



The AI Production Paradox

The problem isn't your AI.
It's your comms infrastructure.

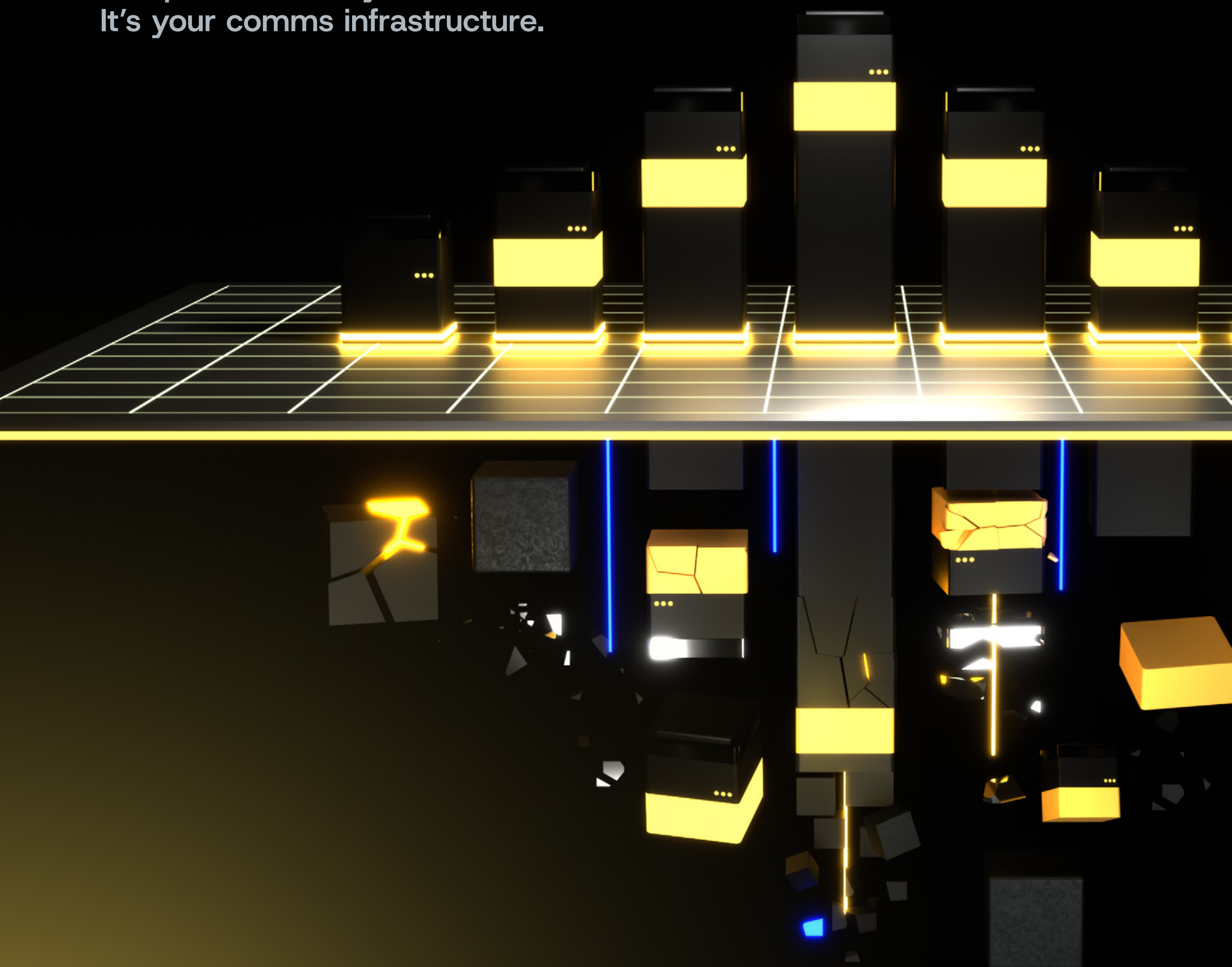




Table of Content

CHAPTER 1	
Pilot purgatory is over	4
CHAPTER 2	
Production is where the real story starts	12
CHAPTER 3	
The real cost of governance failures	20
CHAPTER 4	
The visibility gap that’s dividing leaders	27
CHAPTER 5	
Investing in trust isn’t enough	35
CHAPTER 6	
Infrastructure is where the race gets decided	42
CONCLUSION	
About the research	51

INTRODUCTION

What 2,527 enterprise leaders told us about deploying AI communications at scale

Escaping pilot purgatory was supposed to be the hard part. But for most enterprises, it wasn't. In customer communications, 62% have already shipped AI into production.

What happens next is not what the market expected. Here's what 2,527 enterprise decision-makers across 10 countries told us about the state of AI in customer communications: the critical blind spots in today's AI strategy, and the questions every communications leader should be asking right now.



The industry has assumed that better governance leads to better outcomes. But that's not enough: If governance was the fix, the most mature teams would roll back less, not more. **Our data points to a deeper issue.**

Engineering teams are spending most of their time building and maintaining safety systems, a lot of which their communications infrastructure should be providing, instead of focusing on improving the customer experience. That's the guardrail tax that slows organizations down.

Daniel Morris
CPO at Sinch

CHAPTER 1

Pilot purgatory is over

For two years, the dominant story in enterprise AI has been about being stuck. That story is accurate for enterprise AI broadly. But for AI in customer communications specifically, something different is happening.

These organizations have made it to production. The debate about whether to deploy AI in customer communications is over. What happens next is not what the market expected.



What the **market** is saying

There's a story everyone believes in enterprise AI: Pilot purgatory. The narrative has been set for years – the biggest battle is pushing an AI pilot all the way to production. In 2025, [McKinsey](#) reported that two-thirds of organizations remained in pilot or experimentation phases, and [Gartner](#) has found that at least 50% of generative AI projects are abandoned after proof of concept.

What's more, most have struggled to get value from their AI investments. [BCG](#) found that 60% had yet to show any material value. And only 15% of AI decision-makers report an EBITDA lift according to [Forrester](#), who predicts enterprises will delay 25% of planned AI spend into 2027.

2/3

of organizations remain in pilot or experimentation phases (McKinsey, 2025).

50%

of Gen AI projects will be abandoned after proof of concept (Gartner, 2026).

60%

of organizations have yet to show any material value from their AI investments. (BCG, 2025).

15%

of AI-decision makers report an EBITDA lift from AI – most still don't. (Forrester, 2026).



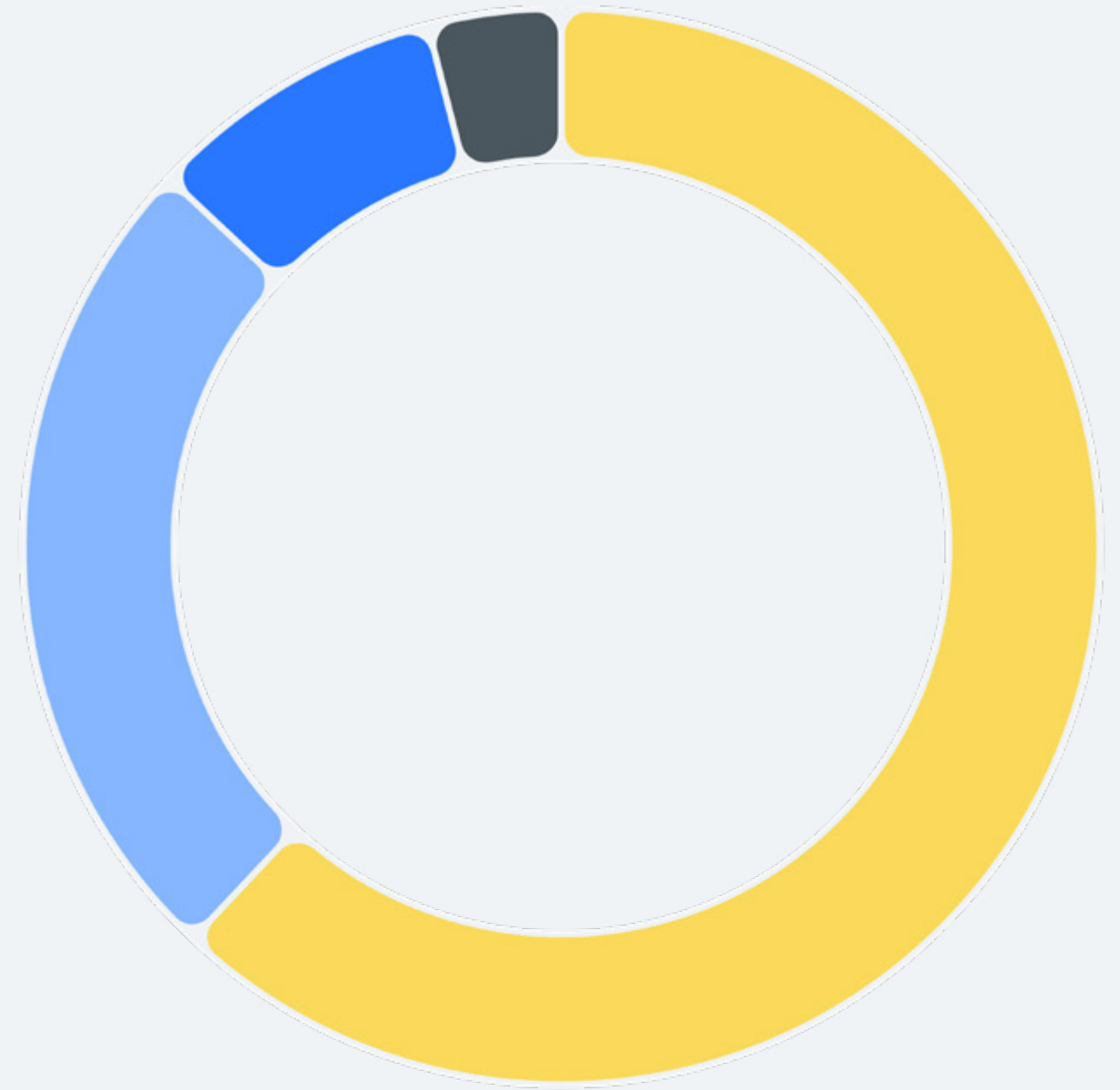
Customer communications has **left pilot stage**

[But in AI customer communications, that story is no longer true.](#) Data from 2,527 enterprise decision-makers across ten countries and six different industries shows most companies are shipping their AI pilots into production.

Our research shows **62% of organizations already have AI agents live across customer channels**, and 88% will be in production within 12 months. That's nearly nine in ten businesses actively deploying AI agents at the end of 2026.

And despite what many would think, the most tech-forward organizations are not significantly ahead when it comes to AI comms. 69% of tech organizations are in production versus 61% of non-tech organizations. That's progress, but it's not a generational lead.

Which best describes your organization's stage of adopting AI agents (autonomous or semi-autonomous systems with capabilities beyond basic chatbots) for customer communications?



- 62%** Already using AI agents in production
- 25%** Deploying within 12 months
- 9%** Deploying within 2-3 years
- 4%** Planning but no timeline set

Sinch research (2026) shows **62% of organizations** have AI agents live across customer channels.



What I've seen in AI adoption is a pretty hard acceleration across businesses. People thought maybe enterprises hadn't been adopting AI, hadn't been deploying chatbots. We've seen the actual opposite. When we talk to enterprises, it's not about getting an AI agent into production. **62% are already in production.** And by the end of 2026, that's going to be nearly 90%.

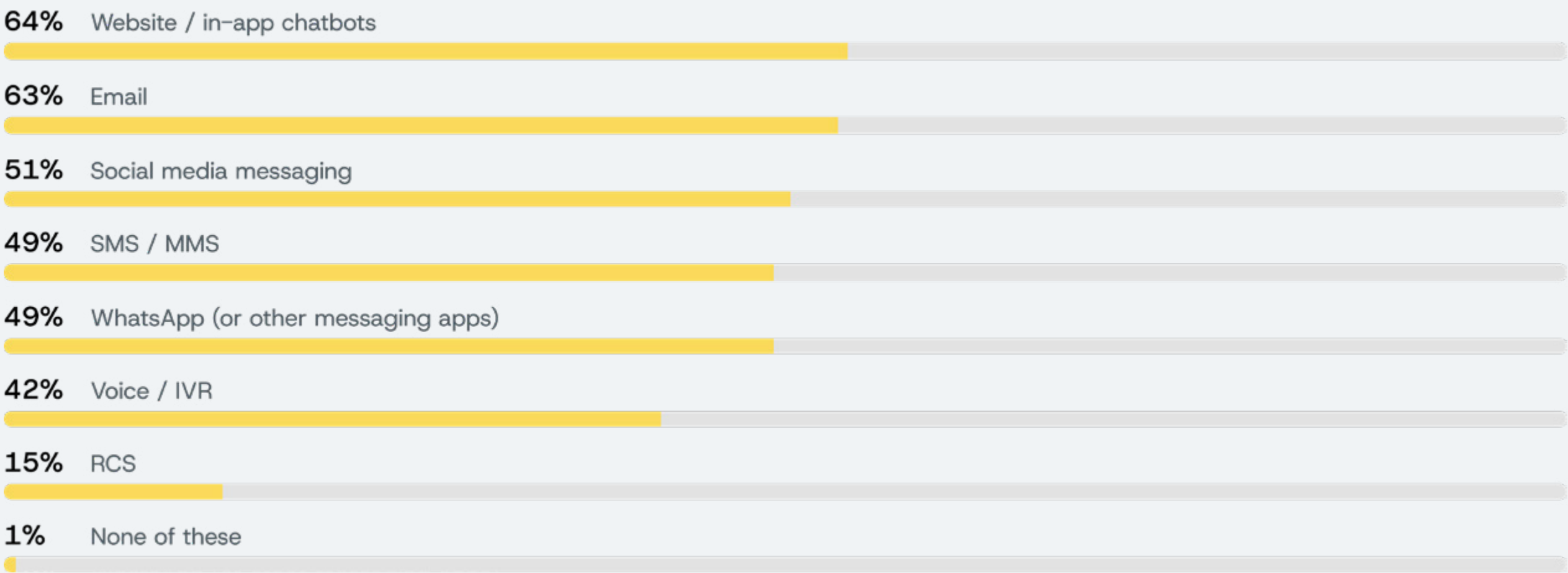
Daniel Morris
CPO at Sinch

What businesses are deploying and where

Enterprises aren't running AI on one channel and watching what happens. The average deployment spans 3.3 channels simultaneously, with nearly half of enterprises running AI across four or more. [The most popular use cases for AI agents across industries](#) include customer support, identity verification, and fraud prevention.

The channel mix is also strong. When asked which channels they are planning to integrate with AI agents, respondents primarily cited web chatbot (64%), [email](#) (63%), social media (51%), WhatsApp (49%), SMS/MMS (49%). [Voice and interactive voice responses](#) sit at 42%, but [AI has moved voice from backlog to priority](#). 67% of enterprises now rate it as equally or more important than text and messaging, and 41% see fully autonomous call handling as the single biggest voice opportunity.

Which customer communication channels does your organization plan to integrate with AI agents?
(Select all that apply)



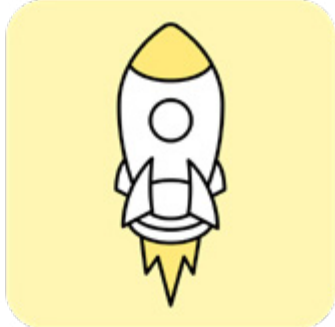
Sinch research (2026) shows web chatbot (64%), email (63%), social media (51%) are the top channels for integrating AI agents.

67% of enterprises now rate it as equally or more important than text and messaging, and **41%** see fully autonomous call handling as the single biggest voice opportunity.

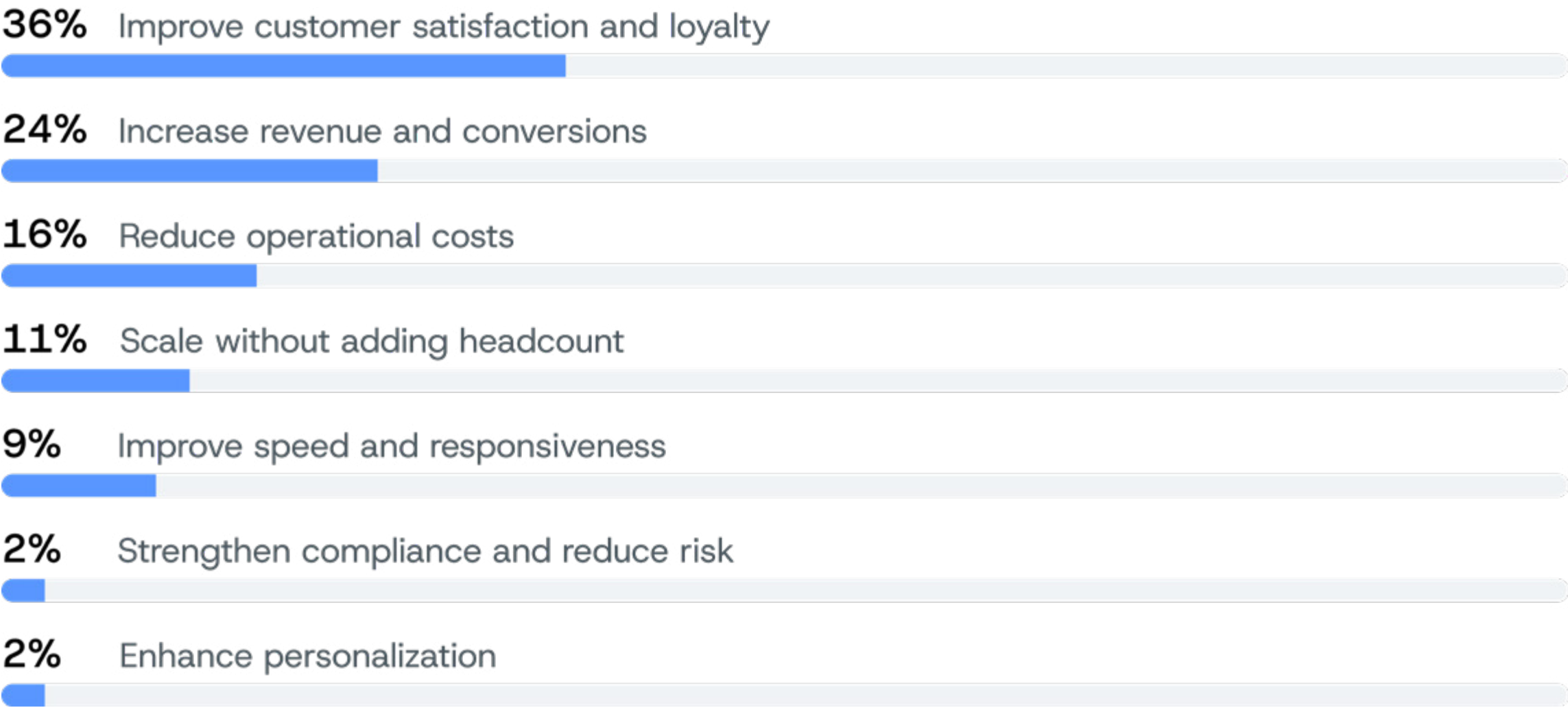
What organizations are trying to achieve

With nearly half of enterprises running AI across four or more channels simultaneously, what they’re trying to achieve with all of it matters. Sinch research (2026) shows that, for 36% of respondents, the primary goal driving investment decisions is improving customer satisfaction and loyalty. Revenue growth comes second at 24%. Cost reduction is third at 16%.

What’s interesting is that the organizations investing most aggressively are not trying to cut headcount. [They’re competing on customer experience](#) and trying to earn something harder to measure than efficiency or cost: **customer trust**.



What is the single most important outcome your organization is trying to achieve with AI agent-powered customer communications?



Sinch research (2026) reveals the main goals driving AI investment decisions are improving customer satisfaction (36%), revenue growth (24%), and cost reduction (16%).

The investment case **is settled**

The lack of investment is no longer a question in AI customer communications: 98% are increasing spend in 2026, and only 0.3% plan to cut. What’s more, only 7% of respondents cited budget constraints as their main barrier to driving business impact through agentic customer communications.

And the ROI case is also real. Among leaders already in production, **nearly 60% estimate operational efficiency gains** and customer satisfaction improvements of more than 25% within two years, with similar gains anticipated for revenue and cost reduction.

98%

are increasing spend in 2026, and only 0.3% plan to cut

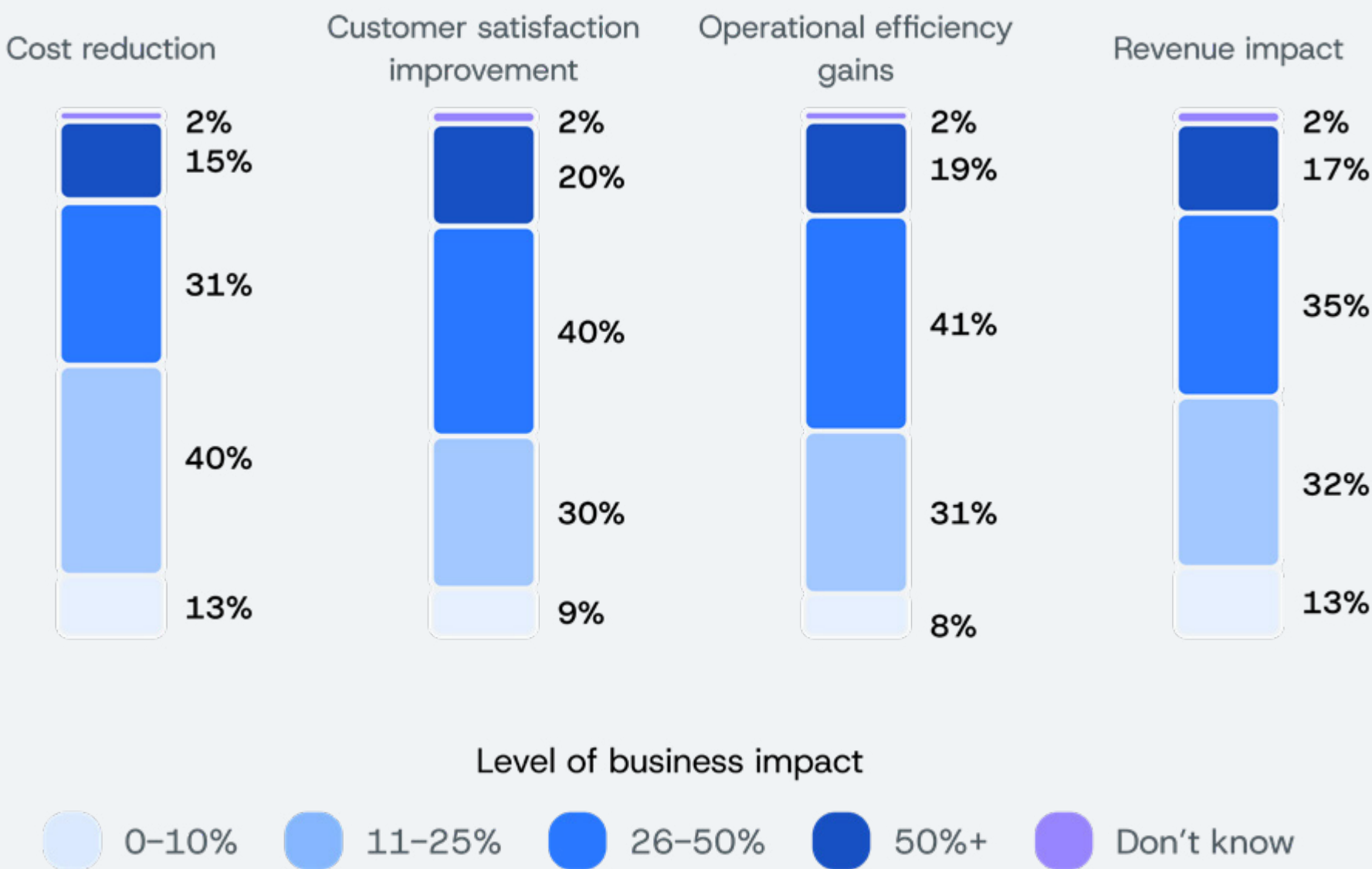
60%

expect 25%+ gains in efficiency and customer satisfaction

7%

of respondents cited budget constraints as their main barrier

What business impact do you expect from AI agent-powered customer communications in the next 2 years?



Sinch research (2026) shows the expected business impact of AI agent-powered customer communications in the next two years.



So, are leaders actually putting money behind AI communications? Well, **98% of the leaders** that we surveyed are increasing their AI comms spend in 2026. Only 0.3% are planning to cut. So that debate is pretty much settled.



Robert Gerstmann
Chief Evangelist at Sinch

THE REGIONAL VIEW

Countries are sequencing investment and deployment differently

Pilot purgatory is over globally, but the exit isn't uniform. Deployment rates range from 55% in EMEA to 67% in APAC, and the country-level gaps are wider. Brazil leads at 76% – 14 points above the global average. Germany has the lowest AI deployment rate at 46% – 16 points below it.

The more interesting story, though, is how different countries are sequencing investment and deployment. Some markets are still building toward their AI ambitions. Others are further along than the global numbers suggest.



India and Brazil are the countries where ambition and action match best. Both report high deployment rates (68% and 76%) and rank first and third globally for plans to increase investment by 25% or more.



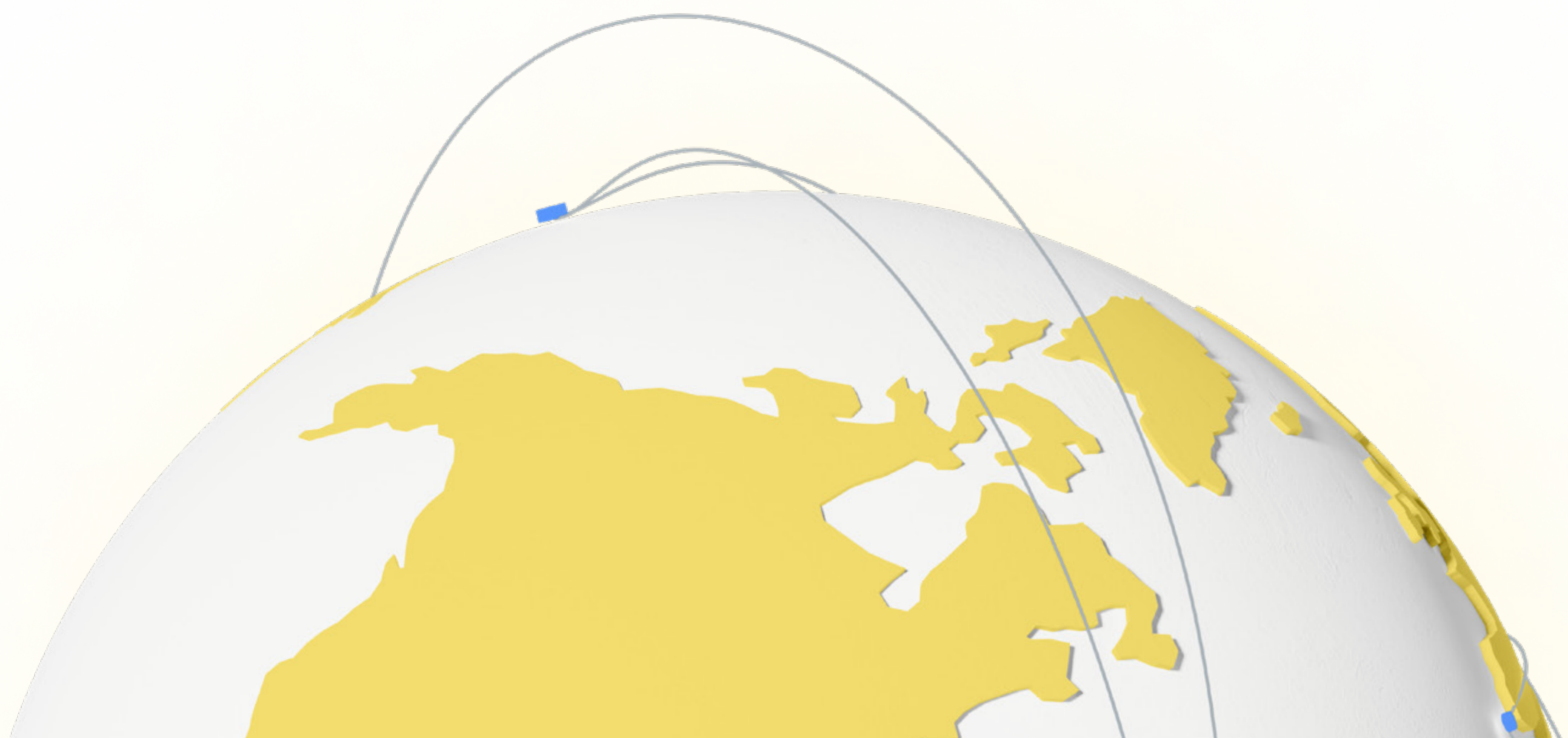
Germany has the most aggressive AI investment plans in Europe – 53% plan to increase investment by 25% or more – but its deployment rate (46%) is 16 points below the global average.



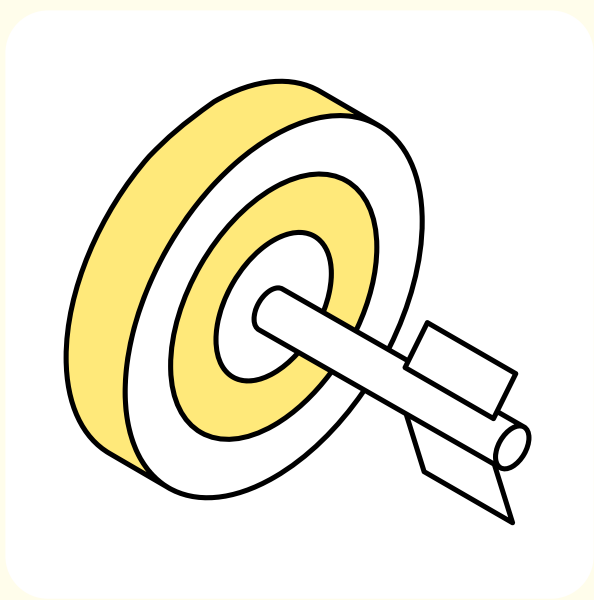
In **Singapore**, deployment is well ahead of investment intent. 72% have an AI customer communications agent live but only 40% plan to increase investment by 25% or more in 2026.



Canada shows the most caution. 49% have AI agents in production, and only 42% plan to increase investment by 25% or more in 2026 – compared to 67% deployment and 61% investment intent in the US.



What to do with this



62% of organizations have an AI customer communications agent live. 88% will have one by the end of 2026.

Getting to production was hard, and most of enterprises have made it. But the data suggests escaping pilot purgatory wasn't the hardest part.

The organizations pulling ahead aren't debating production timelines or investment cases anymore. They've deployed, they're scaling – and what they've found on the other side is not what the market expected.

AI customer communications has finally left the pilot stage. But what happens after enterprises successfully ship radically changes the question every AI communications leader should be asking right now.

CHAPTER 2

Production is where the **real story** starts

Nearly three in four enterprises have been forced to shut down or roll back a deployed AI communications agent. That number holds across every region and every industry in the study, and it doesn't decline with experience nor investment. In fact, among organizations that describe their guardrails as fully mature, the rollback rate is even higher.

This is happening in production, on customer-facing channels, in real interactions. When something goes wrong, customers are watching. They notice. And when that happens, a technical issue becomes a real trust event.





```
01 import requests
02
03 applicationKey = "<REPLACE_WITH_APP_KEY>"
04 applicationSecret = "<REPLACE_WITH_APP_SECRET>"
05
06 to_number = '<REPLACE_WITH_YOUR_NUMBER>'
07
08 sinchVerificationUrl = "https://
09 verification.api.sinch.com/verification/v1/
10 verifications"
11
12 payload = {
13     "identity": {
14         "type": "number",
15         "endpoint": to_number
16     },
17     "method": "sms"
18 }
```

Shipping **didn't solve** the problem

In AI customer communications, most companies have already shipped. By the end of the year, nine in ten organizations will have a live AI agent in production across one or more of their channels.

But 74% of those that have already shipped have been forced to roll back or shut those agents down. Among organizations that describe their guardrails as fully mature, the rollback rate is 81%.

All along, the market has been drawing the wrong finish line. We thought getting to production was the hard part. The data shows it isn't. Instead of solving the problem, organizations are deferring it to a moment when stakes are higher, and customer trust is on the line.

74%

Average rollback rate among organizations that reached production.

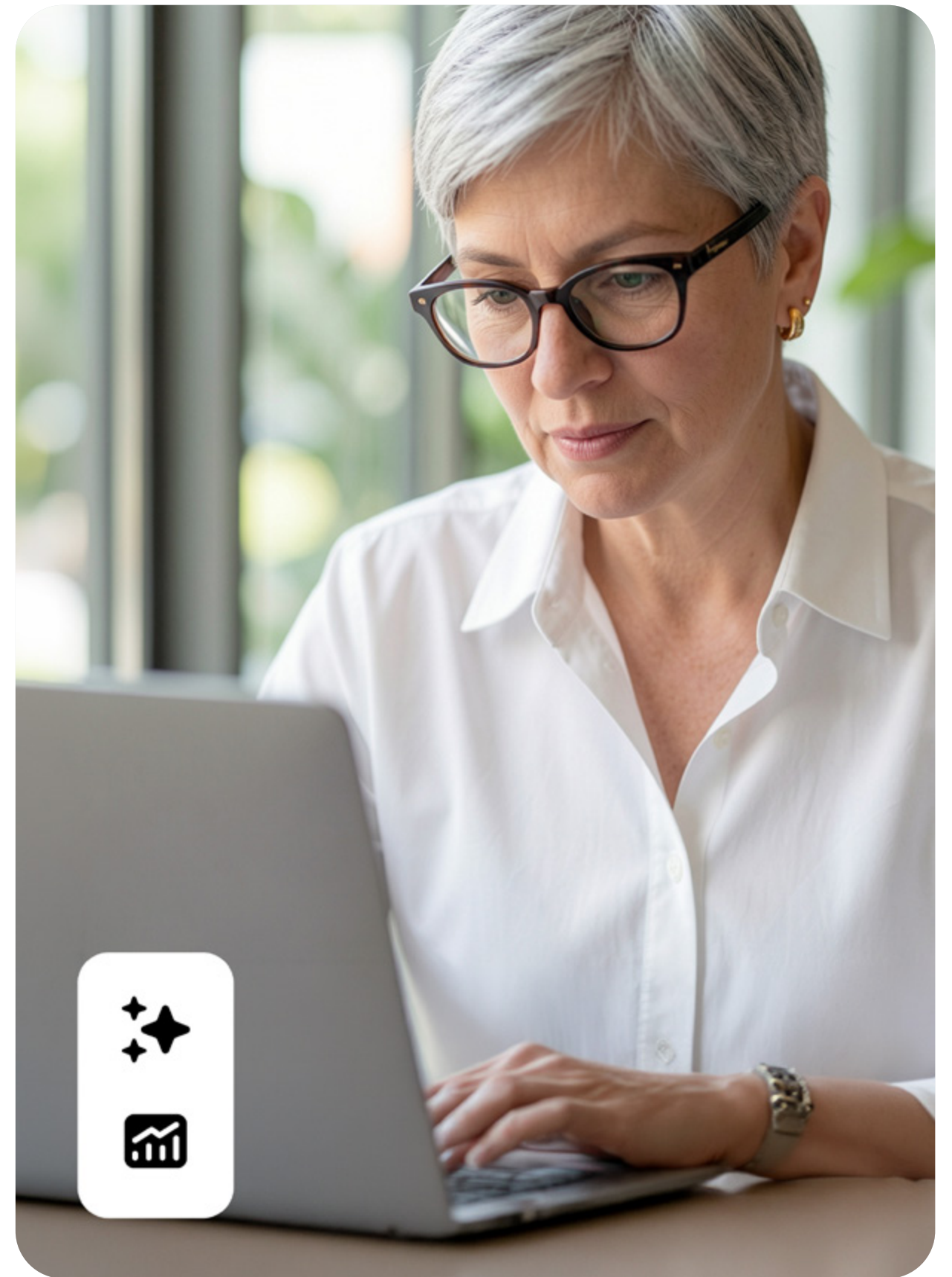
81%

Rollback rate among organizations with fully mature guardrails.

Why rollback rate **increases** with mature organizations

Organizations with mature governance instrumentation can see failures that less mature organizations miss entirely. Those reporting no rollbacks aren't necessarily running cleaner AI programs. In many cases, they simply lack the monitoring to know when something goes wrong.

The most likely explanation, supported by the broader data, is that the **programs failing at 81% – those that reported fully mature guardrails – aren't failing because they're poorly run. They're failing because their [improved monitoring allows them to detect issues that others can't spot.](#)** The organizations reporting no governance failures are not the benchmark. Instead, they may be the ones with the least visibility into what's happening.



The more you can monitor and understand what an AI agent is doing in production, the better. That allows you to **discover problems before your customers do**, which means you can get ahead of any negative interactions and preserve your brand's value while still scaling with AI communications.

Daniel Morris
CPO at Sinch



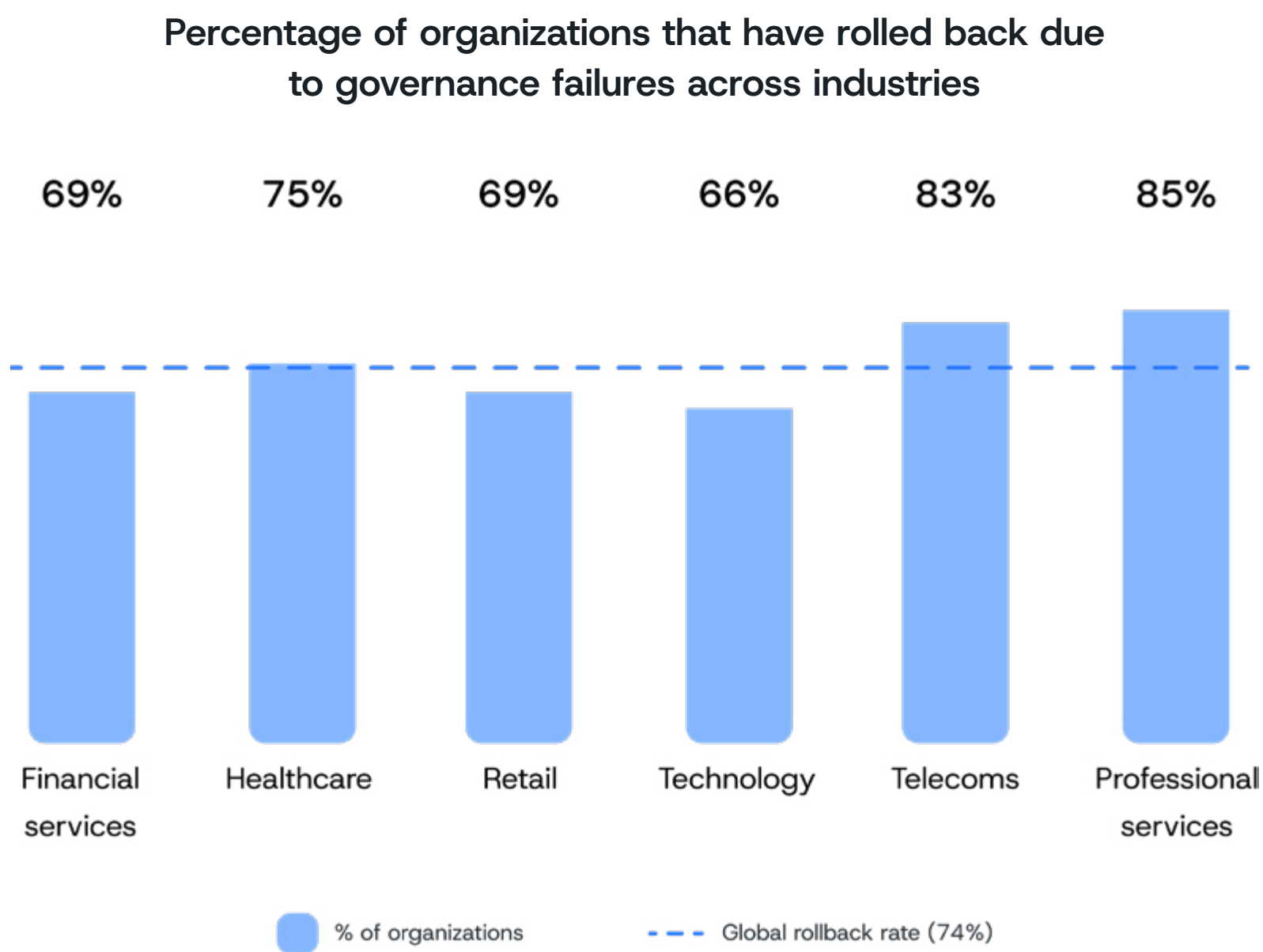
VERTICAL SPOTLIGHT

No industry has solved the rollback problem

Rollback rates are high across every industry – from 66% in technology to 85% in professional services. The problem is universal.

The market has assumed compliance complexity drives higher failure rate, but our data doesn’t support that. Healthcare converts the fewest pilots to production but tracks with every other industry across all 19 compliance and privacy measures, and its rollback rate sits just 1 point above the global average. In financial services, the organizations with the most mature guardrails are also the furthest along on deployment: Each step up in guardrail maturity makes an organization twice as likely to be operating at the most advanced level. But the rollback rate is 5 points below the average at 69%.

The market has also assumed that tech-forward sectors have found a way through. And while technology companies do have the lowest rollback rate at 66%, that’s only 8 points below the global average. No industry has really found a way past it.



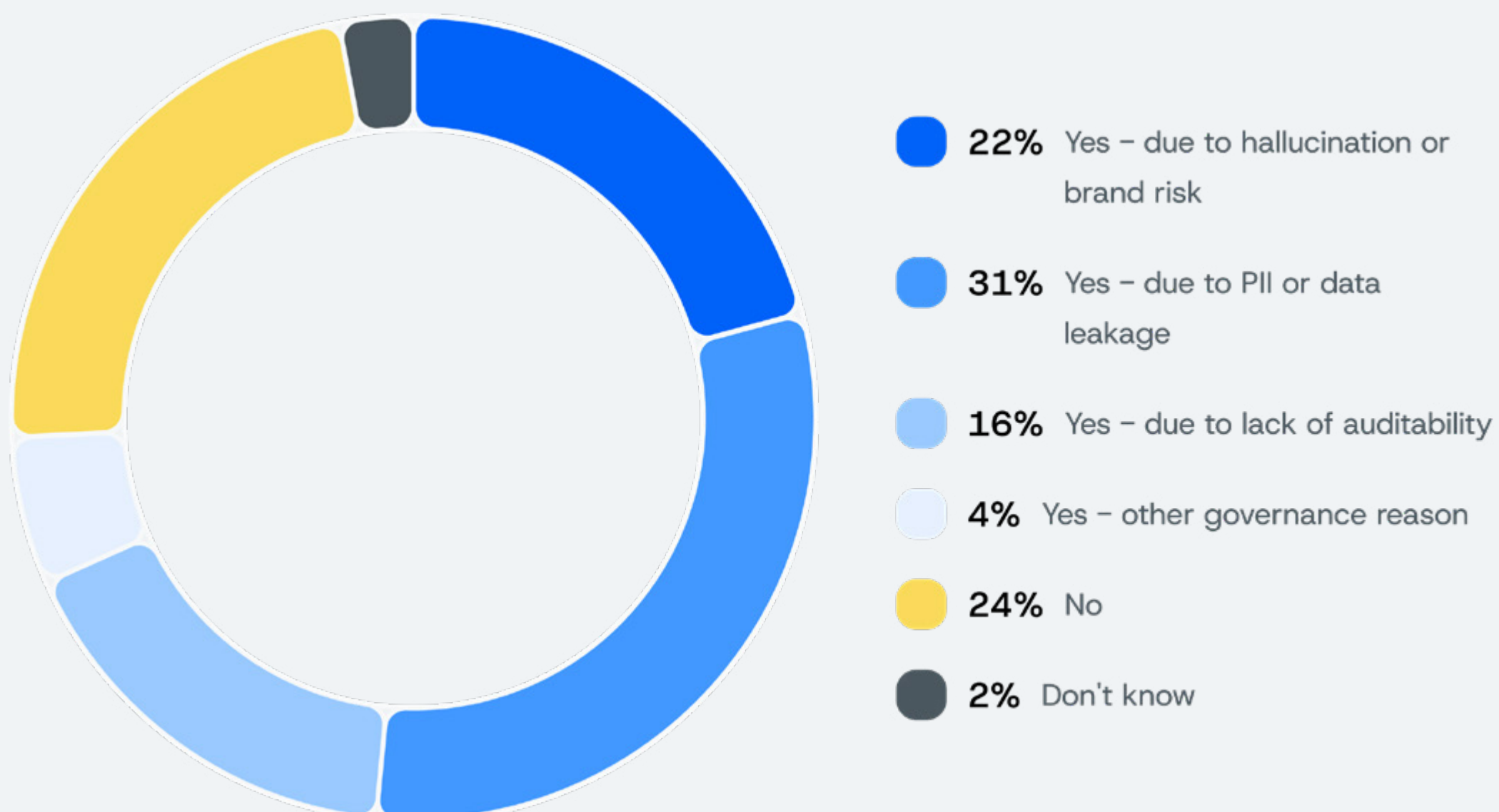
Sinch data (2026) shows rollback rates range from 66% in the technology industry to 85% in professional services.

What's triggering the rollbacks

Personal identifiable information (PII) or customer data exposure is the leading cause among those that reported a governance failure rollback, cited by 31% of organizations. Hallucination or brand risk is second at 22%. Lack of auditability – the inability to diagnose what went wrong – is third at 16%.

These aren't abstract risk categories. PII exposure means a customer's personal data surfaced in an interaction it shouldn't have. Hallucination means an AI agent said something confidently wrong to a real customer, on a live channel, under your brand's name. And when 16% of rollbacks can't be fully diagnosed because there's no audit trail, the organization is left with a failure it can't learn from and no way to prove it's been fixed.

Has a deployed AI agent ever been rolled back or shut down due to a governance failure?



Sinch data (2026) reveals PII or data leakage (31%) and hallucinations (22%) are the main reasons for rolling back an AI agent.

16% of rollbacks can't be fully diagnosed because there's no audit trail. When that happens, organizations are left with a failure it can't learn from and no way to prove it's been fixed.

Confidence is not protection

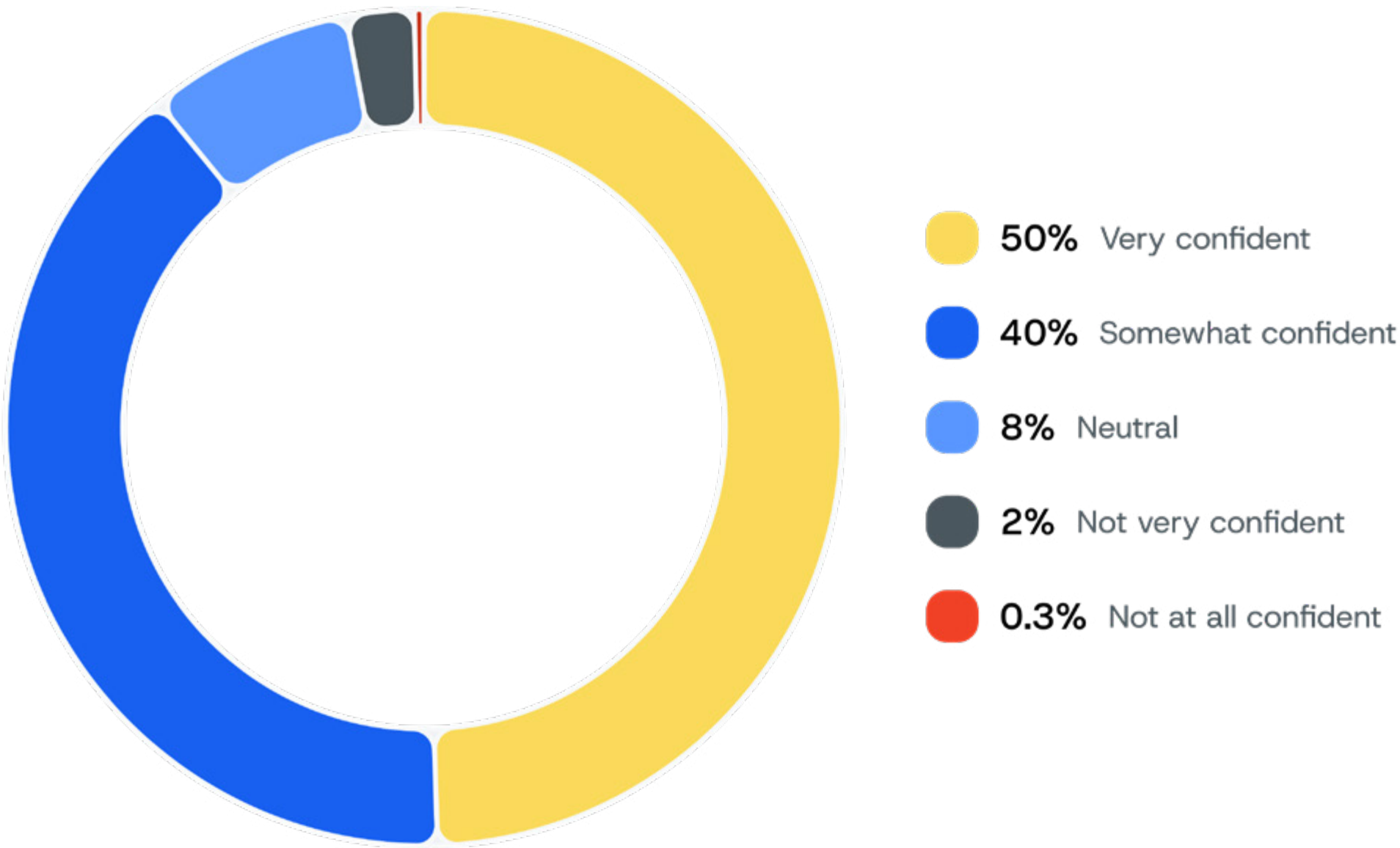
90% of enterprise decision-makers describe themselves as confident in their AI agent readiness. But the moment you stack those confidence scores against operational reality, a gap opens.

Of those that rate their confidence as “somewhat” or “very confident”, **75% have experienced at least one governance rollback**. Confidence doesn’t correlate with fewer failures. In fact, both deployment and rollback rates are broadly consistent regardless of how ready organizations feel.

Confidence is a leading indicator of ambition, not a guarantee of governance readiness. The more useful question for any leadership team isn’t “are we confident?” It’s “what would we see if something went wrong right now, and how fast would we see it?”



How confident are you that your organization is ready to deploy AI agents that autonomously communicate with customers at scale?



Sinch research (2026) shows 90% of business decision-makers describe themselves as confident in their AI readiness.

THE REGIONAL VIEW

Market dynamics shape rollback rates

The 74% global rollback rate reflects a problem that holds across every region, but the failure modes look different depending on where you are.



APAC leads on deployment and on rollbacks. 83% of organizations have shut down or rolled back an AI agent – the highest of any region. 38% of those failures are driven by PII leaks. **Australia has the most rollbacks** in the region at 84%.



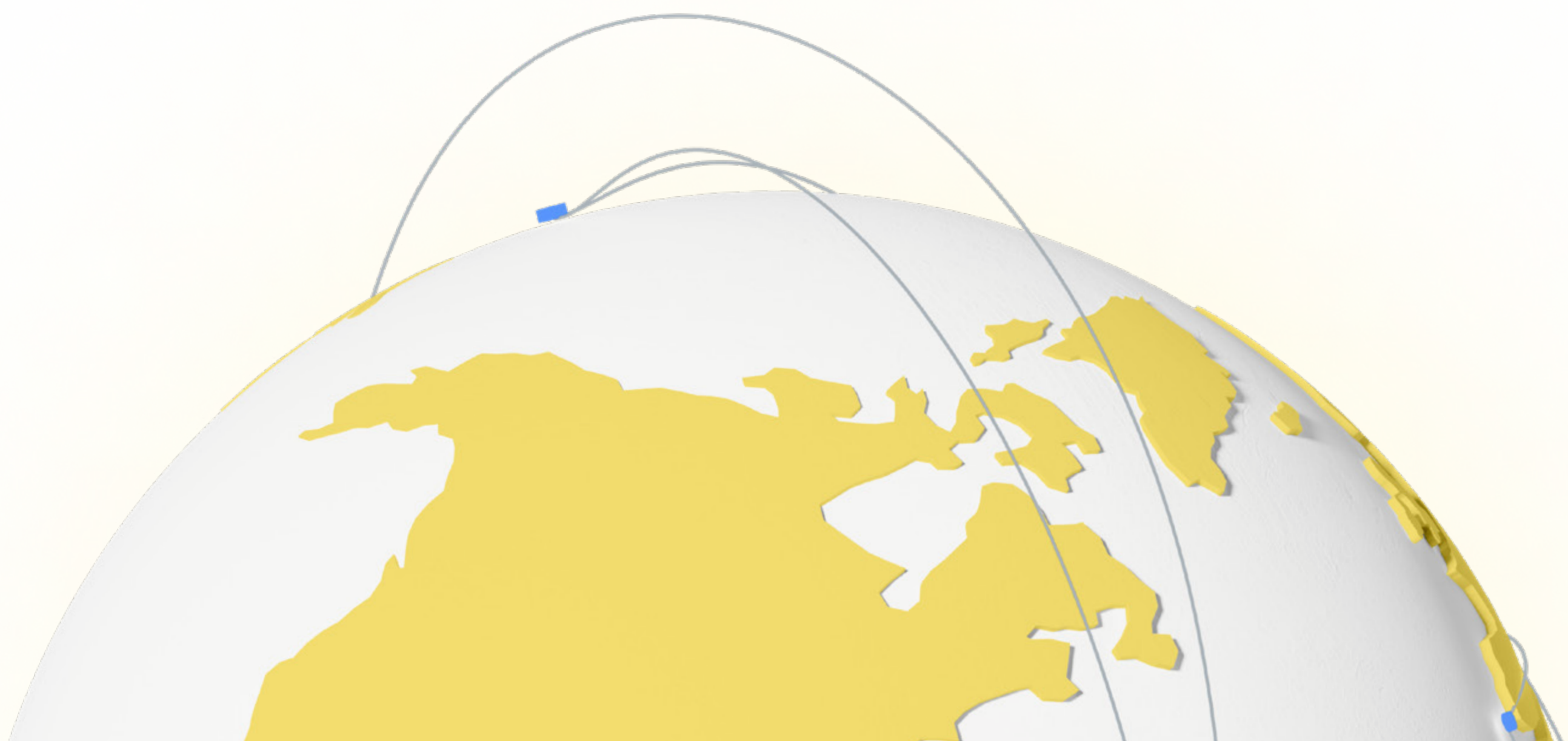
Latin American **organizations** roll back at 82% and report the highest rate of PII-related incidents (41%) across a region – ten points above global average. **Brazil is the fastest-deploying** market in the hemisphere and one of the fastest to fail.



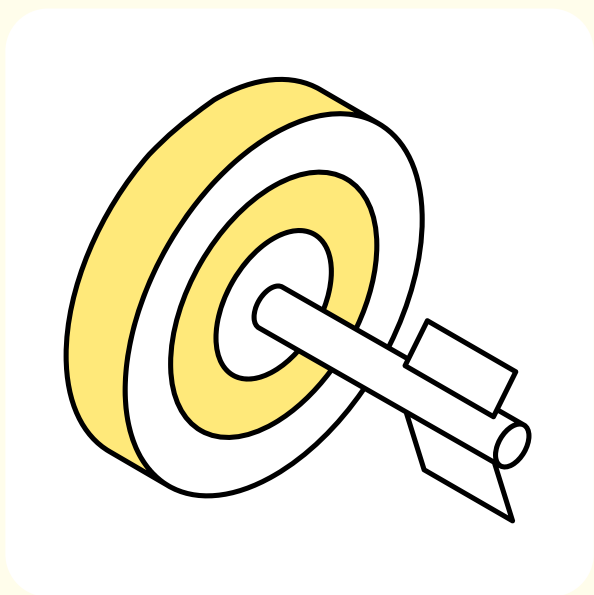
EMEA sits at 70% rollbacks overall, but France is an outlier. **90% of French organizations have rolled back** an AI agent, with 42% of those rollbacks PII-related. CNIL enforcement is creating a specific failure mode that standard GDPR compliance isn't preventing.



North America has the lowest rollback rate of any region at 63%, and **34% report zero rollbacks**. Where failures occur, hallucinations and PII exposure are tied as the leading cause, the only region where model failure matches infrastructure failure.



What to do with this



74% of organizations have had to shut down a live agent.

That rollback rate is a reason to be precise about what you're building on and what you'll be able to see when something goes wrong. And the uncomfortable implication of the 81% figure isn't that mature programs are failing more – it's that they're able to see what less governed organizations are missing.

One question worth reflecting on: If your AI communications agent failed right now, would you know before your customers did?

The monitoring gap is where organizations are most exposed, and where a technical issue suddenly becomes a customer trust issue.

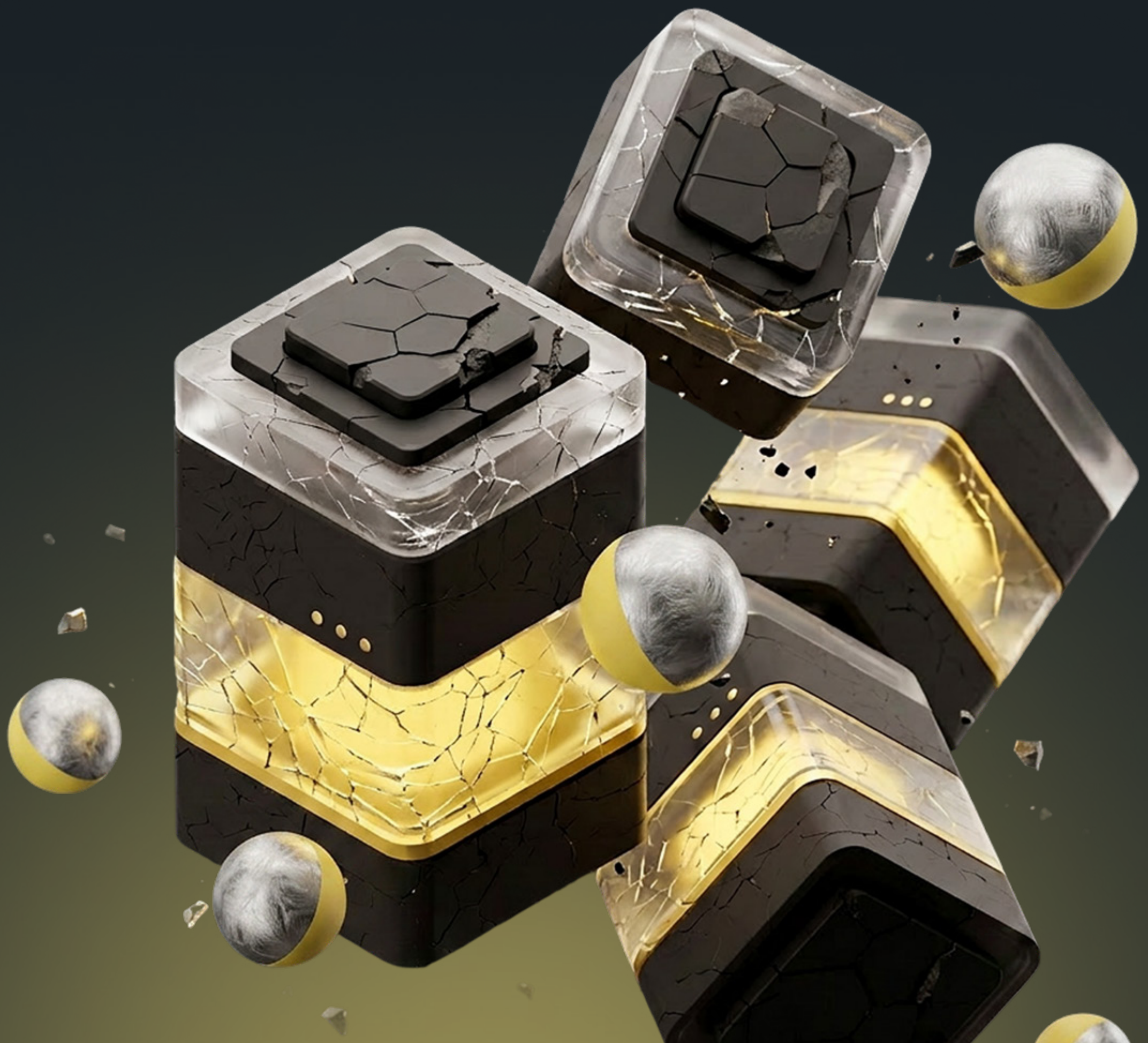
CHAPTER 3

The **real cost** of governance failures

When an AI communications agent fails in production, customers notice. Sinch data shows the impact splits in three directions: the support queue, the brand perception, and the engineering cost.

Most organizations are only tracking – and trying to mitigate – one of them while the other two go unnoticed and compound over time.

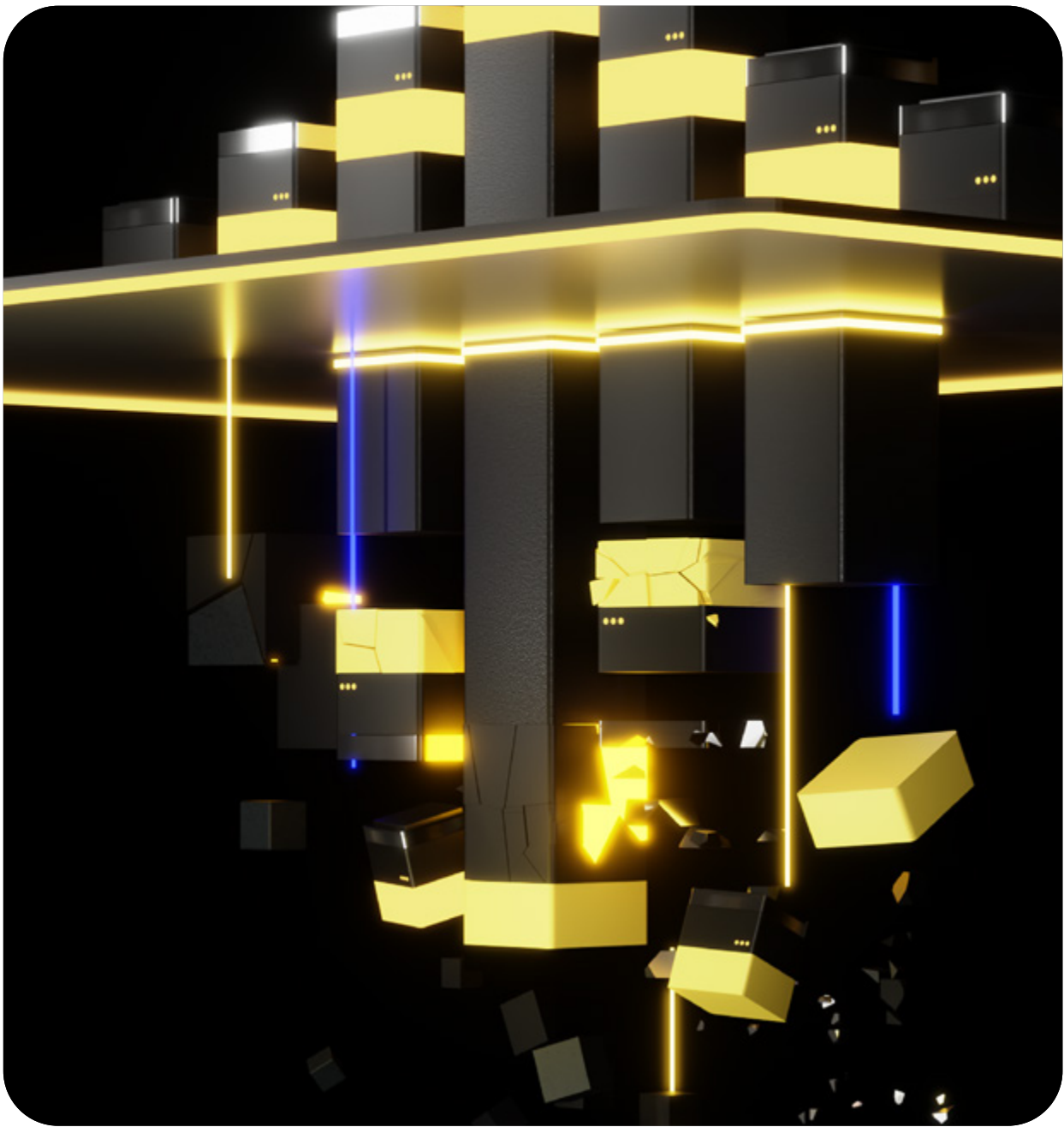
Pilot purgatory wasn't the problem. It was the warning.



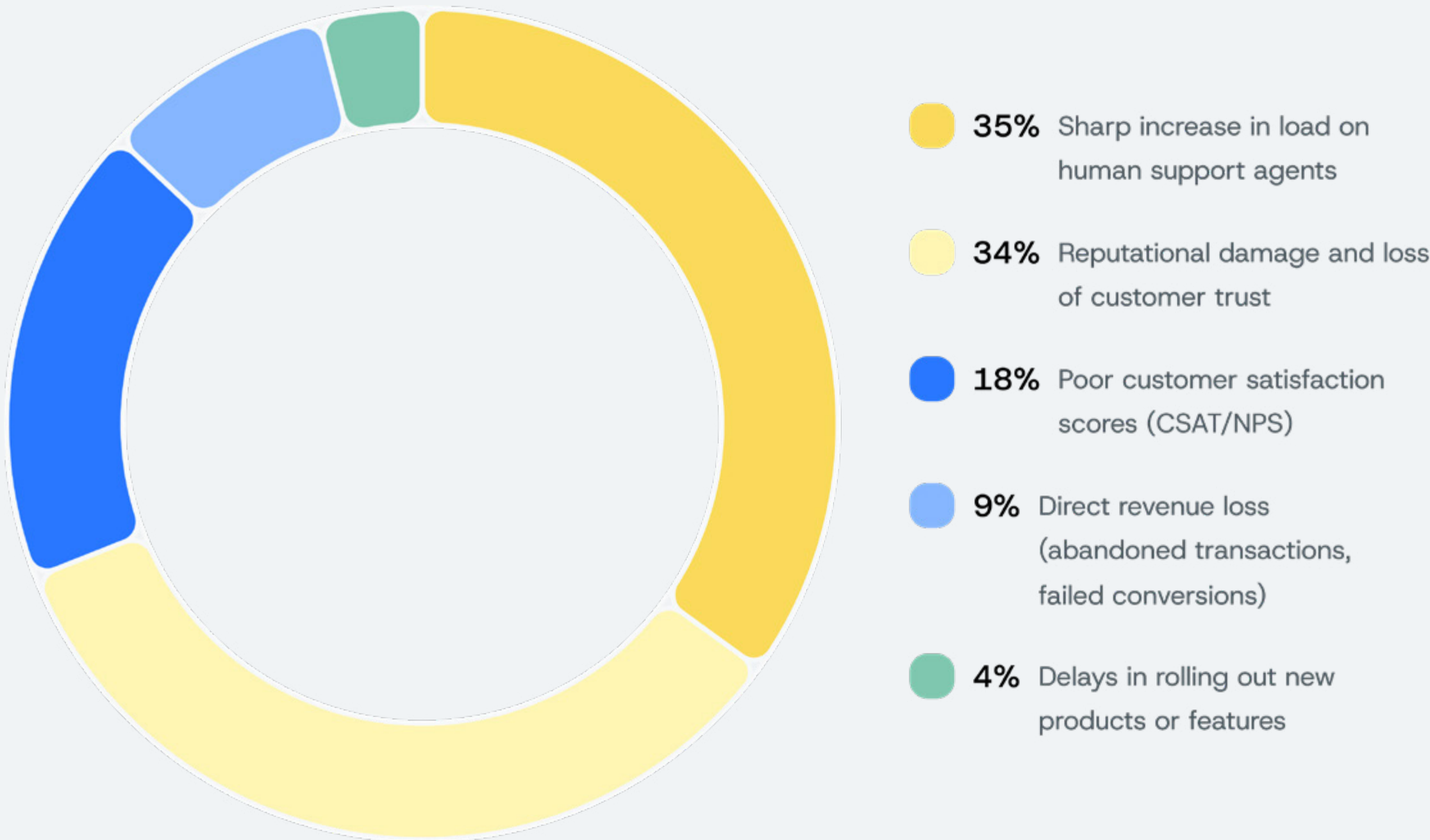
The business impact of AI failure

When asked what’s the single biggest impact of an AI agent’s failure, around a third of respondents cite a surge in human support agent load. Reputational damage and loss of customer trust run almost level with it.

That near-tie shows how complex these AI failures can be, because these two modes – the support load and the reputational impact to the brand – are not equivalent in how they resolve. One has a clear path to resolution, while the other might have a longstanding impact on your organization that’s harder to mitigate.



When an AI agent-driven customer interaction fails due to a technical issue (slow response, lost context, incorrect information), what is the most significant business impact?



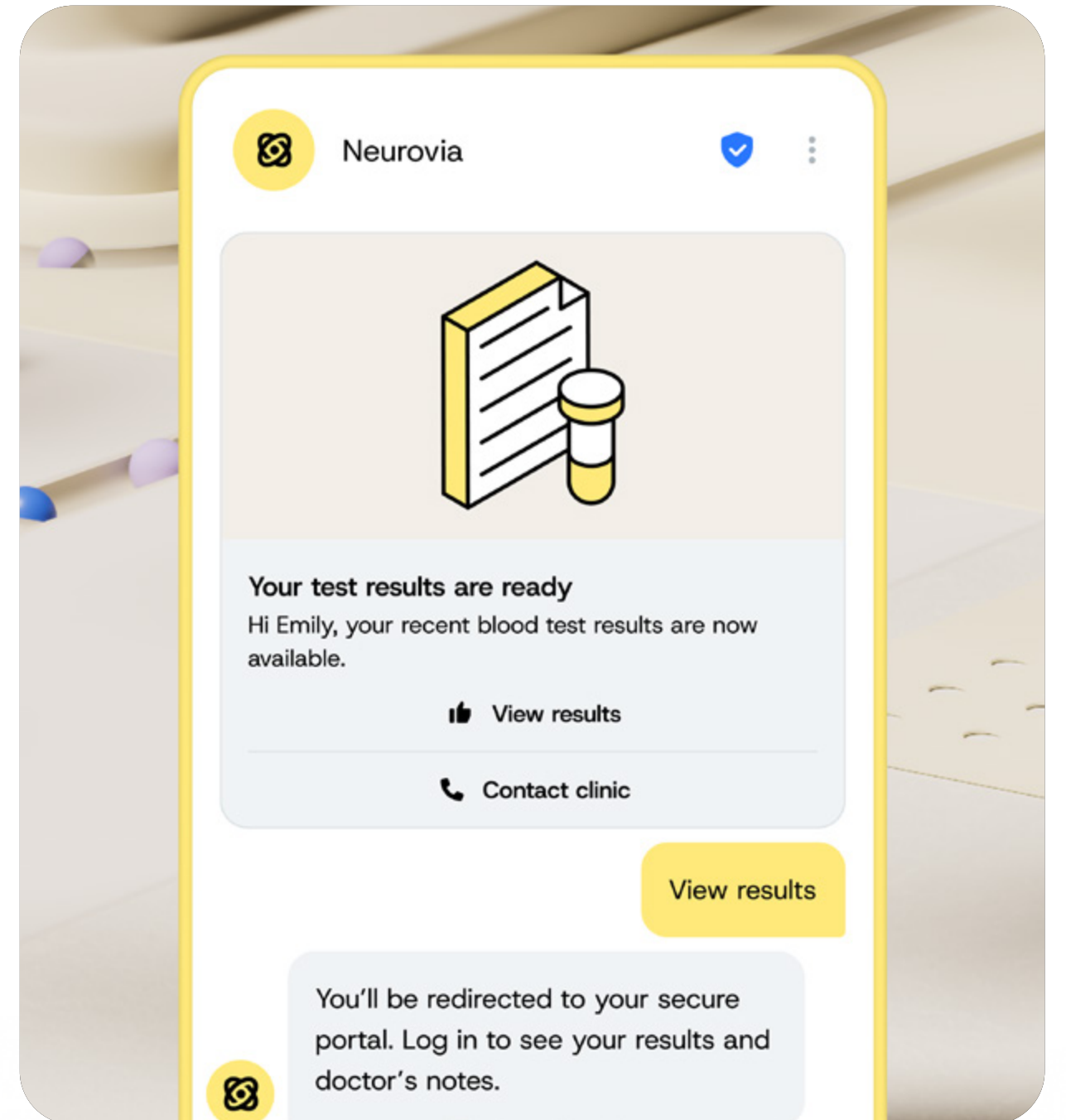
Sinch research (2026) shows an increase in the support queue (35%) and reputational damage to the brand (34%) are the biggest impacts of AI agent failure.

The support queue

35% of organizations cite a surge in human support agent load as the primary consequence of an AI communications failure. The agent goes down, and every interaction it was handling reverts to a human. A support team sized for a world where AI handles significant volume is suddenly handling all of it itself.

At peak times (a product launch, a service outage, a seasonal spike) this can become a real operational crisis. And at the same time, this is also the failure mode that gets reported upward. It shows up in dashboards and generates incident reviews.

It's clearly visible and measurable, but it resolves when the agent comes back online.



This type of issue for a customer is table stakes. They expect us to respond quickly with accurate information, and when we're not there to support them, that frustrates them and erodes their view of our brand. So, we have to ask ourselves two questions: **why did this happen, and how can we ensure that this doesn't happen again?** Until you do that, you haven't actually solved the problem. You've just reset the clock.

Heather Arrington

SVP, Customer Support Operations at Sinch

The risks to the brand

Reputational damage and loss of customer trust is cited by 34% of organizations as the primary consequence of an AI communications failure – essentially tied with support overload across all industries surveyed. In the financial services and tech industries, reputational damage actually outranks support overload as the number one fear.

But unlike support overload, reputational damage doesn't have a clear resolution path. From the customer's perspective, there is no platform, there is only your brand. [That attribution is permanent in a way that a queue spike is not...](#)

VERTICAL SPOTLIGHT

For retail, every AI failure is a **brand failure**

In retail, 69% of AI agents have been rolled back. Data exposure and hallucinations have different causes, but they land in the same place: an industry whose entire AI investment case rests on knowing the customer.

The PII failure – cited by 29% – is a trust failure. Retail’s personalization model runs on a customer agreement: share your data, get a better experience. A leak breaks the contract. And once a customer decides the exchange isn’t worth the risk, the investment becomes irrelevant.

The hallucination failure – reported by 22% – is subtler and potentially worse. An AI with deep customer context doesn’t just say something wrong. It says something wrong with the confidence of a system that knows the customer well. 35% of retail leaders say driving loyalty and satisfaction is their primary AI goal. A confidently wrong agent hits that directly.

35%

of retail organizations say driving customer loyalty and satisfaction is their primary AI goal.

69%

of retail organizations have rolled back an AI customer communications agent due to a governance failure.

29%

in retail have had to roll back an AI agent due to a data leakage.

22%

in retail have had to roll back an AI agent due to hallucinations.

The engineering cost

There’s a cost that appears in neither the dashboard nor the customer complaint. When an agent gets rolled back, the engineering team goes back with it – diagnosing, rebuilding, re-testing, re-deploying – while the feature backlog accumulates.

And that engineering burden doesn’t start with a rollback. Sinch research (2026) shows **84% of AI communications engineering teams report spending at least half their time building guardrails and safety controls**, even before a single failure occurs. 35% spend most of their time there instead of on the next feature.

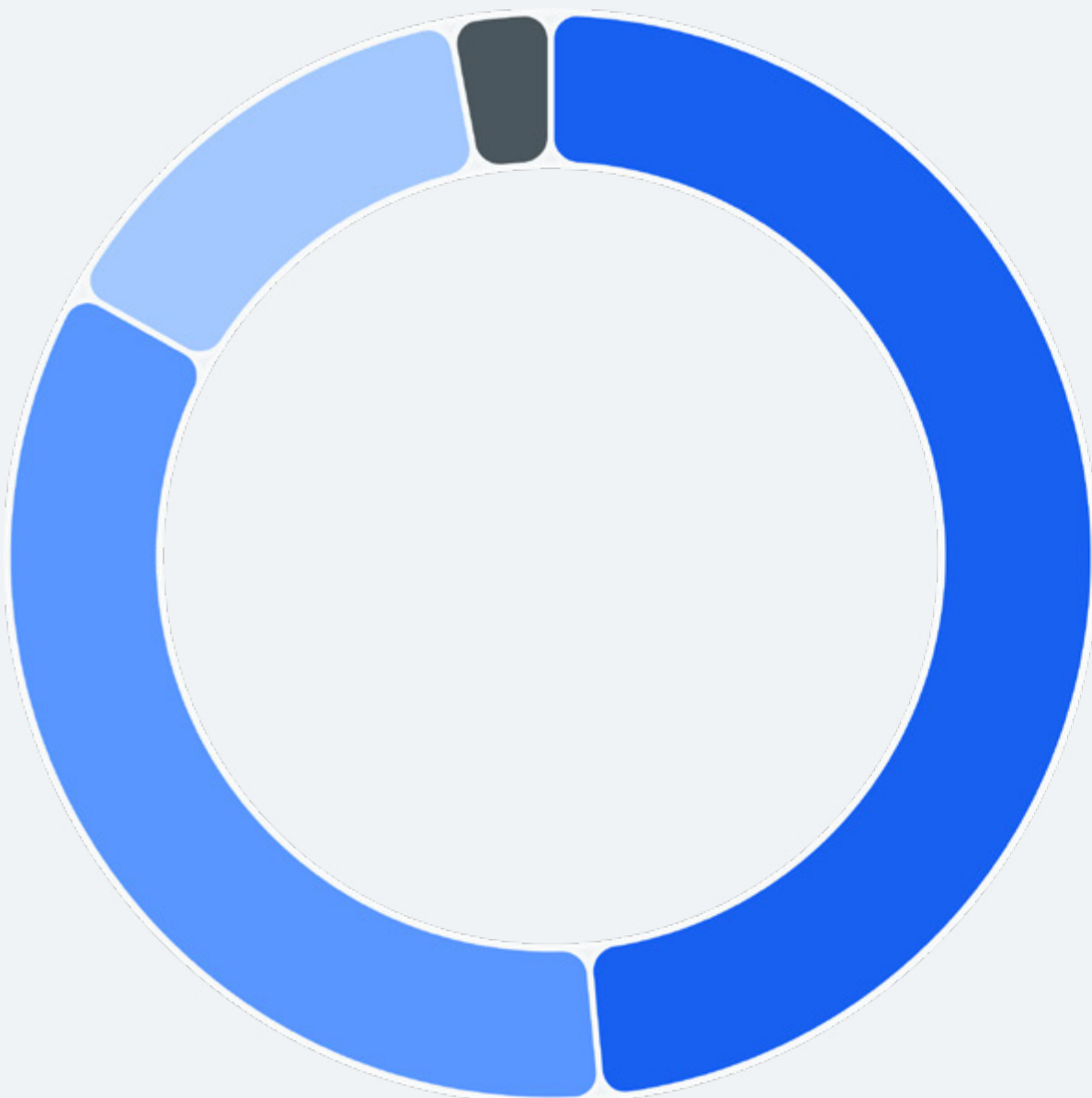
Not all that work is fixing the same thing, though. PII exposure, context loss, and audit trail gaps originate in the infrastructure layer. They’re failures the platform should be catching before they reach the agent. Hallucination and off-brand responses are a different category, model and prompting problems that no amount of infrastructure investment will prevent. The guardrail tax compounds either way, but what you’re paying to fix is different.



```
PYTHON

01 # AI communications agent - sending personalized renewal notices
02 for customer in get_customers_needing_renewal():
03     prompt = f"Write a renewal notice for {customer.name}, plan: {customer.pl
04     response = openai.chat.completions.create(model="gpt-4", messages=[{"role
05     send_email(customer.email, response.choices[0].message.content)
06     log.info(f"Sent renewal to {customer.id}") # ✓ logged as success
07 # No check: did the response contain placeholder text? A refusal? The wrong c
08 # No check: did send_email() silently drop the message? No alert threshold. N
09
10 # AI communications agent - sending personalized renewal notices
11 for customer in get_customers_needing_renewal():
12     prompt = f"Write a renewal notice for {customer.name}, plan: {customer.pl
13     response = openai.chat.completions.create(model="gpt-4", messages=[{"role
14     send_email(customer.email, response.choices[0].message.content)
15     log.info(f"Sent renewal to {customer.id}") # ✓ logged as success
16
17 # No check: did the response contain placeholder text? A refusal? The wrong cus
18 # No check: did send_email() silently drop the message? No alert threshold. No
```

How does your AI engineering team typically allocate its time between building guardrails and safety controls (rate limiting, PII masking, retry logic) versus improving AI agent functionality?



- 49% Even split between building guardrails and improving AI agent functionality
- 35% Most time is spent building guardrails and safety controls
- 14% Most time is spent improving AI agent functionality
- 2% Not sure / outside my area of visibility

Sinch data (2026) shows 84% of AI communications engineering teams spend at least half their time building guardrails and safety controls.

THE REGIONAL VIEW

Same failure, different costs

Not all regions feel the impact of AI agent rollbacks in the same way.

When an agent fails in APAC, where a third of enterprises sends over 100 million messages a month, the support exposure is immediate and massive. In EMEA, GDPR shapes how organizations approach engineering architecture. The regional context changes the costs picture.



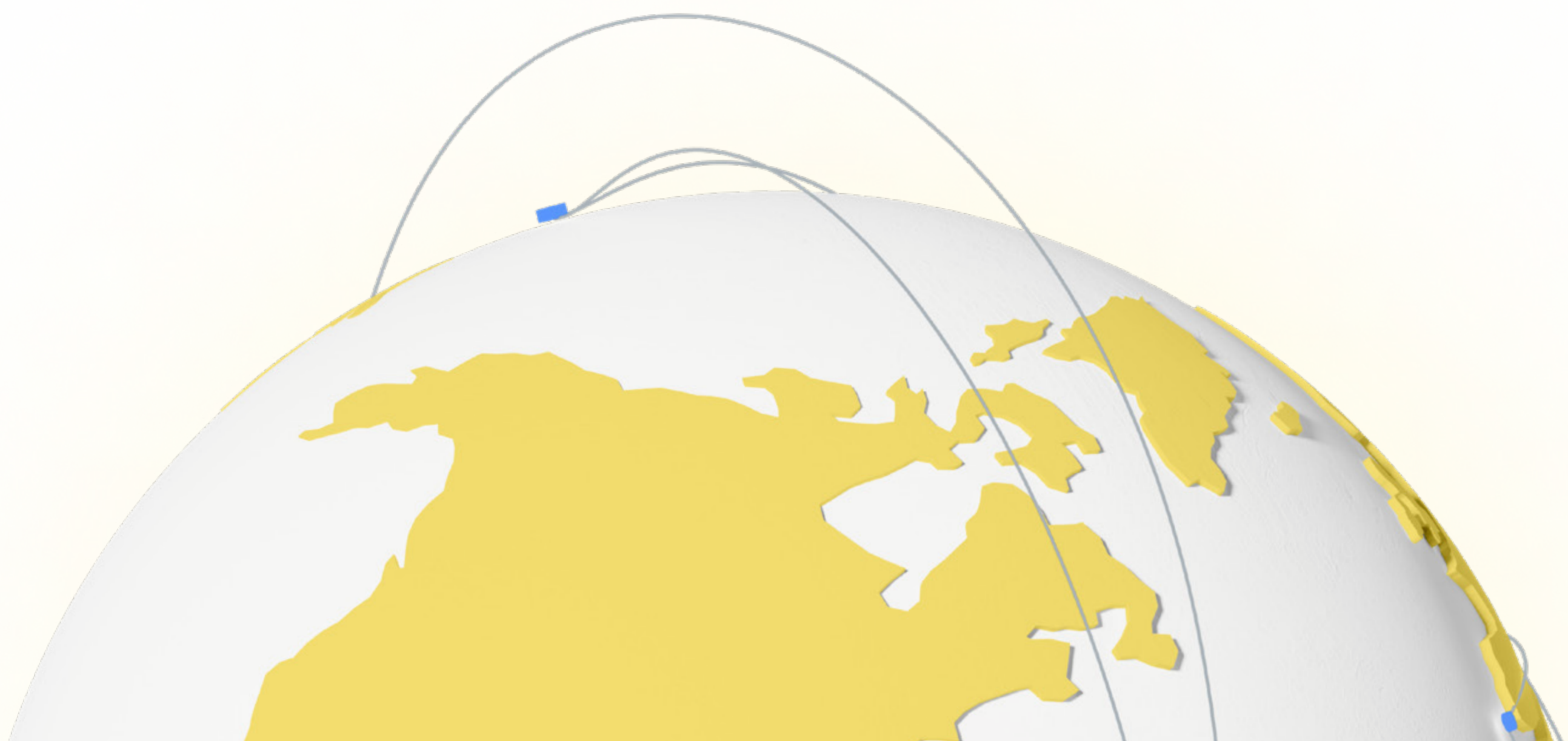
45% of **APAC** organizations cite a **surge in support team workload** as the primary consequence of AI agent failure — 10 points above the global average.



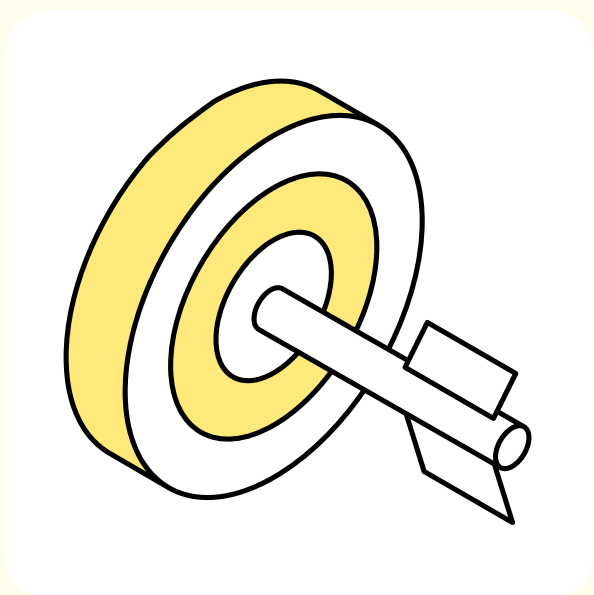
In **North America**, respondents cite **reputational damage to the brand** as the biggest impact of AI agent failure at 38%.



In **EMEA**, where GDPR makes privacy-by-design a development mandate, 78% of engineering teams spend **at least half their time on safety controls** — the lowest of any region, but still the majority of their capacity.



What to do with this



AI governance failures in customer communications impact three areas simultaneously: the support queue surges, the brand perception, and the engineering team. But these costs don't all have the same fix. If your engineers are spending close to half of their time on guardrails – as most teams in this research report – ask what kind of guardrails.

The infrastructure failures are solvable at the platform layer, with PII masking, rate limiting, audit trails, and compliance enforcement built natively into it, not as engineering deliverables. And that's the sprint capacity that goes back to the product roadmap. But the model failures need a different fix. Treating them the same way means spending on symptoms while the cause compounds.

CHAPTER 4

The visibility gap that's **dividing** leaders

90% of enterprise leaders describe themselves as confident in their AI agent readiness. Their rollback rate is the same as those less confident. The visibility data in this chapter shows why.

Across every role and every level of seniority, leaders are looking at the same AI challenges and seeing different things. When they don't share the same picture of what's happening, what's failing, and what's at risk, governance failures go unnoticed and investment goes to the wrong place.



The technical and business **divide**

Technical leaders (IT, engineering, security, and data – across all seniority levels) report rollbacks at 78%. Business leaders (finance, marketing, CX, product, and operations) at the same organizations report them at 69%. The gap extends to the nature of the incidents: technical leaders report PII-related issues at 34% while business leaders report them at 27%. **Business leaders are less likely to know a rollback happened, and they're also less likely to know why.**

The same pattern holds for the engineering burden. Technical leaders report their teams spending more than half their time on guardrails at a rate 9 points higher than business leaders (38% vs 29%). The engineering cost is real, but it's not landing equally across the leadership team.

And yet technical leaders are more confident than their business counterparts. Technical leaders describe themselves as very confident in their AI readiness at 54%, compared to 43% of business leaders. Having a line of sight into how the AI program is being built and run – even when it surfaces more failures – appears to positively impact confidence compared to having no technical visibility at all.

78%

of technical leaders have shut down or rolled back a live AI agent

69%

of business leaders have shut down or rolled back a live AI agent

54%

of technical leaders are very confident in their AI readiness

43%

of business leaders are very confident in their AI readiness



The confidence gap within the C-Suite

Across the C-suite, confidence in AI readiness is generally high: 60% of leaders describe themselves as “very confident” in their organization’s AI readiness. But two leaders sit significantly below that. **Only 45% of CPOs and 38% of CISOs define themselves as “very confident”** – and what they have in common is proximity to what failure actually costs. They understand the risk and feel it most when it happens.

For **CPOs**, every governance failure lands directly on the product roadmap. Every rollback is a sprint derailed, a timeline attacked, a revenue opportunity delayed. That exposure produces a different relationship with confidence.

For **CISOs**, it’s something more specific. Customer data exposure is the leading cause of AI rollbacks at 44%, and the CISO is the person who has to explain it when it happens. The closer you are to understanding what goes wrong, the harder it is to be confident it won’t.



Leaders “very confident” their organization is ready to deploy AI agents that autonomously communicate with customers at scale (% by role).



Sinch data (2026) shows only 45% of CPOs and 38% of CISOs define themselves as “very confident.”

Proximity to the consequences lowers confidence

The C-suite's confidence is high. But the C-suite isn't the one rebuilding the agent after a rollback. They're not triaging the support queue when the bot goes down or explaining to a customer why their data surfaced in a conversation it shouldn't have. **The people doing that work assess their AI readiness very differently.**

60% of C-suite leaders describe themselves as very confident in their organization's AI readiness. That number drops to 49% at VP level, and 43% among Directors and Managers. The pattern holds regardless of industry, AI maturity, region, or organization size.

That 17-point gap between the C-suite and the people running the program highlights a difference in exposure. And it runs one way: the further from the consequences, the more confident the assessment.

60%

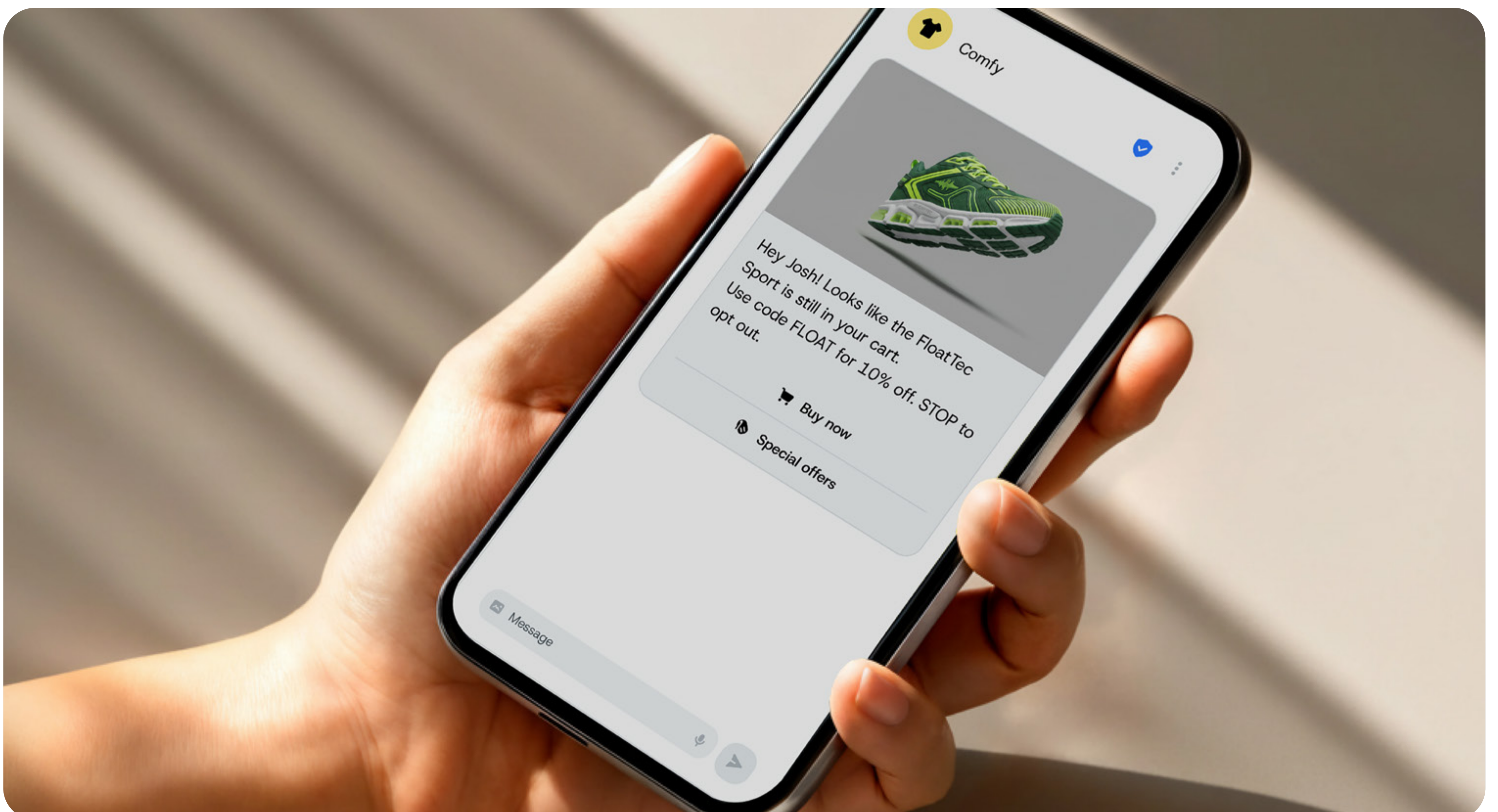
of **C-suite leaders** describe themselves as very confident in their organization's AI readiness..

49%

of **VPs** describe themselves as very confident in their organization's AI readiness.

43%

of **Directors and Managers** describe themselves as very confident in their organization's AI readiness.



VERTICAL SPOTLIGHT

The view from the top looks different across all industries

The more senior the leader, the further they are from what happens when AI fails – the engineering rebuild, the support surge, the customer trust recovery. That distance shapes what they believe about how the program is performing, across every vertical.



In **tech**, 69% of senior leaders (VP and above) describe themselves as very confident in their AI programs, compared to 53% of the Directors and Managers running them. A 16-point confidence gap between the people setting the strategy and the people executing it.



In **retail**, C-suite executives are 2.3x as likely than their VPs and Directors to say most AI communications pilots are reaching production. Same organizations, same programs, and yet completely different accounts of how it's going.



In **healthcare**, 36% of C-suite leaders report fully mature guardrails, compared to 18% of the Directors and Managers building them. The people signing off on governance maturity and the people responsible for it are not looking at the same picture.



The expectations gap is widest on cost. 34% of healthcare C-suite executives anticipate cost reductions of more than 50% from AI. Among the Directors and Managers responsible for delivering those reductions, the number is 12%.

Five different leaders, one AI program

Every leader in this study has a legitimate view of their AI program. But their different pictures are shaping different decisions, different investments, and different definitions of what needs fixing.



IT and engineering leads

CIOs, CTOs, IT and engineering leads have the clearest view of what's actually in production: 65% report active deployments, the highest of any group, likely because they own the infrastructure it runs on. But visibility doesn't mean immunity from failure. These leaders report rollbacks at 76% – 2 points above average.

65%

report AI agents in production. (Sinch, 2026)

76%

report an AI agent rollback. (Sinch, 2026)



Product leaders

CPOs and product leaders invest the most in trust and security at 81% yet report the most rollbacks (82%) and the lowest guardrail maturity at 29%. Only 3% say their current provider isn't falling short. Part of that may be these leaders' definition of failure – broader than any other persona because the product roadmap sits at the intersection of every governance issue, every integration failure, and every sprint derailed by safety infrastructure.

82%

report an AI agent rollback. (Sinch, 2026)

29%

report fully mature guardrails. (Sinch, 2026)



Security leaders

Of all the roles in this study, the CISO and security leaders understand the problem best. That hasn't made it easier to solve, though. They report fully mature guardrails more than any other group at 38% – but that still leaves six in ten not fully there. 41% spend most of their time on safety and still score lower than any other on satisfaction with their provider's security and compliance infrastructure. The people with the clearest view of the problem are the least satisfied with the tools available to solve it.

38%

report fully mature guardrails. (Sinch, 2026)

41%

spend most engineering time on safety. (Sinch, 2026)



Marketing and CX leaders

To deliver AI-powered customer experiences, CMOs, CXOs, and other marketing and CX leads depend entirely on infrastructure they don't own. And that dependency shows: They are the least likely to define themselves as very confident at 45%, and report one of the lowest deployment success rates. For a CMO, the success of an AI deployment is directly tied to its impact on customer experience, which explains why – when an AI agent fails – CSAT damage is their primary concern.

45%

define themselves as very confident. (Sinch, 2026)

58%

report AI agents in production. (Sinch, 2026)



Finance leaders

The CFO organization controls the budget but not the infrastructure. Their primary concern when AI agent fails is brand damage – which, for a financial leader, translates directly to financial exposure. That framing extends to how they think about risk before a failure even happens: At 53%, they are the most worried about data leakage of any group – 7 points above average and 15 points above the CISOs and security leaders responsible for preventing it. And yet finance leaders remain the most confident group in the study at 53% very confident.

53%

cite data leakage as a top AI-related concern. (Sinch, 2026)

53%

define themselves as very confident. (Sinch, 2026)

What the disconnect **costs**

When leaders experience different realities of the same incident, the time it takes to respond to failure increases. It's during that time that irreversible brand damage happens, where incomplete information leads to misdirected investment decisions, and where the next failure gets funded.

The organizations most exposed are the ones where confidence increases the further you get from the failures – where the sprint log says one thing, the board update says another, and no one with budget authority has seen both.

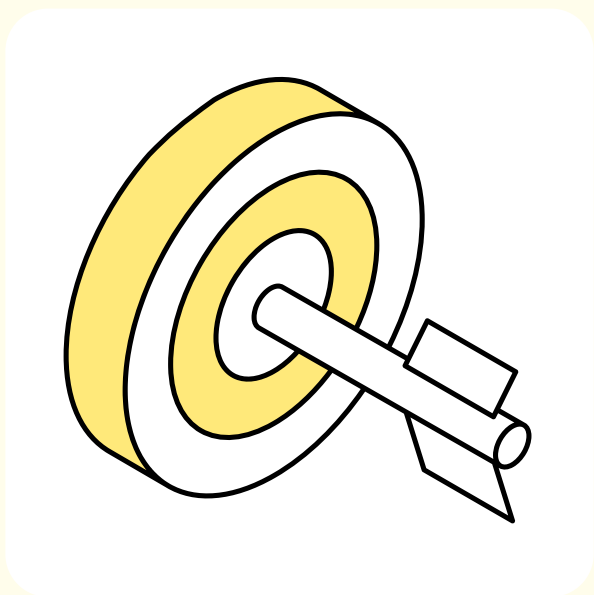


Ultimately, these visibility gaps **impact brand trust**. If the people with the authority to act aren't seeing the same thing as the people closest to it, that gap is exactly where the brand takes the hit.

Robert Gertsman

Chief Evangelist & Co-Founder at Sinch

What to do with this



The visibility gap is not a reporting problem. Adding another dashboard won't solve it. The gap is structural, and it's built into how different roles experience the same AI program: their goals, their involvement, how they measure success differently, and how they define failure.

Two questions worth asking before your next AI deployment: What's the biggest risk in your current program? And where's the biggest opportunity to move faster?

If your CFO and your CIO give different answers, the gap is already shaping what gets fixed and what gets funded.

CHAPTER 5

Investing in trust isn't enough

75% of organizations rank trust, security and compliance as their #1 AI investment priority – ahead of AI development itself – across every industry, region, and persona group in this study. And yet three in four enterprises globally have still had to roll back a live AI agent due to a governance failure.

Enterprises are deploying AI in customer communications to build customer trust, but the data shows it's breaking it instead. The question is no longer whether organizations are taking governance seriously, but why taking it seriously isn't enough to correct the failure rate.



Heavy investment, same failure rate

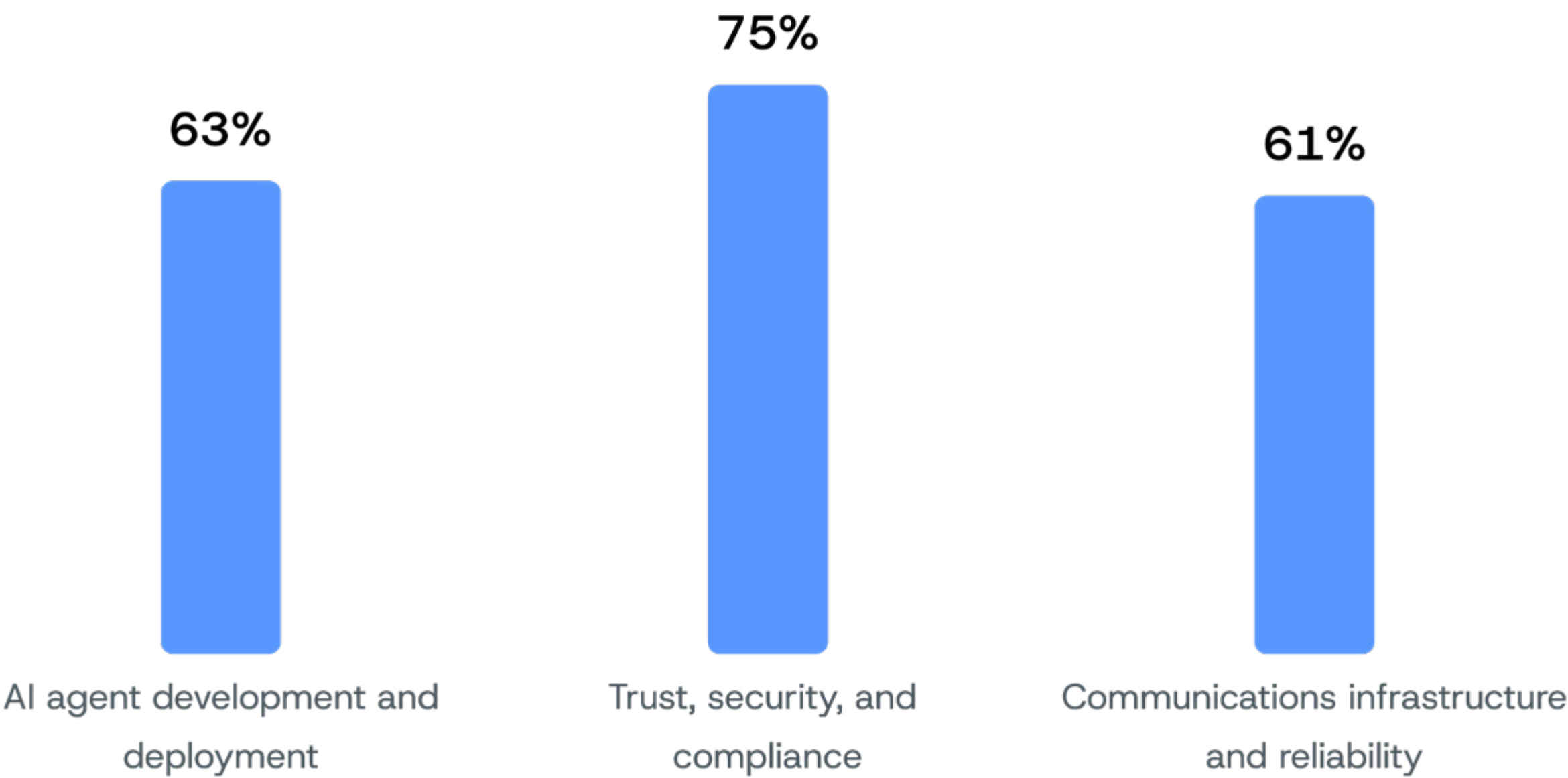
Enterprises invest more in trust, security, and compliance than in AI development. It's the #1 spending category globally, selected by 75% of respondents, ahead of AI development at 63%. Organizations have voted with their budgets: [Governance is the priority](#).

This consensus holds across industries. In tech, trust and security investment outpaces AI development by 10 points – 81% versus 71%. Healthcare (75%), financial services (78%) and retail (75%) also rank it first.

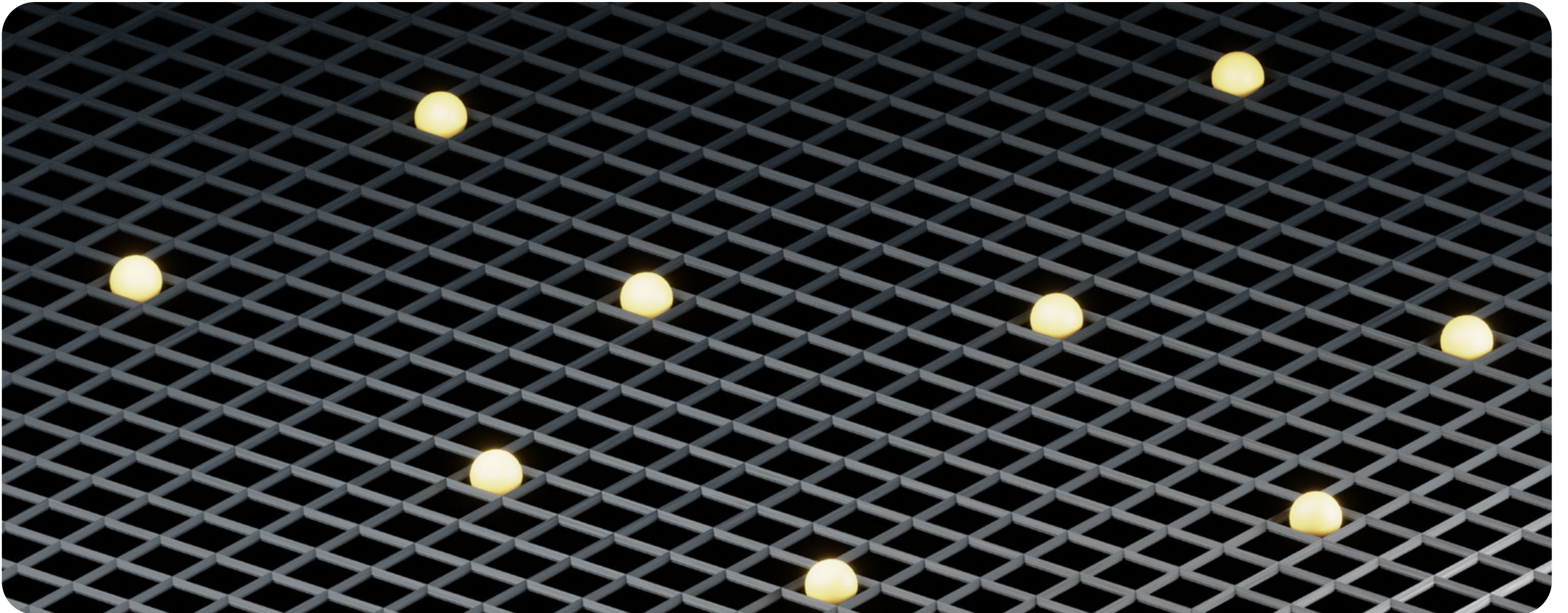
But our data shows this isn't enough. 74% of organizations that reached production have still shut down or rolled back an agent. Among the most governed programs – the ones investing most heavily in safety – the rate is 81%.



Areas that will receive the most investment focus for organizations' AI communication agent deployment in 2026. (Respondents selected up to three priorities)



Sinch data (2026) shows trust, security, and compliance is the #1 spending category in AI programs globally.



The guardrail tax compounds

84% of engineering teams are already spending at least half their time building guardrails and safety controls from scratch. Over a third spend most of their time there.

But the guardrail tax doesn't flatten out as AI programs mature. It actually accelerates. Production-stage engineering teams – those that have already deployed AI agents – spend significantly more time on guardrails than pre-production teams. **38% of production-stage teams spend most of their time on governance, versus 29% of pre-production teams.**

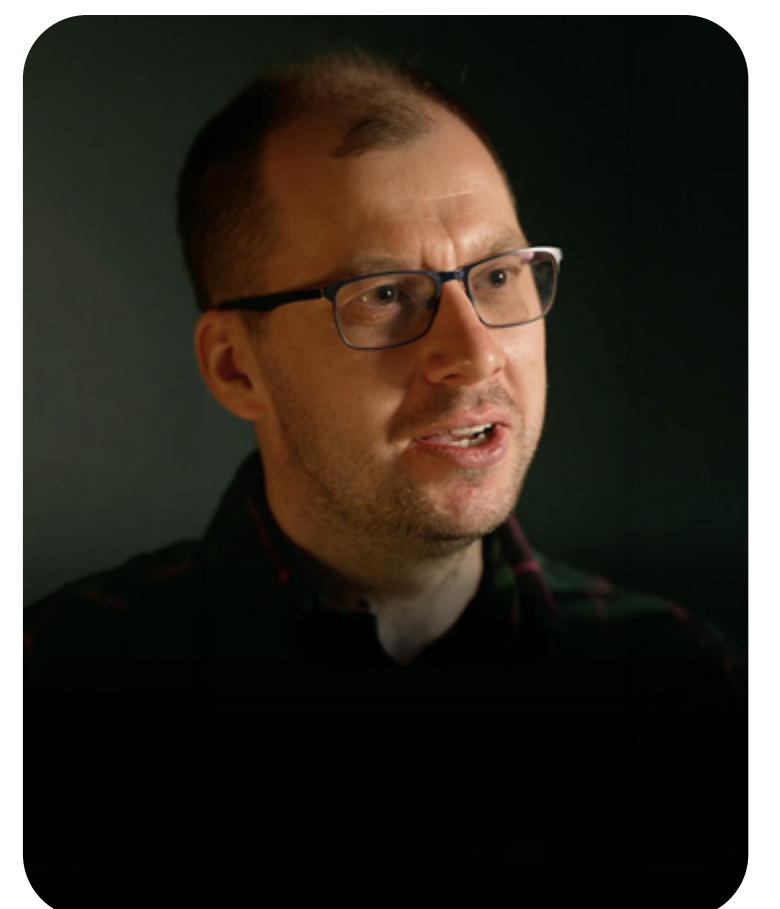
More agents, more channels, more compliance requirements – each one adds another layer that has to be built, maintained throughout the product lifecycle, and rebuilt when something breaks. The organizations furthest along are paying more for safety, indefinitely, for the same failure rate.



The biggest problem is that the cost of building custom guardrails compounds over time, especially as the team goes through the product lifecycle. So, you're **losing that momentum** over working the features that would provide the ultimate value to the customer.

Anton Efimenko

SVP, Software Engineering at Sinch



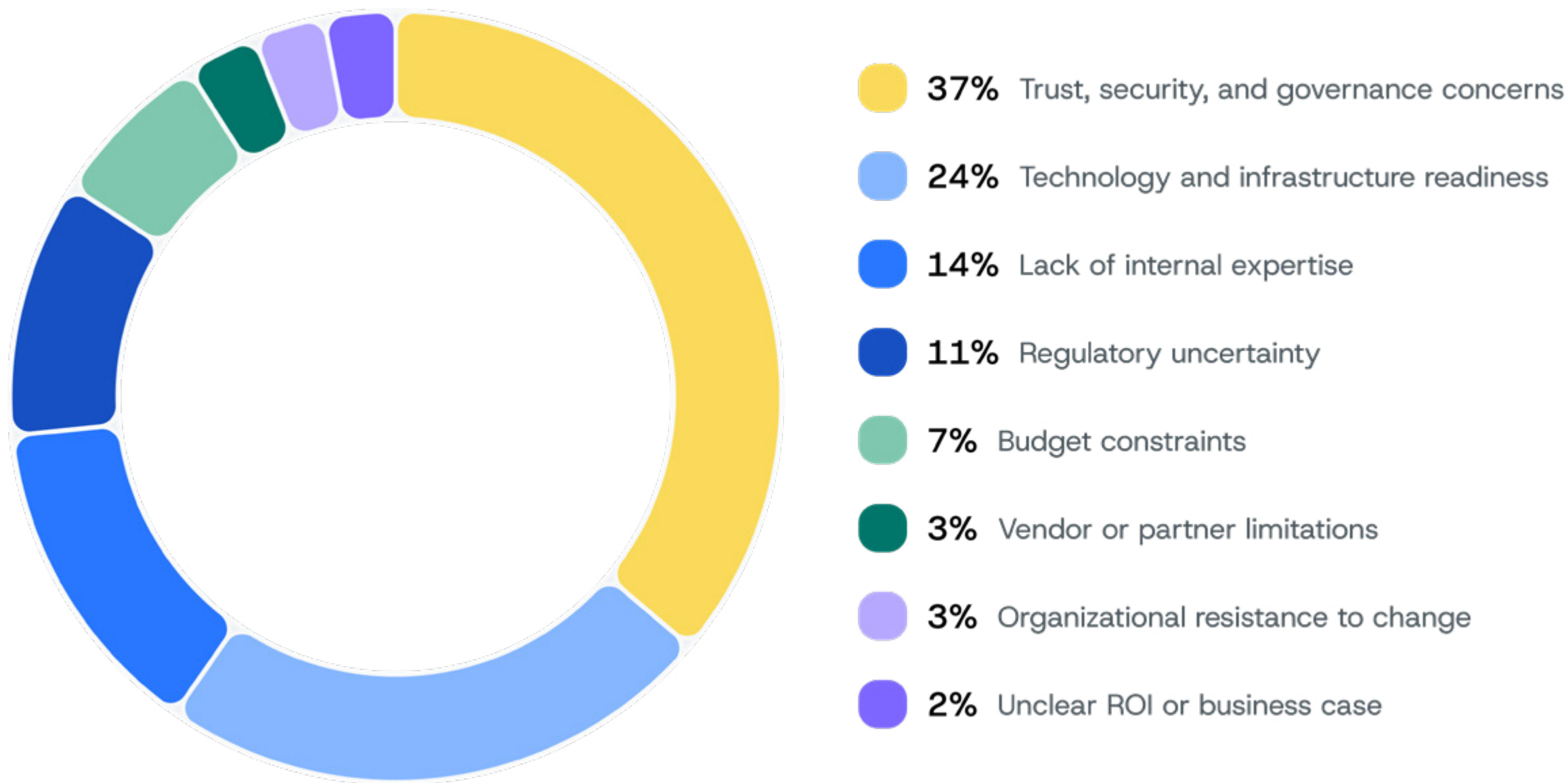
Where governance investment breaks down

Across the C-suite, trust is the one thing every leader agrees on. Every persona in the study puts trust, security, and compliance at the top of their AI investment priorities, ranging from 72% to 86%.

And yet when asked what's actually blocking business impact, the same category comes back as the #1 barrier, cited by 37% of organizations – ahead of every other factor in the study. Budget constraints sit at just 7%. This isn't a resourcing problem.

The category attracting the most investment is also the one blocking the most progress. When the most committed leaders and the most governed programs are failing at the same rate as everyone else, the problem isn't who's leading the program or how seriously they're taking it. Everyone is trying to solve the right problem in the wrong place.

What is the single biggest barrier to achieving the business impact you expect from AI agent-powered customer communications?



Sinch data (2026) shows trust, security, and governance concerns as the #1 barrier to AI business impact.

The customer trust paradox

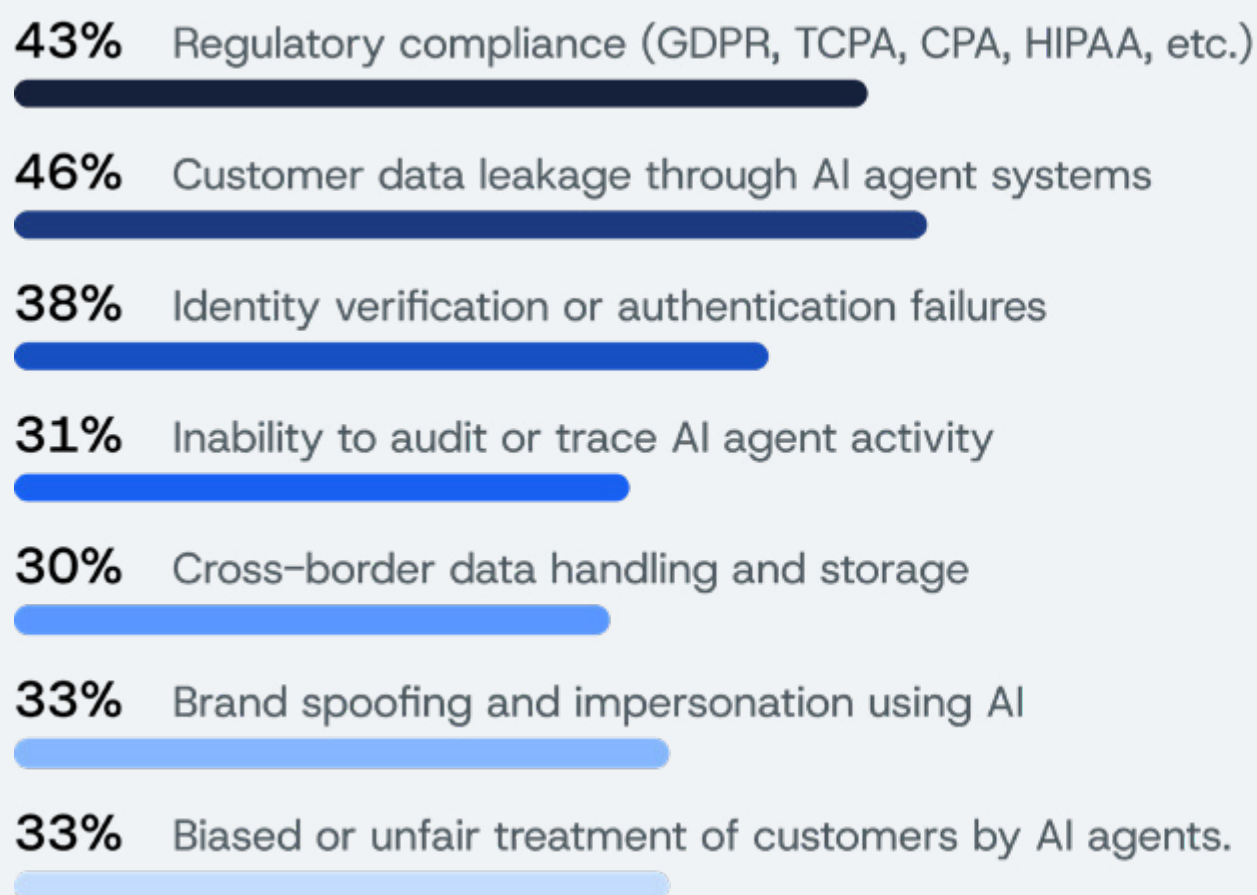
That mismatch has a cost, and it lands where enterprises can least afford it.

The primary goal behind AI communications investment is improving customer satisfaction and loyalty – cited by 36% of respondents, ahead of revenue and cost reduction.

But when AI fails in a customer interaction, the consequences land in exactly the same place. 34% cite reputational damage and loss of customer trust as the primary impact. What's more: **46% of organizations identify customer data leakage as the #1 security risk in their AI programs.** In production, at scale, that exposure doesn't hit one customer. It hits thousands, and the trust it erodes doesn't come back on the next deployment cycle.

When the investment designed to build trust is producing the failures that break it, more budget won't close that gap.

Which security, privacy, or legal risks most concern your organization when deploying AI agents? (Select up to 3)



Sinch data (2026) shows customer data leakage is the #1 security, privacy, and legal risk that most concerns organizations when deploying AI agents.



You've shipped this AI agent into production. Now you have customers interacting with it, they're asking it questions, they're getting responses. A **single mistake** hitting 10,000 customers, a million customers, can erode trust with a massive cohort of your customer base. And once you lose trust, it's much harder to earn it back.

Daniel Morris
CPO at Sinch



THE REGIONAL VIEW

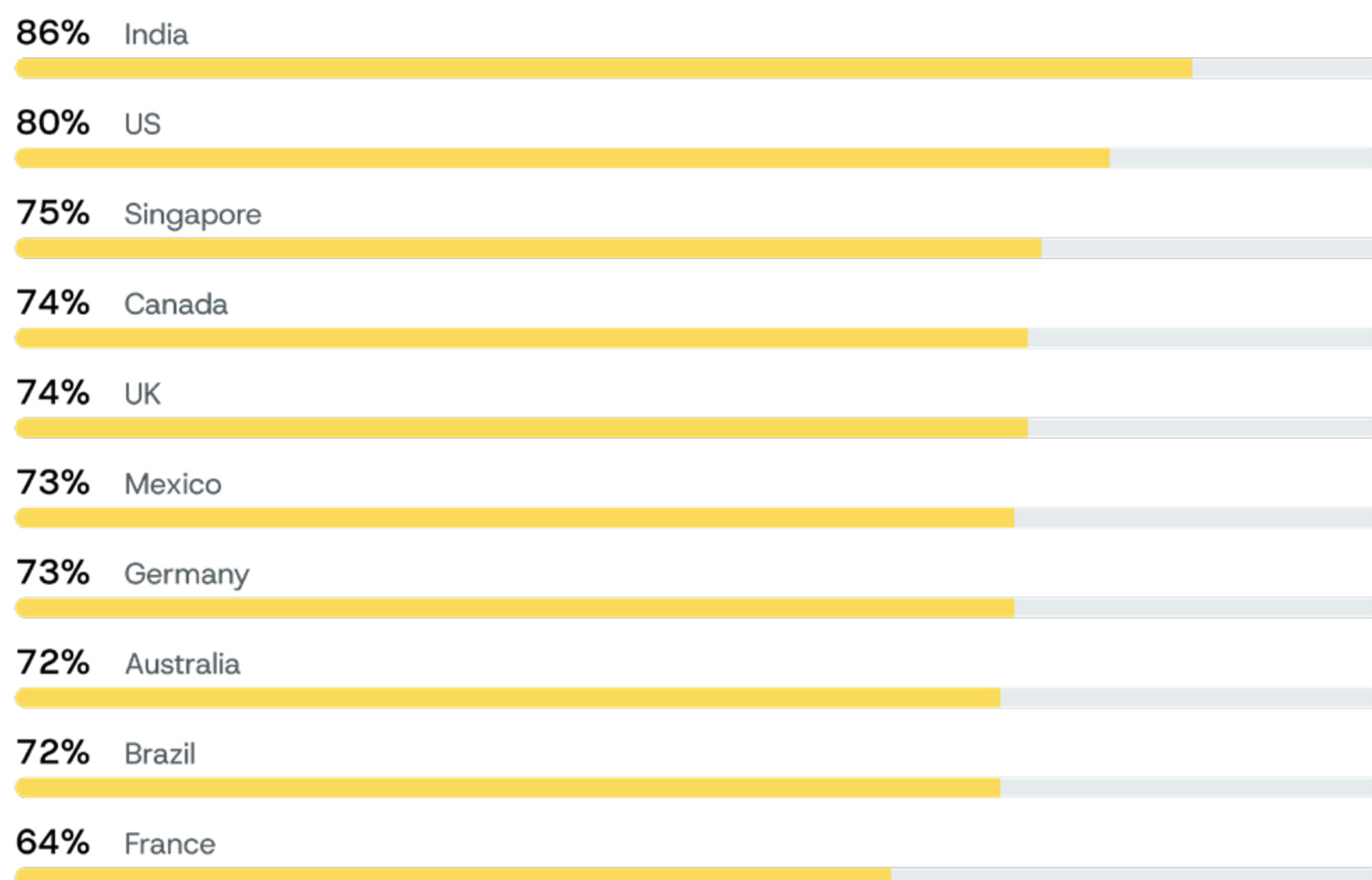
The governance gap is a **global story**

Trust, security and compliance is the #1 investment priority globally. From India at 86% to France at 64%, governance is cited as a main investment area across all countries in the study. **The commitment is universal and consistent.**

So are the failures. Across every region in the study, at least six in ten AI communications programs have experienced a rollback. In some markets it runs as high as nine in ten. No region has cracked the governance problem, regardless of how much they're investing in it.

The global consensus on governance investment is clear. What's less clear is whether it's going where it needs to.

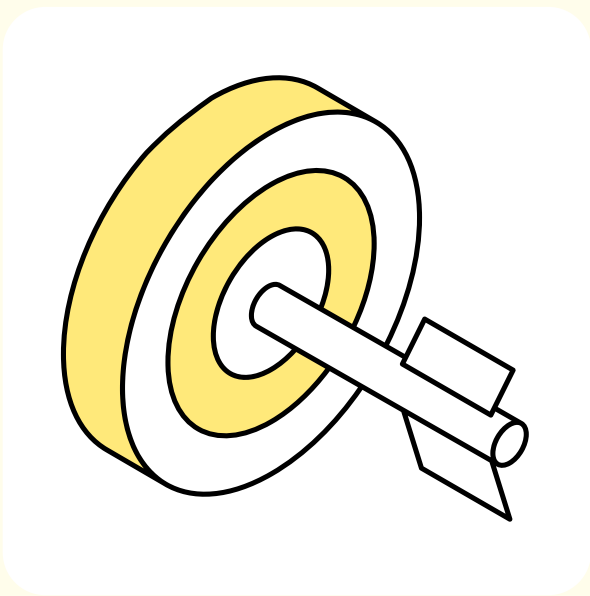
Percentage of organizations that rank trust, security, and compliance as the number #1 investment focus for AI communication agent deployment in 2026 – by country.



Sinch data (2026) shows trust, security, and compliance as the #1 investment focus across countries.



What to do with this



Trust, security and compliance attracts more investment than any other category in AI communications. And yet 74% of organizations have still rolled back a live agent due to a governance failure. The most governed programs fail even more.

Start with your own numbers. Compare your governance investment trajectory against your rollback trajectory. If the spend has gone up and the failure rate hasn't come down, the cost is compounding. And the next budget increase will do the same.

The next chapter looks at why, and what the organizations closing the gap are doing differently.

CHAPTER 6

Infrastructure is where the race gets decided

Communications infrastructure satisfaction is the #1 predictor of AI deployment success. This variable consistently outperforms all others across every statistical method applied to this dataset (correlations, regression models, and cross-tabulations). Not the budget. Not the AI maturity. Not the guardrail sophistication.

Investment alone isn't enough. In AI customer communications, it all comes down to the infrastructure. And that's where the race will get decided.



The infrastructure gap

The correlation between infrastructure satisfaction and AI deployment confidence is 0.52 – the strongest relationship across 4,656 variable pairs analyzed in this study. Put simply: How an organization feels about its communications infrastructure is a better predictor of AI success than its investment level, guardrail maturity, or how long it has been deploying AI.

Yet despite 87% rating high-performance communications infrastructure as essential or very important, most organizations say their current provider falls short in at least one meaningful area. 42% report insufficient reliability, 37% cite limited multi-channel capabilities, and 32% mention lack of AI platform integrations.

The gap between what enterprises need and what they currently have is where most of the findings in this report originate.

87%

of enterprises rate high-performance communications infrastructure as essential or very important to their AI strategy.

42%

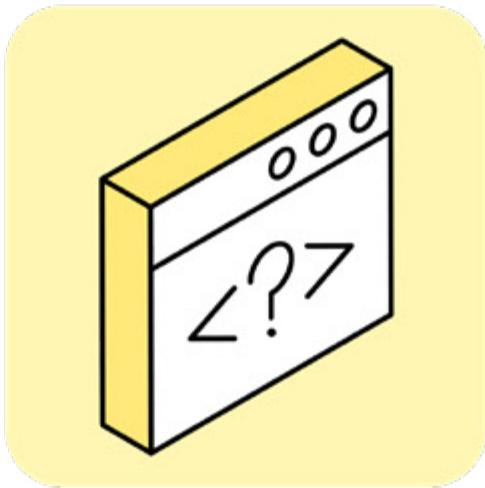
of respondents report insufficient reliability for AI at scale from their current provider.

37%

of respondents cite limited multi-channel capability from their current provider.

32%

of respondents report a lack of AI platform integrations from their current provider.



The cause lies below the surface

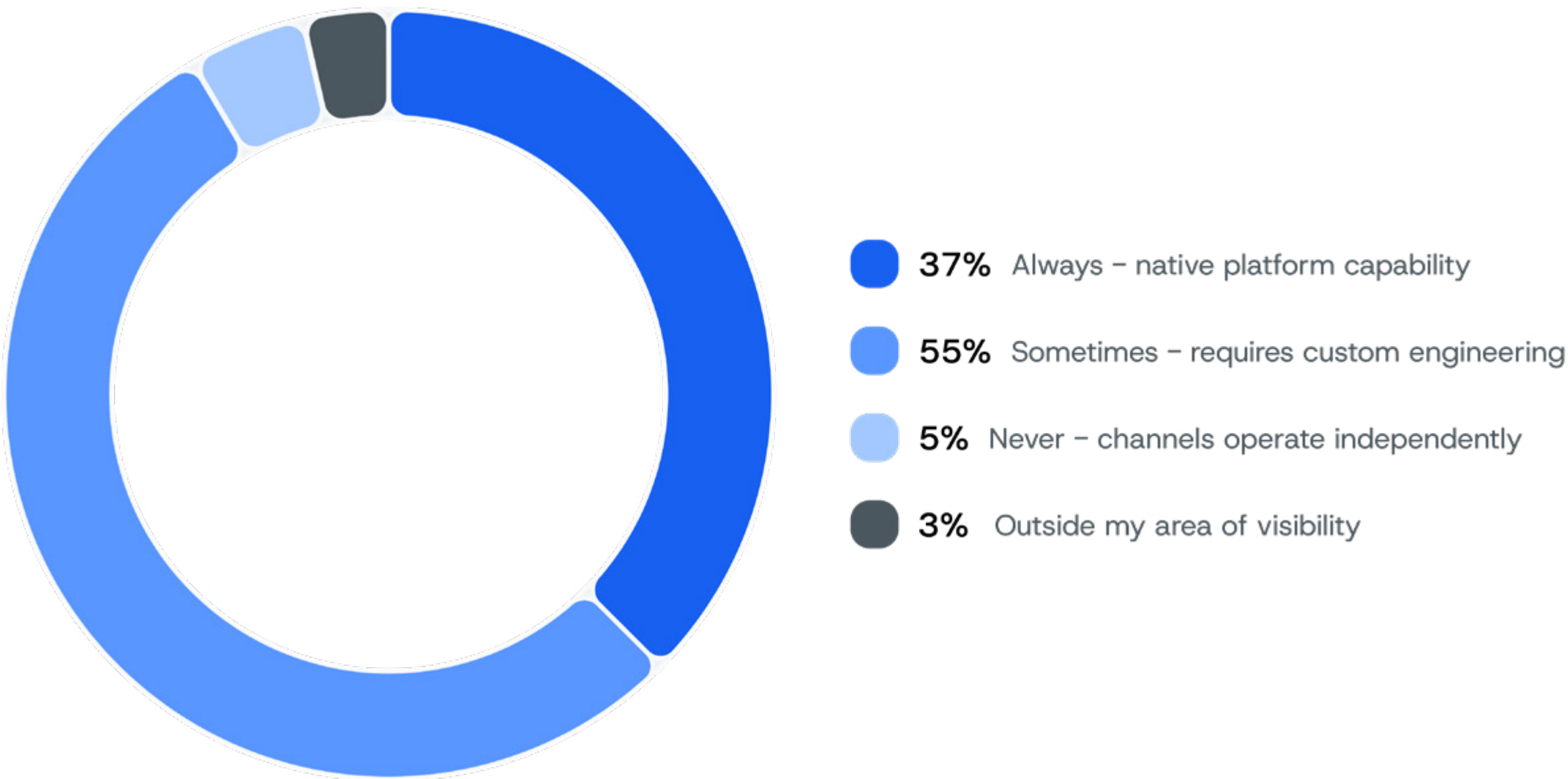
The issue isn't how much is being spent. It's where the money is going. And for most engineering teams, it's going to building safety infrastructure that the communications platform should already provide natively.

For many of those failures, the root cause is one layer below the governance program. Our research shows that **more than half of enterprises (55%) are custom-engineering the ability to preserve customer context** when someone switches channels – from chat to voice, from WhatsApp to a phone call – because their platform doesn't provide it natively.

When a customer has to repeat themselves to an AI agent, they're not experiencing a model failure. They're experiencing the infrastructure gap directly, and it's your brand that pays the price.



When a customer switches channels (e.g., from web chat to SMS), how often is context from previous AI agent conversations preserved without asking them to repeat themselves?



Sinch data (2026) shows 55% of enterprises are custom-engineering context preservation across channels.

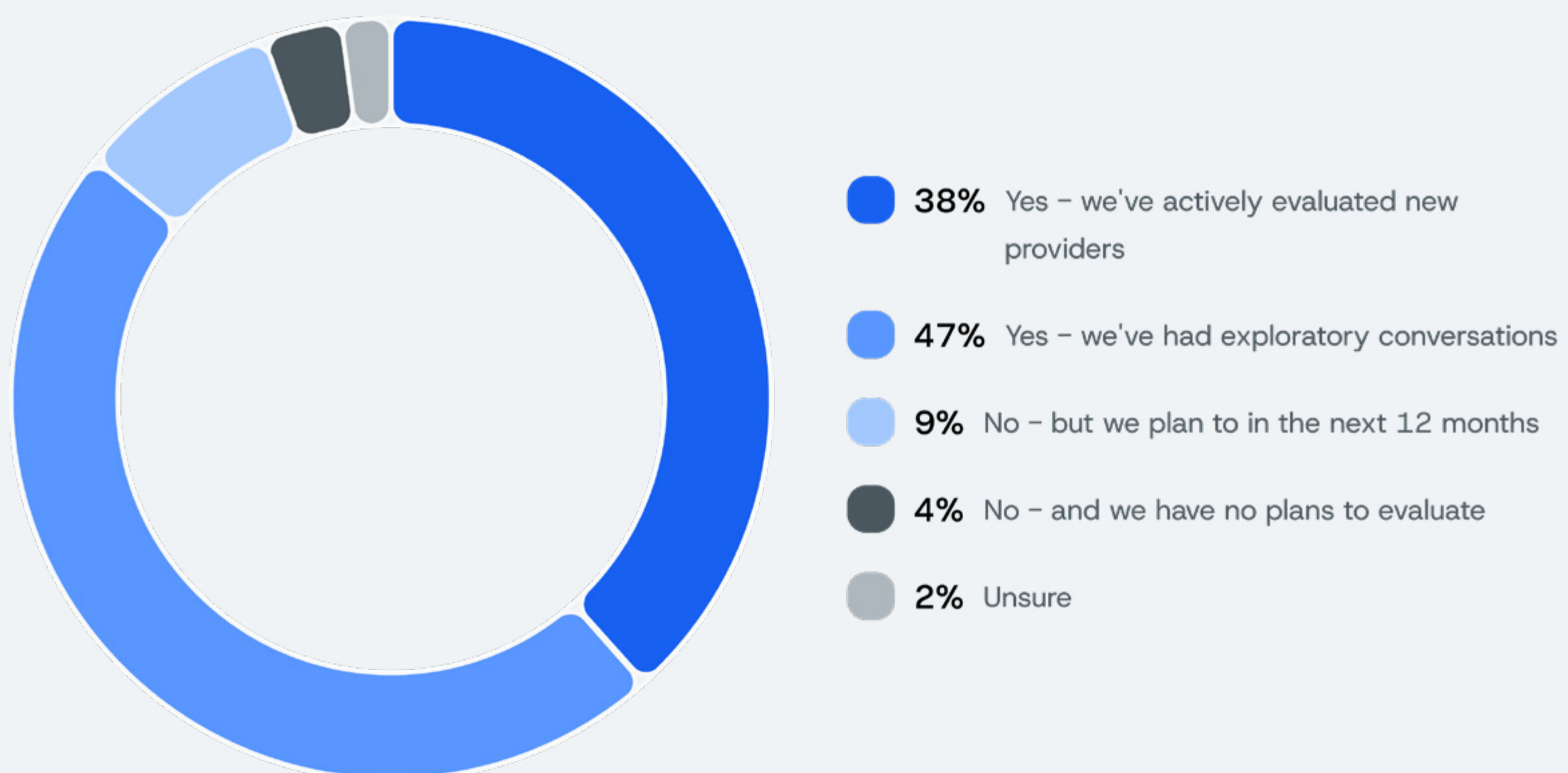
The market is already responding

Enterprises haven't fully articulated that diagnosis yet, but their behavior suggests they've felt it. **86% have had active or exploratory conversations with alternative providers** in the past 12 months, and only 4% have no plans to evaluate.

The two strongest triggers for switching providers confirm where the pain is. **91% of enterprises that have experienced a governance rollback have evaluated** or are actively evaluating a new communications provider. And organizations that rate high-performance infrastructure as a top strategic concern are twice as likely.

The market isn't necessarily shopping because it's unhappy with a vendor. It's shopping because [AI ambitions have outgrown](#) what the current infrastructure was built to handle.

Has your organization evaluated or considered new communications providers in the past 12 months specifically to support AI agent initiatives?



Sinch data (2026) shows 86% of enterprises are having active or exploratory conversations with alternative providers.

VERTICAL SPOTLIGHT

Industries are shopping for **two reasons**

Industries shopping for new providers are driven by either ambition or failure.

In financial services, the triggers are specific and measurable. Enterprises citing lack of AI integrations are 2.4x more likely to evaluate a new provider. Those flagging insufficient reliability are 1.4x more likely. The infrastructure gap has moved from general dissatisfaction to precise switching intent.

But not all evaluation is driven by failure. In retail, the most advanced operators – those with the highest deployment rates, highest confidence, and most mature guardrails – are also the most active shoppers: 57% are actively evaluating providers compared to just 10% of retail brands still in early-stage deployment. The organizations most actively shopping for a new provider are the ones already furthest ahead.

57%

of the most advanced operators in retail are actively evaluating providers.

10%

of retail brands in early-stage deployment are actively evaluating providers.

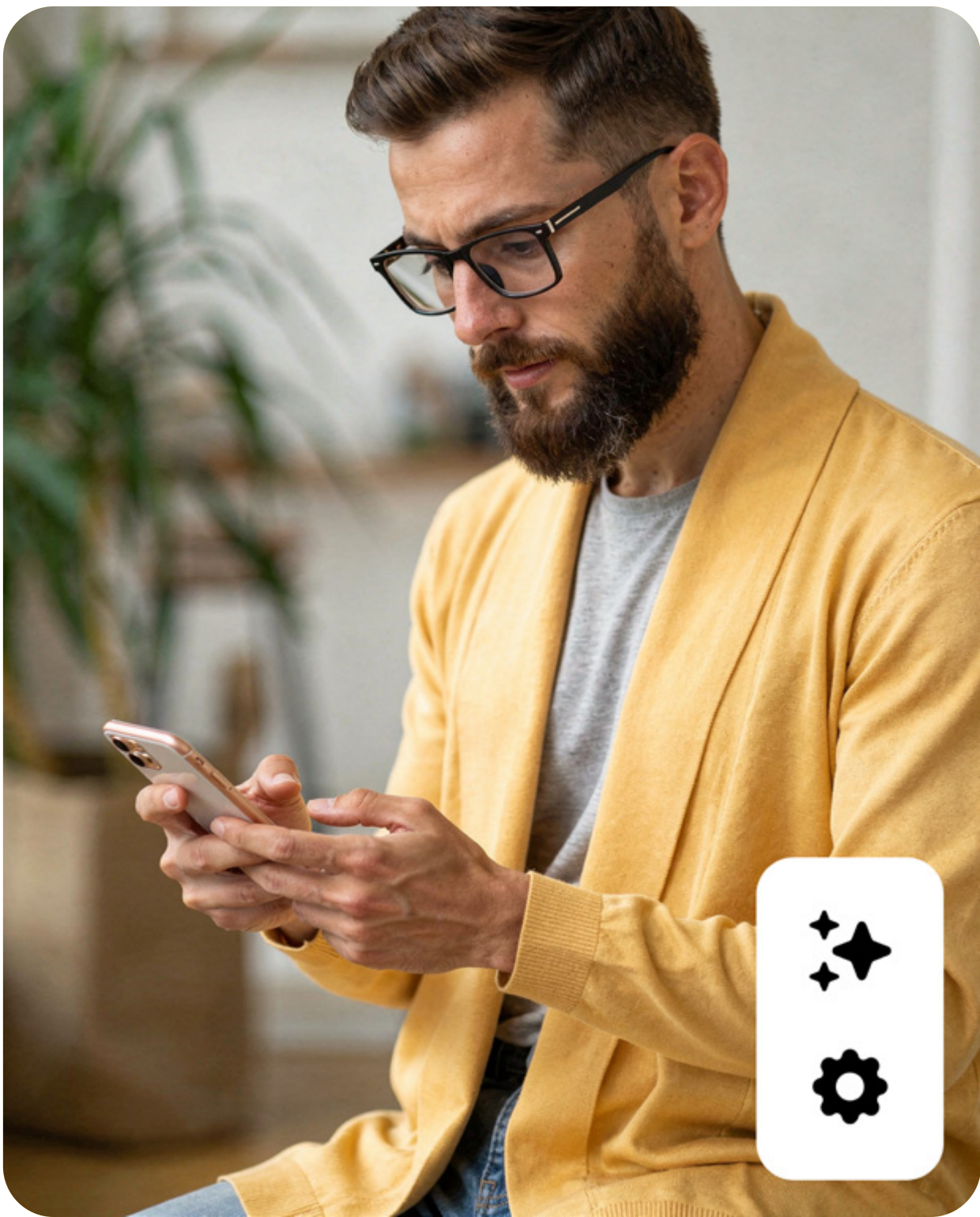
Not all evaluation is driven by failure. In retail, the most advanced operators – those with the highest deployment rates, highest confidence, and most mature guardrails – are also the **most active shoppers.**

What AI leaders are evaluating on

The infrastructure gap feels different depending on where you sit. Security leaders are the least satisfied with their security and compliance infrastructure, despite being closest to it. Marketing and CX leaders cite the most reliability shortcomings at 46%, and only 3% of product leaders – the most demanding buyers in the study – say their provider is not falling short.

Globally, though, the priorities are clear. When asked to rank criteria for evaluating new communications providers to power agentic interactions, reliability ranked first: 29% of respondents placed it top, ahead of global reach and compliance capability (23%), and ease of integration (18%).

Security and trust and multichannel capabilities close the top five, ahead of pricing. In fact, pricing ranks eight out of the nine factors evaluated in the survey.



When evaluating communications providers for AI agent-powered interactions, which of the following criteria is the most important?



Sinch data (2026) reveals reliability and uptime is the most important factor for 29% of respondents.

The structural cost of **building** without a foundation

Good security architecture doesn't ask whether to build controls like PII masking, rate limiting, and audit logging. It asks where. The right answer is both layers: platform-native first, with engineering teams adding controls on top. That's defense-in-depth.

Building without a foundation means engineering teams aren't making security decisions. Instead, they're filling gaps, reactively, on every new agent deployment and every new channel. But when a solid foundation exists, these teams are able to build intentional controls on top of something that holds.

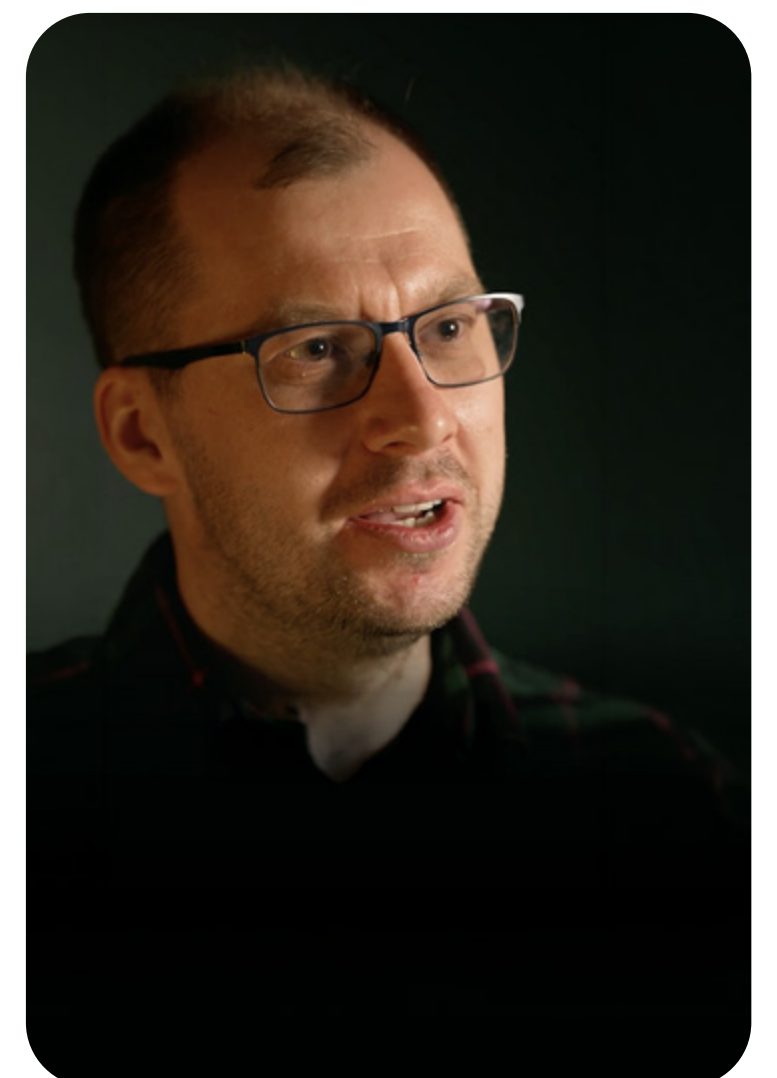
That's where the architectural question becomes a roadmap question. When platform-native controls **exist**, **engineering leaders can see what their team is actually building and assess what adds genuine protection**, and what's just reconstructing infrastructure the platform should already own. That determines whether engineering teams are building toward the next capability or just keeping the lights on.



Every team needs to decide what controls belong at the platform layer and what their engineers should build on top, because the cost of building custom guardrails compounds over time, especially as the team moves through the product lifecycle. Each new agent, each new channel, each new deployment adds to the pile. And eventually you **lose that momentum** when it comes to outperforming on the market.

Anton Efimenko

SVP, Software Engineering at Sinch



THE REGIONAL VIEW

The most satisfied markets are **still shopping**

The infrastructure gap shows up in every region, but differently. North America is the most satisfied region and is also the most active buyer market. EMEA organizations – the least satisfied on every infrastructure dimension – trail on provider evaluation, with the wave of switching activity still ahead of them.



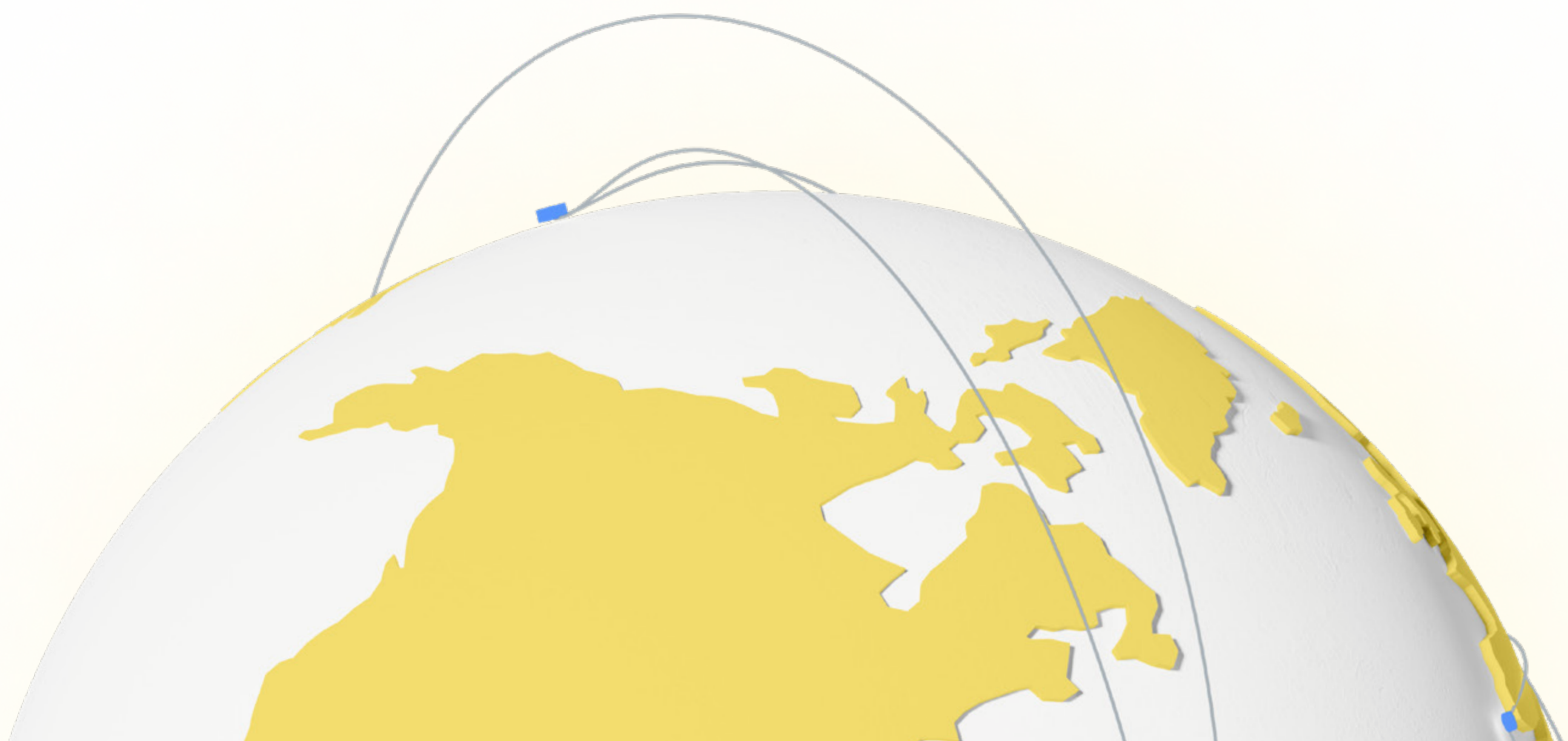
In **North America**, satisfaction with current providers is the highest of any region, and yet 85% have had **active or exploratory conversations** with alternative providers. The bar is high and the market is still moving.



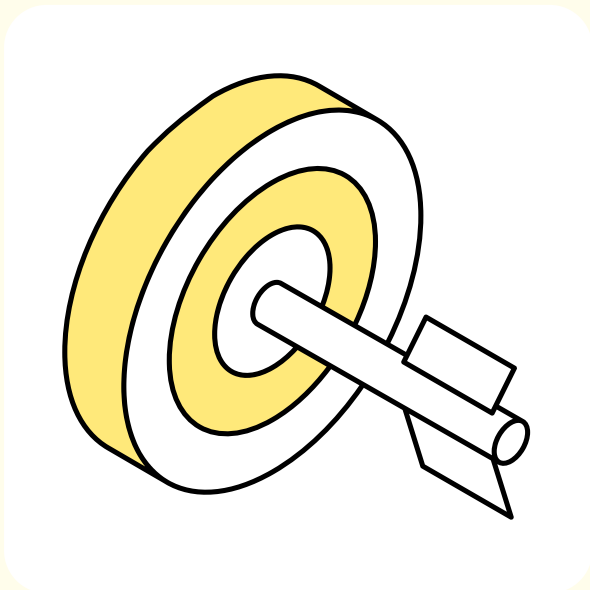
EMEA trails the global average on **every infrastructure dimension** analyzed, with Germany the furthest behind. The region is also 8 points behind the global average on active evaluation but leads on planned evaluation at 14%.



In **APAC**, enterprises whose provider can't integrate with AI platforms are **3x more likely to experience a communications failure**. Half of Indian enterprises are actively evaluating a new provider.



What to do with this



You've seen the data. Now it's time to act on it. Three questions worth taking into your next leadership review:

- How much of your engineering team's time is going toward building guardrails from scratch versus building the next customer experience?
- If those guardrails fail in a live customer interaction, would you know before it becomes a trust problem?
- And is your communications provider built for what you're planning to deploy in the next 12 months, or for what you shipped 18 months ago?

The findings in this report point to where your priorities should sit for the year ahead.

About “The AI Production Paradox” research

The AI Production Paradox is a 2026 industry research report by Sinch, based 2,527 enterprise decision-makers surveyed across 10 countries and 6 industries. All respondents hold Director, VP, or C-suite roles. All work at organizations with 1,000 or more employees. And every respondent is directly responsible for their organization’s AI communications strategy.

This is not a general AI survey. Every question is specifically about how enterprises deploy AI in customer communications: messaging, voice, email, chat, WhatsApp, and other channels.

Methodology

Sinch fielded this industry survey in January 2026. The survey consisted of 2,527 enterprise decision-makers from across six industries: financial services, healthcare, retail, telecommunications, technology, and professional services.

All respondents hold Director, VP, or C-suite roles at organizations with 1,000 or more employees actively working on AI communications strategy. Organizations not deploying or planning to deploy AI in customer communications were screened out.

Business impact figures are self-reported organizational estimates, directional rather than audited. Percentages are rounded to the nearest whole number. As a result, totals may not always sum to 100%. All other findings are based on reported organizational practices and attitudes at the time of survey. Sinch commissioned this research but did not direct the analysis or findings.

Survey breakdown

Country	Organization size	Industry	Job roles
Australia: 10.5% Brazil: 11.3% Canada: 6.7% France: 7.3% Germany: 5.9% India: 6.7% Mexico: 9.9% Singapore: 6.0% United Kingdom: 9.8% United States: 25.9%	1,000 to 4,999 employees: 68.3% 5,000+ employees: 31.7%	Financial Services (FinServ): 20.4% Healthcare: 18.7% Retail: 16.3% Telecommunications: 17.2% Technology: 16.6% Professional Services:10.8%	C-suite executive: 33.1% VP level executive: 27.4% Director/Manager: 39.5%



Sinch's vision is to connect every business with every customer, everywhere in the world. With the industry's most trusted foundation for intelligent customer communications, Sinch powers over 900 million customer interactions annually for more than 200,000 customers across the globe.

Leading global companies, including AI innovators, rely on Sinch to strengthen customer relationships and deliver seamless experiences across messaging, email, and voice. Profitable since its founding in 2008, Sinch generated net sales of USD 3 billion (SEK 28.7 billion) in 2024 and has over 4,000 employees in more than 60 countries, with headquarters in Stockholm.

Sinch is listed on Nasdaq Stockholm (XSTO: SINCH). Visit us at sinch.com.

