



6th May, 2024 - Scatter Gather

Minutes of MeetingDate: May 6, 2024

Attendees: Aravind Narayanan, Prof. Javed Mostafa

Subject: Development of Scholars Database and Integration with PATTIE for DCDI Conference

Agenda:

1. Discussion on setting up a relational database for DCDI.
2. Integration and expansion of the PATTIE platform.

Minutes:

1. Database Access and Setup:

- A relational database will be used to store data about known authors, scholars, and researchers, specifically for the DCDI. This includes names, titles, affiliations, and research areas.

2. Data Ingestion:

- The initial setup will focus on ingesting basic details into the database and modifying that part of the code in PATTIE. After ingestion, the system will categorize data based on keywords related to digital humanities, such as historical data, anthropology, and literature.

3. Digital Methods and Functionality Expansion:

- The functionality of PATTIE will be expanded by starting with author data and gradually including more detailed topics. The topics from the database will be mapped onto an ontology for effective topic modeling.

- Use Friends of Friends data structure to start with.

4. Handling Noise in Data:

- While the author's webpages will be utilized to determine their areas of interest, the potential noise in data will be addressed by initially creating a large corpus and then refining it using an ontology of key terms relevant to digital humanities.

5. Scholars Database (Scholars DB) Development:

- Will explore native options on Google Cloud for the relational database (RDB) to ensure seamless integration and avoid unnecessary code rewriting.

6. PATTIE UI Enhancements:

- Review the existing PATTIE UI code to explore possibilities for normalizing data into a 2D space for better interpretation of proximity and relationships among scholars, even at the expense of visual appeal.

7. Emphasis on Mapping and Distance:

- It is crucial to emphasize the accurate mapping of distances between topics and scholars to enhance the utility of the PATTIE platform in understanding the digital humanities landscape.

Action Items:

- Set up a relational database for ingesting data relevant to digital humanities.
- Develop a strategy for data ingestion and refinement using an ontology.
- Assess and modify the PATTIE UI for better data visualization and interpretation.

Next Meeting:

- Next two weeks Aravind will have to travel so updates will be given through google workspace with primary work being done on reading than coding.

Notes:

- The team should ensure that all developments are in line with the objectives of the DCDI Conference and the broader goals of digital humanities research.
- Do not rewrite code unless really necessary.