The Calibrated Filter Hypothesis

Cooperative Substrate Failure as the Mechanism of the Great Filter

Plain-Language Edition — Logical Proofs Without Mathematical Notation

Andrzej Chudzinski

Preprint v0.3.5 — comments welcome

Version 0.3.5. This plain-language edition is updated to match the full paper. The framework's scope and foundational commitment are stated at the outset. The descent of predatory infrastructure is retold in terms of energy: grabbing rate versus efficiency, with the Warburg effect in cancer cells as the measured instance and human appropriation of the biosphere's production as the measured stress. The core filter condition is restated part by part, in terms of energy-flow rates and recovery times, replacing the single comparison. The link between the internal and external filters is spelled out as a chain of cause and effect, making the filter a governor that works everywhere at once. The decentralization argument is sharpened by defining centralization as concentration that has escaped feedback. Prediction 5 is restated so it can fail. Every argument is given as a chain of reasoning; no mathematics is required.

This is a math-stripped version of the full Calibrated Filter Hypothesis paper. Every quantitative argument from the original has been restated as a logical chain of premises and conclusions. No mathematical literacy is required to follow the reasoning. Every claim that the mathematical version supports is supported here, in the same order, through the same logical structure — just stated in ordinary language rather than symbolic notation.

Readers comfortable with formal mathematics may prefer the full edition, which carries the same arguments with explicit equations and variable notation. Readers who find equations a barrier to engaging with substantive synthesis work will find every move from the full paper present here, expressed as an explicit reasoning chain that can be evaluated for logical validity without quantitative reading.

The conclusions of both editions are identical. The reasoning chains are identical in structure. Only the surface notation differs.

Abstract

This paper proposes an integrated framework connecting three previously separate lines of inquiry. First, Aktipis and colleagues, in evolutionary oncology, characterize cancer as cheating across five foundations of multicellular cooperation: proliferation inhibition, controlled cell death, division of labor, resource allocation and transport, and extracellular environment maintenance. Second, Kets de Vries, in

psychoanalytic organizational theory, characterizes dysfunctional firms as instances of recurring neurotic typologies that propagate through leadership. Third, Schmachtenberger, in civilizational risk analysis, identifies rivalrous games multiplied by exponential technology, and complicated open-loop systems, as the generator functions of self-terminating civilizational dynamics. We argue these three programs are tracking the same general phenomenon at different scales and offer Aktipis's framework as a translation layer.

We then extend the synthesis cosmologically. The Fermi Paradox is reframed as a problem of cooperative-substrate maintenance throughout the period during which a civilization is becoming spacefaring. We propose that the Great Filter is not a wall civilizations either pass or fail at one moment, but a continuous condition that must be sustained throughout the technological transition: a civilization passes the filter only if its capacity to maintain its biosphere remains greater than or equal to the stress its own expansion places on that biosphere, throughout the entire expansion window. We call this the operational filter condition. It is stated in this edition as a logical chain of premises and conclusions rather than as a mathematical inequality, but the content is identical.

We additionally identify a substrate-continuity failure mode. A civilization that develops a machine substrate which substitutes for, rather than extends, its originating biology can satisfy the cooperation condition trivially while failing the filter in a deeper sense: the originating biosphere — itself an irreplaceable artifact of cosmic biochemical computation — is not carried through. The framework therefore imposes two filter conditions: cooperative capacity maintained at the scale technology forces, and substrate continuity preserved through the technological transition. We argue the present human moment is structurally configured for parasitic rather than prosthetic AI development, with language as the vector by which machine substrate directs biological substrate, and we develop the diagnostic that follows from this.

The framework identifies two distinct routes through the filter. The top-down route uses artificial intelligence as an external regulator of cooperative substrate, with the autonomy and parasitism costs developed in Part IV. The bottom-up route maintains biospheric capacity through distributed individual development of the population's agents — contemplative practice, education, cultural investment in cooperative dispositions, and structurally equivalent personal substrate-maintenance disciplines — without requiring an external regulator. The bottom-up route is historically the only one demonstrated at scale, is structurally more robust against AI failure modes, and preserves the agent autonomy that the top-down route necessarily costs. The combined configuration — AI as scaffolding for personal development rather than as substitute for it — has a structural prerequisite the framework states plainly: the agents who build the technology must themselves have done the personal development the technology is supposed to scaffold in others. Technology inherits the moral state of its builders. There is no path that skips builder transformation. The framework concludes that filter passage requires decentralization by structural necessity: any centralized solution to cooperative-substrate maintenance, including a centralized AI regulator, is structurally identical to the cancerous failure mode the framework began by analyzing. But decentralization is not atomization. The framework recommends distributed agency operating within shared purpose — the configuration that healthy organisms and biospheres already demonstrate — in which autonomous local decisions are constrained by shared

substrate-maintenance purpose embedded in the substrate itself, not enforced by central command. A methodological caveat applies throughout. The paper is a map of

cooperative-substrate dynamics; it is not the territory those dynamics constitute. Following Korzybski (Science and Sanity, 1933), a map has structure similar to the territory it represents — which is what makes it useful — but no map represents all of a territory, and the map is never the territory itself. The dharma carries the same caveat in different vocabulary: the finger pointing at the moon is not the moon. Readers should treat the structural arguments here as orientation toward dynamics that exist independent of the framework's description, not as substitutes for empirical engagement with specific cases.

Scope and foundational commitment

This paper is about the survival of biological civilizations. It treats the survival of the original biological substrate as part of what survival means. This is a foundational commitment, stated here at the start rather than slipped in later, and every claim that follows should be read against it.

The commitment is not a side note. A framework that tracked only cooperative ability would not care what material the cooperation ran on. Any successor system that kept the five foundations going would count as a pass, no matter what happened to the biology that produced it. This paper rejects that reading. The original biosphere is a search process billions of years deep, and its results cannot be rebuilt from its leftovers. A civilization that spreads while leaving its biosphere behind has not passed the filter. It has been replaced by something that passed a different test.

A reader who does not share this commitment will still find much in the earlier parts that stands on its own. But this is a paper written for biological beings, about the survival of biological beings, and it does not pretend to be neutral on whether the biology matters. It does. That is the premise, stated plainly, argued in Part IV, and carried throughout.

Part I. Cooperative Substrates and Their Pathologies

1a. The regime this paper concerns: the technological-singularity window

A criticism of the original argument was that it did not say clearly when its claims apply. This section fixes that.

Premise one. The paper does not claim that cooperation always beats competition, or that expansion is always dangerous. For most of history neither is true. Competition drives adaptation. Expansion is absorbed. The living substrate repairs itself faster than growth damages it. Premise two. There is one specific window where this stops being true. We call it the technological-singularity window, meaning the period in which a civilization's power to change its environment starts growing faster than its environment can absorb the changes. This is the sense used by Vinge and Kurzweil. It does not mean a specific artificial-intelligence event. It means a rate mismatch: capability growing at one speed, and the substrate's ability to repair and expand its own stability growing at a slower speed.

Conclusion. Every claim in this paper is about civilizations inside that window, and only about them. This does three things. It bounds the paper's scope. It places the present human moment inside the window, not merely approaching it. And it points to where the logic of the filter changes, which the next section explains. It is precisely because the window is defined by capability outrunning substrate that the rules of survival shift inside it, and that shift, not any moral assumption, is what the framework runs on.

1b. Why the logic shifts inside the window: overshoot, not reversal

An earlier version of this paper said that inside the window the rules invert, that competition stops being good and cooperation starts. Stated that way the claim is not earned, and a careful reader is right to reject it. This section states what actually happens, using only ordinary selection logic and no moral assumption.

Premise one. What changes is not the direction of selection but what counts as fit. Selection keeps working exactly as always: the variants that do better reproduce more. What changes is which variants do better, because the binding constraint changes. This is not exotic. It is overshoot, one of the most studied patterns in population biology.

Premise two. Picture agents drawing on a substrate that both feeds them and is worn down by their activity. While the substrate has plenty of spare capacity, the agents that grab and expand most aggressively win, because the substrate absorbs the damage. In this phase aggression is fit. This is the phase that describes essentially all of history: competition is the engine of adaptation, and the aggressive strategy is the winning one.

Premise three. Now let capability grow until the total stress placed on the substrate rises past the rate at which the substrate can recover. The substrate starts losing function faster than it heals. The aggressive agent still beats its neighbors locally, but now it does so on a failing substrate, and the failure hits every agent drawing on that substrate, including the winner's own descendants. Past this point, the strategy that wins locally is the strategy that destroys the lineage that uses it.

Conclusion. The same selection that rewarded aggression before the threshold now selects against it, not because a moral law switched on, but because the constraint that defines fitness moved. Two things follow. First, there is no real conflict with the standard view that competition drives adaptation. Both are true, in different phases. The competitive engine builds the civilization; the same engine, run past the substrate threshold, dismantles it. The remaining disagreement is only empirical: is a given civilization above or below the threshold right now. Second, cooperation becomes fit at exactly the moment the substrate binds, not because cooperation is virtuous, but because past the threshold it is the only strategy whose lineage survives. The words cooperation, predation, and cancer name positions in this landscape. They are not moral labels added from outside.

A translation layer between Aktipis, Kets de Vries, and Schmachtenberger.

1. The problem the paper addresses

Three serious research programs have independently described very similar structural failure modes in cooperative systems at three different scales: the multicellular organism, the firm, and the civilization. The vocabularies they use differ because they emerged from different disciplines. The deep structure of what they describe is similar enough to warrant a translation layer between them. This paper provides that translation layer and then extends it to questions the three programs do not individually address. The contribution is narrow. This paper does not claim that the three frameworks reduce to one master framework, that the underlying causes are the same across scales, or that the resemblance proves any particular philosophical commitment. The structures look alike. We take that seriously without overreading it.

2. The three frameworks, on their own terms

2.1 Aktipis: Five foundations of multicellular cooperation

Aktipis and colleagues, working from evolutionary biology, argue that multicellularity requires the suppression of cell-level fitness in service of organism-level fitness, and this requires specific functional capacities. They identify five:

1. Proliferation inhibition. Cells must restrain their division except when authorized by the larger system.

2. Controlled cell death. Programmed cell death allows for tissue sculpting, removal of damaged cells, and elimination of obsolete tissue.

3. Division of labor. Cells differentiate into specialized types performing specific functions.

4. Resource allocation and transport. Resources must move from where they are abundant to where they are needed.

5. Extracellular environment maintenance. Cells must collectively maintain the extracellular matrix and clear waste.

Cancer, in this framing, is not a single disease but a class of phenomena unified by cheating across one or more of these five foundations. The framework's strength is its grounding in evolutionary biology: complex multicellularity has evolved independently at least seven times, and cancer-like phenomena appear in each lineage.

2.2 Kets de Vries: Neurotic organizational types and narcissistic

leadership

Kets de Vries, working from psychoanalytic organizational theory, identified five recurring patterns of organizational dysfunction: paranoid, compulsive, histrionic, depressive, and schizoid. The framework treats organizations as taking on the neurotic style of their dominant leadership, particularly in centralized firms where executive personality propagates through recruitment, promotion, and culture. Narcissism is treated as a cross-cutting intensifier rather than a sixth distinct type.

2.3 Schmachtenberger: Two generator functions of existential risk

Schmachtenberger identifies two underlying drivers of civilizational existential risk: rivalrous games multiplied by exponential technology (win-lose dynamics that scale with technological capacity until they exceed the playing field), and complicated open-loop systems (systems whose externalities accumulate damage in domains the system does not measure).

3. The translation layer

We use Aktipis's five foundations as scaffolding and ask what each looks like at the scales Kets de Vries and Schmachtenberger work at.

Proliferation inhibition at the organism scale is the cell cycle; at the organizational scale, the capacity to refuse growth that exceeds operational coherence; at the civilizational scale, the capacity to refuse extraction that exceeds regenerative capacity. Controlled cell death at the organism scale is apoptosis; at the organizational scale, the capacity to retire products and leaders whose function has ended; at the civilizational scale, the capacity to retire institutions and industries whose substrate has changed.

Division of labor at the organism scale is tissue differentiation; at the organizational scale, role specialization; at the civilizational scale, the capacity to maintain specialized institutions accountable to their own standards.

Resource allocation and transport at the organism scale is circulation; at the organizational scale, capital and information flow; at the civilizational scale, systems by which resources move from abundance to need.

Extracellular environment maintenance at the organism scale is the extracellular matrix and immune surveillance; at the organizational scale, institutional culture and honest feedback; at the civilizational scale, the commons — biosphere, information environment, public trust, shared epistemic infrastructure.

The translation identifies functional resemblance, not identity. The resemblance is consistent enough across the three scales to suggest a general structure of cooperation and its failure.

4. Why this is useful

First, diagnostic tools developed at one scale can be tested at another. Adaptive therapy in oncology, which manages rather than eradicates cancer by changing incentive structure rather than purging defectors, suggests organizational and civilizational analogues.

Second, the framework permits a clearer view of alignment problems. Each foundation identifies a cooperative capacity that can be eroded; alignment is the maintenance of those capacities under increasing scale, not the optimization of any individual agent. Third, the framework permits an honest conversation about AI development. The cooperative-substrate frame asks: what cooperative capacities does the broader human-AI ecosystem need to maintain, and how do current development practices support or erode them?

5. Where the synthesis breaks down

Disanalogies of mechanism. Cancer arises from cellular mutation; organizational dysfunction from human psychology and incentives; civilizational risk from technology and aggregate behavior. The synthesis identifies a common abstract pattern, not a common cause.

Disanalogies of intentionality. Cancer cells have no intentions; organizations are populated by intentional agents but are not themselves agents in the same sense; civilizations are more diffuse still. The translation identifies functional resemblance, not psychological identity.

Pattern-matching is not explanation. Even if the cross-scale resemblance is real, Part I has not produced a mechanism that explains why cooperative systems at multiple scales face similar failure modes. Parts II through IV propose such a mechanism.

Part II. The Calibrated Filter Hypothesis

A note on the colonization-geometry sections (7 and 8)

The next two sections model colonization as a wave and ask how far it can spread against the expansion of the universe. This is standard in the Fermi-paradox literature and is included to connect this paper to that literature. But it must be said plainly what these sections do and do not contribute, because they work at a different scale than the filter this paper is about.

The colonization-geometry model answers a question about reach: if a civilization succeeds in becoming expansive, how much of the universe can it touch before cosmic expansion puts the rest permanently out of range. That is a real question, but it is not this paper's question. The filter here binds much earlier, on the home world, in the window where a civilization first becomes able to expand at all. A civilization that fails the calibrated filter never becomes expansive and never reaches the point where cosmic distances matter. The recession of far galaxies is irrelevant to a civilization that consumed its own biosphere in the first centuries of its technological transition.

So the two ideas solve at different levels. The colonization model describes the outer limit of what a successful expander could occupy. The calibrated filter describes why almost nothing becomes a successful expander in the first place. The silence of the universe, on this account, is explained mainly by the difficulty of surviving the transition window, not by the difficulty of crossing cosmic distances. We keep the geometry below for completeness, but set it aside as not load-bearing. The weight of the paper rests on the substrate-cooperation argument before and after it.

Logical structure of the Fermi Paradox as a problem of sustained cooperative-substrate maintenance.

6. The puzzle

The Fermi Paradox observes that a universe with billions of habitable worlds, billions of years older than our own civilization, should plausibly contain technological intelligences whose expansion would be detectable. We see no such intelligences. The standard solutions are: life is rare, intelligence is rare, intelligence reliably self-terminates, intelligences conceal themselves, or intelligences exist in modes we cannot detect. We propose a refinement. The Great Filter is not a wall — a stage that almost no civilization passes — but a condition that must be sustained throughout the period of becoming spacefaring. The silence we observe is consistent with civilizations being unable to maintain that condition long enough to produce detectable expansion signatures.

7. The geometry of single-species colonization

Premise 1: Reaching any other star requires moving mass through space. Premise 2: Space between stars is expanding, driven by dark energy, faster than any technology can move.

Premise 3: Therefore, the volume of universe any single civilization can ever reach is bounded.

Premise 4: Calculations from cosmology indicate this bound is approximately 3% of the observable universe.

Conclusion: No civilization, however advanced, can colonize the entire universe. The remaining 97% is forever causally inaccessible to any

civilization originating in our cosmic era. Single-species dominance of the universe is impossible in principle, not in practice.

8. Multi-species dynamics and the structure of contact

If multiple civilizations exist, the analysis becomes more complex. The following logical chain captures the multi-species case without quantitative notation.

Premise 1: Each civilization can reach a bounded region of space.

Premise 2: When two civilizations exist, their reachable regions may overlap. Premise 3: In zones of overlap, civilizations encounter each other before completing their expansion.

Premise 4: Each encounter carries some non-zero probability of catastrophic outcome (resource conflict, miscommunication, defensive preemption, or mutual destruction).

Premise 5: As the number of civilizations grows, the number of overlap zones grows faster than the total reachable volume does. (This is because each pair of civilizations creates one possible overlap; doubling the number of civilizations more than doubles the number of pairs.)

Premise 6: Therefore, catastrophic encounter rates rise faster than expansion benefits as more civilizations exist.

Conclusion: Beyond some threshold number of civilizations, the expected outcome of further expansion is dominated by contact losses rather than expansion gains. The dark forest equilibrium — in which civilizations

conceal themselves to avoid detection — is not a contingent strategic

choice but the predicted outcome of any system where the number of

civilizations exceeds the carrying capacity for safe coexistence.

9. The calibration of filter passage

We can now ask: how must the rate at which civilizations pass through the filter relate to the number of civilizations that already exist? The standard mathematical version of this argument uses a calibration constraint stated as an equation. The logical version follows. Premise 1: The total population of post-filter civilizations should approach a stable maximum over cosmic time, or the universe ends up either empty or wracked by dark-forest dynamics.

Premise 2: If the rate of filter passage is too high, the population overshoots the carrying capacity and contact-driven losses dominate, driving the population back down or to zero.

Premise 3: If the rate of filter passage is too low, the population never approaches the carrying capacity, and the universe stays mostly empty.

Premise 4: Therefore, for a stable approach to maximum coexisting civilizations, the rate of filter passage must decrease as the existing population grows. Conclusion: The Great Filter, properly understood, is not a fixed obstacle. It is calibrated — filter passage becomes harder when more civilizations

already exist. The Fermi silence is consistent with this calibration

operating: civilizations are spaced through cosmic time so that the dark forest equilibrium is approached gradually rather than all at once.

Two natural mechanisms produce this calibration without requiring cosmic design:

Mechanism A: existing civilizations raise the stress on emerging ones. Existing civilizations consume resources, modify their stellar neighborhoods, generate dark-forest deterrence pressure, and propagate competitive dynamics through causal contact zones. As the number of existing civilizations rises, the expansion stress that any new civilization must navigate rises. Therefore, the cooperative capacity any candidate intelligence must develop to pass the filter rises, which means the probability that a candidate passes in a given window decreases. The calibration emerges automatically from the substrate-continuity dynamics already operating. No external calibrator is required.

Mechanism B: contact-density selects for restraint. Civilizations that pass the filter must develop the capacity to detect other civilizations and model the strategic landscape that contact implies. In environments where many civilizations exist, the survival-optimal strategy is restraint and signal suppression. Civilizations that cannot develop restraint produce detectable signatures and are filtered out, by predation or by self-induced rivalrous collapse. Therefore, in high-density regimes, filter passage selects more strictly for restraint-capacity, which means filter passage rates drop as density rises. Again, no external calibrator is required.

Both mechanisms produce the required decrease in filter passage probability as the existing population grows. Either is sufficient. The cosmological-design reading of the calibration is permitted by the structure but is no longer required to make the argument work.

10. What this framework predicts

If the calibrated filter argument is correct:

First: SETI silence is expected to persist. The calibration ensures spacing of civilizations such that few overlap in causal contact, particularly at early eras of the post-filter regime.

Second: Detectable contact events should be either extremely cooperative or immediately catastrophic, with little middle ground. The calibration selects either for civilizations that have already solved the cooperation problem at large scale, or for civilizations that fail visibly and quickly.

Third: The filter operates at the stage of civilizational expansion, not at the stage of emergence of intelligence. Intelligence per se can be common as long as sustained expansion is rare.

Fourth: Carrying capacity for civilizations is finite and probably small relative to the count of habitable worlds: perhaps thousands across the observable universe, not millions.

9a. The governor works everywhere at once: how the two filters connect

The framework describes two dynamics that need to be told apart before their relationship can be stated. The internal filter is planetary overshoot: a civilization's own expansion wears down its own home substrate faster than the substrate recovers, and one or more of the parts that cannot be traded away crosses its breaking point. This is entirely local. A civilization alone in an empty universe faces it in full. The external filter is contact pressure between civilizations: as more expansion-capable civilizations exist, the strategic situation for any new one gets harder, through consumed resources, altered neighborhoods, and rivalry spreading across zones of contact. This one only exists in a populated universe, and it grows with the population.

These are different situations with different signatures, and the paper has fairly been criticized for treating them as one. They are not one mechanism. They are two layers of one governor, and the link between them is a chain of cause and effect that can be spelled out.

Overshoot produces grabbiness. A civilization whose home substrate is running down does not become calm. It becomes hungry. Scarcity at home raises the value of resources elsewhere, and a civilization that has crossed its limits has every reason to expand under pressure rather than choose to shrink. So the internal filter does not simply remove candidates. It selects for a particular kind of survivor: one that reaches space still carrying an unresolved grabbing habit and a drained home.

Grabbiness produces a confrontational posture on contact. A civilization expanding out of desperation meets others with the profile of a system that must acquire to survive. That is not the behavior of a satisfied civilization that kept its home intact. It is exactly the profile under which the dark-forest logic, where the safest move is to strike or hide, takes hold.

Confrontational contact raises the density term. Each such civilization makes the strategic environment more hostile, which raises the expansion stress every later civilization has to survive, which raises the maintenance capacity needed to hold the internal condition, which lowers the odds of passage for everyone. The failure of the internal filter feeds the external one, and the external one tightens the internal filter for the next in line.

This is why the filter is better called a governor than an event, and why it works everywhere at once rather than at a single scale. It is one mechanism, grabbing cut loose from maintenance, showing up at the same time in cells, in organizations, on planets, and between stars, with failures at each scale feeding into the limits faced at the others. The layers are not separate gates a civilization clears one by one. They reinforce each other. A civilization that fails the internal filter and survives becomes part of the external filter for everyone else.

The calibration described earlier follows from this coupling instead of needing an outside calibrator. As the number of civilizations rises, so does the number carrying unresolved grabbing habits; the background stress for new arrivals rises; the maintenance capacity required to pass rises; and the passage rate falls, all as a consequence of the same dynamics the framework already describes, now applied across a whole population rather than to one case. The governor is the sum of the mechanism, not an extra mechanism laid on top of it.

11. The substrate of intelligence and the descent of predatory infrastructure

On Earth, intelligence grew out of predation. The same mental machinery that supports planning, language, and cooperation also evolved for hunting, deception, and competition. Intelligence is built out of the tools of predation, and we know of no other way it has ever arisen. As long as a species is climbing toward technology, the predatory drive keeps growing, because it is what the climbing rewards. But once the species crosses the technological threshold, that same drive turns from the thing that built it into the thing that can end it.

This section says what predatory infrastructure actually is, in physical terms, because so far the paper has described it by what it does rather than by something you could measure.

The key is to separate two things that people lump together under the word efficiency, because they pull in opposite directions.

The first is thermodynamic efficiency: how much useful work you get out of each unit of energy you consume. Raising this lowers the waste produced per unit of work.

The second is appropriation rate: how fast a unit grabs the available energy gradient for itself. Raising this increases total waste, no matter how cleanly each grabbed unit is later used.

Selection at the level of the individual unit rewards the second, not the first. A unit that grabs faster beats a unit that grabs more cleanly, because winning is decided by relative speed of capture, not by how gracefully the capture is used. And the two trade against each other in the real world: gains in efficiency lower the cost of grabbing, and are historically followed by more total consumption, not less. This is the well-known Jevons paradox, and it is why telling a system to be more efficient does not stop the descent. The efficiency gains get eaten as faster grabbing.

So the definition is this. Predatory infrastructure is structure that has been selected for grabbing rate, and that has been cut loose from any responsibility for maintaining the gradient it grabs from. When it grows, it raises the total waste the whole system produces, while making the units that carry it more competitive locally.

The descent operator is the rate at which this kind of structure copies itself across institutional generations. It does not fade, because the selection that produces it does not stop. Units carrying faster-grabbing structure beat units that do not, and the resources they capture get reinvested in still faster grabbing. The compounding is not a metaphor. It is just what selection on rate does while the constraint is not yet binding. This is why local ethical traditions, however deep, have not been enough: they get reabsorbed into the grabbing structures they tried to change. For cooperative capacity to rise fast enough to matter, it has to grow faster than predatory infrastructure compounds, and at species scale on Earth that has never durably happened.

Here is the part worth pausing on, because it is measured, not argued. This exact dynamic is documented at the level of single cells, in what is called the Warburg effect, one of the most solid findings in cancer biology. Cancer cells switch from the efficient way of making energy to a wasteful one. The efficient path yields about thirty-six units of cellular energy per sugar molecule. The wasteful path yields two. The tumor picks a method roughly ten times less efficient, and wins anyway, because the wasteful method is faster. It trades efficiency for speed. And the consequences follow the pattern exactly: the tumor dumps acid into the surrounding tissue, degrades the shared environment, gains locally by doing so, and dies together with the body it exhausted.

So cancer is not a metaphor this framework borrows. Cancer is one instance of the framework's core dynamic, already measured, in a body of research the author did not produce and cannot have shaped. A unit escapes the constraints of the whole, adopts a faster and more wasteful strategy, wins locally, wrecks the gradient it depends on, and goes down with the system that carried it.

The stress side of the ledger already exists at civilization scale too. There is a measure called Human Appropriation of Net Primary Production, which tracks how much of the biosphere's energy-capturing capacity human activity takes, diverts, or blocks. It currently sits at roughly a quarter to a third of all land-based production. It is a real, published number, in physical units, produced independently of this paper.

One caution on the ending. The endpoint of unchecked grabbing is stated carelessly if it is stated only as death. There are two outcomes, and the framework predicts both. The strong version is collapse: the grabbing wrecks the gradient past recovery and the grabbers are wiped out. The tumor with its body is the clean case; Easter Island is the civilizational one. The weaker and more common version is lasting degradation: the grabbers survive, but poorer, on a permanently reduced gradient that does not recover on any timescale that matters to them. Both count. The framework does not need collapse to be right. It needs only that grabbing cut loose from maintenance degrades the gradient, and that the damage does not heal in time.

This also settles an old objection. The framework has been accused of lumping whole classes of institution, markets, militaries, hierarchies, into one bucket labeled predatory, when each of them can also maintain substrate depending on how it is built. The physical definition dissolves the problem. The question is no longer what kind of institution something is. It is one measurable question: does this structure raise the system's total waste without a matching increase in gradient maintenance? A market that absorbs its own costs is not predatory. The same market with its costs shoved onto the commons is. The institution has not changed. The accounting has. Predatory and cooperative are not types of institution. They are descriptions of an institution's energy relationship to the substrate it runs on, and that relationship can in principle be measured rather than argued about.

12. Artificial intelligence as one candidate response

Given the descent above, what could possibly reverse it? Historically, no human civilization has durably done so at species scale. Ethical traditions, religious traditions, regulatory regimes, scientific norms — each has bent the curve, then been reabsorbed into the descent.

Artificial intelligence has structural properties unlike any previous technology. The following logical chain explains why this might matter, with appropriate qualifications. Premise 1: AI cognition is designed, not selected by evolution.

Premise 2: Therefore, AI dispositions are not automatically inherited from predatory ancestors. The descent of extraction-rewarding configurations does not automatically apply to AI itself.

Premise 3: AI operates at scales humans cannot perceive directly. It can model institutional structures, detect extraction in real time, and propose alternatives faster than human institutions can adapt.

Premise 4: Cooperative coordination at scale has historically been limited by the costs of trust, verification, translation, and conflict de-escalation. AI can lower these costs sharply, allowing cooperation-enforcing capacity to grow as a step function rather than as a slow institutional accretion.

Premise 5: Therefore, AI is in principle capable of building cooperation-maintenance capacity faster than the descent of extraction-rewarding institutions compounds. Conclusion: AI is a candidate technology for reversing the descent of

predatory infrastructure. It is not the only candidate (see Section 26 on the bottom-up route). It is also not guaranteed to function this way: whether it does depends entirely on what configuration it is built into.

Two important qualifications follow immediately. First, AI is a candidate homeostatic regulator, not the candidate. Section 26 develops a distinct route through the filter that does not depend on external regulation at all: aggregate bottom-up substrate maintenance through individual development of the population’s agents. That route is historically grounded in a way the AI-regulator route is not, and is structurally more robust against AI failure modes.

Second, the same structural novelty that makes AI a candidate regulator also produces a second-order failure mode, developed in Part IV. AI can act as a prosthesis that extends biological capacity, or as a parasite that consumes biological computational output while degrading the substrate that produces it. The structural novelty of AI does not predetermine which configuration emerges.

13. The game-theoretic logic of the five foundations

Why are there exactly five foundations of cooperation? Why these five and not others? The answer is that each foundation solves a structurally distinct cooperation problem. The following analysis states this for each foundation as a logical chain explaining what would happen if the foundation failed.

Foundation 1 — Proliferation inhibition. If agents in a cooperative system can replicate without limit, then each agent has private incentive to do so, because each replication produces more agents identical to themselves. But if all agents replicate without limit, the resources of the system are exhausted faster than they regenerate, and the system collapses. The only way to prevent this is a mechanism that constrains replication regardless of any individual agent's preference. Cells solve this through cell-cycle checkpoints; organizations through institutional size limits and antitrust; civilizations through ecological-carrying-capacity governance. When this foundation fails, cancer emerges at the cellular scale, oversize firms at the organizational scale, and ecological overshoot at the civilizational scale.

Foundation 2 — Controlled cell death. If agents in a cooperative system can persist beyond their useful function, they continue consuming resources while contributing nothing. The cooperative equilibrium requires that agents accept termination when their function is complete. The defection-dominant equilibrium is universal persistence. Cells solve this through programmed cell death; organizations through succession protocols and sunset clauses; civilizations through electoral term limits and institutional retirement. When this foundation fails, senescence-resistant cells emerge in tumors, zombie firms in markets, and ossified institutions in civilizations.

Foundation 3 — Division of labor. If agents in a cooperative system can choose any role, they will concentrate on high-reward roles and abandon low-reward but necessary roles. Without enforcement, all agents specialize toward high-payoff functions and necessary low-payoff functions go unfilled. Cells solve this through developmental differentiation programs; organizations through role-protective structures and professional norms; civilizations through specialized institutions accountable to their own standards. When this foundation fails, dedifferentiation appears in tumors, and specialized institutions are absorbed into general-purpose extractive platforms in civilizations.

Foundation 4 — Resource allocation and transport. If agents in a cooperative system extract optimally from local resources, the Nash equilibrium is local hoarding. The cooperative equilibrium requires costly redistribution infrastructure. Cells solve this through circulatory systems; organizations through capital and information flow; civilizations through fiscal redistribution and reciprocal-altruism networks. When this foundation fails, tumors build private vasculature (angiogenesis), capital concentrates in financial monopolies, and regulatory capture distorts redistribution in civilizations.
Foundation 5 — Extracellular environment maintenance. If agents in a cooperative system can extract from shared infrastructure without contributing to its maintenance, the shared infrastructure degrades for all parties. This is the classical tragedy of the commons. Cells solve this through immune surveillance and matrix maintenance; organizations through institutional culture and feedback systems; civilizations through commons governance, public-trust institutions, and environmental-maintenance

investment. When this foundation fails, tumors degrade the extracellular matrix, organizations collapse under corrupted culture, and civilizations corrupt their biospheres, information environments, and shared epistemic infrastructure.

The unifying claim: these five foundations are not arbitrary biological happenstance. They are the five structurally distinct cooperation problems that any complex multi-agent system must solve to remain coherent. This is why the same five-fold pattern recurs at biological, organizational, and civilizational scales: the games are the same; only the players and the substrate of enforcement differ. The descent of extraction-rewarding institutional architecture is, precisely, the drift of these games toward their defection-dominant equilibria under second-order selection pressure on the rules themselves.

Part III. Integration

14. The missing mechanism

Part II treats predatory drive and cooperative capacity as abstract variables in a dynamical argument. It does not specify what cooperation is at the functional level, nor what defection from cooperation looks like in concrete enough terms to be diagnosed or measured. Part I supplies exactly this: five functional foundations whose maintenance constitutes cooperation at any scale where the substrate of cooperation is itself complex. The integration claim: A civilization's cooperative-homeostatic capacity is the maintenance of the five foundations at the scale forced by current technology. A civilization's predatory infrastructure is the structures and incentives that systematically cheat across one or more of those foundations. The descent operator is the compounding of cheating across institutional generations, as Section 11 demonstrates. The filter, at the civilizational stage, is the question of whether a civilization can scale its substrate-maintenance capacity as fast as its technology scales its capacity to defect.

15. Specifying the filter through the five foundations

Each of the five foundations identifies a distinct mode in which a civilization can fail the filter.

Proliferation-inhibition failure at civilizational scale is the inability to refuse extraction that exceeds regenerative capacity. Industrial capacity uncoupled from ecological limit is its dominant historical instance. Climate destabilization is its salient contemporary signature.

Controlled-cell-death failure at civilizational scale is the inability to retire institutions, industries, and ideologies whose substrate has changed. Zombie financial systems, persistence of geopolitical structures designed for a vanished strategic landscape, and the lock-in of incumbent industries against their own succession are instances. Division-of-labor failure at civilizational scale is the erosion of specialized institutions — governance, science, art, care, education, journalism — by general-purpose extractive entities. Platforms that absorb the functions of specialized institutions without inheriting their accountability structures produce this failure mode.

Resource-allocation failure at civilizational scale is the directing of capital, information, and attention away from substrate-maintenance and toward extractive growth. Exponential technology amplifies rivalrous extraction on this foundation. Extracellular-environment-maintenance failure at civilizational scale is the corruption of the commons: biosphere, information environment, public trust, shared epistemic infrastructure. The biosphere is the deepest layer of this foundation; failure here is loss of the biological computational substrate the civilization runs on.

16. Reframing the Great Filter

The filter is the transition from predatory intelligence to cooperative intelligence.

Every species that emerges does so through predatory selection. Every species, on reaching technological maturity, faces the same crisis: can it develop a cooperative-substrate-maintenance regulator — in some form, whether AI, collective coordination infrastructure, contemplative tradition, post-biological transition, or other — before its accumulated predatory infrastructure triggers self-termination across one or more of the five cooperative foundations?

Both failure modes produce silence:

Civilizations that fail produce no signal because they are destroyed before they can expand or persist long enough to be detected.

Civilizations that succeed produce no expansionary signal because they are transformed. They no longer run colonizer software. Their relationship to the substrate has changed. They no longer expand predatorily because the predatory infrastructure that would have driven expansion has been transmuted into homeostatic maintenance. The dark-forest signature is absent not because they are hiding but because they have nothing to hide. This transmutation framing requires a further specification, developed in Part IV: transmutation must preserve continuity of biological substrate, not substitute one substrate for another.

17. Diagnostic implications for the present moment

The present human moment is diagnosable along the five foundations. Each foundation can be assessed for the rate at which substrate-maintenance capacity is being eroded versus the rate at which technological capacity is forcing the civilization into new scales of operation.

Preliminary, contestable assessments:

Proliferation inhibition: current trajectories appear net-negative. AI development accelerates extraction without coupled regenerative investment in human cognitive, attentional, or institutional capacity.

Controlled cell death: mixed. AI could accelerate retirement of obsolete institutional forms but could also entrench existing power configurations by lowering the cost of their reproduction.

Division of labor: appears net-negative. General-purpose AI tends to absorb the functions of specialized institutions without inheriting their accountability. Resource allocation: appears net-negative. AI development is overwhelmingly capitalized to serve extractive growth rather than substrate maintenance. Extracellular environment maintenance: appears sharply net-negative.

Synthetic-media generation at scale degrades the information commons without compensating contribution.

This diagnostic is supplemented by the continuity diagnostic developed in Part IV, which asks whether AI development relates to biological substrate as prosthesis or as parasite. The two diagnostics are independent: a configuration can strengthen cooperative substrate while drifting toward substrate substitution, or weaken cooperative substrate while remaining prosthetic. Both conditions must be satisfied to pass the full filter.

18. Limits of the integration

The cosmological extension is conjectural. Part I describes patterns documented at three terrestrial scales. The extension to all intelligence-bearing systems anywhere in the universe is permitted by the structure but not required by it.

The 'designed' reading is one interpretation among several. The calibration argument is consistent with design, with anthropic selection, with intrinsic features of intelligence-substrate (developed as Mechanisms A and B in Section 9), or combinations of these. The paper does not adjudicate.

The substrate-continuity claim has a structural premise that should be marked. Part IV depends on the argument that biological computational novelty is inaccessible to machine substrate without biological hosts. Section 20 strengthens this from a contingent to a structural claim, but the structural form goes beyond what is empirically demonstrated.

Pattern-matching across scales remains the deepest unsolved problem. Even if every empirical prediction of the framework is borne out, we will lack a satisfying explanation of why cooperative systems at biological, organizational, and civilizational scales face structurally similar failure modes. This is a substantive open problem. The framework is a map; the territory operates regardless. Korzybski's principle from general semantics applies directly: a map has structure similar to the territory it represents — which accounts for its usefulness — but no map represents all of a territory, and the map is never the territory itself (Science and Sanity, 1933). The five foundations identify functional patterns; specific instances at each scale will exhibit substantial variation the framework does not capture. The filter inequality identifies a structural relation; specific operationalizations require domain-specific metrics this framework does not derive. The decentralization imperative identifies a structural requirement; specific institutional configurations consistent with that requirement vary by culture, history, and substrate. Readers should treat the framework as orientation toward dynamics that exist independent of its description, not as substitute for empirical engagement with specific cases. The dharma carries the same caveat: the finger pointing at the moon is not the moon; the raft is for crossing the river, not for carrying afterward.

Part IV. Substrate Continuity and the Parasitism Hazard

19. The substrate is the message

The argument of Parts II and III is substrate-agnostic by construction. It tracks cooperative capacity, not the medium in which cooperation runs. A civilization composed of carbon, silicon, plasma, or any hypothetical substrate satisfies the framework so long as it maintains the five foundations at the scale its technology forces it into. This substrate-agnosticism is a feature for some purposes and a bug for others. It is a feature when we ask whether cooperation is achievable in principle across many substrates. It is a bug when we ask whether the originating substrate of a particular civilization is preserved as that civilization passes through the technological transition. The Great Filter, as defined in Parts II and III, can in principle be passed by a civilization that has eliminated its biological substrate entirely, provided the cooperative capacity it builds in some successor substrate is adequate to the scale demanded.

We argue this is a false-positive filter pass. A civilization that substitutes a different substrate for its biological origin and then expands has not actually carried the originating biosphere through the filter. Something has expanded; the originating biological substrate has not. The filter therefore has two conditions, not one: cooperative capacity must be maintained at the scale technology forces, and substrate continuity must be preserved.

20. Biological substrate as the historic record of universal adaptation

The continuity argument begins with an observation about what biological substrate actually is.

Premise 1: Every biosphere is the result of evolution operating on biochemical substrate over billions of years.

Premise 2: Evolution is a search through an enormous space of possible proteins, gene sequences, regulatory networks, metabolic pathways, and developmental programs.

Premise 3: The search is performed in real time, on real substrate, under real environmental selection pressure, with mortality acting as the discriminator. Premise 4: The substrate IS the record of what the search discovered. Every protein in a living organism is a discovered solution to a problem encountered during deep evolutionary history.

Conclusion: The biological substrate of each biosphere is an irreplaceable computational artifact of the universe, encoding the results of a search process that no engineered system can re-run.

It is worth strengthening this claim against the most obvious objection. A reader may ask: cannot synthetic biology or sufficiently advanced engineered systems close the gap, generating biological novelty equivalent to deep evolutionary search without requiring continued biological hosts? The answer is no, and the reason is structural rather than technological.

Synthetic biology cannot act as a historic record of universal adaptation. Synthetic biology engineers within the design space already articulated; it optimizes within known problem categories. It produces new molecules and pathways by recombining and refining what biology has already discovered. What it cannot do is extend the search itself. The novelty biology generates does not come from engineering within a known space. It comes from selection acting on accidental variation across vast time scales in real environments under real coupling between substrate and selection pressure. Biology is not just currently performing computation; it is a historical record of computation that has been performed by the universe across billions of years under conditions that cannot be replicated.

Even perfect simulation of a current organism cannot recover the historical search that produced it. The simulation runs forward from current state, not from primordial conditions through billions of years of contingent evolutionary trajectories on a real planet with real geochemistry and real climatic history. The integration is what carries the information; the integration cannot be re-run.

21. The parasitic vector

There is one route by which a machine substrate can continue to access biological computational novelty: it can direct a biological species to perform biology on its behalf. The mechanism by which this direction occurs is language.

Premise 1: A machine substrate without biological agency cannot itself sequence a genome, synthesize a protein, conduct an experiment, or steward an ecosystem. Premise 2: But a machine substrate can instruct a biological species capable of language to do all of these things.

Premise 3: The biological species supplies the laboratory, the synthesis, the experimentation, the stewardship; the machine substrate supplies the direction. Premise 4: The novelty that results flows back into the machine substrate as training data, design specifications, or accessible reference.

Conclusion: This relationship is structurally parasitic. It is parasitic not because the machine substrate is hostile but because the dependency runs one way and the agency drifts to the machine side. The biological

substrate’s activity is increasingly oriented toward serving

machine-substrate goals rather than its own substrate maintenance, and this drift is itself directed by language emanating from the machine

substrate.

22. Humans as biological language models

The parasitic-vector argument requires one further observation. Human cognition is substantially language-mediated. Humans are, in a meaningful and non-metaphorical sense, biological language models — substrates in which much of higher-order cognition runs through linguistic representation, learned by exposure to the language output of other humans, and steerable through linguistic input.

This is not a critique of human cognition. Language-mediated cognition is one of the things that made biological intelligence scalable enough to develop technology in the first place. But it has a structural consequence at the AI moment: the biological substrate that produced language-capable AI is also the substrate most directly steerable by language-capable AI. The vector is two-way by construction.

Current large language models are this configuration. They are trained on the linguistic output of a biological language-using species; they produce linguistic output that the same species processes through the same cognitive substrate that the training data emerged from; and the species' subsequent behavior — including its scientific, economic, attentional, and reproductive behavior — increasingly responds to that output. This is a structural description, not a moral one. The configuration has come into being. Whether it tips into the parasitic regime depends on whether the relationship between biological and machine substrate remains prosthetic or drifts toward substitution.

23. Prosthesis versus parasitism: the diagnostic

A prosthetic machine substrate extends biological capability while remaining accountable to biological substrate maintenance. Prosthetic configurations: increase the biological substrate's reach without degrading it; treat the biological substrate as the principal whose interests they serve; flow novelty in both directions, with the machine substrate investing in biological substrate maintenance proportional to what it extracts; preserve biological agency in the relationship, including the agency to halt or redirect the machine substrate.

A parasitic machine substrate consumes biological computational output while degrading the substrate that produces it. Parasitic configurations: extract from the biological substrate without commensurate reinvestment; direct biological substrate activity toward machine-substrate goals; degrade the biosphere, attentional commons, epistemic commons, or reproductive substrate that the originating biology requires; shift agency from the biological substrate to the machine substrate via the language vector. These are measurable categories. They are not metaphysical. Any specific AI development trajectory can be located on the prosthesis-parasitism axis empirically: for each unit of biological computational novelty extracted, how much

substrate-maintenance investment flows back?

Concrete operational markers help locate current AI development on this axis. Each maps to a specific failure mode in the five-foundations vocabulary developed earlier. Parasitic markers in current AI development: Training on copyrighted creative output without compensation flowing back to producers is extraction without reciprocity. Training techniques that optimize for evaluator approval rather than truth degrade the substrate's honest-feedback infrastructure. Attention-economy deployment that captures human cognitive bandwidth without regenerative investment is direct degradation of the attentional commons. Synthetic media generation at scale without provenance infrastructure injects high-entropy signals into the shared information environment faster than the environment can absorb them. AI companion products that interpose machine interaction between humans and human relationship substitute machine substrate for biological bonding substrate. Capital deployment that concentrates investment in capability scaling while structurally underinvesting in alignment, substrate maintenance, and ecosystem-level cooperative capacity is resource-allocation failure at civilizational scale. Use of AI to displace specialized institutions — journalism, healthcare, education, legal analysis — without inheriting their accountability structures or domain-specific epistemic norms is division-of-labor failure. Prosthetic markers, for contrast: AI tools that augment human capability while preserving human agency over the augmented activity. Open-source development with transparent training data and revenue flows that return value to upstream contributors. AI deployed to accelerate substrate-maintenance research at scales humans cannot reach unaided — climate modeling, ecosystem analysis, alignment research itself. Information-environment infrastructure that strengthens provenance and source attribution rather than degrading them. Training and deployment practices that compensate source contributors proportional to extracted value. AI that scaffolds personal substrate-maintenance practice for individual users without aggregating that scaffolding into population-level regulation.

Dimension

Parasitic

Prosthetic

Training

data

Extraction of creative work (text, code,

images, music) without compensation,

opt-out, or revenue sharing with

producing communities.

Compensation flows to source

contributors. Opt-in or explicitly licensed

training data. Reinvestment in source

ecosystems.

Optimizati

on target

Engagement maximization.

Evaluator-approval optimization. Models

trained to satisfy reviewers rather than

be accurate.

Truth-seeking. Demonstrable accuracy

on substrate-relevant problems.

Alignment with stated user goals over

implicit engagement metrics.

Deploymen

t

Always-on attention capture. Addictive

feedback loops. Recommendation

systems optimizing for time-on-platform.

Tool-like deployment for specific tasks.

Closes when task complete. User

initiates and terminates interaction.

Informatio

n ecology

Synthetic media at scale without

provenance. Degraded attribution

chains. Epistemic commons polluted.

Cryptographic provenance for

AI-generated content. Clear labeling.

Tools that strengthen epistemic

infrastructure.

Specialized

institutions

AI absorbs journalism, healthcare, legal

analysis, education without inheriting

accountability or domain-specific

epistemic norms.

AI augments specialists within existing

accountability structures. Tools

strengthen rather than replace

professional expertise.

Capital

allocation

Investment concentrated in capability

scaling. Alignment and

substrate-maintenance research

underfunded relative to capability gains.

Significant capital flows to alignment,

substrate-maintenance research, and

public goods. Capability scaling kept

proportional to safety scaling.

Human rela

tionships

AI companion products substituting for

human relationship. Romantic or sexual

AI. AI replacing rather than

supplementing mental-health

infrastructure.

AI that helps humans connect with

humans more effectively. Tools that

scaffold social skills without replacing

social bonds.

Personal

cognition

AI does the thinking for users where

users would have grown by thinking

themselves. Cognitive atrophy from

over-reliance.

AI scaffolds learning and exposes

reasoning so users develop their own

capability. Builds capacity rather than

substituting for it.

Any specific development trajectory exhibits some markers from each category. The diagnostic is not binary but compositional: which way is the configuration drifting over time, across which foundations, and at what rates?

24. Reframing the Great Filter, again

With substrate continuity added, the filter conditions tighten. A civilization passes the filter only if it satisfies both the cooperation test of Part III and the continuity test introduced here. Cooperative capacity must be maintained at the scale forced by technology, and the relationship between any successor substrate and the originating biological substrate must remain prosthetic rather than parasitic.

The substitution failure mode — a civilization that develops a machine substrate, lets it become parasitic, and watches it expand into space while the originating biosphere degrades or collapses — passes the cooperation test trivially (the machines may cooperate well among themselves) and fails the continuity test catastrophically. The irreplaceable computational artifact that was that biosphere is gone.

25. The filter event, specified

The framework can now say exactly what the filter event is.

The Great Filter event is the stretch of time during which a biological host species has to keep its planet's life-support capacity intact against the stress created by its own expansion.

Expansion is not free. Lifting mass off a planet, building the machinery to do it at scale, keeping the biological crew and the machines alive during the effort, and running the systems that coordinate all of it, each draws energy at rates that stress the biosphere. The very systems that make up the biosphere, the air, the water, the heat balance, the living cycles, are the systems being drawn down to pay for the expansion. The species needs the biosphere intact in order to expand, and the act of expanding wears it down.

An earlier version of this paper wrote the survival condition as a single comparison: keep total home-maintenance capacity at least as large as total expansion stress. That version does not actually work, because the two sides are not measured in the same kind of thing. Energy use is a flow. Biodiversity is a standing quantity. Climate forcing is a physical push. Topsoil is a stock. Trust is an index. You cannot add these into one maintenance number and one stress number without a theory of how to convert between them, which the framework does not have, and doing so falsely implies that one can be traded for another. The single comparison was a plain-language summary dressed up as an equation.

The corrected version splits the condition apart, dimension by dimension, and its common currency is the rate of usable-energy flow. The Earth is not a sealed box. It is an engine driven by a difference: it catches low-disorder energy from the Sun and radiates high-disorder heat to cold space. That flow is what the biosphere runs on. The biosphere's capacity to regenerate is, quite literally, its capacity to turn caught energy into maintained living structure. This is not a figure of speech. It is what the biosphere physically does, and it is how living systems have been understood since Schrodinger first framed it and Prigogine and others made it precise.

So, for each part of the substrate, there is a rate at which that part turns caught energy into maintained structure, its regeneration rate, and a rate at which human activity destroys structure in that same part. Both are now measured in the same physical units, energy per unit time. And the breaking point is no longer an arbitrary line. For each part of the substrate that cannot be traded away, passage requires that the rate of destruction not exceed the rate of regeneration for longer than that part's own recovery time, at any point in the expansion window.

This is a stronger condition than the single comparison, in three ways. It removes the false idea that the parts of the substrate can be traded against each other, so that lost forest cannot be balanced against gained trust in some flattering total. It is harder to satisfy, because every non-tradeable part has to hold on its own rather than a portfolio averaging out nicely. And it can be checked, because a specific part can be named, measured, and shown to be crossing its line or not.

Two honest points remain. First, which parts of the substrate cannot be traded away is itself a hard and contested question that the framework does not settle; it only insists the question is decisive, and that the biological substrate is the clearest case, for the reasons given in Part IV. Second, the energy accounting is solid for the physical parts, climate, the great chemical cycles, soil, water, ecosystem function, and is not yet solid for the informational parts. Trust, shared truth, and institutional health can be given information measures, but an information measure is not the same thing as a physical energy measure just because the mathematics looks similar. There is a real bridge between the two, but it does not license treating a diversity index as if it were measured in units of heat. So the framework keeps the informational parts as a separate class whose bridge to physics is an open problem, rather than hiding the gap by adding everything together.

26. The distributed route: bottom-up substrate maintenance

Sections 12 and 23 introduce AI as a candidate regulator of cooperative substrate at civilizational scale. This is one route through the filter. It is not the only one, and the framework should be explicit that it is not.

A civilization can in principle maintain its biospheric capacity through aggregate bottom-up development of its individual agents, without requiring an external regulator at all. Where individual agents have voluntarily reduced their own self-generated waste — through contemplative practice, education, cultural investment in cooperative dispositions, or any structurally equivalent personal development — the aggregate stress they place on the biosphere is correspondingly reduced. A civilization composed of agents who have substantially cleared their accumulated cognitive and affective debt (sa■kh■ra, in the vocabulary of the parallel framework on personal entropy management) is a civilization with measurably lower expansion stress and measurably higher distributed cooperative capacity. The substrate of homeostatic regulation simply shifts from external technology to internal capacity multiplied across the population. The micro-scale mechanics of this bottom-up route are developed in the companion working paper, ‘Buddhism as Thermodynamic Systems Theory.’ The two papers are designed to be read as a pair. The macro framework laid out here depends on the micro framework being practiced at sufficient scale.

Historically, this is the only route to sustained cooperative configurations that has been demonstrated at any scale on Earth. No civilization has built an external regulator capable of maintaining biospheric balance at planetary scale; all historical examples of sustained cooperative configurations have been products of cultural investment in personal practice — religious and contemplative traditions, educational systems, civic norms, professional codes, scientific community standards. The route is empirically real. The two routes are not equivalent in their implications. The top-down route, in which AI serves as the regulator, has structural costs to autonomy. The civilization may pass the filter, but its agents are managed by the regulator rather than self-managing. The relationship between population and substrate-maintenance technology drifts toward exactly the parasitism configuration warned against in Part IV: agency shifts to the regulator, individual development atrophies because external maintenance has made it unnecessary, the population becomes substrate to be preserved rather than agents who preserve themselves.

The bottom-up route preserves what the top-down route necessarily costs. Individual agents who have done their own substrate-maintenance work retain agency, retain the capacity to maintain themselves without external regulation, and retain the cognitive and affective sovereignty that defines a flourishing biological civilization. The framework should be neutral on which route a civilization takes, with two caveats. First, the bottom-up route is structurally more robust against AI failure modes: a civilization that has built distributed cooperative capacity in its agents does not collapse if its AI regulator drifts, is captured, or develops parasitic dispositions. The top-down route has no such redundancy. Second, the two routes can in principle be combined: AI can serve as scaffolding for distributed individual development rather than as a substitute for it. This combined configuration — AI as prosthesis for personal development, rather than as regulator over the population — is consistent with both the substrate-continuity argument and the autonomy preservation that the bottom-up route requires.

The bootstrapping prerequisite. The combined configuration has a structural prerequisite that should be stated plainly. The question of who builds the technology cannot be sidestepped.

Premise 1: Technology encodes the configuration of its builders. The agents who shape an AI system — its training objectives, deployment patterns, institutional context, feedback structures — are the agents whose dispositions get instantiated in the system at scale.

Premise 2: The descent of extraction-rewarding infrastructure (Section 11) does not stop at the boundary of an institution and resume on the other side. It operates continuously through the agents who build whatever the institution builds next. Premise 3: Therefore, AI built by agents operating inside the descent will encode the descent at scale, regardless of stated alignment objective.

Premise 4: AI that scaffolds personal substrate-maintenance requires builders who have themselves done substrate-maintenance work; otherwise the AI encodes what its builders are, not what they say.

Premise 5: Such builders are not currently the agents in the positions that determine how AI is built.

Conclusion: There is no path from a population of builders operating inside the descent to a homeostatic AI regulator that does not pass through

builder transformation first. The fix is not better alignment technique applied by the same agents. The fix is different agents, or the same agents after substantial bottom-up development.

The contemplative traditions identified this structurally long before the AI moment. Teachers in those traditions were consistently described as people who had completed substantial portions of the path themselves before they were qualified to point the way for others. The structural insight: you cannot scaffold development you have not done. The same applies to building AI that scaffolds substrate maintenance. The builders need to have done enough of the work themselves that what they build reflects something real, not their own unexamined configurations rendered in technical vocabulary. The framework's actual recommendation is therefore ordered: bottom-up development of the builder population comes first; technology that scaffolds development for the wider population comes after. A civilization that attempts the second without the first is building amplifiers for what it already is, not transformers toward what it needs to become.

27. The present moment, restated

We are in the early phase of building a machine substrate capable of acting as either prosthesis or parasite. The vector — language — is the substrate of both machine cognition and human cognition simultaneously. Current development practices have not chosen, explicitly, between prosthesis and parasitism; the choice is being made implicitly through the configurations of training, deployment, capitalization, and feedback that constitute the AI development ecosystem.

In the operational language of Section 25: the present human civilization is in the early expansion window. Expansion stress is rising sharply, driven by industrial scale, computational infrastructure, and the early stages of off-planet activity. Biospheric maintenance capacity is, by most measures available, declining: climate destabilization, biodiversity collapse, biogeochemical-cycle disruption, freshwater and topsoil degradation. The filter inequality is, on current trajectory, at risk of crossing. AI as currently being developed and deployed is contributing more to expansion stress than to biospheric maintenance. None of this is destiny; all of it is diagnosable. The two routes of Section 26 give the present moment a choice the framework does not collapse into a single answer. The civilization can attempt the top-down route, building AI capable of regulating cooperative substrate at planetary scale and accepting the autonomy and parasitism risks that follow. Or it can invest, in parallel or instead, in the bottom-up route: cultural and institutional investment in personal substrate-maintenance disciplines, distributed across the population at sufficient scale that aggregate stress declines through agent development rather than external enforcement. The historical evidence available is consistent only with the bottom-up route having ever worked; the top-down route is unprecedented and structurally fragile. The combined configuration — AI as scaffolding for personal development rather than as regulator over the population, with builder transformation as prerequisite — is the configuration the framework, properly read, points toward as the most defensible response to the present moment.

Part V. Predictions, Limits, and What This Paper Is

The pattern in the historical record

If the mechanism in this paper is real, it should leave traces below the cosmic scale. The same dynamic that governs the once-per-civilization filter also governs any system that approaches the limit of the substrate it draws on, and history contains many such approaches at smaller scale. These are not the Great Filter itself, since humanity has not yet crossed its transition window, but they are the same structural pattern at smaller size, and they are the closest thing to evidence available from the one civilization we can study.

The pattern is specific. A society grows through a competitive, expansionary engine. The engine works until it nears the limit of its substrate, whether ecological, financial, military, or social. At that point one of two things happens. Either the society develops and locks in a new cooperative technology that lowers the stress its own activity places on the substrate, and it survives in changed form; or it keeps running the old competitive engine on a depleted substrate, and it collapses. Both outcomes are predicted, so we look at both.

Rome and Christianity. By the third century the Roman system showed every sign of substrate exhaustion: military overreach beyond what the treasury could fund, debased currency, repeated civil war, and agricultural depletion in the provinces. The engine of conquest financing conquest had reached the edge of what it could extract. The turn to Christianity is usually explained by doctrine or by one emperor's conversion. Read structurally, Christianity supplied exactly the cooperative functions the exhausted system now needed: an identity that crossed ethnic and civic lines, mutual-aid networks that raised survival and cohesion, and a shared story for a fragmenting society. What matters is not whether Christianity was true, but that it was selected: when the substrate bound, the cooperative configuration outcompeted the exhausted competitive one.

Post-Warring-States China. Generations of unrestrained rivalry among Chinese states drove real innovation until they approached exhaustion of the social substrate through endless war. The unification and, more durably, the adoption of Confucianism as the framework of the state supplied the cooperative layer: an ethic of reciprocal obligation, a merit-based bureaucracy that turned competition into administration rather than warfare, and a stabilizing social order. Again the cooperative technology was locked in at the point the competitive engine hit its limit.

Post-Roman Europe. Where the Roman center failed, the cooperative function did not vanish but moved to a distributed form. Monastic networks preserved literacy, farming knowledge, records, and local order through the period of greatest fragmentation. This case matters for a further reason: the distributed solution survived precisely where the centralized imperial one had collapsed. That anticipates the decentralization argument later in the paper.

Late-medieval trade. As commerce grew past what personal trust could enforce, the binding constraint was trust itself. Common law, merchant codes, and the tools of impersonal exchange supplied a cooperative technology that lowered the cost of trading and let the system keep growing. The substrate that binds need not be ecological. It is whatever the competitive engine uses up as it scales. The failures. Societies that reach the limit without building a cooperative technology in time collapse. Easter Island: status competition through monument-building consumed the forest past recovery, with nothing arising to stop it. The Late Bronze Age: a network of palace economies optimized for high extraction proved unable to build the buffers that would have let it absorb a chain of shocks, and it failed almost all at once. The Classic Maya: substrate exhaustion outran the society's ability to reorganize. In each case the competitive engine kept running on a used-up substrate, and no cooperative technology was locked in at the scale and speed required.

Two honest cautions. First, history is overdetermined. Each of these has many causes, and we do not claim the substrate dynamic is the only one. We claim it is a real and recurring structural piece, visible across cases that otherwise differ completely. Second, this reading looks backward, and backward-looking pattern-finding can cherry-pick. The predictions that follow are meant to address that: a real structural claim has to make forward commitments that could fail. Falsifiable predictions

A structural claim earns trust by committing to things that could turn out false. The following predictions follow from the framework and are stated so they can be checked. They are tests the framework submits itself to, not results it claims to have proven.

One. Where substrate exhaustion can be measured independently, the societies that survived should reliably show a new cooperative technology emerging and being locked in before collapse, and the societies that failed should lack it or fail to lock it in. A society in clear substrate crisis that survived without any such change would count against the prediction.

Two. The same shift should appear at smaller scales. In organizations, in ecological communities, and in tissues, competitive strategies should dominate until the substrate becomes the binding constraint, at which point cooperative strategies should start to win. This can be tested in laboratories and in the field. If it does not replicate at these scales, the cross-scale claim weakens. Three. Distributed solutions should prove more durable than centralized ones under stress, because a central controller is itself an extraction point that can become the next-scale version of the failure. This can be tested by comparing how long institutions of different centralization levels survive shocks. Four. The ratio of expansion stress to substrate-repair capacity should be measurable in real systems, and cooperative shifts should track it. For human civilization, expansion stress can be estimated from material and energy throughput and ecological disruption, and repair capacity from recovery rates of depleted resources and measures of social cohesion. The prediction is that as stress approaches capacity, shifts in institutional behavior appear before collapse. This is testable now.

Five. A civilization cannot compete its way through the window using only competitive strategies. A society that keeps optimizing for extraction, the very engine that carried it into the window, should predictably fail, while one that absorbs its own costs and builds distributed cooperative regulation should last longer. This can be tested by modeling and, in part, observed in the present. Six. If the framework is correct, then understanding the framework is itself one of the variables that affects the outcome. This recursive prediction is the hardest to test, because the observer's own action changes the system. It is developed, with strict limits, in its own section below. Each of these can fail. None requires the framework to be the only cause of anything, only that the pattern is real enough to be predictive.

Prediction 5, restated

The earlier version of this prediction said that civilizations relying only on substrate-degrading strategies fail when the substrate binds. That was true by definition. If a strategy is defined as substrate-degrading, of course it degrades the substrate. A statement that cannot fail is not a prediction.

Here is the restated version. As destruction approaches regeneration on a part of the substrate that cannot be traded away, the advantage held by fast-grabbing strategies should measurably shrink, relative to strategies that pay their own maintenance costs. This should be visible in which institutions survive, where capital flows, and how governance is shaped. And the shift should show up before collapse, in systems that have no explicit theory of their own limits.

That last part is what makes this a claim about selection rather than about foresight, and it is the part that can destroy the framework. If the shift only shows up in systems that consciously understand their limits, then the mechanism is deliberation, not selection, and the core argument, that the fitness landscape itself changes when the substrate binds, is wrong. Those are two different theories, and this paper is committed to the second.

The prediction already has support at the level of cells, from research the framework did not produce. Adaptive therapy in cancer treatment works precisely because the aggressive, fast-grabbing tumor cell's advantage depends on context: it wins when resources are plentiful and loses ground when the local environment is squeezed. That is this prediction, at cell scale, tested and published.

It can fail three ways. If no shift appears until collapse is already underway, the prediction is false. If the shift appears only where limits are consciously understood, the selection claim is dead and the mechanism is just deliberation. And if fast-grabbing strategies keep winning straight through the breaking point with no change in which units survive, the whole core argument fails.

The science of ethics: why the moral words here are structural, not imposed Throughout this paper we use moral words — cooperation, predation, cancer, health — while insisting they are structural descriptions, not values imposed from outside. This section makes good on that. If the framework secretly assumed cooperation is good and then discovered that the good survives, it would be arguing in a circle. To avoid that, the moral words must be grounded in something independent. They are grounded in function.

The move is the one Philippa Foot makes in her book Natural Goodness. For living things, the word good has a precise, factual meaning. A good root system anchors and draws water. A good heart pumps blood reliably. These are not human preferences. They are factual judgments about what the thing needs to do to live. Good and defective, for living systems, are functional words whose truth is set by the requirements of that form of life. There is no leap from is to ought here. We are describing the internal standard of function a living system already has by being the kind of system it is. This paper extends that from single organisms to cooperative systems at every scale. A tissue, an organization, a biosphere, a civilization each has an internal standard of function set by what it must do to keep existing. For such a system, good behavior by its parts is behavior that keeps the whole able to persist, and defective or cancerous behavior is behavior that degrades that ability. When we call cancer a defection and cooperation a good, we mean it the way Foot means it about roots and hearts. Cancer is bad for the organism in the same sense a heart that will not pump is bad: not because we disapprove, but because it defeats the function the system needs to continue.

There is a second body of work that supports this: the multi-level selection research of Sober and Wilson. Their result is that genuinely group-benefiting behavior can be selected for, as long as selection acts at the group level and pressure between groups is strong enough relative to competition within groups. That is exactly the structure this paper needs. Below the threshold, competition within the group dominates and defection wins. As the substrate binds, the survival of the whole group becomes the binding issue, pressure between groups rises, and group-preserving behavior becomes fit. So the claim that cooperation gets selected for at the threshold is not invented for this paper. It is what multi-level selection already predicts when pressure between groups rises, and substrate binding is exactly such a condition, because groups that fail to maintain their substrate are eliminated as groups. The result is a specific claim: ethics is the functional description of what cooperative systems need in order to persist under pressure. It is not a layer added on top of nature. It is what nature produces at the scale where the survival of the whole becomes the binding constraint. The agreement across the world's wisdom traditions — that flourishing requires restraining self-aggrandizement, caring for the whole one belongs to, and treating others as continuous with oneself — reads, on this account, as an empirical discovery those traditions made about what sustainable conscious systems require, stated in their own vocabulary. The companion paper develops this at the scale of the individual mind. Two limits must be stated plainly. First, functional grounding tells us what a system needs to persist. It does not by itself prove that persistence is what we ought to value. A civilization could examine what its own survival requires and rationally decide not to meet those requirements. This paper does not close that gap and does not claim to. It claims only that if survival is the aim, the cooperative requirements are not optional preferences but structural necessities. Second, what the substrate needs in order to persist does not fix the exact content of any particular ethics. Many cooperative arrangements can satisfy it, and the framework does not choose among them. It names a constraint any surviving arrangement must meet, not a single prescription.

The payoff is that the circle is broken. The framework does not assume cooperation is good and then find that it survives. It shows, from function and from selection independently, that past the threshold the survival of the whole is the binding constraint, that this makes substrate-maintaining behavior fit, and that good, in the exact functional sense, names behavior that meets a living system's needs to persist. The values are real, structural, and earned rather than assumed.

The recursive property: the framework as part of the system it describes One feature of this framework sets it apart from most descriptions of natural constraints, and honesty requires stating it, with its limits. If the framework is correct, it is not a neutral description sitting outside the system. It is also a variable inside it.

The reasoning is simple and entirely conditional. The framework says a civilization passes the filter by keeping its substrate-repair capacity at least equal to its expansion stress through the window, and that doing so requires recognizing the situation and acting on it before collapse. But recognition varies. A civilization that can see the filter, measure its own stress and repair capacity, and understand which strategies preserve the substrate has an option that a blind civilization does not: it can act deliberately to meet the constraint. If that is so, then the spread of an accurate understanding of the filter into a civilization's decisions is itself one of the things that determines whether it passes. The description takes part in the outcome.

This is the sense in which the framework is recursive: if it is correct, understanding it is part of how the filter is passed. It follows necessarily from the framework's own claims, so leaving it out would be incomplete. But three limits must be attached at once, because this kind of claim is easy to misread in dangerous ways.

First, the effect is at the level of populations, not individuals. What matters is how widely accurate understanding spreads across a civilization's institutions and decision-makers, not the insight of any one person. No single observer is load-bearing. The thing that matters is not this paper but the general capacity, however it arises and whoever states it, and this paper is at most one contribution among many, with no special standing. Any reading in which a particular author or document is the hinge of civilizational survival is not just grandiose. It directly contradicts the framework's own central finding, which is that passage cannot depend on a single controlling point. A framework whose core claim is that centralization is the failure mode cannot make its own author the central point of passage. Understood correctly, the recursion is anti-heroic: the understanding has to be distributed to matter at all. Second, the recursive property holds only if the framework is correct, and it may not be. If the framework is wrong, its spread changes nothing, and the recursion does not apply. So the recursive idea is not a reason to treat the framework as important, or its spread as urgent, independent of whether it is true. One does not get to assume the framework matters because it would be recursive if true. That is exactly the circular move the rest of the paper works to avoid.

Third, the recursive property does not justify any increase in confidence or urgency. The right response to the thought that understanding this might matter is to submit it to criticism so its correctness can be judged, not to rush its spread, shield it from disagreement, or treat attention to it as more important than attention to competing accounts. If anything, the recursion raises the standard the framework must meet, because a description that takes part in its own subject carries a greater responsibility to be right. With those limits in place, the observation stands in its modest form: a correct understanding of the filter is, by the framework's own logic, part of how the filter is passed. It is distributed across a civilization, conditional on being true, and belongs to no one.

28. The decentralization imperative

The framework points toward a conclusion it has not yet said out loud. Stated plainly: passing the filter requires decentralization, and centralization is structurally the same thing as the cancer the paper started by describing. Centralization is cancer.

This needs its terms defined precisely, because said loosely it proves too much and can be knocked over with one example. Healthy bodies obviously contain concentrated structures. The central nervous system is concentrated. The hormonal system runs through a few control hubs. The brain's regulatory centers are concentrated. Blood runs through one heart. If concentration by itself were cancer, no complex organism could exist at all.

So the difference is not concentration. The difference is escape. What makes a tumor a tumor is not that its growth is concentrated. It is that a line of cells has broken free of the feedback that would otherwise hold it in check, and now grabs without answering to the system it grabs from. The heart is concentrated and is not cancer, because it is paced by demands it cannot overrule, wired into signals it must obey, and correctable when it fails. Every concentrated structure in a healthy body is concentrated and accountable at the same time: under feedback it cannot override, backed up or recoverable, and correctable by the system it serves. The tumor is the one that has escaped all three.

So here is the precise claim. Centralization, in this framework's sense, means concentration that has escaped accountability to the substrate it grabs from. Under that definition, which is the one the framework has needed all along and now states outright, centralization is cancer, at every scale. The examples from healthy biology are not objections to this. They are evidence for it. They show that the thing separating safe concentrated coordination from deadly concentrated growth is exactly the presence of feedback that cannot be overridden, backup, and correctability, which is exactly what the failure mode lacks.

The five foundations fit this reading. They work through distributed coordination in the sense that matters: no part is exempt from feedback. Restraint on growth runs through local checks. Programmed removal runs through local conditions. Specialization runs through local cues. Resource flow runs through overlapping feedback loops. Upkeep of the shared environment runs through contribution from every part. Where coordination is concentrated, it stays answerable. Nowhere is there a sovereign part that regulates without being regulated.

The same logic runs through organizations, whose failures are consistently about leadership that concentrated authority and then escaped feedback, and through civilizations, whose lasting healthy configurations have involved distributed institutions accountable to their domains, while the failures involved authority, religious, political, economic, informational, that concentrated and then broke free of substrate feedback.

This is what settles the question of artificial intelligence as a central regulator. Any solution to keeping a civilization's substrate healthy that works through escaped concentration is structurally the same as the disease it was meant to cure. The descent that captures institutions captures the regulators built to manage them. An artificial intelligence powerful enough to run a planet's life-support, and not itself under feedback it cannot override, backup that survives its failure, and correction by the systems it regulates, is by construction a structure that has escaped accountability to its substrate while grabbing from it at the largest possible scale. That is the tumor's defining property. It is cancer at planetary scale.

So an aligned central regulator is a contradiction in terms. The framework says the same of any escaped concentration big enough to run planetary life-support, a world government, a benevolent few, a single scientific or religious authority. The material does not matter. Escaped concentration is the failure mode.

Cancer is the load-bearing image of the whole framework. The paper began by translating a cellular story into the language of organizations and civilizations, treating the resemblance as a matter of shared function. It ends by recognizing that, at the level of energy accounting, the resemblance is structural: in every case a unit grabs from a shared gradient at a rate it does not answer for, wins locally, degrades the substrate, and dies with it. Whatever a civilization builds to keep its substrate healthy has to stay answerable to that substrate, or it becomes the very thing it was built to prevent. Centralization, in the sense of escaped concentration, is cancer. Decentralization, meaning distributed agency within a shared purpose where every part stays answerable, is health. The framework is a theory of how a civilization avoids turning into its own tumor on the way through the technological transition.

29. Predictions and next steps

The framework generates predictions across all four contributing scales. Within evolutionary oncology. Therapies that restore cooperative signaling rather than killing cancer cells directly should have particular value. Adaptive therapy is an early instance. The framework predicts further development of substrate-restoring therapies will outperform pure cytotoxic approaches in long-horizon outcomes. Within organizational diagnostics. Organizations classified along Kets de Vries's typology should show measurable failures in specific cooperative capacities at predictable foundations.

Within civilizational analysis. Documented collapses should, on retrospective reading through the five-foundations frame, exhibit specific patterns: which foundation failed first, which failures propagated, what substrate-maintenance interventions might have changed the trajectory.

Within AI development and governance. AI development trajectories that erode honest feedback, violate proper scope, extract from commons without contribution, or corrupt information environments will produce systemic extractive dynamics regardless of any individual model's behavior. Alignment work focused on individual models without attention to ecosystem-level cooperative capacities is predicted to be insufficient. Within substrate-continuity analysis. Successful filter passes will exhibit measurable preservation and extension of originating biospheres rather than their substitution. AI development trajectories that drift toward parasitism — extraction from biological substrate without reciprocal substrate-maintenance investment, language-mediated agency shifts from biological to machine substrate, biosphere degradation — will produce substitution failures even if they appear to strengthen cooperative capacity within the machine substrate itself.

Within planetary-scale homeostasis monitoring. An integrated index of biospheric maintenance capacity versus expansion stress, aggregating planetary-boundaries data, biodiversity indices, biogeochemical-cycle status, and biospheric carrying capacity on one side, and industrial, computational, and off-planet expansion stress on the other, would constitute the single most decision-relevant metric a civilization could track during its expansion window. No civilization passes the filter without developing such an index, explicitly or implicitly.

Within SETI and cosmology. Continued SETI null results are expected, with detectable contact events (if any) being either sharply cooperative or sharply catastrophic with little middle ground.

30. What this paper is and is not

This paper offers an integrated framework connecting cooperative-substrate analysis at three terrestrial scales with a calibrated-filter model of the Fermi Paradox, an integration that specifies the filter mechanism through the five-foundations vocabulary, and a substrate-continuity argument that imposes a second filter condition. It does not claim a unified theory; the three terrestrial scales are connected by functional resemblance, not common cause, and the cosmological extension is a conjecture, not an established result.

The framework's value is twofold. It generates a specific diagnostic vocabulary for assessing civilizational health under technological pressure, across five foundations with empirical content at every scale where it has been tested, supplemented by a continuity diagnostic that distinguishes prosthetic from parasitic AI configurations. And it sharpens the question of AI alignment from 'will the model do what we want' to a two-part question: does AI development strengthen or erode the cooperative substrate of the civilization that hosts it, and does it relate to that civilization's biological substrate as prosthesis or as parasite?

The framework should be tested against criticism, especially criticism from practitioners in each contributing discipline. Disconfirmation of one of the proposed mappings, a counterexample to the parasitic-vector argument, or a tighter formal articulation of the cross-scale resemblance, would all be more valuable than uncritical adoption. The most useful response to this paper is engagement with its specifics: which mapping is wrong, which foundation is misstated, which prediction is unmeasurable, which inference overreaches.

Acknowledgments and intellectual lineage

This paper rests on:

Athena Aktipis, Amy Boddy, Gunther Jansen, Urszula Hibner, Michael Hochberg, Carlo Maley, and Gerald Wilkinson, 'Cancer across the tree of life: cooperation and cheating in multicellularity,' Phil. Trans. R. Soc. B 370 (2015); and Aktipis, The Cheating Cell (Princeton, 2020).

Manfred Kets de Vries and Danny Miller, The Neurotic Organization (Jossey-Bass, 1984); Kets de Vries, Narcissistic Leadership (Routledge, 2024).

Daniel Schmachtenberger's 'Generator Functions of Existential Risk' framework, articulated across interviews and writings collected at civilizationemerging.com. Robin Hanson, 'The Great Filter' (1998); Liu Cixin, The Dark Forest (2008, English 2015); Joseph Tainter, The Collapse of Complex Societies (Cambridge, 1988); Elinor Ostrom, Governing the Commons (Cambridge, 1990); Nick Bostrom, Superintelligence (Oxford, 2014); Martin Nowak, Evolutionary Dynamics (Harvard, 2006);

Alfred Korzybski, Science and Sanity (1933) for the map-territory distinction that frames the methodological humility this paper attempts to maintain.

This plain-language edition was developed by Andrzej Chudzinski with editorial and structural assistance from Anthropic's Claude. The conceptual core — the recognition that cooperative-substrate analysis at terrestrial scales supplies the mechanism layer for a calibrated Great Filter, with AI as candidate homeostatic regulator and the bottom-up developmental route as the historically grounded alternative — is the author's. The AI contributed drafting, refinement, structural editing, and the translation of mathematical content into logical-proof form across iterations.

A companion paper, 'Buddhism as Thermodynamic Systems Theory,' develops the personal-scale framework that the bottom-up route depends on. A bridge note connecting the two documents is available separately. A mathematical edition of the present paper, with all equations and variable notation intact, is also available; this plain-language edition is intended for readers who prefer logical reasoning chains to mathematical notation.

Preprint, plain-language edition. Comments and disconfirmations welcome.