

Programmable Incentives: Aligning Behavior in AI Strategy

Executive Summary

Soumitra Dutta¹ and Yves Doz²

An agentic supply-chain system detects a disruption. It sees rising demand in one region, delayed shipments from a supplier, tightening logistics capacity, and early signs of inventory pressure. It reasons through the alternatives and recommends a response: shift inventory, ask suppliers for updated data, reroute shipments, reward partners who can add capacity. The system may even initiate part of the response. But the outcome still depends on behavior. Suppliers must share accurate data. Logistics partners must reprioritize. Store managers must accept substitutions. Employees must trust the system enough to act. External partners must see a reason to cooperate.

This is the alignment problem at the heart of AI strategy. As firms deploy agentic systems, success is no longer determined only by the quality of the intelligence those systems generate. It is determined by whether the humans, organizations, and AI agents around those systems actually behave in ways that turn intelligence into outcomes.

This paper argues that competitive advantage in the AI era depends on two capabilities that are complementary and mutually dependent. The first is intelligence generation: building the data, models, decisions, workflows, and interfaces that make up the AI Intelligence Stack. The second is behavioral orchestration: ensuring that all actors — employees, partners, suppliers, customers, developers, and AI agents — behave in ways that convert that intelligence into results at scale. Most organizations have invested heavily in the first. Very few have invested systematically in the second. Programmable incentives are the mechanism through which behavioral orchestration is built.

Why Traditional Incentives Are No Longer Sufficient

For most organizations, the incentive toolkit has not changed fundamentally in decades. Bonuses reward annual performance. Supplier contracts specify delivery terms. Loyalty programs reward transactions. These instruments were designed for a world of periodic decisions, bounded relationships, and relatively stable operating conditions. They are not equipped for the continuous, distributed, granular behavioral alignment that agentic organizations require.

The problem is not only speed, though timing matters. The deeper issue is visibility. Traditional incentives reward outcomes that are easy to measure — revenue, delivery, clicks, cost. They are blind to behaviors that AI systems can now detect but traditional metrics ignore: a supplier who proactively shares early-warning data; an employee who contributes accurate customer insight even when it does not close a deal; a developer who writes secure, maintainable code rather than maximizing output; a logistics partner who maintains buffer capacity at personal cost to absorb system-level disruptions. When these behaviors go unrewarded, they go unprovided.

¹ Director, Portulans Institute, Washington DC and Former Founding Dean, SC Johnson College of Business, Cornell University, New York, USA

² Solvay Chaired Professor of Technology Innovation Emeritus, INSEAD, France

A new capability is now both possible and necessary: programmable incentives — adaptive, software-based mechanisms that dynamically reward, recognize, prioritize, or grant access based on real-time contribution. They can use money, tokens, credits, reputation scores, rankings, access rights, preferred status, or governance participation. They can be issued automatically, adjusted as conditions change, and calibrated differently for human employees, ecosystem partners, and AI agents acting on behalf of firms.

The Architecture of Behavioral Alignment

Programmable incentives are not simply better rewards. They are a new architecture. To work, six layers must operate together.

The contribution layer defines what behavior creates value. Not all activity is equally valuable. Data accuracy may matter more than data volume. Resilience contribution may matter more than cost efficiency. Code security may matter more than code quantity. This layer is strategic: leaders must decide what behaviors they want to shape, before they can design incentives to shape them.

The measurement layer determines how contributions are observed. This is where AI changes the incentive problem. Firms can now see who shares useful early-warning signals, whose code is secure and reusable, whose participation strengthens community trust, and whose actions improve system resilience. Measurement converts invisible behavior into visible contribution.

The incentive unit layer determines what carries the reward. Sometimes it is money. Often it is tokens, credits, reputation scores, rankings, access rights, governance votes, or status markers — digital units that translate contribution into recognition, influence, or access. Google Local Guides uses points, levels, and early access to encourage users to contribute reviews and local knowledge. GitHub Sponsors enables communities to financially support open-source contributors. Neither is a cryptocurrency system, but both illustrate the principle: contribution can be encoded into digital units that accumulate, compound, and carry real strategic value.

The allocation logic layer is where incentives become executable. Ride-sharing platforms use dynamic pricing when supply and demand diverge. Energy demand-response programs compensate customers for shifting consumption during grid stress. In both cases, a system detects imbalance and adjusts the incentive in real time. Participants respond. The system learns and recalibrates.

The governance layer decides who controls the rules. Programmable incentives shape behavior directly, so organizations must decide who can adjust the rules, how gaming is detected and corrected, how disputes are resolved, and when human judgment must override system logic.

The feedback layer ensures the system improves. If incentives produce gaming, they must be redesigned. If they crowd out intrinsic motivation, they must be recalibrated. If agents optimize locally while damaging system-wide performance, the objectives must change.

Inside the Firm and Across Ecosystems

Inside the organization, programmable incentives address the dual alignment challenge of agentic AI: aligning both human employees and the AI agents working alongside them. Consider sales. A firm deploys AI agents to improve forecasting, identify opportunities, and prioritize accounts. If salespeople are rewarded only for short-term revenue, they underinvest in accurate customer data.

If the AI agent is optimized only for conversion probability, it steers attention toward easy wins at the expense of strategic relationships. A programmable approach aligns both — rewarding employees for data quality, forecast accuracy, and long-term customer value, while recalibrating AI agents whose recommendations improve short-term metrics but weaken strategic priorities.

The greater opportunity lies beyond the firm. Modern organizations depend on contributors they do not employ: suppliers, developers, creators, users, data providers, logistics partners, and autonomous agents acting on behalf of others. AI makes it possible to observe and reward ecosystem behavior at a new level of granularity. Supply chains can reward partners who share real-time disruption signals, not merely those who deliver on time. Open-source platforms can recognize developers who maintain security and documentation, not merely those who commit the most code. Energy grids can compensate households, businesses, and automated devices that shift consumption during periods of peak stress. In each case, distributed infrastructure becomes more resilient when contributions are measured and incentives are programmable.

The Risks of Designing Badly

Programmable incentives carry real risks. If every action is measured and scored, participants may feel monitored rather than motivated — and intrinsic commitment can erode. If incentives reward what is easy to measure, participants optimize for metrics rather than outcomes. If the rules are opaque, participants lose trust in the system. And if AI agents are given objectives that optimize narrowly, they can satisfy every metric while undermining the purpose the metrics were meant to capture.

These risks argue not against programmable incentives but for more careful design. Effective systems combine economic incentives with social, reputational, and governance mechanisms. They reward contribution to genuine value, not merely observable proxies for it. They preserve space for intrinsic motivation — purpose, belonging, identity, and craft — that no reward system can fully substitute. And they maintain human oversight: some behaviors should not be scored, some decisions should not be automated, and some values should not be made negotiable.

Strategy as Behavioral Infrastructure

The strategic implication is direct. Organizations that invest only in AI capabilities — building better models, improving predictions, automating decisions — will find that intelligence without aligned behavior does not produce competitive advantage. The supply chain that detects a disruption but cannot mobilize suppliers to respond. The recommendation system whose insights go unused because employees are rewarded for other behaviors. The platform whose AI improves continuously but whose community drifts toward disengagement because no one designed the incentives to keep contributors aligned.

Programmable incentives are not a feature of compensation design or platform management. They are the foundational behavioral infrastructure of the AI-enabled organization. Intelligence generation determines what the firm can perceive and decide. Behavioral orchestration determines whether the people, partners, and agents around that intelligence act in ways that turn it into competitive outcomes. In the age of agentic AI, both capabilities are required, and neither is sufficient without the other. The organizations that build both — and integrate them coherently — will compete in a fundamentally different register from those that build only one.

Disclosure: AI was used for research and copy editing of the article