

The Trilemma Test — A Pre-Scenario Diagnostic for Structural Constraints

SCENARIO PLANNING METHODOLOGY SERIES

Mike Edholm · edholms.com

There is a step scenario planning usually skips: identifying whether a system's own stated objectives are actually compatible before building scenarios around them. This essay develops that step into a transferable diagnostic, the Trilemma Test, using a 2008 scenario exercise conducted in Burundi for the Swedish Trade Council as its origin case. That exercise produced three scenarios, Doubt, Dogma, and Trust. What made it methodologically distinctive, though, was not the scenarios themselves. It was identifying, in advance, why those three were the right scenarios to build. The underlying constraint was a genuine incompatibility between security, social cohesion, and economic efficiency. The essay argues this constraint-identification step applies well beyond Burundi. It tests that claim against a current case, artificial intelligence, where capability, safety and security, and speed of deployment appear to form the same kind of trilemma. Scenarios built on a correctly identified constraint are more durable than ones built without it, because they describe the shape of a trade-off rather than the content of a surprise.

I.

There is a step that scenario planning methodology describes imprecisely and practitioners perform inconsistently. It comes before the driving forces, before the critical uncertainties, before the axes are drawn and the scenario spaces named. It is the step in which someone in the room asks: what is this system actually trying to do, and are those things compatible?

Most scenario exercises skip it. The question feels preliminary, almost managerial, the kind of thing executives settle before calling in the foresight team. In practice it is neither preliminary nor settled. It is the most important analytical move the exercise will make, and leaving it unexamined is the single most common reason that scenario sets produce polished narratives

that do not change how anyone thinks.

I came to understand this in Burundi, which is not the place most corporate strategy practitioners expect to find a methodological insight. But the circumstances there made visible something that more comfortable environments conceal.

II.

Burundi, 2008

The assignment was to produce a scenario set for the Swedish Trade Council, evaluating the investment environment in Burundi over a horizon running to 2018. The fieldwork began in 2008, later than originally planned, delayed by the practical complications that anyone who has worked in fragile states will recognize: travel logistics, visa difficulties, access negotiations that take longer than schedules allow. The horizon to 2018 remained. A slightly compressed timeframe, but deemed workable. The country was several years out of a civil war that had killed an estimated 300,000 people over a decade. The Arusha Peace Agreement had held, but barely. The 2005 elections had produced a government, but the government's authority over its own territory was incomplete and contested. The economy was among the poorest on the continent. The regional environment, spanning the Great Lakes region, the East African Community, and the shadow of the Democratic Republic of Congo to the west, was fractious and unpredictable.

The standard scenario approach would have been to identify the critical uncertainties, such as political stability, regional integration, FDI flows, and commodity prices, and build a matrix from the two most important. That approach would have produced serviceable scenarios. It would not have produced the right ones.

Before the scenario work began, several weeks were spent in country. The interviewees included government ministers and their deputies, UN representatives managing the poverty reduction program, World Bank and IMF officials overseeing structural adjustment conditions, and business people at multiple levels, from the small traders who animated whatever informal economy existed to the handful of larger operators trying to build something durable in an environment that had not yet decided what rules it wanted to play by.

What those interviews revealed, collectively, was a society trying simultaneously to do three things that could not all be done at once. The UN's own Poverty Reduction Strategy Paper for Burundi, published in January 2007, named them explicitly: Security, Social Cohesion, and Efficiency. A country at peace with itself, capable of controlling population growth and managing ethnic tension. A society reducing poverty and expanding access to education, healthcare, and basic services. An economy modernized, diversified, competitive, and integrated into re-

gional and global trade.

These were the right objectives. They were also, structurally, in conflict with each other.

III.

The Trilemma

A trilemma is not a list of competing priorities. Competing priorities are familiar and manageable: you rank them, allocate resources accordingly, and accept the trade-offs with reasonable transparency. A trilemma is different in kind. It names a situation where pursuing any two of three objectives simultaneously makes the third materially harder to achieve, not because of resource constraints alone, but because of structural incompatibilities built into the system itself.

The clearest illustration from economics is the impossible trinity: a national government cannot simultaneously maintain a fixed exchange rate, free movement of capital, and an independent monetary policy. It can have any two. Choosing all three does not produce a compromise among them. It produces instability, as a series of currency crises have demonstrated when governments tried anyway. The trade-off is not a political choice about priorities. It is an architectural feature of how the system works.

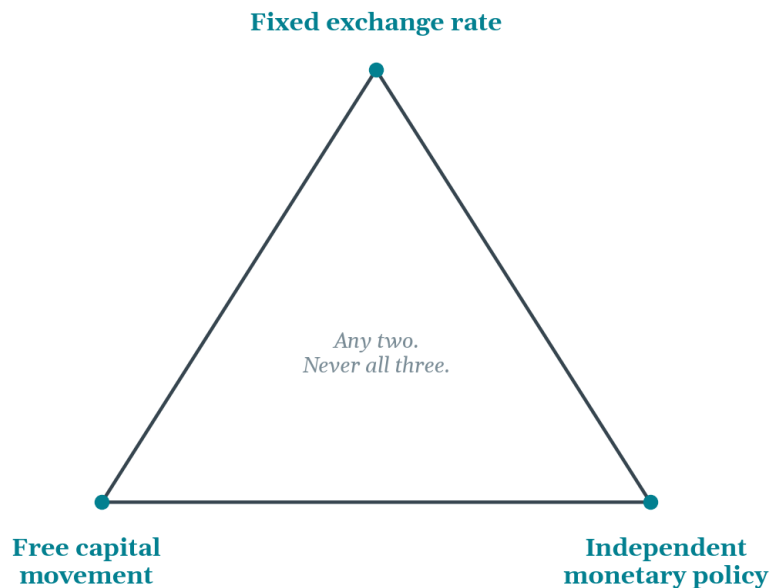


Figure 1. The impossible trinity, the canonical trilemma. A government can hold any two vertices at once; holding all three is not merely difficult but structurally unstable.

Burundi's developmental trilemma had the same structure. Prioritizing Security above all, the logic of a state that has just survived civil war, meant strong central authority, controlled borders, limits on civil society, and tolerance for institutional opacity as the price of stability. This was inimical to Efficiency: a modernized, competitive, externally integrated economy requires transparency, rule of law, property rights, and freedom of capital movement, all of which a security-dominant government finds threatening. Prioritizing Social Cohesion, meaning poverty reduction, access to services, and ethnic reconciliation, required redistribution and inclusive institutions. Those institutions competed directly with the investment climate that Efficiency demanded, and with the centralized control that Security required.

No configuration could satisfy all three. Every plausible Burundian future was, at bottom, a different answer to the same question: which constraint would bind most tightly, and what would follow from that binding?

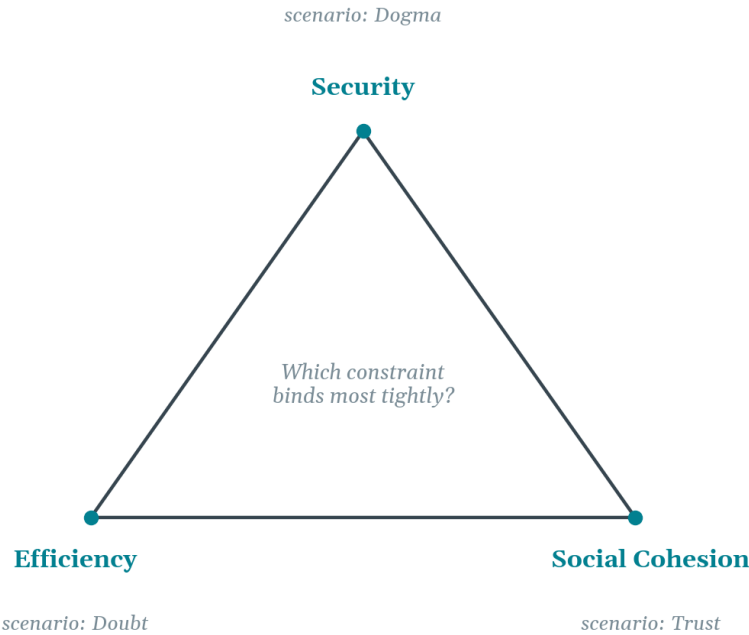


Figure 2. Burundi's developmental trilemma, 2008. Each scenario is the structural consequence of one vertex winning: the scenario space defined itself once the constraint was visible.

Once that question was visible, the scenario space defined itself. Three scenarios, one for each dominant axis. Not three points in an arbitrary matrix, but three structurally coherent end-states, each internally consistent precisely because it had resolved the trilemma in a different direction.

IV.

Doubt, Dogma, Trust

The three scenarios were named Doubt, Dogma, and Trust, chosen not as labels for the objectives but as descriptions of the social and institutional character each dominant logic would produce.

Doubt was the Efficiency-adjacent scenario gone wrong. Not the clean version of efficiency, with transparent markets, clear property rights, and competitive dynamics rewarding track record and innovation, but efficiency captured by regulatory dysfunction. Rules and regulations multiplied across overlapping jurisdictions. Lawyers and accountants ensured compliance with frameworks that conflicted with each other. Transaction costs were high, favoring scale and vertical integration over scope and agility. Capital moved toward short-term arbitrage rather than long-term investment. In business, people did not act. They assigned blame and reacted. Burundi's growth curve would reach considerably below country potential.

Dogma was the Security-dominant scenario carried to its institutional conclusion. National champions declared in key sectors. FDI regulations becoming a protectionist tool. The East African Community losing coherence, Chinese and Indian demand declining, pan-African trade organizations losing influence. Factions gaining control as central government authority weakened. A regional environment perceived as fragmented and polarized. Things moved ever slower as trade decreased and people became isolated. In business, people trusted only those with whom they could identify. Absence of market incentives compounded fragmentation.

Trust was the Social Cohesion scenario achieving its best-case institutional expression. The 2010 election producing a stable majority. Structural reforms successfully implemented. The East African Community and the Great Lakes Region harmonizing their regulatory frameworks. FDI regulations promoting entrepreneurship rather than incumbency protection. Transaction costs falling, innovation and reputation becoming keys to competitive success. The government ruling by incentive rather than by control. Information flowing freely across boundaries. In business, people were proactive and cooperative.

Trust was explicitly identified as the most desirable scenario, for investors, for the government, and for the Burundian population. The ongoing reform initiatives at the time of the exercise pointed toward it. The 2010 election was identified as the single most important near-term trigger: the scenario the election reinforced would likely govern the decade that followed.

What made these three scenarios analytically credible, and not merely serviceable, was not the richness of the narrative development. It was the structural derivation. Each scenario was the logical consequence of a specific resolution of the underlying trilemma, not a creative invention assembled from driving forces. The methodology had discovered its scenarios rather

than constructed them.

V.

The Diagnostic Step

The Trilemma Test, as a transferable method, runs on four questions. Each is short enough to state in a sentence, each is harder to satisfy than it looks, and each builds directly on the one before it.

The first question is objective definition: what does each candidate objective actually require, institutionally, economically, and behaviorally? Not the strategic priorities of the client organization, which are a different thing, but the objectives the system itself is organized around, whether those are explicitly stated or have to be inferred from observed behavior. In Burundi, these requirements were spelled out in the UN's own Poverty Reduction Strategy Paper with a clarity that only a country newly emerged from a decade of civil war could bring to its own diagnosis: Security meant strong central authority and controlled borders; Efficiency meant transparency, rule of law, and free capital movement; Social Cohesion meant redistribution and inclusive institutions. Naming these requirements precisely, before asking whether they conflict, is what keeps the next question from becoming guesswork.

The second question is the pairwise test: for each pair of objectives, what specific mechanism makes the third one harder to achieve, not merely more expensive, but structurally undermined? This is the question that separates a genuine trilemma from an ordinary competition for resources. Burundi's answer was concrete rather than abstract. A security-dominant government's tolerance for institutional opacity was not a matter of degree that more budget could fix; it was the direct negation of the transparency Efficiency required. Redistribution aimed at Social Cohesion did not merely cost money that Efficiency's investment climate would have preferred spent elsewhere; it required inclusive institutions that a centralizing, security-focused state found threatening to its own logic. If the answer to this question is that one objective would simply cost more, the system likely has competing priorities rather than a trilemma, and the standard scenario-planning toolkit is probably the right one after all.

The third question is the stress test: under what shock or scarcity condition does the incompatibility actually become binding? A trilemma can sit dormant for years under conditions of relative slack, when enough resources or goodwill exist to paper over the underlying incompatibility, and an analyst who stops at the pairwise test alone can be lulled into thinking a system has more room than it does. Burundi's incompatibilities were binding constantly, given the country's fragility, but a more resourced or stable environment might mask the same structural conflict until a genuine shock, a fiscal crisis, a security incident, a commodity price collapse,

removes the slack and forces a choice. Naming the stress condition in advance is what turns the trilemma from an interesting observation into an early-warning system.

The fourth question is the scenario implication: which observable, trackable signals would indicate that one objective is becoming dominant over the other two? This is where the test hands off to the scenario work it exists to discipline. In Burundi, the 2010 election was identified in advance as the clearest such trigger. Its outcome would most likely determine which of the three scenarios, Doubt, Dogma, or Trust, the following decade would resemble. A trilemma that cannot name its own observable triggers has correctly identified a structural constraint but has not yet done the work needed to make that constraint useful for planning.

What the Trilemma Test adds to standard scenario methodology is a prior constraint on the scenario space. It does not replace driving-force analysis or critical uncertainties identification. It precedes them, and it disciplines them. Driving forces that do not connect to the resolution of the governing trilemma are secondary at best. Critical uncertainties that do not bear on which trilemma constraint will bind are real uncertainties, but they are not the structural ones. The Trilemma Test separates the structural from the contingent before the scenario work begins.

VI.

Transfer to AI

The Opaque Surprise essay closed by suggesting that trilemma models might offer more durable guidance than narrative scenarios under conditions of opaque surprise, and proposed that question as one of two left open. This essay attempts an answer.

The governing trilemma for AI development, proposed here not as settled analysis but as a diagnostic hypothesis, places three objectives in structural tension: Capability, Safety and Security, and Speed of Deployment.

Capability is what the technology can do: the scope and depth of cognitive tasks it can perform, the range of domains it can operate in, the degree to which it can pursue complex goals over extended sequences of action. The competitive and scientific incentives driving AI development are almost entirely organized around capability advancement. The frontier labs measure themselves against capability benchmarks. Investors evaluate capability trajectories. Governments assess national security implications in capability terms.

Safety and Security covers two related but distinct concerns that current discourse is increasingly treating as inseparable. Safety is the degree to which AI systems behave as their designers intend across the full range of conditions they encounter, including conditions not anticipated during training and deployment. It includes alignment with human values and intentions in edge cases, robustness to adversarial inputs, interpretability sufficient to detect and correct

failures, and the institutional capacity to identify unexpected behaviors before they propagate. Security is the degree to which AI systems, and the infrastructure on which they run, are resistant to hostile exploitation: deliberate manipulation of model behavior, theft of model weights, adversarial use of capability by state and non-state actors. The two concerns are connected. A system whose behavior is not fully understood by its builders is also a system whose behavior under adversarial conditions cannot be fully guaranteed. As the Opaque Surprises essay noted, specific emergent capabilities have repeatedly surprised the labs producing them. Both safety and security work are, in part, attempts to catch up with a capability frontier that understanding of the system's behavior has consistently trailed. The distinction between the two is real and worth preserving analytically, but for the purposes of the trilemma, they pull in the same direction: toward slower, more deliberate development and deployment than competitive dynamics encourage.

Speed of Deployment is the rate at which systems move from research into commercial and institutional use. It is driven by competitive dynamics: any organization that deploys later than its competitors loses market position and, critically, the feedback from real-world use that is itself a driver of capability development. Regulatory environments have generally not slowed deployment meaningfully; in most jurisdictions, the frameworks governing AI deployment either do not yet exist or were designed for earlier generations of the technology.

The structural incompatibilities are not hard to identify. Advancing Capability faster than Safety and Security research can keep pace means deploying systems whose behavior in edge cases and under adversarial conditions is not fully understood. That is the current situation, by the admission of the labs themselves. Prioritizing Safety and Security rigorously enough to resolve the alignment and robustness questions currently open would require slowing capability advancement and deployment to a rate that research could actually track, which the competitive dynamics of Speed of Deployment make nearly impossible without coordinated action that has not materialized. Maximizing Speed of Deployment delivers the feedback and market position that sustain the resources required for Capability and Safety and Security work, but compresses the time available for evaluation before systems are operating at scale in consequential environments.

No configuration currently achieves all three. Every plausible AI trajectory is, at bottom, a different answer to which constraint will bind most tightly, and what follows from that binding.

VII.

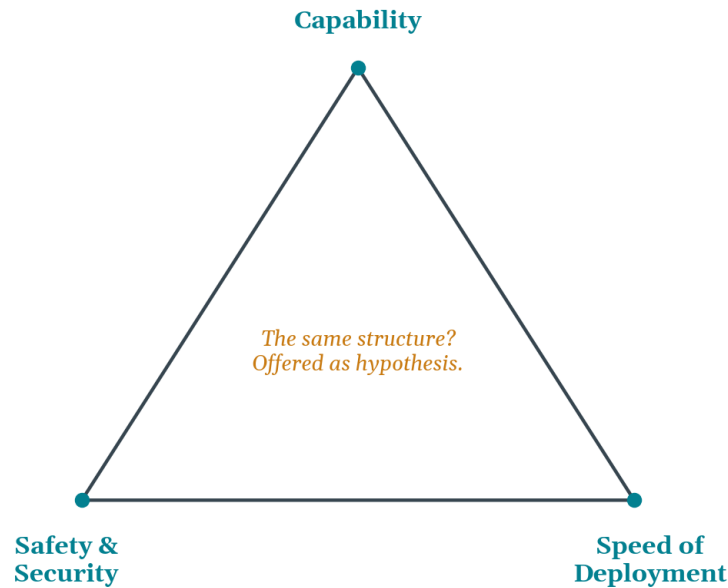


Figure 3. The same test, applied forward. Whether this is the correct three-way framing for AI is deliberately left open; the discipline is identifying the constraint before narrating the scenarios.

Three AI Scenarios

The scenario space the AI trilemma defines has a structure analogous to Burundi's, not because the domains are similar, but because the underlying trilemma logic is the same.

A Capability-dominant scenario is one in which the competitive pressure to advance what AI systems can do continues to outrun both safety research and deployment governance. Emergent capabilities continue to arrive as surprises. Deployment continues at the speed competitive dynamics require. Safety work remains reactive, catching up after capabilities have already been discovered, sometimes after systems are already deployed. The business environment rewards incumbents with the compute scale and data access to stay at the capability frontier. Smaller organizations and most national governments become consumers of capability produced elsewhere, with limited ability to influence its direction or govern its effects. This is not a scenario in which AI development goes catastrophically wrong in a single event. It is a scenario in which the gap between what AI systems can do and what anyone can reliably predict about what they will do widens continuously, without a specific crisis forcing the question.

A Safety and Security-dominant scenario is one in which a forcing event shifts the competitive and institutional equilibrium toward safety research, security hardening, and governance. That forcing event could be a high-profile deployment failure, a serious adversarial exploit of a

frontier system, a coordinated regulatory intervention, or a voluntary pause by the major labs. The triggering event need not be a single dramatic incident. A sustained pattern of smaller failures accumulating public and regulatory pressure could produce the same reorientation. Capability advancement slows, or at minimum is conditioned on evaluations that have not previously been required. Deployment is gated by institutional review processes. The feedback loop that Speed of Deployment provides to Capability research is partially broken. The business environment rewards organizations that can demonstrate interpretability, alignment, and security maturity, rather than those that can simply deploy the most capable systems fastest. Progress becomes more legible, and more negotiable, than in the Capability-dominant scenario, but the competitive position of organizations that cannot afford the overhead of rigorous evaluation deteriorates.

A Deployment-dominant scenario is one in which the competitive dynamics of Speed of Deployment become the organizing principle, but unlike the Capability-dominant scenario, the driver is not frontier capability research. It is the rapid proliferation of existing capability into every available application context, at the fastest rate infrastructure and adoption will permit. The distinction matters. A Deployment-dominant scenario is not primarily about what AI can do. It is about how completely it is woven into consequential institutional processes before those processes have adapted to accommodate it. The business environment in this scenario rewards integration speed and distribution reach over frontier capability. The safety questions are not answered; they are absorbed into operational practice in ways that are difficult to reverse.

These three scenarios are structurally derived, not narratively invented. Each is the logical consequence of a specific trilemma resolution. Each has a different business environment logic, a different competitive dynamic, and a different set of trigger events worth monitoring. The 2010 election was the key trigger for Burundi. For AI, the analogous triggers are the ones that would shift which constraint binds. A major deployment failure or security incident could change public and regulatory tolerance. A breakthrough in interpretability research could make safety evaluation tractable at scale. Or a competitive event, a new architectural approach, a new entrant with a different capability profile, could reshape the deployment dynamics itself.

A note on what the trilemma as framed does not address. Governance, the institutional and regulatory architecture through which AI development is overseen, is a distinct question, and an important one. So is the problem of bias: the ways in which AI systems can encode, amplify, or generate discriminatory outcomes that neither their builders nor their deployers intended. These are not captured by the Capability / Safety and Security / Speed of Deployment frame, because they operate at a different level of the system, less about the structural dynamics of the development race, more about the design choices and institutional accountabilities that shape what systems do once deployed. They deserve their own treatment, and will receive it

in a separate essay.

VIII.

What the Test Does Not Resolve

Two limitations deserve explicit acknowledgment, in the same spirit of methodological clarity the Opaque Surprises essay tried to model.

The first is that the AI trilemma proposed here is a hypothesis, not a finding. The Burundi trilemma was derived from weeks of fieldwork and a close reading of the system's own stated objectives. The AI trilemma proposed above was derived by applying the same diagnostic logic to publicly available information about how the frontier labs describe their own situation. It has not been tested against the same quality of primary evidence. It may be that a different set of three objectives better captures the structural constraint. Access, who gets to use capable AI, on what terms, may create a trilemma with Capability and Speed that Safety and Security does not. Or the safety and security dimensions may simply pull in different enough directions to warrant separate treatment. The hypothesis is offered as a starting point, not a conclusion.

The second limitation is more fundamental. The Trilemma Test is a tool for structuring the scenario space. It does not resolve the opaque surprise problem identified in the Opaque Surprises essay. A trilemma model specified without knowing which AI capability will emerge next is more durable than a narrative scenario that assumes a particular capability profile, because it describes the shape of the trade-off rather than the substance of the surprise. But it does not make the substance of the surprise visible. A Capability-dominant scenario that is structurally coherent still cannot specify which emergent capabilities will arrive, in what sequence, or with what consequences for any particular domain. The model frames the constraint. It does not narrate the content.

This is not a failure of the model. It is a feature of the condition. The appropriate response to opaque surprise is not a better narrative but a clearer structure, one that says clearly what it knows and what it does not. The Trilemma Test offers exactly that: a structure derived from what is knowable about the system's governing constraints, without pretending to know what the system's emergent behavior will look like. That is the most durable form of scenario architecture available under these conditions.

Whether it is durable enough is a question the next decade will answer.

Mike Edholm has spent three decades advising corporations and institutions on scenario planning and strategic foresight, with assignments spanning telecommunications, computing, industrial automation, and international development. This essay is part of a series on scenario planning methodology.