

# Citation Graph

## Research

Samantha Shultz, Sultan Alteneiji, Aaron Van De Brook, Luis Mora, Corrina Del Greco  
Department of Electrical Engineering and Computer Science  
Embry-Riddle Aeronautical University  
Daytona Beach, FL 32114

A software development team at Embry-Riddle Aeronautical University in Daytona Beach is designing a web-based tool with which a user can input research papers, view graphs of the citations, and add necessary info. The system will accept research papers that contain references in the LaTeX/BibTex format in order to output a visual dependency graph that shows the dependencies between the paper and its references. The purpose of the project is to create a tool which a user can use to see the dependencies between a paper and its references. The user will be able to view the dependencies between the papers and will be able to add to and modify information about the papers. An important step to this process is gaining knowledge in the field of this type of software and see what other developers have come up with. This paper will analyze two found open-source projects and discuss the similarities and differences from the group's own project and will discuss what new ideas they can now take for their citation graph development.

The first found project comes from Lanston Chu, a graduate student in Computer Science from the University of Wisconsin-Madison (Chu, 2020). Chu created a Python and a Java version very similar to the idea that the group originally designed. This design allows the user to input .bib files into a table where they are also able to edit and remove papers. The display gives a unique ID to that paper and displays its title, author, year, and references. The program also allows the user to generate a graph of these papers with various filters. These filters include displaying certain papers from specific years, showing certain authors, and showing papers based on their unique ID.

There is a vast amount of information that can be taken from this paper to aid in the group's research. The Embry-Riddle team will not have as sophisticated of a graphical user interface (GUI) in terms of flexibility with moving graph objects around freely and presenting status reports of all actions. The team does plan to use a table-like format to display all of the reports and also does intend to have a presentable graph displaying the papers that the user wants. Looking into this citation graph, the team was able to finally see their ideas in a real-life project that aided the group with coming together to see a final end goal.

The second project comes from Peter Prescott who is a PhD Researcher at the University of Liverpool (Prescott, 2019). Prescott created a Python version of the citation graph that also took in BibTeX files to create the graph with the use of the Tika toolkit and Zotero to generate the .bib files. This project is a lot more to the bare bones than the previous one was but does provide solid ideas of extracting script from PDFs and displaying a citation network of references.

The project from Prescott provides valuable information that the team can learn from even though their goal is to create a more sophisticated approach. Similarly, both projects allow the user to obtain information about the author's name, paper name, and date. The user is able to highlight over a specific paper to get the information that they need. However, in the group's project, the user will be able to update and edit information for each of the papers. There also will be security protocols put into place so that no one except for the development team and approved users will be allowed to make changes, which is something not shown in either of the found citation graphs.

Both found citation graphs gave the Embry-Riddle team a picture of where they want their project to be. Their ideas lie basically in the middle between the advancements of the two projects. The biggest take away from these projects is that security is not something that most of these projects consider and that the program needs to be very easy to use for the user. The team plans to take these two projects into consideration when making designs of their own project and to make something safer and more accommodating.

## References

- Chu, L. (2020). Automatic generation of academic citation graph (Java version). Retrieved February 15, 2021, from <https://lanstonchu.wordpress.com/2020/08/10/automatic-generation-of-academic-citation-graph-java-version/>
- Prescott, P. (2019). Citation graph of scholarly literature¶. Retrieved February 18, 2021, from <https://citation-graph.readthedocs.io/en/latest/>