



## Eixo 3 Segurança Pública - Descritivo do Projeto

### Sistema Inteligente de Transcrição e Análise de Áudio para Segurança Pública

#### 1. Contextualização

No âmbito do Plano de Trabalho do Convênio nº 01/2024- CIIADF, coordenado pelo Instituto Hardware Br – HBR foram desenvolvidas soluções inovadoras para modernização da segurança pública no Distrito Federal utilizando Inteligência Artificial.

No contexto da segurança pública, o atendimento de ocorrências por meio de chamadas de voz (centrais de atendimento, despacho policial, emergências) exige rapidez, precisão e capacidade de registro estruturado das informações.

Atualmente, grande parte dessas interações é registrada manualmente, o que pode gerar:

- Perda de informações relevantes;
- Inconsistência no preenchimento de boletins de ocorrência;
- Aumento do tempo de resposta operacional;
- Sobrecarga dos agentes responsáveis pelo atendimento.

Diante desse cenário, o uso de Inteligência Artificial aplicada à transcrição e interpretação de áudio surge como uma oportunidade estratégica para automatizar e qualificar esse processo.

#### 2. Objetivo do Projeto

Esta Prova de Conceito (PoC) tem como objetivo demonstrar a viabilidade de um sistema inteligente capaz de:

- Transcrever chamadas de áudio em tempo real com alta acurácia em português brasileiro;
- Identificar múltiplos falantes em uma conversa;
- Extrair automaticamente informações estruturadas relevantes para boletins de ocorrência;
- Integrar todo o pipeline (áudio → transcrição → análise → armazenamento) em uma arquitetura escalável.

#### 3. Metodologia

O desenvolvimento da solução seguiu uma abordagem experimental e incremental, típica de projetos de Prova de Conceito (PoC), combinando técnicas de processamento de sinais, aprendizado de máquina e engenharia de software para validação da viabilidade técnica em ambiente controlado.

A metodologia foi estruturada nas seguintes etapas:

3.1. Definição de Requisitos e Arquitetura: Inicialmente, foram levantados os requisitos funcionais e não funcionais do sistema, considerando o contexto de uso em segurança pública.

3.2. Implementação do Pipeline de Áudio: Foi desenvolvido um pipeline completo para



processamento de áudio, composto por: Captura e ingestão de áudio via requisições HTTP e WebSocket; Detecção de atividade de voz (VAD) para segmentação inteligente dos trechos com fala; e Pré-processamento do áudio, garantindo qualidade adequada para transcrição.

3.3. Transcrição Automática (ASR): A transcrição foi realizada por meio de modelos de reconhecimento automático de fala (ASR) de última geração, com foco em: Avaliação da acurácia para português brasileiro; Testes com diferentes durações e características de áudio; Validação da capacidade de operação em tempo quase real.

3.4. Diarização de Falantes: Foi implementado um mecanismo de identificação de múltiplos falantes

3.5. Análise Semântica e Extração de Dados: Após a transcrição, foi aplicada uma camada de Inteligência Artificial baseada em modelos de linguagem para gerar resumos automáticos das ocorrências, extrair informações estruturadas relevantes (nome, endereço, tipo de ocorrência, pontos de referência) e transformar texto livre em dados organizados.

3.6. Desenvolvimento da API e Integração: Foi construída uma API backend para orquestração de todo o fluxo, garantindo a interoperabilidade da solução.

3.7. Persistência e Gerenciamento de Dados: Os dados gerados (transcrições, resumos e informações estruturadas) foram armazenados em banco de dados relacional.

3.8 Containerização e Ambiente de Execução: A solução foi containerizada, possibilitando reprodutibilidade do ambiente e facilidade de implantação.

#### 4. Resultados Alcançados

A implementação da PoC demonstrou a viabilidade técnica da solução proposta, com os seguintes resultados:

- Transcrição de áudio em português brasileiro com boa acurácia, mesmo sem ajustes finos específicos;
- Extração automática de dados estruturados a partir do conteúdo transcrito, com funcionamento adequado em cenários reais;
- Implementação de fallback heurístico para garantir funcionamento mesmo na indisponibilidade do modelo de linguagem;
- Processamento em tempo real via WebSocket com latência compatível com uso operacional;
- Pipeline completo funcional, integrando todas as etapas desde o áudio até o armazenamento dos dados;
- Ambiente totalmente containerizado, facilitando implantação e replicação.

Os resultados indicam que a solução é tecnicamente viável e possui potencial para aplicação em ambientes reais de segurança pública.

#### 5. Evolução Esperada da PoC para Produto

A startup selecionada deverá evoluir a PoC de transcrição e análise de áudio para uma solução operacional demonstrável, capaz de apoiar centrais de atendimento, despacho e registro



estruturado de ocorrências em segurança pública. Considerando o ciclo de 180 dias e o aporte equity-free de R\$ 100 mil, a expectativa é que a startup acelere a adaptação de uma solução madura ao blueprint técnico já validado, com foco em acurácia, baixa latência, robustez operacional, segurança, integração e evidências para contratação. A PoC original já demonstrou viabilidade técnica com pipeline completo de áudio, transcrição em português brasileiro, diarização de falantes, extração semântica, API backend, persistência em banco relacional e ambiente containerizado.

Ao final dos 180 dias, a solução deverá demonstrar capacidade de processar chamadas ou áudios representativos, gerar transcrição confiável, identificar falantes, extrair informações relevantes e estruturar dados úteis para boletins, despacho ou análise posterior. A entrega esperada não é apenas “transcrever áudio”; é reduzir perda de informação, acelerar registro, melhorar qualidade do atendimento e criar rastreabilidade operacional.

A evolução esperada deve contemplar, no mínimo:

#### **i. Pipeline operacional de áudio em tempo real ou quase real**

A startup deverá entregar pipeline funcional para ingestão de áudio, segmentação, transcrição, diarização, análise semântica, armazenamento e consulta. O sistema deve demonstrar latência compatível com uso operacional e capacidade de processar áudios com ruído, sotaques, interrupções e múltiplos interlocutores.

#### **ii. Transcrição calibrada para português brasileiro e contexto de segurança pública**

A solução deverá ser avaliada em áudios representativos do contexto real, incluindo termos comuns de emergência, localização, tipos de ocorrência, nomes próprios, números, endereços, referências geográficas e linguagem espontânea. O objetivo é medir acurácia útil para operação, não apenas desempenho em áudio limpo.

#### **iii. Extração estruturada de dados críticos**

A solução deverá extrair campos relevantes para atendimento e registro, como tipo de ocorrência, local, nomes, telefone, endereço, ponto de referência, envolvidos, risco percebido, urgência, armas, vítimas, veículos, horário, resumo da ocorrência e eventuais ações recomendadas. Cada campo extraído deve preservar vínculo com o trecho da transcrição que o originou, para permitir conferência humana.

#### **iv. Interface de revisão humana e validação operacional**

A solução deverá incluir mecanismo para revisão, correção e aprovação das transcrições e dados extraídos. Em segurança pública, a IA não pode ser tratada como fonte final sem validação. O sistema deve apoiar o operador, reduzir digitação e perda de informação, mas manter rastreabilidade e controle humano.



## **v. Logs, auditoria e cadeia de custódia da informação**

A solução deverá registrar eventos relevantes: áudio recebido, processamento realizado, modelo utilizado, transcrição gerada, alterações humanas, usuário responsável, data/hora e versão dos dados. A rastreabilidade é crítica para confiança institucional, auditoria e eventual uso administrativo ou jurídico da informação.

## **vi. Segurança da informação e controle de acesso**

A startup deverá apresentar arquitetura segura, com criptografia em trânsito e em repouso, segregação de perfis, autenticação, autorização, política de retenção, controle de acesso e tratamento adequado de dados sensíveis. O contexto de segurança pública exige padrões mais rígidos que uma aplicação SaaS genérica.

## **vii. Integração com sistemas existentes ou camada intermediária**

A startup deverá demonstrar como a solução pode se integrar a sistemas de atendimento, despacho ou registro existentes, ainda que inicialmente por API, exportação estruturada, webhook, fila de eventos ou integração assistida. O ponto de chegada deve ser interoperabilidade, não uma aplicação isolada.

## **viii. Evidência de desempenho e custo operacional**

A startup deverá entregar métricas de acurácia, latência, taxa de erro por tipo de áudio, estabilidade, consumo computacional e custo estimado por chamada ou minuto processado. Para procurement público, custo unitário e previsibilidade operacional são tão importantes quanto sofisticação técnica.

## **ix. Pacote de contratação e implantação**

Ao final do ciclo, a startup deverá entregar documentação técnica e operacional: arquitetura, requisitos de infraestrutura, integrações, segurança, SLAs sugeridos, plano de implantação, plano de treinamento, matriz de riscos, custos estimados, limites conhecidos e roadmap de evolução.

### **Marcos sugeridos de evolução**

- **Até 60 dias:** arquitetura-alvo validada, amostras de áudio ou cenários representativos definidos, baseline de acurácia/latência estabelecido, integração-alvo mapeada e primeira versão do pipeline operacional preparada.
- **Até 120 dias:** transcrição, diarização e extração estruturada testadas em cenários representativos, interface de revisão implementada, métricas preliminares consolidadas e controles de segurança definidos.



INSTITUTO  
HARDWARE BR

- **Até 180 dias:** demonstração operacional completa realizada, evidências de desempenho documentadas, integração ou simulação de integração validada, plano de implantação preparado e pacote de contratação entregue.

#### **Critério de chegada**

A PoC será considerada evoluída com sucesso quando houver uma solução demonstrável, segura, auditável e tecnicamente documentada, capaz de transformar áudio em informação operacional estruturada, com desempenho medido, integração planejada e evidências suficientes para subsidiar contratação ou piloto ampliado.

#### **6. Considerações Finais**

O projeto demonstra o potencial da Inteligência Artificial para transformar processos críticos na segurança pública, automatizando tarefas operacionais e ampliando a capacidade analítica das instituições.

A solução proposta representa um passo relevante na modernização dos sistemas de atendimento e registro de ocorrências, com ganhos em eficiência, precisão e escalabilidade.