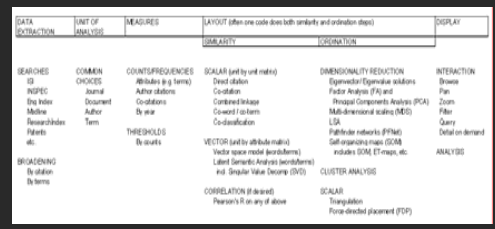
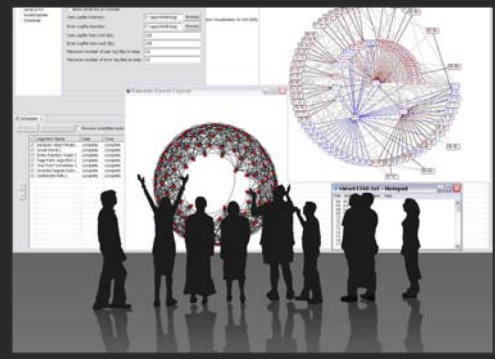
Cybertools that Support the Study of Science

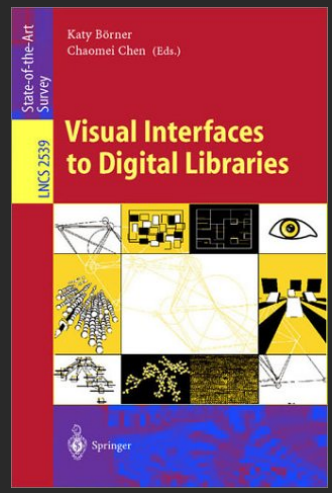


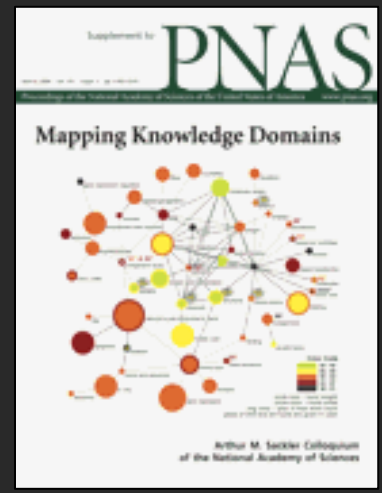


Dr. Katy Börner
Cyberinfrastructure for Network Science Center, Director
Information Visualization Laboratory, Director
School of Library and Information Science
Indiana University, Bloomington, IN
katy@indiana.edu

Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China
March 27th, 2008

Computational Scientometrics: Studying Science by Scientific Means





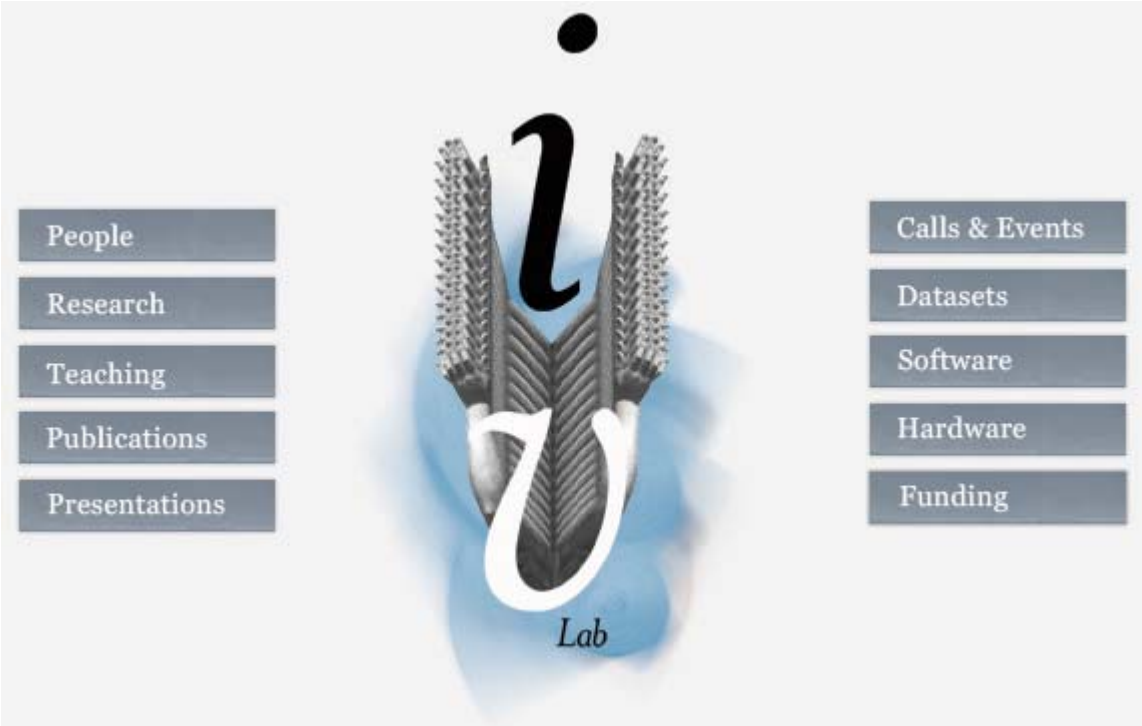
- *Börner, Katy, Chen, Chaomei, and Boyack, Kevin. (2003).* ***Visualizing Knowledge Domains.*** *In Blaise Cronin (Ed.), Annual Review of Information Science & Technology, Medford, NJ: Information Today, Inc./American Society for Information Science and Technology, Volume 37, Chapter 5, pp. 179-255.* 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.), Annual Review of Information Science & Technology, Information Today, Inc./American Society for Information Science and Technology, Medford, NJ, Volume 41, Chapter 12, pp. 537-607.* http://ivl.slis.indiana.edu/km/pub/2007-borner-arist.pdf

- ***Places & Spaces: Mapping Science*** *exhibit, see also* *http://scimaps.org*.

Data Acquisition for Comprehensive Analysis

Lab/Center Management System vs. Spacebook and MS Famulus

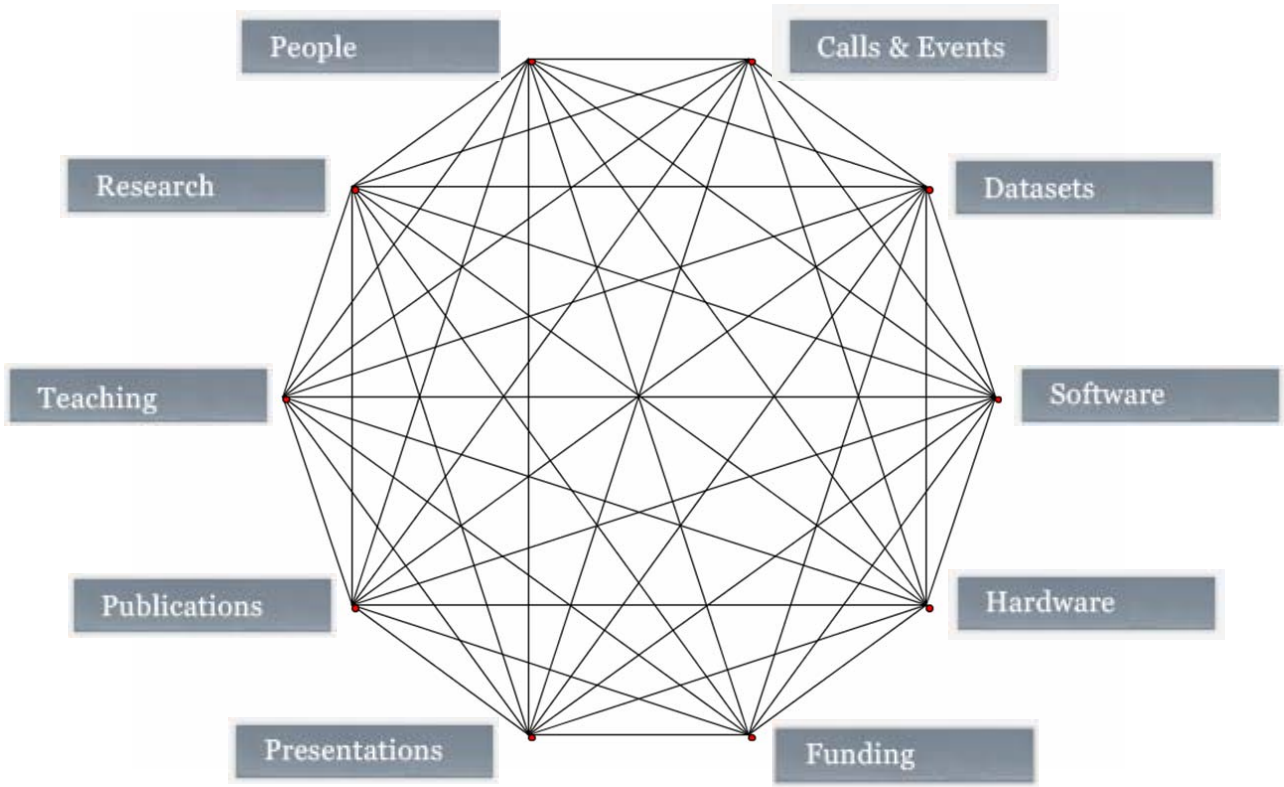
Designed to track, manage, and make use of data relevant for the daily operation of a medium size research team.



http://ivl.slis.indiana.edu

Data Entities and Interlinkages

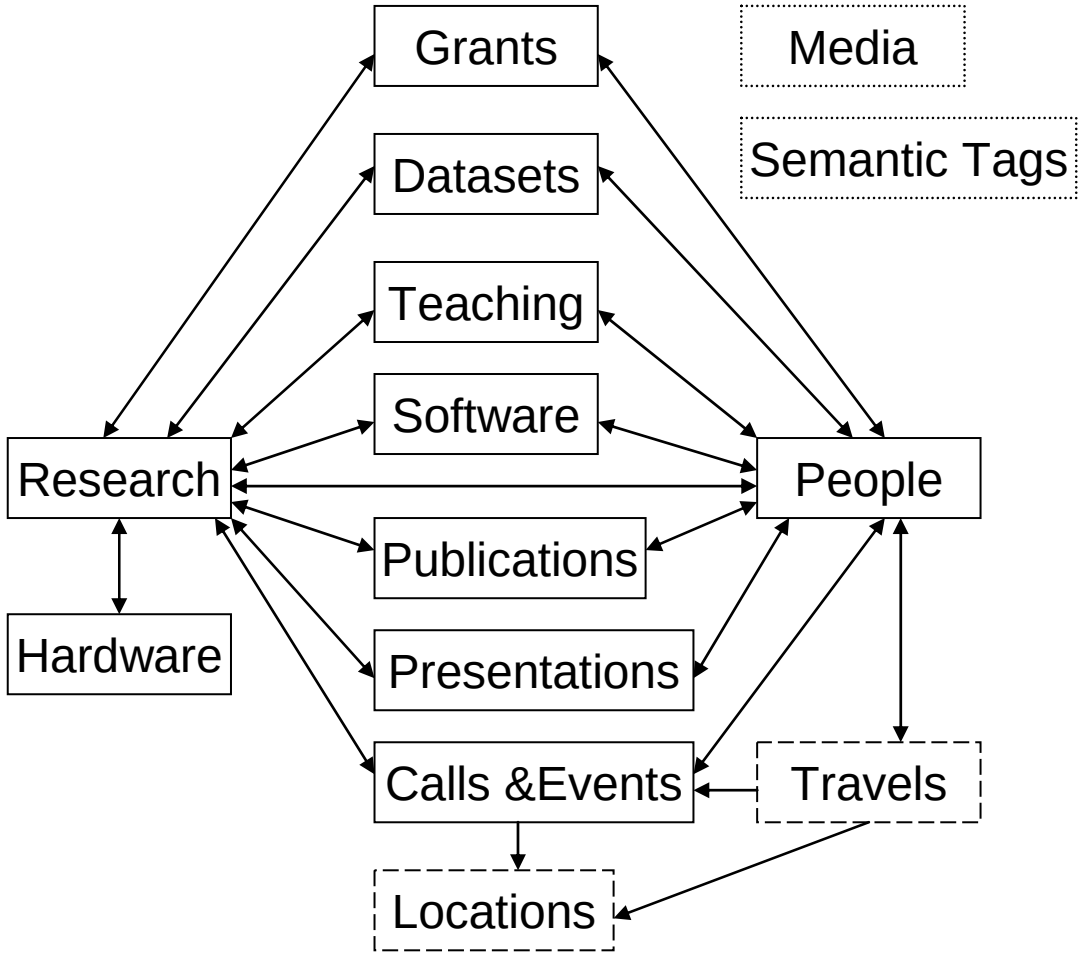
Designed for team leads, members, IT admins but also for external scholars and funding agencies.



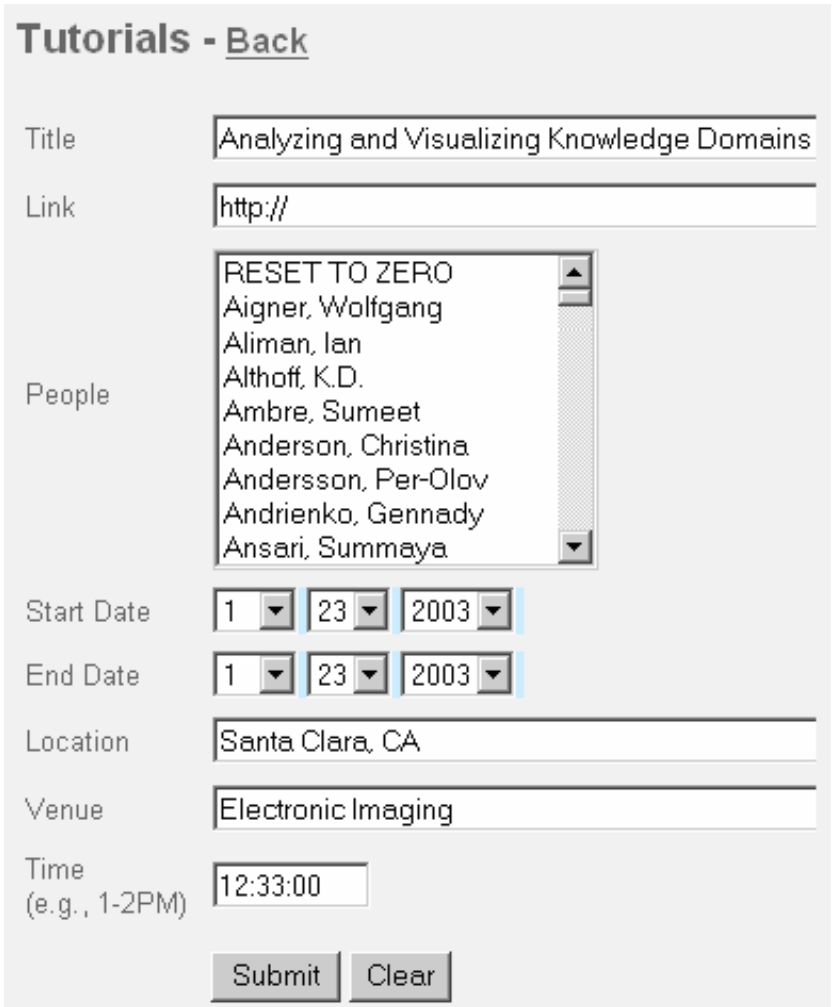
Not covered:
- Queries
- Workflows
- Protocols

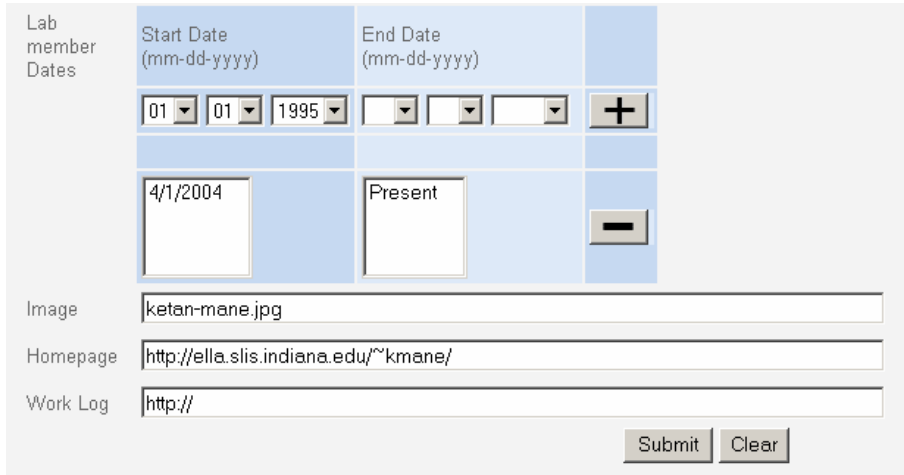
- Comments
- Bookmarks
- Ratings

Simplified representation of the IVL database schema

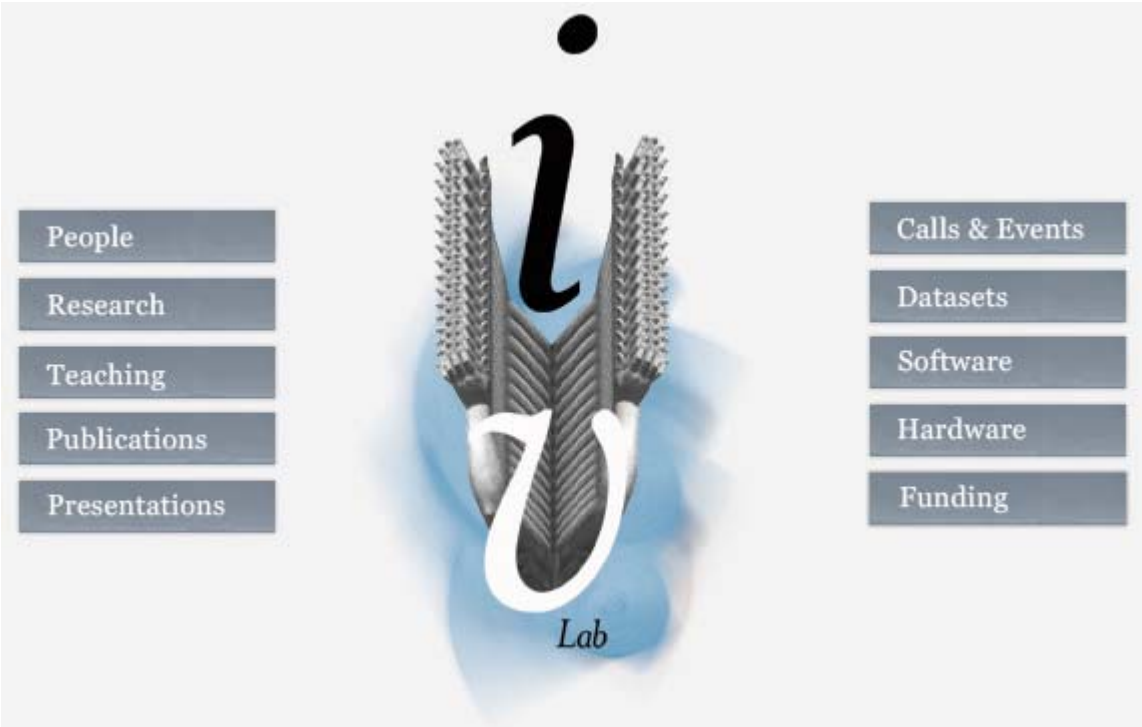


Data Entry


Tutorials - Back
Title
Analyzing and Visualizing Knowledge Domains
Link
http://
People
RESET TO ZERO
Aigner, Wolfgang
Aliman, Ian
Althoff, K.D.
Ambre, Sumeet
Anderson, Christina
Andersson, Per-Olov
Andrienko, Gennady
Ansari, Summaya
Start Date
1
23
2003
End Date
1
23
2003
Location
Santa Clara, CA
Venue
Electronic Imaging
Time
(e.g., 1-2PM)
12:33:00
Submit
Clear

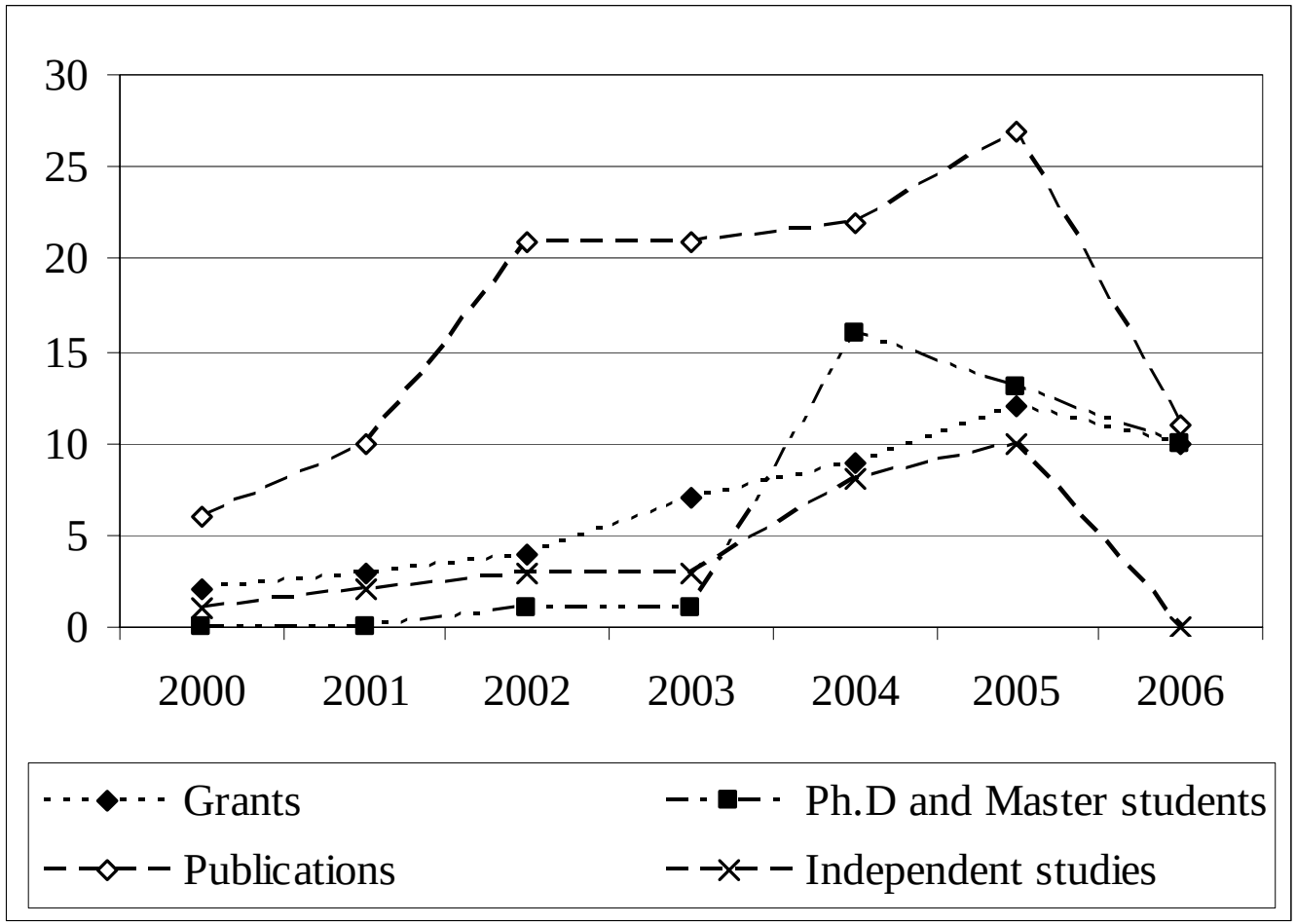

Lab member Dates
Start Date (mm-dd-yyyy)
End Date (mm-dd-yyyy)
01
01
1995
4/1/2004
Present
Image
ketan-mane.jpg
Homepage
http://ella.slis.indiana.edu/~kmane/
Work Log
http://
Submit
Clear

Demo

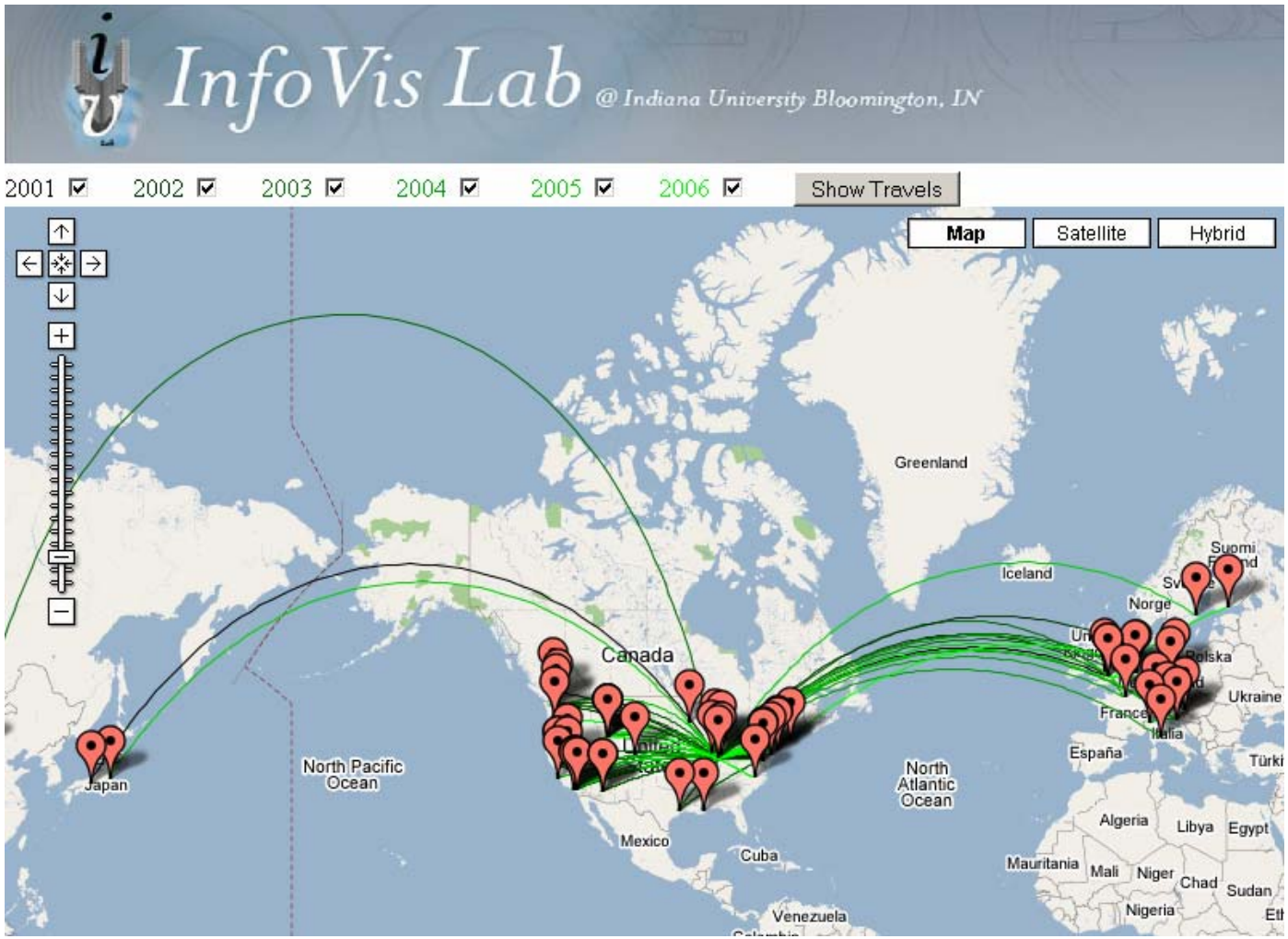


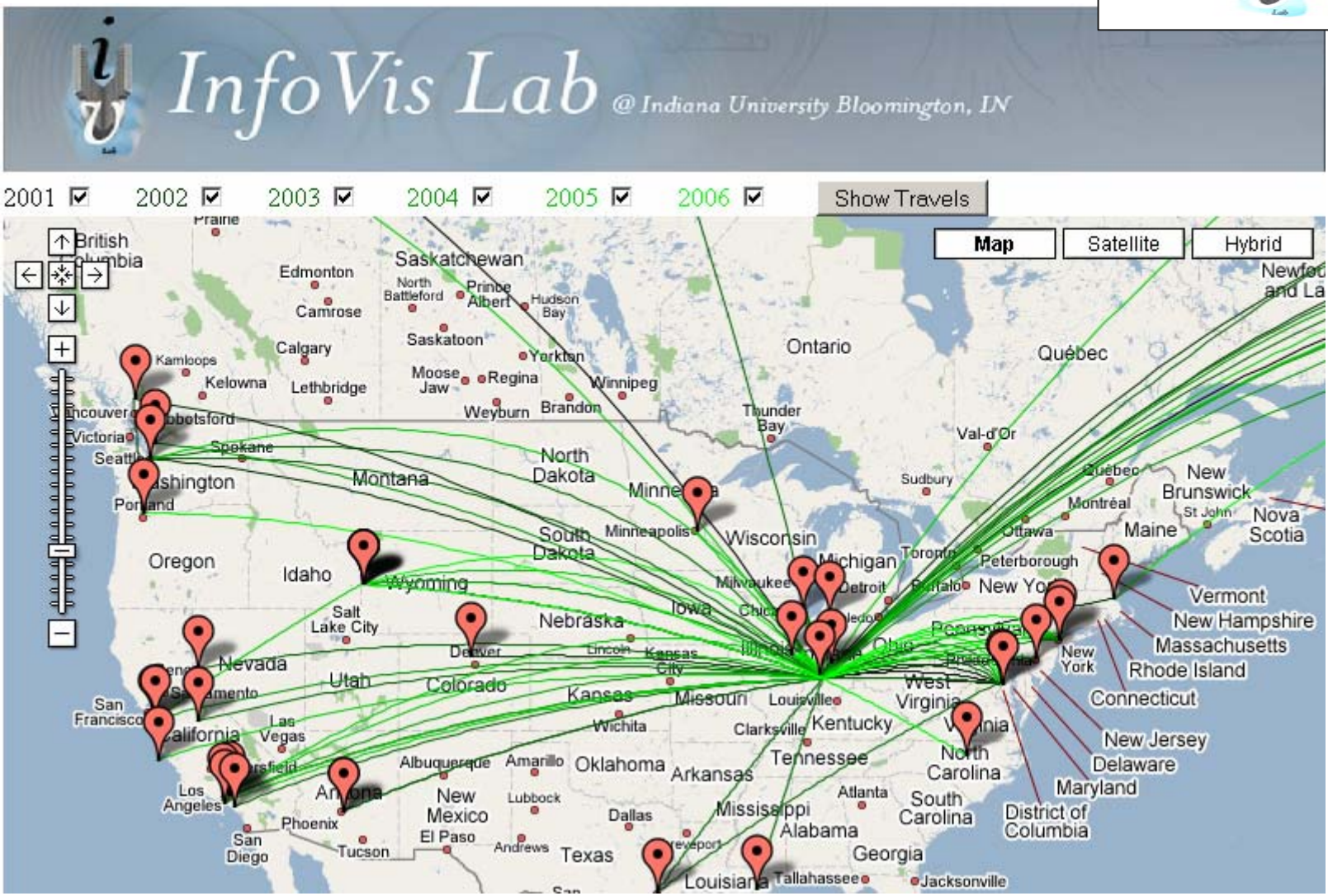
http://ivl.slis.indiana.edu

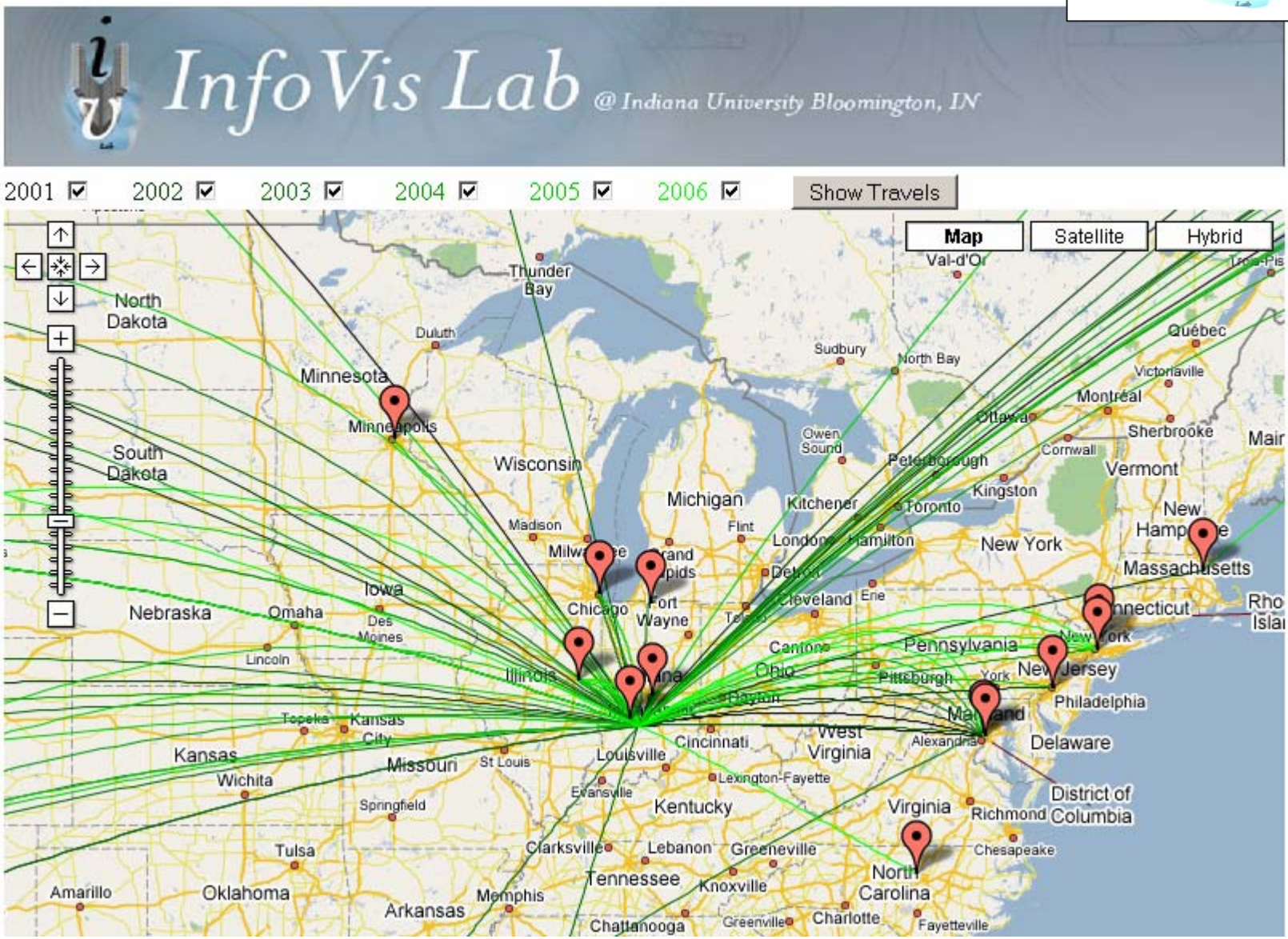
Time series analysis & visualization

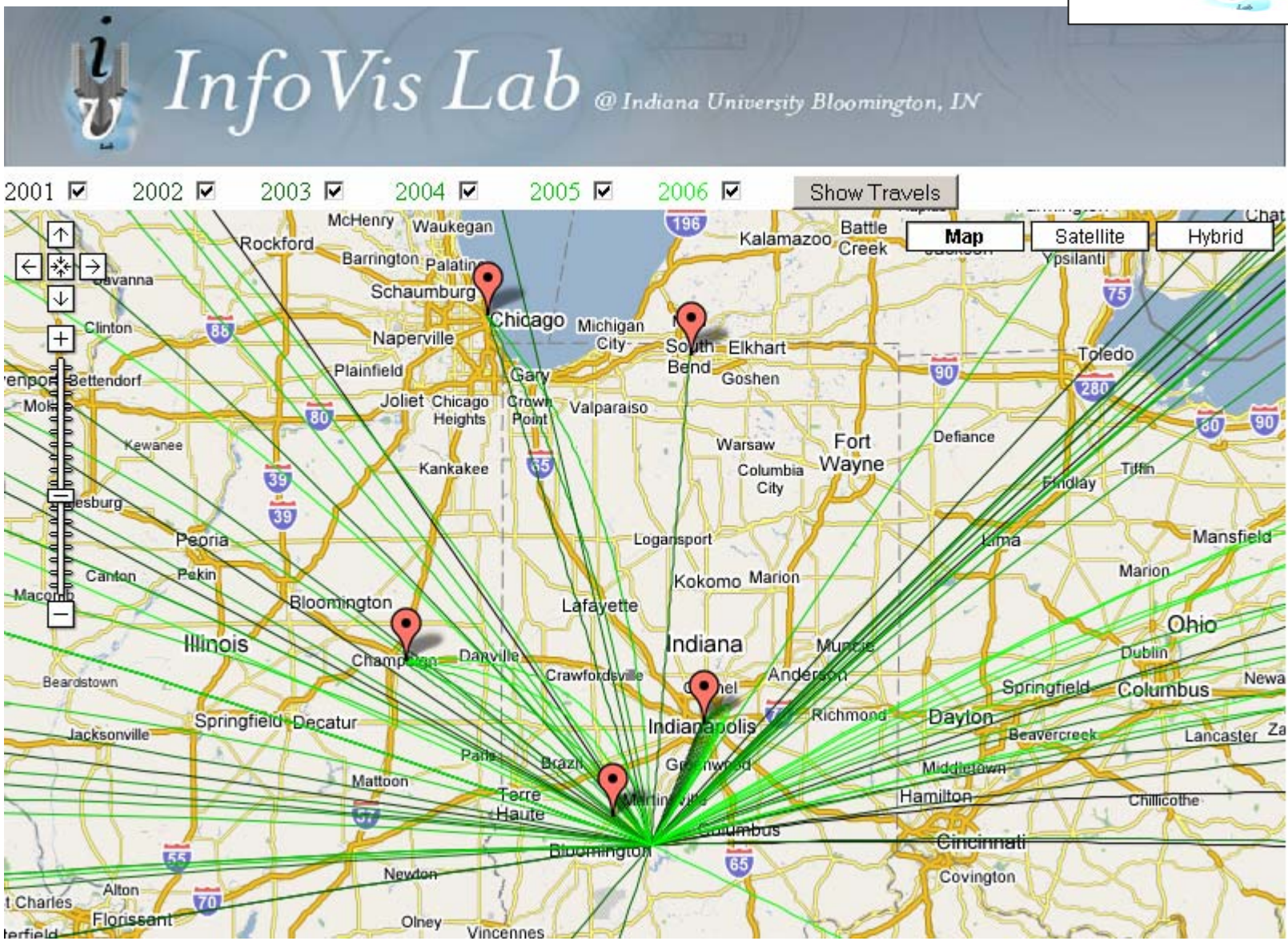


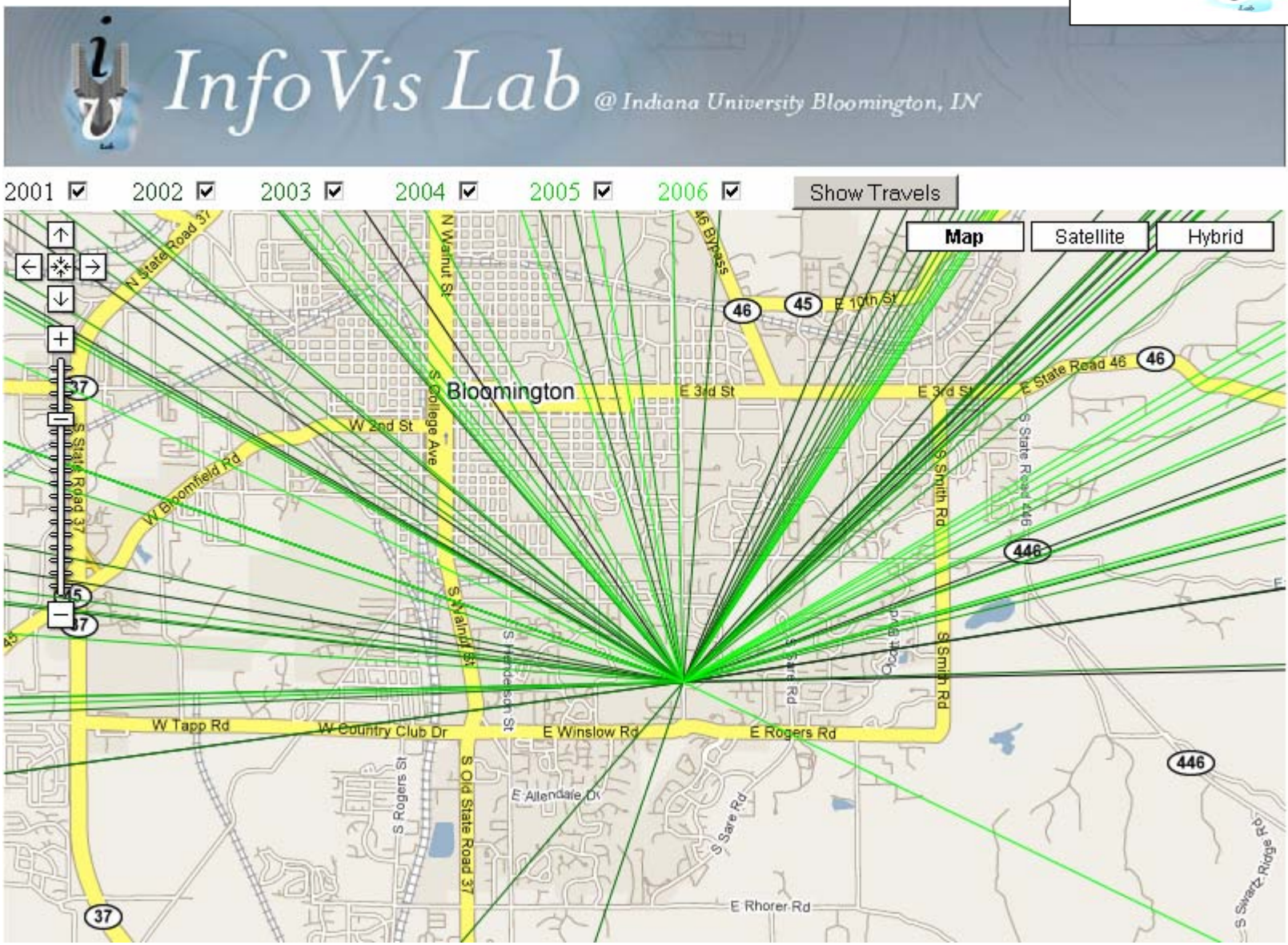
Katy's Travels in 2000-2006

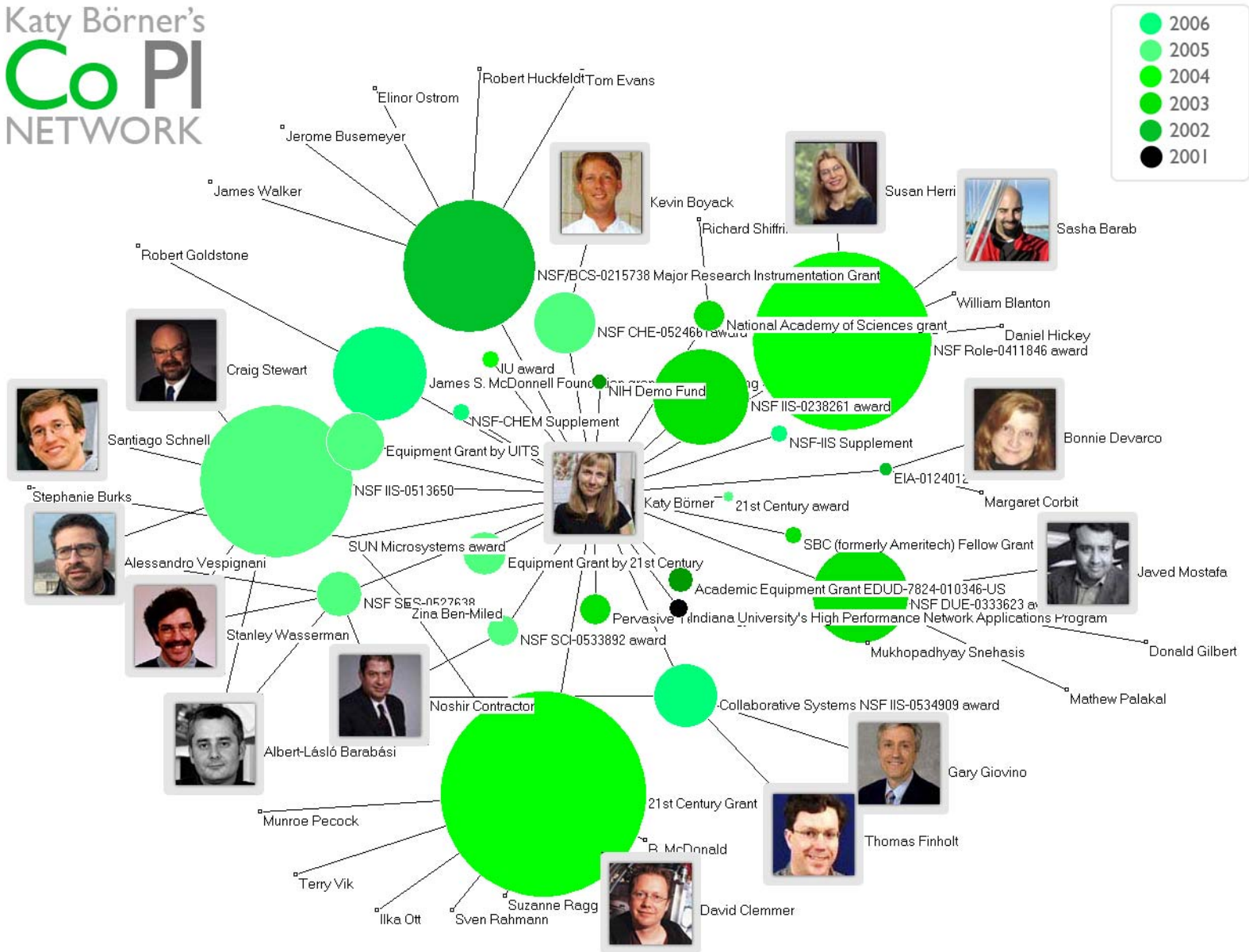



InfoVis Lab @ Indiana University Bloomington, IN
2001
2002
2003
2004
2005
2006
Show Travels
Map
Satellite
Hybrid
British Columbia
Edmonton
Saskatchewan
Camrose
Saskatoon
Calgary
Kamloops
Kelowna
Lethbridge
Regina
Winnipeg
Brandon
Weyburn
Ontario
Québec
Thunder Bay
Val-d'Or
Sudbury
Montréal
Ottawa
Toronto
Peterborough
Maine
New Brunswick
Nova Scotia
Vermont
New Hampshire
Massachusetts
Rhode Island
Connecticut
New Jersey
Delaware
Maryland
District of Columbia
Spokane
Seattle
Portland
Montana
North Dakota
South Dakota
Minneapolis
Wisconsin
Michigan
Detroit
Idaho
Wyoming
Nebraska
Iowa
Salt Lake City
Denver
Nevada
Utah
Colorado
Kansas
Missouri
Louisville
Kentucky
Tennessee
West Virginia
North Carolina
South Carolina
Georgia
Alabama
Mississippi
Arkansas
Oklahoma
Texas
Louisiana
New Mexico
Arizona
Albuquerque
Amarillo
Lubbock
Dallas
El Paso
Phoenix
Tucson
San Diego
Los Angeles
San Francisco
Las Vegas
Atlanta
Jacksonville
Tallahassee


InfoVis Lab @ Indiana University Bloomington, IN
2001
2002
2003
2004
2005
2006
Show Travels
Map
Satellite
Hybrid
North Dakota
South Dakota
Minnesota
Minneapolis
Duluth
Thunder Bay
Wisconsin
Michigan
Madison
Chicago
Fort Wayne
Flint
Kitchener
Toronto
London
Hamilton
Sudbury
North Bay
Owen Sound
Peterborough
Kingston
Ottawa
Montreal
Québec
Victoriaville
Sherbrooke
Cornwall
Vermont
New York
Massachusetts
Connecticut
Pennsylvania
New Jersey
Philadelphia
Delaware
Maryland
Alexandria
Pittsburgh
Cleveland
Erie
Canton
Ohio
Iowa
Nebraska
Omaha
Des Moines
Lincoln
Illinois
Kansas City
Topeka
Kansas
Wichita
Missouri
St Louis
Springfield
Tulsa
Oklahoma
Amarillo
Arkansas
Memphis
Louisville
Cincinnati
West Virginia
Lexington-Fayette
Evansville
Kentucky
Clarksville
Lebanon
Greeneville
Tennessee
Knoxville
Chattanooga
Greenville
Charlotte
Fayetteville
North Carolina
Virginia
Richmond
Chesapeake
District of Columbia

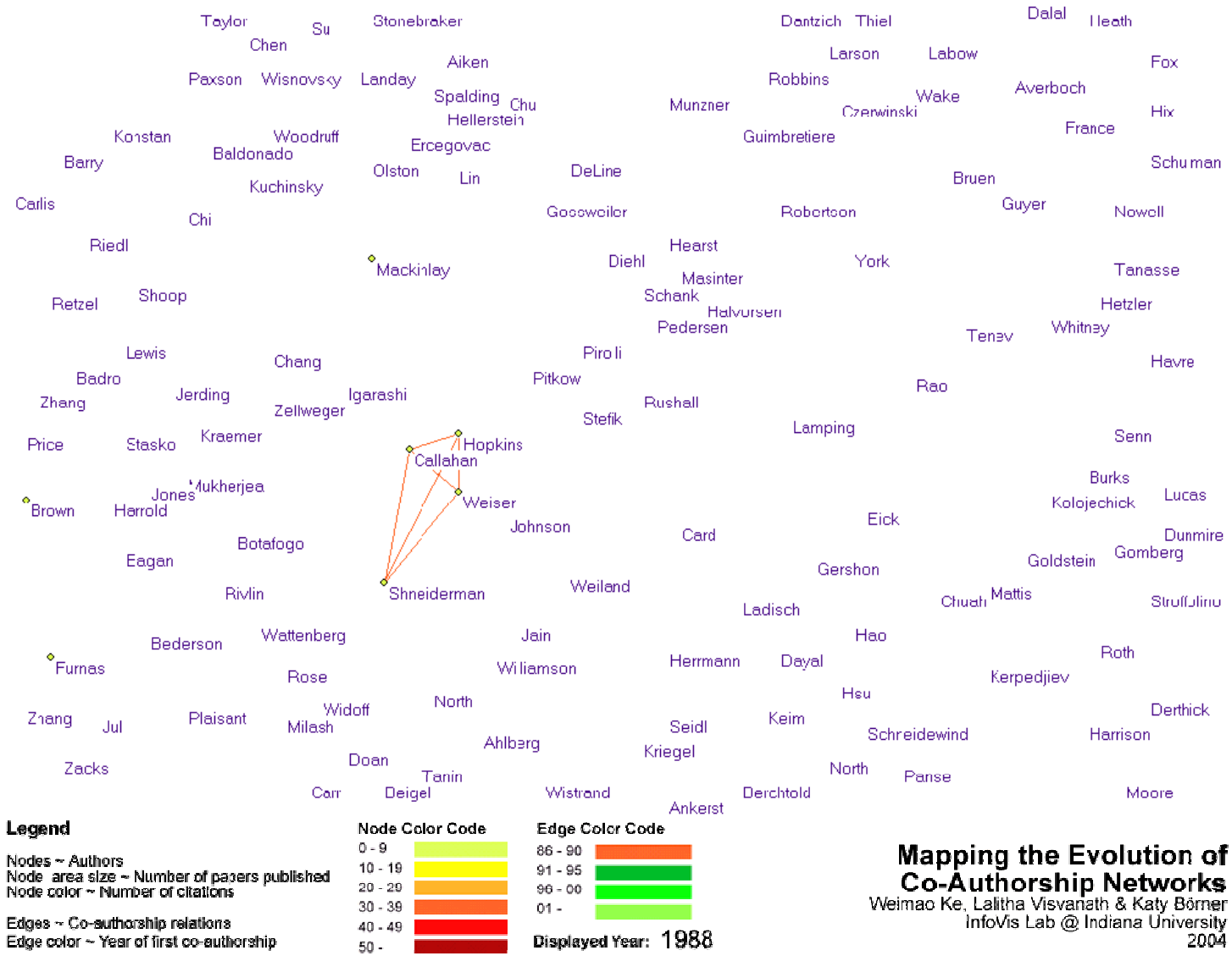

InfoVis Lab @ Indiana University Bloomington, IN
2001
2002
2003
2004
2005
2006
Show Travels
Map
Satellite
Hybrid
Chicago
Schaumburg
Naperville
Plainfield
Joliet
Gary
Michigan City
South Bend
Elkhart
Goshen
Kalamazoo
Battle Creek
Toledo
Fort Wayne
Warsaw
Columbia City
Defiance
Findlay
Tiffin
Mansfield
Marion
Kokomo
Logansport
Lafayette
Bloomington
Champaign
Illinois
Indiana
Ohio
Indianapolis
Dayton
Springfield
Columbus
Richmond
Anderson
Muncie
Terre Haute
Cincinnati
Hamilton
Middletown
Springfield Decatur
Jacksonville
Peoria
Pekin
Canton
Rockford
Waukegan
Barrington
Clinton
Kankakee
Valparaiso
Crown Point
Chicago Heights
Crawfordsville
Brazil
Mattoon
Newton
Olney
Vincennes
Florissant
Alton
Covington
Chillicothe
Lancaster
Beavercreek
Dublin


InfoVis Lab @ Indiana University Bloomington, IN
2001
2002
2003
2004
2005
2006
Show Travels
Map
Satellite
Hybrid
Bloomington
N State Road 37
N Walnut St
S College Ave
S Walnut St
S Rogers St
S Old State Road 37
S State Road 37
W 2nd St
W Bloomfield Rd
E 3rd St
E 10th St
E State Road 46
S State Road 446
S Smith Rd
S Sare Rd
W Tapp Rd
W Country Club Dr
E Winslow Rd
E Rogers Rd
E Allendale Dr
E Rhorer Rd
S Swartz Ridge Rd
46
45
37
446


Katy Börner's
Co PI
NETWORK
2006
2005
2004
2003
2002
2001
Robert Huckfeldt
Tom Evans
Elinor Ostrom
Jerome Busemeyer
James Walker
Kevin Boyack
Susan Herri
Sasha Barab
Richard Shiffri.
Robert Goldstone
NSF/BCS-0215738 Major Research Instrumentation Grant
William Blanton
National Academy of Sciences grant
Daniel Hickey
NSF Role-0411846 award
Craig Stewart
NSF CHE-052466 award
JU award
James S. McDonnell Found
NIH Demo Fund
NSF IIS-0238261 award
NSF-CHEM Supplement
Santiago Schnell
Equipment Grant by UITS
NSF-IIS Supplement
Bonnie Devarco
EIA-0124012
Stephanie Burks
NSF IIS-0513650
Katy Börner
21st Century award
Margaret Corbit
SBC (formerly Ameritech) Fellow Grant
SUN Microsystems award
Equipment Grant by 21st Century
Alessandro Vespignani
Javed Mostafa
Academic Equipment Grant EDUD-7824-010346-US
NSF SES-0527638
Zina Ben-Miled
Pervasive T
Indiana University's High Performance Network Applications Program
Stanley Wasserman
NSF SCI-0533892 award
NSF DUE-0333623 a
Mukhopadhyay Snehasis
Donald Gilbert
Noshir Contractor
Collaborative Systems NSF IIS-0534909 award
Mathew Palakal
Albert-László Barabási
Gary Giovino
21st Century Grant
Munroe Pecock
Thomas Finholt
R. McDonald
Terry Vik
Suzanne Ragg
David Clemmer
Ilka Ott
Sven Rahmann

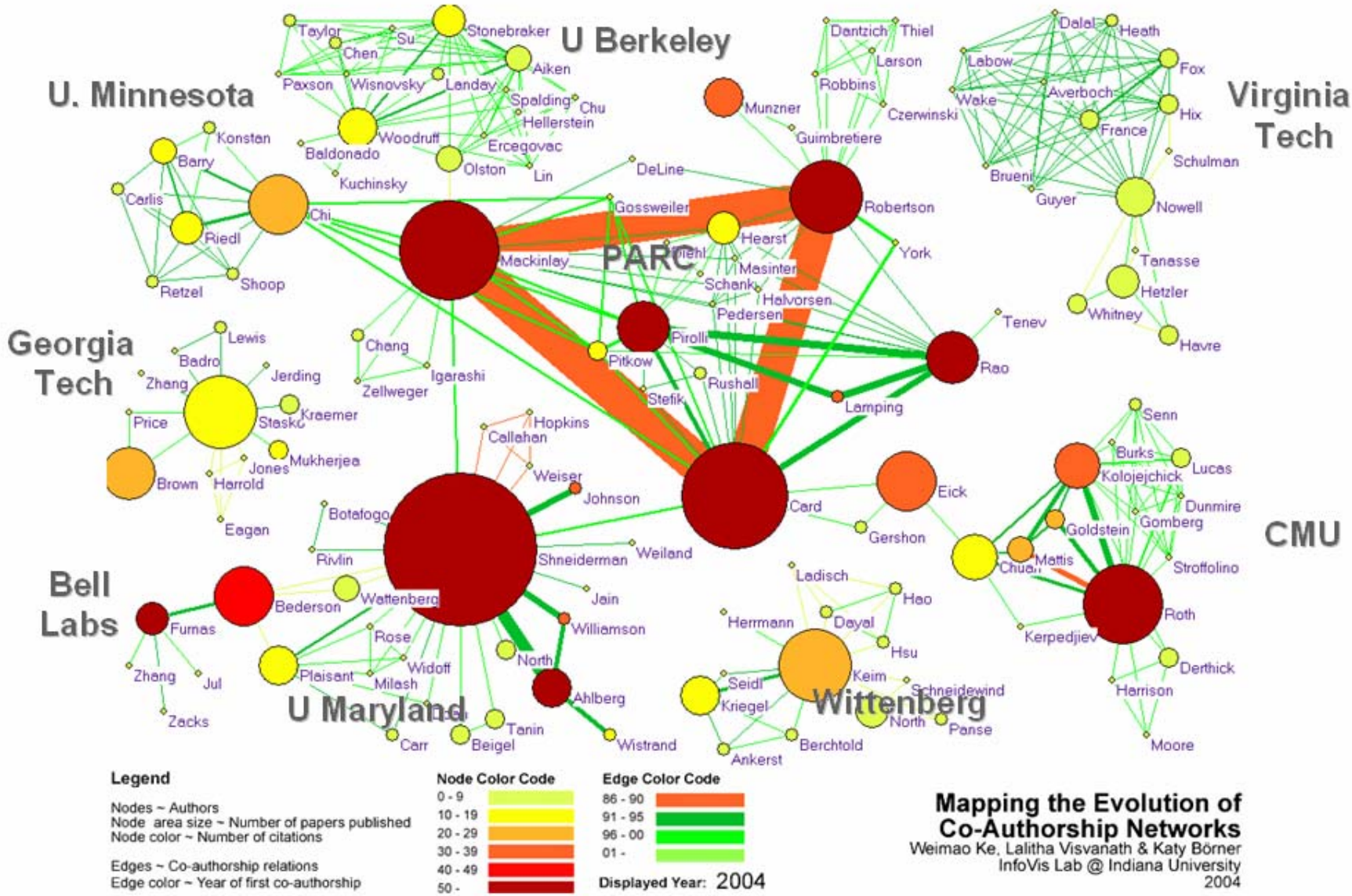
Mapping the Evolution of Co-Authorship Networks

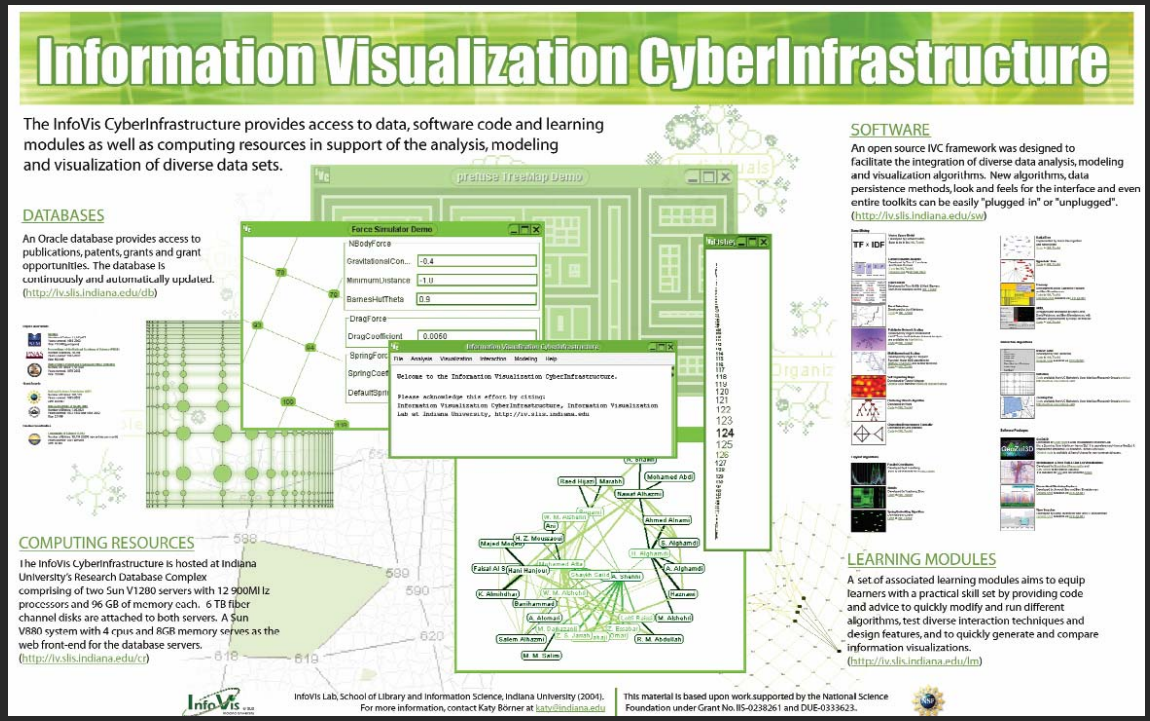
Ke, Visvanath & Börner, (2004) Won 1st price at the IEEE InfoVis Contest.



Mapping the Evolution of Co-Authorship Networks

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





CAREER: Visualizing Knowledge Domains. NSF IIS-0238261 award (Katy Börner, $451,000) Sept. 03-Aug. 08.
http://iv.slis.indiana.edu/

CIShell
cyberinfrastructure shell

SDB | SCHOLARLY DATABASE

Scholarly Database
http://sdb.slis.indiana.edu

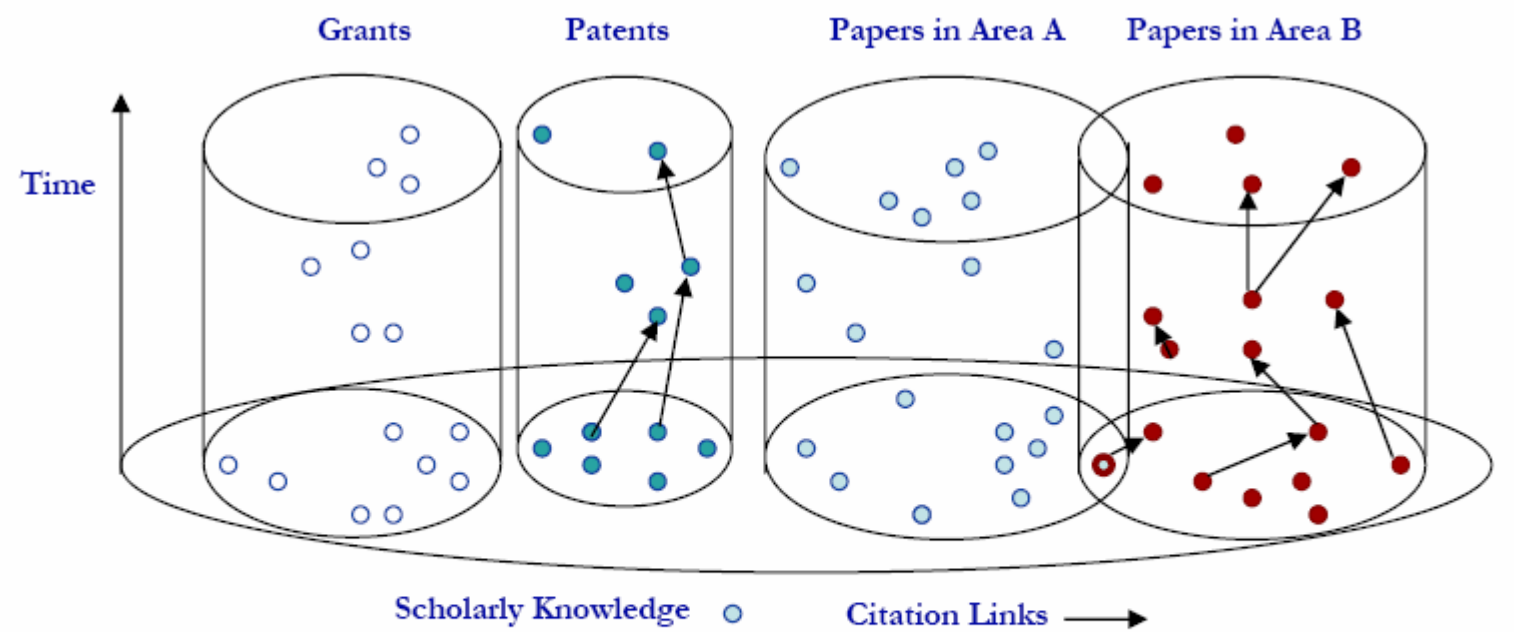
NetworkWorkbench
A Workbench for Network Scientists

SEI: Network Workbench: A Large-Scale Network Analysis, Modeling and Visualization Toolkit for Biomedical, Social Science and Physics Research. NSF IIS-0513650 award (Katy Börner, Albert-Laszlo Barabasi, Santiago Schnell, Alessandro Vespignani & Stanley Wasserman, Eric Wernert (Senior Personnel), $1,120,926) Sept. 05 - Aug. 08. http://nwb.slis.indiana.edu

NSF SBC Sun microsystems
HPNAP High Performance Network Applications Program
INDIANA 21st Century Research & Technology Fund

SDB | SCHOLARLY DATABASE

Challenges - Interlink $ Input & Publication/Patent Citation Output



Need to interlink

- Grants and papers/patents.
- Grants/papers/patents and their PIs/authors/inventors, etc.

Use resulting networks to

- Count #papers, #citations, etc.
- Determine strength of co-PI/author/inventor relations, etc.

SDB | SCHOLARLY DATABASE

SCHOLARLY DATABASE

SEARCH INTERFACE: https://iv.slis.indiana.edu/db/
DOCUMENTATION: http://iv.slis.indiana.edu/db/

DB PROJECT LEAD
Gavin LaRowe
glarowe@indiana.edu

DB DEVELOPER
Sumeet Ambre
sambre@indiana.edu

PROJECT MANAGER
Katy Börner

STATUS
as of 06.08.28

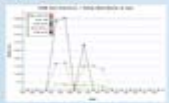
Information Visualization Laboratory
Cyberinfrastructure for Network Science Center
School of Library and Information Science
Indiana University
Bloomington, IN 47405, USA

PAPERS

SDB MEDLINE

SDB PHYSREV

SDB PNAS



SDB JCR

PATENTS

SDB USPATENTS

KNOWLEDGE WEBS

SDB WIKI

GRANT AWARDS

SDB NSF

SDB NIH

DOCUMENT TABLE

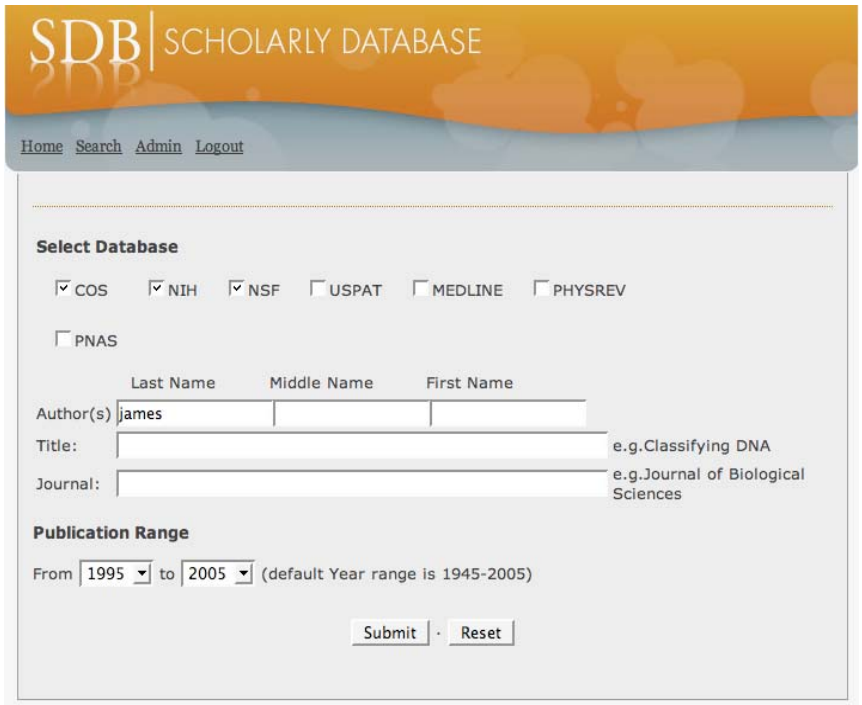
FUNDING OPPORTUNITIES

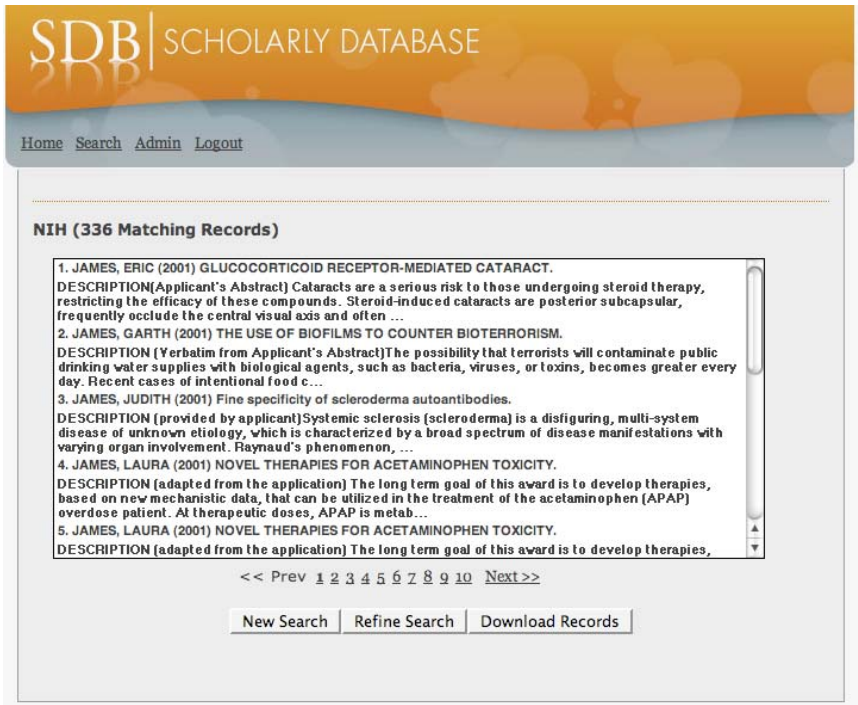
SDB COS

DESIGN BY ELISHA HARDY

Scholarly Database: Web Interface

Search across publications, patents, grants.
Download records and/or (evolving) co-author, paper-citation networks.





Register for free access at https://sdb.slis.indiana.edu.

Scholarly Database: # Records & Years Covered

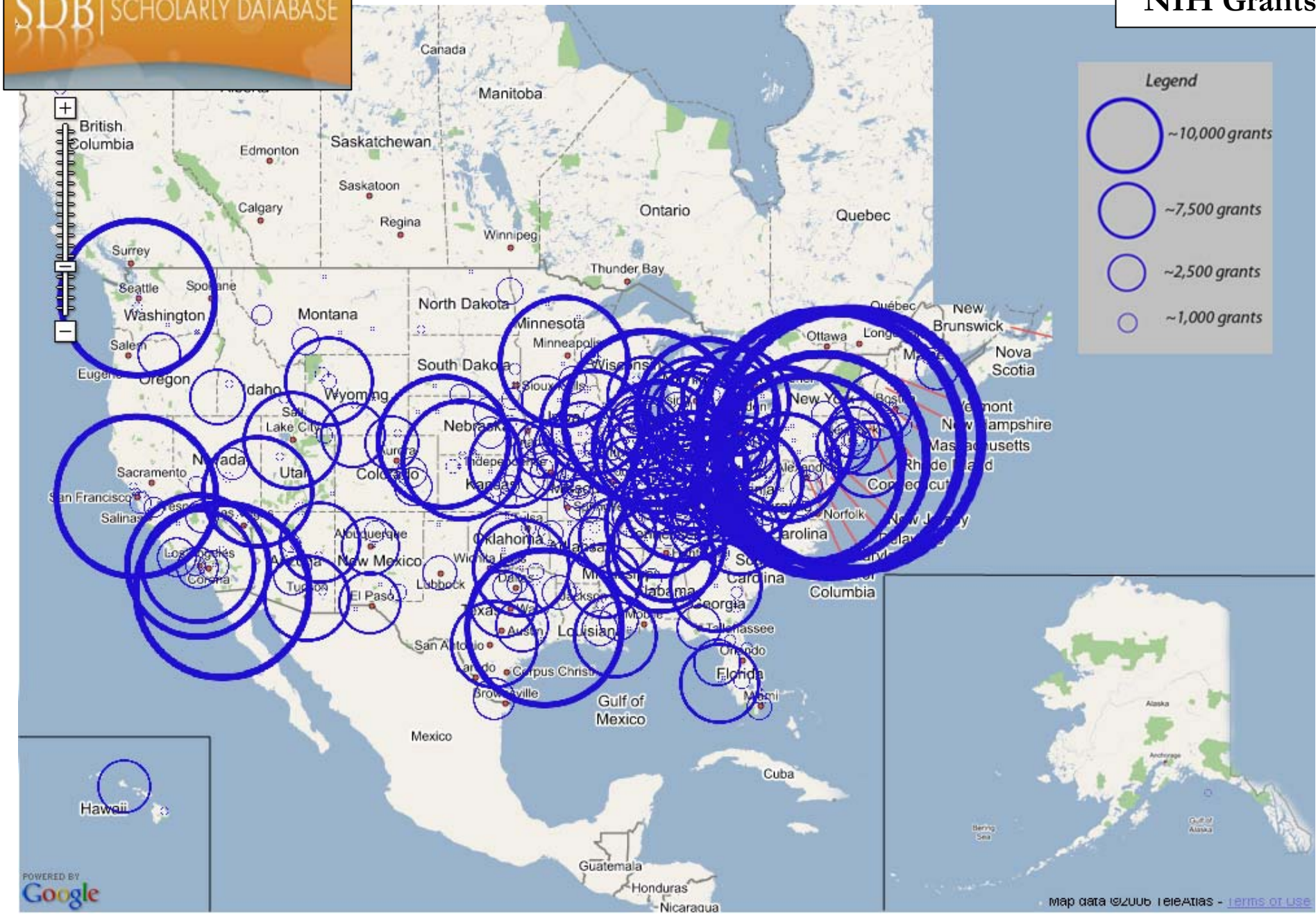
Datasets available via the Scholarly Database (* future feature)

Dataset	# Records	Years Covered	Updated	Restricted Access
Medline	13,149,741	1965-2005	Yes	
PhysRev	398,005	1893-2006		Yes
PNAS	16,167	1997-2002		Yes
JCR	59,078	1974, 1979, 1984, 1989 1994-2004		Yes
USPTO	3,179,930	1976-2004	Yes*	
NSF	174,835	1985-2003	Yes*	
NIH	1,043,804	1972-2002	Yes*	
Total	**18,021,560**	**1893-2006**	**4**	**3**

Aim for comprehensive time, geospatial, and topic coverage.

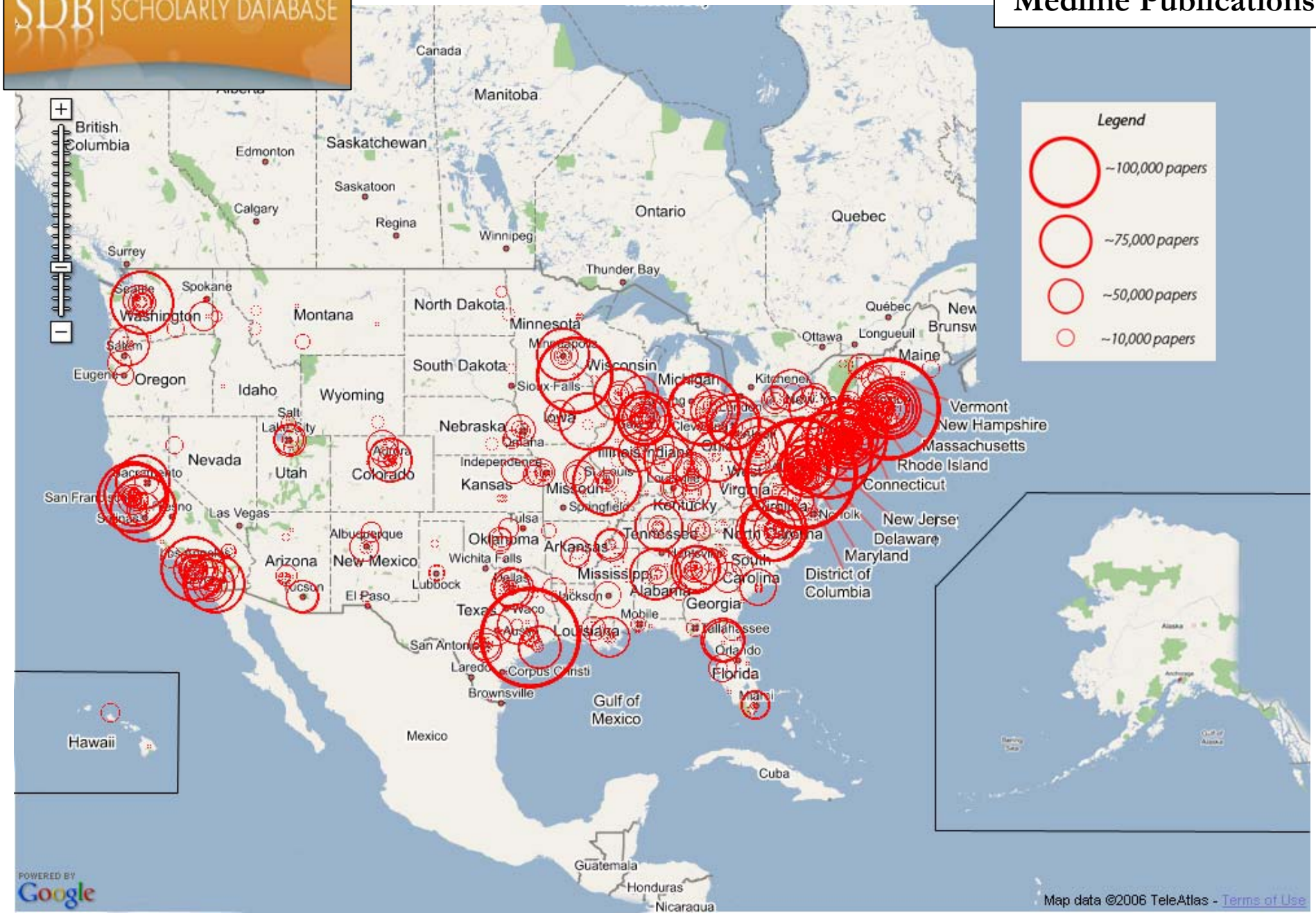
SDB | SCHOLARLY DATABASE

NIH Grants



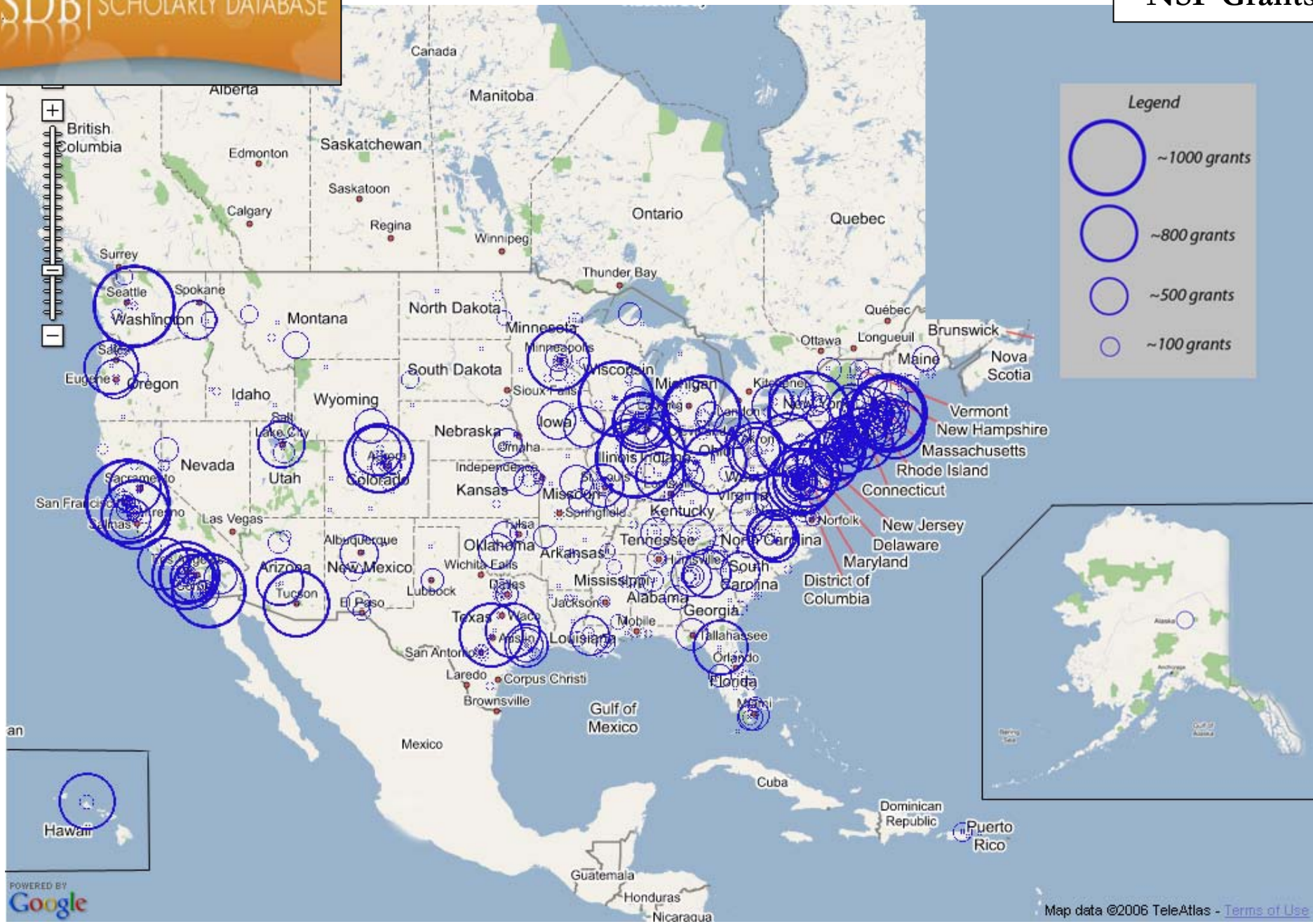
SDB | SCHOLARLY DATABASE

Medline Publications



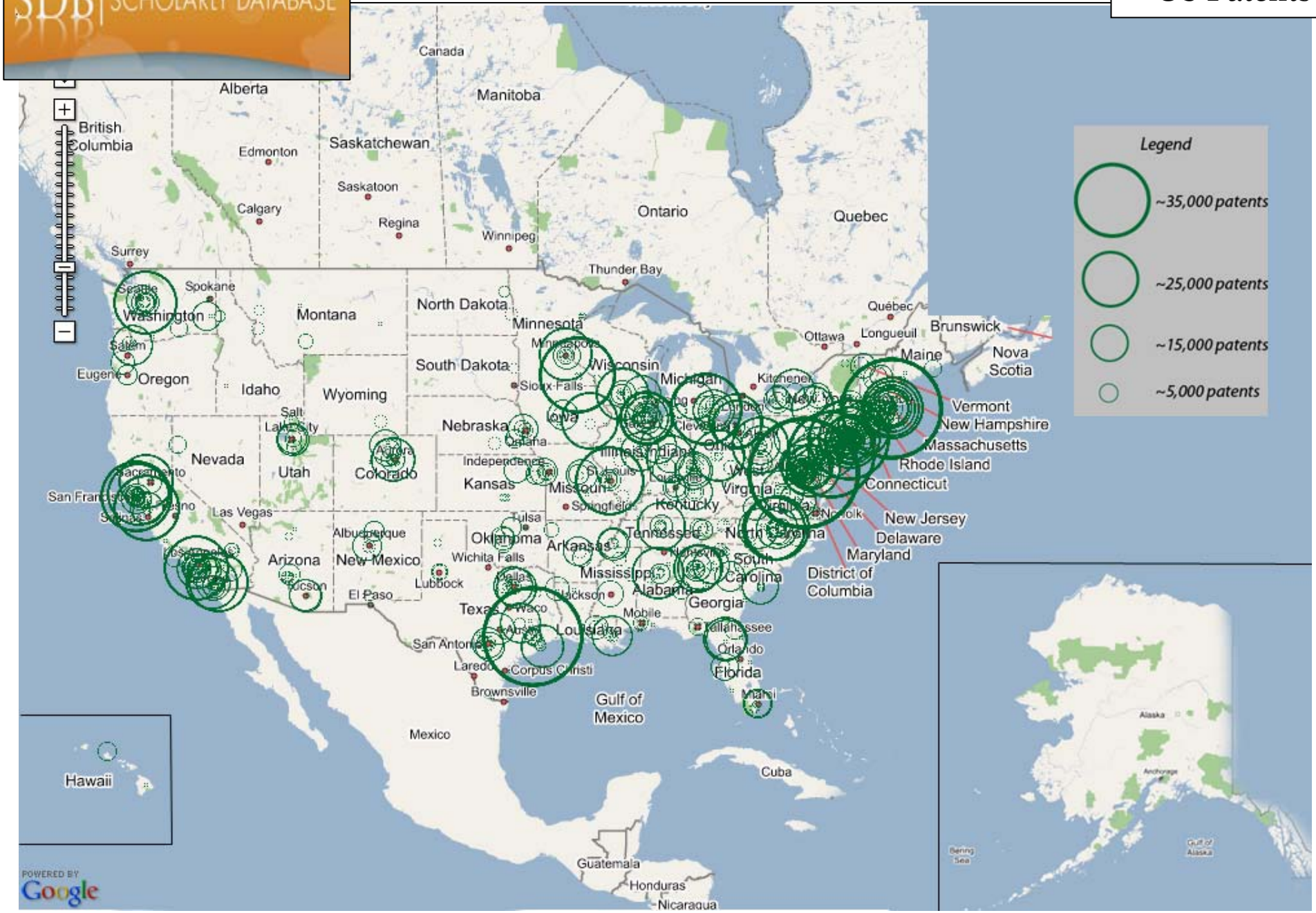
SDB | SCHOLARLY DATABASE

NSF Grants



SDB | SCHOLARLY DATABASE

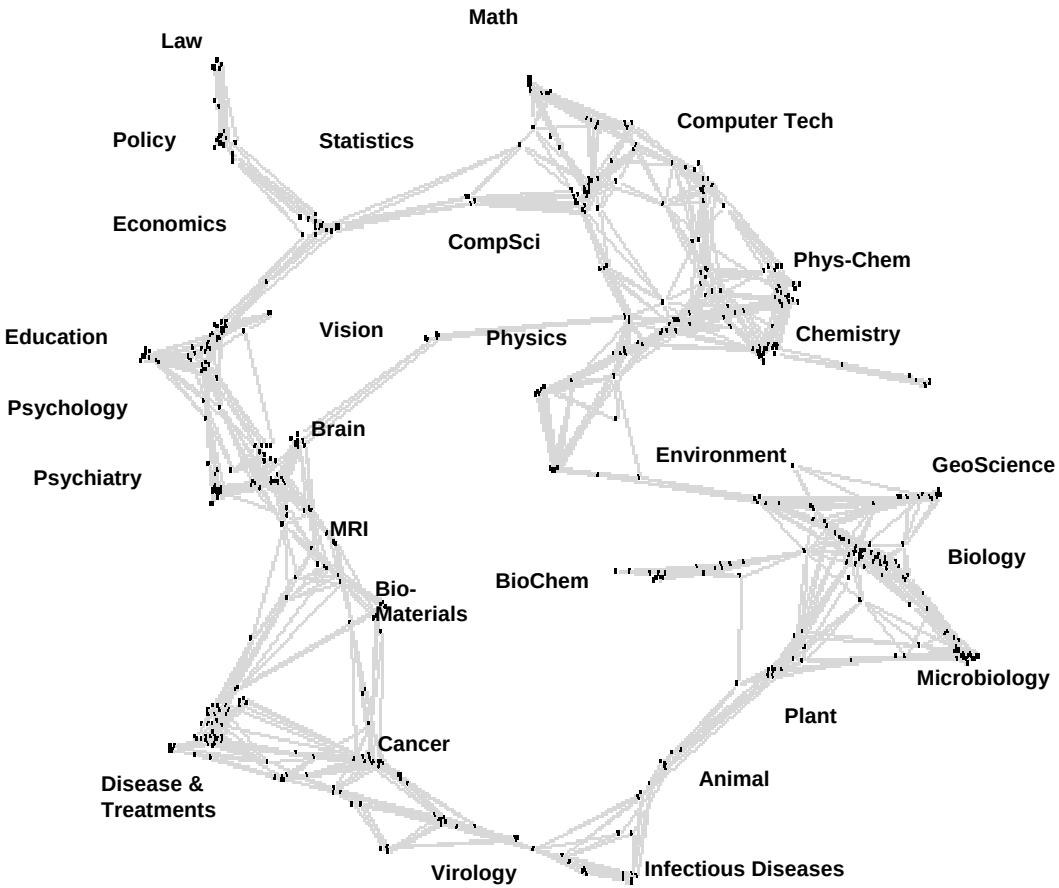
US Patents



Latest 'Base Map' of Science

Kevin W. Boyack & Richard Klavans, unpublished work.

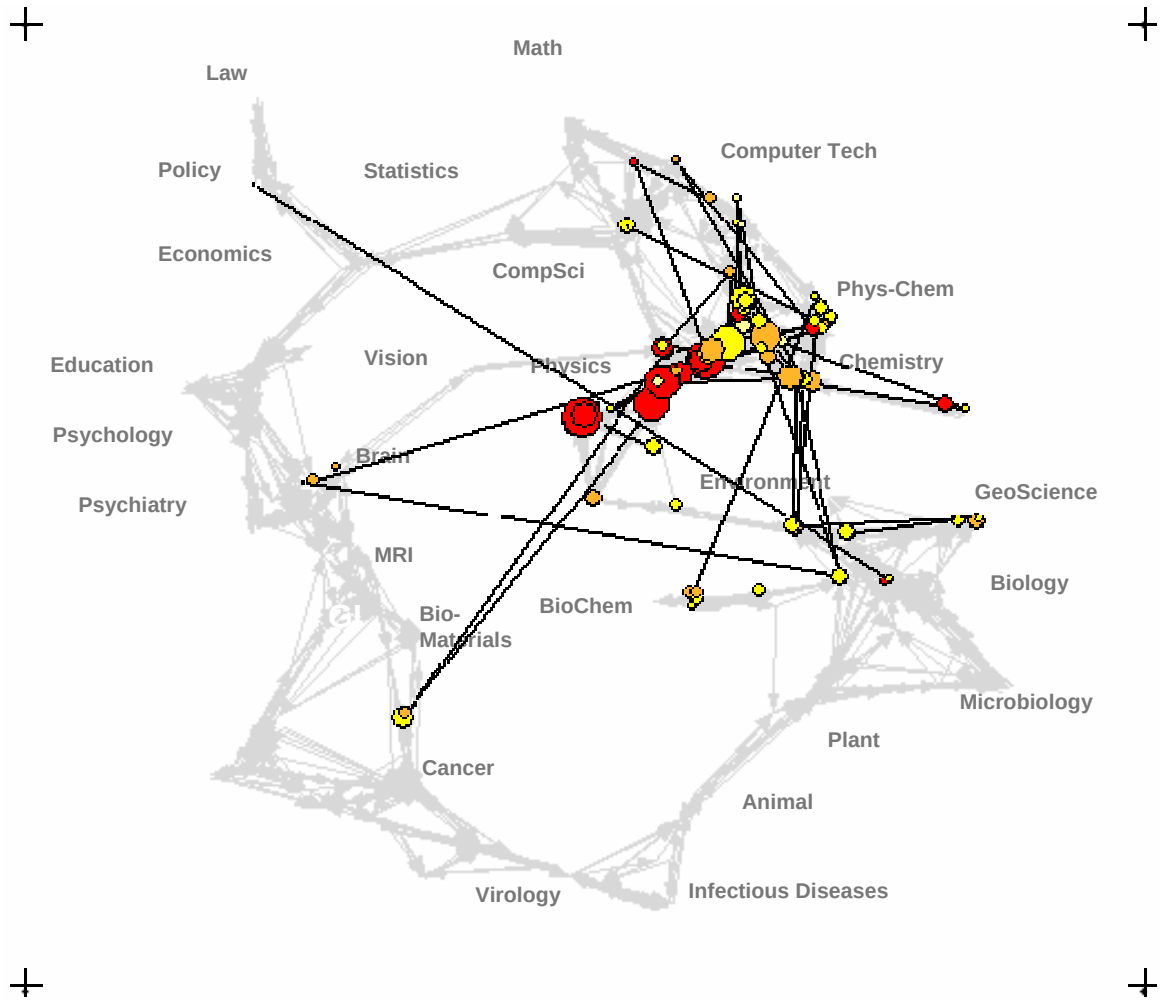
- Uses combined SCI/SSCI from 2002
  - 1.07M papers, 24.5M references, 7,300 journals
  - Bibliographic coupling of papers, aggregated to journals
- Initial ordination and clustering of journals gave 671 clusters
- Coupling counts were reaggregated at the journal cluster level to calculate the
  - (x,y) positions for each journal cluster
  - by association, (x,y) positions for each journal



Science map applications: Identifying core competency

Kevin W. Boyack & Richard Klavans, unpublished work.

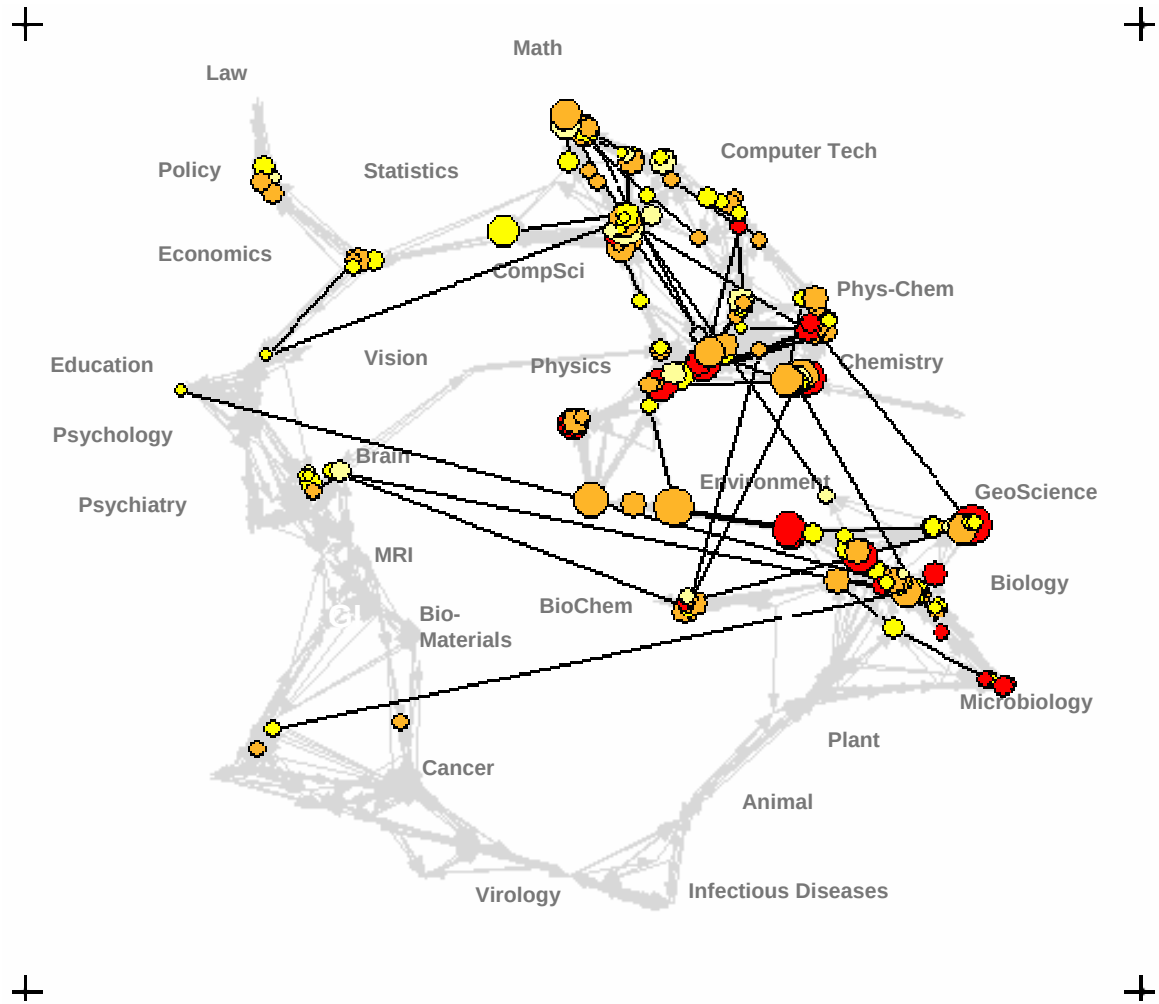
Funding patterns of the US Department of Energy (DOE)



Science map applications: Identifying core competency

Kevin W. Boyack & Richard Klavans, unpublished work.

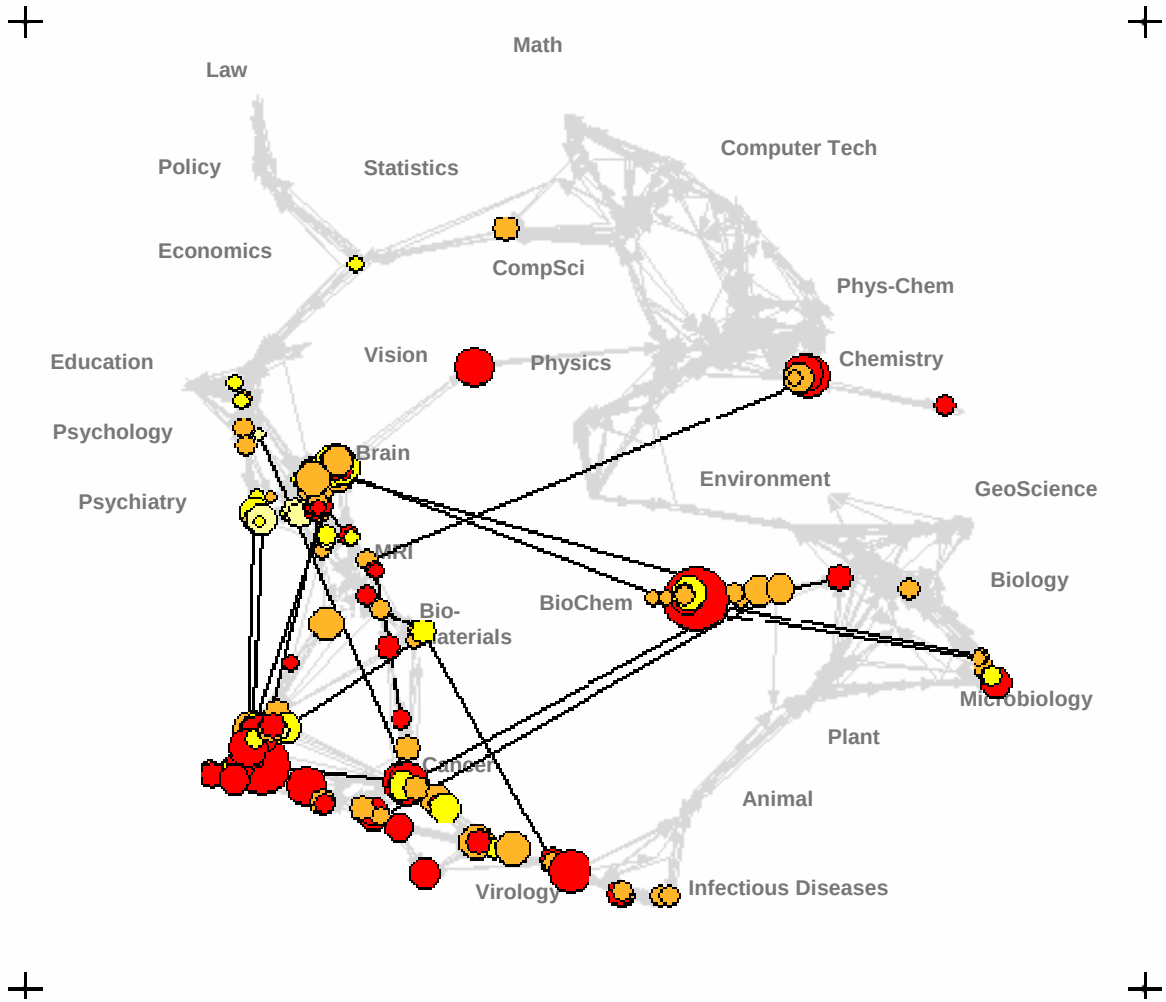
Funding Patterns of the National Science Foundation (NSF)



Science map applications: Identifying core competency

Kevin W. Boyack & Richard Klavans, unpublished work.

Funding Patterns of the National Institutes of Health (NIH)



Building Market Places not Cathedrals

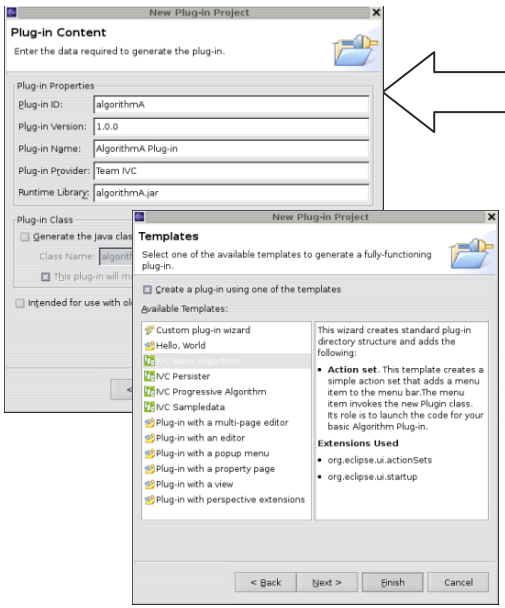
- 'Software glue' has to interlink datasets and algorithms written in different languages using different data formats.
- The smaller the glue or 'CI Shell', the more likely it can be maintained.

CIShell – Serving Non-CS Algorithm Developers & Users

Developers

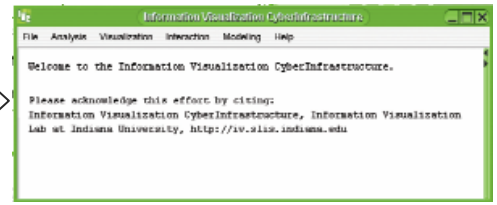
Users

CIShell Wizards

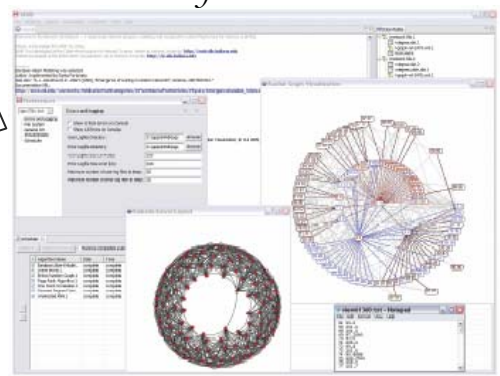


CIShell

IVC Interface



NWB Interface



CIShell – Build on OSGi Industry Standard

CIShell is built upon the Open Services Gateway Initiative (OSGi) Framework.

OSGi (http://www.osgi.org) is

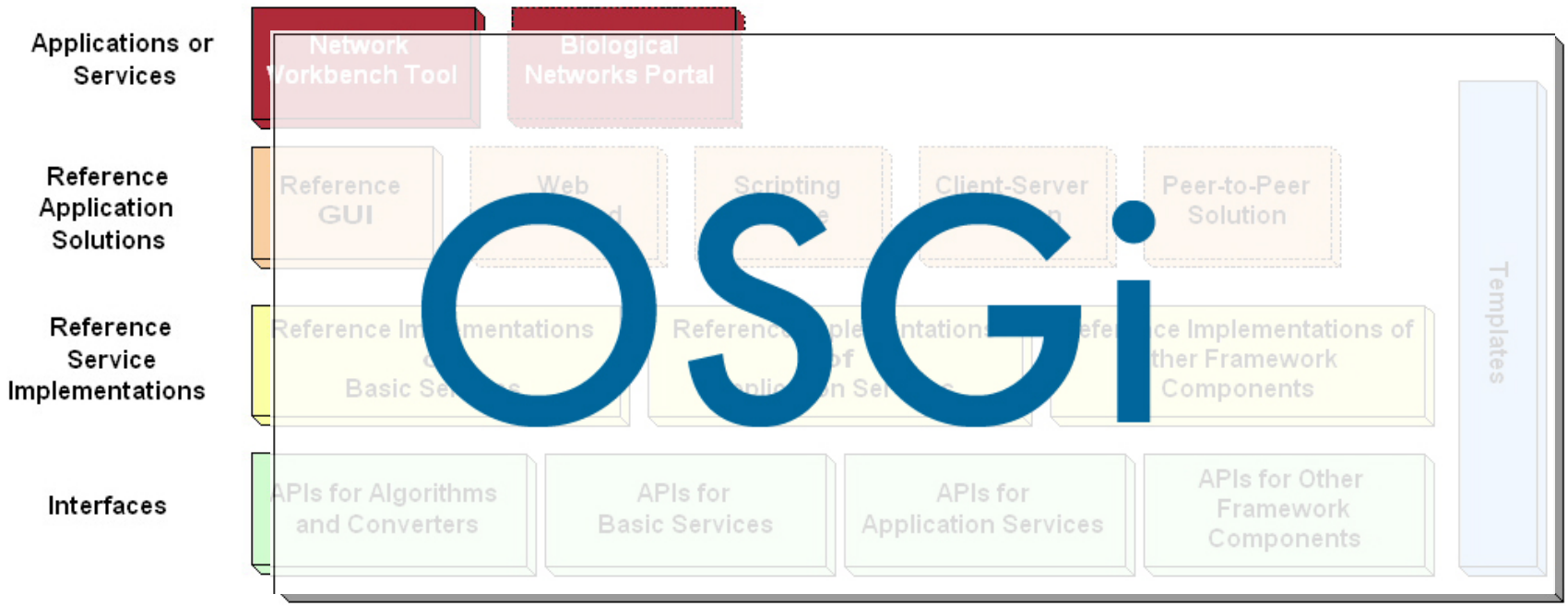
- A standardized, component oriented, computing environment for networked services.
- Successfully used in the industry from high-end servers to embedded mobile devices since 7 years.
- Alliance members include IBM (Eclipse), Sun, Intel, Oracle, Motorola, NEC and many others.
- Widely adopted in open source realm, especially since Eclipse 3.0 that uses OSGi R4 for its plugin model.

Advantages of Using OSGi

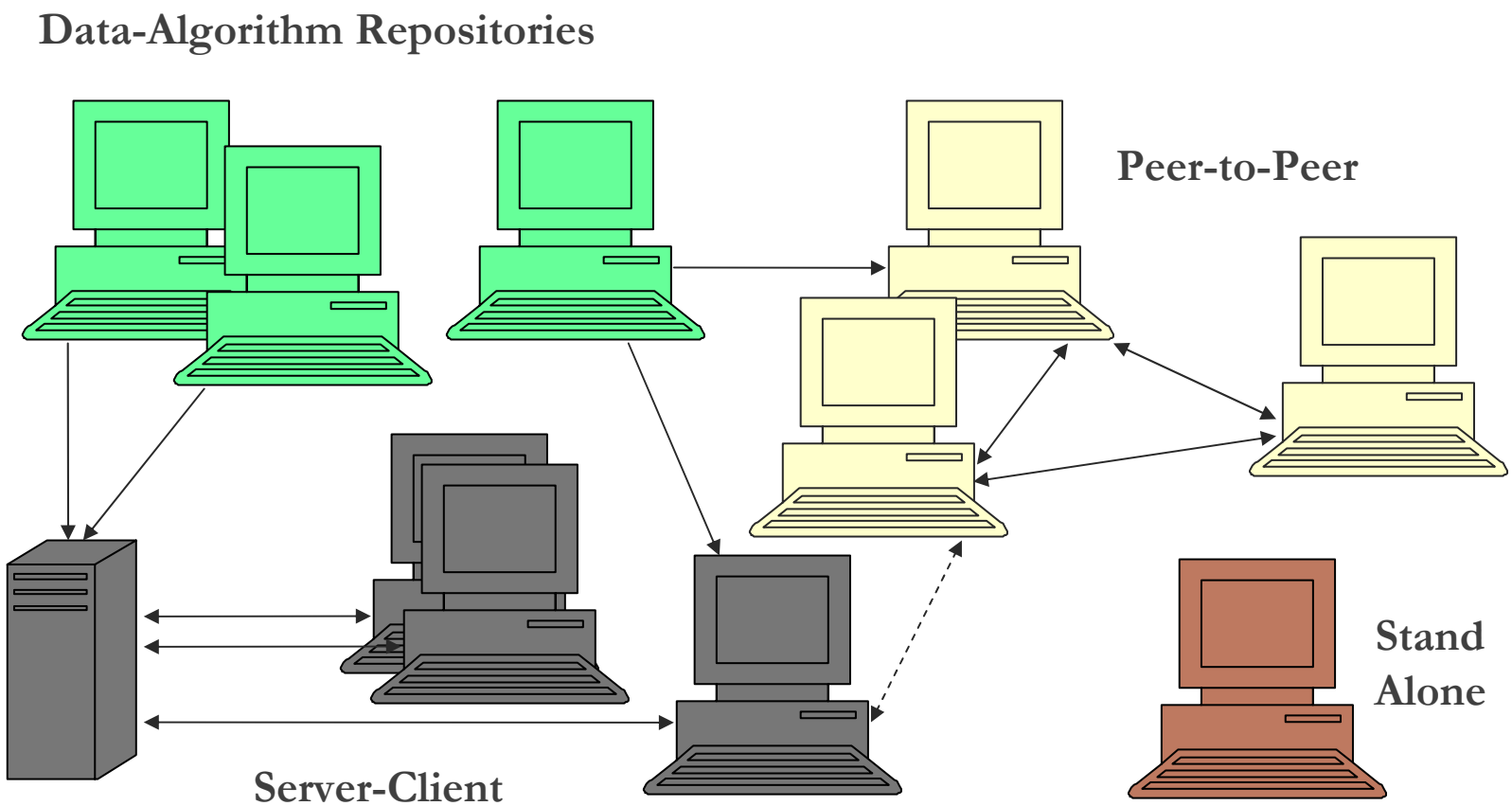
- Any CIShell algorithm is a service that can be used in any OSGi-framework based system.
- Using OSGi, running CIShells/tools can connected via RPC/RMI supporting peer-to-peer sharing of data, algorithms, and computing power.

Ideally, CIShell becomes a standard for creating OSGi Services for algorithms.

CIShell – Layer Cake



CIShell – Deployment



CIShell applications can be deployed as distributed data and algorithm repositories, stand alone applications, peer-to-peer architectures, and server-client architectures.

NWB Tool: Interface Elements

http://nwb.slis.indiana.edu

Load Data

Select Preferences

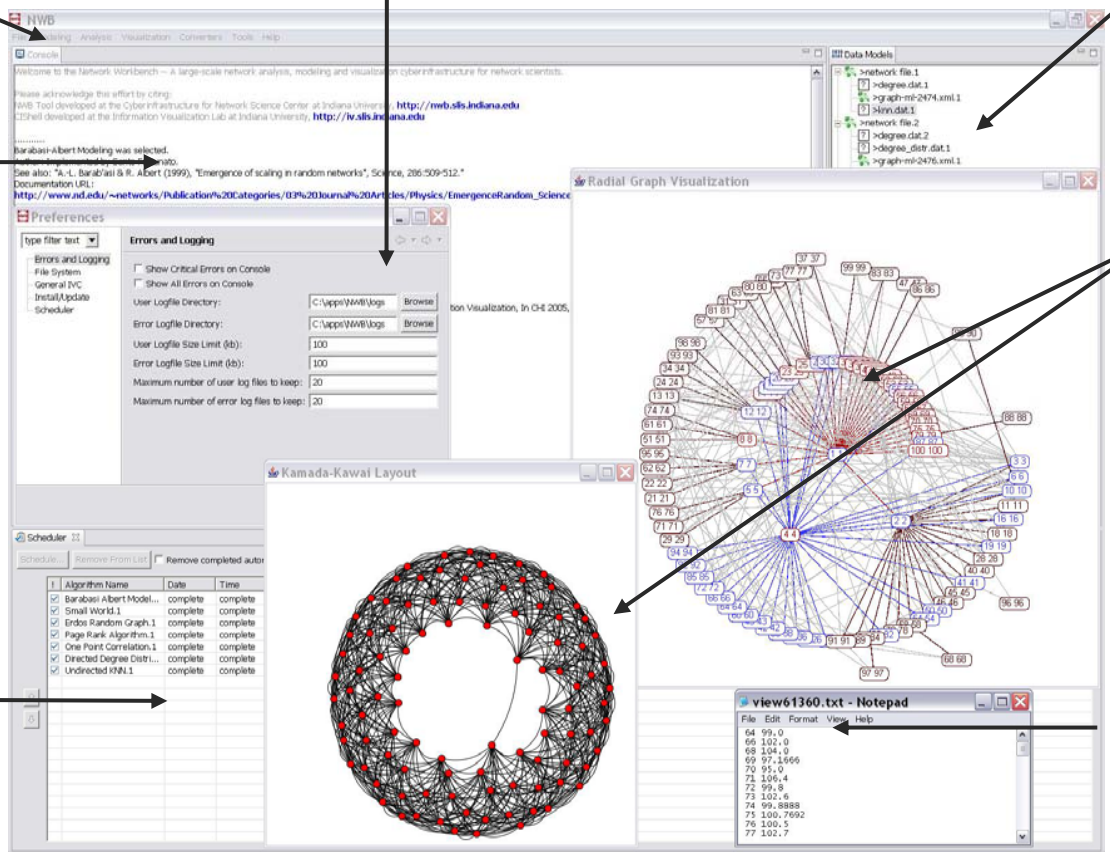
List of Data Models

Console

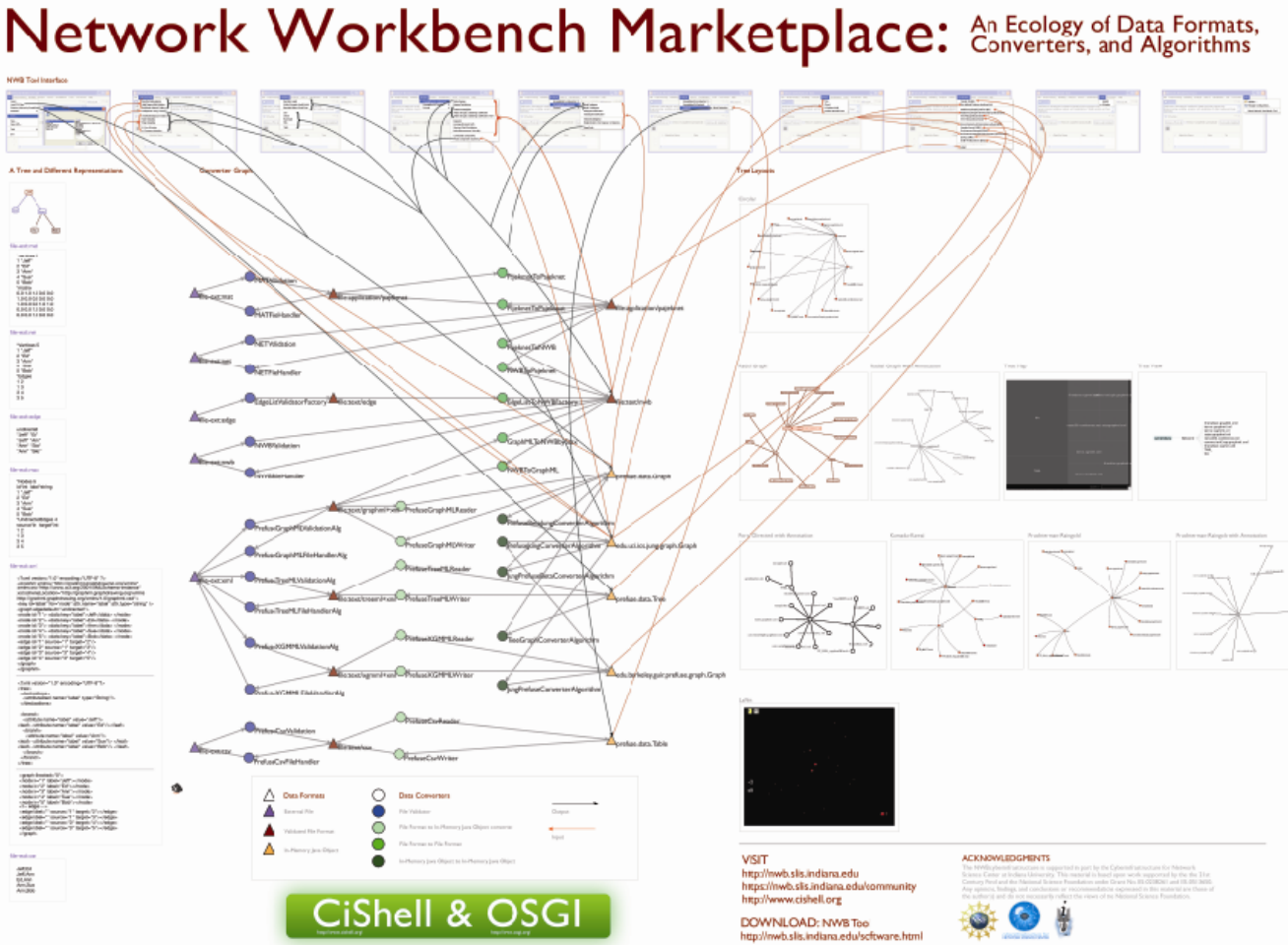
Visualize Data

Scheduler

Open Text Files



NWB Community Wiki



https://nwb.slis.indiana.edu/community/

Places & Spaces: Mapping Science

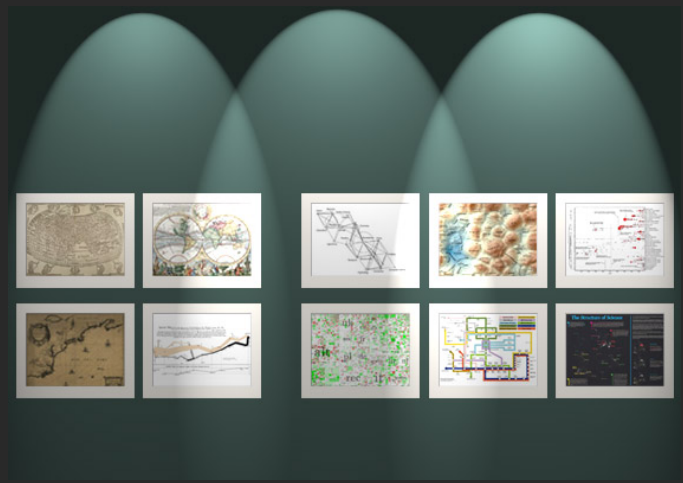
a science exhibit that introduces people to maps of sciences, their makers and users.

http://scimaps.org.

Exhibit Curators: Dr. Katy Börner & Elisha Hardy

The Power of Maps

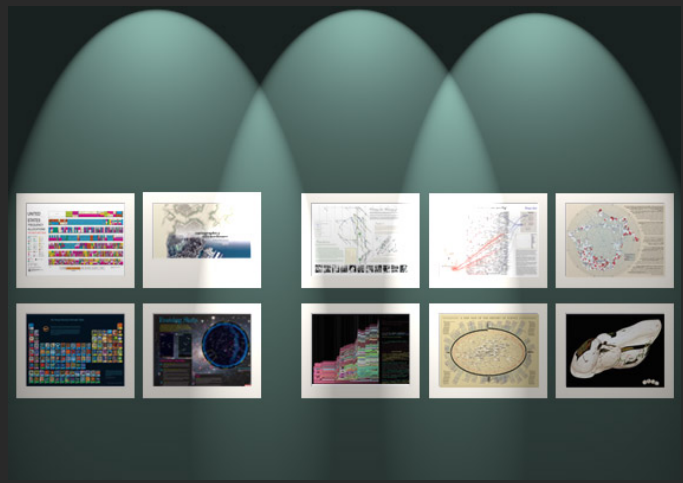
Four Early Maps of Our World
VERSUS
Six Early Maps of Science



(1st Iteration of Places & Spaces Exhibit - 2005)

The Power of Reference Systems

Four Existing Reference Systems
VERSUS
Six Potential Reference Systems of Science



(2nd Iteration of Places & Spaces Exhibit - 2006)

Impact

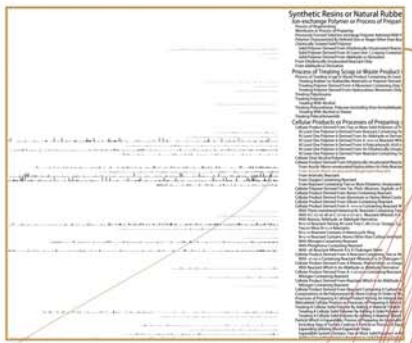
US Patent Hierarchy

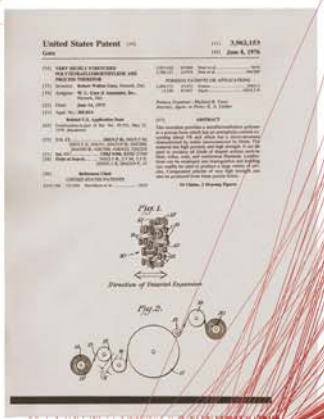
Prior Art

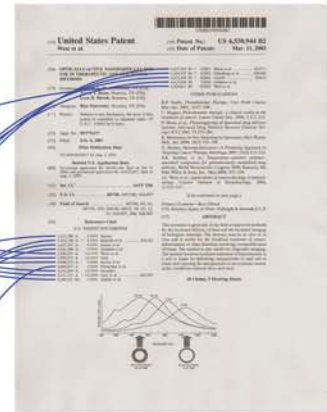
The United States Patent and Trademark Office does scientists and industry a great service by granting patents to protect inventions. Inventions are categorized in a taxonomy that groups patents by industry or use, proximate function, effect or product, and structure. At the time of this writing there are 160,523 categories in a hierarchy that can get as deep as 15 levels. We display the first three levels (13,529 categories) at right in what might be considered a textual map of inventions.

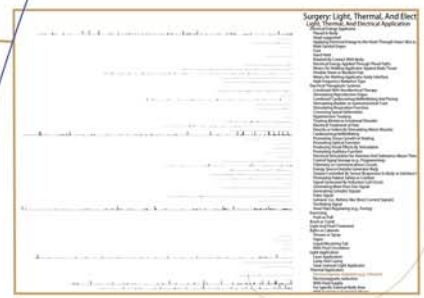
Patent applications are required to be unique and non-obvious, partially by revealing any previous patents that might be similar in nature or provide a foundation for the current invention. In this way we can trace the impact of a single patent, seeing how many patents and categories it affects.

The patent on Goretex—a lightweight, durable synthetic fiber—is an example of one that has had significant impact. The box below enlarges the section of the hierarchy where it is filed, and the red lines (arranged to start along a time line from 1981 to 2006) point to the 130 categories that contain 182 patents, from waterproof clothing to surgical cosmetic implants, that mention Goretex as prior art.









New patents often build on older ideas from many categories. Here, blue lines originate in sixteen different categories that contain the patents cited as prior art for a patent on "gold nanoshells." Gold nanoshells are a new invention: tiny spheres (with a diameter ten million times smaller than a human hair) that can be used to make tumors more visible in infrared scans, and have even helped cause complete remission of tumors in tests with laboratory mice. The blue lines show that widely separated categories provided background for this invention.

Keeping categories understandable is an important part of maintaining any taxonomy, including the patent hierarchy. Categories are easier to understand, search, and maintain if they contain elements (patents in this case) that fit well within the definition of the category. The box above shows a tiny bar chart, part of a "Taxonomy Validator" that helps people decide whether categories are good ones.

Categories can be redefined or combined, and sometimes need to be split when they becomes too large; a constant problem shared by many classifications systems in this information-rich century. But how can we determine exactly where to split a category in two, for example--if there are hundreds or thousands of elements in it?

The Taxonomy Validator measures a "distance to prototype:" how far each element is from an idealized "prototype" element for each bucket. This can be based on statistics, computational comparisons of words, or even human judgement. A simple bar chart can then show how good a category is. A good category has lots of small bars; a generally ragged category is one that might need scrutiny or reorganization; while one that has only one or two tall bars may just mean that one or two elements don't belong. Even simple visualizations like this can ease knowledge work by showing the eye much more than can fit into memory as words; focusing people on just the right issues, and providing a vastly broader background to support more informed judgements.


Extruding a solid polymer containing material to
Synthetic Resins or Natural ...
Coating Processes
Rotary Shafts, Gudgeons, ...
Data Processing: Structural ...
Synthetic Resins or Natural ...
Measuring, Testing
Data Processing: Speech ...
Data Processing: Financial, ...
Insulative Material Deposited
Foraminous Product Produced
Organic Coating Applied
Threaded, Headed Fastener, ...
Jewelry, Symbolic
Synthetic Resins or Natural ...
Data Processing: Artificial ...
Data Processing: Transportation ...
Synthetic Resins or Natural ...
Electrical Computers: Arithmet
Amusement Devices
Electrical Computers And ...

Impact

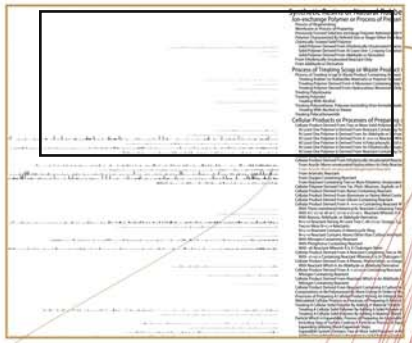
US Patent Hierarchy

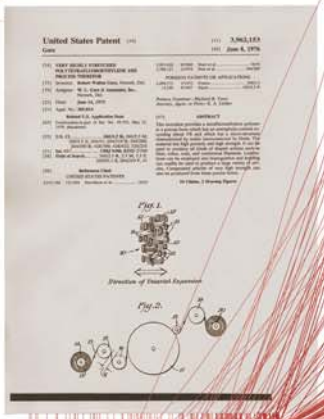
Prior Art

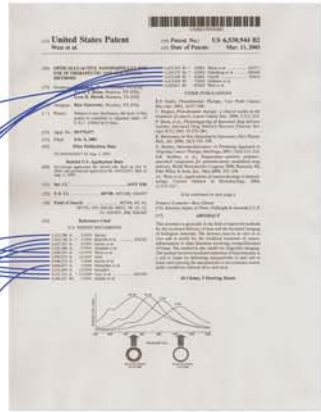
The United States Patent and Trademark Office does scientists and industry a great service by granting patents to protect inventions. Inventions are categorized in a taxonomy that groups patents by industry or use, proximate function, effect or product, and structure. At the time of this writing there are 160,523 categories in a hierarchy that can get as deep as 15 levels. We display the first three levels (13,529 categories) at right in what might be considered a textual map of inventions.

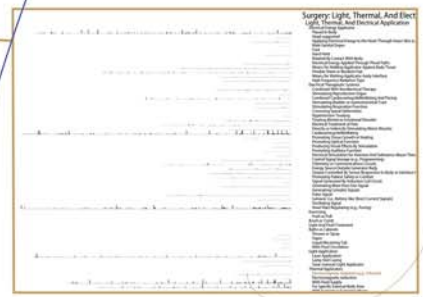
Patent applications are required to be unique and non-obvious, partially by revealing any previous patents that might be similar in nature or provide a foundation for the current invention. In this way we can trace the impact of a single patent, seeing how many patents and categories it affects.

The patent on Goretex—a lightweight, durable synthetic fiber—is an example of one that has had significant impact. The box below enlarges the section of the hierarchy where it is filed, and the red lines (arranged to start along a time line from 1981 to 2006) point to the 130 categories that contain 182 patents, from waterproof clothing to surgical cosmetic implants, that mention Goretex as prior art.







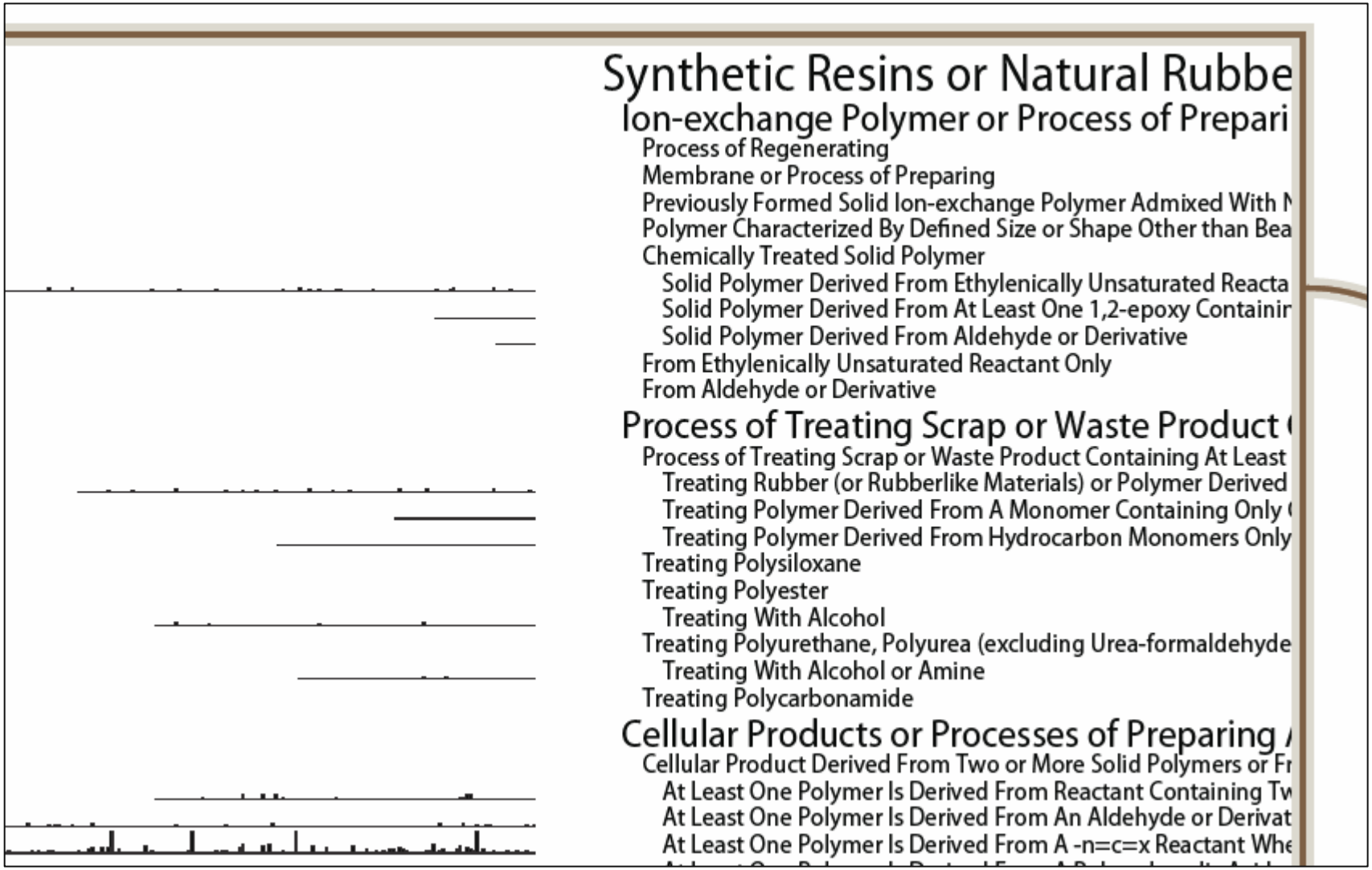


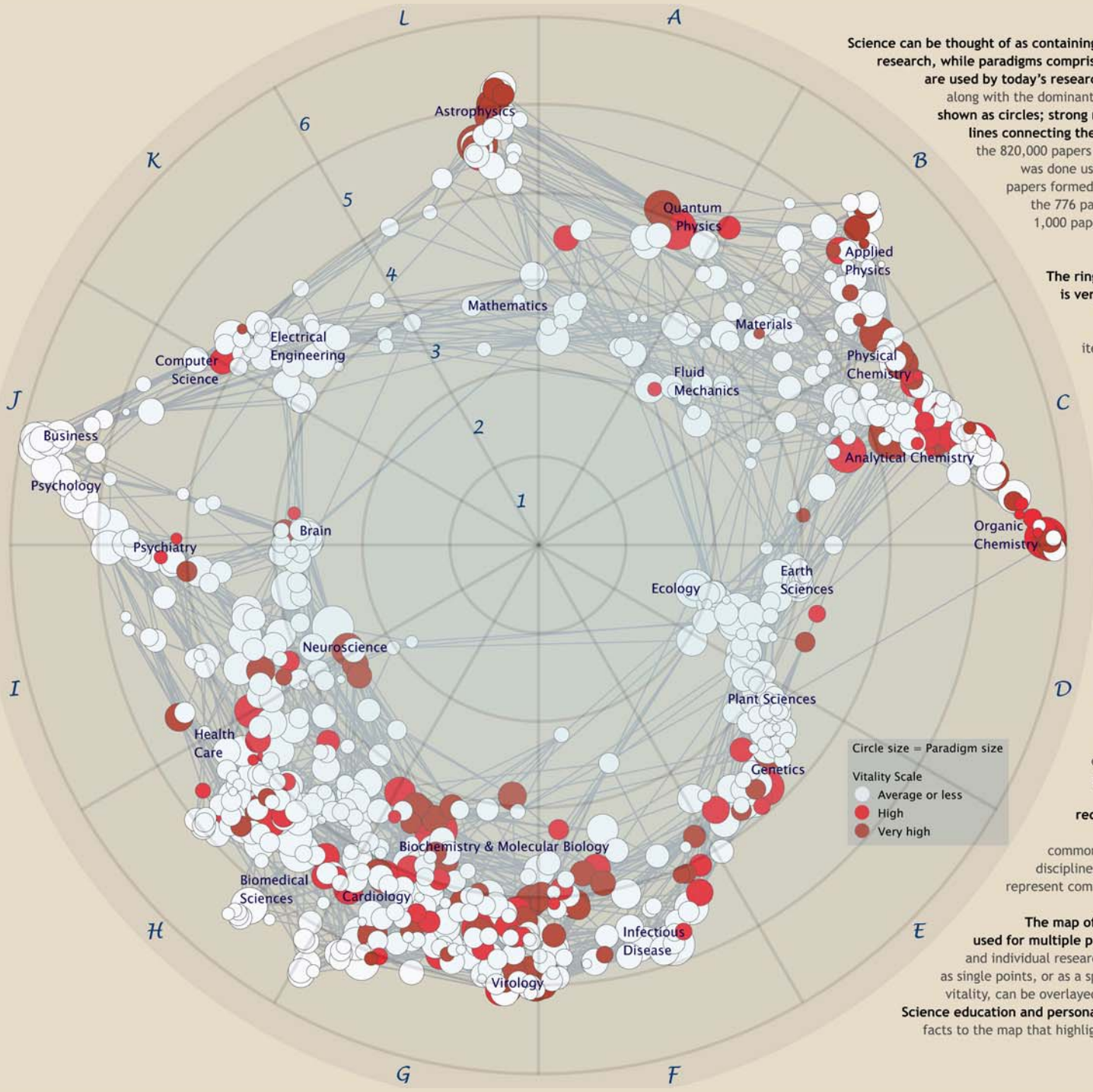
New patents often build on older ideas from many categories. Here, blue lines originate in sixteen different categories that contain the patents cited as prior art for a patent on "gold nanoshells." Gold nanoshells are a new invention: tiny spheres (with a diameter ten million times smaller than a human hair) that can be used to make tumors more visible in infrared scans, and have even helped cause complete remission of tumors in tests with laboratory mice. The blue lines show that widely separated categories provided background for this invention.

Keeping categories understandable is an important part of maintaining any taxonomy, including the patent hierarchy. Categories are easier to understand, search, and maintain if they contain elements (patents in this case) that fit well within the definition of the category. The box above shows a tiny bar chart, part of a "Taxonomy Validator" that helps people decide whether categories are good ones.

Categories can be redefined or combined, and sometimes need to be split when they becomes too large; a constant problem shared by many classifications systems in this information-rich century. But how can we determine exactly where to split a category in two, for example--if there are hundreds or thousands of elements in it?

The Taxonomy Validator measures a "distance to prototype:" how far each element is from an idealized "prototype" element for each bucket. This can be based on statistics, computational comparisons of words, or even human judgement. A simple bar chart can then show how good a category is. A good category has lots of small bars; a generally ragged category is one that might need scrutiny or reorganization; while one that has only one or two tall bars may just mean that one or two elements don't belong. Even simple visualizations like this can ease knowledge work by showing the eye much more than can fit into memory as words; focusing people on just the right issues, and providing a vastly broader background to support more informed judgements.


Synthetic Resins or Natural Rubbe
Ion-exchange Polymer or Process of Prepari
Process of Regenerating
Membrane or Process of Preparing
Previously Formed Solid Ion-exchange Polymer Admixed With N
Polymer Characterized By Defined Size or Shape Other than Bea
Chemically Treated Solid Polymer
Solid Polymer Derived From Ethylenically Unsaturated Reacta
Solid Polymer Derived From At Least One 1,2-epoxy Containir
Solid Polymer Derived From Aldehyde or Derivative
From Ethylenically Unsaturated Reactant Only
From Aldehyde or Derivative
Process of Treating Scrap or Waste Product (
Process of Treating Scrap or Waste Product Containing At Least
Treating Rubber (or Rubberlike Materials) or Polymer Derived
Treating Polymer Derived From A Monomer Containing Only (
Treating Polymer Derived From Hydrocarbon Monomers Only
Treating Polysiloxane
Treating Polyester
Treating With Alcohol
Treating Polyurethane, Polyurea (excluding Urea-formaldehyde
Treating With Alcohol or Amine
Treating Polycarbonamide
Cellular Products or Processes of Preparing A
Cellular Product Derived From Two or More Solid Polymers or Fr
At Least One Polymer Is Derived From Reactant Containing Tw
At Least One Polymer Is Derived From An Aldehyde or Derivat
At Least One Polymer Is Derived From A -n=c=x Reactant Whe



Science can be thought of as containing themes and paradigms. Themes are areas of current research, while paradigms comprise the dominant tool sets and existing knowledge that are used by today's researchers. This map shows **776 major paradigms in science** along with the dominant relationships between these paradigms. **Paradigms are shown as circles; strong relationships between paradigms are indicated by the lines connecting the circles.** The map was created by recursively clustering the 820,000 papers referenced most often in 2003. Clustering at each level was done using VxOrd, a force-directed graph layout routine. These papers formed 53,000 clusters, 6,100 higher-level clusters, and finally the 776 paradigms. Although each paradigm contains, on average, 1,000 papers, some are larger and some are smaller, as shown by different sized circles on the map.

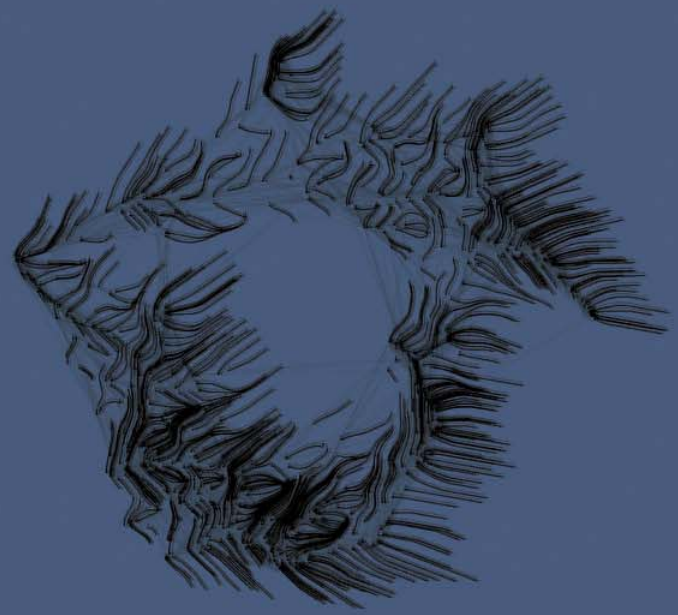
The ring-like structure that is formed by scientific paradigms is very robust. We find similar structures for different years, and for maps generated from scientific journals. "The Structure of Science", a galaxy map shown in the first iteration of Places & Spaces, is a map based on clustering of scientific journals, with superimposition of papers on the journal structure, whereas this map was generated directly from highly-cited papers. "The Structure of Science" shows current science in a disciplinary context, while this map can show the breadth of disciplines that contribute to single paradigms.

Because of the robust nature of the structure of science and its paradigms, we have placed our 776 scientific paradigms within a **reference system containing 12 radial slices and 6 rings.** This allows the position of each paradigm to be codified and available for lookup; for instance ***Fluid Mechanics* paradigms are in grid B3**.

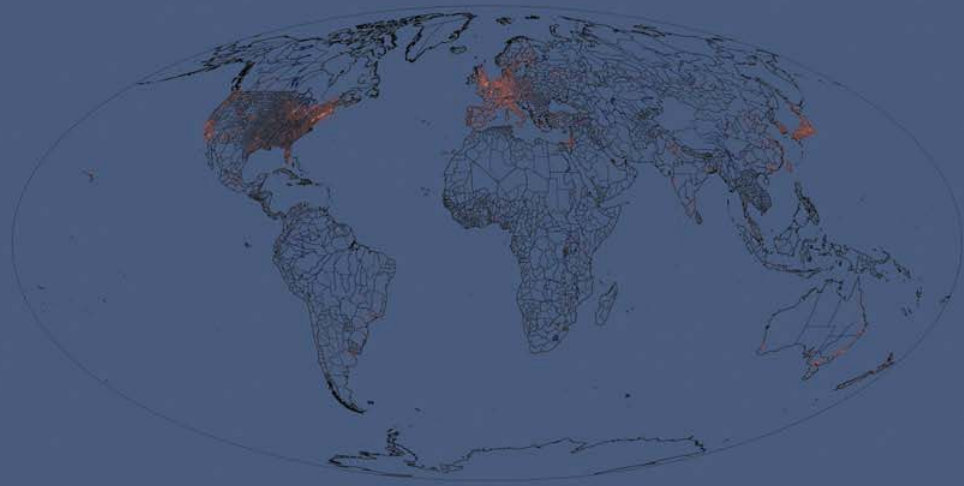
We have also calculated and displayed the vitality of each paradigm. ***Vitality* is a measure of the speed at which a group of researchers reaches consensus about major improvements.** Paradigms are constantly being improved, but it usually takes years to reach consensus about which improvements are major. The white circles represent communities where consensus is reached relatively slowly. This is a common phenomenon in the social sciences, ecological sciences, computer sciences, and mathematics disciplines. **The red circles represent communities of researchers where consensus is reached relatively rapidly.** This is more common in physics, chemistry, biochemistry, and many medical disciplines. Very dark circles (such as those in *Astrophysics, L5-6*) represent communities where consensus is reached extremely quickly.

The map of scientific paradigms and its reference system can be used for multiple purposes. Countries, industries, companies, universities, and individual researchers can all **locate themselves within the map**, either as single points, or as a specific collection of paradigms. Various metrics, such as vitality, can be overlayed on this reference system to highlight specific impacts. **Science education and personal discovery** can also be enhanced by linking stories and facts to the map that highlight scientific history, current advances and relationships between scientific paradigms.

Topic Map: How Scientific Paradigms Relate



Geographic Map: Where Science Gets Done



You may run your finger over each of these maps to control the lighting on the other: touching a place on the world map will light up topics studied in that place; touching a paradigm on the topic map will light up the places that study that topic.

Nanotechnology

This overlay shows the distribution of nanotechnology within the paradigms of science. The majority of current work in nanotechnology takes places in physics, chemistry, and materials science, at the upper right portion of the map. However, an increasing amount of nanotechnology is being applied in the biological and medical sciences, at the lower right.

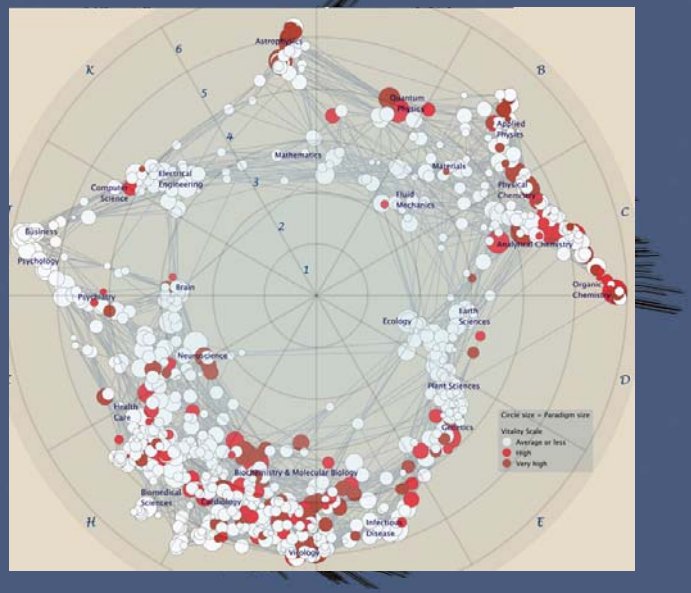
All Topics *Sweep through all 776 scientific paradigms*	**Nanotechnology** *Science on the tiny scale of molecules*
Sustainability *The science behind our long-term hopes*	**Biology & Chemistry** *The interface between these two vital fields*

We sweep slowly through adjoining related topics, lighting up the places in the world that study each topic. You may select a subset of the topics that deal with these three interesting subjects by touching it.

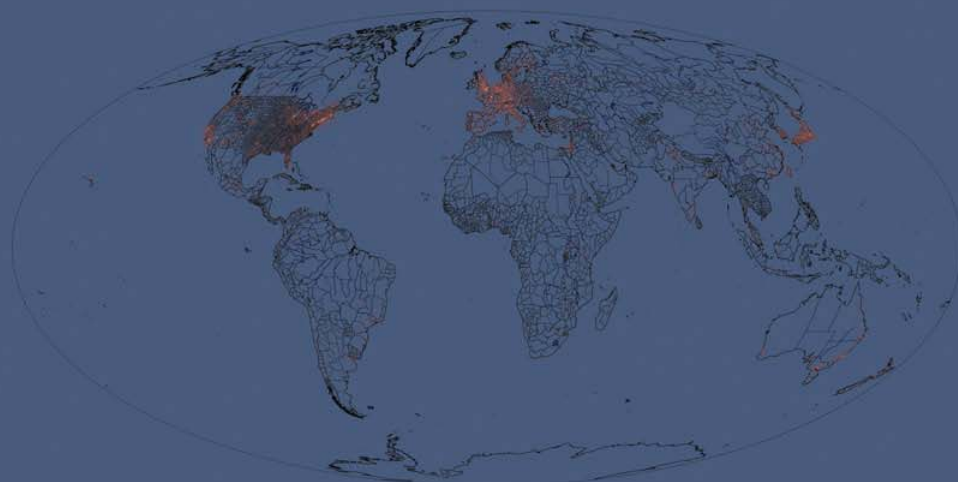
Francis H. C. CRICK *Co-discovered DNA's double helix*	**Albert EINSTEIN** *Revitalized physics with Relativity theories*	**Michael E. FISHER** *Models critical phase transitions of matter*	**Susan T. FISKE** *Connects perception and stereotypes*
Joshua LEDERBERG *Pioneer in bacterial genetic mechanisms*	**Derek J. de Solla PRICE** *Known as the "Father of Scientometrics"*	**Richard N. ZARE** *Uses laser chemistry in molecular dynamics*	**About this display** *People & organizations that helped create it*

A single person's spreading influence is shown as a series of four snapshots. First, we light only topics and places relating to that person's papers—papers that are still highly cited today. The second lights everything that cites that original work. Note that this first-generation impact extends to far more topics than did the original work. The third shapshot lights science that cites the second; and the fourth lights science that cites the third.

Topic Map: How Scientific Paradigms Relate



Geographic Map: Where Science Gets Done



You may run your finger over each of these maps to control the lighting on the other: touching a place on the world map will light up topics studied in that place; touching a paradigm on the topic map will light up the places that study that topic.

Nanotechnology

This overlay shows the distribution of nanotechnology within the paradigms of science. The majority of current work in nanotechnology takes places in physics, chemistry, and materials science, at the upper right portion of the map. However, an increasing amount of nanotechnology is being applied in the biological and medical sciences, at the lower right.

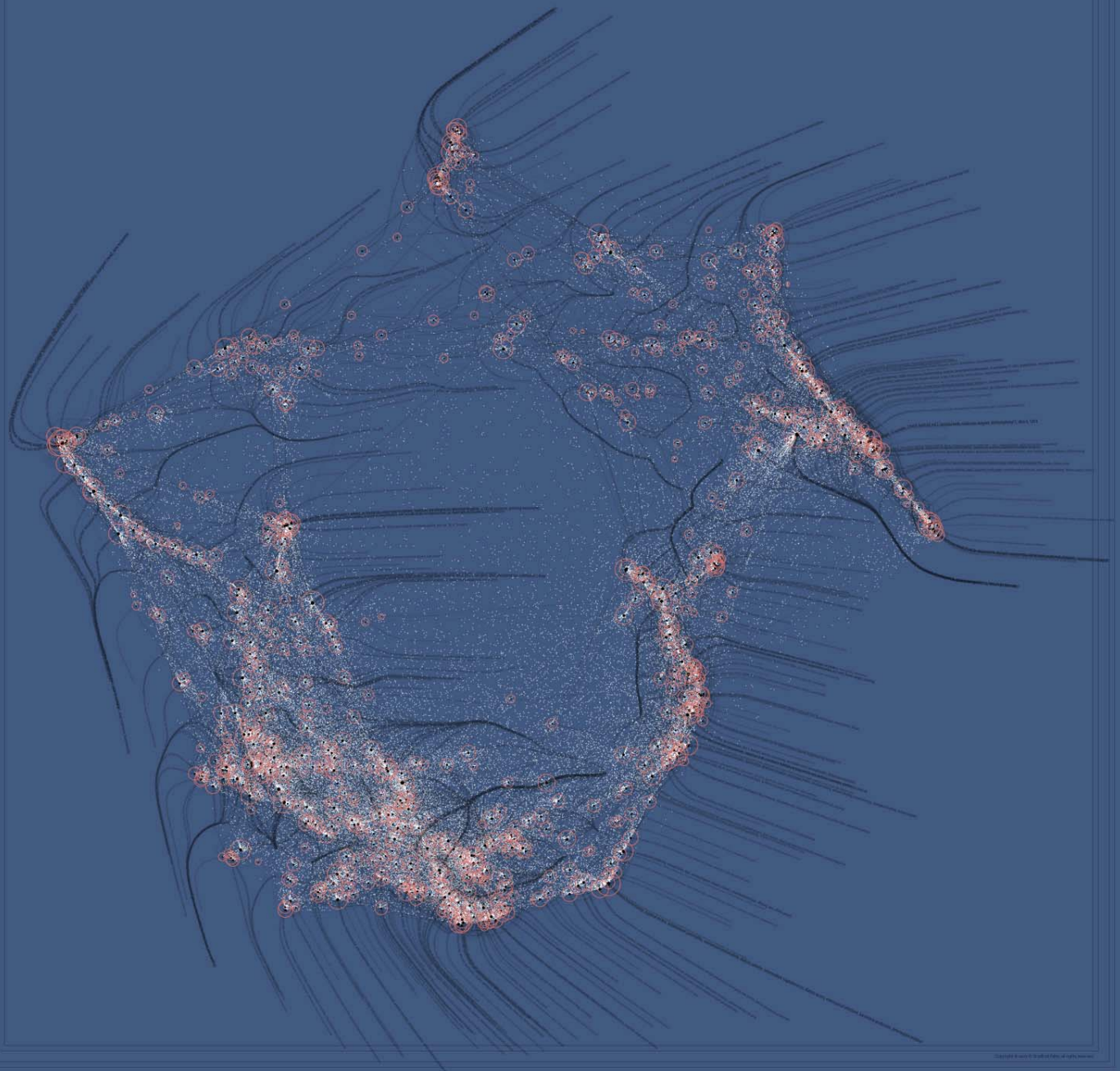
All Topics *Sweep through all 776 scientific paradigms*	**Nanotechnology** *Science on the tiny scale of molecules*
Sustainability *The science behind our long-term hopes*	**Biology & Chemistry** *The interface between these two vital fields*

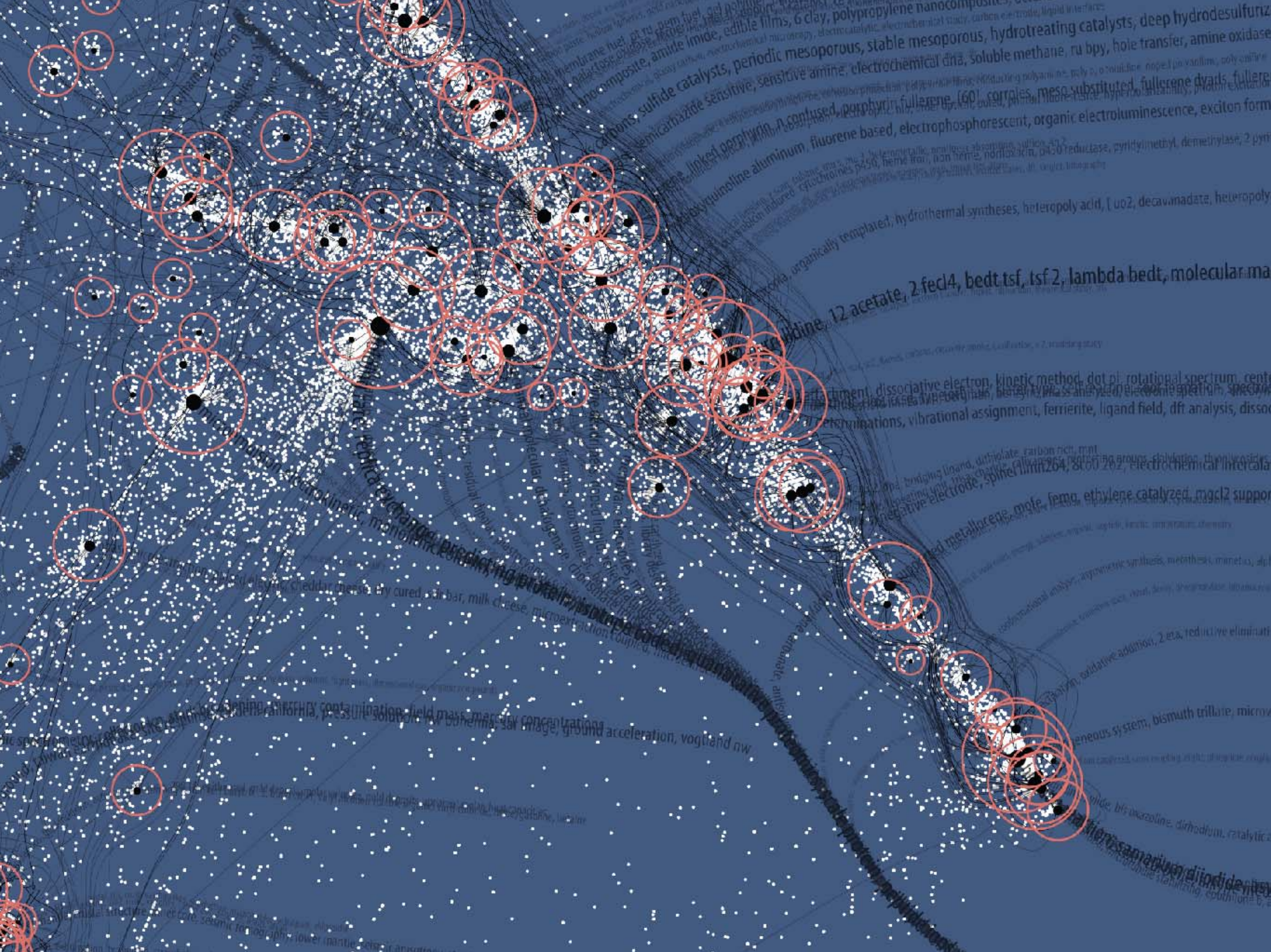
We sweep slowly through adjoining related topics, lighting up the places in the world that study each topic. You may select a subset of the topics that deal with these three interesting subjects by touching it.

Francis H. C. CRICK *Co-discovered DNA's double helix*	**Albert EINSTEIN** *Revitalized physics with Relativity theories*	**Michael E. FISHER** *Models critical phase transitions of matter*	**Susan T. FISKE** *Connects perception and stereotypes*
Joshua LEDERBERG *Pioneer in bacterial genetic mechanisms*	**Derek J. de Solla PRICE** *Known as the "Father of Scientometrics"*	**Richard N. ZARE** *Uses laser chemistry in molecular dynamics*	**About this display** *People & organizations that helped create it*

A single person's spreading influence is shown as a series of four snapshots. First, we light only topics and places relating to that person's papers—papers that are still highly cited today. The second lights everything that cites that original work. Note that this first-generation impact extends to far more topics than did the original work. The third shapshot lights science that cites the second; and the fourth lights science that cites the third.

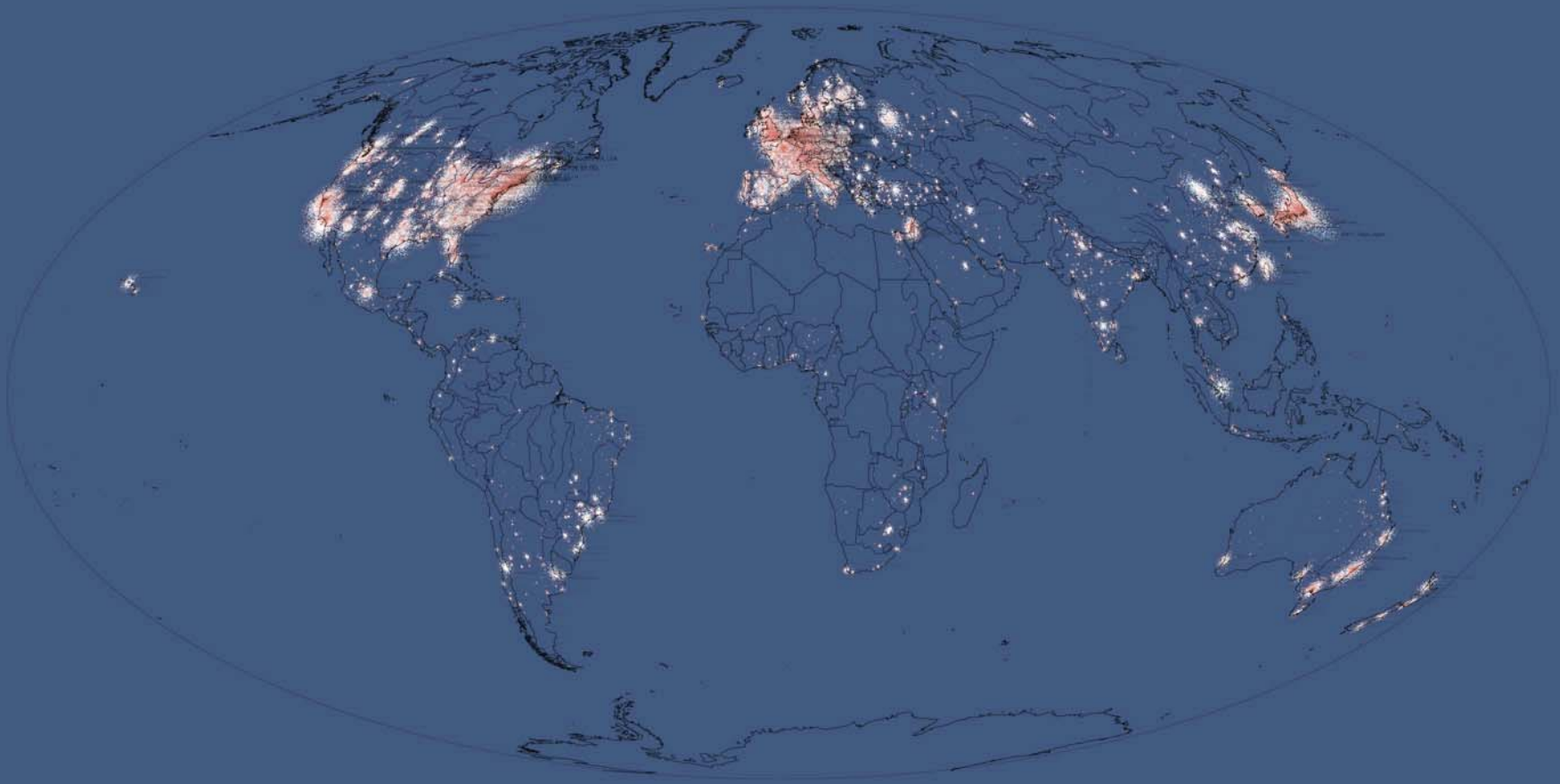
Topic Map: How Scientific Paradigms Relate

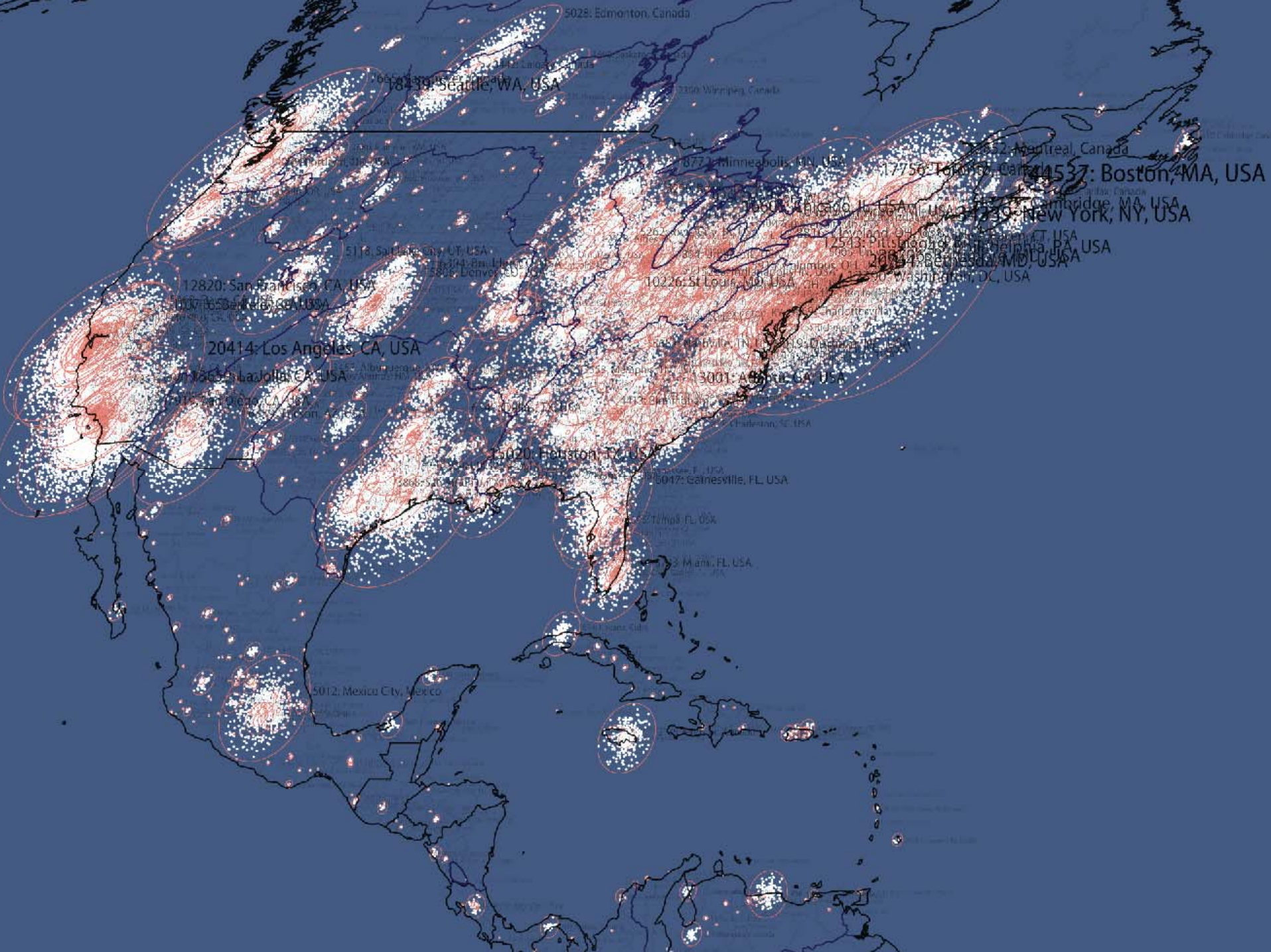


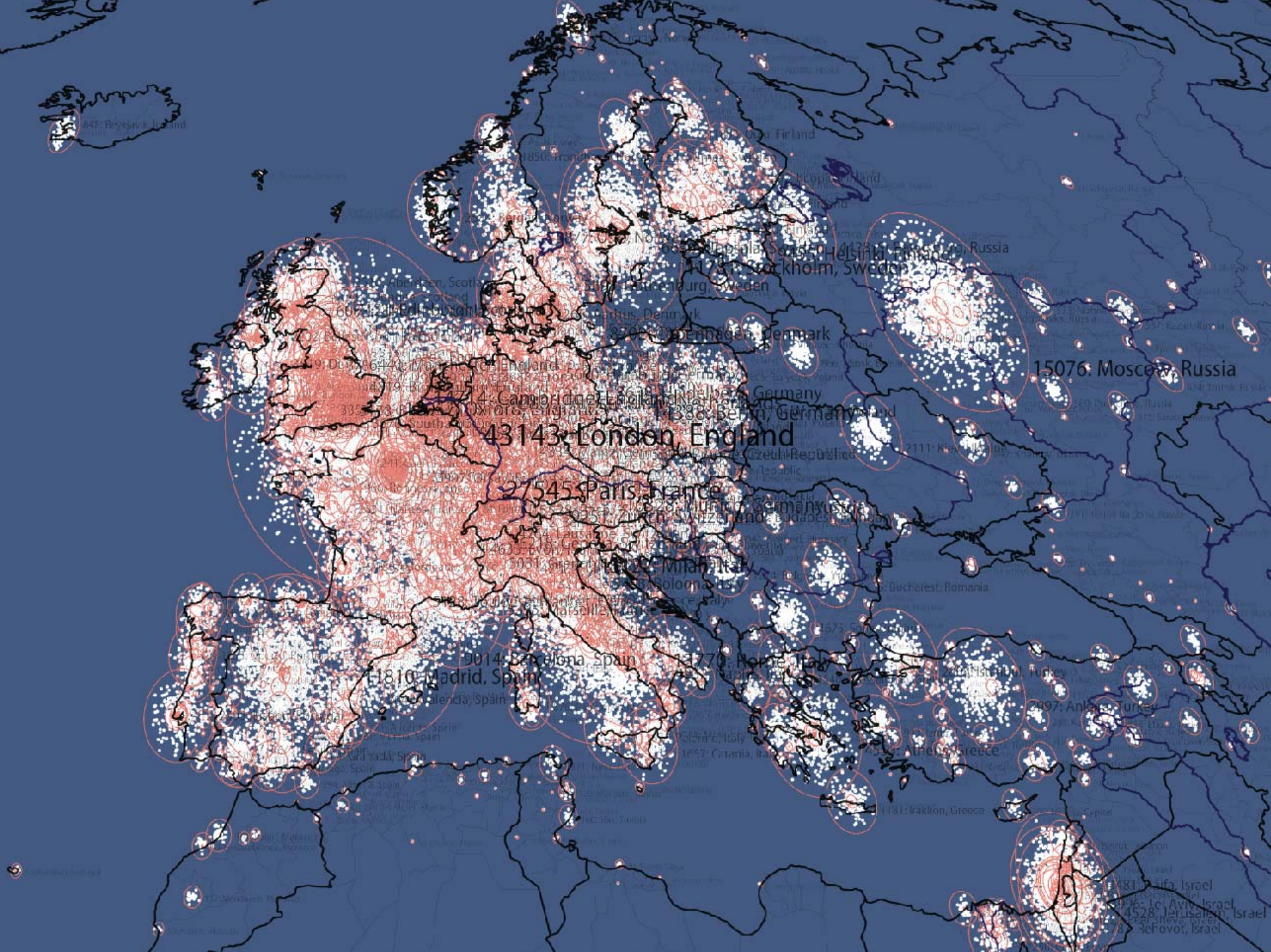


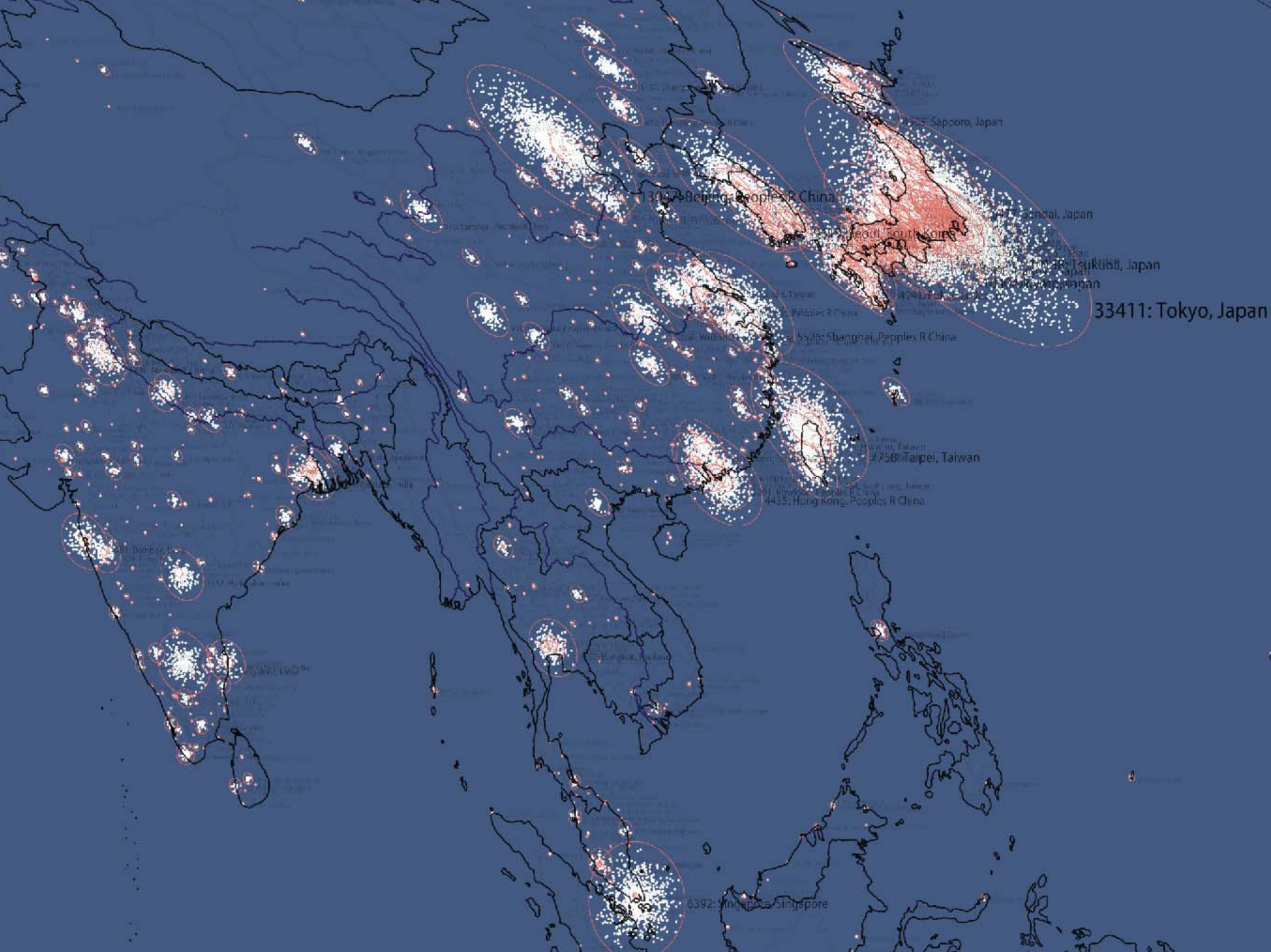
nanocomposite, amide imide, edible films, 6 clay, polypropylene
sulfide catalysts, periodic mesoporous, stable mesoporous, hydrotreating catalysts
soluble methane, ru bpy, hole transfer, amine oxidase
n confused, porphyrin fullerene, [60], corroles, meso substituted, fullerene dyads
fluorene based, electrophosphorescent, organic electroluminescence, exciton
hydrothermal syntheses, heteropoly acid, [uo2, decavanadate, heteropoly
12 acetate, 2 fecl4, bedt tsf, tsf 2, lambda bedt, molecular
dissociative electron, kinetic method, dot pi, rotational spectrum
vibrational assignment, ferrierite, ligand field, dft analysis
mole, femo, ethylene catalyzed, mgcl2
cheddar cheese, dry cured
mercury contamination, field mass, mercury concentrations
ground acceleration, vogtland nw
bismuth triflate
samarium diiodide

Geographic Map: Where Science Gets Done




5028: Edmonton, Canada
18439: Seattle, WA, USA
8771: Minneapolis, MN, USA
17756: Toronto, Canada
14537: Boston, MA, USA
12820: San Francisco, CA, USA
20414: Los Angeles, CA, USA
10226: St Louis, MO, USA
12513: Pittsburgh, PA, USA
13001: Atlanta, GA, USA
5020: Houston, TX, USA
5017: Gainesville, FL, USA
5012: Mexico City, Mexico


15076: Moscow, Russia
43143: London, England
27545: Paris, France
11810: Madrid, Spain

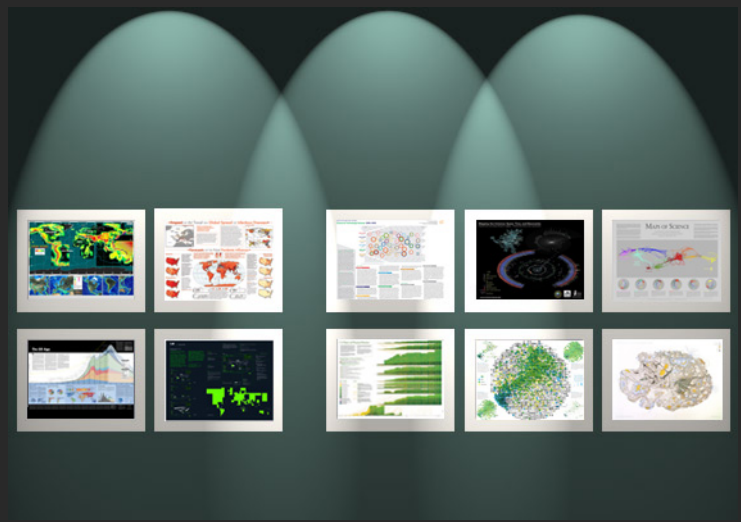

33411: Tokyo, Japan

Illuminated Diagram Display

http://www.youtube.com/watch?v=bXABcOABG4E

The Power of Forecasts

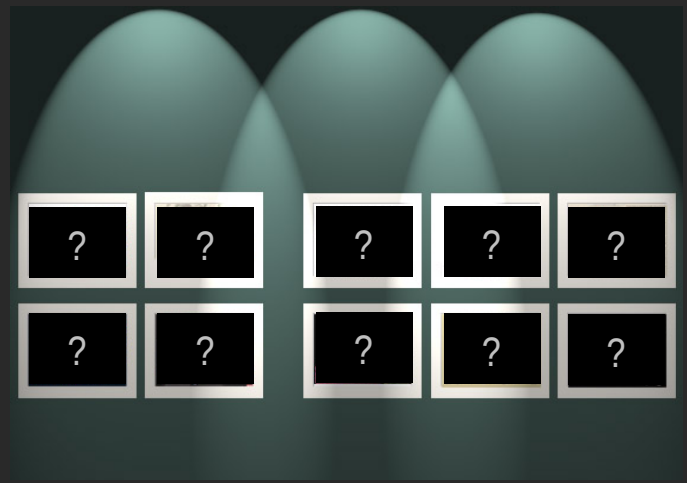
Four Existing Forecasts
VERSUS
Six Potential Science 'Weather' Forecasts



(3rd Iteration of Places & Spaces Exhibit - 2007)

Science Maps for Economic Decision Making

Four Existing Maps
VERSUS
Six Science Maps



(4th Iteration of Places & Spaces Exhibit - 2008)

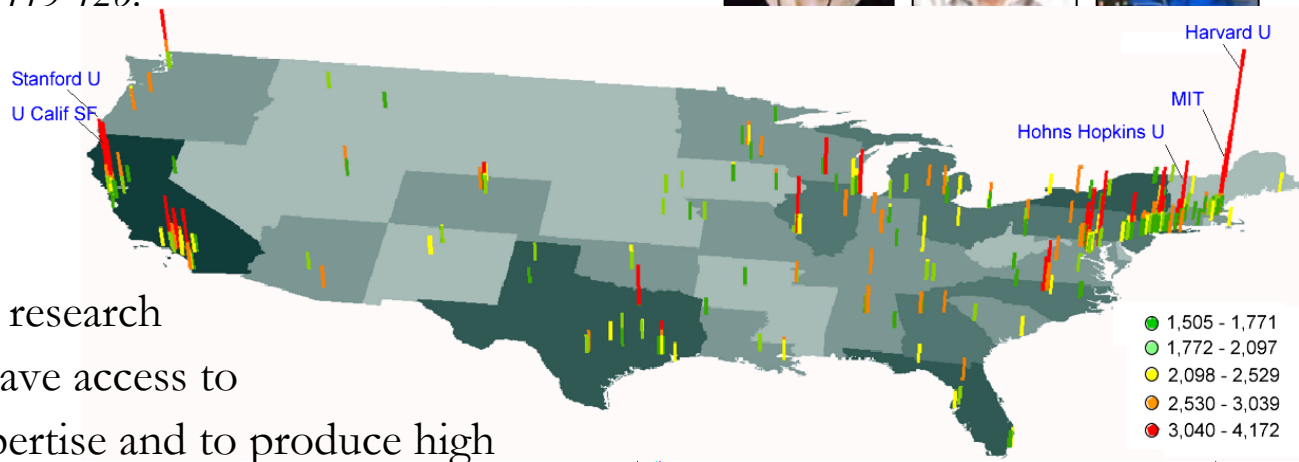
Science Maps in Action

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.

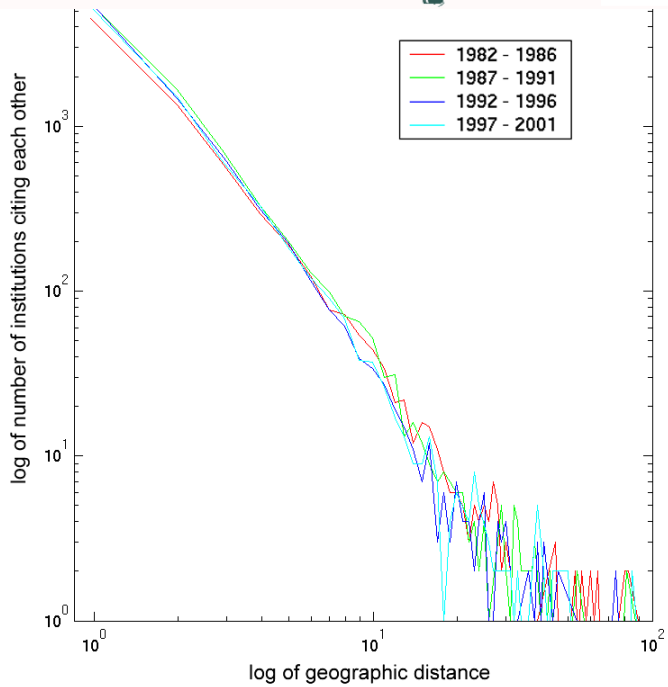
Research questions:

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.



Mapping Topic Bursts

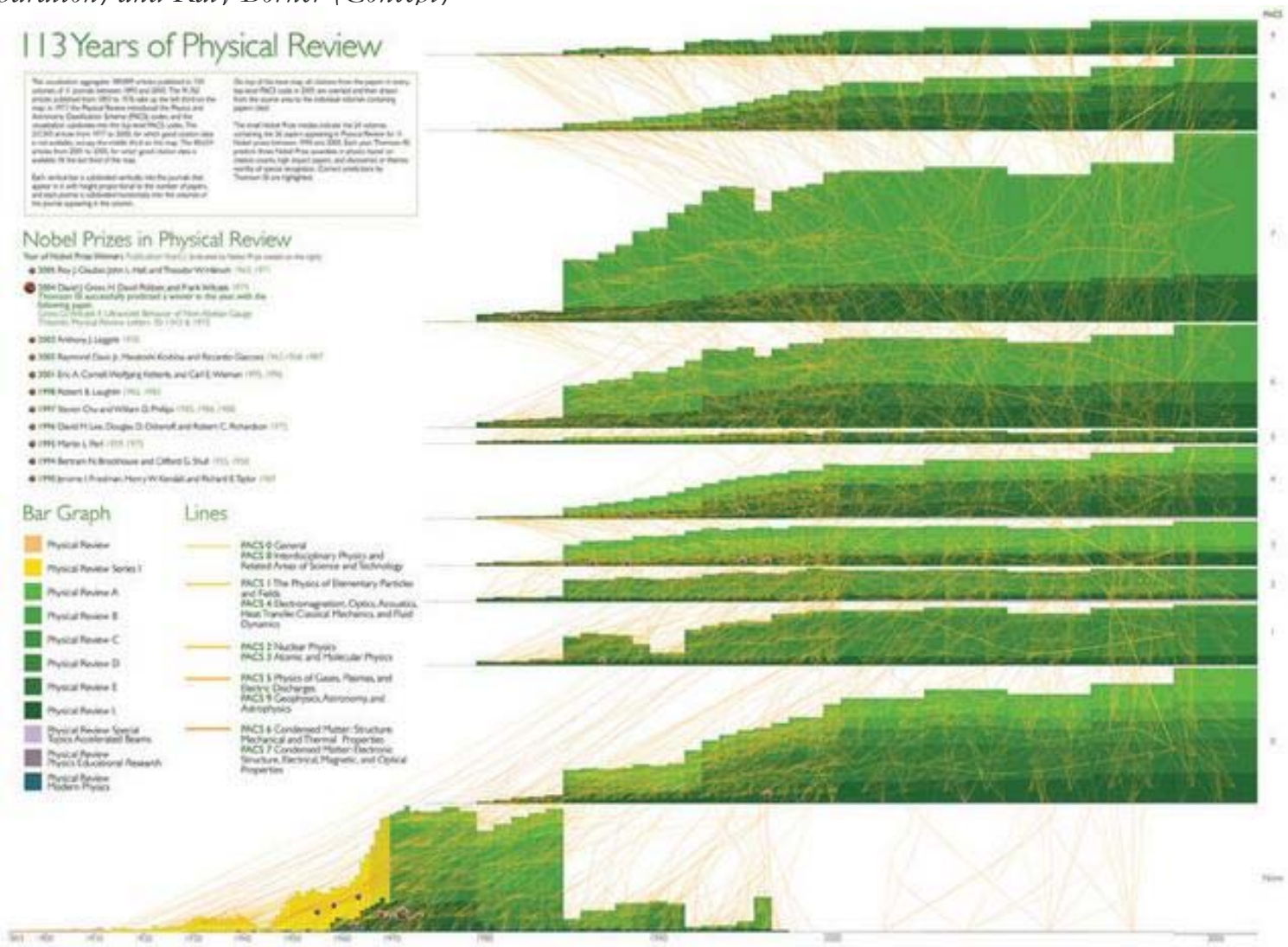
Co-word space of the top 50 highly frequent and bursty words used in the top 10% most highly cited PNAS publications in 1982-2001.

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

113 Years of Physical Review

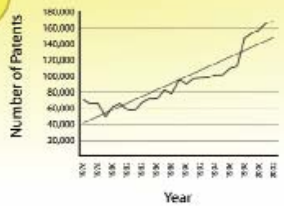
http://scimaps.org/dev/map_detail.php?map_id=171

Bruce W. Herr II and Russell Duhon (Data Mining & Visualization), Elisha F. Hardy (Graphic Design), Shashikant Penumarthy (Data Preparation) and Katy Börner (Concept)



Examining the Evolution and Distribution of Patent Classifications

1 Patents Granted Over the Last 20 Years



Top Classes 1978 - 1982

Class	Title	Patents
260	Chemistry of Carbon Compounds	13,348
424	Drug, Bio-Affecting and Body Treating Compositions	9,621
252	Compositions	5,868
128	Surgery	5,404
428	Stock Material or Miscellaneous Articles	4,984
73	Measuring and Testing	4,895
123	Internal-Combustion Engines	4,142
340	Communications: Electrical	4,043
364	Electrical Computers and Data Processing Systems	3,700
29	Metal Working	3,664
	Total	55,705

Top Classes 1998 - 2002

Class	Title	Patents
514	Drug, Bio-Affecting and Body Treating Compositions	18,778
438	Semiconductor Device Manufacturing Process	17,775
435	Chemistry, Molecular Biology and Microbiology	17,424
424	Drug, Bio-Affecting and Body Treating Compositions	13,697
428	Stock Material or Miscellaneous Articles	13,316
257	Active Solid-State Devices (e.g., Transistors, Solid-state Diodes	12,024
395	Information Processing System Organization	9,955
345	Computer Graphics Processing, Operator Interface Processing, and Selective Visual Display Systems	9,510
359	Optical: Systems and Elements	9,191
365	Static Information Storage and Retrieval	8,392
	Total	140,910

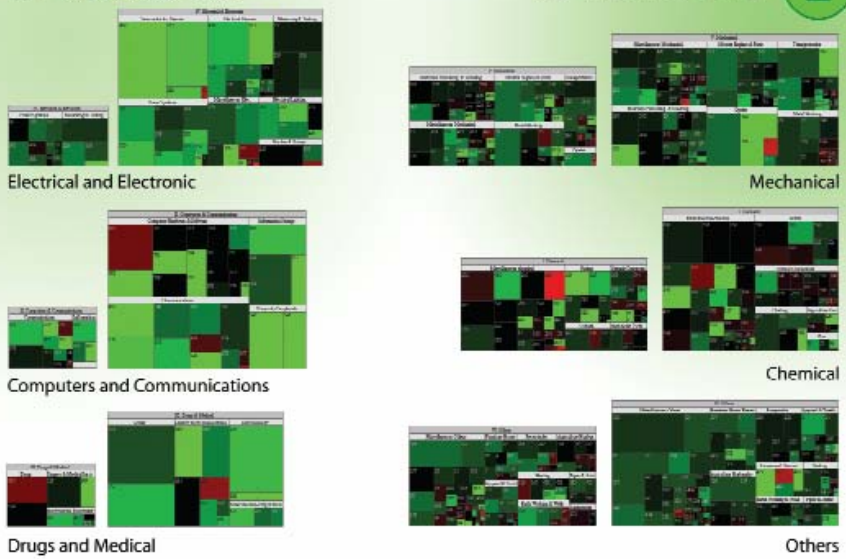
In the United States, each patent gets assigned to one out of more than 450 classes covering broad application domains. An examination of the size and growth of patent classes provides insight about patenting trends.

Treemaps, a space filling technique developed in the HCI Lab at the University of Maryland, are used to communicate major results. Treemaps represent a tree structure as nested rectangles with each rectangle representing a node. A rectangular area is first allocated to hold the representation of the tree, and this area is then subdivided into a set of rectangles that represent the top level of the tree. This process continues recursively on the resulting rectangles to represent each lower level of the tree. The parent-child relationship is indicated by enclosing the child rectangle by its parent rectangle. Typically, the size of each rectangle corresponds to the size of the node. Additional information about a node, e.g., its age or value, can be represented by the color of the respective rectangle.

Fast Growth Domains
1983 - 1987 / 1998 - 2002

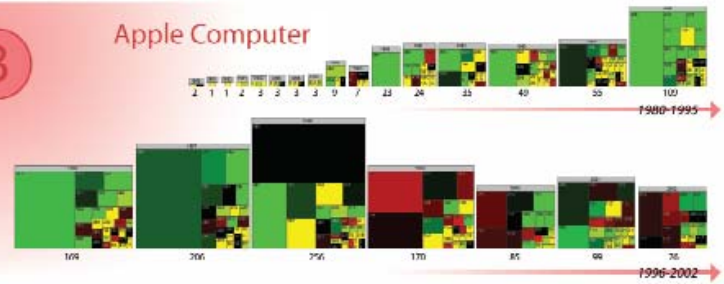
Slow Growth Domains
1983 - 1987 / 1998 - 2002

2



Shown is a comparison of the patent class space for 1983 to 1987 and 1998 to 2002. There is a predominance of growth in the 1998 to 2002 patent space, which correlates to the increase in patent grants during this period. By comparing the growth in categories, one can distinguish between domains that have been receiving a larger amount of patent grants.

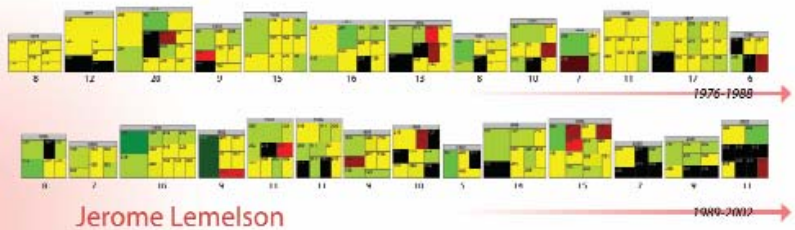
3 Apple Computer



Depicted above is how Apple Computers' portfolio has changed in yearly increments from 1980 to 2002.

Lemelson's patent holdings below show a more even distribution over multiple classes. No class dominates over a majority of the years for granted patents; instead they are distributed more broadly over the intellectual space.

Legend:
Green = Increase in number of patent grants in particular class.
Red = Decrease in number of patent grants in particular class.
Yellow = No patents granted in that class in the past five years.
Size = Number of patent grants in a particular class.

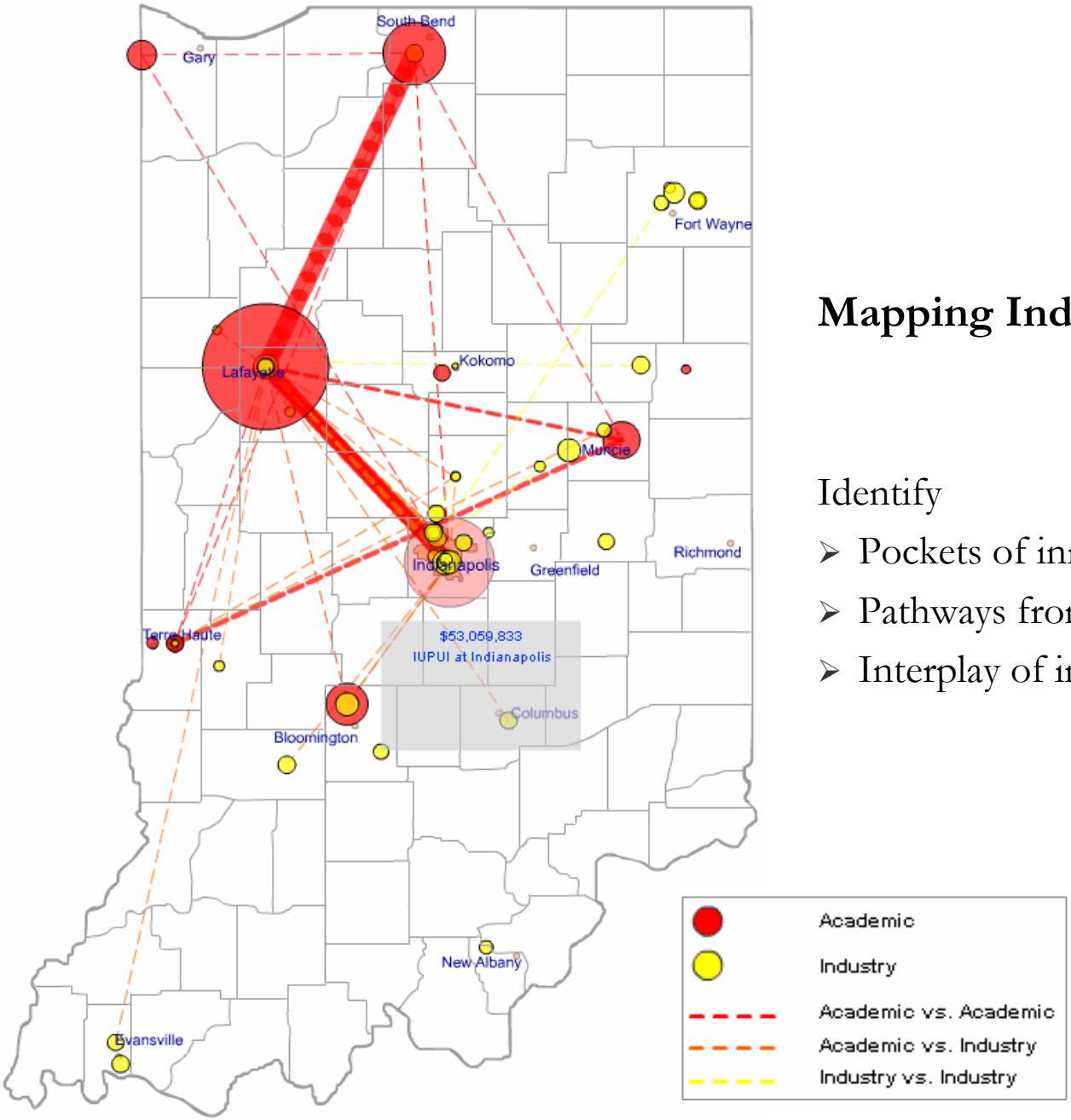


Jerome Lemelson

Kutz, Daniel O. Examining the Evolution and Distribution of Patent Classifications. Accepted for the Information Visualization Conference, London, UK, July 2004.

The material is based upon work supported bythe National Science Foundation under Grant No. IIS-0238261.

For more Information, contact Katy Borner at katy@indiana.edu.



Mapping Indiana's Intellectual Space

Identify

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

Wikipedian Activity

Studying large scale social networks such as Wikipedia

Vizzards 2007 Entry

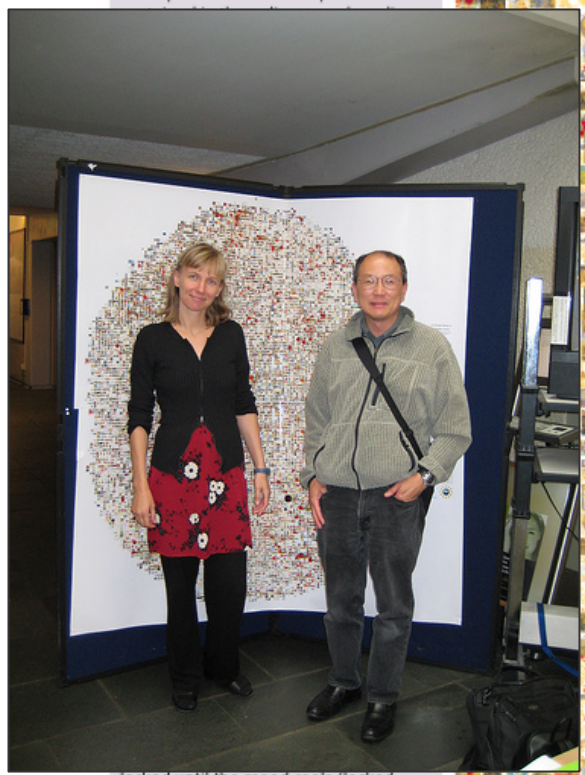
Second Sight: An Emergent Mosaic of Wikipedian Activity, The NewScientist, May 19, 2007

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



locked until the mood cools (locked 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.

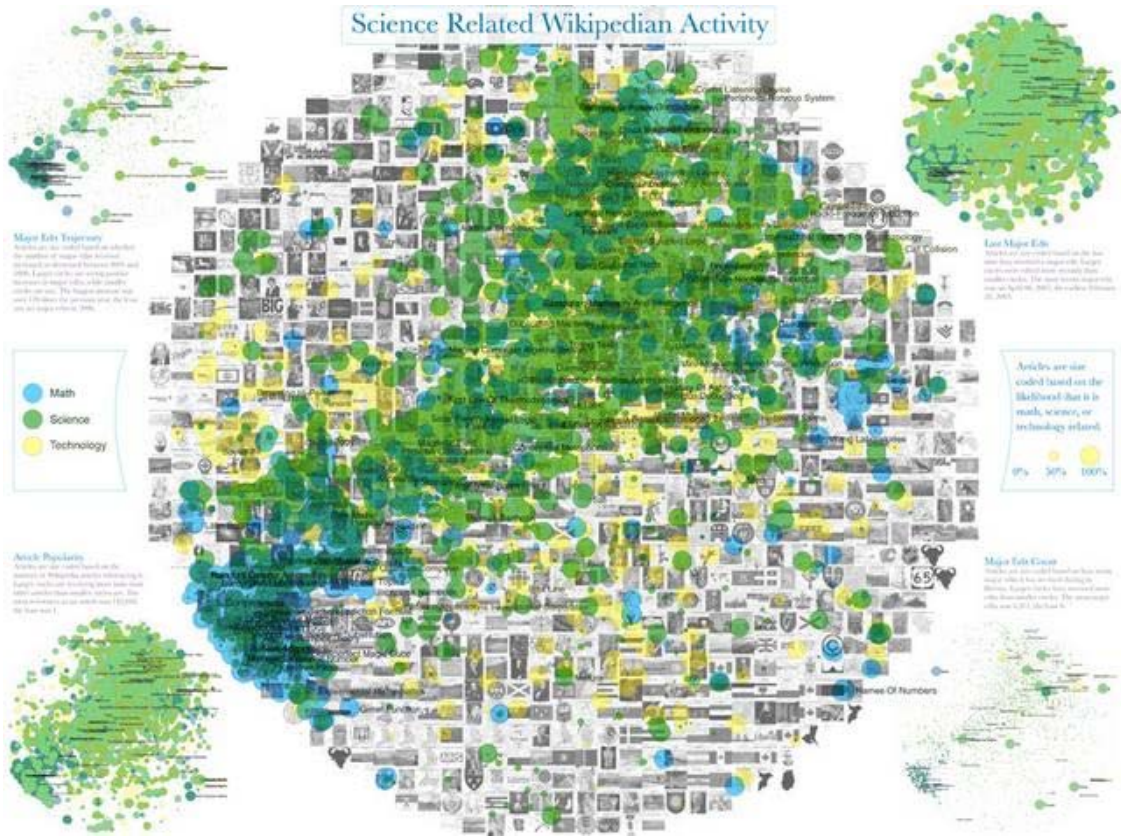
Science Related Wikipedian Activity

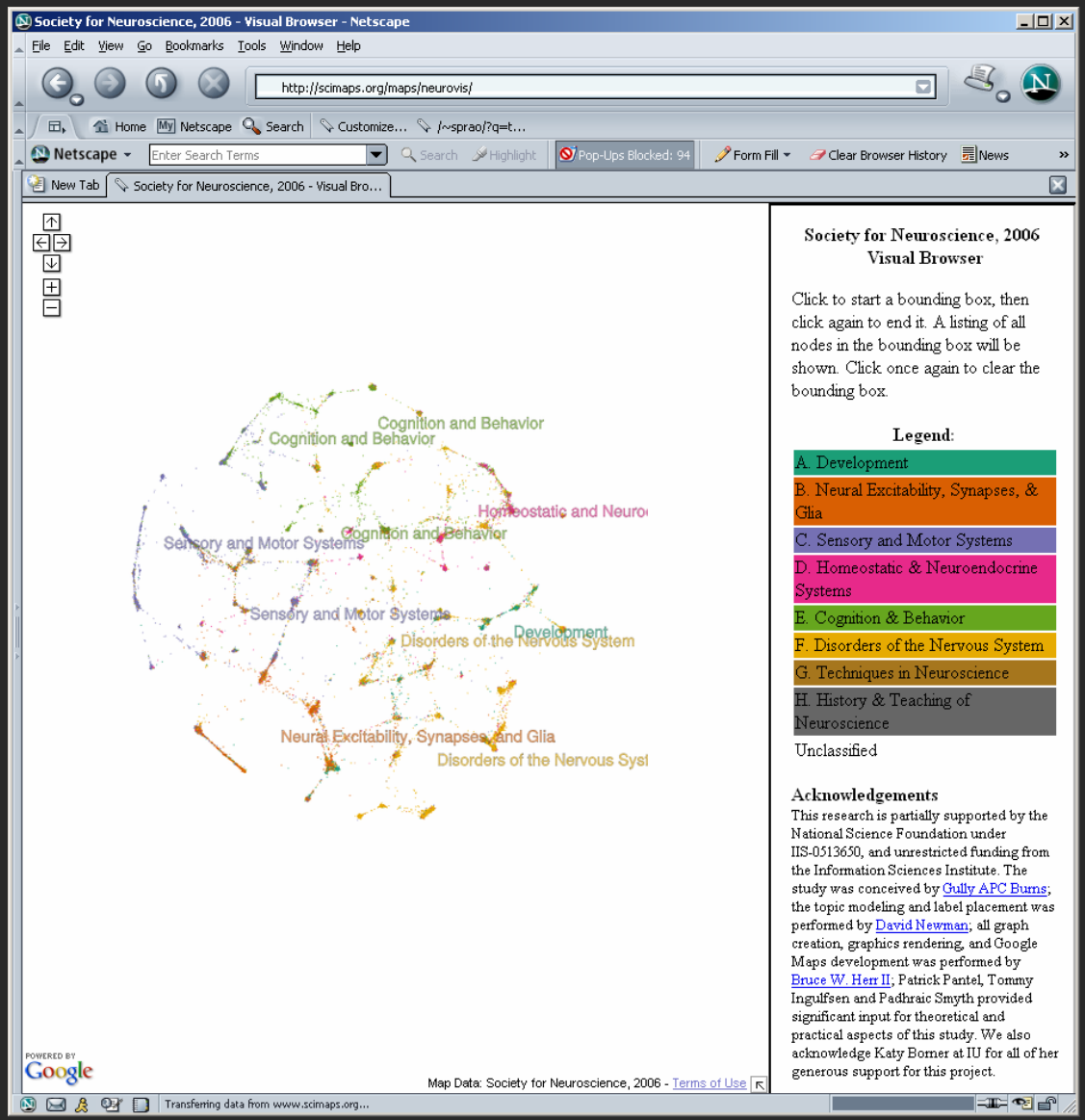
http://scimaps.org/dev/map_detail.php?map_id=165

Same base map.

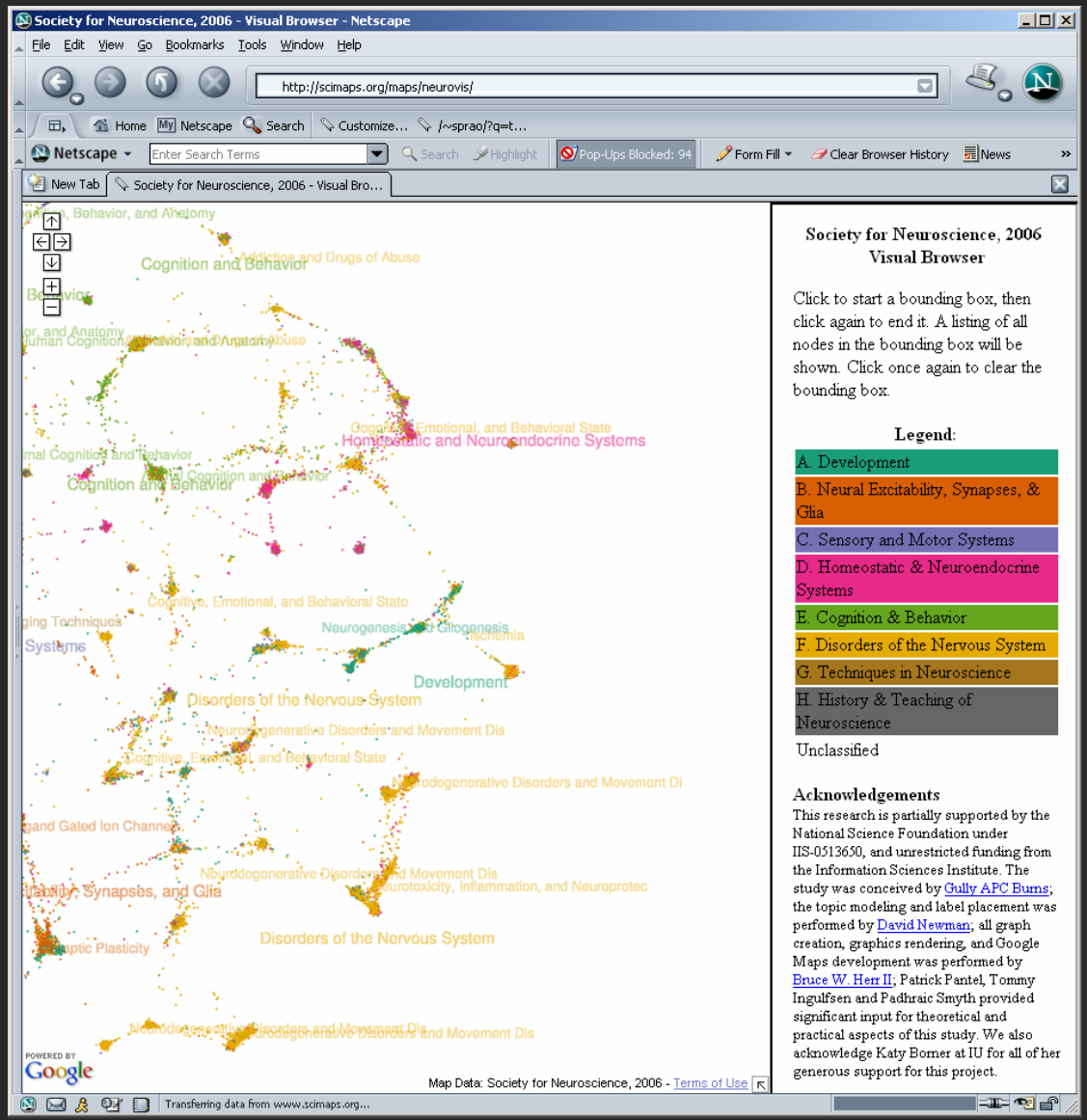
Overlaid are 3,599 math (blue), 6,474 science (green), and 3,164 technology relevant articles (yellow).
All other articles are given in grey.

Corners show articles size coded according to
-article edit activity (top left),
- number of major edits (top right),
- number of bursts in edit activity (bottom, right)
- indegree (bottom left).

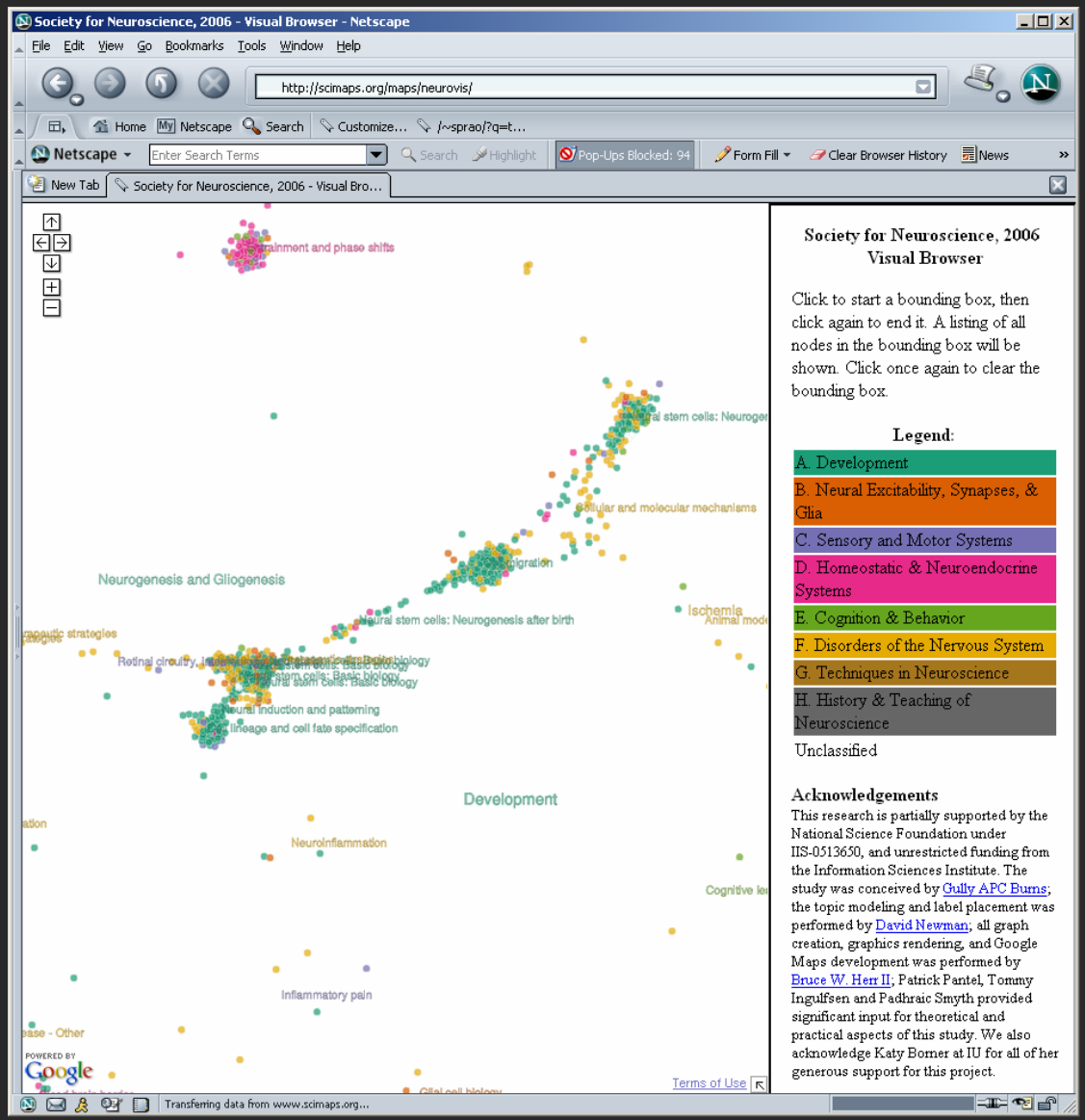




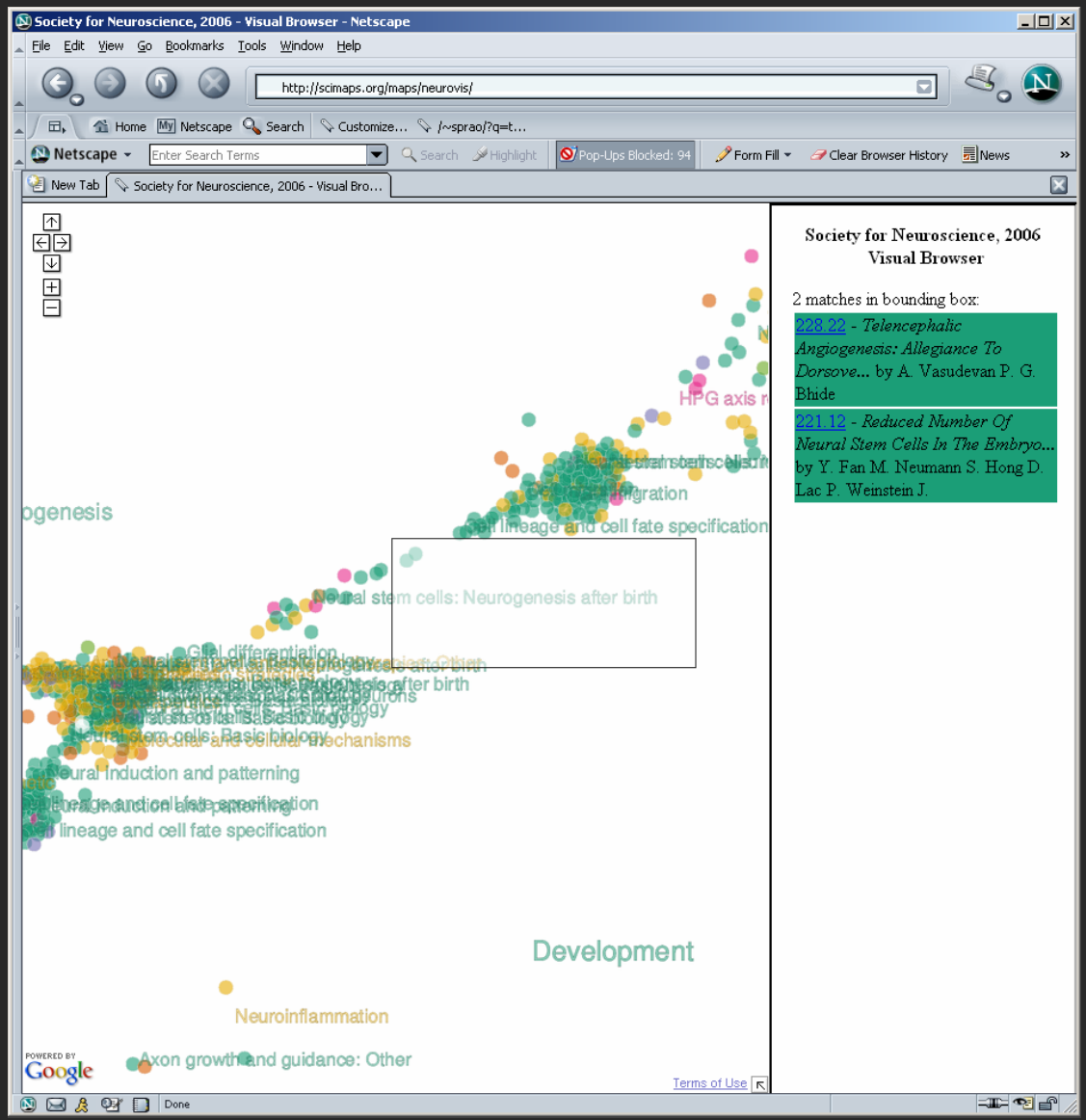
Bruce W. Herr II, Gully Burns (USC), David Newman (UCI), Society for Neuroscience, 2006 Visual Browser, 2007, http://scimaps.org/maps/neurovis/



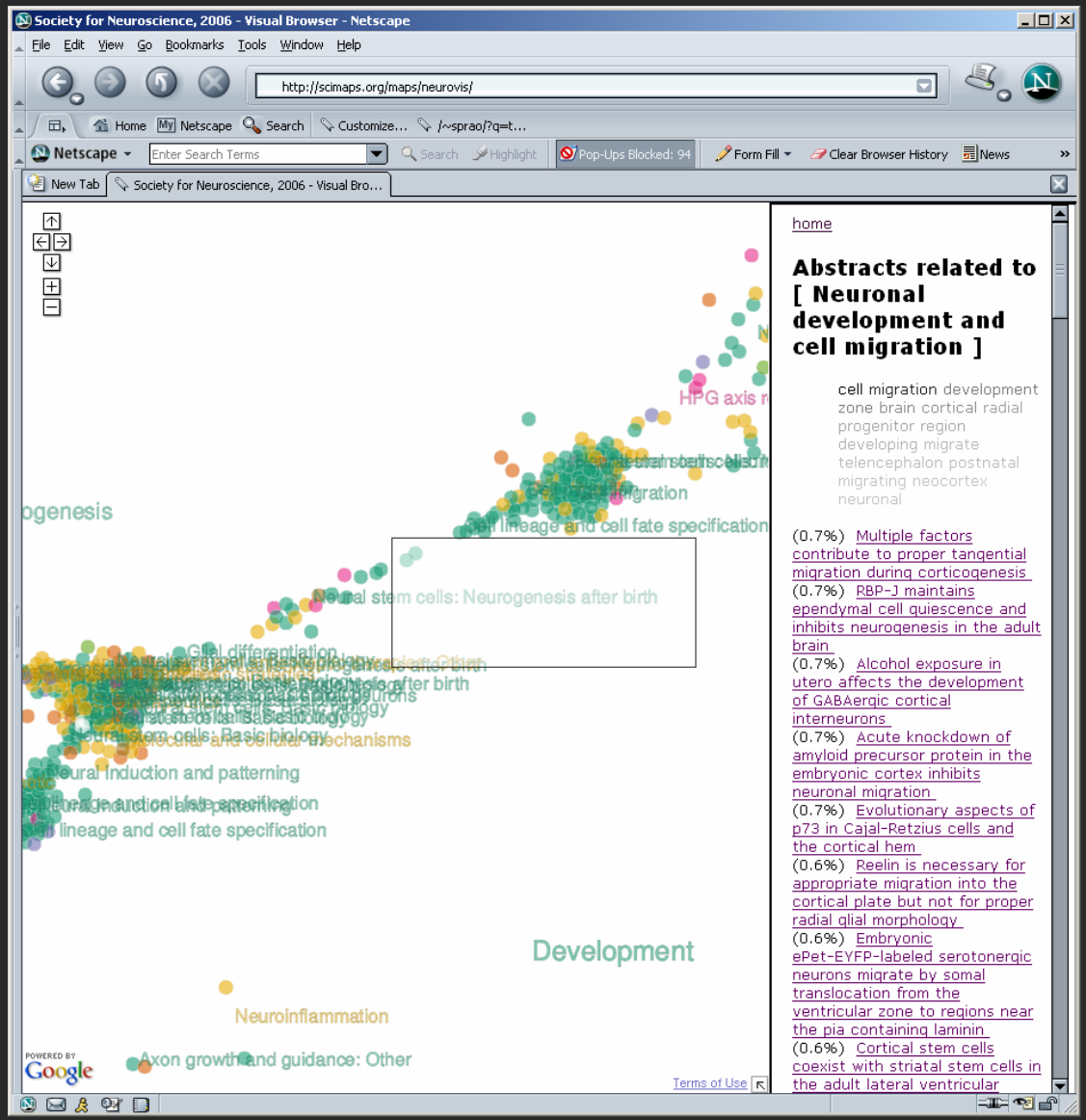
Bruce W. Herr II, Gully Burns (USC), David Newman (UCI), Society for Neuroscience, 2006 Visual Browser, 2007, http://scimaps.org/maps/neurovis/



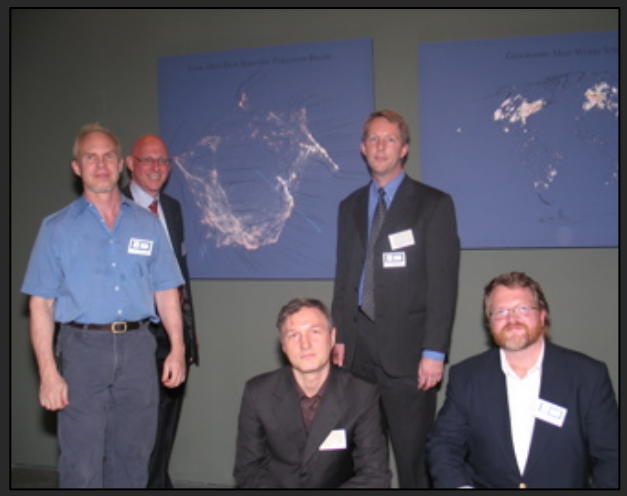
Bruce W. Herr II, Gully Burns (USC), David Newman (UCI), Society for Neuroscience, 2006 Visual Browser, 2007, http://scimaps.org/maps/neurovis/



Bruce W. Herr II, Gully Burns (USC), David Newman (UCI), Society for Neuroscience, 2006 Visual Browser, 2007, http://scimaps.org/maps/neurovis/

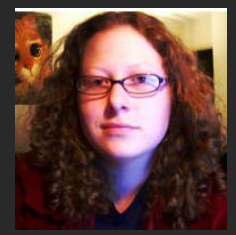


Bruce W. Herr II, Gully Burns (USC), David Newman (UCI), Society for Neuroscience, 2006 Visual Browser, 2007, http://scimaps.org/maps/neurovis/



Interested to get your own science map?
Contact the map makers!

katy@indiana.edu



The End.