

1. Nome do Projeto

- Nome original: STELLAR: A Structured, Trustworthy, and Explainable LLM-Led Architecture for Reliable Customer Support.
- Nome traduzido: STELLAR: Arquitetura Estruturada, Confiável e Explicável Baseada em LLMs para Suporte Inteligente ao Cliente.

2. Sobre o Projeto

2.1. Descrição Sucinta do Projeto

STELLAR (*Structured, Trustworthy, and Explainable LLM-Led Architecture for Reliable Customer Support*) é uma arquitetura modular inovadora para sistemas inteligentes de suporte ao cliente, projetada para orquestrar Grandes Modelos de Linguagem (*Large Language Models* - LLMs) de forma confiável e explicável. Sua base é uma estrutura de Grafo Acíclico Direcionado (*Directed Acyclic Graph* - DAG), que define fluxos de interação controlados entre seus componentes, garantindo previsibilidade e rastreabilidade.

O sistema é composto por 9 módulos especializados:

- Módulo 1 (Classificador de Intenção): Encaminha a consulta inicial do usuário para o caminho correto.
- Módulo 2 (Recuperação de FAQ - RAG): Responde a perguntas gerais usando *Retrieval-Augmented Generation* (RAG) híbrida com o método de *re-ranking*.
- Módulo 3 (Recuperador de Informações Diretas): Extrai informações específicas (ex: contatos) de um corpus definido via *in-context learning*.
- Módulo 4 (Gerenciador de Contexto para Agente Humano): Prepara o contexto, calcula urgência e atribui agentes humanos.
- Módulo 5 (Análise de Sentimento): Avalia o tom emocional da interação.
- Módulo 6 (Coletor de Feedback): Coleta e classifica automaticamente o feedback do usuário.
- Módulo 7 (Construtor de Base de Conhecimento): Sugere rascunhos de novas FAQs (Perguntas frequentes) com base em consultas não resolvidas, para revisão humana.
- Módulo 8 (Verificador de Resolução): Confirma com o usuário se a resposta automatizada resolveu o problema.
- Módulo 9 (Verificador de Conformidade): Avalia as respostas geradas por LLMs com base em critérios de qualidade e segurança antes da exibição.

A interação sempre se inicia com o Módulo 1 roteando a consulta para um dos três caminhos principais: Resposta via FAQ (M2), Informação Direta (M3), ou Escalonamento Humano (M4/M5), conforme ilustrado esquematicamente na **Figura 1**. Pontos de decisão chave (M8 - validação do usuário, M9 - conformidade) e regras de fallback definidas garantem que a interação siga um dos 11 fluxos de trabalho predefinidos, detalhados na **Figura 2**. Essa estrutura DAG garante previsibilidade e controle sobre a execução, utilizando LLMs para tarefas específicas dentro de um fluxo gerenciado e auditável.

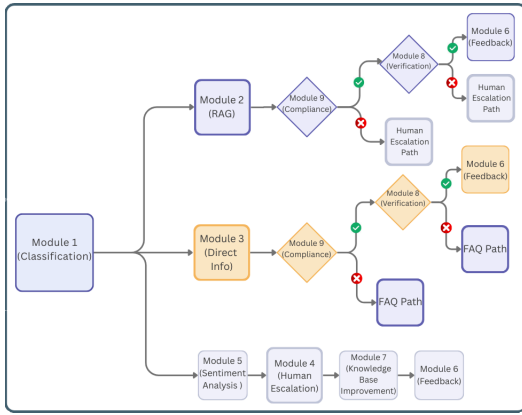


Figura 1: Diagrama simplificado da arquitetura com os 3 caminhos iniciais possíveis.

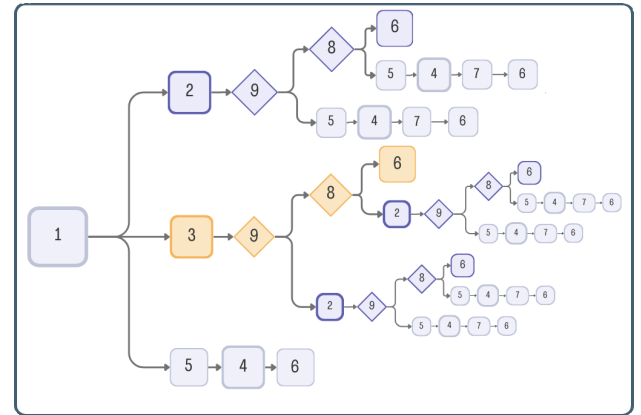


Figura 2: Diagrama completo da arquitetura, com todos os fluxos de trabalho possíveis.

2.2. Oportunidade/Problema

A oportunidade reside na crescente demanda por automação eficiente e confiável no atendimento ao cliente (*Customer Support* - CS). Empresas buscam incessantemente maneiras de otimizar suas operações de suporte, liberando agentes humanos para lidar com casos complexos e urgentes, enquanto tarefas rotineiras e informacionais são resolvidas automaticamente. Uma automação bem-sucedida resulta em respostas mais rápidas para o cliente, redução de custos operacionais e melhor alocação de recursos humanos qualificados.

No entanto, a introdução de tecnologias avançadas como LLMs nesse cenário, embora promissora, expõe um problema crítico: a falta de confiabilidade e controle em sistemas de CS automatizados baseados puramente em Inteligência Artificial (IA) generativa. Especificamente:

- **Modelos de linguagem monolíticos:** Quando usados diretamente, LLMs podem apresentar inconsistências, gerar informações factualmente incorretas (alucinações), e ter dificuldade em seguir instruções complexas ou políticas específicas da empresa de forma consistente. Isso representa um risco significativo em interações críticas com clientes, onde precisão e consistência são essenciais.
- **Sistemas Agênticos Complexos:** Ferramentas que tentam orquestrar múltiplos LLMs ou "agentes" frequentemente priorizam flexibilidade em detrimento da previsibilidade. Isso pode levar a fluxos de interação complexos, não determinísticos e difíceis de depurar. A falta de uma estrutura rigorosa compromete a confiança necessária para um CS de alta qualidade.

Essa lacuna entre a demanda por automação inteligente e a ausência de soluções de IA robustas e confiáveis é a principal dificuldade que a STELLAR busca resolver. Muitas empresas evitam a adoção ampla de IA devido ao risco de experiências negativas para o cliente, ineficiências operacionais ocultas e falhas de conformidade.

A STELLAR enfrenta esse desafio com uma arquitetura baseada em um DAG, que combina o poder dos LLMs com controle, previsibilidade e explicabilidade. Assim, viabiliza-se uma automação de CS eficaz, confiável e gerenciável.

2.3. Mercado Potencial

O mercado para automação inteligente do atendimento ao cliente, onde STELLAR se insere, apresenta um potencial global significativo e está em rápida expansão, com o Brasil destacando-se como um polo de crescimento chave na América Latina.

- Mercado Global: O mercado global de IA para atendimento ao cliente foi avaliado em 12,06 bilhões de dólares em 2024 e espera-se que alcance 47,82 bilhões de dólares até 2030^[1]. Dados complementares do mercado de IA para *call centers* reforçam essa tendência, com estimativas de crescimento de 2,3 bilhões de dólares em 2024 para uma estimativa de 12,8 bilhões de dólares até 2033^[2]. Esse crescimento é impulsionado pela busca contínua das empresas por eficiência operacional, disponibilidade 24/7 e experiências personalizadas para os clientes, indicando uma oportunidade mundial substancial para soluções inovadoras.
- Mercado Brasileiro: O Brasil lidera a adoção de IA conversacional na América Latina, com um mercado estimado em 532,2 milhões de dólares em 2024 e estimativa de 1,893 bilhão de dólares em 2030^[3].

A arquitetura STELLAR é versátil, e, conseqüentemente, facilmente aplicável a qualquer setor com alto volume de atendimento. Alguns dos setores que demonstram maior adoção de serviços com IA e apresentam potencial imediato para uso da STELLAR no Brasil são: Bancário e Financeiro: Setor pioneiro na adoção, utilizando IA para gestão de contas, detecção de fraudes e atendimento inicial; Varejo e E-commerce: Responsável por uma parcela significativa do mercado de IA conversacional no Brasil^[3], usando-a para recomendações, suporte a vendas e pós-venda; Saúde: Crescente adoção pós-pandemia para agendamentos, triagem inicial e suporte a pacientes; Telecomunicações: Utilização para resolução de problemas técnicos, visando à redução do tempo de atendimento.

Embora o projeto STELLAR esteja atualmente em fase de protótipo validado em laboratório, sua proposta de servir como um modelo arquitetural o torna inerentemente aplicável em escala global, visto que ele se propõe a resolver desafios universais na implementação de IA para CS. O tamanho expressivo e a taxa de crescimento dos mercados global e brasileiro, juntamente com a necessidade explícita por soluções mais confiáveis e explicáveis que as abordagens atuais, indicam um mercado potencial robusto e receptivo para a proposta de valor da arquitetura STELLAR.

2.4. Principais Funcionalidades

A arquitetura STELLAR habilita um conjunto de funcionalidades essenciais projetadas para fornecer um suporte ao cliente inteligente, confiável e eficiente:

- Classificação Inteligente e Roteamento Inicial: Analisa a consulta do usuário e a direciona automaticamente para o fluxo de resolução mais adequado (FAQ, informação direta, escalonamento humano).
- Respostas Confiáveis a FAQs: Utiliza um sistema de *Retrieval-Augmented Generation* (RAG) Híbrido (semântico + BM25) com *re-ranking* por LLM para fornecer respostas confiáveis baseadas na base de conhecimento existente.
- Verificação de Conformidade e Qualidade: Avalia automaticamente as respostas geradas por LLMs com relação a critérios predefinidos (ex: confidencialidade, tom, completude) antes de serem apresentadas ao usuário, garantindo segurança e adequação.

- Escalonamento Priorizado e Contextualizado: Quando a automação não é suficiente, prepara um resumo do caso, avalia a urgência (combinando sentimento e tipo de problema) e direciona o cliente para o agente humano mais adequado ou para uma fila de espera priorizada.
- Melhoria Semi-Automática da Base de Conhecimento: Identifica consultas não resolvidas pela automação e utiliza LLMs para gerar rascunhos de novas entradas de FAQ, que são então validadas por humanos, criando um ciclo de aprimoramento contínuo do sistema RAG.

Em conjunto, estas funcionalidades visam oferecer ao cliente um atendimento rápido, personalizado e de alta qualidade, ao mesmo tempo em que otimizam a operação de suporte e garantem a confiabilidade das interações automatizadas.

2.5. Diferenciais da Solução

A arquitetura STELLAR se diferencia das abordagens atuais para automação de suporte ao cliente através de quatro pilares fundamentais:

1. Arquitetura DAG Estruturada para Previsibilidade: Ao contrário de:
 - LLMs Monolíticos: Que operam como "caixas-pretas" com comportamento muitas vezes imprevisível e difícil de controlar para tarefas complexas de CS.
 - Plataformas de Agentes Genéricos (ex: LangChain, Crew AI): Que, embora flexíveis, podem levar a interações excessivamente complexas, não determinísticas e difíceis de auditar.
2. Foco Explícito e Integrado em Confiabilidade:
 - Módulos Dedicados: A inclusão do Módulo 9 (Verificador de Conformidade) como um passo de segurança obrigatório antes de apresentar a resposta ao cliente, e do Módulo 8 (Verificador de Resolução) para validar o sucesso da interação com o usuário, são mecanismos proativos de garantia de qualidade e segurança não integrados de forma tão central em muitas outras soluções.
 - Design Geral: Toda a arquitetura, com seus *fallbacks* definidos e fluxos controlados, é projetada para minimizar respostas incorretas ou inadequadas e garantir que falhas na automação levem a um escalonamento gerenciado.
3. Modelo de Arquitetura Aberta e Adaptável: STELLAR é proposta como um modelo arquitetural, não como uma solução fechada. Sua natureza modular permite que:
 - Empresas adaptem ou substituam módulos específicos (ex: usar um modelo de RAG diferente, integrar com seu próprio sistema de CRM para escalonamento).
 - A arquitetura seja ajustada para diferentes bases de conhecimento, políticas de negócio e até mesmo modalidades (como voz, adicionando módulos de STT/TTS).

Isso oferece uma flexibilidade controlada, permitindo customização sem sacrificar a estrutura e a confiabilidade centrais do sistema.

Esses diferenciais posicionam STELLAR como uma solução única, que busca unir a inteligência dos LLMs com a robustez e controle necessários para o ambiente crítico do suporte ao cliente.

2.6. Maturidade Tecnológica (TRL)

A maturidade tecnológica atual do projeto é: "solução tecnológica demonstrada em ambiente relevante (TRL 6)".

A classificação como TRL 6 baseia-se na existência de um protótipo funcional completo da arquitetura STELLAR. Seus módulos componentes foram individualmente desenvolvidos e validados através de experimentos controlados em ambiente de laboratório, demonstrando alta performance em suas respectivas funções. A integração desses módulos e a execução simulada dos 11 fluxos de trabalho definidos demonstraram a funcionalidade e a viabilidade da arquitetura completa em um ambiente relevante (simulado). O protótipo está, portanto, tecnologicamente demonstrado e pronto para a próxima fase de validação em ambiente operacional real (TRL 7).

3. Modelo de Negócio

A seguir, apresenta-se um modelo de negócio teórico para a arquitetura STELLAR, utilizando a ferramenta *Business Model Canvas*. Este modelo representa uma visão inicial, considerando que o projeto está em estágio TRL 6 e ainda não foi implementado em produção.

- Proposta de Valor: Oferecer uma arquitetura/plataforma para automação de suporte ao cliente (CS) que seja confiável, explicável e eficiente.
- Segmentos de Clientes: Empresas de médio e grande porte com alto volume de interações de CS. Setores prioritários: Bancos, Seguradoras, Telecomunicações, Varejo/E-commerce.
- Canais: Venda direta B2B e parcerias estratégicas com consultorias de tecnologia, CX (Customer Experience) e integradores de sistemas.
- Relacionamento com Clientes: Suporte técnico especializado para implementação e operação; serviços de consultoria para adaptação da arquitetura (customização de módulos, integração com sistemas legados).
- Fontes de Receita: Licenciamento de versões adaptadas da arquitetura a empresas e setores específicos, e receita de serviços profissionais (consultoria, implementação, customização, treinamento).
- Recursos Chave: Conhecimento do Código Fonte do protótipo funcional (base para desenvolvimento de versões adaptadas a necessidades específicas) e experiência técnica em IA, LLMs, Engenharia de Software e processos de CS.
- Atividades Chave: Desenvolvimento e manutenção contínua da arquitetura/plataforma; Implementação e Suporte a clientes; Consultoria para adaptação e otimização.
- Parcerias Chave: Plataformas de *Cloud Computing* (AWS, Azure, GCP) para infraestrutura; Integradores de Sistemas e Consultorias de TI/CX para alcance de mercado e implementação.
- Estrutura de Custos: Infraestrutura: Custos de *Cloud* e uso de APIs de LLM (potencialmente significativos); Custos administrativos gerais.

4. Impacto e Resultados

Impacto Potencial (Sociedade e Negócios):

A implementação de sistemas baseados na arquitetura STELLAR tem o potencial de gerar impactos positivos significativos tanto para as empresas quanto para os consumidores e a sociedade em geral:

- Para os Clientes: Experiências de suporte mais rápidas e consistentes para consultas comuns, redução do tempo de espera, respostas automatizadas mais confiáveis e seguras, e acesso mais ágil a agentes humanos para problemas complexos, resultando em maior satisfação geral.

- Para as Empresas: Redução significativa do custo operacional do suporte ao cliente através da automação eficiente de um grande volume de interações. Aumento da consistência e qualidade do atendimento, fortalecendo a imagem da marca. Otimização do trabalho dos agentes humanos, direcionando-os para tarefas de maior valor agregado e complexidade, o que pode levar a maior satisfação e retenção desses profissionais.
- Para a Sociedade: Padronização e elevação do nível de qualidade no atendimento ao consumidor em diversos setores. Promoção do uso ético e responsável da Inteligência Artificial em interações críticas, através da ênfase em explicabilidade.

Resultados Esperados (em Operação):

Quando implementado em um ambiente operacional real, espera-se que o STELLAR gere resultados mensuráveis, incluindo:

- Aumento da taxa de resolução no primeiro contato (*First Contact Resolution* - FCR) para interações automatizadas.
- Diminuição do custo médio por interação de suporte.
- Redução do Tempo Médio de Atendimento (TMA) global, considerando as interações automatizadas rápidas.

Resultados Já Obtidos (Fase de Protótipo - TRL 6):

Os experimentos realizados durante o desenvolvimento e validação do protótipo STELLAR já demonstram a viabilidade técnica e a eficácia dos componentes chave da arquitetura, fornecendo evidências concretas do seu potencial:

- Classificação Confiável de Consultas: O Módulo 1 (Classificador de Intenção) alcançou 100% de acurácia em um conjunto de 125 casos de teste diversos, garantindo o roteamento inicial correto e eficiente das interações.
- Recuperação Eficaz de Informação (RAG): O Módulo 2 (Recuperação de FAQ) demonstrou excelente desempenho de recuperação com Hit-Rate@5 de 99% em um experimento com uma base de conhecimento de 100 FAQs e 100 casos teste.
- Extração Precisa de Dados Diretos: O Módulo 3 (Recuperador de Informações Diretas) obteve 100% de acurácia em 50 casos testes de extração de informações específicas (ex: números de telefone) presentes em um corpus definido, validando a abordagem de *in-context learning*.
- Análise de Sentimento Precisa: O Módulo 5 (Análise de Sentimento) classificou corretamente o sentimento predominante em 96% (48/50) dos casos de teste, fornecendo um input confiável para a avaliação de urgência.
- Verificação de Conformidade Robusta: O Módulo 9 (Verificador de Conformidade) demonstrou 100% de acurácia tanto na identificação de respostas como conformes/não-conformes quanto na classificação correta do tipo de violação em 50 casos de teste, validando sua função crítica como camada de segurança e qualidade.

Estes resultados obtidos em ambiente controlado e simulado fornecem forte suporte à proposta de valor do STELLAR, indicando que seus componentes fundamentais são capazes de operar com alta precisão e confiabilidade, pavimentando o caminho para os impactos positivos esperados em uma implementação real.

5. Possibilidade de Startup

O projeto STELLAR apresenta um forte potencial para originar uma startup focada em soluções de IA confiáveis para suporte ao cliente, dada a relevância do problema que aborda e a inovação de sua arquitetura. No entanto, no estágio atual, o foco principal da equipe está na validação tecnológica aprofundada, na divulgação científica dos resultados e conceitos através de publicações e eventos (como este concurso). A constituição formal de uma startup é uma possibilidade considerada para o futuro.

6. Parceiros

O desenvolvimento do projeto STELLAR contou com a colaboração e o apoio dos seguintes parceiros:

- Universidade Estadual de Campinas: Instituição onde a pesquisa e o desenvolvimento do protótipo foram realizados, vinculados a trabalho de Iniciação Científica.
- Semantix: Empresa que forneceu apoio financeiro para a realização desta pesquisa.
- Prof. Dr. Hélio Pedrini (Unicamp): Orientador da pesquisa, fornecendo supervisão acadêmica e direcionamento técnico ao longo do desenvolvimento do projeto.

7. MVP

O protótipo funcional da arquitetura STELLAR, que demonstra a interação entre os módulos e a execução dos fluxos de trabalho definidos, está disponível publicamente no seguinte repositório GitHub:

[https://github.com/Matheus-F-Scatolin/STELLAR MVP Selo de Inovacao](https://github.com/Matheus-F-Scatolin/STELLAR_MVP_Selo_de_Inovacao).

Informações sobre como testar a arquitetura estão no arquivo de descrição do repositório.

Observação Importante para Avaliação: O código implementa a arquitetura STELLAR e seus módulos principais. Para fins de demonstração e testes, o protótipo foi configurado utilizando como base de conhecimento dados publicamente disponíveis da empresa Bradesco Seguros (extraídos do *website*, como FAQs e informações de contato) no momento do desenvolvimento. O conjunto de FAQs públicas utilizado neste protótipo é limitado em escopo e variedade. Como resultado, muitas perguntas formuladas pelos avaliadores podem não ter uma resposta direta na base de conhecimento do Módulo 2 (RAG). Nesses casos, o comportamento esperado do STELLAR é ativar os mecanismos de fallback e chamada de atendentes humanos, direcionando a interação para o Módulo 4 (Atendimento Humano), após passar pelas verificações apropriadas (Módulo 9 e/ou Módulo 8). Este comportamento demonstra o funcionamento correto da lógica de controle e das salvaguardas da arquitetura para lidar com consultas não resolvidas automaticamente, o que é parte integral do design do STELLAR.