Rethinking Strategy in a World of Superintelligence:

Soumitra Dutta and Yves Doz
Former Founding Dean, SC Johnson College of Business, Cornell University
Solvay Chaired Professor of Technology Innovation Emeritus, INSEAD

On September 9, 2026, a swarm of 10,000 AI agents produced a 165-page proof of one of the hardest unsolved problems in mathematics. The Navier-Stokes existence and smoothness problem had defeated the world's best mathematicians for decades. The AI swarm solved it in 88 hours. The result was mechanically verified as logically correct. The mathematical community is still working out whether the proof is fully correct.

But there is a deeper question, and it is the one that matters for business leaders. Suppose the proof is correct. Can any mathematician understand it well enough to teach it? To build on it? Or does it sit in the literature like a display item—verified, acknowledged, and inaccessible?

This question is no longer confined to mathematics. AI systems are producing frontier knowledge across drug discovery, semiconductor design, and materials science at a pace and scale that human experts cannot match. The question every leader in a knowledge-intensive industry must now ask is not "can AI produce valuable knowledge?" It plainly can. The question is: "what happens to our people's ability to understand, verify, and build on what the AI produces?" The answer will define which organizations thrive over the next decade—and which accumulate knowledge they cannot use.

Two Ways of Working at the Frontier

On the same September day, another team had been quietly working on the Navier-Stokes problem for months. NYU mathematician Tristan Buckmaster and an AI researcher worked together intensively, using AI assistance to extend their own mathematical engagement. Their approach was slower and less dramatic. But the mathematical understanding they developed was genuine and transmissible. They could explain what they had found, teach it, and build on it.

These two approaches define a distinction that is now strategically decisive. In the first, humans and AI work at the frontier together: the AI extends what humans can do while the humans maintain and deepen their understanding. In the second, AI agents work largely autonomously: humans set up the system, point it at a problem, and verify the output. The AI produces far more, far faster—but the humans' own expertise may or may not grow through the process, and often does not.

Most organizations are being pushed, by efficiency pressure and competitive urgency, toward autonomous AI production. The problem is not autonomous production itself—it is deploying it without maintaining the collaborative mode alongside it. Organizations that do this are building a hidden liability that will not appear on any balance sheet until it is too late.

The Knowledge You Don't Own

An organization deploys AI agents to produce drug candidates, chip designs, or scientific analyses faster than any human team could. The outputs accumulate. Knowledge capital grows. Every dashboard looks good. Then one of three things could happen.

The agentic system is updated or replaced. The new version produces different outputs. Nobody in the organization has the domain depth to evaluate whether the change is an improvement. They are dependent on the continued operation of a system they do not fully understand.

A novel problem arises that the AI cannot handle. The problem is genuinely new, requiring the kind of creative judgment that emerges from deep expertise. The human team no longer has that expertise. They have been generating outputs with AI for years without building their own capability.

A regulator, a client, or a board member asks the organization to explain and be accountable for a key decision. No one can. The AI produced the knowledge. No human can reconstruct the reasoning or stand behind it.

This is the knowledge capital / human capital divergence. Knowledge capital—documented methods, verified results, predictive models—is growing rapidly. Human capital—the tacit expertise, judgment, and domain understanding that makes knowledge usable—is not keeping pace. In the short run, this divergence is invisible. In the medium run, it is existential. The knowledge that your organization cannot explain, verify, teach, or extend is not really yours.

Four Roles That Cannot Be Automated Away

The answer is not less AI. It is more deliberate investment in the human roles that make AI-produced knowledge usable, teachable, and trustworthy. Four roles are essential—and all four are currently underdeveloped in most organizations.

The most important and most neglected is the **Interpreter**. When a 165-page AI-generated proof is published, or an AI-designed drug molecule enters trials, someone must work through the output deeply enough to explain what it demonstrates, why this approach works, what it opens for future research, and how it connects to what the field already knows. Without this work, the proof and the molecule exist as verified facts that nobody can build on. The Interpreter is not a science communicator. They are a domain expert whose primary function is bridging the gap between what AI produces and what humans can absorb and use. Developing one takes years. The role cannot be delegated to whoever is available.

The Orchestrator directs the agentic system toward the right problems. This requires genuine domain expertise—not just familiarity with AI tools. An Orchestrator without deep pharmacological knowledge cannot meaningfully specify what an AI drug discovery system should be optimizing for. They are a domain expert who has learned to leverage autonomous AI systems.

The Verifier checks AI outputs for correctness, validity, and real-world applicability. Mechanical verification tools can confirm logical consistency but cannot confirm significance, generalizability, or freedom from subtle errors in framing. In drug discovery, the Verifier checks whether an AI-designed molecule is biologically plausible, not only computationally attractive.

The Verifier's value comes precisely from the domain expertise that autonomous AI deployment, over time, threatens to erode.

The Applier deploys AI-produced knowledge in real-world contexts, and most importantly knows when not to use the AI's output. That judgment requires the kind of contextual understanding that lives in professional communities, not in any individual or a system.

What This Looks Like in Practice

Three sectors illustrate both the opportunity and the risk.

Insilico Medicine's rentosertib became the first drug in which both the biological target and the therapeutic compound were discovered using generative AI. Phase 2a results published in June 2025 showed improved lung function in 71 IPF patients; the drug has since entered Phase 3 trials in China. The four roles are clearly in play: human pharmacologists directing the system, interpreting its molecular outputs, verifying them against experimental biology, and navigating the clinical trial process with the human judgment that regulatory frameworks require. The risk is that the developmental pathway through which the next generation of medicinal chemists builds expertise—years of structure-activity experiments, binding assays, and drug-likeness intuition—is being compressed. A field that cannot fill its Interpreter and Verifier roles will not be able to evaluate the next AI drug system when it produces outputs that conflict with its predecessor.

In semiconductor engineering, Google's AlphaChip designed the chip layouts for three generations of its AI accelerators, producing configurations that outperform what human engineers would design—in hours rather than months. The configurations violate heuristics that human chip designers have developed over decades. Google's engineers are now primarily Orchestrators and Verifiers of AI-designed layouts. The Interpreter challenge is acute: AlphaChip's non-intuitive configurations represent knowledge about what makes great chip design whose underlying principles are not yet understood in human terms. Knowledge capital in the chip designs is growing. Human capital in chip design is not keeping pace.

In materials science, DeepMind's GNoME predicted the stability of 2.2 million new crystal structures—equivalent, by DeepMind's own estimate, to 800 years of traditional discovery progress. There are now far more AI-predicted stable materials than the materials science community can characterize and develop in the near term. The interpretive challenge is not verifying that the predictions are correct. It is understanding why specific materials are structured as they are and what the predictions reveal that human expertise had not previously seen.

What Leaders Must Do Now

The difference between organizations that compound human expertise alongside AI-produced knowledge and those that deplete it is not primarily a technology decision. It is a leadership decision about what to protect, what to measure, and what to invest in.

The most fundamental protection is maintaining human-AI frontier collaboration alongside autonomous AI production. The collaborative mode—slower, more expensive, less impressive by output metrics—is the only reliable developmental pathway for the Orchestrators, Interpreters, and Verifiers your agentic systems need. Organizations that eliminate it in the name of efficiency are farming the soil without replanting. The harvest looks good for several seasons. Then the field is exhausted.

The most important measurement is also the least common. Standard performance metrics capture what AI can also capture: outputs, publications, products, patents. They do not capture whether your people can understand, verify, teach, and extend what your AI systems produce. Leaders need a second set of indicators: periodic assessments of unassisted human performance on calibrated domain tasks; measures of Verifier community depth relative to agentic output volume; and honest evaluation of how much of the organization's knowledge base any human could actually reconstruct or explain. Treat the gap between your knowledge capital and your human capital as a strategic risk indicator. Report it to the board alongside revenue.

The most consequential investment is in the Interpreter role as a career, not a task. Identify the individuals in each domain who have both deep expertise and the ability to bridge between AI outputs and human comprehension. That combination is rare. Build career paths around it, protect their time, and create structured programs through which AI-produced knowledge is worked through deeply before it is deployed. This will feel like a cost. It is actually insurance against the day when your agentic systems change, fail, or need to be governed.

When AI Becomes Smarter Than Everyone: The Leadership Stakes

The governance challenge described above is urgent today. It becomes existential as AI capabilities expand. Leaders who understand only the immediate challenge will be unprepared for what is coming.

AI systems are already superhuman in specific domains. In mathematics, AI solved four of six International Mathematical Olympiad problems in 2024 and claimed a Millennium Prize Problem in 2026. In biology, AlphaFold solved the protein folding problem that human researchers had worked on for fifty years—solving problems humans had not been able to solve, not merely solving them faster. In games, the era when human-plus-AI collaboration exceeded AI alone has effectively ended: current chess and Go engines are so far beyond human capability that human input adds noise rather than signal.

The trajectory suggests that systems capable of exceeding human performance across most cognitive domains simultaneously may arrive within this decade. When they do, the question of what humans contribute to strategy, decision-making, and governance will be the defining leadership challenge of the era.

Here is the strategic paradox that leaders need to understand now. The more capable AI becomes, the more expertise your organization needs to govern it well—deeper Orchestrators, more skilled Interpreters, more expert Verifiers. But the more you deploy AI in a way that replaces rather than develops human capability, the less governance expertise your organization accumulates. More capable AI demands more governance capacity; substitutive deployment provides less of it. This is what we call the knowledge sovereignty trap: the organizations most aggressive in deploying AI may arrive at the superintelligence threshold least equipped to govern it.

The trap closes through a mechanism that is already visible. Each generation of researchers and practitioners trained primarily on AI outputs, rather than through genuine engagement with frontier problems, knows somewhat less than the generation before it. Not because information is less available—it is more available than ever—but because the productive struggle through which deep expertise is built has been replaced by AI-generated answers. Over two or three professional generations, this produces a population of domain practitioners skilled at using AI but lacking the

depth required to evaluate, redirect, or govern it when it matters most. By that point, rebuilding the expertise base is not an organizational investment problem. It is a generational reconstruction project.

The implication for business strategy and leadership is specific. At the superintelligence horizon, the five-layer AI Intelligence Stack—data, models, decisions, workflows, interfaces—becomes largely self-managing. Models self-improve. Decisions self-calibrate. Workflows self-redesign. Human competitive differentiation migrates upward to a layer above the current stack: what we call the Purpose and Values Layer. This layer asks what the intelligence stack should be doing, what it must never do, and whose interests it serves.

Consider what this means for the CEO's job. When AI can generate better strategic options than any human team, the CEO's primary strategic contribution shifts from formulating strategy to specifying what the organization is trying to optimize for, what values must never be violated, and what kinds of success matter. This is not a simpler job—it is a harder one. The vagueness that allowed traditional strategy to accommodate competing interests and evolving priorities becomes, at superintelligence, a specification error. Organizations that have never had to articulate their values with precision will find themselves governing AI systems that optimize for measurable proxies rather than what they actually care about.

The board's role transforms in the same direction. Board oversight of AI was, until recently, primarily risk management: what could go wrong, and how do we monitor it? At the superintelligence horizon, board oversight becomes the primary mechanism through which purpose and values are specified and enforced. A board that can articulate clearly what the organization values, what it must not do, and what it would refuse to do even if it were profitable, will be able to govern superintelligent systems effectively. A board that cannot will find itself approving systems that optimize for what can be measured while pursuing objectives nobody deliberately chose.

The competitive divide that is opening now will determine which organizations can navigate this transition. On one side are organizations that invest, during this window, in developing the human governance capacity that superintelligence will require: deep domain expertise in the Orchestrator, Interpreter, Verifier, and Applier roles; clarity about organizational values; and leadership capable of specifying purpose rather than merely approving strategy. On the other side are organizations optimizing for output during this period, systematically depleting the human expertise and values clarity that their governance capacity will need. The gap between these two groups will not be visible for several years. It will be decisive within a decade.

The Choice Belongs to Leaders

Knowledge-intensive organizations are at an inflection point. The AI systems available today can produce valuable frontier knowledge at scale. The AI systems coming in the next decade will produce knowledge that exceeds human capability across more and more domains. Whether that knowledge compounds human understanding or gradually displaces it is not a technological question. It is a leadership question about what to protect, what to measure, and how to develop people in a world where AI can produce most of the answers.

The organizations that thrive will be those that treat AI-produced knowledge not as a product to ship but as material to engage with, interpret, verify, and build upon. They will invest in the

Interpreters who make that engagement possible. They will protect the collaborative human-AI pathways through which domain expertise continues to develop. They will measure the gap between what their AI systems know and what their people understand—and they will take that gap seriously as a governance risk, not just a training problem.

Most importantly, they will use this period to develop the values clarity and governance depth that superintelligent systems will require. That is not a technical challenge. It is a leadership one. And unlike the AI capabilities that are arriving whether or not organizations are ready for them, the human capacity to govern those capabilities is something that only deliberate investment will build.

Disclosure: AI was used for research and copy-editing in the production of this paper.