Phronetic Leadership

A Virtue Ethics Approach to Strategic Decision-Making with Generative AI

Abstract

Generative AI (GenAI) is becoming embedded in strategic decision-making (SDM), shaping how organizational actors frame problems, generate options, and prepare recommendations. Existing research has largely emphasized performance gains, efficiency benefits, and technical risks such as bias, opacity, and hallucination. Less attention has been paid to a deeper problem: how persuasive AI outputs may erode the human judgment of those who remain accountable for strategic choices. This conceptual paper addresses that problem by asking how individual decision makers can preserve judgment and responsibility under conditions of increasing algorithmic mediation. Drawing on the Aristotelian concept of *phronesis* (practical wisdom), the paper argues that SDM cannot be delegated to GenAI because strategy requires context-sensitive deliberation, the weighing of competing goods, and accountable human endorsement. Methodologically, the paper develops its argument through conceptual analysis of literature on SDM, AI-assisted work, and virtue ethics, illustrated by a real-life case of fabricated references in a strategic government advisory report. The paper's central contribution is to reframe GenAI in strategy not primarily as a problem of technical substitution, but as a problem of judgment preservation. It then offers a practical contribution in the form of three phronetic disciplines for everyday use by strategists: own the judgment, pause before commitment, and stand behind the choice. These disciplines are proposed as simple, repeatable habits that help individual decision makers preserve critical reflection, evaluative ownership, and answerability in AI-assisted strategic work.

Keywords: Generative AI; Strategic Decision-Making; Phronesis; Virtue Ethics; Judgment Preservation

Introduction

In 2025, a researcher noticed something in a Deloitte report prepared for Australia's Department of Employment and Workplace Relations (DEWR) that should never have appeared in high-stakes advisory work: references that did not exist. The report was later found to contain fabricated academic citations, incorrect references, and a made-up quotation attributed to a Federal Court judgment, revealing multiple AI-generated errors. The key detail is not simply that the report contained mistakes - errors occur in professional work -, but that they appeared in work commissioned precisely for its judgment. Deloitte was being paid to provide reliable strategic advice in a politically sensitive public-sector setting. In such environments, trust is not ornamental but foundational, especially when clients rely on strategic consultants to interpret ambiguity, weigh evidence, and justify consequential decisions (Abgeller et al., 2025). When invented references enter this kind of work, the damage is simultaneously factual, reputational, and institutional. The case also matters because it was not uncovered through an internal quality check but because an external academic noticed that the report cited knowledge that did not exist. One red flag was a supposed book attributed to a University of Sydney professor that was not real. The problem was not buried in technical model behavior; it was visible in the final professional artifact. AI-generated fluency had already crossed into advisory authority.

This episode illustrates a broader shift in contemporary knowledge work. Highly capable Generative AI (GenAI) is now widely accessible for managerial and strategic tasks. Unlike earlier rule-based systems, GenAI produces fluent outputs through large-scale pattern

prediction, a strength that can mask limitations such as hallucinations, bias, and opacity (Feuerriegel et al., 2024). Humans have long extended cognition through writing and digital tools, but GenAI is distinctive because it presents this extension in an interactive, seemingly deliberative form that can enter directly into strategic reasoning. The question is no longer only how cognition is extended, but whether critical judgment is being quietly outsourced.

In management research, GenAI is widely welcomed as a transformative tool, enhancing analytics, digital transformation, and decision-making (Brynjolfsson et al., 2025). Critical voices, however, warn of reinforced biases, weakened reasoning, and the automation of flawed or unethical choices (Gerlich, 2025). Yet most research still focuses on performance and productivity. Underexplored is how repeated reliance on GenAI may reshape the judgment of strategists themselves. As strategically relevant thinking becomes mediated by persuasive systems, a critical question arises: how can humans maintain judgment, reflection, and accountability under conditions of growing algorithmic mediation? The Deloitte case shows why this matters: GenAI need not replace strategists to distort strategy; it only needs to be persuasive enough for decision-makers to lower their guard.

This article reframes GenAI in strategy not as a problem of capability substitution but of judgment preservation. While strategic decision-making research has documented the cognitive limits of reasoning under uncertainty (Eisenhardt, 1989), it has not addressed how human judgment is safeguarded when machine-generated reasoning shapes the strategic process itself. Drawing on Aristotelian phronesis, the article argues that leaders must cultivate the practical wisdom required to protect their judgment. Although AI can support synthesis, option generation, and scenario building, it cannot perceive moral salience, adjudicate competing goods, or assume responsibility for outcomes. Accordingly, the article contributes by reframing GenAI as a challenge of judgment under algorithmic mediation, identifying mechanisms through which AI fluency erodes authorship and accountability, and proposing three practical disciplines to preserve practical wisdom in AI-assisted strategic work.

Why Fluent Output is Dangerous in Strategy

Research on SDM has long emphasized the cognitive limits of human judgment under conditions of uncertainty. Foundational theories such as bounded rationality (Simon, 1955), heuristics and biases (Tversky & Kahneman, 1974), ecological rationality (Gigerenzer, 2001), and dual-process theory (Kahneman, 2011) provide valuable insights into how individuals reason under uncertainty. More recent frameworks, such as the Search-Representation-Aggregation (SRA) model (Csaszar et al. 2024), offer a micro-level account of how alternatives are generated, problems represented, and judgments aggregated. Together, these perspectives illuminate the cognitive architecture of SDM, but they leave its ethical core comparatively underdeveloped.

GenAI intensifies this problem not simply by adding more information, but by introducing a new mode of cognitive pressure. Human judgment is already vulnerable to shortcuts, fluency effects, and overconfidence. GenAI amplifies those vulnerabilities by delivering polished, rapid, and apparently coherent outputs that resemble deliberation without actually being judgment. In this sense, the threat is not that GenAI thinks like a human strategist. The threat is that it can make statistically generated responses appear sufficiently reasoned that human strategists reduce the level of scrutiny they would otherwise apply.

Kahneman's (2011) distinction between fast and slow cognition is especially relevant here. Under uncertainty, people often rely on intuitive, fluent processing and only selectively engage slower, effortful evaluation. Kahneman´s work further shows that when information is presented fluently and confidently, people are more likely to trust it, even when that trust is unwarranted. GenAI is therefore consequential in strategic settings because it does not merely provide content; it presents content in a form that invites premature assent. Its outputs can

look balanced, informed, and decision-ready even when they are weakly grounded, context-insensitive, or false. The risk is not only error in the output, but diminished critical engagement in the user.

This matters because GenAI differs from human intelligence in ways that are highly consequential for strategy. Human cognition is embodied, affective, and socially situated (Korteling et al., 2021). It integrates sensory experience, emotional attunement, tacit knowledge, memory, and normative commitment. These features make possible self-reflection, contextual discernment, and ethical deliberation, capacities that are indispensable in strategic work. GenAI, by contrast, represents a form of disembodied intelligence. Its outputs are fundamentally syntactic and statistical rather than experiential and evaluative. It operates without lived experience, moral horizon, or social accountability; its outputs are products of probabilistic generalization, not acts of intentional reasoning or responsible judgment (Korteling et al., 2021). Its apparent competence is also highly sensitive to context and task framing (Feuerriegel et al., 2024). What appears to a user as stable judgment may in fact be the temporary success of pattern prediction under a favorable prompt. In short, GenAI can generate plausible strategy talk, but it cannot exercise practical wisdom.

Within organizations, AI has typically been described either as a support tool that augments human judgment or as an automated agent embedded in workflows (Iansiti & Lakhani, 2020). In strategic contexts, however, the more pressing risk is not replacement but deference. As GenAI moves upstream in organizational processes, shaping how problems are framed, how information is synthesized, and how strategic options are imagined, its fluent outputs can create a cognitive illusion of understanding that diverts attention from reflection and verification (Gerlich, 2025). Repeated reliance may therefore weaken the habits of critique, deliberation, and accountable choice on which strategic judgment depends.

The theoretical problem, then, is not simply whether GenAI improves strategic analysis. It is whether fluent AI output can erode the human capacity to judge well while leaving formal accountability in place. The Deloitte case is not just an anecdote about citation failure. It is a concrete example of what happens when persuasive output, institutional credibility, and time-saving technology converge before judgment has fully done its work.

Virtue Ethics as Theoretical Lens

Once the problem is framed in these terms, the central question shifts. It is no longer only whether GenAI can generate options or synthesize information, but what sustained reliance on these systems cultivates in the strategist: discernment or passivity, accountability or dependence, genuine judgment or mere endorsement.

Classical philosophical concepts, especially those from virtue ethics, offer a productive way to approach this concern. Rooted in Aristotle's work (2009), virtue ethics emphasizes the development of moral character and practical wisdom (phronesis): the capacity to deliberate well about what is good and fitting in particular circumstances. Rather than prescribing universal rules or optimizing for outcomes, virtue ethics focuses on the qualities of the agent: the dispositions that enable sound judgment when situations are ambiguous, complex, and uncertain. This is precisely where GenAI raises the hardest questions: not whether a tool is accurate, but what repeated use of that tool is doing to the quality of human judgment itself.

Virtue ethics is therefore better suited to this problem than frameworks oriented solely toward compliance or outcomes. Rules and results remain important, but they do not fully capture what is at stake. The issue is not only whether an AI-assisted decision follows procedure or yields desirable effects; it is whether increasing reliance on GenAI reshapes the strategist's capacities for discernment, responsibility, and moral agency. Virtue ethics moves

the focus from the decision's output to the formation, and potential erosion, of the decision-maker's judgment.

At the center of this lens is phronesis. Phronesis is not cleverness or technical optimization. It is context-sensitive practical wisdom: the ability to perceive what matters in a situation, balance competing considerations, and commit to a choice that can be publicly justified. In strategic work, this includes recognizing when seemingly efficient options conflict with organizational purpose, stakeholder dignity, or long-term responsibility. Phronesis clarifies why SDM cannot be delegated to GenAI. A system may support analysis, synthesis, and option generation, but it cannot perceive moral salience, weigh goods in light of human purposes, or bear responsible judgment.

Much of the AI governance literature focuses on organizational structures, such as policies, controls, and oversight mechanisms (Papagiannidis et al., 2025). That work is essential, but it does not address the moment when an AI-shaped recommendation becomes something an individual strategist must personally endorse. The Deloitte example makes this visible: even with contracts and controls in place, the decisive failure was a failure of judgment.

This Aristotelian framing also enables movement from diagnosis to prescription. If strategic judgment is a matter of phronesis, then the practical task for leaders using GenAI is to preserve the human capacities on which phronesis depends. The disciplines developed below, e.g. own the judgment, pause before commitment, and stand behind the choice, should therefore be understood not as generic managerial slogans, but as applied practices of practical wisdom under conditions of increasing algorithmic influence.

Three Practical Disciplines for Strategists Using AI

If the central danger of GenAI in strategic work is the erosion of judgment through overreliance, then the response should be practical guidance for the strategist who remains accountable for the final output. If phronesis is the capacity to deliberate well about what ought to be done in a particular situation, then its practical value in AI-assisted strategy lies in habit formation. The aim is not to provide decision makers with another complex evaluative framework, but with a small set of repeatable disciplines that help preserve critical reflection, practical wisdom, and accountability under conditions of speed, fluency, and cognitive pressure. The three disciplines proposed here are therefore intended as everyday habits of phronetic practice for individual strategists using GenAI. They are deliberately simple: each attaches to a distinct moment in the decision process and each can be enacted through a brief self-check before strategic endorsement.

1. Own the judgment. GenAI may suggest, summarize, and synthesize, but it cannot judge on the strategist's behalf. This first discipline requires the decision maker to treat AI output as input rather than conclusion. The practical task is simple: do not ask whether the output sounds strong, but whether you have independently assessed its assumptions, relevance, and implications. In Aristotelian terms, this is the discipline of keeping practical judgment with the agent rather than with the tool. In practical terms, this means using GenAI most cautiously in areas where one's own expertise is thin, checking the evidence on which the recommendation rests, and treating the system's fluency as something to interrogate rather than trust. The point is not to reject AI assistance, but to ensure that endorsement follows critical evaluation rather than mere acceptance.

2. Pause before commitment. Because GenAI compresses the time between question and answer, it can also compress the time between draft and endorsement. This discipline inserts a brief deliberative pause before commitment. Such a pause is not an inefficiency; it is a safeguard against the collapse of reflection into convenience. In Aristotelian terms, pause creates the practical space in which deliberation about particulars and ends can still occur.

Actionably, this means resisting immediate acceptance and separating generation from endorsement. Do not move directly from AI output to recommendation. Restate the issue in your own words. Compare at least one alternative. Ask what may be wrong, missing, overstated, or insufficiently challenged. Where the decision is consequential, create some distance between draft and commitment. The purpose of pause is not delay for its own sake, but better judgment.

3. Stand behind the choice. Strategic judgment remains human only if the final recommendation can still be owned by the person making it. This discipline requires the strategist to preserve answerability: the ability to explain why this option was chosen, what trade-offs it involves, and why it should be defended. Where such answerability is absent, responsibility becomes diffuse and strategic accountability weakens. A leader may use AI-generated material, but must still be able to say, in a fully human sense, that the judgment is their own. In practice, this means not presenting a strategic recommendation, memo, or slide deck that one cannot explain without the tool open. It also means being able to reconstruct the chain from evidence to conclusion in one's own words. If a strategist cannot do that, the output may still look polished, but judgment has already thinned.

Figure 1. Three phronetic disciplines for everyday AI-assisted strategic work

Discipline	Phronetic aim	When to use it	Daily habit	Quick self-check	What it protects against
Own the judgment	Keep final judgment with the human decision-maker	As soon as GenAI provides an answer or recommendation	Treat AI output as just one input among many	Do I believe this because I tested it or because it sounded right?	Passive acceptance, automation bias, misplaced certainty
Pause before commitment	Create space to think before endorsing AI-generated work	When moving from draft to decision	Add a brief pause, restate the issue in your own words	What might be missing, or too quickly assumed here?	Rushed decisions, reduced scrutiny, over-reliance on AI shortcuts
Stand behind the choice	Preserve ownership, accountability, and moral authorship	Right before approving, presenting, or acting on the recommendation	Explain your rationale publicly	Can I justify this choice in detail to someone who will be affected?	Diffused responsibility, shallow authorship, decisions no one fully owns

Together, these three disciplines are best understood as habits of practical wisdom for individual decision makers working with GenAI. They are designed to be lightweight rather than elaborate: not a new evaluative architecture, but a set of repeatable self-checks that can become part of ordinary strategic work. Their purpose is to ensure that, even when GenAI accelerates analysis and produces persuasive recommendations, the strategist retains the critical distance necessary for sound judgment.

When Fluent Output Outruns Practical Wisdom

With these disciplines in view, the Deloitte episode becomes more analytically useful. What failed was not only citation accuracy. What failed was the practical chain by which advisory work becomes endorseable. Official correspondence shows that DEWR expected expert advice underpinned by clear evidence and later required Deloitte to check and verify the report's references, citations, and statements. This material also shows that the use of a ChatGPT-based toolchain had been explicitly discussed and approved as part of the engagement. In other words, the case was not a simple story of prohibited technology use. It was a story about what happens when authorized tool use outruns disciplined judgment.

A phronetic reading sharpens the lesson. To own the judgment would have required that AI-assisted output remain provisional until its legal references, academic support, and inferential claims had been critically interrogated. To pause before commitment would have required more than light-touch review of polished prose; it would have required enough cognitive and temporal distance for verification, challenge, and comparison. To stand behind the choice would have required the final report to be defensible, in fully human terms, by those submitting it. Deloitte therefore illustrates the precise problem this article identifies: under conditions of speed, fluency, and institutional pressure, strategists can remain formally accountable while becoming practically deferential.

The case also strengthens, rather than weakens, the focus on individual discipline. Nothing in the public record suggests that GenAI autonomously issued the report. Humans remained in the loop throughout. Yet that did not prevent fabricated references from entering a document sold as expert advice. That is precisely why the problem cannot be reduced to organizational governance alone. The final burden still falls on the individual strategist who chooses to accept, revise, challenge, or endorse what the system has produced.

Why the Burden Keeps Falling on Individuals

This is where the broader structural tension in contemporary AI adoption becomes visible. Organizations are pushing speed, scaling adoption, and expecting workers to integrate GenAI quickly into daily work. McKinsey reports that leaders and employees want to move faster with AI (Mayer et al., 2025), while Microsoft's 2025 Work Trend data shows that 53% of leaders say productivity must increase even though 80% of the workforce says they lack enough time or energy to do their work; Microsoft also reports that 81% of leaders expect agents to be integrated into company AI strategy within 12-18 months (Microsoft, 2025).

At the same time, support remains uneven. In the University of Melbourne-KPMG global study, only 60% of organizations reported providing responsible AI training and only 34% reported having organizational policy or guidance for generative AI tools; the study also found widespread inappropriate and complacent use of AI at work, including reliance on AI output without evaluation and mistakes in work due to AI (Gillespie et al., 2025).

This mismatch matters because many governance conversations are organizational in form, while the lived burden is individual in practice. Leaders are expected to move faster, adopt aggressively, and produce strategically useful output under mounting time pressure. Yet the disciplines required to do this well are often assumed rather than taught. Deloitte is therefore not merely a cautionary tale about one consulting firm. It is a public illustration of a wider contradiction: organizations want the speed and scale of GenAI, but they often leave the hard work of discernment, verification, and accountable endorsement to individual decision-makers without equivalent investment in judgment-preserving practice.

This also implies that the preservation of judgment is not solely a matter of individual virtue, however important that remains. Organizations shape the conditions under which strategists use AI by structuring workflows, incentives, time horizons, data access, and norms of accountability. Recent managerial accounts of AI transformation stress that technology adoption creates value only when accompanied by process redesign, broad digital fluency, and organizational integration rather than siloed experimentation (Neeley & Leonardi, 2022). This article contends that process redesign must not only increase productivity; it must also preserve spaces for reflective assent, justification, and moral authorship. Otherwise, organizations may become more AI-capable while becoming less judgment-capable.

This is where virtue ethics adds something that cognitive and managerial accounts alone do not. Research on bounded rationality, heuristics, bias, and cognitive offloading explains important features of human and AI-assisted decision-making. But those traditions do not fully address the normative question of what kind of decision-maker repeated GenAI

use is helping to form. By foregrounding Aristotelian phronesis, this paper argues that strategic judgment is not exhausted by information processing or procedural correctness. It is also a matter of character, self-governance, and the preservation of a non-delegable human capacity for responsible assent.

The practical implication is not that organizations should exclude GenAI from strategic work, but that they should resist patterns of use that erode authorship, critique, and reflective pause. The three disciplines advanced here are deliberately modest in form but ambitious in purpose: they do not constitute a full governance regime, but they do offer a portable discipline for keeping strategic judgment human under conditions of accelerating AI use.

Conclusion

Deloitte is a useful example because it makes the practical problem difficult to ignore. A report sold as strategic expert advice contained fabricated references and false legal support, and it took external academic scrutiny to expose what internal controls had not prevented. In such settings, the danger of GenAI is not simply hallucination. It is the speed with which fluent output can acquire institutional authority before human judgment has fully engaged. This article has argued that the central challenge of GenAI in SDM is not technical substitution alone, but the preservation of human judgment under increasing algorithmic mediation. By foregrounding Aristotelian phronesis, the paper shifts attention from what GenAI can generate to what repeated dependence on it may gradually shape in those responsible for strategic choice. The deeper risk is not only factual inaccuracy, but diminished scrutiny, weakened evaluative ownership, and reduced answerability when persuasive output makes endorsement easier than deliberation.

The paper's main theoretical contribution is therefore to recast AI-assisted strategy as a problem of practical wisdom. SDM cannot be reduced to information processing, option generation, or analytical speed, because strategy also requires context-sensitive judgment about what ought to be done, why, and with what consequences. GenAI may support these activities, but it cannot bear judgment on behalf of the strategist.

Its practical contribution is to translate this insight into three simple disciplines for individual decision makers: own the judgment, pause before commitment, and stand behind the choice. These disciplines are intentionally modest. They function not as a governance model but as everyday habits of reflection designed to counteract the pull of speed and fluency. Their value lies in their usability: they help ensure that strategic recommendations remain humanly judged even when AI shapes their production.

This paper is conceptual and does not empirically test the proposed argument. Its claims are therefore primarily normative and analytical. Future research should examine whether the three disciplines meaningfully influence decision quality, ethical outcomes, and accountability in AI-assisted strategic work. Experimental and field studies could clarify when GenAI strengthens judgment and when it begins to displace it, as well as how repeated exposure shapes leaders' habits and capacities over time.

Ultimately, the argument is not that GenAI should be excluded from strategic work. It is that SDM remains irreducibly human in one crucial respect: someone must still judge what ought to be done and be able to stand behind that judgment. If GenAI is to strengthen rather than weaken strategy, the preservation of practical wisdom must remain central.

References

Abgeller, N., Saunders, M. N., Bachmann, R., & Mishra, A. (2025). Identifying trust exchange dynamics and constituents of employee trust within management consulting. *Work, Employment and Society, 39*(2), 336-360.

Aristotle. (2009). *The Nicomachean ethics* (D. Ross, Trans.). Oxford University Press.

Brynjolfsson, E., Li, D., & Lindsey Raymond, L. (2025). Generative AI at work. *The Quarterly Journal of Economics, 140*(2), 889-942. DOI: 10.1093/qje/qjae044

Csaszar, F. A., Ketkar, H., & Kim, H. (2024). Artificial intelligence and strategic decision-making: Evidence from entrepreneurs and investors. *Strategy Science, 9*(4), 322-345. DOI: 10.1287/stsc.2024.0190

Eisenhardt, K. M. (1989). Making fast strategic decisions in high-velocity environments. *Academy of Management Journal, 32*(3), 543-576. DOI: 10.2307/256434

Feuerriegel, S., Hartmann, J., Janiesch, C., et al. (2024). Generative AI. *Business & Information Systems Engineering, 66*, 111-126. DOI: 10.1007/s12599-023-00834-7

Gerlich, M. (2025). AI tools in society: Impact on cognitive offloading and the future of critical thinking. *Societies, 15*(1), 6. DOI: 10.3390/soc15010006

Gigerenzer, G. (2001). *Simple heuristics that make us smart*. Oxford University Press.

Gillespie, N., Lockey, S., Ward, T., Macdade, A., & Hassed, G. (2025). *Trust, attitudes and use of artificial intelligence: A global study 2025* [Report]. The University of Melbourne. DOI: 10.26188/28822919.v1

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.

Korteling, J. E., van de Boer-Visschedijk, G. C., Blankendaal, R. A. M., Boonekamp, R. C., & Eikelboom, A. R. (2021). Human-versus artificial intelligence. *Frontiers in Artificial Intelligence, 4*, Article 622364. DOI: 10.3389/frai.2021.622364

Mayer, H., Yee, L., Chui, M., & Roberts, R. (2025, January 28). *Superagency in the workplace: Empowering people to unlock AI's full potential*. McKinsey & Company.

Microsoft. (2025, April 23). *2025: The year the Frontier Firm is born*. Microsoft WorkLab.

Neeley, T.,& Leonardi, P. (2022). Developing a digital mindset: How to lead your organization into the age of data, algorithms, and AI. *Harvard Business Review, 100*(3), 50-55.

Papagiannidis, E., Mikalef, P., & Conboy, K. (2025). Responsible artificial intelligence governance: A review and research framework. *The Journal of Strategic Information Systems, 34*(2), Article 101885. DOI: 10.1016/j.jsis.2024.101885

Shotter, J., & Tsoukas, H. (2014). In search of phronesis: Leadership and the art of judgment. *Academy of Management Learning & Education, 13*(2), 224–243. DOI: 10.5465/amle.2013.0201

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics, 69*(1), 99-118. DOI: 10.2307/1884852

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124-1131. DOI: 10.1126/science.185.4157.1124