

**UNIVERSIDADE NOVE DE JULHO
DIRETORIA DOS CURSOS DE INFORMÁTICA**

Antonio Augusto Lopes Marianno - RA 324102840

Davi Anjo Gabriel Brandino Ribeiro - 2221200912

Douglas de Sousa Rodrigues -RA 424110441

Felipe Trivia Koshimizu - RA 923205456

Felipe Norberto Malveira- RA 2221201409

Igor Alves da Silva Pereira - RA 419202553

Juliano César Thomé Marin - RA 320102190

Luan de Oliveira Ramos - RA 3021201889

Renan Nunes Marques dos Santos - RA 421202108

Renan de Oliveira Silva - RA 421201286

Echo Mind

Antonio Augusto Lopes Marianno - RA 324102840

Davi Anjo Gabriel Brandino Ribeiro - 2221200912

Douglas de Sousa Rodrigues -RA 424110441

Felipe Trivia Koshimizu - RA 923205456

Felipe Norberto Malveira- RA 2221201409

Igor Alves da Silva Pereira - RA 419202553

Juliano César Thomé Marin - RA 320102190

Luan de Oliveira Ramos - RA 3021201889

Renan Nunes Marques dos Santos - RA 421202108

Renan de Oliveira Silva - RA 421201286

Echo Mind

Trabalho apresentado à Universidade Nove de Julho, UNINOVE, em cumprimento parcial às exigências da disciplina de Projeto Echo Mind, sob orientação do Prof. Dr. Edson Melo de Souza.

**SÃO PAULO
2025**

RESUMO

O projeto EchoMind é um sistema de inteligência artificial interativo que permite aos usuários conversar com personalidades virtuais baseadas em figuras reais ou fictícias, cada uma com seu estilo de linguagem, comportamento e humor.

Palavras-chave: Inteligência Artificial , Personalização, Ollama, React, Streaming de Dados.

ABSTRACT

The EchoMind project is an interactive artificial intelligence system that allows users to converse with virtual personalities based on real or fictional figures, each with their own language style, behavior, and mood.

Keywords: Artificial Intelligence, Personalization, Ollama, React, Data Streaming.

SUMÁRIO

RESUMO	3
ABSTRACT	4
1. OBJETIVOS	6
2. DESCRIÇÃO DA EMPRESA	7
3. MISSÃO, VISÃO E VALORES DA EMPRESA	8
3.1 MISSÃO	8
3.2 VISÃO	8
3.3 VALORES	8
4. EchoMind	9
4.1 Ollama	9
4.2 Fluxo	10
4.3 Tecnologias usadas	11
5. Próximos Passos / Evolução	12
6. DEFINIÇÃO DA EQUIPE, DIVISÃO DE PAPÉIS E TAREFAS, E CRONOGRAMA DE DESENVOLVIMENTO.	12
Componentes da Equipe e Divisão de Papéis e Tarefas:	16
7. Cronograma de Desenvolvimento	20
CONCLUSÃO	22
REFERÊNCIAS	23

1. OBJETIVOS

O projeto EchoMind tem como objetivo geral desenvolver um sistema interativo de inteligência artificial que permita aos usuários personalizar e interagir com personalidades virtuais baseadas em figuras reais ou fictícias, apresentando estilos de linguagem, comportamento e humor próprios.

- Simular interações realistas com diferentes personalidades, mantendo coerência no estilo e comportamento.
- Demonstrar a possibilidade de customização de personagens.
- Proporcionar uma experiência de usuário fluida e natural através de uma interface de chat simples e intuitiva.
- Disponibilizar um catálogo inicial de personalidades para interação.
- Validar a viabilidade técnica da solução utilizando tecnologias como Ollama para processamento de IA local e React para o desenvolvimento do front-end.
- Coletar feedbacks de usuários sobre o engajamento e a qualidade da experiência para futuras melhorias.

2. DESCRIÇÃO DA EMPRESA

A CogniConnect Labs nasceu em 2024, em São Paulo, da paixão de um grupo de engenheiros e designers de software por inteligência artificial e experiências digitais imersivas. Observando a crescente demanda por interações mais personalizadas e envolventes no ambiente digital, a equipe vislumbrou um futuro onde a comunicação com IAs transcendia a mera funcionalidade, tornando-se uma forma de entretenimento e conexão emocional.

A fundação da CogniConnect Labs foi motivada pela crença de que a IA pode ir além de assistentes genéricos, permitindo que as pessoas interajam com "ecos" digitais de personalidades que admiram ou que as inspiram. A empresa investiu inicialmente em pesquisa e desenvolvimento focado em tecnologias de linguagem natural e processamento local, buscando soluções que oferecessem privacidade e controle aos usuários, sem a necessidade de depender exclusivamente de grandes infraestruturas de nuvem.

Com o advento de ferramentas como o Ollama, que possibilitam a execução de modelos de linguagem diretamente em dispositivos locais, a CogniConnect Labs encontrou a tecnologia ideal para materializar sua visão. A empresa focou no desenvolvimento de uma plataforma que não apenas utilizasse essa capacidade de processamento local, mas que também oferecesse uma interface de usuário intuitiva e visualmente atraente, utilizando o React.js para garantir uma experiência fluida e interativa.

O EchoMind é o projeto carro-chefe da CogniConnect Labs, representando o primeiro passo da empresa em direção a um ecossistema de personalidades digitais interativas. A PoC (Prova de Conceito) do EchoMind foi desenvolvida com o objetivo de validar a viabilidade técnica e o apelo do conceito de interação com personalidades virtuais, buscando simular conversas realistas e demonstrar a customização dos personagens. A empresa está comprometida em expandir as funcionalidades do EchoMind, incluindo integração de voz, versões mobile e interações mais proativas, consolidando sua posição como líder em experiências digitais personalizadas e imersivas.

3. MISSÃO, VISÃO E VALORES DA EMPRESA

3.1 MISSÃO

- Democratizar a interação com inteligências artificiais personalizadas, oferecendo experiências digitais imersivas e envolventes que transcendam a funcionalidade e promovam o entretenimento e a conexão emocional.
- Capacitar os usuários a conversar com "ecos" digitais de personalidades que admiram, utilizando tecnologias de IA local para garantir privacidade, controle e uma experiência fluida.

3.2 VISÃO

- Ser a principal referência global em plataformas de interação com personalidades virtuais, expandindo continuamente o catálogo de personagens e as funcionalidades, como integração de voz e interações proativas, tornando-se sinônimo de inovação em entretenimento digital e comunicação baseada em IA.

3.3 VALORES

- **Inovação:** Buscar constantemente novas tecnologias e abordagens para aprimorar a interação entre humanos e IA.
- **Personalização:** Acreditar que a experiência digital deve ser única e adaptada aos interesses de cada usuário.
- **Privacidade e Controle:** Priorizar a segurança dos dados e o controle do usuário sobre suas interações digitais.
- **Engajamento e Diversão:** Criar produtos que gerem entretenimento, nostalgia e uma conexão genuína com as personalidades virtuais.
- **Acessibilidade:** Desenvolver soluções que sejam fáceis de usar e acessíveis a um público amplo, independentemente do sistema operacional ou dispositivo.

4. EchoMind

4.1 Ollama

O Ollama é uma ferramenta de código aberto que permite a execução de modelos de linguagem avançada (LLMs) diretamente em um computador, eliminando a necessidade de depender de servidores na nuvem. Isso confere maior controle sobre os dados, privacidade aprimorada e a capacidade de rodar modelos como Llama 2, Mistral, Code Llama e LLaVA localmente.

A execução local de modelos de LLMs pelo Ollama oferece vantagens como a privacidade dos dados, que não saem do seu computador, e a não necessidade de conexão com a internet após o download do modelo. Além disso, proporciona um baixo tempo de resposta, pois não depende da latência de servidores externos.

O gerenciamento de modelos no Ollama é simplificado por um sistema de linha de comando. Com um único comando, é possível baixar um modelo (ex: `ollama run llama2`), iniciar uma sessão de chat com ele, listar os modelos baixados (`ollama list`), remover um modelo (`ollama rm modelo`) ou verificar o tamanho em disco (`ollama list -v`).

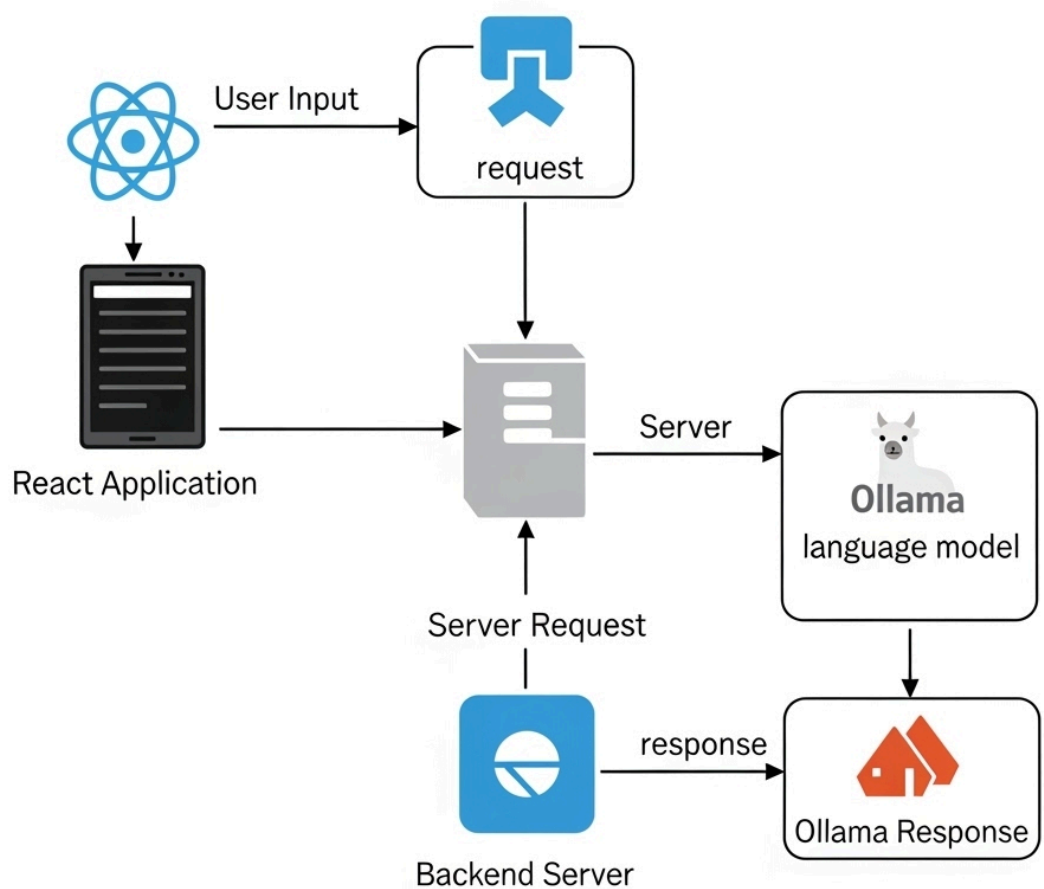
Em termos de compatibilidade, o Ollama oferece suporte multiplataforma, abrangendo macOS, Linux e Windows. Essa característica permite a fácil integração do Ollama aos fluxos de trabalho existentes, independentemente do sistema operacional utilizado. Vale ressaltar que o suporte para Windows está atualmente em fase de prévia.

O Ollama também se destaca pelo suporte a múltiplos modelos de LLMs, que podem ser utilizados prontos para uso ou personalizados para atender a requisitos específicos de projetos. Entre os modelos mais populares com suporte nativo, estão:

- **LLAMA 3(Meta):** Um modelo versátil para tarefas de Processamento de Linguagem Natural (PNL), como geração de texto, resumo e tradução automática. É uma opção popular para o desenvolvimento de chatbots e sistemas de IA conversacionais devido à sua capacidade de entender e gerar texto semelhante ao humano. Usado em nosso projeto.
- **Mistral / Mixtral:** Ideal para desenvolvedores que trabalham com plataformas de programação orientadas por IA, lidando com geração de código e análise de dados em larga escala. Seus recursos de reconhecimento de padrões permitem automatizar processos de codificação e identificar bugs.
- **Gemma (Google):** Uma família de modelos de IA generativa leves e abertos do Google, baseados na mesma pesquisa e tecnologia dos modelos Gemini. São projetados para diversas tarefas de geração, como resposta a perguntas, resumo e raciocínio.
- **Phi:** O Phi-3 é voltado para aplicações em ciência e pesquisa, sendo treinado em grandes conjuntos de dados acadêmicos, o que o torna útil para revisões de literatura, resumo de dados e análise científica.
- **Codellama:** Excelente para tarefas relacionadas à programação, como escrever e revisar códigos, automatizando fluxos de trabalho para aumentar a produtividade de desenvolvedores.

- **Nous Hermes:** Um modelo de instrução altamente refinado, construído sobre Mistral ou LLaMA.
- **OpenHermes:** Um modelo de linguagem de código aberto baseado na arquitetura LLaMA, desenvolvido pela comunidade OpenAccess AI Collective, buscando ser uma alternativa aberta a modelos proprietários como o ChatGPT ou Gemini.

4.2 Fluxo



O diagrama ilustra o fluxo de comunicação no projeto EchoMind, onde um usuário interage com um sistema de inteligência artificial que utiliza o Ollama.

O processo começa com o **"User Input"** (entrada do usuário): O usuário digita sua mensagem ou pergunta na **"React Application"** (aplicativo React), que é a interface do usuário do EchoMind.

A partir da aplicação React, essa entrada é enviada como uma **"request"** (solicitação) para um **"Server"** (servidor). O servidor atua como um intermediário, recebendo as requisições do front-end.

O servidor, então, encaminha essa solicitação ao **"Ollama language model"** (modelo de linguagem Ollama). O Ollama, que executa modelos de linguagem avançada localmente, processa a entrada do usuário e gera uma resposta.

Finalmente, a **"Ollama Response"** (resposta do Ollama) é enviada de volta ao **"Backend Server"** (servidor de back-end). O servidor, por sua vez, retransmite essa resposta para a **"React Application"**, onde o usuário pode visualizar a resposta da IA. Todo esse fluxo é projetado para ocorrer de forma contínua e em tempo real, utilizando streaming de dados para uma experiência de interação dinâmica.

4.3 Tecnologias usadas

- **Ollama:**

- Motivo da Escolha: O Ollama é uma ferramenta de código aberto que permite executar modelos de linguagem avançada (LLMs) diretamente na máquina do usuário, sem a necessidade de depender de servidores na nuvem. Isso foi crucial para o projeto EchoMind por proporcionar maior controle sobre os dados e privacidade aprimorada, além de possibilitar a execução de modelos como Llama 2 e Mistral localmente. A execução local também garante um baixo tempo de resposta, uma vez que não depende da latência de servidores externos.
- Funcionalidades Utilizadas: Gerenciamento local de modelos de IA, suporte a múltiplos modelos de LLMs, customização e criação de modelos próprios via Modfiles, e uma API local compatível com OpenAI que facilita a integração.

- **React com TypeScript:**

- Motivo da Escolha: O React.js é uma biblioteca JavaScript declarativa e baseada em componentes, criada para facilitar a construção de interfaces de usuário interativas e eficientes. Sua escolha para o front-end do EchoMind se deveu à sua capacidade de criar aplicações web escaláveis e reutilizáveis, otimizar a performance via Virtual DOM, e possuir uma comunidade ativa e um ecossistema rico.
- A adição do TypeScript (TS) ao React foi motivada pela necessidade de maior segurança e robustez em um projeto em crescimento. O TypeScript, como um superset do JavaScript que adiciona tipagem estática, ajuda a evitar erros

comuns durante o desenvolvimento, melhora o suporte em IDEs com sugestões e autocompletar mais inteligentes, aumenta a clareza sobre os dados (props e estados), e facilita a manutenção em projetos grandes ou com múltiplos desenvolvedores.

5. Próximos Passos / Evolução

Para as próximas fases do projeto, os seguintes aprimoramentos estão planejados:

- Integração de voz sintetizada com timbres similares aos personagens.
- Versão mobile para maior acessibilidade e praticidade.
- Geração de avatares baseados no usuário, usando seu estilo de escrita, voz ou dados comportamentais.
- Interações proativas dos personagens, como quizzes ou desafios.
- Treinamento personalizado (fine-tuning) com falas reais dos personagens.
- Modo educativo e filtros infantis para uso responsável por públicos jovens.
- Expansão do catálogo, incluindo personagens históricos, influenciadores e criações da comunidade.

6. DEFINIÇÃO DA EQUIPE, DIVISÃO DE PAPÉIS E TAREFAS, E CRONOGRAMA DE DESENVOLVIMENTO.

Igor Alves da Silva Pereira (419202553) - Desenvolvedor Full-stack Líder (Front-end & Back-end)

- **Função:** Arquiteto e principal desenvolvedor de todo o sistema EchoMind, abrangendo desde a interface do usuário no front-end até a lógica de inteligência artificial e a comunicação com o Ollama no back-end. Responsável pela integração e otimização de todas as camadas do projeto.
- **Tarefas:**

- a. Desenvolvimento e otimização dos Modfiles para personalização de personagens.
- b. Implementação da API de comunicação entre o Ollama e o front-end.
- c. Garantia da coerência e persistência do contexto da conversa.
- d. Desenvolvimento dos componentes React para a interface de chat.
- e. Implementação da exibição de respostas via streaming.
- f. Monitoramento do desempenho e otimização do uso de recursos (CPU/GPU) do Ollama.

- **Definição da Metodologia:** Inicie explicando a abordagem geral que será adotada no projeto, como qualitativa, quantitativa ou mista. Descreva também se o estudo é experimental, exploratório, descritivo, correlacional, ou outro tipo. **Exemplo:** "Este projeto utiliza uma abordagem qualitativa para analisar [tema do projeto], com foco em entrevistas semiestruturadas e análise de conteúdo."
- **Ferramentas e Tecnologias:** Liste as ferramentas e tecnologias que serão utilizadas durante o projeto. Isso pode incluir softwares, equipamentos ou plataformas específicas. Exemplo: VSCode, linguagem Java, equipamentos, etc.
- **Cronograma de Atividades:** Inclua um cronograma que apresenta as etapas do projeto ao longo do tempo, especificando prazos e prioridades.

Antonio Augusto Lopes Marianno (324102840) - Arquiteto de Soluções & Padrões

- **Função:** Responsável pela arquitetura geral do sistema, definindo padrões de desenvolvimento, garantindo a escalabilidade e segurança do projeto. Atua na pesquisa e implementação de melhores práticas.
- **Tarefas:**
 - Definição das melhores práticas de desenvolvimento e padrões de código para o projeto.
 - Revisão de código para garantir qualidade e aderência aos padrões.
 - Avaliação e integração de novas tecnologias e frameworks.
 - Planejamento da escalabilidade futura do EchoMind.

Felipe Trivia Koshimizu (923205456) - Especialista em Pesquisa & Otimização de Modelos de IA

- **Função:** Focado na pesquisa e seleção dos modelos de linguagem mais adequados para o Ollama e na otimização de seu desempenho. Atua na exploração de técnicas de quantização e customização de modelos.
- **Tarefas:**
 - Pesquisa e seleção de modelos de linguagem (LLMs) da biblioteca Ollama.
 - Exploração de opções de quantização para modelos (Q4, Q5, Q8) para otimização de recursos.
 - Estudo de customização e criação de modelos próprios (Modfiles).
 - Análise de desempenho dos modelos em diferentes configurações de hardware.

Felipe Norberto Malveira (2221201409) - Analista de Requisitos & Design de Interação

- **Função:** Responsável pela coleta e análise dos requisitos do projeto, garantindo que o produto final atenda às necessidades dos usuários. Atua no design da experiência do usuário (UX) e na definição das interações com a IA.
- **Tarefas:**
 - Levantamento de requisitos junto aos stakeholders.
 - Definição das características e funcionalidades de interação com as personalidades virtuais.
 - Criação de fluxos de usuário e wireframes para o sistema.
 - Análise de feedback de usuários para identificar oportunidades de melhoria na interação.

Renan Nunes Marques dos Santos (421202108) - Analista de Qualidade (QA) & Testes de Padrões

- **Função:** Responsável por garantir a qualidade do software, identificando bugs e inconsistências. Focado na validação da aderência aos padrões de desenvolvimento e na coerência das respostas da IA.
- **Tarefas:**
 - Criação e execução de planos de teste funcionais e não-funcionais.

- Realização de testes de integração e validação da comunicação via streaming.
- Verificação da coerência das respostas da IA com as personalidades definidas.
- Relato e acompanhamento de bugs.
- Garantia da conformidade do produto com os requisitos e padrões estabelecidos.

Renan de Oliveira Silva (421201286) - Gerente de Projeto & Documentação

- **Função:** Responsável pelo planejamento, execução e monitoramento do projeto. Gestão do cronograma, recursos e riscos. Garante a comunicação eficaz e a documentação completa de todas as fases.
- **Tarefas:**
 - Definição e acompanhamento do cronograma do projeto.
 - Gestão das tarefas e delegação de responsabilidades.
 - Condução de reuniões de acompanhamento e status.
 - Identificação e mitigação de riscos.
 - Elaboração e manutenção de toda a documentação do projeto, incluindo requisitos, especificações e manuais.

Juliano Cesar Thome Marin (320102190) - Especialista em Integração & Ferramenta

- **Função:** Focado na pesquisa e implementação de ferramentas e bibliotecas que facilitem o desenvolvimento e a integração de componentes. Atua na otimização de processos e na automação de tarefas.
- **Tarefas:**
- Pesquisa de ferramentas e bibliotecas para otimizar o desenvolvimento (ex: LangChain, LlamaIndex).
- Configuração de ambientes de desenvolvimento e teste.
- Automação de processos de build e deploy.
- Suporte à equipe com problemas de integração e configuração de ferramentas.

Luan de Oliveira Ramos (3021201889) - Analista de Conteúdo & Prompt Engineering

- **Função:** Colabora na criação e refinamento dos "prompts de sistema" para moldar o comportamento e a personalidade das IAs. Responsável por garantir que as interações sejam autênticas e envolventes.

- **Tarefas:**

- Criação e teste de prompts para cada personagem.
- Curadoria e expansão do catálogo de personalidades.
- Análise e aprimoramento das interações da IA.
- Garantia da qualidade e adequação do conteúdo gerado pela IA.

A equipe da CogniConnect Labs é composta por talentos diversos, com especialidades que abrangem desde o desenvolvimento de IA e front-end até a gestão de projetos e experiência do usuário. A seguir, detalhamos os componentes da equipe, seus papéis e tarefas, e um cronograma de desenvolvimento para o projeto EchoMind.

Componentes da Equipe e Divisão de Papéis e Tarefas:

Igor Alves da Silva Pereira (419202553) - Desenvolvedor Full-stack Líder (Front-end & Back-end)

- **Função:** Arquiteto e principal desenvolvedor de todo o sistema EchoMind, abrangendo desde a interface do usuário no front-end até a lógica de inteligência artificial e a comunicação com o Ollama no back-end. Responsável pela integração e otimização de todas as camadas do projeto.
- **Tarefas:**
 - Desenvolvimento e otimização dos Modfiles para personalização de personagens.
 - Implementação da API de comunicação entre o Ollama e o front-end.
 - Garantia da coerência e persistência do contexto da conversa.
 - Desenvolvimento dos componentes React para a interface de chat.
 - Implementação da exibição de respostas via streaming.
 - Monitoramento do desempenho e otimização do uso de recursos (CPU/GPU) do Ollama.

Antonio Augusto Lopes Marianno (324102840) - Arquiteto de Soluções & Padrões

- **Função:** Responsável pela arquitetura geral do sistema, definindo padrões de desenvolvimento, garantindo a escalabilidade e segurança do projeto. Atua na pesquisa e implementação de melhores práticas.
- **Tarefas:**

- Definição das melhores práticas de desenvolvimento e padrões de código para o projeto.
- Revisão de código para garantir qualidade e aderência aos padrões.
- Avaliação e integração de novas tecnologias e frameworks.
- Planejamento da escalabilidade futura do EchoMind.

Felipe Trivia Koshimizu (923205456) - Especialista em Pesquisa & Otimização de Modelos de IA

- **Função:** Focado na pesquisa e seleção dos modelos de linguagem mais adequados para o Ollama e na otimização de seu desempenho. Atua na exploração de técnicas de quantização e customização de modelos.
- **Tarefas:**
 - Pesquisa e seleção de modelos de linguagem (LLMs) da biblioteca Ollama.
 - Exploração de opções de quantização para modelos (Q4, Q5, Q8) para otimização de recursos.
 - Estudo de customização e criação de modelos próprios (Modfiles).
 - Análise de desempenho dos modelos em diferentes configurações de hardware.

Felipe Norberto Malveira (2221201409) - Analista de Requisitos & Design de Interação

- **Função:** Responsável pela coleta e análise dos requisitos do projeto, garantindo que o produto final atenda às necessidades dos usuários. Atua no design da experiência do usuário (UX) e na definição das interações com a IA.
- **Tarefas:**
 - Levantamento de requisitos junto aos stakeholders.
 - Definição das características e funcionalidades de interação com as personalidades virtuais.
 - Criação de fluxos de usuário e wireframes para o sistema.
 - Análise de feedback de usuários para identificar oportunidades de melhoria na interação.

Renan Nunes Marques dos Santos (421202108) - Analista de Qualidade (QA) & Testes de Padrões

- **Função:** Responsável por garantir a qualidade do software, identificando bugs e inconsistências. Focado na validação da aderência aos padrões de desenvolvimento e na coerência das respostas da IA.
- **Tarefas:**
 - Criação e execução de planos de teste funcionais e não-funcionais.
 - Realização de testes de integração e validação da comunicação via streaming.
 - Verificação da coerência das respostas da IA com as personalidades definidas.
 - Relato e acompanhamento de bugs.
 - Garantia da conformidade do produto com os requisitos e padrões estabelecidos.

Renan de Oliveira Silva (421201286) - Gerente de Projeto & Documentação

- **Função:** Responsável pelo planejamento, execução e monitoramento do projeto. Gestão do cronograma, recursos e riscos. Garante a comunicação eficaz e a documentação completa de todas as fases.
- **Tarefas:**
 - Definição e acompanhamento do cronograma do projeto.
 - Gestão das tarefas e delegação de responsabilidades.
 - Condução de reuniões de acompanhamento e status.
 - Identificação e mitigação de riscos.
 - Elaboração e manutenção de toda a documentação do projeto, incluindo requisitos, especificações e manuais.

Juliano Cesar Thome Marin (320102190) - Especialista em Integração & Ferramentas

- **Função:** Focado na pesquisa e implementação de ferramentas e bibliotecas que facilitem o desenvolvimento e a integração de componentes. Atua na otimização de processos e na automação de tarefas.
- **Tarefas:**
 - Pesquisa de ferramentas e bibliotecas para otimizar o desenvolvimento (ex: LangChain, LlamaIndex).
 - Configuração de ambientes de desenvolvimento e teste.
 - Automação de processos de build e deploy.
 - Suporte à equipe com problemas de integração e configuração de ferramentas.

Luan de Oliveira Ramos (3021201889) - Analista de Conteúdo & Prompt Engineering

- **Função:** Colabora na criação e refinamento dos "prompts de sistema" para moldar o comportamento e a personalidade das IAs. Responsável por garantir que as interações sejam autênticas e envolventes.
- **Tarefas:**
 - Criação e teste de prompts para cada personagem.
 - Curadoria e expansão do catálogo de personalidades.
 - Análise e aprimoramento das interações da IA.
 - Garantia da qualidade e adequação do conteúdo gerado pela IA.

Davi Anio Gabriel Brandino Ribeiro (2221200912) - Analista de Suporte ao Produto & Feedback

- **Função:** Responsável por coletar e analisar o feedback dos usuários, transformando-o em insights para melhorias contínuas. Atua no suporte ao produto e na comunicação com a comunidade de usuários.
- **Tarefas:**
 - Gerenciamento de canais de feedback (ex: formulários, redes sociais).
 - Análise e categorização do feedback dos usuários.
 - Criação de relatórios de feedback para a equipe de desenvolvimento.
 - Fornecimento de suporte e respostas a dúvidas dos usuários.
 - Identificação de tendências e necessidades emergentes dos usuários.

Douglas de Sousa Rodrigues (424110441) - Designer Gráfico & UI/UX

- **Função:** Responsável pela criação da identidade visual do EchoMind e pelo design da experiência do usuário (UX). Garante que o sistema seja visualmente atraente, intuitivo e fácil de usar.
- **Tarefas:**
 - Criação de wireframes, protótipos e mockups da interface.
 - Definição da paleta de cores, tipografia e elementos visuais.
 - Criação de ícones e ilustrações para o sistema.
 - Garantia da consistência visual em toda a aplicação.

- Colaboração com o Igor para implementação do design no front-end.

7. Cronograma de Desenvolvimento

Fase 1: Concepção e Planejamento (28 de Março - 10 de Abril)

- **Semanas 1-2 (28 de Março - 10 de Abril):**
 - Definição da visão, missão e valores da CogniConnect Labs.
 - Brainstorming e definição do conceito inicial do EchoMind.
 - Análise de requisitos e elaboração da Prova de Conceito (PoC) (Renan D.A. Ribeiro, Felipe N. Malveira).
 - Design da arquitetura inicial do sistema (Antonio Augusto).
 - Definição da stack tecnológica (Ollama, React) (Igor, Antonio Augusto).
 - Criação da documentação detalhada da ideia do projeto e seus objetivos (Renan de O. Silva).

Fase 2: Desenvolvimento da PoC (11 de Abril - 1 de Maio)

- **Semanas 3-4 (11 de Abril - 24 de Abril):**
 - Configuração do ambiente de desenvolvimento (Ollama, Node.js, React) (Igor, Juliano C. Thome Marin).
 - Desenvolvimento do core da integração Ollama-React (comunicação via streaming) (Igor).
 - Criação da primeira versão da interface de chat básica (Igor, Douglas S. Rodrigues).
 - Testes iniciais de comunicação e resposta da IA (Renan N. Marques).
- **Semanas 5-8 (25 de Abril - 1 de Maio):**

- Desenvolvimento dos primeiros perfis de personalidade e prompts (Igor, Luan de O. Ramos, Felipe N. Malveira).
- Refinamento da interface de usuário com base em feedbacks iniciais (Igor, Douglas S. Rodrigues).
- Implementação de persistência de contexto em sessões de chat (Igor).
- Testes de funcionalidade e usabilidade da PoC (Renan N. Marques).
- Apresentação interna da PoC e coleta de feedbacks da equipe.

Fase 3: Aprimoramento e Expansão (2 de Maio - 16 de Maio)

- **Semanas 9-12 (2 de Maio - 16 de Maio):**

- Melhorias de desempenho e otimização do uso de recursos do Ollama (Igor, Felipe T. Koshimizu).
- Expansão do catálogo de personagens e aprimoramento dos prompts (Luan de O. Ramos, Felipe N. Malveira).
- Início da pesquisa e prototipagem para integração de voz sintetizada (Felipe T. Koshimizu).
- Estudo de viabilidade para versão mobile (Antonio Augusto).

Fase 4: Implantação, Testes e Operação (17 de Maio - 30 de Maio)

- **Semanas 13-16 (17 de Maio - 30 de Maio):**

- Desenvolvimento de funcionalidades de personalização de avatar (se viável na fase) (Igor, Douglas S. Rodrigues).
- Início do planejamento para interações proativas dos personagens (Felipe N. Malveira, Davi A. G. B. Ribeiro).
- Testes de integração das novas funcionalidades (Renan N. Marques).
- Pesquisa de filtros de conteúdo e modo educativo (Igor, Luan de O. Ramos).
- Refinamento da experiência do usuário com base nos testes (Davi A. G. B. Ribeiro, Douglas S. Rodrigues).
- Preparação para a primeira fase de testes externos (alpha/beta).

CONCLUSÃO

O projeto EchoMind, desenvolvido pela CogniConnect Labs, representa um avanço significativo na interação entre usuários e inteligências artificiais personalizadas. Ao utilizar o **Ollama** para processamento de IA local, o projeto garante maior controle sobre os dados e privacidade, além de otimizar a performance. A escolha do **React.js** para o front-end assegura uma interface de usuário intuitiva e responsiva, com a comunicação via streaming proporcionando uma experiência de conversação fluida e em tempo real.

A equipe multidisciplinar da CogniConnect Labs, com seus papéis e tarefas bem definidos – desde o desenvolvimento full-stack liderado por Igor, passando pela arquitetura e padrões, pesquisa de modelos de IA, design de interação, qualidade, gestão de projetos, até a análise de conteúdo e suporte ao produto – é fundamental para o sucesso do EchoMind. A metodologia ágil e o cronograma detalhado, embora ambiciosos para o período de 28 de março a 30 de maio de 2025, demonstram o compromisso com a entrega de valor contínuo e a adaptação às necessidades do usuário.

A Prova de Conceito (PoC) já validou a viabilidade técnica e o apelo do EchoMind em simular interações realistas com personalidades virtuais. As futuras etapas, que incluem a integração de voz, versões mobile e interações mais proativas, prometem expandir ainda mais as possibilidades do projeto. O EchoMind se posiciona como uma plataforma inovadora, capaz de transformar a maneira como interagimos com a IA, oferecendo entretenimento, nostalgia e uma nova forma de conexão digital.

REFERÊNCIAS

OLLAMA. *Run large language models locally*. Disponível em: <https://ollama.com>. Acesso em: 28 maio 2025.

META AI. *Introducing LLaMA: Open and Efficient Foundation Language Models*. 2023. Disponível em: <https://ai.meta.com/llama/>. Acesso em: 28 maio 2025.

GOOGLE. *Gemma: Open models built from the same research and technology as Gemini*. 2024. Disponível em: <https://ai.google.dev/gemma>. Acesso em: 28 maio 2025.

MICROSOFT. *Introducing Phi-3: the next generation of small language models*. 2024. Disponível em: <https://www.microsoft.com/en-us/research/blog/introducing-phi-3/>. Acesso em: 28 maio 2025.

OPENACCESS AI COLLECTIVE. *OpenHermes – an open-source instruction-tuned model*. 2024. Disponível em: <https://huggingface.co/openaccess-ai-collective>. Acesso em: 28 maio 2025.

FACEBOOK. *React – A JavaScript library for building user interfaces*. 2013. Disponível em: <https://reactjs.org>. Acesso em: 28 maio 2025.

MICROSOFT. *TypeScript: JavaScript with syntax for types*. Disponível em: <https://www.typescriptlang.org>. Acesso em: 28 maio 2025.

GITHUB. *EchoMind Project – igorcalves/EchoMind*. [S.l.]: GitHub, 2025. Disponível em: <https://github.com/igorcalves/EchoMind?tab=readme-ov-file>. Acesso em: 28 maio 2025.