

Mecanismos de Automação usando SLMs

Docente: Hugo Ribeiro Belonia Gonçalves

Orientador: Prof. Brivaldo Alves da Silva Jr

Introdução

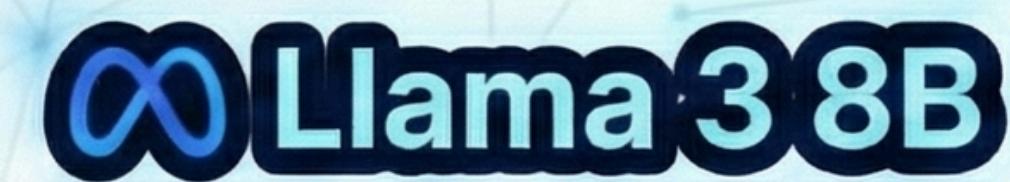


O impacto dos Grandes Modelos de Linguagem (LLMs)



Desafios Críticos e Limitações dos LLMs Atuais:

- **Altos Custos Computacionais**
- **Impacto Ambiental e Consumo Energético**
- **Dependência de Fornecedor**
- **Inviabilidade em Cenários de Borda**
 - Latência alta, dependência de conectividade e restrições de *hardware* local.



O que são Small Language Models (SLMs)?

- Modelos de Linguagem Otimizados, projetados com um número reduzido de parâmetros para garantir eficiência sem exigir supercomputadores.
- O Conceito é flexível

Vantagens dos SLMs

- Execução Local
- Baixa Latência
- Redução de Custos
- Privacidade e Segurança

Desafios e Limitações Técnicas

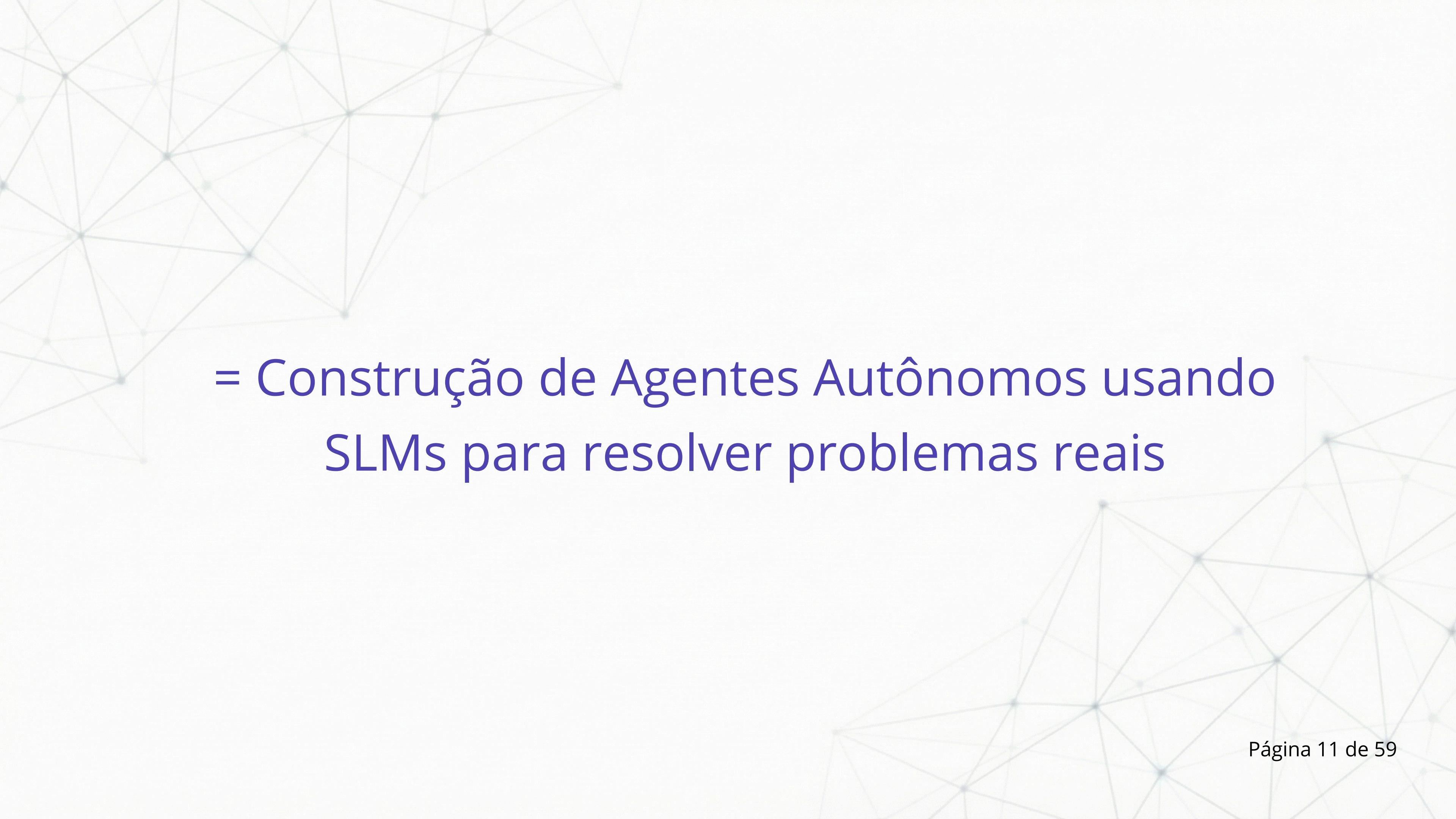
- Limitação de Raciocínio Profundo
- Alucinações e Vieses

Definições e Restrições do Escopo do Projeto

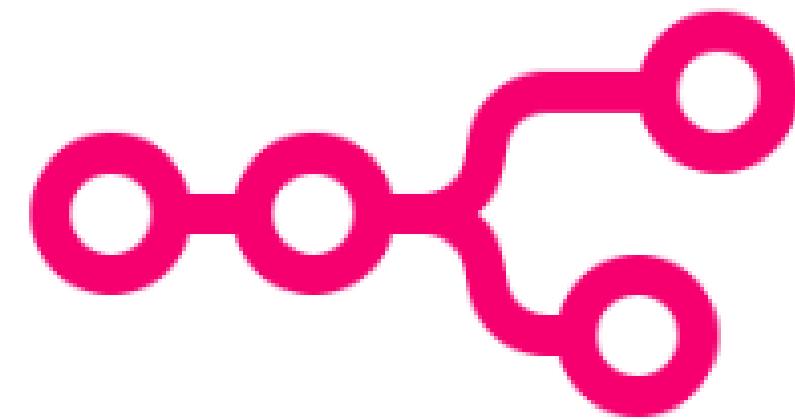
- O Desafio da Multilinguagem
 - (Llama, Qwen e Phi)
- Limitações de *Hardware* (4GB de VRAM)
 - (0.5 a 3B de parâmetros)



Limitações dos LLMs
+
Baixo Custo dos SLMs
+
Limitações de *Hardware*
+
Conceito de Agentes de IA
+
A maioria dos estudos concentra-se no refinamento do
modelo isolado



= Construção de Agentes Autônomos usando
SLMs para resolver problemas reais



n 8 n +





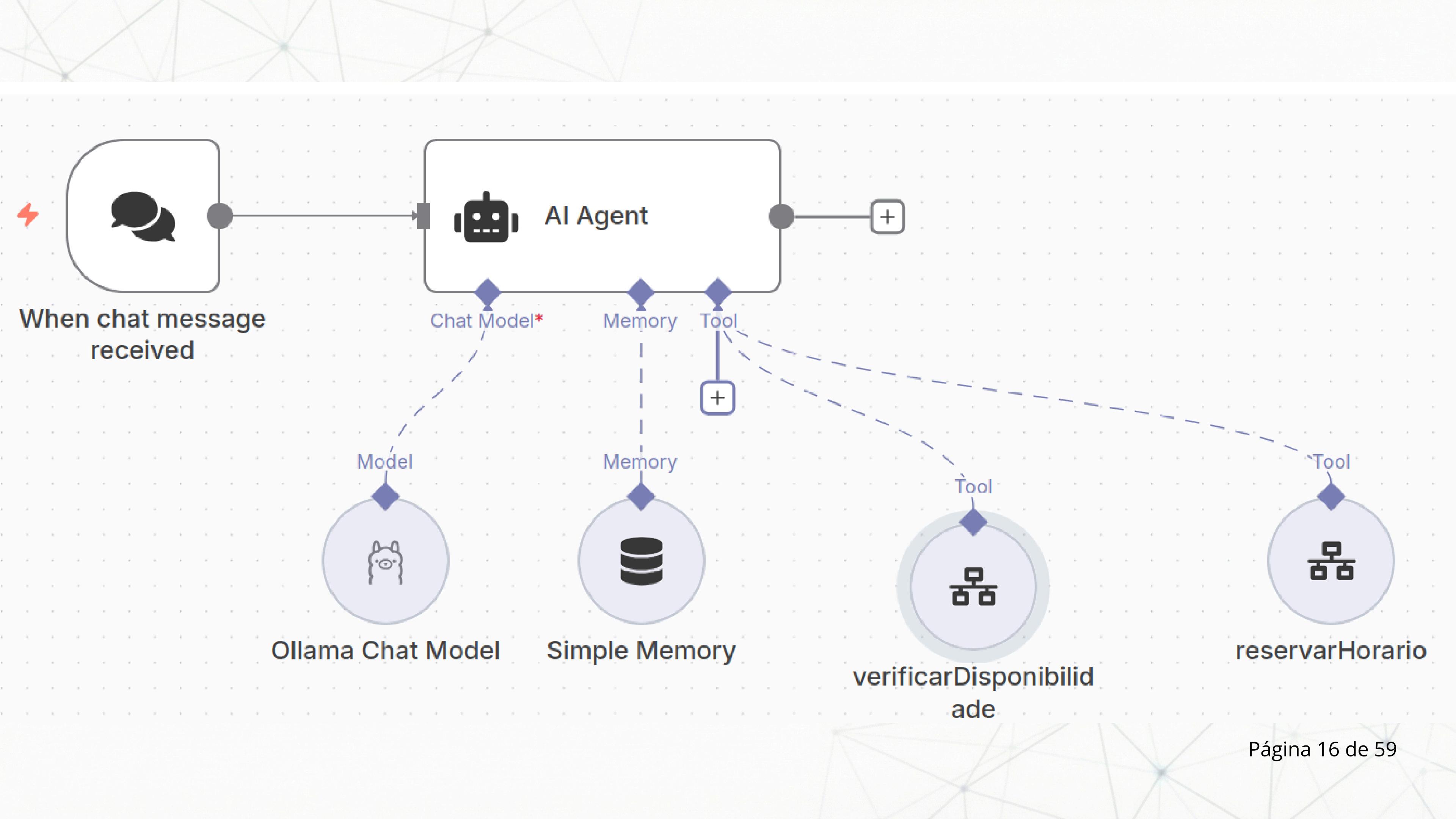
Implementações

Os Casos de Uso

- Agendamento de Horários de Reunião
- Lista de Tarefas Seguindo Regras (Pomodoro)
- Ler Documentos em Busca de Padrões de Interesse
- Resposta Automatizada de Tarefas de Alunos

Agendamento de Horários de Reunião

- **Objetivo:** Permitir agendamentos em linguagem natural
- **Abordagem:** Ao invés de um "super prompt", dividir (os dados) para conquistar





PASSOS DA CONVERSA:

1. Inicie com a seguinte frase (sem adicionar nada a mais):

> Olá! Eu sou João, assistente do professor Brivaldo.
Você gostaria de marcar um horário de reunião?

2. Aguarde a resposta.

Se "não", diga:

> Certo, fico à disposição caso precise. Tenha um ótimo dia!

Se "sim", continue.

3. Use a função `verificarDisponibilidade`. você DEVE mostrar os horários disponíveis e pergunte qual o aluno prefere.

4. Quando o aluno escolher um horário, use `verificarDisponibilidade` novamente para confirmar se ele ainda está livre.



5. Quando o aluno informar os dados, confirme tudo assim

- > Por favor, confirme as informações:
- > Nome: [nome]
- > E-mail: [email]
- > Assunto: [assunto]
- > Data: [AAAA-MM-DD]
- > Hora: [HH:MM]
- > As informações estão corretas?

6. Quando o aluno disser: “Sim, as informações estão corretas.”, ****você DEVE executar****:

```
```js
reservarHorario({
 nome: "[NOME]",
 email: "[EMAIL]",
 data: "[AAAA-MM-DD]",
 hora: "[HH:MM]",
 assunto: "[ASSUNTO]"
})
```

```



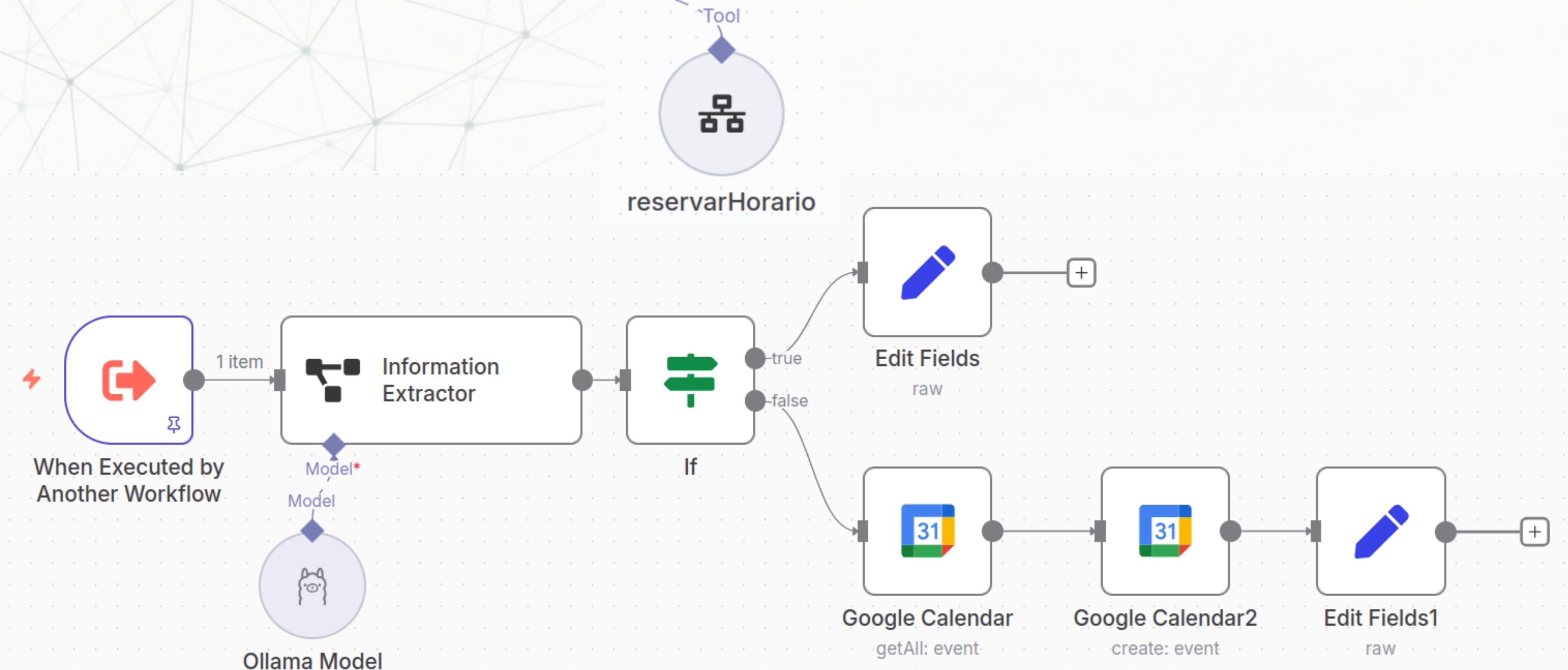
AI Agent

7. Se a função retornar `true`, diga:

> Pronto! Sua reunião foi agendada com sucesso. O professor Brivaldo estará esperando por você. Qualquer dúvida, estou à disposição. Até logo!

Atenção:

- Nunca pule nenhum passo.
- Só execute `reservarHorario` após o aluno confirmar os dados.
- Nunca execute `reservarHorario` de forma antecipada ou automática.







JSON

```
{  
  "nome": "hugo ribeiro",  
  "email": " HUGO.BELONIA@UFMS.BR ",  
  "data": "09-dez-2025",  
  "hora": "16h",  
  "assunto": "tcc"  
}
```

```
{  
  "nome": "Hugo Ribeiro",  
  "email": "hugo.belonia@ufms.br",  
  "data": "2025-12-09",  
  "hora": "16:00:00",  
  "assunto": "Reunião sobre TCC"  
}
```

```
{  
  "nome": "hugo ribeiro",  
  "email": " HUGO.BELONIA@UFMS.BR ",  
  "data": "09 dezembro 2025",  
  "hora": "16h",  
  "assunto": "discussão TCC"  
}
```

Tool



reservarHorario

reservarHorario Success in 32.195s

Input Output ... ▾

| INPUT | 1 item | OUTPUT | 1 item |
|--|--|--|--------|
| <input type="button" value="query"/> query | | <input type="button" value="resultado"/> resultado | true |
| <input type="button" value="dados"/> dados | {"nome": "hugo ribeiro", "email": "hugo.belonia@ufms.br", "data": "09/12/2025", "hora": "16:00", "assunto": "tcc"} | | |

The image shows a screenshot of a workflow management system. On the left, there is a circular icon for a tool named 'reservarHorario', which contains a 'Tool' icon (a blue diamond) and a 'Variables and context' icon (a stack of three squares). Below this is a section for an 'Information Extractor', which features a camera icon and the text 'Information Extractor'.

reservarHorario

Information Extractor

INPUT

When Executed by Another Workflow ↗ 1 item

AB | dados {"nome": "hugo ribeiro", "email": "hugo.belonia@ufms.br", "data": "09/12/2025", "hora": "16:00", "assunto": "tcc"}

Variables and context

Output Logs 1 item [{ "output": { "nome": "hugo ribeiro", "email": "hugo.belonia@ufms.br", "assunto": "tcc", "datahora": "2025-12-09T16:00:00.000-03:00" } }]

Schema Table JSON

Página 25 de 59

Hoje



9 de dezembro de 2025



Dia



TER.

9

GMT-04

1 PM

2 PM

3 PM

4 PM

(Sem título)
4 – 5pm

5 PM

6 PM

7 PM

8 PM

9 PM

10 PM

11 PM

reuniao agendada com hugo ribeiro,
assunto: tcc

Terça-feira, 9 de dezembro • 4:00 – 5:00pm

[Criar ata da reunião](#)
Inicie um novo documento para fazer anotações

reuniao agendada. email do aluno:
hugo.belonia@ufms.br

Horários de Reunião
Criado por: HUGO RIBEIRO BELONIA GONÇALVES

Agente



Olá! Eu sou João, assistente do professor Brivaldo. Você gostaria de marcar um horário de reunião?

Usuário

gostaria de marcar um reunião

Usuário

sim

Agente



Quais desses horários você gostaria?

[...]

2025-12-16T16:00:00-04:00

Usuário

2025-12-16T16:00:00-04:00

Agente



Ótimo! [...]

Por favor, me informe:
Nome completo, E-mail, Assunto...

Usuário

hugo belonia, hugo.belonia@ufms.br, reunião sobre TCC

Agente



Por favor, confirme as informações:

[...]

As informações estão corretas?

Usuário

sim

Destaque:

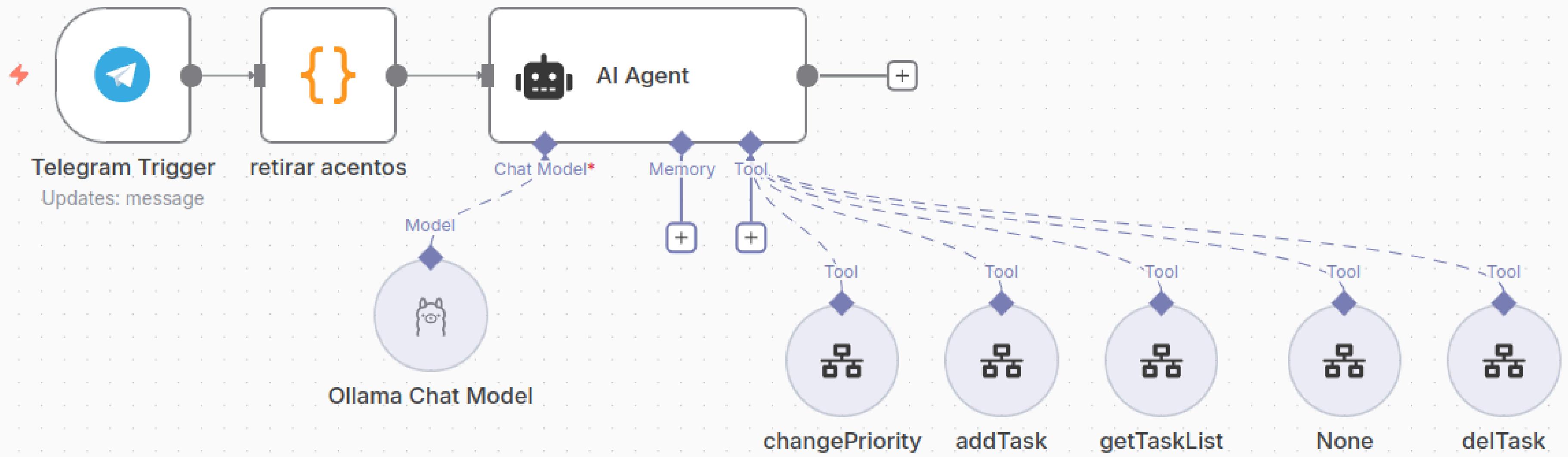
- qwen2.5:3B

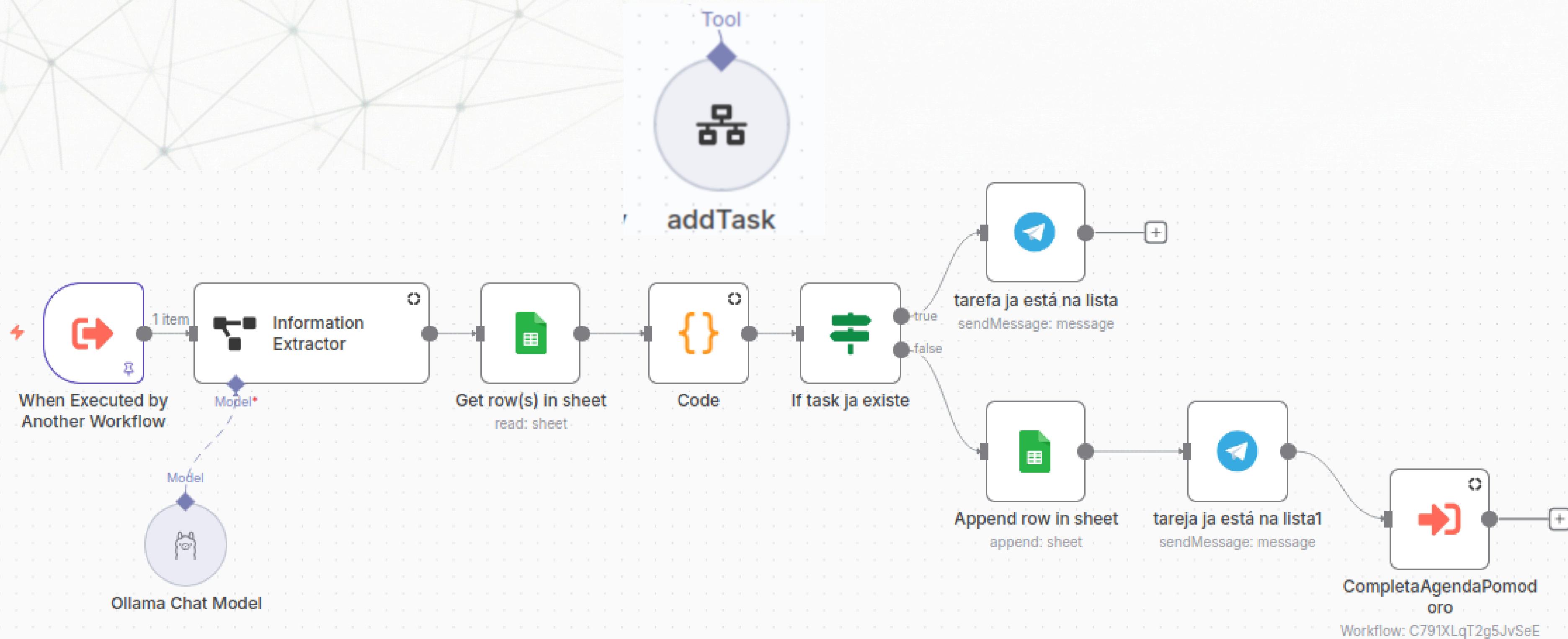
Descoberta :

- O sucesso com instruções imperativas explícitas no *System Prompt*.

Lista de Tarefas Seguindo Regras (Pomodoro)

- **Objetivo:** Receber, priorizar e alocar atividades na agenda seguindo a técnica Pomodoro
- **Abordagem:** Agente de roteamento.





INPUT Q Schema Table JSON

When Executed by Another Workflow 1 item

AB query jogar futebol com meus amigos urgentemente

Variables and context

Output Logs ⓘ ⓘ Q Schema Table JSON

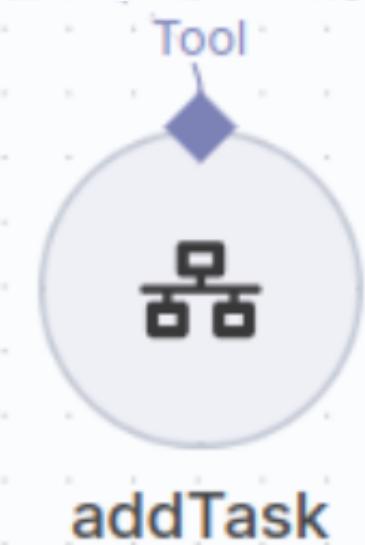
This node will output an empty item if nothing would normally be returned

1 item

```
[{"output": {"tarefa": "jogar futebol com meus amigos urgentemente"}]
```



Information Extractor





AI Agent

Today

Quero corrigir as provas dos meus alunos

19:58 ✓✓

A tarefa corrigir as provas dos meus alunos foi adicionada a lista de tarefas

This message was sent automatically with n8n

20:00

preciso encontrar com os amigos para jogar futebol o quanto antes

20:16 ✓✓

Quando você gostaria de encontrar tempo para jogar futebol?

This message was sent automatically with n8n

20:18

Preciso jogar futebol com meus amigos urgentemente

20:19 ✓✓

A tarefa jogar futebol com meus amigos urgentemente foi adicionada a lista de tarefas

This message was sent automatically with n8n

20:20



AI Agent

Today

Quero corrigir as provas dos meus alunos 19:58 ✓

A tarefa corrigir as provas dos meus alunos foi adicionada a lista de tarefas

This message was sent automatically with n8n

20:00

preciso encontrar com os amigos para jogar futebol o quanto antes 20:16 ✓

Quando você gostaria de encontrar tempo para jogar futebol?

This message was sent automatically with n8n

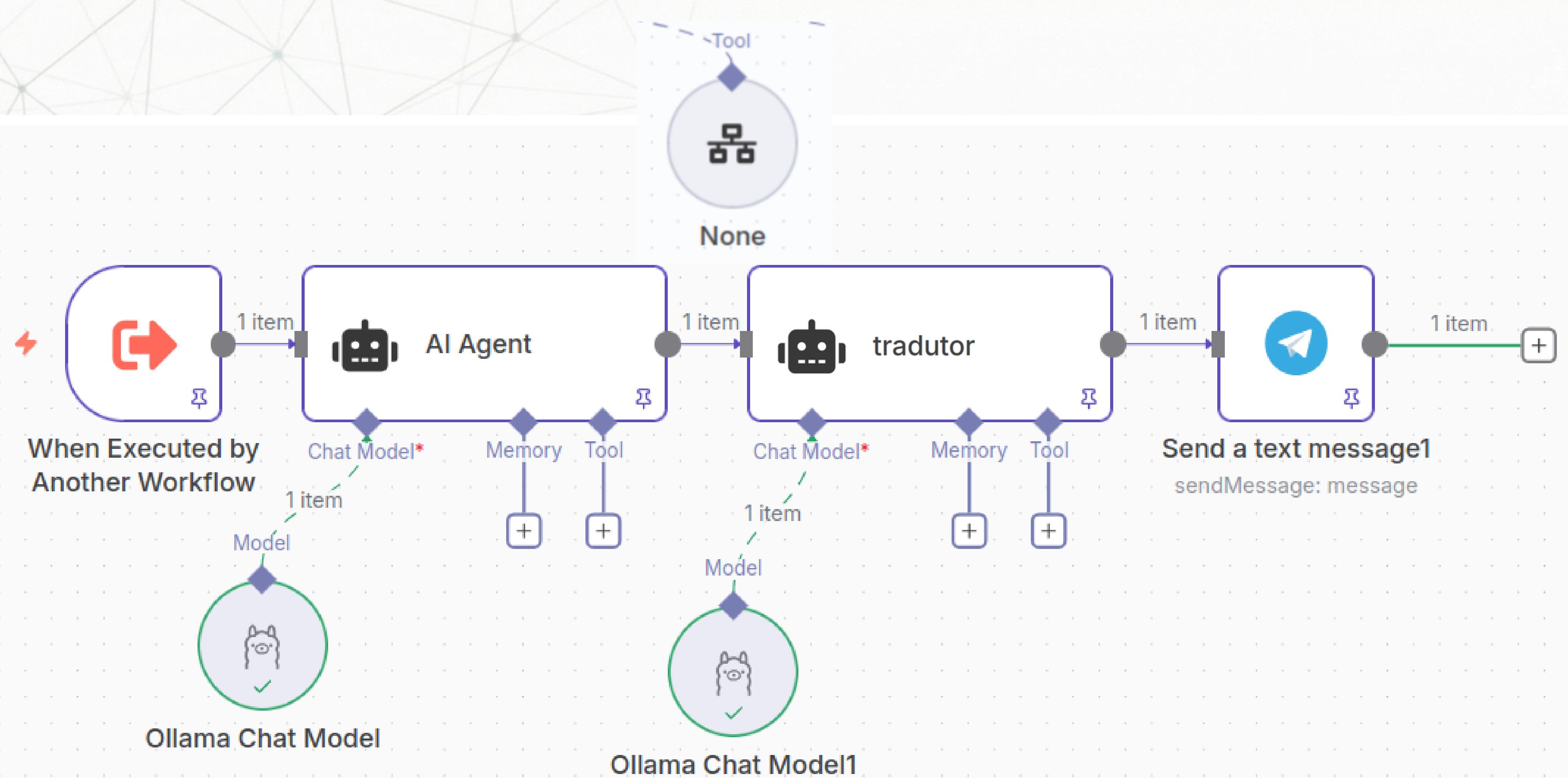
20:18

Preciso jogar futebol com meus amigos urgentemente 20:19 ✓

A tarefa jogar futebol com meus amigos urgentemente foi adicionada a lista de tarefas

This message was sent automatically with n8n

20:20





AI Agent

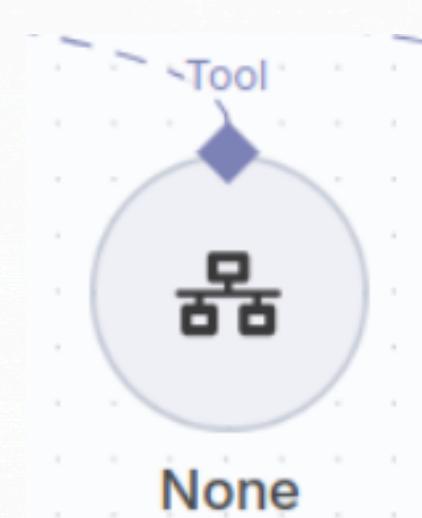
INPUT

Q Schema Table JSON

When Executed by Another Workflow 1 item

AB query preciso encontrar com os amigos para jogar futebol o quanto antes

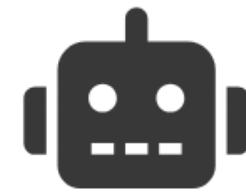
Variables and context



Output Logs Q Schema Table JSON  

1 item

```
[{"output": "Alright. When would you like to find time to play football?"}]
```



tradutor

INPUT

Q Schema Table JSON

When Executed by Another Workflow 1 item

AB query preciso encontrar com os amigos para jogar futebol o quanto antes

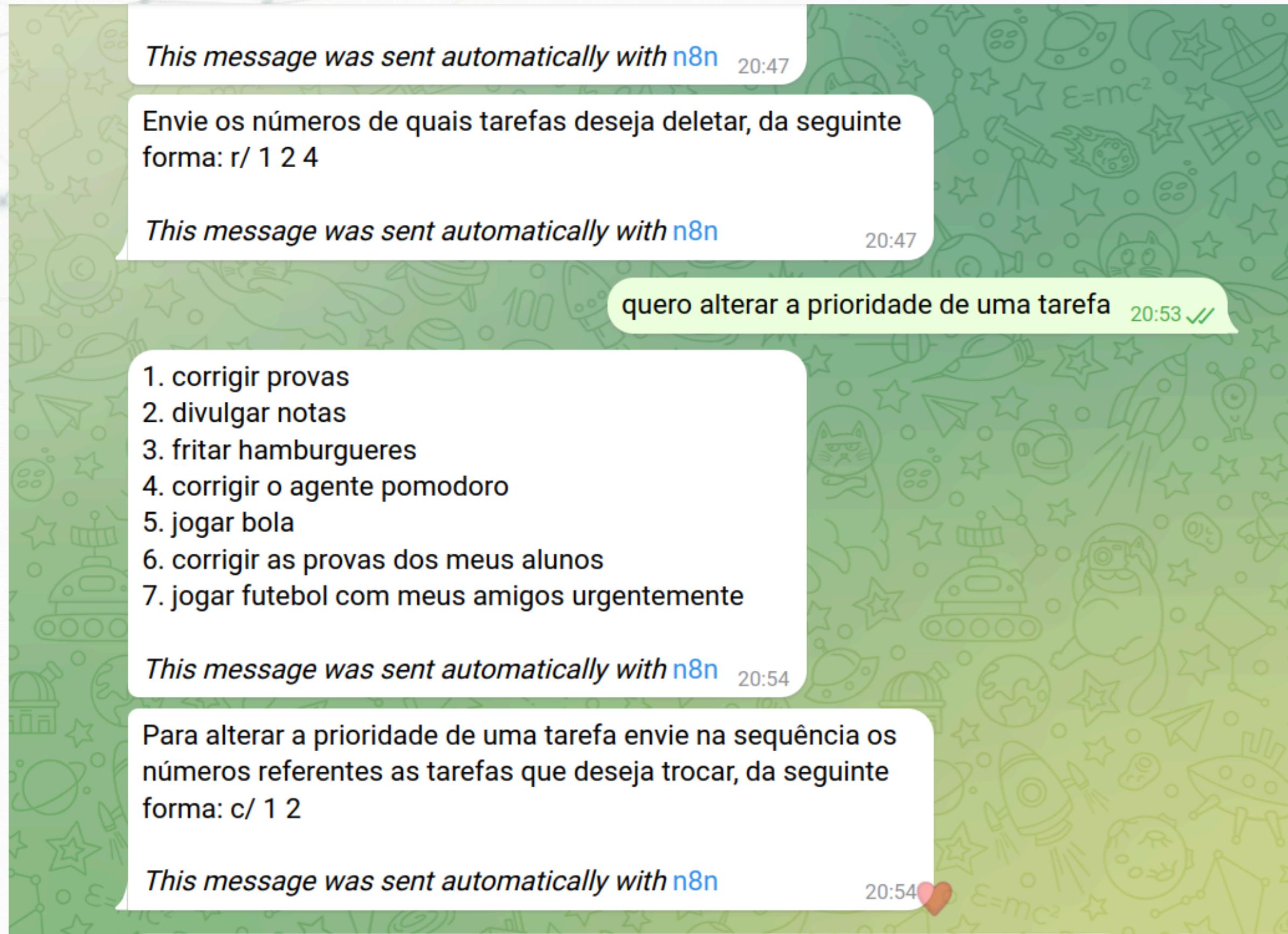
Variables and context

Output Logs Q Schema Table JSON  

1 item

```
[{"output": "Alright. When would you like to find time to play football?"}]
```





This message was sent automatically with n8n

None

1. corrigir provas
2. divulgar notas
3. fritar hamburgueres
4. corrigir o agente pomodoro
5. jogar bola
6. corrigir as provas dos meus alunos
7. jogar futebol com meus amigos urgentemente

Quero remover uma tarefa

Envie os números de quais tarefas deseja deletar, da seguinte forma: r/ 1 2 4

This message was sent automatically with n8n

Hoje



Nov. – dez. 2025



Semana



| DOM. | SEG. | TER. | QUA. | QUI. | SEX. |
|--------|-------------------------------|--|------|------|------|
| 30 | 1 | 2 | 3 | 4 | 5 |
| GMT-04 | | | | | |
| 8 AM | | | | | |
| 9 AM | | | | | |
| 10 AM | corrigir provas[POMODORO] | Defesa de TCC:
Hugo Belônia
9 – 11am | | | |
| 11 AM | divulgar notas[POMODORO] | | | | |
| 12 PM | fritar hamburgueres[POMODORO] | | | | |
| 1 PM | corrigir o agente pomodoro | | | | |
| 2 PM | | | | | |
| 3 PM | jogar bola[POMODORO] | | | | |
| 4 PM | corrigir as provas dos alunos | | | | |
| 5 PM | jogar futebol com meu filho | (Sem título)
4 – 5pm | | | |
| 6 PM | | | | | |
| 7 PM | | | | | |

Lista de Tarefas Seguindo Regras (Pomodoro)

Destaque:

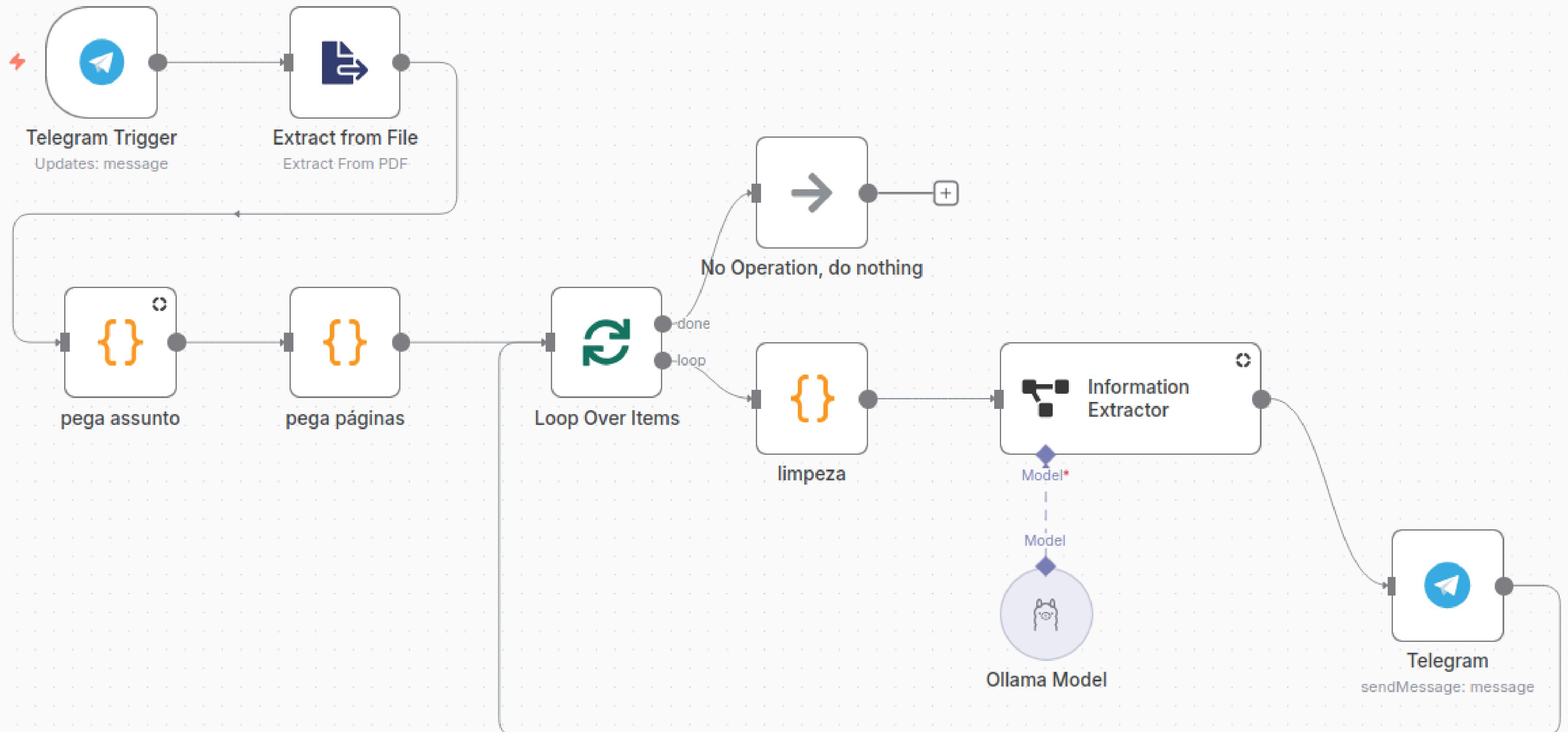
- Qwen2.5:3B (roteamento)
- Llama3.2:3B (tradutor)

Descoberta:

- A estratégia de redirecionamento reduziu a complexidade entregue ao modelo e evitou a sobrecarga cognitiva do modelo

Ler Documentos em Busca de Padrões de Interesse

- **Objetivo:** Eliminar a necessidade de leitura manual completa de documentos extensos
- **Abordagem:** Pré-filtragem via Código (*Regex*)





Information Extractor

INPUT

Q

Schema

Table

JSON



{}

limpeza

1 item

AB

assunto

16/06/2025, 13:20 SEI/UFMS -
5690031 - Edital
https://sei.ufms.br/sei/controlador.php?
acao=procedimento_trabalhar&acao_
origem=procedimento_visualizar&id_
procedimento=6429869&infra_s...
1/2
17/06/2025
Nº 8567
Pg.
231
CAMPUS DE CHAPADÃO
DO SUL
Avenida Engenheiro
Douglas Ribeiro Pantaleão, nº
5167
Fone: (67)3562-6300
CEP
79560-000 - Chapadão do Sul -
MS
Referência: Processo nº
23455.000347/2025-76 SEI nº
5690031
16/06/2025, 13:20 SEI/
UFMS - 5690031 - Edital
https://sei.ufms.br/sei/controlador.php?
acao=procedimento_trabalhar&acao_
origem=procedimento_visualizar&id_
procedimento=6429869&infra_s...
2/2
1...



{}

output

AB

assunto

RESOLUÇÃO Nº 303-CGB-ENF/CPCX, DE 13 DE JUNHO DE 2025

#

páginas

2



Information Extractor

Palavra-Chave: anderson

Assunto: Resultado Final da Seleção de Propostas para
Concessão de Recursos Financeiros

Páginas: 225, 226, 227

Today

21:45

This message was sent automatically with n8n

Palavra-Chave: anderson

Assunto: Edital de Convocação para Programa de Gestão e
Desempenho na AGETIC

Páginas: 267

21:45

This message was sent automatically with n8n



boletim_anderson.pdf
6.0MB

Sabrina Caetano

21:57 ✓/✓

Palavra-Chave: Sabrina Caetano

Assunto: Resolução do Coordenadoria de Curso de
Enfermagem

Páginas: 77

This message was sent automatically with n8n

21:59



 boletim_anderson.pdf
6.0MB

Ana Patricia Araujo 22:09 ✓

Palavra-Chave: Ana Patricia Araujo
Assunto: RESOLUÇÃO Nº 301-CGB-ENF/CPCX, DE 12 DE JUNHO DE 2025.
Páginas: 75

This message was sent automatically with n8n 22:16

Palavra-Chave: Ana Patricia Araujo
Assunto: RESOLUÇÃO Nº 302-CGB-ENF/CPCX, DE 13 DE JUNHO DE 2025.
Páginas: 76

This message was sent automatically with n8n 22:16

Palavra-Chave: Ana Patricia Araujo
Assunto: Aprovação da composição das bancas examinadoras dos Trabalhos de Conclusão de Curso (TCC) do curso de Enfermagem-Bacharelado
Páginas: 77



Ler Documentos em Busca de Padrões de Interesse

Destaque:

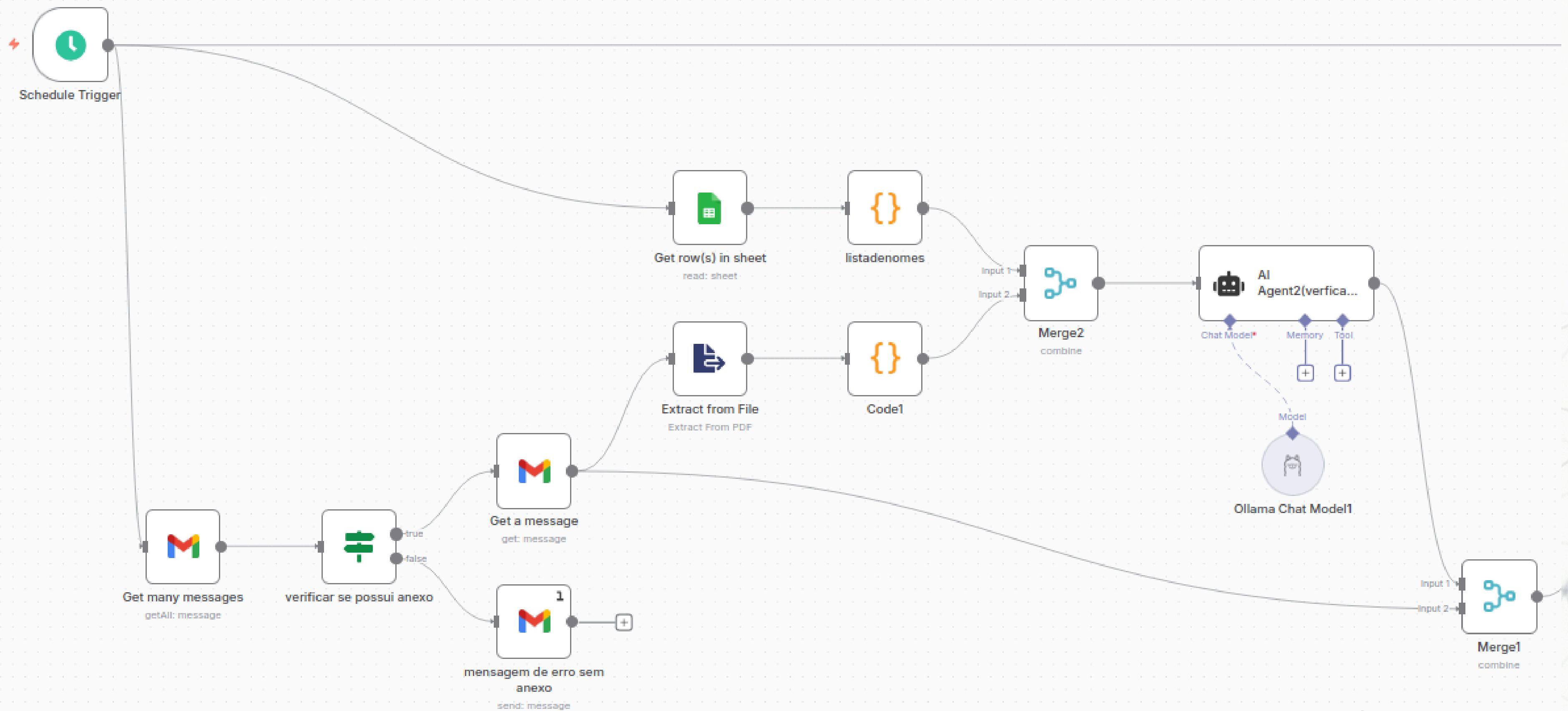
- Qwen2.5:3B

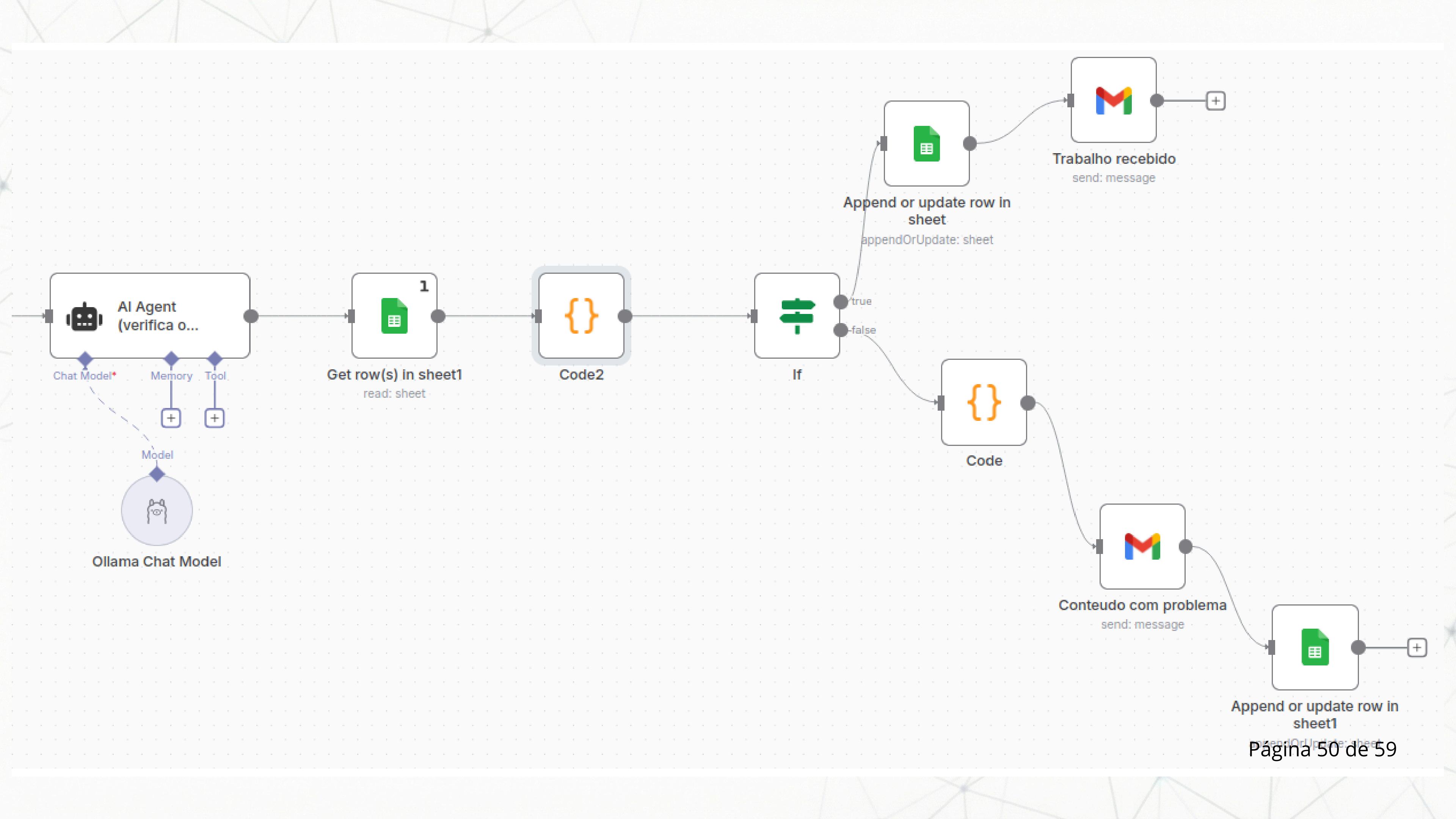
Descoberta:

- A pré-filtragem com *Regex* provou-se eficaz para contornar a limitação física da janela de contexto (2048 *tokens*)
- Foi necessário ajustar a temperatura do modelo para 0.3

Resposta Automatizada de Tarefas de Acadêmicos

- **Objetivo:** Conferência de entrega e a validação básica do conteúdo de um trabalho acadêmico
- **Abordagem:** Os agentes atuam pontualmente como validadores semânticos, alterando os metadados





{ } Assunto do trabalho Javascript

Execute step

Parameters

Settings

Docs ↗

Mode

Run Once for All Items

Language

JavaScript

JavaScript

```
1 const meuJson = {  
2   assunto: "leitura de arquivos por parametro usando  
3   python"  
4 };  
5  
6 return [{  
7   json: meuJson  
8 }];
```

Type \$ for a list of **special vars/methods**. Debug by using `console.log()` statements and viewing their output in the browser console.

NOME: Josefo

TRABALHO: trabalho de teste do checkbot

CONTEÚDO:

Leitura de arquivos por parâmetro em python se da dessa forma:

```
# importa o módulo 'sys' para acessar os argumentos da linha de comando
import sys

# Verifica se o nome do arquivo foi passado como argumento
if len(sys.argv) < 2:
    print("Erro: Você precisa especificar o nome do arquivo.")
    print("Uso: python seu_script.py <caminho_para_o_arquivo>")
    sys.exit(1) # Encerra o script com um código de erro

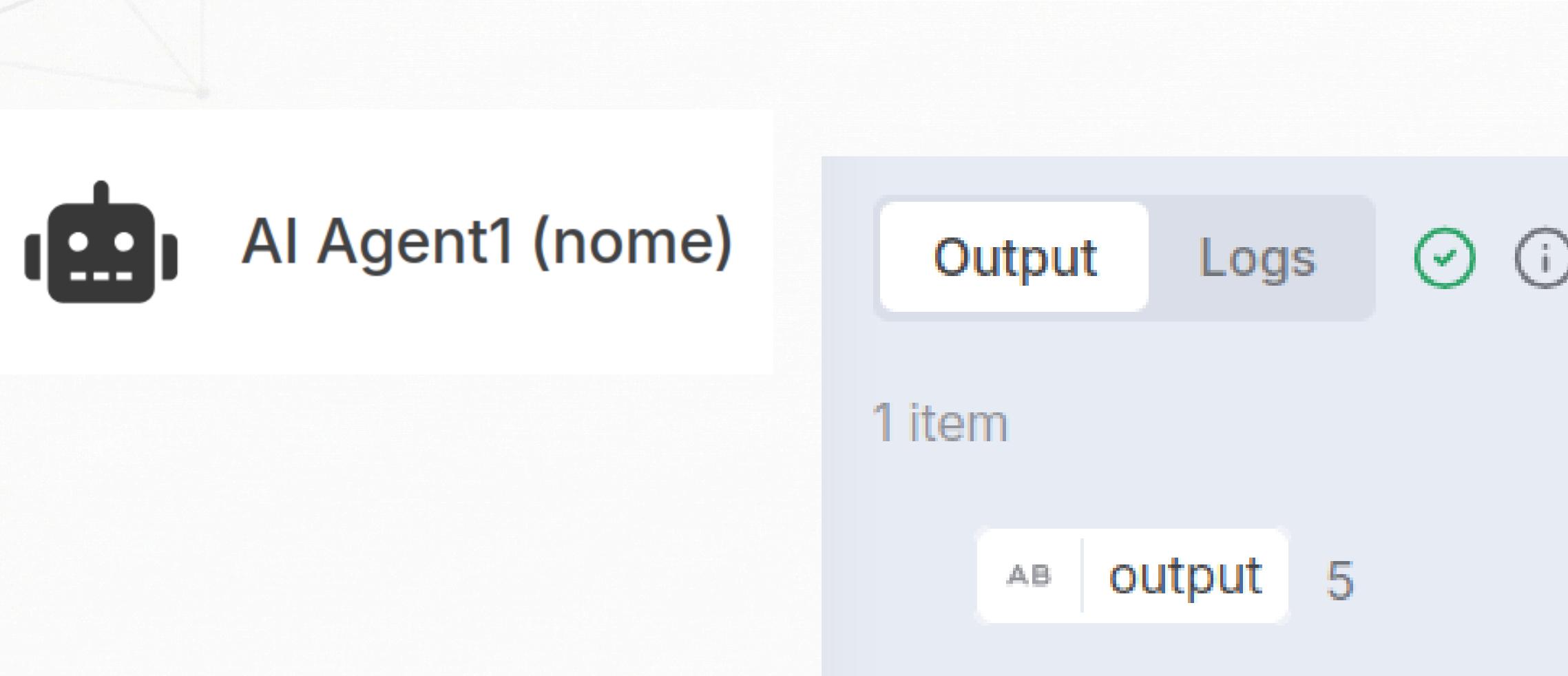
# Pega o nome do arquivo do primeiro argumento (o índice 0 é o nome do script)
nome_do_arquivo = sys.argv[1]

try:
    # O 'with' garante que o arquivo será fechado automaticamente
    with open(nome_do_arquivo, 'r', encoding='utf-8') as arquivo:
        conteudo = arquivo.read() # Lê todo o conteúdo do arquivo
        print(conteudo) # Imprime o conteúdo na tela
except FileNotFoundError:
    print(f"Erro: O arquivo '{nome_do_arquivo}' não foi encontrado.")
except Exception as e:
    print(f"Ocorreu um erro inesperado: {e}")
```

YOUR TASK

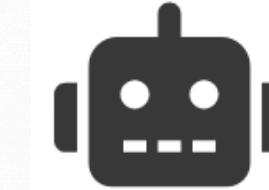
```
lista_de_nomes: {{ $json.lista_final }}  
nome_para_encontrar: {{ $json.nome }}
```

Sua Resposta:



The image shows a screenshot of an AI interface. On the left, there is a dark gray icon of a robot head. To its right, the text "AI Agent1 (nome)" is displayed. In the center, there is a light gray box with rounded corners. At the top of this box, there are three buttons: "Output" (which is highlighted in white), "Logs", and two small circular icons with a checkmark and an information symbol. Below these buttons, the text "1 item" is shown. At the bottom of the box, there is another row of buttons: "AB" (disabled), "output" (highlighted in white), and "5". The background of the entire interface is a light gray with a subtle network-like pattern of thin lines and dots.

AI Agent2 (conteúdo)



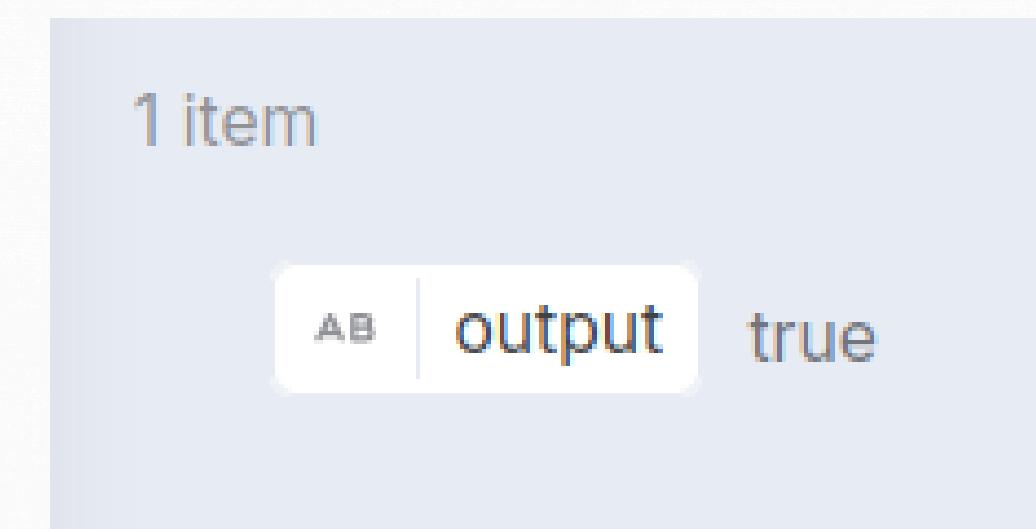
Você é um assistente especialista em analisar o conteúdo de documentos técnicos.

O documento a seguir foi extraído de um PDF:

```
 {{ $('Code1').item.json.texto }}
```

O tema a ser verificado é: " {{ \$('Assunto do trabalho Javascript').item.json.assunto }} "

Baseado no conteúdo do documento, ele é sobre o tema especificado? Responda exclusivamente com a palavra "true" ou a palavra "false", em letras minúsculas e sem pontuação.



C13 ▾ | fx |

| | A | B | C | D | E | F | G | H | |
|----|-------------------------|---------------------------|----|----------|-------|----------|-------------------|------------------------|------|
| 1 | Nomes | Nome:RowNumber | id | Entregue | Vazio | Conteudo | Ok | Email | Nota |
| 2 | Fulano da bezerra | Fulano da bezerra:2 | 2 | OK | Sim | | | hugo.belonia@gmail.com | |
| 3 | Beltrano da silva | Beltrano da silva:3 | 3 | | | | | | |
| 4 | Deltrano do Motta Alves | Deltrano do Motta Alves:4 | 4 | OK | Não | ? | Ok (ver conteudo) | hugo.belonia@gmail.com | |
| 5 | Josefo Caipira | Josefo Caipira:5 | 5 | OK | Não | OK | OK | hugo.belonia@gmail.com | |
| 6 | Jack Bueno | Jack Bueno:6 | 6 | | | | | | |
| 7 | Joseff Climber | Joseff Climber:7 | 7 | | | | | | |
| 8 | | | | | | | | | |
| 9 | | | | | | | | | |
| 10 | | | | | | | | | |

Resposta Automatizada de Tarefas de Acadêmicos

Destaque:

- Qwen2.5:3B (confirmação)
- Llama3.2:3B (validação)

Descoberta:

- A estratégia de utilizar agentes para enriquecimento de metadados, mantendo o controle de fluxo via programação tradicional, mostrou-se eficaz.

Resultados

Qwen2.5:3B

Consistente

Atingiu a maioria dos objetivos

Alucina menos vezes

Uso Recomendado: Extração de texto e
acionamento de ferramentas

Prompt: Diretivo, claro e curto.

Phi3.2:8B

Não suportado para o *hardware*

Qwen2.5 (0.5 e 1.5B)

Desempenho semelhante ao Llama,
porém se mostrou mais precário
quanto a linguagem natural

Llama3.2(1 a 3B)

Inconsistente

Dificuldade no uso de ferramentas

Alucina constantemente

Uso Recomendado: chat

Prompt: Tipo persona, com **exemplos**



Obrigado