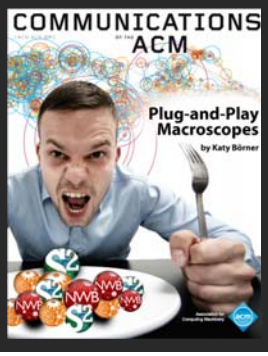
Visual Design Principles

Katy Börner
Cyberinfrastructure for Network Science Center, Director
Information Visualization Laboratory, Director
School of Library and Information Science
Indiana University, Bloomington, IN
katy@indiana.edu

With special thanks to the members at the Cyberinfrastructure for Network Science Center; the Sci2, NWB, and EpiC team; and the VIVO Collaboration

3rd International Meeting on Visualizing Biological Data (VIZBI 2012)
EMBL, Heidelberg, Germany
http://vizbi.org/2012

March 8, 2012

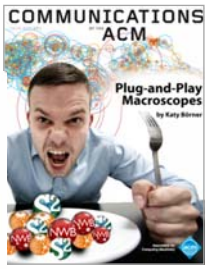


Overview

1. **Motivation:** Design informative and visually pleasing visualizations that make a difference.

2. **Theory:** Learn from and combine approaches from psychology, cartography, computer science, information visualization, statistics, graphic design.

3. **Practice:** Sample visualizations designed for experts and a general audience + plug-and-play macroscope tools that commoditize data mining and visualization.



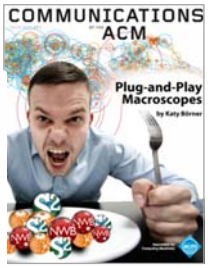
cyberinfrastructure for NETWORK SCIENCE CENTER
School of Library and Information Science | Indiana University Bloomington

Overview

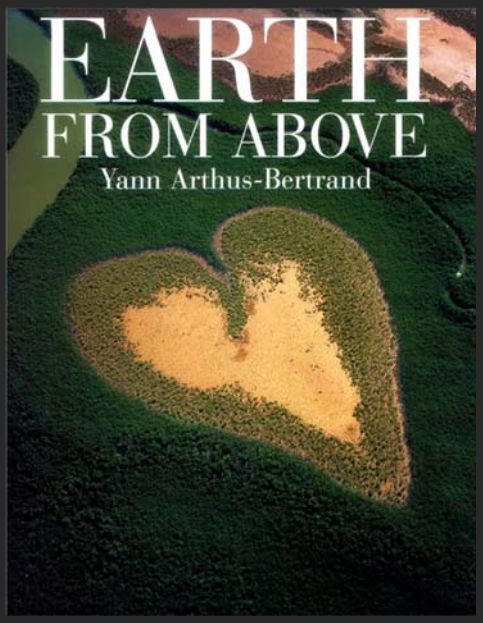
1. **Motivation:** Design informative and visually pleasing visualizations that make a difference.

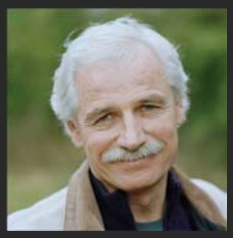
2. **Theory:** Learn from and combine approaches from psychology, cartography, computer science, information visualization, statistics, graphic design.

3. **Practice:** Sample visualizations designed for experts and a general audience + plug-and-play macroscope tools that commoditize data mining and visualization.

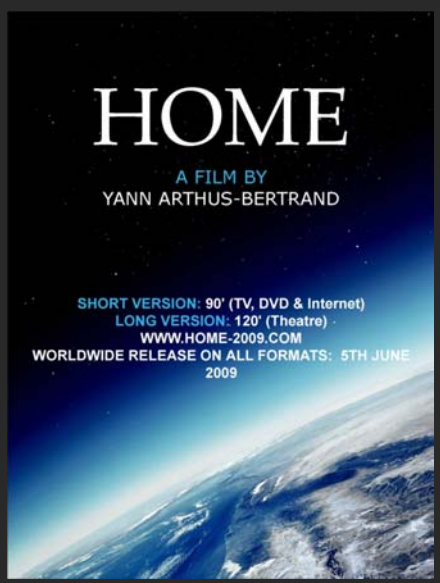


cyberinfrastructure for NETWORK SCIENCE CENTER
School of Library and Information Science | Indiana University Bloomington

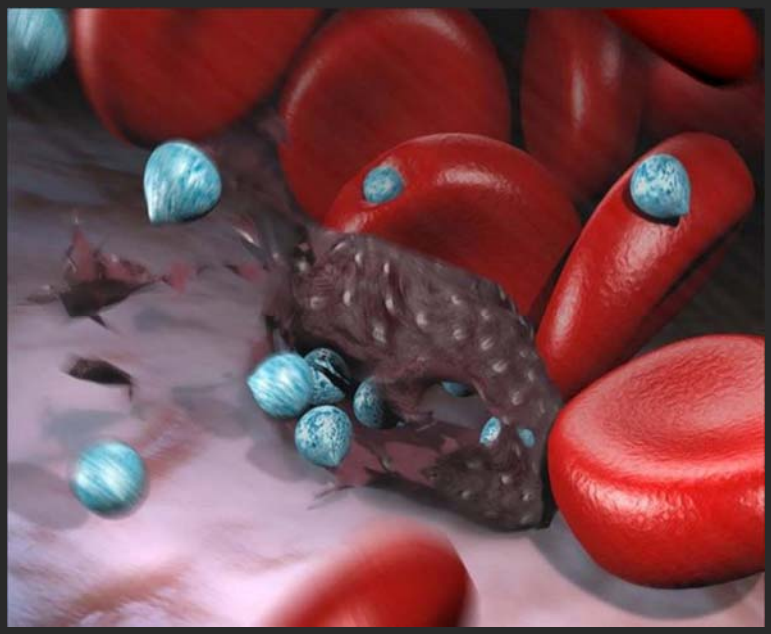




Yann Arthus-Bertrand



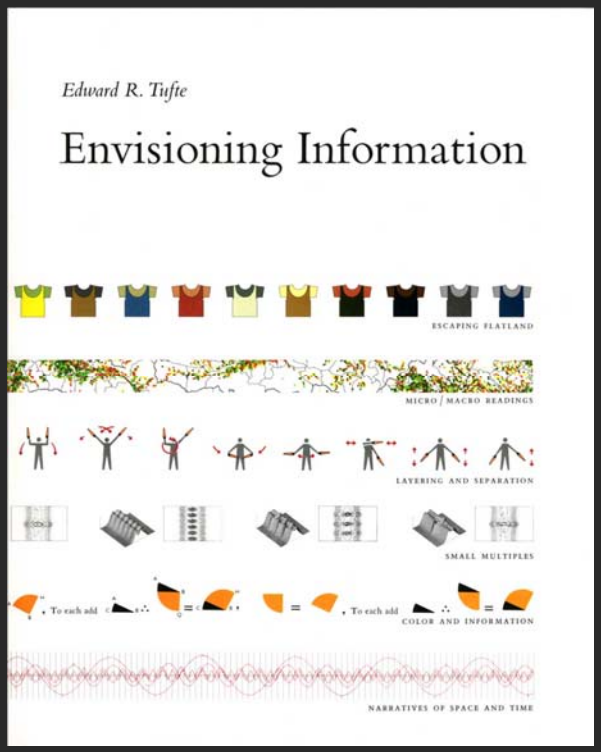
http://www.home-2009.com



http://www.malarialifecycle.com

The Whole Brain Catalog:
http://wholebraincatalog.org

Drew Berry



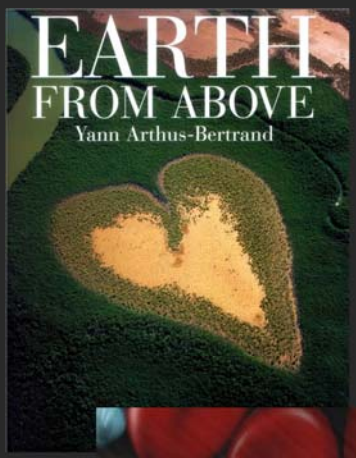
Edward R. Tufte

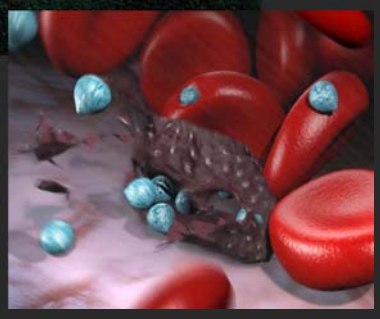
All three care deeply about

1. Data,
2. Existing expertise and insight needs, and
3. Are able to acquire the resources it takes to

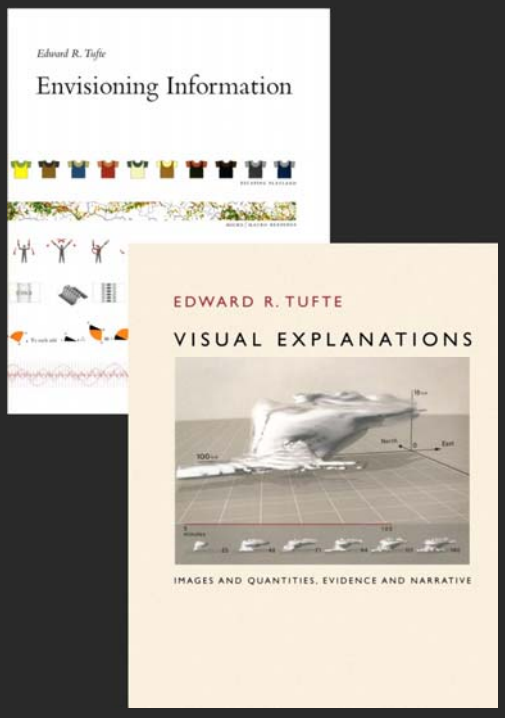
- Spent months/years to deeply understand the problem/possible solutions.
- Render data optimally for the human perceptual-cognitive system – given our current understanding of human perception/cognition and technology.

The result are insightful yet perceptually stunning, intellectually stimulating, and emotion provoking imagery.

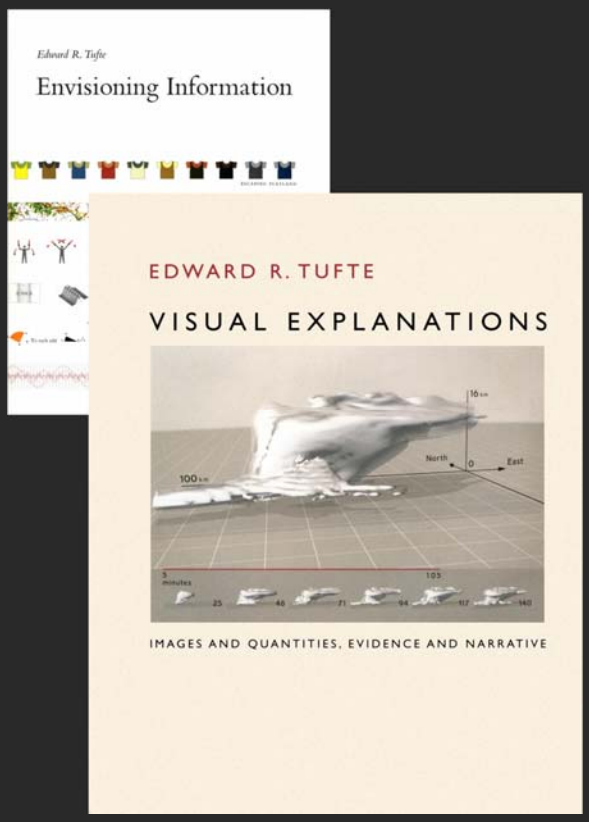




Photography, scientific visualization, animation— x-y-z positions exist.



Data visualization, data graphics— Mostly without x-y-z positions.



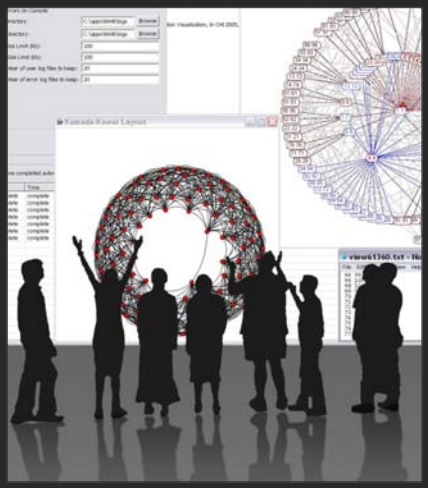
Today's massive amounts of streaming data cannot be rendered by hand.

How can computers be used to render informative and visually pleasing visualizations that make a difference?

Need BIG data analyses and visualizations that conform to human visual perception and cognitive processing.

Microscopes

Telescopes



Macroscopes

Hubble Telescope

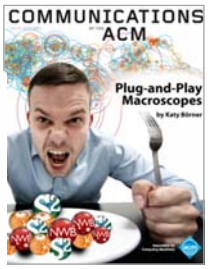
Original total cost estimate: US$400 million
Construction costs: US$2.5 billion
Cumulative costs: US$4.5 and $6 billion with €593 million EU contribution
http://www.spacetelescope.org/about/faq

Overview

1. **Motivation:** Design informative and visually pleasing visualizations that make a difference.

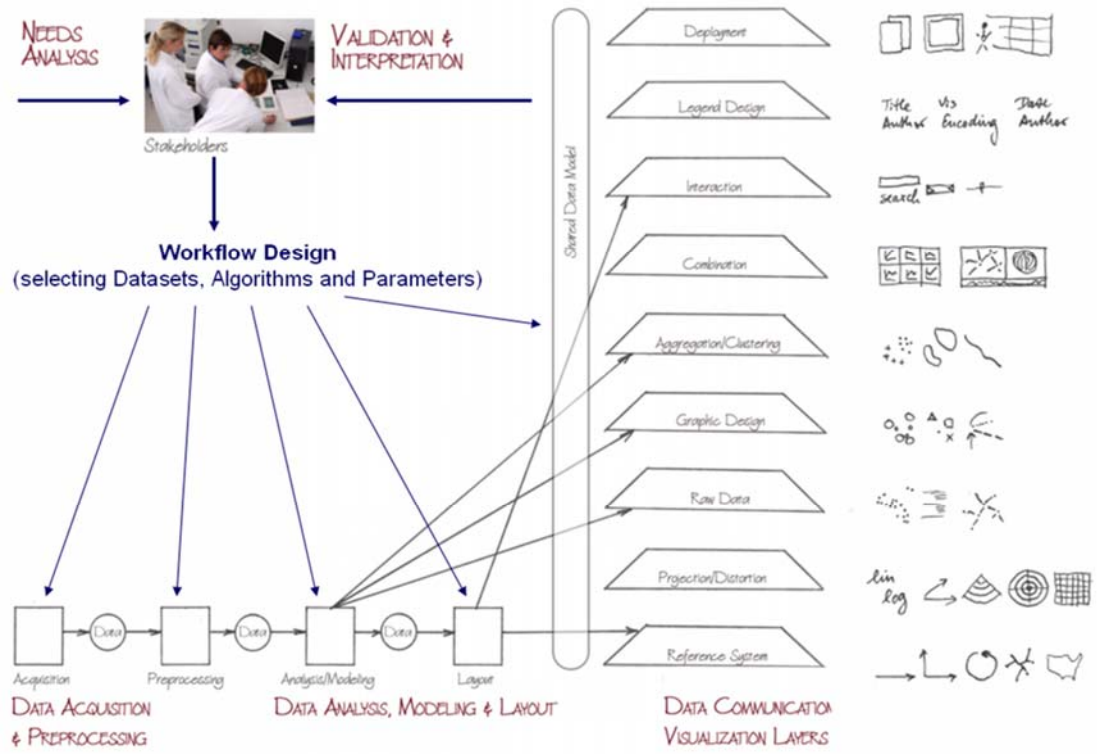
2. **Theory:** Learn from and combine approaches from psychology, cartography, computer science, information visualization, statistics, graphic design.

3. **Practice:** Sample visualizations designed for experts and a general audience + plug-and-play macroscope tools that commoditize data mining and visualization.



cyberinfrastructure for NETWORK SCIENCE CENTER
School of Library and Information Science | Indiana University Bloomington

Needs-Driven Workflow Design using a modular data acquisition/analysis/ modeling/ visualization pipeline as well as modular visualization layers.



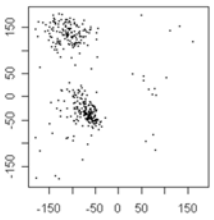
Börner, Katy (2010) Atlas of Science. MIT Press.

Visualization Layers: Reference System (samples)

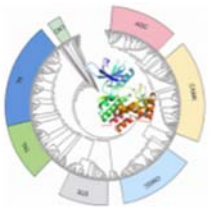
1D

Aravind
Baldauf
Becerra
Benachenhou
Brinkmann
Brocks
Brown
Caetano-Anolles

2D



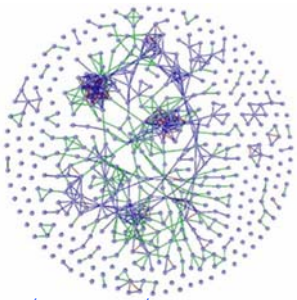
Circular



Tabular

Tree

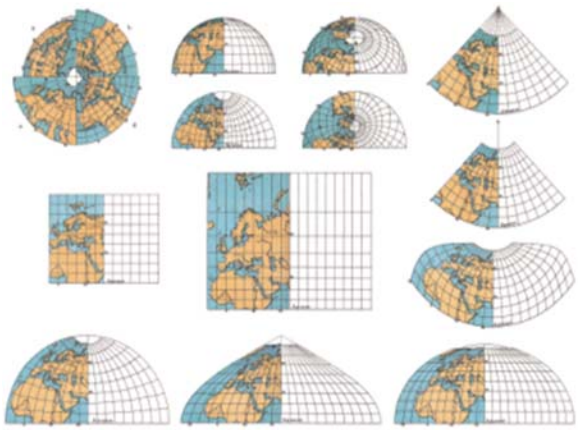
Network

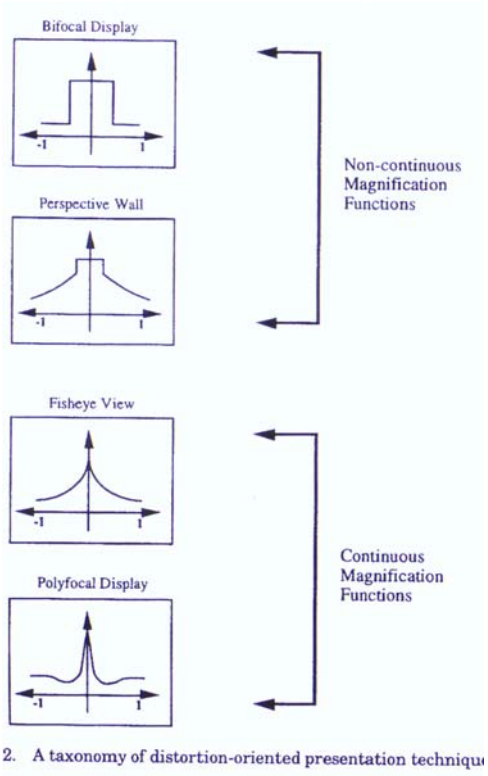


Map

Examples taken from http://vizbi.org/Posters/2012

Visualization Layers: Projection/Distortion

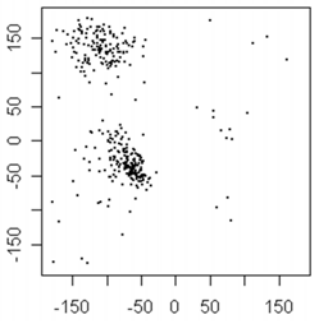


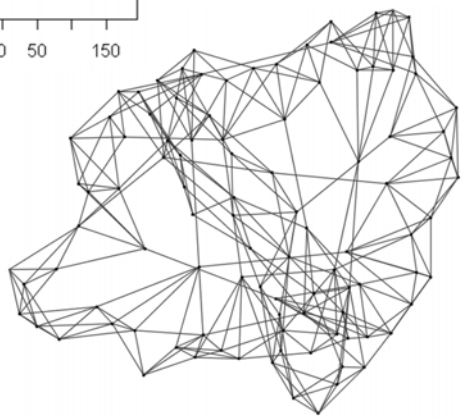


2. A taxonomy of distortion-oriented presentation techniqu

Leung, Y. and Apperley, M. (1994) A Review and Taxonomy of Distortion-Oriented Presentation Techniques, ACM ToCHI, 1 (2), pp. 126-160.

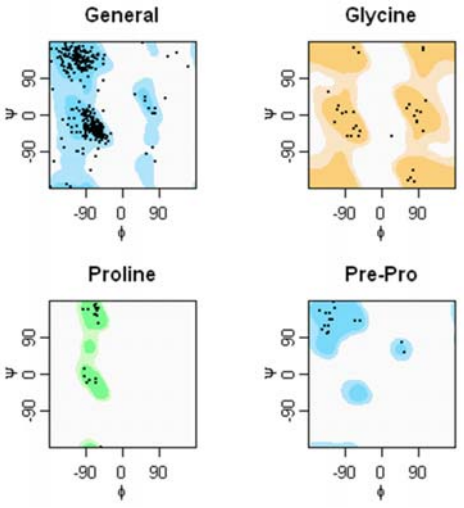
Visualization Layers: Raw Data

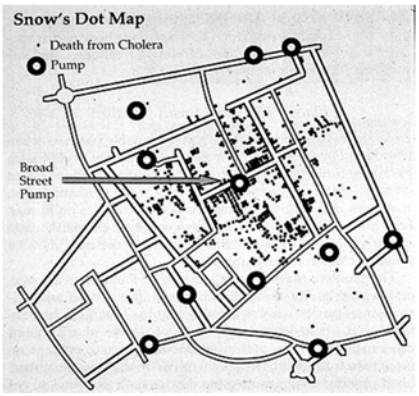


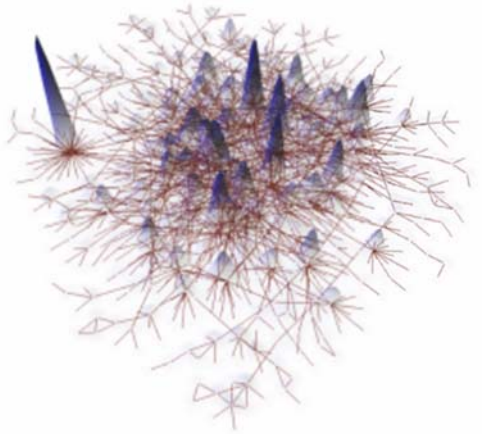


Examples taken from Wikipedia and
http://www2.warwick.ac.uk/fac/sci/moac/people/students/peter_cock/r/ramachandran/

Graphic Design





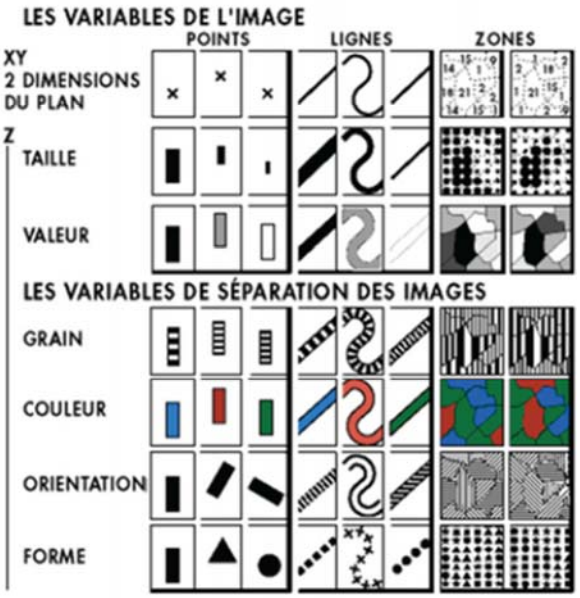


Shaping our digital future

Institute for Informatics & Digital Innovation

Semiology of Graphics

- visual encoding
  - points, lines, areas
    - patterns, trees/networks, grids
  - positional: XY
    - 1D, 2D, 3D
  - retinal: Z
    - size, lightness, texture, colour, orientation, shape,
  - temporal:
    - animation



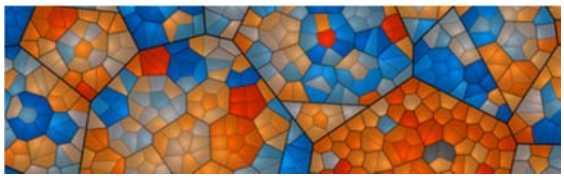
Semiology of Graphics. Jacques Bertin, Gauthier-Villars 1967, EHESS 1998

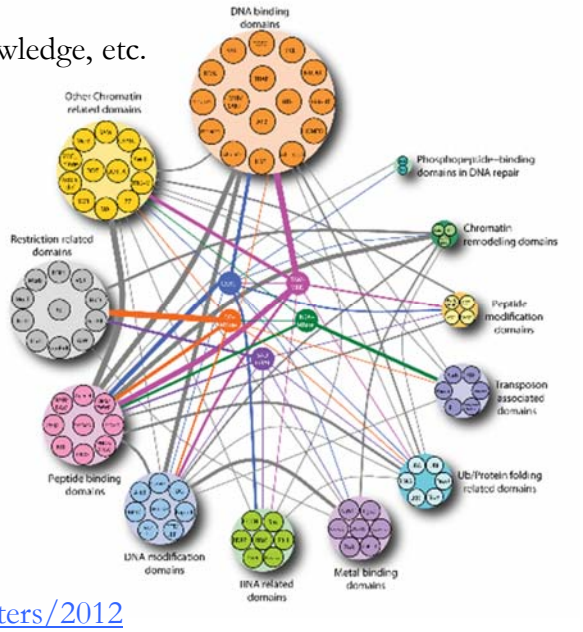
See Jessie Kennedy's Tutorial slides on "Visualization Principles"
http://mkweb.bcgsc.ca/vizbi/2012

Visualization Layers: Aggregation/Clustering

Aggregate over time, (geo) space, semantic similarity, network structure

Cluster using
- data mining on data itself.
- exogenous classifications, ontologies, knowledge, etc.
- Idiographic data classification using
  - natural breaks
  - quantiles
  - means and standard deviations
  - equal intervals
  - etc.

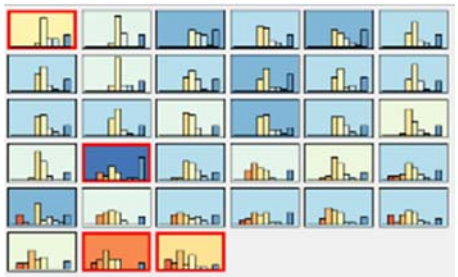




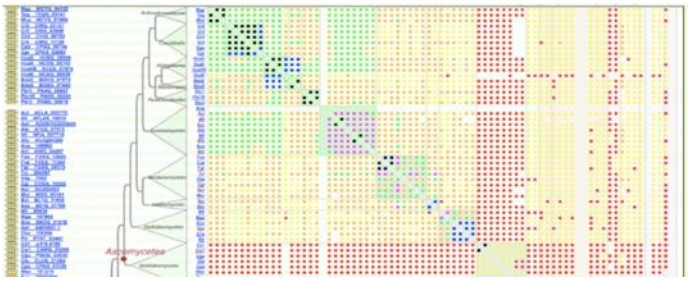
Examples taken from http://vizbi.org/Posters/2012

Visualization Layers: Combination

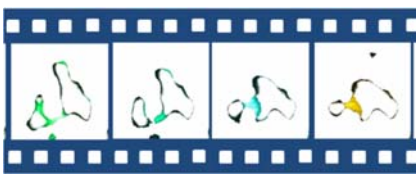
Small Multiples

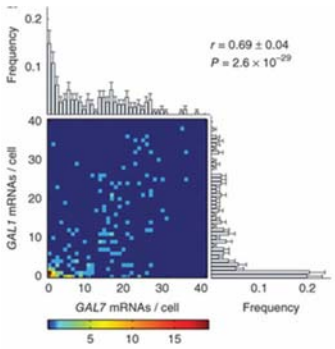


Multiple (coupled) windows



Change over time





Examples taken from http://vizbi.org/Posters/2012

Visualization Layers: Interaction

Search, filter, select

Zoom, Pan

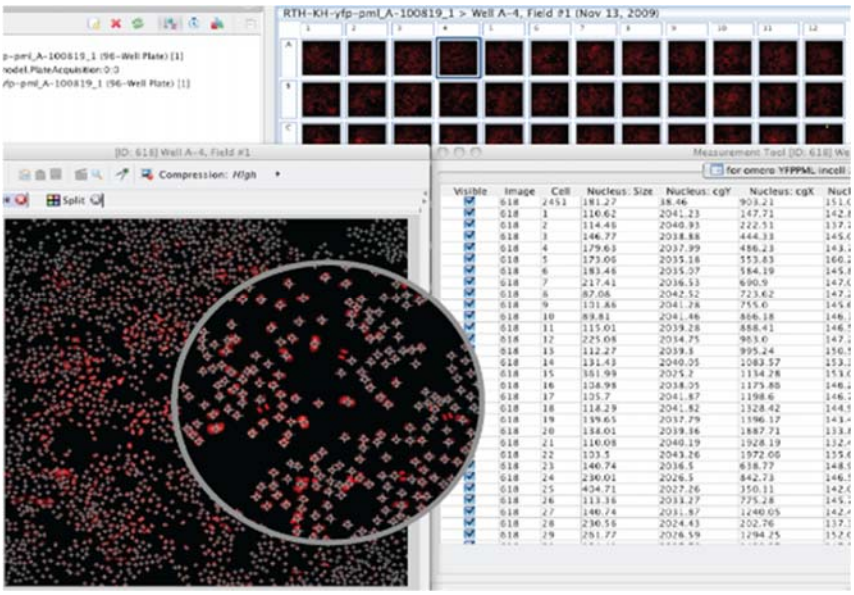
Pruning

Brushing

Details on demand

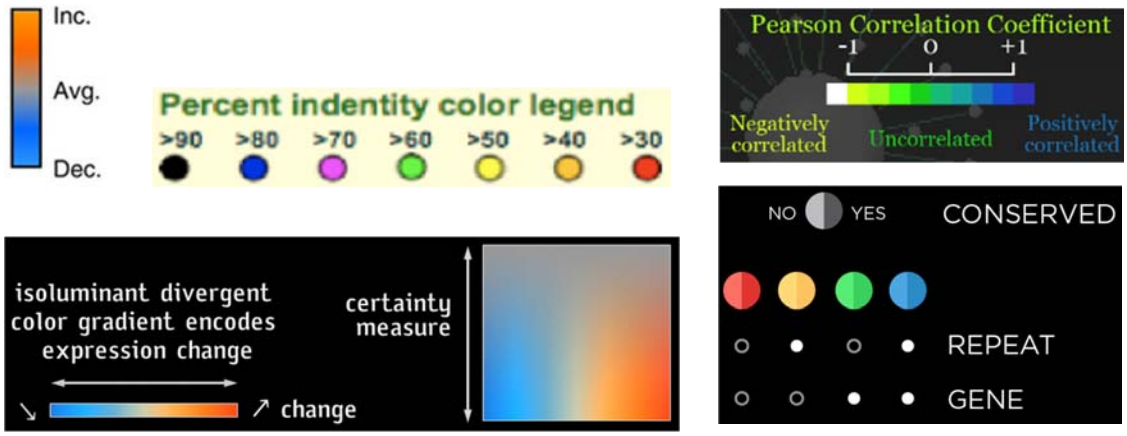
Focus & context

User status



Examples taken from http://vizbi.org/Posters/2012

Visualization Layers: Legend Design



Examples taken from http://vizbi.org/Posters/2012

Show scale for objects that have a real-world size.
Grid lines to help people visually track data.
Info on "How to read the map" and major insights to take away.

Data details, analysis+modeling algorithms and parameters, credits, author names.

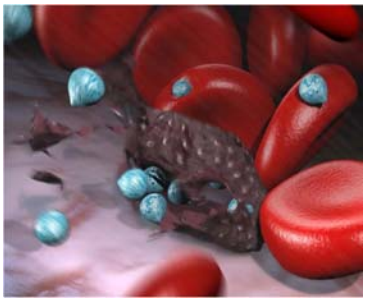
Visualization Layers: Deployment

Paper printout
(static, high resolution)

Interactive displays
(dynamic, low resolution)

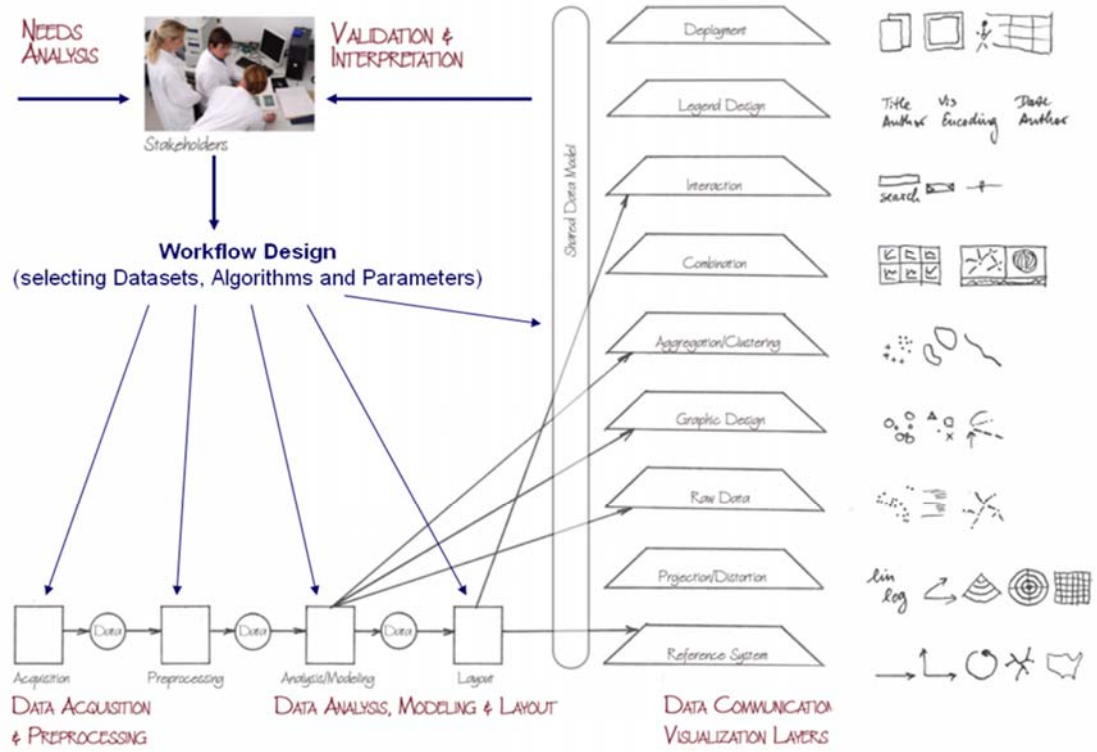
Illuminated Diagram

Animations



Hands-on exercises

Needs-Driven Workflow Design using a modular data acquisition/analysis/ modeling/ visualization pipeline as well as modular visualization layers.



Börner, Katy (2010) Atlas of Science. MIT Press.

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 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

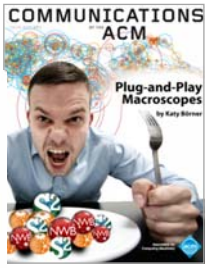
cyberinfrastructure for NETWORK SCIENCE CENTER
School of Library and Information Science | Indiana University Bloomington

Overview

1. **Motivation:** Design informative and visually pleasing visualizations that make a difference.

2. **Theory:** Learn from and combine approaches from psychology, cartography, computer science, information visualization, statistics, graphic design.

3. **Practice:** Sample visualizations designed for experts and a general audience + plug-and-play macroscope tools that commoditize data mining and visualization.



cyberinfrastructure for NETWORK SCIENCE CENTER
School of Library and Information Science | Indiana University Bloomington

Different Stakeholder Groups and Their Needs

Funding Agencies

- Need to monitor (long-term) money flow and research developments, identify areas for future development, stimulate new research areas, evaluate funding strategies for different programs, decide on project durations, funding patterns.

Scholars

- Want easy access to research results, relevant funding programs and their success rates, potential collaborators, competitors, related projects/publications *(research push)*.

Industry

- Is interested in fast and easy access to major results, experts, etc. Influences the direction of research by entering information on needed technologies *(industry-pull)*.

Advantages for Publishers

- Need easy to use interfaces to massive amounts of interlinked data. Need to communicate data provenance, quality, and context.

Society

- Needs easy access to scientific knowledge and expertise.

Scholars Have Different Roles/Needs

Researchers and Authors—need to select promising research topics, students, collaborators, and publication venues to increase their reputation. They benefit from a global view of competencies, reputation and connectivity of scholars; hot and cold research topics and bursts of activity, and funding available per research area.

Editors—have to determine editorial board members, assign papers to reviewers, and ultimately accept or reject papers. Editors need to know the position of their journals in the evolving world of science. They need to advertise their journals appropriately and attract high-quality submissions, which will in turn increase the journal's reputation.

Reviewers—read, critique, and suggest changes to help improve the quality of papers and funding proposals. They need to identify related works that should be cited or complementary skills that authors might consider when selecting project collaborators.

Teachers/Mentors—teach classes, train doctoral students, and supervise postdoctoral researchers. They need to identify key works, experts, and examples relevant to a topic area and teach them in the context of global science.

Inventors—create intellectual property and obtain patents, thus needing to navigate and make sense of research spaces as well as intellectual property spaces.

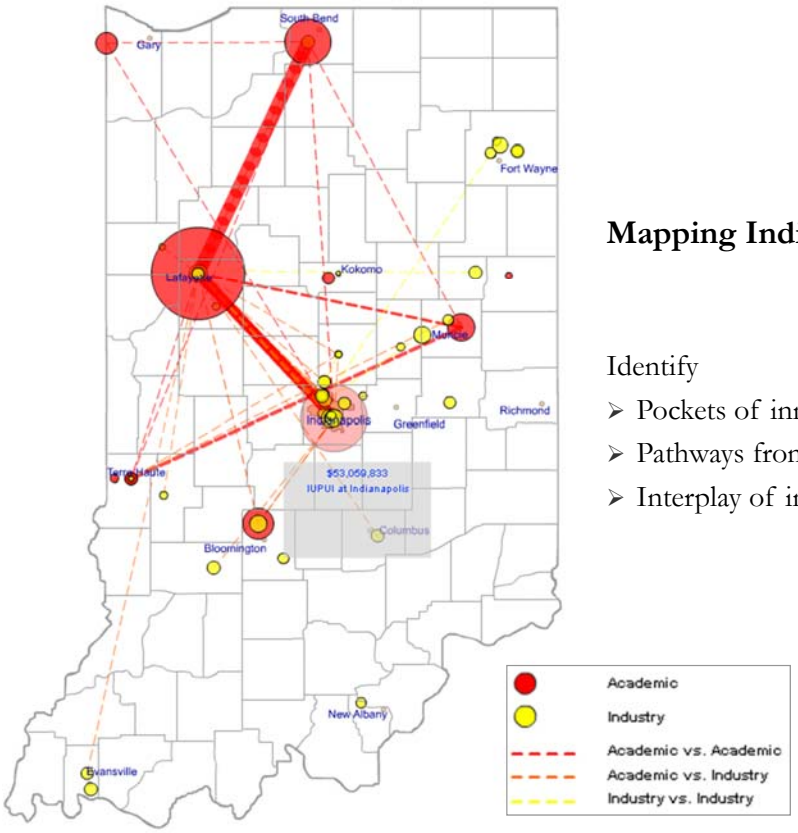
Investigators—scholars need funding to support students, hire staff, purchase equipment, or attend conferences. Here, research interests and proposals have to be matched with existing federal and commercial funding opportunities, possible industry collaborators and sponsors.

Team Leads and Science Administrators—many scholars direct multiple research projects simultaneously. Some have full-time staff, research scientists, and technicians in their laboratories and centers. Leaders need to evaluate performance and provide references for current or previous members; report the progress of different projects to funding agencies.

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 NS… SA, all of sci…
Temporal Analysis (When)	Funding portfolio of one individual	…ic bursts …f PNAS	113 Years of P… Research
Geospatial Analysis (Where)	Career trajectory of one individual	Mapping a s… intellectual l…	PNAS …
Topical Analysis (What)	…s.	…research	VxOrd/Topic r… NIH funding
Network Analysis (With Whom?)	NSF … work of one …	…	NIH's … cy

cyberinfrastructure for NETWORK SCIENCE CENTER
School of Library and Information Science | Indiana University Bloomington



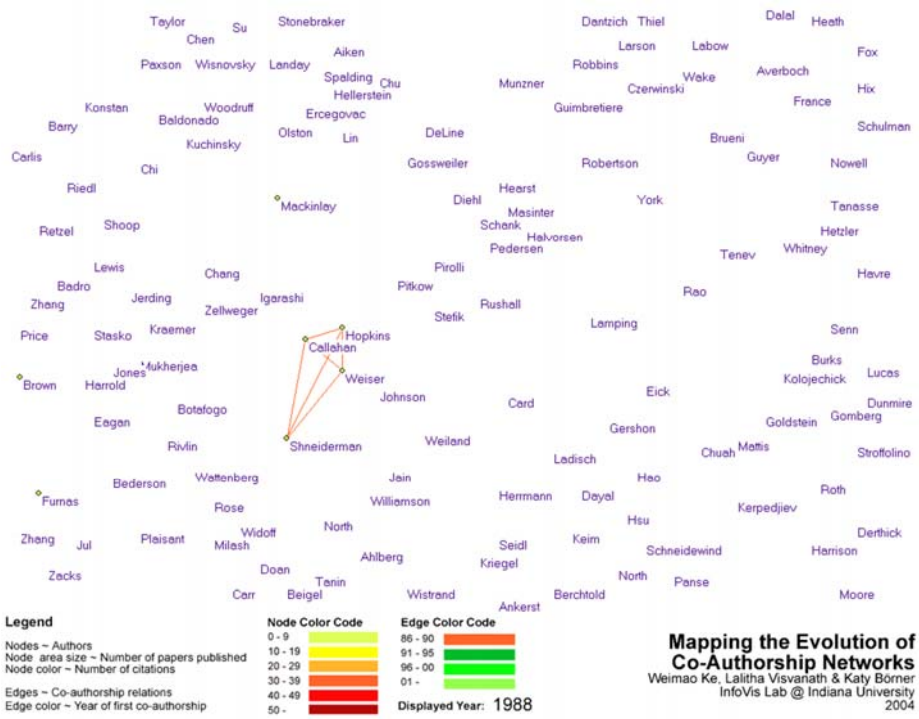
Mapping Indiana's Intellectual Space

Identify

- Pockets of innovation
- Pathways from ideas to products
- Interplay of industry and academia

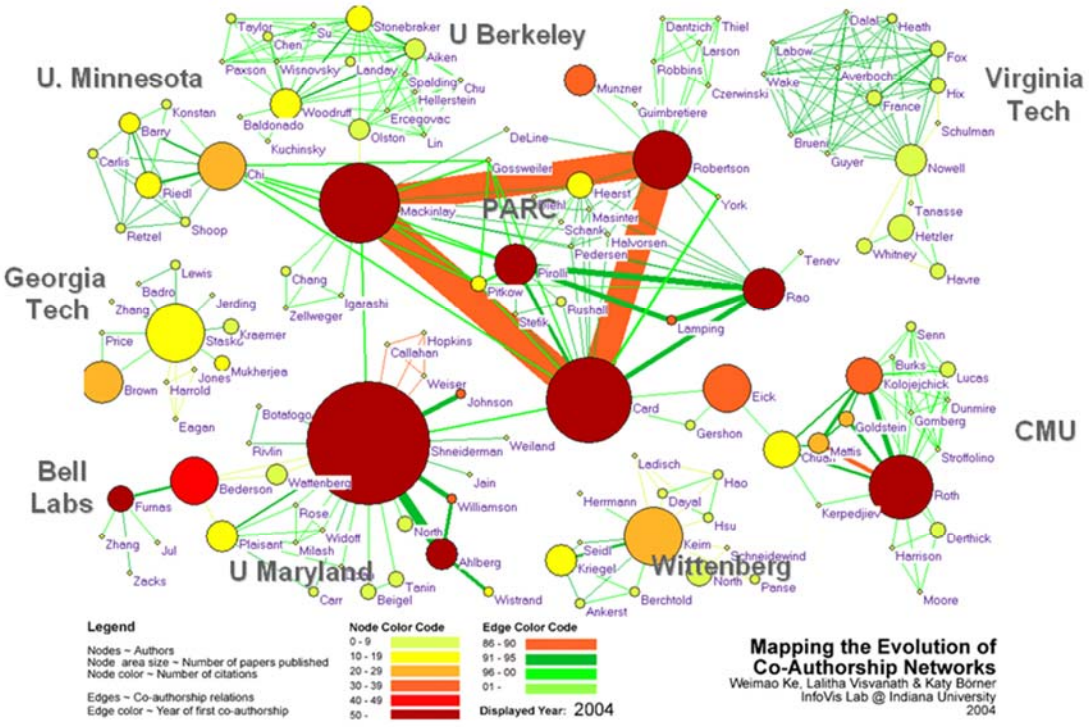
Mapping the Evolution of Co-Authorship Networks

Ke, Visvanath & Börner, (2004) Won 1st price at the IEEE InfoVis Contest.



Mapping the Evolution of Co-Authorship Networks

Ke, Visvanath & Börner, (2004) Won 1st price at the IEEE InfoVis Contest.

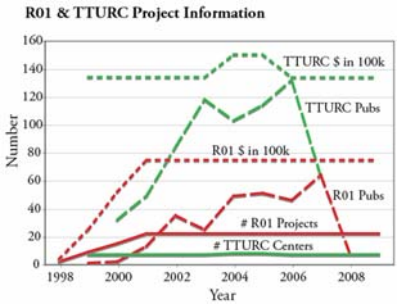


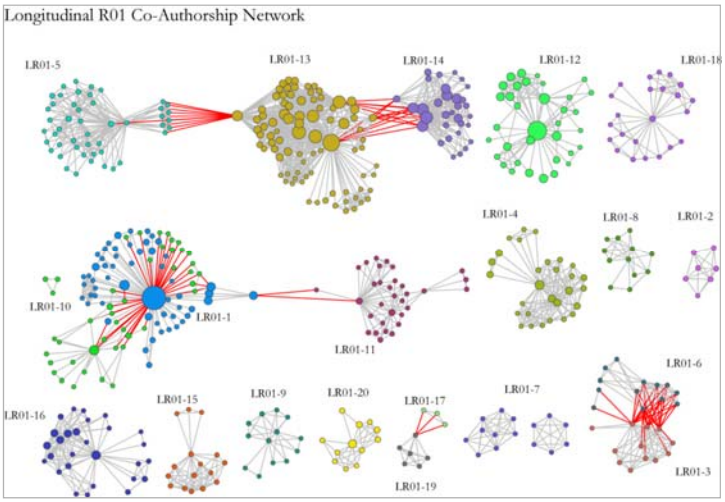
Mapping Transdisciplinary Tobacco Use Research Centers Publications

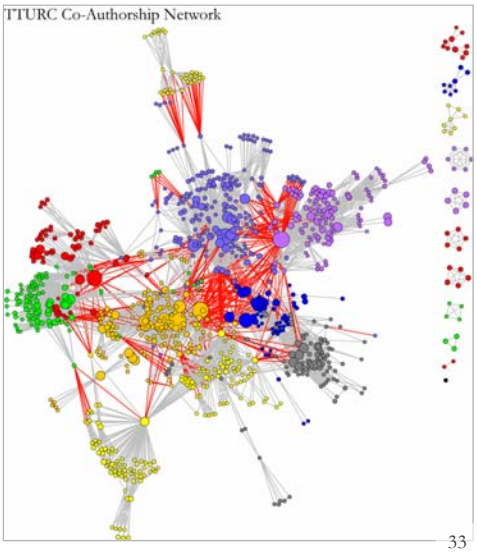
Compare R01 investigator based funding with TTURC Center awards in terms of number of publications and evolving co-author networks.

Zoss & Börner, forthcoming.

Supported by NIH/NCI Contract HHSN261200800812







Mapping Topic Bursts

Co-word space of the top 50 highly frequent and bursty words used in the top 10% most highly cited PNAS publications in 1982-2001.

Mane & Börner. (2004) PNAS, 101(Suppl. 1): 5287-5290.

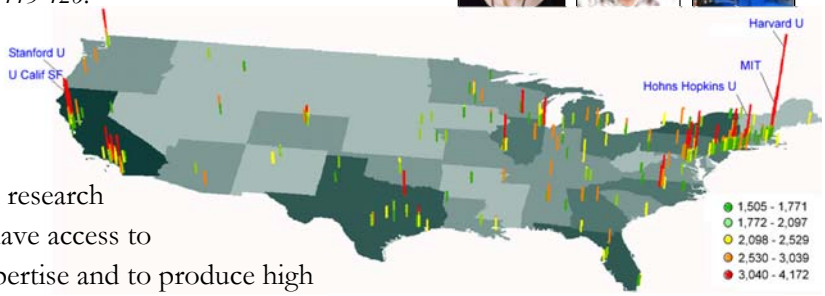
Spatio-Temporal Information Production and Consumption of Major U.S. Research Institutions

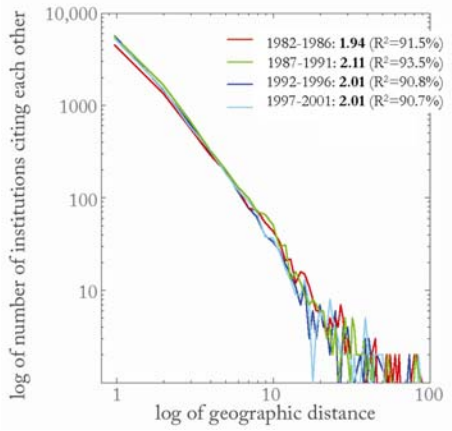


Börner, Katy, Penumarthy, Shashikant, Meiss, Mark and Ke, Weimao. (2006) Mapping the Diffusion of Scholarly Knowledge Among Major U.S. Research Institutions. Scientometrics. 68(3), pp. 415-426.

Research questions:

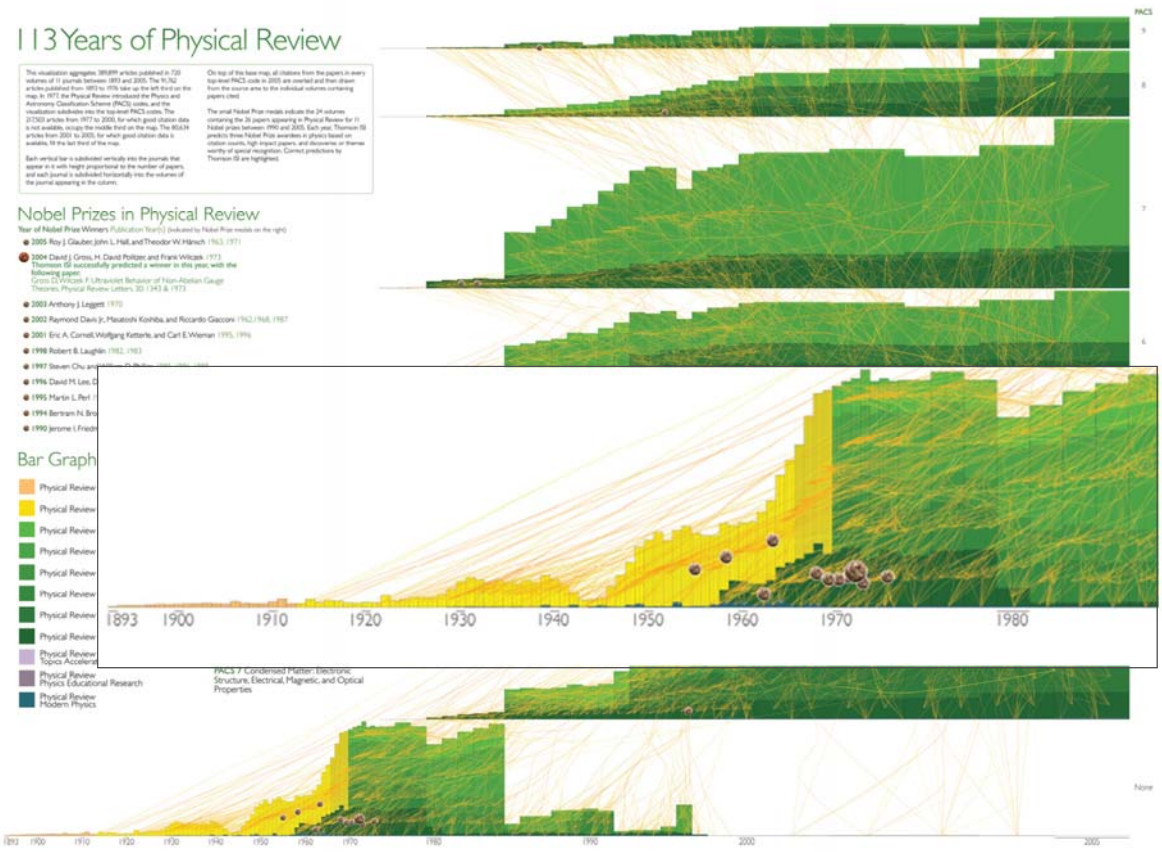
1. Does space still matter in the Internet age?
2. Does one still have to study and work at major research institutions in order to have access to high quality data and expertise and to produce high quality research?
3. Does the Internet lead to more global citation patterns, i.e., more citation links between papers produced at geographically distant research instructions?





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.

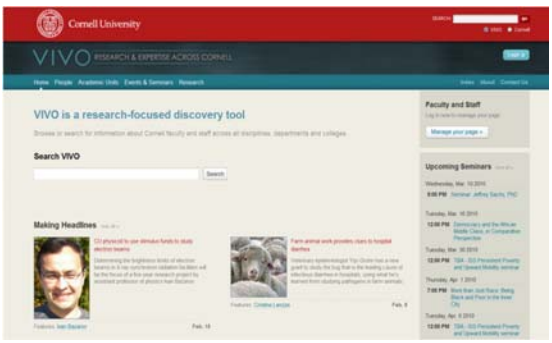


VIVO International Researcher Network

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.

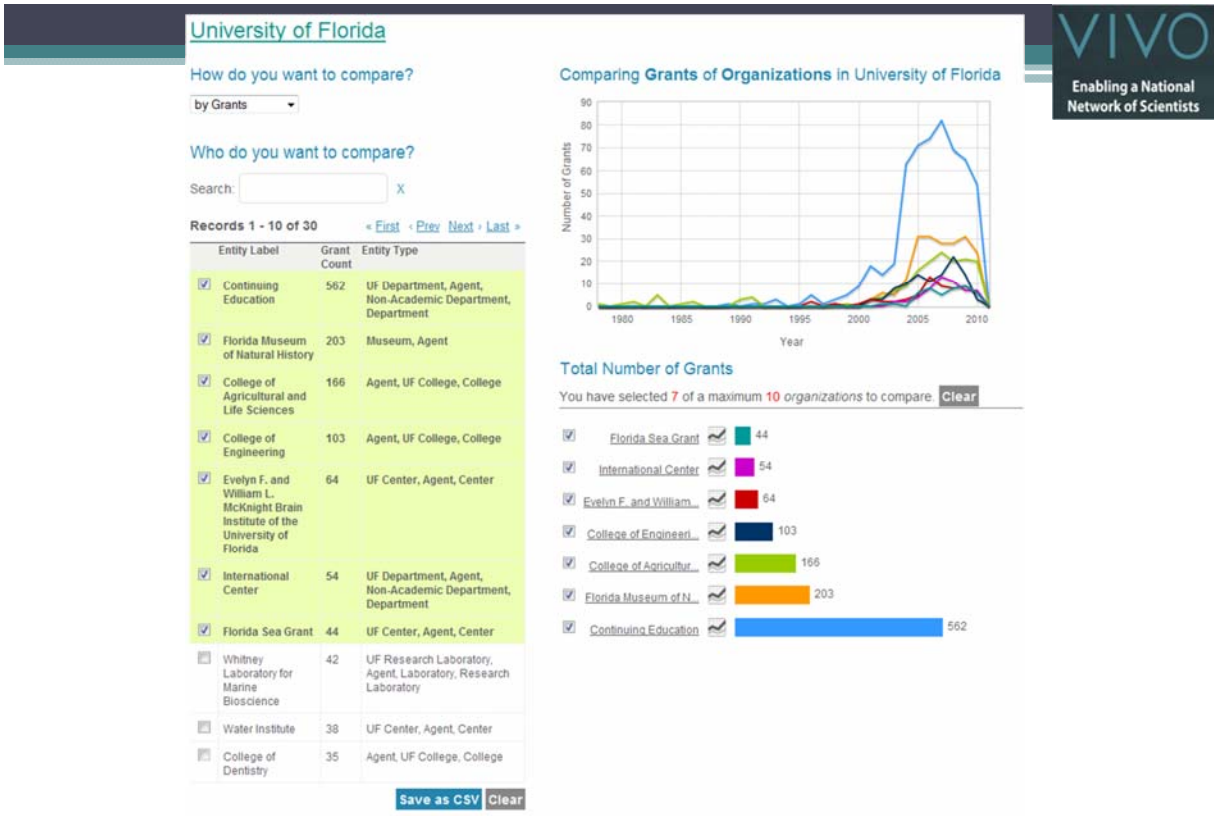


Funded by $12 million NIH award.

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.

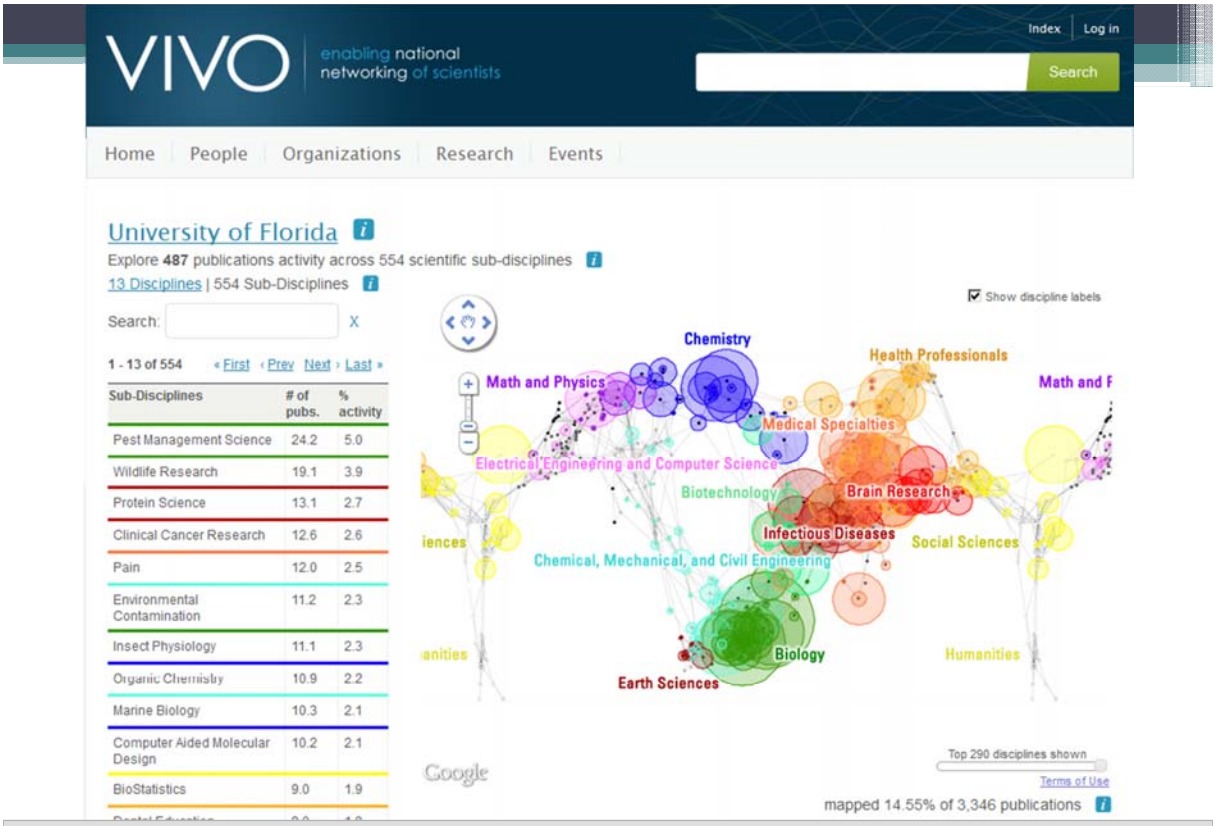
UF Clinical and Translational Science Institute
UNIVERSITY of FLORIDA

39

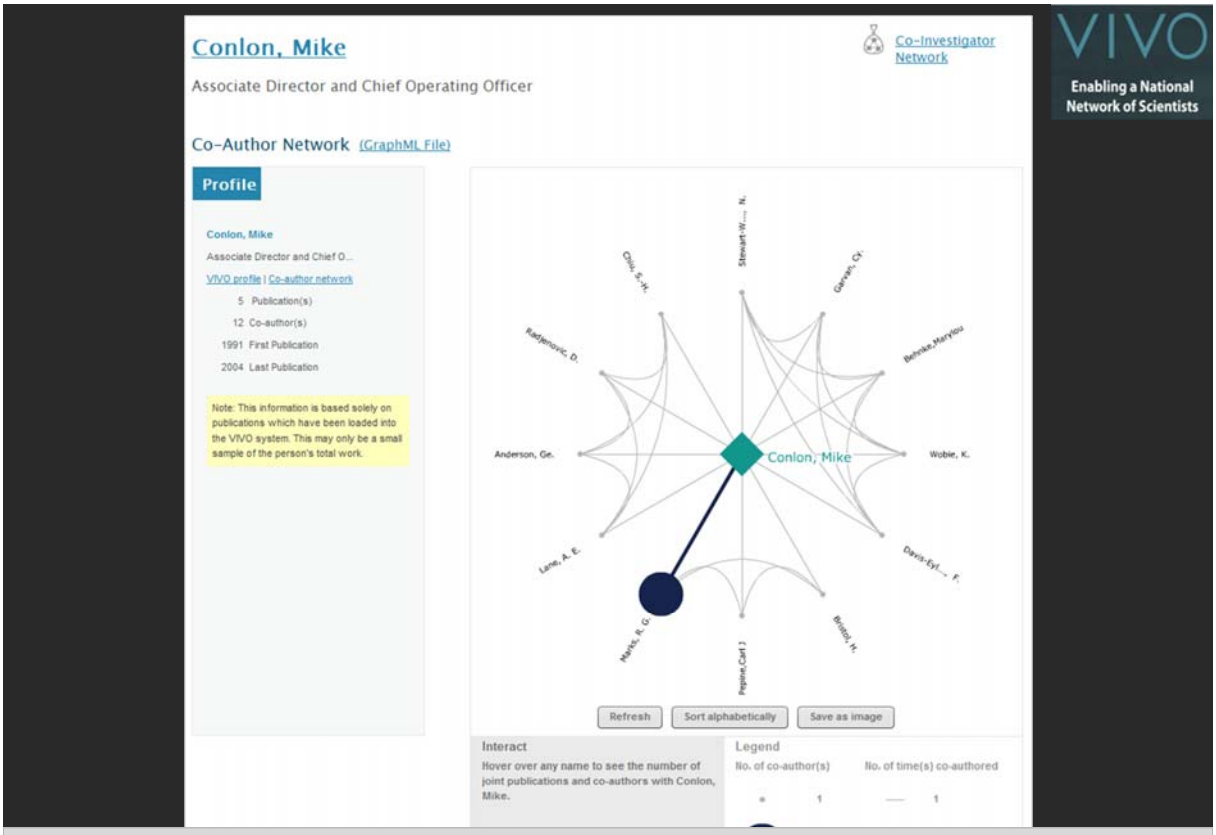


Temporal Analysis (When) Temporal visualizations of the number of papers/funding award at the institution, school, department, and people level

40



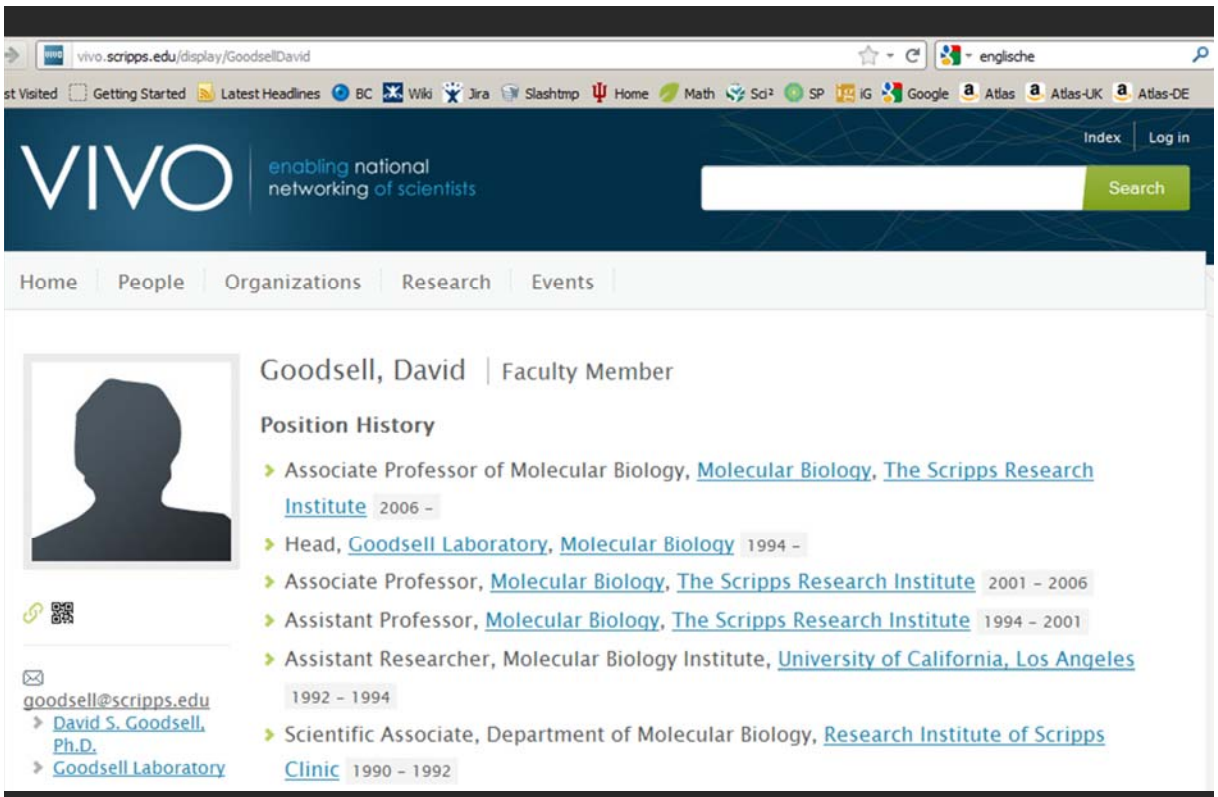
Topical Analysis (What) Science map overlays will show where a person, department, or university publishes most in the world of science. (in work)



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.



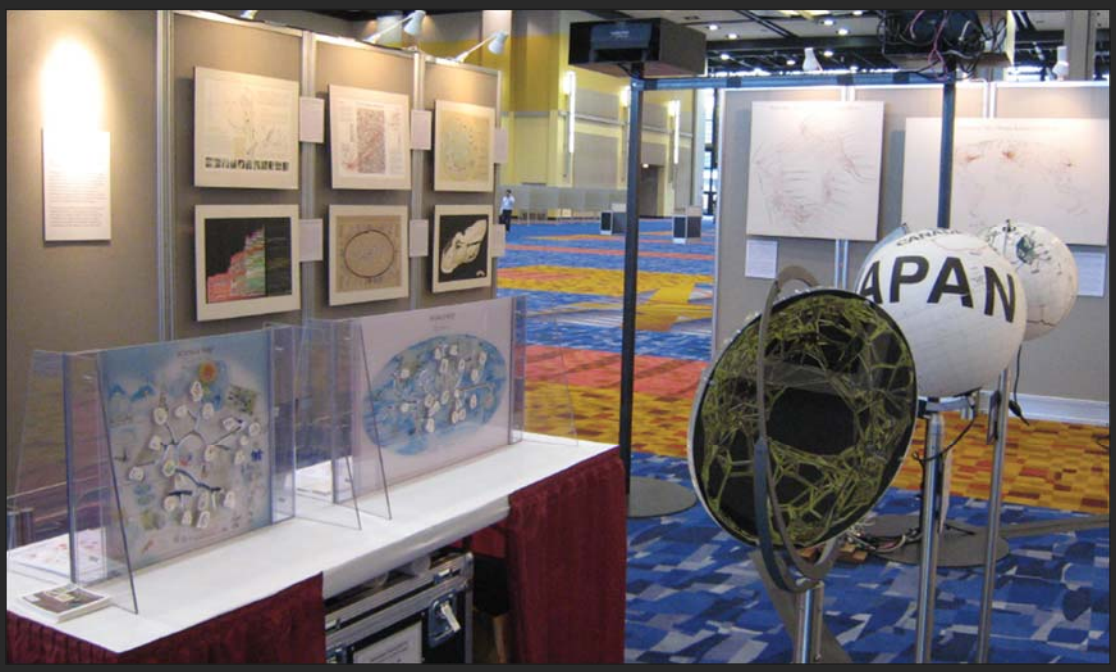
http://vivo.scripps.edu/display/GoodsellDavid

VIVO On-The-Go

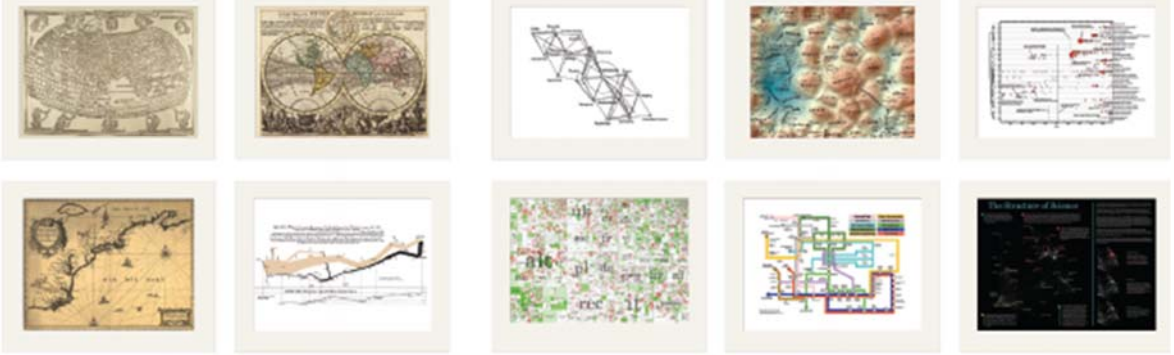
Overview, Interactivity, Details on Demand
come to commonly used devices and environments

Mapping Science Exhibit – 10 Iterations in 10 years

http://scimaps.org/



The Power of Maps 2005



The Power of Reference Systems 2006

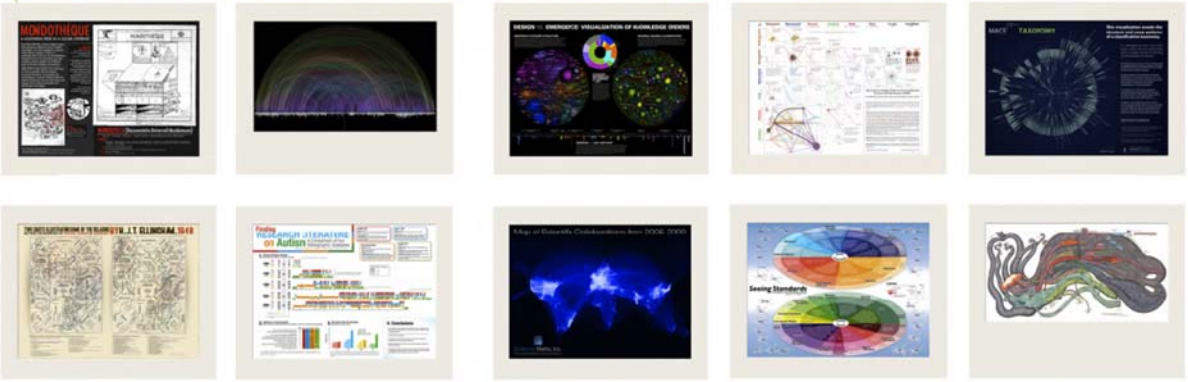
The Power of Forecasts 2007

Science Maps for Economic Decision Makers 2008

Science Maps for Science Policy Makers 2009

Science Maps for Scholars 2010

Science Maps as Visual Interfaces to Digital Lieraries 2011



Read about and zoom into maps at http://scimaps.org/exhibit_info

Call for Maps for the 8th Iteration of the Places & Spaces: Mapping Science Exhibit on "Science Maps for Kids" (2012)

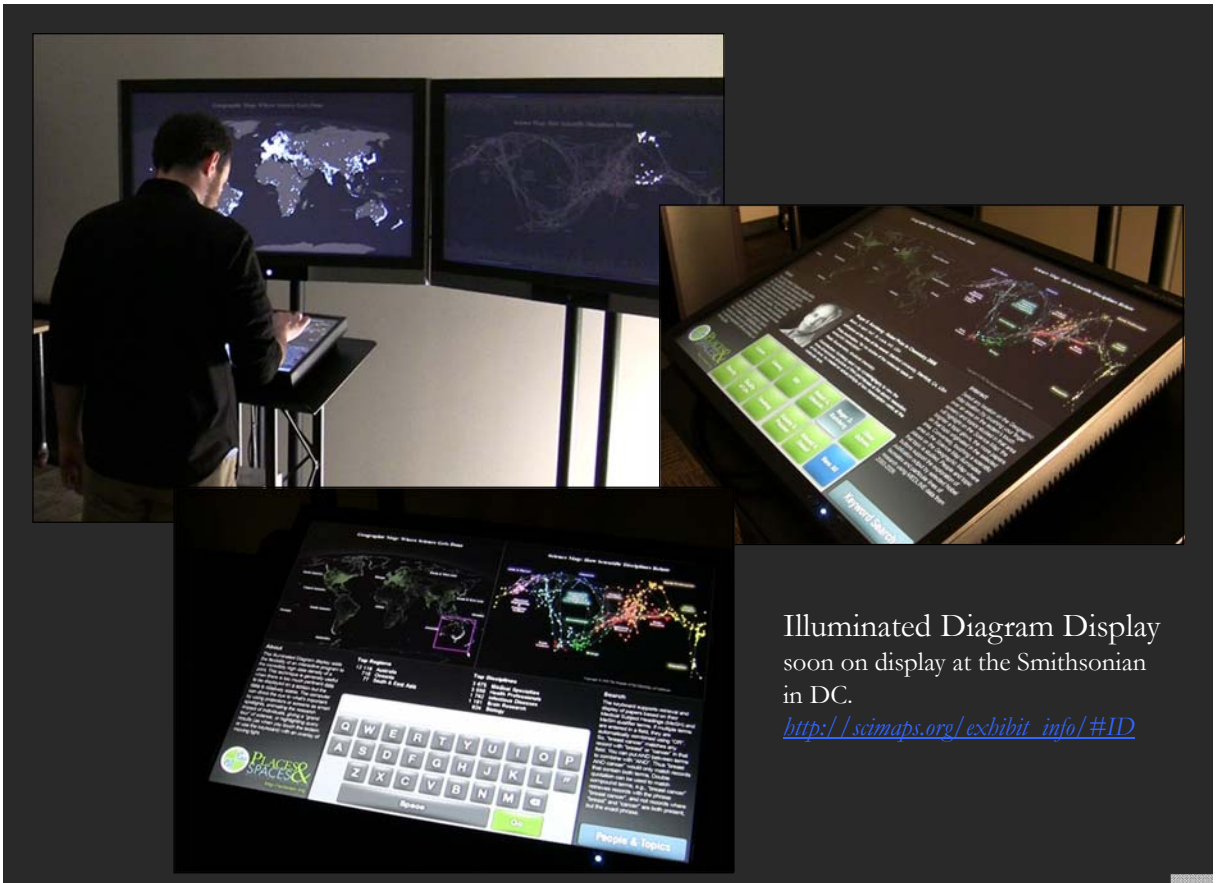
http://scimaps.org/call

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*

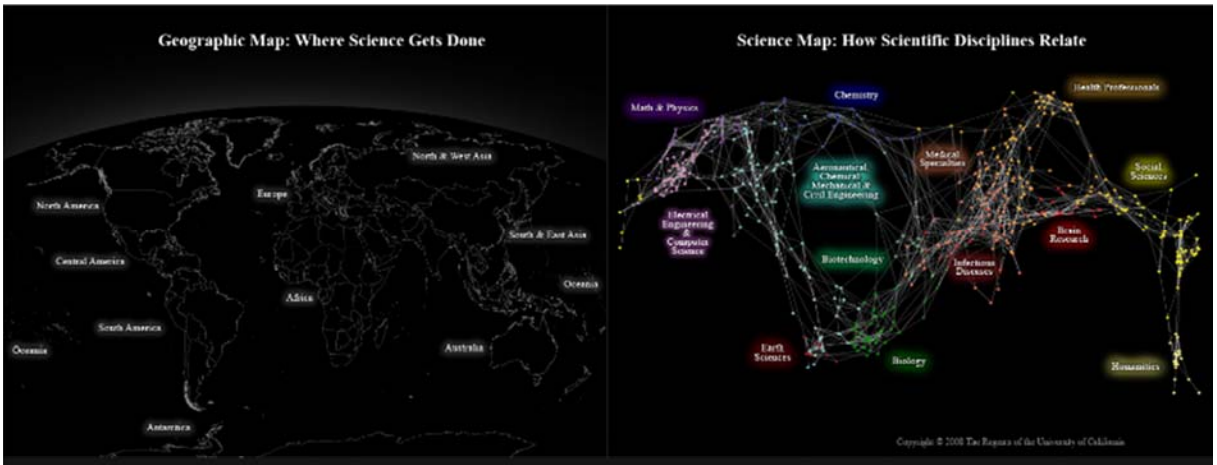
51

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

52



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



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.

Cloning

Cloning is the process of making a genetically identical copy. Cloning can refer to the technique of producing a genetic copy of an organism by replacing the nucleus of an unfertilized ovum with the nucleus of a body cell from the organism. The reconstructed egg containing the DNA from a donor cell must be treated with chemicals or electric current in order to stimulate cell division. Once the cloned embryo reaches a suitable stage, it is transferred to the uterus of a female host where it continues to develop until birth.

The first adult mammal cloned was Dolly the Sheep in 1997.

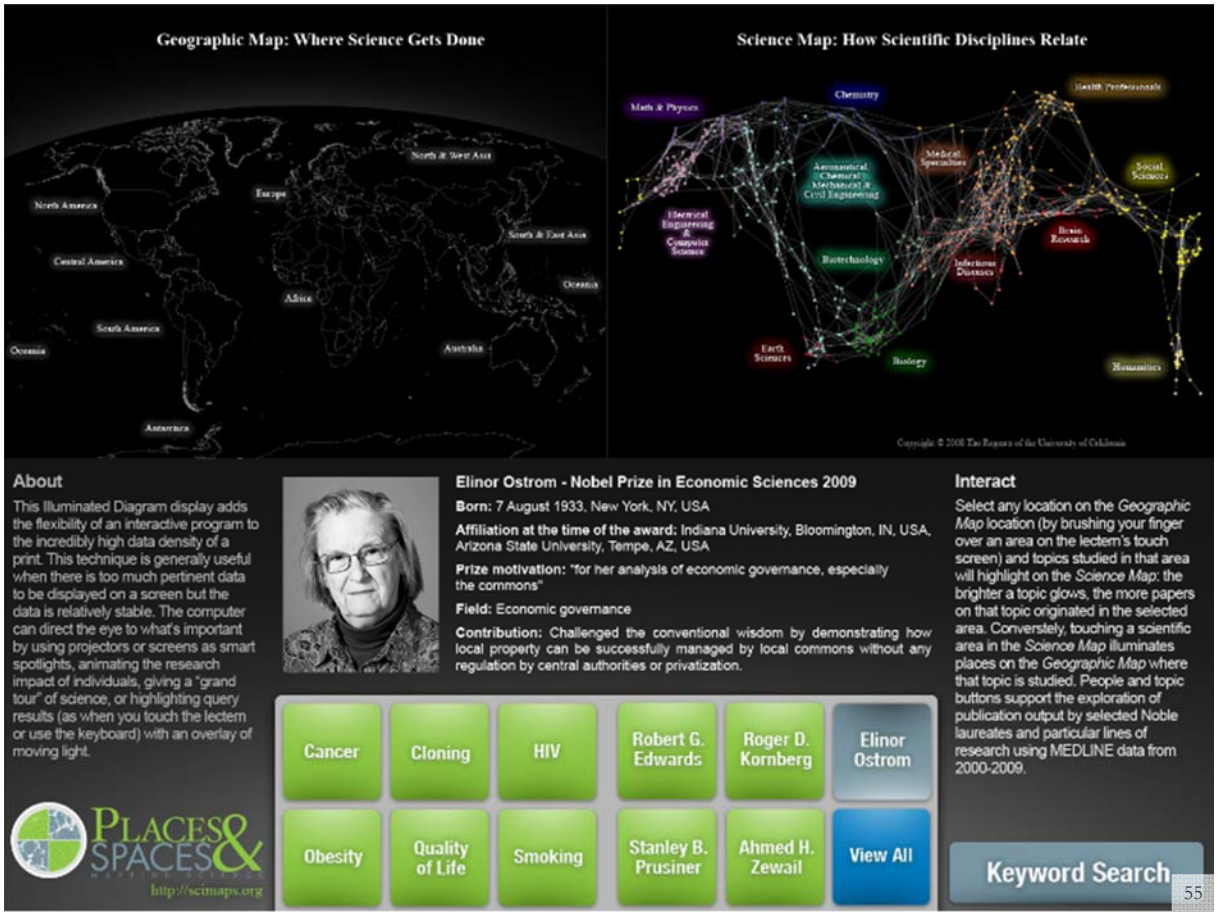
Interact

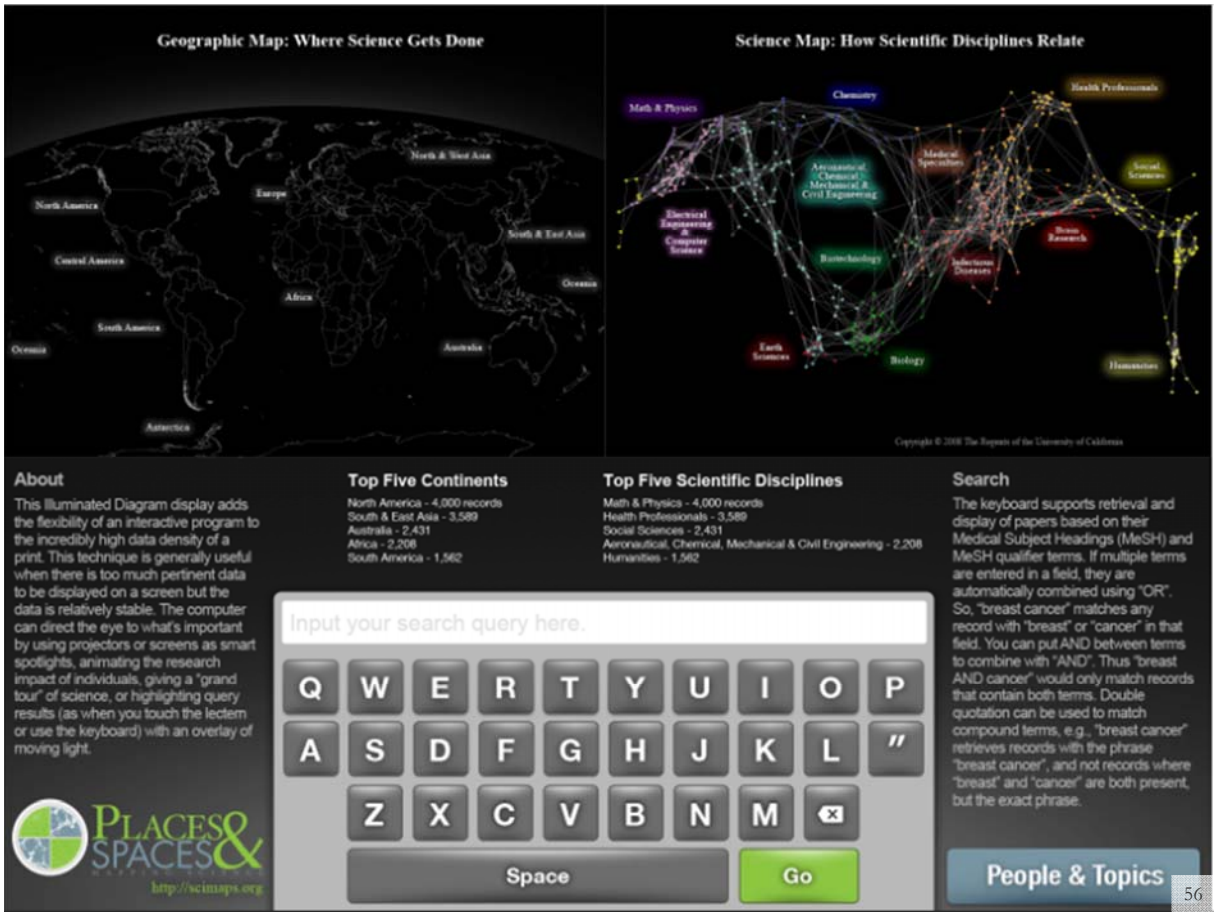
Select any location on the *Geographic Map* location (by brushing your finger over an area on the lectern's touch screen) and topics studied in that area will highlight on the *Science Map*: the brighter a topic glows, the more papers on that topic originated in the selected area. Conversely, touching 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 Noble laureates and particular lines of research using MEDLINE data from 2000-2009.

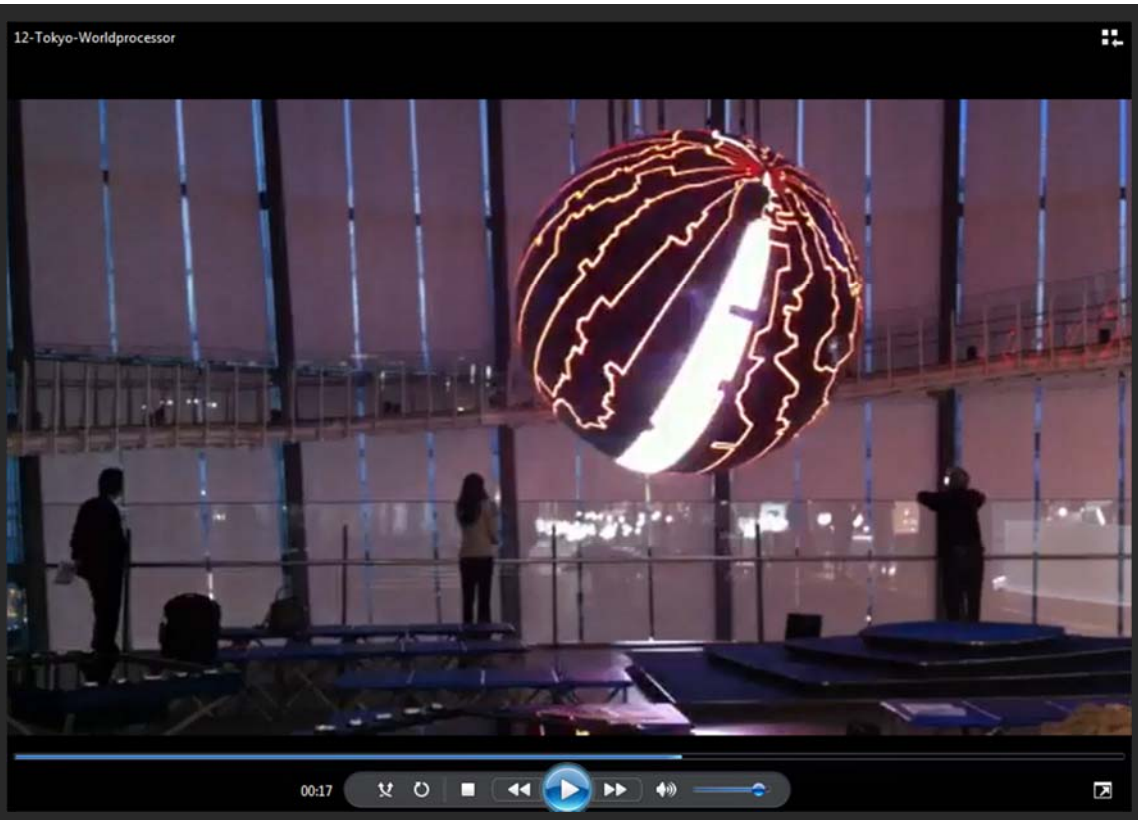
PLACES & SPACES
http://scimaps.org

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


Geographic Map: Where Science Gets Done
North America
Europe
North & West Asia
South & East Asia
Central America
Africa
Oceania
South America
Australia
Antarctica
Science Map: How Scientific Disciplines Relate
Math & Physics
Chemistry
Health Professionals
Medical Specialties
Social Sciences
Aeronautical, Chemical, Mechanical & Civil Engineering
Electrical Engineering & Computer Science
Brain Research
Biotechnology
Infectious Diseases
Earth Sciences
Biology
Humanities
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.
PLACES & SPACES
http://scimaps.org
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.
Cancer
Cloning
HIV
Robert G. Edwards
Roger D. Kornberg
Elinor Ostrom
Obesity
Quality of Life
Smoking
Stanley B. Prusiner
Ahmed H. Zewail
View All
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 Noble laureates and particular lines of research using MEDLINE data from 2000-2009.
Keyword Search


Geographic Map: Where Science Gets Done
North America
Europe
North & West Asia
South & East Asia
Central America
Africa
Oceania
South America
Australia
Antarctica
Science Map: How Scientific Disciplines Relate
Math & Physics
Chemistry
Health Professionals
Medical Specialties
Social Sciences
Aeronautical, Chemical, Mechanical & Civil Engineering
Electrical Engineering & Computer Science
Brain Research
Biotechnology
Infectious Diseases
Earth Sciences
Biology
Humanities
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.
PLACES & SPACES
http://scimaps.org
Top Five Continents
North America - 4,000 records
South & East Asia - 3,589
Australia - 2,431
Africa - 2,208
South America - 1,562
Top Five Scientific Disciplines
Math & Physics - 4,000 records
Health Professionals - 3,589
Social Sciences - 2,431
Aeronautical, Chemical, Mechanical & Civil Engineering - 2,208
Humanities - 1,562
Input your search query here.
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
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.
People & Topics



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

References

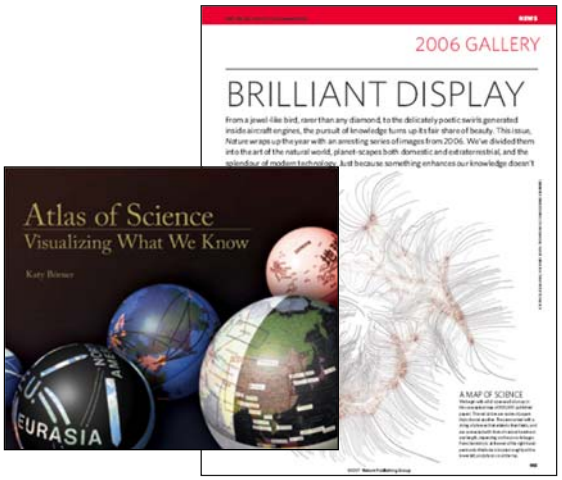
Börner, Katy, Chen, Chaomei, and Boyack, Kevin. (2003). **Visualizing Knowledge Domains.** In Blaise Cronin (Ed.), *ARIST*, Medford, NJ: Information Today, Volume 37, Chapter 5, pp. 179-255.
http://ivl.slis.indiana.edu/km/pub/2003-borner-arist.pdf

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/vol101/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.

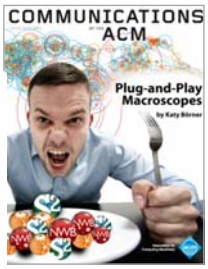


Overview

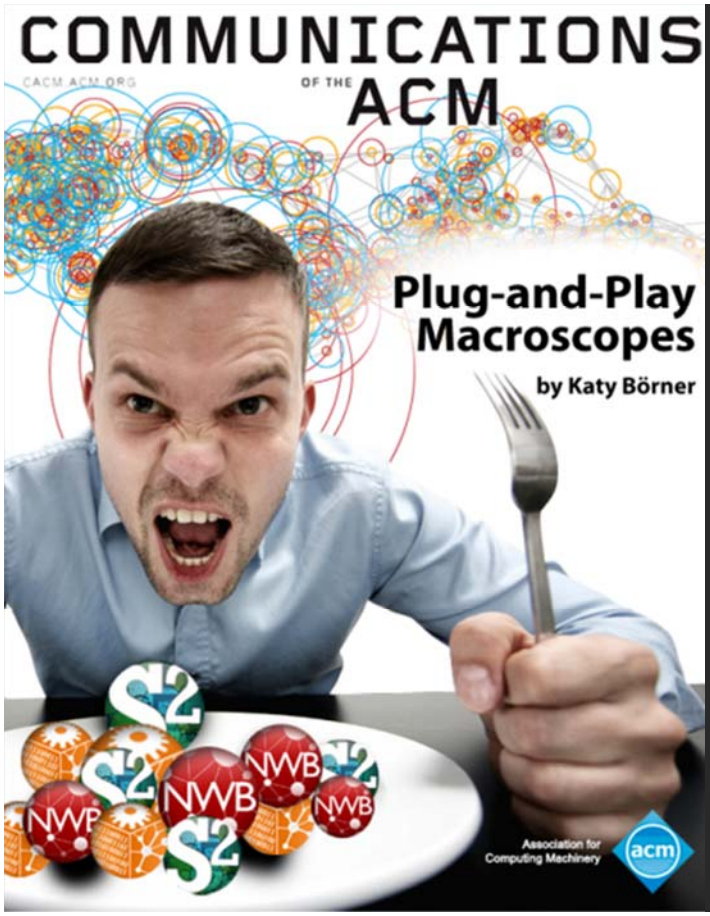
1. **Motivation:** Design informative and visually pleasing visualizations that make a difference.

2. **Theory:** Learn from and combine approaches from psychology, cartography, computer science, information visualization, statistics, graphic design.

3. **Practice:** Sample visualizations designed for experts and a general audience + plug-and-play macroscope tools that commoditize data mining and visualization.



cyberinfrastructure for NETWORK SCIENCE CENTER
School of Library and Information Science | Indiana University Bloomington



Börner, Katy. (March 2011). Plug-and-Play Macroscopes. *Communications of the ACM, 54(3), 60-69.*

Video and paper are at
http://www.scivee.tv/node/27704

Designing "Dream Tools"

Many of the best micro-, tele-, and macroscopes are designed by **scientists keen to observe and comprehend what no one has seen or understood before.** Galileo Galilei (1564–1642) recognized the potential of a spyglass for the study of the heavens, ground and polished his own lenses, and used the improved optical instruments to make discoveries like the moons of Jupiter, providing quantitative evidence for the Copernican theory.

Today, scientists **repurpose, extend, and invent new hardware and software** to create **"macroscopes"** that may solve both local and global challenges.

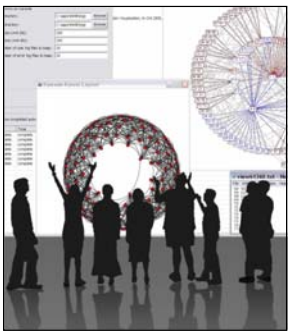
Plug-and-play macroscopes **empower** me, my students, colleagues, and 100,000 others that downloaded them.

Macroscopes

Decision making in science, industry, and politics, as well as in daily life, requires that we make sense of data sets representing the structure and dynamics of complex systems. Analysis, navigation, and management of these continuously evolving data sets require a new kind of data-analysis and visualization tool we call a macroscope (from the Greek macros, or "great," and skopein, or "to observe") inspired by de Rosnay's futurist science writings. Macroscopes provide a "vision of the whole," helping us "synthesize" the related elements and enabling us to detect patterns, trends, and outliers while granting access to myriad details. Rather than make things larger or smaller, **macroscopes let us observe what is at once too great, slow, or complex for the human eye and mind to notice and comprehend.**

Microscopes

Telescopes



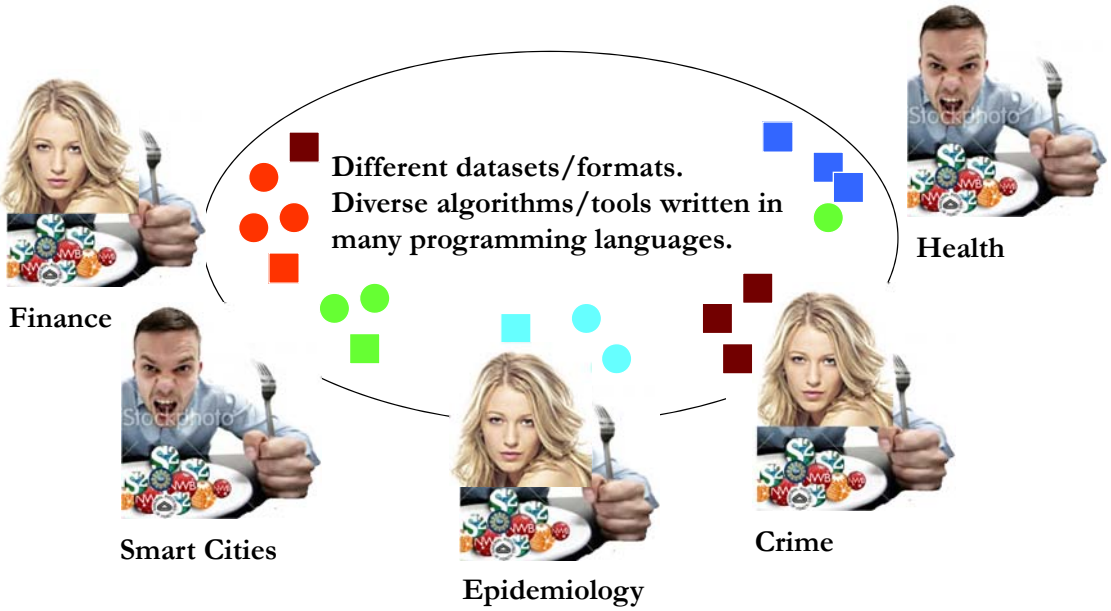
Macroscopes

Macroscopes cont.

While microscopes and telescopes are physical instruments, **macroscopes resemble continuously changing bundles of software plug-ins.** Macroscopes make it easy to select and combine algorithm and tool plug-ins but also interface plug-ins, workflow support, logging, scheduling, and other plug-ins needed for scientifically rigorous yet effective work.

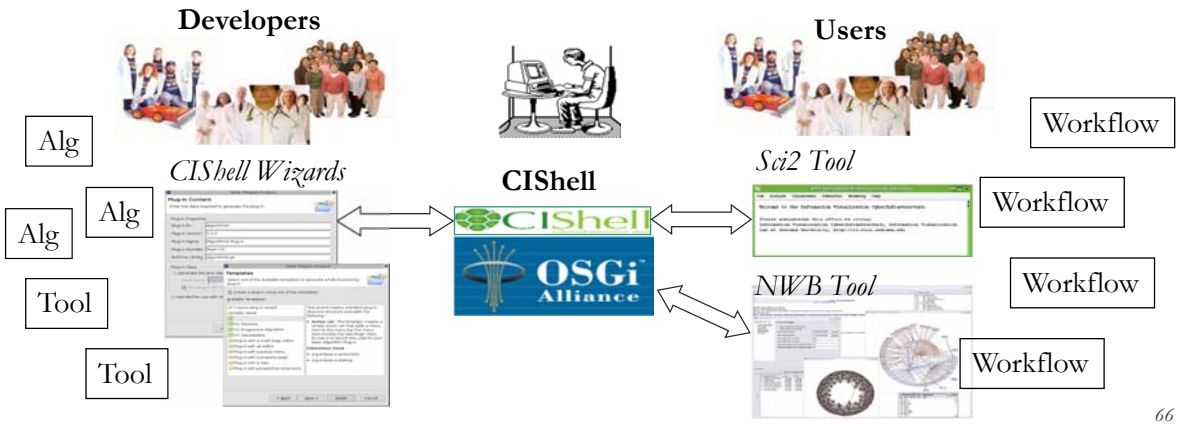
They make it easy to share plug-ins via email, flash drives, or online. To use new plugins, simply copy the files into the plug-in directory, and they appear in the tool menu ready for use. No restart of the tool is necessary. **Sharing algorithm components, tools, or novel interfaces becomes as easy as sharing images on Flickr or videos on YouTube. Assembling custom tools is as quick as compiling your custom music collection.**

Macroscopes Serve International, Interdisciplinary Scholars

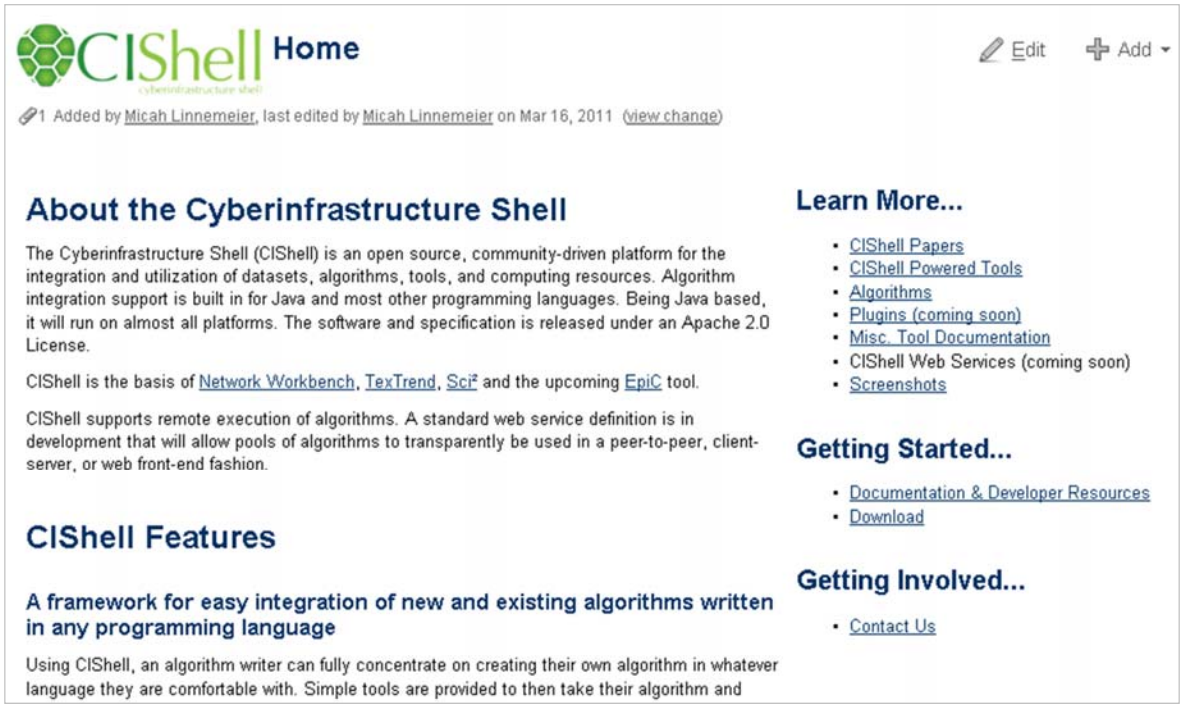


OSGi & CIShell

- CIShell (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, CIShell provides "sockets" into which existing and new datasets, algorithms, and tools can be plugged using a wizard-driven process.



CIShell Developer Guide
(http://cishell.wiki.cns.iu.edu)



CIShell Portal (http://cishell.org)

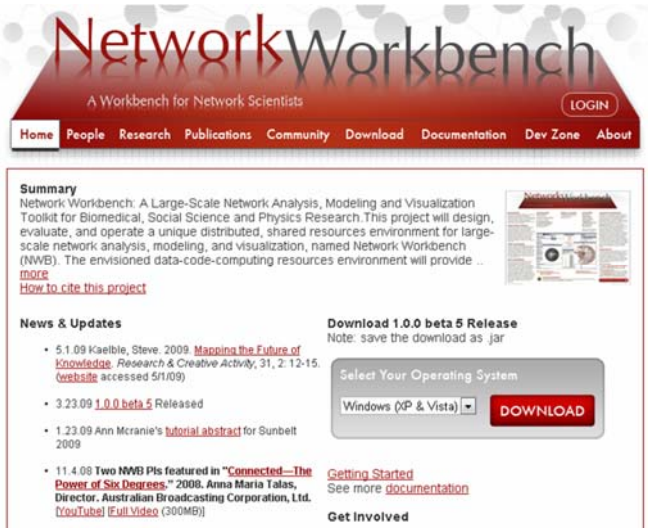
Network Workbench Tool

http://nwb.cns.edu

The Network Workbench (NWB) tool supports researchers, educators, and practitioners interested in the study of biomedical, social and behavioral science, physics, and other networks.

In February 2009, the tool provides more 169 plugins that support the preprocessing, analysis, modeling, and visualization of networks.

It has been downloaded more than 110,000 times since December 2006.



Herr II, Bruce W., Huang, Weixia (Bonnie), Penumarthy, Shashikant & Börner, Katy. (2007). Designing Highly Flexible and Usable Cyberinfrastructures for Convergence. In Bainbridge, William S. & Roco, Mihail C. (Eds.), Progress in Convergence - Technologies for Human Wellbeing (Vol. 1093, pp. 161-179), Annals of the New York Academy of Sciences, Boston, MA.

Computational Proteomics

What relationships exist between protein targets of all drugs and all disease-gene products in the human protein–protein interaction network?

Yildrüm, Muhammed A., Kwan-Il Goh, Michael E. Cusick, Albert-László Barabási, and Marc Vidal. (2007). Drug-target Network. Nature Biotechnology 25 no. 10: 1119-1126.

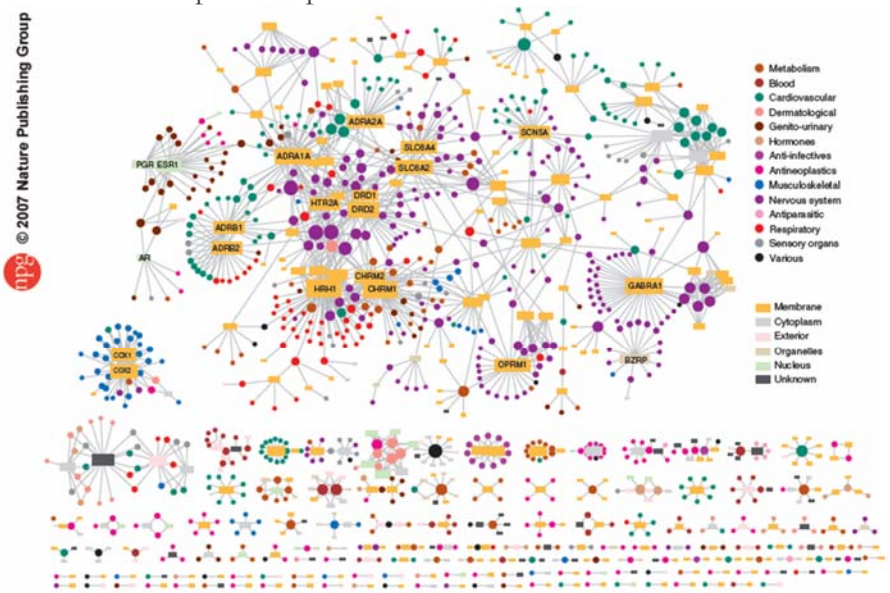


Figure 2 Drug–target network (DT network). The DT network is generated by using the known associations between FDA-approved drugs and their target proteins. Circles and rectangles correspond to drugs and target proteins, respectively. A link is placed between a drug node and a target node if the protein is a known target of that drug. The area of the drug (protein) node is proportional to the number of targets that the drug has (the number of drugs targeting the protein). Color codes are given in the legend. Drug nodes (circles) are colored according to their Anatomical Therapeutic Chemical Classification, and the target proteins (rectangular boxes) are colored according to their cellular component obtained from the Gene Ontology database.

Computational Economics
Does the type of product that a country exports matter for subsequent economic performance?

C. A. Hidalgo, B. Klinger, A.-L. Barabási, R. Hausmann (2007) The Product Space Conditions the Development of Nations. Science 317, 482 (2007).

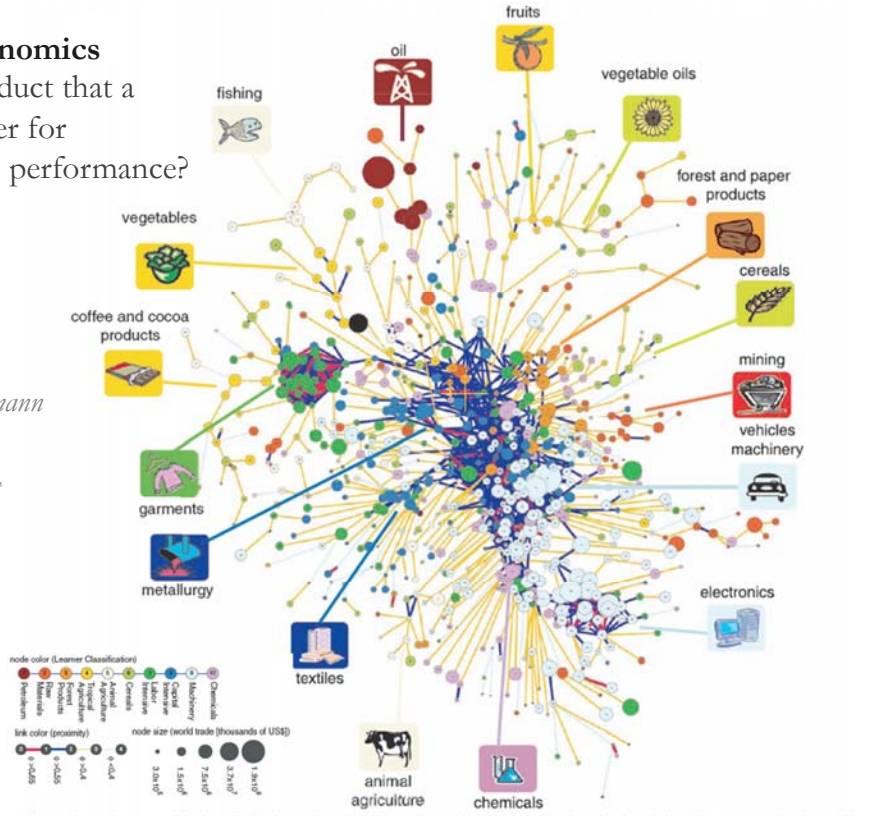


Fig. 1. The product space. **(A)** Hierarchically clustered proximity (φ) matrix representing the 775 SITC-4 product classes exported in the 1998–2000 period. **(B)** Network representation of the product space. Links are color coded with their proximity value. The sizes of the nodes are proportional to world trade, and their colors are chosen according to the classification introduced by Leamer.

Computational Social Science
Studying large scale social networks such as Wikipedia

Second Sight: An Emergent Mosaic of Wikipedian Activity, The NewScientist, May 19, 2007

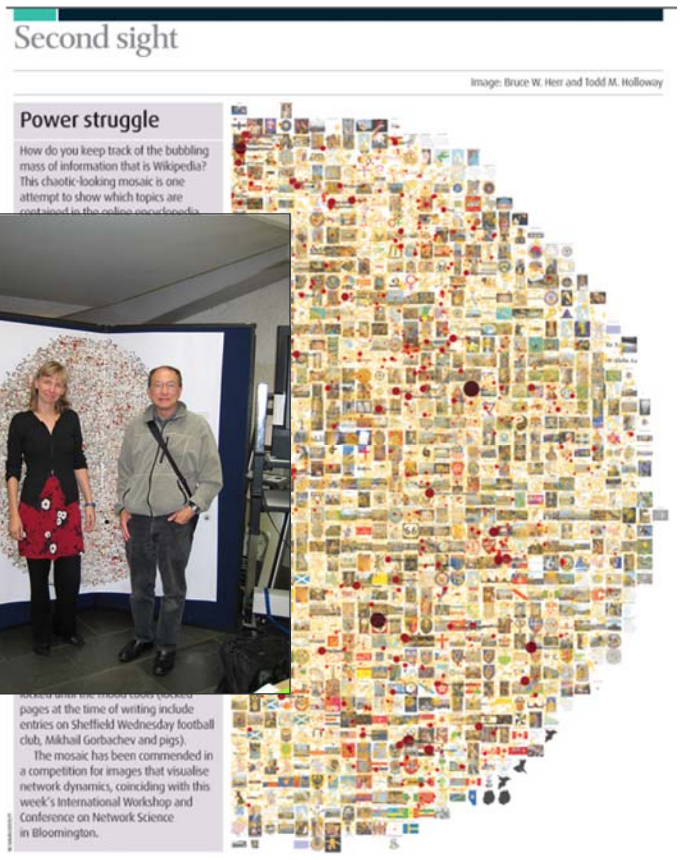

Second sight

Image: Bruce W. Herr and Todd M. Holloway

Power struggle

How do you keep track of the bubbling mass of information that is Wikipedia? This chaotic-looking mosaic is one attempt to show which topics are

pages at the time of writing include entries on Sheffield Wednesday football club, Mikhail Gorbachev and pigs).

The mosaic has been commended in a competition for images that visualise network dynamics, coinciding with this week's International Workshop and Conference on Network Science in Bloomington.

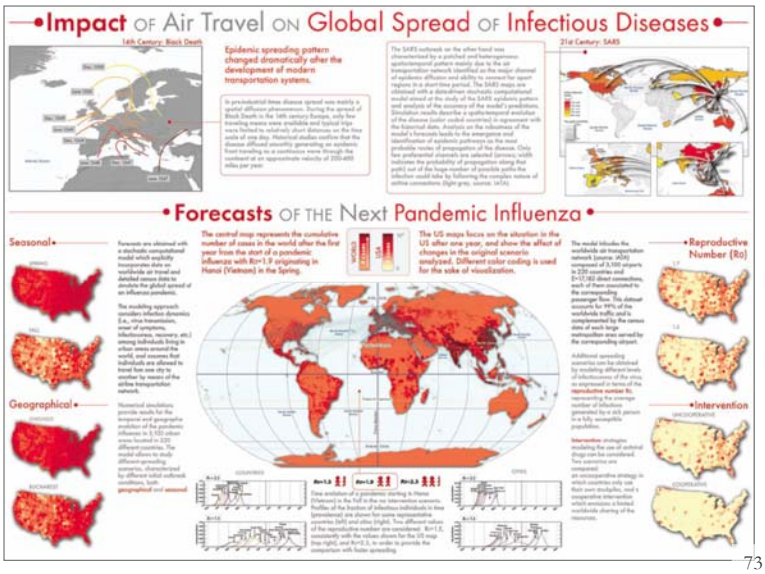
www.newscientist.com

Computational Epidemics

Forecasting (and preventing the effects of) the next pandemic.

Epidemic Modeling in Complex realities, V. Colizza, A. Barrat, M. Barthelemy, A.Vespignani, Comptes Rendus Biologie, 330, 364-374 (2007).
Reaction-diffusion processes and metapopulation models in heterogeneous networks, V.Colizza, R. Pastor-Satorras, A.Vespignani, Nature Physics 3, 276-282 (2007).

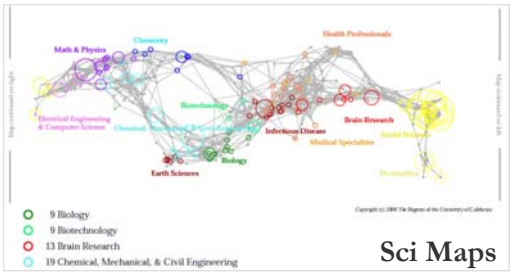
Modeling the Worldwide Spread of Pandemic Influenza: Baseline Case and Containment Interventions, V. Colizza, A. Barrat, M. Barthelemy, A.-J. Valleron, A.Vespignani, PloS-Medicine 4, e13, 95-110 (2007).



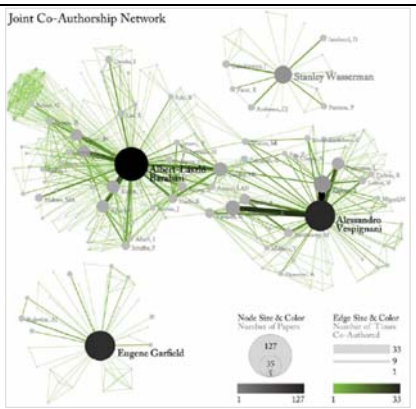
Sci² Tool – "Open Code for S&T Assessment"

Users come from 40+ countries

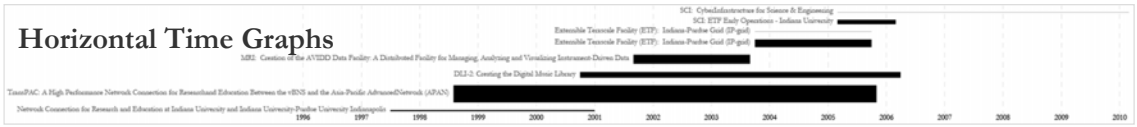
OSGi/CIShell powered tool with NWB plugins and many new scientometrics and visualizations plugins.



Sci Maps

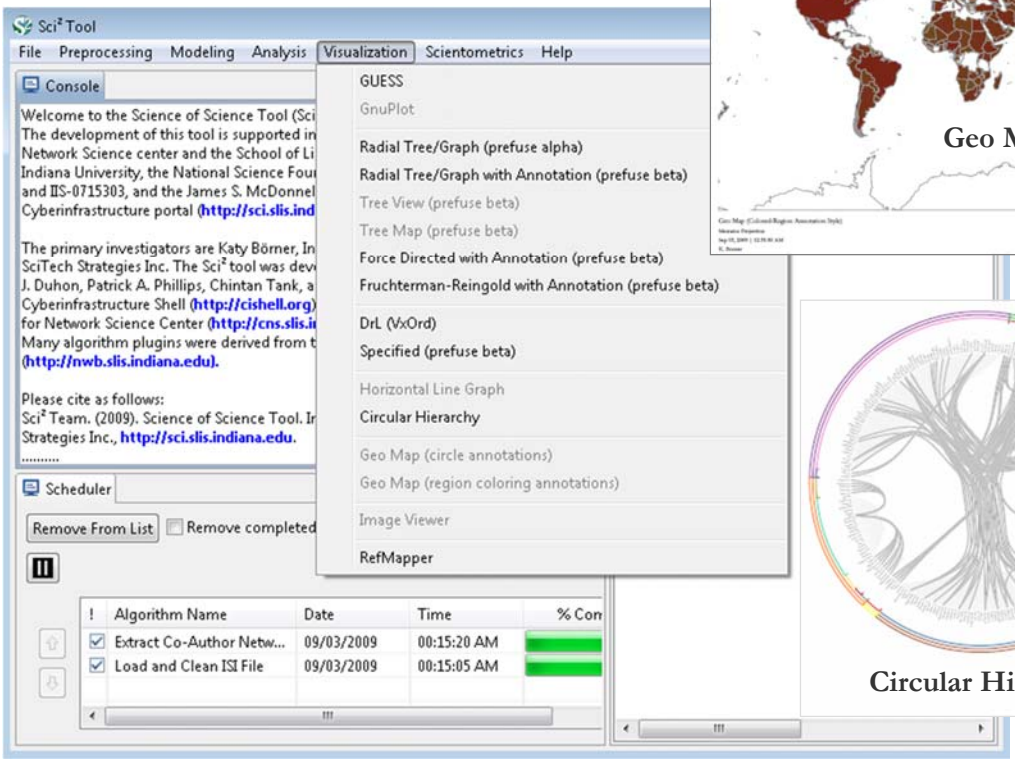


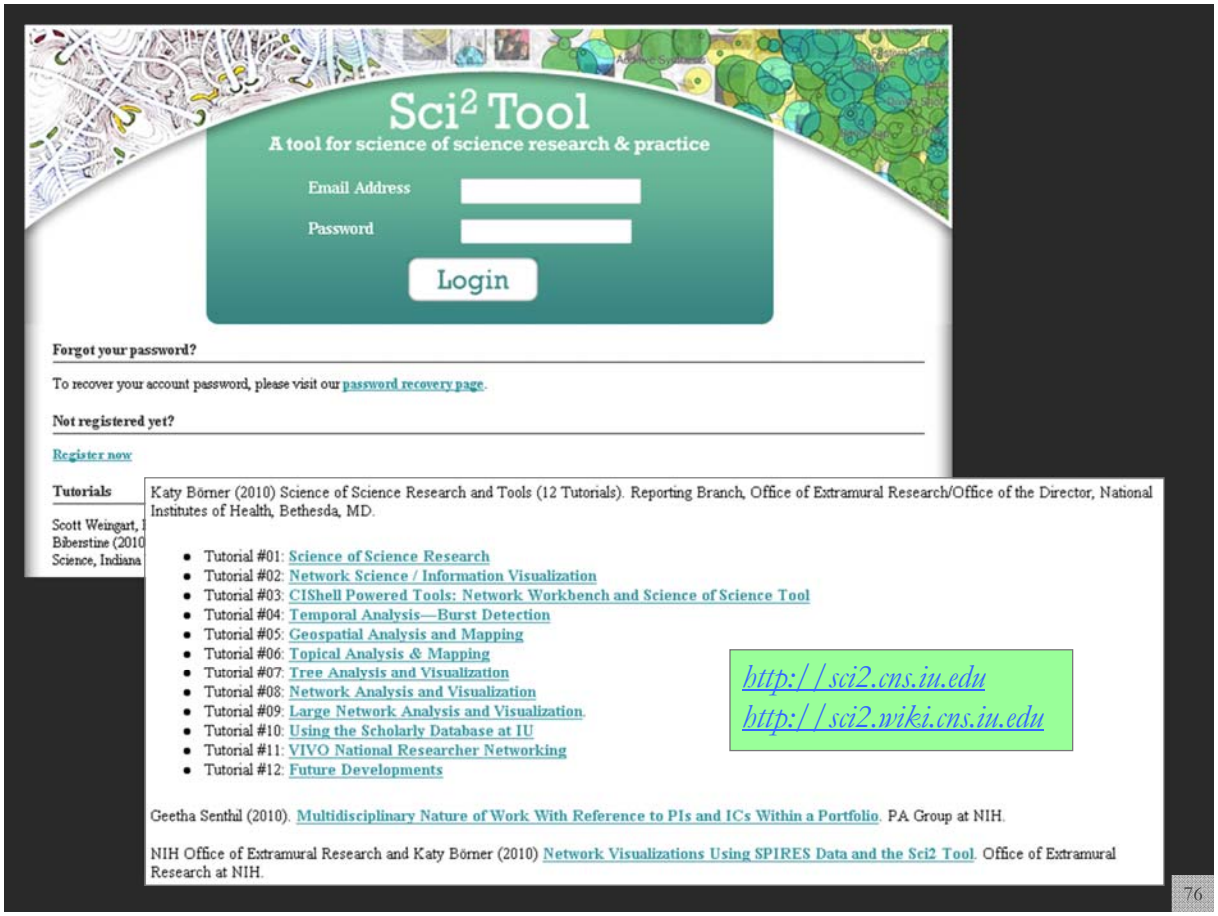
GUESS Network Vis

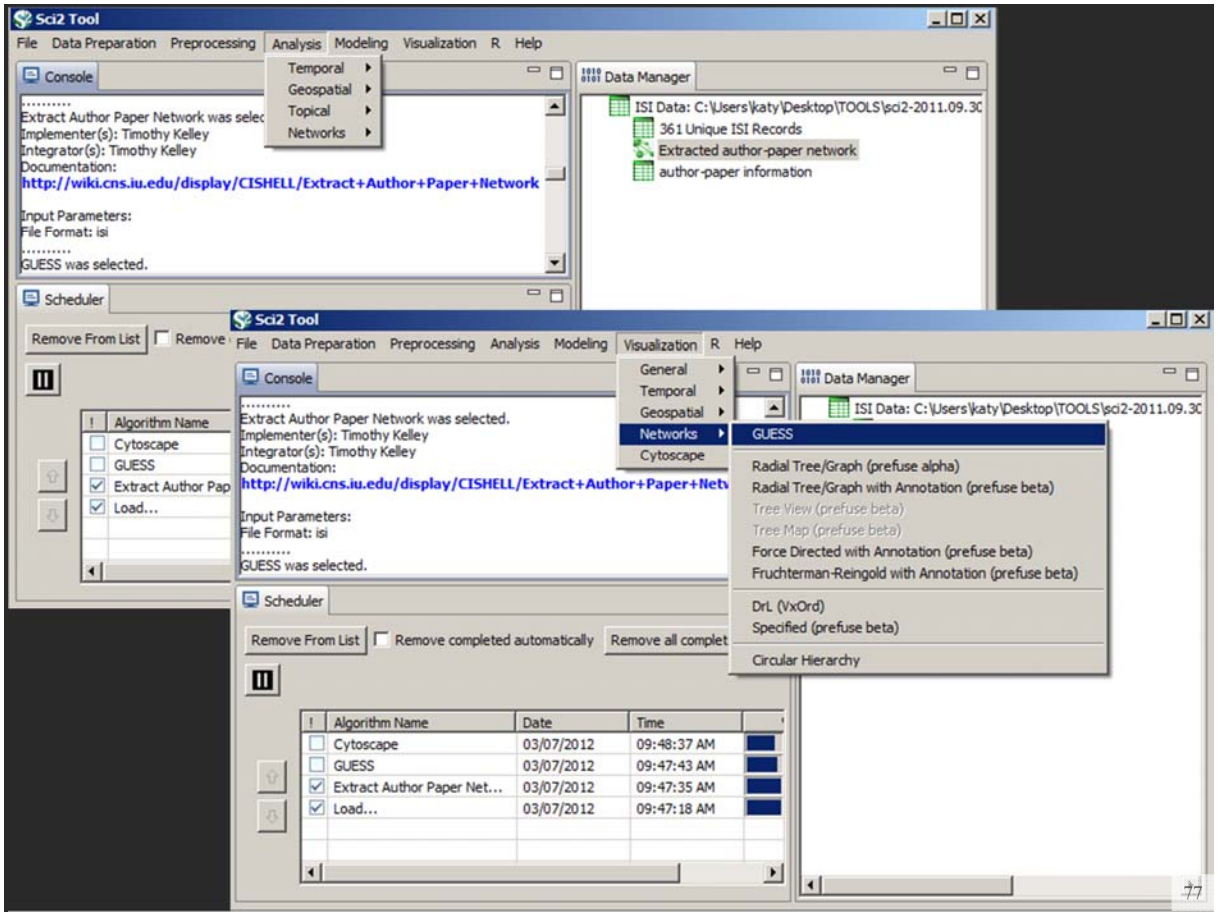


Börner, Katy, Huang, Weixia (Bonnie), Linnemeier, Micah, Duhon, Russell Jackson, Phillips, Patrick, Ma, Nianli, Zoss, Angela, Guo, Hanning & Price, Mark. (2009). Rete-Netzwerk-Red: Analyzing and Visualizing Scholarly Networks Using the Scholarly Database and the Network Workbench Tool. Proceedings of ISSI 2009: 12th International Conference on Scientometrics and Informetrics, Rio de Janeiro, Brazil, July 14-17 . Vol. 2, pp. 619-630.

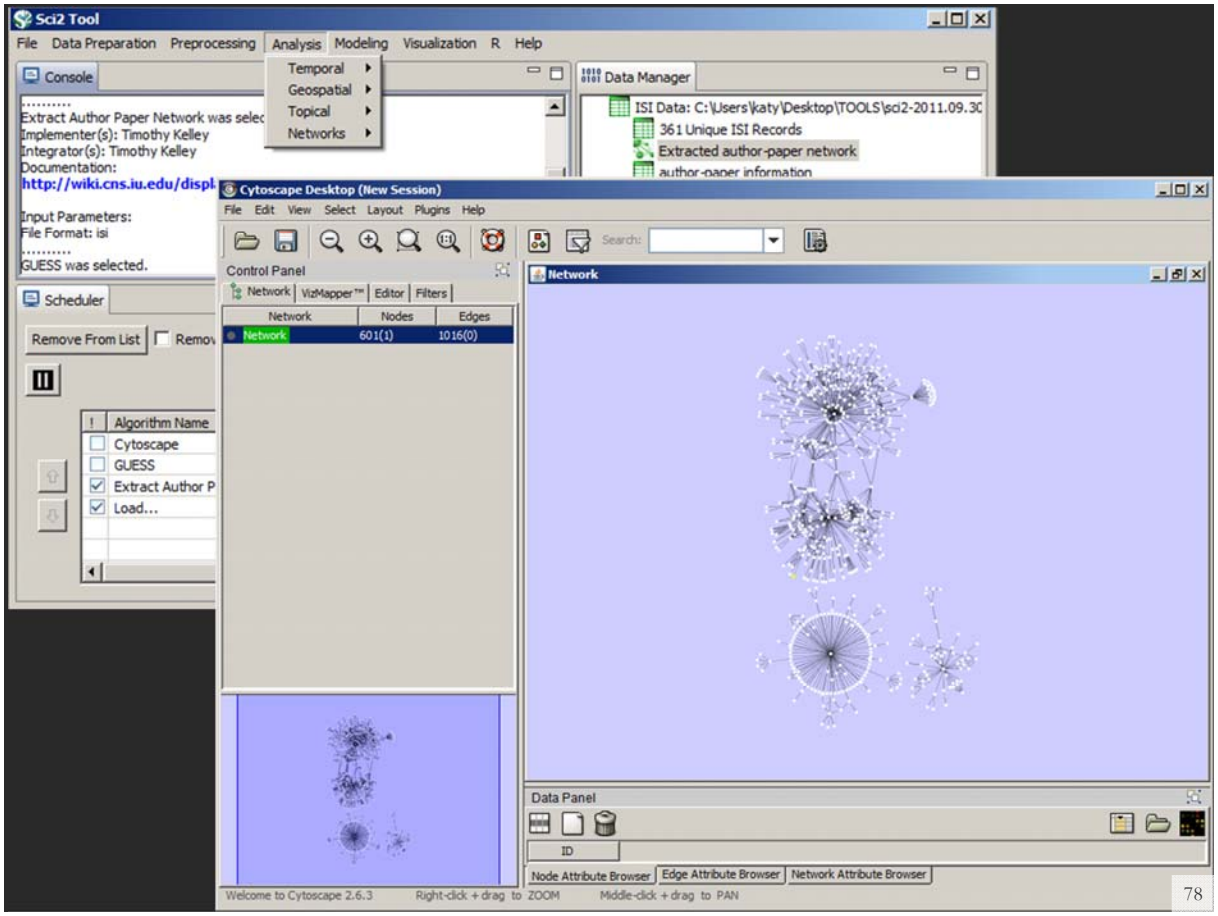
Sci² Tool Vis cont.



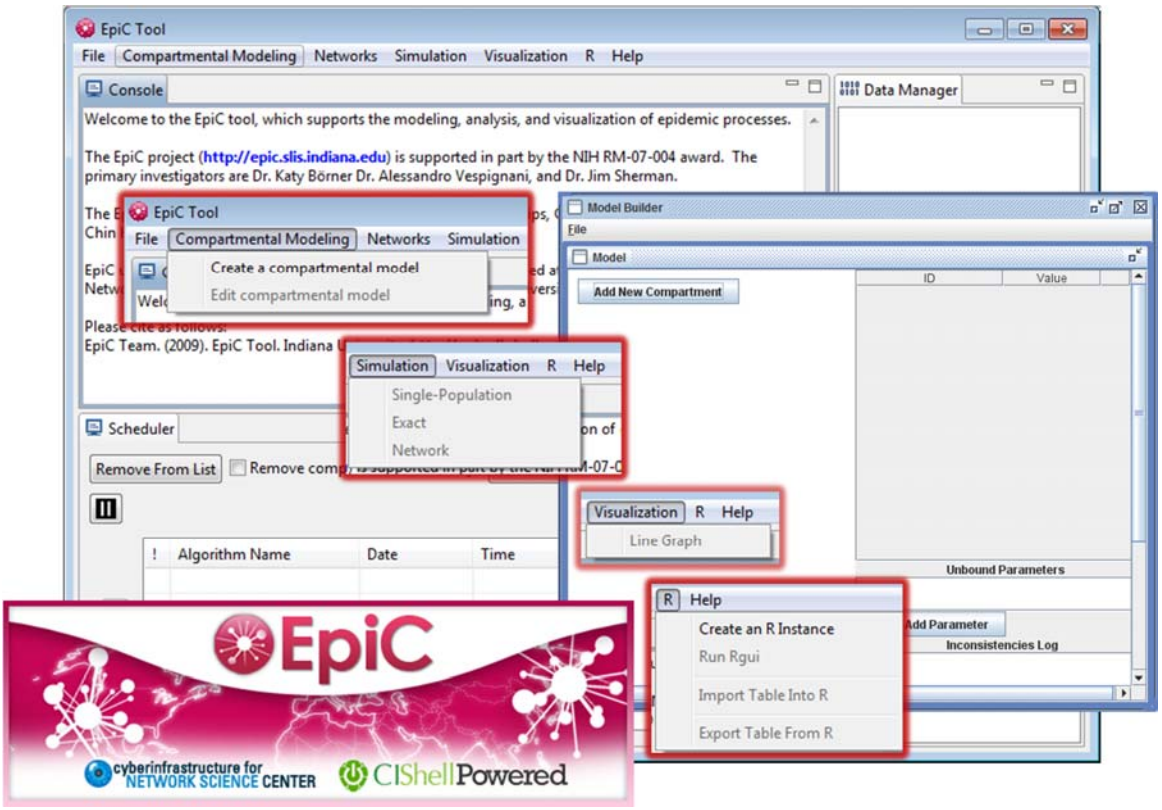



Sci2 Tool
File Data Preparation Preprocessing Analysis Modeling Visualization R Help
Temporal
Geospatial
Topical
Networks
Console
Extract Author Paper Network was selec
Implementer(s): Timothy Kelley
Integrator(s): Timothy Kelley
Documentation:
http://wiki.cns.iu.edu/display/CISHELL/Extract+Author+Paper+Network
Input Parameters:
File Format: isi
GUESS was selected.
Data Manager
ISI Data: C:\Users\katy\Desktop\TOOLS\sci2-2011.09.30
361 Unique ISI Records
Extracted author-paper network
author-paper information
Scheduler
Remove From List
Remove
Algorithm Name
Cytoscape
GUESS
Extract Author Pap
Load...
Sci2 Tool
File Data Preparation Preprocessing Analysis Modeling Visualization R Help
General
Temporal
Geospatial
Networks
Cytoscape
GUESS
Radial Tree/Graph (prefuse alpha)
Radial Tree/Graph with Annotation (prefuse beta)
Tree View (prefuse beta)
Tree Map (prefuse beta)
Force Directed with Annotation (prefuse beta)
Fruchterman-Reingold with Annotation (prefuse beta)
DrL (VxOrd)
Specified (prefuse beta)
Circular Hierarchy
Console
Extract Author Paper Network was selected.
Implementer(s): Timothy Kelley
Integrator(s): Timothy Kelley
Documentation:
http://wiki.cns.iu.edu/display/CISHELL/Extract+Author+Paper+Netv
Input Parameters:
File Format: isi
GUESS was selected.
Data Manager
ISI Data: C:\Users\katy\Desktop\TOOLS\sci2-2011.09.30
Scheduler
Remove From List
Remove completed automatically
Remove all complet
Algorithm Name
Date
Time
Cytoscape
03/07/2012
09:48:37 AM
GUESS
03/07/2012
09:47:43 AM
Extract Author Paper Net...
03/07/2012
09:47:35 AM
Load...
03/07/2012
09:47:18 AM

77


Sci2 Tool
File Data Preparation Preprocessing Analysis Modeling Visualization R Help
Temporal
Geospatial
Topical
Networks
Console
Extract Author Paper Network was selec
Implementer(s): Timothy Kelley
Integrator(s): Timothy Kelley
Documentation:
http://wiki.cns.iu.edu/displ
Input Parameters:
File Format: isi
GUESS was selected.
Data Manager
ISI Data: C:\Users\katy\Desktop\TOOLS\sci2-2011.09.30
361 Unique ISI Records
Extracted author-paper network
author-paper information
Cytoscape Desktop (New Session)
File Edit View Select Layout Plugins Help
Search:
Control Panel
Network
VizMapper™
Editor
Filters
Network
Nodes
Edges
Network
601(1)
1016(0)
Network
Scheduler
Remove From List
Remo
Algorithm Name
Cytoscape
GUESS
Extract Author P
Load...
Data Panel
ID
Node Attribute Browser
Edge Attribute Browser
Network Attribute Browser
Welcome to Cytoscape 2.6.3
Right-click + drag to ZOOM
Middle-click + drag to PAN

78



OSGi/CIShell Adoption

A number of other projects recently adopted OSGi and/or CIShell:

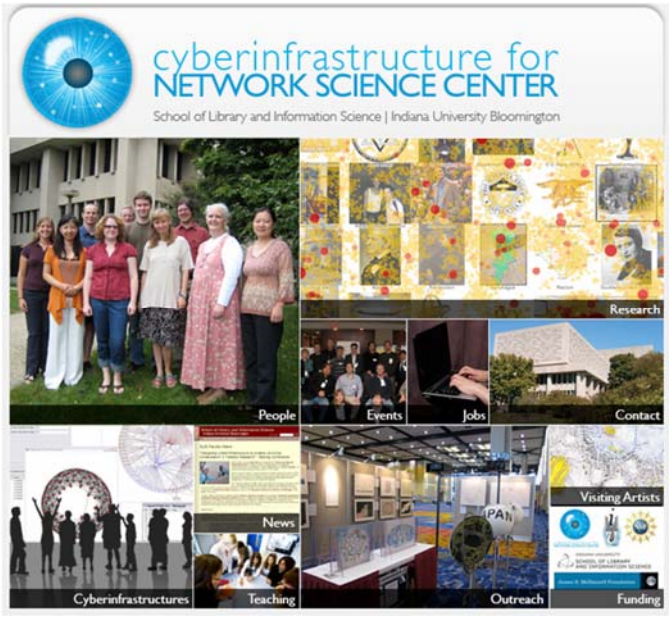
USA

- *Cytoscape (http://cytoscape.org)* Led by Trey Ideker at the University of California, San Diego is an open source bioinformatics software platform for visualizing molecular interaction networks and integrating these interactions with gene expression profiles and other state data (Shannon et al., 2002).
- *MAEviz (https://wiki.ncsa.uiuc.edu/display/MAE/Home)* Managed by Jong Lee at NCSA is an open-source, extensible software platform which supports seismic risk assessment based on the Mid-America Earthquake (MAE) Center research.

Europe

- *Taverna Workbench (http://taverna.org.uk)* Developed by the myGrid team (http://mygrid.org.uk) led by Carol Goble at the University of Manchester, U.K. is a free software tool for designing and executing workflows (Hull et al., 2006). Taverna allows users to integrate many different software tools, including over 30,000 web services.
- *TEXTrend (http://textrend.org)* Led by George Kampis at Eötvös Loránd University, Budapest, Hungary supports natural language processing (NLP), classification/mining, and graph algorithms for the analysis of business and governmental text corpuses with an inherently temporal component.
- *DynaNets (http://www.dynanets.org)* Coordinated by Peter M.A. Sloot at the University of Amsterdam, The Netherlands develops algorithms to study evolving networks.
- *SISOB (http://sisob.lcc.uma.es)* An Observatory for Science in Society Based in Social Models.

As the functionality of OSGi-based software frameworks improves and the number and diversity of dataset and algorithm plugins increases, the capabilities of custom tools will expand.



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