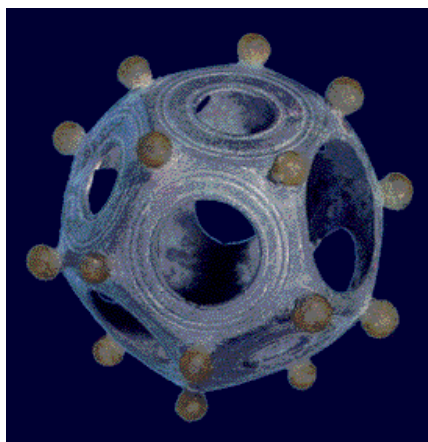




Indiana University

Project Report – Information Visualization



MAPPING THE KNOWLEDGE LANDSCAPE OF DESIGN

Temporal growth and topical analysis to map evolution of collaboration networks in areas of design

Project Team

Name	Email	Role
Apurva Singhvi	asinghvi@indiana.edu	Project Lead
Akansha Deshpande	akandesh@iu.edu	Analyst
Vinita Chakilam	vinichak@iu.edu	Technical Analyst
Karthik Vegi	kvegi@iu.edu	Technical Analyst
Pavan Jaggampudi	pkjaggam@iu.edu	Business Analyst

1. Project Description

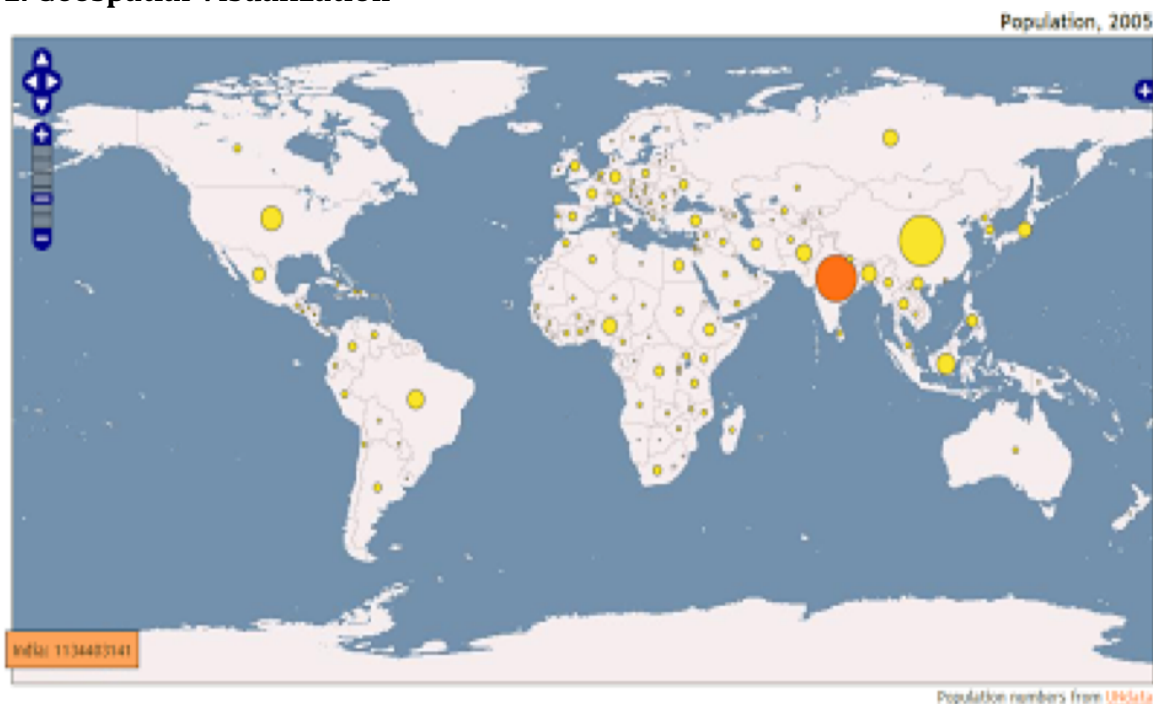
We trace the history of research and debate over the last three decades in areas of “design”. Design being a very vast subject, the work has both practical and scientific value to the diverse community to see the scope of design issues and research. This would allow the community to possibly come together. Students are invited to analyze and visualize the temporal growth and bursts of activity, to map evolution of collaboration networks, to overlay the data on a map of science so that changes in topical coverage can be understood and communicated.

Goals:

- Understand how research in areas of design has changed over time
- Identify the bursts of activity that happened in the last three decades
- Understand the collaboration of networks over time
- To map cluster of authors working on different topics, across topics and locations
- To map design research topics and their interrelationships
- To map geographic location of different research topics
- To be able to map the design research by specific journal collections.
- To map how much cross fertilization is taking place across these journals and domains.
- Interactive visualizations which can be viewed through webpage with options for different possible lenses

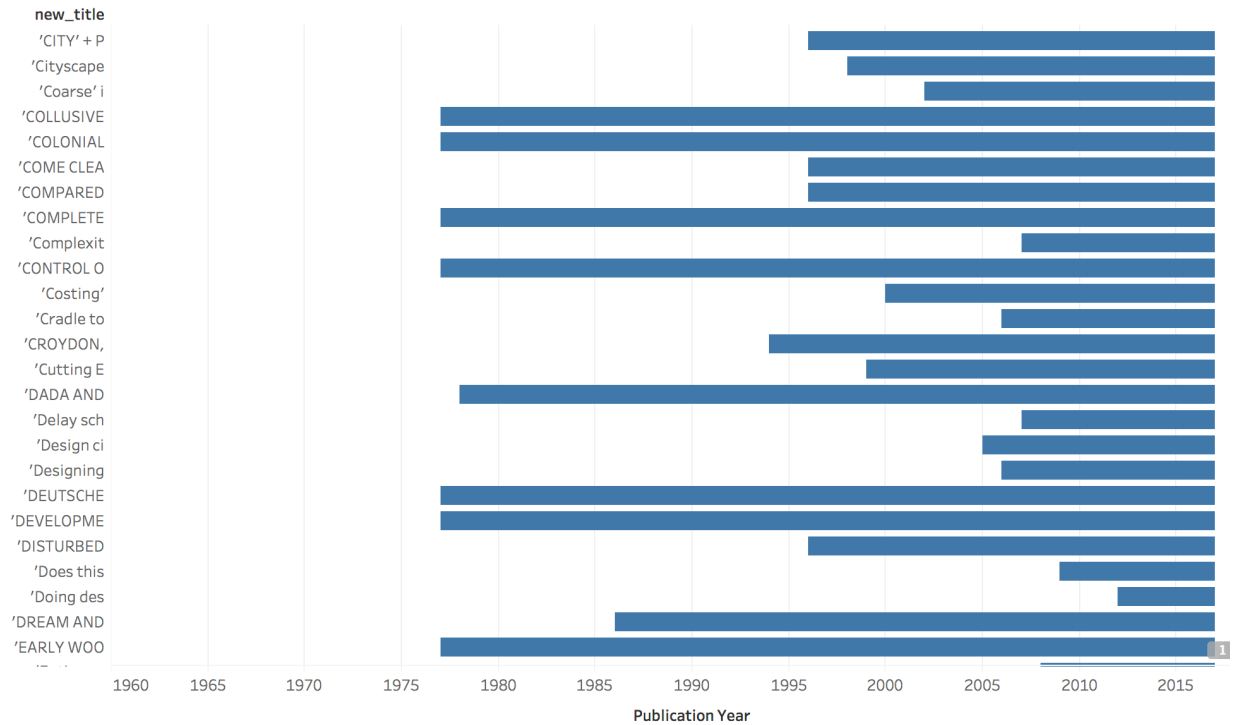
2. Envisioned Visualizations

1. Geospatial Visualization

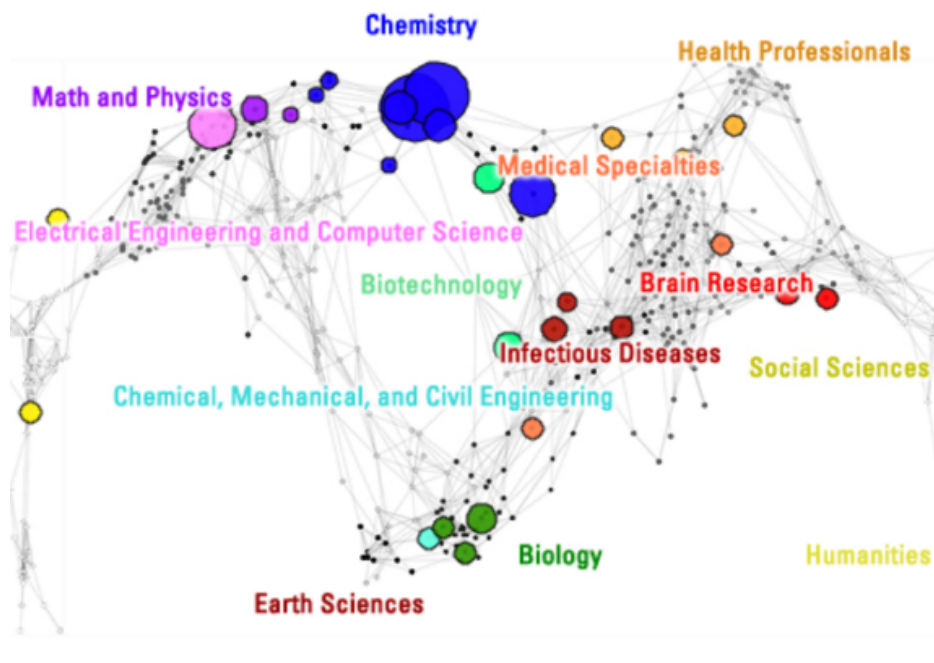


2. Burst Analysis and temporal growth

Temporal growth and topical analysis to map evolution of collaboration networks in areas of design

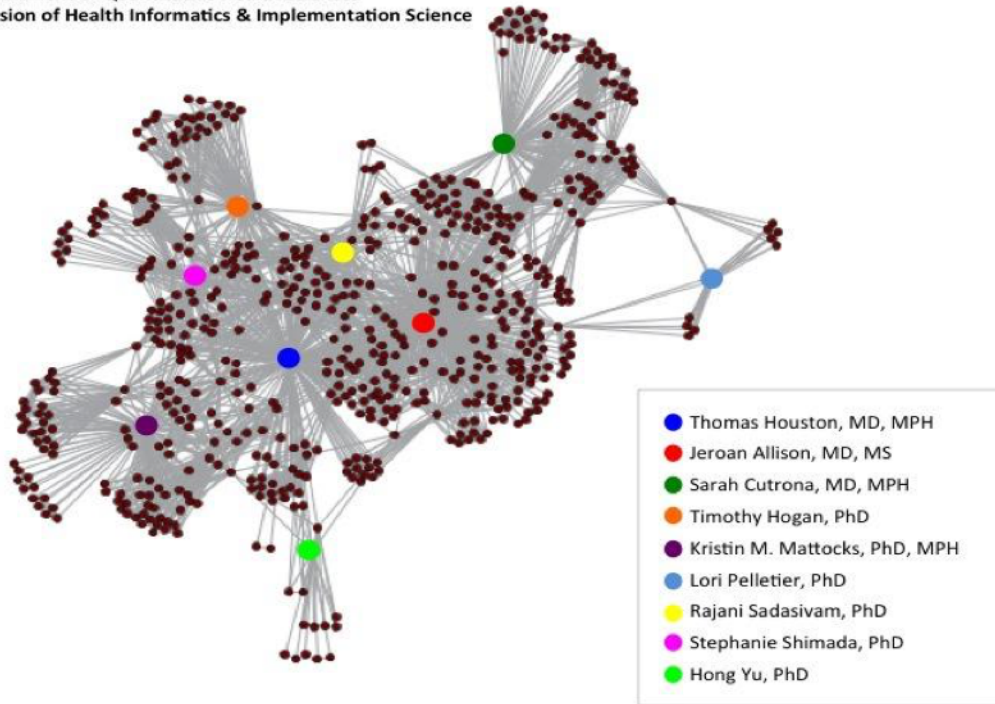


3. Collaboration networks over map of science



4. Co-author network

Department of Quantitative Health Sciences
Division of Health Informatics & Implementation Science



Importance of Visualizations:

- Geospatial visualization will help us to understand the locations of research happening in design across the world
- Temporal burst analysis will help us understand the topics in research that were prominent and the bursts of activity
- Collaboration network over a map of science will help us identify cluster of topics on which the research in design is focused
- Co-author network analysis will help us identify the cluster of authors collaborating

3. Related work

We studied the related work given by the client – ‘Mapping the Knowledge Structure of Research on Patient Adherence’ which was to understand the domain visualization based co-word analysis and social network analysis

Link: <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0034497>

We looked at the description, visualizations, metrics and the references to get a clear picture of what is expected out of this project. This helped us with the envisioned visualizations and the initial project requirements to get started with.

Additional related work referenced:

Reference Systems

- World Map as reference systems
- Use Map of Science as the existing reference system

Visualizations

- Proportional Symbol Map
- Co Author Network
- Burst Analysis

4. Tools Used

Given below are the tools that we have used for all the activities in this project:

- Microsoft Excel
- Sci2
- Gephi
- Tableau
- Bing geo-coding API
- R (maps package)

5. Dataset Statistics

- Publications are collected from 32 design journals - IUNI Web of Science dataset.
- Total dataset size: **54110** records
- Significant list of attributes:
 - WOS_ID (Unique Identifier)
 - Authors
 - Abstracts
 - Cited references
 - Count of cited references
 - No. of times cited
 - Document Type
 - Publication type
 - Journal Name
 - Language
 - Publisher
 - Publication date
 - Publisher address
 - Researcher id
 - Research address

6. Algorithms applied and resulting visualizations

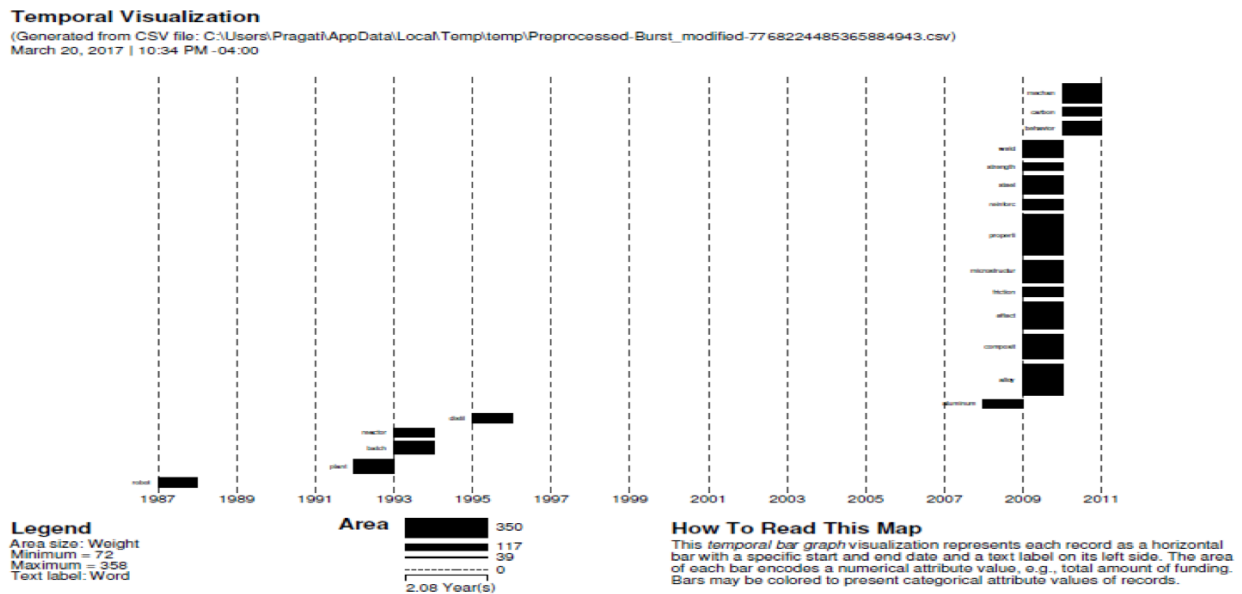
Some of the algorithms that were used during the project for pre-processing, analysis, validation and visualization are:

- Burst Analysis

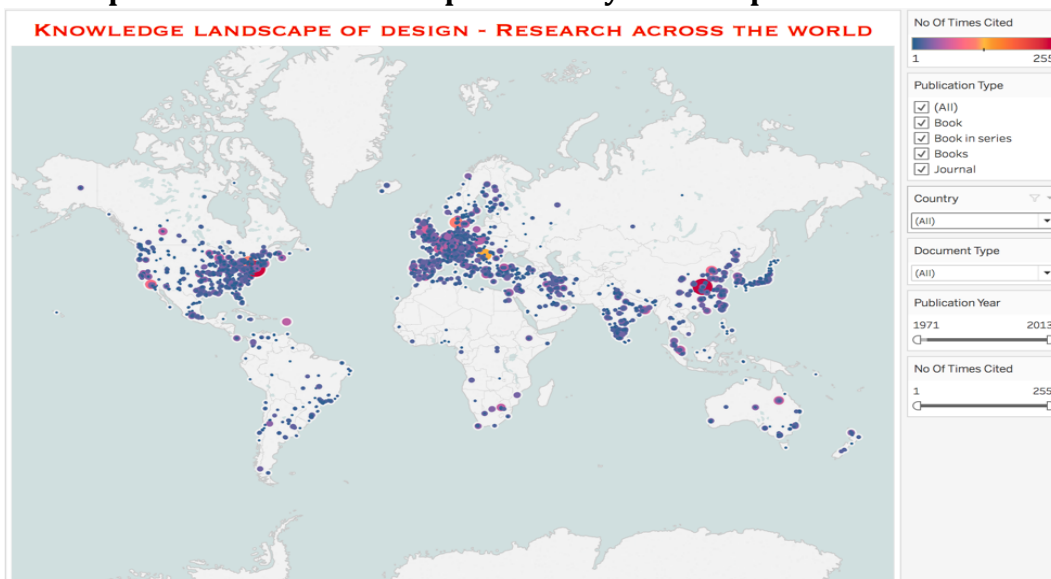
- Co-occurrence network
- Word co-occurrence network
- Proportional symbol map
- Network Analysis Toolkit
- Topical – tokenizer
- Geospatial – Bing geocoder

7. Resulting Visualizations

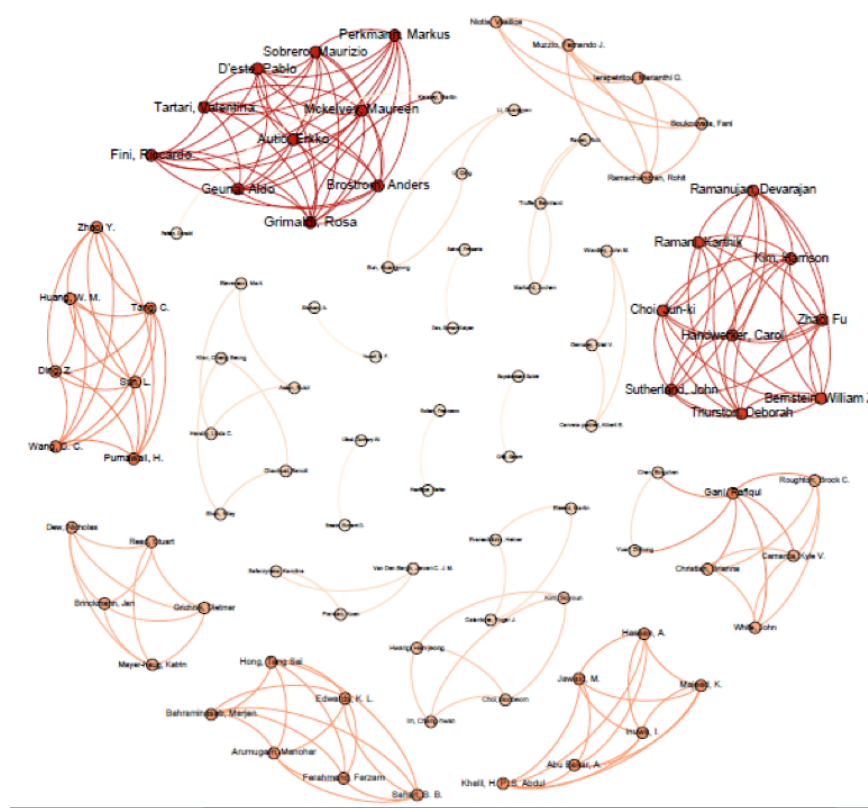
1. Temporal – Burst Analysis



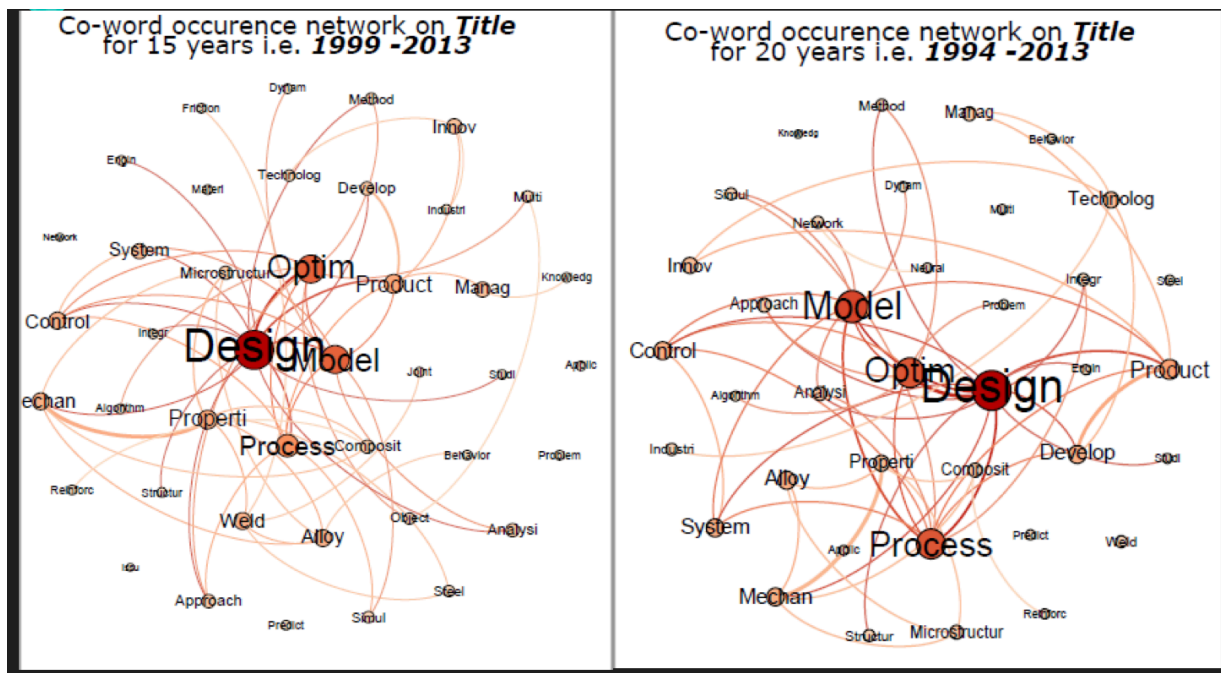
2. Geospatial – Interactive Proportional symbol map



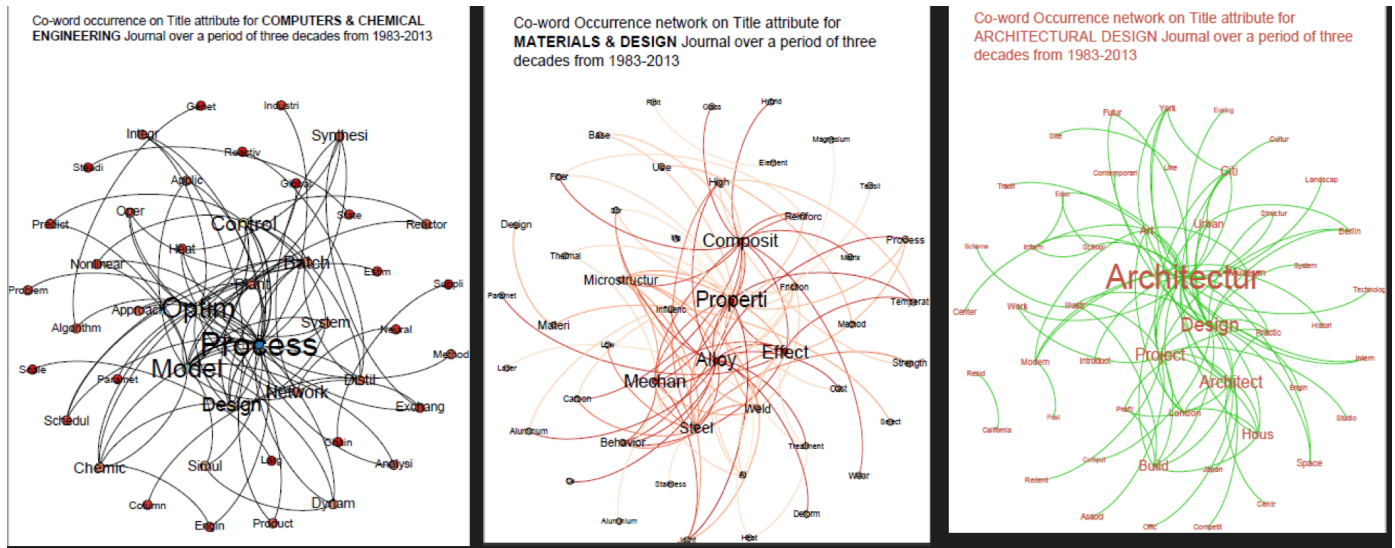
3. Co-author network



4. Co-word network based on titles



5. Co-word network based on journals



8. Key Insights gained from visualizations

Some of the insights gained from the initial visualizations are:

- Burst analysis has given a change of design landscape from the last three decades from robotics in 1987 to carbon recently
- Co word network has shown a strong correlation between design, process, model, product and optimization
- Proportional symbol map has shown a good amount of research done in USA and Germany in Europe. Dynamic visualization provided the research done in different lenses

9. Validation and Re-design

- Client wanted an interactive geospatial visualization, so we switched from Sci2 to Tableau
- We redesigned the proportional symbol map to make it interactive with a lot of data filters
- **Link:** <https://public.tableau.com/views/knowledge-landscape/geospatial>
- Client has asked for a steam graph and we are looking at the possibility of using D3.js for this visualization

10. Challenges & Opportunities

- Sci2 could not extract zipcodes of countries other than USA. This gave us an opportunity to look at other options like parsing the address field using **R** package *maps* and interactive visualizations in tableau
- **Data cleansing**
 - Unique Identifier is null

- Duplicate records
- Title contains numbers
- Removing the noise words like 'New', 'Paper' while making word occurrence network
- **File Loading Issues**
 - Not enough Heap space (87 MB)
 - File Format (IUNI Dataset)

11. Acknowledgements

Firstly, we thank all our group members for taking ownership of the tasks assigned to them and coming up with the visualization and insights. We thank our client, Dr. Eswaran Subrahmanian for his valuable feedback, insights and incredible patience while answering our questions. We also thank our instructor Prof. Katy Borner, and our co-instructors Andreas Buckle and Michael Ginda for their continuous support and valuable feedback. We truly appreciate your endless help and knowledge.

12. References

- Proportional Symbol Map. Retrieved from <http://blog.mastermaps.com/2008/04/proportional-symbol-mapping-with.html>
- Map of Science. Retrieved from <https://sciencenode.org/visualization/mapping-publications-science.php>
- Sci2 Team. (2009). Science of Science (Sci2) Tool. Indiana University and SciTech Strategies, <http://sci2.cns.iu.edu>

Tableau

- <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0034497>
- http://onlinehelp.tableau.com/current/pro/desktop/en-us/maps_build.html
- <https://www.theinformationlab.co.uk/2014/12/11/show-symbol-maps/>
- <https://www.tableau.com/learn/tutorials/on-demand/maps-tableau>

Reference project

- <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0034497>