The Metacrisis as a Degradation of Evolvability: Evolutionary Governance for Reducing Systemic and Global Catastrophic Risk

Matt Boyd[a*,#], Nick Wilson[b]

[a] Adapt Research Ltd, Reefton, New Zealand

[b] Department of Public Health, University of Otago Wellington, Wellington, New Zealand

[#] Current Address: Adapt Research Ltd, 14 Broadway, Reefton, New Zealand

*Corresponding author E-mail: matt@adaptresearchwriting.com

Word count: 9,999 (excluding abstract, tables and references)

Acknowledgments: The authors express great thanks to Nick King and Ajay Gambhir for helpful feedback on an earlier draft, also to attendees at the Society for Risk Analysis conferences in Christchurch (Jan 2026) and Alicante (May 2026) for discussions that helped improve the work.

Abstract

Global catastrophic risks are amplified by systemic risks, which have become entangled as a polycrisis. Exposure to global risk is rising despite decades of awareness and available mitigations. Lacking is second-order understanding of why solutions are not adopted, and at what level to intervene. We identify five scholarly and policy communities, ranging from disaster risk reduction to behavioural science, that each describe parts of this problem. These communities have been converging, often unawares, on the explanatory framework of evolutionary biology in its extended form. We make three claims: (1) Fragile risk producing architectures will recur wherever certain evolutionary dynamics obtain. (2) A deeper metacrisis exists, in the form of degradation of evolvability conditions in the systems meant to mitigate and govern global risks. Evolvability depends on variation, modularity, selective-environment stability, transmission fidelity, and outlaw (defector) suppression. These conditions are degraded by dynamics including multilevel selection, arms races, transmission bias, path dependence and lock-in. Each condition is observable, giving a diagnosable state. (3) Because these dynamics typically reward locally fit but collectively destructive strategies, governance must identify hazards but also reshape the selection environment, including fitness defined in terms of biosphere persistence and human wellbeing, and maintain effective outlaw suppression. Cancer biology supplies the paper's central model of how a higher-level cooperative system can mitigate global risk only by continually constraining lower-level strategies that benefit themselves at the expense of the whole. Biology supplies global risk governance designs, most tractable at regional level. The framework yields falsifiable predictions and pilotable interventions.

Keywords

Evolution, Evolvability, Global Catastrophic Risk, Governance, Metacrisis, Polycrisis, Systemic Risk

I. Introduction: The Monks and the Elephant

Humanity faces a risk landscape more interconnected and volatile than at any point in its history, challenging global risk governance. Direct disaster losses run to an estimated US$200 billion a year, and cascading effects likely push the true figure an order of magnitude higher (UNDRR, 2025). Global organisations acknowledge catastrophic and existential risk (UN General Assembly, 2024), yet by many measures the world grows more exposed, not less (Department for Environment Food Rural Affairs, 2026; Intergovernmental Panel on Climate Change, 2023; Madhav et al., 2023; World Economic Forum, 2026).

The puzzle is not that risk is rising but that it rises despite known mitigations. Major hazards were identified decades ago and their fixes formalised (Meadows, Meadows, Randers, & Behrens, 1972; United Nations, 1997). These include nuclear disarmament, greenhouse gas emissions reduction, health-security investment, lower agricultural intensity, rewilding, and limits to growth and extraction (Nebel, Kling, Willamowski, & Schell, 2024). Humanity is not ignorant of what threatens us or of what would reduce the threat, yet we fail to act at the scale and speed required, and the gap widens. Lacking is second-order causal understanding, *why* we consistently fail to act, and *what* level to intervene.

E. O. Wilson diagnosed humanity's core problem as "Paleolithic emotions, medieval institutions, and godlike technology" (Wilson, 2009). Beneath this mismatch lies a generative process, where institutions, technologies, and norms evolve under selection pressures that reward what succeeds locally and immediately, while modern technology turns once-local costs global. Selection drives fragile, risk-generating architectures and technological scale catastrophically amplifies failures. This is a structural diagnosis, not a moral one.

A second problem is fragmented global risk scholarship (Jehn et al., 2025). Like the parable of blind monks examining an elephant, each risk silo diagnoses a different problem and infers a different solution. National risk assessment and disaster risk reduction address the symptoms of hazard, global catastrophic risk studies establish the magnitude of the tail, systemic-risk and polycrisis scholarship map interconnections and attractor states, and behavioural and evolutionary science has begun to explain the human dynamics beneath them all. Each is partly right, what is missing is integration to reveal the whole elephant, and an account of *why* it exists.

Fragmentation is not merely inconvenient because it drives inadequate response. A framework blind to the mechanisms shared across hazards cannot address them and so treats symptoms one at a time while the mechanisms generate new risks faster than old ones resolve. Risk governance must understand upstream generative dynamics.

This paper argues that at least five fields are converging, independently and informally, on one explanatory framework, namely evolutionary biology in its extended, contemporary form. The convergence shows in shared vocabulary of adaptation, tipping points, and co-evolution, but it has not been systematised, meaning diagnostic and practical responses go unrealised.

We take global catastrophic risk as our primary target; systemic risk names the conditions under which such risk cascades, polycrisis the entanglement of risk, and metacrisis the deeper failure that lets both persist. "Metacrisis" has been used mostly in essays and talks rather than peer-reviewed work (metacrisis.org, 2026). We adopt the term but define it more precisely (see below, also Glossary Table 1, and Extended Glossary in Supplemental File).

This paper makes three contributions: (1) we argue that global risk problems and evolutionary biology share deep structural features. Common dynamics generate risk in domain after domain, supplying the *why* beneath the *what*. (2) We define a metacrisis precisely, as the degradation of the conditions for evolvability, a system's capacity to generate, retain, and select for adaptive solutions, in the systems that govern global risk. (3) We argue that intervention must occur at the level of those conditions, particularly in the suppression of "outlaw" (defector) strategies that gain locally while degrading the whole. Furthermore, we show that biology's account of defector suppression, cancer's evolutionary biology in particular, supplies not just designs governance can miss but a unifying rationale for suppression, important selection architecture, and a repertoire for generating new interventions. For futures studies, it is important not just to imagine and prepare for a range of futures, but to see anticipatory capacity as a feature of governance that is rooted in evolution, ensuring capabilities to generate, retain, and select adaptive responses, that is, ensuring societies remain capable of navigating uncertainty.

This is a theoretical, integrative contribution to the explanatory and governance problems of global catastrophic and systemic risk. We present a framework, illustrate it, and propose directions rather than prescribing policy or formalising models. Each evolvability condition is observable, the framework yields predictions about which interventions may and may not work, and the governance directions derived are open to piloting. The three contributions build in sequence, the evolutionary dynamics that generate risk (Sections III and IV), the metacrisis as degraded evolvability (Sections V and VI), and the suppression-design agenda from biology (Sections VII and VIII).

II. Five Fields, One Problem

At least five scholarly and policy communities address global risk (Table 2). This section briefly characterises what each contributes and what it misses. We argue the fields are converging, often unawares, on a single explanatory need.

National risk assessment (NRA) and disaster risk reduction (DRR) anchored in the Sendai Framework and delivered through national emergency-management agencies (UNDRR, 2015), applies all-hazards methods, likelihood–impact matrices, scenario sets, and national risk registers, to risks of regional and national significance (OECD, 2017). It typically has a historical bias toward natural hazards, response over prevention, near-term and local hazard horizons, and exclusion of cross-border, cascading, and catastrophic tail risks (M. Boyd & Wilson, 2023; Kohler, 2023; UNDRR, 2025). NRA embeds implicit value judgements and often justifies existing arrangements rather than exposing vulnerabilities. Most consequentially, NRA appears institutionally incapable of capturing risks that governance systems generate rather than receive, namely endogenous, emergent, cascading risk (Arnscheidt et al., 2025). The Sendai Midterm Review conceded that disaster impacts are rising despite these frameworks (UNDRR, 2023) and resilience investment remains disproportionate to risk despite strong return-on-investment evidence (UNDRR, 2025).

Global catastrophic and existential risk studies address risks that could overwhelm civilisational response, including nuclear war, engineered pandemics, AI-related catastrophe, and massive volcanic events (Bostrom & Cirkovic, 2008; Ó hÉigeartaigh, 2025). Institutional recognition is beginning, in the UN Pact for the Future (UN General Assembly, 2024), the US Global Catastrophic Risk Management Act (United States Congress, 2022), and the UK House of Lords (House of Lords Select Committee, 2021). The field's strength is its long horizon and willingness to confront tail risks with annualised harm greater than common disasters (Graham, Boyd, Sadler, & Noetel, 2025). Identified weaknesses include hazard-by-hazard analysis (Jehn et al., 2026), uncertain probabilities, techno-solutionism/techno-optimism (Cremer & Kemp, 2023), and thin integration with political and governance realities.

Systemic risk analysis treats the global system as a generator of hazards, not merely a receiver. Systemic risk is characterised by feedback-rich causal networks and nonlinear dynamics (Liu & Renn, 2025). The Cascade Institute's stress–trigger–crisis model separates slow stresses that erode systemic resilience from faster triggers that tip systems into crisis, singling out homogenisation and hyper-connectivity as mutually reinforcing drivers of synchronous system failure (Lawrence et al., 2024). Related work presents a catalogue of 14 trackable stresses across nine global systems (Lawrence, Shipman, & Collins, 2025). Arnscheidt et al. (2025) argue that catastrophic risk arises from emergent properties, interconnectedness, limited diversity, and technological dependence, and name actors extracting local advantage at systemic cost as a primary risk driver, which Section VII examines as the "outlaw" problem. Scheffran formalises a coupled-systems framework in which cascades become self-sustaining once thresholds are crossed (Scheffran, 2025). Other work has operationalised "watchpoints" warning of systemic risk cascades (Oliver et al., 2026), and normative and ethical principles for systemic risk governance (Richardson et al., 2026). What none of this yet systematises is *why* humans keep producing these risk-generating conditions, *why* actors generate destabilising externalities, and *why* suppression mechanisms erode. These are questions about generative mechanisms, not system states.

Polycrisis scholarship appreciates causal interrelatedness and polycrisis has been defined as the causal entanglement of crises across global systems in ways that degrade humanity's prospects (Lawrence et al., 2024). Three synchronising pathways include common stresses, domino effects, and inter-systemic feedbacks. One polycrisis review considered 2,299 publications and analysed 59, synthesising component crises, and recurring systemic drivers (Rakowski et al., 2025). Systemic risk research specifies the conditions under which cascading failure becomes possible, polycrisis research tracks the cascades once entangled (Liu & Renn, 2025), and neither yet explains why governance consistently fails to adapt. The dynamics of polycrisis actively undermine the needed capacities (Søgaard Jørgensen et al., 2026), with calls for major reconfiguration of institutional risk architecture (Gambhir et al., 2026). The field increasingly reaches for dynamical, quasi-evolutionary structure, framing the polycrisis as a cybernetic "complexity gap," a challenge–capacity framework that is, in effect, a functional analogue of evolvability conditions (Sections IV–V) (Hämäläinen, 2025), or model how sociopolitical events and planetary boundaries co-evolve (Bertolami & Elísio, 2026), while empirical evidence shows crises spread non-uniformly and regionally, through "Anthropocene traps" (Delannoy et al., 2025). Structural-demographic pressures decide whether shocks yield resilience or collapse, with climate resilience requiring a science of cultural evolution (Hoyer et al., 2025), due to the co-evolutionary struggle between resilience capacities and the polycrisis (Søgaard Jørgensen et al., 2026). The field's signature advance is recognising the polycrisis as an entangled system with attractor states, the Polycrisis CORE model's "Illiberal Decline" and "Mad Max" basins (Cascade Institute, 2025). The field's signature gap is mapping the elephant precisely without explaining why it is there, or why it persists. Needed are frameworks of the core dynamics, integrated with socio-ecological systems and evolutionary theory (Søgaard Jørgensen et al., 2026).

Behavioural and evolutionary science speaks the needed language. The global ecological crisis is characterised as a “human behavioural crisis” of growth, energy extraction, marketing, and pronatalism rooted in human psychology, where rationality is often the exception (Hagens, 2020; Merz et al., 2023). Little is said about why that psychology persists. King and Jones classify growth, capitalism, and governance failure as proximate drivers emergent from deeper distal traits, invoking Mayr’s proximate/ultimate distinction (Section III) (King & Jones, 2025). Evolutionary game theory explains why locally adaptive strategies become globally destructive, through prestige competition, defection and multipolar traps (Maynard Smith, 1982), popularised as Alexander’s “Moloch” (Alexander, 2014). Cultural-evolution theory treats technologies, norms, and institutions as genuine “evolving populations” (R. Boyd & Richerson, 1985; Henrich, 2016). Others propose humanity is undergoing a major evolutionary transition toward a panhuman superorganism (Hagens, 2020; Lior & McNamara, 2025); Waring et al. describe the Anthropocene as an evolutionary ratchet driven by cultural group selection (Waring, Wood, & Szathmáry, 2023), arguing that the population structure needed to evolve global cooperation is unavailable on one planet. These valuable insights are scattered, not integrated into global risk governance.

A vocabulary pattern runs through these five fields, we separate two kinds of terms: (1) generic complex-systems language (resilience, feedback loops, tipping points), which signals shared systems-thinking; (2) specifically evolutionary language (selection, adaptation, co-evolution, traps and ratchets), recurring across fields with no evolutionary self-conception, signalling latent Darwinism and convergence on evolution (Table 2). When systemic-risk scholars reach for biomimetic design, and suggest many candidate designs be “tested by natural selection” (Liu & Renn, 2025), they approach the right logic but stop at stress-testing existing designs, not at the conditions under which adaptive designs accumulate.

Each field does valuable work, yet none systematises why the same crises, often identified decades ago and technically soluble, keep amplifying. Answering that requires moving from *what* to *why*, and the discipline to distinguish those questions is evolutionary biology.

III. Evolution as Explanatory Framework

Our argument depends on a conception of evolution that is more detailed than school biology and which rejects the loose, damaging metaphors of “social evolution” and Social Darwinism that historically equivocated evolutionary processes with moral merit (Hofstadter, 1944). This section sets it out for a general reader. The extended evolutionary framework is a live, rigorous research programme (Acerbi, 2019; Mesoudi, 2011), and explanatory precision matters for everything that follows.

Evolution by natural selection is not just a theory about genes or organisms. It is an abstract process, arising wherever three conditions hold: heritable variation, differences in success that depend on that variation, and differential reproduction or persistence as a result (Dawkins, 1976; Godfrey-Smith, 2009). Where they hold, a population accumulates the variants that succeed, with no need for foresight, intention, or a designer. This is what Dennett called the substrate neutrality of selection. Evolutionary logic is indifferent to what the variants are made of (Dennett, 1995). Genes are one substrate, organisms another. But firms, institutions, legal codes, professional norms, technologies, and ideas can also be substrates to varying degrees. When we say these "evolve," we are not reaching for analogy but observing the same fundamental process on a different substrate, which is why the norms of a risk-governing agency, or the design lineage of a weapons system, are as open to evolutionary analysis as any biological organism.

Two qualifications prevent overreach, (1) instantiation comes in degrees and continua (Acerbi & Mesoudi, 2015). Godfrey-Smith's (2009) "Darwinian populations" treats paradigm cases, high variation, faithful inheritance, a tight link between trait and success, as one end of a continuum, with marginal cases loosely satisfying the criteria at the other end. Diverse human social, cultural, and technological systems sit at many points, and how strongly they exhibit evolution by natural selection is an empirical, testable, falsifiable question rather than a rhetorical one. Bicycle design is a clean demonstration, where design lineages show the adaptive-radiation signature of biological clades, early diversification, extinction, convergence on dominant forms, that no ultimate designer intended (Lake & Venti, 2007); (2) because human systems are also deliberately designed, telos and the invisible hand of selection operate simultaneously. Since no actor can intend outcomes everywhere all the time, and since the world of human systems is incalculably complex, much of what results is evolved, rather than designed (Acerbi & Mesoudi, 2015; Henrich, 2016).

What makes the framework "extended," and not a story about genes alone (Richerson & Boyd, 2005), is its account of inheritance. Boyd and Richerson (1985) established culture as a second inheritance system with its own transmission dynamics. Jablonka and Lamb widened this to four channels, genetic, epigenetic, behavioural, and symbolic, operating at different speeds and fidelities (Jablonka & Lamb, 2005). These channels run at different rates, we see symbolic inheritance (laws, standards, and institutional DNA), can change within a single political administration, while the psychological dispositions of the humans involved may shift far more slowly. Additionally, evolutionary niche construction theory shows that organisms modify their environments, with ramifications for selection, and bequeath this ecological inheritance (Odling-Smee, Laland, & Feldman, 2003). For example beaver dams are frequently maintained and reoccupied, modifying the behaviour of subsequent generations. In human society, institutions are components of such inherited niches, often selecting for their own continuation, and niches shape the development of psychological and societal traits while evolving themselves (Sterelny, 2012).

Selection may operate at several levels at once (Sober & Wilson, 1998). A strategy adaptive for a gene, cell, individual, or firm may be destructive for the group, ecosystem, or civilisation containing it, the structure behind the "tragedy of the commons" and, at planetary scale, behind atmospheric carbon overload, overfishing or planetary boundary transgressions. Cultural group selection describes how group-level traits are favoured when groups compete (R. Boyd & Richerson, 1985), the mechanism behind much cooperative human achievement and some of its largest failures. The strength of group selection is contested, but evolutionary game theory (Nowak, 2006) and institutional economics reach related conclusions.

Crucially, not all cultural change is adaptive. Drift through change driven by transmission biases such as prestige copying, or conformity, rather than any fitness advantage (Henrich & Gil-White, 2001; Mesoudi, 2011), means the norms of risk-producing and risk-governing institutions can spread or decay for reasons unrelated to their merit. By 'merit' in the global risk context, we mean the tendency to produce adaptations that reduce risks to the biosphere and support human wellbeing. Fast cultural change is not the same as adaptive change and speed can discard hard-won adaptations even while producing new ones.

Several further terms recur below (Glossary in Table 1; Extended Glossary in Supplemental File): the evolutionarily stable strategy, a configuration that resists supplantation once established even when destructive for the whole (Maynard Smith & Price, 1973), as in a nuclear standoff difficult to exit unilaterally; the local fitness optimum, a trap from which improvement requires first getting worse (Wright, 1932), as in carbon lock-in; and the 'evolutionary suicide', a lineage that persists until it destroys the conditions of its own survival (Rankin, Bargum, & Kokko, 2007). Each characterises an outcome no participant would rationally choose.

The framework is descriptive, not normative. Being selected says nothing about whether a strategy is good. Cancer can be evolutionarily successful for those cancer cells that arise inside the body, and that success is precisely the problem for the organism as a whole. We see this with extractive industrial practices or viral disinformation in social media. Instead, selective success is the thing to be explained, and often contained. Also, it is not reductionist. We do not claim humans are their genes, or that institutions evolve identically to organisms. The claim is that a higher-order structure, variation, selection, differential persistence, are shared to degrees across biological and social systems, and that recognising that is explanatorily powerful.

To be very clear, sometimes the evolutionary framework describes, in different terms, phenomena independently developed in other fields (eg organisational theory, socioecological resilience, behavioural economics). These fields each identify *aspects* of the full evolutionary trajectory of human society. The evolutionary approach integrates them and connects them across boundaries in theory, coordinating the complementary and multidisciplinary nature of global risk studies.

With these tools in place, we ask our central question, not what is going wrong to produce a global polycrisis, but why, and why so consistently?

IV. What Keeps Going Wrong: Evolutionary Dynamics as the Driver of Global Risk

The five fields describe what is happening to global systems and the costs. None systematises why humanity keeps generating these risks, or why the mechanisms meant to reduce them keep failing. Fossil carbon emissions have risen since the Paris Agreement (Friedlingstein et al., 2026), and key nuclear arms-control treaties have lapsed despite earlier non-proliferation success (VCDNP, 2026).

The relevant distinction is Mayr's between proximate and ultimate causation (Mayr, 1961), supplemented by Tinbergen's four questions (Tinbergen, 1963), which separate mechanism and development from function and evolutionary origin. The five fields operate largely at the proximate level, mapping events, responses, and non-linear dynamics. This section turns to the ultimate level, the selection pressures, transmission dynamics, and game-theoretic structures that make proximate failures commonplace. Section III licenses the move: because selection is substrate-neutral, "ultimate" here does not mean gene-level causation but the level of selective dynamics acting on whatever replicates, genes, organisms, firms, institutions, technologies, or ideas. King and Jones reach a related distinction independently in their review of behavioural traits and the polycrisis (King & Jones, 2025).

There are two driving forces, (1) the invisible hand of selection, where variants accumulate because they are locally successful, irrespective of specific intent, (2) deliberate strategy, where actors who understand the game act to entrench their advantage or degrade suppressive mechanisms. This distinction has implications for intervention. Where risk drivers are selected you can redesign incentives, but where actors actively degrade redesign, suppression must be designed (Section VII).

Critically, if global crises are predictable outputs of global-level evolutionary dynamics, responses that do not operate at the level of those dynamics, ultimately tend to be overwhelmed, whatever their technical merit or political backing.

In the following, we explain what has gone wrong to generate the serious global risks humanity faces, grouping by the driving dynamics rather than by hazards themselves. Common forces generate apparently unrelated crises, and a response calibrated to climate or geopolitical risk alone will be outpaced by the dynamics generating both. These dynamics explain anthropogenic, incentive-driven risk, where some actor's local advantage results in the wider hazard. For exogenous hazards such as a pandemic of natural origin, no one's advantage produces the event, and the framework explains not the hazard but the failure to build resilience to it (Section V).

Dynamic 1: Multilevel selection. Selection acts at several levels at once, and what is adaptive for an individual actor is often destructive for the group or commons that sustains it (Sober & Wilson, 1998). For example, in artificial selection, when hens are bred individually for egg yield, the top producers prosper partly by suppressing cage mates, and flocks feather-peck and cannibalise themselves into heavy mortality, whereas selecting whole groups for productivity yields cooperative, low-mortality flocks that produce more eggs (Muir, 1996). The same structure can scale catastrophically. States, firms, and consumers face little selection pressure against emitting carbon or other pollutants. The benefit is local and immediate, the cost distributed and deferred. Cooperation fails here not from weakness of will but because it runs against selection at every level at once, and fossil-fuel firms are rewarded for maximising extraction even when run by people who have grasped the adverse consequences for decades. The same logic drives overfishing, aquifer depletion, and antimicrobial overuse. Evolutionary suicide is the limiting case, an evolutionary form of the tragedy of the commons (Rankin et al., 2007), where a lineage degrades the very substrate sustaining it, the biosphere or the multilateral order, finding later it is unable to bear the cost of the new circumstances (Arnscheidt et al., 2025).

Historically, many such failures (Sterelny's "SNAFUs") arose because the link from behaviour to consequence was opaque (Sterelny, 2007). That generally no longer holds with existing scientific knowledge. The tragedy of the global commons is now less a failure of understanding, than a failure of selection-environment design at the governance level, and above all the absence of outlaw/defector suppression at the right levels (Hardin, 1968; Ostrom, 1990), and at effective thresholds.

Dynamic 2: Prestige competition and arms races. Competitive selection produces Red Queen dynamics, escalatory races in which every participant must invest ever more merely to hold position, causing collectively negative returns and no unilateral exit (Van Valen, 1973). In nature, the co-evolution of newt tetrodotoxin and garter-snake resistance have driven both sides to costly physiological extremes (Brodie & Brodie, 1990). Human institutions reproduce the structure, and critically, modern technology raises the stakes from local to global.

Frontier AI is a contemporary and clearly anticipated case (Armstrong, Bostrom, & Shulman, 2016; Bostrom, 2014), where safety investment is individually costly and competitively disadvantageous if rivals do not match it, so the field races to deploy insecure models. This is not because developers wish to be reckless but because non-reckless actors are simply outcompeted. In cyber conflict offence outpaces defence, barriers to entry are low, and attribution issues obscure feedback, so investment flows to capability rather than restraint. Another example is nuclear weapon stockpiling as a multipolar standoff in which it is hard for states to disarm unilaterally and the stable strategy is both individually rational and a collective global risk (Schelling, 1960). Great-power competition broadly exhibits the same security dilemma, in which defensive measures are often read as offensive by others and drive unwanted escalation (Jervis, 1978). Gain-of-function and other dual-use research carries the same arms-race structure, competitive and prestige incentives, or malice, driving capability ahead of containment. The erosion of arms-control institutions amplifies the problem, and degrading the suppression mechanism is itself a move in the game.

Dynamic 3: Lock-in, path dependence, and local fitness traps. Evolutionary systems can settle into configurations that are stable under current pressures but globally suboptimal. This is particularly where exit requires first overcoming barriers around a fitness valley that no actor can achieve alone (Kauffman, 1993; Wright, 1932). Human systems have no automatic means of attaining distant 'fitness peaks', so the trap holds unless one of two levers is pulled deliberately, (1) the selective environment is reshaped so the pathway to the better optimum is rewarded, or (2) the system's evolvability is raised so the search widens, the counterpart of what cybernetic governance describes as variety amplification (Hämäläinen, 2025). Fossil-fuel infrastructure, energy grids, and urban form are accumulated local optima that are costly to leave, and the cost is often self-reinforcing, since each addition to sunk investment and complementary technology can increase the reward of staying (Arthur, 1989), while 'temporary' price relief is structurally hard to reverse (Mahdavi & Ross, 2026). Incumbents build this trap, laying down infrastructure and coalitions that a fossil-fuel economy needs and thereby reshaping the selective environment to favour their continuation (Unruh, 2000), the human system analogue of evolutionary niche construction (Odling-Smee et al., 2003). The financial system's requirement for economic growth to meet interest demands year-on-year is another difficult to escape lock-in producing externality pressures on the biosphere and human systems (Jackson & Victor, 2015).

Governance is subject to the same dynamic and institutions optimised for a founding problem can lack response variety that new disruptions demand. This widens Hämäläinen's (2025) "complexity gap" between the variety of disruptions faced and responses available, necessitating broad anticipatory governance (M. Boyd & Wilson, 2025). This is a trade-off of institutional robustness vs evolvability where optimisation for past conditions produces present stability but limited adaptability (Duit & Galaz, 2008). For example, there is a reactive-governance trap, exemplified in New Zealand, where roughly 97% of central-government natural-hazard spending since 2010 went to response and recovery and only about 3% to risk reduction, despite evidence that a dollar of prevention saves several downstream (IAG/Sapere, 2025). Each disaster consumes the capacity that appropriate prevention needs, while the underlying stresses keep accumulating.

Dynamic 4: Transmission bias and degrading the epistemic commons. Cultural evolution need not select for the accuracy of what is transmitted. It can favour variants that are more readily learned, remembered, and passed on (R. Boyd & Richerson, 1985; Mesoudi, 2011; Richerson & Boyd, 2005). This is mainly selection rather than deception, though deliberate exploitation also occurs. Evolved biases help shape those selection pressures, for example people readily attend to moral-emotional, novel, and out-group-animosity-laden content (Brady, Wills, Jost, Tucker, & Van Bavel, 2017; Rathje, Van Bavel, & van der Linden, 2021; Vosoughi, Roy, & Aral, 2018), use prestige as a heuristic for whom to copy (Henrich & Gil-White, 2001), and treat familiar claims as true, a bias that repeated messaging can exploit (Dechêne, Stahl, Hansen, & Wänke, 2010; Hasher, Goldstein, & Toppino, 1977). In engagement-optimised environments (eg social media), such content out-competes accurate, nuanced material (Acerbi, 2019). Platforms competing for attention are themselves selected toward whatever maximises engagement. This results in difficult to avoid degradation of the epistemic commons. The historically recent arrival of digital and social media, and AI, has accelerated this self-reinforcing dynamic, with AI use potentially causing a tragedy of the cognitive commons (Lovett, 2026).

We call this cognitive niche destruction. The shared map of reality on which collective action depends erodes as transmission favours divergent, emotionally loaded representations over convergent, evidence-based ones. This meta-risk limits ability to respond to all other risks. Reality-tracking information is, however, favoured when misinformation produces failure, yet global connectivity and distant causation can masks or defer the relevant feedback. Maladaptive information spreads, correction is deferred, and adaptive governance cannot evolve. Where the corruption is deliberate, to obscure or delay recognition of a harm, it is strategy exploiting these channels (Oreskes & Conway, 2010).

These four major dynamics, along with others, explain the puzzle of global risk. The crises of the polycrisis era are not just random accidents, the product of individual bad decisions, or failures of knowledge or will. They result from dynamics such as multilevel selection, arms races, lock-in, and transmission bias, operating at several scales at once, mostly beyond governance attention, and ratcheting the world toward catastrophe. Appealing to the rationality or good will of individual competitors cannot ultimately change this, only altering the selection environment and suppressing system-threatening defection can.

Where the selection environment has been deliberately re-engineered, commons and arms-race problems have been solved or contained. The internationally agreed Montreal Protocol reversed a potential global commons tragedy (destruction of the ozone layer) by redirecting rather than destroying the offending industry. Also, nuclear arsenals were reduced under verified arms-control regimes from the late 1980s to 2010. Sulphur controls also curbed acid rain. Ostrom (1990) documented local commons governed durably for generations. Each proves the programme proposed here, and each shares a common thread, namely a suppression or coordination mechanism was constructed and protected, converting locally advantageous defections into losing moves at system level.

That these phenomena are structural not accidental is corroborated. The structural-demographic record shows the same signature in societies reaching sufficient complexity: overproduction of elites, erosion of collective governance, and power-holders doubling down rather than adapting (Hoyer et al., 2025; Kemp, 2025). The human group-selection ratchet has driven environmental exploitation, namely groups that extract more tend to dominate. And humanity cannot readily evolve a cooperative global counterpart, because the competition between planet-scale entities that cultural group selection would require does not exist. These are precisely the reasons why deliberate structural intervention is needed (Waring et al., 2023).

All the above means that responses calibrated to proximate symptoms of global risk will be overwhelmed by the ultimate dynamics that generate them. The level at which intervention must occur is the evolvability of risk-reducing systems. Humanity must engineer the conditions under which adaptive, cooperative solutions can be generated, retained, and selected for against the pressures described here, and against the two main failure modes, (1) defection against the higher-level system, and (2) selection that is mis-aligned to global risk reduction. Those conditions are the subject of the next section.

V. Evolvability: What Distinguishes Resilient from Fragile Systems

Evolvability is the capacity to generate, transmit, and retain adaptations. In biology, evolvability has become a recognised meta-concept. It scales from within-population variation to major evolutionary transitions, and must be built into a system rather than being assumed (Nuño de la Rosa, 2023; Pigliucci, 2008; Wagner & Altenberg, 1996). That scalability makes it the right organising concept for global risk reduction, because the same property can be asked of a gene network, a firm, or a system of global governance.

Evolvability is value-neutral. Culture and institutions evolve, but evolution is no guarantee of wisdom, and a system can be highly evolvable toward catastrophe, as a cancer is evolvable inside the body before destroying it. So evolvability must be paired with a fitness function worth sustaining at the right level of organisation. In the context of global risk this means governance must guide evolvability of the mechanisms that need to change to reduce risk, while pairing with selection toward resource, biosphere and planetary boundary stability, as well as human societal wellbeing at global scale. This is where the ethics, principles, and fairness flagged in the literature need to bite.

Additionally, the dynamics of Section IV are pervasive in time and space, technologies, institutions, and norms are always under selection. The question is not about escaping evolution, it is whether the conditions that let beneficial adaptations accumulate remain intact. Fast change can discard accumulated adaptation just as it clears obstacles. German scientific forestry offers a cautionary case. From the eighteenth and nineteenth centuries, foresters increasingly replaced complex mixed woodland with simplified, often even-aged stands dominated by commercially valuable conifers and managed chiefly for predictable timber production. The first rotations could perform extremely well, but over subsequent rotations some pure stands suffered declining productivity, degraded nutrient cycling and increased vulnerability to pests and storms. The redesign captured exceptionally well what planners could count and optimise, while discounting ecological relationships whose importance became apparent only later (Scott, 1998).

Importantly, selection acts only on what is expressed. In biology a mutation is invisible to selection unless it changes function. In human systems the equivalent code, the laws, syllabi, regulations, norms, and standards that reconstruct institutions over time, is inert to selection until it changes behaviour. An unenforced statute, a norm professed but not practised, is a "mutation" that never reaches differences in function upon which selection acts. Adaptive governance must not merely rewrite the code but ensure this propagates into conduct.

In biology, Wagner and Altenberg (1996) located evolvability not in any single trait but in the architecture that maps genotype to phenotype. Kirschner and Gerhart's theory of facilitated variation identified features permitting random change to produce viable variants rather than broken ones, namely conserved core processes, weak regulatory linkage, exploratory mechanisms, and modularity (Kirschner & Gerhart, 2005). Development is not a passive readout either and West-Eberhard showed that phenotypes can change first and be stabilised by selection afterwards. Genes are often the followers, reminding us that in human systems practice can shift ahead of formalised encoding (West-Eberhard, 2003). Evolvability, on these accounts, is itself an evolved achievement, selected at the level of the lineage.

Evolvability is also a cultural achievement and Sterelny demonstrates cumulative cultural adaptation needs the same triad, suppression of defectors who corrupt the shared information commons, inheritance faithful enough for small improvements to compound, and modularity that lets domains vary independently (Sterelny, 2006). Constructed, information-rich environments have traditionally done much of this work. This is the niche construction of Section IV, now read as the scaffolding on which risk reducing evolvability rests.

Functioning institutions and stable informational environments are constitutive of evolvability rather than background to it. In governance, complex adaptation needs two things at once, a developmental environment reliable enough that a policy or practice matures the same way from one cohort, or location, to the next, and a selective environment stable enough that selection acts on reliably developing variants rather than noise.

However, evolvability is not maximised by maximising any single condition but occupies an interior optimum. Too much variation destroys the inheritance structures cumulative adaptation depends on; too much stability entrenches whatever is currently selected, including maladaptation. The target is a productive combination, not the extreme of any one condition, a point we use to diagnose a metacrisis in Section VI.

Five conditions for appropriate evolvability can be abstracted from these accounts:

1. **Variation**: the supply of diverse candidate solutions;
2. **Modularity**: semi-independent components that can change without triggering system-wide failure;
3. **Stability of the selective environment**: pressure consistent enough for adaptation to accumulate;
4. **Fidelity of transmission**: the reliable inheritance of what works;
5. **Outlaw/defector suppression**: the containment of entities that gain locally by breaking cooperative rules (such actors are also variously described as ‘cheaters’, ‘exploiters’, or ‘free riders’ across various contexts).

These conditions are broadly the mirror of Section IV’s failure modes that generate global risk, though the correspondence is many-to-many. Section VI sets out each in turn together with the evidence for its displacement.

Outlaw suppression is particularly important. At each transition to a higher-level unit, the evolvability of the lower level must be restricted and that of the new unit enhanced, because the two levels disagree about which adaptations are beneficial. Mediating that conflict is what makes the higher-level unit a unit of selection at all (Michod, 1999). Outlaw/defector suppression is therefore constitutive of evolvability rather than merely something nice to have to support cooperation, and this holds at the cell, organism, group, societal, national, and global level alike. It is a decisive argument against non-interventionist, purely market solutions, and Section VII develops it in full.

Our positive thesis varies by risk type. For anthropogenic risks, calibrating these conditions can address both the generation of risk and the response to it. For exogenous hazards such as a natural pandemic, solar storm, or large volcanic eruption, governance changes may not prevent the event, and the contribution of tuning evolvability is in developing latent, generalised, threat-agnostic resilience. Features like modularity, redundancy, and reserve capacity pay off whatever the shock.

These evolvability conditions are demanding, and the domains that matter most for global risk meet them least well. Governance scores poorly with low trialability, low observability, heavy confounding, and strong incentives to hoard information (Sterelny, 2007), which is why technology adapts visibly while governance lags. Governance is not readily evolvable by default and change must be deliberately scaffolded. Because invisible-hand selection runs continuously, even firms whose leaders try to engineer their culture mostly fail. The scaffolding must be structural, not intentional alone. This points to two levels at which governance can act, (1) it can shape the selection pressures that decide what wins (fitness function/incentives), or (2) shape the evolvability conditions that determine what adaptive change is possible at all. Risk-governing institutions, on this view, are referees of competition and architects of the search for adaptations.

VI. The Formal Metacrisis: A Degradation of Evolvability

The term "metacrisis" has circulated in discussions of global systemic risk and the polycrisis (Hedlund, 2023; Nørgaard, Hedlund, & Meglin, 2025; Rowson, 2023), but has not been operationalised in a form amenable to risk analysis. Existing treatments locate it in cultural, epistemic and institutional conditions (eg failure of "sense-making"). This paper proposes a precise operational definition.

Definition: The metacrisis is the simultaneous degradation of the conditions of evolvability that let risk-reducing adaptations accumulate, persist, and be selected for: a crisis not of any particular risk but of the capacity to adapt to risk at all.

The definition explains not merely that global risk governance continues to fail but why it fails consistently across domains. If the five evolvability conditions degrade together, no effort aimed at individual hazards can keep pace, because the machinery that converts effort into durable adaptation is itself being dismantled.

Evolvability occupies an interior optimum: variation can be too high as well as too low, stability too rigid as well as too volatile, so "degradation" does not mean monotonic loss. Degradation means displacement from that balance in either direction, scaffolding that should be stable becoming brittle uniformity, variation that should be structured becoming either monoculture or noise. Two targets are salient, (1) evolvable elements inherent in human systems that can mitigate system risk, and (2) the risk-governing systems themselves. The metacrisis is an asymmetry between the risk generation of Section IV and the capacity to evolve away from it.

We now illustrate each evolvability condition. In each case we observe object-system facts and read these as proxies for relevant governance failures, a sign that the institutions able to hold the corresponding condition have failed to do so.

Variation. Systems with more competing models have more adaptive potential, though the relationship between competition and innovation is an inverted U rather than a straight line (Aghion, Bloom, Blundell, Griffith, & Howitt, 2005). The variation that matters is structured toward viable forms and depends on robustness, or canalisation, the capacity to absorb perturbation and still function. Variation also results from recombination of existing variants. The failure is not when systems converge on useful shared standards, but emerging harmful homogenisation. The global food system has concentrated onto a few staple crops, input suppliers, and trade chokepoints, raising the risk of correlated or simultaneous failure (Monbiot, 2023; Puma, Bose, Chon, & Cook, 2015). What has eroded is the capacity of competition authorities, seed banks, and strategic reserves to mandate variety and hold buffers. Defection makes things worse, as lobbies secure subsidies by persuading regulators that only they can supply what is needed (Clapp, 2021). A subtler instance appears to be diversity, where ostensibly grassroots movements are increasingly coordinated top-down, with coordinated slogans propagated across borders by a few actors (Walker, 2014), so an apparent proliferation of radical political opinion is actually the spread of a few engineered variants. In both cases, variation is lost while looking like variation gained.

Modularity. The basis of modularity is weak linkage, where components are coupled by signals rather than by flows of matter or energy, so one can change state without dragging its neighbours along (Kirschner & Gerhart, 2005). A failure mode is where weak linkage hardens into strong, through hyper-specific statutes or prescriptive mandates. This looks efficient in the short term but undermines evolvability. Separation of powers, federalism, and institutional independence are some governance analogues. Where previously there were semi-independent, buffered arrangements, tight coupling creates systems in which local failures propagate globally in technological systems (Perrow, 1999), or hyper-connected banking (Haldane & May, 2011). The 2011 Tōhoku earthquake produced harms cascading through just-in-time supply chains, the 2021 Suez blockage by one ship halted a tenth of world trade, and the 2026 Hormuz crisis sequestered critical petroleum products. These are the visible symptoms but beneath them lies a regulatory and corporate-governance regime that rewarded stripped buffers and hardening of coupling in the name of efficiency. No institution was mandated to preserve slack, and free market logic put little value on resilience. Toyota's response to Tōhoku actually shows the same shock generating the adaptive counter-move. Toyota recombined buffer stocks and diversified suppliers into a more resilient design, which spread (replicated) as competitors imitated it (DiMaggio & Powell, 1983; Matsuo, 2015). One event thus exhibits both the failure of coupling and the recombination that repairs it, the difference lying in whether any institution was positioned to take effective action.

Stability of the selective environment. Adaptation accumulates where selective pressure is consistent enough and sustained long enough. The choice of fitness function is critical and what is actually selected in fast moving cultural evolution is rarely human or ecological wellbeing but profit, votes, engagement, or military advantage. A stable environment selecting for the wrong metric can entrench maladaptation efficiently. Biology also shows that a lineage commits to directed, long-term adaptation only when its environment is predictable, and otherwise hedges through short-term diversification in proportion to how often conditions change (Kussell & Leibler, 2005). Change in AI, bioengineering, and digital platforms now outstrips the rate at which institutions can adapt incrementally, so no stable gradient can form. This is the Collingridge dilemma in the evolutionary context, by the time a new technology's effects are understood it is too entrenched to steer (Collingridge, 1980). "Move fast and break things" is, in these terms, an anti-evolvability strategy. Policy reversal with each electoral cycle compounds this, since capital flows into long-horizon projects, such as green infrastructure, only where policy is stable enough for the investment to compound. Volatility does not merely slow the energy transition, it selects against the structural adaptation the transition requires, a case Section VII develops.

Fidelity of transmission. Institutional memory, accumulated norms and apprenticeship are the ecological inheritance of governance. Transmission must be coherent and directional enough for selection to compound improvements. Fidelity also needs error-correction, the detection and repair of maladaptive drift before it spreads, without which faithful transmission merely entrenches the errors it carries. In contrast to the generation of variants above, information scaffolding for successful adaptations must be transmitted intact, not undermined. Fidelity is degraded by institutional churn, by the erosion of professional norms as patronage and conflict-of-interest breaches normalise, and by the fragmentation of the shared informational environment on which collective action depends. Churn exists at the political and strategic level, where direction-setting of priorities resets, the United Kingdom has had seven prime ministers 2015–2026. Pandemic preparedness is a clear case of such information loss. The SARS outbreak of 2003 and the influenza pandemic of 2009 generated exactly the adaptations the framework predicts, in national plans, stockpiles, dedicated coordinating units, and simulation exercises whose findings identified specific gaps. Much of that inheritance had degraded before the Covid-19 pandemic in 2020. The adaptation was acquired, transmitted poorly, and largely lost before the next selection event arrived.

Outlaw suppression. A system may satisfy the four conditions above and still fail, because it fails to contain actors who gain local advantage by breaking cooperative rules or exploiting shared resources. This means higher-level adaptations cannot hold and Section V sets out why such containment is constitutive of evolvability rather than an optional addition to it. Weak enforcement, regulatory capture, and jurisdictional arbitrage let those who profit from homogenisation, tight coupling, instability, and informational corruption continue unchecked, and increasingly attack the suppression mechanisms themselves (Stigler, 1971). Structural-demographic dynamics describe the deeper pattern of a wealth pump of elite overproduction, capture, and popular immiseration that then degrade suppression at several layers at once (Turchin, 2023). This is why piecemeal anti-corruption reform is reliably outpaced. Outlaws need not be conscious defectors (lawbreakers or cheats), because underlying dynamics mean that those who happen to exploit through legal means, wield power, or seek resources tend to multiply and degrade systems regardless of intent. That is precisely what necessitates a well-calibrated higher-level suppression architecture.

A recent case illustrates this. At the height of the United States opioid epidemic, wholesale distributors were flooding pharmacies with prescription narcotics. Rather than evade enforcement, the industry disabled the Drug Enforcement Administration through sustained lobbying, a revolving door in which a former agency attorney helped draft the relevant bill, and the Ensuring Patient Access and Effective Drug Enforcement Act of 2016 (Higham & Bernstein, 2017). The suppression mechanism was not evaded but recruited and degraded. This is the governance analogue of a tumour co-opting the immune cells sent to control it.

Two independent bodies of empirical work corroborate these evolvability concerns. Søgaard Jørgensen et al.'s 14 "Anthropocene traps" and the Cascade Institute's catalogue of systemic stresses, including homogenisation and hyper-connectivity (Lawrence et al., 2024; Søgaard Jørgensen et al., 2024). Søgaard Jørgensen et al. find 45% of pairwise interactions between their traps reinforcing and only 3% dampening, evidence that the conditions generating global risk typically compound rather than cancel (Søgaard Jørgensen et al., 2024). Their later work identifies “five groups of evolutionary processes”, though without naming a clear theoretical grounding in the evolvability concept, and omitting outlaw suppression entirely (Søgaard Jørgensen et al., 2026). They name the co-evolutionary struggle, but supply no suppression architecture. These frameworks present important factors discovered inductively, providing a partial description, whereas evolvability and the extended evolutionary paradigm illuminates by specifying precisely the complete framework in advance.

One might object that charges of degraded evolvability point merely to history happening, the ordinary churn of institutions rising and falling. We contend that simultaneous displacement across all five conditions, driven by the identifiable dynamics of Section IV, is not random fluctuation, but pattern with traceable causes, and therefore a tractable governance target. Worryingly, the present constellation of competitive dynamics may itself be an evolutionarily stable strategy from which exit is genuinely hard. Other framings name the dynamics without the apparatus to diagnose or change them, but the evolvability account supplies both and identifies leverage. If the metacrisis is largely a degradation of suppression, then well-designed restoration of suppression is where intervention must begin, though it runs uphill against the forces described here. That is the subject of Section VII.

VII. The Outlaw Problem: Defector Suppression as a Key Evolvability Condition

Section VI identified the erosion of outlaw suppression as a degradation that disables the other four evolvability conditions. This is particularly problematic when outlaws accumulate great power. Suppressing outlaw strategies is a governance problem but also a precondition for evolvability and therefore humanity’s adaptive potential in the face of global systemic and catastrophic risk. Biology, having solved this defection problem repeatedly at every major evolutionary transition (Maynard Smith & Szathmáry, 1997), offers a design template that risk-governance theory lacks. Cultural evolution studies show that past solutions include reputational systems and ostracisation, incentives for third party sanctions, culturally transmitted mental frames (eg disgust), and taboos (Henrich, 2016; Kemp, 2025). Practices and beliefs in historical and prehistorical societies often supply these functions implicitly, even if the actors involved don’t understand the evolved mechanisms (Henrich, 2016). Indigenous and traditional societies likely harbour a repertoire of evolved structural solutions to outlaw problems, stabilised by time (selection).

The problem is named in the global-risk literature but left unsolved. Arnscheidt et al. (2025) and Søgaard Jørgensen et al. (2024) both treat actors taking local advantage at systemic cost as drivers of catastrophe, but neither supplies a suppression architecture, ie how such actors are detected, monitored, and contained before their strategies entrench. That is what a multi-level reading of evolution provides.

An outlaw strategy gains local fitness, under selection within a cooperative system, by violating the norms or over exploiting the shared resources that sustain the system. This needs no malice, firms externalising costs, states departing from treaties, and platforms amplifying misinformation are responding to legitimate incentives, not necessarily conspiring, so the object of intervention is the selection environment. Outlaws can be operating within the law, but the law allows behaviour that runs against collective interest. Technological leverage and nesting amplify the problem, a single actor can now move the whole system toward a planetary boundary (thanks to network connectivity and exponential technology), and outlaws sit nested from individuals through firms to states, so suppression must be designed for multilevel structure. The potential to contribute to planetary boundary transgressions, deplete resources before sustainable alternatives are established, or to destabilise human systems must be identified and suppressed before agents develop power enough to effect these impacts, which sum across aggregate actors to generate large risks.

As an example, cancer cells and risk-generating actors are tokens of one type: both emerge when natural selection within a cooperative system selects strategies that are locally successful (eg regulatory capture) and collectively destructive at the larger level (eg harmful externalities). Cancer cells are biology's characteristic betrayal, violating five foundations of multicellular cooperation (Aktipis, 2020; Aktipis et al., 2015). Table 3, Panel A illustrates how these correspond to human system failures.

Borrowing of theory must be disciplined by where the parallels break. Institutions, unlike tumours, can potentially foresee consequences, redesign themselves, and have no single discrete heritable genome. But each theoretical importation below rests only on the shared structure, selection within a cooperative system producing locally-fit, globally-destructive strategies, not on those disanalogous features. Foresight and deliberate redesign are in fact advantages the human body lacks, meaning governance can in principle do better than multicellularity, not merely imitate it.

The evolutionary reading therefore does not supply designs governance lacks in principle; several of the strongest suppression interventions were arrived at with no evolutionary input. What it adds is a unifying rationale, why suppression is constitutive of evolvability (Michod, 1999), rather than one possible policy among many. It also adds a method to find more, for example other catalogues of evolutionary solutions include ways biology solves the tragedy of the commons (Rankin et al., 2007). Four therapeutic principles from cancer's evolutionary biology transfer directly, developed in Table 3, Panel B:

- *Redundant, multi-layer suppression*
- *Calibrate suppression, not maximise it*
- *Protect the suppressors*
- *Suppress early, before entrenchment (ie before outlaws become too powerful)*

The aim of suppression is to prevent catastrophic cheating (which can take the form of 'death by 1,000 cuts' eg summing global carbon emissions) while preserving legitimate diversity. Additionally, "outlaw" must not be defined by powerful incumbents to suppress challengers. The fitness function being defended must be explicit and accountable, grounded in ethics and fairness, and the suppression architecture subject to transparency.

Design is important, as it is not clear how often the natural evolutionary dynamics driving use of the commons lead to collapse and extinction rather than sustainability (Rankin et al., 2007). Pairing the drivers of planetary boundary transgressions with approvals processes or taxation offers one example, governance could identify and limit any agent or process that contributes to transgression and reward those that effect retreat from boundaries. Ostrom (1990) supplies an institutional backbone through her principles of monitoring, graduated sanctions, and nested enterprises, which are conditions under which defector suppression makes a risk free higher-level unit possible.

We suggest that possible interventions cross two dimensions. On the internal axis, supply chain examiners modelled on bank examiners could embed system-stability oversight (spawning anti-fragility mechanisms) inside firms whose lean optimisation is individually rational but produces a fragile collective system. Planetary boundary examiners would be another option, ensuring that organisation externalities don't exceed a level proportionate to the firm's share of a market calibrated to sustainable or regenerative limits.

On the external axis, energy policy shows the evolutionary reading yielding a recommendation complexity science does not. Capital commits to long-horizon adaptation only under a predictable environment, so the prescription is not larger subsidies but durable policy contracts insulated from electoral cycles or leadership whim, since volatile subsidy actively selects against structural adaptation. This becomes concrete at regional scale, where, absent a bloc-wide energy plan (eg to distribute generation and load fluctuations), member states face a multipolar trap and defect to national optimisation. The needed intervention is downward regulation from the bloc to stabilise the shared environment. Cooperation is more readily evolvable at bloc and regional scale than global, where the competition between planet-scale entities that would bring internal cooperation to each does not exist (Waring et al., 2023).

Other generative additions the evolutionary framework supplies might include modular societal structures, whether "transition towns", or revitalised rural towns expressing variation (Brooks & Agosta, 2024), or refuges, namely semi-autonomous, self-reliant nodes able to preserve capability through a global catastrophe and reseed adaptation afterwards, less as insurance against a named hazard than as protection for evolvability itself (M. Boyd & Wilson, 2019; King & Jones, 2021).

Evolutionary expertise embedded inside risk governing institutions is another obvious implication.

The common thread is a shift in the object of governance, from how to regulate and comply to how to shape what selection rewards and manage the conditions that give rise to solutions. Institutions, on this view, are architects of the search, and unless they deliberately make cooperative, long-horizon strategies evolutionarily stable, markets and states will keep falling into the dynamics that built our fragile world. How risk governance might pilot this is the subject of our final discussion.

VIII. Discussion

This paper has made three moves, (1) define the metacrisis formally, not as mere upstream causes, an epistemic failure, or a name for many crises, but as the simultaneous insufficiency of five evolvability conditions; (2) it reframes outlaw suppression as constitutive of evolvability rather than instrumental to cooperation, so outlaw suppression becomes a precondition for adaptive governance rather than one goal among many; (3) it treats evolutionary biology, with examples from cancer biology, not as metaphor but as a transferable source of suppression principles and a procedure for generating governance designs for mitigating global risk generating mechanisms. Three implications follow:

1. Epistemologically, researchers across the fields of global risk now have grounds to treat evolutionary biology as workable theory rather than illustration, with a structured map of which concepts bear on which governance problems.
2. Diagnostically, because each evolvability condition is independently observable, the five together form the basis for a dashboard tracking not just the state of risk-

governance capacity but its direction of travel, an early-warning signal current approaches cannot supply, though see Oliver et al. (2026) for steps to adapt in this direction.

3. Therapeutically, the cancer biology-derived principles of Section VII: calibrate rather than maximise, build redundant layers, and protect the suppressors, supply a common design logic that unifies interventions, which existing frameworks reached in isolation.

We can now develop designs suggested by the transfer of the extended evolutionary framework to risk governance that are not already in the risk mitigation repertoire. Perhaps supply chain or planetary boundary examiners nurturing system-stability or overreach prevention from inside organisations (Section VII), or modular societal refuges preserving variation and capability through catastrophe, or other analogues in Table 3.

Predictions. The framework doesn't just interpret, it makes directional predictions that are falsifiable. First, on matched collective-action problems, interventions that alter the selection environment should ultimately outperform those that alter actors' knowledge or goals. Second, governance domains scoring low on the five evolvability conditions should show measurably weaker adaptive response to comparable shocks, testable against polycrisis datasets (Delannoy et al., 2025; Rakowski et al., 2025). Third, the framework predicts the direction of dynamic response at each level, ie a lower level harbours defection unless downward-regulated, and the stronger the entrenchment or niche construction, the greater the inertia and the more capture attempts that a given intervention provokes. These distinguish the account from mere redescription and set an empirical agenda.

Objections. The serious objections are methodological. That the framework is unfalsifiable is answered by the predictions above. That there is no planetary population for selection to act on, so global evolvability is ill-defined, is correct, and fixes the scope: the Darwinian conditions can hold among firms, professions, and national and regional institutions, and only marginally at the global tier, which must be governed by construction rather than selection (Section V). That complexity science and commons research already describe this understates the differences, complexity science explains how systems become fragile, the evolutionary account why fragile architectures keep being selected despite full awareness, and the two readings diverge in practice. That the cancer evolution analogy overreaches is met by restricting every transfer to the shared selective structure, not the disanalogous features (Section VII). The weaker objections resolve quickly. Reflective agents cannot simply choose differently, because the market does not stop selecting when participants understand the game. And the framework is not fatalist but the reverse, by locating failure in selection pressures rather than ethical deficits, leverage points move from critique and blame to redesign. The successes in Section IV, from Montreal to arms control, show the dynamics can be beaten where suppression is built and protected.

A forward agenda. (1) An empirical strand would operationalise the five conditions into indicators and test our degradation claim against the growing polycrisis datasets. (2) An institutional strand would pilot the suppression designs now, embedded examiners, ring-fenced prevention budgets, and systemic-impact assessment at market entry, each with measurable outcomes. The cheapest and most overlooked is embedding people trained in evolutionary thinking inside risk institutions, itself an instance of the variation principle. (3) A theoretical strand could take up the major-transitions problem. If humanity is undergoing a transition toward a higher-level unit (Lior & McNamara, 2025), the inter-group competition that stabilised prior transitions is unavailable at planetary scale, an open question is what could constitute such a unit from within. Simulation can potentially inform the answer, modelling planet-level systems with different regional blocs, trade and pollution-control groupings, or different international-governance architectures, to explore the scale at which competing entities best sustain cooperation without external selection on the whole.

However, cooperation is likely more evolvable at regional than global scale. Building suppression and resilience across many connected local and regional units, with sufficient regional-level governance of evolvability, is therefore both more tractable and more robust than striving for global agreement, and weak global bodies are a symptom of this structural fact.

The risk community has spent great effort cataloguing what is going wrong. An evolutionary reading names, monitors, and builds the conditions that determine what can be redesigned to go right. Deliberative governance must make the invisible hand of the extended evolutionary paradigm visible, and direct its action.

IX. Conclusion

Five fields have each grasped part of the problem of global systemic and catastrophic risk governance. The part they share but have infrequently named is evolution. Already in their vocabulary, naming it converts a catalogue of what is amplifying global risk into an account of *why*, and an account of *why* directs intervention.

The mechanisms that mitigate global risk, whether system features or governance structures, must be both evolvable and selected under a fitness function aligned to biosphere protection and human wellbeing. Evolvability without aligned selection can yield rapid adaptation toward catastrophe; aligned selection without evolvability yields intentions that cannot generate solutions.

There are three requirements: incentives under which fitness tracks biosphere persistence and human wellbeing, curated evolvability so that solutions can be generated at all, and suppression of outlaws, whose “success” undermines systems for their own ends.

The solution to global risk is deliberate design of the substrate on which an undesigned evolutionary process runs. Selection operates continuously and works in humanity’s favour only where that right substrate is in place. The task is not to micro-manage outcomes but to set the larger forces, internal and external, sectoral and institutional, then monitor and adjust them so the system does not ratchet on unattended. Intervene where evolvability is weakest and selection most misaligned, and do it regionally, across many connected units, rather than awaiting global agreement.

Table 1. Glossary of Key Terms (see extended glossary in Supplemental File)

Term	Definition
Adaptation	A trait, strategy or arrangement retained because it performs relatively well under the relevant selection pressures; in this paper, the term can apply to human social systems (eg, institutions and practices) as well as biological traits.
Adaptive capacity	The ability of a system to modify its behaviour, organisation or practices in response to changing conditions.
Canalisation	The capacity of a biological or analogous system to absorb some variation or disturbance while continuing to produce a reliable functional outcome.
Catastrophic risk / global catastrophic risk (GCR)	Risk of harm on a scale large enough to severely damage human civilisation or overwhelm its capacity to respond.
Co-evolution	Reciprocal evolutionary change in interacting populations or systems, each altering the selection pressures acting on the other.
Cognitive niche	The socially and informationally constructed environment that supports human learning, reasoning and knowledge.
Cultural evolution	Change in socially transmitted behaviours, technologies, institutions, beliefs and norms through processes including selection, inheritance, innovation and transmission bias.
Cultural group selection	Evolutionary change in which traits of groups spread because some groups persist, grow or outperform other groups.
Darwinian population	A population that exhibits the conditions for natural selection to some degree; the concept allows social and technological systems to be more or less strongly Darwinian.
Defection	Pursuing individual advantage by violating or withdrawing from a cooperative arrangement.
Developmental environment	The conditions within which a variant develops into its expressed form; reliable development is needed before selection can consistently distinguish among variants.
Differential persistence / reproduction	The process by which some variants last or reproduce more successfully than others and therefore become relatively more common.
Ecological inheritance	Environmental modifications produced by previous generations that alter the developmental or selective environment encountered by later generations.
Endogenous risk	Risk generated from within a system by its own structures, incentives or interactions, rather than imposed from outside.
Entrenchment	The process through which an actor, institution or technological arrangement accumulates sufficient infrastructure, influence and complementary systems to become difficult to displace.
Evolvability	The capacity of a system to generate, transmit and retain adaptive variants over time.

Term	Definition
Exogenous hazard	A hazard whose originating event is external to the incentive dynamics of anthropogenic systems, such as a naturally occurring pandemic, volcanic eruption, space weather event or asteroid impact.
Extended evolutionary framework	The paper's broad evolutionary perspective in which inheritance and evolutionary processes are not restricted to genes but can include cultural, behavioural, symbolic, ecological and other transmitted structures.
Externality	A cost or benefit created by an actor's behaviour that is borne by others and is not fully reflected in the actor's own incentives.
Fitness	Relative success at surviving, reproducing or persisting under a particular selection environment; it does not mean moral goodness or overall social benefit.
Fitness function	The criterion that effectively determines which variants or behaviours succeed in a system—for example profit, votes, engagement or, normatively in the paper's proposal, biosphere persistence and human wellbeing.
Fitness landscape	A representation of possible variants or strategies in terms of their relative fitness, usually imagined as peaks, valleys and alternative routes through a space of possibilities.
Inheritance system	A mechanism through which information is transmitted over time; the paper discusses genetic, epigenetic, behavioural and symbolic inheritance.
Local fitness optimum / local optimum	A configuration that performs better than nearby alternatives but is inferior to another possible configuration that cannot be reached without first accepting lower performance or significant transition costs.
Major evolutionary transition	A transition in which previously independent lower-level entities become integrated into a new higher-level cooperative unit, as cells did in the evolution of multicellular organisms.
Maladaptation	A trait, practice or system arrangement that is poorly suited to the relevant higher-level or changed environment, even if it may have been successful under other conditions.
Metacrisis	Specifically defined in this paper as the simultaneous degradation of the conditions that allow risk-reducing adaptations to arise, persist and be selected for: a crisis in the capacity to adapt to risk itself.
Modularity	Organisation into semi-independent components that can change or fail without forcing corresponding change or failure throughout the whole system.
Multilevel selection	Theory in evolutionary biology stating that natural selection operates simultaneously at multiple levels of a hierarchy—such as genes, cells, individuals, groups, firms, or states.
Natural selection	The process through which heritable variants associated with greater reproductive success or persistence become more common, without requiring foresight or deliberate design.

Term	Definition
Niche construction	The process by which organisms or other actors modify their environments and thereby alter the selection pressures acting on themselves and others.
Outlaw strategy	The paper's umbrella term for a strategy that achieves local fitness by breaking, circumventing or exploiting cooperative rules in ways that damage the larger system.
Phenotype	The observable or expressed characteristics of an organism; used by analogy in the paper for the expressed behaviour or functioning produced by institutional "code".
Planetary-boundary transgression	Movement beyond one of the proposed safe operating thresholds in major Earth-system processes.
Polycrisis	The causal entanglement of crises across different systems such that they interact, reinforce one another and jointly worsen humanity's prospects.
Proximate causation	Explanation in terms of immediate mechanisms and processes producing an outcome.
Scaffolding	Stable structures, institutions or informational environments that make learning, transmission and adaptive development possible.
Selection environment / selective environment	The surrounding incentives, constraints and conditions that determine which traits, behaviours or institutional forms succeed.
Selection-environment design	Deliberate modification of rules and incentives so that socially beneficial strategies become more successful and harmful strategies less so.
Socio-ecological system	A coupled system in which human societies, institutions, economies and ecosystems interact and co-develop.
Substrate neutrality	The principle that natural-selection logic does not depend on what the competing variants are physically made of; genes, organisms, groups, technologies, institutions and ideas may all instantiate it to differing degrees.
Suppression architecture	The paper's term for the interconnected institutions and mechanisms that detect, constrain and prevent locally advantageous but system-damaging strategies.
Systemic risk	Risk produced or amplified by the structure and interconnections of a system, such that failures may propagate widely rather than remain local.
Telos	A deliberately intended purpose or goal; the paper contrasts purposeful human design with the unintended "invisible hand" of evolutionary selection.
Tragedy of the commons	Degradation of a shared resource when individuals gain by exploiting it while the costs of depletion are spread across the wider group. See also "free rider" above.
Transmission bias	A systematic tendency for certain cultural variants to be copied more readily than others because of properties such as prestige, familiarity, emotional salience or conformity.

Term	Definition
Ultimate causation	In evolutionary explanation, the deeper selection history or selective dynamics explaining why a recurring trait or pattern exists; in this paper the term is extended beyond genes to institutions, technologies and other evolving systems.
Unit of selection	An entity at a particular organisational level whose traits can be acted on by selection; major evolutionary transitions create new higher-level units of selection.
Variation	Differences among candidate traits, practices, strategies or designs on which selection can act.
Variety amplification	In cybernetic terms, increasing the range of responses available to a governing system so that it can match the diversity of disturbances it faces.
Weak linkage	Connection between components that permits communication or coordination without making each component strongly dependent on the others; an important basis of modularity.

Table 2. Five fields converging on a shared explanatory need. Each addresses a genuine part of the global-risk problem and reaches, in its own vocabulary, for evolutionary concepts, but none supplies the generative account that would unify them.

Field	Core contribution	What it does not yet systematise	Examples of evolutionary vocabulary in use	Representative sources
National risk assessment (NRA) and Disaster risk reduction (DRR)	All-hazards methods and national registers; explicit objective functions; a mature resilience vocabulary, anchored in the Sendai Framework.	Siloed; structurally unable to capture endogenous, emergent, cascading risk. Practice favours response over prevention; near-term, local hazard horizons; excludes catastrophic tail risks, resilience defined by treaty as restoration of prior structure, excluding adaptive and transformative change by definition.	Build back better, climate adaptation; detailed 'fitness function' (preserving wellbeing, economy, etc) without theory of selection, 'evolving practice' of NRA.	(UNDRR, 2015; 2023; 2025; OECD, 2017; Boyd & Wilson, 2023; Kohler, 2023)
Global catastrophic and existential risk (GCR / xRisk)	Long horizon; confronts the magnitude of the risk tail; gaining institutional recognition.	Evolutionary argument at field founding, dropped subsequently. Tends to hazard-by-hazard analysis and techno-solutionism; thin on the endogenous processes that generate the tail.	Great Filter, major transitions, anthropic shadow, extinction, natural selection, value drift, curtailment of humanity's potential (evolvability).	(Bostrom & Ćirković, 2008; House of Lords, 2021; Ó hÉigeartaigh, 2025; Jehn et al., 2026)
Systemic risk analysis	Treats the global system as a generator of hazards; identifies constitutive properties and a stress–trigger–crisis model; names risk-creating actors.	Explains system states, not why actors keep producing risk-generating conditions or why suppression erodes.	Experiments, Natural selection, Darwinian, evolutionary strategy, biomimetic, surviving systems, deep leverage.	(Liu & Renn, 2025; Lawrence et al., 2024, 2025; Scheffran, 2025; Arnscheidt et al., 2025; Oliver et al., 2026)
Polycrisis scholarship	Maps crisis entanglement; foundational definition and synchronising pathways; systematic review; attractor-state models.	Drivers "asserted rather than analysed"; maps the elephant but not why it persists; no unified generative account.	Adaptive cycles, maladaptation, Anthropocene traps, ratchets, co-evolutionary struggle, socio-ecological resilience.	(Lawrence et al., 2024; Rakowski et al., 2025; Delannoy et al., 2025; Hämäläinen, 2025; Søgaard Jørgensen et al., 2026; Gambhir et al., 2026)
Behavioural and evolutionary science	Proximate/ultimate causation, game-theoretic traps, cultural evolution, major transitions, behavioural crisis.	Seldom global risk-specific, small relative to the problem, and not integrated into risk governance.	Ecosystems & ecological economics, k-selection, proximate/ultimate distinction, multilevel selection, defection, evolvability, major transitions, cultural evolution.	(King & Jones, 2025; Waring, Wood & Szathmáry, 2023; Lior & McNamara, 2025; Merz et al., 2023; Hagens, 2020)

Table 3. Defector suppression: from cancer's evolutionary biology to global risk governance design. Cancer and risk-generating actors are instances of one type, natural selection within a cooperative system selecting strategies that are locally successful and collectively destructive, so biology's solutions transfer as design templates rather than mere metaphors. Panel A maps the breakdown of cooperation; Panel B maps its suppression.

Panel A. The breakdown: how cancer cells cheat and multicellular cooperation fails (Aktipis, 2020; Aktipis et al., 2015)

Foundation of multicellular cooperation	Cheating in cancer	Example analogues relevant to global risk governance design
Proliferation inhibition	Uncontrolled proliferation	Hypergrowth: growth-at-all-costs economic models that externalise systemic cost
Controlled cell death (apoptosis)	Evasion of programmed cell death	Self-regulation of organisations can be unreliable and there is regulatory evasion and arbitrage: firms restructuring or relocating to a permissive jurisdiction to avoid sanction
Resource allocation	Resource diversion (nutrients) and monopolisation	Excessive and unsustainable extraction and capture: monopoly rent-seeking and enclosure of shared resources
Division of labour	Deregulated differentiation (loss of specialised function)	Abandonment/degradation of mandate: institutions shedding their specialised public good protecting function for self-perpetuation or rent extraction
Maintenance of the shared extracellular environment	Environmental degradation	Pollution of the physical and informational commons: concealed emissions; engagement-optimised disinformation

Panel B. The therapy: how outlaw strategies/defection are suppressed.

Suppression principle	How the human body achieves cancer suppression	Global risk governance corollary	Illustrative governance intervention
Redundant, multi-layer suppression	Limitation of cellular defection at three layers: controls within the cell, signals from neighbouring cells, and systemic immune surveillance (Aktipis, 2020). The TP53 network, a master regulator that halts or destroys defecting cells, integrates these signals rather than acting as a tripwire, calibrated against both false positives (chilling legitimate cell proliferation eg for growth in children or in wound repair) and false negatives (letting "outlaw" cancer cells spread).	Organisational analogues are internal compliance, civil-society and media scrutiny, and regulation, national and international law, and enforcement. These must operate together and integrate multiple indicators, avoiding the vulnerability of single-point suppression.	Beneficial-ownership transparency and anti-money-laundering gatekeeping (Findley, Nielson, & Sharman, 2014)
Calibrate suppression, do not necessarily maximise it	Maximal suppression/therapy can select for more resistant defection, as aggressive chemotherapy can eliminate less resistant and select for more resistant tumours. Adaptive therapy instead manages the burden by reshaping the selective environment to disadvantage aggressive cell lines (Aktipis, 2020).	Graduated, incentive-altering enforcement is more durable than maximal deterrence, which can provoke organised evasion; eradicate only where feasible and the residual risk is intolerable – smallpox and rinderpest	Responsive regulation and the enforcement pyramid (Ayres & Braithwaite, 1992)

Suppression principle	How the human body achieves cancer suppression	Global risk governance corollary	Illustrative governance intervention
	Calibration is not always restraint and suppression and eradication are two settings of one dial, with eradication optimal when a target is discrete, hard to regenerate, and carries intolerable residual risk.	were eradicated, the WWII Axis militarist states dismantled.	
Protect the suppressors	Tumours do not merely evade the immune system; they can recruit and repurpose it (Aktipis, 2020). Disrupted cell-signalling allows tumours to evade suppression (Hanahan & Weinberg, 2011). These mechanisms must be acknowledged and suppression mechanisms alert to them.	Defend oversight from deliberate weakening, capture, and manufactured doubt by industry seeking to avoid regulation (Oreskes & Conway, 2010); Independence, transparency, and revolving-door limits keep suppressors from becoming infrastructure that sustains outlaws. Risk reducing governance may need to disrupt relevant industry coordination, cabals, or lobbying mechanisms.	Anti-capture design (Carpenter & Moss, 2014); countering manufactured doubt (Oreskes & Conway 2010); What made historically durable suppressors capture-resistant, bank examiners, the Montreal regime, independent central banks, was design: independence, transparency, revolving-door limits, and insulation from short-term pressure.
Suppress early, before entrenchment	Cancer turns particularly dangerous once cells coordinate to invade and metastasise. Treatment is easiest before that threshold is crossed.	Identify and constrain systemically significant (powerful) actors before lock-in or becoming "too big to fail", prevent them having too much influence, treat the system as an interdependent ecology rather than firms policed individually; Internal examiners – approvals based on advance/retreat to planetary boundaries.	Macroprudential designation of systemically important institutions (Hanson, Kashyap, & Stein, 2011); Stringent approvals at market entry, calibrated to fair proportionality of commons use. Mechanisms for removal of powerful but misaligned actors. Additionally, governance should not feed risk creating actors, for example subsidies on nitrogenous fertiliser can entrench application rates above agronomic optimum in already nutrient-saturated regions, compounding pressure on the biogeochemical flows boundary (FAO, 2021; Steffen et al., 2015).

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Author contributions statement

MB: Conceptualization, methodology, investigation, project administration, validation, writing – original draft, writing – review and editing; NW: validation; writing – review and editing.

Conflict of interest statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work the authors used Claude Opus 5 and Claude Fable 5 as well as ChatGPT 5.6 Sol, in order to help identify relevant source material, convert author notes and drafts into smoothed text, rewrite sections to reframe or shift emphasis in accordance with author direction, condense word count, improve grammar and readability, and provide criticism of drafts and claims. After using these tools/services, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

References

Acerbi, A. (2019). *Cultural Evolution in the Digital Age*. Oxford: Oxford University Press.

Acerbi, A., & Mesoudi, A. (2015). If we are all cultural Darwinians what's the fuss about? Clarifying recent disagreements in the field of cultural evolution. *Biology & Philosophy, 30*(4), 481-503. doi:10.1007/s10539-015-9490-2

Aghion, P., Bloom, N., Blundell, R., Griffith, R., & Howitt, P. (2005). Competition and innovation: an inverted-U relationship. *The Quarterly Journal of Economics, 120*(2), 701-728.

Aktipis, A. (2020). *The Cheating Cell: How Evolution Helps Us Understand and Treat Cancer*. Princeton, NJ: Princeton University Press.

Aktipis, A., Boddy, A. M., Jansen, G., Hibner, U., Hochberg, M. E., Maley, C. C., & Wilkinson, G. S. (2015). Cancer across the tree of life: cooperation and cheating in multicellularity. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 370*(1673). doi:10.1098/rstb.2014.0219

Alexander, S. (2014). Meditations on Moloch. *Slate Star Codex*. Retrieved from https://slatestarcodex.com/2014/07/30/meditations-on-moloch/

Armstrong, S., Bostrom, N., & Shulman, C. (2016). Racing to the precipice: a model of artificial intelligence development. *AI & Society, 31*(2), 201-206.

Arnscheidt, C. W., Beard, S. J., Hobson, T., Ingram, P., Kemp, L., Mani, L., . . . Wunderling, N. (2025). Systemic contributions to global catastrophic risk. *Global Sustainability, 8*, e19. doi:10.1017/sus.2025.20

Arthur, W. B. (1989). Competing technologies, increasing returns, and lock-in by historical events. *The Economic Journal, 99*(394), 116-131.

Ayres, I., & Braithwaite, J. (1992). *Responsive Regulation: Transcending the Deregulation Debate*. New York: Oxford University Press.

Bertolami, O., & Elísio, R. (2026). Towards a theory of coupled sociopolitical events–planetary boundaries, crises, polycrisis and Earth System syndromes. *Physics Letters A*. doi:10.1016/j.physleta.2026.130911

Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford: Oxford University Press.

Bostrom, N., & Cirkovic, M. (Eds.). (2008). *Global Catastrophic Risks*. Oxford: Oxford University Press.

Boyd, M., & Wilson, N. (2019). The Prioritization of Island Nations as Refuges from Extreme Pandemics. *Risk Analysis, 40*(2), 227–239. doi:10.1111/risa.13398

Boyd, M., & Wilson, N. (2023). Assumptions, uncertainty, and catastrophic/existential risk: National risk assessments need improved methods and stakeholder engagement. *Risk Analysis, 43*(12), 2486-2502. doi:10.1111/risa.14123

Boyd, M., & Wilson, N. (2025). From Disaster Response to Anticipatory Governance: why Aotearoa New Zealand's long-term resilience thinking must address global catastrophic risk and systemic vulnerabilities. *Policy Quarterly, 21*(4). doi:10.26686/pq.v21i4.10337

Boyd, R., & Richerson, P. J. (1985). *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.

Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A., & Van Bavel, J. J. (2017). Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences, 114*(28), 7313-7318.

Brodie, E. D., III, & Brodie, E. D., Jr. (1990). Tetrodotoxin resistance in garter snakes: an evolutionary response of predators to dangerous prey. *Evolution, 44*(3), 651-659.

Brooks, D., & Agosta, S. (2024). *A Darwinian Survival Guide: Hope for the Twenty-First Century*. Cambridge MA.: MIT Press.
Carpenter, D., & Moss, D. A. (2014). *Preventing Regulatory Capture: Special Interest Influence and How to Limit It*. New York: Cambridge University Press.
Cascade Institute. (2025). *Polycrisis CORE Model v2.5: Technical Synopsis of Results*. Retrieved from Victoria, BC: https://cascadeinstitute.org/wp-content/uploads/2025/08/Cascade-Institute-Polycrisis-Core-Model-v2.5-technical-synopsis-of-results-Aug-11-2025.pdf
Clapp, J. (2021). The problem with growing corporate concentration and power in the global food system. *Nature Food, 2*(6), 404-408. doi:10.1038/s43016-021-00297-7
Collingridge, D. (1980). *The Social Control of Technology*. London: Frances Pinter.
Cremer, C., & Kemp, L. (2023). Democratising Risk: In Search of a Methodology to Study Existential Risk. In S. Beard & T. Hobson (Eds.), *An Anthology of Global Risk*: Open Book Publishers.
Dawkins, R. (1976). *The Selfish Gene*. Oxford: Oxford University Press.
Dechêne, A., Stahl, C., Hansen, J., & Wänke, M. (2010). The truth about the truth: a meta-analytic review of the truth effect. *Personality and Social Psychology Review, 14*(2), 238-257.
Delannoy, L., Verzier, A., Bastien-Olvera, B. A., Benra, F., Nyström, M., & Søgaard Jørgensen, P. (2025). Dynamics of the polycrisis: temporal trends, spatial distribution, and co-occurrences of national shocks (1970–2019). *Global Sustainability, 8*, e23. doi:10.1017/sus.2025.10008
Dennett, D. C. (1995). *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. New York: Simon & Schuster.
Department for Environment Food Rural Affairs. (2026). *Nature security assessment on global biodiversity loss, ecosystem collapse and national security*. Retrieved from London, United Kingdom: https://www.gov.uk/government/publications/nature-security-assessment-on-global-biodiversity-loss-ecosystem-collapse-and-national-security
DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: institutional isomorphism and collective rationality in organizational fields. *American Sociological Review, 48*(2), 147-160.
Duit, A., & Galaz, V. (2008). Governance and complexity—emerging issues for governance theory. *Governance, 21*(3), 311-335.
FAO. (2021). *A multi-billion-dollar opportunity: Repurposing agricultural support to transform food systems*. Retrieved from https://openknowledge.fao.org/items/7872fdc7-cb4f-4c76-aab2-84b69b5d127f
Findley, M. G., Nielson, D. L., & Sharman, J. C. (2014). *Global Shell Games: Experiments in Transnational Relations, Crime, and Terrorism*. Cambridge: Cambridge University Press.
Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Bakker, D. C. E., Hauck, J., . . . Zeng, J. (2026). Global Carbon Budget 2025. *Earth Syst. Sci. Data, 18*(5), 3211-3288. doi:10.5194/essd-18-3211-2026
Gambhir, A., Hendel-Blackford, S., Albert, M. J., Asipovich, H., Benini, L., Doe, S. S. P., . . . Murphy, K. (2026). The urgency of building systemic risk capacity in a polycrisis world. *One Earth, 9*(4), 101655. doi:https://doi.org/10.1016/j.oneear.2026.101655
Godfrey-Smith, P. (2009). *Darwinian Populations and Natural Selection*. Oxford: Oxford University Press.

Graham, J., Boyd, M., Sadler, G., & Noetel, M. (2025). Mapping Australia's risk landscape: A comparative analysis of global catastrophic risks and traditional hazards [Preprint]. *SSRN*. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5253625
Hagens, N. J. (2020). Economics for the future – Beyond the superorganism. *Ecological Economics, 169*, 106520. doi:https://doi.org/10.1016/j.ecolecon.2019.106520
Haldane, A. G., & May, R. M. (2011). Systemic risk in banking ecosystems. *Nature, 469*(7330), 351-355. doi:10.1038/nature09659
Hämäläinen, T. J. (2025). Polycrisis – root causes and governance solutions. *World Futures: The Journal of New Paradigm Research, 81*(8), 1-33. doi:10.1080/02604027.2025.2572007
Hanahan, D., & Weinberg, Robert A. (2011). Hallmarks of Cancer: The Next Generation. *Cell, 144*(5), 646-674. doi:10.1016/j.cell.2011.02.013
Hanson, S. G., Kashyap, A. K., & Stein, J. C. (2011). A macroprudential approach to financial regulation. *Journal of Economic Perspectives, 25*(1), 3-28. doi:10.1257/jep.25.1.3
Hardin, G. (1968). The tragedy of the commons. *Science, 162*(3859), 1243-1248.
Hasher, L., Goldstein, D., & Toppino, T. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning and Verbal Behavior, 16*(1), 107-112.
Hedlund, N. (2023). *Visionary Realism And The Emergence Of A Eudaimonistic Society: Metatheory In A Time Of Metacrisis.* (PhD). University College London,
Henrich, J. (2016). *The Secret of Our Success*: Princeton University Press.
Henrich, J., & Gil-White, F. (2001). The evolution of prestige: freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior, 22*(3), 165-196.
Higham, S., & Bernstein, L. (2017, Oct 15). The drug industry's triumph over the DEA. *The Washington Post*. Retrieved from https://www.washingtonpost.com/graphics/2017/investigations/dea-drug-industry-congress/
Hofstadter, R. (1944). *Social Darwinism in American Thought*. Philadelphia: University of Pennsylvania Press.
House of Lords Select Committee. (2021). *Select Committee on Risk Assessment and Risk Planning Report of Session 2021–22: Preparing for Extreme Risks: Building a Resilient Society*. Retrieved from https://publications.parliament.uk/pa/ld5802/ldselect/ldrisk/110/110.pdf
Hoyer, D., Bennett, J. S., Whitehouse, H., François, P., Reddish, J., Davis, D., . . . Turchin, P. (2025). Crises averted: how a few past societies found adaptive reforms in the face of structural-demographic crises. *Cliodynamics, 16*, 1-50. doi:10.21237/C7clio.38365
IAG/Sapere. (2025). *Natural hazards-related public spending in New Zealand: Tracking costs over time by the nature of spending*. Retrieved from https://www.iag.co.nz/newsroom/news-releases/iag-sapere-report-nz-natural-hazard-spending
Intergovernmental Panel on Climate Change. (2023). *Climate Change 2023: Synthesis Report*. Retrieved from Geneva, Switzerland: https://www.ipcc.ch/report/ar6/syr/
Jablonka, E., & Lamb, M. J. (2005). *Evolution in Four Dimensions: Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life*. Cambridge, MA: MIT Press.
Jackson, T., & Victor, P. A. (2015). Does credit create a 'growth imperative'? A quasi-stationary economy with interest-bearing debt. *Ecological Economics, 120*, 32-48. doi:10.1016/j.ecolecon.2015.09.009

Jehn, F. U., Engler, J. O., Arnscheidt, C. W., Wache, M., Ilin, E., Cook, L., . . . Kemp, L. (2025). The state of global catastrophic risk research: a bibliometric review. *Earth Syst. Dynam., 16*(4), 1053-1084. doi:10.5194/esd-16-1053-2025

Jehn, F. U., Rössler, M., Kemp, L., Cassidy, M., Kallenborn, Z., García Martínez, J. B., . . . Boyd, M. (2026). No Place to Hide? Regional Resilience and Vulnerability to Global Catastrophic Risk. *Global Challenges, 10*(8), e70146. doi:https://doi.org/10.1002/gch2.70146

Jervis, R. (1978). Cooperation under the security dilemma. *World Politics, 30*(2), 167-214.

Kauffman, S. A. (1993). *The Origins of Order: Self-Organization and Selection in Evolution*. New York: Oxford University Press.

Kemp, L. (2025). *Goliath's Curse*: Penguin Books.

King, N., & Jones, A. (2021). An Analysis of the Potential for the Formation of 'Nodes of Persisting Complexity'. *Sustainability, 13*, 8161. doi:10.3390/su13158161

King, N., & Jones, A. (2025). Human behavioural traits and the polycrisis: a systematic review. *Sustainability, 17*(4), 1495. doi:10.3390/su17041495

Kirschner, M. W., & Gerhart, J. C. (2005). *The Plausibility of Life: Resolving Darwin's Dilemma*. New Haven, CT: Yale University Press.

Kohler, K. (2023). *National Risk Assessments of Cross-Border Risks*. Retrieved from https://www.preventionweb.net/publication/national-risk-assessments-cross-border-risks

Kussell, E., & Leibler, S. (2005). Phenotypic diversity, population growth, and information in fluctuating environments. *Science, 309*(5743), 2075-2078.

Lake, M., & Venti, J. (2007). Quantitative Analysis of Macroevolutionary Patterning in Technological Evolution: Bicycle Design from 1800 to 2000. In S. Shennan (Ed.), *Pattern and Process in Cultural Evolution*: UCLA Press.

Lawrence, M., Homer-Dixon, T., Janzwood, S., Rockström, J., Renn, O., & Donges, J. F. (2024). Global polycrisis: the causal mechanisms of crisis entanglement. *Global Sustainability, 7*, e6. doi:10.1017/sus.2024.1

Lawrence, M., Shipman, M., & Collins, C. (2025). *Global Systemic Stresses: Understanding the Drivers of Polycrisis*. Retrieved from Victoria, BC: https://cascadeinstitute.org/wp-content/uploads/2025/11/global-systemic-stresses-report.pdf

Lior, Y., & McNamara, P. (2025). Major evolutionary transitions in culture and cognition: the Anthropocene and techno-biotic cognition. *Evolutionary Behavioral Sciences*. doi:10.1037/ebs0000380

Liu, H., & Renn, O. (2025). Polycrisis and systemic risk: assessment, governance, and communication. *International Journal of Disaster Risk Science, 16*(4), 526-549. doi:10.1007/s13753-025-00636-3

Lovett, N. (2026). The Tragedy of the Cognitive Commons: How AI Could Disrupt the Regeneration of Professional Expertise. *Human Resource Development Review*, 15344843261470602. doi:10.1177/15344843261470602

Madhav, N., Oppenheim, B., Stephenson, N., Badker, R., Jamison, D., Lam, C., & Meadows, A. (2023). *Estimated Future Mortality from Pathogens of Epidemic and Pandemic Potential*. Retrieved from https://www.cgdev.org/publication/estimated-future-mortality-pathogens-epidemic-and-pandemic-potential

Mahdavi, P., & Ross, M. L. (2026). The worst energy policy in the world. *Science, 393*(6814), 845-845. doi:doi:10.1126/science.aej2018

Matsuo, H. (2015). Implications of the Tohoku earthquake for Toyota's coordination mechanism: supply chain disruption of automotive semiconductors. *International Journal of Production Economics, 161*, 217-227.

Maynard Smith, J. (1982). *Evolution and the Theory of Games*. Cambridge: Cambridge University Press.
Maynard Smith, J., & Price, G. R. (1973). The logic of animal conflict. *Nature, 246*(5427), 15-18.
Maynard Smith, J., & Szathmáry, E. (1997). *The Major Transitions in Evolution*: Oxford University Press.
Mayr, E. (1961). Cause and effect in biology. *Science, 134*(3489), 1501-1506.
Meadows, D. H., Meadows, D. L., Randers, J., & Behrens, W. W. (1972). *The Limits to Growth*. New York: Universe Books.
Merz, J. J., Barnard, P., Rees, W. E., Smith, D., Maroni, M., Rhodes, C. J., . . . Sutherland, R. (2023). World scientists' warning: The behavioural crisis driving ecological overshoot. *Science Progress, 106*(3), 00368504231201372. doi:10.1177/00368504231201372
Mesoudi, A. (2011). *Cultural Evolution: How Darwinian Theory Can Explain Human Culture and Synthesize the Social Sciences*. Chicago: University of Chicago Press.
metacrisis.org. (2026). Meta-Crisis Meta-Resource. Retrieved from https://metacrisis.org/META-CRISIS/03.+%F0%9F%8C%8E+Projects/The+Consilience+Project
Michod, R. E. (1999). *Darwinian Dynamics: Evolutionary Transitions in Fitness and Individuality*. Princeton, NJ: Princeton University Press.
Monbiot, G. (2023). *Regenesis: Feeding the world without devouring the planet*. London: Allen Lane / Penguin Books.
Muir, W. M. (1996). Group selection for adaptation to multiple-hen cages: selection program and direct responses. *Poultry Science, 75*(4), 447-458.
Nebel, A., Kling, A., Willamowski, R., & Schell, T. (2024). Recalibration of limits to growth: An update of the World3 model. *Journal of Industrial Ecology, 28*(1), 87-99. doi:https://doi.org/10.1111/jiec.13442
Nørgaard, B., Hedlund, N., & Meglin, C. (2025). *Mapping an Ecology of Integrative Approaches to Addressing the Metacrisis*. Retrieved from https://polycrisis.org/resource/mapping-an-ecology-of-integrative-approaches-to-addressing-the-metacrisis/
Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science, 314*(5805), 1560-1563.
Nuño de la Rosa, L. (2023). A History of Evolvability: Reconstructing and Explaining the Origination of a Research Agenda. In T. Hansen, C. Pelabon, D. Houle, & M. Pavlicev (Eds.), *Evolvability: A Unifying Concept in Evolutionary Biology?* : MIT Press.
Ó hÉigeartaigh, S. (2025). Extinction of the human species: What could cause it and how likely is it to occur? *Cambridge Prisms: Extinction, 3*, e4. doi:10.1017/ext.2025.4
Odling-Smee, F. J., Laland, K. N., & Feldman, M. W. (2003). *Niche Construction: The Neglected Process in Evolution*. Princeton, NJ: Princeton University Press.
OECD. (2017). *National Risk Assessments: A Cross Country Perspective*. Retrieved from https://www.oecd-ilibrary.org/governance/national-risk-assessments_9789264287532-en
Oliver, T. H., Doherty, B., Dornelles, A. Z., Greenwell, M., Harrison, L., Jones, I., . . . Gilbert, N. (2026). Appraising cascading systemic risks, 'watchpoints', and interventions –methodological reflections. *Global Sustainability, 9*, e4. doi:10.1017/sus.2025.10038

Oreskes, N., & Conway, E. M. (2010). *Merchants of Doubt: How a Handful of Scientists Obscured the Truth on Issues from Tobacco Smoke to Global Warming*. New York: Bloomsbury Press.
Ostrom, E. (1990). *Governing the Commons: The evolution of institutions for collective action*. Cambridge: Cambridge University Press.
Perrow, C. (1999). *Normal Accidents: living with high risk technologies* (2nd ed.). Princeton, NJ: Princeton University Press.
Pigliucci, M. (2008). Is evolvability evolvable? *Nature Reviews Genetics, 9*(1), 75-82.
Puma, M. J., Bose, S., Chon, S. Y., & Cook, B. I. (2015). Assessing the evolving fragility of the global food system. *Environmental Research Letters, 10*(2), 024007. doi:10.1088/1748-9326/10/2/024007
Rakowski, J. J., Schaan, L. N., van Klink, R., Herzon, I., Arth, A., Hagedorn, G., . . . Pe'er, G. (2025). Characterizing the global polycrisis: a systematic review of recent literature. *Annual Review of Environment and Resources, 50*, 159-183. doi:10.1146/annurev-environ-111523-102238
Rankin, D. J., Bargum, K., & Kokko, H. (2007). The tragedy of the commons in evolutionary biology. *Trends in Ecology & Evolution, 22*(12), 643-651. doi:10.1016/j.tree.2007.07.009
Rathje, S., Van Bavel, J. J., & van der Linden, S. (2021). Out-group animosity drives engagement on social media. *Proceedings of the National Academy of Sciences, 118*(26), e2024292118.
Richardson, R., Hendel-Blackford, S., Benini, L., Donges, J., Gibbons, B., Jácome-Polit, D., . . . Gambhir, A. (2026). Principles for just and effective systemic risk governance. *Global Sustainability, 9*, e21. doi:10.1017/sus.2026.10057
Richerson, P., & Boyd, R. (2005). *Not by Genes Alone: How culture transformed human evolution*. Chicago: University of Chicago Press.
Rowson, J. (2023). Prefixing the World: Why the polycrisis is a permacrisis, which is actually a metacrisis, which is not really a crisis at all. Retrieved from https://perspecteeva.substack.com/p/prefixing-the-world
Scheffran, J. (2025). Systemic risks and governance of the global polycrisis in the Anthropocene: stability of the climate–conflict–migration–pandemic nexus. *Global Sustainability, 8*, e39. doi:10.1017/sus.2025.10026
Schelling, T. C. (1960). *The Strategy of Conflict*. Cambridge, MA: Harvard University Press.
Scott, J. C. (1998). *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*. New Haven, CT: Yale University Press.
Sober, E., & Wilson, D. S. (1998). *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Cambridge, MA: Harvard University Press.
Søgaard Jørgensen, P., Delannoy, L., Maniatakou, S., Folke, C., Moore, M.-L., & Olsson, P. (2026). Both adaptive and transformative capacities are necessary to navigate global polycrisis. *Global Sustainability, 9*, e16. doi:10.1017/sus.2026.10053
Søgaard Jørgensen, P., Jansen, R. E. V., Avila Ortega, D. I., Wang-Erlandsson, L., Donges, J. F., Österblom, H., . . . Crépin, A. S. (2024). Evolution of the polycrisis: Anthropocene traps that challenge global sustainability. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 379*(1893), 20220261. doi:10.1098/rstb.2022.0261
Steffen, W., Richardson, K., Rockström, J., Cornell, S. E., Fetzer, I., Bennett, E. M., . . . Sörlin, S. (2015). Planetary boundaries: Guiding human development on a changing planet. *Science, 347*(6223), 1259855. doi:10.1126/science.1259855
Sterelny, K. (2006). The evolution and evolvability of culture. *Mind & Language, 21*(2), 137-165. doi:10.1111/j.1468-0017.2006.00284.x

Sterelny, K. (2007). SNAFUS: an evolutionary perspective. *Biological Theory, 2*(3), 317-328.
Sterelny, K. (2012). *The Evolved Apprentice: How Evolution Made Humans Unique*. Cambridge MA.: MIT Press.
Stigler, G. J. (1971). The theory of economic regulation. *The Bell Journal of Economics and Management Science, 2*(1), 3-21. doi:10.2307/3003160
Tinbergen, N. (1963). On Aims and Methods in Ethology. *Zeitschrift fur Tierpsychologie, 20*, 410-433.
Turchin, P. (2023). *End Times: Elites, Counter-Elites, and the Path of Political Disintegration*. New York: Penguin Press.
UN General Assembly. (2024). *The Pact for the Future (A/RES/79/1)*. Retrieved from https://docs.un.org/en/A/RES/79/1
UNDRR. (2015). *Sendai Framework for Disaster Risk Reduction 2015-2030*. Retrieved from Geneva: https://www.undrr.org/publication/sendai-framework-disaster-risk-reduction-2015-2030
UNDRR. (2023). *Report of the Midterm Review of the Implementation of the Sendai Framework for Disaster Risk Reduction 2015-2030*. Retrieved from Geneva: https://www.undrr.org/publication/report-midterm-review-implementation-sendai-framework-disaster-risk-reduction-2015-2030
UNDRR. (2025). *Global Assessment Report 2025: Resilience Pays: Investing and Financing for Our Future*. Retrieved from https://www.undrr.org/gar/gar2025
United Nations. (1997). *Kyoto Protocol to the United Nations Framework Convention on Climate Change*. Retrieved from https://digitallibrary.un.org/record/250111
United States Congress. (2022). *Global Catastrophic Risk Management Act of 2022, S. 4488, 117th Cong*. Retrieved from https://www.congress.gov/117/bills/hr7776/BILLS-117hr7776enr.pdf#page=1290
Unruh, G. C. (2000). Understanding carbon lock-in. *Energy Policy, 28*(12), 817-830.
Van Valen, L. (1973). A new evolutionary law. *Evolutionary Theory, 1*, 1-30.
VCDNP. (2026). *End of New START: Short- and Medium-Term Options*. Retrieved from https://vcdnp.org/end-of-new-start/
Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science, 359*(6380), 1146-1151. doi:10.1126/science.aap9559
Wagner, G. P., & Altenberg, L. (1996). Perspective: complex adaptations and the evolution of evolvability. *Evolution, 50*(3), 967-976.
Walker, E. T. (2014). *Grassroots for Hire: Public Affairs Consultants in American Democracy*. Cambridge: Cambridge University Press.
Waring, T. M., Wood, Z. T., & Szathmáry, E. (2023). Characteristic processes of human evolution caused the Anthropocene and may obstruct its global solutions. *Philosophical Transactions of the Royal Society B: Biological Sciences, 378*(1893), 20220259. doi:10.1098/rstb.2022.0259
West-Eberhard, M. J. (2003). *Developmental Plasticity and Evolution*. New York: Oxford University Press.
Wilson, E. (2009). *Debate at the Harvard Museum of Natural History*. Cambridge, Mass.
World Economic Forum. (2026). *The Global Risks Report 2026* (978-2-940631-60-5). Retrieved from Geneva, Switzerland: https://www.weforum.org/publications/global-risks-report-2026/
Wright, S. (1932). *The roles of mutation, inbreeding, crossbreeding, and selection in evolution.* Paper presented at the Proceedings of the Sixth International Congress of Genetics.