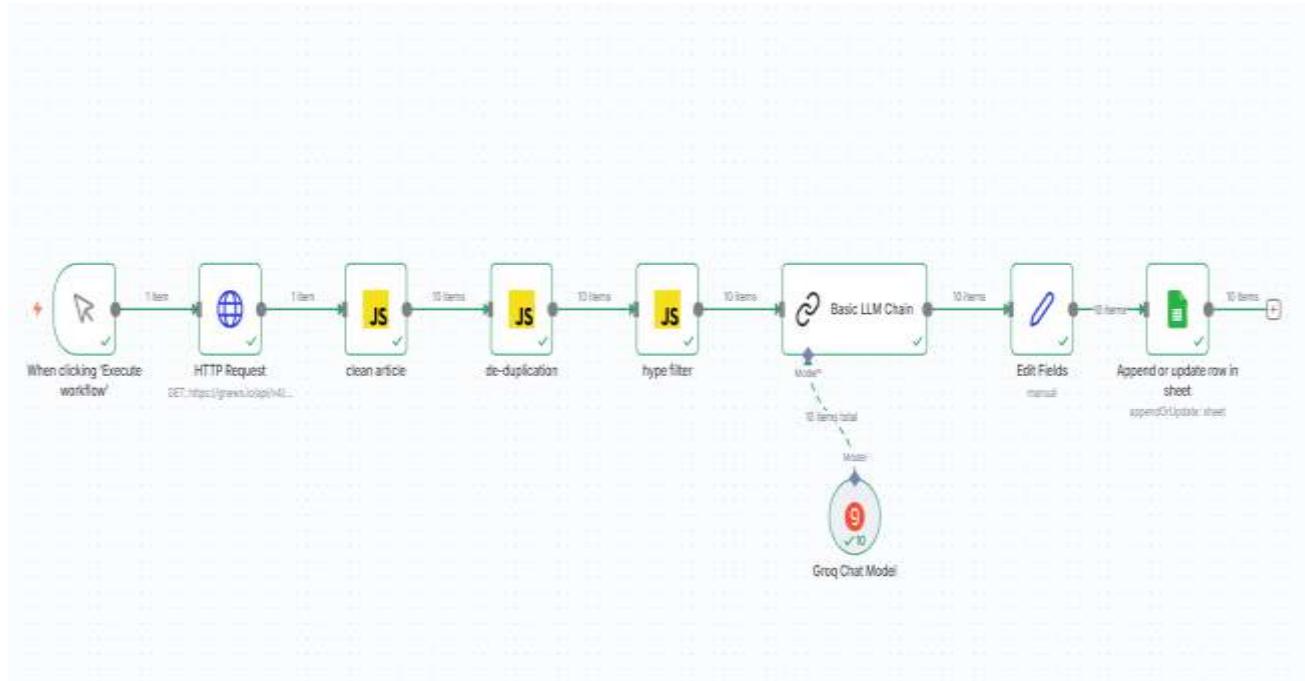


AI Analyst Agent – Workflow Diagram



1. HTTP Request – Fetch AI Startup News (This step connects to the GNews API and retrieves the latest news articles related to AI startups.
It gives us a raw JSON list of articles containing title, description, content, and other metadata.)



2. Clean Article Data – Extract Useful Fields (This step cleans the raw API response by keeping only essential fields like the title, description, content, URL, source, and published date. It also includes error handling so the workflow continues smoothly even if the API limit is hit or some fields are missing.)



3. De-duplication Logic – Remove Repeated Articles (This step checks every article's title and removes duplicates so that only one unique version of each news event is kept. It prevents repeated or similar articles from entering the workflow. Error handling is added to ensure the workflow continues even if titles are missing or all items are duplicates.)



4. Hype Filter – Remove Low-Information Articles (This step filters out articles that contain too many marketing or hype words and keeps only the ones with real technical or useful information. It also removes very short or low-quality articles and includes error handling to manage missing data or cases where all articles are filtered out.)

5. LLM Structured Extraction – Convert Article into JSON (This step sends each filtered article to the **Groq LLM**, which extracts structured fields such as the company name, AI category, sentiment score, and whether the article is funding news. Groq returns this information in clean JSON format, making the data easy to store and analyze. Error handling ensures the workflow continues smoothly even if the model output is incomplete.)



6. Edit Fields – Parse and Format the LLM Output (This step takes the JSON text produced by the Groq LLM and converts it into clean, structured fields for storage. It parses the LLM's output safely and extracts the final values for company name, category, sentiment score, and funding status. This prepares the data in a consistent format for Google Sheets.)



7. Google Sheets Output – Store the Final Structured Data (This step saves the extracted and formatted information into a Google Sheet for easy tracking and analysis. Each processed article is added as a new row containing the company name, category, sentiment score, and funding status. This provides a clean, searchable record of AI startup news.)

Created by: Adhithya
MP
n8n Workflow: Analyst Agent (Option 2)