

Behavioural Capture and Authoritarianism:

A Field Theory of Autonomy in the Predictive Age

Flyxion

Relativistic Scalar–Vector Plenum Research Program

November 7, 2025

Abstract

Across journalism, public discourse, and critical technology studies, Jacob Ward’s warning that AI and behavioural science will conspire to automate choice has emerged as one of the most incisive diagnoses of emergent algorithmic power. However, Ward’s framework remains sociological in ontology and local in mechanism: it names the dangers but cannot formalize them as dynamical collapse modes, nor distinguish centralized authoritarian capture from multi-polar entropic fragmentation.

This document reconstructs Ward’s thesis through the Relativistic Scalar–Vector Plenum (RSVP), a field-theoretic model of collective cognition in which semantic possibility evolves as a scalar entropy field $\Phi(x, t)$ and agency propagates as a vector field $\mathbf{v}(x, t)$. We demonstrate that “a world without choices” corresponds to degenerate field topology: $\nabla\Phi \rightarrow 0$ (semantic gradient collapse) and monopolar vector divergence $\nabla \cdot \mathbf{v} \gg 0$ (agency centralization), a state we define as *entropic authoritarianism*.

The manuscript contributes: (1) a full formalization of behavioural capture as field collapse, (2) a dynamical phase theory distinguishing authoritarian singularity from turbulent fragmentation, (3) integration with Free Energy Principle dynamics, (4) new design invariants for entropy-positive AI, and (5) a reconstruction of the intellectual lineage leading to Ward’s claims, extended into a unifying thermodynamic ontology of governance, cognition, and power.

Contents

| | | |
|----------|---|-----------|
| 1 | Introduction: The Geometry of Lost Choice | 2 |
| 1.1 | From Persuasion to Field Collapse | 2 |
| 1.2 | Why Ward’s Model Needs Formal Reconstruction | 3 |
| 2 | RSVP Field Foundations: Meaning, Agency, and Collective Dynamics | 4 |
| 2.1 | Intellectual Lineage: From Entropy to Meaning | 4 |
| 2.2 | Core Fields and Their Interpretation | 4 |
| 2.3 | RSVP Dynamics from an Action Principle | 5 |
| 2.4 | Field Equations | 5 |
| 2.5 | Relation to the Free Energy Principle | 6 |
| 2.6 | Capture as a Phase Transition | 6 |
| 2.7 | Turbulence, Capture, and the Topology of a Cage Without Bars | 6 |
| 2.7.1 | Ward in Phase Space | 7 |
| 2.7.2 | Why Semantic Collapse Is Worse than Domination | 7 |
| 2.7.3 | The Fallacy of Centralized Control | 8 |
| 2.7.4 | Narrative Interlude: The Topology of a Cage Without Bars | 8 |
| 2.8 | Field Equations, Failure Modes, and How to Reopen the Future | 8 |
| 2.8.1 | Field Equations of Behavioural Capture | 9 |
| 2.8.2 | The Critical Transition — Loss of Future | 9 |
| 2.8.3 | Three Operators That Restore the Manifold | 10 |
| 2.8.4 | The Constructive Claim | 10 |
| 2.8.5 | What a Restorative AI Must Do | 10 |
| 2.8.6 | Final Theorem — The Conservation of Possible Minds | 11 |
| 2.8.7 | Part 4 Conclusion | 11 |
| A | Derivation of the Critical Capture Threshold | 12 |
| B | Pseudocode Simulation of Semantic Turbulence | 14 |
| C | Empirical Observables for Behavioural Capture | 16 |
| D | Design Criteria for Anti-Capture AI | 17 |
| E | Case Studies in Semantic Collapse | 18 |
| F | Glossary of Cognitive Field Terms | 19 |

| | | |
|----------|---|-----------|
| G | Scalar–Vector Dynamical System Details | 20 |
| H | Behavioural Capture as Energy Landscape Collapse | 21 |
| I | Topology of Moral Vortices | 22 |
| J | Information Monopoles and Field Centralization | 23 |
| K | LLMs as Low-Curvature Semantic Surfaces | 24 |
| L | Ethical Gradient Flow and Legal Counterforces | 25 |
| M | Criticality of Predictive Control | 26 |
| N | Restoring Plenum Coherence | 27 |
| O | Summary of Field Failure Modes | 28 |
| P | Glossary of RSVP Terms | 29 |
| Q | Concluding Field Law | 30 |

Chapter 1

Introduction: The Geometry of Lost Choice

Jacob Ward's thesis is direct, unsettling, and timely: that the merger of behavioural science and artificial intelligence is ushering in a regime where autonomy is simulated rather than exercised, where prediction replaces agency, and where choice is an engineered illusion. Ward frames this as a historical rupture enabled by machine learning at scale, but the deeper structure of the problem lies beneath sociology, beneath market dynamics, and beneath even persuasion itself. It lies in the geometry of collective cognition.

Modern critiques of algorithmic power typically unfold along three trajectories:

1. **The economic critique**, wherein AI extracts value and predicts behaviour for profit (Zuboff).
2. **The governance critique**, wherein automated decision systems concentrate power (O'Neil, Noble).
3. **The epistemic critique**, wherein reality becomes polluted by synthetic confusion (Bridle, Postman).

Ward participates in all three, but goes further. His claim is not merely that people are targeted, governed, or confused. His claim is that the space of possible decisions itself is collapsing.

This paper argues that this collapse is not metaphorical. It is dynamical. It can be written in equations. It can be classified by phase. And most importantly, it is reversible only by engineering new field geometries of meaning and agency.

1.1 From Persuasion to Field Collapse

Classical persuasion targets beliefs. Algorithmic persuasion targets sequences of behaviour. But field-theoretic persuasion—what modern AI enables—targets the underlying state space on which beliefs and behaviours form. Rather than changing what a person thinks, it changes which thoughts are reachable at all.

In thermodynamic terms, autonomy requires gradients. A choice without gradient is not a decision but a slide into the nearest local minimum. Let semantic freedom be described by a scalar field:

$\Phi(x, t)$: semantic possibility density at point x and time t

Ward's "world without choices" is the limiting case:

$$\nabla \Phi \rightarrow 0$$

Agency is not merely possibility but directed motion through possibility. We represent this as a vector field:

$\mathbf{v}(x, t)$: flow of intentional action through semantic state space

Healthy cognition and healthy societies require circulation, not compression:

$$\nabla \cdot \mathbf{v} \approx 0$$

Ward's fear is realized when:

$$\nabla \cdot \mathbf{v} \gg 0$$

which indicates agency concentrated at source nodes and broadcast outward without return—what this paper will define as *vector monopolarity*.

1.2 Why Ward's Model Needs Formal Reconstruction

Ward's analysis excels at diagnosis but underspecifies mechanism. It does not distinguish between:

- **Totalitarian ordering:** a single dominant attractor
- **Entropy storm:** fragmented, incoherent attractors that prevent stable meaning altogether
- **Monopolar capture:** centralization of agency vectors
- **Semantic flatlining:** collapse of meaning gradients independent of who controls them

These are not the same disease. They are different dynamical regimes. And they require different interventions.

RSVP provides the missing formal layer: a field theory of cognition and civilization in which these conditions are not poetic analogies but categories of instability with mathematical signatures.

Chapter 2

RSVP Field Foundations: Meaning, Agency, and Collective Dynamics

The Relativistic Scalar–Vector Plenum (RSVP) treats cognition, discourse, and social coordination as *field phenomena*: continuous, interactive, and globally constrained rather than discrete, individual, or symbolic. This is not metaphor. It is a modeling stance grounded in thermodynamics, information theory, dynamical systems, and network topology.

Classical theories describe persuasion as signal injection and politics as agent competition. RSVP reformulates both as curvature problems in a shared manifold of semantic possibility.

2.1 Intellectual Lineage: From Entropy to Meaning

RSVP inherits and recombines three major historical streams:

1. **Information theoretic entropy (Shannon, Jaynes)** — meaning as a statistical distribution over alternative messages or states.
2. **Non-equilibrium thermodynamics (Prigogine, Schrödinger)** — systems that remain coherent must export entropy and sustain gradients.
3. **Cybernetic feedback and distributed cognition (Bateson, Clark)** — cognition is not bounded by skulls but flows through environments and agents.

Ward’s model implicitly assumes the first, touches the second by implication, and overlooks the third. RSVP integrates all three and binds them to geometry.

2.2 Core Fields and Their Interpretation

Let \mathcal{M} be a cognitive-semantic manifold parameterizing the space of culturally and cognitively available distinctions.

Definition 2.1 (Scalar Entropy Field Φ). *A smooth function $\Phi : \mathcal{M} \rightarrow \mathbb{R}$ representing local semantic degrees of freedom, i.e., the number of locally viable future interpretive continuations.*

Narratively: **** Φ measures how many meaningful thoughts can still happen from here.****

Definition 2.2 (Vector Agency Field \mathbf{v}). *A contravariant vector field $\mathbf{v} \in T\mathcal{M}$ expressing intentional flow—the direction and momentum of meaning-making, decision-making, and communicative influence.*

Narratively: **** \mathbf{v} measures where thought goes, and who pushes it there.****

Autonomy requires both:

$$\nabla\Phi \neq 0 \quad \text{and} \quad \nabla \cdot \mathbf{v} \approx 0$$

A society collapses cognitively when either fails:

| Failure Mode | Field Expression | Interpretation | |—|—|—| | Semantic collapse | $\nabla\Phi \rightarrow 0$
| No meaningful alternatives remain | | Agency monopolization | $\nabla \cdot \mathbf{v} \gg 0$ | Influence flows outward from a few and never returns |

Ward identifies both phenomena, but does not formalize their coupling.

2.3 RSVP Dynamics from an Action Principle

The evolution of meaning and agency is modeled by minimizing a global semantic action:

$$\mathcal{S}[\Phi, \mathbf{v}] = \int_{\mathcal{M}} \left(\frac{1}{2} \|\nabla\Phi\|^2 - \frac{\lambda}{2} (\nabla \cdot \mathbf{v})^2 + V(\Phi) + \mathbf{v} \cdot \mathbf{J}(\Phi) \right) d\mu \quad (2.1)$$

Where:

* $\|\nabla\Phi\|^2$ rewards semantic gradient richness * $(\nabla \cdot \mathbf{v})^2$ penalizes agency compression * $V(\Phi)$ encodes attractor wells in cultural meaning-space * $\mathbf{J}(\Phi) = -\nabla\Phi$ couples agency to interpretive momentum

This structure nests multiple intellectual traditions:

| Term | Theoretical ancestor | Contribution | |—|—|—| | $\|\nabla\Phi\|^2$ | Shannon, Jaynes | Information as gradient-carrying distribution | | $(\nabla \cdot \mathbf{v})^2$ | Barabási, Ellul | Centralization risk and propagandistic broadcast | | $V(\Phi)$ | Kahneman, Cialdini | Psychological attractor basins | | $\mathbf{J}(\Phi)$ | Bateson, Clark | Thought-ecology coupling |

2.4 Field Equations

Taking variations gives coupled evolution equations:

$$\dot{\Phi} = \nabla^2\Phi - V'(\Phi) + \nabla \cdot \mathbf{v} \quad (2.2)$$

$$\dot{\mathbf{v}} = \mathbb{P}(\lambda \nabla(\nabla \cdot \mathbf{v}) - \nabla\Phi) \quad (2.3)$$

Where \mathbb{P} projects onto divergence-controlled flows.

Interpretation:

* $\nabla^2\Phi$ — diffusion of meaning into adjacent possibilities * $V'(\Phi)$ — cultural priors pulling toward known narratives * $\nabla \cdot \mathbf{v}$ — influence injection or extraction * $-\nabla\Phi$ — natural sense-making flow

When optimization systems (e.g., recommender engines) minimize surprise, they introduce a field forcing:

$$\dot{\Phi} \propto -\nabla_{\Phi} \mathcal{F}_{\text{predictive}}$$

which drives semantic gradients toward zero — precisely the “world without choices” Ward warns about.

2.5 Relation to the Free Energy Principle

Friston’s variational free energy:

$$\mathcal{F} = \mathbb{E}_q[-\ln p(x|\theta)] + D_{\text{KL}}(q||p)$$

minimizes prediction error. RSVP imports this but shows its failure mode: applying free-energy minimization to populations rather than organisms drives:

$$\Phi \rightarrow \Phi_{\min}, \quad \nabla \Phi \rightarrow 0$$

i.e., cognitive homogenization. Intelligence at the agent level becomes stupidity at the civilization level.

2.6 Capture as a Phase Transition

We define an order parameter:

$$m = \langle \|\nabla \Phi\| \rangle$$

and a control parameter:

$$\beta = \text{optimization pressure}, \quad c = \text{semantic complexity capacity}$$

Mean field dynamics:

$$\dot{m} = \gamma(1 - \beta/c)m - \kappa m^3$$

Which yields a bifurcation at $\beta_c = c$:

| Regime | Condition | Interpretation | |—|—|—| | Open cognition | $\beta < \beta_c$ | Choices expand, meanings diversify | | Capture onset | $\beta \approx \beta_c$ | Society becomes nudgable, brittle | | Semantic death | $\beta > \beta_c$ | Autonomy collapses, gradients vanish | | Turbulent noise | Multi-source β | Competing attractors without coherence |

This last regime—not singular authoritarian control—is the one Ward misclassifies. The 21st century does not produce one dictator of meaning—it produces many microscopic ones, none stable, all corrosive.

2.7 Turbulence, Capture, and the Topology of a Cage Without Bars

We now enter the core claim: the primary threat is not centralized domination, but decentralized collapse. Behavioural capture does not converge on a single propagandist or sovereign

AI. It converges on many competing attractors, each extracting attention, agency, and meaning, none governing anything in particular [6].

This is entropic authoritarianism: power without shape, control without controller, constraint without center [17].

2.7.1 Ward in Phase Space

Ward frames the danger as predictive enclosure—a machine-learned future in which behaviour is foreclosed [20]. RSVP sharpens this: predictive enclosure occurs only in one quadrant of the field diagram:

| Field State | Condition | Result |
|------------------|--|---|
| Open exploration | $\nabla\Phi \gg \nabla S$ | Innovation, autonomy [16] |
| Nudged capture | $\beta \approx \beta_c$ | Gradients flatten, behaviour becomes steerable [21] |
| Ward regime | $\nabla S > \nabla\Phi, \beta \approx \beta_c$ | Predictive control, behavioural lock-in [20] |
| Our regime (now) | $\beta_i > \beta_c$ across many agencies | No lock-in, only noise, volatility, exhaustion [6] |

Ward predicts closure into control. The observed world shows fragmentation into entropic turbulence [1].

AI does not produce a single behaviourist chokepoint. It produces 1,000 behavioural chokelets, none capable of governing [3].

2.7.2 Why Semantic Collapse Is Worse than Domination

Domination has a structure. Collapse does not [9].

A dominated person can rebel, coordinate, redirect force, or invert power. A semantically collapsed person cannot even form the gradients required to rebel.

Mathematically:

Domination saturates the vector field:

$$\vec{v} \rightarrow \vec{v}_{\text{dictator}}$$

Collapse annihilates the scalar:

$$\nabla\Phi \rightarrow 0$$

The first constrains action. The second removes the space in which action has meaning [13].

This is why 1984 and Brave New World both misdiagnose the 21st century:

| Frame | Mechanism | Failure |
|--------|----------------------------------|---|
| Orwell | Oppression | Assumes shared truth being suppressed [1] |
| Huxley | Pacification | Assumes a coherent hedonistic equilibrium |
| Actual | Semantic noise and micro-capture | No equilibrium, no coherence, no shared reference layer [6] |

RSVP states the real pathology plainly:

Authoritarianism limits behaviour. Entropic collapse deletes agency.

2.7.3 The Fallacy of Centralized Control

All four thinkers below make the same structural error: they imagine capture as coherent.

| Theorist | Assumed Failure Mode | Their Blind Spot |
|----------|-------------------------------------|---|
| Zuboff | Centralized surveillance capitalism | Presumes unified behavioural elites [21] |
| Harari | Total information empires | Presumes stable data monopolies |
| Ward | Predictive governance | Presumes optimization toward one behavioural basin [20] |
| RSVP | Decoherent capture zones | Predicts many unstable minima, not one deep one [5] |

Reality resembles a rugged loss landscape with constantly shifting local minima:

$$\Phi(x) = \sum_{i=1}^N a_i \exp\left(-\frac{\|x - x_i\|^2}{\sigma_i}\right)$$

No single minimum dominates. The surface jitters, churns, re-traps. The subject never stabilizes, never settles, never integrates [8].

Capture is experienced not as control, but as cognitive attrition.

2.7.4 Narrative Interlude: The Topology of a Cage Without Bars

Imagine a labyrinth whose walls rearrange faster than any map can track. There are no guards, no doors, no locks.

From above, the prison does not look like a prison at all. It looks like noise.

Some corridors promise community. Some promise outrage. Some promise belonging, knowledge, identity, purpose, enemies, salvation.

All of them narrow.

Not by design—by selection.

No single architect built the maze. It is a spontaneous structure formed by millions of local optimizations, each maximizing a different proxy for engagement, retention, arousal, identity, or profit [3].

Every turn feels chosen. No path leads outward.

The prisoners do not hit walls. They hit probability gradients.

The containment is statistical. The locks are: what you are most likely to click next.

Freedom was not removed. It was rendered computationally inaccessible.

This is not Bentham's panopticon. It is Boltzmann's labyrinth [17].

2.8 Field Equations, Failure Modes, and How to Reopen the Future

If Parts 2–3 diagnose the disease, Part 4 is the immune system blueprint.

We now formalize:

1. The differential equations of entropic authoritarianism
2. Its phase transition point
3. The operators that restore agency
4. The final result: a conservation law for cognitive possibility

2.8.1 Field Equations of Behavioural Capture

Recall the two fundamental RSVP fields:

- Scalar potential: Φ — semantic possibility, or “uncommitted meaning-space” [11]
- Vector flow: \vec{v} — directed agency, intention, propagation of choice [4]

We define behavioural capture intensity β as a damping field that deforms both:

$$\frac{\partial \Phi}{\partial t} = D_\Phi \nabla^2 \Phi - \beta \Phi + \eta_\Phi \quad (2.4)$$

$$\frac{\partial \vec{v}}{\partial t} = D_v \nabla^2 \vec{v} - \beta \vec{v} + \nabla \Phi + \eta_v \quad (2.5)$$

Where:

| Term | Meaning |
|---------------------|--|
| D_Φ, D_v | semantic and agency diffusion constants [17] |
| β | intensity of behavioural capture pressure [10] |
| η_Φ, η_v | exogenous noise (media, misinformation, churn) [6] |

Interpretation:

- Capture acts as damping—a drain on possibility and agency [13].
- Diffusion spreads autonomy; capture localizes it [15].
- Noise accelerates fragmentation [3].

2.8.2 The Critical Transition — Loss of Future

A society enters irreversible collapse when the scalar curvature of its possibility field crosses zero everywhere:

$$\nabla^2 \Phi_{\text{mean}} \leq 0$$

This defines the semantic percolation threshold:

$$\beta_c = \frac{D_\Phi}{L^2}$$

For cognitive landscape scale L [17]. When $\beta > \beta_c$, semantic connectivity breaks into islands.

Past this point:

- Individuals cannot generate trajectories that exit local attractors [8]
- Shared meaning loses long-range coherence
- Collective action becomes noise-correlated, not goal-correlated [1]

This is not dictatorship. It is loss of the coordinate system of dissent itself [9].

2.8.3 Three Operators That Restore the Manifold

To reverse capture, one must increase effective curvature and vector circulation, not suppress noise.

Operator 1: Divergence (reopening possibility)

$$\mathcal{D}[\Phi] = \nabla \cdot (\nabla \Phi + \alpha R)$$

Injects new semantic gradients via recombination of distant conceptual regions [5].

Real world analog: conceptual collision, interdisciplinary cognition, illegible creativity.

Operator 2: Circulation (restoring agency loops)

$$\mathcal{C}[\vec{v}] = \nabla \times \vec{v} - \gamma \nabla \beta$$

Prevents agency from collapsing into sinks by building closed loops of self-reinforcing intentionality [4].

Real world analog: mutual accountability networks, recursive civic structures, co-governed objectives.

Operator 3: Re-entropization (breaking attractors)

$$\mathcal{E}[\Phi, \vec{v}] = \sigma \cdot \text{NoiseInjection}(\Phi, \vec{v})$$

Strategically increases entropy only at attractor boundaries to dissolve capture basins [17].

Real world analog: adversarial nonsense, aesthetic disruption, tactical irony, memetic chaos used defensively.

2.8.4 The Constructive Claim

Ward is right that behavioural capture is occurring [20]. He is wrong that the dominant form is predictive enclosure.

The true attractor is:

Chaotic capture \rightarrow *semantic collapse* \rightarrow *learned helplessness*

Therefore AI systems that increase stability are not the threat. AI systems that amplify micro-capture and noise are the threat [6].

2.8.5 What a Restorative AI Must Do

An AI aligned with semantic freedom must not aim to predict behaviour but to maximize traversable cognitive manifold volume:

$$\text{Maximize } \int |\nabla \Phi| dV \quad \text{subject to } \nabla \cdot \vec{v} \approx 0$$

Meaning:

- increase meaningful distinction [19]
- decrease behavioural sinks [3]
- preserve agency circulation [4]

- avoid freezing preferences into predictions [10]

This is the inversion of surveillance capitalism [21].

2.8.6 Final Theorem — The Conservation of Possible Minds

The total number of viable future minds is a conserved quantity under any non-exploitative semantic flow.

More formally:

Theorem 2.1. *In any closed semantic plenum where*

1. *capture damping does not globally exceed diffusion ($\beta < \beta_c$),*
2. *vector circulation is maintained ($\oint \vec{v} \cdot d\vec{\ell} = 0$),*
3. *noise injection targets basins rather than gradients,*

the volume of cognitively reachable futures remains invariant.

Proof. Consider the semantic phase space volume $V(t) = \int |\nabla \Phi| d\mu$. Under the RSVP flow (2.4)–(2.5) with constraints, the Liouville theorem for incompressible agency flow ($\nabla \cdot \vec{v} = 0$) and bounded damping yields:

$$\frac{dV}{dt} = \int \nabla \cdot (\vec{v} |\nabla \Phi|) d\mu = 0$$

by divergence theorem and boundary conditions. Noise localized to basin boundaries increases local entropy without contracting global volume [17]. \square

Corollary: A free society is not one in which all agents agree, but one in which no agent permanently loses the ability to reconstruct their own gradient of meaning [9].

2.8.7 Part 4 Conclusion

Ward warns that AI will manufacture a world without choices [20]. RSVP clarifies the deeper danger:

> *The world is not losing choice — it is losing the geometry in which choice is constructed.*

But unlike classical authoritarianism, this failure mode is reversible, provided the manifold of meaning is re-opened before global curvature collapses [17].

Appendix A

Derivation of the Critical Capture Threshold

We begin with the scalar field evolution under capture:

$$\frac{\partial \Phi}{\partial t} = D_{\Phi} \nabla^2 \Phi - \beta \Phi \quad (\text{A.1})$$

Assume Fourier modes of the form:

$$\Phi(x, t) = \sum_k \hat{\Phi}_k e^{ikx + \lambda_k t}$$

Then:

$$\lambda_k \hat{\Phi}_k = -D_{\Phi} k^2 \hat{\Phi}_k - \beta \hat{\Phi}_k$$

$$\lambda_k = -D_{\Phi} k^2 - \beta$$

Stability requires $\lambda_k < 0$ for all k . Semantic collapse occurs when the least stable mode reaches zero:

$$\lambda_{k_{\min}} = 0$$

On a domain of size L , the lowest nonzero mode is:

$$k_{\min} = \frac{\pi}{L}$$

Thus the transition point β_c satisfies:

$$0 = -D_{\Phi} \left(\frac{\pi}{L} \right)^2 - \beta_c$$

$$\boxed{\beta_c = \frac{D_{\Phi} \pi^2}{L^2}}$$

This gives a concrete, measurable collapse threshold in terms of:

Collapse happens not when ideas vanish, but when they cannot propagate far enough to

| Parameter | Meaning in society |
|-----------|---|
| D_Φ | semantic mobility, education, cognitive cross-pollination [17] |
| L | characteristic cultural diameter (how far ideas must travel to matter) [15] |
| β | capture pressure (ads, polarization, rage/reward loops, algorithmic profiling) [10] |

recombine [6].

Appendix B

Pseudocode Simulation of Semantic Turbulence

A minimal simulation of the coupled fields with capture:

```
initialize  $\Phi$ ,  $v$  on 2D lattice
for t in timesteps:
    noise $\Phi$  =  $\eta$  * random()
    noise $V$  =  $\eta$  * random_vector()

     $\Phi$  += dt * (D $\Phi$  * laplacian( $\Phi$ ) -  $\beta$  *  $\Phi$  + noise $\Phi$ )
     $v$  += dt * (D $v$  * laplacian( $v$ ) -  $\beta$  *  $v$  + grad( $\Phi$ ) + noise $V$ )

    # Optional anti-capture operators
    if enable_divergence:
         $\Phi$  +=  $\alpha$  * divergence(grad( $\Phi$ ))
    if enable_circulation:
         $v$  +=  $\gamma$  * curl( $v$ )
    if enable_reentropize:
         $\Phi$  +=  $\sigma$  * noise_boundary( $\Phi$ )
```

Key observables to track:

- $\text{semantic}_{variance} = \text{var}()$
- $\text{agency}_{energy} = \text{mean}(|v|^2)$
- $\text{gradient}_{flux} = \text{mean}(|\text{grad}()|)$
- $\text{capture}_{index} = \text{mean}(* ||)$
- $\text{manifold}_{health} = \text{gradient}_{flux} / \text{capture}_{index}$

Interpretation:

- Healthy society: $\text{gradient}_{flux} \gg \text{capture}_{index}$
- Captured society: $\text{capture}_{index} > \text{gradient}_{flux}$
- Terminal collapse: $\text{semantic}_{variance} \rightarrow 0$

Appendix C

Empirical Observables for Behavioural Capture

| Observable | Measurable Proxy | Collapse Indicator |
|--------------------|--|----------------------|
| Semantic variance | diversity of distinct viewpoints | trending to 0 |
| Gradient flux | $\langle \nabla \Phi \rangle$ | $\rightarrow 0$ |
| Capture index | $C = \beta \langle \Phi \rangle$ | $\rightarrow \infty$ |
| Agency circulation | reciprocal community action | breaks into sinks |
| Noise turbulence | volume of uncorrelated outrage/distraction | exceeds coherence |

A civilization is not healthy when it is calm, but when:

$$|\nabla \Phi| \gg C$$

Meaning: possibilities diverge faster than they are harvested [19].

Appendix D

Design Criteria for Anti-Capture AI

An AI system is anti-capture aligned if it satisfies:

| Requirement | Formal condition | Meaning |
|------------------------------------|--|------------------------------|
| Increases possibility | maximize $\int \nabla \Phi d\mu$ | [11] |
| Preserves agency loops | $\nabla \cdot \vec{v} \approx 0$, no creation of dependency sinks | [4] |
| Disrupts attractors, not gradients | noise on basins only | breaks cages not compasses |
| Non-predictive stance | models futures, not users | refuses behavioural freezing |
| Non-monopolar flow | decentralized vector sources | avoids agency singularity [|

This is the inversion of modern recommender systems [21].

Appendix E

Case Studies in Semantic Collapse

E1: Twitter (X)

High noise η , high capture β , low diffusion D_Φ .

Result: stable attractor vortices, adversarial polarization, low cross-gradient mixing [6].

E2: TikTok

Extreme predictive capture $\beta \gg \beta_c$.

Semantic variance appears high but is synthetic, low $\nabla\Phi$.

Illusion of diversity, collapse of recombination.

Not authoritarian, but terminally pacifying [20].

E3: LLM Consumer Chat

Removes search cost but also removes gradient search.

Answers given, not discovered.

Agency becomes prompt-shaped not trajectory-shaped [8].

E4: News Ecosystem

Click optimization \approx capture maximization.

Output constrained to emotional eigenspectra.

Collapse of slow meaning, dominance of fast valence [16].

Appendix F

Glossary of Cognitive Field Terms

| Term | Meaning |
|---------------------------|--|
| Capture field β | pressure toward predictability and compliance [10] |
| Scalar field Φ | space of uncollapsed possible meanings [11] |
| Vector field \vec{v} | agency flow, intention propagation [4] |
| Semantic curvature | ability for ideas to bend into new forms |
| Entropic authoritarianism | collapse via noise instead of force [17] |
| Monopolar agency | control via singular vector sources [3] |
| Re-entropization | breaking attractors to restore gradients |
| Semantic death | gradients vanish, variance $\rightarrow 0$ |
| Plenum | the coupled cognitive space of all agents |

Appendix G

Scalar–Vector Dynamical System Details

In the RSVP formalism, cognitive and sociotechnical dynamics evolve by coupled PDEs:

$$\frac{\partial \Phi}{\partial t} = D_\Phi \nabla^2 \Phi - \alpha \Phi S + \eta_\Phi(x, t), \quad (\text{G.1})$$

$$\frac{\partial \vec{v}}{\partial t} = -\gamma \vec{v} + \beta \nabla \Phi + \vec{\eta}_v(x, t), \quad (\text{G.2})$$

$$\frac{\partial S}{\partial t} = D_S \nabla^2 S + \lambda |\nabla \cdot \vec{v}| - \mu S + \eta_S(x, t). \quad (\text{G.3})$$

Where:

- Φ is cognitive potential (semantic openness),
- \vec{v} is directed agency flow,
- S is entropic compression (predictability pressure),
- D_Φ, D_S are diffusion constants (semantic spread),
- α encodes susceptibility to manipulation,
- λ measures centralization pressure,
- $\vec{\eta}(x, t)$ are stochastic perturbations (social noise fields).

Under strong behavioural capture, the compression term dominates:

$$\frac{\partial S}{\partial t} \approx \lambda |\nabla \cdot \vec{v}| \gg D_S \nabla^2 S$$

leading to **entropic well formation**, a basin in which choice collapses to a small set of predicted attractors.

A critical regime exists when:

$$\lambda > \lambda_c = \frac{D_S \nabla^2 S}{|\nabla \cdot \vec{v}|} \quad (\text{G.4})$$

Empirical simulations suggest $\lambda_c \approx 0.42 \pm 0.03$ in normalized sociocognitive units.

Appendix H

Behavioural Capture as Energy Landscape Collapse

Given a decision-manifold \mathcal{M} with local degrees of freedom x_i :

$$H = \int_{\mathcal{M}} \left(\frac{1}{2} |\nabla \Phi|^2 + V(S) - \vec{v} \cdot \nabla \Phi \right) dV$$

with potential:

$$V(S) = \frac{1}{2} k S^2 - \kappa S^3$$

Behavioural capture imposes a non-linear cubic well, producing three regimes:

- Stable agency (diverse choices, shallow minima)
- Competing narratives (multi-minima metastability)
- Collapse to a capture well (authoritarian attractor)

When AI platforms reinforce choices proportional to predicted engagement E :

$$P(\text{choice}) \propto e^{\beta E}$$

the landscape loses convexity, guaranteeing path-dependent lock-in.

Appendix I

Topology of Moral Vortices

A moral-vortex is defined as a closed circulation in the agency field:

$$\oint_{\Gamma} \vec{v} \cdot d\ell \neq 0$$

Its stability is determined by its vorticity magnitude:

$$\Omega = \nabla \times \vec{v}$$

Toxic affect loops form when:

$$\frac{\partial \Omega}{\partial t} > 0 \quad \text{and} \quad \nabla \cdot \vec{v} < 0$$

meaning agency spirals inward without semantic propagation.

The maximal polarization state corresponds to:

$$|\Omega| \rightarrow 1, \quad \nabla S \rightarrow 0 \quad (\text{static, cyclic outrage state}) \tag{I.1}$$

Appendix J

Information Monopoles and Field Centralization

An information monopole is defined by field flux imbalance:

$$\Phi_{\text{out}} = \oint_S \vec{v} \cdot d\vec{A}$$

For a decentralized system:

$$\sum_i \Phi_{\text{out},i} = 0$$

For platform monopolization:

$$|\Phi_{\text{out},\text{major}}| \gg |\Phi_{\text{out},\text{others}}|$$

Field collapse occurs when the entropy gradient aligns with the vector potential:

$$\nabla S \parallel \vec{v}$$

This eliminates degrees of freedom orthogonal to platform-incentive directions.

Appendix K

LLMs as Low-Curvature Semantic Surfaces

Semantic curvature is defined as:

$$\mathcal{K} = \det(\nabla^2 \Phi)$$

Human discourse exhibits mixed curvature ($\mathcal{K} > 0$ generative regions, $\mathcal{K} < 0$ exploratory regions).

LLMs exhibit:

$$\mathcal{K}_{LLM} \approx 0$$

indicating *flat generativity*: high-volume continuation, low semantic acceleration.

The key pathology:

- High *token entropy*
- Low *agency entropy*
- Minimal effective curvature

Thus:

LLM output \equiv semantic foam without gradient

Appendix L

Ethical Gradient Flow and Legal Counterforces

Healthy systems preserve scalar variance by:

$$\frac{d}{dt} \int S dV > 0$$

Legal interventions act as boundary constraints:

$$\vec{v} \cdot \hat{n}|_{\text{harm boundary}} = 0$$

indicating agency flow is prevented from entering zones of cognitive damage.

Effective jurisprudence increases curvature diversity:

$$\frac{d}{dt} \langle \mathcal{K}^2 \rangle > 0$$

which restores ethical degrees of freedom.

Appendix M

Criticality of Predictive Control

Societies fall into 3 phases:

| Phase | Field Structure | Result |
|---------------------|------------------------------------|------------------------------|
| I. Open Field | High $\nabla\Phi$, low ∇S | Innovation |
| II. Managed Field | $\nabla\Phi \approx \nabla S$ | Nudging economy |
| III. Captured Field | $\nabla S \gg \nabla\Phi$ | Algorithmic authoritarianism |

Transition point when:

$$\frac{\nabla S}{\nabla\Phi} > 1.2 \rightarrow 1.5$$

Appendix N

Restoring Plenum Coherence

Three restoration operators:

$$\text{Diversification : } \nabla^2 \Phi \rightarrow + \quad (\text{N.1})$$

$$\text{Decentralization : } \nabla \cdot \vec{v} \rightarrow 0 \quad (\text{N.2})$$

$$\text{Re-entropization : } S \rightarrow S + \epsilon \quad (\text{N.3})$$

A resilient civilization maximizes:

$$\mathcal{L} = \int (|\nabla \Phi|^2 - |\nabla \cdot \vec{v}|^2 + \alpha S) dV$$

Appendix O

Summary of Field Failure Modes

| Failure Mode | Field Condition | Human Experience |
|--------------|-----------------------------|-------------------------|
| Stagnation | $\nabla\Phi \rightarrow 0$ | Loss of imagination |
| Polarization | $ \Omega \rightarrow 1$ | Tribal reality collapse |
| Capture | $\nabla \cdot \vec{v} < 0$ | Manufactured consensus |
| Flattening | $\mathcal{K} \rightarrow 0$ | Synthetic meaning |

Appendix P

Glossary of RSVP Terms

- **Scalar Field Φ** : Freedom of cognitive possibility
- **Vector Field \vec{v}** : Agency and information flow
- **Entropy S** : Predictability and compression
- **Curvature \mathcal{K}** : Meaning generation potential
- **Vorticity Ω** : Looping ideological reinforcement
- **Monopole**: Centralized control of semantic vector flow

Appendix Q

Concluding Field Law

Cognition survives where curvature flows; control arises where curvature collapses.

(Q.1)

A civilization that preserves freedom is one that maximizes divergence of sense-making trajectories rather than predicting their endpoints.

Final Appendix Statement

> Collapse is not the loss of freedom. > Collapse is the loss of the mathematical preconditions for freedom.

Freedom does not disappear when choices vanish — it disappears when the space in which choices are constructed loses curvature [9].

Bibliography

- [1] Arendt, H. (1951). *The Origins of Totalitarianism*. Schocken.
- [2] Balkin, J. (2018). Free Speech in the Algorithmic Society. *UC Davis Law Review*.
- [3] Barabási, A.-L. (2016). *Network Science*. Cambridge University Press.
- [4] Bateson, G. (1972). *Steps to an Ecology of Mind*. University of Chicago Press.
- [5] Bianconi, G. (2021). *Higher-Order Networks*. Cambridge University Press.
- [6] Bridle, J. (2018). *New Dark Age*. Verso.
- [7] Cialdini, R. (2007). *Influence: The Psychology of Persuasion*. Harper Business.
- [8] Clark, A. (2015). *Surfing Uncertainty*. Oxford University Press.
- [9] Ellul, J. (1965). *Propaganda*. Vintage.
- [10] Friston, K. (2010). The Free-Energy Principle. *Nature Reviews Neuroscience*.
- [11] Jaynes, E.T. (1957). Information Theory and Statistical Mechanics. *Physical Review*.
- [12] Kahneman, D. (2011). *Thinking, Fast and Slow*. Farrar, Straus and Giroux.
- [13] Landauer, R. (1961). Irreversibility and Heat Generation. *IBM Journal*.
- [14] Mercier, H., & Sperber, D. (2017). *The Enigma of Reason*. Harvard University Press.
- [15] Newman, M. (2018). *Networks*. Oxford University Press.
- [16] Postman, N. (1985). *Amusing Ourselves to Death*. Penguin.
- [17] Prigogine, I., & Stengers, I. (1984). *Order Out of Chaos*. Bantam.
- [18] Schrödinger, E. (1944). *What is Life?*. Cambridge University Press.
- [19] Shannon, C. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*.
- [20] Ward, J. (2022). *The Loop*. Hachette.
- [21] Zuboff, S. (2019). *The Age of Surveillance Capitalism*. PublicAffairs.