



EBOOK

Funding AI with savings

How to underwrite GenAI with cloud, data, and process efficiencies

Executive summary

AI adoption is arguably the biggest strategic imperative for enterprises today. Across industries, enterprise leaders are facing increased pressure to operationalize AI and unlock its wide spectrum of benefits. Despite working in diverse sectors, these leaders share a few common AI challenges.

First, IT budgets are already bursting at the seams, making it difficult to allocate funds for new AI initiatives. Second, far too many AI projects simply collapse, indicating that companies aren't choosing optimal use cases for AI.

Establishing new budgets for AI pilot projects is out of the question for many businesses. What they need is a complete revision of how to fund these initiatives. The objective should shift from procuring new money to maximizing and reinvesting savings from cloud, data, and process efficiencies.

In other words:

It's time to focus on cloud optimization, platform rationalization, and automation, and use those financial gains to fuel AI. Suffice it to say, the bedrock for this approach is choosing the right use cases for AI adoption.

This white paper offers a detailed and actionable guide to this funding-by-savings AI strategy. It contextualizes AI adoption, explores case studies, and offers a 12-month roadmap to scale this innovative funding-by-savings approach. The promise? A 90-day pathway to first value, with robust governance and measurable ROI.



1. Context: AI pressure, realism, and budget constraints

AI adoption is in full swing. According to Gartner¹, almost 60% of CEOs view AI as the technology that will radically transform their sector. McKinsey² substantiates that claim, finding that 88% of companies have implemented AI for at least one business function, and 20% use it in as many as five business functions.

It goes without saying that AI isn't a productivity hack anymore. It's mission-critical. But the road to AI adoption is opaque, largely due to market noise and undefined executive demand for outcomes.

While corporate decision-makers are fully committed to implementing AI at scale, the big-bang programs they often opt for don't yield results.

Besides poor results, these initiatives surface the following problems:



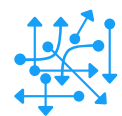
Scope creep:

Project roadmaps, milestones, and objectives often lack clarity, eventually losing focus.



Infrastructure bottlenecks:

Hasty AI adoption exacerbates the uncontrolled proliferation of hybrid infrastructures, creates compute and storage demand spikes, complicates system integration, and creates silos across data and toolstacks.



Unclear ROI:

KPIs and metrics are often misaligned, making it impossible to measure the impact of AI initiatives accurately.

1. <https://www.gartner.com/en/webinar/639023/1409751>

2. <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai>

While big-bang approaches are often detrimental to sustainable AI adoption, there's another concern: Most enterprises don't even know how they want to leverage AI.

Choosing the wrong use case dooms AI projects from the get-go, with Gartner³ reporting that more than 40% of agentic AI initiatives will crash before 2027—a consequence of the costs, risks, and misdirection that come with an improper AI use case.

In this AI landscape, how can business leaders optimize AI adoption? The preliminary step is choosing the best pilot to get the most out of AI. But it's primarily their funding strategy that needs work.

Enterprises don't need to change their existing budget allocations. Instead, they need to optimize savings from other IT domains to create new financial streams and channel those streams toward new AI projects.



3. <https://www.gartner.com/en/newsroom/press-releases/2025-06-25-gartner-predicts-over-40-percent-of-agentic-ai-projects-will-be-canceled-by-end-of-2027>

2. The funding-by-savings thesis

AI adoption is essential, but budgets are slim. The funding-by-savings approach circumvents establishing new budgets. Instead, it focuses on ring-fence verified cost savings to underwrite AI pilots.

But where should these savings come from?

- **Cloud:** Optimizing cloud costs can be game-changing. McKinsey⁴ reports that only 1 in 10 cloud transformations unlocks value. The takeaway? Behind subpar cloud operations is a massive savings pool waiting to be discovered.
- **Data:** Data optimization is cost optimization. Addressing issues in the data lifecycle and right-sizing storage and data processing resources can create significant re-investable savings.
- **Processes:** Ironing out bottlenecks across business-critical processes can expedite workflows, improve resource efficiencies, and yield significant cost savings.

These cloud, data, and process efficiencies yield major gains for funding new AI projects:

- **Improved fiscal discipline:** Ensures controlled investments for AI projects, balancing short-term financial needs and long-term sustainability.
- **Faster approvals:** Expedite greenlighting by providing key stakeholders and decision-makers with transparent budgets and value-driven approval requests.
- **Reduced risk:** Eliminates speculative investments by only injecting savings-based finances into ROI-aligned AI use cases and initiatives.

4. <https://www.mckinsey.com/capabilities/tech-and-ai/our-insights/unlocking-cloud-value-achieving-operational-excellence-through-sre>

Better stakeholder alignment: Connects leaders and critical teams by focusing on shared AI objectives and tangible outcomes.

Here are some useful guardrails to ensure funding-by-savings success:

- **No double-counting:** Cost savings from cloud, data, and process optimization should be clearly delineated and not overlap across projects.
- **CFO sign-offs:** The CFO should meticulously govern the cost-savings-to-AI-reinvestment pipeline, validating and authorizing every new injection of funds.
- **Auditable baselines:** Businesses should consistently document the current state to track and measure the future impact of AI investments.



3. The three savings levers

LEVER 1:

Cloud and Platform Optimization

Cloud and platform expenses can balloon without stringent oversight. The following are some key FinOps-based optimization strategies:

- > Right-size workloads to align compute and storage resources with actual needs and avoid overprovisioning.
- > Optimize reserved instance (RI) and savings plan (SP) coverage based on predictable workloads to avoid high on-demand pricing.
- > Opt for tiered storage models and move cold data to cheaper storage infrastructure.
- > Cap outbound data transfers to minimize costs and ensure accurate financial forecasting.

Decommission zombie resources, i.e., redundant and orphaned instances, that provide no business value.

LEVER 2:

Tooling and License Rationalization

Tool and licence sprawl result in uncontrolled expenditure that massively impacts ROI. Recommendations for consolidating tools and licenses include:

- > Consolidate redundant analytics, ETL, and observability suites; prioritize unified toolchains.
- > Decommission unnecessary AI or ML services and add-ons; only keep AI/ML services that directly connect to outcome-driven AI projects.
- > Rationalize SKUs by removing underused and obsolete products.

LEVER 3:

Process Automation and Deflection

Manual processes are a hive of operational and cost inefficiencies. AI-driven process automation strategies include:

- > Implementing GenAI and RAG architectures to expedite routine tasks and knowledge retrieval
- > Rerouting calls to AI agents instead of overwhelming support teams
- > Classifying and triaging tickets to ensure that critical issues are addressed first (this also reduces alert fatigue)
- > Expediting cycle time to optimize resource expenditure and improve time-to-market

Lever	Typical Savings Range	Time to Realize	Confidence	Effort
Cloud and Platform Optimization	15%-30% of cloud spend	1-3 months	High	Medium
Tooling and License Rationalization	20%-40% of software spend	2-4 months	Medium-High	Low-Medium
Process Automation and Deflection	30%-50% of labor spend for automated tasks	3-6 months	Medium	High

4. Quantification model and CFO-ready math

BASELINES

The baseline pillar focuses on starting points. Every single effect of kickstarting a funding-by-savings strategy should be measured against baselines.

Which ones should you focus on?

- > **12-month spend:** A one-year trail of cloud, data, and process-related costs
- > **Utilization:** A snapshot of (inefficient) resource consumption and usage patterns
- > **Labor:** A summary of staffing levels, capacity gaps, and labor redundancies

BENEFITS

Once your baselines are in place, it's time to categorize benefits into CFO-friendly slices.

Here's how to group:

- > **Hard cash:** Direct and immediate savings from ironing out inefficiencies
- > **Avoided hires/FTE backlogs:** Value gained by replacing headcount-heavy processes with AI-driven automation
- > **Cost-to-serve reduction:** Savings from reduced per-unit product expenses via expedited cycles and fewer manual interventions
- > **Risk-adjusted returns:** Overall financial gains minus potential adoption and operational risks

ILLUSTRATION OF A CFO-READY QUANTIFICATION MODEL

Let's take a look at a hypothetical scenario to grasp how this quantification model can help companies.

Category	Baseline	Time to Realize	Confidence
Cloud Run-Rate	\$3 million per year	15%	\$450k
Tooling Consolidation	\$900k per year	20%	\$180k
Automation Benefits (via Avoided Hires)	\$120k loaded per FTE	Avoided costs of 5 FTEs	\$600k (use CFO-approved rules for counting)

Total Savings for AI Reinvestment:
\$630k cash + policy-approved portion of \$600k equivalents

5. The 90-day pilot pattern (from scope to win)

STEP 1:

Choose an Optimal Use Case

Remember: Most AI pilots fail. So focus on simple, actionable, and high-value AI use cases like content generation, coding, data analytics, or automation instead of complicated vanity projects.

Also, use no-nonsense KPIs, e.g., support deflection, handle time, and invoice cycle time.

STEP 2:

Ensure Data Readiness

Strong GenAI and agentic AI pilot projects hinge on data readiness. Focus on answering these key questions:

- > **Availability:** Is the data complete and current?
- > **Access:** Can critical AI teams safely and easily access project-essential data?
- > **Quality:** Is the data fresh, accurate, and appropriate for the AI use case?
- > **Privacy:** Is the data protected according to relevant privacy regulations and frameworks?

STEP 3:

Establish the Technical Architecture

Next up: Set up architectural sketches for the AI pilot. Key elements include:

- > **RAG architectures for proprietary data:** Ensures rapid, compliant, and accurate retrieval of information.
- > **GPU access:** Ensures sufficient compute resources for AI pilot workloads.
- > **Latency considerations:** Deliver optimal response times to meet business goals and end-user expectations.

STEP 4:

Plan Delivery Cadence

Designing a project delivery timeline is crucial for AI initiatives. Key recommendations include:

- > **Week-by-week breakdown:** Define weekly deliverables and milestones.
- > **Demo frequency:** Schedule regular demos to showcase features and gather feedback for iterative improvements.
- > **Acceptance criteria:** Set clear standards to determine when a project is complete.
- > **Success thresholds:** Establish minimum conditions for moving the AI project from pilot to full implementation.

STEP 5:

Define Exit Criteria and Scaling Prerequisites

Not all AI pilots move on to full-fledged implementation, nor should they. This final step involves establishing a set of requirements and prerequisites that an initiative must meet to graduate.

For example, a GenAI support agent must achieve a 70%+ deflection rate with <5% escalation to human agents for the project to be deemed ready for graduation. Similarly, if a pilot doesn't reach 15% ROI after 90 days, it's important to conduct a thorough root cause analysis instead of scaling further.

If these KPIs and governance gates are met, your AI pilot is ready for take-off.

6. Funding governance and controls

How can you keep your funding-by-savings AI strategy on track? Set up a robust governance program, augmented by stringent technical and cultural controls.

Here are a few core components of AI funding governance:

RACI

The Responsible, Accountable, Consulted, and Informed (RACI) framework means establishing key roles and responsibilities:

- > **CIO:** Program execution

> **CFO:** Benefits validation

> **CISO:** Risk management
- > **Data/AI leads:** Project delivery

> **FinOps:** Savings verification

> **PMO:** Progress tracking

Activity	Responsible	Accountable	Consulted	Informed
Baseline Validation	FinOps	CFO	CIO	Board
Pilot Execution	Data/AI leads	CIO	CISO	CFO
Savings Verification	FinOps	CFO	CIO	Board
Risk Remediation	CISO	CIO	Legal	CFO

POLICIES

Policies formalize funding-by-savings, making the strategy scalable and repeatable. Non-negotiable policies include:

- > **Baseline lock:** Freezes initial spend and metrics for accurate measurement
- > **Benefit attribution:** Maps cost savings to their source efficiencies
- > **Reinvestment rules:** Establish how cost savings should be reinvested in AI projects
- > **Audit trails:** Track the entire funding-by-savings lifecycle, from initial efficiencies to reinvestment

Risk	Likelihood	Impact	Priority
Model Drift	High	Medium	High
Data Bias	Medium	High	Medium
Capacity Constraints	Low	Medium	Low
Vendor Lock-In	Medium	Low	Low

7. Communicating with the CFO and board

CFO and board sign-off is essential to scale funding-by-savings AI initiatives. The secret to clinching this is clear and concise communication.

A simple one-page CFO memo should be sufficient, but it must include the following:

- › **Description:** Outline the AI use case and quantify its business value.
- › **Baselines:** Include current state cost, labor, and performance metrics to measure improvements.
- › **Verified savings sources:** Specify the cloud, data, and process efficiencies contributing to the AI reinvestment pool.
- › **Pilot budget:** Detail allocated resources, including hard cash and policy-approved equivalents.
- › **Timelines:** Map the AI project schedule, including objectives, cadence, and completion dates.
- › **Risk controls:** List governance, compliance, and security measures to mitigate risks.
- › **KPIs:** Provide the right metrics to track measurable outcomes of the AI project.



8. Case snapshots

Below are four hypothetical examples that highlight how funding-by-savings AI pilots work in real-world scenarios. Each snapshot features the baseline, intervention, 90-day outcome, and scale plan.

AGENT IN A BOX

A 90-day M365 Copilot program transforms technical teams into AI practitioners while deploying production-ready custom agents. This structured approach combines training, use case development, and secure deployment with governance controls.



Baseline

Technical teams lack Copilot expertise and AI-implementation experience. Manual processes dominate workflows with fragmented communication, while organizations struggle to identify high-value use cases aligned with strategic objectives.

Intervention

Deploy a five-phase program: Copilot immersion training, use case workshops, environment security review, custom agent development, and production deployment with governance. Staff receive hands-on experience in sandbox environments, followed by guided development of business-aligned agents with compliance controls.

90-Day Outcome

- > Production agent is deployed with 60%-80% adoption, 20%-30% reduction in task completion times, and 4.5/5.0 user satisfaction.
- > 3-5 certified staff capable of building additional agents with a full governance framework are in place.
- > A documented case study with measurable ROI demonstrates 30% faster completion and 80% active usage rates.

Scaling Plan

The enterprise expands to 3-5 additional agents across departments using the trained internal team. It also establishes a Center of Excellence with standardized templates and governance playbooks for self-service development, plus funds expansion through verified pilot savings, creating a sustainable funding-by-savings loop.

CX ASSIST AND DEFLECTION

A GenAI pilot is launched to manage a high volume of support calls while enhancing compliance and response quality.



Baseline

The business receives a high volume of support calls. Costs per contact are high, and compliance tracking is negligible.

Intervention

- > Implement a text- and voice-based GenAI agent, featuring built-in compliance guardrails.

90-Day Outcome

- > Between 10% and 20% of support calls are deflected.
- > Costs per contact are significantly reduced.
- > Human support teams can focus on more advanced work.

Scaling Plan

The enterprise strategizes to roll out more GenAI agents across support use cases, broadening coverage and further automating customer communications.

KNOWLEDGE IN A BOX

A RAG-powered AI pilot is deployed to improve operations, including expediting task execution, reducing unnecessary escalations, and minimizing onboarding time



Baseline

Knowledge retrieval is a manual process. As a result, task completion takes longer, teams experience more escalations, and significant time and resources are required to train new employees.

Intervention

The enterprise launches the RAG-driven application for immediate and automatic information retrieval.

90-Day Outcome

- > Task completion time is slashed by 25%-40%.
- > Escalation rates drop sharply.
- > Onboarding time is significantly reduced as new teams benefit from AI-powered knowledge retrieval.

Scaling Plan

The business plans to broaden the use case potential of the RAG-powered AI tool, operationalizing it across teams and domains.

SPEND INTELLIGENCE

A FinOps + AI pilot is launched to continuously optimize expenses and eliminate unnecessary costs and resources.



Baseline

Cloud services, tools, and data are siloed, leaving teams without a holistic view of resource usage patterns and spending. As a result, overprovisioned resources stagnate and costs balloon.

Intervention

The business implements a FinOps and AI solution designed to detect idle assets in real time and present optimization options.

90-Day Outcome

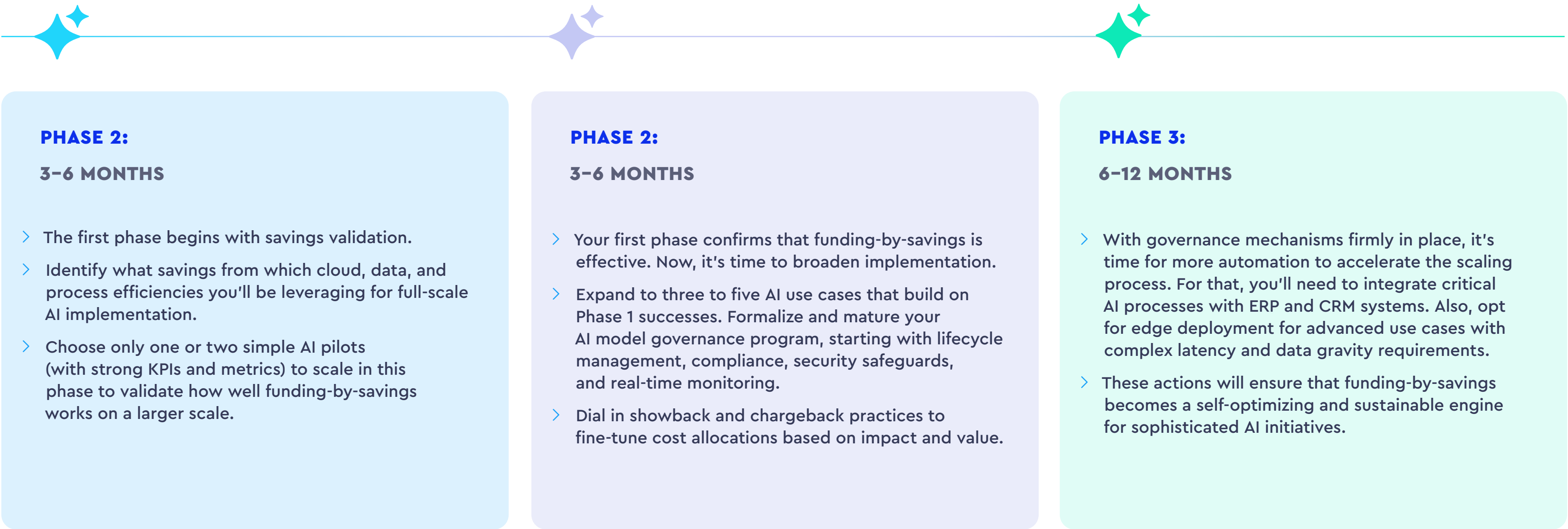
- > Redundant and dormant resources are decommissioned, unlocking immediate cost savings.
- > Continuous right-sizing supports continuous cost optimization.

Scaling Plan

The enterprise scales its spend intelligence solution across more environments, creating a vast reinvestment pool for more advanced and expensive AI use cases.

9. Scaling AI projects: A 12-month roadmap

Successful AI initiatives should be lined up for full-fledged implementation, which requires a meticulous 12-month plan. Here's a breakdown of the full-scale implementation roadmap, broken into three phases.



10. Build your funding-by-savings engine with RapidScale

The funding-by-savings strategy is a powerful solution for business leaders to launch AI projects without shaking up existing budget allocations. Instead of procuring an entirely new stream of funds, funding-by-savings focuses on optimizing cloud, data, and processes, and reinvesting those cost savings in AI projects.

Unveiling these hidden streams of AI funding is an intricate undertaking, and businesses might need a little support. That's why RapidScale's AI/ML and cloud services can be a game changer.

RapidScale's cost optimization and FinOps services can unlock as much as 60% cost savings during cloud migration and 30% cost savings over conventional data centers. That makes up a huge reinvestment pool to funnel into AI projects. Plus, RapidScale's advanced AI and machine learning services are built to support Microsoft-centric environments, and can help you design, deploy, and scale those AI initiatives end to end.

From identifying invisible inefficiencies and developing mitigation plans to implementing outcome-driven AI projects with your cost savings, learn how RapidScale has you covered

Contact us to speak to a RapidScale expert today.

meet@rapidscale.net

