AI USAGE & PERFORMANCE ASSESSMENT

Exploratory Analysis and KPI Framework Design

ESFERA AI

February 2026

CONTEXT OF THE COMPANY

This case study analyzes an internal AI assistant implemented in a Latin American enterprise software company.

The organization operates in Spanish and serves regional clients in construction and project management.

All interactions analyzed in this project were originally generated in Spanish.

The objective of this study is to design a scalable analytical monitoring framework applicable to global enterprise environments.

ANALYSIS CONTEXT

The ESFERA AI system generates a growing volume of user interactions but lacked clear metrics to evaluate:

- What the AI is used for
- How well it responds
- How technically stable it is
- How costly its operation is

This analysis aims to transform raw system logs into actionable insights.

PROJECT OBJECTIVES

- Define clear and measurable KPIs for AI usage and performance
- Assess the current system status based on observed log data
- Build an analytical foundation prepared for continuous monitoring
- Design actionable visualizations for technical and product teams

DATA SOURCES AND METHODOLOGICAL APPROACH

- ESFERA AI system logs
- Metric processing and computation in BigQuery
- Exploratory analysis and validation in Kaggle
- Visualization and dashboard design in Tableau

The approach combines exploratory analysis with metric design built to scale over time.

DEFINED KPIs

1. Satisfactory Response Rate
2. Reformulation Rate
3. Average Queries per Session
4. Explicit System Error Rate
5. AI Response Time
6. Cost per Query

These KPIs cover usage, quality, technical stability, and operational efficiency.

ANALYSIS CRITERIA

- **Session:** a set of queries from the same user without inactivity gaps longer than 10 minutes.
- **Real Query:** a user message expressing a clear intent or specific need (greetings, automated replies, and trivial messages are excluded).
- **Internal Data:** the analysis includes internal team usage and test logs, considered part of real system behavior during this early stage.

DASHBOARD 1: SYSTEM STATUS AND USAGE

This dashboard provides a consolidated, high-level view of the current state of the AI:

- What it is primarily used for
- How response quality behaves
- How technically stable the AI is
- What costs are associated with its usage

It is designed as the **main dashboard for overall monitoring**.

KEY INSIGHTS

- The AI is primarily used for **finance, reporting, and project progress tracking**.
- The **satisfactory response rate** is **moderate to low**, with clear room for improvement.
- The **reformulation rate is high**, suggesting friction in the user experience.
- Explicit technical errors are **low (<1%)**, indicating strong technical stability.
- The **average cost per query is currently controlled**, with potential for further optimization.

DASHBOARD 2: TECHNICAL EXPLORATION

This dashboard presents a demonstrative visualization of AI response time (**TTFT vs. total duration**), designed as an example of how technical performance metrics could be monitored over time.

It is not considered statistically conclusive due to the current low data volume, but it establishes the necessary structure for future technical metrics and potential exploratory analyses.

ANALYTICAL FOUNDATION AND SYSTEM MONITORING

- **Defined and validated KPIs**

Clear metrics were established to evaluate system usage, response quality, and technical performance.

- **Dashboards designed for continuous monitoring**

The visualizations enable tracking the evolution of key indicators over time.

- **Technical monitoring with alert potential**

Technical metrics allow expected ranges to be defined and deviations to be detected.

- **Analytically measurable and scalable system**

The constructed structure provides a solid foundation for continuous monitoring and operational control.

IDENTIFIED IMPROVEMENT OPPORTUNITIES

- **Explicit user feedback**

Currently, there is no direct signal that allows us to understand how users perceive the quality of each response.

Incorporating simple feedback mechanisms would complement the existing KPIs and help validate the analytical findings.

Examples:
- Simple reactions such as 👍 / 👎
- Short evaluation messages after each response
- Occasional questions about usefulness or clarity of the interaction

IDENTIFIED IMPROVEMENT OPPORTUNITIES

- **High reformulation rate as a friction signal**

The observed reformulation rate indicates that, in many cases, users need to adjust their query in order to obtain the expected response.

This suggests opportunities to assist users during query formulation and reduce friction in the interaction.

Examples:

- Automatic reformulation suggestions when a query is unclear
- Short clarification prompts such as “Did you mean…?”
- Improvements in onboarding guidance or example queries

CONCLUSION

This project transforms system logs into a clear and actionable analytical foundation, enabling a deeper understanding of ESFERA IA’s usage, quality, and performance.

Enabling:

- Continuous system improvement
- Informed product decision-making
- Sustained technical monitoring over time