

Scientometric Mapping of Interpreting Evolution of topics, authors and collaboration's over time and space

Chandan Uppuluri
Indiana University

<mailto:chanuppu@indiana.edu>

Vinay Kumar Ranganath Babu
Indiana University

<mailto:vinranga@uemail.iu.edu>

Nisha Chadwani
Indiana University

<mailto:nchandwa@iu.edu>

Shruthi Ramakrishnan
Indiana University

shrurama@uemail.iu.edu

Abstract:

Interpreting is the domain of translating speech or signs from one language to another [1]. Although science mapping is relatively established in other domains, the field of “interpreting” has not looked at itself from a Scientometric perspective. The project is a first step toward gaining first insights into this academic field by analyzing the academic articles published in the main journal in the field, Interpreting, using Scientometric analysis methods and information visualization.

Description and Goals:

The client needs visualizations that gives insights about “domain of interpreting” by analyzing the articles that have been published related to this domain. We will be creating visualizations which gives insights about the evolution of topics, authors, collaborations and citations across time (when) and space (where: geo-spatial). Listed below are the various visualization we would be implementing for this project:

1.) Topic Analysis – Word Cloud:

The goal is to create a chart in the form of a world cloud. We will overlay topics data. Here is an example visualization:



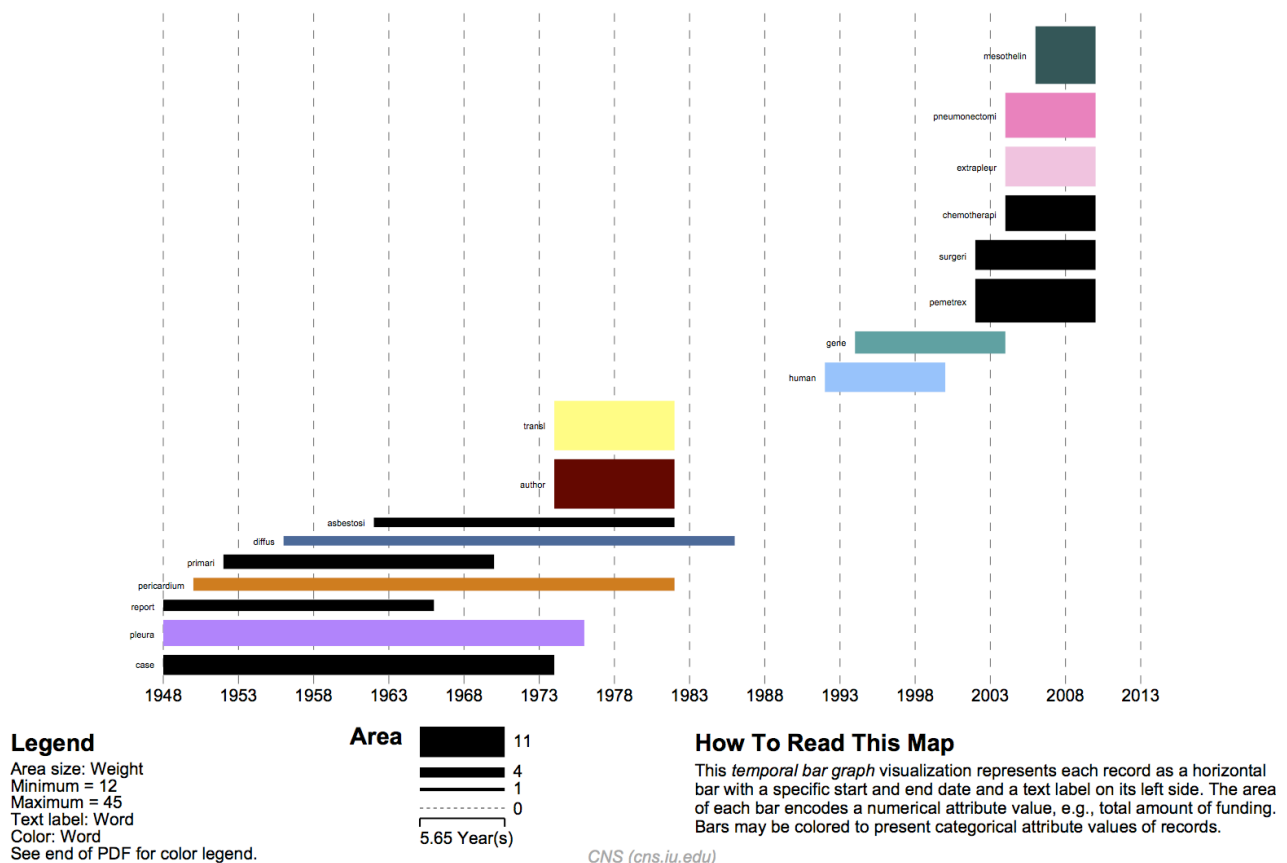
Figure 1: Word cloud example [2]

2.) Topic Analysis: Burst of terms across time

In this visualization, we can see the burst of topics in the domain of interpreting across time. Below is an example visualization:

Temporal Visualization

(Generated from CSV file: /var/folders/36/p3kwfcdx61b766km8_gp3gd40000gn/T/temp/Preprocessed-burst_gamma_3.3-615527217027125300.csv)
January 21, 2017 | 4:42 AM EST

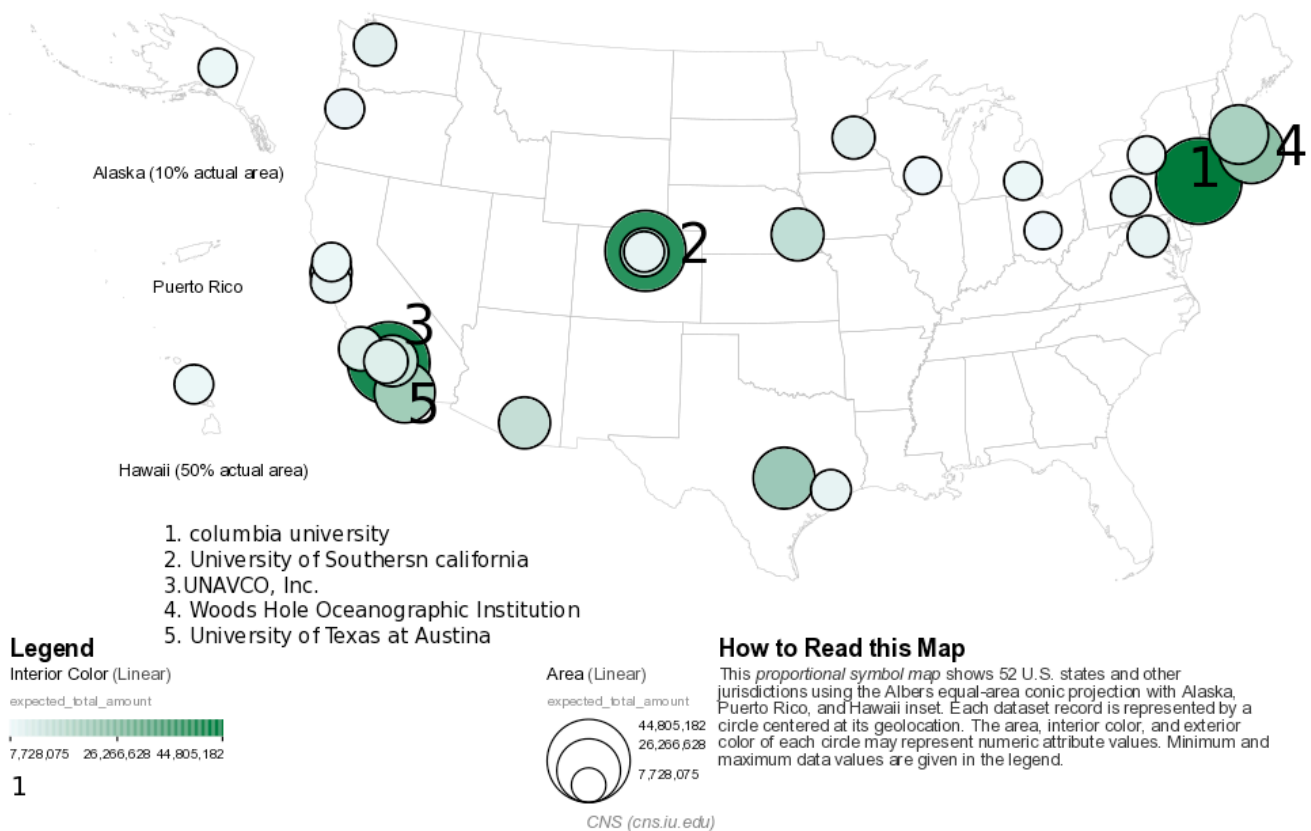


3.) Geospatial Location of Lead Authors:

The reference system will be a world map or any other map. We will be using proportional symbol map to overlay the data of number of authors that have contributed towards the domain of interpreting. An example proportional symbol map would look something like this: (It shows the universities that receive the most amount of funding. And size of the symbol is proportional to the amount of funding. In our case this will be the number of authors that have contributed from that university or the location of lead authors i.e the authors who have contributed the most.)

Geospatial Visualization (Proportional Symbol Map)

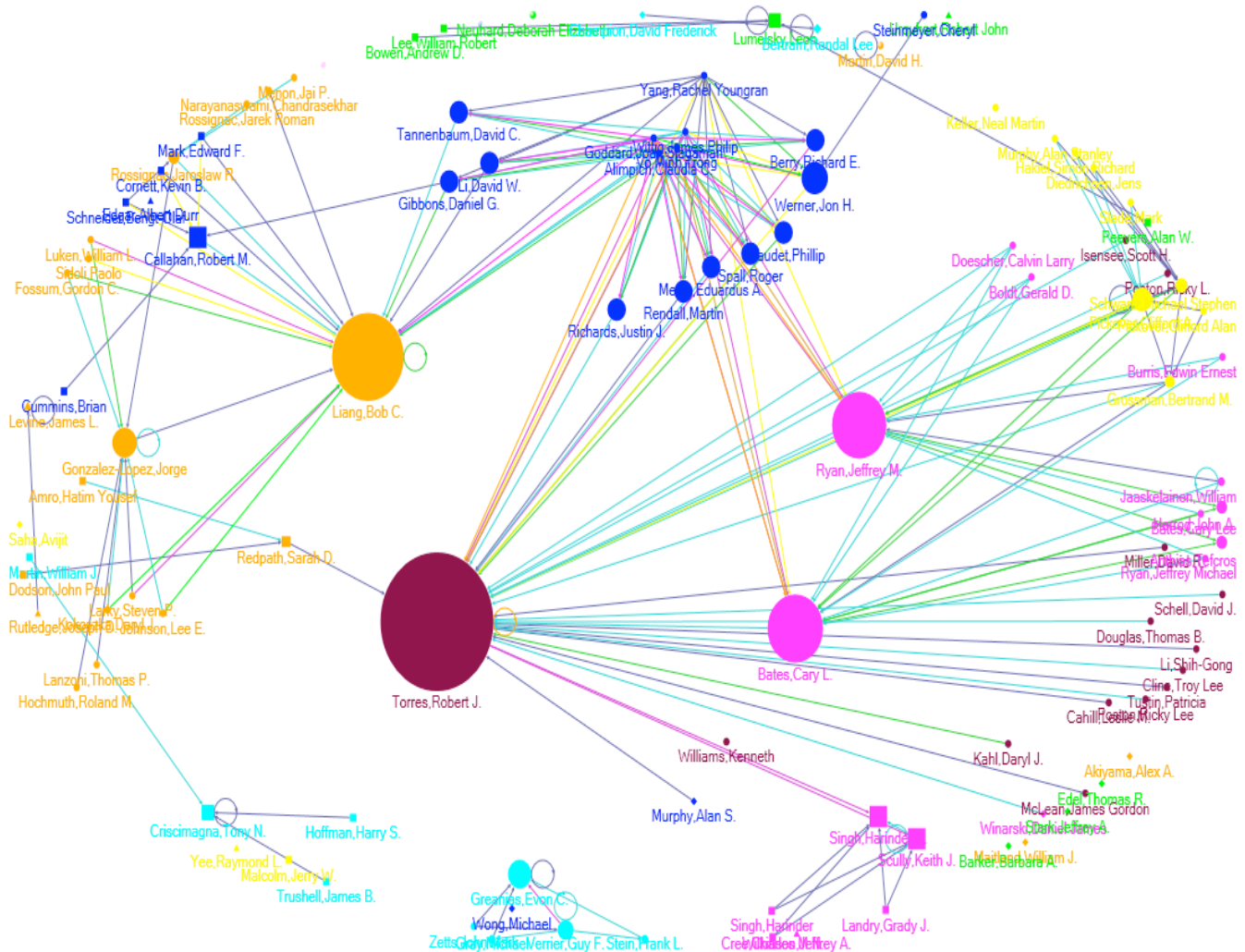
Generated from Top 30 row(s) (descending order) (based on expected_total_amount) of Aggregation performed using unique values in 'institution_name' column.
Jan 29, 2017 | 03:20:24 PM EST



4.) Paper Citation Network:

The reference system that will be used here is a network graph that shows the citations network of the authors. This depends on the dataset that we have and how much information it holds as the dataset is still tentative and not yet finalized. But here is an example of Patent Citation network of Computer Graphics Processing Group [3]

Number of Patent Citations among companies with images



Created with NodeXL (<http://nodexl.codeplex.com>)

5.) GeoSpatial location of topics across time:

This uses world map or Country maps as a reference system. We will overlay lead topics that have evolved at various locations across the world. We can also include one more factor which is the time. We can produce series of maps for various time frames.

Interactive Visualizations:

Though client is looking just for visualization's that just gives insights, a better visualization would always be the one with which we can interact. We can use the Shiny web app with plotly/ D3 or any other program for this purpose to implement interactive visualizations.

Various levels of interactions can be as follows:

- 1.) Time Toggle affect: With this we can select the time / year to see how the topics/authors have evolved during the particular time frame.
- 2.) Filters: Using filters we should be able to sub select authors, Countries, Topics etc
- 3.) Data Overlay: Based on the time/Filters used the appropriate data must be overlaid.

About the Data:

The required data for this project is obtained from the Scopus database. It has research papers related to various domains including the field of interpreting.

The search query is as follows: "interpreting studies" OR "conference interpreting" OR "court interpreting" OR "medical interpreting" OR "sign language interpreting" OR "community interpreting" OR "simultaneous interpretation" OR "consecutive interpretation".

This gave us around 3000 records of data. But it has a lot of noise. It Should be filtered to exclude publications related to statistics, remove non-articles such as books etc.

The following fields are present in the data:

- Author
- Title
- Year
- Cited by
- Affiliations
- Authors with affiliations
- Abstract
- Author Keywords

- **References**

For the geo-spatial mapping, we can use the Affiliation's filed. It has the university location address. This address can be used for geocoding to obtain the latitude and longitude which can then be used for spatial mapping.

Related Work:

There is not much of a work done in the Scientometric analysis of the interpreting studies domain. But we have found very few works related to this . In Xu and Pekelis_Chinese Interpreting Studies A Data Driven Analysis_Peerj 2015 paper [5] it discusses about how we can identify various trends in the interpreting studies using quantitative and qualitative analysis. But this study focusses on mainly the Chinese works. Doors and Gambier Measuring relationships in Translation Studies Perspectives 2015 [6], in this study they mainly focus on the geographical spread of the Translation and Interpreting Research using Affiliations.

Discussion of challenges and opportunities:

- 1.) The major challenge is to obtain a clean data set as there is a lot of possible noise. But by refining query terms in Scopus database we can obtain a better dataset.
- 2.) Another challenge is to implement an interactive visualization's. Using appropriate tools such as shiny web app/plotly/D3 etc should help us build interactive visualizations easily.

Acknowledgements:

Lluís Baixauli-Olmos, Assist. Prof., Dept. of Classical and Modern Languages at the university of Louisville has helped with us in understanding the scope of the project and any other queries that we had. We have worked with LLuis closely to obtain the necessary dataset from the Scopus database.

References:

- [1] Interpretation or Interpreting wikipedia : https://en.wikipedia.org/wiki/Language_interpretation
- [2] World Cloud example image: <http://www.jsquaredanalytics.com/word-clouds/>
- [3] Citation Network wiki: https://wiki.cs.umd.edu/cmssc734_11/index.php?title=File:Citation_IBm.png
- [4] Sci2 Team. (2009). Science of Science (Sci2) Tool. Indiana University and SciTech Strategies, <http://sci2.cns.iu.edu>.
- [5] In Xu and Pekelis_Chinese Interpreting Studies A Data Driven Analysis_Peerj 2015
- [6] Doors and Gambier_Measuring relationships in Translation Studies_Perspectives 2015.

