



4th July, 2024 - Scatter Gather

Minutes of MeetingDate: July 4, 2024

Attendees: Aravind Narayanan, Prof. Javed Mostafa

Minutes of Meeting:

1. Discussion on Previous Week's Work:

- **Data Migration:** Successfully ported the entire data into PostgreSQL.
- **Data Ingestion Workflow:** Development in progress.
- **Backend Development:** Significant testing completed for PostgreSQL connectivity, with some testing still pending to ensure seamless background operation.
- **PATTIE Deprecated Functions:** Identified issues with deprecated functions in PATTIE, need to discuss further with Nibras.

2. Application Potential:

- Discussed the potential of the application and the use of dcdi.laihub.com.
- Considered the conference as a starting point for developing an Archix-type hub.

3. Coordination with Nibras:

- Need to discuss strategies with Nibras to avoid redundant work on PATTIE's newer versions.

4. Frontend Improvements:

- Discussed enhancements, particularly the author options.

- Suggested adding a "show authors" option under topic clusters similar to the "Show documents" feature.

5. Topic Clustering and ChatGPT Integration:

- Explored expanding topic clustering into phrases using ChatGPT.
- Emphasized the need for prompt engineering to utilize GPT effectively across multiple sections.

6. Database Automation:

- Mentioned the potential of automating database population and its inclusion in the paper.

Action Items to be done and reviewed in the upcoming weeks:

- Complete backend application functionality.
- Obtain the updated version from Nibras.
- Finalize the data ingestion workflow.
- Write an update for the author toggle to show authors instead of papers.
- Begin work on topic modeling after the author version is set up.
- Modify PATTIE to reference PostgreSQL instead of PubMed.
- Use LLMs to improve cluster labels and expand descriptions.