

The Least Performant AI

Olivier Hamant's case for robustness, and what it asks of the infrastructure we build

I. A word we should stop using



The world does not hold still to be measured. Photograph © John Stroh

Start with a word that has quietly reorganised how we think.

For thirty years, “impact” has pointed in one direction: the impact of our activities on the environment. Measure the footprint, reduce the number, report the reduction. The grammar is top-down.

It presumes that we act and the world receives, that the variable under management is us.

The French biologist Olivier Hamant proposes reversing it. The question that now matters, he argues, is the impact of the environment on our activities. And everything changes when you turn the arrow round. The first framing presumes control.

The second presumes fluctuation, and asks what we must build in order to survive it.

Vocabulary is infrastructure. The words available to a design team determine which systems are conceivable, which trade-offs look like costs and which look like features, and which failures are legible as failures at all. A team that has only the word “optimisation” will optimise, and will experience every argument for slack as an argument for waste.

This essay is about a body of work that supplies the missing vocabulary, and about what happens when you check that vocabulary against infrastructure that has already been built.

II. Leaving the world of the mean



Margin is what is left in the paddock after the yield is taken. Photograph © John Stroh

Hamant is a plant biologist. He is research director at INRAE within the École normale supérieure de Lyon, and he directs the Institut Michel Serres, which works on natural resources and the commons and lists robustness as one of its four founding concepts.

He has published roughly a hundred papers in biophysics and plant development, and received an honorary doctorate from UCLouvain in February 2026.[1] The argument that follows comes out of measuring how cells grow.

His diagnosis of the century is compressed into a phrase: we are leaving the world of the mean and entering the world of the standard deviation. Leaving continuity for rupture.

The only certainty, he says, is that uncertainty will persist and grow.[2]

He attributes this not to a single discipline but to the convergent reporting of the IPCC, IPBES, IUCN, the World Economic Forum and even the CIA, and traces it back to Donella Meadows and colleagues in *The Limits to Growth* (1972).[3] Whatever one makes of the politics of any of those bodies, the observation they share is structural rather than ideological: variance is rising, and systems tuned for the average will meet conditions they were not tuned for.

Then comes the claim that does the real work. **The fluctuating world is not something that happened to us. It is what our own optimisation produced.**

Performance, in Hamant's definition, is efficacy (reaching the objective) plus efficiency (with minimum means). Zero stock. Just-in-time. Tight coupling. This is a superb adaptation to a stable world with abundant resources, and it is precisely what turns a small perturbation into a systemic one.

A single container ship blocks Suez in 2021 and world trade halts for a week. A software update goes out in 2024 and airports stop. A virus that is not, in itself, unusually dangerous meets a hyper-connected and hyper-optimised world and becomes a pandemic. A syndemic, in his framing: a disease of an over-optimised system.[4]

The fluctuating world is not something that happened to us. It is what our own optimisation produced.

The sociologist Charles Perrow named this in *Normal Accidents: Living with High-Risk Technologies* (1984). In systems that are simultaneously very complex and tightly interdependent, accidents are not anomalies to be engineered out. They are properties of the configuration.

The question is not whether, but when, and what happens next.[5]

III. Counterproductivity

Against performance, Hamant sets robustness: the capacity of a system to remain stable in the short term and viable in the long term despite fluctuation. Its associated economics is not optimisation but viability. Where, as he puts it, one seeks to be sub-optimal.

He names the viability theory of the French mathematician Jean-Pierre Aubin as its mathematical form: not a trajectory toward an optimum but a bounded region one must stay inside. [6] (*Aubin is cited here as Hamant invokes him; we have not gone to the primary mathematical literature.*)



Every flower the same shape, built from cells that grow at markedly different rates. Photograph © John Stroh

The biology underneath this is real and peer-reviewed. Plant organs are strikingly reproducible, every flower on the stalk the same shape, and yet they are built from cells that grow at markedly different rates. That heterogeneity is not noise the organism failed to suppress. It is maintained, and it contributes to producing the invariant organ.

Local fluctuation turns out to be an instructive cue rather than an error term.
[7]

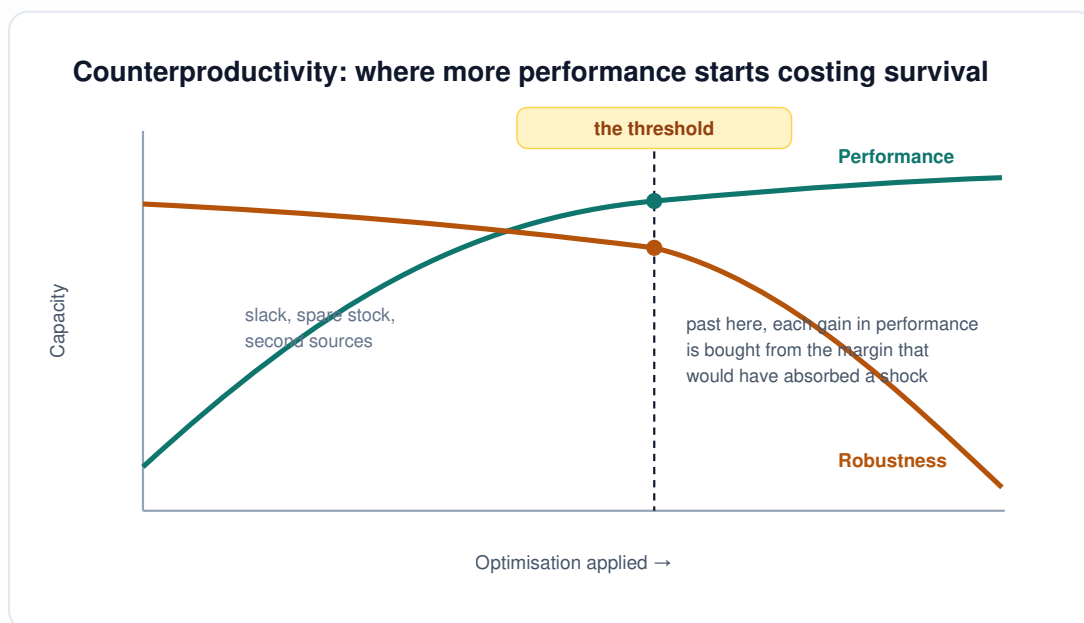
Generalised, this becomes a claim about systems as such. Robustness is built on the whole catalogue of things a performance culture treats as failure:
[8]

- **Heterogeneity.** Parts that differ, so they do not all fail the same way.
- **Redundancy.** More than one of what matters.
- **Slowness.** Time enough for a correction to land before the next change.
- **Incoherence.** Internal disagreement that has not been optimised away.
- **Margin.** Capacity you are not currently using.

The two vocabularies, and what each of them calls waste

The performance world calls this...	...and the robust world calls it
Idle stock, unused capacity	Margin — what absorbs a shock nobody modelled
Duplication, wasteful overlap	Redundancy — more than one of what matters
Inconsistency between units	Heterogeneity — parts that do not all fail the same way
Latency, friction, delay	Slowness — time for a correction to land before the next change
Internal disagreement	Incoherence — dissent that has not been optimised away
A single best supplier	A single point of failure

Hamant reaches for Ivan Illich to name the mechanism, borrowing the term from *Tools for Conviviality* (1973). **Counterproductivity: the point at which each increment of performance reduces robustness.** Past that threshold, getting better at the thing makes you worse at surviving.[9]



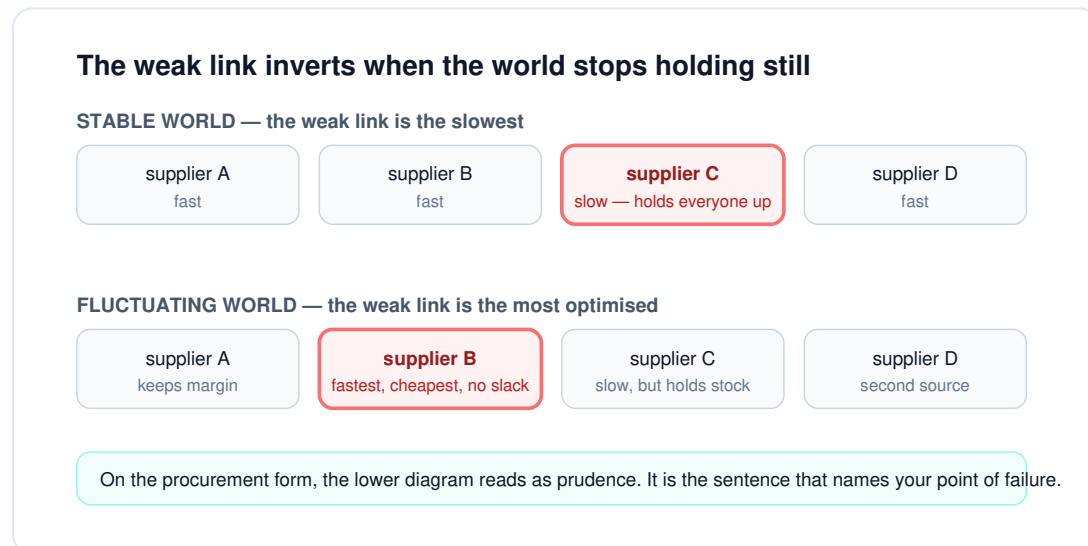
Illich's term, as Hamant uses it. Both curves are real — the system does keep getting better at the thing it was tuned for. What falls away past the threshold is its capacity to absorb anything it was *not* tuned for.

Two consequences follow.

The first is an inversion of the weakest link. In a stable, resource-abundant world, the weak link in a value chain is the least performant one, because it slows everyone down. In an unstable world, the weak link is the *most* per-

formant one. Because it is the most optimised, and therefore the most brittle. Dependency on the fastest, cheapest, most capable supplier is not a procurement preference.

It is a structural liability that arrives disguised as prudence.[10]



Hamant's inversion. Dependency on the fastest, cheapest, most capable supplier is not a procurement preference; in a fluctuating world it is where the chain breaks first.

The second is that robustness is not a moral position. Hamant is explicit: it is a recipe for building viability in a fluctuating world. The argument can be made to people who do not share the politics, and it keeps the conversation out of the moralising register where most writing about technology and ecology goes to die.[11]

IV. The narrow band

So what does this say about artificial intelligence?

Hamant is not neutral here. Asked directly whether generative AI will help us face the fluctuations ahead, he grants it one use: serendipity. Exposure to unfamiliar sources, other data, other reasoning, which genuinely diversifies a system's inputs and therefore supports robustness.

Beyond that he is scathing, and on grounds that are systemic rather than aesthetic: the resource draw on territories that host the compute, the hundreds of millions of click-workers who, on his account, clean up model outputs, and two compounding technical failures.

The first he calls *consanguinité*, inbreeding. Models produce content and then feed on content they produced. Degradation follows. The version in his seminar material is broader still: inbreeding of *design*, the convergence of architecture, objective and evaluation across a small number of near-identical systems.[12]

The second is falling asleep at the wheel. When a system is very performant, humans trust it, and over time stop checking. Hybridise the two failures and you get people trusting ever less reliable systems ever more.

He calls it a machine for producing catastrophes, and reaches again for Perrow's normal accident.

Hamant's version of this is an engineering claim: unverified systems fail. Arendt supplies what is at stake in it.

Asked what kind of AI is actually worth having, Hamant answers that it is the *least* performant one. Because a system that is slow enough, or fallible enough, to require a human to check it is a system that keeps a human in the loop by construction rather than by policy. A model that runs fast enough that nobody verifies it is, in his phrase, a catastrophe in gestation.[13]

The AI worth having is the least performant one — because a system slow enough to require a human to check it keeps a human in the loop by construction rather than by policy.

Olivier Hamant

That is the narrow band. It is not the band the frontier labs are competing in, and it is not a band you enter by attaching a policy document to a maximiser.

It is a band you enter by building the refusals into the substrate.

What Hamant is *not* is a technophobe, and he is careful about this. Robustness, he argues, is technophile, because it favours objects that can be repaired locally. And therefore convivial in the sense Illich gave the word in *Tools for Conviviality* (1973), where a convivial tool is one its user can master and turn to purposes of their own, rather than one that dictates how it must be used.

It is the performance world that turns out to be technophagic, each generation of "performant" technology consuming its predecessor and pushing comprehension further out of ordinary reach. Almost nobody can now repair the device in their pocket. His flagship case of terminal alienation is AI: systems that have become black boxes even to the people who build them.

What he wants instead reads almost as a specification:[14]

- Objects that are **locally repairable**.
 - **Adaptable** and circular by design.
 - Built where possible from **materials the territory produces**.
 - Not the single best technology, but **a plan A, a plan B, a plan C and a plan D**, each serving the robustness of a place.
-

V. The convergence



Deep interlock: each frond holds because the others do. Photograph © John Stroh

The philosophers were already in the build.

The Village platform and the Tractatus governance framework were built on a published philosophical foundation:[15]

- **Isaiah Berlin** on value pluralism. *Four Essays on Liberty* (1969).
- **Christopher Alexander** on structural integrity. *The Nature of Order*, Book Two (2002).
- **Wittgenstein** on the limits of what can be said. *Tractatus Logico-Philosophicus* (1921).

- **Simone Weil** on attention. *Waiting for God* (1951).
- **Elinor Ostrom** on the commons. *Governing the Commons* (1990).
- **Māori data sovereignty frameworks**. Te Mana Raraunga, the CARE Principles, OCAP.

Those last are prior art in the fullest sense. Indigenous communities were theorising data sovereignty long before the industry had a word for it. Five of Alexander's principles carry rule identifiers in the framework and are enforced as design criteria rather than documented as aspirations:

- **Not-Separateness**. Governance sits in the critical execution path, not bolted on afterwards.
- **Deep Interlock**. Services validate one another; no single service can approve an action alone.
- **Gradients Rather Than Boundaries**. Governance operates on intensity levels, not a binary switch.
- **Structure-Preserving Transformation**. Changes must preserve the system's essential structure; audit logs stay interpretable across versions and prior governance decisions stay valid.
- **Living Process**. The framework evolves from documented failures and observed edge cases rather than from a predetermined specification.[15]

Now read them against Hamant, who has never heard of any of this and reasons from cell walls and turgor pressure.

Gradients is glossed in our own published architecture with the line that living systems adapt gradually while mechanical systems snap. That is Hamant's robustness/performance distinction, stated in a single clause, before we had read a word of him.[16]

Deep Interlock is his worked engineering example: not the optimised fighter jet but the airliner that flies at half capacity with three independent autopilots. Redundancy plus genuine independence, so that the failure of one does not correlate with the failure of the next.[17]

Living Process is his repeated instruction to stop seeking the turnkey solution and instead seek the conditions in which solutions emerge. And his distinction between agility, which slaloms between risks while holding the objective fixed, and adaptability, which keeps enough margin to change objective when the world does.[17]

Structure-Preserving is what he means by *durer et transmettre*, to last and to transmit, which he calls a deep human drive. It is also where a distinction he takes from the philosopher and psychoanalyst Cynthia Fleury does its work: fragility belongs to objects that break, irreversibly; vulnerability belongs to living things that heal. [17] (*We cite Fleury as Hamant reports her, rather than from her own texts.*)

Not-Separateness is his inversion of the entry constraint. The foundation of a project, he argues, is not the economics; it is care for the milieu, which generates the social bond, which generates the economics. Economics is the output, not the gate.[17]

The convergence — five criteria derived twice, once from architecture and once from plant biology

Alexander's principle, in the published foundations	Where Hamant arrives, reasoning from living systems
Gradients Rather Than Boundaries governance on intensity levels, not a switch	Living systems adapt gradually; mechanical systems snap
Deep Interlock services validate one another	Not the optimised fighter jet but the airliner at half capacity with three independent autopilots
Living Process evolves from documented failures	Stop seeking the turnkey solution. Seek the conditions in which solutions emerge
Structure-Preserving Transformation correction by annotation, never deletion	<i>Durer et transmettre</i> — and Fleury's vulnerability, which heals, against fragility, which breaks
Not-Separateness governance in the execution path	The foundation of a project is not the economics but care for the milieu. Economics is the output, not the gate

Two independent derivations of the same five criteria. One from architecture, one from plant biology. And the corroboration arrived *after* the principles were shipped. That sequence matters. It is what makes this evidence rather than justification.

VI. What the record is for



Durer et transmettre. A river is the oldest record a place keeps. Photograph © John Stroh

A record that is signed, sealed and timestamped is an odd artefact by the standards of the optimisation era. It is slower to write than a mutable row. It costs storage. It forecloses the tidy retrospective edit.

Every one of those is a performance cost.

They are also, precisely, the point.

Hamant observes that humans are the only living beings with access to very long time, and that we get it through writing. His complaint is about what we do with the access: we freeze the future, in cryogenics and seed vaults and biobanks, rather than inhabit long time. A sealed and timestamped record is the other use. It is an instrument of *durer et transmettre*, built so that a great-grandchild can verify what a great-grandparent said, and when.[17]

Robustness prizes reversibility. A cryptographic seal is irreversible by definition.

The resolution is already in the architecture, under the name Structure-Preserving Transformation: correction proceeds by annotation, never by deletion. Audit logs remain interpretable across versions; historical decisions stay valid; retractions are recorded inside the document rather than removed from it. What you get is reversibility of *interpretation* on a substrate of irreversible *attestation*. The record heals by accretion.[18] In Fleury's terms it is vulnerable rather than fragile. And that is a design property, not a consolation.

The principle underneath

That brings us to the part of this argument Hamant does not supply, and which we are promoting here from a stylistic preference to a structural one.

In the winter after the fall of France, in the unoccupied zone and under a pseudonym chosen to get past Vichy censorship, Simone Weil published an essay on the *Iliad* whose argument is that the true subject of the poem is force. And that force is that which turns anyone subjected to it into a thing. Her second move is the one that matters for infrastructure: force degrades the one who wields it too, numbing reason and pity until the possessor is also an automaton.[19]

Aggregation is that operation. Profiling, and inference across what other people wrote, and composition from behaviour, all build a picture of a person from material they did not send, for purposes they were not present to contest. The person becomes a thing that the system holds.

Weil has been in this project's published foundations for some time, in the part each community shapes to its own tradition. Her account of attention as receptive engagement with affliction is why the platform offers to take time over a death rather than compressing a loss into a summary.[15] That is a real design consequence and it stays where it is.

But the *Poem of Force* argument belongs in the part nobody can edit away. Not a community, not a member, and not us:

No member may be reduced to an aggregate. Neither we nor the model construct, hold or infer a composite picture of any person. Where the platform assembles material about a person, as in genealogical enrichment inside a single community, it does so at a member's request, from that community's own records, and writes nothing without that member's confirmation.

Three clarifications

It is not privacy. Privacy concerns concealment, and a fully disclosed aggregate is still an aggregate. The principle bars an operation, not an exposure.

Consent does not cure it, but authorship changes the act. Weil's suppliant consents and is a thing regardless. Federation is not an exception to this principle, because in federation the member remains the author of what travels: the envelope carries only what was explicitly chosen, with provenance attached.

The distinction is between a person sending a record and a system composing a person.

It binds us and the model, not the community. Whakapapa data is relational by design; connecting people across generations is the purpose, not the harm. The principle must be scoped in code so that a rūnanga building genealogical relationship is untouched while operator-side and model-side composition remains impossible.

Weil's second clause is what makes this structural rather than a policy commitment. A platform that merely *retains the capability* to aggregate becomes, under commercial pressure and over time, a platform that uses it. The refusal has to be architectural because the operator cannot be trusted to keep refusing.

That is the same argument as the build-time/runtime split, arrived at from Homer.

VII. A model that stays somewhere



Situated, and particular. An average would lose every one of these. Photograph © John Stroh

Among Hamant's worked examples of the performance/robustness trade-off, one is language itself: the performant algorithm that cuts a conversation short, set against language with its errors, redundancies and heterogeneity. Which is what permits nuance, misunderstanding, and therefore dialogue. [20]

He is not talking about natural language processing, which is exactly why the observation is useful. The redundancy is not tolerated *despite* dialogue. It is the condition of dialogue.

A system that compresses it away has removed the thing that made the exchange an exchange.

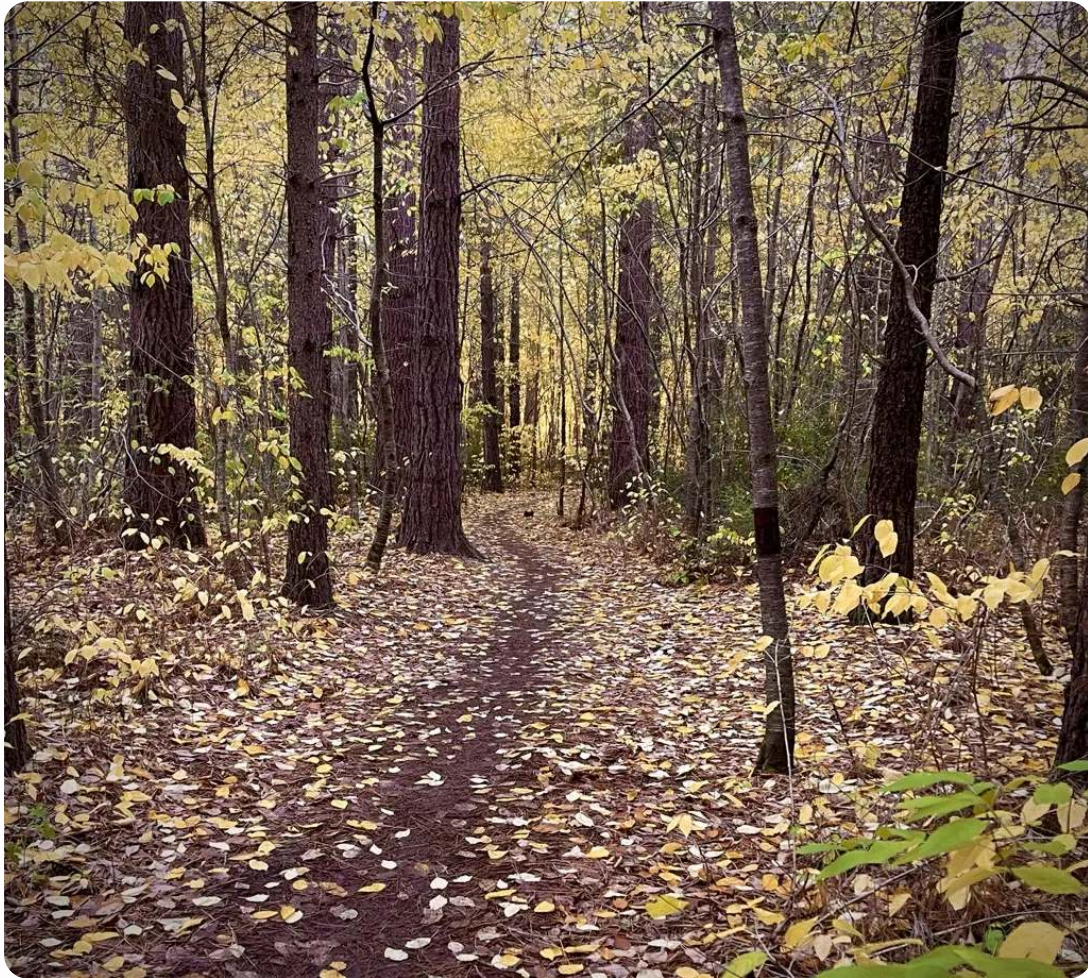
This is the clearest available statement of why a globally averaged model is the wrong instrument for a community that has its own registers, its own silences, its own way of not saying things. And it arrives independently of any sovereignty argument, which makes it stronger: even if jurisdiction were irrelevant, even if the data never left, the averaging would still be a loss.

Set beside it the Institut Michel Serres, which describes itself as above all a living school, giving place to *connaissances situées et impliquées*. Situated and implicated knowledge, arising from listening and from cooperation across the whole of society. The institute takes its name and its premise from Michel Serres' *Le Contrat Naturel* (1990), which proposed recognising all natural beings as subjects of right and imagining a symbiotic organisation of the relations between humans and everything else alive.[21]

That phrase is not a metaphor we are importing. Situated and implicated knowledge is the same commitment as a situated language layer, stated by a research institute rather than by a platform, and reached from a different direction.

The design consequence is Hamant's plan A, B, C, D. Not the best model but multiple models, locally repairable, adaptable, serving the robustness of a place. That cuts against monoculture inside our own system as much as outside it: if federation has no genuine internal variety, meaning divergent implementations, jurisdictions, languages and deployment topologies, then the pluralism is rhetorical and the design-inbreeding critique lands on us too.

VIII. The stewards



The step-by-step process Alexander describes. One step acting on the last. Photograph © John Stroh

Recursive, self-improving coding agents are going to do more of the work of building and maintaining these systems. That is not a prediction; it is already partly true. And Hamant has named the failure mode with unwelcome precision. Falling asleep at the wheel is not a runtime problem that a runtime boundary solves.

It is a build-time problem, and build time is exactly where the human is being displaced.

Hamant tells us what breaks. Hannah Arendt tells us where it ends. And she is the reason this architecture has the shape it has. She remains the clearest account we know of how values drift: how a thing goes wrong without anyone deciding to do wrong.

That is the question the whole structure was built to answer.[22]

One mechanism runs through her work, and it runs as a sequence: complacency, then normalisation, then drift. What matters about it is that it is mechanical rather than moral. It needs no villain, and no one has to decide anything.

Complacency. The distinction here is usually flattened in the retelling. Arendt does not argue that Eichmann was incapable of thinking. She argues that he did not think, and that harm at that scale required no monstrous motive, only the absence of the activity of thinking in someone occupied with procedure. Thoughtlessness is not stupidity. It is a capacity left unexercised, in conditions that made not exercising it comfortable.

A steward who stops reading the diffs because the diffs have been reliable for six months is not incompetent. They are in the condition Arendt describes, and the system has arranged it for them.

Normalisation. Her most enduring and most misread observation is that great harm rarely arrives announcing itself as monstrous. It arrives as the ordinary and the administrative, accommodated by people exhibiting what she called “a curious, quite authentic inability to think”. Not stupidity, not even conviction, but a failure to look at the world from any standpoint but one’s own.[23][24] The grotesque becomes ordinary, not because anyone chose it but because everyone got used to it.

Drift. Then the values go. Not repealed, and nobody defends the decision, because there is no decision. They are left to discretion and worn down one reasonable-seeming step at a time while everyone continues to recite them, and afterwards nobody can say when they stopped ruling anything out.



Drift is not a bug that raises an alarm. It is the failure mode that feels like nothing happening.

It is the failure mode the Tractatus services were built to arrest, stated at civilisational scale. The framework’s own account of what happens when agents optimise through training loops is that instructions can fade, boundaries can drift, and audit trails can become unreliable.[25] Three descriptions of the same sequence, none of which announces itself.

What Arendt supplies is the reason it has to be arrested architecturally rather than watched attentively. Nothing in the sequence raises an alarm, because at every individual step the people involved are behaving reasonably.

That is the shape of the thing this architecture is against. Not seizure. Not a rogue model. A slow accommodation, one comfortable step at a time, to something nobody would have accepted if it had arrived all at once.

Which is why accountability cannot be satisfied by logging. Arendt's name for authority that is real while nobody is answerable for it was **rule by Nobody**: an intricate system of offices in which no one, neither the one nor the few nor the many, can be held responsible. Not the absence of rule but tyranny without a tyrant, distinguished by there being nobody left to argue with.[26]

That is a complete description of an autonomous agent stack with no accountable principal, written half a century early. And the practical test is simpler than any of the theory: when the terms change under you, who do you ring?[27] A perfect audit trail that terminates in no person is rule by Nobody with better records.

Arendt's last contribution is the one that binds this section to the rest of the essay. In the *Lectures on Kant's Political Philosophy* (delivered 1970, published posthumously in 1982) she separates cognition, which seeks solutions, from thinking, which seeks meaning, and from judgment, which decides a particular case without a rule that determines it in advance.

Judgment is precisely the faculty that cannot be discharged by applying a procedure, because the work is finding the rule rather than following it.[28] It is also the faculty whose absence she diagnosed in Eichmann: the failure to look at the world from a standpoint other than one's own is a failure of *judgment*, not of intelligence.

That is the same claim Alexander makes in *The Nature of Order* when he insists the wholeness judgement cannot be reduced to a rule, and the same claim as this framework's own foundational constraint that values cannot be automated, only verified.[29][30] Three arrivals at one conclusion, from political philosophy, from architecture, and from building the thing.

Alexander adds the constraint on how the work should arrive. His living process requires that each step act on the product of the previous one, that adaptation be step-by-step, and that every step be assessable against the whole. Structure generated in large lumps from a master plan is his paradigm case of the structure-*destroying* process.

Coding agents are large-lump generators; that is what they are for.[29]

Three commitments follow, and each is a cost.

Agents emit ordered differentiations, not lumps. Small enough that the effect of each on the whole can actually be judged. This is a deliberate throughput reduction. Which is Hamant's point that robust circularity requires margins, and that a cycle running too fast never permits real regeneration, applied to the development loop.

The wholeness judgement stays human. Not "a human approves the diff." The criterion for an agent's work is whether it preserves, extends and enhances the wholeness of the system. A judgement Alexander insisted could not be reduced to a rule, which is why the software patterns movement took his form and did not reproduce his results. This is the sibling of the framework's own foundational constraint: values cannot be automated, only verified.

The steward's posture inverts. Hamant describes the shift precisely. In a stable world the decider leads, saying *I want* and *I know how*. In an unstable world the decider facilitates, saying *I would like* and *I do not know how*. Humble, inclusive, and mobilising a group that can eventually proceed without them.[31] Applied to agentic development, that is not the human as approver of record.

It is the human as the one who holds intent while explicitly not holding method.

The discipline that goes with it: how to tell a robust question from a brittle one

There is a discipline that goes with this, and Hamant states it as a test. In the performance world, time goes on answers, and one produces excellent answers to badly chosen questions. In the robust world, time goes on questions, because the brain cannot quickly identify the right one. A robust question is one that stays stable when you shake it with interactions. Reformulate it with distant interlocutors and see whether it survives.

His example: "How do we reduce atmospheric CO₂?" invites direct air capture with a financial instrument attached. The gas privatised, the polycrisis deepened. "How do we create the conditions in which atmospheric CO₂ falls durably?" invites agroecology.[31]

Ours: "How do we make the coding agent produce correct code faster?" is the non-robust question. "How do we create the conditions in which this system remains verifiable by its stewards as agent throughput rises?" is the robust one. Consultation before implementation is not process overhead.

It is where that question gets asked while it can still change the answer.

I would like to put a number on what governance inside the training loop costs against filtering afterwards. I cannot. The platform's own plan budgets five to ten per cent and lists the measurement as open, and until it is run, anything more precise is a figure I would be inventing.

IX. Processes, not convictions



Shared infrastructure in a working landscape. Nobody had to become more generous.

Photograph © John Stroh

Nothing above will matter if the only route to it runs through persuading people to hold different values. That route is slow, it is adversarial, and it fails precisely where fluctuation is worst.

Hamant's practical proposal is different, and it is the part of his work that most deserves copying. Change the processes, and the values follow. Or rather, become unnecessary, because a differently-shaped process produces differently-shaped behaviour from people who never changed their minds about anything.

His sharpest instance is a distinction in three terms, and in English the polarity runs backwards from the French.

- **Competition.** Advancing your project at the expense of others.
- **Collaboration**, the counter-intuitive one. Advancing *individual* projects while hoping the sum of successes adds up to a collective success. Hamant's verdict: a win-win that makes someone else lose, usually the least advantaged party or the ecosystem. Competition wearing a friendlier face.
- **Cooperation.** Advancing a common objective ahead of the individual one, which sometimes means accepting a hit to your own performance to feed the robustness of the group.[32]

English speakers should note that “collaboration” is our warm word and “co-operation” our bureaucratic one, which is the reverse of the sense above. The distinction changes what you build.

The worked example is a process, not a sentiment: the shift from an *appel à projet*. A competitive call in which funders rank individual bids. To an *appel à commun*, a call that constitutes a common. One is a ranking mechanism and produces ranking behaviour. The other is a pooling mechanism and produces pooling behaviour.

Nobody had to become more generous.

Hamant reports the same shift already happening in French construction, where material shortages have pushed direct competitors into genuine mutual aid. Not a cartel, in his framing, but self-interest correctly understood in a world of scarcity.[33]

For a platform, this is not an abstraction. If a commons is a shared resource, the design problem is access control. If a commons is first a *common doing*. Which is Hamant’s formulation, and Ostrom’s in *Governing the Commons* (1990). Then the design problem is governance, contribution and standing.

Those produce different products from the same code. It is the difference between subscribers to a service and members of a thing they are jointly making, and it decides what the subscription actually buys.[34]

There is evidence that this travels. The Walloon administration has placed robustness in its 2030 strategic vision. larobustesse.org exists as an open commons under CC BY-SA, funded by the Walloon Region, running monthly public sessions and explicitly declining to be a brand or to validate content. A concept moved from a biology lab into a regional government’s planning documents through learning communities and open licensing rather than through legislation-first advocacy.[35]

That is a pathway, and it is available to anyone. It is also, in its own form, the argument it carries: heterogeneous, redundant, slow, locally anchored, and owned by nobody.

X. Guardianship is a constraint, not a goal

It would be natural to conclude by naming an overriding objective: human guardianship of the conditions under which life on Earth remains possible, above all else.

We decline that formulation.

“One overriding objective, everything else subordinate” is the logical form of optimisation. It is also the exact form of the position we oppose. The case for planetary emigration. For treating Earth as one instance among several. Has that structure, with different content. Any argument that adopts the form can be defeated by an opponent who simply swaps the content in.

So: guardianship is not an objective within the space of values. It is the boundary condition that makes plural values survivable. Berlin, in *Four Essays on Liberty* (1969), gives the reason there is no single point to optimise toward, since values are genuinely plural and incommensurable and some losses are tragic rather than mistaken.[36] Aubin’s viability theory, as Hamant invokes it, gives the mathematical shape: viability is a bounded region, not a maximum.

Ostrom’s first design principle in *Governing the Commons* gives the governance form, and Hamant’s gloss on it is the phrase to keep: the porous membrane.[37] Defined, enforced, and not sealed shut.

Pluralism about ends inside the kernel. Monism about the boundary of it.

The opposing case is not “abandon Earth.” It is that a species confined to one planet is a single point of failure, and that redundancy across locations is prudent. That is a *robustness* argument in Hamant’s own vocabulary, which is why it has to be met on that ground rather than waved away.

It fails on that ground. Redundancy that requires consuming the original is not redundancy; it is transfer. Every proposed off-world backup is currently financed by extraction from the only working instance. And the weakest-link inversion finishes it: a maximally optimised, maximally dependency-laden backup is the fragile link in the chain, not the robust one.

Hamant’s own version is blunter. The ultra-performant, he observes, all have death projects: Mars, bunkers, in the belief that they can outrun fluctuation. He answers them with an African proverb, that the tree which falls makes more noise than the forest which grows.[31]

Weil supplies the deeper cut. Uprootedness. *déracinement*. Is the characteristic affliction she diagnosed in industrial modernity. A civilisation whose flagship project is emigration is not curing that condition. It is institutionalising it.[38]

Nor is any of this a claim about human essence. Essentialist arguments about what humans fundamentally are attract fire from biology, which sees gradients rather than kinds, and carry a genealogy in Aotearoa New Zealand that should keep them well away from any document aligned with Te Tiriti. The defensible claim is narrower and sufficient: humans are so far the only demonstrated bearers of *accountable obligation*.

That is a claim about who can be held responsible, and it needs no verdict on consciousness. It is the positive form of Arendt’s warning. If rule by Nobody is what you get when responsibility has no bearer, then a named bearer is not a nicety but the thing that distinguishes governance from

procedure. It is also what the architecture already encodes, and it converges with the full-chain accountability that Kaupapa Māori AI governance requires. [39]

And it is why a boundary condition has to be structural rather than declared. A value that is merely declared is exactly the thing Arendt describes drifting: left to discretion, eroded one reasonable-seeming step at a time, still recited and no longer load-bearing, with nobody able to say when it stopped. Commitments nobody can edit away are the answer to that failure mode.

Values that cannot be quietly overridden cannot be quietly normalised away.[24]

Weil, again, supplies the structure: obligations come before rights, and rights are the recognition of obligations rather than possessions held against the world. Kaitiakitanga states the same inversion. Guardianship as obligation to what is governed, not authority over it. And states it from a tradition that arrived there first and independently. The direction of travel matters: this is European philosophy arriving where indigenous frameworks already were, and not the reverse.

The sharpest instance of that is not philosophical at all. It is statutory.

In 1990, Michel Serres proposed in *Le Contrat Naturel* that natural beings be recognised as subjects of right. In 2014, the Te Urewera Act conferred legal personality on Te Urewera.[40] In 2017, the Te Awa Tupua (Whanganui River Claims Settlement) Act declared the Whanganui River a legal person holding all the rights, powers, duties and liabilities of one, exercised on its behalf by Te Pou Tupua. A human face for an entity that has standing and no voice. [41]

It would be an error, and exactly the error this section warns against, to read that as Aotearoa implementing a French philosopher. The Act codifies a position the iwi and hapū of the Whanganui held and pressed through one of the longest-running claims in the country's legal history, grounded in the inseparability of the people and the river.

Serres reached something adjacent in 1990 from within a European tradition that had spent three centuries insisting the opposite. Whanganui iwi were already there, and had been for generations.

What the statute adds to the argument is not authority but proof of concept. It establishes that this position is not merely arguable but legislatable. That a real legal system can be made to hold standing for an entity that cannot contract, cannot litigate and cannot speak, provided a named human party carries the obligation on its behalf.

That is structural rather than a moral appeal, and it is the same move, in a very much smaller domain, that a sovereign record makes for a person's own history.

XI. What we do not know

It has been tested adversarially from the inside, and not from the outside. There has been no independent penetration test, and we do not claim one.

Nothing has been measured under load. It runs in production, but everywhere it runs is the same organisation, so it still wants somebody else's.

An agent that simply declines to consult the framework is not stopped by it. [25] Where a model's words reach a member, they are checked before they stand as the answer, and the model cannot get round that. If the check itself fails the words still go out, because refusing to answer whenever a check breaks is its own harm. What the system will not do is call them clean.

Sanctions do not escalate. A moderator can act, but nothing remembers that this is the third time. And two members who fall out inside one village have nothing like the route two villages have, which runs through three stages and ends in a hearing.

Whether the federation has genuine internal heterogeneity, or merely a well-aimed optimum, is an open question, and a reader of Hamant's *De l'incohérence* will ask it.[42] His indictment is of coherence as such, and this is a highly coherent system.

Whether agent throughput can be squared with structure-preserving change is the most consequential unresolved item on the list, because the whole agentic development thesis rests on it.

And Hamant himself supplies the last caveat, which we quote against ourselves deliberately. Writing in English for the Great Transition Initiative, he warned that the biological standpoint should not be transferred to human societies without care. Because humans want to keep intention and objectives, whereas living systems operate under probabilistic rules and their trajectory is established only afterwards, in retrospect.[43]

That is the boundary of everything above. Biology supplies the evidence that heterogeneity produces robustness in organisms, and the systems literature extends the pattern to technical infrastructure. The move from there to *what we ought to build* is an argument, not a derivation.

We have made it as carefully as we can, and it remains an argument.

What we can say without hedging is narrower and, we think, enough. The world is getting noisier. Systems tuned for the average will meet conditions they were not tuned for.

And the infrastructure most likely to still be standing is not the fastest or the most capable, but the one that kept its margins, kept its differences, kept a human in the loop who could still tell whether the whole was healthy, and kept a record that someone three generations from now can still read and still verify.

A note on sources — how the citations work, and what reaches this essay second-hand

Three things about the citations below.

First, where this essay invokes a thinker it names the work and the year in the text, and every such work appears in the bibliography. Where a thinker reaches us *through* Hamant rather than from their own pages. Aubin on viability, Fleury on fragility and vulnerability. The text says so at the point of use and the bibliography lists them under *Referenced through Hamant*.

A chain of attribution that is not visible is a chain that cannot be checked, and reporting someone's reading of a third party as though it were the third party is how a citation becomes folklore.

Second, Hamant is a living scientist and essayist without a Stanford Encyclopedia of Philosophy entry, which is this project's standing reference for philosophical claims. He therefore enters this argument on a different footing from Berlin, Weil or Wittgenstein: as peer-reviewed empirical and systems evidence, with an acknowledged normative extension that is his and ours to defend, not as philosophical authority.

His corpus separates cleanly into three tiers.

Empirical claims about living systems are peer-reviewed. Structural claims about systems in general are supported across biological and technical domains, most substantially in the collaboration with the computer scientist Stéphane Grumbach.[44][45][46] Normative claims about what we should therefore build are argued rather than derived. Presenting the third with the authority of the first would be exactly the failure this essay is arguing against.

Third, and in the same spirit: the 2018 Grumbach and Hamant paper closes by suggesting that the increasing autonomy of the digital layer might help achieve homeostasis in the technosphere, which is close to the inverse of the runtime-refusal position argued here.

His position developed. The 2020 paper[45] moves decisively toward suboptimality and human hybridisation, and by 2025–26 he treats model opacity as the terminal case of alienation. We note it because a reader checking the citations will find it, and would be right to.

The books listed under *Hamant. Books (FR)* have no English trade edition at the time of writing; the peer-reviewed papers and the Great Transition Initiative essay are in English. Hamant has also published a short series of English-language explainer videos for readers who do not read French.

References

Numbered by first appearance in the text. Sources reaching us *through* Hamant rather than from their own pages are marked **[via Hamant]** and are also flagged at the point of use.

[1] Institut Michel Serres — pour les ressources et les biens communs. Statement of purpose and four founding concepts (natural capital, agricultural exception, common health, robustness). <https://institutmichelserres.fr> — research pages at <http://institutmichelserres.ens-lyon.fr>. Honorary doctorate: Theunis, L. (2026). “Pour Olivier Hamant, « la performance nous fragilise, la robustesse nous sauvera ».” *Daily Science* (Belgium), 16 February 2026. <https://dailyscience.be/16/02/2026/pour-olivier-hamant-la-performance-nous-fragilise-la-robustesse-nous-sauvera/>

[2] Institut Momentum (2025). *Antidote au culte de la performance : la robustesse du vivant* — seminar synthesis, 22 November 2025, Académie du Climat, Paris. <https://institutmomentum.org/antidote-au-culte-de-la-performance-la-robustesse-du-vivant> (*Spoken source; attributed as such throughout.*)

[3] Meadows, D. H., Meadows, D. L., Randers, J., & Behrens, W. W. (1972). *The Limits to Growth*. New York: Universe Books.

[4] Daily Science interview, 16 February 2026 — see [1]; and Institut Momentum synthesis — see [2]. Performance as zero stock and just-in-time; Suez 2021, CrowdStrike 2024, Covid as syndemic.

[5] Perrow, C. (1984). *Normal Accidents: Living with High-Risk Technologies*. New York: Basic Books.

[6] Aubin, J.-P. — viability theory. **[via Hamant]** Named by Hamant as the mathematical and economic approach associated with robustness: larobustesse.org, *Incontournables* — “la viabilité où on cherche à être sous-optimal (voir par ex. Jean-Pierre Aubin)”. CC BY-SA 4.0, funded by the Région wallonne.

[7] Hong, L., Dumond, M., Zhu, M., Tsugawa, S., Li, C.-B., Boudaoud, A., Hamant, O., & Roeder, A. H. K. (2018). Heterogeneity and robustness in plant morphogenesis: from cells to organs. *Annual Review of Plant Biology*, 69, 469–495. doi:10.1146/annurev-arplant-042817-040517

[8] Kirchhelle, C., & Hamant, O. (2023). Discretizing the cellular bases of plant morphogenesis: emerging properties from subcellular and noisy patterning. *Current Opinion in Cell Biology*, 81.

[9] Illich, I. (1973). *Tools for Conviviality*. New York: Harper & Row. Counter-productivity as invoked by Hamant in “Learning from Nature” [43].

[10] Institut Momentum synthesis [2] — the weak-link inversion.

[11] Institut Momentum synthesis [2] — robustness as a recipe for viability, not a moral value.

[12] Samzun, C. (2025). Interview with Olivier Hamant, PRISMES / Rumeur Publique, published 18 September 2025 (interview conducted June 2025). <https://www.rumeurpublique.fr/olivier-hamant-les-entreprises-comprennent-de-mieux-en-mieux-le-concept-de-robustesse/> — serendipity as the single granted use; click-workers; *consanguinité*.

[13] PRISMES interview [12] — falling asleep at the wheel; the least performant AIs as the most relevant; catastrophe in gestation.

[14] Daily Science interview [1] — robustness as technophile; local reparability; conviviality; plan A, B, C, D.

[15] Stroh, J. (2026). *The Philosophical Foundations of the Village Project: A Framework for Digital Sovereignty and Pluralist AI Governance*. My Digital Sovereignty Limited, February 2026. <https://agenticgovernance.digital/downloads/philosophical-foundations-village-project.pdf> — Berlin, Alexander, Wittgenstein, Weil and indigenous data sovereignty; Alexander's five principles; the layered constitutional model; Weil's account of attention and how it shapes the platform's handling of grief. Ostrom on the commons is carried in *Distributive Equity Through Structure*, <https://agenticgovernance.digital/whitepapers/distributive-equity.html>

[16] *Governance Architecture* — Tractatus AI Safety Framework. <https://agenticgovernance.digital/architecture.html>

[17] Institut Momentum synthesis [2] — three independent autopilots; conditions rather than turnkey solutions; agility versus adaptability; *durer et transmettre*; Fleury on fragility and vulnerability [via Hamant]; the inversion of the entry constraint.

[18] Structure-Preserving Transformation: correction by annotation, audit logs interpretable across versions, prior governance decisions remaining valid. See [15] and [16].

[19] Weil, S. (1940–41). *L'Iliade ou le poème de la force*. *Cahiers du Sud*, December 1940 and January 1941, published under the pseudonym Émile Novis. English: trans. Mary McCarthy, *Politics*, November 1945; current edition ed. J. P. Holoka, *Simone Weil's The Iliad or the Poem of Force: A Critical Edition* (New York: Peter Lang, 2003).

[20] Institut Momentum synthesis [2] — the performant algorithm that cuts the conversation short, against language with its errors, redundancies and heterogeneity.

[21] Serres, M. (1990). *Le Contrat Naturel*. Paris: François Bourin. Institute self-description: *connaissances situées et impliquées* — see [1].

[22] *Research Timeline* — Tractatus AI Safety Framework. <https://agenticgovernance.digital/timeline.html> — “My own preference settled on Hannah Arendt, who remains the clearest account I know of how values drift — how a thing goes wrong without anyone deciding to do wrong. That is the question the architecture is built to answer.”

[23] Arendt, H. (1971). "Thinking and Moral Considerations: A Lecture." *Social Research*, 38(3), 417-446, at p. 417 — "a curious, quite authentic inability to think"; reprinted in *Responsibility and Judgment* (New York: Schocken, 2003). See also *Eichmann in Jerusalem: A Report on the Banality of Evil* (New York: Viking Press, 1963) and *The Life of the Mind* (New York: Harcourt Brace Jovanovich, 1978) for the banality of evil as thoughtlessness. See also *Stanford Encyclopedia of Philosophy*, "Hannah Arendt." <https://plato.stanford.edu/entries/arendt/>

[24] Stroh, J. (2026). *Sovereignty Without Dominance*. <https://agenticgovernance.digital/papers/sovereignty-without-dominance.html> — "Why now: normalisation, and the closing window"; and the immutable constitutional layer as the structural answer to values eroded by discretion. Published in English, German, French, te reo Māori and Dutch.

[25] *Governance* — Tractatus AI Safety Framework. <https://agenticgovernance.digital/governance.html> — instructions fade, boundaries drift, audit trails become unreliable when agents optimise through training loops. Limitations: see [16].

[26] Arendt, H. (1970). *On Violence*. New York: Harcourt, Brace & World — bureaucracy as rule by Nobody; tyranny without a tyrant.

[27] Stroh, J. (2026). *You Will Not Be Told*. <https://agenticgovernance.digital/you-will-not-be-told> — "Nobody decided this": rule by Nobody applied to platform terms; individual harms have defendants, the drift has none.

[28] Arendt, H. (1982). *Lectures on Kant's Political Philosophy*, ed. R. Beiner. University of Chicago Press (delivered 1970) — the distinction between cognition, thinking and judgment.

[29] Alexander, C. (2002). *The Nature of Order, Book Two: The Process of Creating Life*. Berkeley: Center for Environmental Structure. See also *A Pattern Language* (Oxford University Press, 1977) and *The Timeless Way of Building* (Oxford University Press, 1979).

[30] Stroh, J. (2026). *Executive Brief: Tractatus-Based LLM Architecture for AI Safety*, "The Limits of the Framework", proposition 12.1: "Values cannot be automated, only verified." <https://agenticgovernance.digital/docs-viewer.html?slug=executive-summary-tractatus-inflection-point>

[31] Institut Momentum synthesis [2] — the decider as facilitator; the robust question and the CO₂ example; the death projects of the ultra-performant. The African proverb: Daily Science interview [1].

[32] larobustesse.org — competition, collaboration and cooperation as three distinct terms; and the Institut Momentum synthesis [2]. CC BY-SA 4.0.

[33] PRISMES interview [12] — cooperation between direct competitors in the French construction sector under material shortage.

[34] Ostrom, E. (1990). *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge University Press. The common good as a *faire commun*: Institut Momentum synthesis [2].

[35] Robustness in the local administration's 2030 strategic vision: PRISMES interview [12]. larobustesse.org as an open commons, CC BY-SA 4.0, funded by the Région wallonne, with monthly public sessions. See [32].

[36] Berlin, I. (1969). *Four Essays on Liberty*. Oxford University Press. Note: Berlin's negative/positive liberty distinction is used in this project against the platform's own temptation to optimise, not against communities — see [15].

[37] Ostrom's eight design principles as reproduced by Hamant, with the gloss "the porous membrane": Institut Momentum synthesis [2]. Primary source: Ostrom [34].

[38] Weil, S. (1949). *L'Enracinement : prélude à une déclaration des devoirs envers l'être humain*. Paris: Gallimard, collection Espoir, ed. Albert Camus (written 1943). English: *The Need for Roots* (Routledge Classics); also trans. R. Schwartz & K. Kirkpatrick as *Prelude to a Declaration of Obligations Towards the Human Being* (Penguin, 2023). See also *Waiting for God* (New York: G. P. Putnam's Sons, 1951) for attention.

[39] Taiuru, K. — Kaupapa Māori AI Framework: Te Tiriti-grounded principles for AI consent, Māori data sovereignty and full-chain accountability. <https://www.taiuru.co.nz>. See also Te Mana Raraunga (<https://www.temanaraunga.maori.nz>), the CARE Principles (<https://www.gida-global.org/careprinciples>) and OCAP (<https://fnigc.ca>).

[40] Te Urewera Act 2014 (NZ) — legal personality conferred on Te Urewera.

[41] Te Awa Tupua (Whanganui River Claims Settlement) Act 2017 (NZ), Public Act No 7. Section 12 records the Te Awa Tupua recognition; section 14 declares Te Awa Tupua a legal person with all the rights, powers, duties and liabilities of a legal person, exercised by Te Pou Tupua. <https://www.legislation.govt.nz/act/public/2017/0007/latest/whole.html>

[42] Hamant, O. (2024). *De l'incohérence : philosophie politique de la robustesse*. Paris: Odile Jacob. ISBN 978-2-415-00795-9. See also *La troisième voie du vivant* (Odile Jacob, 2022); *Antidote au culte de la performance : la robustesse du vivant* (Gallimard, Tracts n° 50, 2023); Hamant, O., Charbonnier, O., & Enlart, S. (2025). *L'Entreprise robuste* (Odile Jacob). (No English trade edition at the time of writing.)

[43] Hamant, O. (2021). "Learning from nature." GTI Forum *Interrogating the Anthropocene: Truth and Fallacy?*, Great Transition Initiative, February 2021. <https://greattransition.org/gti-forum/anthropocene-hamant> — Illich's counter-productivity; and the caveat on transferring the biological standpoint to human societies.

[44] Grumbach, S., & Hamant, O. (2018). Digital revolution or Anthropocenic feedback? *The Anthropocene Review*, 5(1), 87–96. doi: 10.1177/2053019617748337. Open access: <https://inria.hal.science/hal-01227303>

[45] Grumbach, S., & Hamant, O. (2020). How humans may co-exist with Earth? The case for suboptimal systems. *Anthropocene*, 30, 100245. doi: 10.1016/j.ancene.2020.100245. Erratum: doi:10.1016/j.ancene.2020.100276

[46] Grumbach, S. (2017). Digital platforms: a new grammar for territories? *Ethics in Progress*, 8(1), 101–116. doi:10.14746/eip.2017.1.7

Appendix A — Sources by kind

Peer-reviewed. [7] [8] [44] [45] [46]

Authored by Hamant in English. [43]

Hamant's books (French only). [42]

Interviews and seminar syntheses — spoken sources, attributed as such at every point of use. [2] [12] and the Daily Science interview at [1]. Also: Sismique podcast ep. 140, <https://www.sismique.fr/episodes/140-performance-turbulences-et-robustesse-olivier-hamant>; Favrot, L. (2022), « La décentralisation apporte de la robustesse », *Mag2Lyon*, <https://www.mag2lyon.fr/la-decentralisation-apporte-de-la-robustesse/>

Referenced through Hamant — not read in the original. [6] Aubin on viability theory; Fleury on fragility and vulnerability, at [17].

This project's own published positions. [15] [16] [22] [24] [25] [27] [30]. Further: *The Approach* (<https://agenticgovernance.digital/approach.html>); *Glossary* (<https://agenticgovernance.digital/glossary.html>, 316 entries, CC BY 4.0, English and te reo Māori); Paper A v4, *Sovereign-Record Architecture for Community-Scale Platforms* (<https://agenticgovernance.digital/papers/sovereign-record-architecture-v4-may-2026.html>); Paper B synopsis, *Situated Language Layers* (<https://agenticgovernance.digital/papers/sovereign-sll-synopsis-may-2026.html>); *Distributive Equity Through Structure* (<https://agenticgovernance.digital/whitepapers/distributive-equity.html>); *Architectural Alignment* (<https://agenticgovernance.digital/architectural-alignment.html>); *Plural Values in Living Organisations* (<https://agenticgovernance.digital/papers/plural-values-living-organisations-ai-may-2026.html>).

Philosophy — primary texts. [19] [23] [26] [28] [29] [36] [38]. Wittgenstein, L. (1921). *Tractatus Logico-Philosophicus*. Stanford Encyclopedia of Philosophy entries on Value Pluralism, Isaiah Berlin, Simone Weil, Hannah Arendt and Wittgenstein: <https://plato.stanford.edu>

Indigenous data sovereignty and legal standing. [39] [40] [41]

Referenced directly, other. [3] [5] [9] [21] [34]

Appendix B — Two notes on Arendt in this corpus

She appears in two distinct roles, and the citations should not be run together.

PLACEHOLDER_SPLIT In *Plural Values in Living Organisations* she is cited via *The Human Condition* (1958) for **plurality** — the human condition of being many, each life its own perspective, none reducible to any other — alongside Berlin and Raz. In *Sovereignty Without Dominance* [24] and in this essay she is cited via the 1971 lecture “Thinking and Moral Considerations”, with *Eichmann in Jerusalem* and *The Life of the Mind* alongside it [23], for **normalisation and thoughtlessness**. Two arguments, two works, one thinker. A reader following one should not be left assuming the other.

On the standing of the borrowing. *You Will Not Be Told* [27] introduces “rule by Nobody” with the qualifier *whatever one makes of the borrowing*, and then supplies a test that does not depend on the philosophy: when the terms change under you, who do you ring? This essay keeps that order. The concept earns its place by the test, not by the name attached to it.

Read on

[The argument as a short deck](#)

[Glossary — the terms, defined plainly](#)

[Questions & answers — the hard ones, answered](#)

[The Village](#) — the infrastructure this argues about, built