Automatic Retrieval of Papers

I'm working on an NLP-based system for automatic retrieval and analysis of research papers, leveraging the Semantic Scholar API as data collection source to retrieve research papers. The project builds a citation network graph based on the data collected using a crawler. Which then builds a citation network graph based on the citations each paper cites to another paper or has, to explore relationships between papers. My focus is on enhancing relevance and usability by:

1. **Improving Paper Ranking:** Using embeddings and citation context analysis to rank papers based on semantic similarity and citation significance.
2. **Citation Network Analytics:** Identifying influential papers and authors, tracking topic evolution over time, and detecting research clusters.
3. **Knowledge Graph Construction:** Extracting entities and relationships from papers to build an interactive knowledge graph for deeper exploration.
4. **Contextual Summarization:** Generating tailored summaries of papers to help users quickly grasp their core contributions.
5. **Retrieve Papers:** Retrieve paper titles based on semantic search, in order to do that we will be embedding the queries and titles to find similar research papers based on that query. (This also includes fine-tuning an embedding model and testing out multiple models to evaluate the results.

The crawler I use is responsible for collecting the data and then creating a Citation Network graph. This citation network graph will be used as a database or a simple DataFrame, where I can analyse the network and identify the relationships between papers.

I'm not far with the results as i'm still analysing the network and relationships, thus cannot visualize anything at this moment. Like the citation network graph itself.