eFuturesCFO Masterclass Series

AI Workflows for the Modern CFO

PART 2

The Discovery

Thirty Days of Conversations at Helix Cloud Systems

Hindol Datta

Chief Content and Product Officer, eFuturesCFO

A Note Before Part 2

Part 1 was the foundation. You now have a working understanding of how large language models operate, how they are accessed, how they integrate with enterprise systems, and where the governance perimeter sits. Part 2 is the application. Specifically, it is the application of that foundation to the most demanding executive context a finance leader will ever face: the first thirty days in a new CFO role at a company that has not yet thought clearly about AI but is already using it.

You will meet John Campbell, who joined Helix Cloud Systems as CFO on a Tuesday in late winter. Helix is a Series B B2B SaaS company with $22.4 million in annual recurring revenue, 142 employees, and a recent quarter that disappointed the board. John inherited a finance function staffed by a competent controller, a thoughtful senior FP&A; manager, and very little else. He inherited an organization that had begun experimenting with AI in an uncoordinated way, with one pilot already running in revenue operations and a general air of executive curiosity that had not yet been channeled into anything productive.

Over his first thirty days, John conducted nine structured conversations across the executive team. He did not announce any initiatives. He did not propose any architecture. He listened. He asked careful questions. He made notes. By the end of the thirty days, he had identified the patterns that would shape the next eighteen months of his tenure.

What follows is a curated and lightly stylized account of those nine conversations. The dialogue is not a verbatim transcript; no executive remembers any conversation verbatim. The dialogue is a faithful reconstruction of the substance and the texture of what was said, organized to highlight the executive skill of listening across functions while carrying a coherent technical and governance framework in your head.

Read it slowly. Watch how John applies the foundation from Part 1 in real time. Notice which concepts surface in which conversations. Notice what he chooses to push on and what he lets pass. The observation is half of the executive practice. The next part, Part 3, is where John converts what he has heard into a written architectural framework. The framework cannot exist without the listening.

> *The first thirty days are not for deciding. The first thirty days are for hearing the system clearly enough that the decisions made later are decisions that can hold.*

Hindol Datta

Contents

Each chapter contains the substance of a single conversation, followed by John's brief reflection. The closing chapter synthesizes what he learned across all nine.

Preface · Helix Cloud Systems and Its Executive Team

Helix Cloud Systems is a workflow automation platform for mid-market operations teams. The product helps customers automate the handoffs between systems that mid-market companies depend on but have never been able to fully integrate. Sarah Chen and Raj Patel co-founded the company in March 2019, after leaving senior roles at a larger enterprise software firm where they had both watched their employer struggle with exactly the integration problem Helix now solves.

The company grew steadily from inception. Series A closed in September 2021 at a $35 million valuation, led by a regional firm in Palo Alto. Series B closed in June 2024, raising $42 million led by Meridian Growth Partners, at a $180 million pre-money valuation. The Series B was raised on the strength of 51 percent year-over-year ARR growth and a clean balance sheet. The company has not raised since.

At the time John Campbell joined, Helix served 410 customers, employed 142 people, and was generating $22.4 million in annual recurring revenue. Headquarters were in SoMa San Francisco, with a small engineering office in Austin and a two-person commercial presence in Dublin. The auditor was Marsh and Henning LLP, a mid-tier firm with strong SaaS practice depth, though Diana Reyes-Okonkwo, the audit committee chair, had begun to wonder whether the company would need to move to a Big Four firm before the next round of capital.

The board consisted of six members. Sarah Chen and Raj Patel represented the founders. Marcus Whitfield held the lead director seat as the Meridian partner who led the Series B. A partner from the Series A firm held a second investor seat. Diana Reyes-Okonkwo was the independent director and audit committee chair, recruited at the time of the Series B from her prior role as CFO of a public SaaS company. A second independent seat remained open.

Q3 of the year John joined had been disappointing. The company missed plan by 8 percent, driven primarily by a $1.2 million enterprise deal that had slipped from Q3 into Q4. The board had been measured in its response, but Diana had asked several pointed questions about pipeline reliability that had not been satisfyingly answered. The Q4 forecast carried meaningful risk. The cash runway, at 22 months, was comfortable but not generous; another soft quarter would compress the window in which the next round of capital could be raised on favorable terms.

Aaron Schultz, the interim CFO consultant who had filled the role for the previous five months, departed the day John arrived. He had been kind in his transition memo, identifying the controller and the senior FP&A; manager as the two strongest individuals in the finance organization, and flagging that the broader business was beginning to experiment with AI in ways that no one was tracking centrally. He recommended that John establish his own view of what was already underway before assuming anything.

The executives John would meet

Sarah Chen, the CEO, was 41 years old, intense, and direct. She had been an engineering product manager before founding the company. She read widely, slept little, and held her executives to a standard that several had described to John as exacting but fair.

Raj Patel, the CTO and Sarah's co-founder, was 44 years old and temperamentally the opposite of Sarah. Where she pressed, he absorbed. He had built the engineering organization from scratch and now led 38 engineers, with a strong technical reputation inside the company and a habit of asking questions in meetings that exposed flawed assumptions without ever sounding accusatory.

Michael O'Brien, the chief revenue officer, had joined fourteen months earlier from a Series D peer company. He was 49, gregarious, and had brought several of his prior sales leaders with him. The sales team had grown from 18 to 34 in his tenure. His relationship with the rest of the executive team was warm but professionally measured. He had personally approved the AI pilot in revenue operations that John had heard about before joining.

Priya Sundaram led customer success. She had been employee number 11 at Helix and had built the customer success function from three people to nineteen. She was thoughtful, slightly anxious, and held more institutional knowledge about the customer base than anyone else in the company. Her function reported into Michael, which Sarah had identified to John as a structure that might need revisiting.

Karen Lindqvist, the chief human resources officer, had joined eight months earlier from a larger company. She was building the people function from a foundation that had been thin. She was measured, deliberate, and had begun to map the company's capacity-to-revenue relationships in a way that suggested she thought systemically about her function.

David Kim, the controller, had been at Helix for two and a half years. He was 38, methodical, deeply competent at the close, and had earned the respect of the audit committee through clean audits in each of the prior two years. He had been the de facto finance leader during much of Aaron's interim tenure.

Elena Vargas, the senior FP&A; manager, reported into David. She was 33, sharp, ambitious, and the person who actually built the forecast each quarter. She was also the person most visibly frustrated by the gap between what the finance function was asked to produce and the tools available to produce it. John had been told by Aaron that Elena was either the future head of FP&A; at Helix or, if not given the right opportunity, the person who would leave first.

Tom Reeves led operations, which at Helix meant procurement, vendor management, infrastructure spend, and a portion of what a larger company would call business operations. He was 52, pragmatic, and the only executive over 50 in a company whose average age was in the early thirties.

Naomi Bridges was the general counsel, working three days a week on a fractional basis. She had previously been in-house at a fintech company and now served three growth-stage clients. She was sharp, slightly impatient with founders who treated legal as an afterthought, and had a particular interest in the

regulatory environment around AI that had been intensifying through 2025 and into 2026.

Wei Zhao led security and IT. He was 35, came from a security background at a public company, and had been hired explicitly because the Series B board had insisted on a real security function before the company grew further.

Diana Reyes-Okonkwo, the audit committee chair, would be the last of the nine conversations. She was 58, had been a public company CFO for seven years before retiring two years earlier, and now served on three boards. She had been the lead voice on the board in raising the bar on financial discipline at Helix, and Sarah had described her to John, with respect, as the toughest interview of any board meeting.

These were the nine people John would meet. The conversations that follow are organized in the order they happened.

Chapter 1

Sarah Chen, Chief Executive Officer

On strategy, the board, and what AI is doing in our company already

Day 2 · Wednesday morning · CEO office · forty-five minutes

Sarah Chen's office was a corner room on the third floor of the Helix headquarters building, sparsely furnished, with a single whiteboard covered in handwritten notes from what looked like an engineering planning session. She was already at her desk when John arrived. She closed her laptop and gestured to the small round table by the window.

SARAH

Good. I have forty-five minutes. I want to spend most of it listening to what you think, but let me start with three things I want you to know, and then we can get into whatever you want.

JOHN

Please.

SARAH

First. The board is not unhappy with the company. They are unhappy with the predictability of the company. There is a difference. Marcus and Diana both made it clear at the last meeting that they can live with a quarter that misses by 8 percent. What they cannot live with is a forecast that says we are going to hit plan and then a phone call two weeks before close that says we are not. That is the issue. The board feels they are being managed by surprise rather than by visibility.

JOHN

How many quarters has that happened?

SARAH

In the last six quarters, we have missed our internal plan three times and missed the board commitment twice. The pattern of being surprised by the miss late in the quarter is what is wearing on them. Not the absolute numbers.

JOHN

I appreciate the framing. The board reaction is rational. Two surprises in six quarters is not bad, but the question is whether anyone on our side of the table is surprised by the same thing the board is surprised by, which is a forecast that does not hold. If we are calibrated and the forecast just has noise in it, that is one problem. If we are not calibrated and the forecast is systematically optimistic, that is a different problem.

Sarah looked at him for a beat before nodding.

SARAH

That is exactly the question Diana keeps asking, and I have not had a clean answer. Michael will tell you the pipeline is fine and the deals slip for reasons no one can predict. That is true some of the time. I do not think it is true all of the time. I want you to figure out which is which. Second thing.

JOHN

Please.

SARAH

We are doing AI. We are doing it badly, but we are doing it. Michael started a pilot in revenue operations about four months ago that uses an AI tool to summarize sales calls and write the follow-up emails. He claims it has improved his team's productivity by something he occasionally quotes at 20 percent and other times at 35 percent. I have not seen the methodology. I have asked twice. He says it is being measured.

JOHN

Anyone else?

SARAH

Raj has been experimenting on the engineering side. Some of it is internal tooling, code review assistants, documentation generation. That I am less worried about because it is contained within engineering and Raj has a real security posture. But there are people in marketing using AI to draft content. There are people in customer success using it to draft customer-facing messages. There is a finance person, I am told, who has been using ChatGPT to draft variance commentary. No one is tracking any of this centrally. I do not know what data is going where. I do not know what we have committed to which vendors. I do not know what we are paying or whether we are paying.

JOHN

When did this start?

SARAH

Some of it has been happening for over a year. Most of it ramped up in the last six months. We have no policy. We have no governance. We have a company that is experimenting with a technology that has real risk attached to it, and we have no one whose job is to look at the experimentation and make a coherent set of decisions about what we should and should not be doing.

JOHN

Is that an explicit assignment for the CFO role?

SARAH

It is now. I do not want it to be the entire job. I want it to be a meaningful slice of the job. I think the rest of it will follow from getting AI right, because the answer to predictability and the answer to visibility and the answer to operational quality all have AI components.

JOHN

I would tend to agree with that. Third thing.

SARAH

Third thing. Cash. We have 22 months of runway. We planned the Series B on the assumption we would not need to raise again until the company was at a genuinely scaled point. Series C economics for SaaS right now favor companies with very predictable growth and very efficient go-to-market. If we are going to raise on terms that do not embarrass us, we need to be that kind of company, not just claim to be. The window is roughly twelve months from now. Maybe fourteen. I would like us to have a forecast that the board trusts and a finance function that the next investor trusts, and I would like both of those to be true by Q4.

JOHN

Understood. That is enough to work with for now. Can I ask you something different?

SARAH

Yes.

JOHN

When you imagine the finance function eighteen months from now, what does it do that it does not do today?

Sarah paused. She looked out the window for a long moment, then turned back.

SARAH

It tells me what is going to happen before it happens. Not just the numbers. The patterns. What is breaking. What is about to break. Which customer is going to churn before they tell us. Which deal is going to slip before Michael calls me. Which expense is going to be twice what we forecasted before the bill comes in. I do not need a finance function that prepares beautiful board decks. I need a finance function that sees around corners. The board decks are the artifact, not the value.

JOHN

That is the right answer.

SARAH

I have read your background. I picked you because of the build at your last company. The AI work. I did not pick you for it, exactly, but it was a meaningful part of why you stood out. I am not asking you to be an AI expert. I am asking you to be a CFO who is fluent enough in this technology to actually govern it. There is a difference.

JOHN

There is. I will not pretend to be the technical authority. I will own the governance and the architecture as it touches finance. Raj will need to be a real partner. So will Naomi.

SARAH

They will be. Both have already been told. Raj is expecting you on Friday. Naomi I can introduce you to next week. Anything you need from me in the meantime?

JOHN

Time. Not yours. Mine. I want thirty days before I propose anything. I want to talk to nine or ten people, including the audit committee chair, and I want to write down what I have heard before I commit to anything. If you need me to put out a fire in week three, I will do it, but I would rather have thirty days of listening first.

SARAH

Take the thirty days. I have waited longer than that for a real CFO. Another month is fine.

She stood. The meeting was over.

SARAH

One last thing.

JOHN

Yes.

SARAH

When you do propose something, I want it in writing. I do not want a slide deck. I want a memo. I want to read your thinking. I have learned that the people whose thinking holds up in prose are the people whose decisions hold up in practice.

JOHN

You will get a memo.

John's reflection that evening

Sarah is direct, which is a relief. She has framed the role correctly. The board wants predictability. The company is experimenting with AI in an uncoordinated way and someone needs to bring coherence to that. The Series C window is twelve to fourteen months out and the bar for that round is rising.

What she said about seeing around corners is interesting and worth holding onto. She is not asking for a better forecast in the conventional sense. She is asking for a forecast that incorporates leading indicators the current process does not capture. That is a different architecture from what is in place. It implies pulling signal from customer success data, from product usage, from pipeline velocity, from spend patterns. Each of those lives in a different system. The integration burden is real.

She asked for a memo, not a deck. That is a serious signal about how she thinks. She wants to read the reasoning. The memo will have to be good.

I should pay attention in the Michael conversation to whether the AI pilot in revenue operations is producing real productivity or producing the appearance of productivity. The two are easy to confuse, and the 20-to-35 percent range Sarah quoted is the kind of range that suggests the measurement is not being done.

Chapter 2

Raj Patel, Chief Technology Officer

On the data stack, the AI environment, and what already exists

Day 4 · Friday afternoon · Engineering wing · ninety minutes

Raj Patel's office sat at the far end of the engineering wing, past three pods of open desks where most of the engineering team worked. His office had no door, by deliberate choice, and was visible from anywhere in the open space. He was at a standing desk when John arrived, reviewing a pull request. He closed the laptop with no apparent urgency and gestured to two chairs in the corner of the room.

RAJ

Sarah told me you wanted to talk about the data stack and the AI work. She also told me you have done some of this before. So let me know where you want to start and I will skip what you already know.

JOHN

Start with the data stack. The honest version. I have been a CFO at two companies and I have read a hundred data architecture diagrams, most of which were aspirational. I would rather hear what actually exists and what does not.

Raj nodded slowly. The question seemed to please him.

RAJ

Okay. The systems of record are the standard set. Salesforce for CRM. We are on the Sales Cloud, plus a limited Service Cloud deployment for support cases. NetSuite for ERP and financials. ADP for payroll. Gusto for benefits. We migrated off Carta a year ago and our cap table is now in Pulley. The HRIS is a smaller product called Rippling. The product itself runs on AWS, primarily us-west-2, with a smaller footprint in eu-west-1 for our European customers. Standard observability stack, Datadog and PagerDuty. Standard code repository, GitHub. Standard CI/CD, GitHub Actions and a small Argo workflows deployment for our internal orchestration.

JOHN

Data warehouse?

RAJ

Snowflake. We started loading data into it about eighteen months ago. There is a Fivetran connector pulling from Salesforce and NetSuite. There is a custom pipeline pulling product event data from our application tier. The warehouse is real and it is being used. The BI tool on top of it is Looker.

JOHN

Quality of the data in the warehouse?

RAJ

Variable. Salesforce data is high-quality at the deal level because Michael's team operates with discipline. It is lower quality at the contact level because nobody has owned that. NetSuite data is clean. Product event data is voluminous but inconsistent across product versions; we have changed our schema twice since we started instrumenting, and the historical data has gaps. The warehouse is good for current-state reporting. It is harder to use for longitudinal analysis without effort.

JOHN

Who owns the warehouse?

RAJ

Officially, a data engineer named Jordan who reports into me. In practice, the warehouse is used by Elena Vargas in FP&A; more than by anyone else. She writes most of the Looker dashboards. Jordan keeps the pipelines running but is not building analytical content.

John made a note. The relationship between Elena and the warehouse was already emerging as a significant fact. The most data-fluent person in the company on a daily basis was a senior FP&A; manager, not someone with the title to match.

JOHN

Let me switch to AI. What is actually deployed?

RAJ

In production, two things. The first is GitHub Copilot for the engineering team. We are on the enterprise tier with the data handling commitments. The second is the pilot Michael runs in revenue operations, which uses a product called Cadenza that wraps OpenAI and does call transcription and email generation. That has been in production for about four months. The license is on their card.

JOHN

Whose card?

RAJ

Michael's. Yes. We have not put it on the corporate AmEx because they would not give us a master agreement on terms we could justify, and Michael decided to expense it month by month rather than wait.

JOHN

How much per month?

RAJ

About 4,200 dollars at the current usage. It has been growing.

JOHN

Beyond those two?

RAJ

Outside of production, a lot. Several engineers, including me, have Claude and ChatGPT subscriptions on personal accounts that get expensed. Marketing has a team subscription to ChatGPT. Customer success has been using Claude through the consumer interface to draft customer messages, with no policy framework. Elena has been using ChatGPT to draft variance commentary. I am told there is at least one person in finance using AI to do reconciliation work, but that is hearsay; I have not confirmed it.

JOHN

Data going where?

RAJ

That is the question. The Claude and ChatGPT consumer interfaces do not, by default, train on our data, but there is no enterprise-grade data handling agreement in place. The marketing team subscription is at least a team account with some commitments. The Cadenza pilot has a data processing agreement; Michael had to agree to it but I am not sure he read it.

JOHN

API keys? Are there any keys outstanding?

RAJ

There are two keys I know about. One is mine, attached to the engineering account at Anthropic, used for the internal tooling experiments. The other is on the Cadenza vendor account, not on ours. Beyond those, I cannot say for certain there are no other keys floating around. We have not done an audit.

JOHN

Connectors?

RAJ

No connectors deployed into our enterprise systems. The Cadenza tool uses its own Salesforce integration. We have not connected any AI directly to Salesforce or NetSuite or the warehouse.

JOHN

MCP?

Raj looked at him sharply. It was the first time in the conversation that John had used a term that surprised him.

RAJ

You know about MCP.

JOHN

I know what it is. I do not pretend to know how to build an MCP server. I know that anything we deploy should probably go through the protocol rather than through custom connectors.

RAJ

I agree. We have not deployed any MCP yet. We have talked about it. There is an obvious use case for an MCP server in front of Salesforce, and another in front of the warehouse, but neither has been built. That is engineering work I would do happily but I have not had the budget or the executive cover to prioritize it.

JOHN

What about security posture for any of this?

RAJ

Wei Zhao runs security. You should talk to him. The short version is that we have basic posture in place: SSO, MFA, least-privilege access for production, audit logging on all access to NetSuite. We have a SOC 2 Type 2 report. The AI work largely sits outside this posture today because it is not deployed in production-grade ways. That is part of why it makes me nervous.

JOHN

It should make you nervous.

RAJ

It does.

John let the silence sit for a moment. Then he asked a different question.

JOHN

If you had executive cover and budget to build the right AI environment for finance, what would you build first?

RAJ

Three things. First, an MCP server in front of the data warehouse, with proper read-only access controls and a row-level security layer. That gives any AI tool we deploy a single, governed door into our data. Second, a structured prompt and workflow registry, where every AI workflow we deploy is registered, versioned, and monitored. So we know what is running, what it is doing, and what it has produced. Third, an audit log that captures every AI call: who made it, what was sent, what was returned, what action was taken. The audit log is the boring foundation that lets everything else be defensible.

JOHN

How long to build all three?

RAJ

Three months for a usable version of all three with one senior engineer dedicated. Six months if we want enterprise-grade. The shorter version would give you enough to start running serious workflows on top of it within the year.

JOHN

You have that senior engineer?

RAJ

I can free one. If you tell me this is a priority and Sarah agrees, I can have someone start in two weeks.

JOHN

I am not committing to anything today. But I would like to hear that proposal in writing. A page or two. What you would build, what it would cost, what it would enable.

RAJ

I will have it to you Monday.

JOHN

Last question. Is there anything I should know that I have not asked?

Raj considered this for a long moment.

RAJ

Two things. The first is that we are about to lose Elena Vargas if we do not give her something bigger to work on. I see her in the warehouse every day. She is doing work that is several tiers above her title and compensation. The second is that I think Michael genuinely believes his AI pilot is working, but the productivity numbers he cites are not coming from any measurement system I am aware of. They are coming from his sense of how his team feels. That is not nothing. It is also not data.

JOHN

Thank you. Both of those are useful.

John's reflection that evening

Raj is what I hoped he would be. Technically deep, honest about gaps, willing to admit uncertainty. The three things he proposed to build are essentially right. An MCP server in front of the warehouse, a workflow registry, and an audit log. If we build those three things, almost every AI workflow we eventually want becomes feasible. Without them, we are accumulating risk faster than we are accumulating capability.

The AI inventory is messier than I expected. One sanctioned tool, Copilot for engineering. One sanctioned pilot, Cadenza for revenue operations, on Michael's personal expense card. A considerable amount of unsanctioned shadow usage across marketing, customer success, and finance. No central inventory, no policy, no data handling framework. That is a governance gap I need to address before anything else, but I need to address it without shutting down experimentation that is producing value.

The Elena observation is significant. Raj has now flagged it too, after Aaron flagged it in his transition memo. Two independent observations of the same thing usually means it is real. I should make a point of meeting Elena directly during the David conversation rather than meeting her separately, so she does not feel singled out. But I should hear her thinking, not just David's thinking.

On Michael's pilot. Raj has confirmed my suspicion that the 20-to-35 percent productivity number is not measured. I will need to ask Michael for the methodology directly, and not let the conversation drift past it.

Chapter 3

Michael O'Brien, Chief Revenue Officer

On the pipeline, the forecast, the Q3 miss, and an AI pilot

Day 6 · Tuesday afternoon · CRO office · sixty minutes

Michael O'Brien's office was on the second floor, next to the sales pod. He was on a video call when John walked in, finished it briskly, and came around the desk with the practiced warmth of an experienced sales leader.

MICHAEL

John. Glad we are doing this. I want to give you a real picture of the commercial side and I want to hear what is on your mind.

JOHN

Appreciate it. Let me start broad. Walk me through how the sales motion works today.

MICHAEL

We have two motions. The first is mid-market inbound, sub-50K ACV, mostly self-serve evaluation, closed by a six-person inside sales team. The second is mid-market plus enterprise outbound, 50K to 250K ACV, run by twelve outside reps, with a small enterprise overlay of three people doing deals north of 200K. Average sales cycle is 47 days inbound, 96 days outbound, 140 days enterprise. Win rate is 28 percent overall, higher on inbound, lower on outbound.

JOHN

Pipeline coverage?

MICHAEL

We aim for 3.5x going into a quarter. We were at 3.1x going into Q3 and we ended at 2.9x going into Q4. That is part of what is making everyone nervous about Q4.

JOHN

Tell me about the Q3 miss.

Michael leaned back. The question was clearly one he had answered before.

MICHAEL

Q3 missed by 1.4 million on plan. The single biggest factor was a deal at Wexford Insurance that was 1.2 million ACV, supposed to close in mid-September. Their CFO went out on medical leave in early September and the procurement review got paused. We have it now in Q4, October close date. The deal is real. They have signed an order form pending one final security review.

JOHN

Is the security review on track?

MICHAEL

It is moving. Wei is on it. I do not have visibility into whether anything has surprised them. The customer is engaged.

JOHN

And outside Wexford?

MICHAEL

Six smaller deals slipped by between two weeks and two months. None individually were large. Aggregate impact was about 200K. Of those six, four have since closed, one is still active, one we lost.

JOHN

Why did Wexford surprise us late?

It was the question Sarah had asked him to push on. Michael paused.

MICHAEL

Honestly? Because the rep believed the close date until the day the CFO went on leave. The customer had given verbal commitment. The order form was sitting in their procurement queue. We were forecasting 90 percent confidence. The thing that happened was outside what we could have forecasted.

JOHN

What confidence do you forecast deals at?

MICHAEL

Reps assign stage probabilities. Stage 1 is 10 percent, stage 2 is 25, stage 3 is 50, stage 4 is 75, stage 5 is 90. Verbal commitment plus order form in procurement is stage 5.

JOHN

How often does a stage 5 deal slip?

A longer pause this time. John could see Michael calculating.

MICHAEL

Honestly, I do not have that number off the top of my head. We have stage 5 deals slipping sometimes. I would have to ask Lisa to pull the data.

JOHN

Lisa is your sales operations lead?

MICHAEL

Yes. Lisa Mahoney. She runs the data side.

JOHN

Can you have her pull the slip rate by stage for the last eight quarters? I would like to look at it.

MICHAEL

Absolutely. I will have it Friday.

John watched Michael for a moment. The CRO was responsive, but something in the cadence suggested he had been asked similar questions before and the data had not quite been pulled previously. The note John made to himself was that the absence of stage-slip data over an eight-quarter window was itself a signal about how the forecast was being managed.

JOHN

Let me ask about the AI pilot.

MICHAEL

Cadenza. Yes. So we onboarded Cadenza about four months ago. It is a sales productivity tool. It records calls, generates summaries, and drafts follow-up emails. The reps can review the email, modify it, and send it. The summary goes into Salesforce automatically. We have rolled it out to the outbound and enterprise teams. About 22 reps are using it actively.

JOHN

Sarah mentioned you have a productivity figure associated with it.

MICHAEL

We have seen meaningful improvement. I have been using a figure of 20 to 25 percent.

JOHN

How is that measured?

Michael smiled slightly. He had been expecting the question.

MICHAEL

It is not measured rigorously. I will be honest with you. The figure comes from rep feedback and from my observation of how the team is operating. I have reps who tell me they are getting an extra hour back per day. I am extrapolating from that. I have not run a controlled measurement.

JOHN

Has the productivity translated to results?

MICHAEL

It is hard to say with confidence. Our outbound team's activity metrics are up. The number of meetings booked is up modestly. The number of opportunities created is roughly flat year over year. The win rate is flat. So if the tool is producing the productivity I think it is, the effect on results has not yet shown up cleanly in the funnel.

JOHN

What is the monthly cost?

MICHAEL

Four-two a month and rising. I have been expensing it.

JOHN

Why expensing rather than corporate billing?

MICHAEL

Their pricing structure for a master agreement would have committed us to a year. I did not want to do that until I was more confident. So I have been paying monthly through expense.

JOHN

That is reasonable, but we will need to put it on a sanctioned footing soon. If we are going to keep using it, we should have a master agreement with proper data handling commitments, and the spend should be on the corporate side rather than on your card. Independent of the productivity question.

MICHAEL

Agreed. I have been waiting for a CFO to push on this. Aaron was not going to.

JOHN

Do you have visibility into what data the tool is sending where? Sales call transcripts contain sensitive customer information.

MICHAEL

They have a DPA. I have it. I did read it. The short version is that they process audio and transcripts through OpenAI under enterprise terms with no training on our data. Audio is retained for 30 days for quality. Transcripts are retained as long as we are using the service.

JOHN

I would like to share that DPA with Naomi.

MICHAEL

Of course.

JOHN

One more topic. The pipeline forecasting question. Sarah is worried about predictability. The board is worried about predictability. If you and I were going to make the forecast more reliable in the next six months, what would we change?

Michael considered the question.

MICHAEL

I would do three things. First, I would tighten the rep discipline on stage advancement. Reps are sometimes advancing deals into later stages based on relationship confidence rather than on specific gating criteria. I would implement stricter gating. Second, I would build better analytics on historical slip rates, so we can apply real probability weights rather than the static stage percentages. Third, I would invest in some form of AI-assisted deal review. There are tools that analyze deal histories and flag deals that look like they are going to slip. I have not deployed one. I would consider it.

JOHN

Do you know of one you would recommend?

MICHAEL

I have looked at three. Some are mature, some are not. I would want to talk to someone in finance about it before committing.

JOHN

I would like to be that someone.

MICHAEL

Done.

John's reflection that evening

Michael is competent and self-aware. He understands the forecast problem. He understands the AI pilot situation. He has been waiting, in his own words, for a CFO to push on governance, and he was glad to be pushed.

The data point I most need is the stage slip rate over eight quarters. I will see what comes back on Friday. If reps are advancing deals into stage 5 based on confidence rather than on objective criteria, and if stage 5 deals are slipping at a meaningful rate, then the forecast is systematically optimistic and that is the source of the board's concern. The Wexford deal might be the surface event, but the underlying pattern is what we need to surface.

The Cadenza pilot is real and is producing some value, but the value has not yet been measured rigorously. We need to either measure it or stop pretending we know what it is. The governance gap on the tool itself is significant: a CRO's personal expense card paying for an AI service that is processing our customers' call audio is a configuration that should not exist in a company moving toward Series C. We will fix that quickly.

Michael mentioned AI-assisted deal review as a future possibility. That is interesting and worth noting. If we can build pipeline intelligence that flags slipping deals before the rep realizes they are slipping, we directly address the predictability problem Sarah is most focused on. That is a workflow worth designing properly.

Chapter 4

Priya Sundaram, Vice President of Customer Success

On churn, expansion, and the early warning signs no one is reading

Day 9 · Friday morning · CS team area · seventy minutes

Priya Sundaram's desk was in the customer success pod rather than in a private office. She had told John she would prefer to meet in one of the small conference rooms, and when he arrived she was already there, two laptops open, with a small notebook beside her.

PRIYA

I have prepared for this conversation more than I probably should have. I apologize in advance.

JOHN

No apology needed. I would rather have a prepared CS leader than an unprepared one.

PRIYA

I have wanted to talk to a CFO about customer success data for a long time. Aaron was not interested. Sarah is interested but does not have the time. I am hoping you have both the interest and the time.

JOHN

Probably. Walk me through the function.

PRIYA

We have 19 people. 13 customer success managers, 4 support specialists, and 2 on the renewals team. We have 410 customers. CSMs carry between 25 and 60 accounts each, depending on tier. We segment into enterprise, mid-market, and growth. Enterprise CSMs carry 25 accounts each. Mid-market CSMs carry 40. Growth CSMs carry 60.

JOHN

NDR?

PRIYA

118 percent overall, last twelve months. Higher in enterprise, around 132 percent. Lower in growth, around 104 percent. Mid-market is in between at 115.

JOHN

Gross retention?

PRIYA

91 percent overall. That is the number I am less happy about. Enterprise is 96 percent. Mid-market is 92. Growth is 84. The growth segment is where our churn lives.

JOHN

Why is growth churn high?

PRIYA

Three reasons that I think are real. First, the growth segment buys on shorter cycles and lower commitment, so the switching cost is naturally lower. Second, our product fits enterprise and mid-market use cases better than it fits growth use cases; the workflow automation we deliver is most valuable when the customer has enough operational complexity to need it, and many growth customers do not. Third, we have not invested in growth-segment customer success at the same intensity. The economics make it hard to justify.

JOHN

Do you predict churn?

Priya hesitated.

PRIYA

I have indicators. I do not have a real prediction model. The indicators are health scoring that we do manually. CSMs rate their accounts each quarter on a green-yellow-red basis. The rating is subjective. It tracks actual outcomes only roughly.

JOHN

How roughly?

PRIYA

Last year, about 60 percent of accounts that churned had been rated green or yellow in their last health review. Forty percent had been red. So we caught roughly 40 percent of churn in advance.

JOHN

Sixty percent of your churn surprised you?

PRIYA

Yes. That is the honest answer.

John let the answer sit for a moment. Priya did not flinch from it; she had clearly come to the meeting determined to be honest, even about the parts of her function that did not work well.

JOHN

What would you need to do better?

PRIYA

I think we have the data. Most of the churn signals are visible in product usage, in support ticket patterns, in the cadence of customer engagement with our CSMs. We are just not aggregating any of it systematically. The health score is being produced by CSMs from their own sense of the relationship, not from data that would expose patterns the CSM might not see.

JOHN

Does the warehouse have the data?

PRIYA

Product event data is there. Salesforce data on customer engagement is there. Support ticket data is partly there. I have asked Elena to pull samples for me, and she has. We have not built it into a regular reporting flow because that is not her job and there is no one else to do it.

JOHN

Have you considered an AI approach?

PRIYA

I have thought about it. I do not have the authority to commission one. I have read about companies using AI to predict churn from multi-source signals. I have not pursued it because, frankly, I did not know how to start.

JOHN

If we did pursue it, what would you want?

PRIYA

A scoring system that runs every week on every account, draws from product usage, support tickets, customer engagement, billing history, and the CSM's recent notes. It would produce a risk score and a brief explanation of what is driving the score. The CSM uses it as a starting point for their own assessment. The system is not replacing the CSM's judgment. It is making sure the CSM's judgment is not missing things the data is showing.

JOHN

That is well-articulated.

PRIYA

I have spent a long time thinking about it.

JOHN

Let me ask about expansion. NDR at 118 percent. How much of that is contractual escalators versus actual expansion bookings?

PRIYA

About 70 percent is true expansion bookings. The remaining 30 percent is contractual escalators plus price increases at renewal. So our true expansion rate is closer to 113 percent than the reported 118.

JOHN

That is a meaningful disclosure for the board.

PRIYA

It is. I have made the point to Aaron. He did not change the reporting.

JOHN

We will change it. I do not need NDR to look better than it is. I need NDR to be honest. Honest NDR at 113 with an expansion mix that is real is far better than reported NDR at 118 that the next investor will dig into and resent us for.

Priya nodded slowly. She had been waiting for this answer for some time.

PRIYA

Thank you.

JOHN

Last question. What about churn already in motion that we do not see yet?

PRIYA

I think we have at least three accounts in the mid-market segment, currently rated yellow, that are likely to non-renew at their next anniversary. I think there are probably another six or seven I do not know about. Some of the indicators I cannot see directly. The CSM might see them but might also be telling themselves the relationship is fine because they have good rapport with the champion.

JOHN

Whose CSMs are those?

PRIYA

Two are mine specifically. The others are spread across the team.

JOHN

Would you mind sending me a quick note about the two you know about? Names, ARR, your read of the situation. I would like to follow them quietly.

PRIYA

I will have it to you by end of day.

John's reflection that evening

Priya is exactly what Aaron and Raj suggested. Thoughtful. Self-aware. Has been waiting for a finance leader who would take the customer health question seriously. The fact that she walked into the meeting knowing the gap between her health scoring and actual churn outcomes is a credit to her.

Sixty percent of churn surprising us is a serious operational finding. If we can move that number down to 30 percent, even 40 percent, by aggregating signals the CSMs are not currently combining, we directly improve NDR and we directly improve forecast predictability. The use case is the same one Sarah asked for at the broader level: seeing around corners.

The NDR disclosure issue is real. We are reporting 118 because we are not separating contractual escalators from real expansion. The next investor will rebuild this number from first principles and discover the gap. I would rather correct it now than have it corrected for us later. Sarah needs to know this before the next board meeting.

The architecture for a customer health scoring system is the same architecture as the forecasting system Sarah wants. Multi-source signal aggregation, model-assisted analysis, human-in-the-loop output, regular cadence. If we build the foundation for one, the other becomes cheap. I should keep this in mind when thinking about sequencing.

Chapter 5

Karen Lindqvist, Chief Human Resources Officer

On hiring, capacity, and the relationship between people and revenue

Day 11 · Tuesday afternoon · CHRO office · sixty minutes

Karen Lindqvist's office was on the third floor near Sarah's. She had set up the meeting with a clear agenda emailed to John in advance: hiring plan review, capacity-to-revenue model, people analytics. She was, John had noticed, the most structured of the executives in her preparation.

KAREN

I would like to walk you through three things, in order. The hiring plan as it currently stands. The capacity model that underpins it. The gaps I am most worried about. Each will take about fifteen minutes. That leaves us fifteen for whatever you want to discuss.

JOHN

Please.

KAREN

Hiring plan first. We are at 142 today. The plan for the year approved at the last board meeting takes us to 187 by Q4. That is a 32 percent increase in headcount against a planned 38 percent ARR growth. The plan front-loads engineering and sales hiring in the first half, and shifts toward customer success and operations in the second half.

JOHN

Specific composition?

KAREN

15 net new in engineering, all in the first half. 12 in sales: 8 quota-carrying reps, 3 SDRs, 1 sales engineer. 8 in customer success: 6 CSMs, 2 support. 5 in operations, finance, and G&A; combined. 5 in product. The total is 45 net new hires, which after expected attrition translates to approximately 55 gross hires.

JOHN

Where are you against plan year to date?

KAREN

We are behind. We have completed 18 hires in Q1 and the first half of Q2 against a plan of 23 for that period. We are behind in engineering, where the labor market is tight, and slightly behind in sales, where we have been more selective than plan assumed.

JOHN

Is the slippage a problem?

KAREN

It is a mixed signal. Lower run rate cost is good for cash. But the engineering shortfall is limiting our product capacity, which is being felt in the roadmap. And the sales shortfall, while smaller, is contributing to the pipeline thinness Michael has been flagging.

JOHN

Tell me about the capacity model.

KAREN

I built this in my first three months. The model maps each function's headcount to its expected output. For sales, that means quota dollars per rep, ramp time, attainment rate. For customer success, it is accounts per CSM and the renewal and expansion rates per segment. For engineering, it is more difficult; I have used a rough proxy of stories per engineer per quarter, which is not great but is the best available. For operations and G&A;, I have used spans of control as a benchmark.

JOHN

And from those, you back into hiring?

KAREN

Exactly. The capacity model tells me how much output we should expect from the current and planned headcount, and I compare that to the revenue plan. Where the math does not work, we either need more people, better tools to extract more output from existing people, or a revenue plan that is more achievable.

JOHN

Where does the math not work right now?

KAREN

Three places. First, in customer success, where our segmentation puts too many accounts on each mid-market CSM and we are seeing service quality erode at the high end of that ratio. Priya has flagged this. Second, in sales enablement, where we have one person supporting 34 reps, which is below industry standard. Third, in finance, where you have inherited two people doing the work of four, and they will burn out before the year is over if we do not address it.

JOHN

I appreciate you stating that directly.

KAREN

I am paid to state it. The board hired me to be the early warning on people issues, and burnout in your two finance people is an early warning.

JOHN

Let me ask about AI in your function.

KAREN

Limited adoption so far. I personally use Claude and ChatGPT for various drafting tasks, on a team-tier subscription that I purchased. Our talent acquisition lead has been experimenting with AI to screen resumes, which I have asked her to pause until we have a policy framework, because the regulatory environment for AI in hiring is now serious in several jurisdictions including ours.

JOHN

You stopped it on regulatory grounds.

KAREN

Yes. The New York City bias audit requirement has been in effect for a while. California has pending legislation. The EU AI Act treats employment screening as a high-risk use case. I was not comfortable proceeding without legal review. Naomi agreed with me.

JOHN

That was the right call.

KAREN

It was an easy call. The more interesting AI question for my function is the capacity model itself. Right now, the model is built from spreadsheets. The relationships between headcount and revenue are static formulas. They do not adjust dynamically as conditions change. I have thought about whether there is a way to build something more responsive, where the model adapts to changes in the underlying productivity metrics. I am not sure where to start.

JOHN

That is an interesting question. The forecasting work that Sarah has asked me to focus on touches this. Capacity feeds revenue, revenue feeds forecasting, and forecasting feeds capacity decisions. If we build the right architecture for one, the others benefit.

KAREN

I would like to be part of that conversation.

JOHN

You will be.

JOHN

Let me ask one more thing. The hiring plan assumes 38 percent ARR growth. Is the hiring plan flexible if ARR growth comes in lower?

KAREN

It is partly flexible. Engineering hiring tends to commit further in advance because of recruiting cycles. Sales hiring can flex more dynamically because we can hold offers, slow down recruiting, or pause SDR class starts. Customer success is in between. If we needed to take 10 hires out of the second-half plan, I could do it with a month of notice. If we needed to take 20, that is harder and would mean rescinding offers that have been extended.

JOHN

What is the trigger you would want to see before pulling back?

KAREN

I would want to see a Q2 result that confirmed the pipeline weakness Michael has been describing was real, not a one-quarter phenomenon. If Q2 is on plan, we keep hiring. If Q2 misses meaningfully and the pipeline going into Q3 is still thin, we slow down.

JOHN

That is a sensible posture. Let us hold to it.

John's reflection that evening

Karen is structured, prepared, and has built a capacity model that is more rigorous than I would have expected. The model is the kind of foundation a finance function should aspire to but rarely owns. The fact that she built it before I arrived is a gift.

Her observation about my two finance people burning out is one I should not ignore. Aaron flagged Elena as a flight risk. Karen has now flagged both Elena and David as a capacity risk. The two findings reinforce each other. We need to either add finance headcount, deploy tools that reduce the work, or both. AI workflows applied carefully to finance operations are exactly the tool category that would help.

Her decision to pause the AI resume screening was right. The regulatory exposure for AI in employment decisions is real and growing. The fact that she paused it without being asked tells me something about her judgment. I should note this as a positive signal about the broader question of who in this company can be trusted to make the right call about AI deployment in their own function.

The capacity model question she raised is interesting. A dynamic capacity model that updates based on observed productivity is not impossible. It would be a workflow rather than an agent, and it would draw from the same underlying data architecture that the forecasting workflow needs. Another argument for getting the foundation right and then deploying multiple workflows on top of it.

Chapter 6

David Kim and Elena Vargas, Controller and Senior FP&A;

On the close, the FP&A; practice, and what is breaking

Day 14 · Friday afternoon · finance conference room · two hours

John had arranged the meeting deliberately. He wanted David and Elena in the same room so that the picture of finance operations could be assembled from both perspectives. He had also wanted Elena there because Aaron, Raj, and Karen had each independently flagged her, and he wanted to hear her thinking directly.

David Kim was at the table when John arrived. Elena Vargas came in shortly after, carrying a laptop and a printed spreadsheet with marked-up notes.

JOHN

Thank you both for the time. I want to use this conversation to understand the finance function as it operates today. Not as it should operate. As it does.

DAVID

Where would you like to start?

JOHN

Walk me through the close. Day by day, what happens.

DAVID

The close is a six-business-day process. Day one, we cut off bookings and complete the bank reconciliations from the prior month. Day two, we close subledgers and run preliminary numbers. Day three, we review and book accruals, allocations, and adjustments. Day four, we do the management review, which I run with Elena. Day five, we finalize the financial statements and the operational reporting. Day six, we distribute the close package.

JOHN

Six days is reasonable for a company of our size. Where in the process does most of the time go?

DAVID

Day three. The accruals and allocations. We have legacy spreadsheet processes for several categories that take Elena most of a day to run.

ELENA

It is closer to a day and a half. The marketing accruals alone are a half day because the spend is fragmented across so many vendors and the invoicing cadence is irregular.

JOHN

How much of day three is judgment versus mechanical work?

ELENA

Maybe 20 percent judgment. The rest is mechanical aggregation, calculation, and entry. I am pulling from spreadsheets, from vendor portals, from NetSuite, from Salesforce, and reconciling them by hand. The judgment piece is identifying which accruals need attention. The mechanical piece is everything else.

JOHN

How long have you been doing this exact set of tasks?

ELENA

Most of them since I joined two years ago. The process has not changed materially.

JOHN

Have you considered automating any of it?

ELENA

Yes. I have tried to. I have built some scripts in Python to pull from APIs where I can. I have used ChatGPT to draft the variance commentary, which I would not call automation exactly but has helped. The bigger pieces require engineering support I have not been able to get prioritized.

DAVID

That is fair. The engineering team has been fully consumed with product work. Internal tooling for finance has not made the priority list.

JOHN

What about the variance commentary? Tell me about that.

ELENA

After the management review on day four, I draft the variance commentary that goes into the close package. It explains why each line item differs from plan and from prior period. The draft used to take me four to five hours. I have been using ChatGPT to produce a first draft from the variance data, which I then edit. It has cut the time to maybe two hours.

JOHN

What data are you giving it?

ELENA

I paste the variance table from Excel and a short context about the period. It produces a first draft. The draft is usually 70 percent of the way there. I edit for accuracy, for tone, and for things the model gets wrong.

JOHN

Are you using a personal account?

ELENA

A personal account. I pay for the subscription myself. I have been careful not to paste anything that would be a problem in the model's training, but I am aware the boundary is not fully clear.

JOHN

That is going to change. I am not asking you to stop using AI. I am asking you to stop using a personal account for company work. We will get you a sanctioned arrangement quickly. Probably this month.

ELENA

Thank you. I have wanted that for a year.

John nodded. He had wanted to acknowledge the issue directly. There would be other shadow uses across the company. Elena's admission was, in fact, easier to handle than most because she had been thoughtful about it.

JOHN

David, where in the close do you most worry about errors?

DAVID

Three places. The marketing accrual, because the data is messy. The deferred revenue calculations, because the contract terms are increasingly complex. The expense categorization at the GL level, because we have too many ambiguous transactions and too few people reviewing them. Anything we miscategorize flows into the wrong P&L; bucket and ends up either misleading management or, worse, generating audit questions.

JOHN

How many ambiguous transactions per month?

DAVID

Roughly 200 to 400. Most are small. Some are meaningful.

JOHN

Who reviews them?

DAVID

I do. About one full day per month.

JOHN

And that is a real day, on top of everything else.

DAVID

Yes.

JOHN

Let me ask about the FP&A; side. Elena, walk me through the forecast.

ELENA

The forecast is a rolling four-quarter view that I update monthly. The revenue forecast is built bottom-up from the Salesforce pipeline, with my own probability weights applied. The cost forecast is built from the headcount plan, the marketing plan, and our run-rate operational costs. I reconcile to the board plan each month.

JOHN

How much time does the monthly forecast cycle take?

ELENA

About three days each month. Plus a fourth day for the board prep when we have a board meeting.

JOHN

Where in those three days does the time go?

ELENA

About a day on data preparation. About a day on building the model itself and running scenarios. About a day on documentation and commentary.

JOHN

How accurate is the forecast?

ELENA

On revenue, the four-week-out forecast misses by about 5 to 8 percent in either direction. On costs, it is tighter, around 2 to 3 percent. The forecast accuracy degrades as we look further out. The thirteen-week-out forecast on revenue is closer to 10 percent.

JOHN

Which is more often the case: forecast too high or forecast too low?

ELENA

Too high. Especially on the four-week-out view, which corresponds to the end of quarter. We over-forecast revenue more often than we under-forecast.

JOHN

By how much on average?

ELENA

Roughly 3 to 4 percent on the four-week-out view, looking at the last eight quarters.

JOHN

That is a meaningful pattern.

ELENA

It is. I have flagged it. The pattern correlates with how aggressively reps stage their late-quarter deals.

JOHN

Have you shared this with Michael?

ELENA

I have shared it with David. David has mentioned it to Michael. Michael has acknowledged it. The reps still stage aggressively.

John made a note. The pattern Michael had been unwilling or unable to fully address in his conversation was now confirmed by the finance side. Elena had quietly built the evidence base. Whether it had been used

at the executive level was a separate question.

JOHN

Elena, if you had a free hand and the right tools, what would you change about the FP&A; function?

Elena paused. She looked at David briefly before answering.

ELENA

I would build a forecasting layer that incorporates leading indicators beyond the pipeline. Customer health signals, product usage trends, hiring pace, macro conditions. The current forecast is essentially a pipeline-derived revenue forecast plus a headcount-derived cost forecast. The signal that explains the most variance is in places we are not looking.

JOHN

Continue.

ELENA

I would also build a variance explanation system that can ingest the underlying data and produce both the commentary and the diagnosis. Right now, my commentary describes what happened. It does not always explain why. The why often requires running additional analyses that I do not have time to run.

JOHN

And I would assume there is a piece on close automation.

ELENA

Yes. The mechanical work in the close. The accrual aggregation, the expense categorization, the reconciliations. There is a lot of routine analysis the model could do that David and I do today.

JOHN

You have thought about this more than I expected.

ELENA

I have had time to think about it. I have not had the authority to act on it.

JOHN

Authority is granted. We will need to discuss what that looks like. Not today. But this conversation is the beginning of it.

JOHN

Last topic. The board package.

DAVID

The board package is the artifact at the end of the close. Twenty-five to thirty slides plus an exhibit pack. I build the financials. Elena builds the operational metrics. Sarah writes her opening commentary. Michael contributes the commercial section. Priya contributes customer success. Aaron used to pull it together. That has fallen to me recently.

JOHN

How long does it take to assemble?

DAVID

About a week. Most of it concentrated in the two days before the board meeting.

JOHN

That is too much time for a board package. Once we have the right architecture, the board package should be a derivative of the close, not a separate construction.

DAVID

That would be welcome.

JOHN

For the next board meeting, we run it the way it has been run. After that, we redesign.

John's reflection that evening

The longest conversation so far. Two hours, two people, and the picture of the finance function is now legible. David is a strong controller. Elena is a strong analyst trapped in a structure that is preventing her from doing the work she is most capable of. They both work too many hours. They both produce good output. They are both at the edge of what sustainable looks like.

The Elena observation Aaron flagged, Raj confirmed, and Karen reinforced is now fully evident. She has been doing work two tiers above her title. She has built informal analytical capabilities that the function would not otherwise have. She has thought about the architecture of the FP&A; function more deeply than anyone in the company. I will need to give her a path that is commensurate. Otherwise we will lose her, and we will deserve to lose her.

The forecast pattern she identified, of systematic over-forecasting at the end of quarter correlating with rep staging behavior, is the second confirmation of what Michael partly acknowledged. The triangulation is now complete. We are systematically optimistic on revenue forecasts in the final weeks of each quarter, and the pattern is driven by rep behavior that has not been disciplined. This is the problem the board is sensing. Fixing it requires both a sales operations fix and a forecasting architecture fix.

The 200 to 400 ambiguous transactions per month that David reviews are a clear automation target. An AI workflow that classifies transactions with high confidence and escalates the genuine edge cases would reclaim a full day of David's time per month and probably improve consistency. This is the kind of unglamorous use case that produces real value.

Elena's informal use of ChatGPT for variance commentary is a sanctioned-AI opportunity. We pull it onto a real arrangement, give her the proper tooling, and the time savings compound across the whole FP&A; function. She has already done the proof of concept. Productionizing it is engineering work, not strategy work.

Chapter 7

Tom Reeves, Director of Operations

On vendors, infrastructure spend, and the unglamorous spend nobody watches

Day 18 · Tuesday afternoon · Tom's office · seventy minutes

Tom Reeves's office was on the first floor near the small operations team. He was the most senior person on his side of the building and the only person John had met who had a printer next to his desk.

TOM

I assume Aaron told you what I do.

JOHN

He said you run procurement, vendor management, infrastructure spend, and a portion of business operations.

TOM

That is the official description. The real description is I am the person who tries to keep us from being overcharged by every vendor in sight, and the person who tries to keep our AWS bill from quietly exceeding our hosting budget by 30 percent every quarter.

JOHN

Tell me about the AWS bill.

TOM

It is currently 312 thousand a month. It was 258 thousand a month a year ago. It is growing faster than our revenue is growing. Some of that is real. We are serving more customers, processing more data, running more compute. Some of it is not real. It is engineers spinning up instances they forget to spin down, over-provisioned services that are running at 15 percent utilization, storage that is accumulating without any retention policy.

JOHN

How much of the bill is real and how much is waste?

TOM

My estimate, which is not rigorous, is that 20 to 25 percent is waste. That would be about 60 to 80 thousand a month. I have raised this with Raj. Engineering has done occasional cleanups. The pattern of growth resumes within two months.

JOHN

Why does it resume?

TOM

Because there is no continuous monitoring with real consequences. Cleanups happen episodically. Engineers spin things up freely. There is no ongoing accountability for AWS spend at the team level.

JOHN

How would you fix it?

TOM

Two ways. One is organizational: assign every AWS cost item to a team and put that cost on the team's ledger. The second is technical: continuous monitoring of utilization with alerts when things drift. Both would help. Both have been talked about. Neither has been implemented.

JOHN

What would the technical version look like with AI in it?

Tom looked at John with mild surprise.

TOM

I have not thought about it that way. But it is a good question. An AI workflow that watches utilization patterns, flags anomalies, identifies waste candidates, and produces a weekly summary for engineering leadership. That would be useful. We have the data. We are not analyzing it.

JOHN

Beyond AWS, what else is in your domain?

TOM

Software licenses. We have 47 sanctioned SaaS tools across the company. I track them. License utilization across the portfolio is about 71 percent. The other 29 percent is licenses we pay for that nobody uses. That is about 180 thousand a year in unused licenses.

JOHN

How do you track utilization?

TOM

Manually, quarterly. I send a survey to each license owner asking who is still using which tools. The response rate is about 80 percent. The accuracy of the responses is variable. Some license owners say everything is being used because they do not want to lose access to anything.

JOHN

That sounds painful.

TOM

It is painful. It is also one of the highest ROI things I do because I find tens of thousands of dollars per quarter in waste.

JOHN

Have you considered automating it?

TOM

I have. Some of the larger SaaS tools provide utilization data through their APIs. Most do not. The smaller tools are the worst because they have the least visibility. I have not had engineering support to build anything automated.

JOHN

Vendor management generally?

TOM

About 140 active vendors across the company. Roughly 20 of them are material in dollar terms, meaning over 50 thousand per year. The other 120 are small but they accumulate. I have a renewal calendar I maintain in spreadsheets. I review pricing on the material ones at renewal. I do not have the bandwidth to review the small ones.

JOHN

What is your read on the vendor portfolio right now? Where do you see the most risk?

TOM

Three places. First, we have several vendors where the contract auto-renews on terms that are not optimal. I have flagged a handful that I want to renegotiate before the next renewal window, but renegotiating takes time and I do not always get to it. Second, we have a few vendors that are critical to the product where we have no redundancy. If they go down or get acquired, we have a problem. Third, and this is a recent one, we are starting to accumulate AI vendor relationships that nobody is tracking centrally.

JOHN

AI vendors specifically?

TOM

Cadenza, which Michael runs. The Claude and ChatGPT consumer subscriptions that various people have. GitHub Copilot on the engineering side. A team subscription to ChatGPT that marketing has. A separate subscription that customer success started and may or may not still be paying for. I have asked twice for a central register of these and have not been given one.

JOHN

I will get you the register.

TOM

Thank you. The reason it matters is that each one of those vendors has terms, has data handling commitments, has security posture, has price. If we do not know what we have, we cannot manage any of it.

JOHN

Last question. If you had AI tools to help with your function, what would you most want?

TOM

Three things, all unglamorous. First, the AWS analyzer I mentioned. Second, a license utilization tracker that aggregates from APIs where available and prompts for confirmation where not. Third, a vendor contract analyzer that reads our master agreements and flags risky terms, auto-renewal traps, and pricing escalators. None of those are exciting. All of them would save real money.

JOHN

They are exactly the kind of use cases I should be paying attention to. Boring is often the right answer in operations.

TOM

I have been waiting for a CFO to say that.

John's reflection that evening

Tom is the kind of operations leader every company should have and few companies properly value. He is methodical, accurate, and he has been finding hundreds of thousands of dollars in waste a year on a function that is staffed for half what it deserves.

The AWS situation is real. 60 to 80 thousand a month in waste is between 720 thousand and 960 thousand a year, against a total spend that is roughly 3.7 million annually. That is not a rounding error. It is a material control problem that the engineering culture has not addressed. An AI workflow that watches infrastructure spend and produces actionable alerts is exactly the kind of thing that pays for itself.

The license utilization problem is similar. 180 thousand annually in unused licenses, found manually, against an incomplete picture. If we automated this, we would probably find more.

The AI vendor accumulation is now confirmed from a third angle. Raj sees it from engineering. I have inferred it from the conversations with Sarah and Michael. Tom is now telling me that procurement is unable to track it centrally. The governance problem is concrete: we have an unknown number of AI vendor relationships, no central register, no policy framework, no audit. This is the first thing to fix.

Tom's three AI use cases (infrastructure, licenses, vendor contracts) are all real, all valuable, all within the supply chain and working capital category. They are not what most companies build first in finance AI. They are what they should build.

Chapter 8

Naomi Bridges and Wei Zhao, Legal and Security

On the governance perimeter and what the regulators are doing

Day 22 · Tuesday morning · executive conference room · ninety minutes

John had asked Naomi and Wei to meet jointly. He wanted to hear the legal perimeter and the security perimeter in the same conversation, because the two were increasingly inseparable in any AI context.

Naomi Bridges had brought a tablet with what looked like a structured outline. Wei Zhao had brought a single page of notes.

NAOMI

I am going to start with the regulatory landscape, because it has shifted enough in the last twelve months that even people who follow it carefully need to update their mental model. Then Wei can cover the security posture. Then we can take whatever questions you have.

JOHN

Please.

NAOMI

The regulatory environment for AI now has several distinct strands. The EU AI Act is the most comprehensive. It came into application in stages, with high-risk system requirements now in force. We are not directly subject to it because we do not currently deploy any system that the Act classifies as high-risk. We have customers in Europe, however, and the Act's reach includes systems placed on the market or put into service in the Union, which could apply to us indirectly if we deploy AI-powered features that operate on our European customers' data.

JOHN

How do we know whether anything we deploy would qualify?

NAOMI

We do not have a clear process today. That is part of what concerns me. The Act distinguishes between AI uses by risk tier. Most enterprise productivity AI is in the limited-risk tier, with transparency requirements but no significant operational burden. High-risk uses include employment screening, creditworthiness assessment, and a few other categories. We are not currently in those categories, but we will need a classification framework before we deploy anything that could touch them.

JOHN

What else?

NAOMI

In the United States, the federal landscape is shifting. The executive order from late 2023 was revoked. There is no federal comprehensive AI legislation. There are sectoral rules emerging: financial services regulators have published guidance, the FTC has been active on deceptive practices in AI marketing, the EEOC has been active on AI in employment. The states are moving faster than the federal government. California has multiple AI bills in various stages. New York City has the bias audit requirement for automated employment decision tools, which we touched on with Karen. Texas, Florida, Colorado have all passed or proposed something.

JOHN

What about contractual obligations?

NAOMI

This is where it gets more pointed for us. Our enterprise customers, particularly the ones in regulated industries, are starting to include AI clauses in their procurement agreements. They want to know what AI we use, what data we process through which models, what training is happening with their data, what governance we have in place. I have seen three such clauses in customer contracts in the last six months. Two were workable. One was not, and we had to negotiate it.

JOHN

What do those customers want, specifically?

NAOMI

They want assurance that we are not using their data to train third-party models. They want a disclosure of which AI subprocessors we use. They want a commitment that any AI-driven decision affecting their data is auditable. They want the right to be notified of material changes in our AI subprocessor list. None of those are unreasonable. Most of them we cannot answer cleanly today.

JOHN

Wei, take it from there.

WEI

On the security side, we have a SOC 2 Type 2 report that is current. We have basic controls: SSO, MFA, least privilege, audit logging on production systems. We have a data classification framework that is partly implemented. We have an incident response plan that is more theoretical than practiced.

JOHN

And on AI specifically?

WEI

On AI specifically, we have less than I would like. We do not have an inventory of AI tools in use. We do not have a data flow map for AI systems. We do not have an approval process for new AI tools. We have informal practice. Practice is not a control.

JOHN

What is the biggest specific risk you see right now?

WEI

Three. The first is data leaving the perimeter through unsanctioned AI use. Someone pasting customer data into a consumer chat interface. Someone uploading a document to a tool whose data handling we do not understand. The second is credential exposure, where API keys end up in places they should not be. The third, more speculative, is prompt injection in customer-facing applications, where a malicious customer manipulates the prompt to extract information or trigger actions. We do not currently have customer-facing AI but we are going to. When we do, that risk becomes real.

JOHN

What is the path to closing the gaps?

WEI

Three things, in order. First, an AI tool inventory with classification by data sensitivity and risk. Second, an approval process for new AI tools that goes through me before any vendor is engaged. Third, a monitoring layer that detects unsanctioned use, particularly data exfiltration into consumer AI tools.

NAOMI

I would add a fourth. A policy framework that we can actually distribute to employees. Right now, the rule is informal. We need to be able to say, here is what you can do, here is what you cannot, here is how you request an exception.

JOHN

Who would write that policy?

NAOMI

I would draft it. Wei would review it. The CFO would own its enforcement.

JOHN

Why the CFO?

NAOMI

Because the consequences of AI failure in this company will hit the CFO's office first, regardless of where the failure originates. A finance leader who does not own AI governance is a finance leader whose function gets blamed for everyone else's decisions.

JOHN

That is the right answer.

JOHN

Let me ask one more thing. If we were to build an AI governance framework from scratch for Helix, what would you each want in it?

NAOMI

A classification system for AI uses by risk tier. An approval process for new AI deployments. A vendor due diligence framework that covers data handling, subprocessor lists, and model training practices. An audit trail requirement for any AI-driven decision that touches customer data or financial reporting. A review cadence that is at least quarterly.

WEI

A technical implementation that makes the policy actually enforceable. SSO-protected access to all sanctioned AI tools. Network-level monitoring for unsanctioned use. Centralized credential management for API keys. Logging of AI tool usage at the user level for incident response.

JOHN

I am taking notes. Both of you will be on whatever working group designs this. I expect to bring something to Sarah and the board within the next sixty days.

NAOMI

Good. The audit committee will want to see it. Diana has already asked me twice when we would have a framework.

JOHN

I will be meeting Diana next week.

John's reflection that evening

Naomi and Wei together produced the most coherent picture of the governance perimeter that I have heard so far. The regulatory environment is shifting fast. Customer contractual demands are emerging. Internal practice is informal where it needs to be formal. The gap is real and it is mine to close.

Naomi's observation that the CFO will be blamed for AI failures regardless of where they originate is precisely right, and it is the argument I would make to anyone who questions why finance should own AI governance. Failures in AI touch the books, touch controls, touch external reporting, touch audit. The CFO is the executive whose authority and accountability sit closest to those consequences.

The three governance pieces that need to come together are: a policy framework that employees can understand, a technical architecture that makes the policy enforceable, and an audit trail that makes the whole system defensible. Each piece is doable. None is trivial. The combination is what will separate Helix from a typical Series B that has not thought about this seriously.

Diana has been asking for the framework. I should come into the meeting with her prepared to describe a path, not a finished document. But the path needs to be specific enough that she can see real thinking behind it.

Chapter 9

Diana Reyes-Okonkwo, Audit Committee Chair

On external accountability, the next round, and what the audit committee expects

Day 28 · Tuesday morning · Diana's office in San Francisco · ninety minutes

Diana Reyes-Okonkwo kept an office in the financial district from which she ran her board work. John had taken the BART downtown and walked the four blocks. She was in her office reviewing a board pack from another company when he arrived. She closed it and gestured to the chair across from her.

DIANA

I have been looking forward to this meeting. Sarah has spoken well of your first three weeks. I want to use this hour to give you a frank picture of what the audit committee expects, and to give you a chance to ask me anything you want.

JOHN

I appreciate it. Let me start with what is on your mind.

DIANA

Three things. First, the forecast discipline. The board has accepted two misses in six quarters. The board will not accept a third unless it comes with a real explanation of why we did not see it coming. The pattern that concerns us is not the misses. It is the consistency of the surprise. Each miss has been a surprise to management before it has been a surprise to us. That tells me the forecasting process is not reading its own signals.

JOHN

I have heard the same observation from several people in the last three weeks.

DIANA

Second, the AI question. Sarah has talked about wanting to do more with AI. The audit committee's position is that we are supportive of AI in finance and across the business, but the bar for governance is going to be high. We are not going to approve deployments that we cannot understand. We are not going to approve deployments that touch our financial reporting without controls equivalent to what we expect for human-prepared reporting. We will be patient with the pace of implementation. We will not be patient with shortcuts.

JOHN

That is the right posture.

DIANA

Third, the next round of capital. Marcus is going to push for a Series C in twelve to fourteen months. The company will need to be ready. Ready means the financial operations have to look like a company three or four times our current size. That is not because we will be three or four times our current size by then. It is because the next investor will be a larger fund with higher standards, and we will not raise on terms we want unless we can show them a finance function that is ahead of where the company is, not behind.

JOHN

What does ahead look like, in your view?

DIANA

Three things, in order. A close that is five days, not six, and clean. A forecast that is reliable to within 3 percent on the four-week-out view, not 5 to 8. A governance framework that an institutional investor can review and not have questions about. The first two are execution. The third is institutional.

JOHN

I will get there on all three.

DIANA

I expected nothing less. Let me ask you something. What is your AI thesis for this company?

John had been waiting for the question. He had drafted his answer the night before.

JOHN

My thesis is that finance at Helix is a multi-source signal aggregation problem that the current architecture is not solving. The forecast misses because the forecast is built from pipeline data alone and the leading indicators of revenue live in customer success, in product usage, in rep behavior patterns, in macro signals, and in spend trends that the forecast does not currently see. The same applies to churn prediction, to capacity planning, to expense control. We are doing analysis on fragments when we should be doing analysis on the whole.

DIANA

Continue.

JOHN

My thesis is that AI is the tool that makes multi-source analysis practical at our scale and stage. We do not have the human bandwidth to build a data science function. We do not need to. We need a small number of well-architected workflows, governed properly, that bring AI to the signals that already exist in our systems. The objective is not to replace anyone. The objective is to give the existing team the leverage that lets them see what they could not see before.

DIANA

What use cases?

JOHN

I am still finalizing. The pattern I am seeing is that five use cases keep emerging as both high value and addressable. A forecasting layer that aggregates leading indicators. A board reporting and commentary system that compresses the reporting cycle. A pipeline intelligence capability that catches deal slippage and rep behavior patterns. A finance operations copilot that handles the mechanical work in the close and the ambiguous transactions. An infrastructure and vendor intelligence capability that catches waste in AWS spend, software licenses, and vendor contracts.

DIANA

Why those five?

JOHN

Because each of them addresses a real and specific pain that I have heard described by the function leader who lives with it. Because each of them produces measurable value within twelve months. Because each of them is appropriately scoped for what AI can reliably do at our stage of sophistication. And because each of them, together, builds the foundation for the broader pattern that Sarah called seeing around corners. I would rather build five workflows we can defend than fifteen we cannot.

DIANA

And governance?

JOHN

Governance comes first. I am writing the framework now. It will be in your hands before the deployments begin. The order is: framework, foundation, workflows. Not the reverse.

DIANA

That is the right order. Most companies do it backwards and then explain the governance after the fact when something has gone wrong.

JOHN

I am aware.

DIANA

I think you and I are going to have a productive relationship. Let me ask one more question. What do you need from the audit committee?

JOHN

Three things. First, your sustained attention. The single biggest risk in building any AI capability is that it becomes a side project that drifts. If the audit committee makes it a standing agenda item, it does not drift. Second, your willingness to challenge the framework when I bring it. I would rather be pushed back early than have you discover gaps later. Third, your patience with the pace. The right pace is slower than the marketing of these technologies suggests. We will move deliberately.

DIANA

You will have all three. The standing agenda item starts at the next meeting.

John's reflection that evening

Diana is what every audit committee chair should be and what most are not. She is sharp, she is demanding, and she is constructive. The bar she has set is reasonable. A five-day close, a 3 percent forecast accuracy, an institutional governance framework. Those are deliverables I can commit to. They are also deliverables that, if achieved, transform what this company looks like to the next investor.

The conversation crystallized the five use cases. They are not a list I assembled deliberately. They are a list that emerged from twenty-eight days of listening. Forecasting. Board reporting. Pipeline intelligence. Finance operations. Infrastructure and vendor intelligence. Each one traces to a specific pain point I heard described by a specific person.

Diana asked the right ordering question. Framework, foundation, workflows. Governance comes first. The audit committee will be a force-multiplier on that discipline if I let them. The standing agenda item she committed to is exactly what I need.

The thirty days of conversations are essentially complete. The synthesis comes next. I have two days before I sit down to write the architectural memo for Sarah. I will spend them re-reading my notes, looking for the patterns that span more than one conversation, and beginning the framework that Part 3 will fully develop.

Closing · The Pattern That Emerged

Thirty days of conversations had produced more than a list of facts. They had produced a pattern. John sat in his office on the evening of day thirty, with the door closed, a printed stack of his meeting notes in front of him, and he tried to articulate the pattern in writing.

What he wrote first, before any framework, before any architecture, before any use case list, was this: Helix Cloud Systems is a company whose finance function is being asked to see signals that live outside finance. The forecast misses because the forecast cannot see customer success signals or rep behavior signals. The board reporting is slow because the assembly of the report is decoupled from the close. The churn surprises Priya because the health scoring system does not integrate product usage. The expense waste accumulates because no one is watching it continuously. The pattern is the same in every direction: fragments of signal, generated by competent people in different functions, never aggregated into a coherent view.

The pattern, in turn, suggested the architecture. The finance function would not solve this by getting better at what it currently does. It would solve it by becoming the place where the fragments came together. The technology that made the aggregation practical, at the company's scale and stage, was AI. Not as a magic solution, but as the tool that compressed analytical work enough to make multi-source synthesis possible with a small team.

The five use cases John would propose to Sarah were not chosen because they were the most exciting. They were chosen because each of them, in its own way, addressed a fragment of the pattern. Forecasting brought customer signals and macro signals into the revenue view. Board reporting compressed the reporting cycle by making the report a derivative of the close. Pipeline intelligence brought rep behavior patterns into the forecast. Finance operations reclaimed time from David and Elena. Infrastructure and vendor intelligence brought continuous monitoring to spend categories where episodic monitoring had been allowing waste to accumulate.

Behind the five use cases sat the governance framework, which would have to come first. Behind the framework sat the technical foundation, which Raj had already proposed. Behind the foundation sat the operating principle that the CFO's office would own AI governance, would set the architectural standards, and would have the authority to approve or block any AI deployment that touched finance or the books.

John had heard the principle articulated, in different words, by every executive he had spoken with. Sarah had asked for predictability and visibility and treated AI as a means to both. Raj had welcomed the governance assignment. Michael had said he had been waiting for a CFO to push. Priya had said no one had ever asked her about customer success data with this seriousness. Karen had been ready with her capacity model. David and Elena had laid out a function that needed exactly the leverage AI could provide. Tom had three boring use cases that would save hundreds of thousands of dollars annually. Naomi had described the governance perimeter with precision. Wei had laid out the technical controls. Diana had set the bar and committed her sustained attention to clearing it.

The thirty days had produced a coherent picture. The next thirty would produce a memo.

◆ ◆ ◆

Helix Cloud Systems is a company whose finance function is being asked to see signals that live outside finance. The fragments are real. The aggregation is what is missing.

End of Part 2

The Discovery

John has now spent thirty days listening. He has spoken with nine executives and built a coherent picture of the company's finance function, its commercial operations, its customer dynamics, its people structure, its operational footprint, its legal and security perimeter, and its board accountability.

In Part 3, you will read what he writes next. The architectural memo to Sarah, the systems map, the inventory of pain, the AI readiness assessment, the architectural principles, the prioritization framework, the selection of the five use cases, the sequencing decision, and the personal commitment. Part 3 is where listening becomes judgment, where judgment becomes writing, and where writing becomes the constitution of how AI will enter the finance function at Helix Cloud Systems.

Before you proceed, take the assessment that follows. It tests not only the substance of the discovery conversations but the executive discipline they demonstrate. The questions are designed to ensure that you have absorbed not just what John heard, but how he heard it.

Appendix · Assessment

Twenty questions on Part 2. Twelve multiple choice, five short answer, three scenario-based. Answer key with explanations follows.

Part I: Multiple Choice

1. What did Sarah Chen identify as the board's primary concern, distinct from absolute financial performance?

(a) The company's growth rate compared to peers.
(b) The predictability of the forecast and the pattern of late-quarter surprises.
(c) The capital efficiency relative to the most recent funding round.
(d) The composition of the executive team.

2. Raj Patel proposed three foundational technical components that should be built before deploying serious AI workflows. They were:

(a) Data warehouse, BI layer, prompt library.
(b) MCP server in front of the warehouse, workflow registry, audit log.
(c) Fine-tuned model, vector database, retrieval system.
(d) Custom connectors, API gateway, monitoring dashboard.

3. Michael O'Brien's 20-to-35 percent productivity figure for the Cadenza AI pilot was:

(a) Calculated from a controlled measurement using the engineering analytics team.
(b) Based on rep feedback and his observation, not on rigorous measurement.
(c) Generated by Cadenza's own analytics dashboard with audit trail.
(d) Independently verified by the finance team using activity logs.

4. Priya Sundaram reported that, of accounts that churned in the prior year, the share that had been rated green or yellow in their last health review was approximately:

(a) 20 percent — most churn was caught in advance.
(b) 40 percent — about half of churn was correctly anticipated.
(c) 60 percent — most churn surprised the CS team.
(d) 90 percent — almost no churn was anticipated.

5. Karen Lindqvist had paused the AI resume screening tool in talent acquisition primarily because of:

(a) Cost concerns about the vendor pricing.

(b) Regulatory exposure under New York City, California, and EU AI Act provisions covering automated employment decision tools.

(c) A complaint from a candidate about the tool's output.

(d) Internal feedback from hiring managers that the tool was inaccurate.

6. Elena Vargas's informal use of ChatGPT for variance commentary was:

(a) Run on a corporate enterprise account with full data handling agreements.

(b) Run on her personal subscription with no formal data governance framework in place.

(c) A pilot officially sanctioned by the prior interim CFO.

(d) Connected to NetSuite directly through a custom integration.

7. Tom Reeves estimated the share of AWS spend that was waste rather than legitimate consumption at approximately:

(a) Less than 5 percent.

(b) 20 to 25 percent.

(c) 40 to 50 percent.

(d) More than 60 percent.

8. Wei Zhao identified three primary security risks from current AI usage at Helix. Which of the following was NOT among them?

(a) Data leaving the perimeter through unsanctioned AI use.

(b) API key credential exposure.

(c) Prompt injection in future customer-facing AI applications.

(d) Adversarial attacks on the underlying foundation models.

9. Naomi Bridges argued that the CFO should own AI governance because:

(a) The CFO has the largest budget to allocate to AI.

(b) The CFO's office will be blamed for AI failures regardless of where they originate.

(c) Other executives have refused to take responsibility for AI.

(d) The audit committee specifically requested it.

10. Diana Reyes-Okonkwo set three concrete deliverables as the bar for Helix's finance function to meet before the next round of capital. They were:

(a) Six-day close, 5 percent forecast accuracy, basic governance.

(b) Five-day close, 3 percent forecast accuracy, institutional governance framework.

(c) Three-day close, 1 percent forecast accuracy, public-company controls.

(d) Same-day close, real-time forecasting, automated governance.

11. The pattern John identified across the thirty days was best described as:

(a) A finance function that needed to be technically modernized but not architecturally redesigned.

(b) A company where competent people in different functions generated signal fragments that were never aggregated into a coherent view.

(c) A company that needed to replace several executives before any improvement was possible.

(d) A company whose primary problem was insufficient AI deployment relative to peers.

12. The five use cases that emerged from John's discovery process were forecasting, board reporting, pipeline intelligence, finance operations, and:

(a) Customer churn prediction.

(b) Hiring and capacity optimization.

(c) Infrastructure and vendor intelligence.

(d) Sales call analysis.

Part II: Short Answer

13. In two or three sentences, explain why Sarah Chen requested a memo rather than a slide deck from John, and what that request signals about the kind of CFO she expects him to be.

14. Elena Vargas had been using ChatGPT on her personal account to draft variance commentary, and Raj had API keys for engineering experimentation, while Michael paid for Cadenza on his personal expense card. In two or three sentences, explain why this pattern of shadow AI usage was a governance problem even though much of the work being done was genuinely productive.

15. Diana Reyes-Okonkwo articulated three deliverables and one ordering principle for AI deployment at Helix. In two or three sentences, describe the ordering principle and explain why it matters.

16. John deliberately met with David Kim and Elena Vargas in the same conversation rather than scheduling them separately. In two or three sentences, explain the rationale for this choice and what executive skill it reflects.

17. In two or three sentences, explain why the customer health scoring architecture Priya described and the forecasting architecture Sarah requested share a common foundation, and what that implies for sequencing.

Part III: Scenario-Based

18. Scenario: You are a newly appointed CFO at a Series B SaaS company. In your first week, your CEO tells you that the CRO has been running an unsanctioned AI pilot for four months on a personal expense card, processing customer call audio through a vendor whose data handling agreement no one has reviewed. The CRO believes the pilot is producing 20 to 35 percent productivity gains but cannot show measurement methodology. In one paragraph of executive prose, describe how you would approach this situation in the first thirty days, balancing the need for governance against the risk of shutting down activity that is producing real value.

19. Scenario: You are conducting discovery conversations with your executive team. A senior analyst in your finance function admits during a meeting that she has been using ChatGPT on her personal subscription to draft financial commentary, pasting variance tables from internal spreadsheets into the consumer interface. She is your strongest individual contributor and a known flight risk. In one paragraph, describe how you would respond in the moment, what you would commit to her, and what broader policy actions her admission should trigger.

20. Scenario: Your audit committee chair has set three deliverables as the bar for the finance function before the next round of capital: a faster close, a more reliable forecast, and an institutional governance framework. She has committed the committee's sustained attention but has asked what you need from them. In one paragraph of executive prose, describe how you would frame the partnership with the audit committee, what specific operating cadence you would request, and what risk you most want them to help you mitigate.

Appendix · Answer Key with Explanations

Multiple Choice Answers

Question 1: (b)

The board was not unhappy with the absolute performance. The board was unhappy with the pattern of being surprised by misses late in the quarter. The distinction between predictability and performance is central to Sarah's framing and to the rest of Part 2.

Question 2: (b)

Raj proposed an MCP server in front of the warehouse with proper access controls, a workflow registry to track and version every deployed AI workflow, and an audit log capturing every AI call. These three together form the technical foundation for any subsequent finance AI work.

Question 3: (b)

Michael admitted that the figure came from rep feedback and his own observation, not from a measurement system. He had been waiting for a CFO to push on this. The honesty of his admission was a positive signal but the underlying measurement gap remained.

Question 4: (c)

Priya reported that 60 percent of churned accounts had been rated green or yellow in their last health review, meaning most churn surprised the CS team. This was the single most important operational finding in Part 2 from the customer dynamics side.

Question 5: (b)

Karen paused the tool because of regulatory exposure under several frameworks covering AI in employment decisions. She made the call without being asked, which John flagged as a positive signal about her judgment.

Question 6: (b)

Elena was running ChatGPT on a personal subscription with no formal data governance. She was being careful, but the boundary was not fully clear. John committed to fixing this within the month.

Question 7: (b)

Tom estimated 20 to 25 percent of AWS spend as waste, roughly 60 to 80 thousand per month, against a total of about 312 thousand per month. The waste was material and persistent.

Question 8: (d)

Wei identified data exfiltration, credential exposure, and prompt injection. He did not raise adversarial attacks on foundation models as a primary current risk for Helix. That category of risk exists but was not his immediate concern.

Question 9: (b)

Naomi's argument was that the consequences of AI failure hit the CFO's office first regardless of where the failure originates. The CFO is the executive whose authority and accountability sit closest to those consequences. John agreed it was the right answer.

Question 10: (b)

Diana set a five-day close, 3 percent forecast accuracy on the four-week-out view, and an institutional governance framework that an institutional investor could review without questions. These three together represented the bar for Series C readiness.

Question 11: (b)

The pattern was not a deficient finance function. It was a system where competent functional leaders produced signal fragments that were never aggregated. The finance function's opportunity was to become the place where the aggregation happened.

Question 12: (c)

The fifth use case was infrastructure and vendor intelligence, emerging from Tom's conversation about AWS waste, license utilization, and vendor contract terms. Customer churn prediction sits within the broader forecasting and customer success workflows.

Short Answer Explanations

13. The memo versus the deck

Sarah requested a memo because she wanted to read the reasoning, not the conclusions. A slide deck compresses thinking into bullets and leaves the reasoning implicit. A memo forces the writer to articulate the argument in prose, exposing the logic to scrutiny. Her request signals that she expects John to be a CFO who thinks in writing and whose decisions are anchored in clearly articulated reasoning rather than in visual summary. Her comment about people whose thinking holds up in prose being the same people whose decisions hold up in practice is the executive philosophy behind the request.

14. Why shadow AI usage was a governance problem

The pattern of shadow AI usage was a governance problem because the company had no central inventory of what AI was being used, what data was flowing where, what contractual commitments existed, what was being paid by whom, and who was accountable for outcomes. Even where the underlying work was productive, the absence of visibility meant that no one could answer the questions an institutional investor, an auditor, or a regulator would ask. Productive shadow use accumulates risk faster than it accumulates value, and the asymmetry becomes visible only when something goes wrong. The fix was not to stop the productive work but to bring it onto a sanctioned footing.

15. Diana's ordering principle

Diana's ordering principle was that governance comes first, then the technical foundation, then the workflows on top. The principle matters because the reverse order is what most companies do and what produces the failures that haunt them later. When workflows are deployed first and governance is added after, the governance must accommodate workflows that may have made design choices incompatible with proper controls. When governance is established first, every subsequent deployment is built with the controls embedded in the architecture from the beginning. John's commitment to this ordering was the most important architectural signal he could send to the audit committee.

16. The joint David-Elena meeting

John met with David and Elena together because the picture of the finance function had to be assembled from both perspectives, and because Elena had already been flagged by three independent sources as both highly capable and at risk of leaving. Meeting them separately would have been more efficient but would have either signaled that Elena was being singled out or denied John the chance to observe how the two of them worked together. The choice reflects the executive skill of designing the discovery process itself for what you need to learn, rather than defaulting to one-on-one meetings as the standard pattern. Listening is partly about what questions you ask. It is also partly about how you structure the conversations to surface what you need to hear.

17. The shared foundation between churn scoring and forecasting

Both architectures depend on multi-source signal aggregation, weekly or monthly cadence, AI-assisted analysis, and human-in-the-loop output. The customer health scoring system pulls product usage, support ticket patterns, engagement cadence, and CSM notes into a single risk score. The forecasting system pulls pipeline data, customer health, hiring pace, and macro signals into a single revenue projection. The data sources overlap. The architectural pattern is the same. Sequencing implication: if the foundation is built well for one use case, the marginal cost of the second use case drops substantially, which argues for foundation work first and use case deployment second rather than building each use case in isolation.

Scenario Discussions

18. The CRO's unsanctioned AI pilot

The right approach is to address the governance problem without treating the productive activity as the enemy. In the first conversation, the new CFO should listen carefully to the CRO's description of the pilot, acknowledge the value the CRO believes it is producing, and avoid making any immediate decision to shut it down. The CFO should then take three actions in parallel. First, ensure the data handling agreement is reviewed by the general counsel within two weeks, and obtain a clean view of what data is flowing where. Second, bring the spend onto a corporate footing, with a master agreement and proper governance, rather than letting it continue on a personal expense card. Third, request a measurement methodology that establishes whether the productivity gains are real. The CFO should communicate to the CRO that the activity is not being stopped, but that it is being brought onto a footing that the audit committee and the next investor will require. The CFO should not announce any policy in the first thirty days; the policy should be drafted, reviewed by legal and security, and tested against the live situation before publication. The CRO's willingness to admit the gap is a positive signal worth preserving, not penalizing.

19. The senior analyst's admission

In the moment, the right response is to thank the analyst for the honesty, acknowledge that the practice is not sustainable, and commit to fixing the underlying problem within a defined and short window. The new CFO should make three specific commitments. First, the analyst will be moved onto a sanctioned AI arrangement with proper data handling and governance within thirty days. Second, the use case the analyst has proven informally — AI-assisted variance commentary — will be developed into a properly architected finance workflow, with the analyst as a contributor to its design. Third, the analyst's authority and recognition for the analytical leadership she has been exercising informally will be addressed structurally, because the absence of structural recognition is part of why she has been improvising in the first place. The broader policy actions her admission should trigger are the publication of an AI usage policy, the creation of a sanctioned AI tool inventory and approval process, and a confidential survey of other functions to surface comparable shadow usage without penalizing the people who admit to it. The principle is that the admission is a gift; punishing the gift teaches the organization to hide its shadow usage rather than surface it.

20. Framing the partnership with the audit committee

The new CFO should frame the partnership as a standing operational relationship rather than a quarterly reporting relationship. The specific operating cadence to request is threefold: first, a standing AI governance agenda item at every audit committee meeting, not just an occasional update; second, a quarterly written memo from the CFO to the audit committee chair on AI deployment status, risks, and changes to the framework; third, an open commitment that any material change to the AI deployment footprint will be flagged to the chair within days, not at the next scheduled meeting. The risk the CFO most wants the committee to help mitigate is the risk of drift, where the AI work loses momentum and becomes a side project. The audit committee's sustained attention is the single most effective antidote to drift, because it forces management to maintain progress against a public commitment. The framing should be that the CFO and the audit committee are co-owners of the governance framework, not that the audit committee is reviewing the CFO's work after the fact. This framing changes the posture of both sides and produces a faster and more rigorous outcome than either side could produce alone.