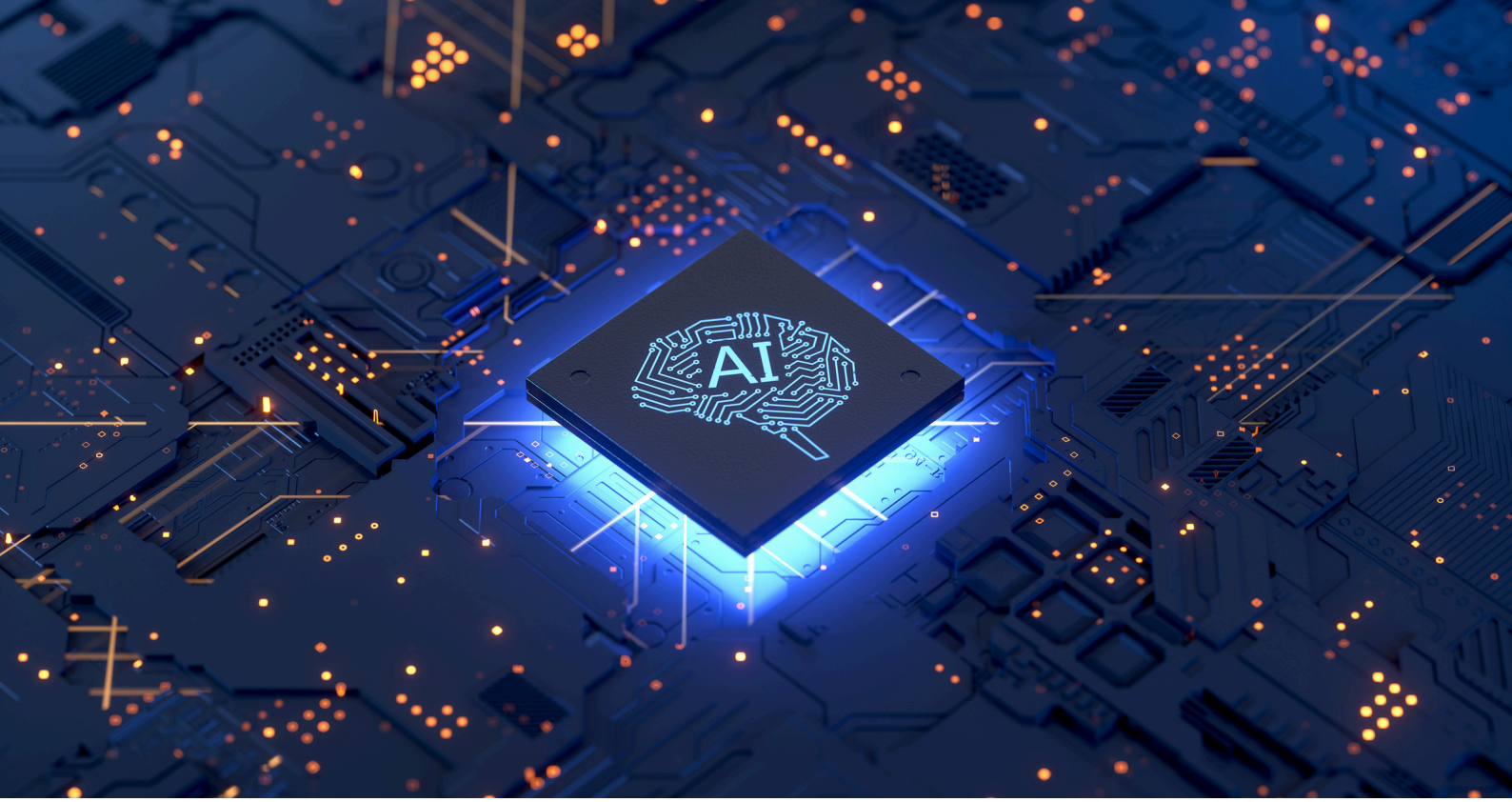
emite
A PROPHECY SOLUTION

A Complete Guide
How to Measure AI Performance in the Contact Centre

Because you can't optimise what you can't measure — and AI is now part of your workforce.



Overview

AI is no longer a future ambition in the contact centre — it is an active participant in daily operations.

It forecasts demand, routes interactions, summarises calls, analyses sentiment, scores quality, assists agents, and automates workflows.

But here's the challenge:
AI is now making decisions traditionally made by humans — yet most organisations have no framework to measure its performance, accuracy, or impact.

This guide gives leaders a complete methodology for measuring AI effectiveness across routing, forecasting, coaching, summarisation, and customer experience — and explains how emite provides the intelligence layer required to govern and optimise AI in real time.

Why AI Measurement Matters

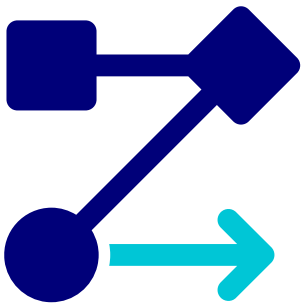
AI is only valuable when it behaves predictably, reliably, and fairly. Without proper measurement:

- AI errors go unnoticed
- Routing decisions become inconsistent
- Forecasts drift out of accuracy
- Summaries misinterpret context
- Sentiment models skew results
- Supervisors lose control over decisions
- CX outcomes become unpredictable
- Compliance exposure increases

Just like agents require performance metrics, AI needs structured oversight. If AI is part of your workforce, it must be measured like one.

The AI Decisions You Must Measure

Modern contact centres rely on multiple AI-driven functions:



1. Routing AI

Predicts intent, assigns skills, adjusts queues, personalises flows.

2. Forecasting AI

Predicts volume, channel mix, patterns, and staffing needs.

3. Agent Assist AI

Provides next-best-actions, knowledge retrieval, recommended replies.

4. Quality & Compliance AI

Scores behaviour, flags risks, identifies gaps, analyses script adherence.

5. Sentiment & Intent Models

Detects emotional tone, predicts churn risk, identifies urgency.

.

6. Summarisation AI

Condenses interactions into CRM notes, case outcomes, follow-up actions.

Each of these decisions must be monitored for accuracy, consistency, and CX impact.

The 12 Metrics That Define AI Performance

1. Accuracy

Did the AI make the correct decision based on the data available?

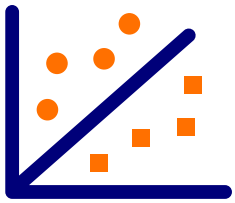
- Correct routing vs. misroutes
- Correct intent classification
- Correct sentiment analysis
- Accurate call summaries

Benchmark: Compare to human decisions and historical patterns.

2. Reliability

Does the model perform consistently over time?

- Accuracy variance
- Frequency of errors
- Consistency across channels
- Consistency per customer segment

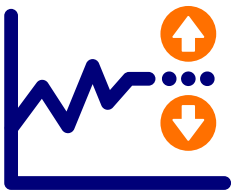


3. Drift Detection (Model Drift)

Are AI decisions becoming less accurate because data patterns changed?

- Routing drift
- Sentiment model drift
- Forecasting variance
- Summarisation deviation

A major risk: AI models degrade silently if not monitored.



4. SLA Contribution

Is AI helping—or hurting—operational outcomes?

Track AI's impact on:

- AHT
- ASA
- FCR
- Transfer rate
- Abandonment

If AI-driven routing or recommendations worsen SLA performance, it must be adjusted.

5. Forecasting Variance

How accurate are AI predictions vs. actuals?

- Intraday variance
- Seasonal variance
- Unexpected spikes
- Channel migration accuracy

Forecasting AI must be continuously recalibrated.

6. Agent Assist Adoption Rate

Are agents using AI-generated suggestions?

- Usage rate
- Decline rate
- Override frequency
- Productivity correlation

Low adoption often signals trust issues or poor relevance.

7. Agent Assist Accuracy

How often do AI suggestions lead to:

- Faster resolution
- Lower escalations
- Higher CSAT
- Reduced handle time

Measure "AI assist success" just like you measure QA scores.

8. Customer Sentiment Impact

Does AI improve or worsen customer emotion across the journey?

- Sentiment before/after interaction
- Sentiment accuracy vs manual QA
- Correlation with churn indicators

9. Bias Detection

Does AI behave differently by:

- Customer profile
- Language
- Demographic
- Channel
- Interaction type

Bias in routing, sentiment or agent scoring creates significant risk.

10. Error-to-Escalation Ratio

How often do AI mistakes require supervisor or human correction?

- Routing errors
- Summarisation inaccuracies
- Misclassified intent
- Incorrect knowledge responses

This determines true AI operational value.

11. Compliance Alignment

Does AI support regulatory requirements?

- PCI redaction accuracy
- PII mask success
- Script adherence alerts
- Risk flagging accuracy

Compliance AI must be measured with 100% scrutiny.

12. Business Value Contribution

The ultimate metric:

- Cost reduction
- Improved CX
- Lower deflection failure
- Reduced repeat contact
- Better first-contact outcomes

If AI doesn't deliver measurable business impact, it's not adding value.

Building an AI Oversight Framework

AI oversight requires structure, not guesswork.
A complete AI governance model includes:

1. A centralised performance dashboard
Cross-system visibility into AI inputs → AI decisions → outcomes.

2. Multi-channel correlation
Connecting voice, digital, CRM, QA and WFM data into one truth.

3. Defined benchmarks for accuracy, variance, and drift
Thresholds for intervention and recalibration.

4. A human-in-the-loop model
Supervisors validate AI on edge cases, exceptions and high-risk outcomes.

5. Continuous model improvement
Ongoing evaluation → retraining → testing → deployment cycles.

6. Executive reporting & risk oversight
AI impact should be visible at operational and board levels.

Why AI Measurement Requires a Unified Intelligence Layer

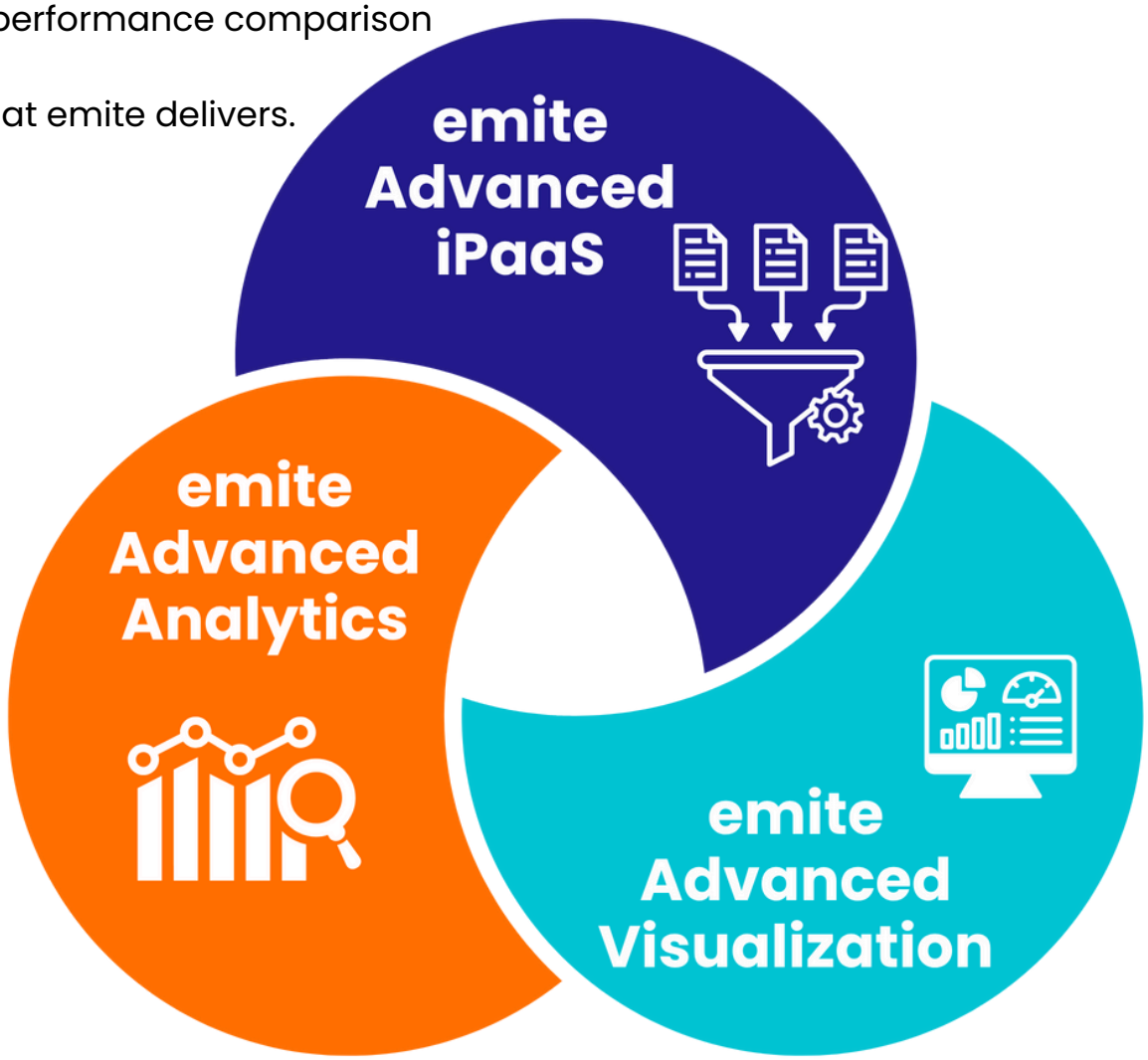
Here's the reality:

You can't measure AI performance if your data is fragmented.
You can't measure AI impact if your journeys aren't connected.
You can't measure AI accuracy if your metrics aren't consistent.

AI performance measurement requires:

- Unified data across all platforms
- Real-time correlation
- Normalised metrics
- Complete customer journey visibility
- Historical context
- AI vs human performance comparison

This is exactly what emite delivers.



How emite Enables AI Governance & Performance Measurement

emite provides the intelligence layer AI depends on:

✓ **Unified, contextualised data across all channels**
Voice, chat, CRM, QA, WFM, sentiment, routing and back office — all connected.

✓ **Real-time data pipelines**
Critical for adaptive AI models.

✓ **Correlated journey timelines**
Understand AI decisions in the context of the full customer experience.

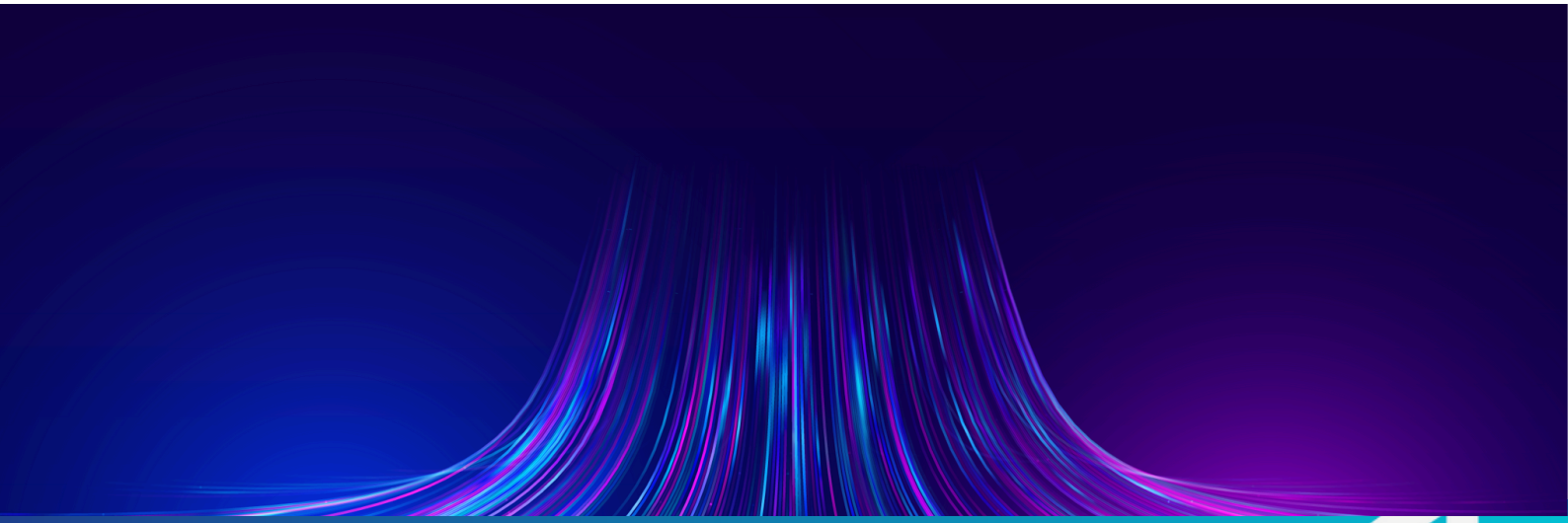
✓ **Metric normalisation**
AHT, SLA, FCR, sentiment — measured consistently across platforms.

✓ **AI vs human outcome comparison**
See the real impact of automation across performance and CX.

✓ **Model drift and anomaly detection**
Identify when AI reliability begins to drop.

✓ **Dashboards for AI governance**
Operational, QA, WFM, and executive-level visibility.

With emite, AI becomes measurable, governable, and trustworthy.



A Final Note

AI is no longer an experiment it is now part of your workforce.

But like any employee, its value depends on oversight, measurement, and continuous improvement.

"One of Gartner's top 10 strategic predictions for 2026 is AI- Driven Decision automation risks catastrophic loss."

The contact centres that win will be those that treat AI performance as seriously as human performance supported by a unified, real-time data foundation.

Explore the emite AI-Ready Data Layer

Book a Demo