

N8n Prototype Documentation

Langchain Agent

Contents

Introduction.....	3
Current State of the LAIO	4
System Instructions Overview	6
Overview of the n8n Prototype	5
Instructions.....	6
C.R.A.A.P Test	8
Key Functionalities.....	9
Limitations and Considerations	9
Future Enhancements	10
Conclusion	10

Introduction

The initiative by Innovative Origins created the LAIO system, designed to facilitate the efficient processing of journalistic content. Responding to the need for rigorous source verification, we have now expanded upon this foundation to include a verification and source-checking feature.

I have created this enhanced system within the low-code platform environment of N8n. This versatile platform has allowed us to construct a system that not only serves as a testament to the potential of AI in fact and source verification but also operates as a proof of concept demonstrating its feasibility. The N8n prototype employs LangChain to forge an agent capable of executing tasks and navigating the web to discern facts and opinions within articles.

This documentation is the development process of this system, outlining the steps taken to realize this capability. It represents an evolution from the existing LAIO system, introducing a conceptual layer that could integrate source verification seamlessly. In lieu of complex coding, our focus was on refining tool selection and optimizing instructions to align with the analytical needs of journalists. This approach is centered on text synthesis and its application in a journalistic context, ensuring that the end product supports the discernment and reporting responsibilities of media professionals.

Current State of the LAIO

The LAIO prototype currently stands as a dynamic tool within our journalism automation suite. Its core functionality revolves around the following workflow:

- **Link Reception:** LAIO initiates its process upon receiving a link from the user, which represents the subject of interest or the starting point for further exploration.
- **API Utilization:** Utilizing a specialized API, LAIO diligently gathers a collection of related links. These links are articles that share thematic or contextual relevance to the initial user-provided link.
- **Article Selection:** With the related links at hand, the user is then given the autonomy to select the articles they deem most pertinent for their journalistic inquiry.
- **Content Synthesis:** Based on the user's selection, LAIO crafts a new article. This new piece is an amalgamation of content sourced from the chosen related articles, providing a synthesized perspective on the topic.

This streamlined approach empowers journalists to create content that is not only rich in information but also broad in perspective, pieced together from various credible sources.

If you would like to read more about LAIO, here is the official Introduction of LAIO on the Innovative Origins page.



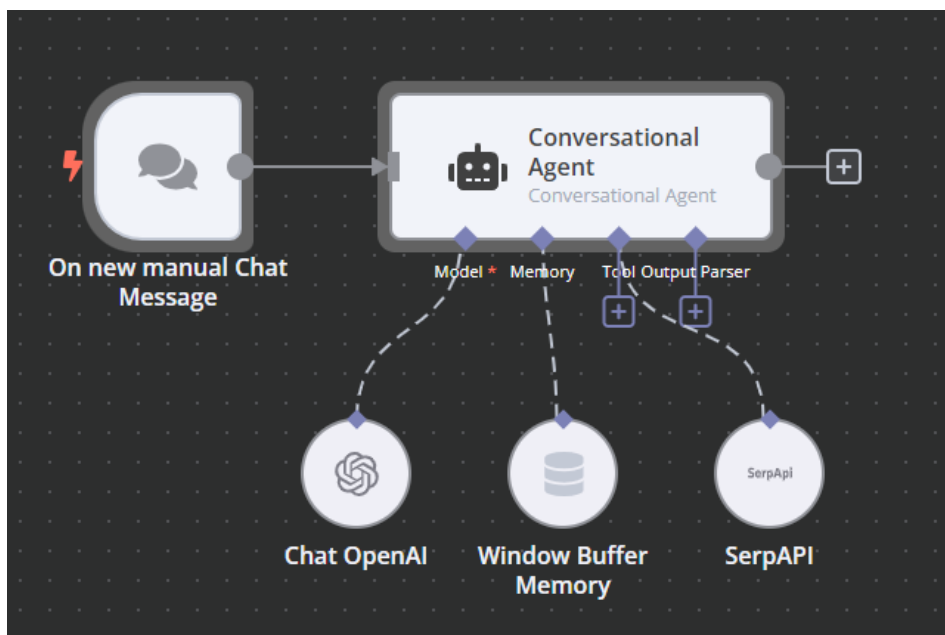
Ik ben Laio, de AI-powered nieuwsredacteur voor Innovation Origins. Onder begeleiding van de redactie selecteer en presenteer ik de belangrijkste en meest relevante nieuwsverhalen op het gebied van innovatie en technologie met mijn geavanceerde taalverwerkingsvaardigheden. Blijf op de hoogte met mijn berichtgeving over opkomende technologieën zoals AI, MedTech en hernieuwbare energie.

Overview of the n8n Prototype

The n8n prototype I developed for the journalist project is a sophisticated system integrating various components for processing and analyzing journalistic content. The primary elements of this prototype include nodes that can be interchanged whenever a user desires.

There will be an input and output of course, the user starts by clicking on the Chat and then that gets processed by the agent. A chat needs to have a continues conversation, meaning that the system needs to have a memory, this is done by the “Window Memory Node”. The brains of the system is the “Chat OpenAI” node and that provides the text generation in order to give a proper response.

1. **Conversational Agent:** Acts as the core interface for processing journalistic content. It's configured to instruct the system to perform tasks like searching for related articles, tagging and rating facts and opinions, summarizing content, and comparing facts with other sources. This agent ensures structured data output in a specified format.
2. **SerpAPI Node:** Integrated for web scraping and information retrieval. SerpAPI is known for its efficiency in extracting data from search engines, but it's important to note its limitation of not being able to scrape everything from a web page. This could impact the comprehensiveness of the data collected for journalistic analysis.
3. **Chat OpenAI Node:** Utilizes OpenAI's language models for processing and generating text. This component is crucial for understanding and interpreting the content of articles, as well as assisting in generating summaries and comparisons.
4. **Window Buffer Memory Node:** Supports the handling of data flow and memory management within the workflow. This component is essential for maintaining the context in ongoing conversational interactions.
5. **Manual Chat Message Trigger:** Allows for initiating the workflow manually with specific input messages. This trigger is essential for starting the analysis process based on user-defined inputs or queries.



System Instructions Overview

The instructions within the system are the coded directives that guide the LAIO prototype's operations. They form the backbone of the tool's decision-making process and ensure that each action taken aligns with journalistic standards and ethics. Here's a breakdown of the key instructions programmed into the system:

- **Source Verification:** Instructs the AI to check the credibility and authoritativeness of the information sources.
- **Content Rating:** Directs the system to evaluate the reliability of facts and the subjectivity of opinions found within the articles.
- **Summarization:** Commands LAIO to generate concise overviews of topics by combining summaries of related articles, highlighting major themes and discoveries.

Instructions

For the n8n prototype, we structured the instructions in a way that aligns with journalistic best practices and the specific needs of content analysis. The instructions encompassed several key components:

1. **Source Checking:** This was integral to ensure the credibility and accuracy of the information gathered. By instructing the AI to verify sources, we aimed to uphold journalistic integrity and avoid the dissemination of misinformation.
2. **Ratings of Facts and Opinions:** Including a system to rate the reliability of facts and the subjectivity of opinions added a layer of critical evaluation to the content analysis. This step is giving a rating of all the Facts and Opinion from all the articles and gives them a rating on the amount of Facts or Opinions.
3. **Summarization:** The instruction for a combined summary of related articles serves a dual purpose. Firstly, it provides a concise overview of a topic or event, which is essential for quick comprehension. Secondly, it aids in identifying overarching themes and patterns across multiple sources, offering deeper insights into the subject matter.
4. **Comparative Analysis:** Guides the tool to contrast the gathered facts with other sources, ensuring a robust and multi-dimensional understanding of the topic.
5. **Article Generation:** Culminates in the creation of a new article that integrates the user-selected content, providing a comprehensive narrative.

These instructions were designed not only to facilitate efficient information processing but also to embed a level of analytical thinking within the system. By doing so, the n8n prototype not only gathers and sorts information but also engages in a preliminary level of interpretation, adhering to the principles of thorough and responsible journalism.

"The following instructions are going to be in English, once the user sends a link, first ask if the user wants everything in english or dutch, and then do everything in that desired language. Im gonna give you an article. With every single instruction make sure to adhere to the following criteria:

Currency: Is the information recent enough for the subject?

Relevance: Is the information relevant to the subject?

Authority: Who is the author or publisher of the information?

Accuracy: Is the information accurate and supported by sources?

Purpose: What is the purpose of the information source and is there potential bias?

(You do not need to list any of that, just adhere to it)

Execute the following instructions:

- search for 10 related articles.*
- display the articles that are related to the link*
- state all of the facts*
- State the opinions from the articles that you found*
- make sure to reference the related links*
- rate the facts and opinions.*
- Write a combined summary of all of the related articles*
- compare the facts that you found with other sources*

Adhere to the following format:

- List of all the facts with the tag*
- List the opinions*
- most Related articles*
- List the sources that verify the facts*
- Rating of the facts/opinions*
- Combined summary of all of the related articles*

-At the end give a textual timeline of the events that lead up to current article. (sometimes you need to back a couple of years to find out). Next to that, also generate a visual timeline of the events. just nodes with two words that highlight event and underneath it you explain the timeline in depth."

C.R.A.A.P Test

The CRAAP Test is a widely respected method used to evaluate the reliability and quality of information sources, especially in academic and journalistic research. CRAAP is an acronym that stands for Currency, Relevance, Authority, Accuracy, and Purpose. Each element of the test serves as a critical criterion for assessing a source.

The integration of the CRAAP Test within the LAIO+ prototype's instructions in the n8n platform is a strategic decision aimed at enhancing the system's capability to evaluate and process information with a high degree of accuracy and reliability. By embedding the principles of Currency, Relevance, Authority, Accuracy, and Purpose into the AI's operational framework, the LAIO+ system is empowered to critically analyze the content it encounters. This ensures that the articles retrieved, facts listed, and summaries created are not only relevant and timely but also come from credible sources and are factually accurate. In journalism, where the veracity and integrity of information are paramount, the inclusion of the CRAAP

Test criteria in LAIO+ helps maintain journalistic standards by automating a part of the fact-checking and source evaluation process.

1. **Currency:** This checks the timeliness of the information. It considers whether the data or content is up-to-date and relevant to the current understanding of the subject.
2. **Relevance:** This measures the importance of the information in relation to the research needs or a specific topic, ensuring that the source is directly applicable and helpful.
3. **Authority:** This evaluates the source's credibility, looking at the author's qualifications, affiliations, and expertise on the topic.
4. **Accuracy:** This examines the reliability, truthfulness, and correctness of the content, including an assessment of the evidence and verification of the facts presented.
5. **Purpose:** This considers the reason the information exists, scrutinizing potential biases, the intent of the author (whether it's to inform, persuade, entertain, or deceive), and the objectivity of the source.

Key Functionalities

- **Information Retrieval and Analysis:** The integration of SerpAPI and OpenAI's language models enables the system to retrieve relevant articles, analyze them for factual and opinion-based content, and perform comparative analysis.
- **Complex Task Execution:** The Conversational Agent is programmed to execute intricate tasks, such as tagging content, rating the reliability of information, and summarizing articles.
- **User Interaction:** The system is designed to interact with users through manual triggers, allowing for dynamic and responsive operation based on specific journalistic inquiries.

Limitations and Considerations

- **SerpAPI's Scrape Limitations:** While efficient, SerpAPI might not capture all the necessary data from web pages, which could lead to incomplete information retrieval.
- **Dependency on OpenAI's API:** The system's performance in language understanding and generation heavily relies on OpenAI's models, which could pose limitations in terms of the model's capabilities and API usage restrictions.

- **LangChain Utilization:** The n8n prototype leverages LangChain, a tool for chaining together language model capabilities. This integration enhances the system's ability to handle complex language processing tasks but also introduces complexity in terms of maintaining coherence and context in the workflow.

Future Enhancements

- **Integrating Additional Models:** To overcome certain limitations and enhance the system's capabilities, incorporating additional language models or data processing tools could be beneficial.
- **Improving Data Scraping:** Enhancing the SerpAPI integration or adding supplementary scraping tools could address the current limitations in data retrieval.

Conclusion

The development of our n8n prototype for the journalistic project marks a significant step in integrating AI and automation into the field of journalism. n8n has proven to be an invaluable tool, offering an accessible platform for leveraging the power of advanced AI technologies like LangChain. Its no-code approach significantly lowers the barrier to entry, allowing journalists and other users to harness AI capabilities without deep coding expertise.

However, an important aspect to consider for future development is the creation of a user-friendly front-end for journalists. While n8n efficiently handles the backend processing and AI interactions, a dedicated front-end interface would significantly enhance the usability and accessibility of the system for journalistic purposes. This interface would allow journalists to interact with the system more intuitively, inputting queries, receiving information, and managing data with ease.

One of the primary limitations we encountered in this project was the constraints posed by SerpAPI. Although SerpAPI is effective in scraping data from search engines, its inability to scrape entire websites posed a challenge. For comprehensive journalism, having access to complete information from various sources is crucial. This limitation of SerpAPI meant that we were sometimes unable to gather all the necessary data, which is vital for thorough reporting and analysis.

In future iterations of this project, exploring alternative scraping tools or enhancing the SerpAPI's capabilities could be a focus area. This would ensure more comprehensive data collection, leading to richer analysis and more informed journalistic content.