29th July, 2024 - Scatter Gather

Minutes of Meeting Date: July 29th, 2024

Attendees: Aravind Narayanan, Prof. Javed Mostafa

Agenda:

1. Overview of Current System

2. Integration with News RSS Feeds

3. Enhancement of Clustering Labels

4. Visualization and Presentation

Discussion Points:

1. **Current System Overview:**

   - The current system, which focuses on clustering and visualizing data, needs to be placed online.

   - The system will be hosted at `dataviz.lairhub.com`.

2. **Integration with News RSS Feeds:**

   - The system will be connected with one or more news RSS feed streams.

   - Initially, the system can be placed online without linking to the news streams, with the integration to be completed later.

   - Suggestions for free RSS feed sites to start with were provided: <u>Zapier Blog on Best RSS Feed Reader Apps</u>.

3. **Enhancement of Clustering Labels:**

   - Use GPT to improve the description and labeling of clusters within the system.

   - The objective is to make the clustering more intuitive and easier to qualify, even if the field is not well understood.

4. **Visualization and Presentation:**

   - Emphasis on creating visually comfortable Python-based visualizations.

   - Focus on how relationships and data clusters can be effectively shown, particularly for news-related data.

   - The visualizations should be generalized to allow for testing with standard academic content as well.

5. **Action Items:**

   - **Aravind:**

     - Place the current system online at `dataviz.lairhub.com` as soon as possible.

     - Begin integration with news RSS feed streams.

     - Continue working on the enhancement of clustering labels using GPT.

     - Share the URL of the site with Javed once the system is online.

     - Continue working on item 4 (Visualization) on your local machine until the news integration is complete.

6. **Next Steps:**

   - Follow up on the integration of RSS feeds and the enhancement of clustering labels.

   - Review visualizations and provide feedback.