

Suporte ao Cliente Multi-Nível: Arquitetura de Agentes Inteligentes

Estudo de Caso: Automação e Empatia no Digital Bank

DISCIPLINA

Arquitetura de Agentes

DOCENTE

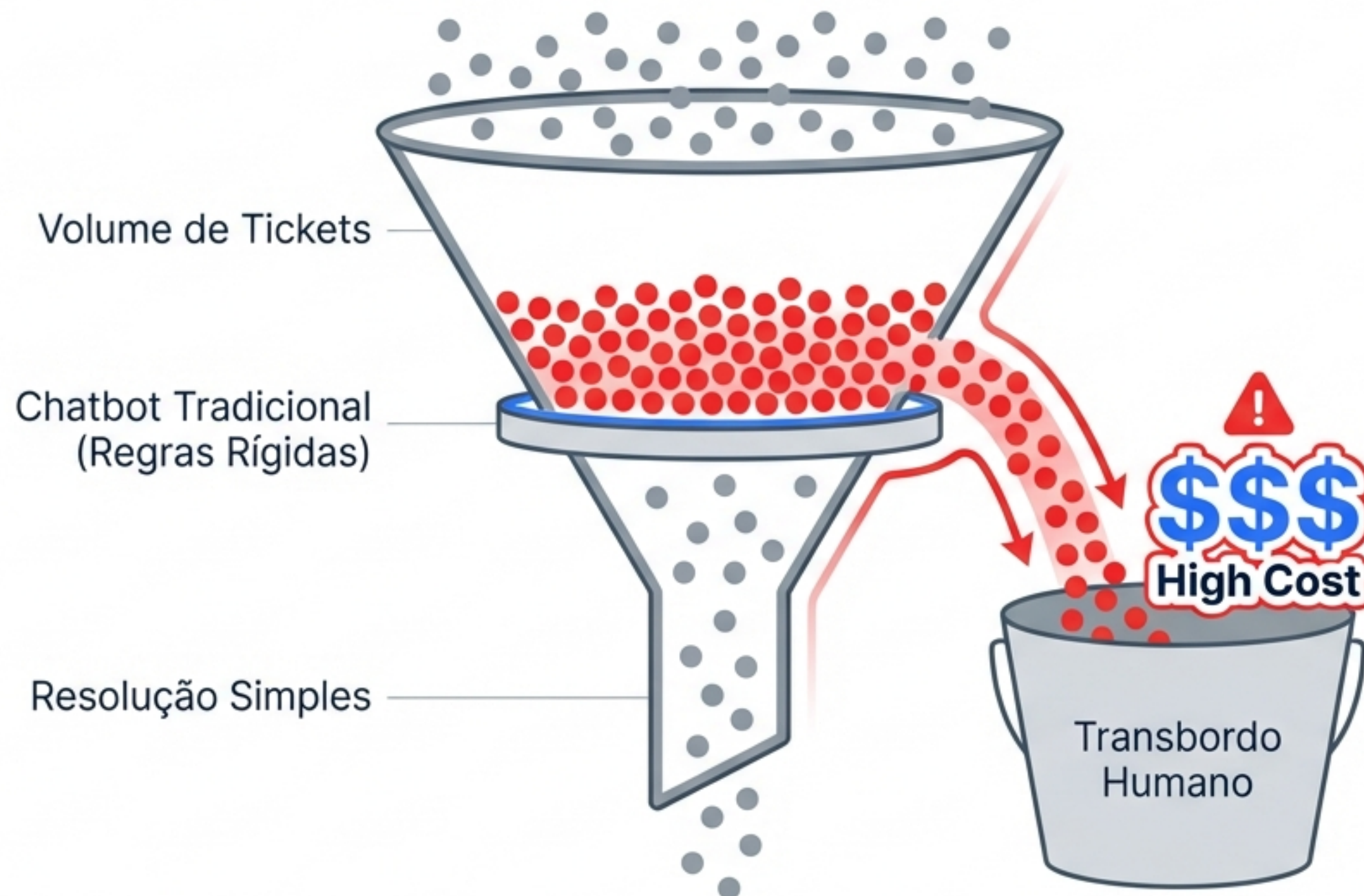
Prof. Carlos Alberto
Stefano Filho

INTEGRANTES

- Luiz Ihara (368484)
- Filipe Fragoso (368500)

- Pedro Martins (368543)
- Daniel Caminha (366740)

Chatbots tradicionais criam barreiras de frustração e geram altos custos de transbordo



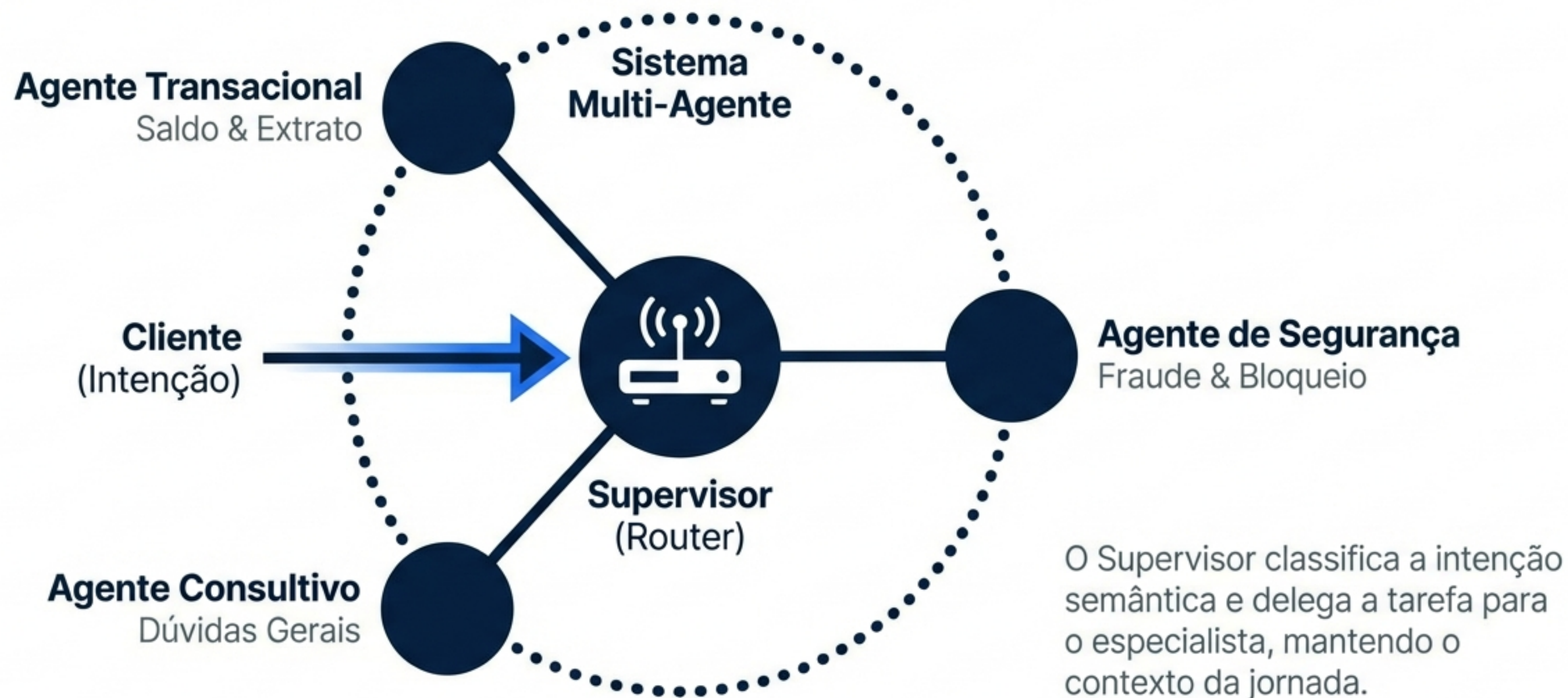
O Problema

Falha em contextos complexos como disputas de fraude.

A Consequência

Ciclo vicioso de frustração do cliente → Transbordo para humano → Alto custo operacional.

A evolução da arquitetura: De um robô solitário para uma equipe de especialistas



O padrão Supervisor/Router garante que a intenção correta acione o fluxo correto



Arquitetura Semântica: Substituição de menus de navegação rígidos por compreensão de linguagem natural.

A hierarquia permite prompts especializados para mitigar riscos e alucinações

AGENTE DE SALDO



Profile: Rápido & Direto

Task: Leitura simples de Database

Risk: Baixo

AGENTE DE FRAUDE



Profile: Cauteloso & Investigativo

Task: Raciocínio complexo, coleta de evidências

Risk: Alto (Requer Prompt de Sistema restritivo)

Takeaway: A modularidade aumenta a segurança ao limitar o escopo de cada agente.

Capacidade de raciocínio para resolver ambiguidades onde scripts rígidos falham



“Meu cartão sumiu, acho que deixei no Uber, mas talvez esteja em casa.”



LLM Reasoning Engine

- > **Inference:** Incerteza sobre roubo vs. perda.
- > **Risk Assessment:** Moderado.
- > **Context Memory:** Cliente viajou para o exterior semana passada.
- > **Decision:** Sugerir Bloqueio Temporário (não cancelamento).
- > **Tool Use:** Consultar API de Geolocalização [ACTIVE].



Entendido. Como há chance de estar em casa, sugiro um bloqueio temporário no app enquanto você procura. Deseja que eu faça isso?

Inteligência Contextual: O sistema avalia nuances para oferecer a solução mais segura e conveniente, evitando atritos desnecessários.

KPIs de impacto direto na eficiência operacional e satisfação do cliente

↑ +25%

FCR (First Contact Resolution)

Resolvido sem intervenção humana

↓ -40%

TMA (Tempo Médio)

Processamento instantâneo de dados

> 60

CSAT / NPS

Qualidade percebida da interação

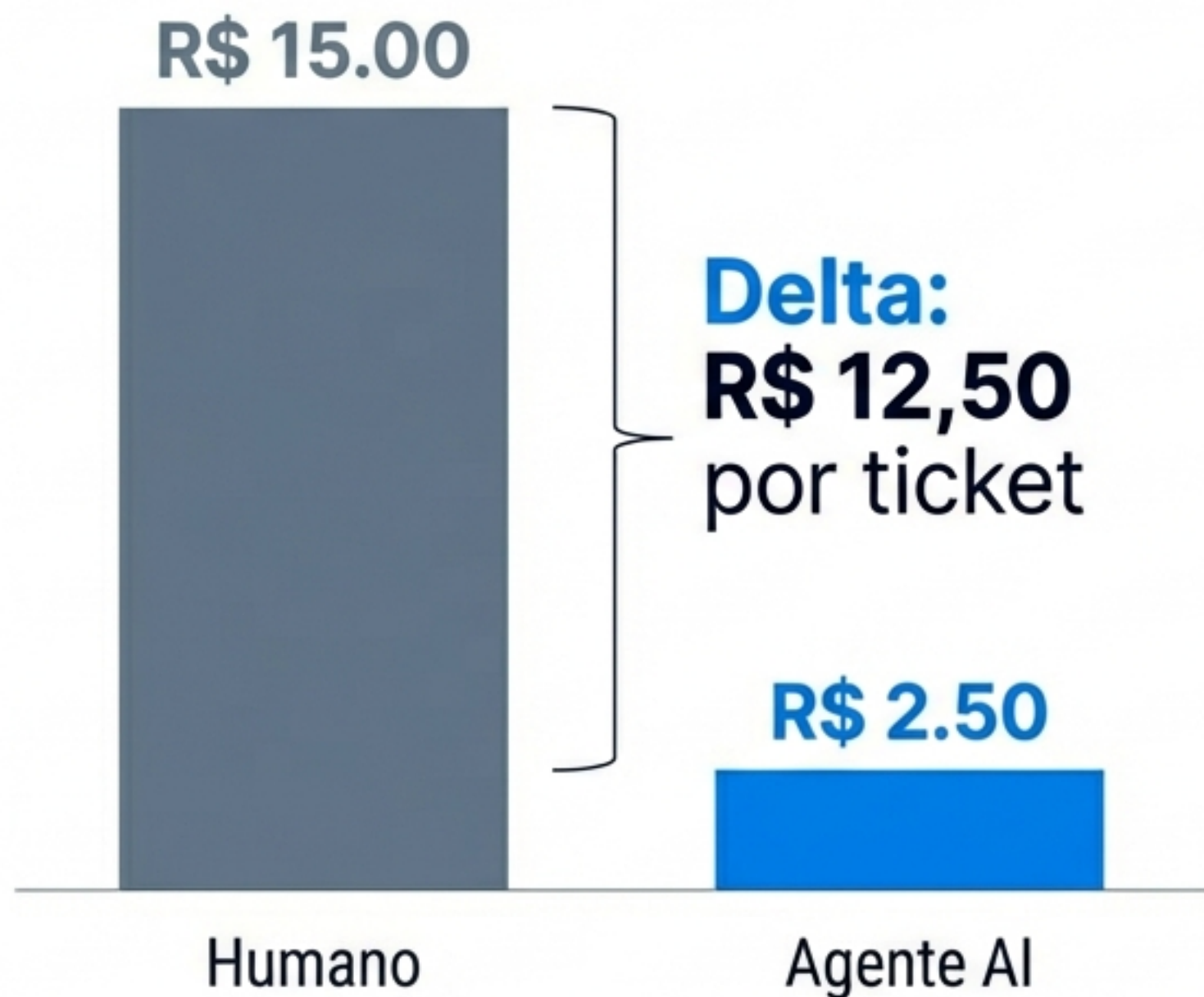
R\$ 2,00 vs R\$ 15,00

Custo por Ticket

AI vs. Operador Humano

Takeaway: Os indicadores de desempenho demonstram um aumento significativo na eficiência e qualidade do atendimento com a implementação de IA.

Uma economia operacional mensal projetada em R\$ 500.000



Cálculo de Economia Mensal

Volume: 100.000 tickets/mês
Taxa de Contenção (AI): 40% (40.000 tickets)

40.000 tickets x R\$ 12,50 economia

R\$ 500.000 / mês (Economia Líquida)

Payback em menos de 1 mês e economia acumulada de R\$ 6 Milhões em um ano

CAPEX Inicial

Setup: R\$ 170k (Dev + Fine-Tuning)

Payback: < 30 dias



Escalabilidade elástica e inteligência de dados acionável



Escalabilidade Infinita

Resposta imediata a picos de demanda (ex: Black Friday) **sem custos** de horas extras ou contratação temporária.



Data Intelligence

Conversas não são perdidas; viram insights estruturados de produto.

“30% dos usuários reclamam da UX do Pix na tela de confirmação”.

Governança e Ética: Os 4 Pilares da "Trust Architecture"

Transparência

Identidade "Sou uma IA" obrigatória e direito à explicação.



Human-in-the-Loop (HITL)

Bloqueios e negações exigem validação humana.

Viés Algorítmico

Monitoramento ativo para evitar discriminação demográfica.



Privacidade & LGPD

Anonimização de PII e infraestrutura privada (Azure/AWS).

FIAP AI Architecture • Visão 2026

Eficiência Operacional + Empatia no
Atendimento + Segurança Bancária

Luiz Ihara • Filipe Fragoso • Pedro Martins • Daniel Caminha