

Engineering Research Center Observatory Visualizations

Katy Börner (@katycns)

Joint work with Gerhard Klimeck, Michael Zentner, Steven Snyder at Purdue University

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

Virtual Presentation to HUBzero community

March 28, 2018



- Dec 09, 2015: Gerhard Klimeck, Michael Zentner, and Katy Börner present <u>Engineering Research Center Observatory</u> at Engineering Research Center Observatory Kick-Off Meeting, Washington, DC.
- Dec 12, 2016: Katy Börner presents <u>Visualizing Nanoscience and</u> <u>Technology</u> at <u>2016 NSF Nanoscale Science and Engineering Grantees</u> <u>Conference</u>, Arlington, VA.
- June 1, 2016: Katy Börner presents <u>Engineering Research Center</u> <u>Observatory Visualizations</u> at Virtual Presentation to NSF EEC Staff, Bloomington, IN.
- Börner, Katy, and Steve Snyder, Gagandeep Singh, Sara Bouchard, Adam Simpson, Gerhard Klimeck, Michael Zentner, and Steve Snyder. 2017. "Engineering Research Center Observatory". Poster at ERC Biennial Conference.



manoHUB PURDUE

· Each node represents an author.

Fill color denotes the year of first publication.

· Two authors are connected if they have authored a paper together.

· Node outline color indicates author type (student, faculty, post-doc, staff, other) Node size and label size denote the number of papers authored.

· Clicking on a node in the network highlights publications by that author.

· Click on usage button in top left to learn how to color code nodes, etc.

· The legend details how the information is being encode

· Edge thickness denotes the number of times two authors appeared on a paper together. · The horizontal bar graph on the right shows the number of papers per author.

· Clicking on an author node in the network or in the bar-graph brings up a listing of publications by that author



Indiana University: Katy Börner, Gagandeep Singh, Sara E. Bouchard, Adam H. Simpson, Scott C. Hutcheson Purdue University: Gerhard Klimeck, Michael Zentner, Steve Snyder



Bistechnolog Humarities

Social Science

Legend

Nodes

10

goo.gl/nd1ojc

Goal: Develop and deploy interactive data visualizations for NSF staff, researchers, and students to increase their understanding of temporal, geospatial, topical, and network patterns and trends in engineering.



Acknowledgements This work is funded by NSF via a supplement to Network for Computational Nanotechnology - Cyber Platform - Engineering Research Center Observatory. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

clusters; understand the backbone of strong co-author linkages that likely serve as main collaboration and knowledge diffusion pathways.

Co-authorship Geospatial Visualization

This interactive visualization shows the co-authorship network overlaid on a geospatial map of the world.

· Each node represents an author.

Zip code: 8528

· Two authors are connected if they have authored a paper together

- · Circle area sizes encode the number of links per node, also called node degree. · Horizontal bar-graph aids in better understanding of data.
- · Clicking on a node highlights the co-authors in the bar graph.

· The legend details how the information being encoded.



Goal, Use Cases, Users, Data, Visualizations



Engineering Observatory: Goal

Facilitate near real-time monitoring of Engineering Research Centers (ERCs) in support of better-informed resource allocation, priority setting, and evaluation.

Data mining and visualization web services will be provided for different stakeholders (NSF staff, researchers, students) to increase their understanding of temporal, geospatial, topical, and network patterns and trends in engineering.

This collaborative work with the nanoHUB team at Purdue University is funded by NSF, Dec 15 – Nov 17.



Engineering Observatory: Use Cases

Use Cases:

- Day-to-day operations
- Strategic decision making
- Prepare for site visits

Initial Power Users:

- Mehmet Ozturk, Nanosystems ERC for Advanced Self-Powered Systems of Integrated Sensors and Technologies (ASSIST)
- Paul Westerhoff, Nanotechnology Enabled Water Treatment Systems (NEWT)
- Greg Carman, Nanosystems ERC for Translational Applications of Nanoscale Multiferroic Systems (TANMS)



Engineering Observatory: Data & Visualizations

Data

• nanoHUB data + bibliography files

Visualizations

- Evolving collaboration networks obtained from bibliography files, network layout or overlaid on geographic map
- Evolving expertise profiles; overlaid onto the UCSD map of science





Interactive Data Visualizations

Importing Data

ere:	
	Drag BibTeX files here
ist-cumulative-April2017.bib 272 entries Author disambiguation helper	
If you find when visualizing your citations that the same author is represented by multiply variations of their name, you can use this search to find all the variations and choose a canonical form for them.	
Chack all Chack pape	
Adams, J. J. 1 citations	
Adams, Jacob J. 1 citations	
Adelegan, U. 2 citations	
Adu, Kon 2 citations	
Agrayazi, 1. 5 cuations	
Agrawal Richa 13 citations	
Abmadiyand, A. 1 citations	
Aitken, R. 2 citations	
Akella, D. 1 citations	
Albe, Virginie 1 citations	
Click above or type to choose preferred format	Combine selected names
Citations	
InProceedings — 2013 IEEE SENSORS Atomic Layer Deposited TiO2 thin films for envir	ronmental gas sensing
Mills, S.	
Unreported Female Male	



Curating Data

After the file is loaded, brief summaries will be shown for each entry. You may click any entry to expand it and give a full, editable view of the properties of that entry

Here, one should make sure that fields that are pertinent to the visualizations being performed are correct and consistent across entries.

	Mills S				
1	Unreported	Female	Male		
	Role				
	Location				
	Lee, B.				
	Unreported	Female	Male		
	Role				
	Location				
	Misra, V.	4000000	1.1.1.1.1		
	Role	remale	Mate		
	Location				
title: booktitle:	Atomic Layer Deposited 1 2013 IEEE SENSORS	TiO2 thin films for env	ironmental gas sensi		
title: booktitle: year:	Atomic Layer Deposited 1 2013 IEEE SENSORS 2013	TiO2 thin films for env	ironmental gas sensi		
title: booktitle: year: month	Atomic Layer Deposited T 2013 IEEE SENSORS 2013 Nov	TiO2 thin films for env	ironmental gas sensi		
title: booktitle: year: month: pages:	Atomic Layer Deposited 1 2013 IEEE SENSORS 2013 Nov 14	TiO2 thin films for env	ironmental gas sensi		
title: bookritle: year: month: pages: keywords:	Atomic Layer Deposited 1 2013 IEEE SENSORS 2013 Nov 1-4 Sensor phenomena and o	TIO2 thin films for env	ironmental gas sensi		
title booktitle year month pages keywords doi:	Atomic Layer Deposited 1 2013 IEEE SENSORS 2013 Nov 1-4 Sensor phenomena and c 10 1109/ICSENS 2013 66	TIO2 thin films for env characterization 588516	ironmental gas sensi		
title: booktitle year month pages keywords: doi:	Atomic Layer Deposited 1 2013 IEEE SENSORS 2013 Nov 1–4 Sensor phenomena and c 10 1109/ICSENS 2013 66	TiO2 thin films for env characterization 688516	ironmental gas sensi		
title: booktitle: year month pages: keywords doi:	Atomic Layer Deposited 1 2013 IEEE SENSORS 2013 Nov 1-4 Sensor phenomena and of 10 1109/ICSENS 2013 66	TiO2 thin films for env characterization 888516	ironmental gas sensi		
title booktitle year month: pagex keywords doi	Atomic Layer Deposited 1 2013 IEEE SENSORS 2013 Nov 14 Sensor phenomena and o 10 1109/ICSENS 2013 60 Hole estlor	TiO2 thin films for env characterization 888516	ironmental gas sensi		
title booktitle year month pages keywords doi	Atomic Layer Deposited 1 2013 IEEE SENSORS 2013 Nov 1-4 Sensor phenomena and o 10 1109/ICSENS 2013 60 Hote Rend Rendo	TiO2 thin films for env characterization 688516	ironmental gas sensi		

Name disambiguation: If the same author is represented by multiple variations of their name in the visualization, use the "Author Search" feature to find all variations and choose a canonical form for them.

Sex: User is provided the choice to select either Female, Male, or Unreported.

Role: User is able to select the role of each author, e.g., Student, Staff, Postdoc.

Geolocation: The location field accepts several types of input including city names and zip codes. After entering that information, either press 'Enter' or move on to the next field and in a few seconds the field will reflect whether geocoding information was found, turning green for success and red for failure.



Downloading and Visualizing Data

At the end of the page there are buttons to download changes as BibTeX or to launch a number of visualizations that highlight different aspects of the data.

Article — European Physical Journal-Special Topics Theoretical and technological building blocks for an innovation accelerator van Harmelen, F. and Kampis, G. and Borner, Katy and van den Besselaar, P. and Schultes, E. and Goble, C. and Groth, P. and Mons, B. and Anderson, S. and Decker, S. and Hayes, C. and Buecheler, T. and Helbing, D.

Article — Journal of Informetrics Approaches to understanding and measuring interdisciplinary scientific research (IDR): A review of the literature Wagner, C. S. and Roessner, J. D. and Bobb, K. and Klein, J. T. and Boyack, K. W. and Keyton, J. and Rafols, I. and Borner, Katy

Article — Scientometrics Mapping interactions within the evolving science of science and innovation policy community Zoss, A. M. and Borner, Katy

Download BibTeX

Visualize 29 journals using map of science

Visualize author geolocation network

Visualize co-authorship network

Visualizations open in a new window, ensure you have enabled popups for the hub.



Co-Author Visualization – Network Layout

Each node in the figure represents an author, and author node area size scales with the number of publications. Author nodes with three publications or more are labelled by the author's name. Two authors are connected if they have authored a publication together and link width scales with the number of joint publications between those authors.







.2

#Papers



Co-Author Visualization – Geomap



Expertise Profile Visualization -- UCSD Map of Science

Maps of science can be used to explore, understand, and communicate the expertise profiles of institutes or nations; to chart career trajectories; to identify emerging research frontiers. They allow us to track the emergence, evolution, and disappearance of topics and help to identify the most promising areas of research.



NEWT (new ERC)



Map updated by SciTech Strategies, OST, and CNS in 2011.

TANMS (most interdisciplinary, *Nature* paper)



2008 The Regents of the University of California and SciTech Strategies. Map updated by SciTech Strategies, OST, and CNS in 2011.

ASSIST (most pubs are in Chem, Mech & Civil Eng, EE and CS, Math and Physics)



2008 The Regents of the University of California and SciTech Strategies. Map updated by SciTech Strategies, OST, and CNS in 2011.

Legend

Circle area: Fractional Journal Count

How To Read This Map

The UCSD map of science depicts a network of 554 subdiscipline nodes that are aggregated to 13 main disciplines of science. Each discipline has a distinct color and is labeled. Overlaid are circles, each representing all records per unique subdiscipline. Circle area is proportional to the number of fractionally assigned records. Minimum and maximum data values are given in the legend.





Visualizations Used in ERC Annual Reports

NEWT Report



Figure A. NEWT Year 1 and Year 2 network-based co-authorship map illustrating publication collaborations. Credit: Cyberinfrastructure for Network Science Center, University of Indiana.

Translational Research Awards Translational research has begun to ramp up in Year 2, below we summarize some highlighted outcomes from those projects.

We investigated membrane fouling mechanisms in Cambrian's hybrid membrane bioreactor-ion exchange membrane process for winery wastewater treatment applications, evaluated and developed fouling control strategies. The recommended fouling control strategy will be used by Cambrian in its full scale system operation.

We collaborated with the nanophotonics materials team at Luna Innovations Inc. to develop and evaluate new photothermal nano materials that can improve the photothermal efficiency of the MD membrane. Titanium nitride nanoparticles, mixture of carbon black and titanium nitride, and carbon coated titanium nitride were evaluated both in suspension and membrane coatings. The Phase (II) proposal was submitted in April for The figures below show the evolving NEWT collaboration network based on coauthorship extracted from bibliography files (Figure A). Legends are included in the bottom-left of each visualization to explain the sizing of author nodes and collaboration links. Each node represents an author, and author node area size scales with the number of publications. Author nodes with two publications or more are labelled by the author's name in the top visualization, while author nodes with three publications or more are labelled on the bottom visualization. Two authors are connected if they have authored a publication together and link width scales with the number of joint 26publications between those authors. Note that size coding differs for the two visualizations.

Network layouts differ structurally as each is spatially optimized for each of the two periods. The network on top shows the 78 authors that published 15 publications in 2015/2016. Note that one publication has 16 authors resulting in a so-called fully connected clique network that is connected to other subnetworks via Pedro J.J. Alvarez. The network below shows 173 authors that published 48 journal articles in 2016/2017. Comparing the two visualizations reveals that NEWT impact in terms of authors and publications has increased considerably—there are about more than twice the number of authors publishing almost four times more publications within 2016/2017 than in previous years. Plus, there are many larger subnetworks that are more interlinked—showcasing intense collaboration and communication. Authors like Pedro J.J. Alvarez, Jorge Gardea-Torresdey, and Paul Westerhoff have not only many publications to their credit but they also interconnect different subnetworks—effectively serving as gatekeepers. Thickness coding of lines supports the identification of major collaboration (and most likely communication) pathways in this rapidly evolving professional network.



NEWT Reporting Year 1







TANMS Report



ERC Web Table 1 provides a quantifiable summary of the TANMS's research productivity related to the three testbeds. Research productivity, in addition to testbed development, is a required reporting metric by NSF to assess the health and quality of a center's program. As can be seen in this table TANMS continues to have a healthy number of journal publications, i.e. 36 publications in the last year from core funding and 35 from associated projects or a total of 229 since center inception. Maybe more complete publication information is provided in the coauthor network visualization provided in Figure 3.2.1-1 for the last couple of years.

Each node in the figure represents an author, and author node area size scales with the number of publications. Author nodes with four publications or more are labelled by the author's name. Two authors are connected if they have authored a publication together and link width scales with the number of joint publications between those authors. The network on top shows the authors network for 2014/2015. The top-three authors with the most publications are Kang L. Wang, Greg P. Carman, and Pedram Khalil Amiri. The network at the bottom of the figure shows publications for 2016/2017. The top-three authors with the most publications are Kang L. Wang, Gregory P. Carman, and Guogiang Yu. Comparing these two visualizations reveals that the TANMS impact in terms of authors and publications was unusually high in the initial years and is increasing continuously—i.e. there are many more unique authors publishing within 2016/2017 than in previous years. Plus, there are large, strongly interlinked subnetworks— showcasing intense collaboration and communication. Authors like Kang L. Wang and Greg P. Carman, to their credit, also interconnect different subnetworks—effectively serving as gatekeepers. Thus, TANMS has a vibrant publication record showing extensive collaboration across all campuses. This brief snapshot clearly shows that TANMS system/team approach representing the corner stone of an NERC is alive and well.









Visualizations Used in ERC Site Visit Presentations

CNS @ Indiana University 2017

#Papers



Muth,**d**ohn F Zhu,**∛**ong

#Co-authored Papers

13 7 1

This network was generated from 271 ASSIST publications. It comprises 48 authors and 128 edges and covers the years 2012 through March 2017. Nodes are sized by the cumulative number of publications associated with an author. Edges are sized by the cumulative number of ASSIST related papers published by an author.

Lach, **J**ohn C.

Calhoun, Benton H.

Wentzloff, David D

Rahn, Ch**ri**stopher D

Trolier-mcki**n**stry, Susan





Trolier-mck**in**stry, Susan



Displayed Year: 2012-2013

CNS @ Indiana University 2017







#Co-authored Papers



This network was generated from 271 ASSIST publications. It comprises 48 authors and 128 edges and covers the years 2012 through March 2017. Nodes are sized by the cumulative number of publications associated with an author. Edges are sized by the cumulative number of ASSIST related papers published by an author.









2017

#Papers

38 18





Displayed Year: 2012-2017



related papers published by an author.











Next Steps & Possible Future Visualizations

Next Steps

- Visualization code is fully integrated in nanoHUB. Decide if visualizations should be added to HUBzero.
- Optimize current visualizations based on user feedback. Interactive visualizations are best for exploration; static visualizations are required for slides, reports, publications.
- Provide guidance on ERC activity data acquisition—what data should be captured in which format to support what kind of decision making.
- Explore visualizations that provide additional insights into ERC usage and the impact of resources on S&T progress, see subsequent two slides.



XDMOD: Sankey Diagram of IT Resource Impact on Research (funded by different NSF project)

Visualization displays the relationship between IT resources, funding, and publications. The width of each line represents grant dollars awarded to researchers at one institution.

Scrivner, Olga, Gagandeep Singh, Sara Bouchard, Scott Hutcheson, Ben Fulton, Matt Link, and Katy Börner. 2018. <u>"XD Metrics on</u> <u>Demand Value analytics:</u> <u>Visualizing the impact of internal</u> <u>information Technology</u> <u>investments on external Funding,</u> <u>Publications, and collaboration</u> <u>networks</u>". *Frontiers Research Metrics and Analytics* 2 (10).



Study and visualize the structure and dynamics of Engineering, particularly the interplay of

- Job market demands
- Education and training (residential and online)
 S&T progress

Communicate S&T dynamics to general audiences via moderated news broadcasts





Science & Technology vs. Education/Training vs. Jobs

Katy Börner, Olga B. Scrivner, Xiaozhong Liu, Indiana University

Need to study the **(mis)match** and **temporal dynamics** of S&T progress, education and workforce development options, and job requirements.

Challenges:

- Rapid change of STEM knowledge
- Increase in tools, AI
- Social skills (project management, team leadership)
- Increasing team size



Science & Technology vs. Education/Training vs. Jobs

Study results are needed by:

- **Students:** What jobs will exist in 1-4 years? What program/learning trajectory is best to get/keep my dream job?
- **Teachers:** What course updates are needed? What curriculum design is best? What is my competition doing? How much timely knowledge (to get a job) vs. forever knowledge (to be prepared for 80 productive years) should I teach? How to innovate in teaching and get tenure?
- Employers: What skills are needed next year, in 5 years? Who trains the best? What skills does my competition list in job advertisements? How to hire/train productive teams?

lobs Courses Science & Technology

What is ROI of my time, money, compassion?

D NATIONAL AG	CADEMY OF SCIEN	ICES		Search Site					
ABOUT THE NAS	MEMBERSHIP	PROGRAMS	PUBLICATIONS	MEMBER LOGIN					
Arthur M. Sa	ckler Colloquia Completed Collog	quia All Completed Colloqui	a	Share					
PROGRAMS Sackler Colloquia	Arthur M. Sackl	er OQUIA							
About Sackler Colloquia									
Upcoming Colloquia	Completed Colloquia								
Completed Colloquia	Below is a list of completed colloquia, beginning with the most recent. Follow the link from the title of each colloquium to view a description of the colloquium and its program, and also find links to videos and PNAS papers as they become available.								
Sackler Lectures									
Video Gallery	2018								
 Connect with Sackler Colloquia 	Creativity and Collaboration: Revisiting Cybernetic Serendipity March 13-14 2018; Washington, D.C. Organized by Ben Shneiderman, Maneesh Agrawala, Alyssa Goodman, Youngmoo Kim, and Boger Malina								
Give to Sackler Colloquia	Grganized by Ben Sineiderman, Maneesh Agrawala, Alyssa Goodman, Youngmoo Kim, and Roger Malina								
Cultural Programs	January 17-18, 2018: Irvine, CA								
Distinctive Voices	Organized by Simon Levin, Stephen Carpenter, Gretchen Daily, Sir Partha Dasgupta, Paul Ehrlich, Geoffrey Heal, Catherine Kling, Jane Lubchenco, and Stephen Polasky								
Kavli Frontiers of Science	Videos								
Keck Futures Initiative	2017								
LabX	Modeling and Visualizing S	cience and Technology Deve	lopments						
Sackler Forum	December 4-5, 2017; Irvine, C Organized by Katy Börner, W Videos	CA illiam Rouse, H. Eugene Stanle	ey, and Paul Trunfio						



http://www.nasonline.org/programs/sackler-colloquia/completed_colloquia/modeling-and-visualizing.html ³⁸

Atlas Trilogy

- Börner, Katy (2010) Atlas of Science: Visualizing What We Know. The MIT Press. <u>http://scimaps.org/atlas</u>
- Börner, Katy (2015) Atlas of Knowledge: Anyone Can Map. The MIT Press. <u>http://scimaps.org/atlas2</u>
- 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





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





Research



Submitted paper: MOOC Visual Analytics

Development



Online version of AcademyScope now available on the National Academies Press website

Videos



Watch Katy Börner's presentation of Humanexus at the IU **CEWiT Faculty Alliance** Salon

Latest News



Börner Appointed to U.S. Department of Commerce - Data Advisory Council

Outreach



CDC Museum to Host Places & Spaces: **Mapping Science** Exhibition

Teaching



Purchase Visual Insights, the IVMOOC companion textbook

Our Products



We work closely with clients to provide custom-made data, visualization, & software solutions