

eFuturesCFO Masterclass Series

AI Workflows for the Modern CFO

PART 9

Infrastructure and Vendor Intelligence

Use Case Five: From OpEx Detail to Cash Runway

Hindol Datta

Chief Content and Product Officer, eFuturesCFO

A Note Before Part 9

Part 9 introduces the fifth and final use case of this masterclass: the Infrastructure and Vendor Intelligence workflow. The workflow is sequenced fifth in John Campbell's deployment plan because it addresses the operational expense side of the business, and the capacity to address operational expense effectively depends on the discipline established by the prior four workflows.

What is different about this workflow is the nature of its output. The Finance Operations Copilot in Part 5 produced clean transaction data and policy adherence. The Board Reporting workflow in Part 6 produced editorial output for an executive audience. The Pipeline Intelligence workflow in Part 7 produced behavioral observations and an adjusted forecast. The Forecasting Engine in Part 8 produced a defensible three-scenario forecast. The Infrastructure and Vendor Intelligence workflow produces something more directly measurable: dollar savings on the operating expense line.

The savings are not transformative in isolation. Aggregate identified savings in the data are approximately one hundred four thousand dollars annually, roughly five percent of non-payroll operating expense. What makes the workflow valuable is not the magnitude. It is the discipline of running it every quarter so that the savings compound over the lifetime of the company. Companies that run this workflow capture roughly twenty percent more operating margin over three years than companies that do not, because the cumulative effect of catching dozens of small inefficiencies before they grow is much larger than the effect of any single optimization.

The workflow is operationally owned by Tom Reeves (Director of Operations) with Raj Patel (CTO) as the technical reviewer for infrastructure recommendations. The finance function is a stakeholder because the savings flow directly into the cash runway calculation that drives Series C timing. The Forecasting Engine in Part 8 made cash runway an explicit measure; the workflow in this part is one of the levers that extends it.

The Use Case Five data package, `Use_Case_5_Data.zip`, accompanies this PDF. It contains seven files: twelve months of AWS spend detail by service, an inventory of forty AWS resources with utilization data, twenty-five SaaS subscriptions with seat-versus-user data, thirty active vendor contracts with renewal dates, a twelve-month payment summary across thirty-three vendors, the spend categorization framework, and the vendor consolidation history document.

Small savings compound. The discipline of finding them every quarter is what separates the finance function that protects cash runway from the finance function that watches cash runway erode quarter after quarter.

Hindol Datta

Contents

- Section 1** The Business Problem
- Section 2** The Eight-Step Methodology Applied
- Section 3** The Data Package Walk-Through
- Section 4** Step-by-Step Tutorial: Building the Workflow
- Section 5** From Tutorial to Production
- Section 6** Expected Outputs and Success Criteria
- Appendix A** Assessment: Twenty Questions
- Appendix B** Answer Key with Explanations

Section 4 contains the eleven-step tutorial. Plan to have the data files accessible in Claude or ChatGPT.

Section 1 · The Business Problem

Helix Cloud Systems spends approximately one point seven million dollars per month on non-payroll operating expenses. The largest category is hosting (AWS), running approximately one hundred sixty-five thousand dollars in April 2026, a thirteen-fold increase over the May 2025 baseline that began the trailing twelve months. The second largest is software subscriptions, approximately twenty-five thousand dollars per month aggregated across roughly two dozen tools. The third category is professional services, marketing programs, and other variable spend.

What John heard in the discovery conversations

Tom Reeves, in his second conversation with John, said three things that established the problem.

First, "we know we are leaking money on AWS, we just have not had time to find it systematically." Tom and Raj had done one informal pass through AWS spend in February 2026, identified a few clearly idle resources, and saved roughly five thousand dollars in monthly spend. Both believed there was meaningfully more, but neither had time to do the full audit.

Second, "we have a lot of software tools whose seats outnumber the people who actually use them, and nobody is responsible for matching those." Tom estimated five to ten percent of software spend was attributable to seats that were paid for but unused. Across roughly three hundred thousand dollars in annual software spend, that was potentially thirty thousand dollars annually.

Third, "the Stratosphere thing has bothered me for months." Tom had identified that Stratosphere Software and Stratosphere Inc were the same underlying vendor billed under different entity names. The finance team was paying both. He had not yet raised it formally because he wanted to verify with the vendor first.

The cost of the current state

The visible costs are individually small. AWS waste of perhaps forty to sixty thousand dollars monthly. Software seat under-utilization of perhaps thirty thousand dollars annually. The Stratosphere duplicate at twenty-four thousand dollars annually. Each individual item is too small to demand executive attention on its own.

The aggregate cost is meaningful. Roughly one hundred thousand dollars annually in identifiable savings in the current data. At the current burn rate of approximately one point seven million monthly, this extends cash runway by approximately two weeks. Two weeks is not transformative, but the Series C window is twelve to fourteen months away, and any extension of the runway buffer matters at the margin.

The compounding cost is the more consequential one. The five percent of non-payroll OpEx that is identifiable savings today will, if not addressed, grow with the company. AWS spend that doubles in the next twelve months brings AWS waste that doubles with it unless the discipline catches it. Software seat under-utilization that exists today will exist again next quarter and the quarter after that, with new tools added to the inventory unless someone is responsible for the audit.

What we will build

The Infrastructure and Vendor Intelligence workflow analyzes four categories of operational spend and produces a structured list of optimization opportunities.

First, AWS spend analysis. The workflow reviews AWS resource utilization, identifies oversized and idle resources, and produces specific resource-level recommendations.

Second, software subscription analysis. The workflow compares seats purchased to actual users, identifies under-utilized subscriptions, and recommends seat reductions or tool elimination.

Third, vendor contract analysis. The workflow reviews upcoming renewal dates, identifies contracts due for renegotiation, and surfaces consolidation opportunities.

Fourth, vendor concentration and risk analysis. The workflow identifies the largest vendors by spend and produces a risk view of vendor concentration.

What this workflow is and is not

The workflow is a structured way of running the operational expense review that Tom and Raj would do if they had time. It does not make decisions; it surfaces opportunities. Every recommendation is reviewed by the operational owner (Tom for vendors, Raj for AWS, the relevant tool owner for software) before any action is taken. The workflow makes the audit possible to run every quarter, not less because the work is easier but because someone has compiled the questions and the data in one place.

Section 2 · The Eight-Step Methodology Applied

Walking through the methodology one final time for this workflow.

Step One: Articulate the business problem

In one sentence: "Helix has identifiable operational savings of approximately one hundred thousand dollars annually that are not being captured because no one has time to systematically audit AWS resources, software seat utilization, and vendor contracts on a recurring basis."

The function that owns this problem is operations (Tom Reeves) with technical input from engineering (Raj Patel) for infrastructure items. The finance function (John Campbell) is the principal stakeholder.

Step Two: Specify the output

The workflow produces four categories of output.

First, the AWS optimization output. A structured list of specific resources flagged for action, organized by recommendation type (right-size, terminate, reserve). Each item has the resource ID, current cost, estimated savings, recommended action, and engineering review required (yes/no).

Second, the software subscription output. A structured list of subscriptions with utilization data, organized by recommendation type (reduce seats, eliminate, consolidate). Each item has the subscription name, current cost, utilization percentage, estimated savings, and recommended action.

Third, the vendor contract output. A structured list of contracts with renewal dates in the next six months, organized by recommendation type (renegotiate, consolidate, terminate). Each item has the vendor, contract value, renewal date, and recommended action.

Fourth, an executive summary covering the total savings identified, the implementation effort required, and the cash runway impact.

Step Three: Classify risk

The workflow is classified Tier Two: Internal Operational. Its outputs are recommendations for operational expense management that do not influence financial reporting directly. Tier Two governance applies: audit trail, named human reviewer, periodic quality monitoring, CFO awareness but not approval required for each invocation.

Why Tier Two rather than Tier Three: the outputs of this workflow do not enter the financial statements or board commentary directly. They produce operational recommendations that, when executed, reduce

future operating expense. The execution of any recommendation has its own approval process appropriate to its magnitude.

Step Four: Map data flow

The workflow reads, in production, the following sources.

Input	Source	Classification
AWS Cost & Usage Reports	AWS Cost Explorer via API	Confidential
AWS Resource Inventory	AWS Config + CloudWatch via API	Confidential
Software Subscription Inventory	Internal procurement system	Confidential
Vendor Contracts	Contract repository	Confidential
Vendor Payment History	NetSuite via MCP	Confidential
Spend Categorization Framework	Internal docs	Internal
Vendor Consolidation History	Internal docs	Internal

Step Five: Design the human review pattern

Review-before-output with operational ownership. The AWS recommendations are reviewed by Raj Patel (CTO) before any engineering action. The software subscription recommendations are reviewed by Tom Reeves with sign-off from the relevant tool owner (e.g., Michael O'Brien for Salesforce, Lisa Mahoney for Outreach). The vendor contract recommendations are reviewed by Tom with John Campbell's sign-off for any recommendation above twenty-five thousand dollars in annual value.

The review pattern reflects that operational expense optimization has distributed ownership. The CFO does not need to approve every small change but does need visibility into the aggregate.

Step Six: Specify the audit trail

Standard thirteen-field audit trail. Retention seven years. Access read by Tom, John, Raj, and the relevant tool owners.

Step Seven: Define the substitution path

Primary model: Claude. Alternative: GPT. Substitution path tested annually.

Step Eight: Approval and registration

Approval obtained from the Governance Working Group. Registered as WFR-005 with version 1.0. The Q3 2026 cycle is the first pilot.

Section 3 · The Data Package Walk-Through

Open Use_Case_5_Data.zip. The new folder is uc5_infra_vendor_intelligence with seven files.

aws_cost_detail.csv

Approximately seventy-two rows showing twelve months of AWS spend broken down by service category (EC2 compute, RDS database, S3 storage, CloudFront CDN, data transfer, other services). Each row has month, service, amount, and notes. Used to establish the AWS spend trajectory and identify which service categories are growing fastest.

aws_resource_inventory.csv

Forty rows covering individual EC2 instances, RDS databases, and other AWS resources. Each row has instance ID, service type, instance type, monthly cost, environment (production/staging/dev), CPU utilization average, memory utilization average, status assessment, and notes. The principal source for AWS right-sizing and idle-resource recommendations.

software_subscriptions.csv

Twenty-five rows covering each SaaS subscription. Each row has subscription name, category, seats purchased, actual users in the trailing thirty days, annual cost, monthly cost, contract term, and renewal date. The principal source for seat optimization recommendations.

vendor_contracts.csv

Thirty rows covering all active vendor contracts with annual value above approximately six thousand dollars. Each row has vendor name, category, annual value, renewal date, contract type, auto-renewal status, and notes. Used to identify contracts due for renegotiation in the next six months.

vendor_payment_summary.csv

Thirty-three rows providing a twelve-month aggregated view per vendor: total paid, payment count, average monthly. The summary view used for vendor concentration analysis.

spend_categorization.md

The four-category optimization framework: Infrastructure, Operational Software, Professional Services and Marketing Programs, and Fixed Operational. Each category has its own optimization approach and risk profile.

vendor_consolidation_history.md

History of past vendor consolidation efforts at Helix, plus the operational principles governing future consolidation decisions. Notes the active Stratosphere consolidation opportunity.

How to look at the files

Spend ten minutes with the files.

Open `aws_resource_inventory.csv`. Filter or sort by the status column. Notice the "over-provisioned" and "idle" rows. Read the CPU percentages: the over-provisioned rows show eighteen percent, twenty-two percent, twelve percent CPU on instances costing seven hundred thirty or eleven hundred twenty dollars monthly. The idle rows show zero to five percent CPU. The pattern is visible at a glance.

Open `software_subscriptions.csv`. Compute mentally for a few rows: seats purchased divided by actual users. Salesforce: thirty-five seats, thirty-two users, ninety-one percent utilized. Gong: twenty-two seats, nine users, forty-one percent utilized. The under-utilized rows are obvious. Notice the Stratosphere Software and Stratosphere Inc rows side by side, with similar profiles, continuing the duplicate vendor pattern from Use Case One.

Open `vendor_payment_summary.csv`. Sort or scan by total twelve-month payment. The top entries are AWS (\$1.98M), SF Office Landlord (\$360K), Google Ads (\$240K), LinkedIn Ads (\$180K), Recruiting Partners (\$120K). The top five represent a meaningful concentration of total spend.

A note about precision

The savings estimates this workflow produces are estimates, not commitments. AWS right-sizing might save the calculated amount, more, or less, depending on whether engineering can execute the change cleanly. Software seat reductions might be partly captured back by users who join in the next quarter. The workflow surfaces opportunity sizes; the operational owner converts opportunity to realized savings.

Section 4 · Step-by-Step Tutorial: Building the Workflow

You will now build the Infrastructure and Vendor Intelligence workflow. The pattern is familiar. Plan for approximately forty minutes.



Tutorial Step 1: Start a fresh conversation

Open your browser. Navigate to claude.ai or chatgpt.com. Start a new conversation.

Why this matters:

Different workflow, different inputs, different audience, different output. Fresh conversation prevents context contamination.

Tutorial Step 2: Set the role and context

Paste the following prompt.

```
You are acting as my Infrastructure and Vendor Intelligence workflow for a Series B B2B SaaS company called Helix Cloud Systems. I am the CFO, John Campbell. The operational owner of this workflow is Tom Reeves (Director of Operations) with Raj Patel (CTO) as technical reviewer for infrastructure recommendations.
```

```
Your job is to analyze our operational expense base (AWS, software subscriptions, vendor contracts) and produce a structured list of optimization opportunities. The output is reviewed by the operational owners before any action is taken.
```

```
The workflow has four output categories:
```

1. AWS optimization: specific resources flagged for right-sizing, termination, or reservation
2. Software subscription optimization: seat reductions, tool eliminations, consolidations
3. Vendor contract optimization: renegotiations, consolidations, terminations
4. Concentration and risk analysis: largest vendors by spend, dependencies

```
For each opportunity, produce:
```

- Annual savings estimate (range, not single point)
- Implementation effort (low, medium, high)
- Risk level (low, medium, high)
- Recommended owner

I will provide seven files: AWS cost detail, AWS resource inventory, software subscriptions, vendor contracts, vendor payment summary, spend categorization framework, vendor consolidation history.

Please confirm the role and ask for the first file.

Send.

Why this matters:

The role prompt emphasizes that the workflow produces ranges rather than point estimates, and that every recommendation is reviewed by an operational owner. These framings shape the output language. The model will produce recommendations with appropriate uncertainty rather than false precision.

Tutorial Step 3: Provide the spend categorization framework

Open `spend_categorization.md`. Copy contents. Paste with this prompt.

Here is our spend categorization framework. Please read and confirm:

1. The four categories
2. The optimization approach for each
3. The risk levels associated with each
4. The convention for workflow output (high-confidence, medium-confidence, investigation)

[Paste contents of `spend_categorization.md` here]

Why this matters:

Loading the framework first establishes the structure the model will use to organize its recommendations. Without the framework, the model would invent its own categorization. With the framework, the recommendations align with how the operational function actually thinks about expense.

Tutorial Step 4: Provide the AWS cost detail

Open `aws_cost_detail.csv`. Copy contents. Paste with this prompt.

Here are 12 months of AWS spend by service category. Please analyze:

1. The total AWS spend in the most recent month vs. 12 months ago (growth rate)
2. Which service categories have grown fastest
3. The current monthly total and trajectory
4. Any service category whose growth rate is materially different from the overall trajectory

[Paste contents of `aws_cost_detail.csv` here]

Why this matters:

The trajectory analysis establishes the magnitude of the AWS spend problem. AWS doubling in twelve months while revenue grows by roughly half that rate is the structural issue. Identifying which categories drive the growth tells the operational owner where to focus.

Expected response: April 2026 total of approximately \$313K versus May 2025 of \$172K, an 82% increase. EC2 grew from \$88K to \$165K (+88%). All categories grew, but EC2 is the largest absolute driver. The trajectory shows steady monthly increase with no sign of natural slowing.

Tutorial Step 5: Provide the AWS resource inventory

Open `aws_resource_inventory.csv`. Copy contents. Paste with this prompt.

Here is the inventory of 40 AWS resources with their utilization data. Please analyze and produce specific recommendations:

1. Identify all resources marked as "over-provisioned" - what is the total monthly cost and what is the estimated savings if right-sized?
2. Identify all resources marked as "idle" - what is the total monthly cost and what is the estimated savings if terminated?
3. For the right-sizing recommendations, what is the recommended target instance type (smaller version of the current type)?
4. For the termination recommendations, are there any that look like they might still be in use (dev/staging) where you would recommend a usage check before termination?
5. Total identified annual savings from AWS resource optimization

[Paste contents of `aws_resource_inventory.csv` here]

Why this matters:

The resource inventory is where the workflow generates the most specific operational recommendations.

Right-sizing recommendations must include the target instance type because that is what engineering needs to act on. Termination recommendations must distinguish between confidently idle and possibly-in-use, because the action is different.

Expected response: four over-provisioned production instances totaling \$3,700/month, right-sizing saves roughly \$2,200/month or \$26K annually. Ten idle resources (8 EC2 + 2 RDS) totaling \$2,418/month, termination saves \$29K annually. Right-sizing targets identified by the model (m5.8xlarge → m5.2xlarge, r5.4xlarge → r5.xlarge, etc.). Total identified annual savings: approximately \$55K from AWS alone.

Tutorial Step 6: Provide the software subscriptions

Open `software_subscriptions.csv`. Copy contents. Paste with this prompt.

Here are our 25 software subscriptions with seats purchased and actual users (trailing 30 days). Please analyze:

1. Compute the seat utilization rate for each subscription (actual users / seats purchased)
2. Identify any subscriptions with utilization below 60% and their annual cost
3. For each under-utilized subscription, what is the estimated annual savings from right-sizing seats to actual users plus a 10% buffer?
4. Identify any subscriptions where the seat count itself looks anomalous (very large or very small relative to company size of 142 employees)

5. Note any potential consolidation opportunities (subscriptions with overlapping functionality)
6. Total identified annual savings from software subscription optimization

[Paste contents of software_subscriptions.csv here]

Why this matters:

Software optimization is largely arithmetic: seats purchased minus actual users equals the optimization opportunity, with a buffer for new joiners. The workflow does the arithmetic for every subscription consistently, which is the work that currently does not happen because no one has time.

Expected response: four subscriptions with utilization below 60% (Gong 41%, Outreach 48%, ZoomInfo 40%, Loom 40%). The Stratosphere Software and Stratosphere Inc rows are flagged as a consolidation opportunity. Total software savings approximately \$35-40K annually including the Stratosphere consolidation.

Tutorial Step 7: Provide the vendor contracts

Open vendor_contracts.csv. Copy contents. Paste with this prompt.

Here are our 30 active vendor contracts with their renewal dates and notes. Please analyze:

1. List all contracts with renewal dates within the next 90 days, in order of renewal date
2. For each near-renewal contract, identify whether the notes flag any specific opportunity (renegotiation, consolidation, under-utilization)
3. Identify the contracts where auto-renewal is enabled but the notes suggest action would be valuable
4. The Stratosphere Software and Stratosphere Inc contracts: confirm the consolidation opportunity and estimate the savings

[Paste contents of vendor_contracts.csv here]

Why this matters:

Contract analysis is calendar-driven. Renewals must be acted on before they happen; afterwards the leverage is gone. The workflow surfaces the contracts approaching renewal so the operational team has time to renegotiate, consolidate, or terminate before auto-renewal triggers.

Expected response: a list of approximately 8-12 contracts renewing in the next 90 days, with the highest-leverage items being the under-utilized subscriptions (Gong August 15, Outreach July 1, ZoomInfo September 30) and the Stratosphere consolidation (July 15 and August 1). The estimated savings from Stratosphere consolidation is roughly \$24K annually.

Tutorial Step 8: Provide the vendor payment summary for concentration analysis

Open vendor_payment_summary.csv. Copy contents. Paste with this prompt.

Here is the 12-month vendor payment summary across our 33 largest vendors. Please:

1. Identify the top 5 vendors by total 12-month payment
2. Compute what percentage of total non-payroll OpEx these top 5 represent
3. Assess the concentration risk: are any of these vendors single points of failure?
4. For the top 5, identify any that warrant specific board-level visibility because of their size or strategic importance
5. The vendor consolidation history mentions historic consolidations - are there other consolidation opportunities visible in this data?

[Paste contents of vendor_payment_summary.csv here]

Why this matters:

Concentration analysis is the strategic view of vendor spend. Operational optimization at the line-item level is one consideration; vendor concentration is a separate strategic consideration that should be visible to the board. AWS at over twenty percent of OpEx is a strategic exposure that the board should know about even if no immediate action is recommended.

Expected response: top 5 vendors are AWS (\$1.98M), SF Office Landlord (\$360K), Google Ads (\$240K), LinkedIn Ads (\$180K), Recruiting Partners (\$120K), totaling approximately \$2.88M out of approximately \$4.4M total annual non-payroll OpEx, roughly 65%. AWS specifically is identified as a strategic dependency. The consolidation history note about Stratosphere is repeated.

Tutorial Step 9: Provide the consolidation history

Open vendor_consolidation_history.md. Paste with this prompt.

Here is our vendor consolidation history and the operational principles governing future consolidation decisions. Please:

1. Note the past consolidations and their savings
2. Confirm the operational principles
3. Apply the principles to the Stratosphere consolidation opportunity: who should be the owner, when should it be executed, what is the documented business case?

[Paste contents of vendor_consolidation_history.md here]

Why this matters:

The consolidation history establishes the principles the company has adopted. Applying those principles to the current opportunity is what makes the recommendation operational rather than theoretical. By forcing the workflow to apply principles explicitly, you produce a recommendation that is ready to execute rather than a recommendation that requires further thought before execution.

Tutorial Step 10: Produce the consolidated savings summary

Now the synthesis. Type this prompt.

Now produce the consolidated savings summary for me and Tom Reeves. The format:

OPPORTUNITY SUMMARY TABLE

A structured table with one row per opportunity, columns for:

- Opportunity name
- Category (AWS, Software, Vendor Contract)
- Annual savings estimate (range)
- Implementation effort (low/medium/high)
- Risk level (low/medium/high)
- Recommended owner
- Recommended target action date

CASH RUNWAY IMPACT

Calculate the total identified annual savings and translate to the cash runway impact at our current burn rate of approximately \$1.7M monthly.

EXECUTIVE NOTE

Two to three paragraphs of prose suitable for inclusion in the next board commentary, framing the operational expense optimization program, the savings identified in this cycle, and the implementation plan.

Why this matters:

The synthesis is what Tom and you will take into the operational planning meeting. The table format makes the opportunities directly actionable. The cash runway translation connects the operational work to the strategic context (the Series C window). The executive note prepares the board communication.

Expected response: a structured table with roughly 8-12 specific opportunities totaling approximately \$100-110K annually. The cash runway impact is approximately 2 weeks of additional runway. The executive note frames the program as ongoing operational discipline rather than a one-time exercise.

Tutorial Step 11: Save and document

Save the conversation as "Helix Q2 2026 OpEx Intelligence Review." Write the workflow registry note: date, files used, prompts that produced the best output, places where the model required iteration.

The Infrastructure and Vendor Intelligence workflow runs quarterly. The pace is right because operational expense patterns do not change weekly; they accumulate over months. A quarterly cadence catches the cumulative drift before it becomes large.

Why this matters:

The fifth and final workflow completes the masterclass. The pattern you have used five times is now internalized. You can apply the same pattern to any new workflow design. The eight-step methodology, the data loading discipline, the role prompt with audience and governance framing, the specific analytical questions with format requirements, the synthesis with explicit output structure, the documentation. The skill transfers. The specific prompts are the application.



What you have just done

You have built the fifth and final AI workflow of the masterclass. You have now seen the pattern applied five times across five functional areas: finance operations, board reporting, sales pipeline intelligence, forecasting, and infrastructure and vendor intelligence. The cumulative effect is that the executive skill of designing AI work for finance is now yours.

The remaining part of the masterclass (Part 10) is not a new workflow. It is the meta-discipline of operating the program over the eighteen-month horizon: how to detect workflow decay, how to refresh data inputs, how to handle model upgrades, how to retire workflows that have served their purpose, and how to evolve the governance framework as the company scales. Part 10 closes the arc that Part 1 opened.

Five workflows built. The pattern is now yours. The discipline that turns the pattern into a program is what Part 10 will cover.

Section 5 · From Tutorial to Production

The production version of this workflow is the simplest of the five to engineer, because the data sources are already well-structured and the outputs are operational rather than editorial.

What changes in production

1. Data flows automatically

AWS data flows from Cost Explorer and Config via the standard AWS APIs. Software subscription data flows from the procurement system (or, until one is selected, from a centralized spreadsheet that Tom maintains). Vendor contract data flows from the contract repository. All without manual paste.

2. The workflow runs quarterly

The cadence is quarterly because the patterns change slowly. Running monthly would produce no additional value and would add review burden. Running annually would miss opportunities that compound across multiple quarters.

3. The output is a structured ticket queue

In production, each opportunity becomes a ticket in the operational backlog with an assigned owner and a target close date. The ticket queue becomes the operational tracking mechanism. The CFO can see at any time which opportunities are open, which have been closed, and what cumulative savings have been realized.

4. The audit trail captures the disposition of each opportunity

The audit trail extends beyond just the workflow output. Each identified opportunity is logged with its disposition: realized (savings captured), partial (some savings captured), declined (operational owner decided not to act, with reason), or pending (still open). The cumulative disposition data informs the workflow's next cycle.

Engineering work to get from tutorial to production

Approximately five to six engineering weeks for the initial production version. This is the smallest engineering investment of the five workflows because the data sources are simpler and the outputs are operational tickets rather than executive narratives.

Two weeks for the data integration across AWS APIs, the procurement system, and the contract repository. One week for the workflow orchestration. One week for the ticket queue integration so that opportunities flow into the operational backlog. One to two weeks for testing across at least one full quarterly cycle.

A note on continuous improvement

The opportunities that this workflow surfaces in any given quarter are a subset of the total opportunity. The remaining opportunities will become visible in future quarters as the data evolves: new resources are provisioned, new subscriptions are added, new vendors are onboarded, contracts come up for renewal. The workflow is not a one-time audit. It is the operational rhythm of expense discipline.

The compounding discipline

Five percent of non-payroll OpEx in identified savings every quarter, even if only partially realized, is the kind of cumulative discipline that, over three to five years, produces operating margin differences of ten to fifteen points relative to peer companies that do not run this kind of workflow. The discipline does not feel transformative in any single quarter. It is transformative over a longer horizon.

Section 6 · Expected Outputs and Success Criteria

Expected outputs

Every quarterly cycle produces five artifacts.

AWS optimization recommendations

A structured list of specific AWS resources flagged for right-sizing, termination, or reservation. Each item has the resource ID, recommended action, estimated savings, and engineering review flag.

Software subscription recommendations

A structured list of SaaS subscriptions with utilization analysis and seat reduction or elimination recommendations.

Vendor contract recommendations

A structured list of upcoming renewals (next 90 days) with recommended actions for each.

Concentration analysis

Vendor concentration view with the largest vendors identified and risk implications discussed.

Cash runway impact summary

Translation of identified savings into cash runway impact, plus an executive note for the board commentary.

Success criteria

Metric	Baseline	Target (12 months)
Identified savings per cycle	Ad-hoc	\$80-120K identified per quarter
Realized savings per cycle	Untracked	60-75% of identified
Cumulative annualized savings	Untracked	\$300-400K annualized
Cash runway extension	No tracking	6-8 weeks cumulative
AWS waste as % of AWS spend	~5-7% estimated	Below 3%

What is being measured, and why

Identified savings per cycle is the workflow's output measure. Approximately one hundred thousand dollars per quarter is what this size of company should expect.

Realized savings per cycle is the operational measure. Not every identified opportunity gets realized. A sixty to seventy-five percent realization rate is healthy; lower suggests the opportunities are not being prioritized; higher suggests the workflow is being too conservative in what it identifies.

Cumulative annualized savings is the strategic measure. Three to four hundred thousand dollars annually, sustained, is meaningful cash runway extension over the Series C horizon.

Cash runway extension is the consequence the CFO cares about. Six to eight weeks of cumulative runway extension over twelve months is material at the Series C decision point.

AWS waste as a percentage of AWS spend is the principal infra measure. Below three percent is the discipline target.



End of Part 9

Infrastructure and Vendor Intelligence

You have built the fifth and final AI workflow of the masterclass. The five workflows together cover the principal time and judgment consumers in a Series B SaaS CFO's working life: transaction review during close, board commentary at quarter end, pipeline forecast reliability, integrated revenue and expense forecasting, and operational expense optimization. Each workflow operates under the same governance framework. Each was designed using the same eight-step methodology. Each was built using the same tutorial pattern.

Part 10 closes the masterclass. It is not a sixth use case. It is the meta-discipline of operating the program over the eighteen-month horizon and beyond: detecting workflow decay, refreshing data inputs, handling model upgrades, retiring workflows that have served their purpose, and evolving the governance framework as the company scales. Part 10 also includes John Campbell's eighteen-month reflection on what the program produced, what it did not produce, and what the next CFO who inherits this discipline should know.

Before proceeding, take the assessment.



Appendix A · Assessment

Twenty questions on Part 9.

Part I: Multiple Choice

1. The Infrastructure and Vendor Intelligence workflow is sequenced fifth because:

- (a) It is the most engineering-intensive.
- (b) It addresses operational expense optimization, where the discipline depends on the framework established by the prior four workflows.
- (c) The CTO requested it as the latest deployment.
- (d) It is the highest-priority workflow.

2. The workflow is classified as:

- (a) Tier One: Internal Exploratory.
- (b) Tier Two: Internal Operational.
- (c) Tier Three: Financial Reporting Adjacent.
- (d) Tier Four: Financial Reporting Direct.

3. The principal operational owner of the workflow is:

- (a) John Campbell (CFO).
- (b) Tom Reeves (Director of Operations).
- (c) Raj Patel (CTO).
- (d) Elena Vargas (Senior FP&A; Manager).

4. The four AWS resources identified as over-provisioned are:

- (a) All idle dev instances.
- (b) Production instances running at 12-22% CPU and oversized for the workload.
- (c) Database instances only.
- (d) No over-provisioned resources exist in the data.

5. The number of idle AWS resources (EC2 + RDS combined) is:

- (a) Four.
- (b) Eight.
- (c) Ten.
- (d) Sixteen.

6. The four under-utilized SaaS subscriptions identified are:

- (a) Salesforce, NetSuite, Rippling, Slack.
- (b) Gong, Outreach, ZoomInfo, Loom.
- (c) Datadog, Snowflake, GitHub, Notion.
- (d) Google Workspace, Adobe, 1Password, PagerDuty.

7. The Stratosphere vendor consolidation opportunity comes from:

- (a) Two separate vendor relationships for the same underlying vendor.
- (b) Two unrelated vendors with similar names.
- (c) A single vendor billing twice in error.
- (d) A new vendor onboarding process gap.

8. The top 5 vendors by 12-month spend represent approximately:

- (a) 25% of total non-payroll OpEx.
- (b) 45% of total non-payroll OpEx.
- (c) 65% of total non-payroll OpEx.
- (d) 95% of total non-payroll OpEx.

9. The total identified annual savings in the data is approximately:

- (a) \$10-15K.
- (b) \$100-110K.
- (c) \$500K.
- (d) \$1.5M.

10. The workflow runs in production at what cadence?

- (a) Weekly.
- (b) Monthly.
- (c) Quarterly.
- (d) Annually.

11. The recommended target savings realization rate (against identified) is:

- (a) 25%.
- (b) 60-75%.
- (c) 95%.
- (d) 110%.

12. The cash runway extension target for 12 months of cumulative discipline is:

- (a) 1 week.
- (b) 6-8 weeks.
- (c) 6 months.
- (d) 12 months.

Part II: Short Answer

13. In two or three sentences, explain why this workflow is classified Tier Two rather than Tier Three, when most of the prior workflows in this masterclass were classified Tier Three.

14. The tutorial loads the spend categorization framework before loading any spend data. In two or three sentences, explain what this order accomplishes.

15. The workflow surfaces approximately \$100K in annual savings opportunities. The cash runway impact is only about 2 weeks. In two or three sentences, explain why this workflow is still worth running quarterly, given the modest individual cycle impact.

16. The workflow has distributed human review (Raj for AWS, Tom for vendors, tool owners for software, CFO sign-off above a threshold) rather than centralized review by one reviewer. In two or three sentences, explain why this distribution is appropriate for this workflow.

17. The production version produces tickets in an operational backlog rather than a narrative report. In two or three sentences, explain why this output form is more useful than a narrative for this particular workflow.

Part III: Scenario-Based

18. Scenario: Three quarters into running the workflow, the realization rate is consistently around 45% (well below the 60-75% target). Tom Reeves tells you the opportunities are valid but his team is consistently overwhelmed by other work, particularly during quarter-end close periods which coincide with when the workflow runs. In one paragraph of executive prose, describe how you would respond, what change to the workflow or to the operational team's capacity you would consider, and what governance step you would take.

19. Scenario: The workflow identifies an AWS optimization that would save approximately \$40K annually but would require a one-week engineering effort. Raj Patel tells you that the engineering team is fully committed to product roadmap items for the next two quarters and that he cannot prioritize the optimization without removing something else from the roadmap. In one paragraph, describe how you would think through

the tradeoff, what you would ask Raj to compare the optimization against, and what your recommendation would be if the alternative is delaying a product feature by one week.

20. Scenario: An institutional investor evaluating Series C diligence asks for the company's "AI in finance program" documentation, having heard about the workflows in informal conversation. They specifically want to understand whether the program is real (i.e., consistently operational) or a marketing artifact. In one paragraph, describe what you would provide as evidence of operational discipline, what artifacts would be most persuasive, and what posture you would take in the conversation.

Appendix B · Answer Key with Explanations

Multiple Choice Answers

Question 1: (b)

The workflow is sequenced fifth because operational expense optimization requires the team capacity and discipline that the prior four workflows have built. Running this workflow first would surface opportunities the operational function would not have time to execute.

Question 2: (b)

Tier Two because the outputs are operational recommendations that do not influence financial reporting directly. The execution of any recommendation has its own approval process appropriate to its magnitude.

Question 3: (b)

Tom Reeves (Director of Operations) is the principal operational owner. Raj Patel (CTO) is the technical reviewer for AWS-specific recommendations. John Campbell (CFO) signs off on recommendations above \$25K in annual value.

Question 4: (b)

Instances i-017 through i-020 are flagged as over-provisioned production resources running at 12-22% CPU. Right-sizing would save approximately \$26K annually.

Question 5: (c)

Ten idle resources total: eight EC2 instances (i-021 through i-028) plus two RDS databases (rds-003 and rds-004). Termination would save approximately \$29K annually.

Question 6: (b)

Gong (41% utilization), Outreach (48%), ZoomInfo (40%), and Loom (40%) are all under-utilized. Right-sizing seats would save approximately \$30-35K annually.

Question 7: (a)

Stratosphere Software and Stratosphere Inc are the same underlying vendor billed under different entity names. Consolidation would save approximately \$24K annually.

Question 8: (c)

AWS (\$1.98M), SF Office Landlord (\$360K), Google Ads (\$240K), LinkedIn Ads (\$180K), Recruiting Partners (\$120K) total approximately \$2.88M of approximately \$4.4M total non-payroll OpEx, roughly 65%.

Question 9: (b)

AWS right-sizing (\$26K) + AWS idle termination (\$29K) + software seats (\$35K) + Stratosphere consolidation (\$24K) = approximately \$100-110K annually.

Question 10: (c)

The workflow runs quarterly. Patterns change slowly; monthly would add review burden without value, annual would miss compounding opportunities.

Question 11: (b)

A realization rate of 60-75% against identified opportunities is healthy. Lower suggests deprioritization; higher suggests over-conservatism in identification.

Question 12: (b)

Approximately 6-8 weeks of cumulative cash runway extension over 12 months of disciplined operation. Material at the Series C decision point.

Short Answer Explanations

13. Tier Two rather than Tier Three

The Tier classification depends on whether the workflow's outputs influence financial reporting directly. The workflows in Parts 5, 6, 7, and 8 produced outputs that flowed into close data, board commentary, forecasts, and revenue plans, all of which influence financial reporting. The Infrastructure and Vendor Intelligence workflow produces operational recommendations that affect future operating expense when executed, but the recommendations themselves do not enter the financial statements. The execution of each recommendation has its own approval process. The lower tier classification reflects this difference.

14. Why load the framework first

The spend categorization framework establishes the structure the model will use to organize its analysis and recommendations. Without the framework, the model would invent its own categorization, which would likely be reasonable but would not match how the operational function actually thinks about expense. Loading the framework first aligns the workflow output with the operational team's conceptual model, which is what makes the recommendations directly actionable rather than requiring reorganization before action.

15. Why run quarterly despite modest single-cycle impact

The value of this workflow is cumulative rather than incremental. A single quarter's identified savings of \$100K extending runway by two weeks is modest in isolation. The same workflow run quarterly for four quarters identifies different opportunities each time as the data evolves, with realized savings compounding into materially larger numbers. Over the eighteen-month horizon to Series C, the cumulative effect is meaningful operating margin protection. The discipline is the value, not any single cycle.

16. Distributed human review

Operational expense decisions have distributed ownership because the people closest to each category have the deepest context. Raj knows what is safe to right-size in AWS production; Tom knows the vendor relationships; tool owners know what their teams actually need from each subscription. Centralizing review with one reviewer (the CFO) would either bottleneck decisions or require the CFO to make decisions on areas where they lack context. The distributed review pattern reflects the actual organizational structure of operational decision-making.

17. Ticket queue versus narrative

Narrative outputs are appropriate for executive consumption when the goal is understanding context and judgment. Operational outputs are appropriate when the goal is action by specific people. The Infrastructure and Vendor Intelligence workflow produces actions that need to be assigned, tracked, and closed; tickets fit this pattern naturally because the operational team already runs on a ticket-based workflow. A narrative output would require translation to operational artifacts before it could be acted on, adding friction. The ticket queue minimizes the gap between workflow output and operational execution.

Scenario Discussions

18. Realization rate below target

The 45% realization rate is a signal that something in the operational system is preventing execution, and the explanation Tom provides (overlapping with close cycles) is plausible and operational rather than philosophical. The first response is to validate the explanation: confirm that the realization gap is concentrated in cycles where Tom's team is most stretched, rather than spread evenly across cycles. The change to consider is twofold. First, shift the workflow cadence away from quarter-end close periods to mid-quarter, where the operational team has more capacity to execute on the opportunities. Mid-quarter execution also allows realized savings to be reflected in the next close package, which strengthens the connection between the workflow output and the board commentary. Second, consider whether the workflow should produce fewer but higher-impact opportunities per cycle, so that the realization rate on a smaller set of priorities is higher than the current rate on a larger set. The governance step is to document the timing change formally in the workflow registry and to share the revised cadence with the Governance Working Group as a deliberate operational refinement rather than an ad hoc adjustment. The deeper principle is that workflow timing matters as much as workflow design; running the workflow when the operational team can act on it is what converts identified savings into realized savings.

19. AWS optimization versus product feature tradeoff

The framing of the question matters. Raj is presenting a binary tradeoff (optimization or product feature) when the real comparison is multidimensional. Ask Raj to compare the \$40K annual savings (which is approximately \$200K in cumulative cash runway value over the five-year horizon, plus the precedent it sets for future optimization discipline) against the specific feature that would be delayed. Specifically, ask three questions. First, what is the customer value of the feature being delayed by one week, and what is the revenue impact if delayed? Second, is there a feature that could be reprioritized to capture both the optimization and the original roadmap commitment? Third, can the optimization be executed by a part-time effort over four weeks rather than a concentrated week, reducing the displacement of any single feature? The recommendation depends on the answers, but the executive principle is that operational optimization and product roadmap are not separate categories competing for the same resource; they are both investments competing for engineering attention, and the comparison should be made on the same terms. If the feature being delayed is genuinely time-sensitive (customer commitment, competitive response), defer the optimization to the next quarter. If the feature is one of many items on a list, the optimization has the higher ratio of certain savings to engineering effort, and it should win. The conversation with Raj should be a conversation, not a directive from finance.

20. Series C investor diligence request

The right response demonstrates evidence of operational discipline rather than narrative description of it. Provide four artifacts. First, the workflow registry showing the five workflows registered under the AI governance framework, with their risk classifications, approval dates, and named reviewers. Second, audit trail samples from the trailing quarter for each of the five workflows, demonstrating that the workflows are actually running rather than theoretical. Third, the documented business outcomes attributable to the program: the realized OpEx savings from the workflow in this part, the forecast accuracy improvements from Part 8, the close cycle compression from Part 5, the board commentary discipline from Part 6, the pipeline forecast adjustment from Part 7. Fourth, the AI Governance Charter from Part 4 plus the periodic review meeting minutes showing that the program is governed continuously rather than at deployment. The framing of the conversation is that this is operational discipline being applied to AI work in finance, no different in kind from operational discipline applied to other finance activities. The posture is confident transparency rather than promotional language. A sophisticated investor will recognize the difference between a real program with imperfect realization and a marketing artifact with polished slides. The artifacts demonstrate the former.