

Empowering Anyone to Envision Science, Technology, and Innovation

Katy Börner

Victor H. Yngve Professor of Information Science
Director, Cyberinfrastructure for Network Science Center
School of Informatics and Computing, Indiana University, USA

Workshop on Visualization for Big Data: Strategies and Principles
The Fields Institute, Toronto, Canada
February 23–27, 2015

Language Communities of Twitter - Eric Fischer - 2012

READINGS

Papers

- Stipelman, Brooke A., Hall, Kara L., Zoss, Angela, Okamoto, Janet, Stokols, Dan, and Börner, Katy (2014) Mapping the Impact of Transdisciplinary Research: A Visual Comparison of Investigator Initiated and Team Based Tobacco Use Research Publications. *The Journal of Translational Medicine and Epidemiology*.
- Bollen, Johan, David Crandall, Damion Junk, Ying Ding, and Katy Börner. 2014. [From funding agencies to scientific agency: Collective allocation of science funding as an alternative to peer review](#). *EMBO Reports* 15 (1): 1-121.
- Mazloumian, Amin, Dirk Helbing, Sergi Lozano, Robert Light, and Katy Börner. 2013. [Global Multi-Level Analysis of the 'Scientific Food Web'](#). *Scientific Reports* 3, 1167.

Books

- Börner, Katy. 2015. *Atlas of Knowledge: Anyone Can Map*. The MIT Press.
- Börner, Katy, and David E. Polley. 2014. *Visual Insights: A Practical Guide to Making Sense of Data*. Cambridge, MA: The MIT Press.
- Scharnhorst, Andrea, Katy Börner, and Peter van den Besselaar, eds. 2012. *Models of Science Dynamics*. Springer Verlag.
- Börner, Katy, Mike Conlon, Jon Corson-Rikert, and Ying Ding, eds. 2012. *VIVO: A Semantic Approach to Scholarly Networking and Discovery*. Morgan & Claypool Publishers LLC.
- Börner, Katy. 2010. *Atlas of Science: Visualizing What We Know*. The MIT Press.

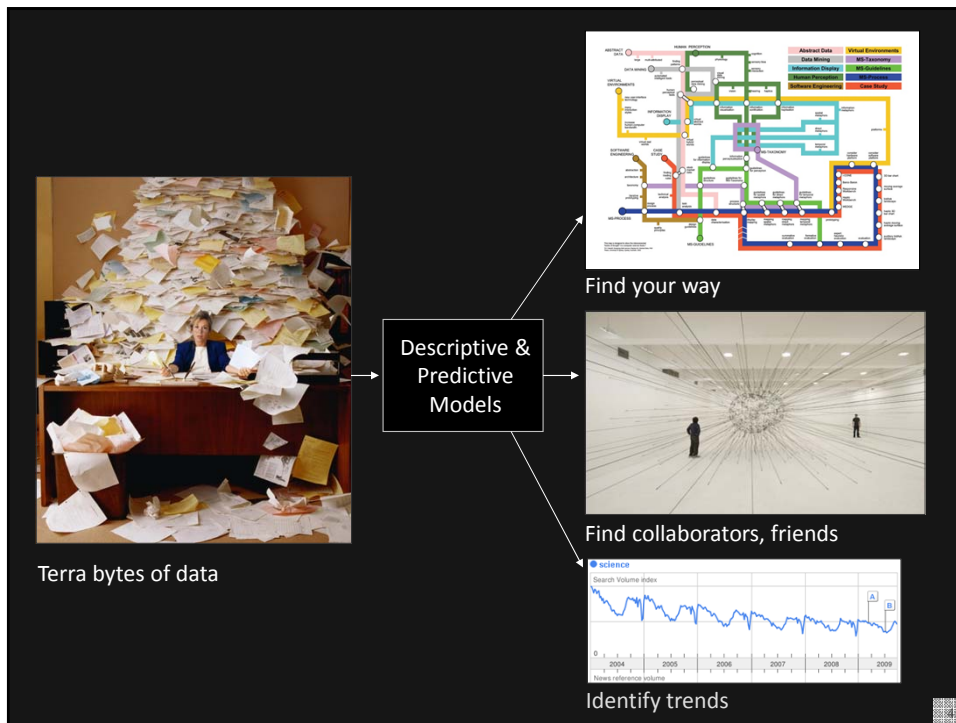
2

Humanexus

[Watch the official trailer »](#)



Producer/Script Writer: Katy Börner, Designer/Artist: Ying-Fang Shen, Sound Artist: Norbert Herber, 2013.
<http://cns.iu.edu/humanexus>



Descriptive Models

Multiple levels: Micro ... Macro

Answering: When? Where? What? With Whom?



Different Levels of Abstraction/Analysis

Macro/Global
Population Level



Meso/Local
Group Level



Micro
Individual Level

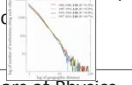






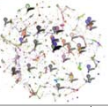

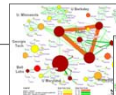
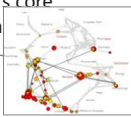


Type of Analysis vs. Level of Analysis

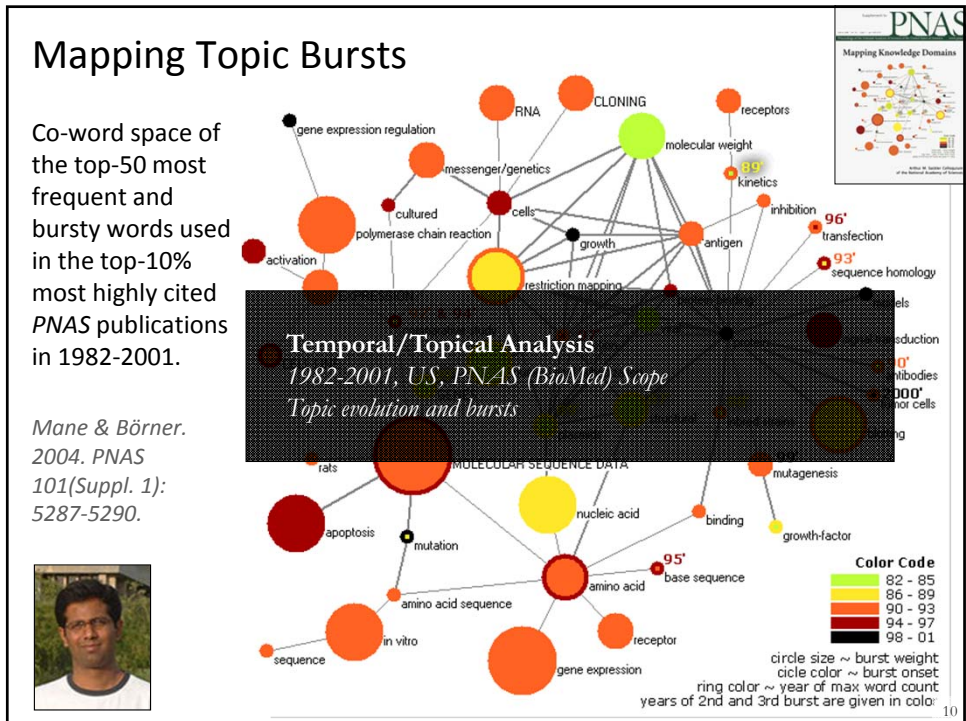
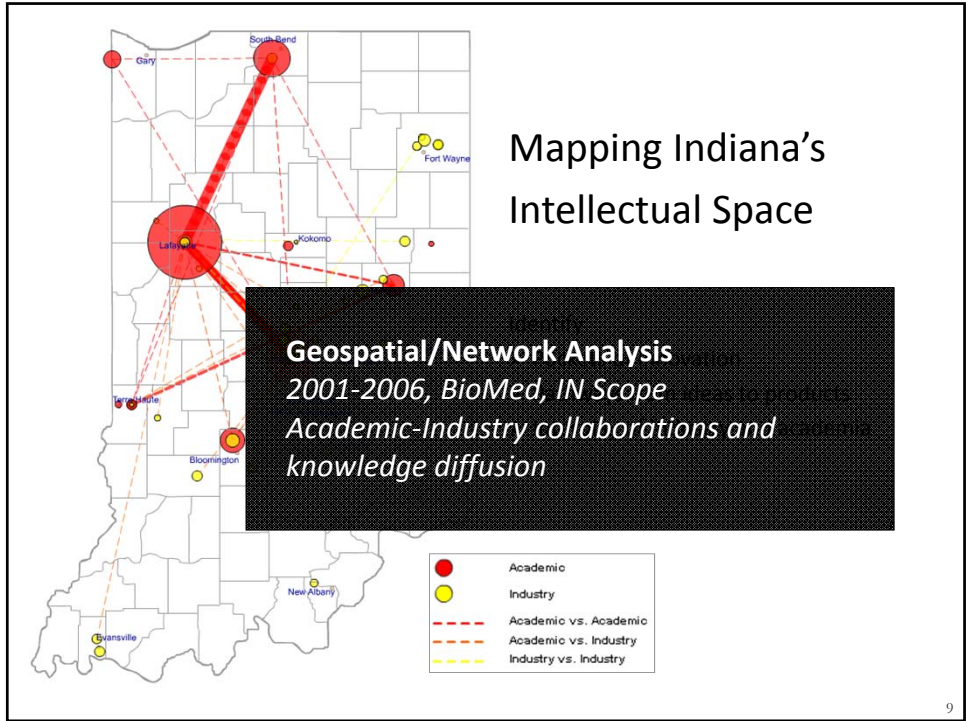
	Micro/Individual (1-100 records)	Meso/Local (101-100,000 records)	Macro/Global (100,000 < records)
Statistical Analysis/Profiling	Individual person and their expertise profiles	Larger labs, centers, universities, research domains, or states	All of NSF, all of USA, all of science.
Temporal Analysis (When?)	Funding portfolio of one individual	Mapping topic bursts in 20-years of PNAS	113 Years of Physics Research
Geospatial Analysis (Where?)	Career trajectory of one individual	Mapping a states intellectual landscape	PNAS publications
Topical Analysis (What?)	Base knowledge from which one grant draws.	Knowledge flows in Chemistry research	VxOrd/Topic maps of NIH funding
Network Analysis (With Whom?)	NSF Co-PI network of one individual	Co-author network	NIH's core competency

7

Type of Analysis vs. Level of Analysis

	Micro/Individual (1-100 records)	Meso/Local (101-100,000 records)	Macro/Global (100,000 < records)
Statistical Analysis/Profiling	Individual person and their expertise profiles	Larger labs, centers, universities, research domains, or states	All of NSF, all of USA, all of science. 
Temporal Analysis (When?)	Funding portfolio of one individual	Mapping topic bursts in PNAS 	113 Years of Physics Research 
Geospatial Analysis (Where?)	Career trajectory of one individual	Mapping a states intellectual landscape 	PNAS publications 
Topical Analysis (What?)	Base knowledge from which one grant draws. 	Knowledge flows in Chemistry research 	VxOrd/Topic maps of NIH funding 
Network Analysis (With Whom?)	NSF Co-PI network of one 	Co-author network 	NIH's core competency 

8



Spatio-Temporal Information Production and Consumption of Major U.S.

Research Institutions

Börner, Penumathy, Meiss, & Ke.
2006. "Mapping the Diffusion of Scholarly Knowledge Among Major U.S. Research Institutions." *Scientometrics* 68 (3): 415-426.



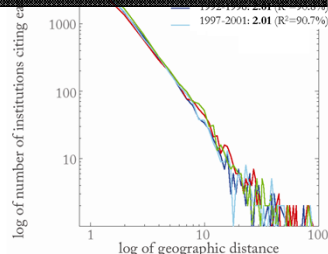
Research questions:

- 1. Does space still matter in the Internet age?
- 2. Does one still have to study and work at research institutions in order to produce high-quality data and high-quality research?
- 3. Does the Internet change research patterns—i.e., more papers produced at geographically distant research institutions?



Temporal/Geospatial Analysis
1982-2001, US, PNAS (BioMed) Scope
Citation impact and knowledge diffusion

1,505 - 1,771
1,772 - 2,097
2,098 - 2,529
2,530 - 3,039
3,040 - 4,172

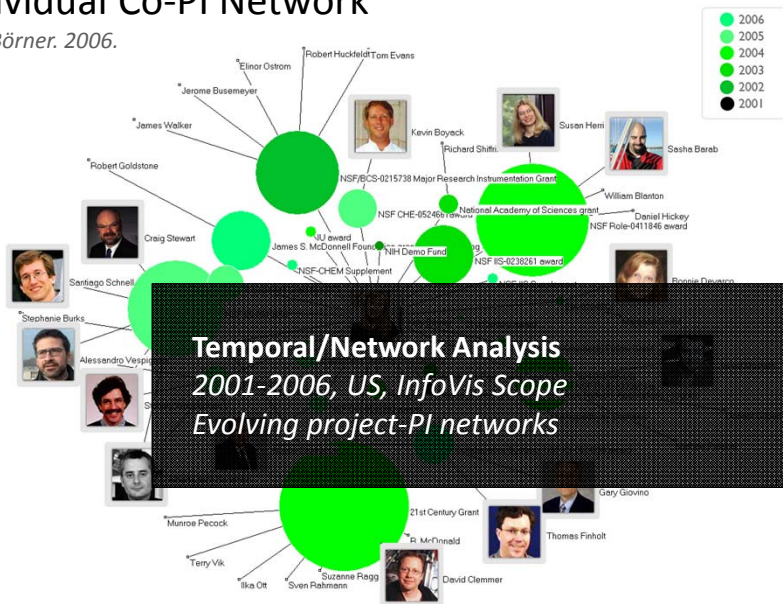


Contributions:

- Answer to Qs 1 + 2 is YES.
- Answer to Qs 3 is NO.
- Novel approach to analyzing the dual role of institutions as information producers and consumers and to study and visualize the diffusion of information among them.

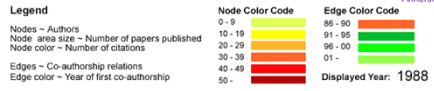
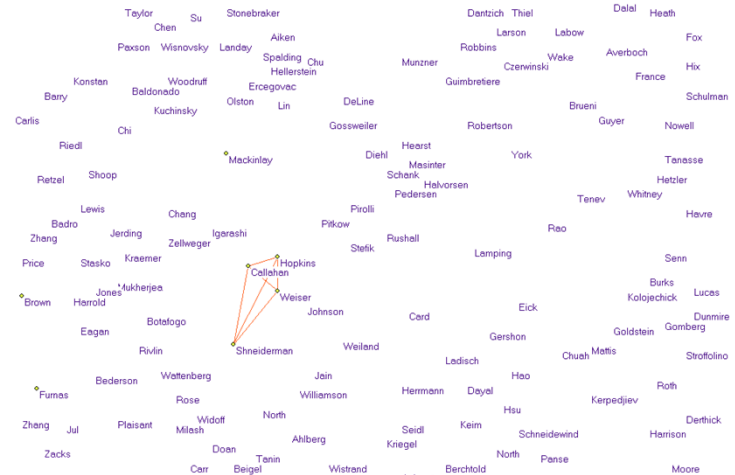
Individual Co-PI Network

Ke & Börner. 2006.



Mapping the Evolution of Co-Authorship Networks

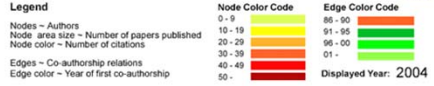
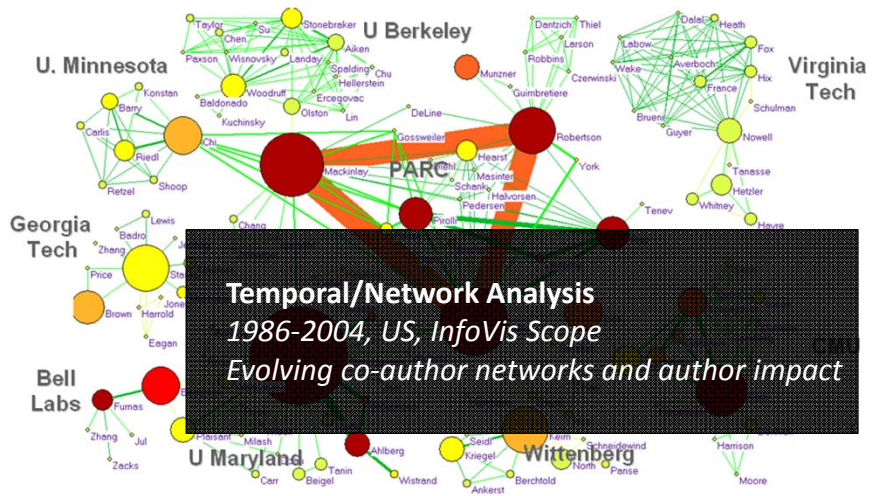
Ke, Visvanath & Börner. 2004. Won 1st prize at the IEEE InfoVis Contest.



Mapping the Evolution of Co-Authorship Networks
Weimao Ke, Lalitha Visvanath & Katy Börner
InfoVis Lab @ Indiana University
2004

Mapping the Evolution of Co-Authorship Networks

Ke, Visvanath & Börner. 2004. Won 1st prize at the IEEE InfoVis Contest.



Mapping the Evolution of Co-Authorship Networks
Weimao Ke, Lalitha Visvanath & Katy Börner
InfoVis Lab @ Indiana University
2004

Studying the Emerging Global Brain: Analyzing and Visualizing the Impact of Co-Authorship Teams

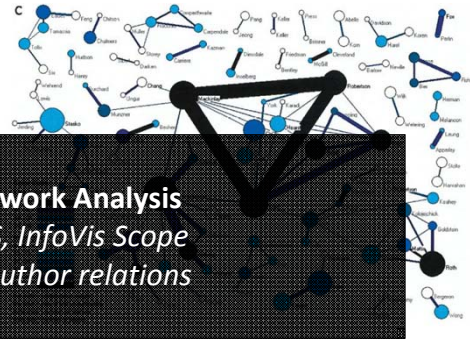
Börner, Dall'Asta, Ke & Vespignani. 2005. *Complexity* 10 (4):58-67.

Research question:

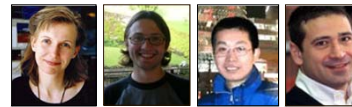
- Is science driven by prolific single experts or by high-impact co-authorship teams?

Contributions:

- New approach to allocate citational credit.
- Novel weighted graph.
- Visualization of the co-author network.
- Centrality measures of impact.
- Global statistical analysis of paper production and citations in correlation with co-authorship team size over time.
- Local, author-centered entropy measure.



Temporal/Network Analysis
1986-2004, US, InfoVis Scope
Impact of co-author relations



15

113 Years of *Physical Review*

http://scimaps.org/dev/map_detail.php?map_id=171

Herr II, Duhon, Hardy, Penumathy & Börner.

113 Years of Physical Review

The magazine aggregates 1000+ articles published in 100 volumes of *Physical Review* (1918-2010). The 1000+ articles are grouped into 100+ categories of Physics and Astronomy. *Physical Review* is the most cited journal in the field of Physics and Astronomy. It has been published continuously since 1918. It is the only journal in the field of Physics and Astronomy that has been published continuously since 1918. It is the only journal in the field of Physics and Astronomy that has been published continuously since 1918.

Nobel Prizes in Physical Review

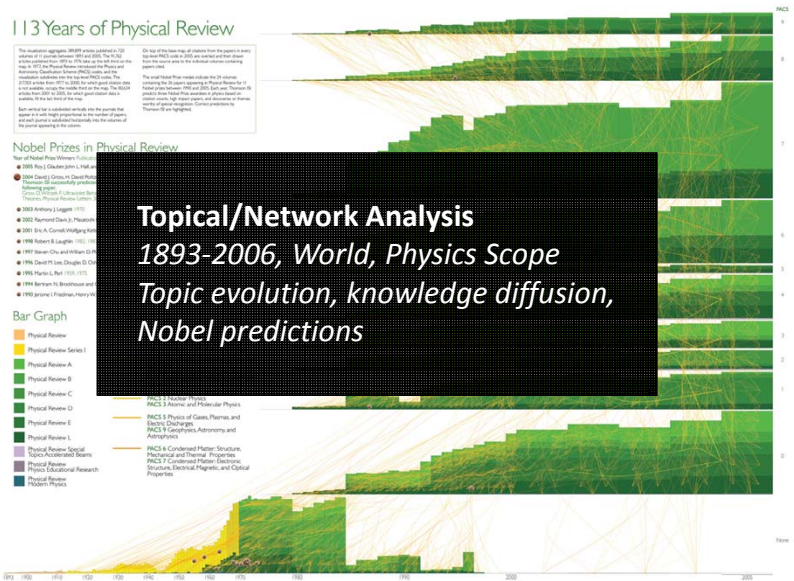
- 2005 Roy Choudhury, L. H. Liu, and...
- 2004 David G. Semakula, David...
- 2003 Anthony J. Leggett (1919)
- 2002 Raymond Davis Jr. (1916)
- 2001 Eric A. Cornell, Wolfgang Ketterle...
- 1998 Robert B. Laughlin (1915)
- 1997 Isaac Newton (1642-1726)
- 1996 David H. Lee, Douglas D. Osheroff...
- 1995 Martin L. Perl (1927-1995)
- 1994 Ben Roy Brattberg and...
- 1993 Jerome I. Friedman, Henry...



Bar Graph

- Physical Review
- Physical Review Series I
- Physical Review A
- Physical Review B
- Physical Review C
- Physical Review D
- Physical Review E
- Physical Review L
- Physical Review Special Topics: Applied Physics
- Physical Review Special Topics: Statistical Physics
- Physical Review Special Topics: Fluid Dynamics
- Physical Review Special Topics: Plasma Physics

Topical/Network Analysis
1893-2006, World, Physics Scope
Topic evolution, knowledge diffusion,
Nobel predictions



16

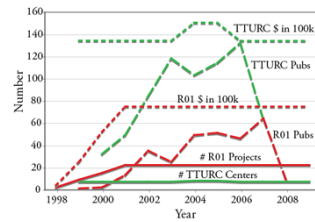
Mapping Transdisciplinary Tobacco Use Research Centers Publications

Compare R01 investigator-based funding with TTURC Center awards in terms of number of publications and evolving co-author networks.

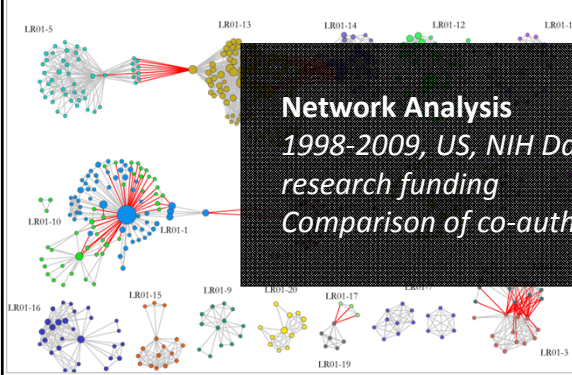
Stipelman, Hall, Zoss, Okamoto, Stokols, Börner, 2014.

Supported by NIH/NCI Contract HHSN261200800812

R01 & TTURC Project Information



Longitudinal R01 Co-Authorship Network



TTURC Co-Authorship Network



Network Analysis
1998-2009, US, NIH Data on Tobacco research funding
Comparison of co-author networks

Research Collaborations by the Chinese Academy of Sciences

Huang, Duhon, Hardy & Börner

北京地区中科院各院所



中科院与世界各地的研究合作关系

黄维霞, Russell J. Duhon, Elisha F. Hardy, Katy Börner, Indiana University, USA

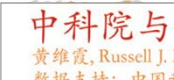
数据支持: 中国科学院国家科学图书馆科学前沿分析中心

金碧辉, 岳峰

甘肃省中科院各院所



上海地区中科院各院所



吉林省中科院各院所



中科院与世界各地的研究合作关系

黄维霞, Russell J. Duhon, Elisha F. Hardy, Katy Börner, Indiana University, USA

数据支持: 中国科学院国家科学图书馆科学前沿分析中心

金碧辉, 岳峰

数据展示了中科院与全球 100 多个国家的 2,954 次合作

合作次数

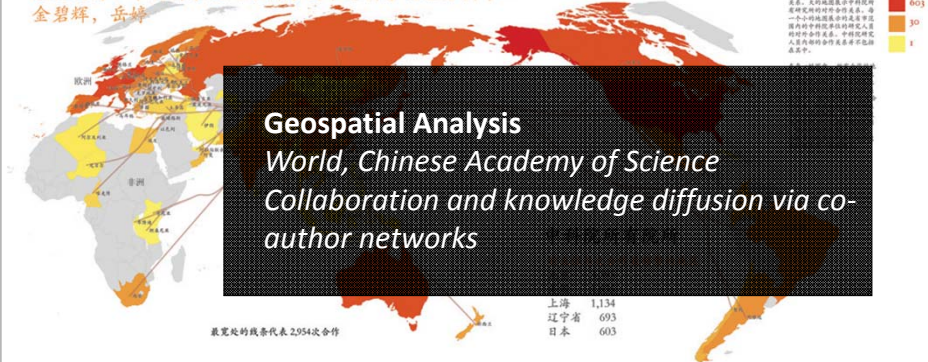
3,395

603

30

1

Geospatial Analysis
World, Chinese Academy of Science
Collaboration and knowledge diffusion via co-author networks



最宽处的线代表 2,954 次合作

上海 1,134
辽宁省 693
日本 603

The Global 'Scientific Food Web'

Mazloumian, Amin, Dirk Helbing, Sergi Lozano, Robert Light, and Katy Börner. 2013. "Global Multi-Level Analysis of the 'Scientific Food Web'". *Scientific Reports* 3, 1167.

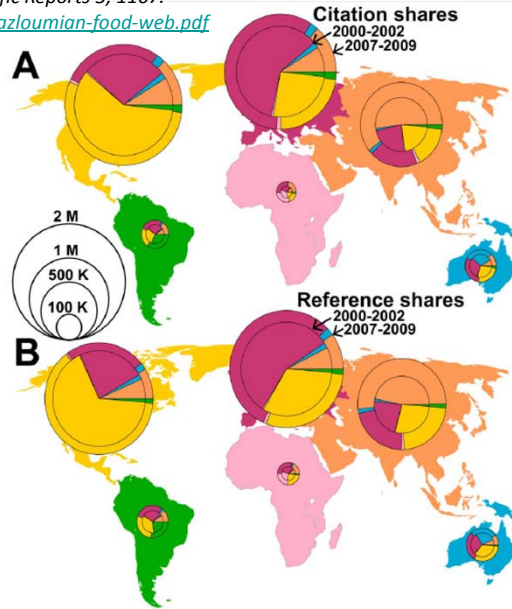
<http://cns.iu.edu/docs/publications/2013-mazloumian-food-web.pdf>

Contributions:

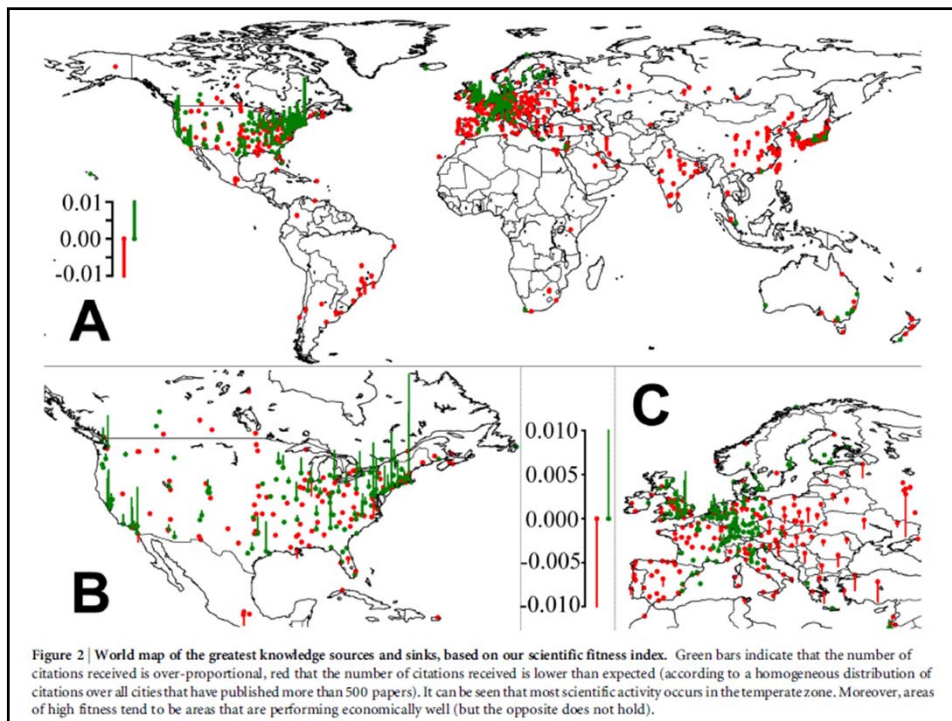
Comprehensive global analysis of scholarly knowledge production and diffusion on the level of continents, countries, and cities.

Quantifying knowledge flows between 2000 and 2009, we identify global sources and sinks of knowledge production. Our knowledge flow index reveals, where ideas are born and consumed, thereby defining a global 'scientific food web'.

While Asia is quickly catching up in terms of publications and citation rates, we find that its dependence on knowledge consumption has further increased.

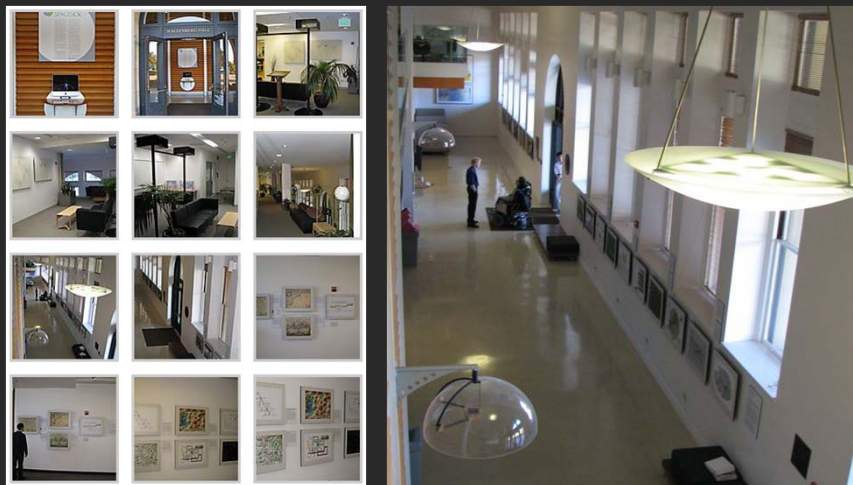


19



Visualizing Science Technology and Innovation (STI)

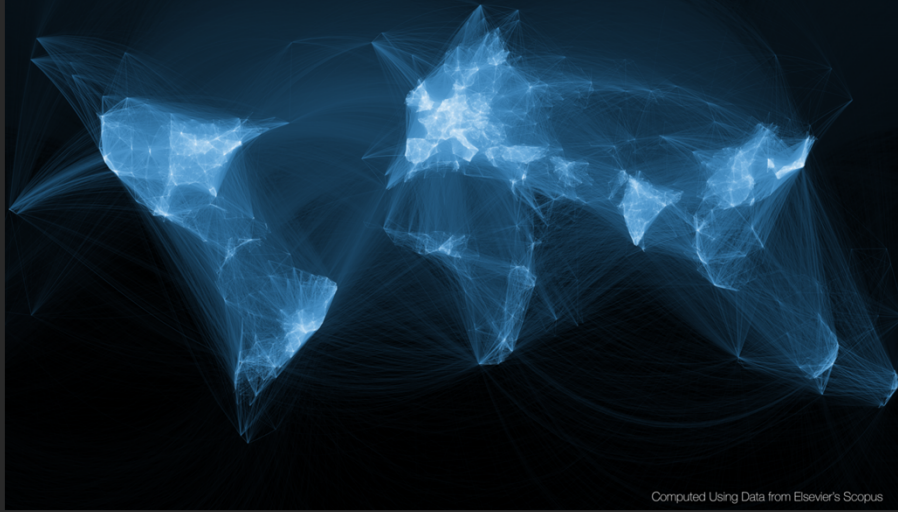
Example: Places & Spaces: Mapping Science Exhibit



Mapping Science Exhibit on display at MEDIA X, Stanford University
<http://mediox.stanford.edu>, <http://scaleindependentthought.typepad.com/photos/scimaps>



Map of Scientific Collaborations from 2005-2009



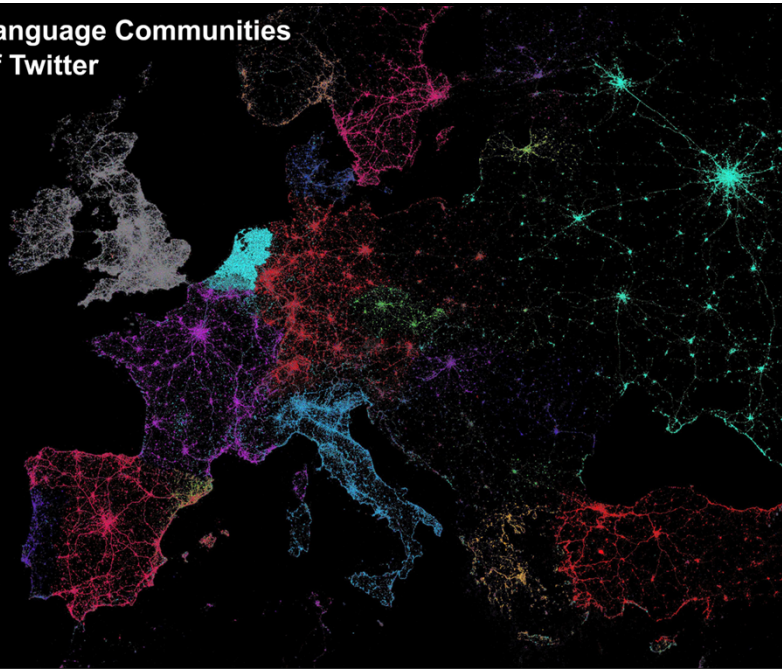
Computed Using Data from Elsevier's Scopus

Olivier H. Beauchesne, 2011. Map of Scientific Collaborations from 2005-2009.



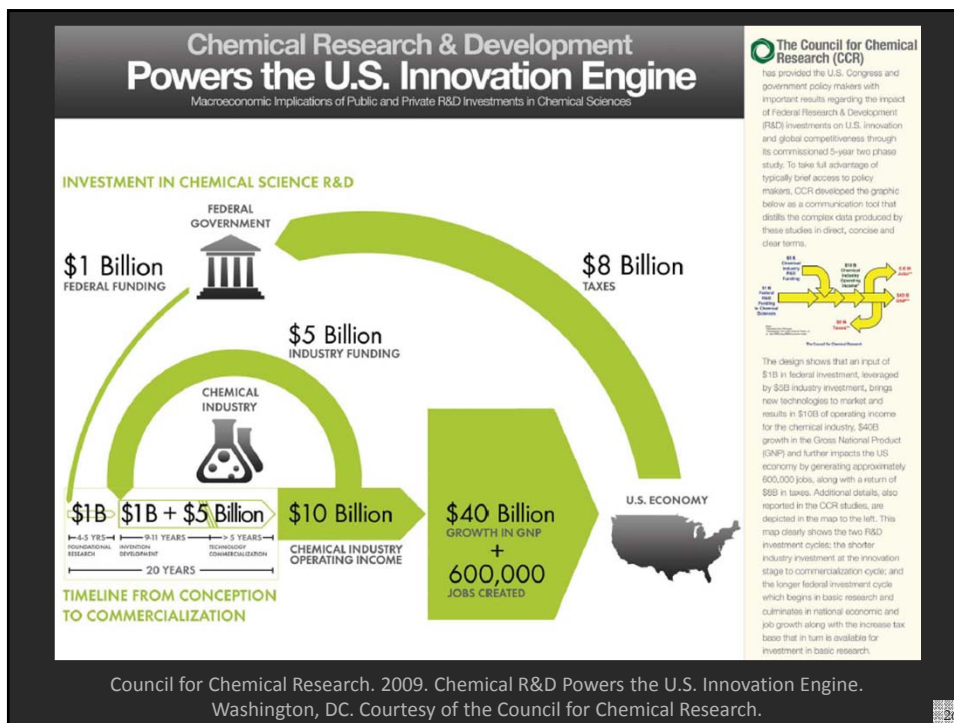
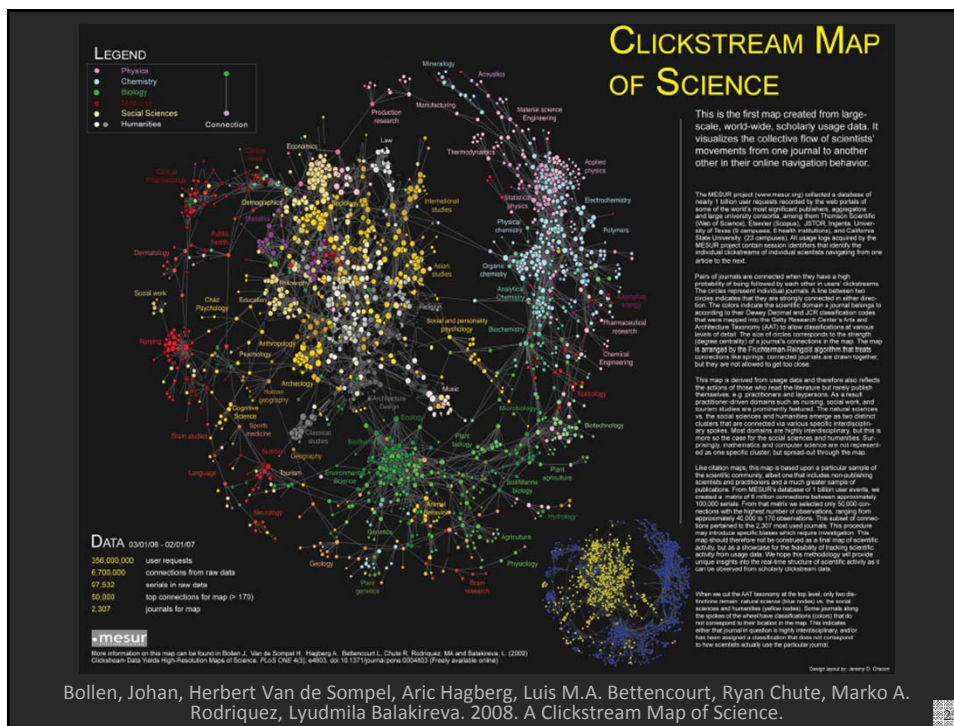
Language Communities of Twitter

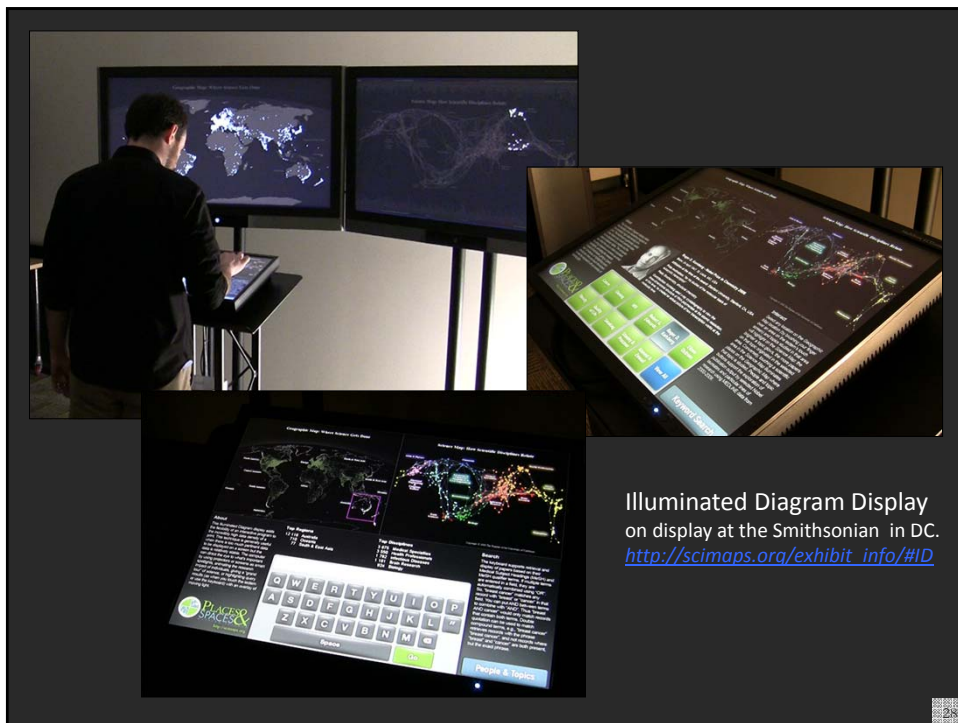
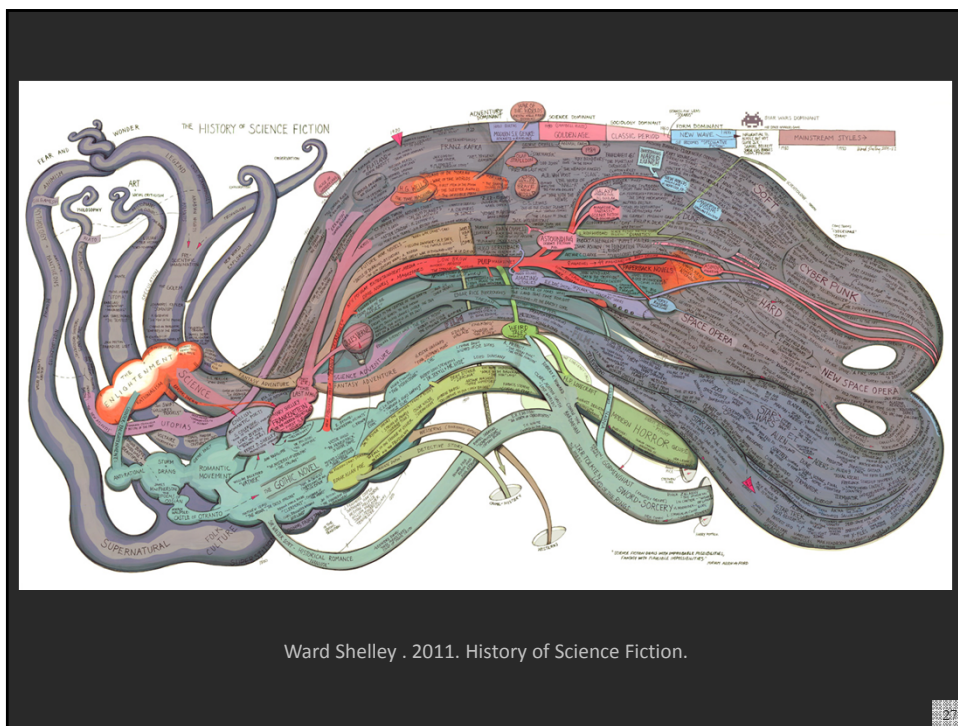
- English
- Portuguese
- Spanish
- Dutch
- Russian
- French
- Italian
- German
- Turkish
- Arabic
- Swedish
- Danish
- Finnish
- Catalan
- Romanian
- Norwegian
- Lithuanian
- Slovak
- Czech
- Greek
- Hungarian
- Polish
- Slovenian
- Albanian
- Latvian
- Galician
- Hebrew
- Croatian
- Bulgarian



Language Communities of Twitter - Eric Fischer - 2012







Geographic Map: Where Science Gets Done

North America, South America, Europe, Africa, Asia, Oceania, Antarctica

Science Map: How Scientific Disciplines Relate

Math & Physics, Chemistry, Health Professionals, Social Sciences, Medicine, Biotechnology, Infectious Diseases, Brain Research, Humanities, Biology, Earth Sciences, Agricultural, Chemical, Mechanical & Civil Engineering

Copyright © 2008 The Regents of the University of California

About

This Illuminated Diagram display adds the flexibility of an interactive program to the incredibly high data density of a print. This technique is generally useful when there is too much pertinent data to be displayed on a screen but the data is relatively stable. The computer can direct the eye to what's important by using projectors or screens as smart spotlights, animating the research impact of individuals, giving a "grand tour" of science, or highlighting query results (as when you touch the lectern or use the keyboard) with an overlay of moving light.

Top Five Continents

- North America - 4,000 records
- South & East Asia - 3,589
- Australia - 2,431
- Africa - 2,206
- South America - 1,562

Top Five Scientific Disciplines

- Math & Physics - 4,000 records
- Health Professionals - 3,589
- Social Sciences - 2,431
- Agricultural, Chemical, Mechanical & Civil Engineering - 2,208
- Humanities - 1,562

Search

The keyboard supports retrieval and display of papers based on their Medical Subject Headings (MeSH) and MeSH qualifier terms. If multiple terms are entered in a field, they are automatically combined using "OR". So "breast cancer" matches any record with "breast" or "cancer" in that field. You can put AND between terms to combine with "AND". Thus "breast AND cancer" would only match records that contain both terms. Double quotation can be used to match compound terms, e.g. "breast cancer" retrieves records with the phrase "breast cancer", and not records where "breast" and "cancer" are both present, but the exact phrase.

Q	W	E	R	T	Y	U	I	O	P
A	S	D	F	G	H	J	K	L	"
Z	X	C	V	B	N	M			

Space Go

[People & Topics](#)

Geographic Map: Where Science Gets Done

North America, South America, Europe, Africa, Asia, Oceania, Antarctica

Science Map: How Scientific Disciplines Relate

Math & Physics, Chemistry, Health Professionals, Social Sciences, Medicine, Biotechnology, Infectious Diseases, Brain Research, Humanities, Biology, Earth Sciences, Agricultural, Chemical, Mechanical & Civil Engineering

Copyright © 2008 The Regents of the University of California

About

This Illuminated Diagram display adds the flexibility of an interactive program to the incredibly high data density of a print. This technique is generally useful when there is too much pertinent data to be displayed on a screen but the data is relatively stable. The computer can direct the eye to what's important by using projectors or screens as smart spotlights, animating the research impact of individuals, giving a "grand tour" of science, or highlighting query results (as when you touch the lectern or use the keyboard) with an overlay of moving light.

Elinor Ostrom - Nobel Prize in Economic Sciences 2009
 Born: 7 August 1933, New York, NY, USA
 Affiliation at the time of the award: Indiana University, Bloomington, IN, USA, Arizona State University, Tempe, AZ, USA
 Prize motivation: "for her analysis of economic governance, especially the commons"
 Field: Economic governance
 Contribution: Challenged the conventional wisdom by demonstrating how local property can be successfully managed by local commons without any regulation by central authorities or privatization.

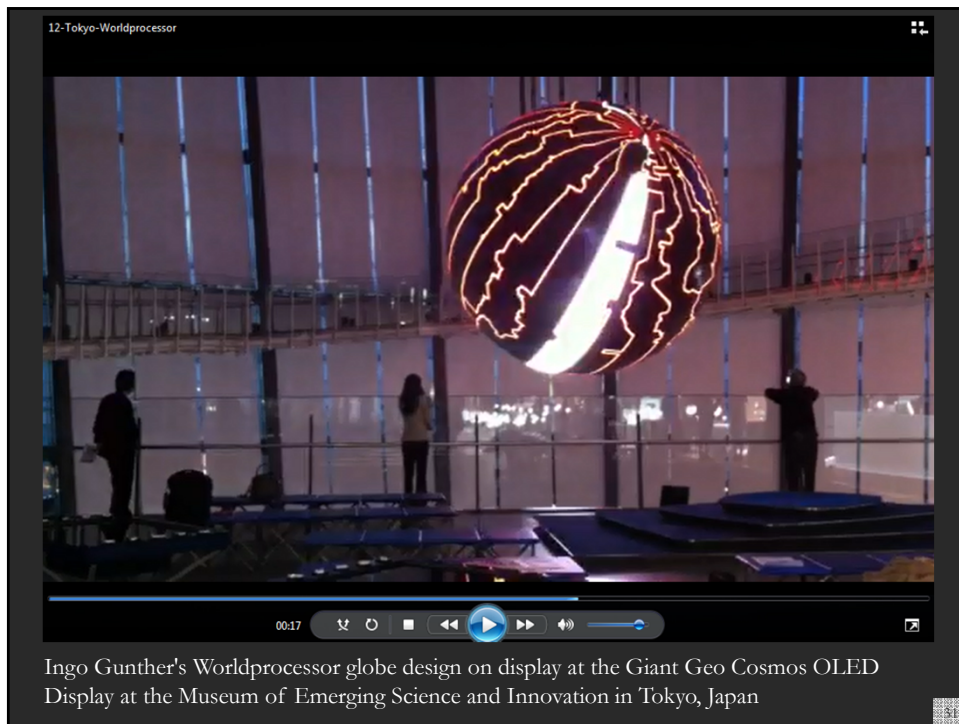
Interact

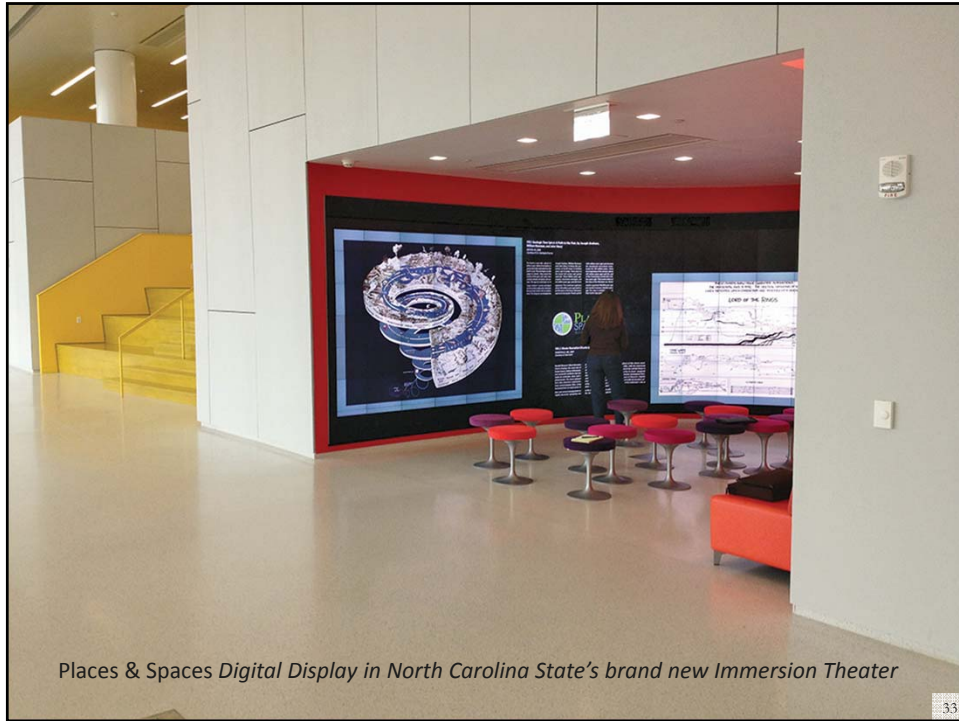
Select any location on the Geographic Map location (by brushing your finger over an area on the lectern's touch screen) and topics studied in that area will highlight on the Science Map: the brighter a topic glows, the more papers on that topic originated in the selected area. Conversely, touching a scientific area in the Science Map illuminates places on the Geographic Map where that topic is studied. People and topic buttons support the exploration of publication output by selected Nobel laureates and particular lines of research using MEDLINE data from 2000-2009.

Cancer	Cloning	HIV	Robert G. Edwards	Roger D. Kornberg	Elinor Ostrom
Obesity	Quality of Life	Smoking	Stanley B. Prusiner	Ahmed H. Zewail	View All

[Keyword Search](#)

[People & Topics](#)





Places & Spaces: Mapping Science Exhibit

<http://scimaps.org>



Maps are available for sale and the exhibit can be hosted by anyone.



Empowering Anyone to Visualize STI

Example: The Information Visualization MOOC



Information Visualization MOOC 2015

INDIANA UNIVERSITY CNS

Information Visualization MOOC 2015

Overview

This course provides an overview about the state of the art in information visualization. It teaches the process of producing effective visualizations that take the needs of users into account.

The course can be taken for three Indiana University credits as part of the [Online Data Science Program](#), as part of the Information and Library Science M.S. program, and as part of the online Data Science M.S. Program offered by the School of Informatics and Computing. Students seeking enrollment information should contact Rhonda Spencer at 812-855-2018, ilsmain@indiana.edu or datasci@indiana.edu.

Among other topics, the course covers:

- Data analysis algorithms that enable extraction of patterns and trends in data
- Major temporal, geospatial, topical, and network visualization techniques
- Discussions of systems that drive research and development.

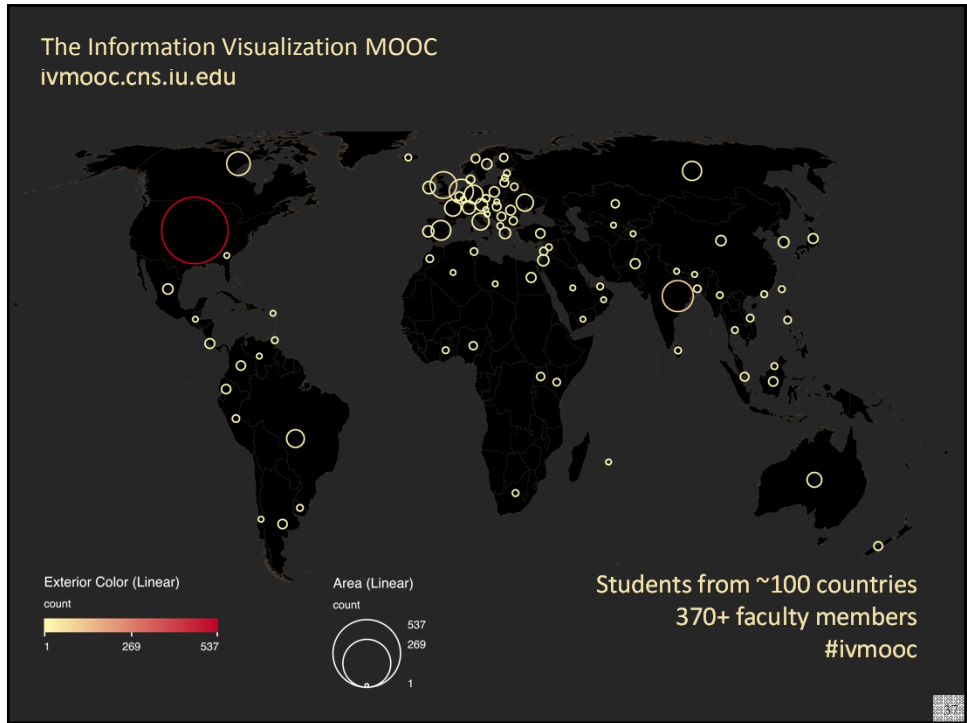
Information Visualization MOOC

[Register for Course](#)

Already registered? [Click here to go to the course.](#)
Forgot your password? [Click here to reset it.](#)

Register for free at <http://ivmooc.cns.iu.edu>. Class restarted in January 13, 2015.

36



Course Schedule

Part 1: Theory and Hands-On

- **Session 1** – Workflow Design and Visualization Framework
- **Session 2** – “When:” Temporal Data
- **Session 3** – “Where:” Geospatial Data
- **Session 4** – “What:” Topical Data

Mid-Term

- **Session 5** – “With Whom:” Trees
- **Session 6** – “With Whom:” Networks
- **Session 7** – Dynamic Visualizations and Deployment

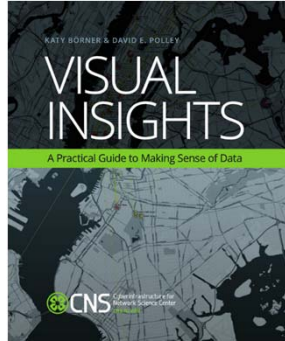
Final Exam

Part 2: Students work in teams on client projects.

Final grade is based on Class Participation (10%), Midterm (30%), Final Exam (30%), and Client Project(30%).

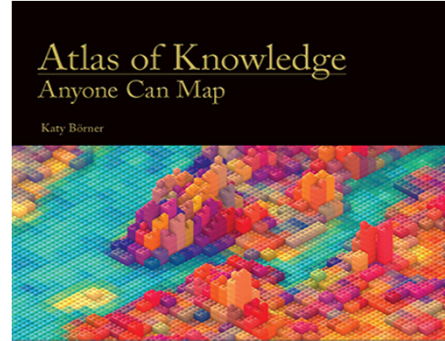


Books Used in the IVMOOC



Teaches timely knowledge:

Advanced algorithms, tools, and hands-on workflows.



Teaches timeless knowledge:

Visualization framework—exemplified using generic visualization examples and pioneering visualizations.

39

Course Schedule

Part 1: Theory and Hands-On

- **Session 1** – Workflow Design and Visualization Framework
- **Session 2** – “When:” Temporal Data
- **Session 3** – “Where:” Geospatial Data
- **Session 4** – “What:” Topical Data

Mid-Term

- **Session 5** – “With Whom:” Trees
- **Session 6** – “With Whom:” Networks
- **Session 7** – Dynamic Visualizations and Deployment

Final Exam

Part 2: Students work in teams on client projects.

Final grade is based on Class Participation (10%), Midterm (30%), Final Exam (30%), and Client Project(30%).



40

Course Schedule

Part 1: Theory and Hands-On

- **Session 1** – **Workflow Design** and Visualization Framework
- **Session 2** – “When:” Temporal Data
- **Session 3** – “Where:” Geospatial Data
- **Session 4** – “What:” Topical Data

Mid-Term

- **Session 5** – “With Whom:” Trees
- **Session 6** – “With Whom:” Networks
- **Session 7** – Dynamic Visualizations and Deployment

Final Exam

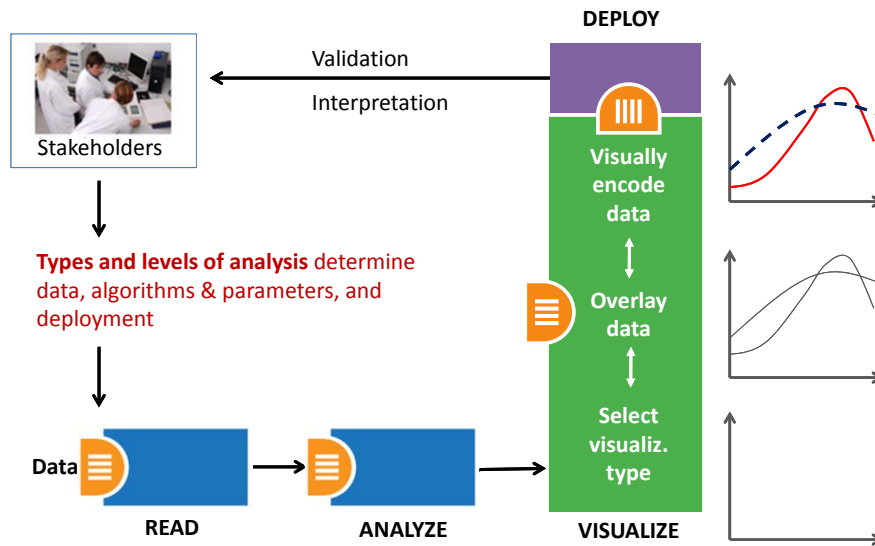
Part 2: Students work in teams on client projects.

Final grade is based on Class Participation (10%), Midterm (30%), Final Exam (30%), and Client Project(30%).

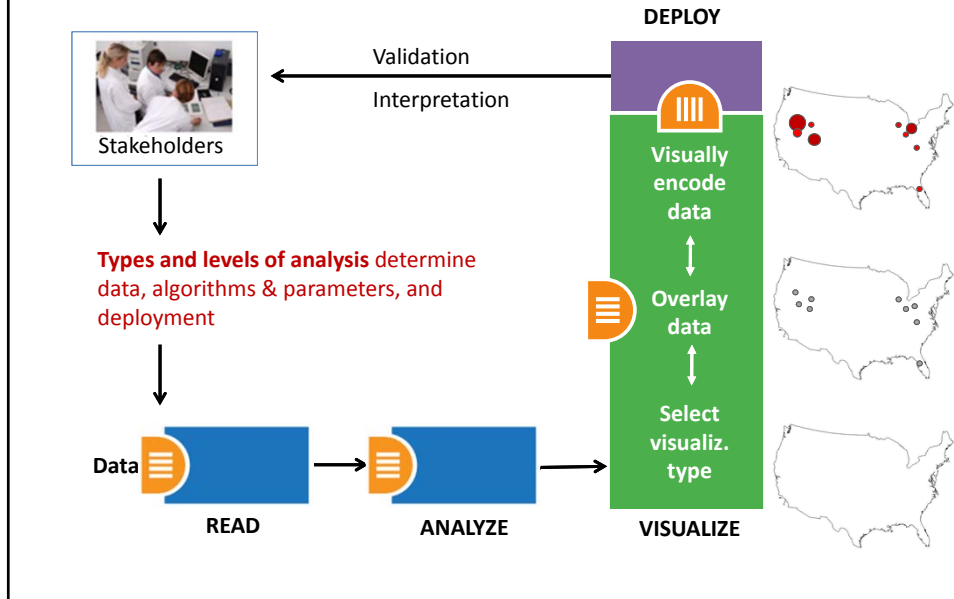


41

Needs-Driven Workflow Design



Needs-Driven Workflow Design

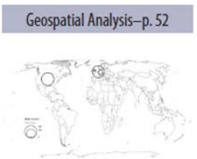
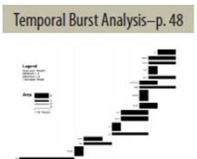


Load **One** File and Run **Many** Analyses and Visualizations

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 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, NJ 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, NJ Falk-Krzesinski, HJ Fiore, SM Hall, KL Keyton, J Spring, B Stokols, D Trochim, W Uzzi, B

Statistical Analysis—p. 44

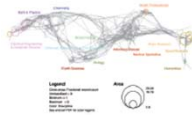
Location	Count	# Citations
Netherlands	13	292
United States	9	318
Germany	11	36
United Kingdom	1	2



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, NJ 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, NJ Falk-Krzesinski, HJ Fiore, SM Hall, KL Keyton, J Spring, B Stokols, D Trochim, W Uzzi, B

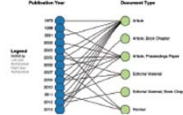
Topical Analysis—p. 56



Paper Citation Network—p. 60



Bi-Modal Network—p. 60



Co-author and many other bi-modal networks.

45

Course Schedule

Part 1: Theory and Hands-On

- **Session 1** – Workflow Design and **Visualization Framework**
- **Session 2** – “When:” Temporal Data
- **Session 3** – “Where:” Geospatial Data
- **Session 4** – “What:” Topical Data

Mid-Term

- **Session 5** – “With Whom:” Trees
- **Session 6** – “With Whom:” Networks
- **Session 7** – Dynamic Visualizations and Deployment

Final Exam

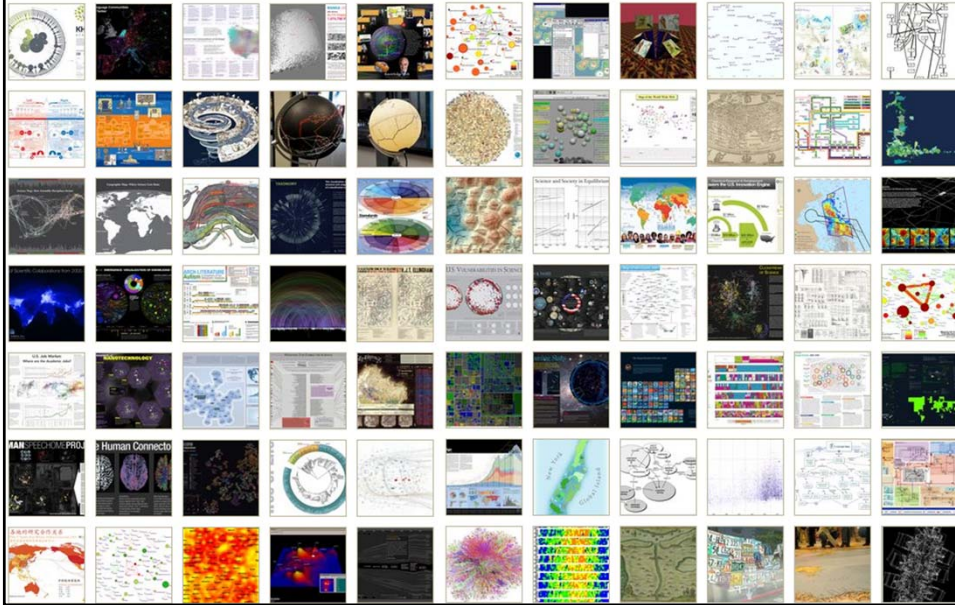
Part 2: Students work in teams on client projects.

Final grade is based on Class Participation (10%), Midterm (30%), Final Exam (30%), and Client Project(30%).



46

Visualization Frameworks



How to Classify Different Visualizations?

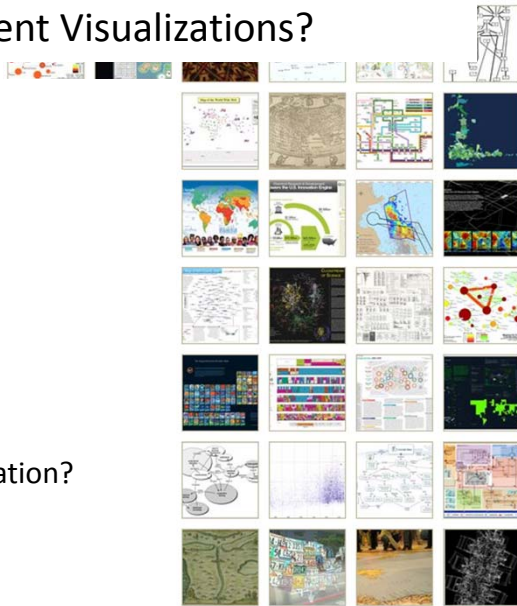
By

- User insight needs?
- User task types?

- Data to be visualized?
- Data transformation?

- Visualization technique?
- Visual mapping transformation?
- Interaction techniques?

- Or ?



Visualization Frameworks

Readings

- Bertin, Jacques. 1983. *Semiology of Graphics*. Madison, WI: University of Wisconsin Press. [Josh Quick and Richard Higgins] [Presentation](#)
- Brehmer, M. and T. Munzner. 2013. "A Multi-Level Typology of Abstract Visualization Tasks." *IEEE Transactions on Visualization and Computer Graphics*. 19 (12): 2376-2385. [Ian Custer] [Presentation](#)
- Engelhardt, Yuri. "Syntactic Structures in Graphics." See also Engelhardt, Yuri. 2002. "The Language of Graphics: A Framework for the Analysis of Syntax and Meaning in Maps, Charts, and Diagrams." PhD diss., University of Amsterdam. [Shambhavi Dhargalkar and David Ebenezer] [Presentation](#)
- Few, Stephen C. 2012. *Show Me The Numbers: Designing Tables and Graphs to Enlighten*. Burlingame, CA: Analytics Press. [Saumya Pandey and Anirudh Ramesh] [Presentation](#)
- Harris, Robert L. 1999. *Information Graphics: A Comprehensive Illustrated Reference*. New York: Oxford University Press. [Ashish Shendure and Manish Mohan Mohapatra] [Presentation](#)
- Horn, Robert E. 1998. *Visual Language: Global Communication for the 21st Century*. Bainbridge Island, WA: MacroVU, Inc.
- MacEachren, Alan M. 2004. *How Maps Work: Representation, Visualization, and Design*. New York: Guilford. 271. [Kartik Adur] [Presentation](#)
- Pirolli, Peter, and Stuart Card. 2005. "The Sensemaking Process and Leverage Points for Analyst Technology as Identified through Cognitive Task Analysis." In Proceedings of the International Conference on Intelligence Analysis, 2-4. McLean, VA: MITRE. [Renuka Deshmukh and John Stewart] [Presentation](#)
- Shneiderman, Ben. 1996. "The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations." In Proceedings of the IEEE Symposium on Visual Languages, 336-343. Los Alamitos, CA: IEEE Computer Society. [Wensi Wang and Jagadish Shrivani Kumar] [Presentation](#)
- Wehrend, Stephen C., and Clayton Lewis. 1990. "A Problem-Oriented Classification of Visualization Techniques." In Proceedings of the 1st Conference on Visualization, 139-143. Los Alamitos, CA: IEEE Computer Society. [Sowmya Achanta and Jae Eun Kum] [Presentation](#)
- Wilkinson, Leland. 2005. *The Grammar of Graphics*. New York: Springer. [Trevor Edelblute] [Presentation](#)
- Yau, Nathan. 2011. *Visualize This: The FlowingData Guide to Design, Visualization, and Statistics*. Indianapolis, IN: Wiley. (ebrary account required to read/download)

Suggestions of other major works are welcome!

We will add Kosslyn, 1989; Kindlmann & Scheidegger, 2014

Tasks	LEVELS		
	MICRO: Individual Level about 1-1,000 records page 6	MESO: Local Level about 1,001-100,000 records page 8	MACRO: Global Level more than 100,000 records page 10
TYPES			
Statistical Analysis page 44			
WHEN: Temporal Analysis page 48			
WHERE: Geospatial Analysis page 52			
WHAT: Topical Analysis page 56			
WITH WHOM: Network Analysis page 60			

See page 5

50

Basic Task Types								
Bertin, 1967	Wehrend & Lewis, 1996	Few, 2004	Yau, 2011	Rendgen & Wiedemann, 2012	Frankel, 2012	Tool: Many Eyes	Tool: Chart Chooser	Börner, 2014
selection	categorize			category				categorize/ cluster
order	rank	ranking					table	order/rank/ sort
	distribution	distribution					distribution	distributions (also outliers, gaps)
	compare	nominal comparison & deviation	differences		compare and contrast	compare data values	comparison	comparisons
		time series	patterns over time	time	process and time	track rises and falls over time	trend	trends (process and time)
		geospatial	spatial relations	location		generate maps		geospatial
quantity	part-to- whole	proportions			form and structure	see parts of whole, analyze text	composition	compositions (also of text)
association	correlate	correlation	relationships	hierarchy		relations between data points	relationship	correlations/ relationships

Atlas of Knowledge
Knowers Can Map

See page 26

Insight Need Types page 26	Data Scale Types page 28	Visualization Types page 30	Graphic Symbol Types page 32	Graphic Variable Types page 34	Interaction Types page 26
<ul style="list-style-type: none"> categorize/cluster order/rank/sort distributions (also outliers, gaps) comparisons trends (process and time) geospatial compositions (also of text) correlations/relationships 	<ul style="list-style-type: none"> nominal ordinal interval ratio 	<ul style="list-style-type: none"> table chart graph map network layout 	<ul style="list-style-type: none"> geometric symbols <ul style="list-style-type: none"> point line area surface volume linguistic symbols <ul style="list-style-type: none"> text numerals punctuation marks pictorial symbols <ul style="list-style-type: none"> images icons statistical glyphs 	<ul style="list-style-type: none"> spatial <ul style="list-style-type: none"> position retinal <ul style="list-style-type: none"> form color optics motion 	<ul style="list-style-type: none"> overview zoom search and locate filter details-on-demand history extract link and brush projection distortion

Atlas of Knowledge
Knowers Can Map

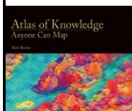
See page 24

Visualization Types (Reference Systems)

1. **Charts:** No reference system—e.g., Wordle.com, pie charts
2. **Tables:** Categorical axes that can be selected, reordered; cells can be color coded and might contain proportional symbols. Special kind of graph.
3. **Graphs:** Quantitative or qualitative (categorical) axes. Timelines, bar graphs, scatter plots.
4. **Geospatial maps:** Use latitude and longitude reference system. World or city maps.
5. **Network layouts:** Node position might depend on node attributes or node similarity. **Trees:** hierarchies, taxonomies, genealogies. **Networks:** social networks, migration flows.

Types

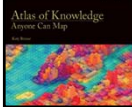
Insight Need Types page 26	Data Scale Types page 28	Visualization Types page 30	Graphic Symbol Types page 32	Graphic Variable Types page 34	Interaction Types page 26
<ul style="list-style-type: none"> • categorize/cluster • order/rank/sort • distributions (also outliers, gaps) • comparisons • trends (process and time) • geospatial • compositions (also of text) • correlations/relationships 	<ul style="list-style-type: none"> • nominal • ordinal • interval • ratio 	<ul style="list-style-type: none"> • table • chart • graph • map • network layout 	<ul style="list-style-type: none"> • geometric symbols <ul style="list-style-type: none"> point line area surface volume • linguistic symbols <ul style="list-style-type: none"> text numerals punctuation marks • pictorial symbols <ul style="list-style-type: none"> images icons statistical glyphs 	<ul style="list-style-type: none"> • spatial <ul style="list-style-type: none"> position • retinal <ul style="list-style-type: none"> form color optics motion 	<ul style="list-style-type: none"> • overview • zoom • search and locate • filter • details-on-demand • history • extract • link and brush • projection • distortion



See page 25

Graphic Variable Types Versus Graphic Symbol Types

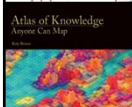
		Geometric Symbols				
		Point	Line	Area		
Spatial	x	quantitative				
	y	quantitative				
	z	quantitative				
Retinal	Form	Size	quantitative	NA (Not Applicable)		
		Shape	qualitative	NA		
		Rotation	quantitative	NA		
		Curvature	quantitative	NA		
		Angle	quantitative	NA		
		Closure	quantitative	NA		
		Value	quantitative			
Color	Hue	qualitative				
	Saturation	quantitative				



See page 36

Graphic Variable Types Versus Graphic Symbol Types

		Geometric Symbols			Color			Text		
		Point	Line	Area	Color	Value	Text	Color	Value	Text
Spatial	x	quantitative								
	y	quantitative								
	z	quantitative								
Retinal	Form	Size	quantitative	NA (Not Applicable)						
		Shape	qualitative	NA						
		Rotation	quantitative	NA						
		Curvature	quantitative	NA						
		Angle	quantitative	NA						
		Closure	quantitative	NA						
		Value	quantitative							
Color	Hue	qualitative								
	Saturation	quantitative								



See pages 36-39

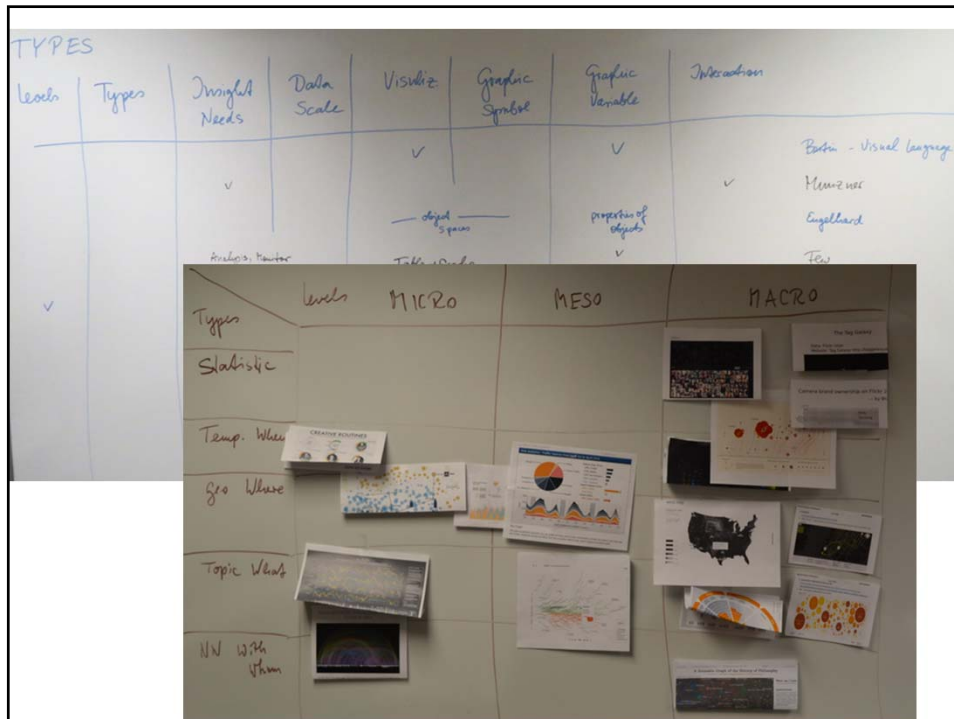
Visualization Frameworks

Readings

- Bertin, Jacques. 1983. *Semiology of Graphics*. Madison, WI: University of Wisconsin Press. [Josh Quick and Richard Higgins] [Presentation](#)
- Brehmer, M. and T. Munzner. 2013. "A Multi-Level Typology of Abstract Visualization Tasks." *IEEE Transactions on Visualization and Computer Graphics*. 19 (12): 2376-2385. [Ian Custer] [Presentation](#)
- Engelhardt, Yuri. "Syntactic Structures in Graphics." See also Engelhardt, Yuri. 2002. "The Language of Graphics: A Framework for the Analysis of Syntax and Meaning in Maps, Charts, and Diagrams." PhD diss., University of Amsterdam. [Shambhavi Dhargalkar and David Ebenezer] [Presentation](#)
- Few, Stephen C. 2012. *Show Me The Numbers: Designing Tables and Graphs to Enlighten*. Burlingame, CA: Analytics Press. [Saumya Pandey and Anirudh Ramesh] [Presentation](#)
- Harris, Robert L. 1999. *Information Graphics: A Comprehensive Illustrated Reference*. New York: Oxford University Press. [Ashish Shendure and Manish Mohan Mohapatra] [Presentation](#)
- Horn, Robert E. 1998. *Visual Language: Global Communication for the 21st Century*. Bainbridge Island, WA: MacroVU, Inc.
- MacEachren, Alan M. 2004. *How Maps Work: Representation, Visualization, and Design*. New York: Guilford. 271. [Kartik Adur] [Presentation](#)
- Pirolli, Peter, and Stuart Card. 2005. "The Sensemaking Process and Leverage Points for Analyst Technology as Identified through Cognitive Task Analysis." In *Proceedings of the International Conference on Intelligence Analysis*, 2-4. McLean, VA: MITRE. [Renuka Deshmukh and John Stewart] [Presentation](#)
- Shneiderman, Ben. 1996. "The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations." In *Proceedings of the IEEE Symposium on Visual Languages*, 336-343. Los Alamitos, CA: IEEE Computer Society. [Wensi Wang and Jagadish Shrivani Kumar] [Presentation](#)
- Wehrend, Stephen C., and Clayton Lewis. 1990. "A Problem-Oriented Classification of Visualization Techniques." In *Proceedings of the 1st Conference on Visualization*, 139-143. Los Alamitos, CA: IEEE Computer Society. [Sowmya Achanta and Jae Eun Kum] [Presentation](#)
- Wilkinson, Leland. 2005. *The Grammar of Graphics*. New York: Springer. [Trevor Edelblute] [Presentation](#)
- Yau, Nathan. 2011. *Visualize This: The FlowingData Guide to Design, Visualization, and Statistics*. Indianapolis, IN: Wiley. (ebrary account required to read/download)

Suggestions of other major works are welcome!

We will add Kosslyn, 1989; Kindlmann & Scheidegger, 2014

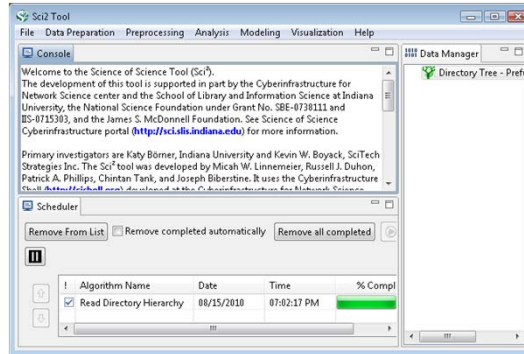


Sci2 Tool – OSGi/CIShell-based Macroscope

Download for free at <http://sci2.cns.iu.edu>

Use

- **Menu** to read data, run algorithms.
- **Console** to see work log, references to seminal works.
- **Data Manager** to select, view, save loaded, simulated, or derived datasets.
- **Scheduler** to see status of algorithm execution.



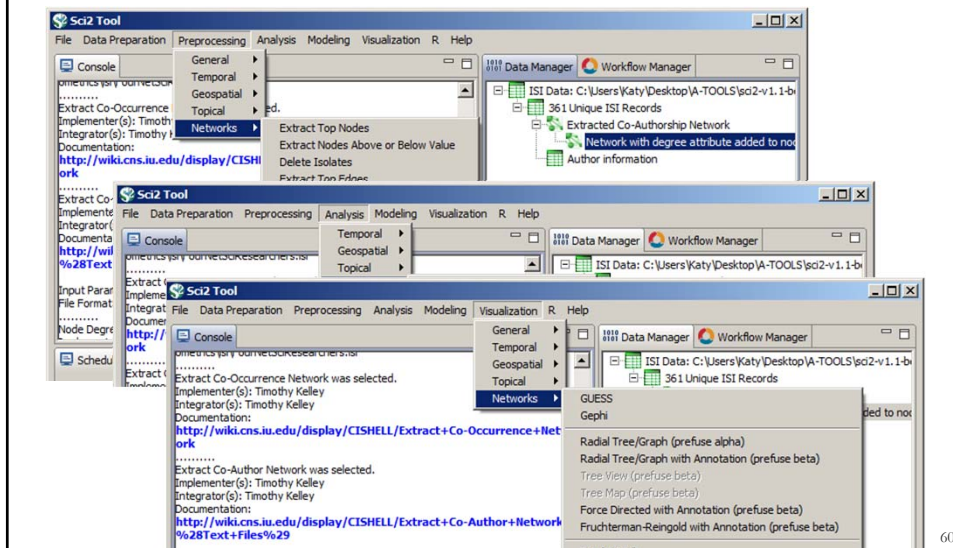
All workflows are recorded into a log file (see /sci2/logs/...), and can be re-run for easy replication. If errors occur, they are saved in a error log to ease bug reporting.

All algorithms are documented online; workflows are given in Sci2 Manual at <http://sci2.wiki.cns.iu.edu>

59

Sci2 Tool Interface Components

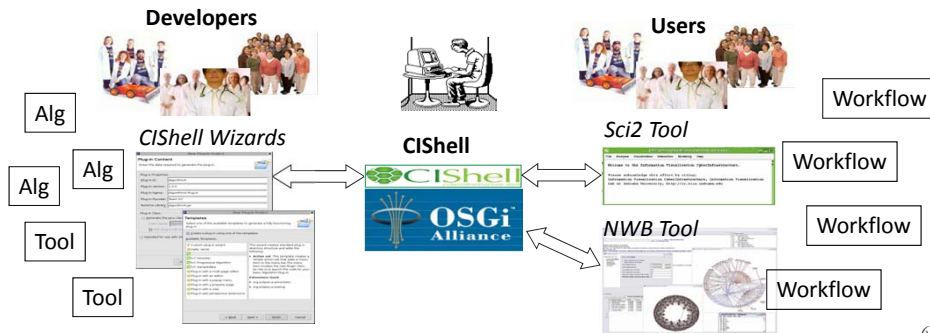
Download for free at <http://sci2.cns.iu.edu>



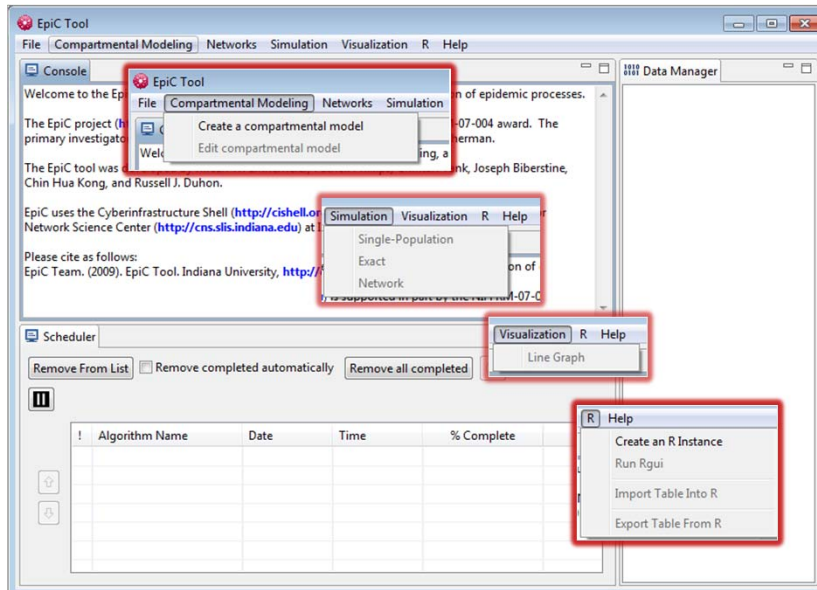
60

OSGi & CShell Support the Plug-and-Play of Algorithms

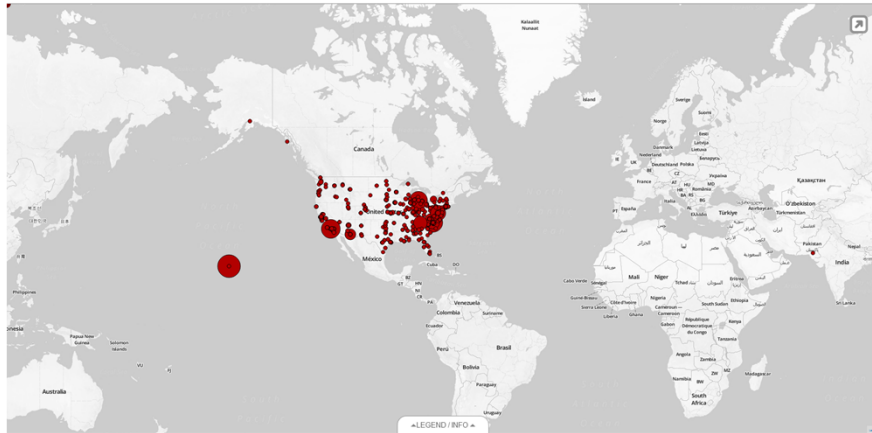
- CShell (<http://cishell.org>) is an open source software specification for the integration and utilization of datasets, algorithms, and tools.
- It extends the Open Services Gateway Initiative (OSGi) (<http://osgi.org>), a standardized, component oriented, computing environment for networked services widely used in industry since more than 10 years.
- Specifically, CShell provides “sockets” into which existing and new datasets, algorithms, and tools can be plugged using a wizard-driven process.



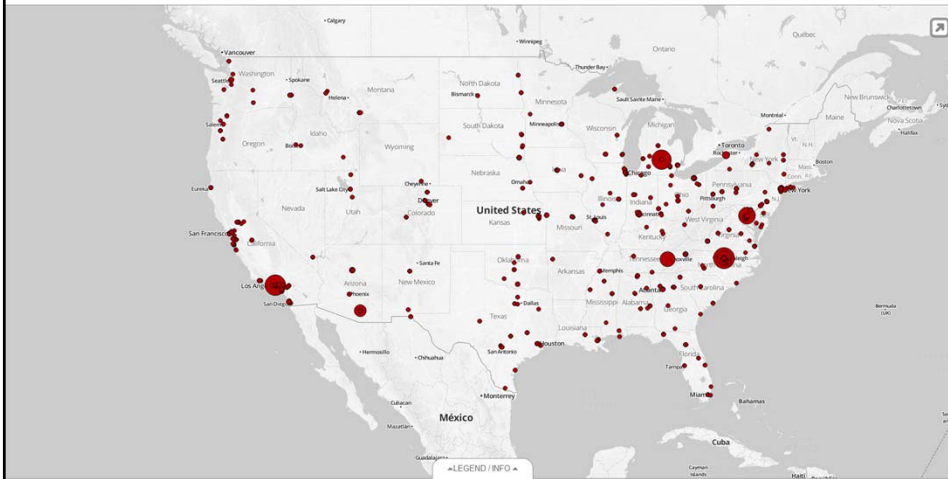
61

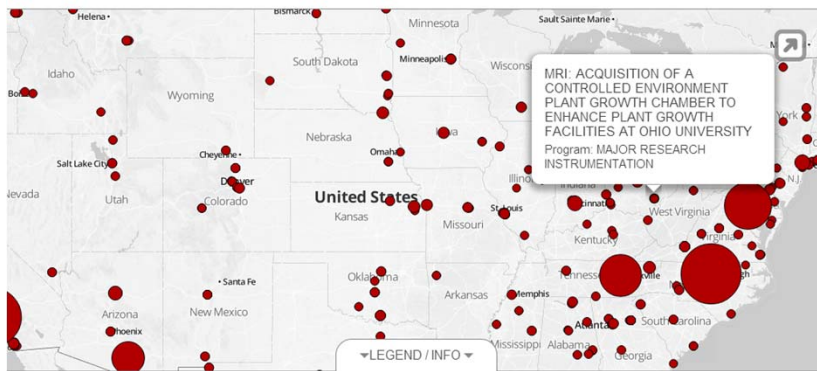
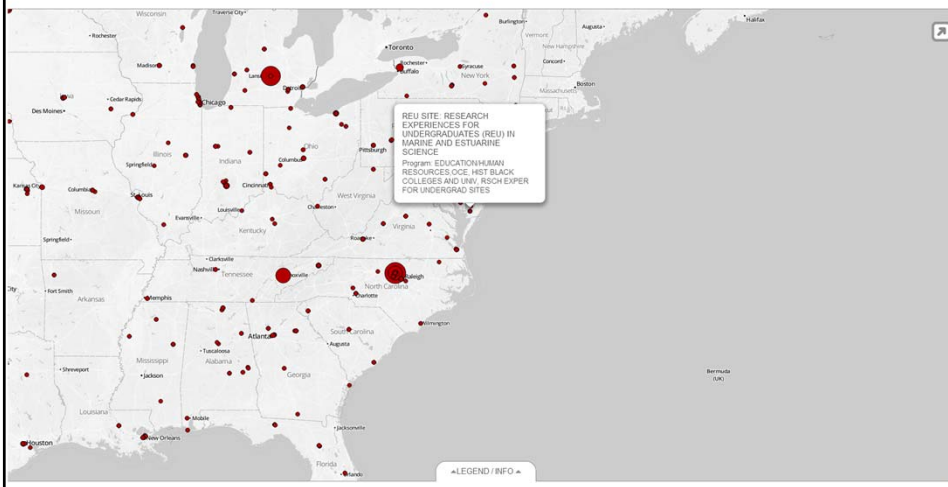


62



Developed for NIH by CNS and NETE. Responsive design.



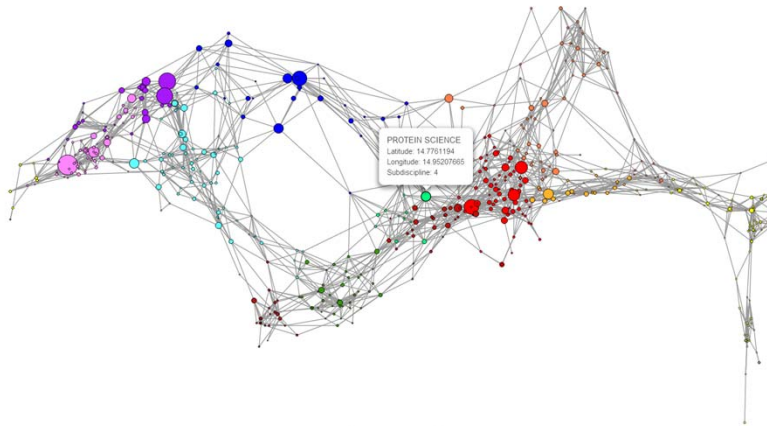


Amount Awarded

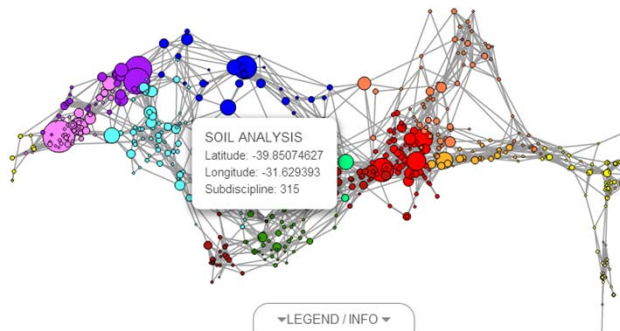


How To Read This Map

This proportional symbol map shows 52 U.S. states and other jurisdictions using the Albers equal-area conic projection with Alaska, Puerto Rico, and Hawaii inset. Each dataset record is represented by a circle centered at its geolocation. The area, interior color, and exterior color of each circle may represent numeric attribute values. Minimum and maximum data values are given in the legend.



LEGEND / INFO



LEGEND / INFO

Total Awards



How To Read This Map

This map is a visual representation of 554 sub-disciplines within 13 disciplines of science and their relationships to one another, shown as points and lines connecting those points respectively. Over top this visualization is drawn the result of mapping a dataset's journals to the underlying sub-discipline(s) those journals contain. Mapped sub-disciplines are shown with size relative to the number of matching journals and color from the discipline.

Course Schedule

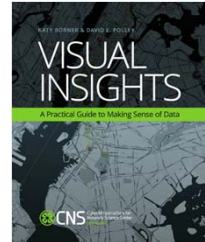
Part 1: Theory and Hands-On

- **Session 1** – Workflow Design and Visualization Framework
- **Session 2** – “When:” Temporal Data
- **Session 3** – “Where:” Geospatial Data
- **Session 4** – “What:” Topical Data

Mid-Term

- **Session 5** – “With Whom:” Trees
- **Session 6** – “With Whom:” Networks
- **Session 7** – Dynamic Visualizations and Deployment

Final Exam







Part 2: Students work in teams on client projects.

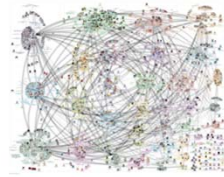
Final grade is based on Class Participation (10%), Midterm (30%), Final Exam (30%), and Client Project(30%).

69

Last 7 Weeks: Students Work in Teams With Clients

Information Visualization MOOC 2015  INDIANA UNIVERSITY  CNS  

Client Projects



AIDS as a Global Media Event

Client Name: Vladimir Cajkovic

Project Description (goal/scientific or practical value):
AIDS has radically transformed the world and become the focus of interdisciplinary study and research from a medical, cultural, and media-historical perspective. Over the past 30 years, the German Hygiene Museum in Dresden has collected numerous items –predominantly posters– which have been used in the media campaign to combat the epidemic. It is the world’s largest collection of AIDS posters with over 9,000 specimens from 147 countries.

The goal of the project is to visualize the distribution of symbols, gestures, and topics addressed in the posters through space and time so that other researchers and members of the public can understand the development of the cultural response to the AIDS epidemic.

<http://ivmooc.cns.iu.edu/clients.html>

Results of 2013 client projects can be found in the *Visual Insights* book [here](#).

Chapter 8 – Case Studies 235

Understanding the Diffusion of Non-Emergency Call Systems 236

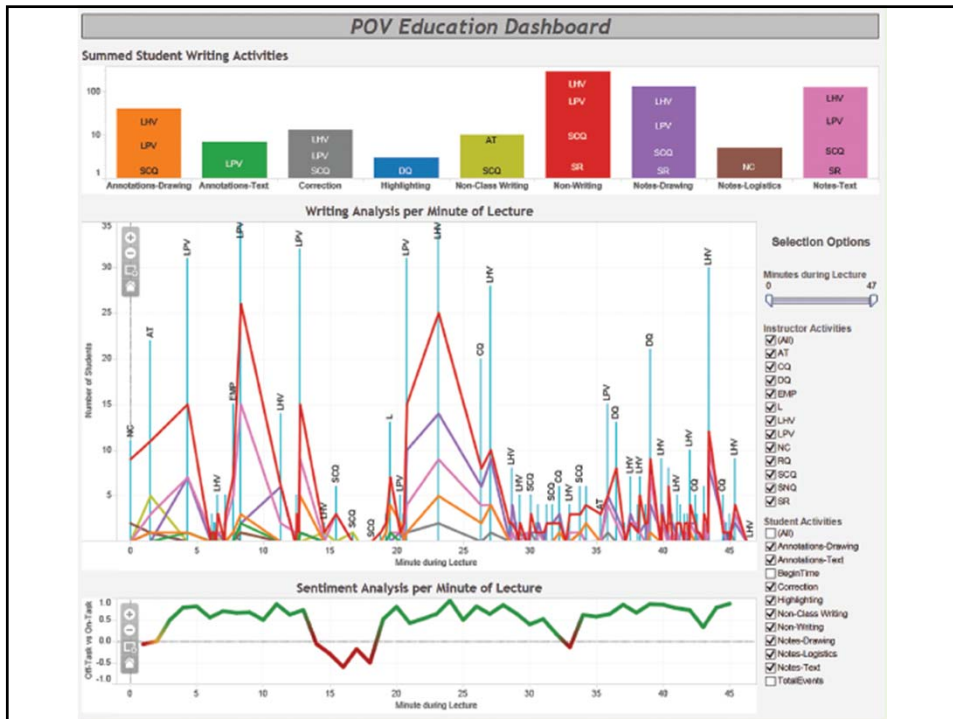
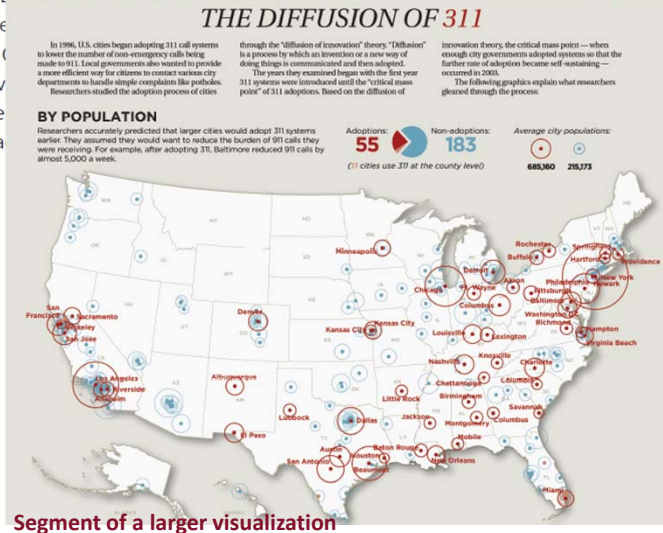
Examining the Success of 311 Call Systems 237

Using Point of View (POV) to Analyze Data 238

Phylot: An Interactive Phylogenetic Tree 239

Isis: Mapping the Geographical Spread of a Disease 240

Visualizing the Impact of a Natural Disaster 241



One Hundred Years of *Isis*

David E. Hubbard (Texas A&M University), Anouk Lang (University of Strathclyde), Kathleen Reed (Vancouver Island University),
Anelise Hanson ShROUT (Davidson College) and Lyndsay D. Troyer (Colorado State University)

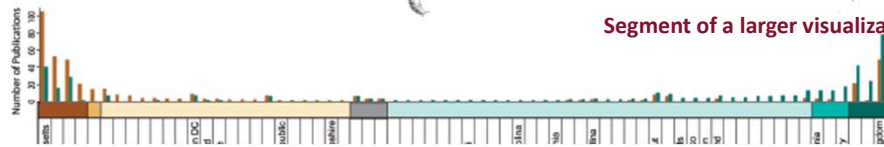
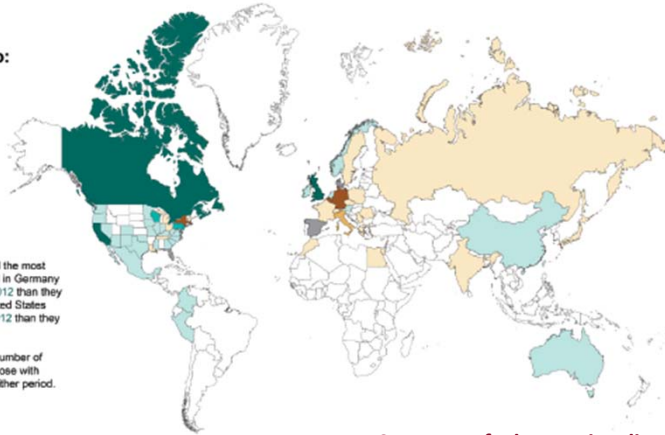
Changes in *Isis* Authorship: 1913-1937 to 1988-2012.

This map and chart show geographic shifts in *Isis* authors' locations across two time periods of 25 years each: 1913-1937 and 1988-2012.

Countries and states tinted brown had a greater number of articles published in the 1913-1937 period, while those tinted green had a greater number published in the 1988-2012 period.

Germany and the United States experienced the most extreme shifts in authorship. Authors located in Germany published 37 fewer articles between 1988-2012 than they did in 1913-1937. Authors located in the United States published 31 more articles between 1988-2012 than they did in 1913-1937.

Countries and states in gray had the same number of published articles in both periods, while those with no color (i.e., white) had no publications in either period.



Segment of a larger visualization

References

Börner, Katy, Chen, Chaomei, and Boyack, Kevin. (2003). **Visualizing Knowledge Domains**. In Blaise Cronin (Ed.), *ARIST*, Medford, NJ: Information Today, Volume 37, Chapter 5, pp. 179-255. <http://ivl.slis.indiana.edu/km/pub/2003-borner-arist.pdf>

Shiffrin, Richard M. and Börner, Katy (Eds.) (2004). **Mapping Knowledge Domains**. *Proceedings of the National Academy of Sciences of the United States of America*, 101(Suppl_1). http://www.pnas.org/content/vo1101/suppl_1/

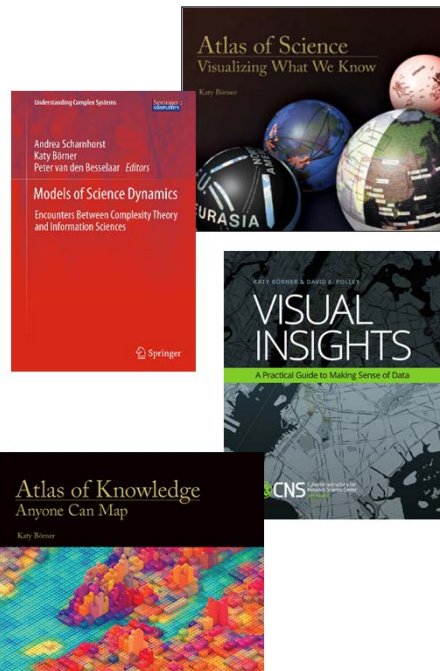
Börner, Katy (2010) **Atlas of Science: Visualizing What We Know**. The MIT Press. <http://scimaps.org/atlas>

Scharnhorst, Andrea, Börner, Katy, van den Besselaar, Peter (2012) **Models of Science Dynamics**. Springer Verlag.

Katy Börner, Michael Conlon, Jon Corson-Rikert, Cornell, Ying Ding (2012) **VIVO: A Semantic Approach to Scholarly Networking and Discovery**. Morgan & Claypool.

Katy Börner and David E Polley (2014) **Visual Insights: A Practical Guide to Making Sense of Data**. The MIT Press.

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



All papers, maps, tools, talks, press are linked from <http://cns.iu.edu>
 These slides will soon be at <http://cns.iu.edu/docs/presentations>
 CNS Facebook: <http://www.facebook.com/cnscenter>
 Mapping Science Exhibit Facebook: <http://www.facebook.com/mappingscience>