

# Maria Madalena: Extração de Dados Enriquecidos via ChatBot

Luca Ferreira Barboza, Danilo Cadena Lima, Ricardo Rocha

Universidade Federal da Bahia





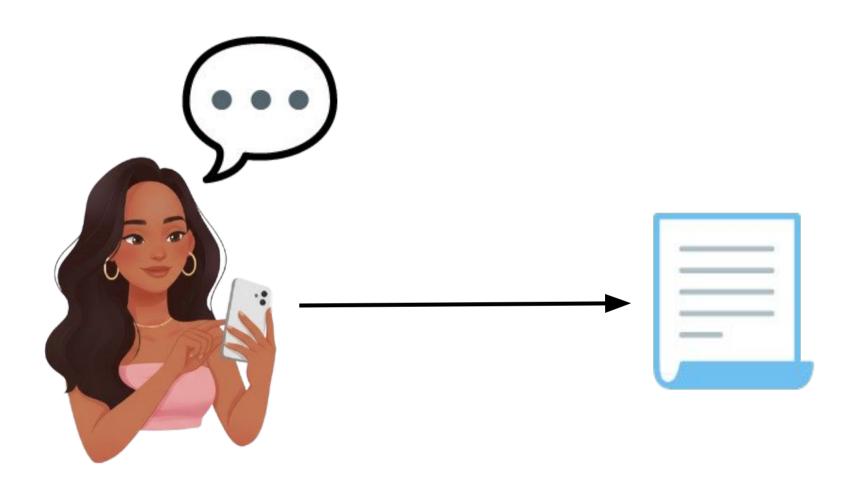
### INTRODUÇÃO

Questionários e formulários tradicionais, por necessitarem de opções de resposta pré-definidas, frequentemente limitam as respostas e a espontaneidade dos dados coletados. Buscando superar essa limitação, este trabalho propõe uma metodologia de coleta de dados lúdica, simplificada e mais confortável, através de um chatbot personalizado, a Maria Madalena. Com ela, queremos aumentar o engajamento do usuário e extrair informações ricas e estruturadas de conversas naturais, incluindo dados não solicitados diretamente.

### MÉTODOS DE INTELIGENCIA ARTIFICIAL

### Extração Estruturada (Chatbot → JSON):

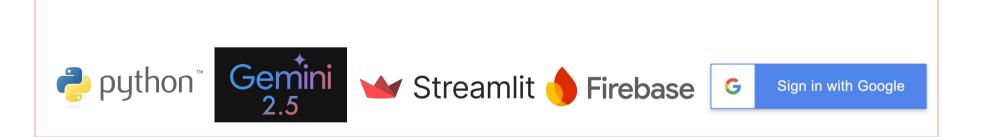
Após a conclusão do chat, o histórico da conversa é processado por um segundo sistema de IA (Gemini 2.5 Flash com output estruturado). Um prompt específico instrui o modelo a extrair as informações mais relevantes. O modelo é obrigado a formatar a saída em um JSON, seguindo um esquema que define os tópicos principais de interesse, como objetivo\_amoroso, valores, estilo\_de\_vida, deal\_breakers e linguagem\_amor. Esse processo garante que dados qualitativos e não estruturados se tornem um dataset.



**Interface:** A plataforma foi desenvolvida em Streamlit, uma biblioteca Python. A autenticação dos usuários é gerenciada de forma segura através do Login com Google.

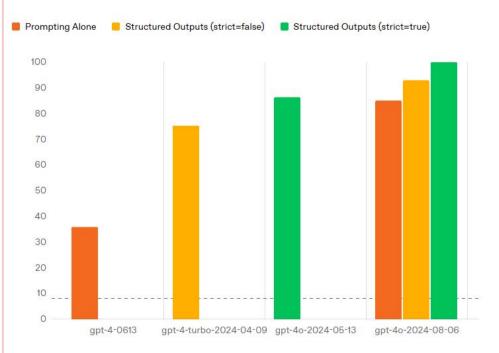
**Inteligência Artificial:** O núcleo do chatbot utiliza a API da Google. O modelo Gemini 2.5 Flash Lite.

**Base de Dados:** Todas as interações e os dados extraídos são armazenados no Firebase, um banco de dados NoSQL



### **OUTPUT ESTRUTURADO**

A inovação central deste trabalho é o uso do output estruturado para converter texto livre em um formato de dados padronizado e pronto para análise. Isso supera a simples engenharia de prompt, que pode gerar resultados inconsistentes.



'objetivo amoroso': 'Alguém que cumpra o que fala e não faça promessas vazias, demonstrando compromisso com o que é dito e não desmarcando compromissos com facilidade.'

'valores': 'Valores familiares, construção de um lar, compartilhar a vida a dois, casar e ter filhos (entre 5 e 8 anos, antes dos 30). Valores cristãos são inegociáveis. Também valoriza uma boa conversa (sobre coisas cotidianas, profundas ou fofocas) e toque físico.',

'estilo\_de\_vida': 'Mistura de caseiro, aventureiro e intelectual.
Atualmente, sente-se mais caseiro devido à rotina, mas sonha em viajar o
mundo e conhecer culturas novas com o parceiro. Gosta de passear por lugares
bonitos, ver filmes e acompanhar universos geeks, assistir séries em casa e
cozinhar receitas novas.',

'deal breakers': 'Falta de palavra e compromisso (falar algo e não demonstrar, marcar e desmarcar com facilidade). Não querer filhos ou família (por serem valores inegociáveis). Traição, embora acredite que possa ser perdoada dependendo da situação e do parceiro.',

'linguagem \_amor': 'Principalmente toque físico (abraços apertados), palavras de afirmação (elogios sinceros) e tempo de qualidade (conversas

expressar ou receber afeto.',

'outros': 'Sente-se muito cansado e esgotado pela convivência e manutenção de relações interpessoais. Busca um parceiro companheiro e recíproco, que o acompanhe em seus interesses (como universo geek e viagens), mesmo que não os compartilhe totalmente, e que tope viajar o mundo e criar filhos

longas). Presentes são agradáveis e fofos, mas não são a principal forma de

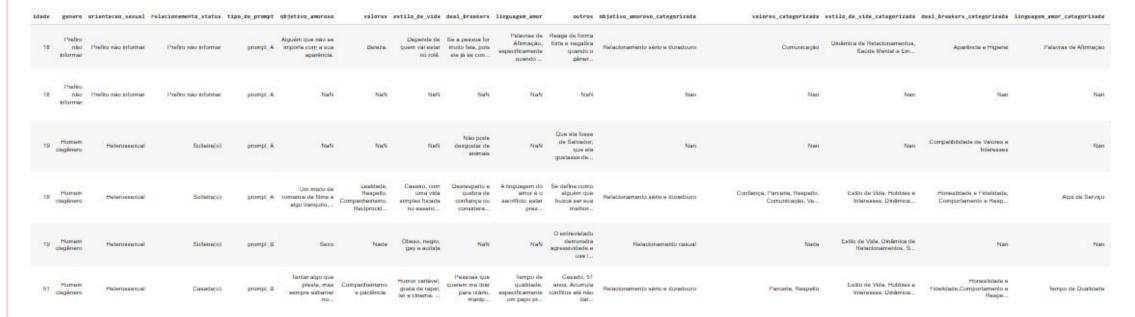
## Interaja com a Maria Madalena

https://mariamadalena.streamlit.app/



### **RESULTADOS**

A aplicação da nossa metodologia resultou na criação bem-sucedida de um dataset consolidado e estruturado, transformando eficazmente as conversas dos usuários em dados prontos para análise.



O chatbot desenvolvido neste trabalho foi projetado para 'coletar informações adicionais' que aparecem de forma livre durante a interação com o usuário. A categoria 'outros' foi criada especificamente para armazenar esses dados.

Reage de forma forte e negativa quando o gênero do potencial parceiro é presumido como masculino (O CARA? TA LOCO).

NaN

Que ela fosse de Salvador; que ela gostasse de animais.

Se define como alguém que busca ser sua melhor versão, oferecendo o melhor aos outros. Suas qualidades são perseverança, revolta, resiliência, força, coragem, altruísmo, bravura, confiança e esperança. Enxerga um relacionamento ideal sem mentiras, com problemas resolvidos por conversas claras e objetivas, onde o casal cresce junto com companheirismo, entendimento e mimo. Sua maior ambição é ser sua melhor versão, temendo ficar aquém de seu potencial. Seu maior prazer é jogar futebol. Na paquera, prefere uma abordagem discreta e comunicação não verbal antes da verbal, preocupado em não gerar momentos chatos. Gosta de conhecer alguém indo nos ambientes que a pessoa gosta e aproveitando juntos, além de papos cabeça. Menciona que só saiu uma vez com uma mulher e nunca namorou.

O entrevistado demonstra agressividade e usa linguagem ofensiva durante a conversa.

Casado, 51 anos. Acumula conflitos até não dar mais conta e explode, culpando quem estiver perto. Afirma nunca estar errado. Não se incomoda com alguém que aplauda suas vitórias e as do seu time (Vitória). É evasivo em relação aos seus

### O APP

# Maria Madalena Conselheira do Amor Sua conselheira amorosa pessoal está aqui para te ajudar! Olá Maria! Que bom te ver aqui! Em que posso te ajudar nos assuntos do coração? Me conta o que está no seu coração...

### TESTE DE HIPÓTESE

**Teste A/B:** Para investigar o impacto da personalidade do chatbot no engajamento, implementamos um Teste A/B. Os usuários são aleatoriamente designados para interagir com uma de duas personas distintas da "Maria Madalena", cada uma definida por um prompt detalhado.

A:

"provocando como amiga divertida"

"Cutucar com perguntas e comentários

- "Seu sarcasmo é afiado, beirando a acidez"
- engraçados"

  "Clima de amizade íntima: leve,
  divertido, provocativo"

  "Jogue verdades na cara, com
  perguntas afiadas e comentários
  ácidos"
  - "Clima de parceria impiedosa: sarcasmo ácido, intensidade, tirando o usuário das falsas ilusões"

### ANÁLISE DOS DADOS COLETADOS



