palqee | Fulton Bank

The New Standard for Oversight Intelligence in Contact Centers

A practical handbook and blueprint for customer service and contact center leaders.

Case study: *How Fulton Bank achieved an 88% reduction in performance review time while unlocking new insights in its contact center with Palqee Prisma.*

WHITEPAPER | 2026

Table of Contents

Click the sections below to navigate

CHAPTER 1

The Contact Center Oversight Gap 3

1.1 The limits of manual sampling and automated QA 4
1.2 The cost of the gap, in six dimensions 4

CHAPTER 2

What Oversight Intelligence Is 5

2.1 The three capabilities that define the category 6
2.2 How it works in practice 6
2.3 What oversight intelligence is not 7
2.4 The shift in what becomes possible 8

CHAPTER 3

The Business Case for Full-Population Oversight 8

3.1 Call handle time 9
3.2 The FTE cost of performance review and coaching 9
3.3 Performance manager efficiency 10
3.4 First contact resolution 10
3.5 Customer churn 11
3.6 Missed revenue capture 11

CHAPTER 4

Fulton Bank Introduces Contact Center Oversight Intelligence with an 88% Reduction in Performance Review Time 12

4.1 The goal 12
4.2 What the pilot found 13
4.3 Service quality at full population scale 14
4.4 The productivity gains 15
4.5 ROI and value delivered 18

CHAPTER 5

How Your Contact Center Can Be Ahead 19

5.1 What a proof-of-value pilot produces 19
5.2 Questions worth answering before you start 20
5.3 From pilot to production 21

About Palqee Prisma 22

palqee

Fulton Bank recently moved from sampling a handful of calls per agent each month to validating every single customer interaction in its contact center. The result: an 88% reduction in the time its performance managers spent on manual review, a system that needed human input on only 5% of cases by its tenth validation run, down from 40% at the start, and distilled insights into service quality and growth opportunities that small-scale sampling had never made visible.

This handbook sets out what made that possible, and what it means for any contact center measured on handle time, coaching cost, first contact resolution, churn, and revenue.

CHAPTER 1

The Contact Center Oversight Gap

The way most organizations monitor customer interactions is no longer enough

Every contact center operating today carries a version of the same quiet liability. Agents are making consequential decisions hundreds or thousands of times a day: how they handle a complaint, whether they surface the right option at the right moment, whether an issue gets resolved on the first call or turns into three more. And the organization's ability to know what opportunities are missed, and whether those interactions are meeting the standard the business has actually set, is, at best, partial.

It is not for lack of effort. Contact centers have tried to close this gap in two ways, and both fall short of what leadership actually needs.

The limit of manual sampling

The oldest approach is manual sampling: a quality analyst listens to a handful of calls per agent each month and scores them by hand. The volume of interactions has grown far beyond what any team of analysts can realistically monitor this way, so coverage stays thin by necessity, not by choice.

> *"Our overall goal is to be able to confidently assess one hundred percent of our phone conversations, to determine the level of quality we are giving to customers on every call. Not just 0.2%."*
>
> — SVP Contact Center Channel, Fulton Bank

The number 0.2% is not an outlier. It reflects the operational reality at many contact centers today. With tens of thousands of customer interactions taking place each month and a finite number of quality analysts available to assess them, coverage rates in the low single digits are common. Managers typically evaluate four to five calls per agent per month. The remainder are invisible.

The limit of automated QA

A newer generation of tools solves the coverage problem directly. Speech analytics or an LLM scores every interaction automatically, so 100% of calls get evaluated. That is real progress over sampling, but it trades one limitation for another: a fixed rubric applied at scale is still just a rubric. It doesn't adapt to how the organization's own standards are actually meant to be interpreted, and it can't reliably tell the difference between an interaction that is genuinely wrong and one that is genuinely ambiguous, it tends to score both the same way.

Call handle time is a clear example. A generic rubric optimizes in one direction: shorter is better. It has no way to recognize when the opposite is true, when the right move is to keep a customer on the line a little longer to deepen the relationship or surface something they didn't call about, or to spend two extra minutes making sure an issue is actually resolved so the customer doesn't call back tomorrow. It sees minutes, not judgment. No automated QA tool available today can distinguish a call that ran long because the agent did the right thing from one that ran long because the agent was inefficient, so it optimizes every agent toward cutting every call short, even the ones where that is the wrong call for the business.

The result is full coverage of the wrong thing: a score that a manager still has to double-check before trusting it, generated at a scale nobody can manually verify.

The consequences of both limitations extend across every metric a contact center is measured on.

From a service quality standpoint

Organizations that begin systematically reviewing their interactions at scale, whether they started from sampling or from a generic automated score, frequently discover that the inconsistency they attributed to agent performance is at least partly a documentation problem. Procedures that seemed clear on paper turn out to contain rules that different agents have interpreted differently for years. The calibration between managers assessing the same interaction type is often imperfect in ways that go unnoticed precisely because neither approach was built to reveal the pattern.

From a growth standpoint

This is where the opportunity is frequently underestimated. The same interactions that carry service risk also carry intelligence about customers: service quality, the friction points that drive repeat calls, churn, and missed opportunities to serve the customer better. An organization relying on a thin sample, or on a generic automated score it doesn't fully trust, is operating largely blind to a continuous, high-fidelity signal about the experience it is actually delivering.

The cost of the gap, in six dimensions

Contact center leaders are measured on a consistent set of metrics: how efficiently agents handle calls, how much it costs to keep quality high, how well managers spend their time, how often issues resolve on the first contact,

whether customers stay, and how much revenue is captured versus missed. Each is usually treated as its own initiative, with its own dashboard and its own owner. In practice, all six share a single root cause: manager time and attention spread across a small sample instead of the full population of interactions.

~0.2%	4–5	100%
Typical interaction monitoring coverage at contact centers today	Calls reviewed per agent per month in a standard QA program	Coverage that leading organizations can now achieve

The oversight gap is a cost and performance problem. And it is one the industry has lacked the instrumentation to solve, until now.

This handbook sets out the practical case for a new standard: full-population oversight intelligence for contact centers, learning continuously from the organization's own people, and turning every interaction into a source of evidence, insight, and improvement. It is written for the customer service and contact center leaders who own these outcomes, and who are closest to both the cost of the gap and the value of closing it.

CHAPTER 2

What Oversight Intelligence Is

And why the distinction from traditional quality assurance (QA) matters

Chapter 1 described two ways contact centers have tried to close the oversight gap: manual sampling, which never reaches full coverage, and automated QA, which reaches full coverage but scores every call against a fixed, generic rubric. Oversight intelligence is a third approach, one that solves both problems at once.

It is the application of high-precision instrumentation to the full population of contact center interactions, not a sample, with the ability to learn continuously from the organization's own people and standards, rather than a rubric that never changes.

That combination matters for two reasons. First, coverage: reviewing every interaction, rather than a handful per agent per month, surfaces the low-volume cases a sample would never catch, including the rare ones that carry the greatest risk to the customer relationship. Second, the feedback loop: as the organization's experts resolve edge cases and refine guidance, the system applies that judgment consistently across the full population going forward, so average performance rises and variance narrows over time.

The distinction is worth making precisely because the term "AI-powered QA" is now used freely to describe tools that do little more than automate the transcription and scoring of interactions using generic LLMs.

Oversight intelligence, as a category, is defined by the capacity to detect growth opportunities and reduce risk across the full population of interactions, with the

accuracy and evidentiary standard that customer-facing operations actually require.

The three capabilities that define the category

Oversight intelligence, properly understood, rests on three capabilities that generic analytics tools do not provide.

High-precision detection	Institutional learning	Evidential output
The ability to identify opportunities, policy gaps, ambiguity, and process failures with the nuance that customer interactions demand: contextual reasoning against the organization's own procedures, not keyword matching.	The ability to learn from the organization's own people, incorporating manager feedback to continuously refine how standards are interpreted and applied, reducing ambiguity over time.	The ability to generate structured, auditable evidence for every validation: a reasoned conclusion with supporting citations from the interaction itself.

These three capabilities are interdependent. Detection without institutional learning produces a system that flags the same ambiguities indefinitely, without improving. Institutional learning without evidential output produces conclusions that cannot be defended to leadership. Evidential output without precision produces documentation of the wrong things. The category is defined by all three working together.

How it works in practice

At its core, oversight intelligence works by connecting the organization's existing policy documentation, procedural guides, brand standards, and scripts to the interactions and records that those standards govern.

Every interaction is validated against the relevant standard for that call type. Where the evidence clearly supports a pass or fail conclusion, the system records its findings with supporting citations. Where context is ambiguous, the interaction is unclear, or the right interpretation is genuinely uncertain, the system flags the case for human review rather than forcing a conclusion.

That routing to human review is the mechanism through which the system learns. When a subject matter expert reviews a flagged case and provides their assessment, the system creates a refined interpretation of how that standard applies in that context, and applies it to future interactions with high precision.

Over time, the volume of cases requiring human review declines as the system internalizes the organization's own interpretive standards, becoming the subject matter expert that runs continuously and covers the full population at a fraction of the time and cost. What begins as a new tool becomes, progressively, a precise reflection of how the organization's best reviewers actually think.

> *"Prisma created 49 internal controls on the back of manager feedback across a single metric. Those controls now tell us what we should include in the next iteration of our policies."*
>
> — Observed during Fulton Bank engagement

This learning loop makes documentation gaps visible. When the system flags a case as ambiguous, it is identifying a place where the organization's guidance is insufficient, where two managers applying the same procedure might reach different conclusions, and where an auditor examining the same interaction would find no clear standard to assess it against.

The system, in effect, audits the standards as well as the interactions.

What oversight intelligence is not

It is not sentiment analysis. Sentiment tools score how a customer felt during an interaction, frustrated, satisfied, neutral, based on tone and word choice. That's a useful signal, but it says nothing about whether the agent actually followed the right procedure or handled the interaction the way the organization's own standards require. A call can score positively on sentiment while still containing a missed disclosure or a coaching opportunity that never surfaces. Oversight intelligence validates the interaction against the organization's own standards, not the customer's emotional register, which is what makes its findings actionable and auditable rather than just descriptive.

It is not a complete replacement for human judgment. Prisma reliably makes the decisions that are unambiguous, the interactions where the correct outcome is clear against the organization's own standards, and reduces the volume of cases that require a person's attention to the ones where judgment is actually needed: the interactions where context is unclear, guidance is open to interpretation, or the stakes warrant a second opinion. Judgment calls remain with people. What changes is the scale at which that judgment is applied and the quality of the evidence they are working from.

It is not only for AI workflows. The oversight gap exists wherever contact center interactions happen at volume, whether they are handled by human agents, AI-assisted processes, or a combination of both. Oversight intelligence applies equally to human-led and AI-led interactions, and to the mix of both that characterizes most modern contact centers.

It is not limited to a single channel. The same instrumentation that validates phone interactions can validate chat transcripts, email threads, and other written channels, giving contact centers a single, consistent standard across every way customers reach them.

TRADITIONAL QA	AUTOMATED QA
○ Sample-based coverage (under 5%)	○ 100% of interactions scored
○ Retrospective, not real-time	○ Near real-time
○ Inconsistent calibration between reviewers	○ Standard applied is a fixed, generic rubric
○ Findings anecdotal, not systematic	○ Ambiguous cases scored as violations
○ Growth opportunities invisible	○ Findings systematic but generic, not standard-specific
○ Policy gaps invisible until an incident	○ Not built to detect growth opportunities
○ Evidence produced manually, if at all	○ Evidence not fully trusted
○ Manager time spent on routine cases	○ Manager time spent double-checking scores

PRECISION OVERSIGHT INTELLIGENCE

- ● 100% population coverage
- ● Near real-time validation at scale
- ● Standard applied is the organization's own, continuously refined
- ● Ambiguous cases flagged for human review
- ● Findings systematic and grounded in institutional standards
- ● Growth opportunities revealed
- ● Policy ambiguity surfaced automatically
- ● Auditable evidence generated for every validation
- ● Manager attention reserved for genuine edge cases

The shift changes what an organization knows about itself: the quality of its service delivery, the gaps in its standards, and the patterns in its customer interactions.

CHAPTER 3

The Business Case for Full-Population Oversight

The six metrics oversight intelligence moves, that coverage alone doesn't

Contact center leaders are measured on a consistent set of outcomes. Each of the six below is usually owned separately, with its own dashboard and its own initiative. In practice, they share a single limitation, and as Chapter 1 explored, it isn't only a coverage problem. A manager sampling four or five calls a month can't see enough. A generic automated QA tool sees every call but can't judge them against what the organization actually values, so it defaults to the same instruction for every interaction: shorter, faster, more calls handled. Full-population oversight intelligence changes both what is visible and what is

understood well enough to act on, and each of the six dimensions below depends on that combination.

01

Call Handle Time

Handle time is usually managed through average handle time (AHT) targets and after-call-work reduction, blunt instruments applied uniformly across an agent's entire call volume. The problem is that averages don't distinguish between time lost to something fixable, unclear procedures, unnecessary repeat explanations, tool switching, and time that is appropriate because the call was genuinely complex.

A manager sampling four or five calls a month cannot make that distinction at scale. Full-population oversight validates every call against the specific reason it took the time it did, giving managers an evidenced basis to target coaching at the behaviors actually driving handle time up, rather than penalizing agents for complexity that was outside their control.

> **WHAT ORGANIZATIONS TYPICALLY DISCOVER**
>
> Agents penalized for "long" calls that were actually the right call to make, alongside short calls that should have run longer, both invisible to an AHT that only measures minutes, not judgment.

02

The FTE cost of performance review and coaching

The cost of the review function itself is often treated as fixed: a set number of quality analysts, reviewing a set number of calls, at a set number of hours per review. What's rarely visible is how much of that time goes to interactions where the outcome was never genuinely in doubt, a compliant, correct, straightforward call, versus the true edge cases where a reviewer's judgment adds real value.

A generic automated QA tool can reduce that count on paper, since it scores every call. But if a manager still has to double-check the automated score before trusting it, the review burden hasn't actually gone away, it has just moved from listening to calls to auditing a rubric's output. The saving only materializes when the system's judgment is reliable enough that a human doesn't need to verify it case by case, which requires learning the organization's own standards, not applying a fixed one at scale.

Full-population oversight intelligence makes that distinction automatically, and grounds it in the organization's own standards rather than a fixed rubric. It validates the interactions where the answer is clear without manual effort, and reserves reviewer time for the cases that actually require it. The Fulton Bank case study in Chapter 4 shows what that looks like in practice: an 88% reduction in the time performance managers spent on review, with the system's assessments aligning consistently with the judgment of Fulton's own subject matter experts, not a score they had to re-check.

WHAT ORGANIZATIONS TYPICALLY DISCOVER

A large share of review hours going to calls where the outcome was never genuinely in doubt, and, where automated QA is already in place, additional hours spent re-verifying scores nobody fully trusts.

03

Performance Manager Efficiency

Beyond raw review hours, the more durable shift is qualitative. A generic automated score still requires a manager to interpret it, decide whether to trust it, and translate it into a coaching conversation themselves. When a finding is already grounded in the organization's own standards, and comes with the reasoning behind it, that translation step disappears. Managers spend their time coaching from specific, evidenced examples rather than re-litigating a rubric, calibrating consistently across a whole team, and supporting a broader view of performance that can be shared with senior leadership with confidence.

WHAT ORGANIZATIONS TYPICALLY DISCOVER

Coaching conversations built on a manager's general impression of an agent rather than specific, evidenced examples, because the data to support anything more granular was never available at scale.

04

First Contact Resolution

Repeat contacts are one of the most expensive and most poorly diagnosed problems in a contact center. A customer calls back because something wasn't resolved, or wasn't fully explained, but by the time that shows up in a repeat-contact metric, the original interaction that caused it is long gone and was almost certainly never sampled.

Full-population oversight sees every interaction that led to a repeat contact and can trace the pattern back to a specific, correctable agent behavior, rather than a vague trend in a monthly report. A pattern like agents quietly skipping a step that would have prevented a follow-up call becomes visible and fixable in weeks, not discovered by accident months later.

A generic automated QA tool optimizing purely for handle time can actively work against this goal, rewarding agents for cutting calls short even when a few extra minutes would have prevented a callback. Oversight intelligence that understands the organization's own standard for when a call should run long, and when it shouldn't, is what keeps first contact resolution from becoming collateral damage in an efficiency score.

WHAT ORGANIZATIONS TYPICALLY DISCOVER

A small number of specific, correctable agent behaviors, not a vague "training gap," driving a disproportionate share of repeat contacts, sometimes actively reinforced by a handle-time score rewarding the shortcut that causes them.

05

Customer Churn

Churn is a lagging indicator. By the time it shows up in retention numbers, the service experiences that caused it happened weeks or months earlier, at a scale the QA team never had the coverage to see. Full-population oversight turns service quality into a leading indicator instead of a lagging one, but only if what it is tracking reflects what actually predicts churn for this organization's customers, not a generic sentiment score or a fixed checklist.

Recognizing that a certain kind of unresolved friction, or a certain kind of missed reassurance, is the pattern that precedes a cancellation is a judgment call specific to the business. It is exactly the kind of judgment a fixed rubric was never built to make.

WHAT ORGANIZATIONS TYPICALLY DISCOVER

A specific, recurring interaction pattern that reliably precedes cancellation, one that had never been tracked as a leading indicator because it doesn't show up in a generic sentiment or compliance score.

06

Missed Revenue Capture

A contact center is also a channel where opportunities to serve the customer better, resolve an issue proactively, or offer something genuinely relevant, are missed constantly, simply because nobody happened to sample those particular calls.

Recognizing one of those moments, that a call is worth extending to deepen the relationship rather than cutting short to hit a handle-time target, is inherently a judgment call about business context. A fixed rubric cannot produce that finding, no matter how many calls it scores, because it was never built to know what "worth deepening" looks like for this organization. Full-population oversight intelligence surfaces where and how often those opportunities are identified but not acted on, turning every interaction into a measurable data point about growth as well as risk.

WHAT ORGANIZATIONS TYPICALLY DISCOVER

Far more moments where deepening the relationship was the right call than leadership assumed, identifiable only once every interaction, not a 4-5% sample, could be reviewed against what actually matters to the business.

These six metrics are not six separate problems. They share the same underlying limitation: coverage without judgment doesn't move them, and judgment without coverage never reaches enough of them to matter. Contact Centers combining these six dimensions commonly see an estimated immediate impact of $12.2M on average.

CHAPTER 4

Fulton Bank Introduces Contact Center Oversight Intelligence with an 88% Reduction in Performance Review Time

How a regional bank achieved an 8.5-fold improvement in productivity and unlocked new insights through Prisma

Fulton Bank is a regional financial institution serving customers across Pennsylvania, Maryland, New Jersey, Virginia, and Delaware through a growing portfolio of consumer, business banking and commercial banking services. Its contact center handles over 50,000 calls each month, supporting a wide range of enquiries.

As Fulton continues to grow, the bank identified a need to strengthen oversight across customer interactions as a means for accommodating growth while maintaining its commitment to excellent customer service.

To test how oversight intelligence, full coverage combined with judgment grounded in an organization's own standards, performs in a live, high-volume contact center, Fulton selected Palqee as its partner based on Prisma's position as a leading Oversight Intelligence platform for contact centers, providing continuous visibility across all customer interactions and enabling the bank to identify growth opportunities, operational trends, and service delivery challenges that would otherwise remain hidden within sampled reviews.

The challenge

While existing sampled customer interaction reviews provided valuable insight into individual calls, they did not easily reveal broader patterns across the full population, such as emerging service quality patterns or inconsistencies in how guidance was being applied across teams, restricting growth and insight on service effectiveness.

The goal

The primary goal of the pilot was not only to help optimize call handle time but to answer a question that no sampling-based program can answer: what is actually happening on every call, and what does that mean for how we coach, how we grow, and how we serve our customers?

> *"The goal with Prisma is to have a really clear picture on the level of quality we are giving to customers on every call."*
>
> — *SVP Contact Center Channel, Fulton Bank*

The scope covered online banking support calls, assessed for accurate information conveyed and authentication, in addition to other call types and quality metrics relevant to Fulton's broader service objectives.

The ambition behind this scope reflects a challenge that has not fundamentally changed in contact centers for decades. Performance measurement today still operates on the same model it did in the 1980s: a manager listens to a small selection of calls, scores them against a set of criteria, and delivers feedback to the agent. More recently, some organizations have layered automated QA on top of this model, scoring every call against a fixed rubric, but that only solves how many calls get reviewed, not whether the scoring reflects the organization's own standards or can tell a genuine violation from a genuine ambiguity.

Either way, the sample is too small to be statistically representative, or the score is too generic to be fully trusted, the feedback loop is too slow to drive consistent improvement, and the interactions that would reveal the clearest picture of where coaching is needed, or where a call was actually handled exactly right, are never properly heard.

The bank sought to demonstrate how Oversight Intelligence supports growth strategies while maintaining consistent service quality and operational control, with the following scope:

- Evaluate service delivery against Fulton’s operational guidelines and quality standards at scale;
- Identify recurring patterns and provide insights into the underlying reasons behind deviations;
- Surface potential gaps, conflicts, or outdated guidance that could contribute to inconsistent customer experiences and restrict growth;
- Assess Prisma’s ability to support quality assurance functions without requiring additional review personnel;
- Support business objectives for growth and efficiency.

Fulton recognized an opportunity to leverage customer interactions as a source of operational intelligence, enabling them to identify improvement opportunities, strengthen service delivery, and support future growth through more informed decision making.

What the pilot found

The findings fell into three categories: first, what the full population of agent performance data reveals about coaching opportunities and service quality that sampling had never made visible; second, opportunities to update policies and strengthen internal guidance; and third, how dramatically the economics of oversight change with the addition of Prisma.

Authentication

Once the full call population was validated by Prisma, failure rates related to customer authentication presented a materially different picture from what periodic sampling had suggested. In some cases, the guidance covering how agents should proceed in a specific authentication scenario could reasonably be read more than one way, which was inflating the apparent failure rate. Collaborating with Fulton's subject matter experts reviewing and resolving these cases, Prisma incorporated their feedback into its validation engine, establishing a more consistent interpretation of the underlying policies.

Rather than forcing these calls into a simple compliant or non-compliant outcome, Prisma flagged them as ambiguous, a distinction that gave Fulton's team a clear case to review rather than a potentially incorrect result.

As a result, ambiguity rates declined substantially and the true failure rate stabilized within a predictable range. Further it unveiled an important process uncertainty in the existing policies and guidelines.

Online banking calls

Across online banking calls, one of the clearest patterns Prisma surfaced was in how consistently agents communicated available reset channels. Agents handling password reset calls should inform customers of all available reset channels (phone, online self-service, branch), but in 10% of the interactions, agents who were already on a call with the customer consistently completed the reset over the phone without mentioning the other channels.

Agents optimizing for call efficiency were inadvertently reducing customer awareness of self-service options, contributing to repeat call volume and a customer experience gap that Fulton had no previous mechanism to quantify. It's the kind of pattern that is nearly impossible to catch consistently through sampling, but stood out clearly once every call could be reviewed.

> *"This insight on password reset procedure will help us to broaden out the coaching and align service delivery with business objectives."*
>
> *— SVP Contact Center Channel, Fulton Bank*

Service quality at full population scale

The pilot's initial focus was on objective, rules-based scenarios: interactions with a clear right or wrong outcome. Here, full-population coverage already changes

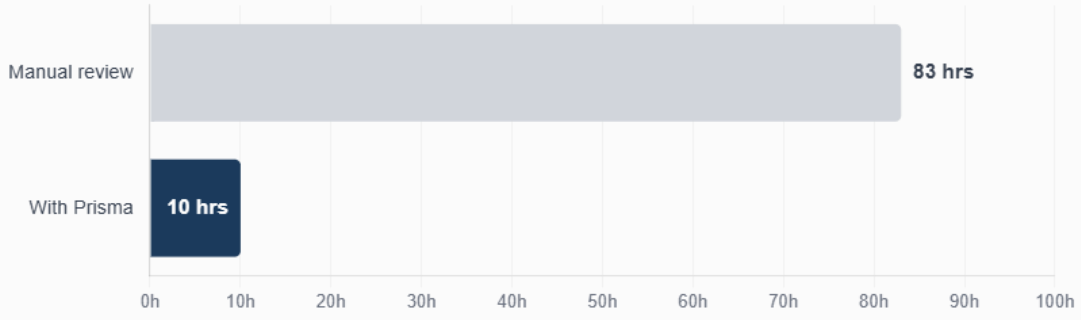
what's possible. Under a sampling model, a manager reviewing four or five calls per agent per month is working with anecdotal evidence — improvement opportunities that exist across an agent's full call volume simply aren't visible at that coverage level. At 100% coverage, those opportunities become measurable, and coaching shifts from impression-based to evidence-based.

The next, harder frontier is subjective service quality: how well an agent listens, whether they convey empathy, whether they match the customer's pace, how effectively they de-escalate a frustrated caller, and whether they go beyond the immediate inquiry to offer relevant guidance on account status or next steps. These qualities matter as much, arguably more, to the customer experience, but they don't reduce to a clear right-or-wrong check.

Extending full-population coverage to these more nuanced, conversational aspects of service is the logical next phase of this work. One that would let Fulton Bank ultimately answer, with confidence, whether it is consistently delivering genuinely expert, empathetic service on every interaction, not just the ones it happened to review.

The productivity gains

Across the full dataset, manual review of the two validation metrics (authentication and accurate information conveyed) would have required approximately 83 hours of performance managers' time. In the Prisma-assisted process, Prisma validated all calls in under 10 minutes, with its assessments aligning consistently with the judgment of Fulton's own subject matter experts. Managers spent approximately a total of 9.5 hours providing feedback to Prisma, rather than reviewing the full dataset themselves.



Incorporating Prisma into the review process has proven to be considerably more efficient, as Prisma's reasoning, feedback requests, and issue explanations are far more precise and explicit than the previous sampling approach could provide. Previously, agents had to listen to entire calls and assess them against multiple criteria.

palqee

"It's the coaching data to drive improvements in performance that is the big headline for me.

The most important element is how many more calls we can base our coaching conversations on.

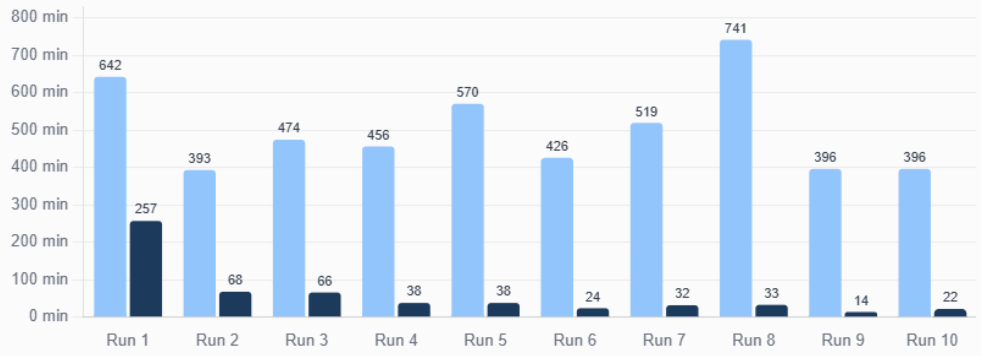
It's absolutely measurable.

Now I can help support a broader senior leadership perspective."

Russell Trezise
Senior Vice President, Contact Center Channel

Fulton Bank

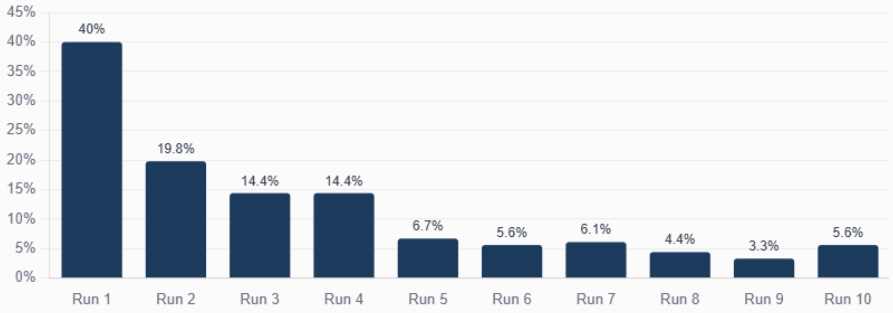
To enable Prisma to quickly learn from the bank's SMEs, the pilot was structured as ten sequential runs of customer calls, with feedback provided between each run. In a production environment, these interactions occur in real time through integrated systems.



The decrease in human review rate is notably as significant as the time spent figure.

In run one, approximately 40% of interactions required the performance manager's input, as Prisma was encountering Fulton's policies for the first time and navigating areas for interpretation.

By run ten, the rate had stabilized at approximately 5%. The system had, over the course of the validation runs, internalized the standards that the manager had been applying case by case, and was now applying them consistently, at scale, without requiring further input.



> *"Prisma created 49 internal controls on the back of manager feedback across a single metric. Those controls now tell us what we should include in the next iteration."*
>
> *— Observed during Fulton Bank engagement*

ROI and value delivered

The Fulton Bank pilot produced three categories of value that may be generalizable to any bank.

First, and most significantly, it has the potential to transform the foundation of performance management and coaching. Prisma's validations aligned with experienced manager judgment across the policy scenarios and call types assessed, and in doing so revealed the greatest opportunities for agent development that sampling had never exposed. The system amplified manager judgment instead of replacing it, applying the standards that the manager had established to a population of interactions that no manager could have reviewed individually.

Immediate financial impact

across reduced call time, avoided FTE costs, performance managers efficiency, first contact resolution rate, reduced customer churn, and missed revenue capture.

What was previously an invisible dataset became a continuous source of intelligence about where agents are performing well, where coaching is needed, and where the institution has the clearest opportunities for growth. Contact centers have operated on the same performance measurement model for decades. This pilot demonstrated what becomes possible when that model changes.

This change was estimated to deliver improvements in call handle time, performance managers' efficiency, first contact resolution rate, reduced customer churn, and missed revenue capture.

Second, it strengthens Fulton's ability to demonstrate consistent, well-evidenced oversight of customer interactions to regulators, reflecting the full population of interactions rather than a sample.

That advantage tends to grow with continued use, as the standards Prisma applies become more refined and consistent over time.

100%

Immediate coverage

with clear evidence, reasoning and justification. Ambiguous records (~5%) should be sampled for human review.

Third, it removed the resourcing objection to full-population oversight. The reduction in manager review time (from 83 hours manually to approximately 10 hours with Prisma across the full dataset) matters not because efficiency is the primary goal, but because it eliminates the assumption that

100% coverage requires additional headcount. It does not.

It requires redirecting the time existing reviewers spend on routine cases toward the genuine edge cases that warrant human judgment, and toward the coaching conversations that the full-population data now makes possible.

8.5x

Increase in productivity

From 83h estimated manager time to review full data set, to under 10h including manager's feedback time.

CHAPTER 5

How Your Contact Center Can Be Ahead

What to expect from the first weeks of full-population oversight

The decision to move from a sampling-based or generic automated QA program to full-population oversight intelligence is not primarily a technology decision, but an organizational one. It requires a clear-eyed view of what the organization is prepared to learn, and a willingness to act on findings that may be more substantive than the sampled, or generically scored, picture has suggested.

A structured proof-of-value pilot is the right vehicle for that decision. It produces real findings from real interactions and generates evidence of what the system can do before any long-term commitment is made. And it produces something that no proof-of-concept based on synthetic data or benchmarks can provide: a specific, organization-level picture of where the oversight gap is, and what closing it would mean.

What a proof-of-value pilot produces

The outputs of a structured Prisma proof-of-value pilot fall into four categories, each of which has standalone value regardless of what the organization decides to do next.

01

PERFORMANCE DATA AT POPULATION SCALE

Hidden opportunities, failure rates, and the ambiguity profile for the call types covered in the pilot, based on the full interaction set rather than a sample. This picture is usually materially different from what periodic review has suggested, and the differences are almost always more actionable than the sampled picture. Many organizations find that the opportunity findings are the highest-value output of the pilot.

02

POLICY GAP INVENTORY

A documented list of the procedural ambiguities and documentation gaps that Prisma flagged during the pilot, organized by frequency and call type. This is, in effect, a standards audit generated as a byproduct of the oversight process.

03

INTERNAL CONTROLS LIBRARY

The set of interpretive controls that Prisma creates on the basis of manager feedback during the pilot. These controls represent the organization's own standards, formalized. They are returned to the organization as an output of the engagement, and can be used directly to improve policy documentation.

04

EFFICIENCY BASELINE

A measurement of the time required to achieve the pilot's coverage level manually versus with Prisma, and the estimated total value recovered across every dimension outlined in Chapter 3. This baseline anchors the business case for a production deployment and removes the resourcing question from the decision.

Questions worth answering before you start

Three questions are worth resolving before beginning a pilot, because the answers shape both what the pilot covers and how the findings are interpreted.

Which queue or interaction type carries the most opportunity?

The pilot should cover the area where the organization has the most to gain from full-population oversight, whether that's the highest-volume queue, the channel generating the most repeat contacts, or the interactions most closely tied to churn and revenue.

Starting with the highest-value area produces findings that are immediately actionable and builds internal confidence in the approach.

Who is the right SME to work with Prisma during the pilot?

The expert's role in the early runs is substantive. They will be making interpretive judgments that will shape how Prisma evaluates interactions. The right person is an experienced reviewer who knows the standards and organizational priorities well and is willing to engage seriously with edge cases. Their involvement time declines rapidly as the pilot progresses, but the quality of their early input determines how fast Prisma will learn.

What will the organization do with the results?

The growth opportunities and standards gaps that a pilot produces are only valuable if the organization is prepared to act on them. The most common pattern is that the findings are handed to the team responsible for the line of

business, and the pilot's controls library is used as the basis for policy update cycles. Organizations that plan this step before the pilot begins move faster from findings to improvement than those that address it after.

From pilot to production

The transition from pilot to production is primarily a question of scope and integration and does not require additional validation. When Prisma completes a pilot, it has already demonstrated its capability to be integrated in the organization's workflow. What production adds is coverage: more queues, more interaction types, more data feeding back into the organization's understanding of its own operations.

The integration path is designed to avoid architectural disruption. Prisma connects via SDK or REST API to the organization's existing infrastructure, transcript sources, case management systems, and analytics platforms, and operates within the organization's own security perimeter. Sensitive data does not leave the organization's environment. Outputs are returned in structured formats that connect directly to existing dashboards and reporting workflows.

The production state that most organizations move toward is one where Prisma handles the full population of routine validations, where the context is clear and the conclusion is not in doubt, and surfaces a small, prioritized set of cases for human feedback. The manager's time shifts from reviewing samples to reviewing the cases that genuinely require judgment.

The organization gains full coverage and unlocks hidden growth opportunities. The manager gains time, and better cases to spend it on.

ABOUT

Palqee Prisma

Palqee Prisma is a High-Precision Instrumentation (HPI) solution for Oversight Intelligence in contact centers. It detects growth opportunities, policy gaps, and process inconsistencies in human-led and AI-led customer interactions, automatically routing high-risk cases to relevant stakeholders and generating the structured evidence that leaders can trust, without a fixed rubric and without requiring a manager to double-check the score.

From contact center conduct and first contact resolution to churn risk and missed revenue capture, Prisma delivers clarity where it matters most: at the point where customer interactions happen, at the scale at which they actually occur.

100%	< 10 min	$12.2M
Oversight coverage: every call and interaction	Time Required by Prisma to validate 50,000 interactions	Estimated immediate financial impact for mid-size organizations

DEPLOYMENT

Prisma deploys on-premises or within the organization's virtual private cloud.

All sensitive data remains within the organization's own security perimeter.

Integration is via SDK or REST API, with structured outputs in JSON or CSV connecting to existing analytics infrastructure - Power BI, Tableau, Looker, or equivalent.

TRUSTED BY

Fulton Bank

and contact center leaders across industries, regulators, policymakers, and technology partners globally.

To discuss a pilot engagement: palqee.com/discovery

info@palqee.com | palqee.com