Final NLP Project Report: AI Agents for Lifelong Learning

The Problem

Lifelong learning requires consistent exposure to relevant, up-to-date content that aligns with individual interests and contributes meaningfully to growth. Despite the abundance of information, there is no solution that effectively curates high-quality, tailored content. Users struggle to sift through irrelevant information or simply fail to consume newsletters and research updates due to the overwhelming volume.

This project emerged from my own experience with this challenge: I wanted a system to deliver filtered, summarized content from selected sources to accelerate my learning process and integrate it seamlessly into my daily routine.

Research on Existing Solutions

During my research, I found no existing tutorial or framework that tackled this specific pain point. While there are tutorials for building agents that scrape data from specific sites or fetch information via search engines like Google, none address the combination of targeted content aggregation, intelligent filtering, summarization, and delivery via accessible channels such as WhatsApp. This gap in available resources highlighted the novelty and necessity of the proposed solution.

The Solution

This project introduces a multi-agent system designed to act as a learning companion, leveraging AI and LLMs to curate, filter, summarize, and deliver content in a personalized and actionable way.

The system was built using:

- **FireCrawl** for LLM content scraping.
- **OpenAI API** for advanced AI filtering and summarization.
- **Wpp-Whatsapp** for WhatsApp integration to send the content.

The agents perform distinct tasks:

1. **Query Content**: The Fetcher Agent dynamically retrieves articles, memos, and updates from high-quality sources like Hacker News and newsletters.
2. **Filter for Relevance**: The Filter Agent identifies the most pertinent content aligned with user-defined learning goals and interests.
3. **Summarize for Precision**: The Summarizer Agent provides concise overviews, highlighting key insights and linking to the original material for deeper exploration.

4. **Notify for Accessibility**: The Notifier Agent sends these summaries directly to the user via WhatsApp.

This system creates a seamless, automated process to integrate continuous learning into users' lives.

How It Works

The multi-agent system operates as follows:

1. **Fetcher Agent** collects data from predefined URLs using FireCrawl.
2. **Filter Agent** processes the fetched data to match user preferences through keyword analysis.
3. **Summarizer Agent** generates concise, actionable summaries using OpenAI's GPT models, focusing on key takeaways, significance, and supporting details.
4. **Notifier Agent** delivers the curated summaries to the user via WhatsApp, ensuring accessibility and ease of use.

This pipeline reduces the cognitive load on users and integrates learning into their routines effectively.

User Interviews and Feedback

To measure the impact, I conducted interviews with two users (asides me) experiencing similar pain points.

User 1: Victor Banzato (Entrepreneur)
Victor struggles to manage multiple subscribed newsletters and wanted a habit-forming solution for morning learning. He reported:
"The system delivered concise and relevant summaries that captured the key insights from 95% of the articles. This will definitely make my mornings more productive and help me establish a learning routine."
Improvement Suggestion: He recommended delivering the insights as a more organized document instead of plain text messages.

User 2: Antonio Peixoto (Multinational Executive)
Antonio, who still reads physical newspapers on weekends, found newsletters buried under work emails. He shared:
"This system will definitely help me stay updated with market trends during busy weekdays. 91% of the delivered content was relevant, and the summaries saved me valuable time."
Improvement Suggestion: He suggested integrating visual aids (images and diagrams) to improve understanding and developing cross-source intelligence to connect related insights.

Conclusions

This project successfully addressed a real-world pain point, creating a system that bridges the gap between abundant information and meaningful, personalized learning. The multi-agent architecture demonstrated its effectiveness in automating the collection, filtering, summarization, and delivery of relevant content. The high relevance accuracy (91%-95%) reported by users validates the system's performance.

Next Steps

Based on user feedback, the following enhancements will be prioritized:

1. **Delivery Format**: Transition from plain text to structured, document-based summaries.
2. **Automation**: Implement a cron job to automate the entire solution, ensuring consistent and timely updates without manual intervention.
3. **Generality Testing**: Conduct extensive testing with additional relevant sources to evaluate the system's ability to perform generically across various content types and domains.

These improvements will further align the system with users' needs and expand its utility as a tool for lifelong learning.