Study Information Block

Visual Analytics of the NSWC Crane Innovation Ecosystem
User Studies
IRB PROTOCOL #1809442778

You are invited to participate in a research study of the effectiveness of a new visualization in development. You were selected as a possible subject because you are an subject matter domain expert. Please read this form and ask any questions you may have before agreeing to be in the study.

The study is being conducted by Dr. Katy Börner and Dr. Olga Scrivner of the School of Informatics, Computing, and Engineering. It is funded by the Naval Surface Warfare Center Crane Division.

STUDY PURPOSE

The proposed work will use large-scale publication and funding data to support the analysis and visualization of key experts, institutions, publications, and funding in strategic areas. This study will help us improve the usability of visualizations developed for this project.

PROCEDURES FOR THE STUDY:

If you agree to be in the study, you will take an online survey that takes approximately 30 minutes to complete. You will be asked to review and provide feedback on visualizations that focus on one of three topics.

RISKS AND BENEFITS

The risk of participating in this research includes discomfort answering questions about unfamiliar visualizations. There is also a risk of loss of confidentiality. Please be aware that you can terminate your participation in the study at any time.

CONFIDENTIALITY

Efforts will be made to keep your personal information confidential. We cannot guarantee absolute confidentiality. Your personal information may be disclosed if required by law. Your identity will be held in confidence in reports in which the study may be published and databases in which results are stored.

Organizations that may inspect and/or copy your research records for quality assurance and data analysis include groups such as the study investigator and his/her research associates, the Indiana University Institutional Review Board or its designees, the study sponsor, and (as allowed by law) state or federal agencies, specifically the Office for Human Research Protections (OHRP), etc. who may need to access your research records.

PAYMENT

There will be no payment for participation in this study.

CONTACTS FOR QUESTIONS OR PROBLEMS

For questions about the study, contact Dr. Olga Scrivner via email obscrivn@indiana.edu or phone (812) 349-8853.

For questions about your rights as a research participant or to discuss problems, complaints or concerns about a research study, or to obtain information, or offer input, contact the IU Human Subjects Office at (812) 856-4242.

VOLUNTARY NATURE OF STUDY

Taking part in this study is voluntary. You may choose not to take part or may leave the study at any time. Leaving the study will not result in any penalty or loss of benefits to which you are entitled. Your decision whether or not to participate in this study will not affect your current or future relations with the School of Informatics, Computing, and Engineering.

Part I: demographics

What is your department name, specialization, or area of focus? What words would you use to describe the focus of your work?

[   ]

What is your native language?

○ English

○ [   ] Other, please specify

Gender

○ Male

○ Female

○ Other

Please indicate your age:

○ <20

○ 21-30

○ 31-40

○ 41-50

○ 51-60

○ >60

The following questions will help us understand your background and experience with visualizations.

Which of these visualization types are you familiar with?

☐ Graphs

- [ ] Geospatial Maps
- [ ] Science (Topic) Maps
- [ ] Tree visualizations
- [ ] Network Visualizations

Click to write the question text

- ( ) Click to write Choice 1
- ( ) Click to write Choice 2
- ( ) Click to write Choice 3

Have you ever taken a seminar, course, or training session on visualization?

- ( ) Yes
- ( ) No

If so, please specify the seminar, course, or training session:

[   ]

Branching Block

In this part of the study, you will view visualizations showing the structure and evolution of a particular research area and answer questions related to these visualizations.
Please select one research area that you would like to explore or that is most relevant to your work:

- ( ) Artificial Intelligence
- ( ) Internet of Things
- ( ) Robotics

Artificial Intelligence Block

In this part of the survey, you will see a series of visualizations that show the structure and evolution of **artificial intelligence** over the last 20 years (1998-2017). Data shown here includes funding provided by the National Science Foundation and publications retrieved from the Web of Science, an online subscription-based scientific indexing service. These databases were used to identify bursts of activity, key experts and their collaboration networks, and changes in topical coverage over time.

Please take a few moments to review each visualization. Then, answer a set of questions and tell us how you might use the visualizations and insights in your daily work.

**Temporal Bursts of Activity**

Burst analysis can be run to determine sudden increases in the frequency of terms in, for example, publication and funding award titles, abstracts, and keyword lists. Note that terms will not be detected as bursting if the frequency with which they appear remains high over time. The top-10 high frequency terms in publications are:

Term in Publication Titles and Abstracts	Frequency
computer science	752
engineering	592
psychology	232
science & technology - other topics	158
neurosciences & neurology	157
physics	109
automation & control systems	106
robotics	86
mathematics	78
materials science	68

The top-10 high frequency terms in awards are:

Term in Publication Titles and Abstracts	Frequency
engineering, electrical & electronic	462
computer science, artificial intelligence	359
computer science, theory & methods	318
neurosciences	155
computer science, hardware & architecture	144
engineering, biomedical	126
computer science, software engineering	115
psychology, experimental	111
automation & control systems	106
computer science	96

Figure 1. Burst of terms in publication (green) and funding awards (blue).
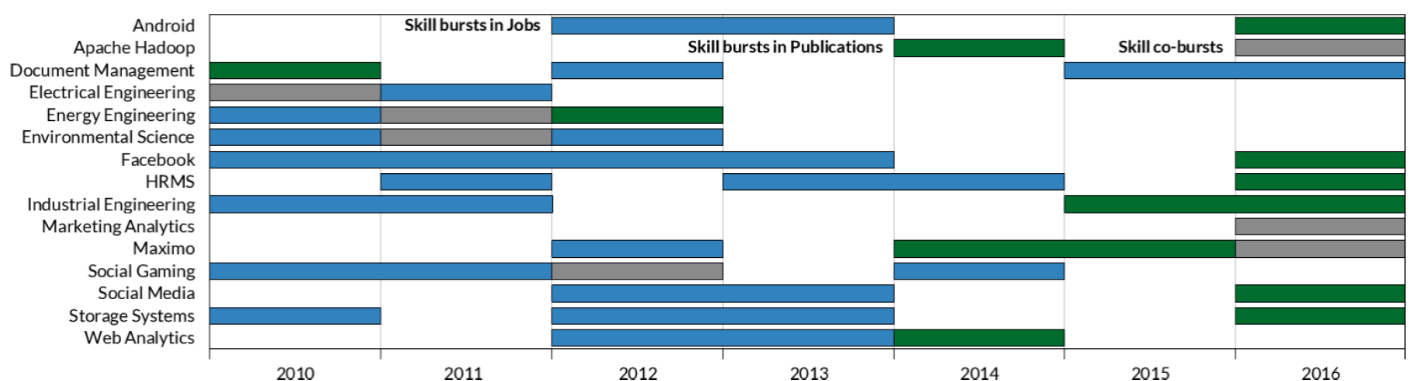
In **Fig. 1** (above), each term that experiences a burst between 1998 and 2017 is rendered as a horizontal bar with a start and an end date; the burst term is rendered on the left. Terms that burst in funding awards are given in blue; terms bursting in publications are in shown in green. Seven skills burst in both datasets during the same years and are shown in gray: Apache Hadoop, Electrical Engineering, Energy Engineering, Marketing Analytics, Maximo, Social Gaming.

What might co-bursting of terms indicate?

Why might there be so few bursts in 2014 and 2015?

Name two skills that burst three times but never in both datasets together.

What might this indicate?

What insights might you gain from this visualization?

How might you use these insights in your daily work?
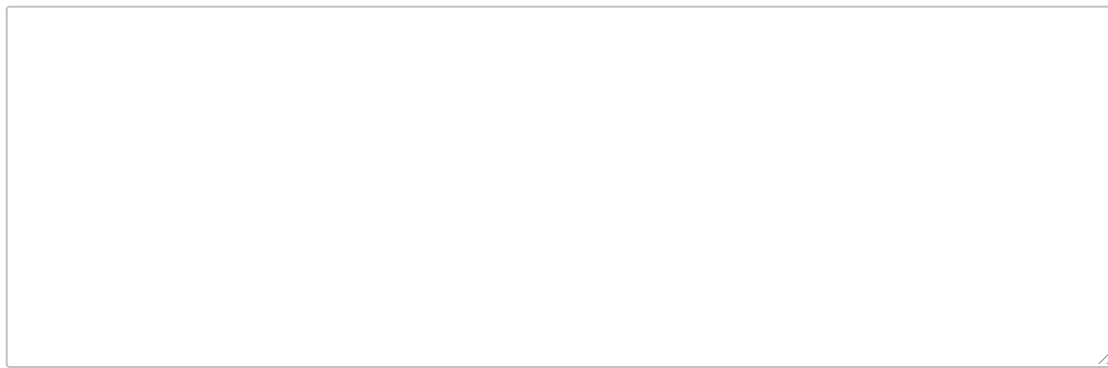
Key Experts and Collaboration Networks (Force Directed Layout)

Figure 2. Co-author network autheors with more than 5 papers. See legend for size and color coding.

Most publications list multiple authors, making it possible to extract the co-author network. Each node in the network represents one author. Two authors are connected if

they co-authored a paper together. In Fig. 2 (above), each node is area size coded by the number of total citations. Citations are counted in full, i.e., if two authors are listed on a publication with 10 citations then each author is credited with 10 citations. Co-author links are thickness coded by the number of times two authors appeared on a paper. Nodes and linkages are color coded by the year of their appearance: 1998-2004 (black), 2005-2007 (dark green), 1908-2014 (green), and 2015-2017 (light green).

Which author interlinks the two major clusters in the network?

Which authors cluster (top left, lower right) has been working in this area for longer?

Which author(s) are most interesting for your decision making and why?

What insights might you gain from this visualization?

How might you use this visualization or these insights in your daily work?

Key Experts and Collaboration Networks (Geospatial Map Layout)
**Figure 3.** Co-author network of authors with identical ZIP codes overlaid on a map of the U.S.
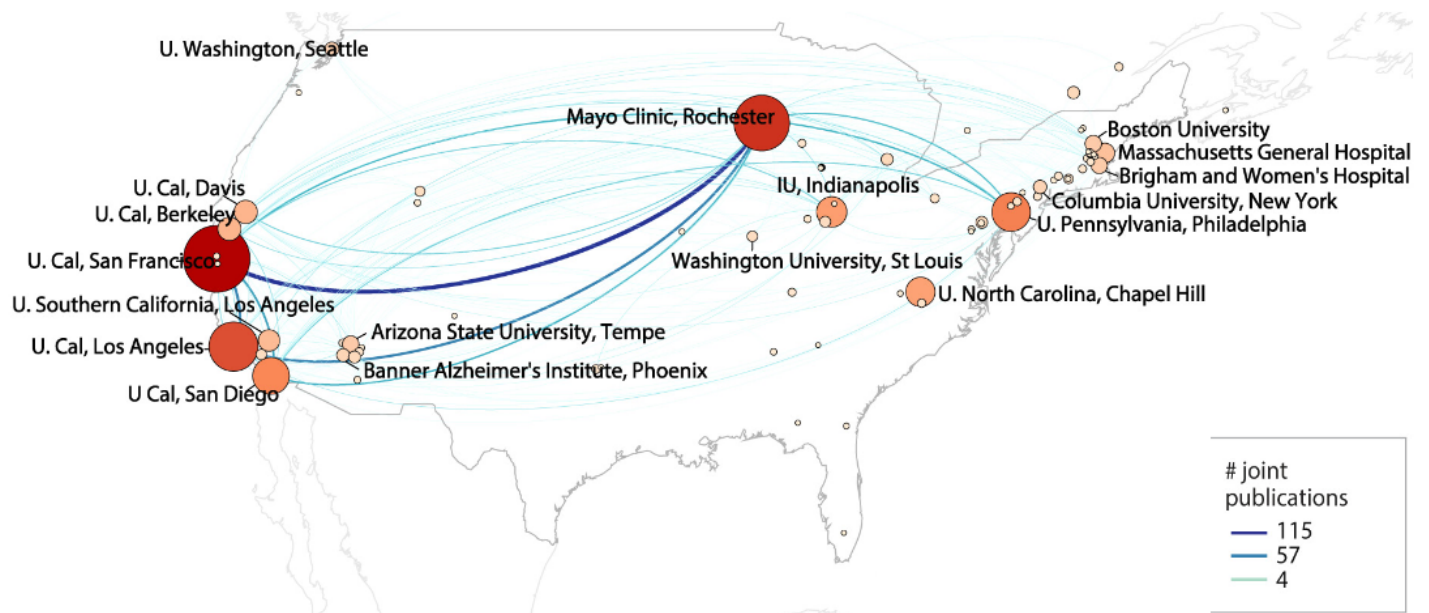
Which ZIP code geolocation interlinks the two major clusters in the network?

Which ZIP code geolocation(s) are most interesting for your decision making and why?
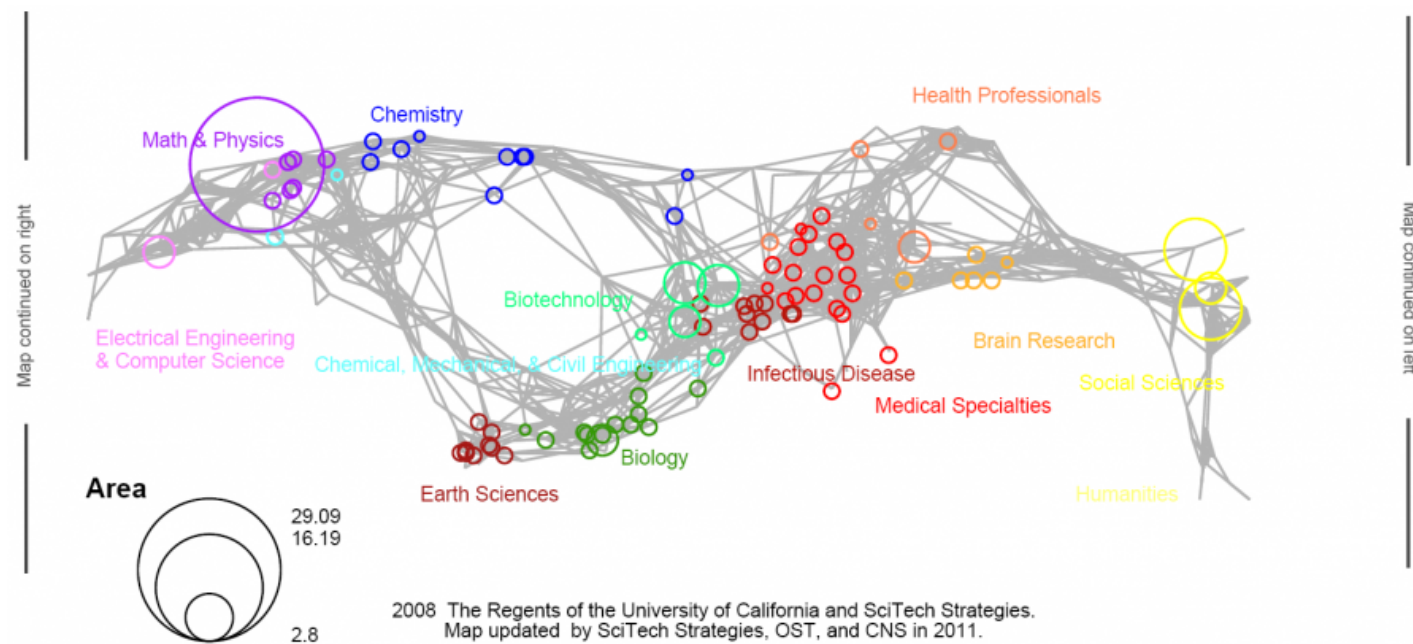
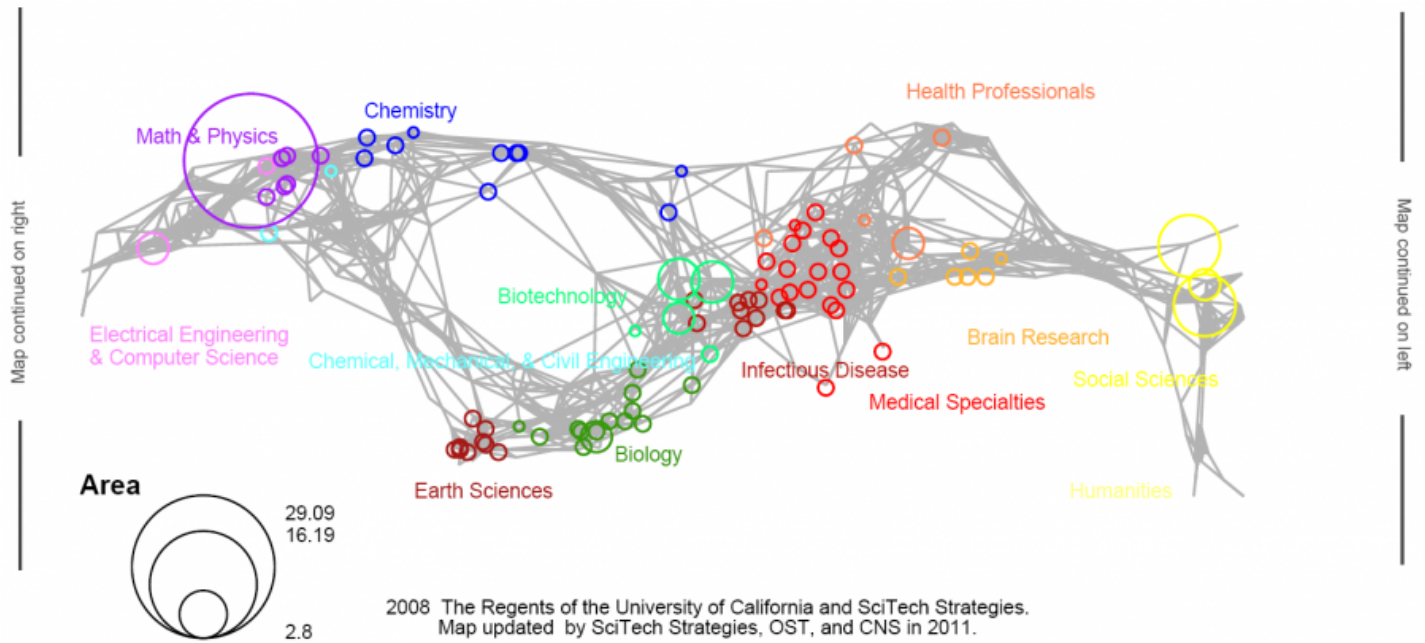What insights might you gain from this visualization?

How might you use this visualization or these insights in your daily work?
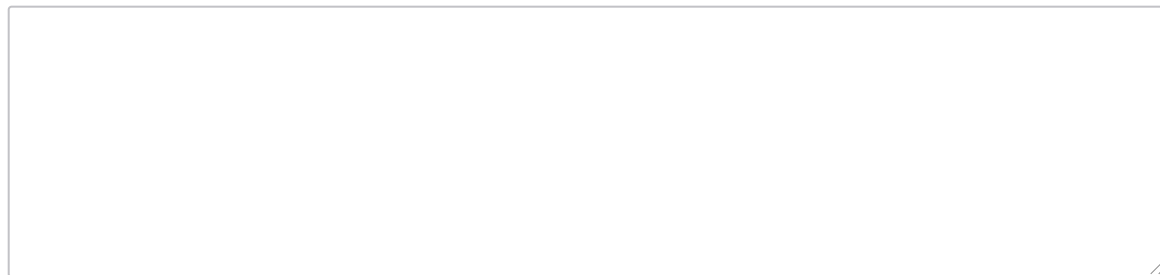
Topical Evolution

Figure 4. Topical coverage of publications for 1998-2007 (top) and 2008-2017 (bottom).

Using publication data for 20 years, the evolution of a research area can be depicted and explored. **Fig. 4** (above) shows the topical coverage of publications for 1998-2007 and 2008-2017 overlaid on the UCSD map of science. Publications (matched by journal name) are aggregated both at the upper level of 13 broad scientific disciplines and at the lower level of 554 subdiscipline clusters. Labels for the 13 scientific disciplines appear on the science map and provide a legend for the color of the subdiscipline clusters appearing on the background tiles.

Which discipline experiences a growth from 1998-2007 to 2008-2017?

Which discipline experiences a decline from 1998-2007 to 2008-2017?

What insights might you gain from this visualization?

[ ]

How might you use this visualization or these insights in your daily work?

[ ]

Part III

Which of the visualizations in this survey were most relevant to your work?

[ ]

Why?

[ ]

What did you **like** about the visualizations and why?

[ ]

What did you **dislike** about the visualizations and why?

[ ]

What questions do you have that were not answered by those visualizations?