POLARIS SHIFT

When to Move Your North Star

A framework for strategic reorientation in the age of AI

Darko Tushev

White paper 01 | October 2026

Executive perspective

AI can improve execution while leaving a more consequential question unanswered: are we still pursuing the right outcome? This paper introduces Polaris Shift, a proposed decision framework for examining the assumptions beneath strategy and choosing whether to stay the course, adjust the route, or move the North Star.

The method connects Signal, Shift, Star, and System: what became possible, what changed because of it, what should matter now, and what must be redesigned. Evidence from action feeds the next cycle.

Written for leaders across products, functions, and organizations, the paper combines research context, two fictional decision rehearsals, and a practical review brief. Its central recommendation is to justify the depth of change with evidence, then align resources, incentives, and measures with the chosen direction.

A conceptual framework and practical decision guide

polarisshift.com

01 The strategic decision

Every strategy depends on assumptions about who needs what, where value is created, and which capabilities make an organization useful. A stable direction helps people coordinate around those assumptions. The difficulty comes when operating performance improves while the assumptions become less credible.

A team can produce more reports without helping leaders make better decisions. A service can resolve requests faster while generating more repeat contacts. A digital business can increase engagement while customers would prefer to finish the job with fewer interactions. These are possible measurement failures, not automatic proof that the strategy is wrong.

Separate purpose from its current expression

In Polaris Shift, a North Star is the guiding outcome used to make strategic trade-offs. A North Star metric is a measurable proxy for progress toward that outcome. The distinction matters: a metric can become misleading while the customer outcome remains valuable.

Layer	Question it answers
Purpose	Why does this work matter and to whom?
North Star	What outcome should guide our trade-offs?
Strategy	Which choices will help us achieve that outcome?
Operating system	How do work, resources, and decisions support those choices?
Measures and guardrails	What shows progress, and what must not deteriorate?

A product, function, business unit, and enterprise can each have a guiding outcome. State the scope before proposing a shift. A team may need to change its own outcome while the enterprise purpose remains intact; a local decision should not silently redefine the whole organization.

Reorientation and rewiring

Reorientation asks which outcome and strategic choices should guide the organization now. Rewiring changes the workflows, capabilities, decision rights, resources, and measures needed to support that direction. Clarifying direction should precede broad commitments to redesign, while small experiments can supply the evidence needed to choose.

The practical starting question is: which assumption must remain true for our present direction to deserve continued investment?

02 What the evidence supports

AI capability, workplace performance, and strategic change are different claims. A demonstration can establish that a task is possible. A field study can test performance in a particular setting. Neither alone establishes that a different strategic outcome is warranted.

Benefits vary by task and setting

In the published version of Generative AI at Work, Brynjolfsson, Li, and Raymond study 5,172 customer-support agents. Access to AI assistance increased issues resolved per hour by 15 percent on average, with substantial differences across workers. Less experienced and lower-skilled agents benefited more. This is evidence about a specific deployment, not a general productivity rate for all AI initiatives. [3]

Dell'Acqua and colleagues examine 758 knowledge workers in a preregistered experiment. They describe a jagged technological frontier: AI improved performance on tasks within its capabilities but reduced correctness on a selected task outside that frontier. The important implication for this paper is that nearby tasks can require different human and AI arrangements. [4]

Counterevidence also matters. METR's early-2025 randomized study found that 16 experienced open-source developers took 19 percent longer on assigned tasks when AI tools were allowed. That result describes those tools, developers, and repositories. In a February 2026 follow-up, METR reported that selection and measurement problems made its newer estimates unreliable as a measure of the current effect. The historical slowdown should therefore not be treated as a verdict on present-day coding tools. [5, 6]

These studies are not directly comparable estimates of one universal effect. Their settings, designs, tools, and participants differ. Together, they support testing local outcomes and preserving the conditions under which an observation was made.

The connection to established strategy research

Polaris Shift builds on established questions rather than claiming to invent adaptation. Teece's work on dynamic capabilities addresses sensing, seizing, and reconfiguring as organizations respond to change. Porter's Five Forces examines industry structure and the distribution of economic value. [1, 2]

The contribution proposed here is a practical sequence that connects an observed change to an assumption, an explicit decision about direction, and an operating response. That sequence has not yet been tested as a distinct intervention. The cited research motivates the questions; it does not validate the framework.

03 Read the signal and diagnose the shift

Signal, Shift, Star, and System organize four different questions. Keeping them separate makes it harder to mistake an impressive capability for a strategic conclusion.

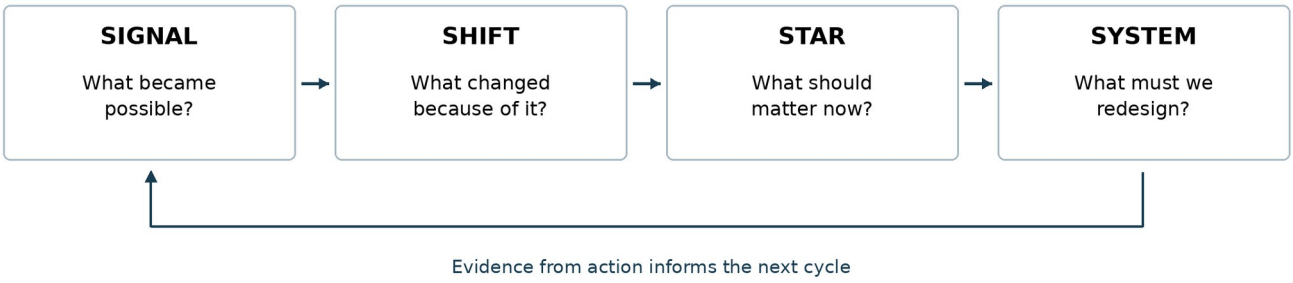


Figure 1 The proposed Polaris Shift sequence and evidence loop

Signal asks what became possible

In an AI context, a signal may be a newly credible capability, a material change in cost or quality, or evidence of different customer behavior. Pair external developments with internal contradictions: a faster process that still disappoints customers can be as relevant as a new model release.

Record the observation, source, date, affected population, and limitations. Distinguish an announcement from demonstrated performance, a pilot from sustained adoption, and stated enthusiasm from actual behavior. A weak signal can justify investigation without justifying deployment.

Shift asks what changed because of it

Translate possibility into a hypothesis about behavior, economics, competition, distribution, or the production of value. Name the inherited assumption that might no longer hold. For example: if buyers delegate routine comparison, our assumption that they will browse every category page may weaken.

Use existing strategy tools where they help. AI might lower an entry barrier, create a substitute, change buyer expectations, or increase dependence on a supplier. These are proposed applications of Porter's framework, not evidence that all five forces have changed in every industry. [2]

Test structural significance through persistence, materiality, observed behavior, and corroboration. Look for alternative explanations such as seasonality, channel mix, or a temporary promotion. Ask what evidence would weaken the interpretation. No single score can replace that judgment.

The output of this stage is a bounded statement: because we observed X, assumption Y may be weakening, which could affect outcome Z. Preserve the uncertainty until the evidence earns a stronger conclusion.

04 Choose the star and align the system

Star asks what should matter now

Decide whether the guiding outcome remains appropriate at the stated scope. A new tool can justify improvement. A changed customer journey can justify redesign. Moving the North Star requires an argument about the outcome itself.

Decision	What stays stable	Likely response
Stay	The guiding outcome and core approach remain credible	Improve execution and continue testing assumptions
Adjust	The guiding outcome still holds	Redesign the route, capabilities, workflow, or measures
Shift	Purpose may endure, but the guiding outcome needs revision	Reorient, then redesign the system around the new outcome

These categories describe the depth of change, not its prestige. An adjustment can be operationally substantial. A shift can begin in one team. State what will remain stable, what will change, and why a less disruptive response would be insufficient.

System asks what must be redesigned

Translate the choice into resource allocation and everyday decisions. Identify an accountable owner, the first bounded action, required capabilities, and what will stop or receive less attention. Reconcile incentives with the chosen outcome. A decision service rewarded only for document volume will continue producing documents.

Define a baseline and evaluate outcomes alongside costs and quality. Depending on the setting, guardrails may include accessibility, privacy, customer control, error severity, repeat work, and employee workload. Specify who can pause an experiment and how affected people can seek human review.

For AI systems, this strategic review should connect to an appropriate risk-management process. NIST's AI RMF 1.0 organizes risk work through Govern, Map, Measure, and Manage, with governance spanning the other functions. Polaris Shift does not replace that work or establish compliance. [7]

Evidence closes the loop

Set a review date and conditions for expanding, revising, or stopping the response. Compare the expected change with what actually happened. A result that contradicts the original interpretation is useful evidence, not a reason to quietly change the success criterion. Feed that learning into the next Signal cycle.

05 Account for the cost of reorientation

Changing too slowly can be costly. So can changing direction before the evidence warrants it. Two proposed Polaris Shift concepts help make those costs discussable without pretending they are already measurable indices.

North Star Debt

North Star Debt is the accumulated cost of continuing to optimize yesterday's definition of success after relevant conditions have changed. It is an interpretive concept, not an accounting liability or a validated performance measure.

It may appear as resources committed to a declining source of value, incentives that preserve an outdated behavior, or dashboards that reward activity disconnected from the intended outcome. Better execution can conceal the problem when the dashboard measures the old definition particularly well.

Diagnose the claimed cost rather than assigning a label after disappointing results. Identify the allegedly outdated assumption, when contrary evidence became available, the decisions it should have affected, and plausible alternative explanations. Opportunity costs are especially uncertain because the unchosen alternative cannot usually be observed directly.

Organizational Mass

Organizational Mass is a working hypothesis about how much interconnected structure must move when direction changes. Relevant factors may include legacy systems, contractual commitments, governance dependencies, customer expectations, specialist knowledge, and entrenched incentives.

Size alone is an inadequate explanation. A small company can be constrained by one platform or customer. A large enterprise can have resources, trusted relationships, and distribution that make a chosen response powerful. The question is which dependencies increase the effort and time required for this specific reorientation.

Make the transition part of the decision

Compare the cost of preserving the present direction with the cost of transition. Examine reversibility, disruption to current customers, required capabilities, and how evidence will arrive during the change. A staged commitment may preserve options while uncertainty is high; an irreversible decision requires a stronger justification.

The governing outcome and the transition plan should be reviewed together. Recognizing a valuable new direction does not prove that the organization can reach it at acceptable cost. Equally, organizational difficulty does not make an outdated direction strategically sound.

06 Rehearse the decision

The following fictional scenarios illustrate distinctions in the framework. They are not client cases, research findings, or claims about an employer.

Adjust the route while preserving the outcome

An internal analysis team aims to help operating leaders make sound inventory decisions. It measures report completion and turnaround time. AI accelerates report production, but leaders still delay decisions because recommendations conflict and key assumptions are unclear.

The signal is the combination of improved output and unresolved decision friction. The challenged assumption is that faster reporting reliably improves decisions. The guiding outcome remains useful, so the team chooses Adjust: replace some recurring reports with concise decision briefs, explicit assumptions, and review of consequential trade-offs.

The team pilots the approach with a limited group. It evaluates decision lead time alongside decision quality, later reversals, unresolved exceptions, and the effort required from users. It compares results with its previous approach and records contextual differences. A faster decision that creates more costly errors would not count as success.

This is workflow and measurement redesign. The ability to generate reports faster did not, by itself, require a new North Star.

Shift the guiding outcome within a defined scope

A product-discovery team currently optimizes for helping shoppers explore relevant options. Its North Star metric is qualified product explorations. Assume that repeated customer research and observed use now show a segment delegating selection and preferring a short, dependable recommendation to extended browsing.

The challenged assumption is that more useful exploration is the best expression of value for this segment. The team proposes shifting its guiding outcome to helping shoppers make a confident, suitable choice with appropriate control. Its broader purpose of helping customers find suitable products remains stable.

The system response might include preference clarification, explanations, comparison of a few alternatives, and an easy route to browse independently. Evaluation would examine successful choices and later regret or returns, alongside accuracy, accessibility, and customer control. Confidence alone would be an insufficient proxy for suitability.

The label depends on the original scope. If the team's North Star had already been successful customer choice, reducing browsing would be an Adjust response. Changing a metric does not automatically constitute a Shift. If the new behavior proves narrow or temporary, the team should retain the broader discovery approach rather than generalize from that segment.

07 Run a practical North Star review

Start with one consequential decision and one assumption that can be examined. Invite the decision owner, people doing the work, and people who experience the outcome. Use the prompts below as a brief, not a new layer of governance.

1 Define the decision and its scope

Name the product, team, function, or business under review. Write the present guiding outcome and the choice that must be made. Identify who has authority to make it.

2 Record the signal and source

What changed, when was it observed, and for whom? Link the evidence. Mark whether it is an announcement, observation, experiment, or inference, and record its limitations.

3 Identify the assumption at risk

What must remain true for the present direction to work? Explain how the signal challenges it. Include at least one plausible alternative explanation and evidence that would change the interpretation.

4 Choose the depth of response

Choose Stay, Adjust, or Shift. State what remains stable, what changes, and what the decision gives up. If proposing Shift, explain why a change in method or measure alone would not address the problem.

5 Design the first test

Specify the smallest useful action, the baseline or comparison, the owner, and the resources required. Define the expected benefit and unacceptable deterioration. Preserve a path to stop or reverse the action where feasible.

6 Schedule reconsideration

Record when the decision will be reviewed and what evidence would justify expansion, adjustment, or stopping. Review on a cadence suited to the stakes and rate of relevant change; avoid a universal schedule.

Keep the disagreement visible

A useful brief preserves contested assumptions rather than manufacturing consensus. Record minority interpretations and unresolved questions. Separate the decision to learn from the decision to scale. A team may reasonably authorize a test while withholding commitment to a new strategic direction.

The desired output is a decision that can be explained, tested, and reconsidered. Completing the brief is not itself evidence that the decision improved.

08 Limits and a research agenda

Polaris Shift is a conceptual framework and practical decision aid. This paper introduces no original empirical study, validated scoring method, or evidence that using the framework causes better organizational outcomes. The fictional rehearsals illustrate reasoning rather than demonstrate effectiveness.

Where judgment remains necessary

The boundaries between Stay, Adjust, and Shift depend on scope and on how the current North Star is defined. Broad outcomes can remain stable while substantial strategic change occurs underneath them. The framework should expose that distinction rather than exaggerate novelty by relabeling every redesign as reorientation.

Formal review can also create costs. It may slow straightforward improvements, encourage excessive reconsideration, or provide a vocabulary that rationalizes a preferred decision. Use the lightest version that helps, and compare its burden with the value of the question being examined.

Uncertainty cannot always be resolved before action. Some evidence arrives only after commitment; customer behavior and technology can change during evaluation. The discipline is to make the uncertainty and consequences explicit, while preserving options where possible.

Questions worth testing

The first empirical question is whether the framework helps teams distinguish a changing outcome from a changing method. Structured decision exercises could compare teams using the framework with teams using their existing review process, with assessors unaware of the condition evaluating the reasoning.

Field applications could then track assumptions, decisions, resources committed, time to reconsideration, and realized outcomes over time. Evaluation should include failed and abandoned applications, adjustment for differences in context, and the effort required to use the framework. Participant satisfaction should be separated from decision quality.

North Star Debt and Organizational Mass require additional work on definitions and observable indicators before numerical scores or predictive claims would be defensible. They should remain research questions until that work supports stronger conclusions.

The leadership practice

A useful North Star provides continuity while leaving its assumptions open to examination. The proposed practice is to investigate meaningful signals, diagnose what changed, choose the appropriate depth of response, and align the system with that choice.

Sometimes that means improving execution. Sometimes it means redesigning the route. When evidence challenges the guiding outcome, it may mean moving the star. The responsibility is to explain which response is warranted and remain willing to learn from what follows.

09 References and author

[1] Teece, D. J. (2007). Explicating dynamic capabilities: The nature and microfoundations of (sustainable) enterprise performance. Strategic Management Journal, 28(13), 1319-1350.
doi.org/10.1002/smj.640

[2] Harvard Business School, Institute for Strategy and Competitiveness. The Five Forces. Institutional explanation of Michael Porter's framework.
www.isc.hbs.edu/strategy/business-strategy/Pages/the-five-forces.aspx

[3] Brynjolfsson, E., Li, D., and Raymond, L. (2025). Generative AI at Work. The Quarterly Journal of Economics, 140(2), 889-942. Figures in this paper use the published version, rather than the earlier working paper.
doi.org/10.1093/qje/qjae044

[4] Dell'Acqua, F., et al. (2026). Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of Artificial Intelligence on Knowledge Worker Productivity and Quality. Organization Science, 37(2), 403-423. First circulated as a working paper in 2023.
doi.org/10.1287/orsc.2025.21838

[5] Becker, J., Rush, N., Barnes, B., and Rein, D. (2025, July 10). Measuring the Impact of Early-2025 AI on Experienced Open-Source Developer Productivity. METR.
metr.org/blog/2025-07-10-early-2025-ai-experienced-os-dev-study/

[6] Becker, J., Rush, N., Cunningham, T., Rein, D., and Mahamud, K. (2026, February 24). We are Changing our Developer Productivity Experiment Design. METR. Read alongside the original study for its follow-up limitations.
metr.org/blog/2026-02-24-uplift-update/

[7] National Institute of Standards and Technology (2023). Artificial Intelligence Risk Management Framework, AI RMF 1.0. NIST AI 100-1. This paper refers to version 1.0; it does not claim alignment with any later revision.
doi.org/10.6028/NIST.AI.100-1

About the author

Darko Tushev is an AI product and strategy practitioner and the creator of Polaris Shift, an independent body of thinking, research, and practical tools for navigating strategic change. This paper presents his independent work and does not represent an employer or endorse a technology vendor.

Polaris Shift | Darko Tushev

Sources reviewed October 2026. Suggested citation: Tushev, D. (2026). When to Move Your North Star: A framework for strategic reorientation in the age of AI. Polaris Shift, White paper 01.