



Saúde Avançada com BigData e IA

Transformando dados em Saúde para Todos



O DESAFIO

Interoperabilidade em saúde

Fragmentação dos dados

Informações de saúde estão distribuídas em sistemas e bases isoladas.

Falta de integração

Dados clínicos, epidemiológicos e socioeconômicos permanecem desconectados.

Impacto na tomada de decisão

A ausência de uma visão integrada dificulta a identificação de padrões, prevenção e planejamento em saúde.



A SOLUÇÃO

Sabiá Data Lake

Integração de múltiplas fontes de dados.

Epidemiológicos

Doenças e vigilância

Clínicos

Prontuários e exames

Genômicos

Dados genéticos

Socioeconômicos

Clima, economia e território

Uma base integrada para análises avançadas e decisões baseadas em evidências.

OBJETIVOS

Sabiá Data Lake

Construir um Data Lake integrado de dados em saúde



Viabilizar análises avançadas e preditivas com Inteligência Artificial

Subsidiar decisões e políticas públicas baseadas em evidências.



Garantir interoperabilidade, governança e segurança dos dados.

ARQUITETURA MEDALHÃO

Sabiá Data Lake

Fontes de dados

Sistemas Hospitalares | e-SUS | Wearables | Outras fontes

Bronze

Dados Brutos

- Ingestão
- Preservação
- Dados no formato original

Prata

Dados tratados

- Limpeza
- Padronização
- Validação
- Integração

Ouro

Dados analíticos

- Modelagem
- Agregações
- Indicadores
- Dados prontos para consumo

Consumo

- IA
- APIs
- Dashboards
- Análises

Camada Transversal de Governança e Segurança

RBAC / LDAP

Linhagem & Audit

Anonimização LGPD

INFRAESTRUTURA

Sabiá Data Lake

Capacidade de Armazenamento

1.2 PETABYTES

Capacidade de processamento

- **512 GB** – Memória RAM
- **192** – Núcleos de Processamento

IA e processamento paralelo

- 2 GPUs NVIDIA RTX A4500
- 40 GB de memória gráfica

INTEGRAÇÃO COM SISTEMAS DE SAÚDE

Para alimentar as bases de dados do Data Lake realizaremos a extração de dados coletados por instituições e hospitais parceiros, sempre priorizando a segurança e privacidade de tais dados, esses dados coletados incluem:

+395.000

registros de
pacientes
atendidos por
hospitais e
instituições
parceiras.

+7.000.000

de registros
clínicos e
diagnósticos
identificados entre
nossos parceiros.

+2.900.000

de exames
realizados e
registrados nos
sistemas de
instituições
parceiras.

+80

bases de dados
brasileiras públicas
mapeadas, com
dados de saúde,
demografia, entre
outros.

TELEMETRIA VIA WEARABLES

A integração de dispositivos vestíveis expande a saúde digital além do histórico estático de exames: sinais vitais viram fluxos de dados em tempo real, capturados de forma contínua e de alta frequência.

Frequência Cardíaca



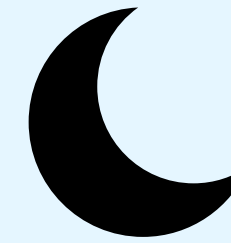
Batimentos monitorados em tempo real

Oxigenação (SpO²)



Saturação de oxigênio no sangue

Padrões de Sono



Qualidade e clicos de descanso



Ingestão na Camada Bronze: esses dados brutos alimentam a Arquitetura Medalhão, permitindo correlacionar métricas contínuas com o histórico do paciente e antecipar eventos de saúde.

INGESTÃO E ANÁLISE PREDITIVA

Pipeline dedicada a tratar, padronizar e estruturar grandes volumes de dados biométricos antes de sua consolidação nas camadas superiores do Data Lake.



O diferencial da arquitetura é o cruzamento em tempo real entre histórico clínico e telemetria recente do wearable.

CATÁLOGO DE SOLUÇÕES



Japiim Health

Chatbot de IA para acesso a informações em saúde.



SynapSpeech

IA para análise de fala e triagem de sinais cognitivos.



Guaraci - Data Lake

Orquestrador para busca e organização de dados públicos.



AutoAI-Pandemics

IA para análises epidemiológicas e bioinformáticas.



Sabiá Vitalis

Monitoramento preventivo em saúde por meio de wearables.



InteliGente

Ecossistema de soluções de IA aplicada à saúde, ciência e inovação.

Guaraci: Diversas Bases; Um caminho.



Um panorama abrangente do Brasil

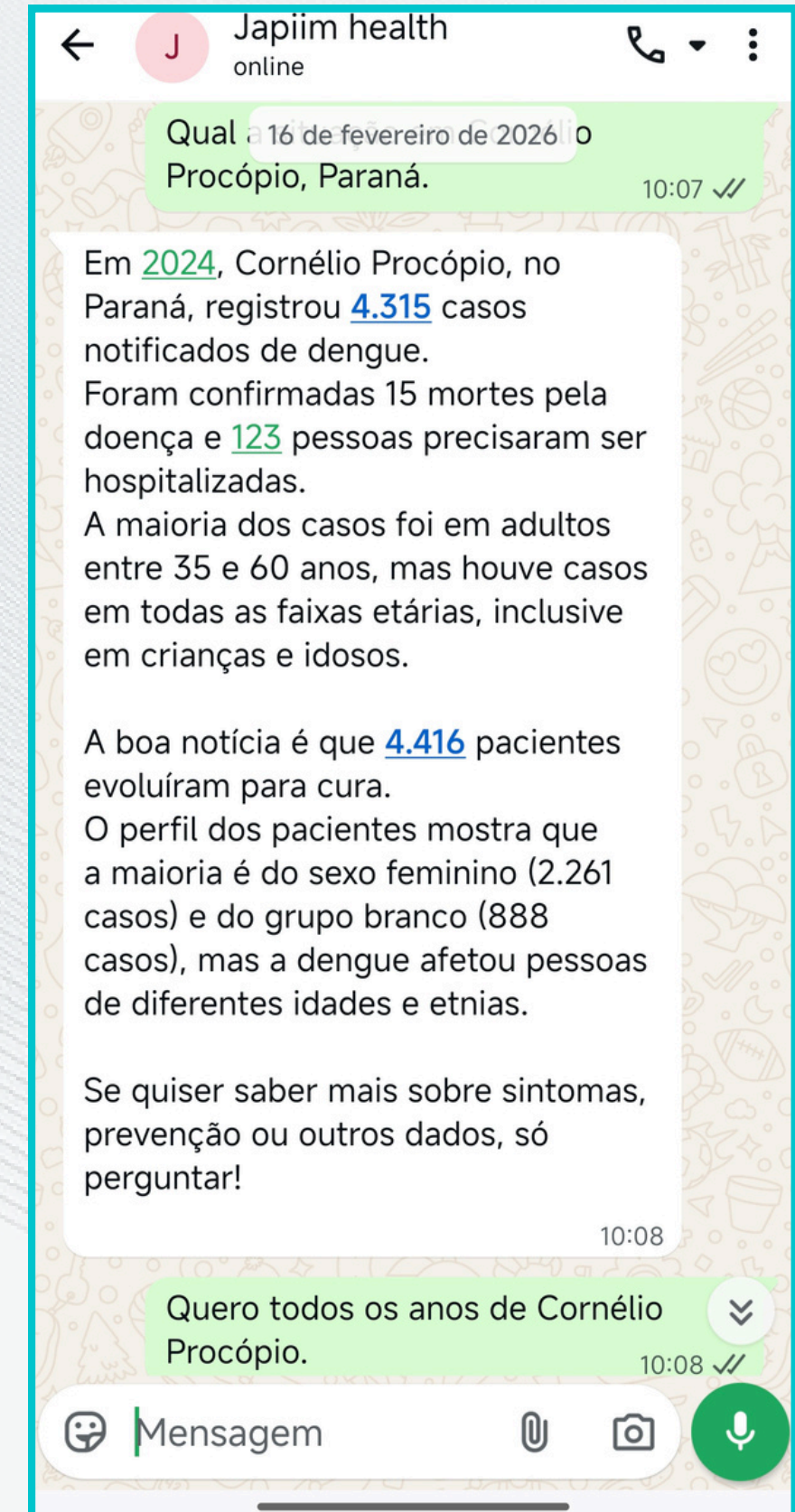
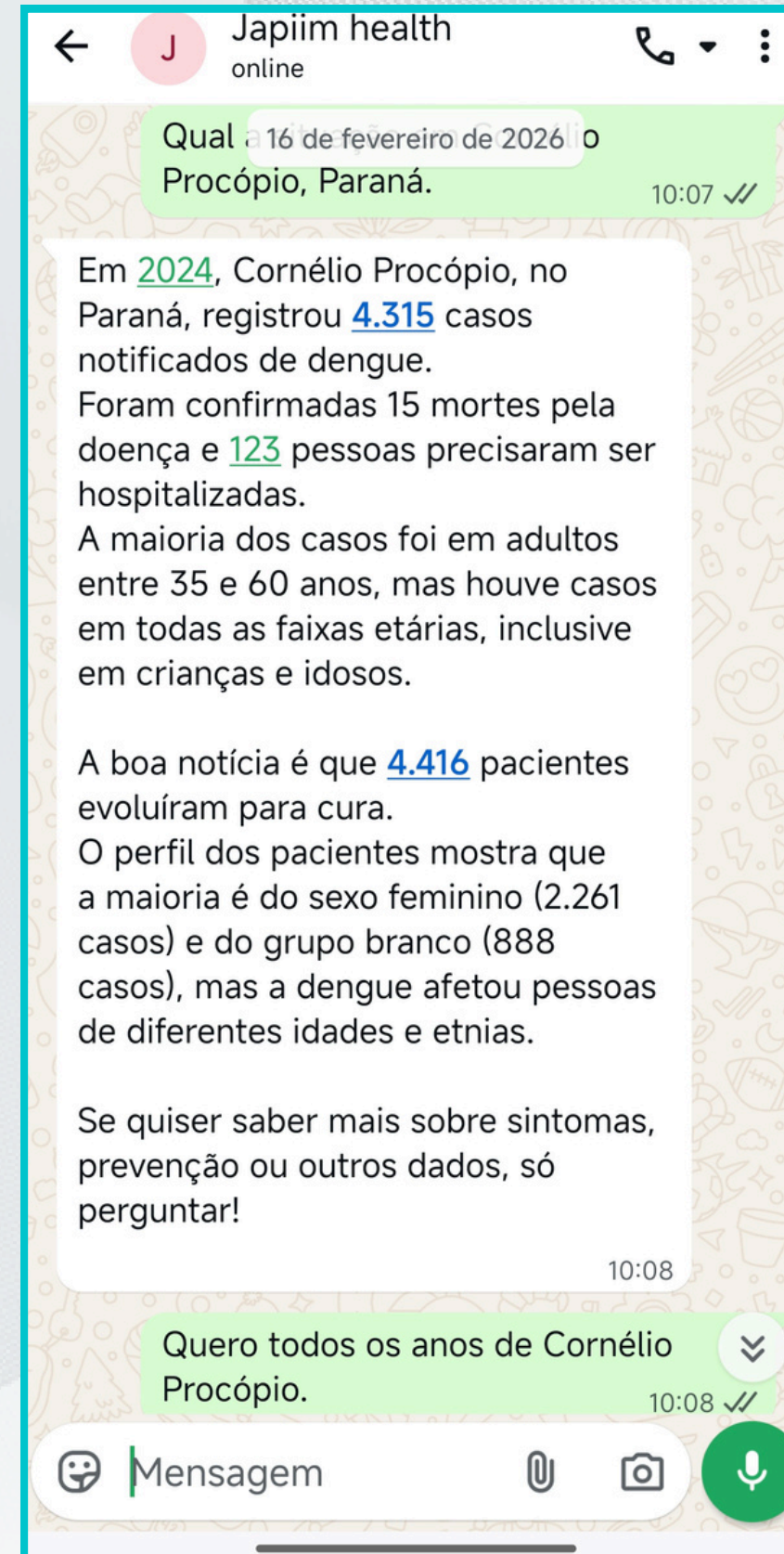
Bilhões de registros de dados de saúde

O Guaraci é um orquestrador de downloads: ele recupera dados públicos do Brasil diretamente de cada fonte oficial e entrega tudo pronto para uso, com uma interface web e fidelidade aos dados originais. Menos tempo buscando bases de dados, mais tempo pesquisando, planejando e criando soluções para o país.

Japiim Health: Saúde pública na palma da sua mão

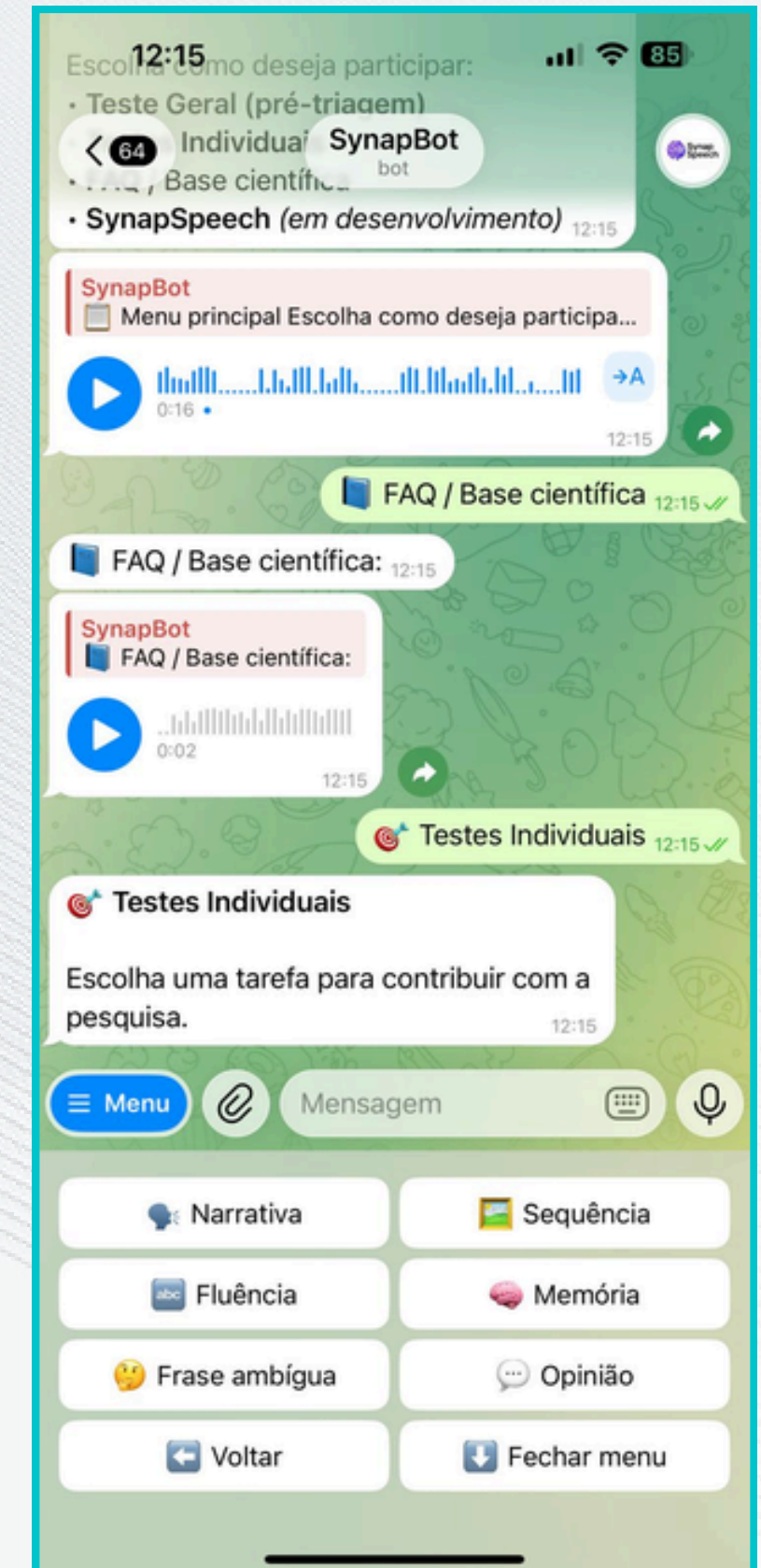
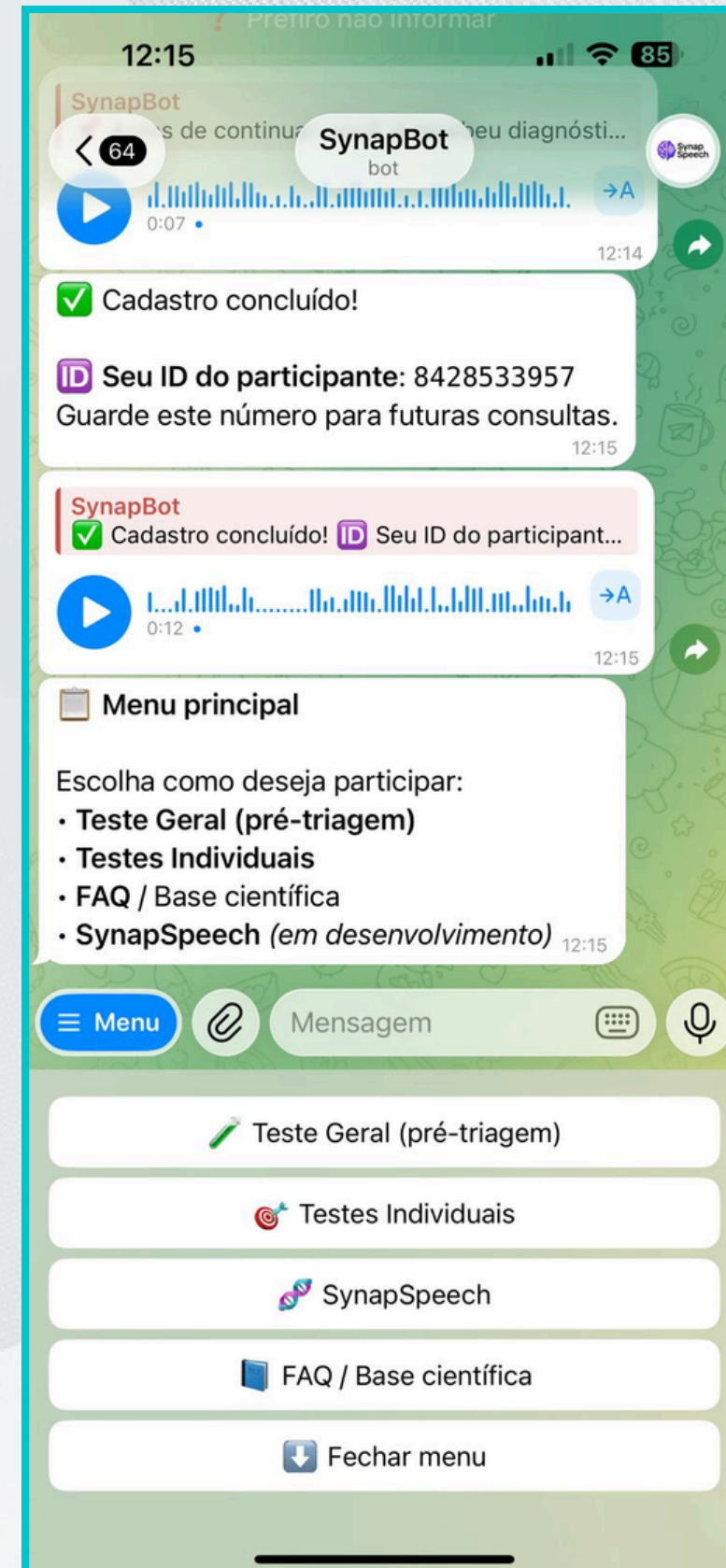


O Japiim transforma dados complexos do SUS em conversas simples no WhatsApp. Informações oficiais, rápidas e acessíveis a todos.



SynapSpeech: Ouvindo sinais por meio da linguagem

Uma solução para triagem e monitoramento de sinais cognitivos associados ao Alzheimer, com base na fala espontânea em português brasileiro.



STATUS DO PROJETO

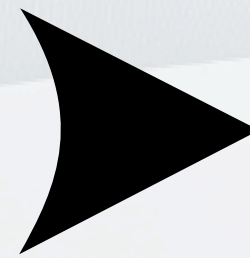
Sabiá Data Lake



VISÃO DE FUTURO

Hoje

Ambiente de armazenamento e governança de dados em saúde



Futuro

Ecossistema preditivo de suporte à decisão em saúde, unindo prontuários históricos e telemetria em tempo real

Detecção Precoce

Identificação antecipada de agravos à saúde

Otimização de Recursos

Melhor alocação de recursos públicos

Insights Acionáveis

Suporte a gestores e profissionais em larga escala

Expansão progressiva das parcerias institucionais e integração contínua de novas fontes de dados.

AGRADECIMENTOS



Página Oficial



MUITO OBRIGADO!