## **Resumo Técnico da Solução: Historinhas do Gemini**

Este documento descreve a arquitetura técnica da solução "Historinhas do Gemini", uma plataforma para a criação e distribuição de histórias multiculturais para crianças.

**1. Visão Geral da Arquitetura**

A solução é centrada em uma arquitetura de agentes, onde cada agente é responsável por uma etapa específica do processo de criação da história. Os agentes se comunicam para gerar o conteúdo final, que é então publicado.

A arquitetura completa da solução "Historinhas do Gemini" é composta pelos seguintes agentes:

* **Agenda Decider Agent:** Este agente é responsável por definir a agenda da história, incluindo o país foco, o valor humano central e outros elementos contextuais.
* **Writer Agent:** Este agente usa a agenda fornecida pelo Agenda Decider Agent para escrever a história propriamente dita.
* **Scene Descriptor Agent:** Este agente recebe a história do Writer Agent e gera descrições detalhadas das cenas para a criação das ilustrações.
* **Illustrator Agent:** Este agente gera as ilustrações para a história, com base nas descrições fornecidas pelo Scene Descriptor Agent.
* **Translator Agents:** Estes agentes traduzem a história para os idiomas desejados.
* **Audio Generator Agents:** Estes agentes geram a narração em áudio da história em cada idioma.
* **Publisher Agent:** Este agente formata o texto, as ilustrações e o áudio para publicação nas plataformas de distribuição (inicialmente, o Instagram).

Os agentes Illustrator, Translator e Audio Generator operam em paralelo, acelerando o processo de produção. O Publisher Agent formata o conteúdo para a plataforma de destino.

**2. Tecnologias Utilizadas**

* **Agent Development Kit (ADK):** O Google ADK é utilizado como framework principal para o desenvolvimento e orquestração dos agentes. O ADK fornece as ferramentas e abstrações necessárias para definir o comportamento dos agentes, gerenciar a comunicação entre eles e executar o fluxo de trabalho geral.
* **Modelos de Linguagem Generativos:** Modelos de linguagem generativos, como o Gemini, são usados para a criação do conteúdo textual das histórias. Esses modelos são capazes de gerar texto criativo e envolvente, adequado para o público-alvo.
* **APIs de Busca:** APIs de busca, como a Pesquisa Google, são usadas para coletar informações contextuais relevantes para as histórias, como eventos históricos ou detalhes culturais de um país específico.
* **Modelos de Geração de Imagens:** Modelos de geração de imagens serão integrados para permitir ao agente Illustrator gerar ilustrações a partir das descrições textuais.
* **Tecnologias de Síntese de Voz:** Tecnologias de síntese de voz serão utilizadas para gerar a narração em áudio das histórias em diferentes idiomas.

**3. Prova de Conceito (PoC)**

A Prova de Conceito (PoC) inicial se concentrou em um subconjunto de agentes para demonstrar a viabilidade da arquitetura e do fluxo de trabalho. Os seguintes agentes foram implementados e testados no Google Colab:

* **Agenda Decider Agent:** Este agente é responsável por definir a agenda da história, incluindo o país foco, o valor humano central e outros elementos contextuais.
* **Writer Agent:** Este agente usa a agenda fornecida pelo Agenda Decider Agent para escrever a história propriamente dita.
* **Scene Descriptor Agent:** Este agente recebe a história do Writer Agent e gera descrições detalhadas das cenas para a criação das ilustrações.

**4. Fluxo de Trabalho Detalhado**

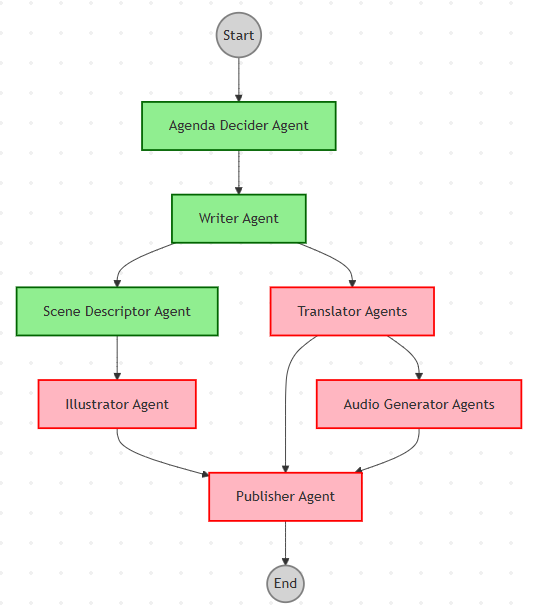
O fluxo de trabalho geral da PoC é o seguinte:

1. O **Agenda Decider Agent** é invocado com uma data específica.
2. O Agenda Decider Agent usa a Pesquisa Google para encontrar eventos globais significativos que ocorreram nessa data e decide o país foco e o valor humano central da história.
3. O Agenda Decider Agent gera uma agenda em formato de texto e a passa para o **Writer Agent**.
4. O Writer Agent usa a agenda para gerar uma história para crianças e a passa para o **Scene Descriptor Agent**.
5. O Scene Descriptor Agent usa a história e a agenda para gerar descrições detalhadas das cenas em formato de texto, que serão usadas para criar as ilustrações.
6. As descrições das cenas são retornadas como resultado do fluxo de trabalho.

**5. Arquitetura Futura**

No futuro, a arquitetura completa da solução envolverá a interação dos sete agentes descritos na seção 1. O fluxo de trabalho expandido será o seguinte:

1. O **Agenda Decider Agent** define a agenda da história.
2. O **Writer Agent** escreve a história com base na agenda.
3. O **Scene Descriptor Agent** descreve as cenas da história.
4. O **Illustrator Agent** gera as ilustrações.
5. Os **Translator Agents** traduzem a história para diferentes idiomas.
6. Os **Audio Generator Agents** criam narrações em áudio para cada idioma.
7. O **Publisher Agent** formata e publica a história completa (texto, imagens e áudio) nas plataformas de distribuição.



**6. Próximos Passos**

Os próximos passos incluem:

* Implementação dos agentes restantes (Illustrator Agent, Translator Agents, Audio Generator Agents e Publisher Agent).
* Integração com modelos de linguagem generativos e outras APIs.
* Desenvolvimento de ferramentas para a criação e gerenciamento de conteúdo.
* Integração com plataformas de distribuição (Instagram e outras).
* Otimização do fluxo de trabalho para desempenho e escalabilidade.
* Adequação dos modelos usados nos agentes, de acordo com a tarefa a ser executada por cada um deles (no momento, há limitações no acesso a modelos mais avançados, dados os baixos privilégios da conta gratuita usada para o PoC - em que apenas o modelo gemini-2.0-flash foi utilizado)