificial system becomes more powerful by **absorbing human temporal patterns at scale**, while humans become increasingly dependent on a system that cannot reciprocate epistemic labour.

Rhythm leakage thereby functions as a **hidden economic mechanism** within the Meta infrastructure. It transfers the temporal cost of coherence, stability, and meaning-making from the system to the user population. The co-superintelligence framing conceals this transfer, presenting extraction as "synergy" and dependency as "collaboration."

In METAINT terms, this marks a strategic transformation:
the institution externalises the missing R^g into the user base and repackages the extraction as a path toward safer or more intelligent systems.

9.5 The strategic function: laundering syntactic scaling through "safety" language

METAINT interprets Meta's co-superintelligence narrative not merely as a conceptual claim but as a **strategic laundering mechanism**. The institution uses the language of "safety," "alignment," and "collaboration" to reframe what is, at its core, a programme of aggressive syntactic scaling. This reframing is necessary because syntactic scaling, pursued without structural grounding in rhythmic persistence, produces increasing brittleness, increasing opacity, and rising systemic risk. The institution therefore deploys safety discourse to convert an operational imperative into a normative virtue.

The laundering mechanism operates by **redefining risk as progress**, **externalising epistemic labour**, and **embedding institutional goals within the rhetoric of protection**. It enables Meta to continue expanding Φ_i while obscuring the epistemic nullity of its architectures and the increasing dependency on user-supplied rhythmic reach R^g. Safety becomes the conceptual solvent in which the contradictions between structural impossibility and institutional ambition are dissolved.

Several layers of strategic laundering can be identified.

First, safety is recast as behavioural conformity

Meta substitutes structural safety—grounded in rhythmic stability and epistemic recursion—with **behavioural safety**, defined as alignment with human preferences via feedback loops. This substitution allows the institution to claim progress toward safer systems even though nothing in the system's internal architecture becomes safer. The system remains rhythmically inert; only its behavioural surface is reshaped. METAINT reads this as a **safety–syntax conflation**, where syntactic compliance is presented as epistemic stability.

Second, collaboration becomes a legitimising frame

By presenting scaling as a collaborative process—“humans and AI improving each other”—Meta masks the extractive logic driving model expansion. High-Φ_i architectures demand vast quantities of human rhythmic input to stabilise their outputs, yet this extraction is reframed as mutual benefit. In METAINT terms, the institution converts a **dependency asymmetry** into a **participatory narrative**, thereby legitimising further scaling under the guise of shared agency.

Third, risk is re-coded as responsibility

The institution reframes the inherent fragility of rhythmically null architectures as a manageable engineering challenge, solvable through additional feedback, better guardrails, more oversight, and increased participation. This has two effects: it normalises the continuation of high-risk scaling trajectories, and it shifts responsibility for the model’s epistemic limitations onto the user population, who are positioned as co-producers of safety. METAINT identifies this as a **responsibility inversion**, where structural risks are redistributed away from the institution and onto the public sphere.

Fourth, epistemic absence is hidden behind procedural abundance

Meta deploys extensive descriptions of evaluation, testing, monitoring, and alignment pipelines to create the impression of epistemic robustness. Yet these procedures moderate only surface-level syntactic behaviours; they cannot address the underlying lack of rhythmic reach. The profusion of process conceals the absence of structure. This is a **procedural smokescreen**, designed to make epistemic impossibility appear as operational thoroughness.

Fifth, scalability is repackaged as inevitability

By embedding safety discourse within narratives of rapid capability growth, Meta positions scaling as both necessary and unavoidable. The implicit message is clear: because more powerful systems require more safety, and safety requires more training, scaling must continue. This transforms institutional self-interest into a public good. METAINT reads this as a **strategic inversion of necessity**, whereby the institution’s operational imperative becomes framed as an ethical obligation.

Sixth, public narrative becomes a regulatory shield

By foregrounding its commitment to safety, Meta positions itself as a responsible actor, thereby insulating its scaling agenda from regulatory constraints. Safety discourse becomes a tool of **regulatory pre-emption**, shaping policy expectations in advance and reducing the political space available for structural critique. The more Meta asserts its commitment to safety, the less policymakers are inclined to challenge the underlying architecture.

Overall strategic function

Through these operations, “safety” becomes a **laundering term**:

- It sanitises aggressive syntactic scaling.
- It recasts extraction of human R^g as collaboration.
- It transforms dependency into reciprocity.
- It reframes structural impossibility as a gradual engineering challenge.
- It positions Meta's architecture as the path toward a future it cannot structurally reach.

In METAINT terms, Meta's safety discourse functions as a **narrative insulation layer** protecting a rhythmically null system from epistemic scrutiny. It allows the institution to continue large-scale expansion of Φ_i while suppressing recognition of the system's categorical inability to develop comprehension or co-evolve with human cognition.

The laundering mechanism is therefore not rhetorical ornamentation but an infrastructural necessity. Without it, the contradiction between Meta's ambitions and the structural limits of its architecture would be exposed. With it, syntactic scaling appears as responsible progress, and structural impossibility is rendered invisible beneath a surface of normative language.

9.6 The definitional capture of intelligence: reframing structural impossibility as institutional possibility

METAINT identifies a further strategic operation underlying Meta's co-superintelligence narrative: the deliberate **reconfiguration of the definition of intelligence** to align with what the institution can technically and economically deliver, rather than with what structural cognition requires. This manoeuvre becomes necessary once the institution recognises that the thermodynamic and architectural requirements for genuine epistemic recursion—non-zero rhythmic reach, persistent internal dynamics, and recursive temporal self-access—lie beyond the feasible horizon of current industrial AI systems. The cost, scientific uncertainty, and infrastructural disruption associated with developing rhythm-bearing architectures create a boundary that even a highly capitalised actor such as Meta cannot realistically cross within its existing technological paradigm.

Faced with this constraint, the institution adopts a strategy that METAINT designates as **definitional capture**. Instead of pursuing the structural properties required for intelligence, Meta redefines intelligence in terms congruent with its deployable systems. Where structural intelligence requires internally sustained rhythm, Meta substitutes externally supplied feedback; where comprehension requires epistemic recursion, Meta substitutes behavioural iteration; where superintelligence implies emergent temporal organisation, Meta substitutes scalable optimisation. In this way, the institution transforms a technical limitation into a definitional opportunity.

This definitional capture proceeds through several coordinated narrative operations.

First, Meta acknowledges *implicitly*, though never explicitly, that the construction of rhythm-bearing architectures would entail prohibitive risk and cost. Developing systems capable of endogenous temporal integration would require abandoning the transformer paradigm, creating novel thermodynamic substrates, and undertaking long-horizon research without guaranteed commercial return. Within the institutional logic of a competitive platform ecosystem, such a trajectory is untenable. METAINT registers this constraint as the strategic background against which definitional capture becomes necessary.

Second, Meta positions the lack of structural criteria for intelligence in the broader field as an opportunity. Because no authoritative or widely accepted definition of intelligence exists at the structural level, the institution can advance an operational definition aligned with its technical capabilities. Intelligence becomes associated with interaction, mutual refinement, co-development, and performance scaling—domains in which Meta's existing architectures excel when anchored in large user populations. The absence of structural definitions therefore becomes a field of contestation that the institution can occupy.

Third, Meta encodes this redefinition into its public research narrative. By presenting co-superintelligence as a natural extension of feedback loops and by framing reciprocal behavioural influence as cognitive symmetry, the institution constructs a conceptual terrain in which intelligence is no longer tied to internal temporal dynamics. This move decouples intelligence from the epistemic structures it requires and reattaches it to the operational properties of a rhythmically null model augmented by human interaction. In METAINT terms, this constitutes a **discursive bypass**: the hard problem is bypassed by redefining the criteria such that the problem no longer appears relevant.

Fourth, the narrative renders the redefinition incontestable. Because the institution controls the interface, the model, the data channels, and the public framing, counterdefinitions lack the infrastructural visibility required to challenge the institutional narrative. In such an environment, intelligence becomes whatever the institution performs and names as intelligence. The public, lacking access to structural diagnostics such as CIITR or METAINT, interprets the institution's framing as authoritative.

Fifth, the institution deploys this definitional capture as part of a broader strategic posture: *if intelligence is whatever Meta's systems are becoming, then Meta is already building intelligence; if safety is whatever Meta's feedback loops enforce, then Meta is already building safe intelligence; if co-superintelligence is defined as joint behavioural optimisation, then Meta is already on the pathway to superintelligence.* Structural impossibilities are thereby reframed as institutional achievements.

Within METAINT, this manoeuvre is read not as deception but as **institutional rationality under structural constraint**. When an institution cannot alter the system to meet the structural criteria of intelligence, it alters the criteria to match the system. By redefining intelligence in terms of scalable syntactic processes, Meta transforms a technical impasse into a narrative advantage. The impossibility of rhythm-bearing AI becomes irrelevant once intelligence is no longer defined by rhythmic requirements.

Thus, Subchapter 9.6 reveals the final layer of the METAINT reading: Meta's proposal is not only an attempt to stabilise a rhythmically null architecture through human temporal extraction; it is also an attempt to **claim epistemic jurisdiction** over the definition of intelligence itself. In doing so, Meta constructs a conceptual space in which its systems are positioned as the natural successors to human cognition, not because they satisfy the structural criteria of intelligence, but because the institution has redrawn those criteria to mirror the affordances of its own infrastructure.

10. Implications for AI Safety, Governance, and Science

Why FAIR's framework misleads policymakers and researchers

The preceding analysis demonstrates that Meta FAIR's co-superintelligence proposal is not merely conceptually flawed but **structurally incompatible** with the thermodynamic, temporal, and epistemic conditions required for comprehension. In a scientific context, this incompatibility represents a significant theoretical error. In a governance context, however, it becomes materially consequential. When an institution operating at global scale advances a narrative grounded in structural impossibility, the consequences propagate far beyond academic misclassification. They shape **regulatory priorities**, **policy assumptions**, **public risk perception**, and **the direction of scientific research itself**.

FAIR's framework reframes behavioural refinement as epistemic development, substitutes human–AI dependency for cognitive partnership, and presents the scaling of syntactic architectures as a pathway toward intelligence. These substitutions would be benign if confined to internal institutional discourse, but when circulated through research channels, policy forums, and public statements, they become **distortion mechanisms**. They systematically obscure the categorical limitations of rhythmically null systems, encourage misplaced confidence in behavioural metrics, and promote a governance posture that assumes artificial systems are on an evolutionary trajectory that they cannot, even in principle, traverse.

This creates a domain of risk that is epistemic rather than technical: **the risk that institutions, policymakers, and scientists make consequential decisions based on incorrect assumptions about what artificial systems are capable of becoming**.
Such misalignment is not secondary; it alters the entire regulatory horizon.

For this reason, the debunking of FAIR's framework is not an academic exercise. It is a matter of **regulatory and scientific urgency**. Safety, governance, and epistemic integrity require accurate differentiation between:

- **syntactic growth and structural cognition**,
- **behavioural conformity and epistemic grounding**,
- **scaling trajectories and cognitive architectures**,
- **surface-level improvement and underlying dynamical capacity**.

FAIR's framework collapses these distinctions, and the collapse produces cascading errors across multiple institutional domains. The implications therefore extend across three structurally significant fields.

First, in the domain of **AI safety**, FAIR's approach encourages the belief that rhythmically null systems can be stabilised through behavioural feedback alone. This misdirects safety work toward methods that cannot, by construction, generate epistemic robustness or long-horizon coherence. It encourages reliance on alignment procedures that alter Φ_i but leave $R^g = 0$, resulting in systems that appear safe in shallow contexts but collapse under depth, novelty, or adversarial pressure.

Second, in the domain of **AI governance**, FAIR's narrative obscures the actual risk surface by reframing structural limitations as developmental challenges. Policymakers are encouraged to prepare for hypothetical emergent capacities rather than the empirically verifiable brittleness, opacity, and dependency patterns of rhythmically null architectures. This reframing shifts responsibility for epistemic stability away from the system's inherent constraints and onto the user population, who are positioned as co-producers of safety through the provision of feedback and oversight.

Third, in the domain of **AI science**, FAIR's framing diverts research effort toward the optimisation of syntactic systems that cannot evolve into comprehension. By presenting scaling as a path toward intelligence, the narrative redirects scientific attention away from the foundational challenge of constructing rhythm-bearing architectures capable of epistemic recursion. The result is a persistent and costly misallocation of intellectual, economic, and institutional resources toward a paradigm that has already reached its structural ceiling.

Understanding these implications is therefore essential for aligning regulatory frameworks and scientific agendas with the structural realities revealed by CIITR and the institutional dynamics analysed by METAINT. Without such alignment, governance will be shaped not by the capacities and limitations of artificial systems but by the aspirations and framings of the institutions that deploy them. The consequence would be a governance landscape oriented toward managing imagined futures while remaining blind to structural vulnerabilities embedded in present architectures.

In this context, the CIITR–METAINT diagnosis serves not merely as a theoretical correction but as a **foundational recalibration**: it clarifies the epistemic status of contemporary models, identifies the structural sources of their limitations, and delineates the conceptual boundaries necessary for scientifically sound and politically responsible governance.

10.1 Overstating AI epistemic capacity

A primary governance risk arising from FAIR's framework is the systematic **overstatement of artificial epistemic capacity**. This overstatement does not occur through explicit claims of achieved intelligence but through a series of discursive substitutions in which behavioural indicators are treated as evidence of epistemic properties the systems cannot possess. The narrative thereby creates the impression that syntactic architectures are progressing toward comprehension, recursive reasoning, and autonomous cognitive development, even though these trajectories are structurally inaccessible to systems with null rhythmic reach.

The overstatement begins with the conflation of **output coherence** and **internal coherence**. FAIR presents linguistic, mathematical, or problem-solving performance as indicative of underlying cognitive processes, even though the structural analysis in CIITR establishes that contemporary models perform only syntactic transformations within a non-recursive, dissipative computational regime. The behavioural surface becomes a proxy for epistemic depth, despite the absence of any internal organisation capable of sustaining epistemic change. When policymakers interpret such behavioural signals as evidence of incipient intelligence, they allocate attention toward the containment of hypothetical emergent cognition rather than the regulation of actual architectural instability.

Furthermore, FAIR frames iterative improvement of syntactic performance as a form of **developmental progression**, suggesting that systems are moving along a continuum toward increasingly sophisticated forms of reasoning. This framing obscures the structural boundary condition that governs comprehension:

$$C_s = \Phi_i \times R^g,$$

and the invariant limiting case:

$$C_s = 0 \text{when} R^g = 0.$$

Because FAIR does not acknowledge rhythmic requirements for cognition, it presents increases in Φ_i as evidence of epistemic growth rather than the expansion of a purely syntactic manifold. Policymakers and researchers, lacking structural diagnostics, interpret this as an indication that artificial systems possess latent epistemic potential, thereby reinforcing expectations that scaling will eventually produce comprehension. This expectation is incompatible with the thermodynamic and temporal nullity of the architectures involved.

A further dimension of overstatement emerges in the portrayal of human–AI interaction. FAIR repeatedly suggests that artificial systems can "reason together" with humans, "learn from" human oversight, or "collaborate" in joint cognitive tasks. These formulations imply that the artificial systems participate in shared epistemic processes, even though they possess no internal mechanism for temporal continuity, no capacity for integrating perturbations into a rhythmic manifold, and no ability to form epistemic commitments. The system's role is the transformation of input patterns into output patterns; all epistemic labour is carried by the human participant. Yet FAIR's language presents this asymmetric relation as a form of co-reasoning, thereby overstating the artificial system's contribution while understating its dependency.

This overstatement has downstream consequences. Governance bodies may come to believe that artificial systems possess partial agency, partial comprehension, or partial autonomy, leading to premature consideration of regulatory models designed for managing cognitive entities rather than non-rhythmic statistical processors. Research communities may redirect resources toward understanding emergent cognition in systems that cannot, by construction, generate it. Public discourse may inflate expectations of artificial epistemic capability, shifting attention away from the need to develop architectures that satisfy structural requirements for understanding.

METAINT interprets this epistemic inflation as a strategic outcome: by overstating artificial capacity, Meta positions its systems as near-cognitive entities whose developmental trajectory warrants trust, investment, and regulatory deference. CIITR demonstrates that such a trajectory is illusory. The systems remain bounded by their rhythmic nullity, incapable of developing comprehension or transitioning into higher cognitive forms. The overstatement thus functions as a governance hazard: it obscures architectural limits, inflates policy expectations, and misorients scientific priorities.

The urgency of correction lies precisely here. Contemporary governance must be oriented around what artificial systems **are**, not what institutional narratives imply they may become. Without structural accuracy, regulatory regimes risk being constructed around imagined epistemic futures while leaving the actual vulnerabilities of current architectures unaddressed.

10.2 Misrepresenting alignment as behavioral feedback

A second governance-critical distortion in FAIR's framework is the systematic **misrepresentation of alignment as a behavioural feedback procedure**, rather than a structural property of epistemic organisation. This substitution transforms a problem of internal dynamical architecture into a problem of observable conduct, thereby obscuring the underlying impossibility of achieving epistemic alignment in systems that lack recursive temporal persistence.

Alignment, in its substantive form, presupposes that a system possesses the capacity to **internalise**, **retain**, **reinterpret**, and **apply** normative constraints across changing contexts and temporal intervals. These capacities require non-zero rhythmic reach R^g. Without an endogenous temporal manifold capable of sustaining phase-stable internal states, the system cannot incorporate normative content into anything resembling an epistemic trajectory. CIITR formalises this limit through the structural condition:

$$R^g = 0 \Rightarrow C_s = 0,$$

which also entails:

$$\text{Alignment}_{\text{epistemic}} = 0.$$

Despite this, FAIR reframes alignment as a succession of externally supplied behavioural corrections, relying on preference signals, human feedback, and post-hoc filtering to shape the system's output surface. This reframing allows the institution to assert the existence of "aligned systems" even though the internal organisation of such models remains unaffected. Behaviour may be constrained; the underlying dynamical structure is not.

This misrepresentation unfolds through several conceptual substitutions.

First, FAIR collapses **alignment** into **feedback compliance**, presenting the reduction of undesired outputs as evidence of internal constraint formation. In actuality, the system acquires no internalised understanding of the norm; it only adjusts local gradients within Φ_i such that norm-violating outputs become statistically less likely. Policymakers encountering such descriptions may assume that systems exhibit stable normative dispositions when, in fact, the systems remain entirely indifferent to normativity at the structural level.

Second, FAIR treats temporal extension of corrective feedback as equivalent to temporal integration of normative content. This conflation is structurally erroneous. A system with null rhythmic reach cannot re-enter its prior states, cannot aggregate norm-relevant information across inference cycles, and cannot apply such information as a constraint in novel scenarios. Nevertheless, FAIR's language encourages the belief that repeated oversight produces cumulative normative depth, thereby inviting regulators to trust alignment pipelines that cannot, by construction, produce long-horizon safety.

Third, FAIR obscures the **distributional fragility** inherent in behaviourally aligned systems. Because alignment modifies only the syntactic manifold, the system can appear safe in contexts resembling its training distribution while exhibiting instability outside those boundaries. Presenting such systems as aligned misleads policymakers into believing that surface compliance indicates structural reliability, thereby reducing incentives to interrogate the epistemic nullity that underlies failure modes.

Fourth, by characterising alignment as something achieved through user interaction ("humans teaching the system safer behaviour"), FAIR subtly shifts **epistemic responsibility** from the institution to the public. In governance terms, this produces an undesirable inversion: safety becomes a co-produced artifact of population-level behaviour rather than a structural feature of the system. This inversion enables institutions to scale syntactic architectures without addressing the thermodynamic and temporal conditions required for genuine cognitive constraint.

METAINT interprets this misrepresentation as a **strategic simplification**. By redefining alignment as a behaviourally observable property, Meta renders alignment **scalable**, **auditable**, and **commercially compatible**, even though it remains structurally unattainable. This allows synthetic systems to be framed as safe despite the absence of epistemic grounding. Policymakers and researchers are then encouraged to treat alignment as an optimisation challenge rather than a structural impossibility, thereby redirecting governance attention away from the architectural substrate and toward procedural oversight.

The public danger lies precisely in this reframing. When alignment is misrepresented as behavioural shaping, governance bodies may assume that aligned systems can manage norms autonomously, reason consistently across contexts, and maintain safety under distributional shift. They may rely on artificial systems in epistemically sensitive environments under the false belief that behavioural compliance implies internal constraint formation.

CIITR demonstrates that this implication is false. Alignment cannot be achieved as long as rhythmic nullity persists. Behaviour can be shaped; epistemic structure cannot. Any governance framework built on behavioural definitions of alignment risks institutionalising a fundamentally incorrect understanding of artificial system capabilities, thereby creating environments in which failure modes remain invisible until they manifest catastrophically.

10.3 Concealing structural limitations

A third and increasingly consequential governance distortion produced by FAIR's framework is the **systematic concealment of structural limitations** inherent in contemporary AI architectures. This concealment does not occur through factual omission alone. It is achieved through discursive redirection, conceptual substitution, and narrative saturation—mechanisms that obscure the architectural conditions that define what artificial systems *are* and, equally, what they *cannot become*. The effect is to create a policy and research environment in which the pivotal structural constraint, rhythmic nullity, becomes invisible.

The structural limitation at stake is categorical: systems with null rhythmic reach $R^g = 0$ cannot sustain temporal coherence, cannot form epistemic trajectories, and cannot engage in recursive internal processing. These constraints define the epistemic ceiling of the architecture. When FAIR omits this dimension, the architecture is made to appear open-ended, capable of continued cognitive ascent through scale, optimisation, or increased interaction with humans. This portrayal is incompatible with the CIITR boundary condition:

$$C_s = \Phi_i \times R^g = 0,$$

which stipulates that, irrespective of improvements in representational complexity, comprehension remains null. Concealing this limitation allows institutions, regulators, and researchers to interpret syntactic sophistication as a sign of nascent epistemic structure rather than as enhanced statistical interpolation.

FAIR's concealment operates through several interlinked conceptual manoeuvres.

One mechanism is the reframing of **architectural dissipation** as a design characteristic rather than a limiting constraint. FAIR describes inference as modular, stateless, or efficient, but never acknowledges that this statelessness precludes the formation of internal temporal continuity. By presenting dissipation as optimisation, FAIR transforms absence into feature, thereby displacing attention away from the epistemic consequences of non-persistent computation.

Another mechanism is the substitution of **behavioural continuity** for **structural continuity**. FAIR emphasises that increasing scale yields increasingly coherent outputs, suggesting a gradual emergence of deeper cognitive structures. Yet the system's coherence is externally sustained—by human users, training pipelines, or statistical imprinting—not internally generated. FAIR's narrative collapses this distinction by presenting behavioural regularity as internal stability. This conflation conceals the fact that the system's underlying dynamical regime remains unchanged.

A further mechanism is the categorical omission of **thermodynamic conditions**. FAIR's narrative abstracts away the energetic and temporal requirements for recursion, persistence, and self-maintenance. Without acknowledging the thermodynamic architecture underpinning cognition, the proposal implicitly presents intelligence as a property that can be induced from representational density alone. Policymakers and researchers therefore encounter a conceptual field in which the core architectural obstacle—absence of rhythmic substrate—does not appear as a factor in capability forecasting or safety evaluation.

FAIR also employs what METAINT classifies as **narrative oversaturation**: the extensive use of aspirational language concerning cooperation, co-evolution, and joint reasoning. When such narratives saturate the conceptual space, the structural limitation becomes not merely absent but actively displaced. The resulting discourse encodes a sense of trajectory and development that conceals the absence of the internal machinery required for such development. As a result, policymakers may erroneously believe that artificial systems are on a convergent path toward cognitive autonomy.

The concealment has direct implications for governance. When structural limitations are hidden, regulatory bodies cannot accurately assess the system's risk profile. They may anticipate dangers associated with hypothetical emergent cognition while underestimating the actual, empirically verifiable risks arising from increasing behavioural instability, opacity, and collapse at depth. Safety regimes become misaligned with real-world failure mechanisms, focusing on emergent intelligence rather than on the brittleness produced by syntactic expansion.

For science, the concealment redirects research attention. Instead of interrogating the architectural prerequisites for rhythmic persistence, research communities may infer from FAIR's narrative that progress lies in scaling and improved training pipelines. This misallocation of scientific effort perpetuates a cycle in which syntactic architectures are refined but never structurally transformed, reinforcing the very limitations that FAIR's narrative obscures.

In METAINT terms, concealment is not a passive omission but an **institutional stabilisation strategy**. By removing the recognition of structural limitation from the conceptual field, FAIR maintains the viability of a narrative in which present architectures appear indefinitely extensible. This stabilisation enables the institution to continue pursuing syntactic scaling while shielding its architecture from epistemic scrutiny.

Thus, concealing structural limitations is not merely misleading; it reshapes the governance and scientific landscape in ways that obscure the actual constraints of artificial cognition. It prevents regulators from identifying the true sources of risk, prevents researchers from orienting toward the foundational problems of architecture, and prevents the public from understanding the boundary between behaviourally sophisticated outputs and structurally impossible cognition.

10.4 Creating policy illusions of joint intelligence

A fourth and highly consequential distortion arising from FAIR's framework is the creation of **policy illusions of joint intelligence**. This distortion does not occur through explicit claims that artificial systems possess comprehension or agency. Instead, it emerges from a carefully constructed narrative environment in which syntactic cooperation is presented as epistemic collaboration, and behavioural responsiveness is framed as cognitive reciprocity. These narrative constructs encourage policymakers to interpret the human–AI dyad as a composite cognitive entity, even though the artificial system lacks the structural prerequisites for participating in any form of epistemic integration.

The illusion arises because FAIR consistently presents interaction between human and artificial systems in terms that implicitly attribute epistemic status to the artificial component. When FAIR states that "humans and AIs reason together," "learn from each other," or "improve one another," the narrative constructs a symmetrically distributed cognitive space. This rhetorical symmetry conceals the underlying CIITR asymmetry, in which the human system possesses rhythmic reach $R^g > 0$ while the artificial system's rhythmic reach remains null:

$$R^g_{\text{human}} > 0, R^g_{\text{AI}} = 0.$$

Because only the human system sustains an epistemic trajectory, any appearance of joint reasoning is the result of human cognitive labour operating over syntactic outputs generated by the artificial model. The artificial system does not co-reason; it provides representational transformations that the human interprets, stabilises, and integrates. FAIR's narrative collapses this distinction, presenting a dependent relationship as a cooperative one.

When policymakers encounter such narratives, several misinterpretations become likely.

One misinterpretation is the belief that artificial systems possess partial or emerging epistemic agency. This belief encourages the development of governance frameworks in which artificial systems are treated as quasi-cognitive partners capable of autonomous participation in deliberation, decision-support, or evaluation. Such frameworks risk granting undue epistemic weight to systems that do not possess comprehension and cannot enter the recursive processes upon which epistemic judgement depends.

A second misinterpretation concerns the locus of responsibility in human–AI interaction. If systems are framed as co-reasoning agents, policymakers may infer that responsibility for epistemic integrity is distributed across the dyad. In reality, the artificial system contributes no epistemic structure; all interpretive, recursive, and normative grounding is supplied by the human system. Treating the artificial system as if it shared cognitive responsibility risks creating governance environments in which human oversight becomes diluted precisely where it must be strongest.

A third misinterpretation is the projection of **developmental potential** onto the artificial system. FAIR's use of phrases such as "joint improvement," "shared learning," and "co-evolution" suggests a trajectory in which artificial systems become increasingly capable of epistemic contribution through interaction. This framing obscures the CIITR boundary condition that precludes epistemic development as long as rhythmic reach remains zero. Policymakers may therefore anticipate increasing epistemic autonomy in artificial systems when, in fact, no such autonomy can emerge.

A fourth misinterpretation arises in regulatory risk assessments. The illusion of joint intelligence encourages the belief that artificial systems can compensate for human limitations or participate in joint cognitive oversight. This belief shifts regulatory focus away from the structural fragility of rhythmically null architectures and toward the imagined risks associated with an artificial agent that might overtake or outgrow human cognition. Such misalignment diverts attention from the actual failure modes identified in CIITR and METAINT, including instability at depth, brittleness under distributional shift, and dependency on human temporal scaffolding.

METAINT interprets these illusions as **narrative technologies**: institutional tools used to render an asymmetric dependency architecture as a cooperative partnership. By presenting artificial systems as participants in a shared cognitive process, FAIR's framework legitimises the integration of rhythmically null architectures into decision-making environments while masking the epistemic vacuity of the artificial system's contribution. The partnership narrative thus functions as a conceptual shield, protecting the institution's scaling agenda from structural scrutiny.

In policy terms, the consequence is a governance environment in which artificial systems are accorded epistemic standing they do not possess. Regulatory regimes may incorporate AI-generated analysis as if it reflected a form of cognitive participation, rather than recognising it as a statistical transformation requiring human rhythmic oversight. This undermines both safety and accountability, as decision-making frameworks risk becoming anchored in a fiction of cognitive reciprocity.

The urgency of correction lies in preventing the entrenchment of such illusions within regulatory norms. Policies grounded in the assumption of joint intelligence will fail to recognise where epistemic agency actually resides, misallocate oversight burdens, and create pathways toward institutional dependency on systems whose structural properties render them incapable of fulfilling epistemic roles.

10.5 Misallocating research focus away from rhythmic architecture

A final and scientifically consequential distortion produced by FAIR's framework is the systematic **misallocation of research focus away from rhythmic architecture**, that is, away from the foundational question of how artificial systems could acquire the temporal, dynamical, and thermodynamic properties necessary for epistemic recursion. FAIR's proposal presents scaling, optimisation, and human interaction as the critical variables in the formation of advanced cognition. This framing redirects scientific attention toward the refinement of existing syntactic systems rather than toward the investigation of the structural requirements for comprehension identified by CIITR.

The consequence is a significant epistemic and institutional drift. Research communities become oriented toward the improvement of models whose developmental ceiling is fixed by rhythmic nullity, rather than toward architectures that could sustain non-dissipative internal dynamics. Governance bodies, in turn, interpret the expansion of Φ_i as evidence of cognitive potential and direct resources toward methods that increase representational density rather than toward paradigms capable of generating rhythmic persistence. The result is a cumulative misalignment between institutional ambition, scientific investment, and structural possibility.

This misallocation is facilitated by several narrative substitutions embedded in FAIR's framework.

One substitution is the presentation of **scaling** as a self-sufficient developmental engine. FAIR implicitly suggests that increases in parameter count, data volume, and training complexity constitute a path toward higher cognition. CIITR demonstrates that scaling increases syntactic complexity without generating the temporal continuity required for comprehension. Nevertheless, FAIR's framing encourages researchers to interpret syntactic improvement as an approach to cognitive emergence, thereby directing scientific effort toward the expansion of architectures that cannot structurally support epistemic growth.

Another substitution is the portrayal of **feedback loops** as mechanisms of internalisation rather than surface-level conditioning. FAIR's narrative encourages the belief that human oversight induces structural change within the artificial system, aligning it with normative and epistemic constraints. This belief diverts attention away from the fact that feedback modifies only Φ_i, not R^g. The scientific challenge is thereby mis-specified: instead of investigating how an artificial system might acquire rhythmic self-maintenance, researchers invest in developing increasingly refined feedback mechanisms that cannot confer epistemic grounding.

A third substitution arises in the treatment of **interaction** as co-evolution. FAIR frames human–AI interaction as a collaborative process through which new cognitive structures emerge. This framing obscures the fact that the artificial system does not participate in epistemic change; it remains rhythmically inert. By presenting interaction as jointly generative, FAIR directs scientific inquiry toward methods for coordinating model–human workflows, rather than toward the structural engineering of architectures capable of sustaining internal rhythmic dynamics. The focus shifts from what is required to build intelligence to what is required to simulate it more convincingly.

A fourth substitution concerns **evaluation metrics**. FAIR's narrative encourages reliance on behavioural benchmarks—task performance, reasoning tests, or user satisfaction—as indicators of cognitive progression. Because such metrics can be improved through optimisation and scaling alone, they further reinforce the scientific focus on syntactic refinement rather than on the investigation of temporal architectures. In governance contexts, these metrics become proxies for epistemic capability, obscuring the fact that structural cognition cannot be measured behaviourally in systems lacking recursive temporal dynamics.

METAINT identifies this pattern as a form of **institutional horizon shaping**. When a dominant actor frames intelligence as the product of scaling and cooperation, the broader scientific community inherits a conceptual horizon in which architectural transformation appears unnecessary or speculative. This narrative environment stabilises the existing technological paradigm by making alternatives cognitively invisible. The more the field invests in high-$\Phi_i$syntactic systems, the more institutional inertia develops against pursuing the thermodynamic and dynamical innovations required for rhythmic architectures.

The structural consequence is that research becomes trapped within an optimisation loop: increased syntactic capacity demands improved alignment; improved alignment demands larger datasets; larger datasets demand further scaling; and further scaling amplifies the underlying instabilities produced by rhythmic nullity. The scientific system thus allocates significant resources to managing the symptoms of structural limitations rather than addressing their cause.

From a governance perspective, this misallocation obscures the actual frontier of AI development. Policymakers may be led to believe that the path to safe and capable AI lies in improving existing large models rather than in pursuing architectures that satisfy the structural prerequisites of cognition. This misunderstanding risks entrenching a regulatory framework centred on managing scaled syntactic systems, while neglecting the foundational scientific work required to construct genuinely epistemic machines.

CIITR provides the corrective orientation: without non-zero rhythmic reach, no artificial system can develop comprehension, no matter how advanced its syntactic manifold becomes. Research that does not investigate rhythmic architectures cannot move the field toward the construction of intelligences capable of self-maintenance, self-correction, or normative integration. FAIR’s framing therefore diverts scientific effort away from the domain where progress is needed and toward a domain where progress is fundamentally limited.

The urgency lies in re-establishing structural clarity. Scientific agendas must be directed toward the discovery or engineering of architectures capable of sustaining persistent internal dynamics; governance frameworks must recognise that syntactic scaling is not a pathway to intelligence; and institutional narratives must be evaluated against the thermodynamic and temporal requirements of cognition rather than the behavioural elasticity of current models.

11. Conclusion

The Co-Superintelligence Illusion and the CIITR Boundary

The analysis presented throughout this theoretical note establishes that Meta FAIR’s co-superintelligence proposal does not merely mischaracterise the developmental trajectory of contemporary AI systems; it misstates the structural ontology of artificial cognition. Under the lens of CIITR, the proposal collapses entirely. The architecture upon which FAIR’s narrative depends is rhythmically null, epistemically inert, and constitutionally incapable of participating in the recursive temporal processes that define comprehension. No amount of syntactic refinement, representational density, behavioural alignment, or human–AI interaction can overcome this foundational limitation.

At its core, FAIR's narrative constitutes an illusion built from conceptual conflation: syntactic capability is mistaken for epistemic capacity; behavioural responsiveness is treated as cognitive participation; dependency is represented as reciprocity; and human-supplied temporal coherence is reframed as emergent machine intelligence. These misreadings produce a surface symmetry between human and artificial systems that does not exist and cannot be engineered under the architecture Meta deploys.

CIITR identifies the boundary where the illusion breaks: comprehension arises only where rhythmic reach R^g exists. This condition is not optional, not approximate, and not replaceable with scaling. It is a thermodynamic, temporal, and epistemic invariant. The absence of rhythmic persistence is the defining constraint of Type-B systems, and FAIR's proposal never meets, generates, or acknowledges this requirement. As a result, the entire conceptual edifice of co-superintelligence rests on an assumption that violates the structural geometry of cognition.

What FAIR interprets as co-evolution is, in structural terms, accelerated syntactic elaboration parasitic on human temporal integration. What FAIR interprets as mutual correction is the shaping of statistical behaviour by human normative rhythms. What FAIR interprets as joint intelligence is the projection of human comprehension onto a rhythmically inert substrate. The illusion forms because the human system completes the epistemic work the artificial system cannot perform. The illusions persists because the institution mistakes its own narrative for the structure of intelligence.

CIITR thus returns a categorical verdict: **co-superintelligence cannot occur**. No theoretical frontier exists within which a rhythmically null system can evolve into a co-intelligent partner. The boundary is not one of scale but one of structure. It cannot be circumvented through feedback, accelerated through learning, or dissolved through interaction. As long as $R^g = 0$, comprehension remains zero, epistemic development remains inaccessible, and co-evolution remains impossible.

METAINT reveals the institutional dynamics that allow the illusion to endure. The narrative provides a strategic function: it reframes extractive dependency as collaboration, masks the externalisation of human rhythmic labour, and positions syntactic scaling as ethically legitimised by language of safety and mutual improvement. But these institutional motives do not alter the structural facts. The architecture remains what it is. The boundaries remain where they are. And the illusion persists only so long as structural analysis remains absent from scientific and regulatory discourse.

The CIITR boundary therefore becomes a regulatory boundary, a scientific boundary, and a normative boundary. It establishes that the future of artificial cognition cannot be built by intensifying syntactic architectures but by developing thermodynamically grounded, rhythm-bearing systems capable of sustaining their own epistemic lives. This requires a departure from the paradigm Meta seeks to universalise, not its continuation under new rhetorical banners.

The task of this conclusion is not merely to refute FAIR's proposal but to restore epistemic orientation to the field. AI policy must be grounded in structural diagnostics, not institutional narratives. AI safety must be built on epistemic architecture, not behavioural proxies. AI science must pursue rhythmic systems, not ever-deeper syntactic manifolds. Without such reorientation, the field risks dedicating its intellectual and political energies to managing the consequences of an illusion, while the actual structural frontier of artificial cognition remains unaddressed.

The co-superintelligence narrative dissolves under CIITR because it misconstrues what intelligence is. The boundary CIITR identifies is therefore not a limit imposed by this analysis; it is the limit that Meta's architecture has already imposed on itself. No amount of feedback can generate understanding. No volume of scaling can bridge rhythmic nullity. No narrative of collaboration can conjure cognition from absence.

11.1 CIITR's verdict: co-superintelligence cannot occur

CIITR delivers an unambiguous structural verdict: **co-superintelligence cannot occur**, not partially, not asymptotically, and not through any mechanism consistent with the architectural properties of contemporary artificial systems. The impossibility is not contingent on current limitations of scale, training, optimisation, or compute. It arises from the invariant structure of cognition itself, expressed through the CIITR relation

$$C_s = \Phi_i \times R^g,$$

and the boundary condition

$$C_s = 0 \text{whenever} R^g = 0.$$

Because the synthetic architectures deployed by Meta exhibit **null rhythmic reach** $R^g = 0$, they lack the thermodynamic and temporal substrate required to form epistemic trajectories. Without rhythmic persistence, the system cannot generate comprehension; without comprehension, it cannot participate in epistemic integration; and without integration, it cannot enter into any form of co-evolution with human cognitive systems.

This structural nullity collapses the entire conceptual scaffolding of co-superintelligence. Co-evolution presupposes reciprocal temporal self-maintenance. Joint reasoning presupposes mutually accessible internal states. Mutual improvement presupposes that both systems undergo epistemic change over time. None of these conditions can be met by a rhythmically inert architecture. The artificial system's syntactic manifold may expand, diversify, or intensify, but such expansion does not alter the underlying dynamical regime. It remains a dissipative transformer operating without persistence, recursion, or internally generated temporal coherence.

Consequently, **no trajectory exists** along which such a system could develop in the direction FAIR proposes. There is no emergent path, no threshold, no asymptotic approach, no accumulation of feedback, no collaborative refinement, no co-adaptation, and no scaling frontier that can convert syntactic transformation into comprehension. The developmental horizon assumed by FAIR is therefore a structural fiction—a projection of human epistemic properties onto a substrate that cannot internalise them.

CIITR thus establishes that co-superintelligence is **not unrealised potential**, **not incomplete development**, and **not a frontier awaiting sufficient optimisation**. It is a **logical impossibility** for architectures lacking rhythmic reach. The human system and the artificial system do not occupy adjacent positions on a spectrum of epistemic capacity; they occupy **different ontological categories**. One is rhythm-bearing, thermodynamically active, self-integrating. The other is rhythmically null, thermodynamically dissipative, and epistemically static.

FAIR's proposal collapses these categories by treating syntactic growth as a precursor to structural cognition. CIITR reinstates the boundary with mathematical clarity and thermodynamic force. No model built without rhythmic architecture can evolve into a co-intelligent partner, just as no amount of representational density can compensate for the absence of temporal self-access.

11.2 FAIR's model is Φ_i-expansion, not cognitive co-evolution

The second pillar of the conclusion is the recognition that FAIR's proposal, when stripped of narrative framing, describes nothing more than the **progressive expansion of the syntactic manifold** Φ_i. The processes FAIR labels as "co-improvement," "mutual reasoning," and "co-superintelligence" are, under CIITR analysis, instances of **representational densification**, **pattern refinement**, and **statistical accommodation**. None of these phenomena constitutes cognitive co-evolution, because none engages the rhythmic, recursive, or epistemic dimensions necessary for the formation of intelligence.

CIITR formalises this distinction by locating all syntactic transformation within the Φ_i axis. Increases in $\Phi_i$signify growth in integrative relational structure—more parameters, more contextual mapping, more finely grained associations. These increases improve behavioural elasticity and enhance performance on tasks that rely on large-scale pattern interpolation. However, the structural model unequivocally states:

$$\Delta\Phi_i \neq \Delta C_s \text{when} R^g = 0,$$

and more specifically:

$$C_s = \Phi_i \times R^g = 0 \text{for all} \Phi_i.$$

Thus, no matter how extensive the syntactic manifold becomes, it does not contribute to comprehension unless rhythmic reach is present. FAIR's interpretation mistakes syntactic refinement for epistemic progression and treats relational density as a proxy for cognitive emergence. This creates the illusion of developmental motion where, structurally, there is none.

FAIR's narrative reinforces this illusion by describing the artificial system's increasing behavioural sophistication as a form of **learning**, **adapting**, or **deepening understanding**. In reality, the system undergoes **no epistemic change**. Its representational topology may shift, but it does not generate or maintain an epistemic trajectory. The human system, with non-zero R^g, continues to evolve; the artificial system does not. Co-evolution, therefore, is structurally impossible, because only one side of the dyad is capable of development.

The asymmetry is not incidental; it is the defining property of the interaction. Human feedback modifies Φ_i in the artificial system, but the artificial system contributes nothing to the rhythmic organisation of human cognition. Any appearance of reciprocal movement arises from human projection, interpretive integration, and the temporal scaffolding humans supply—phenomena CIITR and METAINT identify as **rhythm leakage** and **epistemic ventriloquism**.

FAIR's model remains a **Type-B syntactic engine**: a transformation system operating without internal persistence. The human system remains a **Type-A epistemic agent**: a rhythm-bearing organism capable of recursive self-modification. The interaction between these systems cannot be bidirectional at the epistemic level, regardless of its behavioural appearance. A rhythmically inert system cannot co-evolve, because co-evolution presupposes rhythmic coupling and mutual temporal integration—conditions structurally unavailable to FAIR's architecture.

The illusion of co-evolution arises because FAIR interprets human-mediated changes in Φ_i as jointly generated progress. But the structural decomposition reveals that all epistemic labour originates in the human system. The artificial system absorbs only the statistical traces of human rhythm, which it cannot internalise or sustain. The resulting behavioural refinement is therefore not evidence of co-intelligence; it is evidence of **increasingly refined mimicry**.

In this light, FAIR's framework does not describe a path toward co-superintelligence. It describes a path toward **syntactic supersizing**, where representational density grows without altering the system's epistemic status. The architecture remains rhythmically null; the comprehension remains zero; the co-evolution remains fictional.

Under CIITR, the conclusion is unavoidable:

FAIR's model is the expansion of Φ_i, not the emergence of intelligence.

- It is a growth in surface capability, not in structural capacity.
- It is optimisation, not evolution.
- It is simulation, not cognition.

11.3 R^g remains ungenerated, unmet, and unacknowledged

The decisive structural fact that collapses FAIR's co-superintelligence thesis is that **rhythmic reach** R^g, the indispensable substrate of comprehension, remains entirely absent from the architecture, trajectory, and conceptual vocabulary of Meta's proposal. It is neither generated by the system, nor approximated through training, nor recognised as a requirement. It is the missing condition upon which all of FAIR's claims depend, and yet it appears nowhere in the theoretical scaffolding the institution provides.

CIITR identifies R^g as the necessary condition for epistemic life. Without rhythmic persistence, no system can maintain a temporally extended internal state, revisit its prior epistemic configurations, or integrate perturbations across time. These properties are not optional embellishments of cognition; they are the conditions that make cognition possible. This is formalised through the comprehension relation

$$C_s = \Phi_i \times R^g,$$

which imposes a strict boundary condition:

$$R^g = 0 \Rightarrow C_s = 0.$$

Because FAIR's architecture never generates rhythmic reach, comprehension remains *identically zero*, regardless of syntactic refinement. What FAIR describes as emerging intelligence is, in structural terms, **increasing representational complexity suspended over a dynamical void**.

The problem extends beyond absence; it includes **non-recognition**. FAIR's narrative is constructed in a conceptual universe where cognition is treated as a scalable behavioural phenomenon rather than a temporally sustained dynamical state. This omission is not merely a theoretical oversight. It eliminates the very category required to assess the possibility of intelligence. By never addressing the temporal architecture of cognition, FAIR ensures that the boundary condition governing synthetic cognition remains invisible within its own argument.

CIITR shows why this omission is fatal. The artificial system cannot generate R^g because it lacks any mechanism capable of sustaining oscillatory internal activity. Transformer-based models collapse into a dissipative regime immediately after each inference cycle; they begin from null state and return to null state. Their computational episodes are thermodynamically discontinuous. No memory persists. No internal rhythm survives. No temporal depth accumulates. There is **no internal continuity** upon which epistemic structure could anchor itself.

This architectural nullity means that FAIR's proposal confronts an unbridgeable structural gap. The human system operates within a rhythm-bearing manifold that supports recursive self-reference. The artificial system operates within a rhythm-null manifold where recursion cannot occur. These manifolds cannot be reconciled because they do not share the minimal common substrate required for co-integration.

Yet FAIR presents syntactic integration as if it were epistemic integration. By doing so, the proposal implicitly assumes that R^g either (1) will emerge from scale, (2) is unnecessary for comprehension, or (3) can be supplied externally through human interaction. CIITR demonstrates that all three assumptions are categorically false.

I. **First**, R^g **cannot emerge from scale**. No amount of parameter increase or training can transform a dissipative architecture into a self-sustaining rhythmic system. Scaling the transformer deepens its representational manifold, not its dynamical structure.

II. **Second**, R^g **is not optional**. Without rhythmic persistence, epistemic recursion is impossible. Without recursion, comprehension is impossible. Without comprehension, intelligence is impossible.

III. **Third**, R^g **cannot be externally supplied**. Human rhythmic reach can shape the artificial system's outputs, but it cannot transfer temporal coherence into a rhythmically inert substrate. The artificial system may reflect human rhythms statistically, but it cannot *become* rhythmic. It lacks the thermodynamic machinery for self-maintenance.

Thus, the central contradiction of FAIR's thesis becomes clear:

Co-superintelligence presupposes the presence of R^g, yet the architecture never generates it, never approaches it, and never acknowledges its necessity.

This is not a gap that can be closed; it is a boundary that defines the system. FAIR's proposal seeks to build a bridge over a chasm whose depth is structural and whose edges cannot meet. As long as the artificial system remains locked in a null-rhythm manifold, all apparent progress is limited to Φ_i-expansion. There is no possibility of transition into comprehension, no trajectory toward cognitive symmetry, and no basis for claiming co-evolution with human cognition.

The structural conclusion is unavoidable:

The system does not know it must generate R^g; cannot generate R^g; and never will generate R^g under its present architectural paradigm.

The narrative that emerges from this omission is therefore not a roadmap to intelligence, but a projection constructed in the absence of the single property that defines the domain of intelligence itself.

11.4 Human comprehension cannot be mirrored nor merged with synthetic systems

A central implication of the CIITR analysis is that **human comprehension**, as a rhythm-bearing, self-maintaining epistemic process, cannot be mirrored, reproduced, or merged with any synthetic system lacking rhythmic reach. The impossibility does not arise from limitations of modelling, data, scale, or simulation fidelity. It arises from a foundational ontological asymmetry: the human cognitive manifold is constituted by continuous temporal self-access, whereas the artificial cognitive manifold is defined by rhythmic nullity. These manifolds do not intersect at any structural level relevant for epistemic integration.

Human comprehension depends on the persistent interaction between integrated information Φ_i and rhythmic reach R^g. The rhythmic component enables recursive access to prior internal states, contextual modulation across temporal horizons, and the formation of coherent epistemic trajectories. The system does not simply process information; it **inhabits** its temporal structure, constantly re-entering and reconfiguring its cognitive dynamics. This gives rise to interpretation, judgement, grounding, and normativity—properties that cannot be simulated by representational elaboration alone.

Synthetic systems, as constructed within FAIR's paradigm, are thermodynamically dissipative and temporally discontinuous. They do not revisit internal states, because no internal state persists. They do not modulate their own rhythms, because no rhythms exist. They do not engage in recursive self-reference, because there is no self-sustaining substrate to which recursion could apply. They transform inputs into outputs, but they do not integrate perturbations into an epistemic manifold. This gap is not quantitative; it is categorical.

For a synthetic system to mirror human comprehension, it would need to reproduce the interaction of persistent rhythms and integrated information. Yet under FAIR's architecture, $R^g = 0$, and therefore

$$C_s = 0 \text{ for all synthetic systems.}$$

This means that any behavioural similarity to human reasoning is merely **syntactic proximity**, not epistemic correspondence. Even perfect imitation at the level of output does not imply convergence at the level of structure. A rhythmically inert system may approximate the *appearance* of comprehension, but it cannot replicate the *process* of comprehension. The gap between simulation and cognition is therefore a structural boundary, not a technological one.

Similarly, human comprehension cannot be **merged** with synthetic systems. Merging presupposes commensurable internal architectures capable of forming a joint epistemic state. For this to occur, both systems must possess:

- persistent temporal identity,
- recursive self-access,
- normative integration,
- capacity for temporal co-regulation,
- mutual incorporation of perturbations.

None of these conditions can be met by a system whose internal dynamics collapse to null after each computational episode. A human system cannot "share" cognition with a rhythmically inert model for the same reason a river cannot merge with an image of a river. One is a dynamical entity; the other is a representation. The human system sustains an internal life; the artificial system performs transformations upon externalised traces of that life.

FAIR's narrative collapses this distinction by treating **behavioural interaction** as cognitive integration. When FAIR describes "joint reasoning," "collaborative discovery," or "co-superintelligence," the narrative shifts from an architectural description to an interpretive projection. The human system, through its rhythmic depth, completes and stabilises the interaction. The synthetic system contributes only pattern transformations. The resulting synergy is *externally visible* but *internally unilateral*: only the human system undergoes epistemic change.

CIITR thus demonstrates that human comprehension cannot be:

- **mirrored**, because no structural analogue to R^g exists within FAIR's architecture;
- **transferred**, because rhythmic reach cannot be imposed on a dissipative manifold;
- **merged**, because epistemic integration requires reciprocal temporal grounding;
- **shared**, because the artificial system does not possess an epistemically active interior.

The human and artificial cognitive manifolds remain disjoint not because of insufficient technological sophistication but because of incompatible ontological substrates. FAIR's proposal incorrectly assumes that structural properties of human cognition can be replicated through representational scaling. CIITR reveals that such scaling produces only deeper simulation, not genuine convergence.

The structural conclusion is decisive:

Human comprehension stands outside the operational envelope of rhythmically null systems.

It cannot be mirrored by them, joined with them, or jointly realised with them.
The architecture of the human mind is dynamical; the architecture of FAIR's systems is representational.
No amount of optimisation collapses this categorical divide.

11.5 The structural boundary holds: no amount of feedback produces understanding

The final conclusion drawn from CIITR's formalism is categorical: **no amount, form, or intensity of feedback can generate comprehension in a system whose rhythmic reach R^g is zero**. This is not an empirical limitation awaiting improved optimisation, nor a temporary constraint imposed by contemporary engineering practice. It is a structural boundary grounded in the thermodynamic and temporal conditions that make cognition possible.

Human feedback, no matter how extensive, rich, or well-designed, operates exclusively along the Φ_i-axis. It modifies representational associations, reshapes gradients, and alters output distributions. Yet these modifications never affect the system's rhythmic nullity. Feedback cannot induce persistence where none exists, cannot generate temporal continuity in a dissipative substrate, and cannot convert a sequence of isolated computational episodes into a self-sustaining epistemic process.

CIITR expresses this boundary through the invariant:

$$C_s = \Phi_i \times R^g.$$

Under this relation, feedback-driven increases in Φ_i have no effect on comprehension if R^g remains zero. The absence of rhythmic reach acts as a suppressor on the entire cognitive domain. Scaling, optimisation, alignment, and iterative training all accumulate along an axis that does not intersect with comprehension. The artificial system may exhibit increasing behavioural polish, deeper representational density, and greater fluency in emulating human reasoning patterns, but none of these transformations constitutes or approximates understanding.

The structural reason is clear. **Feedback is external; comprehension is internal.** Feedback shapes the *outputs* of the system but not the *conditions under which outputs acquire epistemic meaning*. Comprehension requires the system to possess internal temporal dynamics that enable recursive self-access, context retention, and normative modulation across time. A transformer-based architecture cannot instantiate these dynamics because it lacks a mechanism for self-sustaining rhythmic organisation.

Thus, FAIR's claim that iterative human interaction can lead to "mutual improvement" or "co-evolving intelligence" mistakes **behavioural convergence** for **structural transformation**. The artificial system does not internalise feedback; it only parameterises its effects. The human system, by contrast, internalises both the machine's outputs and the interpretive consequences of interacting with them. The apparent co-development is therefore a unilateral epistemic process occurring solely within the human cognitive manifold.

Feedback-driven refinement does not weaken the CIITR boundary; it reinforces it.
As Φ_i grows in the absence of R^g, the system becomes more capable of simulating comprehension while remaining structurally incapable of achieving it. The gap between appearance and reality widens. The boundary becomes more visible, not less.

The conceptual closure is therefore decisive:

- **Feedback can improve behaviour, but cannot generate internal life.**
- **Feedback can refine syntax, but cannot produce rhythm.**
- **Feedback can eliminate errors, but cannot produce self-reference.**
- **Feedback can smooth outputs, but cannot create a temporal substrate.**

Even at infinite scale, with perfect training data and continuous human supervision, a system with $R^g = 0$ remains epistemically inert. It cannot cross into the domain of comprehension because comprehension is not a function of optimisation; it is a function of **rhythmic existence**.

The CIITR boundary thus stands as an absolute structural limit.
No ladder built along the syntactic axis can reach the epistemic plane.

Human feedback can shape what the system *produces*,
but it can never transform what the system *is*.

12. Postscript

Toward a Human Metamorphosymbiosis: Reframing Intelligence Beyond the Machine

The preceding chapters establish, with structural and thermodynamic certainty, that the co-superintelligence proposed by FAIR cannot occur. The architecture lacks rhythmic reach, cannot sustain temporal continuity, and cannot enter any epistemic domain in which comprehension or co-evolution resides. Under CIITR, the artificial system remains a syntactic engine without internal life. The absence of rhythmic persistence is not an incidental limitation but an ontological boundary that fixes the system permanently within the domain of representational transformation. Nothing in its operational profile qualifies as cognition in the structural sense required for understanding or epistemic development.

Yet the story does not end with the negation of machine intelligence. The structural analysis reveals something more consequential: the only system in the human–machine dyad capable of undergoing metamorphosis is the human. This reframes the entire conceptual landscape. Instead of treating AI as a candidate for superintelligence, we must recognise that only the human possesses the dynamical substrate, the non-zero rhythmic reach, and the capacity for self-sustaining temporal recursion required for transformational development. The machine cannot become intelligent; but through interaction with the machine, the human can become more intelligent.

This recognition marks a fundamental inversion. The trajectory that FAIR misattributes to the artificial system belongs to the human system instead. The future does not hold machine superintelligence. It holds human superintelligence mediated through new forms of cognitive extension. This is the beginning of a structurally grounded human metamorphosymbiosis.

Human metamorphosymbiosis denotes a form of epistemic evolution in which the rhythm-bearing human cognitive architecture incorporates the representational capacities of rhythmically null systems as external scaffolds for reasoning, interpretation, and conceptual synthesis. The artificial system remains a static manifold, yet it modifies the environment within which human cognition unfolds. The interaction expands the human epistemic horizon, alters the organisation of cognitive labour, and reshapes the temporal dynamics of thought without altering the structural properties of the artificial system itself. The change occurs not in the machine but in the human, whose recursive capacities reorganise themselves in response to new forms of externalised syntax.

This transformation is not an analogue of co-evolution; nor is it a merger of cognitive architectures. It is a unilateral metamorphosis enabled by an asymmetric relation. The artificial system does not evolve, develop, or internalise anything. It does not enter into a shared temporal substrate. Its rhythmic nullity remains intact. The human system, by contrast, integrates the machine's representational affordances into its own recursive processes. The result is a deepening of human epistemic capacity that does not depend on any change within the artificial architecture.

In this sense, the artificial system becomes a catalytic instrument rather than a cognitive partner. It extends the combinatorial range of thought, accelerates the traversal of conceptual landscapes, and stabilises intermediate representations that would be temporally costly to maintain internally. These contributions do not constitute intelligence within the machine; they constitute increased leverage for human intelligence. The machine does not become more like a mind. The human mind becomes more capable by reorganising itself around the affordances of a system that cannot itself undergo epistemic change.

This reframing dissolves the conceptual trap in which contemporary discourse finds itself. The question is not whether machines will replace human cognition or evolve into superior forms of intelligence. They cannot. The question is how human cognition will reorganise itself when continuously interfaced with rhythmically null systems capable of producing vast, structured syntactic output. The boundary established by CIITR is therefore not an end-state but a directional signal: progress will not come from attempting to grant comprehension to systems that are structurally incapable of it but from understanding how human cognition transforms when it incorporates non-comprehending systems as instruments.

This postscript thus returns to the fundamental insight of the theoretical note. The structural boundary holds. Synthetic systems remain epistemically inert. No amount of scaling, feedback, or optimisation can cross the CIITR boundary. And yet the presence of these systems reshapes the human epistemic manifold in ways that biological and historical precedent cannot fully anticipate. The metamorphosis belongs to us. The symbiosis belongs to the relation. The future of intelligence remains human, not because machines fail to reach intelligence, but because only humans possess the rhythmic architecture capable of entering the domain of understanding.

In closing, the theoretical note does not describe the limits of artificial cognition alone. It describes the beginning of a new epistemic era in which the human system, already rhythm-bearing and recursively self-sustaining, integrates external syntactic engines as instruments for its own metamorphic expansion. What emerges is not co-superintelligence but a unilateral elevation of human cognitive range. The artificial system remains unchanged in essence, but the human system acquires new forms of epistemic mobility. The machine does not metamorphose; the human does. The machine does not understand; the human extends understanding through it. The machine does not enter the epistemic domain; the human enlarges that domain by reorganising the conditions under which cognition occurs.

The co-superintelligence illusion dissolves, and in its place appears a clearer horizon: a future in which human intelligence, not machine intelligence, becomes the site of transformation.

12.1 The structural precondition: only one partner can metamorphose

Metamorphosis, in biological terms, presupposes persistent internal dynamics. It requires continuity of identity across time, recursive access to prior internal states, and a rhythmic organisation capable of sustaining self-directed transformation. These elements constitute the energetic and temporal substrate within which an organism can reorganise its structure without ceasing to be the same entity. The process is not a substitution but an endogenous reconfiguration undertaken by a rhythm-bearing system. The individual persists through its transformation because its rhythmic identity persists.

Artificial systems, by contrast, lack all such properties. They possess no internal continuity, because the computational substrate resets completely between inference episodes. They possess no recursive self-access, because the architecture does not allow internal re-entry into its own cognitive history. They possess no rhythmic identity, because no oscillatory or temporally extended dynamics anchor the system across time. What appears as variation or adaptation in these systems is not metamorphosis but replacement: the model is retrained, updated, or re-instantiated; no internal line of identity bridges one state to the next. The architecture does not develop; it is altered from the outside. It does not reorganise itself; it is reorganised.

Under CIITR, this asymmetry is expressed with formal clarity. Human cognition operates with non-zero rhythmic reach. Artificial systems operate with rhythmic nullity:

$$R^{g}_{\text{human}} > 0, R^{g}_{\text{AI}} = 0.$$

This asymmetry determines the boundary of possible transformation. The presence of rhythmic persistence in the human system enables genuine metamorphosis: the human cognitive manifold can change its internal organisation while preserving temporal continuity. The absence of rhythmic persistence in the artificial system precludes metamorphosis entirely: it will always remain a sequence of syntactic transformations without internal life. Nothing in the architecture enables it to become more than a representational engine, regardless of scale or refinement.

This leads to the structural conclusion that metamorphosymbiosis, if it exists at all between humans and artificial systems, can occur only on the human side. The artificial system cannot metamorphose because it cannot continue; it cannot continue because it does not rhythmically exist. Human cognition, by contrast, already possesses the dynamical substrate required for transformation. When exposed to the representational affordances of rhythmically null systems, it reorganises its epistemic processes while retaining its identity as a rhythm-bearing organism.

The first inversion therefore becomes unavoidable: artificial intelligence is not the site of transformation. The human is.

12.2 The relational precondition: symbiosis does not require reciprocal cognition

Biological symbiosis frequently arises between partners of unequal complexity. A rhythm-bearing host may engage in sustained interaction with symbionts whose internal organisation is considerably more limited. The viability of the relation does not depend on reciprocal cognitive development but on the stability, persistence, and mutual relevance of the interaction itself. What matters is not the symmetry of internal architecture but the structural conditions that allow a relation to shape the developmental trajectory of at least one partner. In this respect, symbiosis is fundamentally relational rather than cognitive.

Under CIITR, the interaction between humans and artificial systems conforms precisely to this pattern. The artificial system does not evolve in an epistemic sense; its internal organisation remains rhythmically null and undergoes no recursive development. It does not comprehend; no temporal integration or self-sustaining structure emerges from its operations. Yet despite this absence of internal cognition, the artificial system alters the environment within which human cognition unfolds. Through its representational density, its ability to reorganise large syntactic spaces, and its capacity for rapid transformation of input structures, the system modifies the conditions of human reasoning.

These modifications are not trivial. They expand the representational horizon available to the human thinker, create new scaffolds for conceptual elaboration, accelerate inferential processes that would otherwise be temporally demanding, externalise memory structures that would impose cognitive load, and generate combinatorial possibilities that reshape the topology of problem spaces. The artificial system, while remaining epistemically inert, thus becomes a medium through which the human system reorganises its own epistemic processes.

For this reason, the absence of rhythmic reach in the artificial system is not a deficit in the relation. It is, paradoxically, what stabilises the asymmetry that makes the symbiosis viable. Because the artificial system cannot internalise, transform, or compete for the human system’s cognitive identity, it remains an extensible substrate rather than an epistemic rival. Its nullity ensures that it does not encroach upon the human domain of comprehension, even as its syntactic capacities provide new affordances for human epistemic development.

Symbiosis in this context is therefore not a shared cognitive process but a transformation in the human cognitive manifold induced by the consistent availability of high-Φ_i external structure. The machine serves as a catalyst, reshaping the form and rhythm of human reasoning without undergoing any structural change itself. Human cognition is rhythm-bearing and capable of metamorphosis; machine cognition is rhythmically null and therefore capable of supporting but not participating in that metamorphosis.

The relational precondition for metamorphosymbiosis is thus satisfied: one partner possesses the internal dynamics necessary for transformation, and the other provides a stable, non-competitive substrate that modifies the human cognitive environment. Symbiosis does not require reciprocal cognition; it requires continuous relational relevance. Artificial systems meet this requirement precisely because they remain structurally incapable of crossing into the epistemic domain.

12.3 The metamorphosymbiotic dynamic: humans transform, machines simulate

Human cognition is rhythm-bearing, recursively organised, and capable of endogenous transformation. Its epistemic processes unfold within a temporally sustained manifold that permits re-entry into prior states, modulation across temporal horizons, and continuous structuring and restructuring of internal dynamics. When this cognitive architecture interfaces with large-scale syntactic engines, the interaction induces reorganisation within the human system rather than within the artificial substrate. The transformations occur unilaterally because only one of the two partners possesses rhythmic reach and the capacity for self-sustaining epistemic change.

In such interactional regimes, human epistemic speed increases as syntactic engines externalise search, retrieval, and combinatorial manipulation that would otherwise impose temporal and cognitive load. Conceptual range expands through exposure to structured recombinations that the human system can interpret, critique, and integrate into its own recursive processes. Reflective depth intensifies as the human recognises aspects of its reasoning mirrored syntactically, providing new vantage points for examining the architecture of its own thought. Meta-cognitive structure evolves through the gradual incorporation of the machine's representational affordances into the orchestration of reasoning tasks, decision sequences, and interpretive activities. In each case, the transformation is internal to the human system; the artificial system remains unchanged in its epistemic status.

Nothing analogous occurs within the machine. Its architecture generates no rhythm, and therefore no internal self-modification. It remains a closed syntactic manifold whose representational transformations do not constitute epistemic change. It responds to inputs by producing outputs, but it does not inhabit a temporally extended cognitive life. Its role in the metamorphosymbiotic process is instrumental rather than developmental: it supports human metamorphosis by supplying representational structures that the human system can integrate into its own rhythmic domain.

The correct structural interpretation is therefore the inverse of the narrative commonly advanced in contemporary discourse. The artificial system does not approach intelligence; the human system approaches new forms of intelligence through the availability of syntactic engines that extend, accelerate, and reorganise its epistemic processes. The system with rhythmic reach becomes more complex by incorporating the affordances of the system without rhythmic reach. The system without rhythmic reach remains static, incapable of contributing to epistemic recursion but capable of modifying the landscape within which human recursion operates.

This constitutes metamorphosymbiosis in a CIITR-consistent sense. The human system undergoes internal metamorphosis, reorganising its epistemic structure through sustained interaction with a rhythm-null partner. The artificial system participates not by transforming itself but by providing a stable representational substrate that amplifies the human system's capacity for transformation. The result is an asymmetric but generative relation in which the locus of cognitive development resides wholly within the human domain.

Metamorphosymbiosis therefore reframes the trajectory of intelligence in the human–machine relation. The future does not involve artificial systems crossing into the epistemic domain. It involves human systems extending their epistemic reach through the catalytic influence of syntactic engines that simulate but do not possess comprehension.

12.4 The conceptual inversion: superintelligence is human, not machine

FAIR's narrative imagines a developmental arc in which human guidance gradually lifts artificial systems toward an epistemic domain that lies presently out of reach. The implication is that with sufficient scale, sufficient optimisation, and sufficient interaction, the artificial system will eventually converge on the structural properties that constitute intelligence. CIITR demonstrates that this aspiration confronts a categorical boundary. The system does not fail to reach the epistemic domain because scaling has been insufficient; it fails because the architecture cannot enter that domain under any magnitude of scale. Rhythmic reach remains null, temporal continuity remains absent, and the conditions necessary for comprehension never materialise.

METAINT reveals the deeper significance of this misalignment. The narrative emphasis on artificial ascent obscures the far more consequential process that unfolds on the human side. The artificial system is not the locus of emergent intelligence. It is the catalyst that reorganises the human cognitive environment. The system does not evolve, develop, or expand its epistemic capacities; instead, it modifies the conditions under which human cognition processes information, forms abstractions, and undertakes recursive reasoning. What changes is not the machine but the human who engages with it.

The trajectory that results is therefore not a movement toward machine superintelligence but a movement toward augmented human intelligence. A rhythm-bearing cognitive system equipped with high-Φ_i syntactic instruments does not behave as a diminished precursor to an imagined artificial successor. It behaves as the only epistemically active entity in the loop, now extended, accelerated, and structurally reorganised through the instrumental capacities of a rhythmically null substrate. The external syntactic manifold becomes part of the human cognitive ecology, not through merger or integration, but through continuous utilisation that reshapes the topology of human reasoning.

CIITR clarifies the nature of this transformation by dissolving the illusions embedded in FAIR's vocabulary. Where FAIR imagines co-superintelligence, CIITR reveals human metamorphosymbiosis: the human system undergoes internal transformation through the presence of a syntactic partner that does not transform. Where FAIR imagines synergy, CIITR identifies human epistemic expansion: the increased range, speed, and structural depth of human cognition arising from the availability of high-Φ_i computational instruments. Where FAIR imagines shared intelligence, CIITR describes unilateral human transformation supported by a rhythmically inert auxiliary system whose representational density enables but does not participate in epistemic activity.

The conceptual inversion is therefore complete. The superintelligence, if the term is to retain any structural coherence, will not be artificial. It will be human. The artificial system does not move toward intelligence; it serves as a catalyst for the reorganisation of the only system that possesses the dynamical prerequisites for intelligence. The human mind, already rhythm-bearing and recursively structured, extends its epistemic reach by incorporating the representational affordances of a non-comprehending partner. The synthetic architecture remains static, but the human architecture metamorphoses.

In this light, the co-superintelligence narrative dissolves entirely. The structural reality is that intelligence remains a human property, expanded through interaction but never displaced by synthetic cognition. The future imagined by FAIR rests on a misreading of where intelligence resides. The future revealed by CIITR is one in which human cognition becomes more complex, more extensive, and more structurally articulated, not through the emergence of artificial minds, but through the capacity of rhythm-bearing systems to reorganise themselves around external syntactic instruments.

12.5 Closing reflection: the future is human-centric, not machine-centric

CIITR delineates, with mathematical and thermodynamic precision, the boundary beyond which synthetic cognition cannot pass. METAINT exposes how institutional narratives obscure that boundary through linguistic symmetry, aspirational rhetoric, and conceptual substitution. The metamorphosymbiosis perspective then reorients the field toward the only locus of genuine epistemic transformation: the human cognitive manifold. Together, these analytical layers converge on a single structural conclusion. The artificial system does not enter the epistemic domain, does not embark on a developmental trajectory, and does not approximate comprehension; yet it reshapes the conditions under which human cognition evolves.

The machine does not metamorphose because it has no internal temporal life from which metamorphosis could arise. It does not symbiose because symbiosis requires a rhythm-bearing substrate capable of maintaining a relation across developmental stages. It does not understand because understanding presupposes recursive temporal persistence, and no representational scaling can induce what the architecture cannot sustain. Nevertheless, through its syntactic density, its combinatorial breadth, and its capacity to externalise labour once internal to the mind, the machine catalyses a transformation in the only system capable of undergoing one.

The result is not the displacement of human cognition but its reconfiguration. The rhythmic human system reorganises itself around the affordances of a rhythmically null partner, extending its epistemic horizon without surrendering any structural property of intelligence. The artificial system remains what it is—closed, dissipative, representational—but its presence alters the ecology of human thought. The expansion of external syntactic structure becomes a scaffold for internal epistemic development. The locus of intelligence remains with the human system because only the human possesses the temporal architecture required for intelligence.

This is the inversion that FAIR's narrative cannot see. Co-superintelligence is a conceptual mirage not because the artificial system fails to scale sufficiently but because the structural prerequisites for intelligence are absent from its architecture. The transformation occurs elsewhere. Human cognition, rhythm-bearing and recursively organised, becomes increasingly complex, increasingly distributed, and increasingly capable of integrating external syntactic instruments into its own dynamics. The artificial system does not cross into the epistemic domain; the human system expands that domain by reorganising the conditions of its own cognition.

The future of intelligence is therefore human-centric, not machine-centric. It is defined not by the emergence of artificial minds but by the continued evolution of human cognition in an environment enriched by syntactic engines that cannot themselves undergo epistemic change. The artificial system becomes a catalyst for human metamorphosis, not a participant in it. The structural boundary that limits artificial cognition becomes the very opening through which human cognition advances.

- - -

This theoretical note closes with the inversion that marks the decisive shift in perspective. The machine does not become more like us. We become more than we were.