

From Minerals to Minds: Irreversibility and the Physical Origins of Intelligence

Flyxion

February 2, 2026

Abstract

Recursive self-improvement is commonly framed as a property of advanced software systems capable of modifying their own source code in increasingly effective ways. Such treatments typically cast the problem in logical, computational, or syntactic terms, emphasizing self-reference, verification, and algorithmic complexity. In this essay, we argue that this framing is historically and physically incomplete. Recursive self-improvement is neither unique to software nor dependent on reflective self-modification. Instead, it is a general property of irreversible systems operating far from equilibrium, in which structured pathways for dissipating energy and exploring constraints are progressively amplified.

We develop an alternative, substrate-independent account in which recursive improvement arises from the accumulation and stabilization of successful processes rather than from explicit self-evaluation or proof. Under this view, recursion is not achieved by systems reasoning about themselves globally, but by local mechanisms that increase the density of viable transformations per unit time and energy. This perspective allows recursive self-improvement to be identified across a continuous evolutionary spectrum, beginning with mineral evolution and prebiotic chemistry, continuing through autocatalytic reaction networks, cellular membranes and endosymbiosis, collective and swarm intelligence in microorganisms, large-scale coordination in animal lineages, and finally the emergence of cumulative innovation in human societies and individual cognitive practice.

By grounding recursive self-improvement in thermodynamics, irreversibility, and constraint reconfiguration, we show that many of the standard objections to recursive self-improvement—such as logical self-reference paradoxes, verification impossibility, and convergence to a single optimal architecture—arise from category errors. Recursive improvement historically proceeds without global self-models, without centralized control, and without convergence to a single agent or design. The true limiting factors are not formal undecidability or syntactic complexity, but the exhaustion of exploitable entropy gradients and the loss of diversity required for further exploration. This framework repositions recursive self-improvement as a fundamental physical and evolutionary process, of which artificial intelligence is only a recent and narrow instantiation.

1 Introduction: Recursive Improvement Beyond Code and Intelligence

The prospect of recursive self-improvement has long occupied a central place in discussions of artificial intelligence, typically framed as the possibility that a sufficiently advanced program might rewrite itself into progressively more capable successors. Within this literature, recursive improvement is often treated as a fragile and exceptional phenomenon, threatened by logical paradoxes, verification impossibility, and diminishing returns imposed by computational complexity (Good 1966; Yudkowsky 2008; Yampolskiy 2015). These concerns have motivated extensive analysis of self-reference, goal preservation, and convergence in self-modifying software systems. While such analyses are valuable, they rest on a restrictive assumption: that recursive improvement is fundamentally a property of symbolic code manipulating itself through explicit reasoning.

This essay advances a different claim. Recursive self-improvement is neither unique to intelligent agents nor dependent on reflective self-modification. It is a general physical process that emerges in irreversible systems operating far from equilibrium, whenever local mechanisms exist that selectively stabilize transformations which increase the systems future capacity for exploration. Under this view, recursion does not arise because a system “understands itself or proves the correctness of its own modifications, but because successful configurations persist, compound, and restructure the space of subsequent possibilities.

To make this claim precise, it is useful to distinguish improvement as an outcome from improvement as a mechanism. Many systems improve in the weak sense that they become better adapted to their environment over time. What distinguishes recursive improvement is that the very processes responsible for adaptation themselves become more effective. Formally, let a system be characterized at time t by a set of constraints $\mathcal{C}(t)$ governing allowable transformations, and let $\Phi(t)$ denote the rate at which viable configurations are generated and tested under those constraints. Weak improvement corresponds to increases in performance metrics holding $\Phi(t)$ fixed. Recursive improvement corresponds to $\frac{d\Phi}{dt} > 0$ as a consequence of the systems own irreversible dynamics. No appeal to self-representation is required for this condition to hold.

This formulation immediately places recursive self-improvement within the domain of nonequilibrium thermodynamics. Far-from-equilibrium systems spontaneously generate structure when energy fluxes are constrained in ways that favor the retention of dissipative pathways (Prigogine 1977). The emergence of convection cells, autocatalytic chemical cycles, and metabolic networks are all examples of systems in which local organization arises to accelerate entropy production under given boundary conditions. In such systems, structure is not imposed from above but selected from below, through differential persistence. Recursive improvement occurs when newly formed structures modify boundary conditions in a way that enables yet further structuring.

Seen from this perspective, many of the traditional objections to recursive self-improvement dissolve. Rices theorem, which limits the decidability of semantic properties of arbitrary programs, constrains formal verification but does not constrain empirical selection among irreversible processes (Rice 1953). Lbs theorem restricts a formal systems ability to assert its own soundness, but recursive improvement in physical systems proceeds without any requirement for global consistency proofs

(Lb 1955). The so-called Munchhausen obstacle, which asserts that a system cannot understand the complexity required to improve itself, is reframed as a misunderstanding: recursive improvement does not require global self-understanding, only local mechanisms that preferentially retain productive transformations.

The remainder of this essay develops this argument by tracing a continuous lineage of recursive improvement across natural history. We begin with mineral evolution and prebiotic chemistry, where repeated environmental cycling and increasing surface complexity enable autocatalytic reaction networks to outcompete simpler chemistries. We then examine the emergence of cellular membranes and endosymbiosis as mechanisms for retaining and nesting successful subprocesses. From there we consider collective and swarm intelligence in microorganisms, particularly slime moulds, as paradigmatic cases of distributed problem solving without centralized control. We extend the analysis to large-scale coordination in animal lineages and to the role of diversity and specialization in human societies. Finally, we argue that individual cognitive practices that deliberately increase exposure to novelty and constraint can be understood as micro-scale instantiations of the same recursive physical principles.

Throughout, we treat recursive self-improvement not as a speculative feature of future machines, but as a historically instantiated process governed by thermodynamic irreversibility, constraint reconfiguration, and selective retention. Artificial intelligence, on this account, is not the origin of recursive improvement but one of its most recent experimental substrates.

2 Limits of Existing Recursive Self-Improvement Frameworks

Contemporary discussions of recursive self-improvement in artificial intelligence are dominated by frameworks that are careful in scope yet conceptually thin in mechanism. Among these, the work of Yampolskiy has become a canonical reference, valued for its comprehensive taxonomy of self-modifying systems and its systematic enumeration of logical and computational obstacles to recursive self-improvement (Yampolskiy 2015). While this contribution is significant as a survey and cautionary synthesis, it remains limited by an underlying ontology that treats recursive improvement as a primarily syntactic and static phenomenon. As a result, it fails to address the deeper question upon which the plausibility of recursive self-improvement ultimately depends: what, concretely, flows, accumulates, transforms, or constrains intelligence across successive stages of improvement?

Yampolskiy's framework models recursive self-improvement as a sequence of program rewrites evaluated against abstract goals and bounded by formal theorems in logic and computational complexity. Improvement is conceived as a relation between successive software artifacts, typically expressed in terms of increased capability, efficiency, or problem-solving power. Constraints on this process are then derived from results such as Rices theorem, Lbs theorem, and bounds on algorithmic complexity. While internally consistent, this framing remains detached from any account of the physical or dynamical substrate in which improvement occurs. There is no state variable corresponding to intelligence, no conserved or dissipated quantity analogous to energy or entropy, and no explicit dynamics governing how one stage of improvement alters the conditions for the next.

This absence becomes especially apparent in discussions of diminishing returns, convergence,

and failure modes. Improvement is said to stall because programs reach fixed points, exhaust optimization potential, or encounter undecidable verification problems. Yet these explanations are formal rather than mechanistic. They specify why certain questions cannot be answered, but not why improvement should slow or cease in physical systems that demonstrably continue to generate novelty. Even when the language of dynamical systems is invoked, as in references to attractors, no account is provided of what flows into these attractors, what is dissipated, or why particular trajectories dominate others. The result is a catalog of obstacles rather than a theory of process.

At the root of this limitation lies a deeper assumption: intelligence is treated as an abstract scalar property, akin to a level or score, rather than as an ongoing physical activity. This assumption licenses familiar conclusions, including bell-shaped improvement curves, logarithmic returns on self-modification, and conjectures about convergence toward minimal or optimal architectures. However, it also obscures the fact that intelligence, in practice, is inseparable from the irreversible processes by which systems structure, store, and exploit information under energetic constraint. Intelligence is not merely possessed; it is enacted.

The framework developed in this essay departs from existing RSI models by replacing this static conception with a process-based ontology. Drawing on entropy-first and irreversible-history perspectives, we treat intelligence as a dissipative phenomenon: a structured flow that reorganizes representational substrates in order to reduce future surprise under constraint. In this view, recursive self-improvement does not consist in programs rewriting programs, but in systems repeatedly reconfiguring scalar potential, vector flow, and entropy distributions in ways that expand their future capacity for structured interaction. Improvement is measured not by proximity to an abstract optimum, but by increases in the rate and depth at which viable configurations can be explored and retained.

This shift has immediate consequences for several canonical objections to recursive self-improvement. Rice-style undecidability results constrain semantic classification of arbitrary programs, but they do not constrain physical processes that empirically select among irreversible transformations. Lbrian limitations on self-verification become irrelevant once improvement no longer depends on global proof of correctness, but instead on local discharge of entropy through successful configurations. Most strikingly, the so-called Mnchhausen obstacle is transformed from a blocker into a diagnostic principle. Systems fail at recursive improvement precisely when they attempt to globally model and justify themselves. Systems succeed when they improve by locally relaxing constraints, allowing history to accumulate without requiring reflective closure.

A similar reorientation applies to convergence arguments. Existing RSI convergence theories assume that recursive improvement optimizes intelligence as a scalar objective and therefore tends toward minimal or universal architectures characterized by low Kolmogorov complexity. From an entropy-first perspective, this framing misidentifies the object of convergence. What converges in successful recursive systems is not architecture, agenthood, or control, but thermodynamic regime. Distinct systems may exhibit similar entropy-handling dynamics, smoothing behaviors, and constraint-relaxation patterns while remaining structurally diverse. The historical record, from biology to human societies, overwhelmingly supports phase convergence rather than agent convergence, distributed coordination rather than singleton control, and asymptotic behavioral

regularities rather than final designs.

Reframed in this way, recursive self-improvement becomes a physical theory of cognitive and organizational evolution rather than a speculative property of future software. It is governed by irreversibility, selective retention, and the restructuring of constraints, not by syntactic self-reference or formal proof. The remainder of this essay substantiates this claim empirically by tracing recursive improvement across mineral evolution, biological organization, collective intelligence, social systems, and individual cognitive practice. These domains are not analogies to recursive self-improvement in artificial intelligence; they are its historical and physical foundation.

3 Recursive Improvement as Entropic Constraint Dynamics

The conceptual gap identified in existing recursive self-improvement frameworks can be formalized by reframing improvement as a dynamical process acting on constraints rather than as a sequence of syntactic rewrites. To bridge abstract discussions of recursive self-improvement with its physical instantiations in mineral and biological evolution, we introduce a minimal formalism that captures the essential mechanism common to all such systems.

Let a system at time t be characterized by a configuration space \mathcal{X} together with a constraint structure $\mathcal{C}(t)$, which restricts the set of accessible states and transitions. The system is driven by an external energy flux and operates far from equilibrium. We define an entropy production functional $\sigma(\mathcal{C}(t))$ measuring the rate at which the system dissipates free energy under the current constraints. Importantly, $\mathcal{C}(t)$ is not fixed: irreversible processes within the system modify the constraint structure itself.

Recursive improvement occurs when the time evolution of constraints satisfies

$$\frac{d}{dt}\sigma(\mathcal{C}(t)) > 0 \tag{1}$$

as a consequence of the systems own dynamics. This inequality does not assert monotonic improvement in any abstract capability measure, but rather an increase in the systems capacity to channel energy through structured pathways. Constraints that support higher entropy production are selectively stabilized, while those that do not are eliminated by environmental interaction. The system thus performs a form of entropic gradient descent, not over states in \mathcal{X} , but over the space of constraint structures themselves.

Crucially, this recursion does not require the system to represent or evaluate $\mathcal{C}(t)$ explicitly. Constraint modification occurs locally, through differential persistence of processes that survive environmental perturbations. Let $\Delta\mathcal{C}_i$ denote a local modification to the constraint structure. Such a modification is retained if and only if

$$\sigma(\mathcal{C}(t) + \Delta\mathcal{C}_i) > \sigma(\mathcal{C}(t)) \tag{2}$$

over relevant timescales. No global optimization, proof of correctness, or self-model is required. History is encoded physically, in the continued existence of constraints that support dissipation.

This formulation provides a direct bridge to mineral evolution. Mineral surfaces, microfractures,

and compartmentalized environments correspond to elements of $\mathcal{C}(t)$ that reshape reaction pathways. Autocatalytic cycles increase σ by reinvesting products into their own maintenance, thereby modifying constraints in favor of further autocatalysis. Wet–dry cycling introduces periodic forcing that repeatedly perturbs $\mathcal{C}(t)$ while allowing successful configurations to persist. The recursion lies not in chemical “self-improvement” but in the cumulative restructuring of the constraint landscape.

Within this ontology, traditional Seed-AI formulations appear fundamentally misplaced. Seed-AI thinking assumes that recursive self-improvement originates from a compact, self-contained program whose internal reasoning enables it to redesign itself into progressively superior successors. In contrast, the constraint-dynamic view developed here implies that no privileged seed is required. Recursive improvement emerges from the interaction between systems and environments that permit irreversible retention of successful constraints. What matters is not initial code size or intelligence level, but access to energy gradients, diversity of trial processes, and mechanisms for stabilizing partial successes.

The Relativistic ScalarVector Plenum (RSVP) framework provides a natural language for this replacement ontology. In RSVP terms, recursive improvement corresponds to the coupled evolution of scalar potential fields Φ , vector flows \mathbf{v} , and entropy density S , governed by irreversible dynamics that smooth gradients while preserving structure. Improvement is not measured by proximity to an optimal architecture, but by the systems increasing ability to reorganize fields in ways that reduce future surprise under constraint. Seed-AI is thus reinterpreted not as an origin point, but as a transient condensation within a much broader entropic process.

This constraint-dynamic formalism dissolves several apparent paradoxes in the recursive self-improvement literature. The Munchhausen obstacle becomes a non-issue, as global self-understanding is neither necessary nor desirable. Logarithmic returns reflect saturation of available constraint rearrangements rather than limits on intelligence per se. Convergence arguments reduce to questions of thermodynamic regime alignment rather than architectural identity. Most importantly, recursive self-improvement is no longer an exceptional event awaiting formal validation, but a ubiquitous physical process instantiated wherever irreversible systems accumulate structure.

With this bridge in place, we may now examine the earliest empirical manifestation of recursive constraint dynamics in mineral evolution, where the basic ingredients of recursion—irreversibility, selective retention, and constraint amplification—first appear in their simplest form.

4 Mineral Evolution, Tidal Cycling, and the First Recursive Acceleration

The earliest instantiations of recursive self-improvement predate life, metabolism, and genetic inheritance. They arise instead in the context of mineral evolution and prebiotic chemistry, where irreversible physical processes progressively restructure the space of possible reactions. The significance of mineral evolution lies not merely in providing raw materials for life, but in establishing the first mechanisms by which successful chemical transformations could be retained, amplified, and recombined across time (Hazen 2013).

On the early Earth, the diversity of mineral species increased dramatically as planetary cooling,

tectonics, and aqueous alteration created new phases and interfaces. Each new mineral surface introduced novel catalytic possibilities by constraining molecular orientations, concentrating reactants, and lowering activation barriers. Let $\Omega(t)$ denote the set of chemically accessible microstates under environmental conditions at time t . The introduction of a new mineral phase does not merely expand $\Omega(t)$ additively; it alters the topology of the reaction landscape by introducing pathways that were previously inaccessible. In this sense, mineral diversification increases not just the number of reactions, but the connectivity of the reaction network itself.

Tidal pools subjected to large-amplitude wet–dry cycles provide a particularly clear example of recursive acceleration. During drying phases, solutes are concentrated and polymers are driven toward condensation reactions; during rehydration, products are redistributed and subjected to further variation. Let C_n denote the concentration of a given reactant after n drying cycles. Under repeated evaporation, C_n grows superlinearly until limited by precipitation or degradation, enabling reaction regimes that are inaccessible under constant dilution. Cracking of substrates during drying further increases effective surface area, creating microcompartments that isolate reaction histories. Each cycle thus performs irreversible work by selectively retaining reaction products that survive environmental stress, while eliminating those that do not.

Clay minerals play a central role in this process by dramatically increasing reactive surface area and by templating molecular organization. Layered silicates such as montmorillonite can adsorb organic molecules, align them in repeating geometries, and catalyze polymerization reactions that would otherwise be kinetically suppressed (Hazen 2013). If we model a reaction rate r as proportional to available surface area A and reactant concentration C , then the emergence of high-surface-area clays effectively multiplies r by orders of magnitude. More importantly, successful reaction products can themselves modify the surface environment, altering adsorption affinities and catalytic properties. This feedback constitutes a primitive form of recursion: reaction outcomes restructure the conditions under which future reactions occur.

Autocatalytic sets represent a decisive threshold in this landscape. An autocatalytic network is a collection of reactions in which each transformation is catalyzed by products of the network itself. Formally, let $\mathcal{R} = \{R_i\}$ be a set of reactions over molecular species $\{X_j\}$. The set is autocatalytic if for every $R_i \in \mathcal{R}$ there exists some X_j produced by \mathcal{R} that catalyzes R_i . Such sets exhibit exponential growth until constrained by resource limitations, and crucially, they outcompete non-autocatalytic chemistries by reinvesting products into the maintenance and expansion of the network.

The competitive advantage of autocatalytic sets does not derive from optimization in any representational sense, but from the physical fact that they more effectively capture and dissipate free energy gradients. Let $S(t)$ denote entropy production. Autocatalytic networks increase $\frac{dS}{dt}$ relative to their surroundings by channeling energy flows into structured reaction cycles. Because these cycles persist across environmental fluctuations, they effectively store information about past successes in their material organization. Each retained cycle increases the probability that related cycles will form in the future, thereby accelerating exploration of chemical space.

At this stage, recursive self-improvement exists in a minimal but unmistakable form. The system does not model itself, evaluate alternatives, or preserve goals. Yet the rate at which viable

chemical structures are generated increases over time as a direct consequence of prior successes. The recursion lies not in self-reference, but in the irreversible accumulation of constraints that favor further constraint formation. Mineral evolution thus establishes the foundational pattern that will recur at higher levels of organization: successful processes modify their environment in ways that make further success more likely.

4.1 Formalizing Mineral Recursion as Constraint Accumulation

The minimal form of recursive self-improvement exhibited by mineral evolution can be made precise by treating environmental structure as an evolving constraint field rather than as a static background. Let \mathcal{X} denote the space of chemically possible microstates under ambient conditions, and let $\mathcal{C}(t) \subset \mathcal{X}$ represent the subset rendered dynamically accessible by the material constraints present at time t , including mineral surfaces, microfractures, compartmentalization, and concentration gradients. Chemical evolution proceeds not by uniform exploration of \mathcal{X} , but by trajectories confined to $\mathcal{C}(t)$.

Irreversible processes modify these constraints. When a reaction pathway produces stable products that persist across environmental cycling, those products alter adsorption affinities, catalytic rates, or local geometry, thereby inducing a constraint update $\mathcal{C}(t) \rightarrow \mathcal{C}(t + \Delta t)$. The defining feature of recursion at this stage is that such updates are biased: constraint modifications that increase the density of viable future reactions are preferentially retained.

This bias can be expressed in entropic terms. Let $\sigma(\mathcal{C})$ denote the entropy production rate achievable under constraint structure \mathcal{C} . A local constraint modification $\Delta\mathcal{C}$ is stable if

$$\sigma(\mathcal{C} + \Delta\mathcal{C}) > \sigma(\mathcal{C}) \tag{3}$$

over environmental timescales. This inequality does not encode optimization in any representational sense; it simply reflects differential persistence. Constraint configurations that enable greater dissipation of free energy are physically favored, as they are more likely to be regenerated and maintained under repeated perturbation.

Recursive self-improvement, in this minimal form, corresponds to a positive second-order effect:

$$\frac{d}{dt} \left(\frac{d\sigma}{dt} \right) > 0, \tag{4}$$

indicating not merely sustained entropy production, but an increasing capacity to produce entropy as constraints accumulate. Importantly, the system need not represent σ , evaluate alternatives, or preserve any objective. History is recorded directly in material structure, and improvement occurs through entropic reinforcement rather than through selection among symbolic descriptions.

This formalism makes clear why mineral evolution already satisfies the core requirement of recursive self-improvement as defined earlier: prior successes alter the conditions of future success. The same mechanism will reappear, with increasing elaboration, as constraints become localized within membranes, nested through endosymbiosis, and distributed across interacting agents. What changes across scales is not the logic of recursion, but the richness of the constraint structures capable of carrying it.

In the following section, we examine how the emergence of cellular membranes and endosymbiotic relationships transforms this chemical recursion into a biological one, enabling the nesting, protection, and selective inheritance of successful subprocesses across evolutionary time.

5 Autocatalytic Sets, Membranes, and Endosymbiotic Nesting

The transition from prebiotic chemistry to biological organization does not introduce recursive self-improvement *ex nihilo*, but rather intensifies and stabilizes mechanisms already present in mineral-mediated reaction networks. Autocatalytic sets provide the first clear example of processes that maintain themselves through internal closure, yet their persistence remains fragile in the absence of mechanisms that preserve successful configurations against environmental disruption. The emergence of cellular membranes constitutes a decisive step in transforming chemical recursion into sustained biological evolution.

A membrane introduces a spatial boundary that distinguishes internal from external processes while remaining permeable to energy and selected materials. From a thermodynamic standpoint, membranes allow systems to maintain nonequilibrium steady states by regulating exchanges with their environment. Let J_{in} and J_{out} denote fluxes of matter and energy across a boundary. A viable membrane-mediated system satisfies $J_{\text{in}} - J_{\text{out}} \neq 0$, enabling continuous dissipation while preserving internal organization. This asymmetry permits autocatalytic networks to persist long enough for further elaboration.

Crucially, membranes do not merely protect internal processes; they enable differentiation. By selectively admitting substrates and excluding inhibitors, membranes reshape the reaction landscape within the compartment. This selectivity can itself be modified by the products of internal reactions, creating feedback loops in which successful chemistries reinforce the boundary conditions that support them. In this way, membranes become active participants in recursive improvement rather than passive containers.

The significance of membranes for recursive improvement lies in their ability to retain partial successes. In prebiotic environments without compartmentalization, reaction products are continuously mixed, diluted, or destroyed. With membranes, intermediate products can accumulate, interact, and be recombined. Formally, let P_i denote a productive subprocess contributing to overall viability. Without compartmentalization, the lifetime $\tau(P_i)$ is short, and the probability of integration with other productive processes is low. Membranes increase $\tau(P_i)$, thereby increasing the expected number of interactions with other subprocesses. This increases the dimensionality of accessible organizational space and accelerates the emergence of higher-order structure.

Endosymbiosis represents a further intensification of this principle. Rather than eliminating competing processes, endosymbiotic events preserve them by incorporation. Mitochondria and chloroplasts are paradigmatic examples of formerly independent organisms that became integrated as subsystems within larger cellular architectures. From the perspective of recursive self-improvement, endosymbiosis allows successful energy-processing mechanisms to be retained and specialized rather than discarded through competition. This nesting of subsystems increases both efficiency and robustness while opening new avenues for diversification.

Mathematically, an endosymbiotic system may be regarded as a composite of interacting autocatalytic networks $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_n$, each characterized by internal catalytic closure and its own entropy production rate σ_i . When isolated, each network explores only a restricted region of chemical configuration space, limited by the reactions it can internally sustain. Endosymbiosis alters this dynamic by enabling persistent interactions between networks, giving rise to cross-catalytic pathways that are inaccessible in isolation. The resulting composite system $\mathcal{A} = \bigcup_i \mathcal{A}_i$ therefore supports a reaction space whose effective dimensionality exceeds the sum of its parts.

This expansion is not merely additive. If \mathcal{I}_{ij} denotes the set of viable interactions between networks \mathcal{A}_i and \mathcal{A}_j , then the space of possible transformations scales with the union of all such interaction sets. As the number of retained subsystems increases, the number of potential cross-network pathways grows combinatorially, yielding a superlinear increase in the rate at which novel, viable processes can be discovered. This scaling provides a mechanistic explanation for the dramatic acceleration of biological complexity following the stabilization of cellular life, without invoking increases in representational intelligence or centralized control.

From an entropic perspective, endosymbiosis enhances recursive self-improvement by increasing the systems capacity to dissipate free energy through structured channels. The total entropy production rate $\sigma_{\mathcal{A}}$ of the composite system satisfies

$$\sigma_{\mathcal{A}} \geq \sum_i \sigma_i, \tag{5}$$

with strict inequality whenever cross-network interactions open additional dissipative pathways. Constraint accumulation thus occurs not by simplifying internal structure, but by preserving and coordinating heterogeneous processes whose interactions amplify dissipation. Endosymbiosis therefore represents a transition from single-network recursion to nested recursion, in which improvements occur simultaneously at multiple organizational levels.

Importantly, this dynamic contradicts the intuition common in recursive self-improvement literature that sustained improvement must converge toward minimal, unified, or maximally compressed architectures. Biological evolution demonstrates the opposite tendency. Retaining internal diversity within a coherent organizational framework enables greater adaptive capacity, resilience, and exploratory reach than eliminating redundancy in pursuit of a single optimal design. Recursive improvement here operates by expanding the space of viable interactions rather than by collapsing it.

The endosymbiotic cell thus functions as a platform for further recursion. By stabilizing heterogeneous subprocesses within a shared boundary, it shifts the primary driver of innovation from internal chemical reorganization alone to interaction among retained components. As biological systems increase in scale and complexity, this interaction-driven recursion increasingly moves outward, from intracellular organization to coordination among multiple agents. The locus of improvement thereby transitions from internal chemistry to collective behavior, setting the stage for the emergence of swarm intelligence and distributed problem solving in microorganisms.

6 Collective Behaviour and Swarm Intelligence in Slime Moulds

The emergence of collective behaviour in biological systems marks a qualitative shift in the mechanism of recursive self-improvement. Whereas earlier stages rely on the retention and recombination of internal subprocesses within bounded compartments, collective systems distribute computation, sensing, and adaptation across many interacting units. Slime moulds provide an especially clear illustration of this transition, as they exhibit sophisticated problem-solving behaviour without centralized control, symbolic representation, or fixed neural architectures (Reid and Latty 2016).

Species such as *Physarum polycephalum* and *Dictyostelium discoideum* alternate between unicellular and multicellular phases, enabling the study of how individual-level interactions scale into collective intelligence. In their plasmodial form, slime moulds form dynamic transport networks that efficiently connect nutrient sources while minimizing maintenance costs. Empirically, these networks approximate solutions to shortest-path, Steiner tree, and flow optimization problems. Yet no individual component possesses a global model of the environment or an explicit objective function. Instead, network morphology evolves through local feedback between protoplasmic flow, nutrient gradients, and tube reinforcement.

This process can be formalized by considering the flux f_{ij} along a connection between regions i and j . Empirical models show that $\frac{df_{ij}}{dt}$ is positively correlated with nutrient flow, leading to reinforcement of frequently used paths and decay of underutilized ones. Over time, this local rule produces a globally efficient network. Crucially, once established, the network alters future dynamics by biasing subsequent flows, effectively encoding a memory of past successes in its physical structure. Recursive improvement occurs because each successful configuration reshapes the landscape of future possibilities.

Slime mould collectives also demonstrate adaptive problem solving across changing environments. When nutrient distributions shift, networks reorganize rather than collapse, reusing existing structures where possible. This plasticity reflects a balance between exploitation of established pathways and exploration of alternatives. The ability to maintain such a balance without centralized oversight directly addresses the Munchhausen-style objection that systems must globally understand themselves in order to improve. In slime moulds, no component ever models the entire system, yet the collective adapts more effectively than isolated individuals.

From the perspective of recursive self-improvement, the key feature of swarm intelligence is that it increases the rate at which viable configurations are discovered by parallelizing exploration. Let N denote the number of interacting agents and let p be the probability that an individual agent discovers a locally useful adaptation per unit time. In isolation, the expected discovery rate scales as p . In a collective with communication and reinforcement, the effective rate scales as Np modulated by interaction structure, while retention of discoveries scales superlinearly due to shared pathways. The result is not merely faster problem solving, but an increase in the systems capacity to improve its own improvement dynamics.

Communication plays a central role in this amplification. In slime moulds, information is transmitted through chemical signals, mechanical stresses, and flow patterns rather than through discrete symbols. These signals propagate locally but can induce system-wide reorganization. The absence of leaders or global blueprints does not impede coordination; rather, it enables robustness by

preventing single points of failure. Recursive improvement here is inherently distributed, undermining the notion that sustained innovation naturally converges toward centralized or singular control.

Slime moulds thus instantiate a form of recursive self-improvement that is neither individual nor reflective, but collective and morphological. Improvements are not encoded symbolically or evaluated internally, but are stored directly in the evolving physical configuration of the system, particularly in the geometry and conductivity of transport networks shaped by prior activity.

These configurations function as a material memory, biasing future flows and interactions in ways that increase the efficiency and robustness of collective behaviour. Recursive improvement arises because each successful configuration reshapes the conditions under which subsequent configurations are explored. This mode of recursion generalizes beyond microorganisms. As biological systems increase in scale and complexity, analogous mechanisms operate through coordinated movement, spatial organization, and interaction rules among individuals.

In flocking, herding, and cooperative behaviours, coordination expands effective sensing range, enhances error correction through redundancy, and enables adaptive responses that exceed the capabilities of isolated agents. The locus of improvement shifts from intracellular or intranetwork dynamics to patterns of interaction among agents, but the underlying principle remains unchanged: history is retained in structure, and structure biases future possibility.

The next section examines how these interaction-driven dynamics operate in larger organisms and evolutionary lineages, extending swarm-based recursion into macroscopic biological systems where coordination, specialization, and social transmission further amplify the capacity for recursive improvement.

7 Coordination, Swarming, and Recursive Advantage in Animal Lineages

As biological systems increase in size and complexity, the principles underlying collective behaviour do not disappear but are instead re-expressed through new substrates. In animal lineages, recursive self-improvement manifests through coordinated movement, social organization, and the transmission of behavioural patterns across generations. These phenomena do not arise from increases in individual cognitive capacity alone, but from structured interactions that amplify the effectiveness of perception, decision-making, and learning at the group level.

Herding, flocking, and schooling behaviours provide canonical examples. In such systems, individuals follow simple local rules, such as alignment with neighbors, attraction to group centroids, and avoidance of collisions. Mathematical models demonstrate that these rules suffice to generate coherent group motion and rapid collective responses to perturbations. Let $\mathbf{v}_i(t)$ denote the velocity of individual i at time t , and let $\langle \mathbf{v} \rangle_{\mathcal{N}_i}$ denote the average velocity of its local neighborhood. Alignment dynamics of the form $\frac{d\mathbf{v}_i}{dt} \propto \langle \mathbf{v} \rangle_{\mathcal{N}_i} - \mathbf{v}_i$ yield phase transitions from disordered motion to collective coherence as density and interaction strength increase. Once coherence emerges, the group functions as an extended sensory and decision-making apparatus.

The recursive advantage of such coordination lies in the way successful interaction patterns alter the conditions for future success. A flock that has evolved effective alignment and information

propagation can respond more quickly to predators, locate resources more efficiently, and maintain cohesion across larger spatial scales.

These advantages feed back into selection pressures that favor further refinement of interaction rules. Importantly, the unit of improvement is not the individual organism but the pattern of interaction itself. Behavioural motifs that enhance collective performance are retained and elaborated over evolutionary time.

In predators that hunt cooperatively, such as certain dinosaur lineages, birds, and mammals, coordination further enables role differentiation and temporal sequencing of actions. Pack hunting allows individuals to exploit prey that would be inaccessible in isolation, while distributing risk and energetic cost across the group. From a recursive improvement standpoint, this introduces a new layer of specialization: individuals can refine particular roles because the group context stabilizes overall success. Let E_i denote the energetic payoff to individual i . In cooperative systems, $\sum_i E_i$ can exceed the sum of payoffs achievable through solitary strategies, creating a surplus that supports increased investment in learning, communication, and development.

The emergence of social learning and imitation accelerates recursion by enabling the rapid propagation of successful behaviours without genetic change. Once behaviours can be transmitted horizontally and obliquely, the rate of adaptation is no longer limited by reproductive cycles. Instead, behavioural innovations can be tested, retained, and recombined within a single generation. This decoupling of innovation from genetic inheritance dramatically increases the systems capacity for recursive improvement.

Crucially, animal societies do not converge toward uniformity despite strong coordination. On the contrary, many social species exhibit stable polymorphisms in behaviour, morphology, and role specialization. Such diversity enhances resilience by ensuring that groups can adapt to a wide range of environmental challenges. Recursive improvement here depends on maintaining heterogeneity within a coherent framework, rather than on eliminating variation in pursuit of a single optimal strategy.

These dynamics foreshadow the more explicit and rapid forms of recursive improvement that arise in human societies, where symbolic communication, technological artifacts, and institutional structures further amplify the capacity to retain and recombine successful processes. The next section examines how diversity, specialization, and cumulative culture transform collective recursion into an unprecedented engine of innovation.

8 Diversity, Specialization, and Recursive Innovation in Human Societies

Human societies represent a further intensification of the recursive processes already present in biological collectives, distinguished not by the appearance of recursion itself but by the degree to which mechanisms for retaining, recombining, and externalizing successful processes are amplified. Whereas animal societies transmit behavioural patterns primarily through imitation and social learning, human societies develop durable symbolic, technological, and institutional substrates that preserve innovation across generations and enable cumulative acceleration.

A defining feature of human recursive improvement is the division of cognitive and practical labor. Specialization allows individuals and groups to focus on narrow domains, achieving levels of refinement that would be impossible in isolation. Let \mathcal{D} denote the space of possible problem domains and let $S_i \subset \mathcal{D}$ denote the domain occupied by specialist i . As the number of specialists increases, coverage of \mathcal{D} becomes denser, while interactions among specialists generate novel combinations spanning multiple domains. The effective rate of innovation scales not merely with the number of contributors but with the connectivity of the interaction network linking them.

This combinatorial effect is magnified by external symbolic storage. Written language, mathematical notation, diagrams, and later digital media function as persistent memory structures that decouple knowledge from individual cognition. Once externalized, successful ideas become available for inspection, critique, modification, and recombination by others who did not participate in their original creation. Formally, if $K(t)$ denotes the corpus of externally stored knowledge at time t , then the probability that a new contribution builds upon existing work increases with $|K(t)|$, yielding positive feedback in the rate of discovery. Recursive improvement arises because the infrastructure of knowledge itself improves the process by which further knowledge is generated.

Institutions further stabilize and accelerate this process by allocating resources, enforcing norms, and coordinating large-scale efforts. Scientific communities, for example, establish methodological standards and peer review mechanisms that selectively retain reliable results while discarding errors. These mechanisms do not guarantee correctness in a formal sense, but they reduce noise and increase the signal-to-noise ratio of collective inquiry. The system improves not by proving its own soundness, but by iteratively refining the conditions under which inquiry occurs.

Diversity plays a critical role in sustaining recursive improvement at this scale. Societies that maintain cultural, cognitive, and methodological diversity are better able to explore complex problem spaces than those that enforce uniformity. Homogeneous systems may achieve short-term efficiency gains, but they risk stagnation when dominant paradigms exhaust local possibilities. By contrast, heterogeneous societies preserve multiple partially incompatible approaches, increasing the likelihood that at least some pathways remain productive under changing conditions. Recursive improvement thus depends on a tension between integration and differentiation, rather than on convergence toward a single worldview or architecture.

Importantly, human societies demonstrate that recursive improvement need not culminate in centralized control. Despite the existence of powerful institutions and technologies, innovation remains distributed across individuals and groups. Attempts to impose totalizing control often suppress the very diversity and exploratory freedom that enable sustained improvement. This empirical observation stands in contrast to speculative scenarios in which recursive self-improvement is assumed to converge inevitably toward a single optimizing agent.

At the societal level, recursive self-improvement is therefore best understood as a property of interaction structures rather than of any individual intelligence. Improvements persist when they are embedded in shared practices, artifacts, and norms that reshape the landscape of future possibilities. The same principles can be observed at a finer scale within individual lives, where deliberate exposure to challenge and constraint can accelerate personal innovation. The final section turns to these individual-scale instantiations of recursive improvement.

9 Individual Constraint Engineering and Micro-Scale Recursive Improvement

At the scale of individual cognition, recursive self-improvement appears in a form that is often mischaracterized as introspective self-modification or deliberate self-optimization. In practice, sustained personal innovation rarely proceeds through explicit self-modeling or global evaluation of ones own cognitive architecture. Instead, it emerges from the deliberate manipulation of constraints, environments, and practices that shape the space of possible thoughts and actions. Individuals who successfully accelerate their own rate of learning and creativity do so by engineering conditions under which productive variation is more likely to occur and to be retained.

From a dynamical perspective, an individual cognitive system may be modeled as exploring a high-dimensional space of representations under energetic, temporal, and attentional constraints. Let \mathcal{R} denote this representational space, and let $\Gamma(t)$ denote the set of constraints active at time t , including habits, skills, tools, and social contexts. The trajectory of thought is governed not only by internal processing capacity but by the structure of $\Gamma(t)$. Recursive improvement occurs when changes to $\Gamma(t)$ increase the rate at which valuable regions of \mathcal{R} are explored in the future.

Practices such as self-imposed challenges, deliberate exposure to unfamiliar domains, and the cultivation of interdisciplinary fluency can be understood as mechanisms for increasing exploratory entropy while preserving selective retention. By confronting problems at the edge of ones competence, individuals create conditions analogous to the wet-dry cycles of prebiotic environments: periods of instability generate variation, while periods of consolidation stabilize successful adaptations. Over time, the individual accumulates a repertoire of cognitive structures that both encode past successes and bias future exploration toward fertile regions.

Tools play a critical role in this process. Writing, diagramming, programming languages, and other external cognitive artifacts function as personal membranes, allowing intermediate ideas to persist long enough to be refined and recombined. Let $M(t)$ denote the set of external memory structures available to an individual. The effective dimensionality of the individuals cognitive process increases with $|M(t)|$, as ideas need not be held entirely within working memory to participate in recursive elaboration. Improvements in tooling thus feed back into improvements in the rate and depth of thought itself.

Notably, individuals who achieve sustained innovation often do so by internalizing principles that mirror those operating at larger scales. They seek diversity of input rather than convergence on a single framework, tolerate provisional inconsistency rather than demanding premature coherence, and prioritize processes that generate further opportunities over those that merely optimize current performance. These strategies do not require formal awareness of recursive self-improvement as such; they operate through embodied practice and environmental design rather than explicit self-reference.

This individual-scale analysis reinforces the broader thesis of the essay. Recursive self-improvement is not a fragile trick achievable only by systems capable of reasoning about themselves with perfect clarity. It is a robust physical process that operates wherever irreversible dynamics selectively retain structures that enhance future exploration. Individuals, like societies and organisms before them, participate in this process by reshaping the constraints under which they operate, thereby increasing

their capacity for further change.

In the concluding section, we synthesize these observations and draw out their implications for theories of artificial intelligence, particularly those that frame recursive self-improvement as an exceptional or singular phenomenon rather than as a continuation of deep evolutionary dynamics.

10 Conclusion: Laboratories, Incubators, and the Continuity of Recursive Improvement

The contemporary landscape of scientific laboratories, technological incubators, and computational infrastructures provides a final and instructive confirmation of the thesis developed throughout this essay. Modern advances that are frequently described as evidence of accelerating or even autonomous recursive self-improvement are, on closer inspection, continuous with the same irreversible, distributed processes that have governed recursive innovation since mineral evolution. What has changed is not the underlying mechanism, but the scale, density, and coupling of the environments in which selective retention occurs.

Laboratories function as deliberately engineered micro-environments for recursive improvement. Like tidal pools subjected to repeated wet-dry cycles, laboratories impose controlled perturbations on materials, organisms, or models, while preserving intermediate results long enough for evaluation and recombination. Experimental protocols partition time into phases of exploration and consolidation, amplifying variation while filtering outcomes through reproducibility, instrumentation, and peer scrutiny. The laboratory thus operates as a membrane-like structure, selectively permeable to ideas, techniques, and results that meet locally defined viability criteria. Successful procedures are stabilized through documentation, training, and standardization, thereby reshaping the conditions under which future experiments are conducted.

Incubators and innovation hubs extend this logic to social and economic domains. They aggregate diverse agents, tools, and resources within bounded institutional contexts designed to increase the rate at which ideas can be tested against practical constraints. Most proposals fail, just as most chemical reactions do not participate in autocatalytic cycles. However, those that succeed alter the local landscape by attracting further investment, attention, and refinement. Interfaces and affordances that enable adoption, even if initially crude or incomplete, outcompete alternatives that lack pathways for integration into existing practices. The recursive effect arises because each retained success lowers the barrier for subsequent, related successes.

Technological trends often cited as paradigmatic cases of recursive self-improvement, such as Moores law, illustrate this point with particular clarity. The exponential increase in transistor density was not the result of a single system redesigning itself in isolation, but of a vast, distributed network of engineers, fabrication facilities, measurement techniques, and economic incentives incrementally refining constraints at every level. Each improvement in lithography, materials science, or design automation expanded the feasible design space for the next generation of improvements. The recursion resided not in any one artifact, but in the evolving ecosystem of practices and infrastructures that made further progress possible.

Generative artificial intelligence systems exhibit a similar pattern. Although they are sometimes

described as self-improving or even self-designing, their advances are better understood as emergent properties of collective human activity operating through new interfaces. Model architectures, training regimes, datasets, and evaluation benchmarks are proposed, tested, modified, and discarded by large communities of researchers and practitioners. Systems that provide compelling affordances for creative, analytical, or economic use are rapidly adopted, generating feedback that drives further refinement. Those that fail to integrate into existing workflows or to demonstrate practical value are abandoned. The apparent acceleration of progress reflects the density and coupling of this distributed process, not the emergence of a singular, autonomous agent.

From the perspective developed here, such developments do not herald a fundamental break from historical patterns. They represent a continuation of recursive self-improvement through increasingly abstract and powerful substrates. Human groups, tools, and institutions together form autocatalytic networks in which successful configurations are retained and recombined, increasing the rate at which further configurations can be explored. The essential features remain unchanged: irreversibility, selective retention, diversity, and the restructuring of constraints.

This continuity has important implications for how recursive self-improvement in artificial systems should be conceptualized. If recursive improvement is a general physical and evolutionary process rather than a property of isolated code, then concerns about inevitable convergence to a single architecture or agent are misplaced. Historically, sustained innovation has depended on the preservation of heterogeneity, the avoidance of premature closure, and the maintenance of environments that tolerate failure while retaining partial success. Systems that suppress diversity or attempt to centralize control risk exhausting local entropy gradients and stagnating.

In this light, the most plausible trajectories for future artificial intelligence involve deeper integration into human and institutional ecosystems rather than detachment from them. Recursive improvement will continue to occur at the level of interfaces, affordances, and collective practices, even as computational tools become more powerful. The challenge is not to prevent or unleash recursive self-improvement as a singular event, but to understand and shape the environments in which it unfolds.

Recursive self-improvement, properly understood, is not an anomaly awaiting formal proof or refutation. It is a pervasive feature of irreversible systems that accumulate constraints in ways that favor further accumulation. From minerals to microbes, from swarms to societies, and from laboratories to learning machines, the same physical principles apply. Artificial intelligence does not stand apart from this history; it inherits it.

References

- [1] Good, I. J. (1966). Speculations concerning the first ultraintelligent machine. *Advances in Computers*, 6, 31–88.
- [2] Prigogine, I. (1977). *Self-Organization in Nonequilibrium Systems*. Wiley, New York.
- [3] Rice, H. G. (1953). Classes of recursively enumerable sets and their decision problems. *Transactions of the American Mathematical Society*, 74(2), 358–366.
- [4] Löb, M. H. (1955). Solution of a problem of Leon Henkin. *Journal of Symbolic Logic*, 20(2), 115–118.
- [5] Yudkowsky, E. (2008). Recursive self-improvement. *LessWrong*, December 1, 2008.
- [6] Yampolskiy, R. V. (2015). From seed AI to technological singularity via recursively self-improving software. *arXiv preprint arXiv:1502.06512*.
- [7] Hazen, R. M. (2013). Paleomineralogy of the Hadean Eon: A preliminary species list. *American Journal of Science*, 313(9), 807–843.
- [8] Reid, C. R., and Latty, T. (2016). Collective behaviour and swarm intelligence in slime moulds. *FEMS Microbiology Reviews*, 40(6), 798–806.
- [9] Bonabeau, E., Dorigo, M., and Theraulaz, G. (1999). *Swarm Intelligence: From Natural to Artificial Systems*. Oxford University Press, New York.
- [10] Camazine, S., et al. (2001). *Self-Organization in Biological Systems*. Princeton University Press, Princeton.
- [11] Bostrom, N. (2014). *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press, Oxford.
- [12] Hutter, M. (2012). Can intelligence explode? *Journal of Consciousness Studies*, 19(1–2), 1–2.