Building a Multi-Source Research Agent with the Merge Node in N8N

Overview

This SOP guides you through building a multi-source research agent in N8N using the Merge node. The setup involves aggregating search results from Gemini, ChatGPT, and YouTube, and delivering them through an AI agent. The process combines AI model outputs and YouTube data using structured workflows. By the end of this SOP, you will have a fully functional agent capable of executing multi-source searches and returning combined results in real-time.

Step-by-Step Guide

1. **Set Up the Main Workflow with a Chat Trigger**
    a. Use the `Chat Message` node.
    b. Make the chat trigger public for external access.
2. **Add and Configure the AI Agent**
    a. Use the Gemini 2.5 model with simple memory.
    b. Add the `Call Any10 Tool` node to enable external workflow execution.
3. **Create a Sub-Workflow for Multi-Source Search**
    a. Open a new canvas and use `Execute Sub-Workflow` trigger.
    b. Define workflow inputs:
        - `search_term` (string)
        - `search_engine` (string, accepts: `Gemini`, `GPT`, `YouTube`, `all`)
4. **Add a Switch Node for Engine Routing**

a. Configure routing rules using `.toUpperCase()` on `search_engine`.

b. Route based on conditions: `Gemini`, `GPT`, `YouTube`, `all`.

c. Enable “Send data to all matching options”.

5. **Configure Gemini and GPT AI Nodes**

a. Use `Gemini 2.flash` and `ChatGPT` models.

b. Set system message: “You are a search engine...”

c. Drag `search_term` to prompt.

6. **Use Edit Fields Node to Extract Outputs**

a. Extract only the `text` or `content` from Gemini and GPT outputs.

7. **Set Up YouTube Search**

a. Enable YouTube Data API, Analytics, and Reporting in Google Cloud.

b. Use `Get Many Videos` node with filters:

- `query` from `search_term`
- `order`: `relevance`
- `safeSearch`: `moderate`

c. Extract `title`, `description`, `videoId`.

8. **Aggregate YouTube Outputs**

a. Use `Aggregate` node to combine video metadata.

9. **Merge All Outputs**

a. Use `Merge` node with:

- Mode: `Combine`
- By: `Position`
- Number of inputs: 3
- Options: `Include any unpaired items`, `Use earliest version`, `Deep merge`

10. **Return Merged Result**

11. Ensure the last node in sub-workflow is the `Merge` node for data output.

12. **Link Tool Description in AI Agent**

13. Provide a detailed tool description:
    - Capabilities: search Gemini, GPT, YouTube
    - Inputs: `search_term`, `search_engine`
14. **Validate Agent Tool Awareness**
15. Test the agent with prompt: “What tools do you have?”
16. Confirm it lists the call workflow tool and supported engines.
17. **Test Workflow with Real Search**
18. Send a message to the chat agent with search terms and selected engines.
19. Verify that combined results appear as expected.
20. **Merge Two Google Sheets (Case 1 - Sorted by Index)**
21. Use two `Get Rows` nodes for each sheet.
22. Use `Merge` node with:
    - Mode: `Combine`
    - Match by: `row number`
    - Output: `Keep matches`
23. **Merge Two Google Sheets (Case 2 - Match by Field)**
24. Use `product_id` as key field.
25. In `Merge` node:
    - Mode: `Combine`
    - Match by: `product_id`
    - Options: `Keep all`, `Sort by product_id` (optional)

Warnings and Notes

- Ensure all APIs (YouTube Data, Analytics, Reporting) are enabled in Google Cloud before using the YouTube node.
- When using the Merge node, always select "Include any unpaired items" to prevent workflow hangs.
- Avoid using ambiguous search terms; exact keyword phrasing improves search accuracy, especially for YouTube results.

Visual Diagram

```
graph TD
    A[Start Chat Trigger] --> B[Initialize AI Agent]
    B --> C[Call Any10 Tool Node]
    C --> D[Sub-Workflow Triggered]
    D --> E[Switch on Search Engine]
    E --> F1[Gemini Search]
    E --> F2[GPT Search]
    E --> F3[YouTube Search]
    F1 --> G1[Edit Fields - Gemini]
    F2 --> G2[Edit Fields - GPT]
    F3 --> G3[Aggregate YouTube Videos]
    G1 --> H[Merge Outputs]
    G2 --> H
    G3 --> H
    H --> I[Return Merged Results to Agent]
```