American Physics Society Cross Citations Network Map

CSE 6242: Data and Visual Analytics

Rodrigo Da Costa*, Anthony Sena, Kamal Kishore Kashyap, Ryan J Miller

Summary

The data science and data visualization skills are now more important than ever in any research area specially in physics where navigating the body of knowledge, particularly the citations between publications and area of impact, remains a challenge for researchers.

Here we build a fully functional REST API with the data and performed an analysis to identify correlation between the many publication journals in physics. We hope that this work can help scientists across the world in improving their research skills and finding out critical papers that would be otherwise a challenge to find.

Current State

The current state of navigating though the physics knowledge consists of using free text search on specialized websites.

Researches must navigate a web of citations manually in the selected publications. This necessitates another set of searchers to find all relevant information. The process of performing a literature search has several challenges as identification of citing literature requires manual curation to obtain a full corpus of research. This is a time consuming manual labor.

Our Solution

Our solution provides a web-based explorer that leverages the cross citations network API and the APS data set loaded into MongoDB to provide quick and easy access to publications and their cross citations.

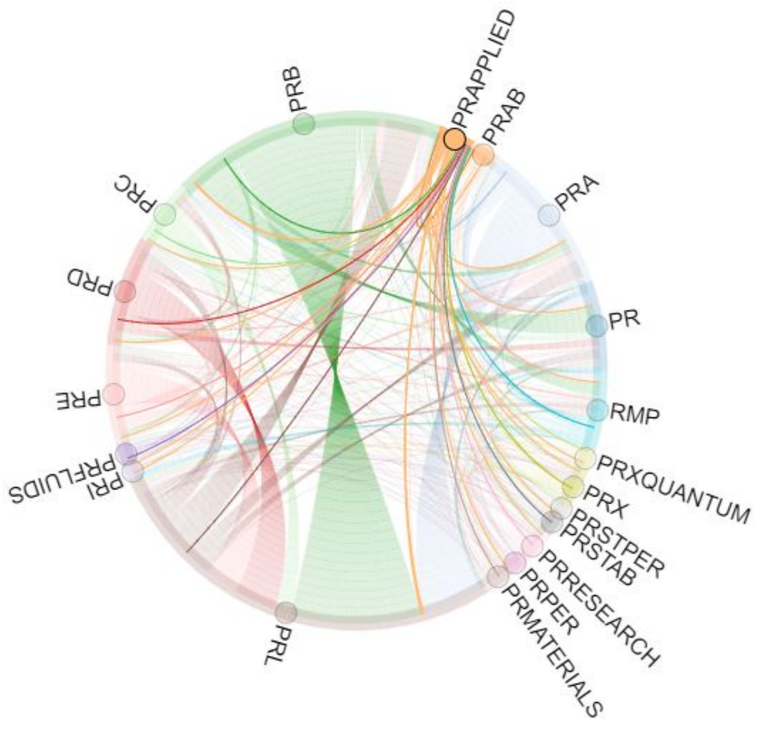
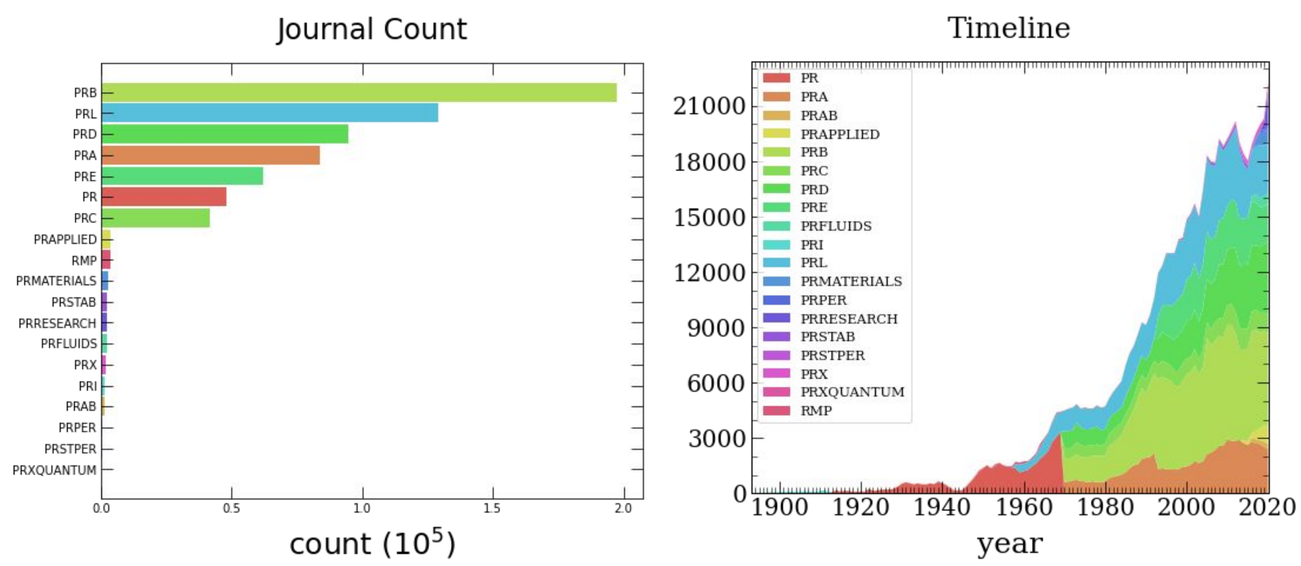


Figure 2 - chord plot that shows the proportion of publications, by journal, and their cross-citations amongst the network.

Data Overview

The American Physics Society (APS) curates a cross-citation map in a machine-readable format (JavaScript Object Notation - JSON) that we utilized to index the full set of physics publications.

- 678,916 nodes/publications since 1913 in **19 journals**.
- 8,850,334 edges/cross references.



(a) (b)

Figure 1 - Data Summary with (a) being the number of publications for each journal and (b) a timeline evolution.

Network API details

Our solution, provides an application programming interface (API) to work with a curated data set provided by the American Physics Society (see Data Source section for more details).

https://phiga-tech.herokuapp.com/api

The API provides programmatic access to digital object identifiers (DOI) entries for each journal publication. Key functions include:

- Load corpus by DOI (with citing and cited info)
- Load DOI list by affiliation (free search)
- Load DOI list by author (free search)

Figure 3 - choropleth map using d3js to visualize countries per contribution (volume of publication). This information is interactive and users can filter by specific journals at https://phiga-tech.herokuapp.com/map

* Corresponding author
email: costa@gatech.edu

Georgia Institute of Technology | APS physics