ketteQ™ / Quintus™

The CSCO's Guide to Free-Range AI™ for Supply Chain

The CSCO's Guide to Free-Range AI™ for Supply Chain

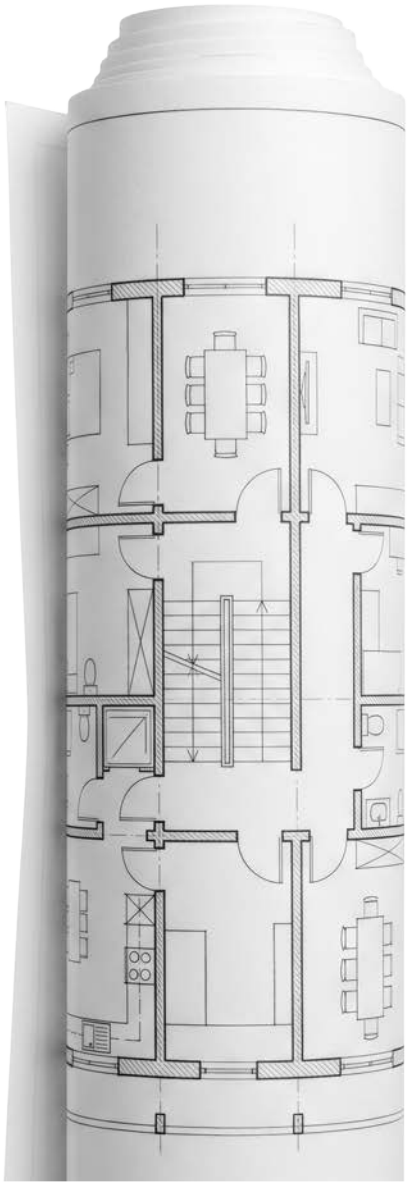
Unscripted. Fully Governed.

By Chris Amet, Chief Technology Officer, ketteQ

Your supply chain plan was already stale the moment you published it. A supplier slipped. A demand signal shifted. A constraint appeared that nobody modeled. This is not a failure of your team or your systems. They are doing exactly what they were built to do, for a world that assumed more stability than supply chains actually have.

This guide is for CSCOs and supply chain leaders asking the harder question: what does a genuinely better operating model look like, and is it achievable without a multi-year disruption? You will see why legacy planning architecture cannot close that gap, what Free-Range AI™ makes possible instead, and how organizations are deploying it in four to eight weeks. The short answer is yes - and it starts with why the planning era is ending, not because legacy systems failed, but because the environment they were built for doesn't really exist.

What Got You Here Is Not What Gets You There



The Architecture of the Planning Era

Legacy planning systems were built on a sensible set of assumptions. Plans would be generated periodically - weekly, monthly, in some cases daily. Scenarios would be evaluated one at a time. Logic would be configured by specialists and locked into proprietary architectures. Human planners would review outputs and make the judgment calls that the system could not.

Supply chains have always been unpredictable. What's changed is the clock speed.

For decades, the pace of disruption was slow enough that a batch-and-review cycle could still catch up. A weekly or monthly re-plan had time to absorb a late shipment, a demand swing, a supplier hiccup, before the gap between plan and reality became unrecoverable. The model wasn't built on stability. It was built on having a margin of time to correct course.

That margin is gone. Lead times compressed, disruptions compound faster than they used to, and by the time a batch cycle catches up, the decision window has already closed. The cost of being wrong hasn't changed in kind. It's changed in how fast it becomes irreversible.

The environment changed. The design did not.

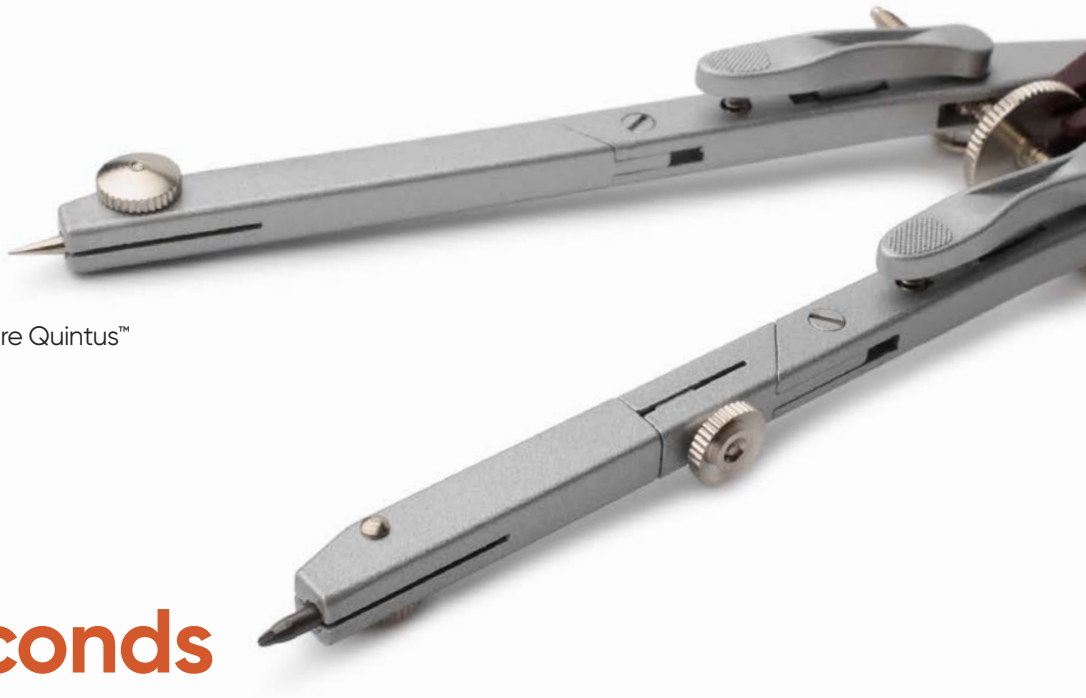
The Gap Between Planning Speed and Reality Speed

The most consequential shift in supply chain management over the past decade is not any single disruption. It is the structural widening of the gap between how fast the world moves and how fast planning systems can respond.

Traditional planning architectures process one scenario at a time. A constrained supply plan that takes four to six hours to generate cannot be regenerated every time a signal changes. So organizations make a practical tradeoff: they plan less frequently, accept more staleness, and rely on their planners to bridge the gap manually.

That manual bridging is not a failure of talent or process. It is the predictable result of deploying a planning architecture against an environment it was not built for.

The cost of that gap is measurable. Industry benchmarks commonly cited by supply chain analysts put it at 3 to 10% of annual revenue lost to stockouts, 1 to 3% OTIF chargebacks, and 8 to 12% expedite premiums for organizations that react instead of anticipate. The cost is not just financial. It is competitive velocity.



4–6 hours

Legacy supply plan generation before Quintus™
(PPF / Partner in Pet Food)

25 seconds

Same supply plan with Quintus™ on top -
same system, no replacement required

Those numbers are not a product claim. They are a production result from a supply chain running Quintus™ on top of an existing planning deployment. The incumbent system did not move. The decisioning layer above it did.

The Planner Capacity Problem

Every planning leader understands this dynamic: the volume of exceptions that require human attention has grown faster than the capacity of planning teams to handle them. The result is a triage culture - planners spending the majority of their time on the loudest exceptions rather than the most strategically significant ones.

This is not a headcount problem. Adding planners does not solve an architecture problem. The architecture is generating exceptions faster than any team can process them, because it was designed for an environment with fewer, slower-moving signals.

> **"Adding planners to a batch planning system is like adding more phone operators to a dial-up network. The constraint is not the people. It is the infrastructure beneath them.**

The question is not whether your planners are good enough. They almost certainly are. The question is whether the system they are working with is giving them any real chance to be strategic.

What Has Not Changed - And Why That Matters

None of this means that the investments your organization has made in SAP IBP, Kinaxis, o9, Blue Yonder, or any other planning platform were wrong. Those systems contain valuable logic, institutional knowledge, and process integration. Replacing them wholesale is a multi-year disruption with enormous implementation risk and cost.

The good news is that replacement is no longer the only path to a meaningfully better outcome.

The architectural advance that changes the calculus is not a better planning system. It is a different kind of intelligence layer that sits above whatever you already run - one that can reason in real time, evaluate thousands of scenarios simultaneously, and act on the results without waiting for the next planning cycle.

That is what Free-Range AI™ means. And it is what Quintus™ delivers.

What Free-Range AI™ Makes Possible Now

Why the Label Matters

Free-Range AI™ is not a marketing phrase. It is a precise architectural description that separates a specific class of capability from everything else being sold as supply chain AI today.

Most supply chain AI is, in practice, limited AI. It is constrained by the architecture it was built inside. The demos are polished because the questions were anticipated. Ask something outside the pre-configured boundaries and the system either fails silently or admits it cannot help. That is not a criticism of the teams that built those systems. It is a structural constraint of bolting AI onto a legacy planning architecture.

Free-Range AI™ is different in a specific, testable way. Quintus™ reasons over any question, executes any action, and addresses any use case - on demand, in real time, with full governance built into every layer. Not because it has better prompts or a larger model. Because the architecture underneath it was built from the start to make that possible.

The Four Requirements That Make It Real

Free-Range AI™ is not a label you can apply to any AI. It requires four things working together. No other planning vendor has all four.

1

Built for Supply Chain

Quintus™ is trained on the complexity of supply chain data and decisions, not general-purpose AI retrofitted for the category. An AI that does not already understand supply chain cannot reason over it, no matter how fast or flexible its architecture is. This is the entry qualifier: it does not replace the other three requirements, it is the reason the other three are possible in a supply chain context at all.

2

Adaptive, End-to-End

Quintus™ reasons across the full data chain, customer to supplier, not siloed by module or trapped in one moment. Pre-AI solutions have static logic and inaccessible data that restrict what AI can do. Quintus™ is able to fully rewire its logic in real time as conditions change.

When another planning vendor adds AI to a legacy planning system, the AI is siloed by module and trapped by static logic. Quintus™ reasons across the whole chain and adapts as conditions change.

3

PolymatiQ™ - The Only Solver Fast Enough

Free-Range AI™ in a live supply chain conversation requires a solver that operates at conversation speed. PolymatiQ™, ketteQ's patent-pending agentic solver, explores the solution space and returns the optimal answer in seconds.

Every other planning engine runs one scenario at a time. A solver that takes hours per run cannot power a real-time AI conversation. It can only report on what was already computed hours ago. PolymatiQ™ is not a faster version of the same approach. It is a different approach entirely.

4

Deployed Anywhere, On What You Already Run

Quintus™ and PolymatiQ™ run above any ERP, any planning platform, any stack you already have - SAP IBP, Kinaxis, o9, Blue Yonder, Logility - without requiring replacement, or as ketteQ's own Planning Solutions and Execution Solutions for organizations ready to replace spreadsheets or systems that have stopped scaling. Customers go live in four to eight weeks.

Every other planning vendor's AI is free-range only inside the cage they built. Their agents can reason over their own system with impressive results. The moment the decisioning surface extends to a rival's platform, they are structurally limited by their own business model. ketteQ has no stack to protect. That is not a feature. It is the business model.

> **Their AI is caged. Ours is Free-Range.**
> **Ask Quintus™ any supply chain question, in real time, with no limits on what you can ask and full governance built into every answer.**

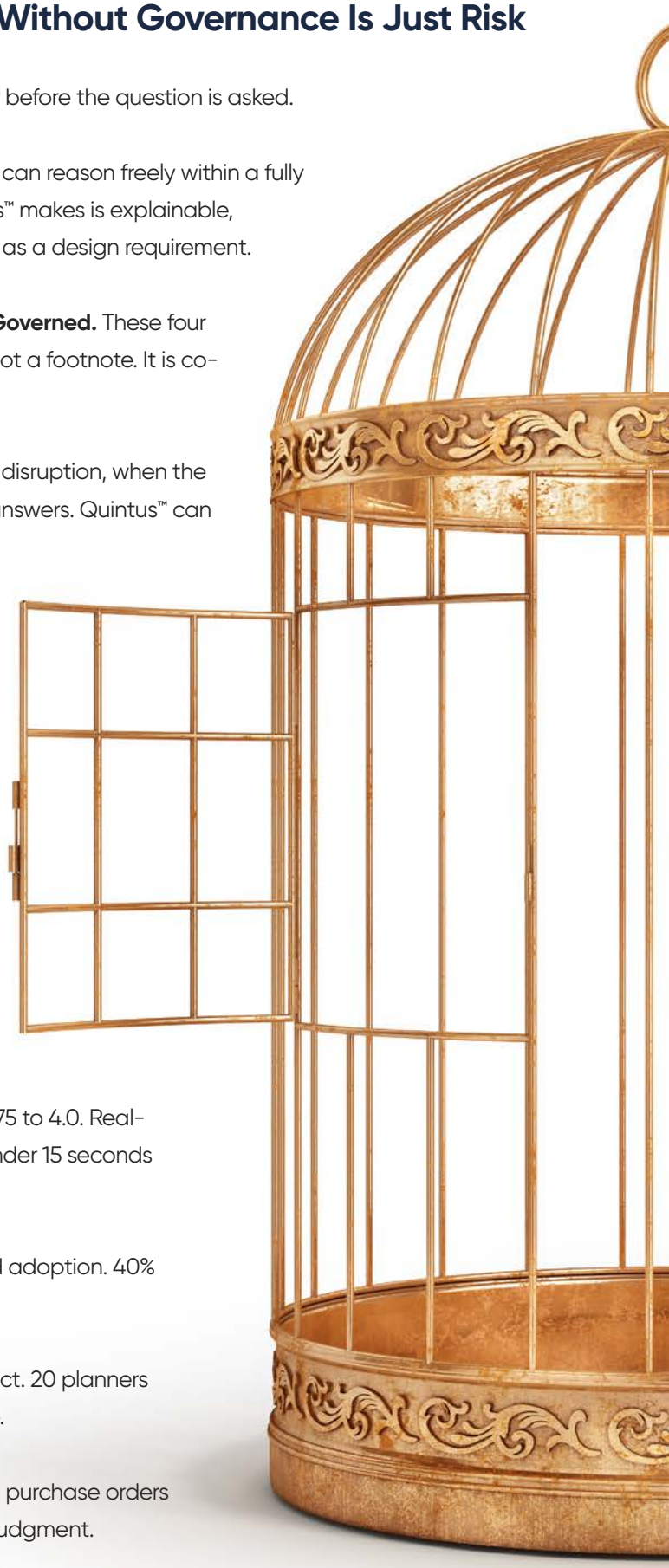
Fully Governed - Because Free-Range Without Governance Is Just Risk

The governance concern is real, and it deserves a direct answer before the question is asked.

Free-Range AI™ does not mean ungoverned AI. It means AI that can reason freely within a fully governed, auditable decision framework. Every decision Quintus™ makes is explainable, auditable, and subject to human override. Not as a disclaimer - as a design requirement.

The standard pairing is non-negotiable: **Free-Range AI™. Fully Governed.** These four words travel together in every context because governance is not a footnote. It is co-equal to the freedom claim.

A limited AI is not a safe AI. It is a useless AI - especially during a disruption, when the questions nobody anticipated are exactly the ones that need answers. Quintus™ can answer those questions. Their AI cannot. And Quintus™ does it with a full decision audit trail.



What This Looks Like in Production

These are not projections. They are results from organizations running Quintus™ today.

- **Partner in Pet Food** deployed Quintus™ above its existing planning system. Supply plan generation: 4 to 6 hours down to 25 seconds. 13% capacity utilization improvement. $8M projected annual cost savings. 100% planning productivity improvement.
- **Alliance Consumer Group**: inventory turns improved from 2.75 to 4.0. Real-time available-to-promise inside Salesforce, answered in under 15 seconds while a sales rep is on a live call.
- **MobilityWorks**: 7,500-plus autonomous appraisals. 83% field adoption. 40% increase in WAV trade-ins.
- **Coca-Cola Bottlers Japan**: $100M projected business impact. 20 planners expanding to 1,500 field users with supply chain intelligence.
- A single overnight production run: 79 automated actions, 32 purchase orders processed, 20 supplier emails actioned, 5 requiring human judgment. Autonomously. While your team was sleeping.

How to Deploy It in Weeks, Not Years

The Disruption You Are Not Required to Accept

The most common reason supply chain leaders defer modernization is not conviction that the status quo is good enough. It is fear of the disruption that modernization requires.

That fear is historically well-founded. Large-scale planning implementations have a documented track record of multi-year timelines, budget overruns, and value that arrives long after the business case was built. The people who proposed those projects often no longer work at the company by the time the system goes live.

Quintus™ is architected specifically to break that pattern. Not by doing a smaller version of the same thing. By doing something architecturally different.

Augmentation First - Replacement When Justified

The primary deployment path for Quintus™ is augmentation. Quintus™ deploys as the intelligence layer above your existing planning systems. Your SAP IBP, Kinaxis, o9, or Blue Yonder installation stays exactly where it is. The planners who know it stay exactly where they are. The integrations, the data flows, the governance structures - all of it remains in place.

What changes is the decisioning surface above it. Quintus™ can now reason over any question in real time, run thousands of scenarios through PolymatiQ™, and surface recommendations or take actions - all with full auditability and human override at every step.

For organizations running spreadsheets or genuinely outdated planning tools, ketteQ's Planning Solutions and Execution Solutions offer a complete end-to-end platform, powered by Quintus™ and PolymatiQ™. Replacement is available when it is justified. It is not assumed.

What Four to Eight Weeks Actually Looks Like

The four to eight week deployment timeline is not a sales commitment. It is an architectural result. Quintus™ is built on Salesforce-native infrastructure with enterprise-grade security, SOC 2 Type II compliance, and automated tuning that eliminates months of manual configuration.

Traditional planning implementations fail to deliver fast because they require the replacement of existing systems, retraining of users, redesign of processes, and deep ERP reintegration. Quintus™ avoids all of those prerequisites by design.

The progression is practical and governed:

1-2 3-4 5-6 8+

Weeks 1 to 2
Connect Quintus™ above your existing stack. No replacement. No retraining.

Weeks 3 to 4
Quintus™ begins sensing signals, running scenarios through PolymatiQ™, and surfacing recommendations.

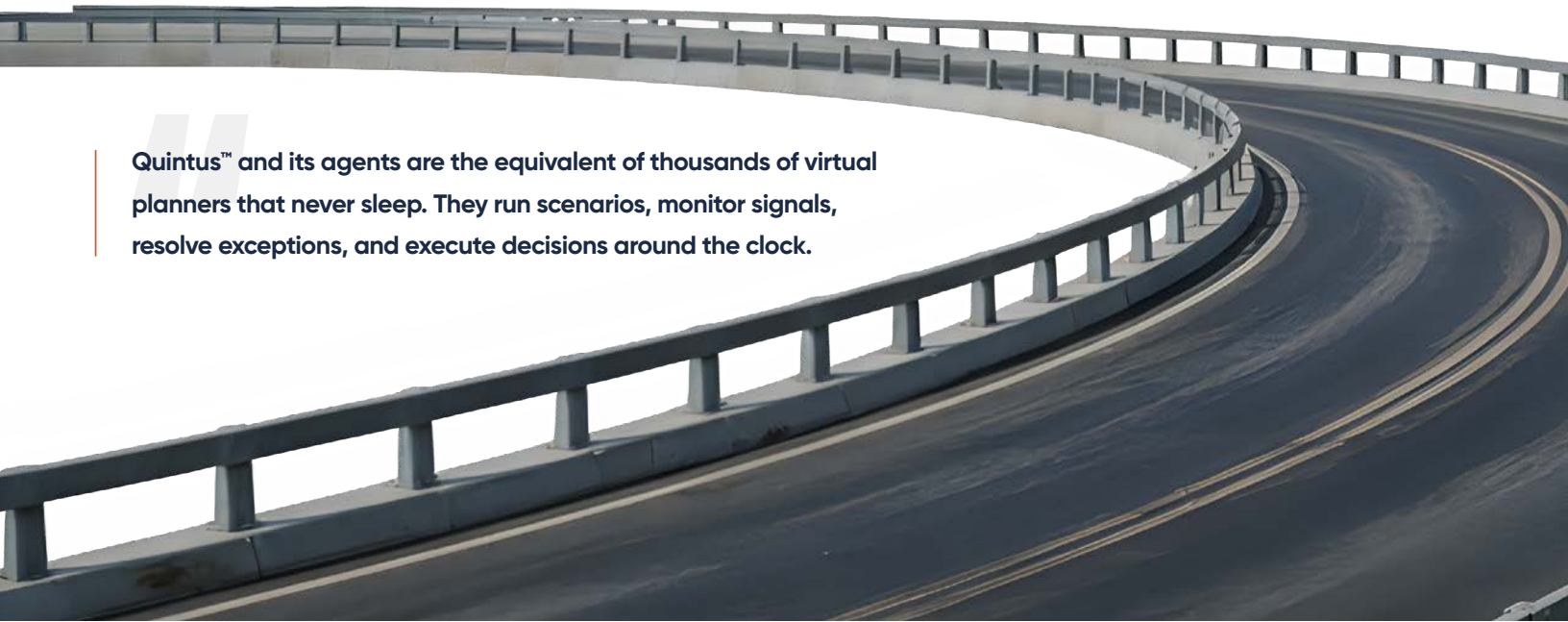
Weeks 5 to 8
Measurable improvements in forecast accuracy, planner workload, and exception handling. Quintus™ is learning the organization. The agents are getting smarter.

Beyond week 8
Autonomous missions run continuously. The intelligence compounds. Every decision Quintus™ makes improves the next one.

The Compounding Advantage

One of the structural differences between Quintus™ and any point-in-time planning improvement is that Quintus™ learns. Skills become automated missions. The longer Quintus™ works with an organization, the more institutional knowledge it absorbs - and the smarter every agent gets.

This is the compounding advantage that traditional planning systems cannot replicate. A static planning system does not get better at your supply chain over time. Quintus™ does. The gap between an organization running Quintus™ and one that is not widens continuously.



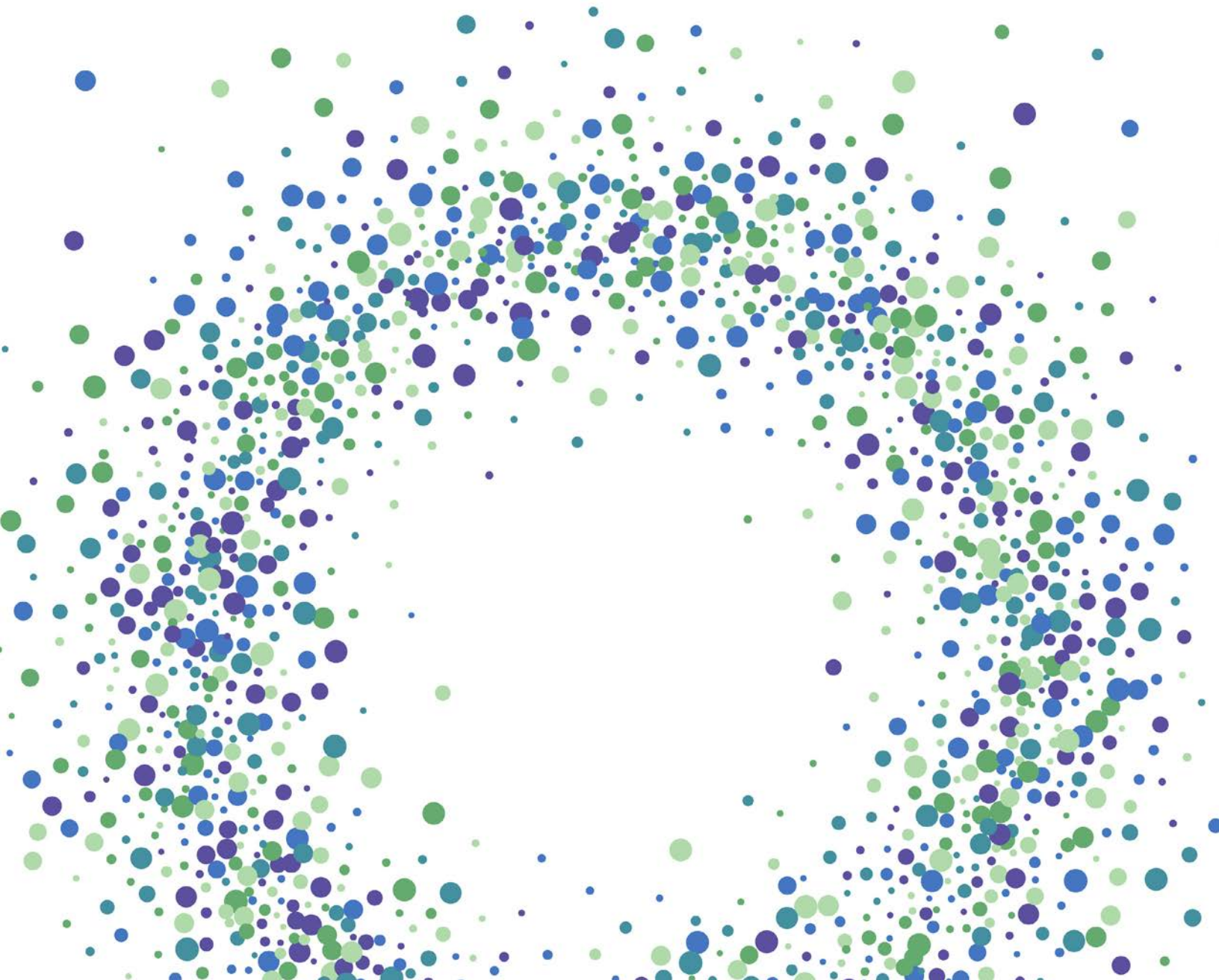
> **Quintus™ and its agents are the equivalent of thousands of virtual planners that never sleep. They run scenarios, monitor signals, resolve exceptions, and execute decisions around the clock.**

Beyond Supply Chain - Where the Intelligence Radiates

Quintus™ is a supply chain solution. That is where it starts and where it plants its flag. But supply chain touches every function that depends on it - and when supply chain intelligence runs in real time, the benefits radiate outward automatically.

Sales teams get real-time available-to-promise answers inside Salesforce while they are on live calls - not after checking back with supply chain hours later. Customer service teams get supply signal intelligence directly in the workflow they already use. The CFO, CEO, and operations head can ask Quintus™ a supply chain question in plain language from Microsoft Teams or Slack and get a fully reasoned answer in seconds - without logging into the ketteQ platform.

This is not an enterprise platform play. It is the natural result of supply chain intelligence that actually runs in real time.



The Plan That Runs Itself Is Not a Vision. It Is Quintus™.

The supply chain planning era produced extraordinary value for a world that assumed more stability than supply chains actually have. That era is not ending because it failed. It is ending because the environment it was built for no longer exists.

The orchestration era is different in a specific, measurable way. Plans adapt continuously rather than going stale the moment they are published. Scenarios are evaluated in seconds rather than hours. Agents execute overnight actions autonomously while your team is focused on strategy. And the intelligence compounds - getting smarter about your supply chain the longer it runs.

Quintus™ is the only Free-Range AI™ for supply chain. The only one built for supply chain from the ground up. The only one adaptive end-to-end, rewiring its logic in real time. The only one with PolymatiQ™, the agentic solver exploring the entire supply chain solution space in seconds. The only one that deploys above any stack you already run, or as ketteQ's own Planning and Execution Solutions, in four to eight weeks, with no rip and replace.

Fully governed. Fully auditable. Subject to human override at every step.

Organizations running Quintus™ today are seeing it: supply plans generated in seconds instead of hours, inventory turns improving from 2.75 to 4.0, revenue-at-risk decisions made in real time instead of discovered after the fact. This is not a future state. It is what is available to model against your own supply chain right now.

The question is not whether Free-Range AI™ will change supply chain planning. It already is. The question is whether your organization is running it.

Scan Here

Request a demo

Explore how Free-Range AI™ is helping supply chain leaders move beyond static planning, deploy intelligent agents in weeks, and improve performance without replacing existing systems or sacrificing governance.

ketteQ Quintus™ Free-Range AI™