

# **AI and the Optimization–Design Mismatch**

*Why making a broken system faster is not progress*

Björn Kenneth Holmström

April 2026

<https://bjorkennethholmstrom.org/essays/ai-optimization-mismatch>

**The Optimization–Design Mismatch Principle:** *When a system's governing metrics and feedback structures are misaligned with its intended purpose, increasing the efficiency of optimization within that system amplifies systemic dysfunction rather than resolving it.*

We are witnessing a textbook case of this principle in real time. Artificial intelligence is being deployed at scale to optimize a global economic operating system whose core metrics—Gross Domestic Product, quarterly earnings, engagement minutes—are structurally misaligned with human and ecological flourishing. The result is not the elimination of drudgery and the unlocking of human potential that techno-optimists promised. It is the acceleration of a broken machine toward a cliff.

This essay diagnoses that mismatch, traces its consequences, and outlines what a realigned deployment of AI—one governed by a higher-fidelity signal—would require.

---

## I. The Principle

Systems theory and cybernetics offer a precise language for this failure mode. Ashby's Law of Requisite Variety states that a controller must have at least as much variety as the system it governs. A thermostat with a single temperature sensor cannot regulate a room with complex thermal gradients. A governance system that observes only monetary throughput (GDP) cannot regulate an economy embedded in ecological limits, social relationships, and human psychology.

When a system's feedback channel is too narrow, optimizing its performance along that channel does not improve the system's overall health. It makes the system more efficient at *ignoring everything the channel cannot see*. This is the **Optimization–Design Mismatch Principle**. It explains why:

- **GDP rises** while soil fertility, social trust, and mental health decline. The metric sees the transaction that repairs the flood damage but not the wetland that prevented the flood.
- **Engagement metrics soar** while public discourse fragments and anxiety rises. The algorithm optimizes for attention, and outrage captures attention more reliably than nuanced truth.
- **Corporate efficiency increases** while worker precarity deepens. The balance sheet sees the cost of labor but not the human cost of burnout or the long-term cost of lost institutional knowledge.

The principle is not a critique of optimization itself. It is a warning about **optimizing the wrong thing**. A car with a broken steering linkage does not become safer if you install a more powerful engine.

---

## II. AI as a Mismatch Amplifier

Artificial intelligence is, at its core, an optimization technology. Machine learning models are trained to minimize a loss function—a mathematical proxy for "error" on a specific task. This makes AI extraordinarily powerful. It also makes it extraordinarily dangerous when the loss function is misaligned with human values or systemic health.

Consider how AI is currently being deployed:

Domain	What AI Optimizes	What the Proxy Ignores	The Amplified Dysfunction
<b>Content Recommendation</b>	Watch time, engagement, ad clicks.	Veracity, psychological impact, civic health.	Algorithmic amplification of outrage, misinformation, and addictive content.
<b>Labor Automation</b>	Labor cost reduction, throughput per hour.	Worker dignity, community stability, purpose.	Accelerated precarity and the "buffer stock of unemployed" logic, even as productivity rises.
<b>Supply Chain Logistics</b>	Delivery speed, inventory turnover.	Carbon footprint, local economic resilience, supplier welfare.	Fragile, just-in-time systems optimized for efficiency but vulnerable to disruption and ecologically destructive.
<b>Financial Trading</b>	Microsecond arbitrage, predictive pricing.	Systemic risk, productive investment, inequality.	A financial system increasingly decoupled from the real economy, extracting rents rather than allocating capital.
<b>Generative AI in Business</b>	Reducing the cost of content creation, customer service, and code.	The erosion of human skill, the homogenization of culture, the displacement of creative labor.	A flood of synthetic content that competes with human expression while eliminating the entry-level jobs that once built careers.

In every case, AI is being used to **scale optimization within a pre-existing, misaligned objective function**. The technology is not the root problem. The problem is that the "steering wheel" of the economy—its governing metrics—is disconnected from the "wheels" of actual human and ecological well-being. AI is simply a more powerful engine pushing the vehicle faster in whatever direction the disconnected steering happens to be pointed.

---

### III. The Consequences: Lock-In and Acceleration

The Optimization–Design Mismatch Principle has a corollary: **the more powerful the optimization technology, the faster the system locks into its misaligned trajectory.** AI does not merely replicate existing dysfunctions; it *hardens* them.

- **Path Dependency:** Once an industry has invested billions in AI systems optimized for a specific proxy (e.g., maximizing ad revenue), it becomes structurally difficult to pivot to a different objective, even if the harms of the current objective become undeniable.
- **Complexity Opacity:** AI models, particularly deep learning systems, are often "black boxes." When a model trained to maximize engagement starts promoting extremist content, its operators can claim they didn't intend it and don't fully understand why. This opacity diffuses accountability and makes systemic correction harder.
- **The Automation of Administration:** As noted in *The Hidden Cost of Money*, a vast portion of the modern workforce is engaged in the administrative overhead of maintaining artificial scarcity—billing, compliance, claims adjustment. AI is being deployed to make this overhead *cheaper*, not to eliminate the need for it. It optimizes the bureaucracy rather than redesigning the system to be less bureaucratic.

The result is an acceleration of what the ecological economist William Rees calls the "global mega-parasite": an economic system that extracts from its host (the biosphere and human communities) with increasing efficiency, measuring the extraction as prosperity and the resulting damage as additional growth when it finally surfaces.

---

### IV. The Alternative: Aligning Optimization with Requisite Variety

The solution is not to halt AI development. It is to **redesign the objective function and the feedback architecture that AI serves.**

If the problem is a low-resolution signal (GDP, profit, engagement) governing a high-resolution system (human lives, ecosystems, communities), the design response must be to **increase the resolution of the signal** and **distribute the capacity to respond.**

This is the work that the Global Governance Frameworks (GGF) project and related initiatives attempt to operationalize. It involves three interdependent shifts:

#### 1. From a Single Proxy to a Plural Dashboard (The LMCI)

The **Love, Meaning, and Connection Index (LMCI)** is a proposed alternative to GDP. It tracks multiple dimensions of flourishing—care relationships, cultural vitality, ecological health, purpose alignment—using

independent data sources rather than a single aggregate. An AI system governed by an LMCI signal would be trained to optimize for *actual well-being*, not just monetary throughput. A content algorithm would be penalized for amplifying outrage; a supply chain AI would be rewarded for increasing local resilience.

## 2. From Centralized Control to Distributed Sensing (Subsidiarity)

A single AI, no matter how powerful, cannot perceive the full variety of local conditions. The GGF's **Bioregional Autonomous Zones (BAZs)** are designed as distributed "observation nodes"—governance units at a scale where the specific needs of an ecosystem and a community can be directly perceived. AI deployed within a BAZ could be tuned to optimize for *that specific bioregion's* LMCI goals, rather than a global corporate objective. This restores the **requisite variety** that centralized optimization destroys.

## 3. From Conditional Scarcity to Unconditional Baseline (AUBI)

AI-driven automation will continue to displace labor within the current system. The GGF's **Adaptive Universal Basic Income (AUBI)** provides the economic architecture for this transition. An unconditional baseline decouples survival from employment, transforming AI from a *threat* to workers into a *liberator* of human capacity. It creates the material security necessary for people to engage in the care work, creative pursuits, and community building that AI cannot automate—and which an LMCI-aligned economy would *value* through mechanisms like the **Hearts** currency.

---

## V. Conclusion: From Fear to Design

The fear that AI is "optimizing a broken system" is not Luddism. It is an accurate diagnosis of the Optimization–Design Mismatch. The response to this fear is not to halt technological progress but to **redirect it toward a more intelligent objective**.

The most important question for the next decade of AI governance is not "How do we make AI safe?" It is "**Safe for what? Optimizing which metrics? In service of whose definition of flourishing?**"

The technology itself is agnostic. It will optimize whatever function we give it. The task before us—the urgent, civilizational task—is to **redesign the function**. That means replacing GDP with LMCI. It means replacing centralized extraction with bioregional stewardship. It means replacing precarious employment with an unconditional foundation for meaningful contribution.

We have built a powerful optimization engine. Now we must build a steering wheel worthy of its power.