Final Year Project

Scholar - Graph

Team :

1. Neha Kundan Sonkamble : 21510007
2. Sourabh Yashwant Chaugule : 21510037
3. Trilok Vikram Kulkarni : 21510119
4. Maria Shaikh : 21510102

Guide : Mr. Shailesh Patil

Content :

Abstract :

- In today's digital age, researchers face information overload with thousands of research papers published daily. This makes literature reviews time-consuming and difficult.
- ScholarGraph is a semi-automatic platform designed to solve this problem. It helps users find, analyze, and visualize research papers using technologies like Natural Language Processing (NLP), web scraping, and knowledge graphs.
- Users can enter a topic, and ScholarGraph automatically retrieves relevant papers, extracts key terms, and analyzes them using methods like TF-IDF, similarity, and co-occurrence matrices. The results are stored in a graph database for clear, interactive visualization.
- This makes it easier for researchers to explore concepts, discover relationships, and quickly identify research gaps, saving time and improving the quality of their studies.

Introduction :

- ScholarGraph is a tool that helps researchers quickly find, analyze, and understand data present in research papers without getting overwhelmed by too much information.
- It uses technology to automatically collect papers on a chosen topic, pick out important terms, and show how these terms are connected using methods like TF-IDF, similarity, and co-occurrence matrices.
- The results are displayed in a clear, interactive graph, making it easy for researchers to explore ideas, see connections, and spot research gaps, saving them time and effort

Literature Review:

1. Open Information Extraction for Scholarly Texts (2015)
   - Platform / Method: ReVerb, OpenIE
   - Key Findings: Improved triple extraction from unstructured text for knowledge bases, QA, summarization.
   - Research Gap: Lacks contextual understanding & domain adaptation. Weak integration with dynamic knowledge graphs.

2. Entity Recognition in Scientific Literature (2018)
   - Platform / Method: BERT-based NER, SpaCy
   - Key Findings: Survey of NER approaches (rule-based, ML, deep learning) for scientific texts.
   - Research Gap: NER models struggle with domain-specific adaptation & noisy data handling.

3. SemOpenAlex (2023)
   - Platform / Method: RDF Knowledge Graphs, AI
   - Key Findings: Created a 26B triple RDF graph for global research landscape, promoting semantic interoperability.
   - Research Gap: Current graphs lack dynamic updates & comprehensive coverage for advanced analytics.

Methodology:

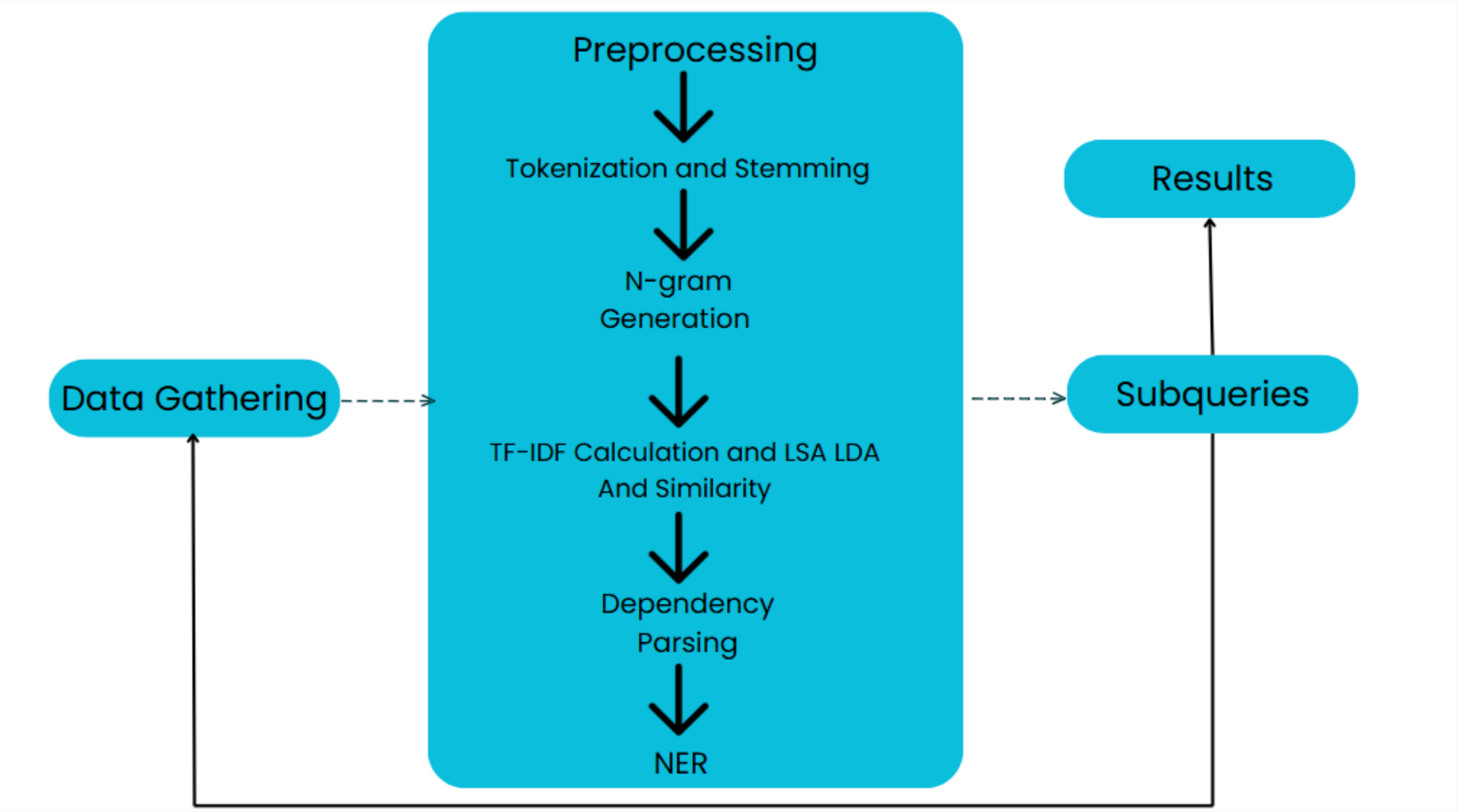


Diagram :

1. Block Diagram :

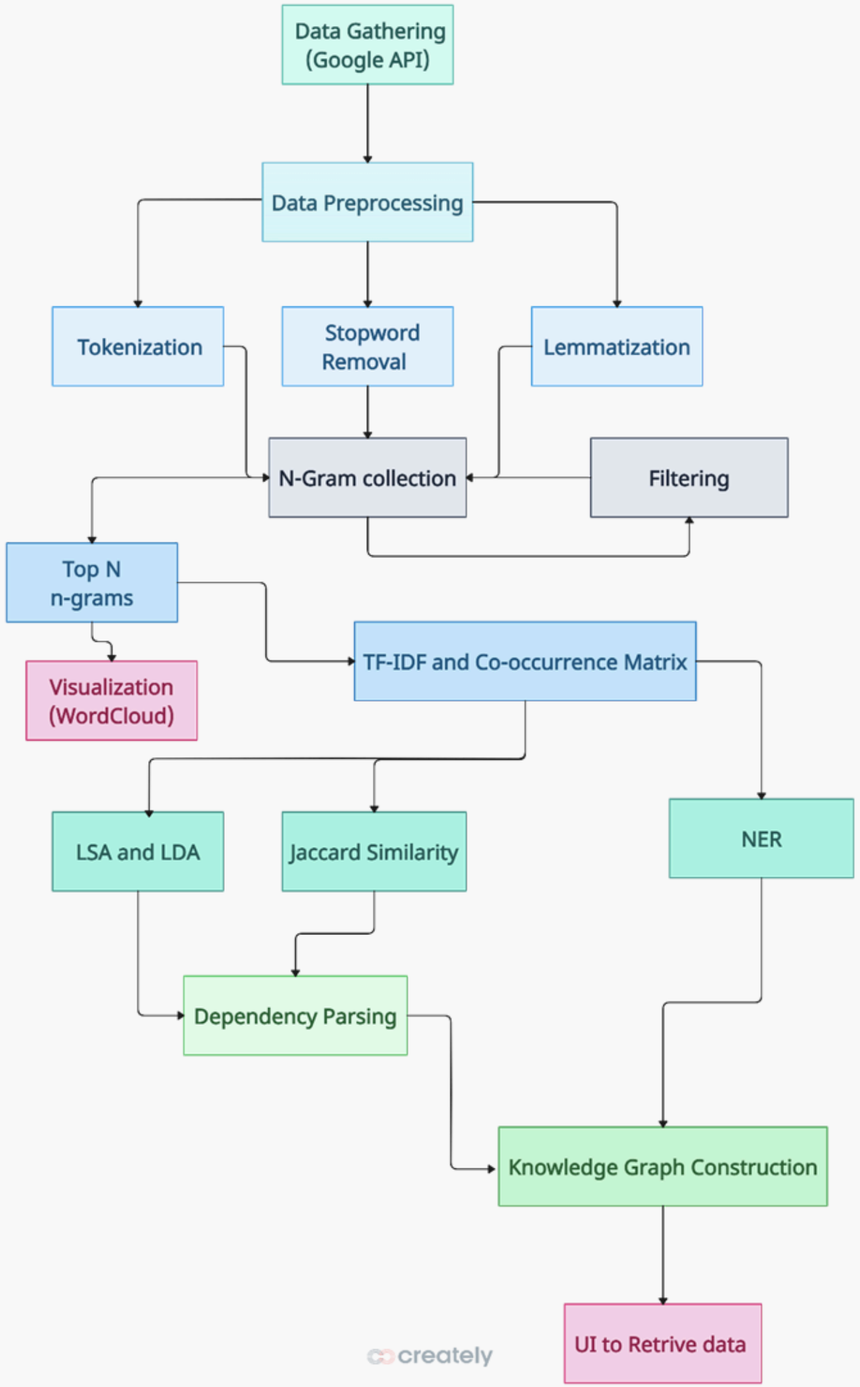
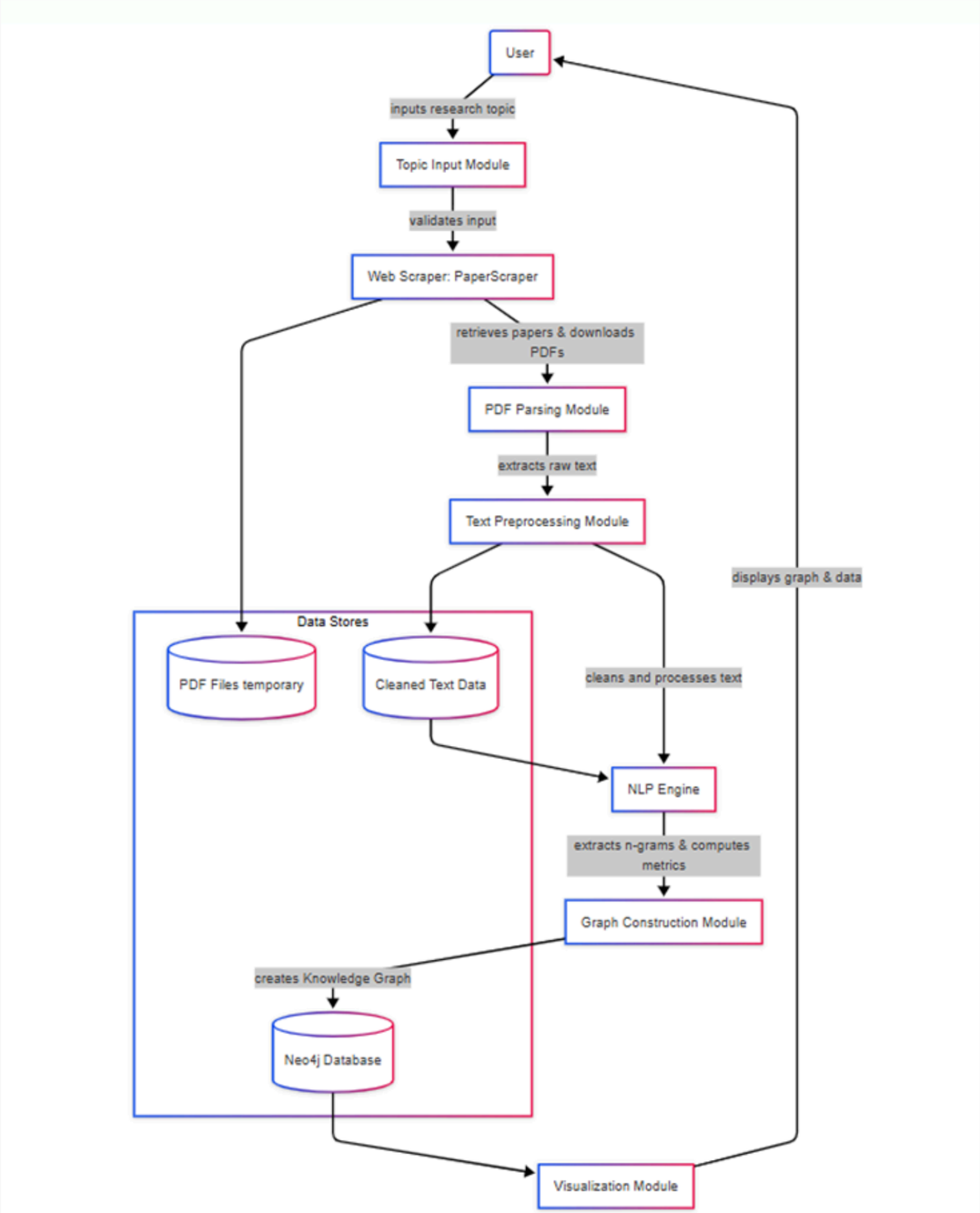


Diagram :

2. Flowchart :



Result and Demonstration

Conclusion and Discussion :

- ScholarGraph is a tool that helps researchers by finding and reading research papers automatically.
- It highlights important ideas and shows how they are connected using a simple graph. This makes complex topics easier to understand and saves time.
- But it has some limits—it works best with English papers, mainly uses Google Scholar, has trouble with poor-quality PDFs, and its graph doesn’t update by itself.
- In the future, it could support more languages, add more paper sources, update the graph automatically, let users customize the graph, and give quick summaries of each paper.


Thank
you!