

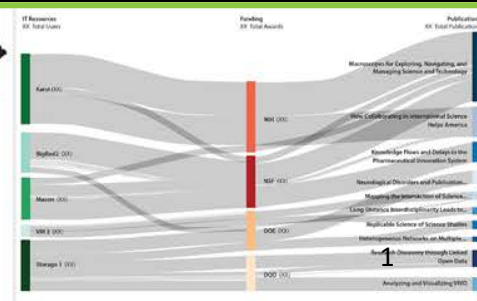
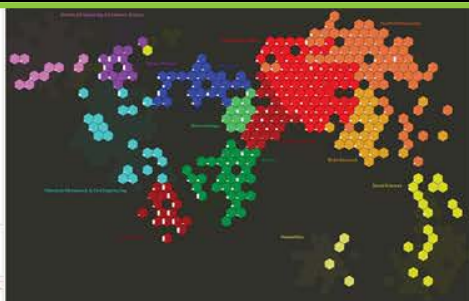
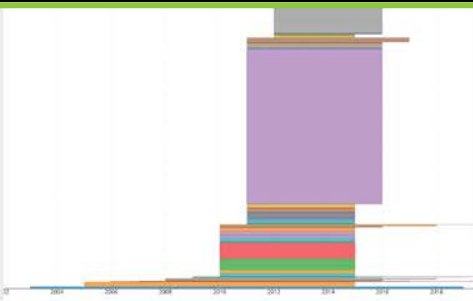
CNS Web Visualization Framework

Adam Simpson, Gagandeep Singh, Katy Börner (@katycns)

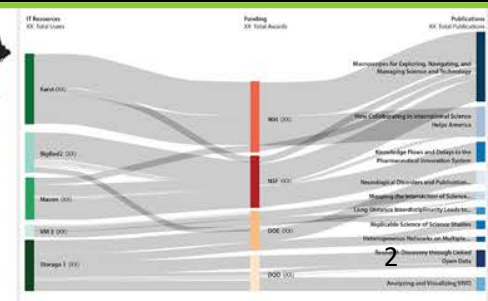
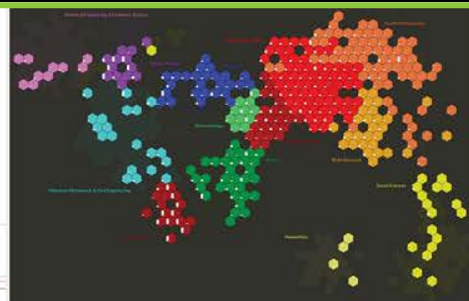
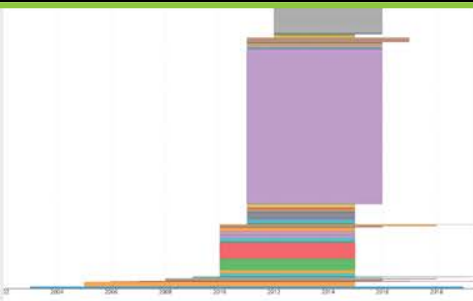
Cyberinfrastructure for Network Science Center
Department of Intelligent Systems Engineering
School of Informatics and Computing and IU Network Science Institute
Indiana University, Bloomington, USA

*Scholars' Commons IQ-Wall (E157H)
Wells Library, Indiana University Bloomington*

April 21, 2017 | 12:00 PM



Defining, Measuring, Improving Data Visualization Literacy

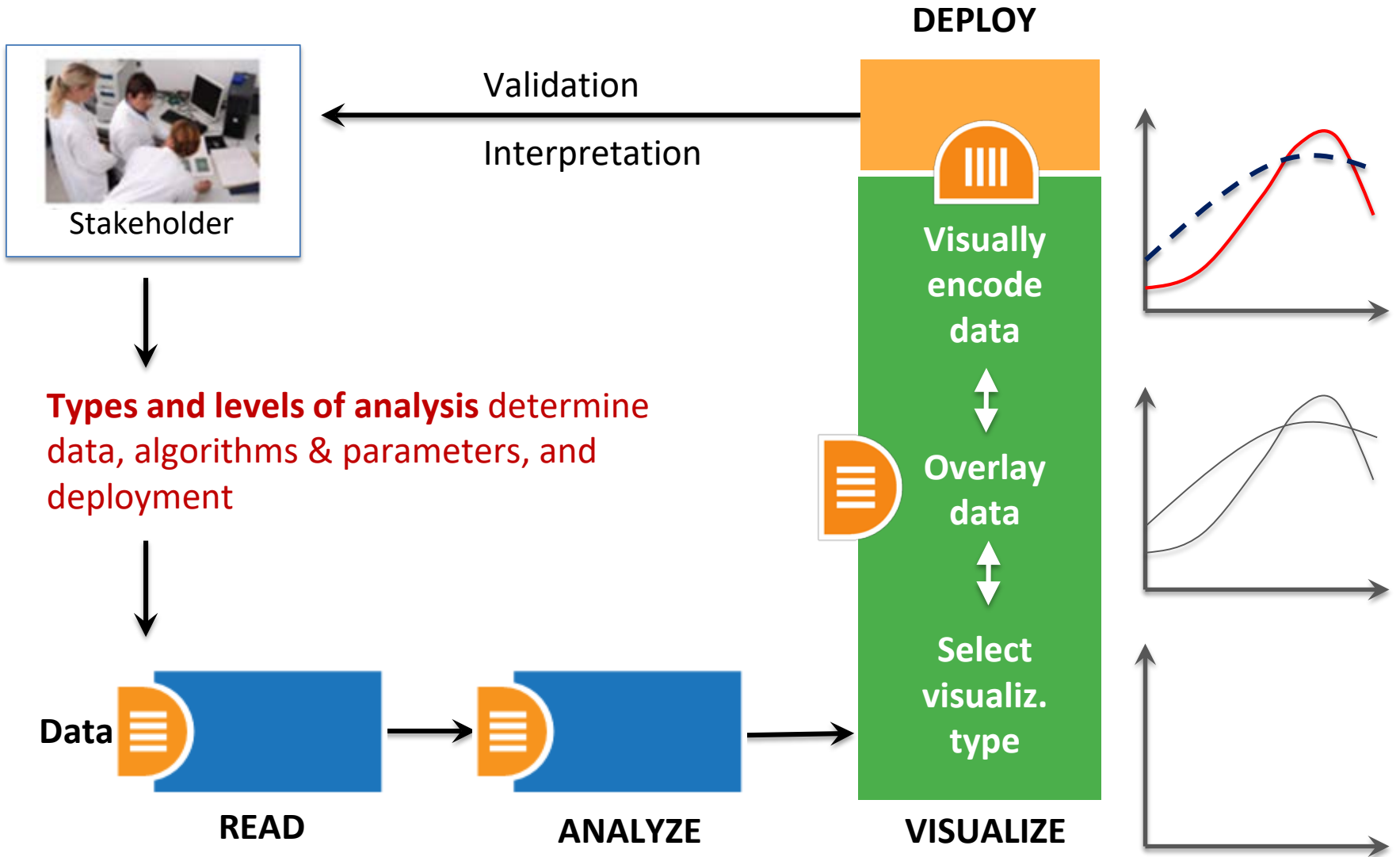


Data Visualization Literacy

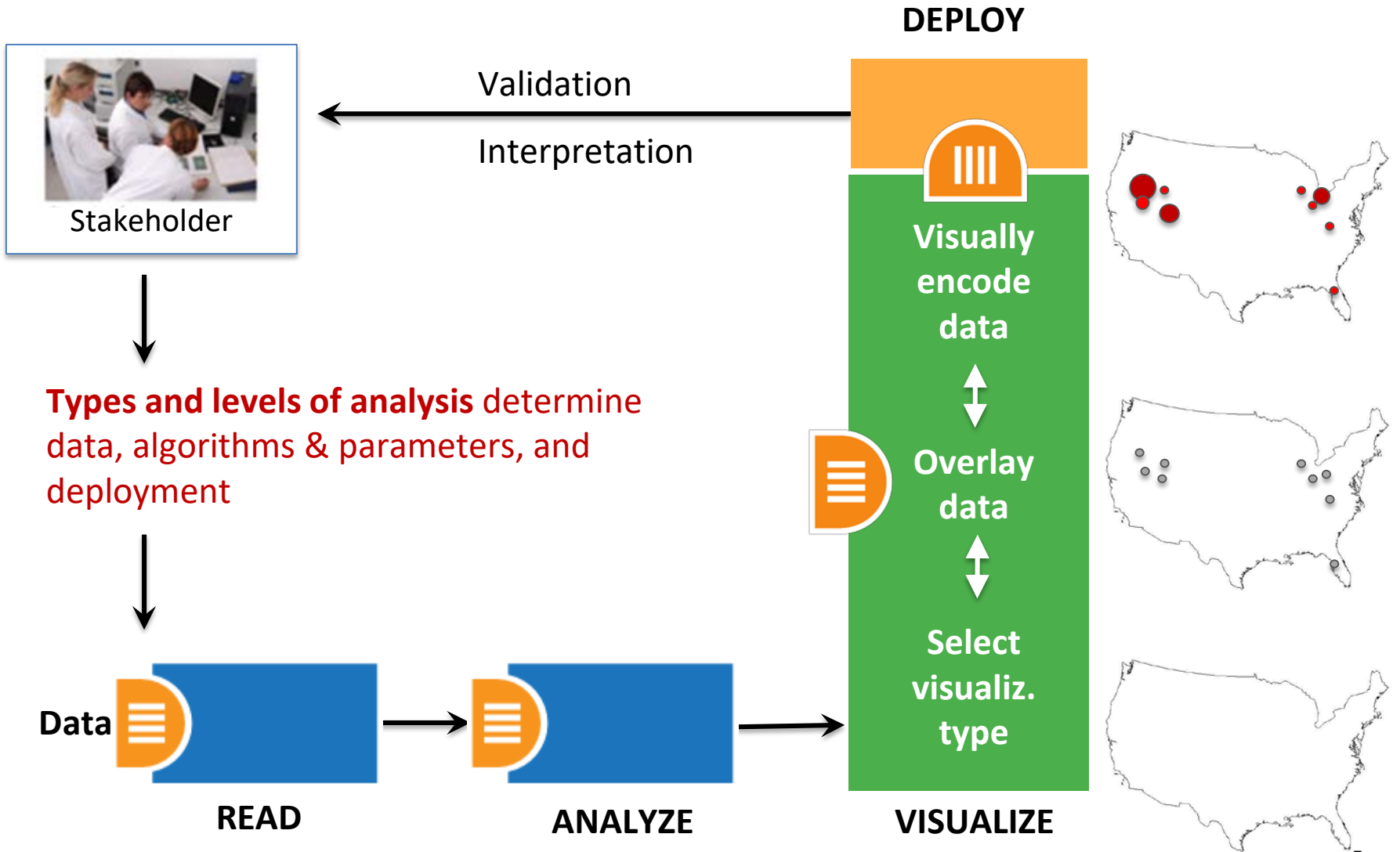
Data visualization literacy (ability to read, make, and explain data visualizations) requires

- *literacy* (ability to read and write text, e.g., in titles, axis labels, legend),
- *visual literacy* (ability to find, interpret, evaluate, use, and create images and visual media), and
- *data literacy* (ability to read, create, and communicate data).

Needs-Driven Workflow Design



Needs-Driven Workflow Design



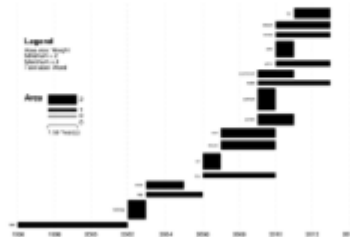
Load **One** File and Run **Many** Analyses and Visualizations

Times Cited	Publication Year	City of Publisher	Country	Journal Title (Full)	Title	Subject Category	Authors
12	2011	NEW YORK	USA	COMMUNICATIONS OF THE ACM	Plug-and-Play Microscopes	Computer Science	Borner, K
18	2010	MALDEN	USA	CTS-CLINICAL AND TRANSLATIONAL SCIENCE	Advancing the Science of Team Science	Research & Experimental Medicine	Falk-Krzesinski, HJ Borner, K Contractor, N Fiore, SM Hall, KL Keyton, J Spring, B Stokols, D Trochim, W Uzzi, B
13	2010	WASHINGTON	USA	SCIENCE TRANSLATIONAL MEDICINE	A Multi-Level Systems Perspective for the Science of Team Science	Cell Biology Research & Experimental Medicine	Borner, K Contractor, N Falk-Krzesinski, HJ Fiore, SM Hall, KL Keyton, J Spring, B Stokols, D Trochim, W Uzzi, B

Statistical Analysis—p. 44

Location	Count	# Citations
Netherlands	13	292
United States	9	318
Germany	11	36
United Kingdom	1	2

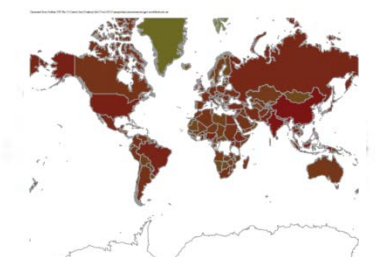
Temporal Burst Analysis—p. 48



Geospatial Analysis—p. 52



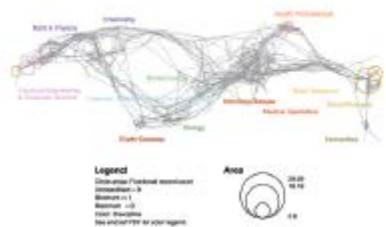
Geospatial Analysis—p. 52



Load **One** File and Run **Many** Analyses and Visualizations

Times Cited	Publication Year	City of Publisher	Country	Journal Title (Full)	Title	Subject Category	Authors
12	2011	NEW YORK	USA	COMMUNICATIONS OF THE ACM	Plug-and-Play Macroscopes	Computer Science	Borner, K
18	2010	MALDEN	USA	CTS-CLINICAL AND TRANSLATIONAL SCIENCE	Advancing the Science of Team Science	Research & Experimental Medicine	Falk-Krzesinski, HJ Borner, K Contractor, N Fiore, SM Hall, KL Keyton, J Spring, B Stokols, D Trochim, W Uzzi, B
13	2010	WASHINGTON	USA	SCIENCE TRANSLATIONAL MEDICINE	A Multi-Level Systems Perspective for the Science of Team Science	Cell Biology Research & Experimental Medicine	Borner, K Contractor, N Falk-Krzesinski, HJ Fiore, SM Hall, KL Keyton, J Spring, B Stokols, D Trochim, W Uzzi, B

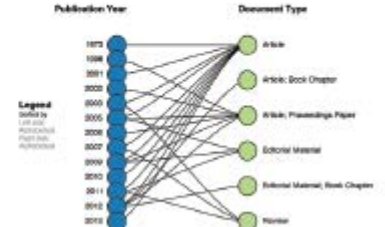
Topical Analysis—p. 56



Paper Citation Network—p. 60



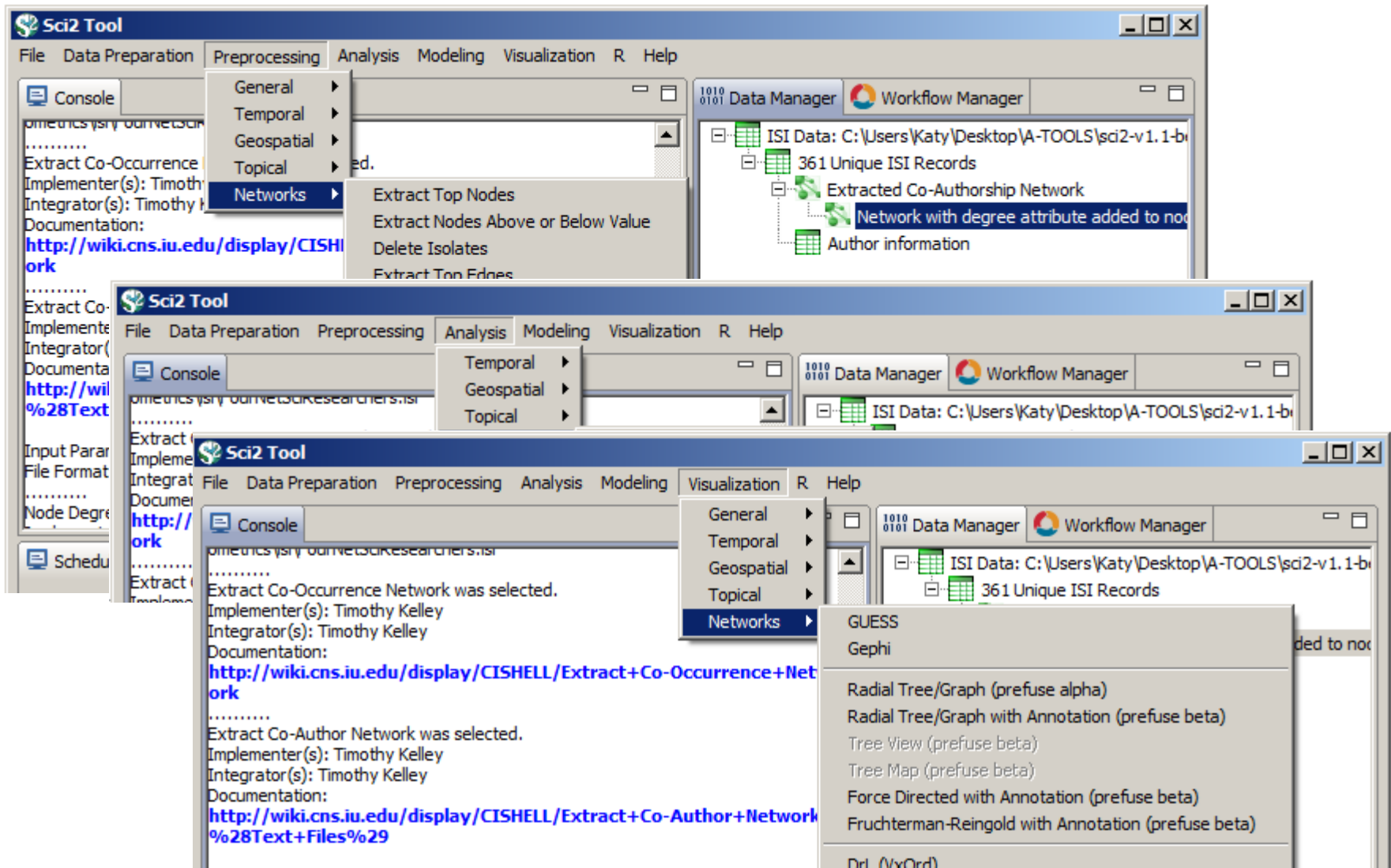
Bi-Modal Network—p. 60



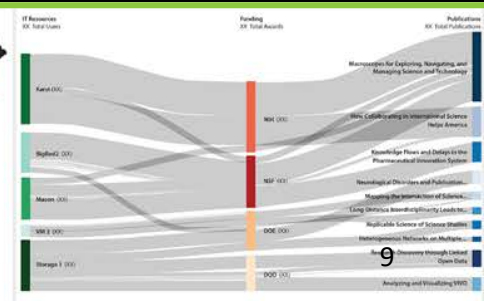
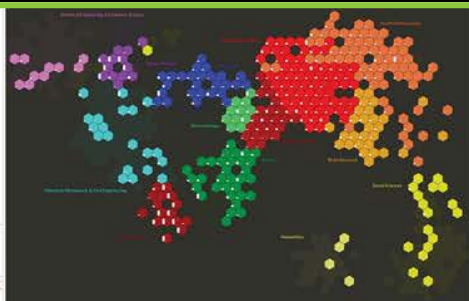
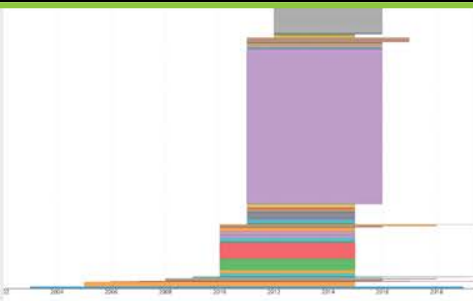
Co-author and many other bi-modal networks.

Sci2 Tool Interface supports A2 Visualization Framework

Download tool for free at sci2.cns.iu.edu



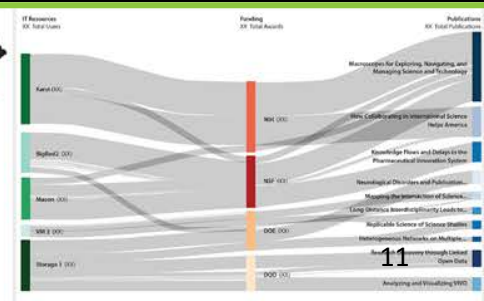
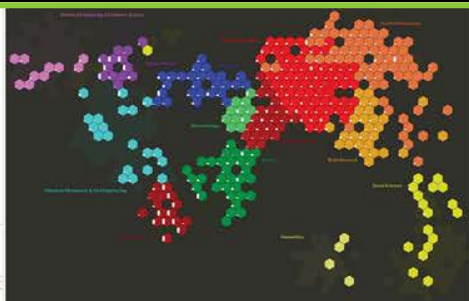
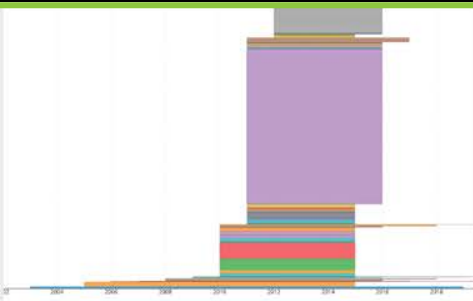
CNS Web Visualization Framework: Functionality



Web Visualization Framework

- Barebones visualizations
- Configuration-based customization
- Multiple stages of rendering
- Instanced visualizations
- Standardization of workflows
- Easy versioned project check-out and deployment

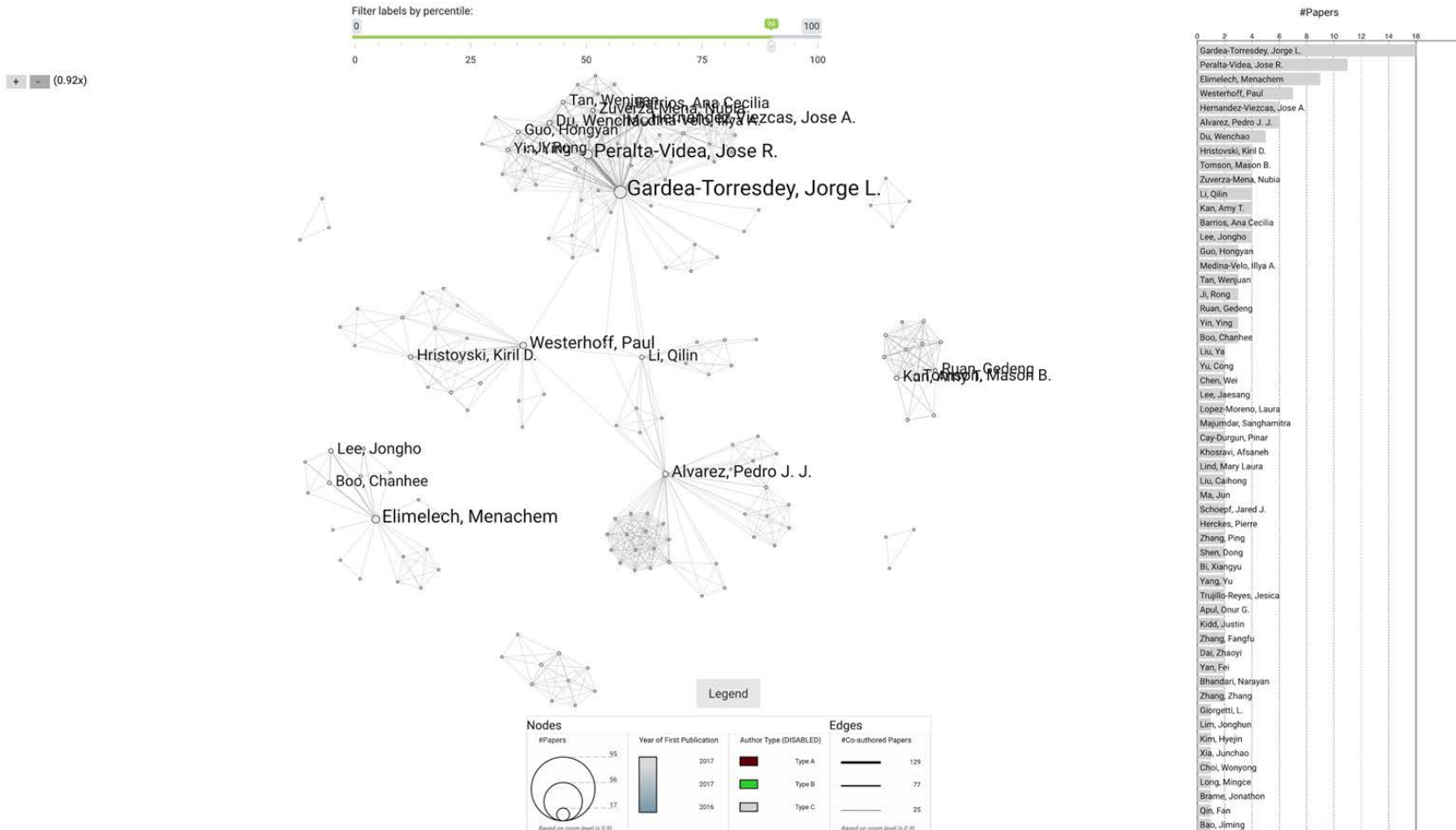
CNS Web Visualization Framework: Examples



Visualization: Co-Authorship Network

Project: ERC

nanohub.org/citations/curate

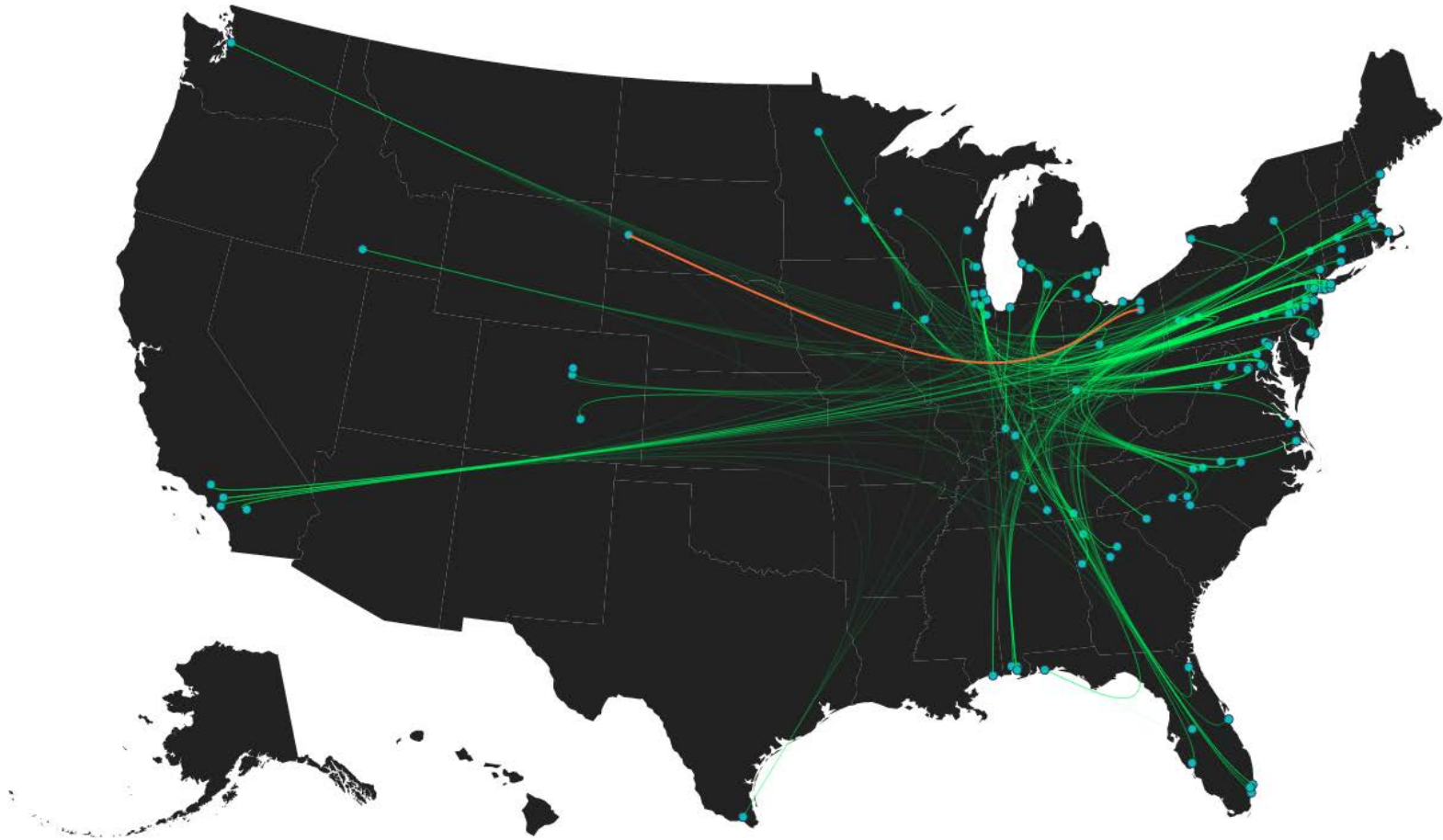


This figure shows the collaboration network of one ERC, based on co-authorship extracted from bibliography files. Each node is an author, and 2 authors are connected if they have authored a publication together.

Visualization: Geographic co-authorship visualization

Project: ERC

nanohub.org/citations/curate

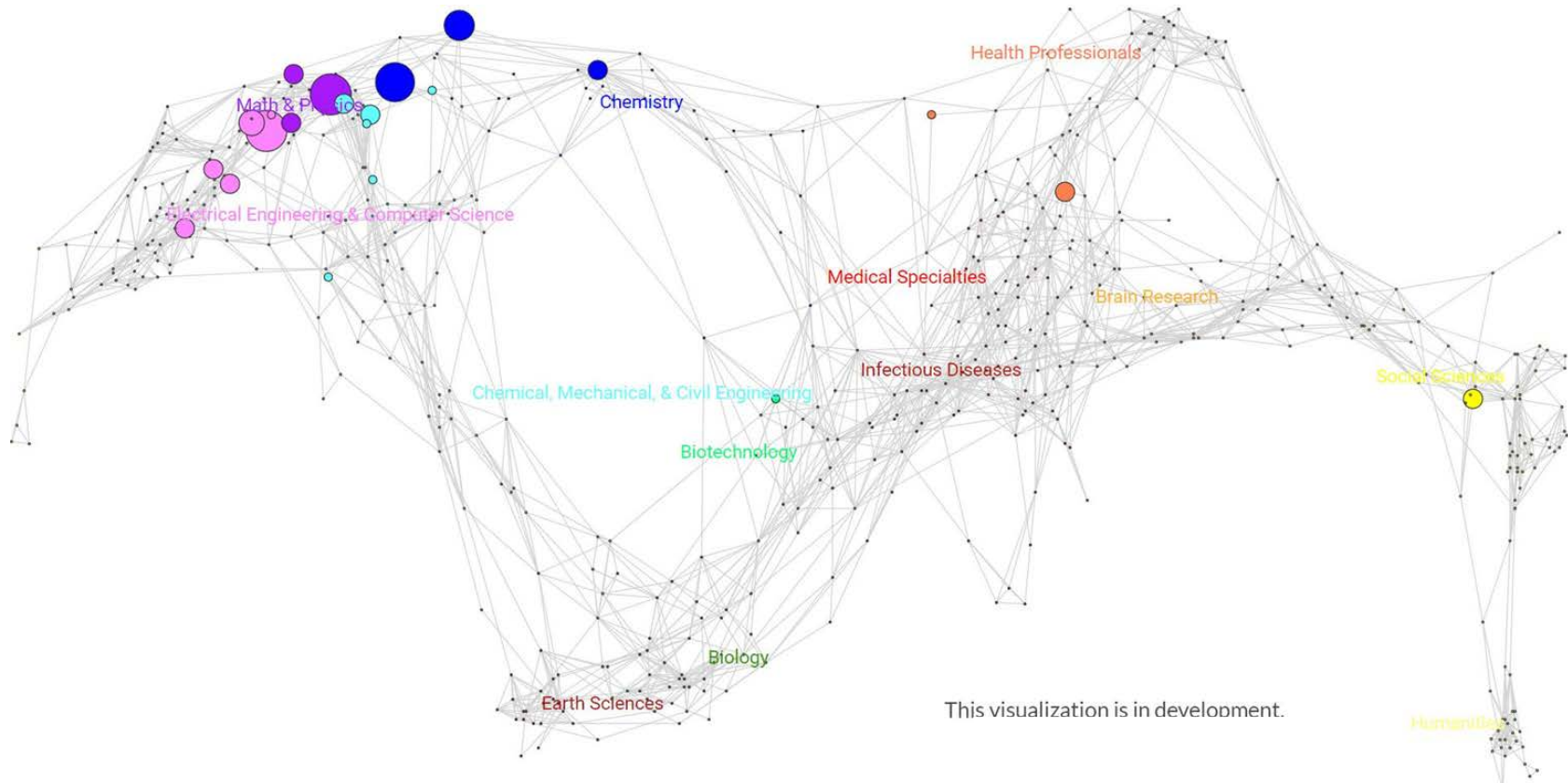


This map shows the co-authorship network overlaid on a geospatial map of the US. each node represents an author and two authors are connected if they have authored a paper together.

Visualization: UCSD Map of Science

Project: ERC

nanohub.org/citations/curate

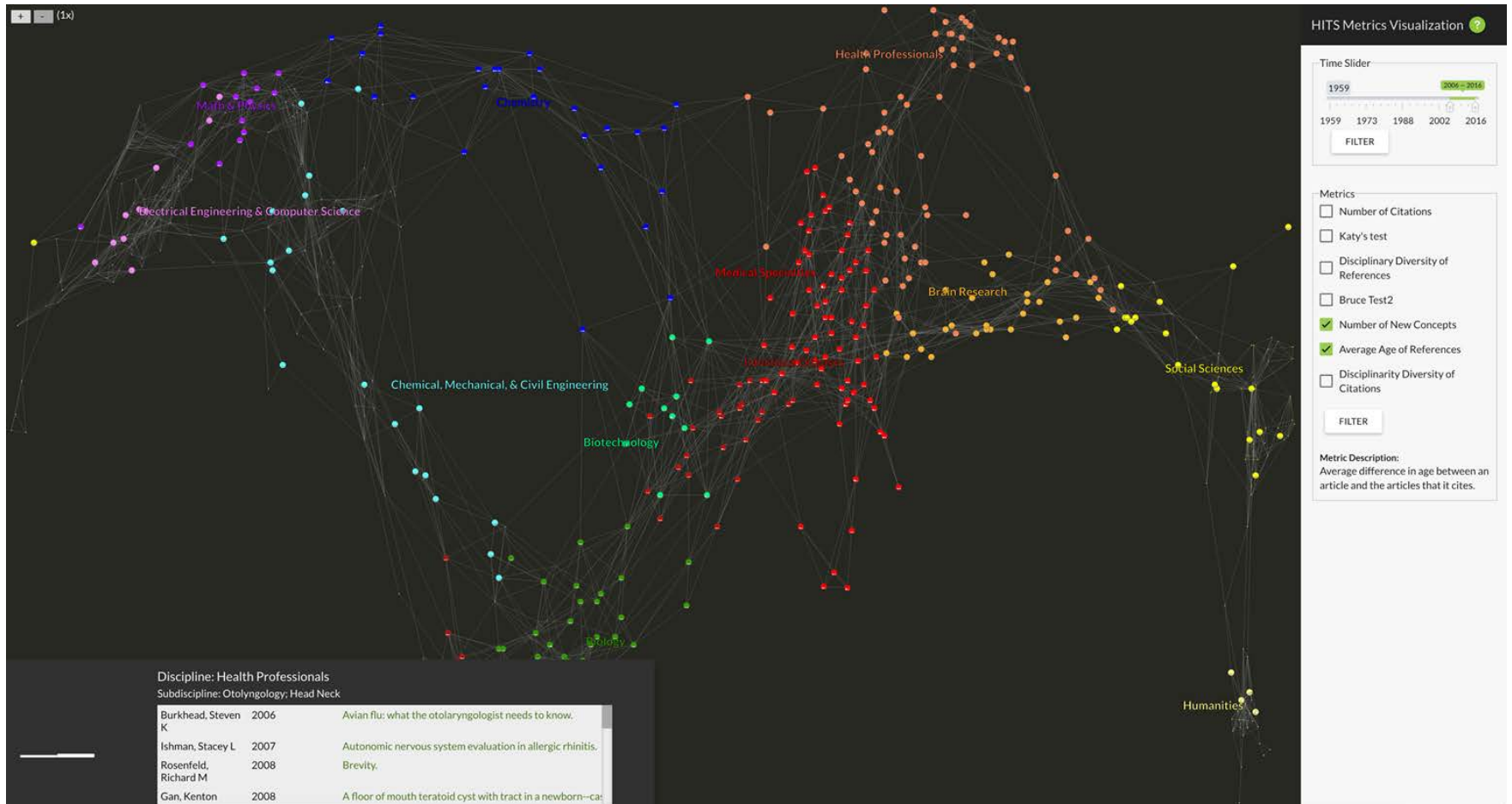


This organizes and visually represents 554 sub disciplines of science and their relationships to one another. Sub disciplines are grouped into 13 overarching disciplines that are color coded (red for medicine, green for biology, etc.) and labelled. Using a journal name based or keyword based mapping process, data overlays can be computed. For example, expertise profiles for an individual or an institution are generated by reading a bibtex or EndNote file with relevant publications, identifying unique journal names, and overlaying geometric symbols such as circles atop the sub discipline(s) that are associated with each journal. This Map of science can be used to explore, understand, and communicate the expertise profiles of an institute or nation.

Visualization: UCSD Map of Science

Project: ECON

demo.cns.iu.edu/client/econ-ucsdmap

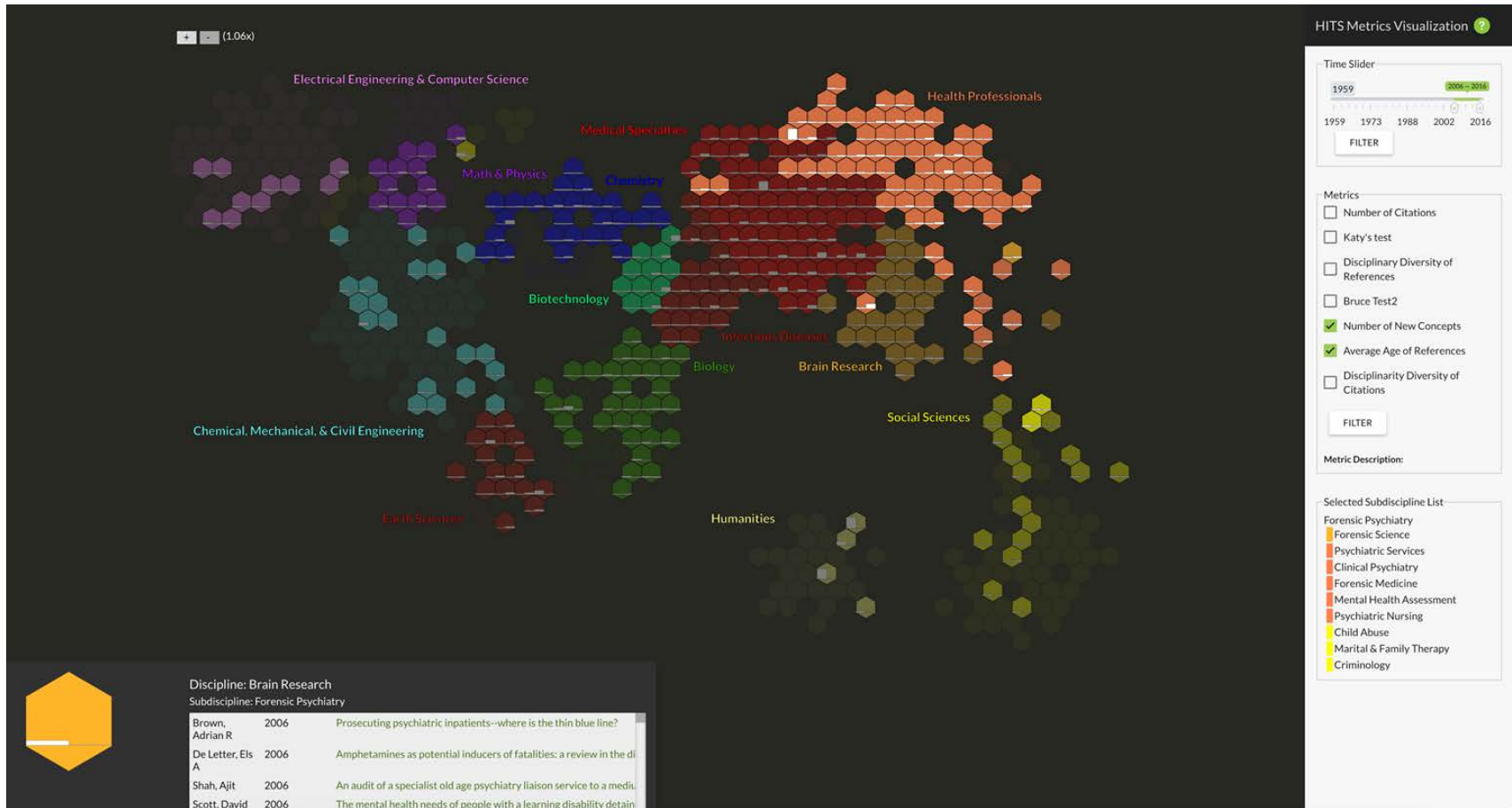


This interactive map supports the exploration and comparison of different high-impact and transformative science (HITS) metrics. Interactive map allows the user to hover over a discipline label to highlight all its sub-disciplines, and also hover over a subdiscipline to see all subdisciplines it is connected to.

Visualization: Hex Map of Science

Project: ECON

demo.cns.iu.edu/client/econ-hexmap

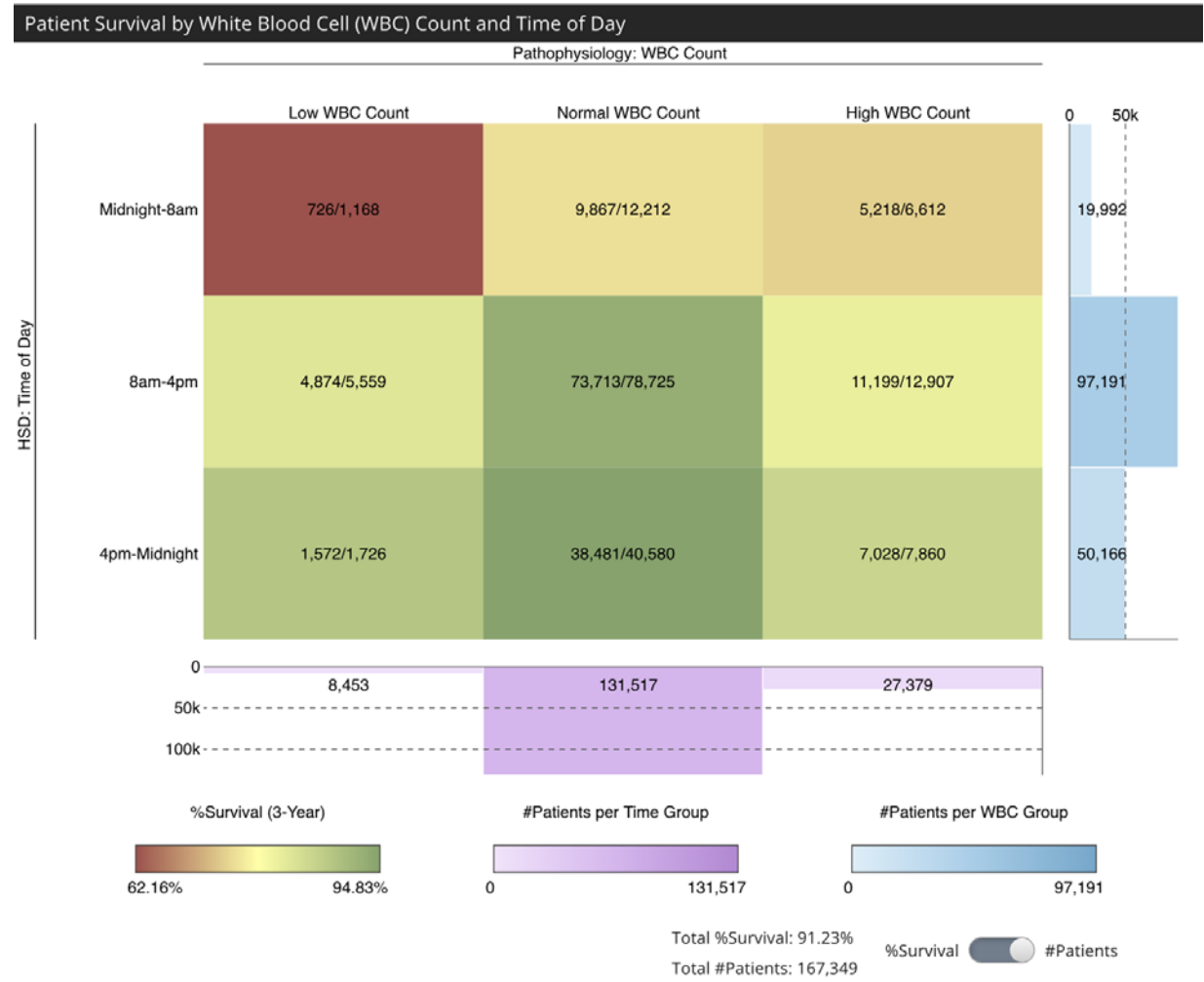


Hex-style rendering of the UCSD map of science allows for easier investigation of metric bars within each hexagon node. This Interactive map allows the user to hover over a discipline label to highlight all its subdisciplines, and also hover over a subdiscipline to see all subdisciplines it is connected to.

Visualization: Heatmap

Project: HSD

demo.cns.iu.edu/client/hsd/static/heatmap_group.html



This visualization shows how white blood cell (WBC) laboratory tests correlate with three-year survival rates. The HSD dimension of the data (rows) is the time of the day of the test; and three-year survival rate (numbers and colors in the boxes) is an outcome variable. **Aggregation level for the HSD time of day are shown—three 8 hour blocks.** The lowest survival rates are for patients with a low WBC value in the morning (specifically at 6am).

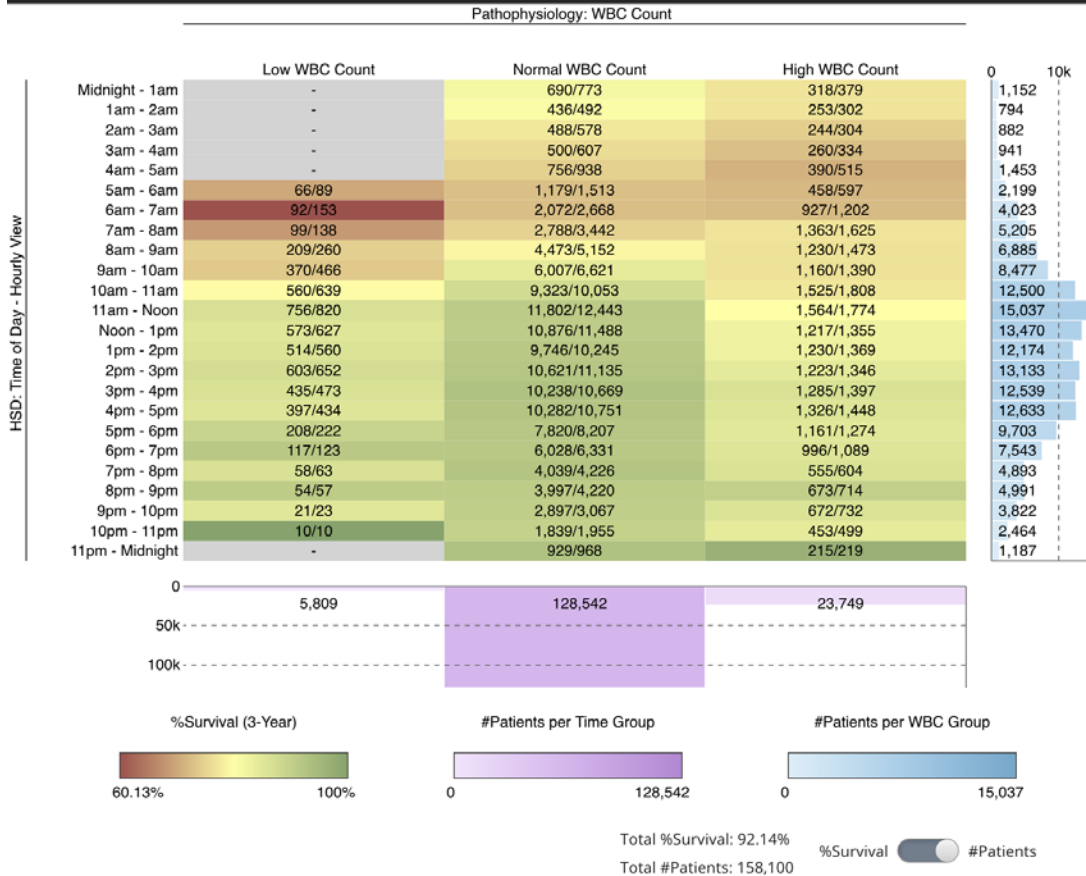
In this project, we created data visualizations to explain HSD to users and to help them incorporate it into their research.

Visualization: Heatmap

Project: HSD

demo.cns.iu.edu/client/hsd/static/heatmap_hour.html

Patient Survival by White Blood Cell (WBC) Count and Time of Day



This visualization shows how white blood cell (WBC) laboratory tests correlate with three-year survival rates. The HSD dimension of the data (rows) is the time of the day of the test; and three-year survival rate (numbers and colors in the boxes) is an outcome variable. **Aggregation level for the HSD time of day are shown— 24 hourly blocks on the right..** The lowest survival rates are for patients with a low WBC value in the morning (specifically at 6am).

In this project, we created data visualizations to explain HSD to users and to help them incorporate it into in their research.

Visualization: Heatmap

Project: Learning Analytics

demo.cns.iu.edu/webvis/learning-analytics/heatmap.html

Learning Analytics

Z637-29374 2015 Student Group Engagement and Scores

	Pre-Course	Week 1	Week 2	Week 3	Week 4	Midterm	Week 5	Week 6	Week 7	Week 8	Week 9	Final	Curr. Score
IVMOOC	26.05%	38.32%	31.32%	29.96%	27.1%	28.34%	31.07%	24.28%	16.86%	18.23%	13.08%	13.41%	20.87%
Z637-29374	33.01%	52.91%	49.89%	59.22%	50.89%	82.56%	65.04%	49.99%	39.59%	61.63%	54.91%	82.25%	82.4%
Z637-32593	25.08%	54.54%	43.58%	50.67%	53.63%	77.67%	65.7%	59.48%	52.19%	65.71%	47.27%	72.59%	75.13%
Z637-33781	29.33%	55.38%	49.26%	62.18%	77.47%	85%	87.4%	69.8%	55.56%	57.6%	45.69%	70.89%	77.94%

Z637-29374 2015 Student Group Engagement for Week 5

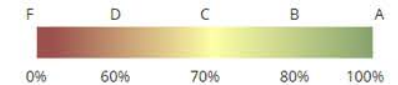
	Discussions	Page Views	Quizzes	Engagement
Student 11	91.7%	98.4%	0%	62.73%
Student 15	81.2%	97.6%	100%	92%
Student 5	76.3%	94.4%	100%	89.33%
Student 14	73.3%	93.5%	100%	88.04%
Student 18	70.7%	94.4%	100%	87.48%
Student 19	68.2%	90.4%	100%	85.34%
Student 4	0%	99.6%	100%	65.87%
Student 12	0%	91.4%	100%	63.16%
Student 13	0%	97.7%	100%	65.24%

Legends

Engagement



Score



Description

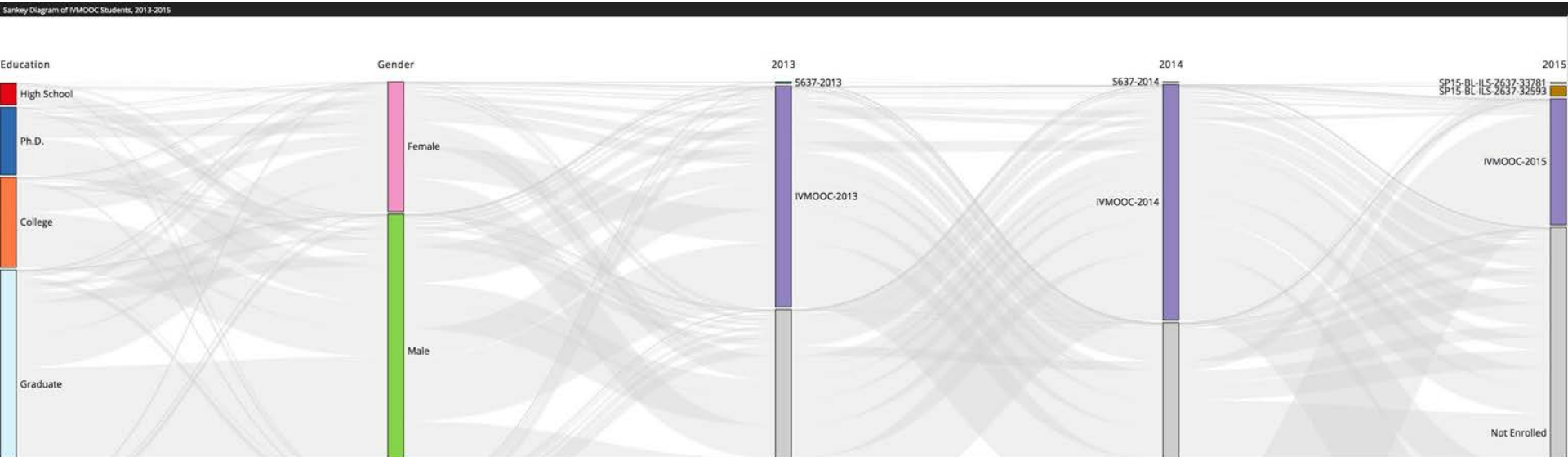
The heat map visualization is a representation of student engagement (magenta to blue color scale) and performance (red to green color scale) throughout a course. The visualization has two levels. The top level provides an overview of engagement and performance for groups of students, while the bottom level provides a detailed breakout of student engagement statistics for individuals with an identified group.

The heat map visualization is a representation of student engagement (magenta to blue color scale) and performance (red to green color scale) throughout a course. The visualization has two levels. The top level provides an overview of engagement and performance for groups of students, while the bottom level provides a detailed breakout of student engagement statistics for individuals with an identified group.

Visualization: Sankey

Project: IVMOOC

demo.cns.iu.edu/webvis/learning-analytics/sankey.html

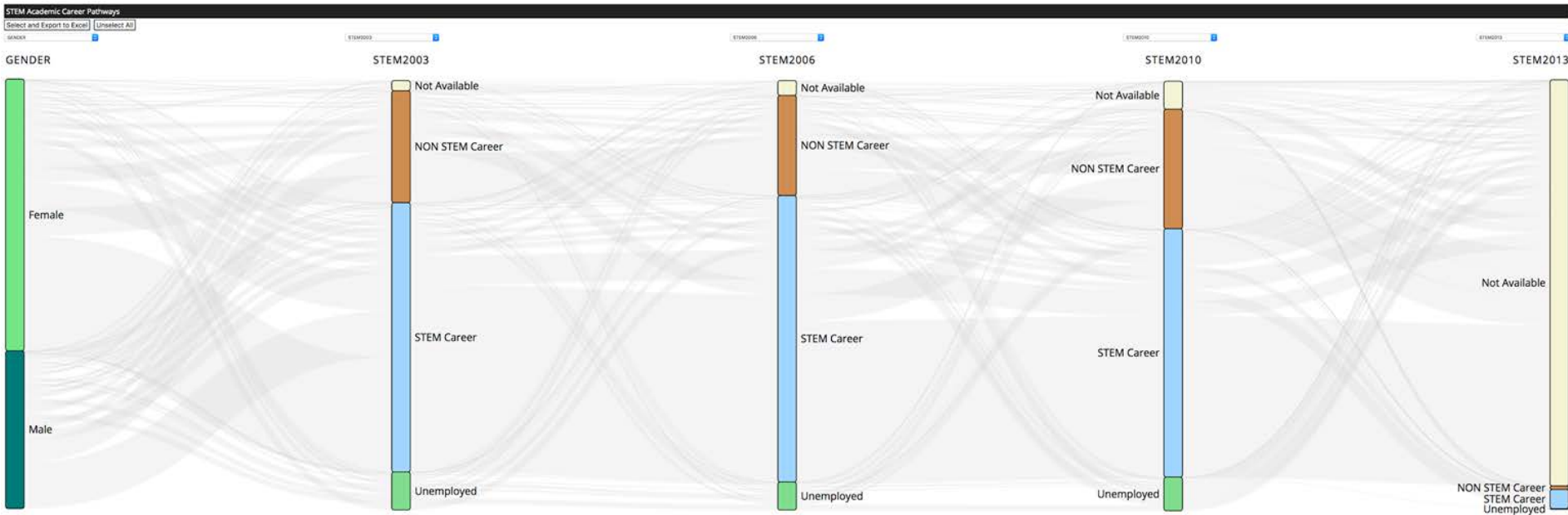


The goal of this visualization is to show the flow of students through Indiana University’s Information Visualization Massive Open Online Course (IVMOOC) from 2013 – 2015. The diagram (see screenshot in Figure 1) was created using D3 with data from learning management systems (LMS), such as Google Course Builder and Canvas, but also student demographic data collected using surveys. The diagram shows demographic information regarding a student’s level of educational attainment and gender together with the year they took the course, broken down further by IVMOOC and IU sections of the course.

Visualization: Sankey

Project: STEM

demo.cns.iu.edu/webvis/stem

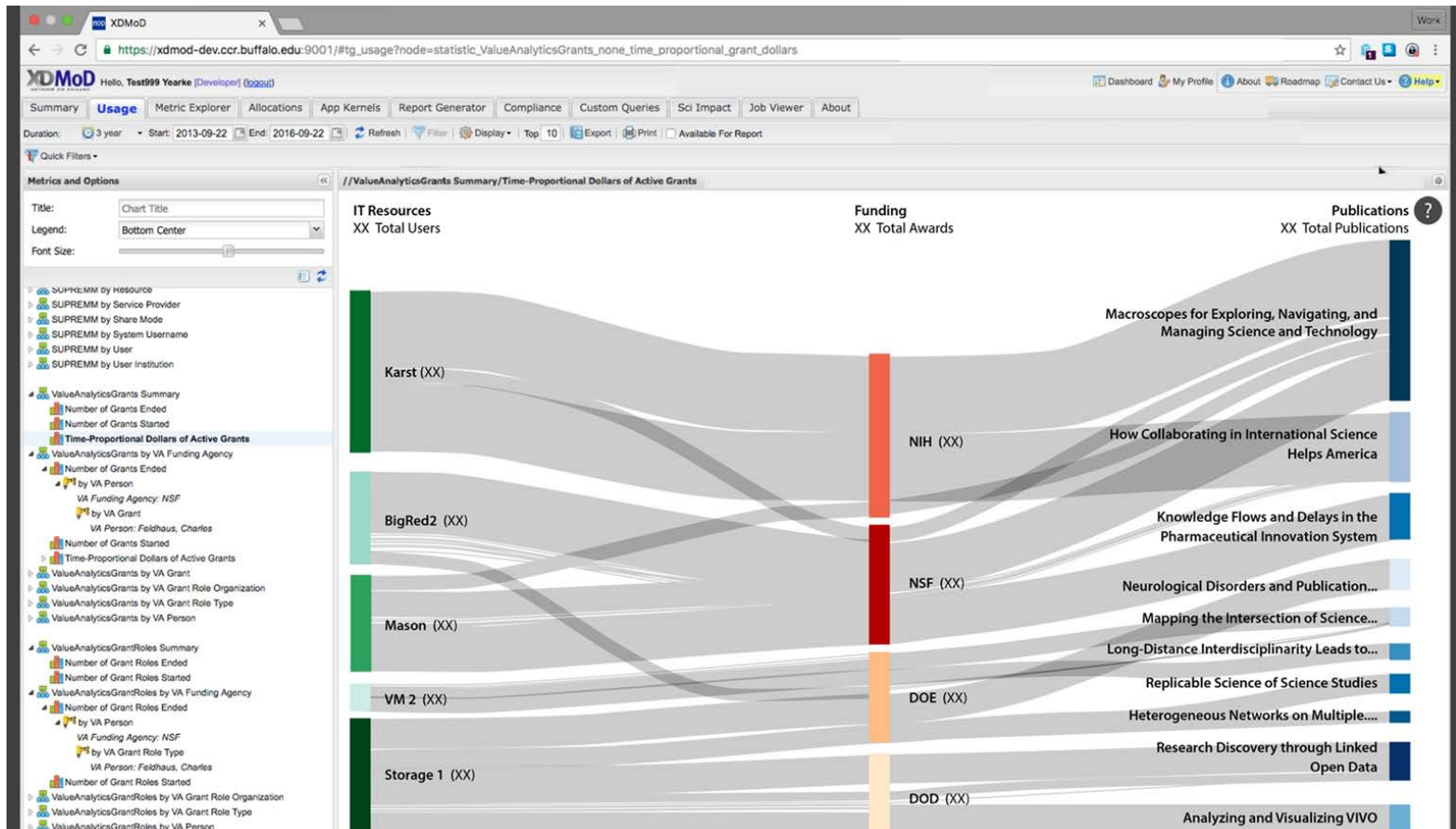


STEM (science, technology, engineering, math) pathways visualization. Data tool that will make it easier for researchers to visualize pathways and identify groups of individuals pursuing specific paths.

The primary goal of Sankey visualization was to understand the growth/decline of STEM career interests in students at various educational levels and across different demographics. Users will be able to precisely analyze and identify the group of students that flow in and out of the STEM career interest areas and in turn perform detailed research on these focus groups to study the reasons and influencing factors for the change in career interests.

Visualization: Sankey Diagram

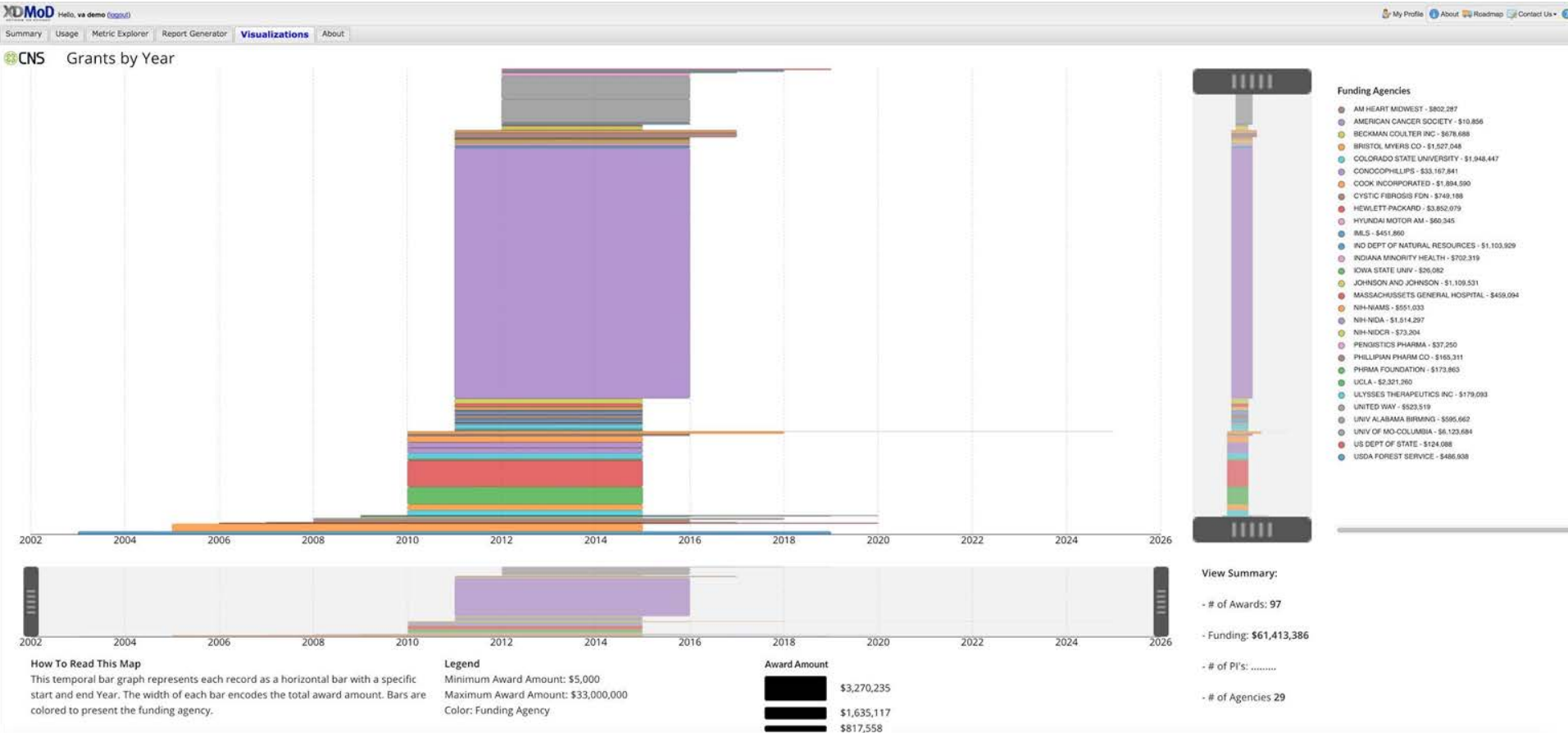
Project: XDMoD



This Sankey diagram displays a multivariate analysis of the relationship between IT resources, funding agencies, and publications. The width of each line represents grant dollars awarded to researchers. The configuration model allows for easy metric switching.

Visualization: Temporal Bar Graph

Project: XDMoD



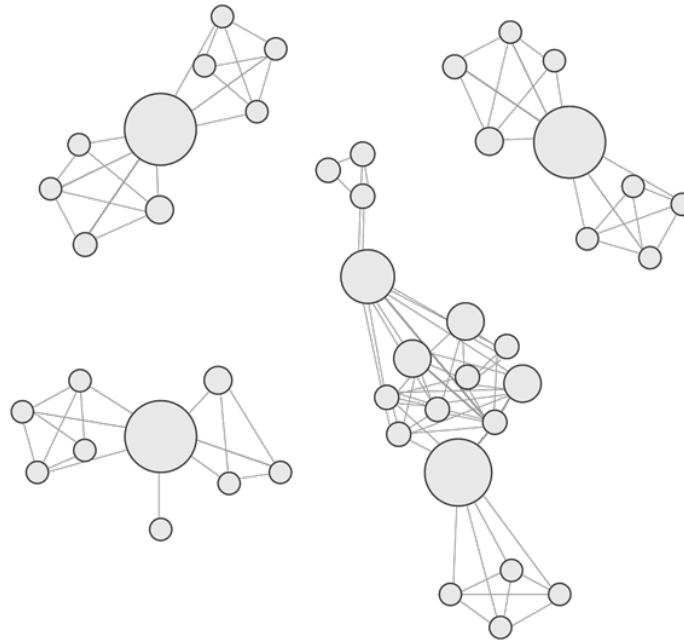
This temporal bar graph represents each record as a horizontal bar with a specific start and end year. The width of each bar encodes the total award amount. Bars are colored to represent the funding agency (legend of funding agencies are listed on the right). This graph shows funding duration, amounts, and types over time.

Visualization: Co-PI

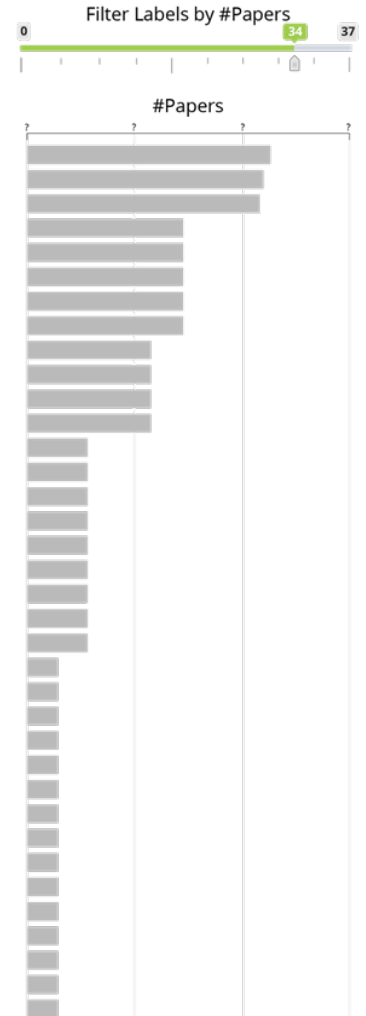
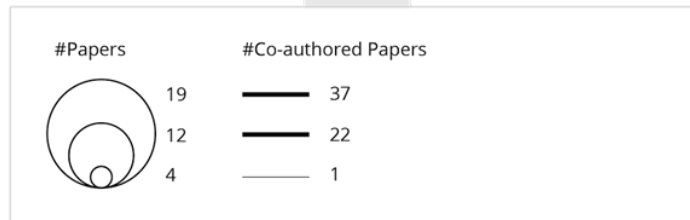
Project: XDMoD

+ -

This network represents collaboration patterns based on NSF funding data. Each node represents a principal investigator (PI or Co-PI). It is labelled by the PI's name, size coded by the total value of all awards for each PI, and color coded by the number of collaborators. An edge between two PIs denotes that they have a grant together with edge thickness denoting the number of times they collaborated and edge color reflecting their success in terms of total sum of all their joint awards. The listing on right rank orders PIs by total dollar amounts.



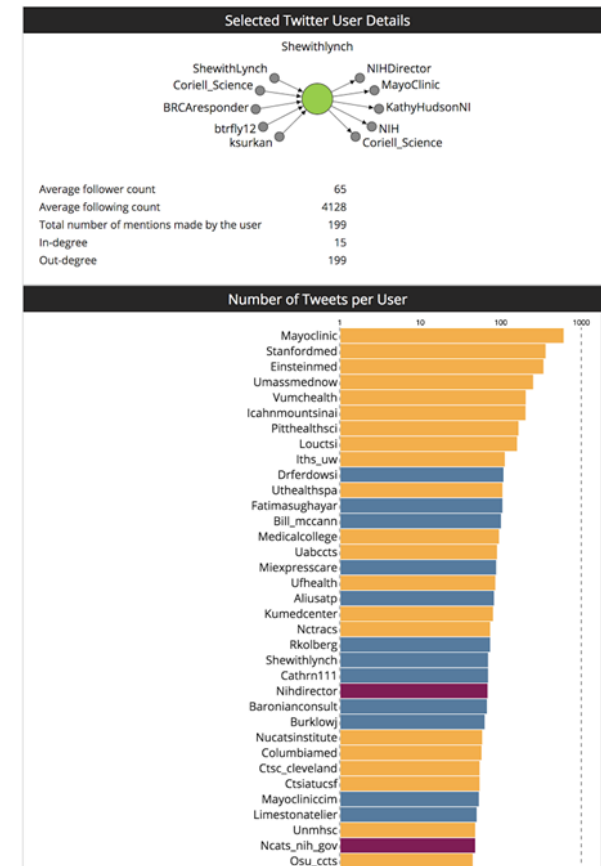
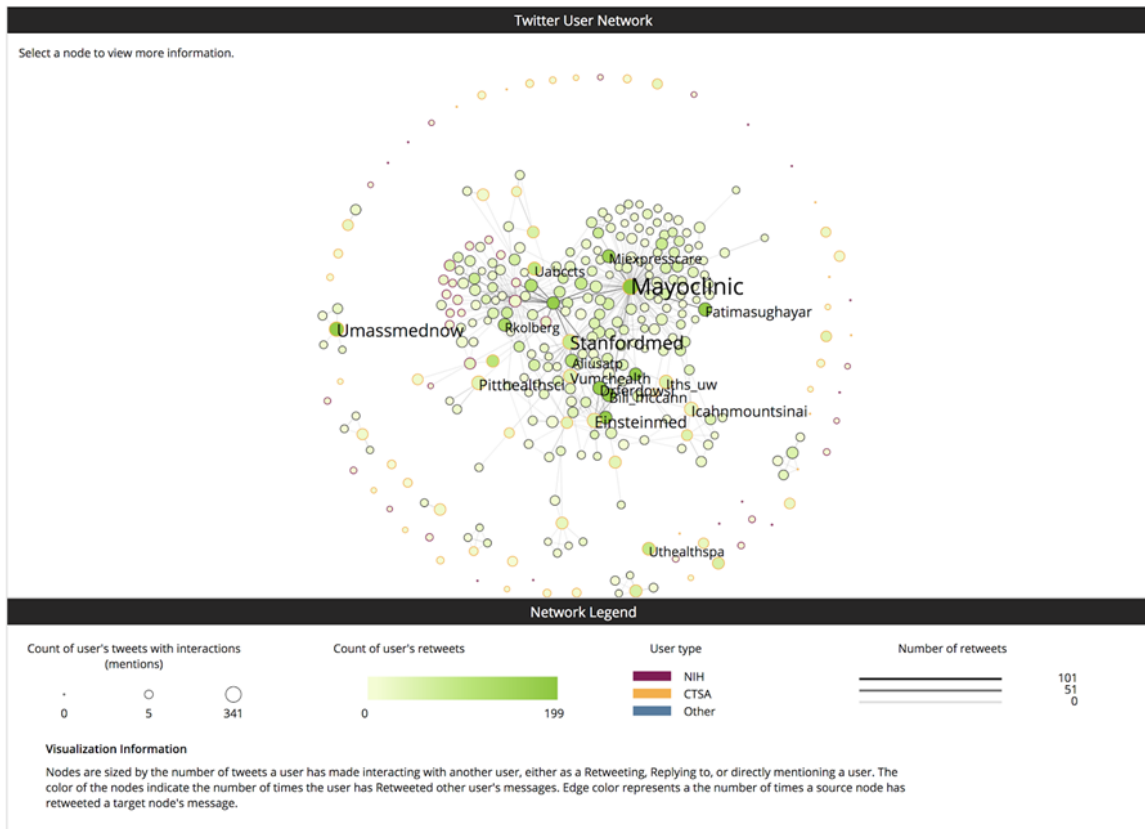
Legend



Visualization: Twitter Network

Project: IAI

demo.cns.iu.edu/client/iai/twitter.html

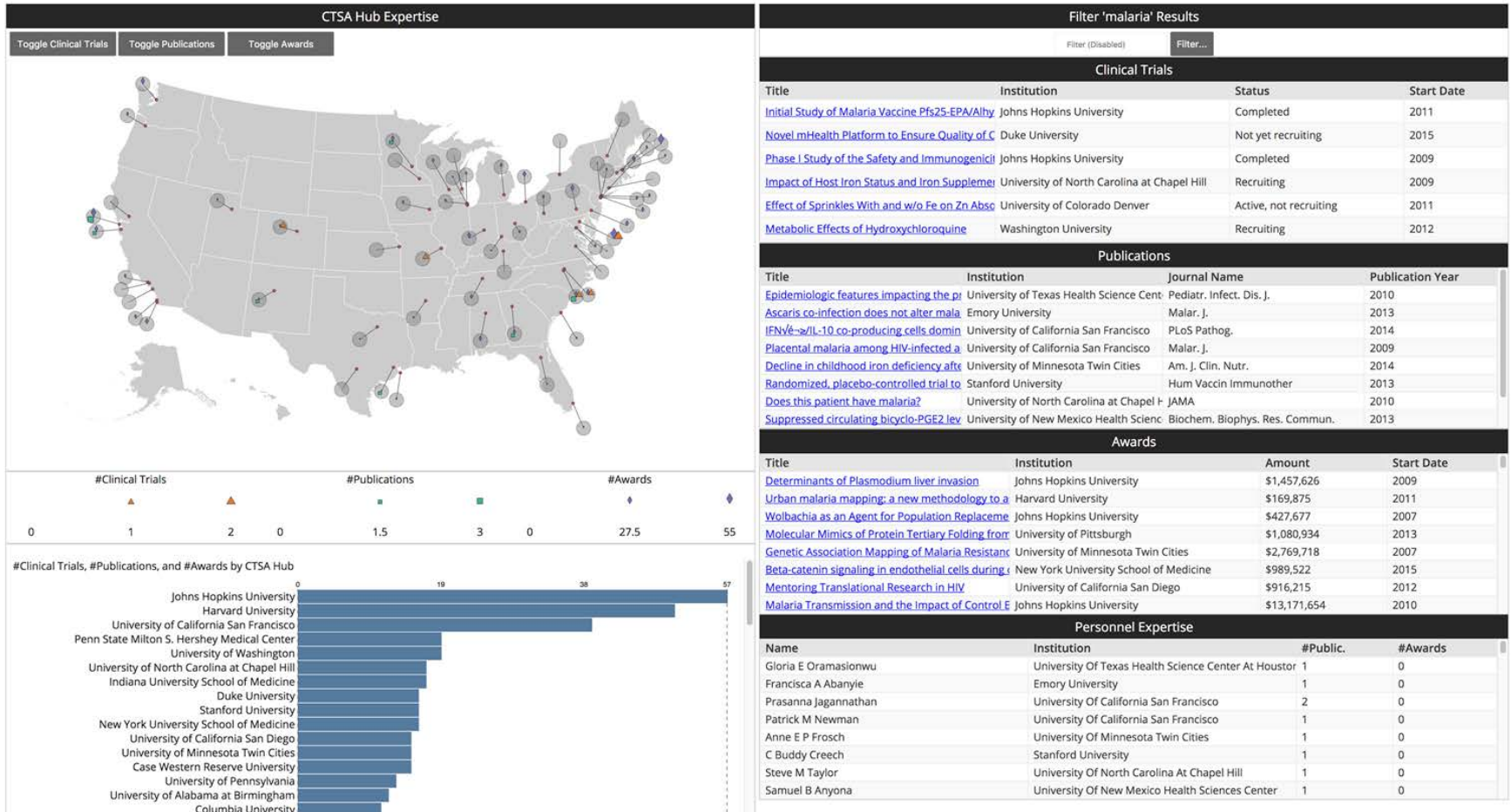


This visualization shows CTSA hub and NIH activities on Twitter. The accounts and tweets associated with CTSA and NIH were collected between August 2015 - Sept. 2015, then processed and analyzed to create a social network based on the interaction behaviors of users. The layout of the user network is force-directed, meaning that nodes that are close to each other have a stronger connection to one another indicating which Twitter users interact with each other most often.

Visualization: IAI Expertise Visualization

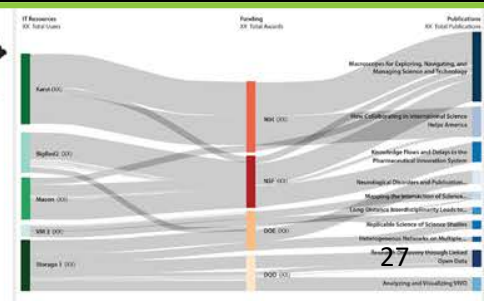
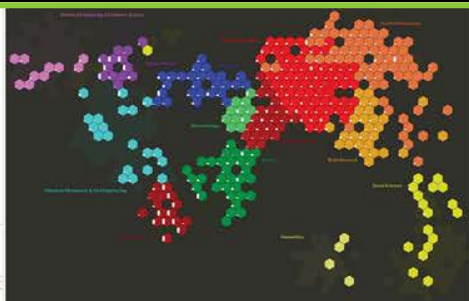
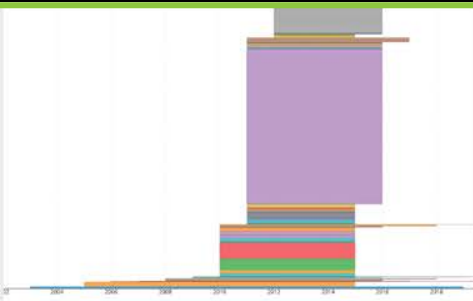
Project: IAI

demo.cns.iu.edu/client/iai/expertise.html?set=malaria



This visualization is based on publication datasets retrieved from the Scholarly Database at IU, and is used to identify relevant experts, publications, clinical trials, and awards that match a search term.

CNS Web Visualization Framework A Boilerplate Example



Activity Overview

To Do - To setup project workspace and generate a sample visualization from the framework.

Steps Involved -

- [Project dependencies set-up](#).
- Including a d3 visualization in the project -
 1. [Part 1](#) - Adding json file to specify git URLs, building and including a viz.
 2. [Part 2](#) - Creating the html page for it.
 3. [Part 3](#) - Calibrating it to read data.
 4. [Part 4](#) - Adding the json data file.
 5. [Part 5](#) - Modifying the configuration javascript file.

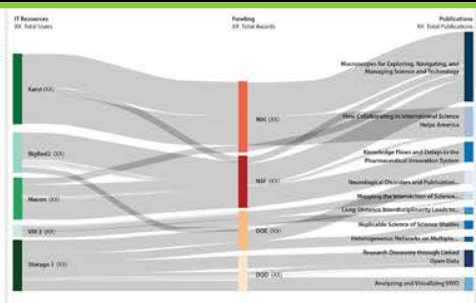
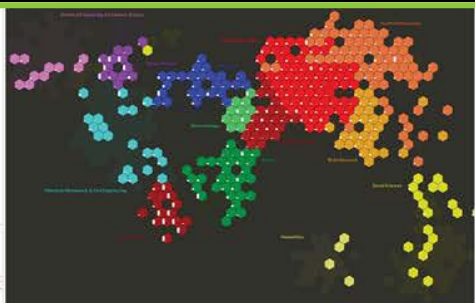
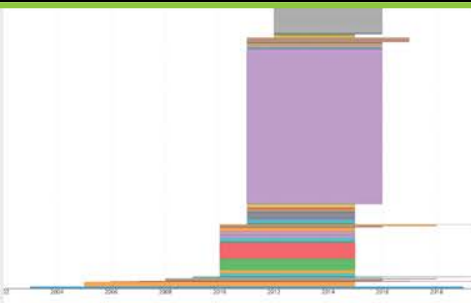
Activity Overview

- Having fun with it -
 1. [Part 1](#) - Adding more html functionality.
 2. [Part 2](#) - Adding the corresponding javascript for it.

Uses -

- Easy to start at an upper layer of abstraction.
- Multiple stages of rendering.
- Multiple instanced visualizations in the same project.
- Easy versioned project check-out and deployment.
- Configuration-based customization.
- Ease of use and extensibility.

Outlook



Outlook

- More robust workflow
- Public-facing code and documentation
- More visualization library support
- Better cross-visualization integration
- Full integration of Angular Material
- More visualization plugins

Atlas Trilogy

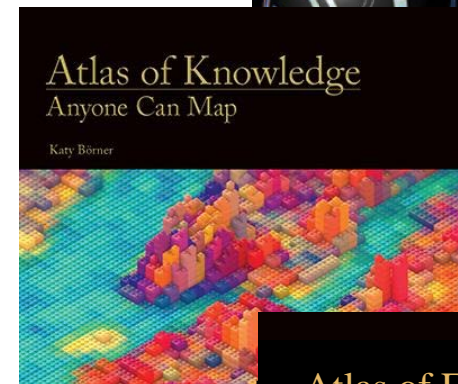
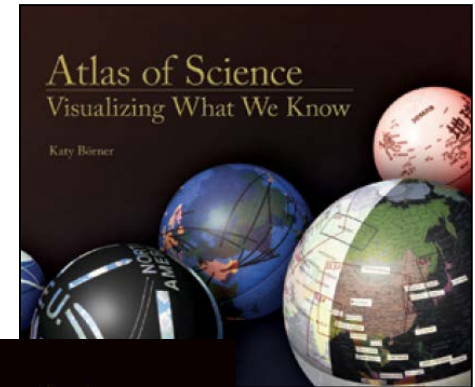
Börner, Katy (2010) **Atlas of Science: Visualizing What We Know**. The MIT Press. <http://scimaps.org/atlas>

Börner, Katy (2015) **Atlas of Knowledge: Anyone Can Map**. The MIT Press. <http://scimaps.org/atlas2>

Börner, Katy (2020) **Atlas of Forecasts: Predicting and Broadcasting Science, Technology, and Innovation**. The MIT Press.

ModSTI Conference slides, recordings, and report are at modsti.cns.iu.edu/report

Upcoming Sackler Colloquium on "**Modelling and Visualizing Science and Technology Developments**" will take place in December 2017 at the Beckman Center, Irvine, CA.



All papers, maps, tools, talks, press are linked from cns.iu.edu
 These slides are at cns.iu.edu/presentations

CNS Facebook: facebook.com/cnscenter

Place & Spaces: Mapping Science Exhibit Facebook: facebook.com/mappingscience



The screenshot shows the CNS website homepage. At the top left is the CNS logo and name. To the right is a search bar with the text "search cns.iu.edu" and a "Search" button. Further right are social media icons for Facebook, Twitter, and Email. Below this is a navigation menu with links: About Us, Research, Development, Teaching, Outreach, Videos, News & Events, and Connect With Us. The main content area features a large image of a group of people in a meeting, with a green overlay on the right containing the text: "We work closely with clients to provide custom-made data, visualization, and software solutions". Below this are several sections: "Research" with a link to "Open Data and Open Code for Big Science of Science Studies"; "Latest News" with a link to "Put your money where your citations are: a proposal for a new funding system (website accessed 9/05/13)"; "Upcoming Events" listing three events: "Katy Börner attends PIUG 2013 Northeast Conference" (OCT 1), "Katy Börner presents Mapping Science Exhibit at WSSF" (10.13), and "Ted Polley & Google Team present IVMOOC at EDUCAUSE" (10.15), plus "Katy Börner presents at the SciELO 15 Years Conference" (10.22); "Development" with a link to "Behind the scenes of the design and development of AcademyScope"; "Outreach" with a link to "See some of the most fascinating data visualizations in the world."; "Videos" with a link to "Watch Katy Börner's full presentation from TEDxBloomington"; "Teaching" with a link to "Successful IVMOOC will be offered again in January of 2014"; and "Our Products" with a link to "We work closely with clients to provide custom-made data, visualization, and software solutions".