Insightful Data Visualizations

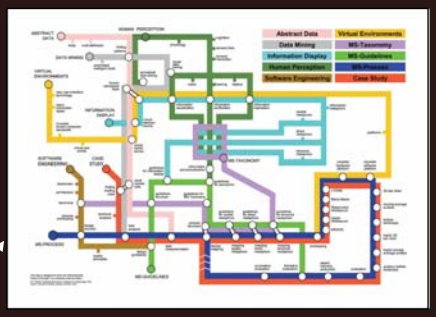
Katy Börner, CNS, LIS, SoIC, Indiana University, Bloomington, Indiana
katy@indiana.edu | http://cns.iu.edu

Data Visualization Guest Lecture
School of Informatics and Computing, Indiana University, Bloomington, IN

September 17, 2013

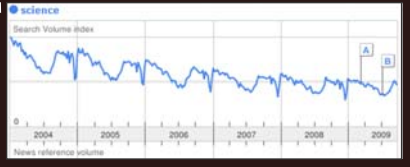
Terra bytes of data

Descriptive & Predictive Models

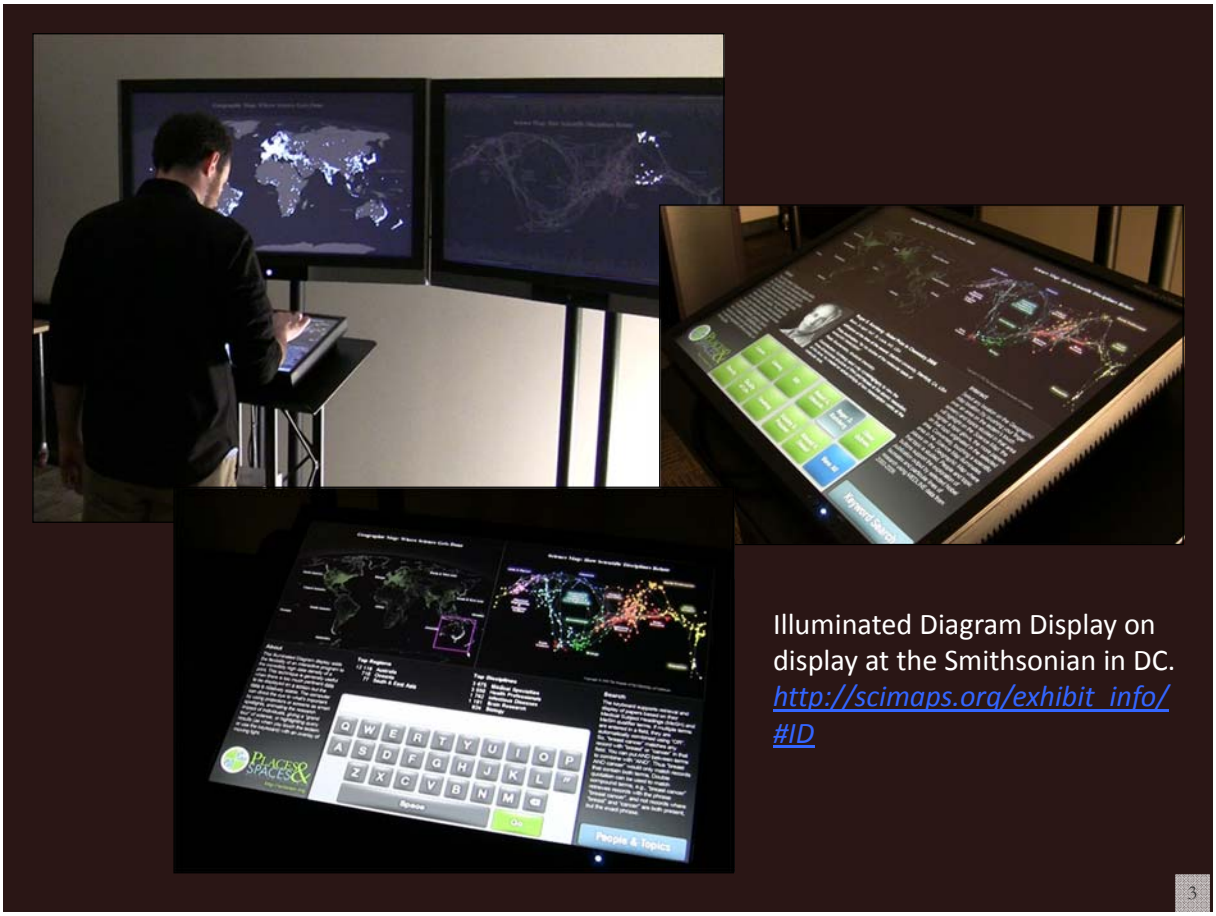


Find your way

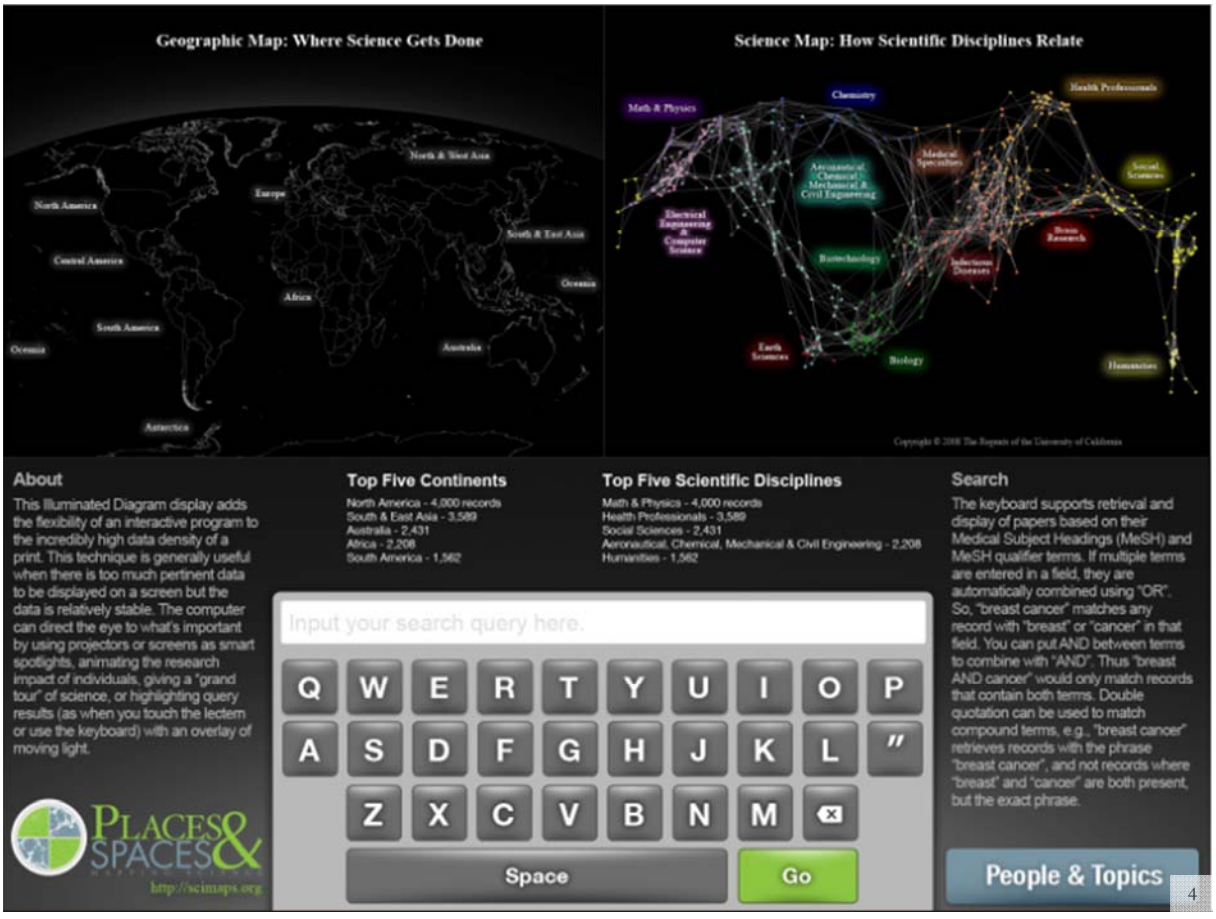
Find collaborators, friends

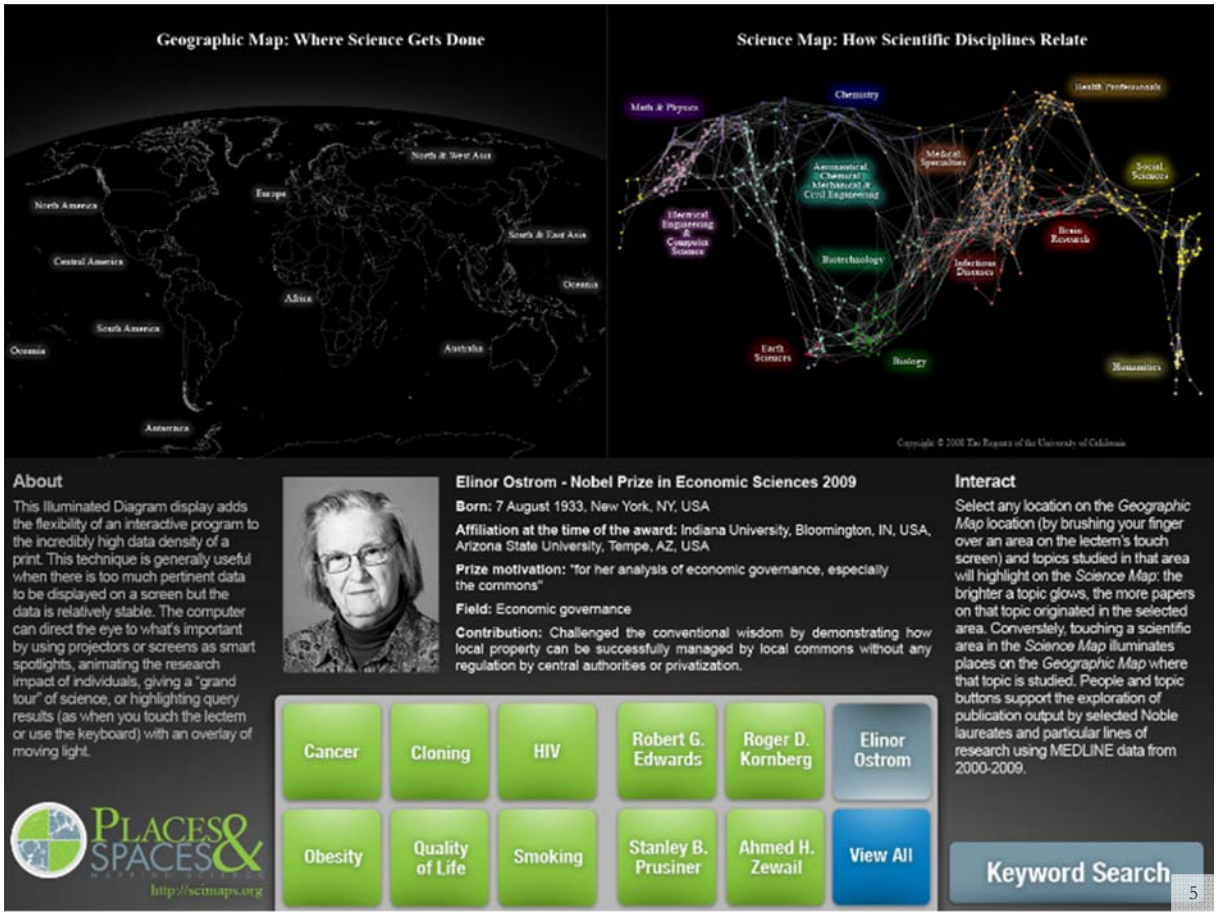


Identify trends

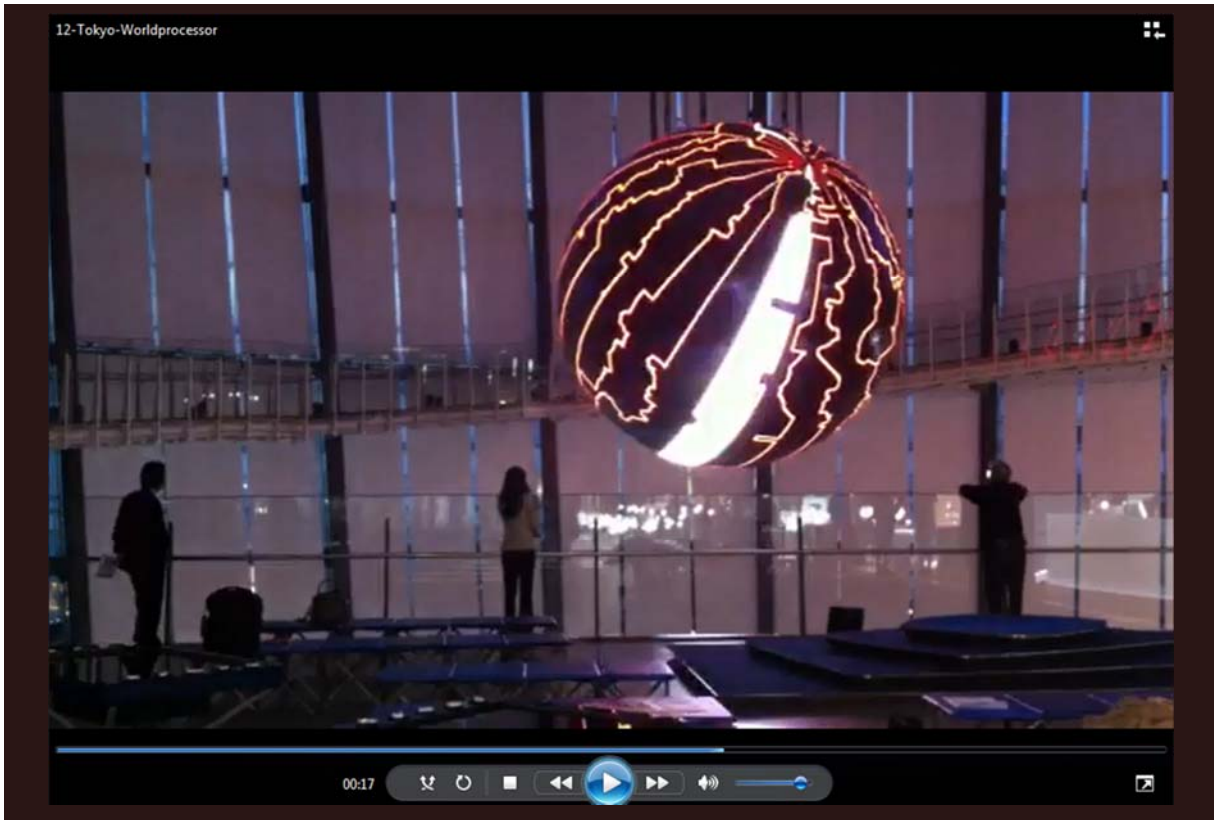


Illuminated Diagram Display on display at the Smithsonian in DC. http://scimaps.org/exhibit_info/#ID





5

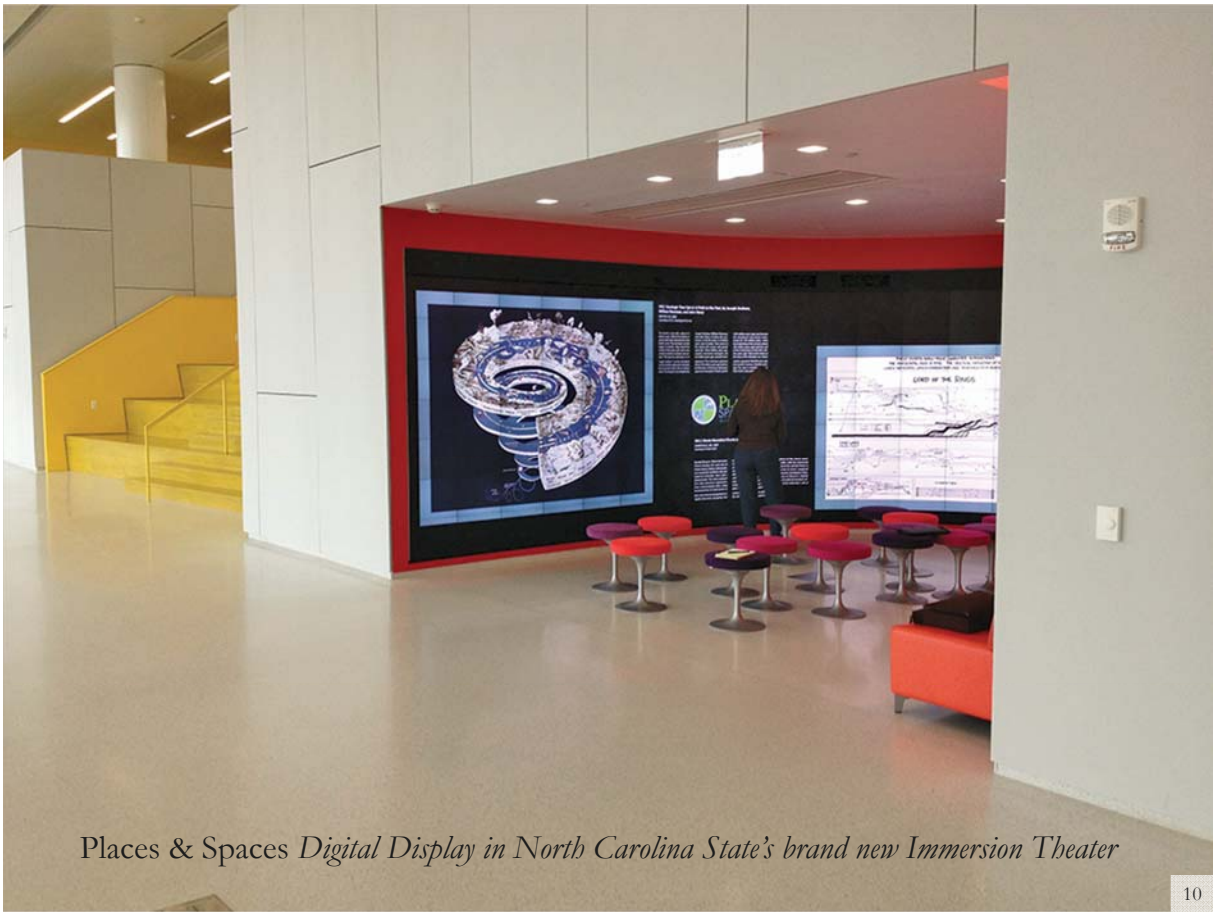


Ingo Gunther's Worldprocessor globe design on display at the Giant Geo Cosmos OLED Display at the Museum of Emerging Science and Innovation in Tokyo, Japan

6

Science Maps in “Expedition Zukunft” science train visiting 62 cities in 7 months 12 coaches, 300 m long Opening was on April 23rd, 2009 by German Chancellor Merkel
http://www.expedition-zukunft.de

Debut of 5th Iteration of Mapping Science Exhibit at MEDIA X was on May 18, 2009 at Wallenberg Hall, Stanford University, http://mediax.stanford.edu, http://scaleindependentthought.typepad.com/photos/scimaps



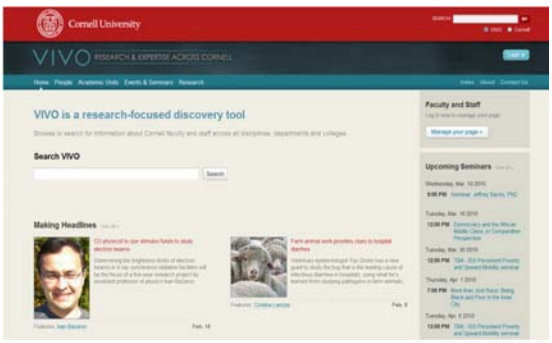
Places & Spaces *Digital Display in North Carolina State's brand new Immersion Theater*

VIVO International Researcher Network

VIVO
Enabling a National Network of Scientists

VIVO: A Semantic Approach to Creating a National Network of Researchers (http://vivoweb.org)

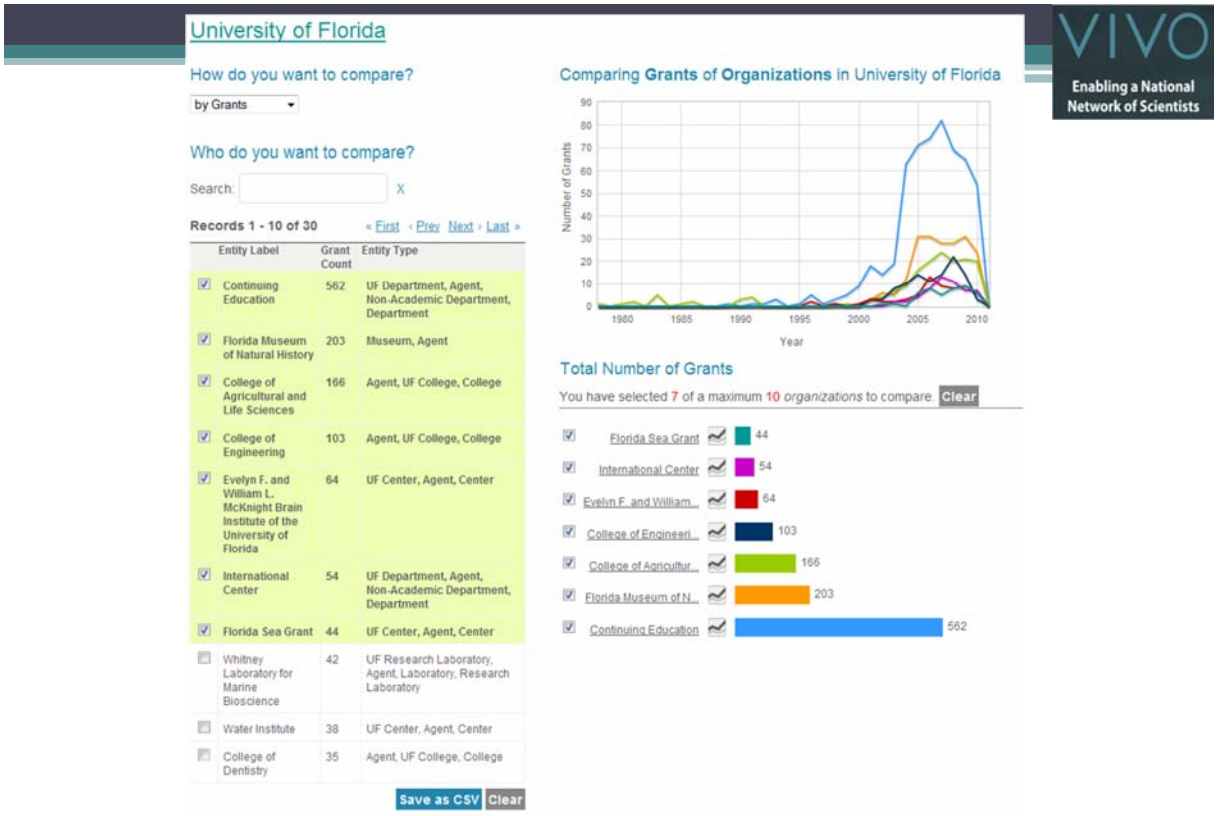
- Semantic web application and ontology editor originally developed at Cornell U.
- Integrates research and scholarship info from systems of record across institution(s).
- Facilitates research discovery and cross-disciplinary collaboration.
- Simplify reporting tasks, e.g., generate biosketch, department report.



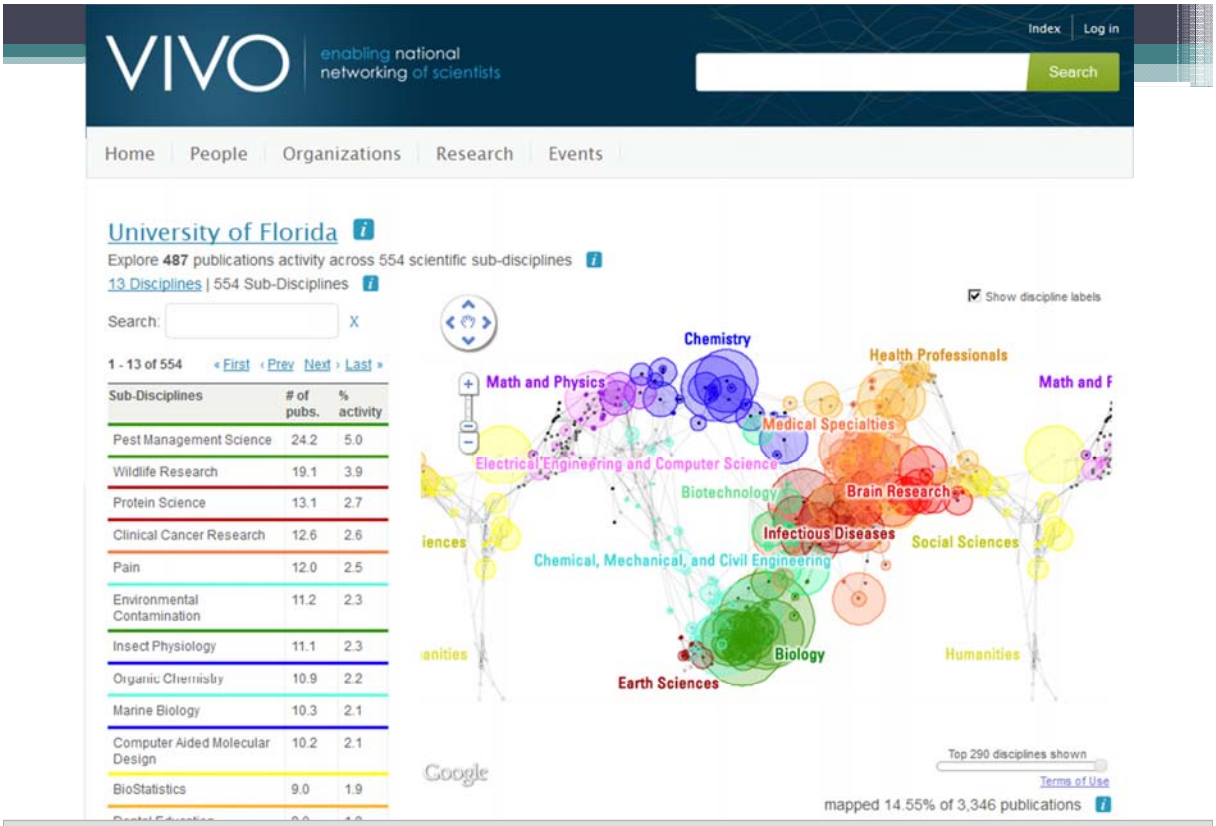
Funded by $12 million NIH award.

Cornell University: Dean Krafft (Cornell PI), Manolo Bevia, Jim Blake, Nick Cappadona, Brian Caruso, Jon Corson-Rikert, Elly Cramer, Medha Devare, John Fereira, Brian Lowe, Stella Mitchell, Holly Mistlebauer, Anup Sawant, Christopher Westling, Rebecca Younes. **University of Florida:** Mike Conlon (VIVO and UF PI), Cecilia Botero, Kerry Britt, Erin Brooks, Amy Buhler, Ellie Bushhousen, Chris Case, Valrie Davis, Nita Ferree, Chris Haines, Rae Jesano, Margeaux Johnson, Sara Kreinest, Yang Li, Paula Markes, Sara Russell Gonzalez, Alexander Rockwell, Nancy Schaefer, Michele R. Tennant, George Hack, Chris Barnes, Narayan Raum, Brenda Stevens, Alicia Turner, Stephen Williams. **Indiana University**: Katy Borner (IU PI), William Barnett, Shanshan Chen, Ying Ding, Russell Duhon, Jon Dunn, Micah Linnemeier, Nianli Ma, Robert McDonald, Barbara Ann O'Leary, Mark Price, Yuyin Sun, Alan Walsh, Brian Wheeler, Angela Zoss. **Ponce School of Medicine:** Richard Noel (Ponce PI), Ricardo Espada, Damaris Torres. **The Scripps Research Institute:** Gerald Joyce (Scripps PI), Greg Dunlap, Catherine Dunn, Brant Kelley, Paula King, Angela Murrell, Barbara Noble, Cary Thomas, Michaeleen Trimarchi. **Washington University, St. Louis**: Rakesh Nagarajan (WUSTL PI), Kristi L. Holmes, Sunita B. Koul, Leslie D. McIntosh. **Weill Cornell Medical College:** Curtis Cole (Weill PI), Paul Albert, Victor Brodsky, Adam Cheriff, Oscar Cruz, Dan Dickinson, Chris Huang, Itay Klaz, Peter Michelini, Grace Migliorisi, John Ruffing, Jason Specland, Tru Tran, Jesse Turner, Vinay Varughese.

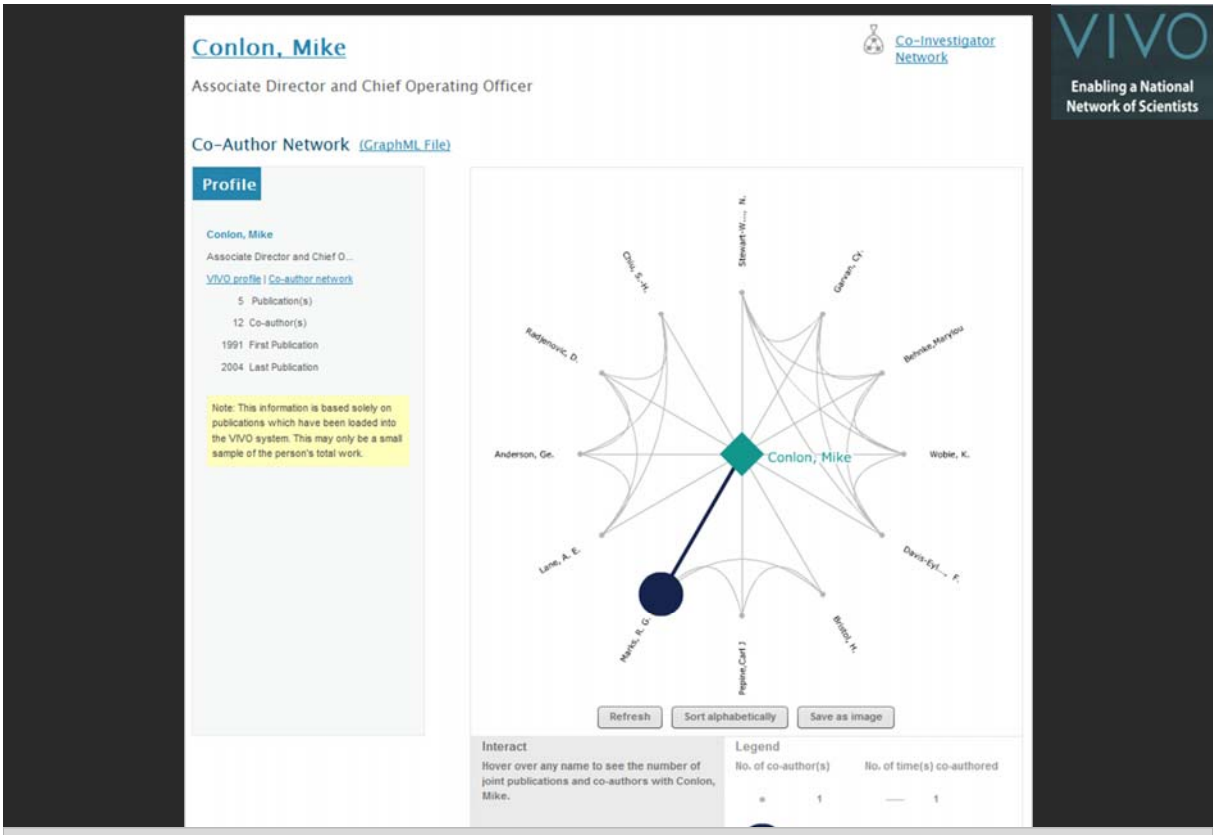
UF Clinical and Translational Science Institute
UNIVERSITY of FLORIDA



Temporal Analysis (When) Temporal visualizations of the number of papers/funding award at the institution, school, department, and people level



Topical Analysis (What) Science map overlays will show where a person, department, or university publishes most in the world of science. (in work)



Network Analysis (With Whom?) Who is co-authoring, co-investigating, co-inventing with whom? What teams are most productive in what projects?

http://nrn.cns.iu.edu

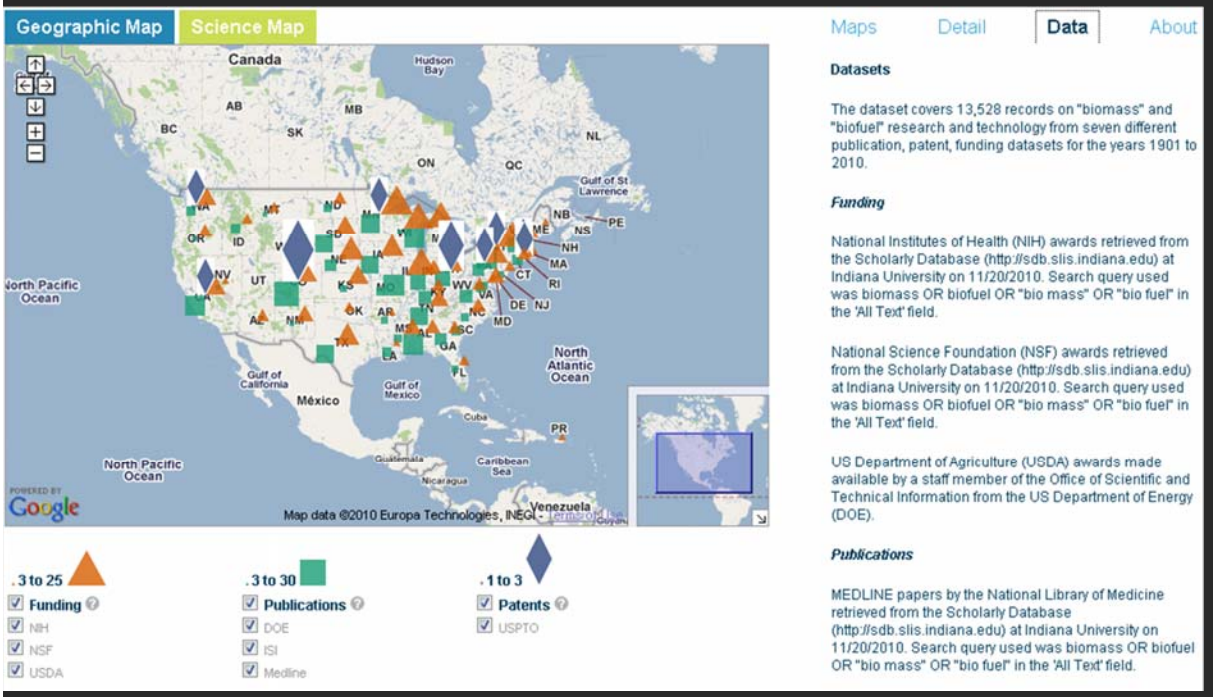
Geospatial Analysis (Where) Where is what science performed by whom? Science is global and needs to be studied globally.

VIVO
Enabling a National Network of Scientists

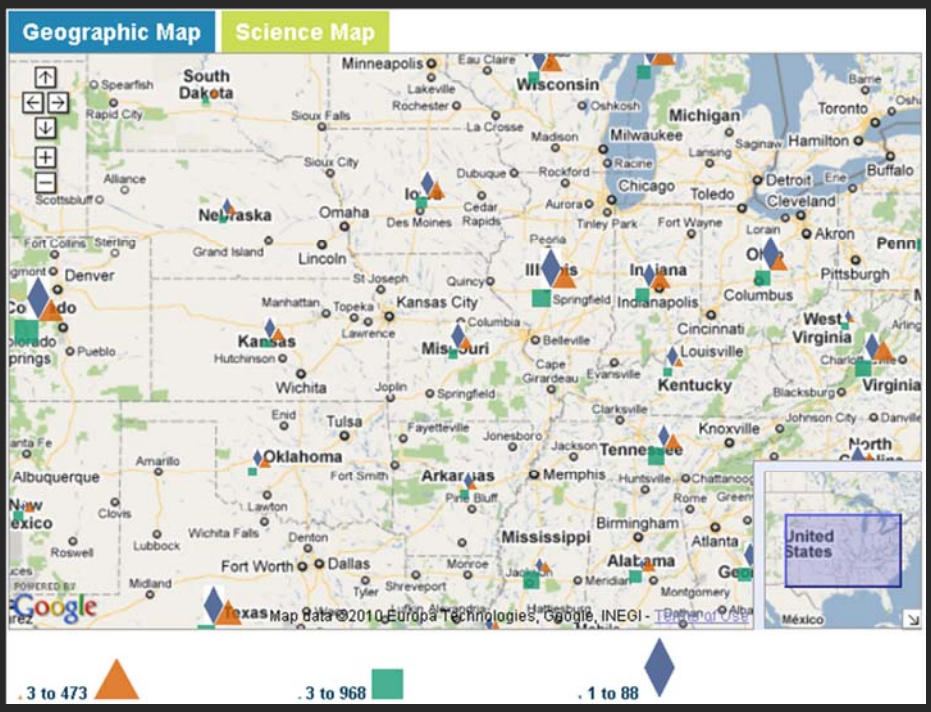
VIVO On-The-Go

Overview, Interactivity, Details on Demand
come to commonly used devices and environments

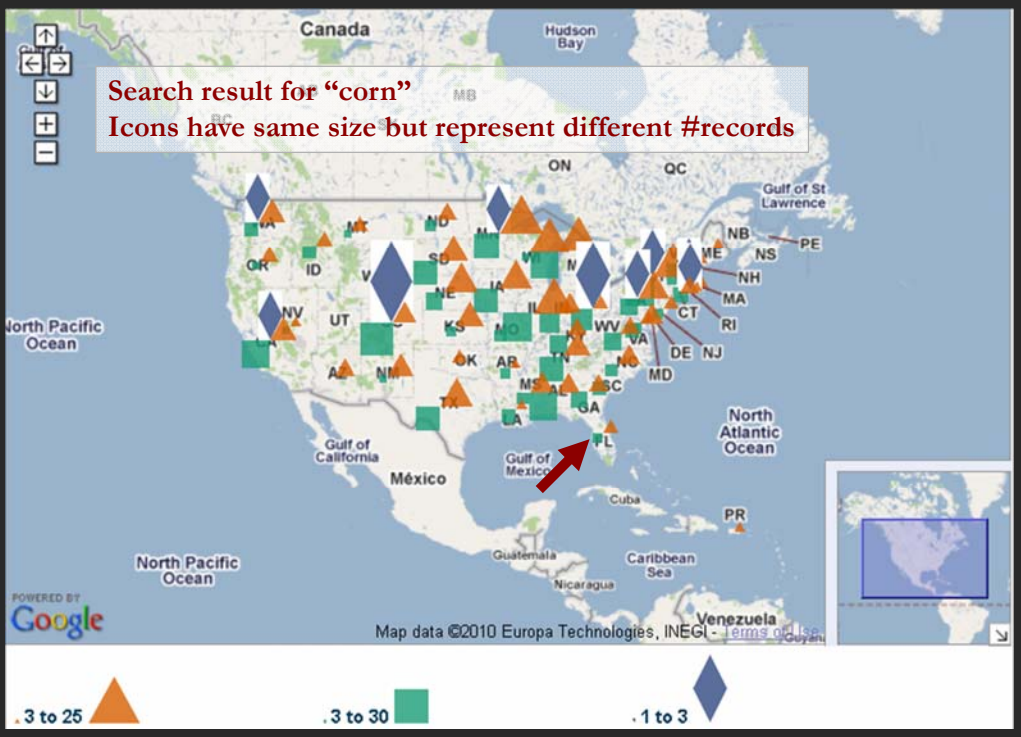
Online Interactive Maps for Sustainability Research and NIH

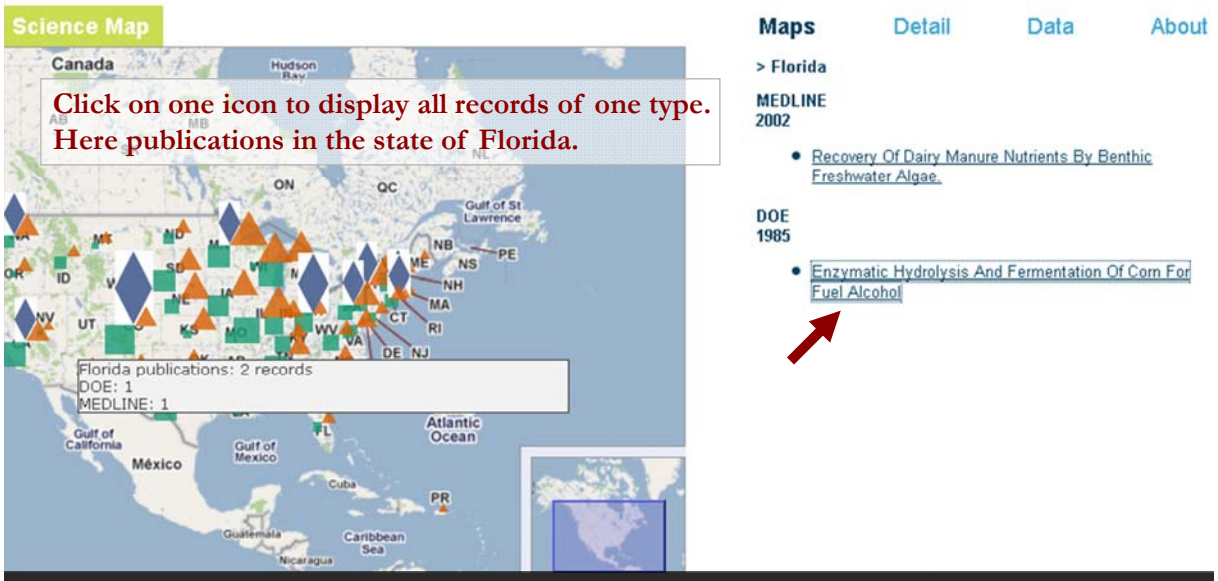


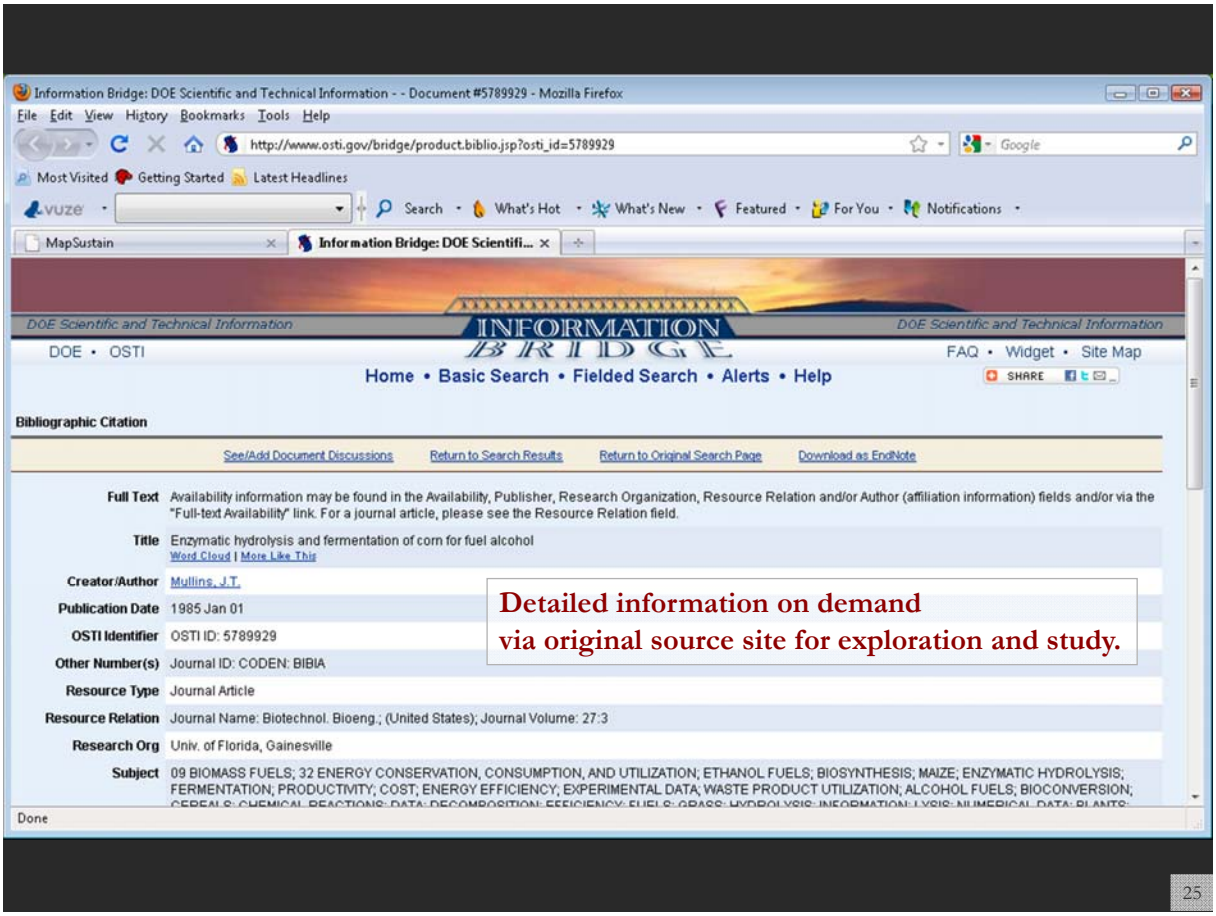
The geographic map at **state level**.

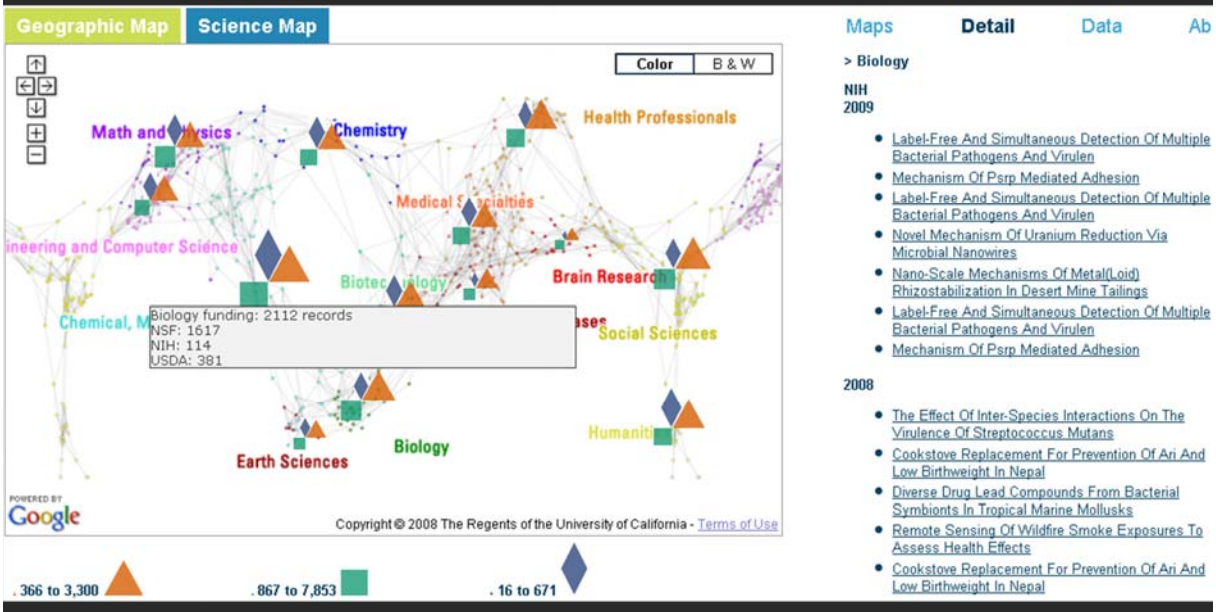


The geographic map at **city level**.

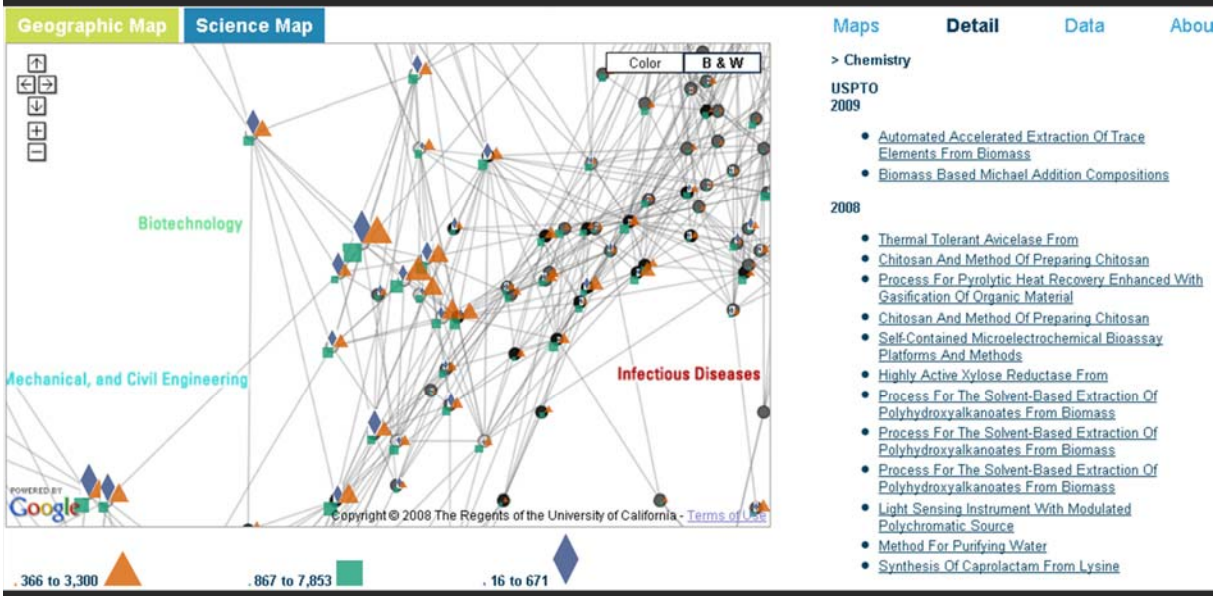

Search result for "corn"
Icons have same size but represent different #records
Canada
Hudson Bay
ON
QC
Gulf of St Lawrence
North Pacific Ocean
North Atlantic Ocean
Gulf of California
México
Gulf of Mexico
Cuba
PR
Guatemala
Caribbean Sea
Nicaragua
Venezuela
Map data ©2010 Europa Technologies, INEGI
.3 to 25
.3 to 30
.1 to 3


Science Map
Maps
Detail
Data
About
> Florida
Click on one icon to display all records of one type.
Here publications in the state of Florida.
Florida publications: 2 records
DOE: 1
MEDLINE: 1
MEDLINE
2002
Recovery Of Dairy Manure Nutrients By Benthic Freshwater Algae.
DOE
1985
Enzymatic Hydrolysis And Fermentation Of Corn For Fuel Alcohol

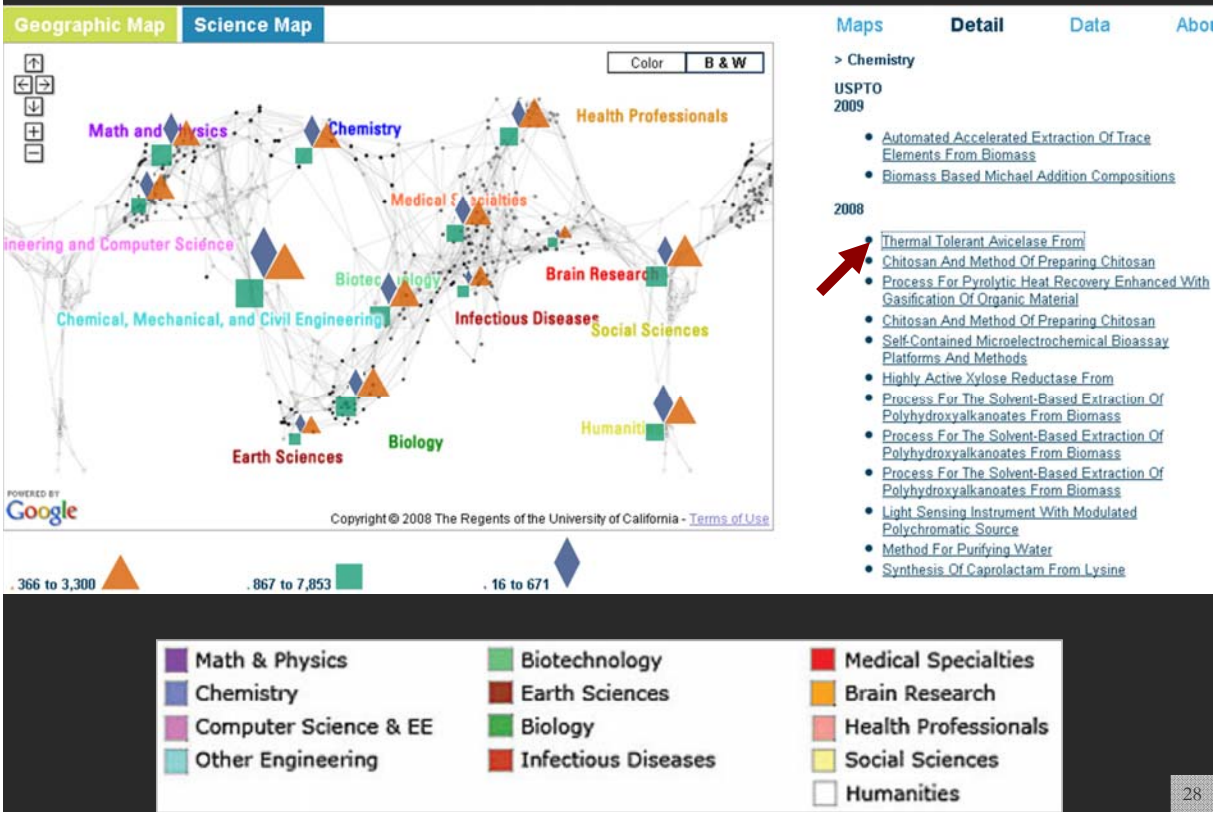


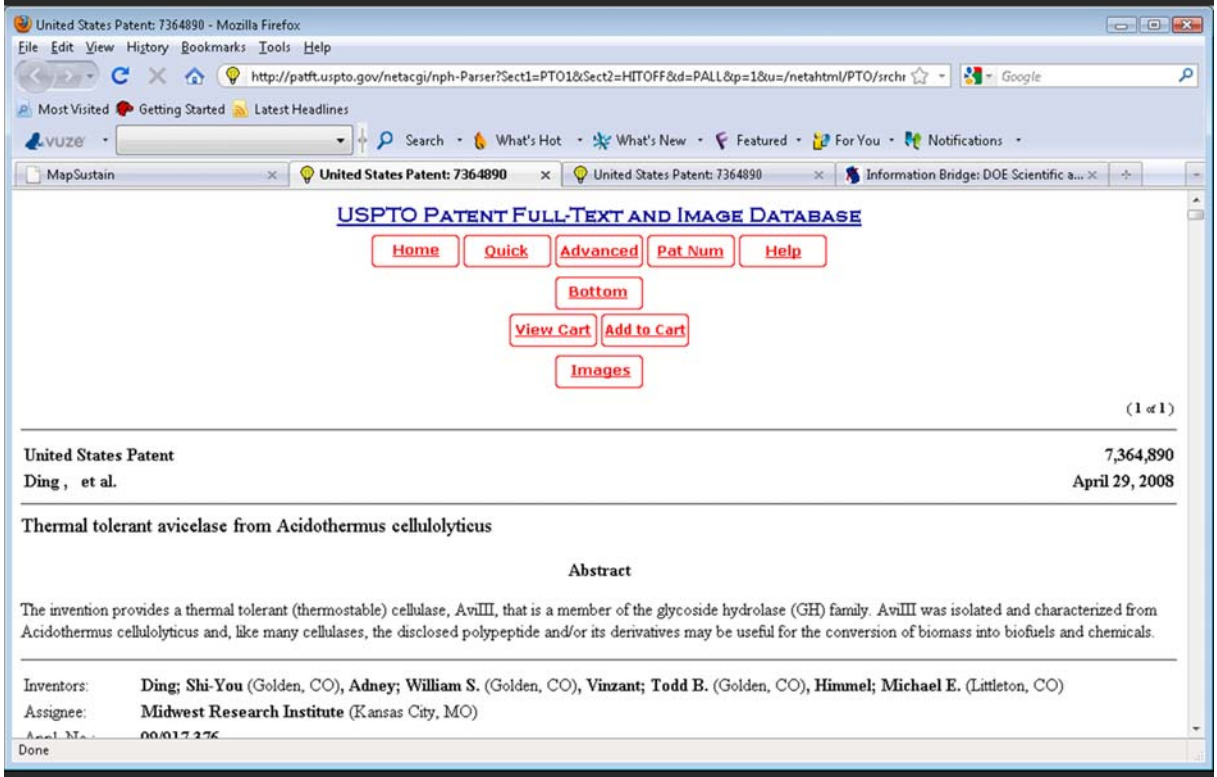


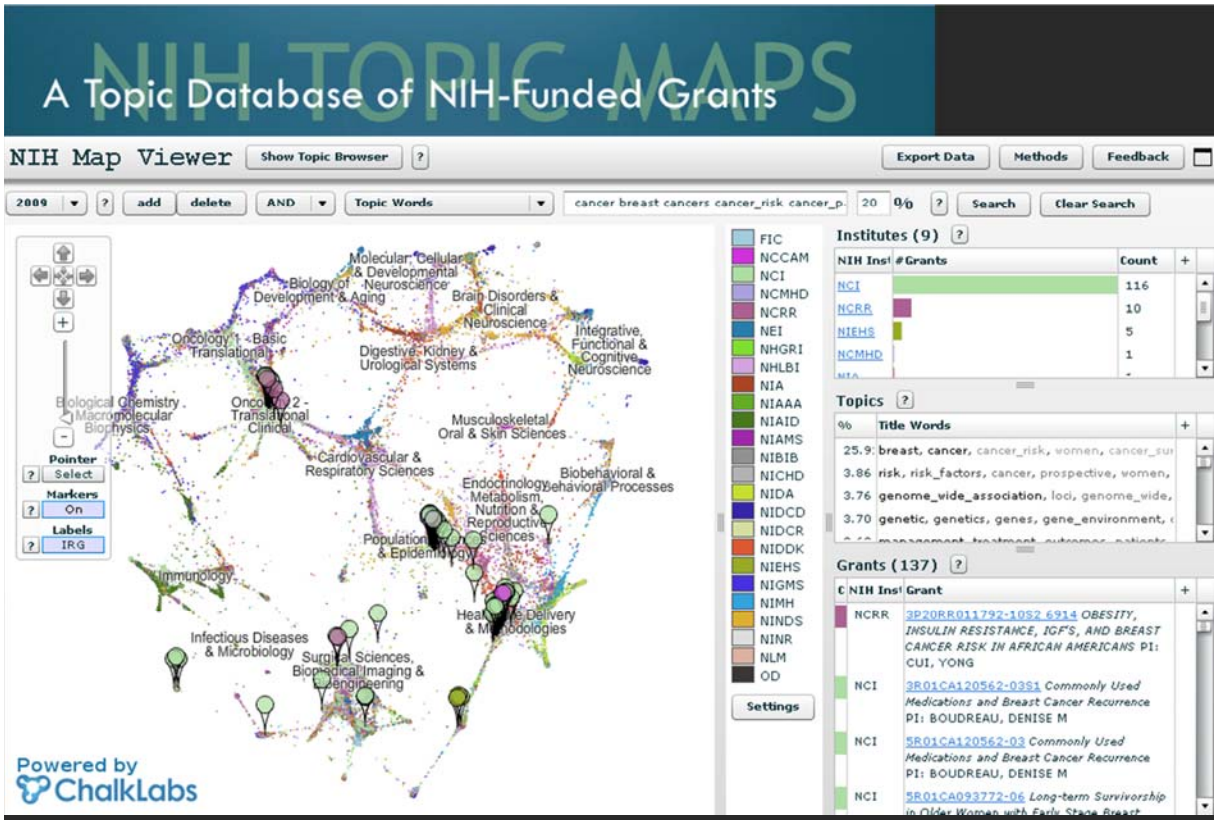
The science map at 13 top-level scientific **disciplines level**.



The science map at 554 **sub-disciplines level**.

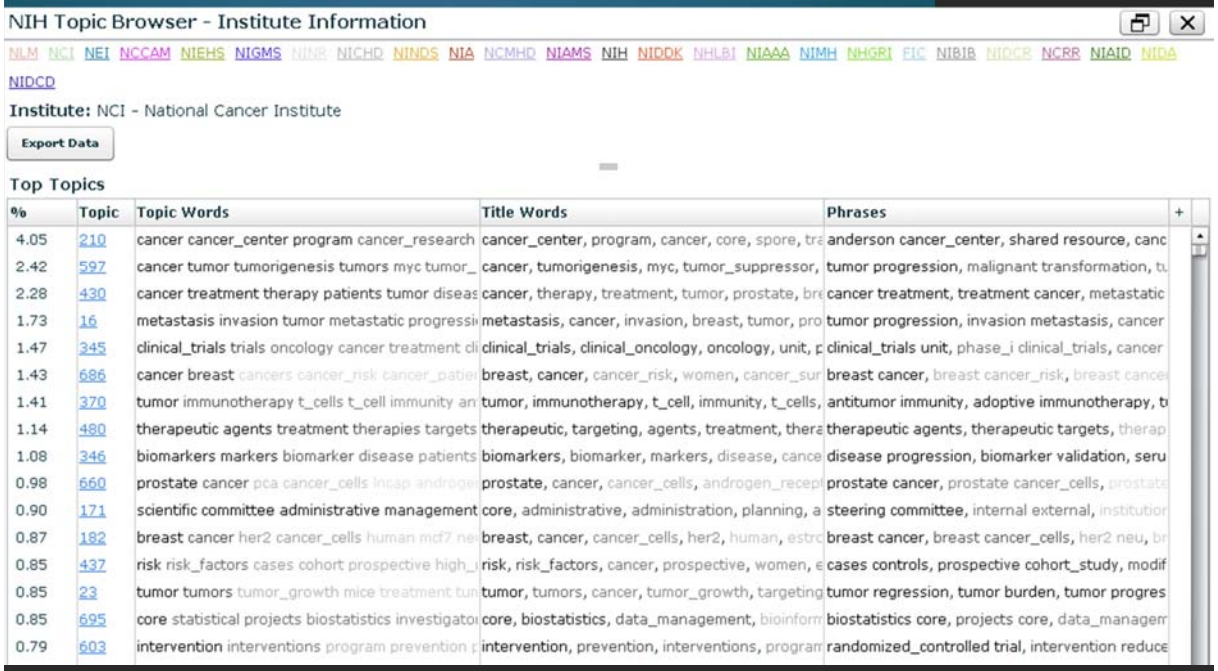






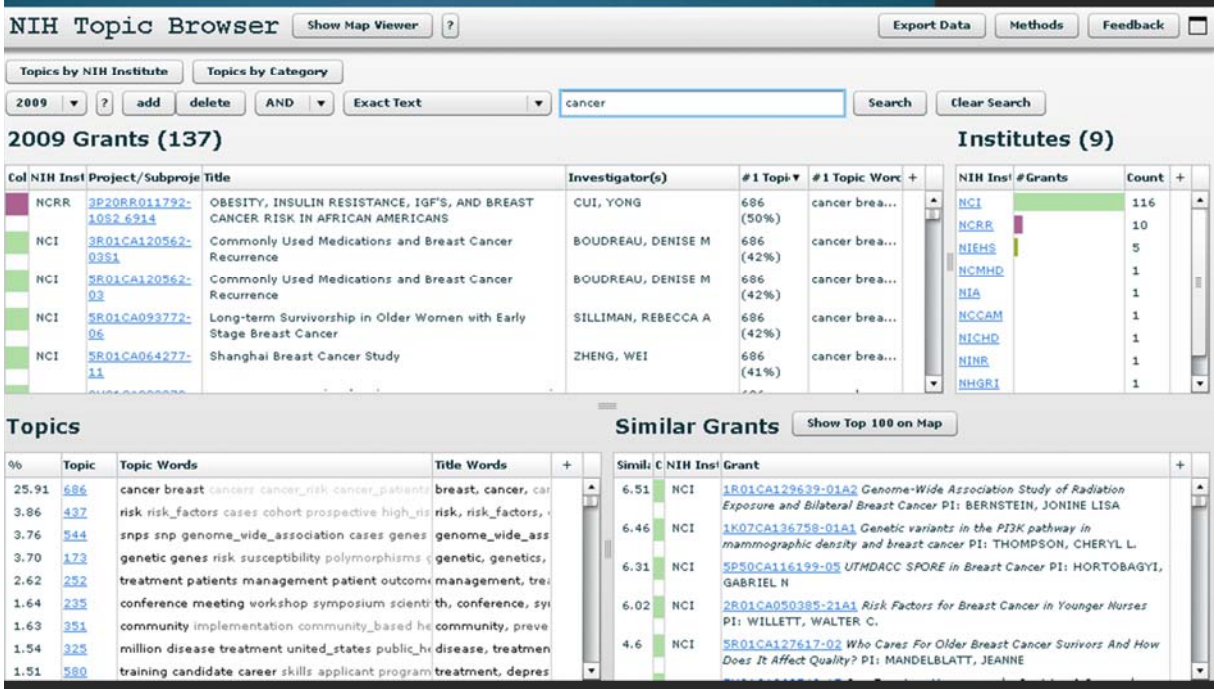
https://app.nihmaps.org

A Topic Database of NIH-Funded Grants



https://app.nihmaps.org

A Topic Database of NIH-Funded Grants

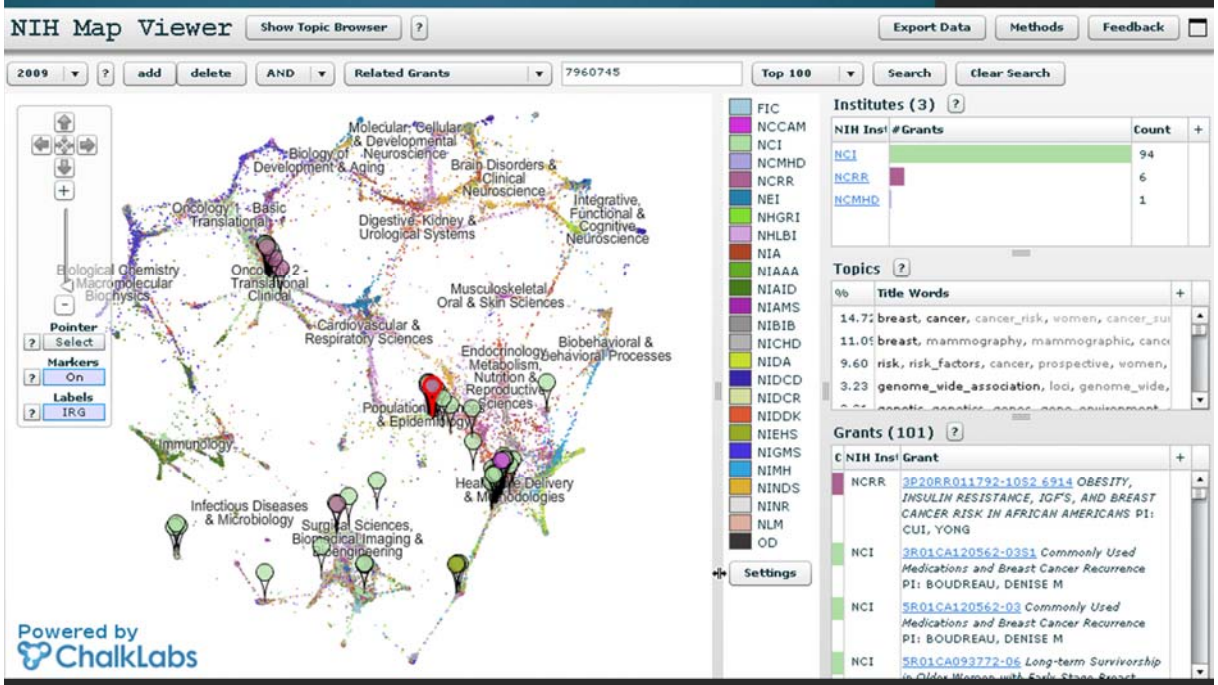


https://app.nihmaps.org

A Topic Database of NIH-Funded Grants

https://app.nihmaps.org

A Topic Database of NIH-Funded Grants

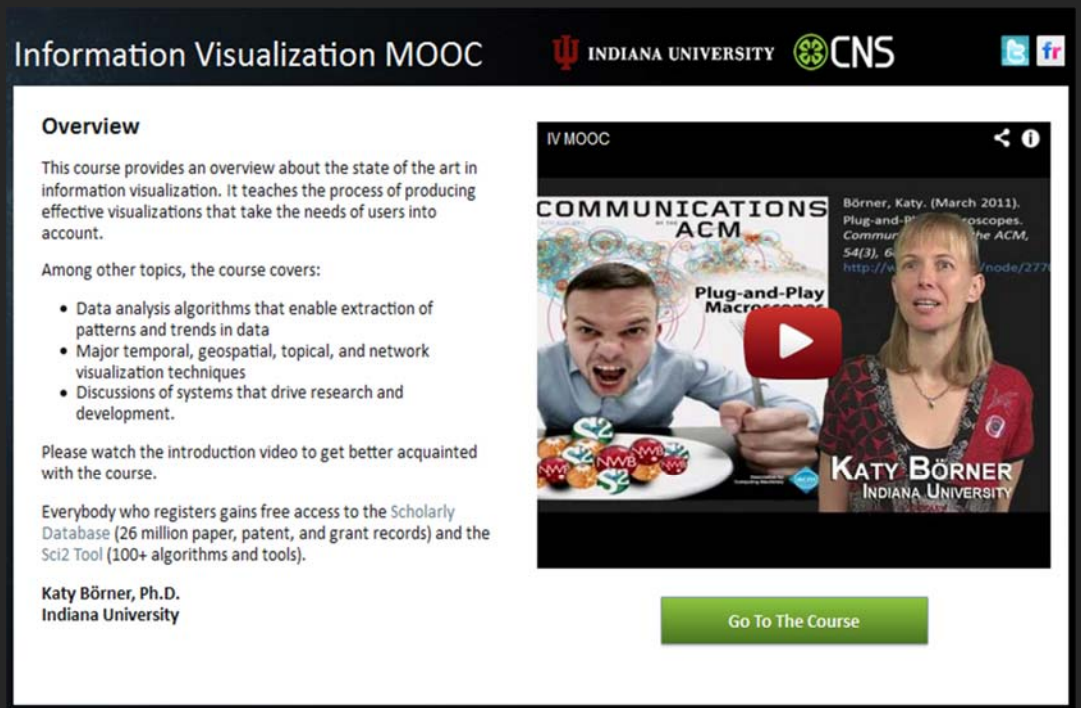


https://app.nihmaps.org

Anyone Can Cook

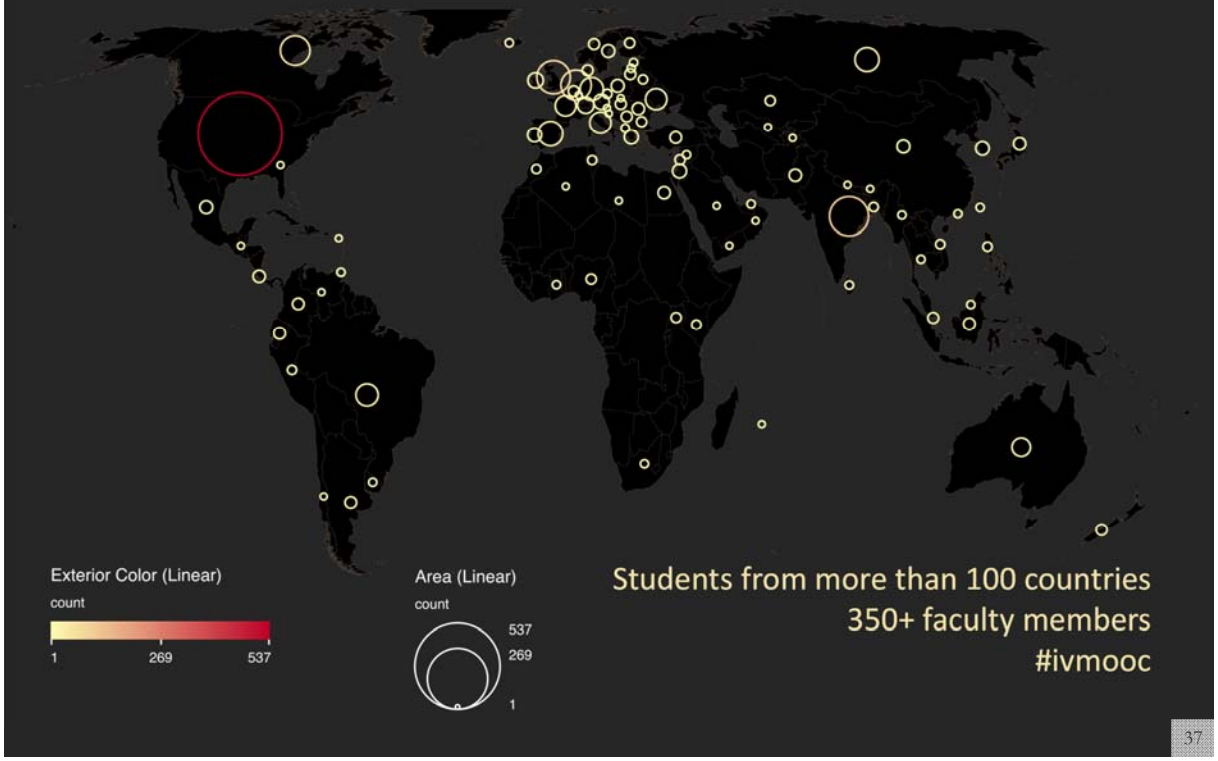
&

Anyone Can Map



ivmooc.cns.iu.edu

The Information Visualization MOOC
ivmooc.cns.iu.edu



37

Instructors

Katy Börner – Theory Parts
Instructor, Professor at SLIS

David E. Polley – Hands-on Parts
CNS Staff, Research Assistant with MIS/MLS
Teaches & Tests Sci2 Tool

Scott B. Weingart – Client Work
Assistant Instructor, SLIS PhD student

Course Schedule

- **Session 1** – Workflow design and visualization framework
- **Session 2** – "When:" Temporal Data
- **Session 3** – "Where:" Geospatial Data
- **Session 4** – "What:" Topical Data

Mid-Term

Students work in teams with clients.

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

Final Exam

Grading

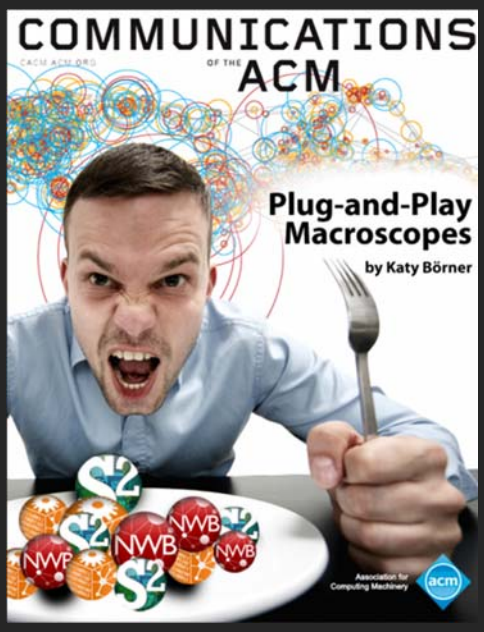
All students are asked to create a personal profile to support working in teams.

Final grade is based on Midterm (**30%**), Final (**40%**), Client Project (**30%**).

- Weekly self-assessments are not graded.
- Homework is graded automatically.
- Midterm and Final test materials from theory and hands-on sessions are graded automatically.
- Client work is peer-reviewed via online forum.

All students that receive more than **80%** of all available points get an official certificate/badge.

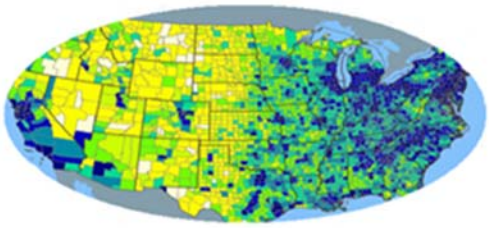
Plug-and-Play Macroscopes
cishell.org



Börner, Katy. (March 2011). Plug-and-Play Macroscopes. *Communications of the ACM, 54(3), 60-69.* http://www.scivee.tv/node/27704

Different Levels of Abstraction/Analysis

Macro/Global
Population Level



Meso/Local
Group Level

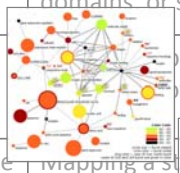
Micro/Individual
Individual Level

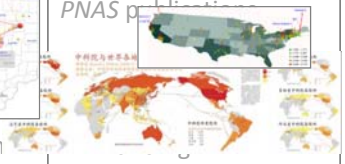
Type of Analysis vs. Level of Analysis

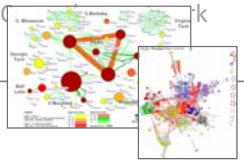
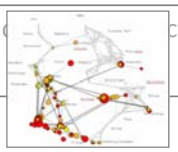
	Micro/Individual (1-100 records)	***Meso/Local (101–10,000 records)***	***Macro/Global (10,000 < records)***
Statistical Analysis/Profiling	Individual person and their expertise profiles	Larger labs, centers, universities, research domains, or states	All of NSF, all of USA, all of science.
Temporal Analysis (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 state's 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

Type of Analysis vs. Level of Analysis

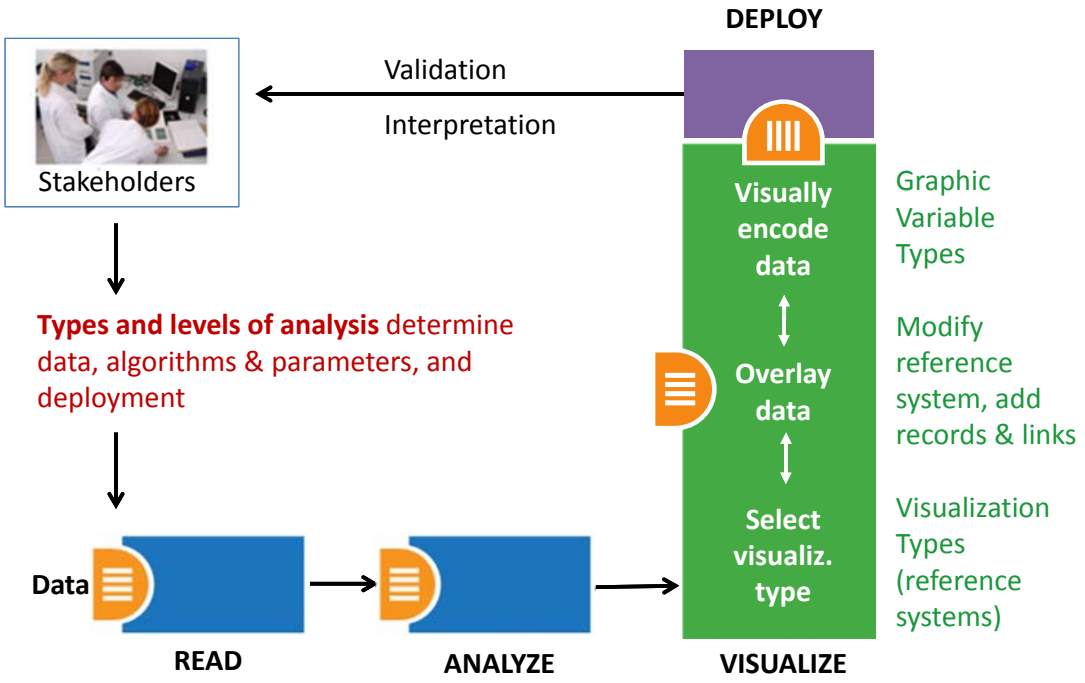
	Micro/Individual (1-100 records)	***Meso/Local (101–10,000 records)***	***Macro/Global (10,000 < records)***
Statistical Analysis/Profiling	Individual person and their expertise profiles	Larger labs, centers, universities, research domains, or states	All of NSF, all of USA, all of science.
Temporal Analysis (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 state's 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



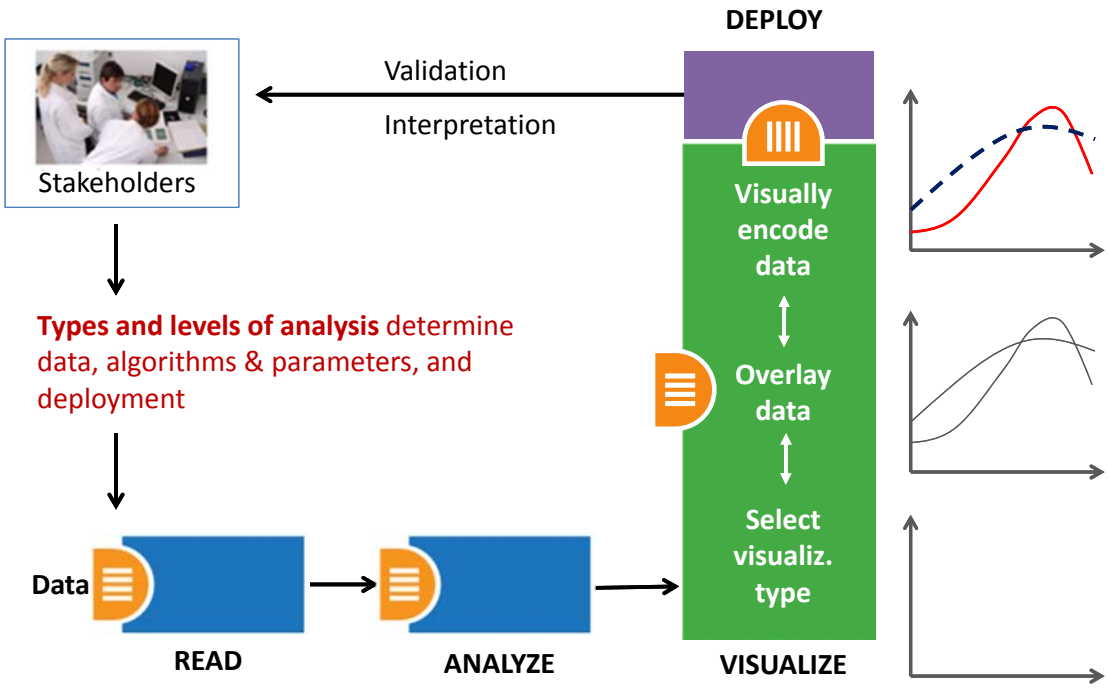


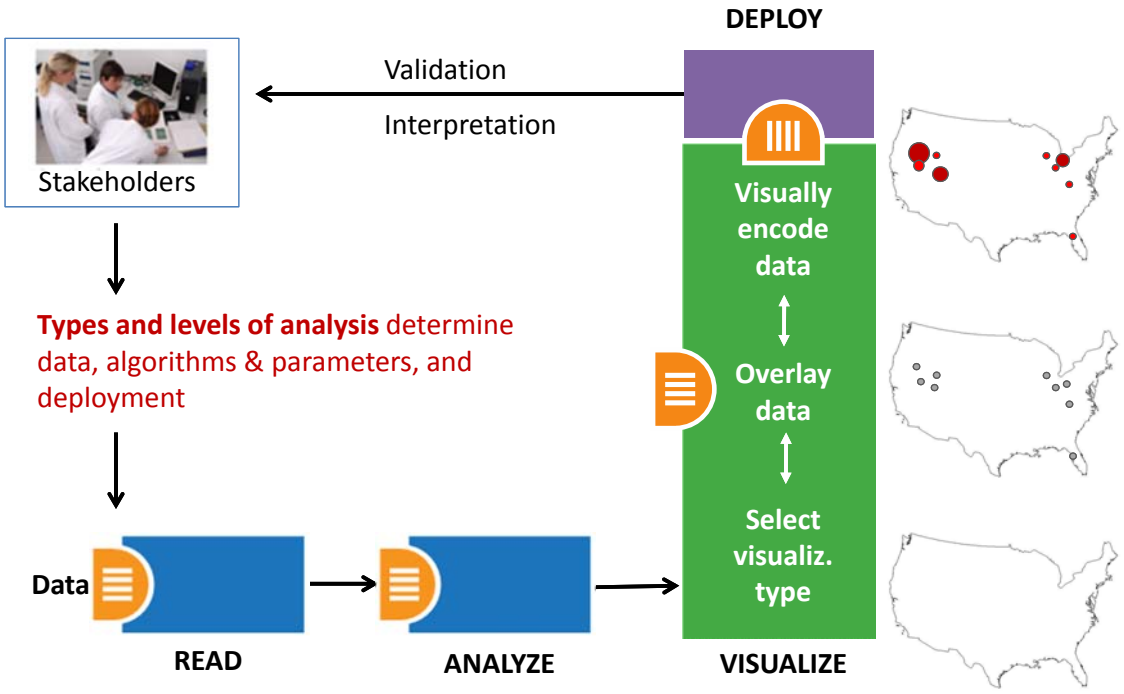
Needs-Driven Workflow Design



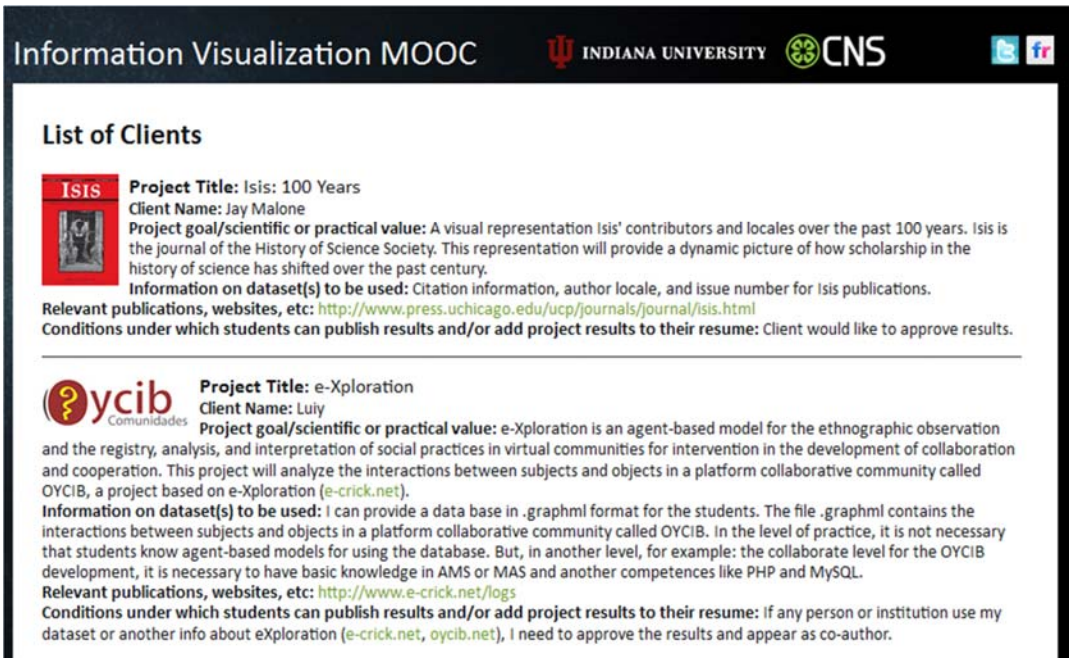
Needs-Driven Workflow Design



Needs-Driven Workflow Design

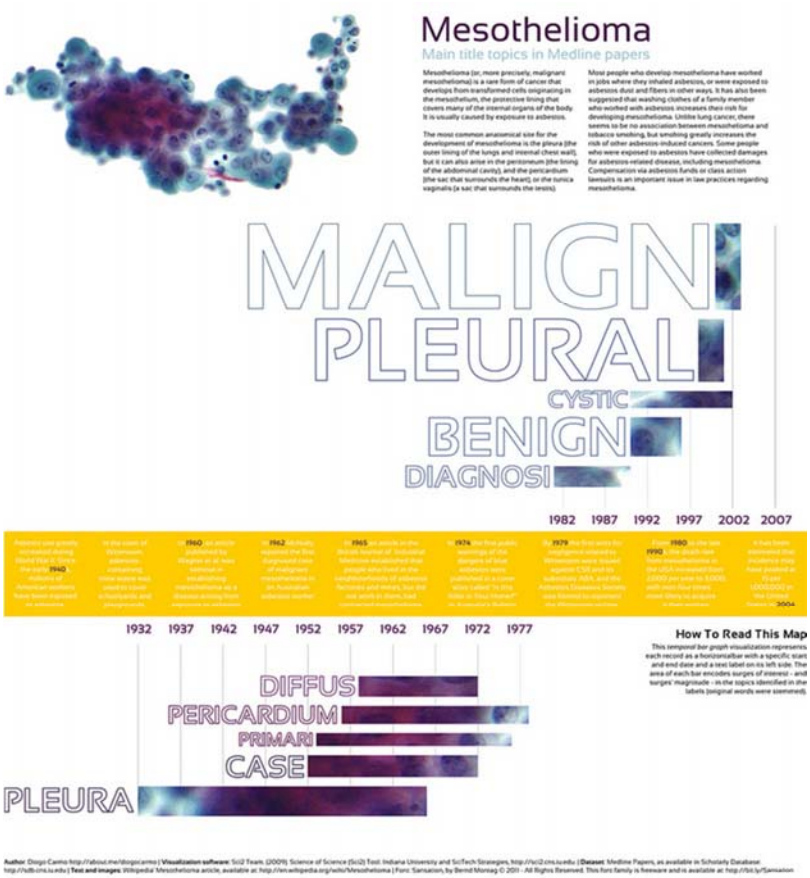


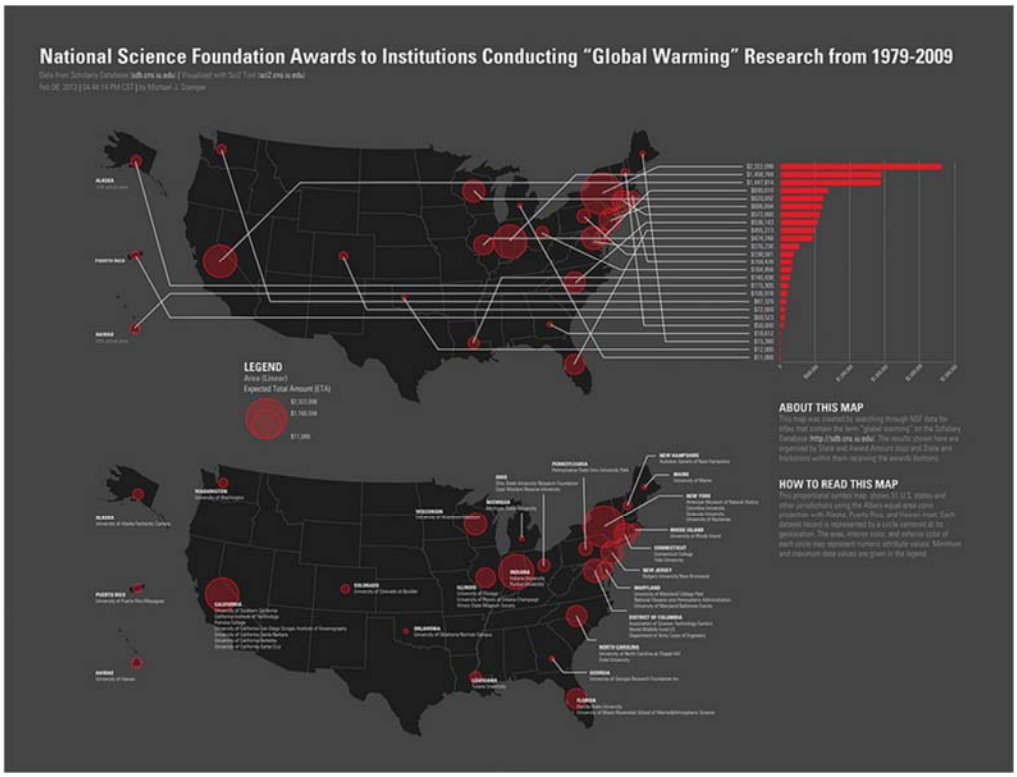
Clients



http://ivmooc.cns.iu.edu/ivmooc_clientprojects.html

Diogo Carmo





mjstamper_ivmooc

References

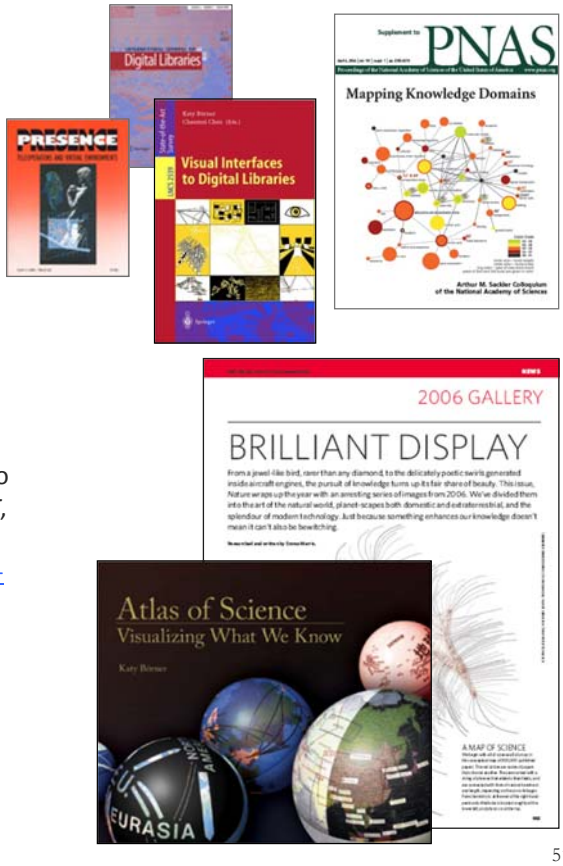
Börner, Katy, Chen, Chaomei, and Boyack, Kevin. (2003). **Visualizing Knowledge Domains.** In Blaise Cronin (Ed.), *ARIST*, Medford, NJ: Information Today, Volume 37, Chapter 5, pp. 179-255. http://ivl.slis.indiana.edu/km/pub/2003-borner-arist.pdf

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).

Börner, Katy, Sanyal, Soma and Vespignani, Alessandro (2007). **Network Science.** In Blaise Cronin (Ed.), *ARIST*, Information Today, Inc., Volume 41, Chapter 12, pp. 537-607. http://ivl.slis.indiana.edu/km/pub/2007-borner-arist.pdf

Börner, Katy (2010) **Atlas of Science**. MIT Press. http://scimaps.org/atlas

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



Acknowledgments

We would like to thank Miguel I. Lara and his colleagues at the Center for Innovative Teaching and Learning for instructional design support, Samuel Mills for designing the web pages, Robert P. Light and Thomas Smith for extending the GCB platform, and Mike Widmer and Mike T. Gallant for adding the Forum. Support comes from CNS, CITL, SLIS, SOIC, and Google.

The tool development work is supported in part by the Cyberinfrastructure for Network Science Center and the School of Library and Information Science at Indiana University, the National Science Foundation under Grants No. SBE-0738111 and IIS-0513650, the US Department of Agriculture, the National Institutes of Health, and the James S. McDonnell Foundation.

Visualizations used in the course come from the Places & Spaces: Mapping Science exhibit, online at http://scimaps.org, and from the *Atlas of Science: Visualizing What We Know*, MIT Press (2010).

INDIANA UNIVERSITY SCHOOL OF LIBRARY AND INFORMATION SCIENCE · NSF · USDA · NATIONAL INSTITUTES OF HEALTH · James S. McDonnell Foundation · CNS Cyberinfrastructure for Network Science Center

All papers, maps, tools, talks, press are linked from http://cns.iu.edu

CNS Facebook: http://www.facebook.com/cnscenter
Mapping Science Exhibit Facebook: http://www.facebook.com/mappingscience