Science of Science Research and Tools
## Tutorial #05 of 12

Dr. Katy Börner
Cyberinfrastructure for Network Science Center, Director
Information Visualization Laboratory, Director
School of Library and Information Science
Indiana University, Bloomington, IN
http://info.slis.indiana.edu/~katy

With special thanks to Kevin W. Boyack, Micah Linnemeier,
Russell J. Duhon, Patrick Phillips, Joseph Biberstine, Chintan Tank
Nianli Ma, Hanning Guo, Mark A. Price, Angela M. Zoss, and
Scott Weingart

Invited by Robin M. Wagner, Ph.D., M.S.
Chief Reporting Branch, Division of Information Services
Office of Research Information Systems, Office of Extramural Research
Office of the Director, National Institutes of Health

Suite 4090, 6705 Rockledge Drive, Bethesda, MD 20892
10a-noon, July 13, 2010

---
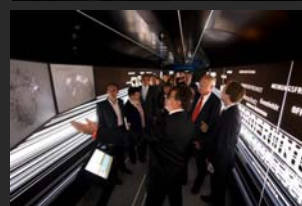
12 Tutorials in 12 Days at NIH—Overview

1.	Science of Science Research	**1st Week**
2.	Information Visualization	
3.	CIShell Powered Tools: Network Workbench and Science of Science Tool	
4.	Temporal Analysis—Burst Detection	**2nd Week**
5.	Geospatial Analysis and Mapping	
6.	Topical Analysis & Mapping	
7.	Tree Analysis and Visualization	**3rd Week**
8.	Network Analysis	
9.	Large Network Analysis	
10.	Using the Scholarly Database at IU	**4th Week**
11.	VIVO National Researcher Networking	
12.	Future Developments	

[#05] Geospatial Analysis and Mapping

➢ General Overview
➢ Designing Effective Geomaps
➢ Sci2-Geomaps With Circle and Colored Region Annotation
➢ Sci2-Animations
➢ Geographic Information Systems (GIS)
➢ Outlook
➢ Exercise: Identify Promising Geospatial Analyses of NIH Data
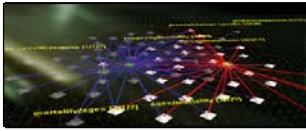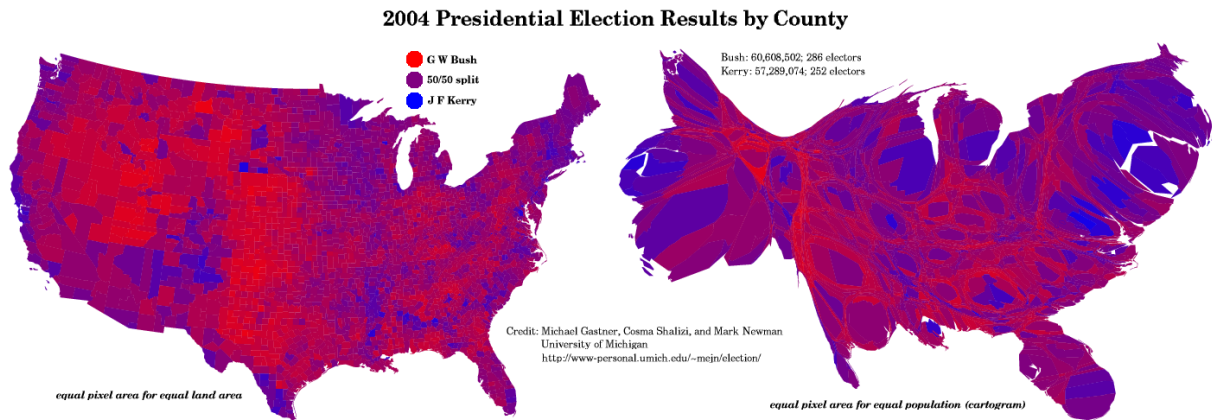
Recommended Reading

➢ NWB Team (2009) Network Workbench Tool, User Manual 1.0.0,
   http://nwb.slis.indiana.edu/Docs/NWBTool-Manual.pdf
➢ Scott Weingart, Hanning Guo, Katy Borner, Kevin W. Boyack, Micah W. Linnemeier, Russell J.
   Duhon, Patrick A. Phillips, Chintan Tank, and Joseph Biberstine (2010) Science of Science
   (Sci2) Tool User Manual. Cyberinfrastructure for Network Science Center, School of Library
   and Information Science, Indiana University, Bloomington.
   http://sci.slis.indiana.edu/registration/docs/Sci2_Tutorial.pdf

[#05] Geospatial Analysis and Mapping

➢ **General Overview**
➢ **Designing Effective Geomaps**
➢ **Sci2-Geomaps With Circle and Colored Region Annotation**
➢ **Sci2-Animations**
➢ **Geographic Information Systems (GIS)**
➢ **Outlook**
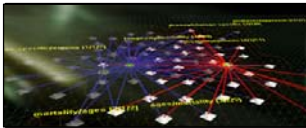➢ **Exercise: Identify Promising Geospatial Analyses of NIH Data**

2004 Presidential Election Results by County

- ● G W Bush
- ● 50/50 split
- ● J F Kerry

Bush: 60,608,502; 286 electors
Kerry: 57,289,074; 252 electors

Credit: Michael Gastner, Cosma Shalizi, and Mark Newman
University of Michigan
http://www-personal.umich.edu/~mejn/election/

equal pixel area for equal land area

equal pixel area for equal population (cartogram)

Information Visualization Course, Katy Börner, Indiana University

Map of human skin colors based on global ultraviolet
radiation intensity and precipitation levels by George Chaplin

Information Visualization Course, Katy Börner, Indiana University
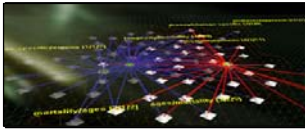
Router density

Population density

(a)

(b)
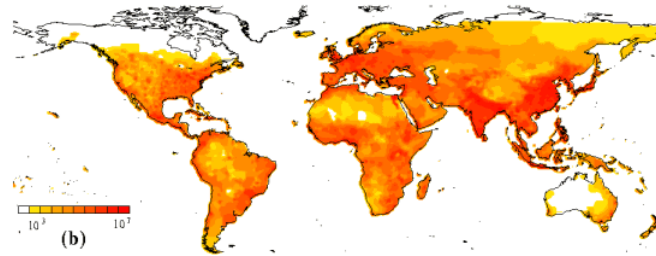
http://www.cybergeography.org/atlas/router_distribution_large.png

Information Visualization Course, Katy Börner, Indiana University
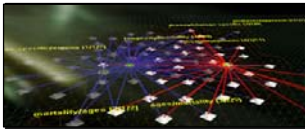
Map Attribute Overlays

http://worldprocessor.com/index_vis.htm

Information Visualization Course, Katy Börner, Indiana University

Red indicates the extent of the radiation cloud on April 27, just after the accident in Chernobyl. Blue, indicates its almost worldwide distribution until the 6th of May.

Dark red circles indicate oil spills and gray-shaded areas indicate sea pollution and land pollution from chemical fertilizers. Shown is only a small part of the entire pollution spectrum in 1988. At an average of every three months this globe becomes obsolete due to yet another major oil spill.
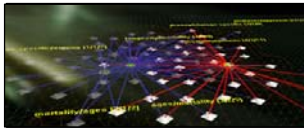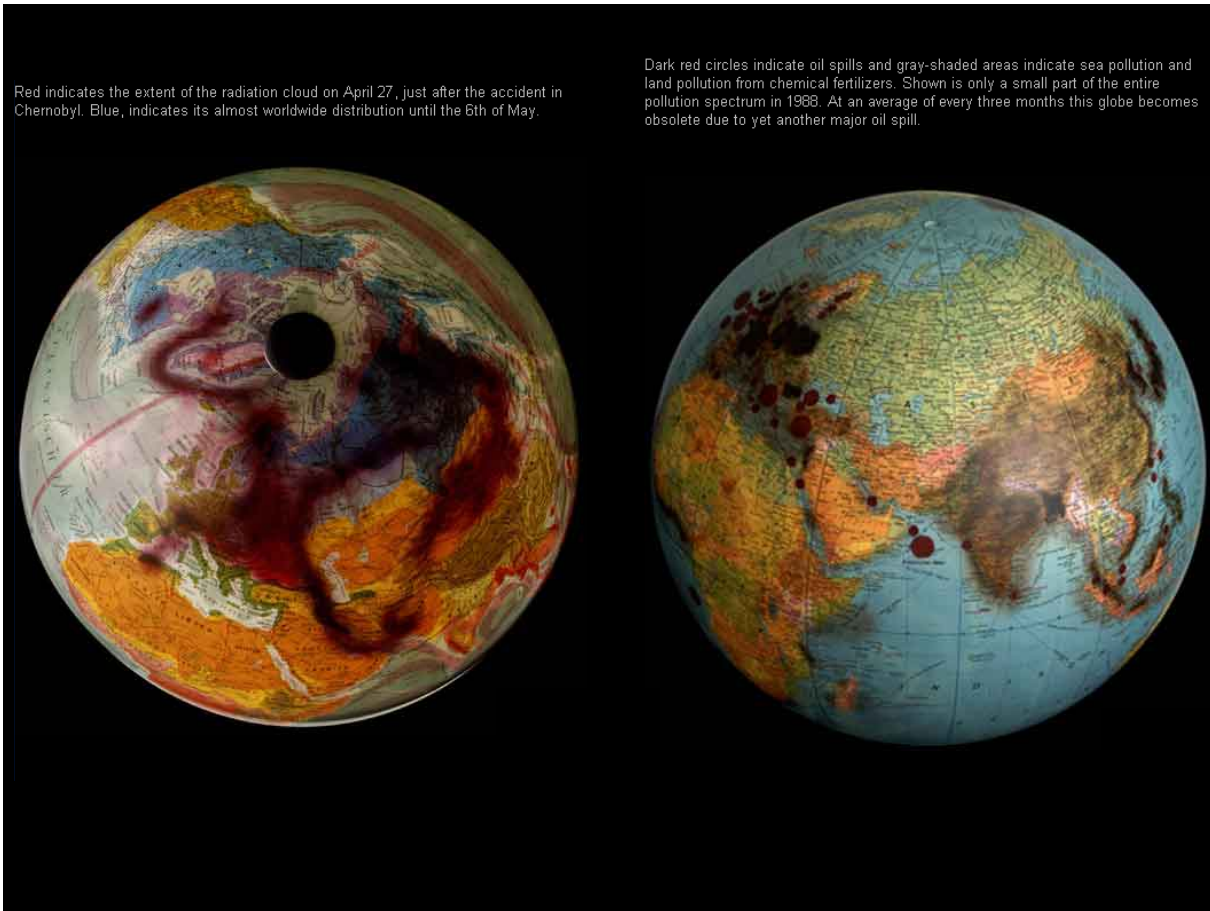
Map Attribute Overlays

Lifelines for visualizing Migrations, Transitions and Trajectories

Figure represents the movements of a person over a single day. Individual starts from the home and visits his workplace, a bank, his workplace and finally a post office, before returning home. The shaded bar at the right identifies periods spent traveling (in black) and at work (cross-hatched).
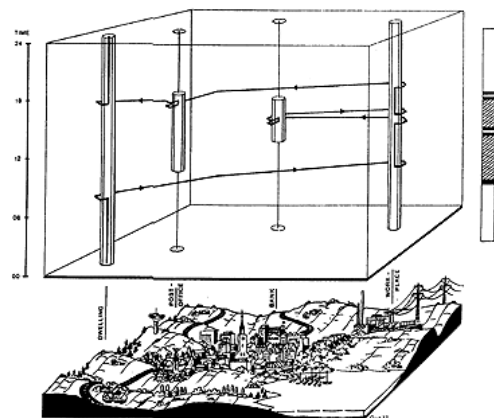
Lenntop's chapter in Carlstein et al.
http://www.geog.port.ac.uk/lifeline/consult/essay.html

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.
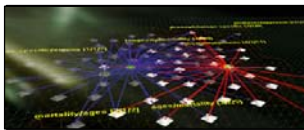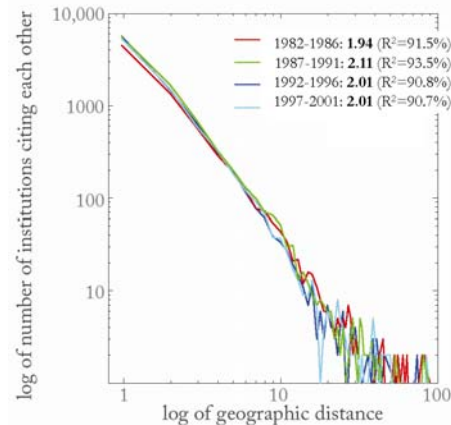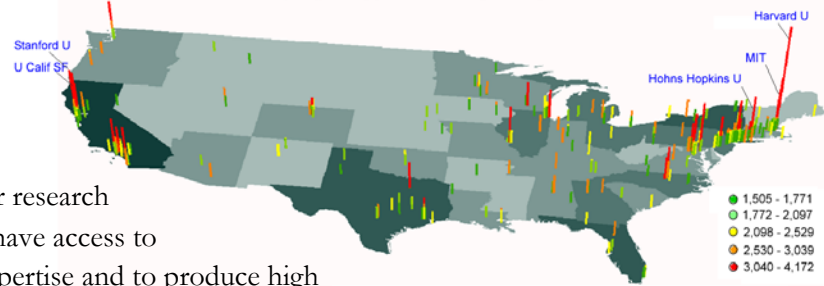
Insights from Geography

1. Places have location, direction, and distance with respect to other places
2. Scale is important--places may be large or small
3. A place has both physical structure and cultural content
4. The characteristics of places develop and change over time
5. Places interact with other places
6. The content of places is rationally structured
7. Places may be generalized into regions of similarities and differences

http://www.csiss.org/learning_resources/content/g5/

Suggested Reading

➢ MacEachren, AM (1995) How Maps Work.. New York, Guilford. Press.

[#05] Geospatial Analysis and Mapping

➢ **General Overview**
➢ **Designing Effective Geomaps**
➢ **Sci2-Geomaps With Circle and Colored Region Annotation**
➢ **Sci2-Animations**
➢ **Geographic Information Systems (GIS)**
➢ **Outlook**
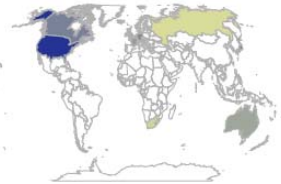➢ **Exercise: Identify Promising Geospatial Analyses of NIH Data**

13

Map Projections - US Map Scope

Eckert IV

Winkel Tripel

Mercator

*Albers Equal-Area Conic
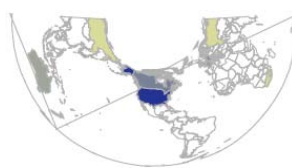
*Lambert Conformal Conic

Geo Maps (region coloring)

Creates a map with colored-region annotations. Re
and colored according to columns in the input table.
be log-scaled before processing.

Map	US States
Projection	Lambert Conformal Conic
Author Name	K. Borner
Region Name	Inst St
Color By	Total Awd Tot $
Color Scaling	Linear
Color Range	Yellow to Blue

* recommended

14

*Eckert IV

*Winkel Tripel

*Mercator

Albers Equal-Area Conic

Geo Maps (region coloring)

Creates a map with colored-region annotations. Re
and colored according to columns in the input table.
be log-scaled before processing.

Map	Countries
Projection	Lambert Conformal Conic
Author Name	K. Borner
Region Name	Inst St
Color By	Total Awd Tot $
Color Scaling	Linear
Color Range	Yellow to Blue

Lambert Conformal Conic

* recommended

15
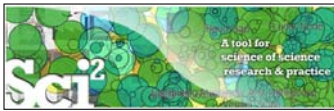
In Sci2 Tool these are called **Colored Region Annotations** and **Circle Annotations**
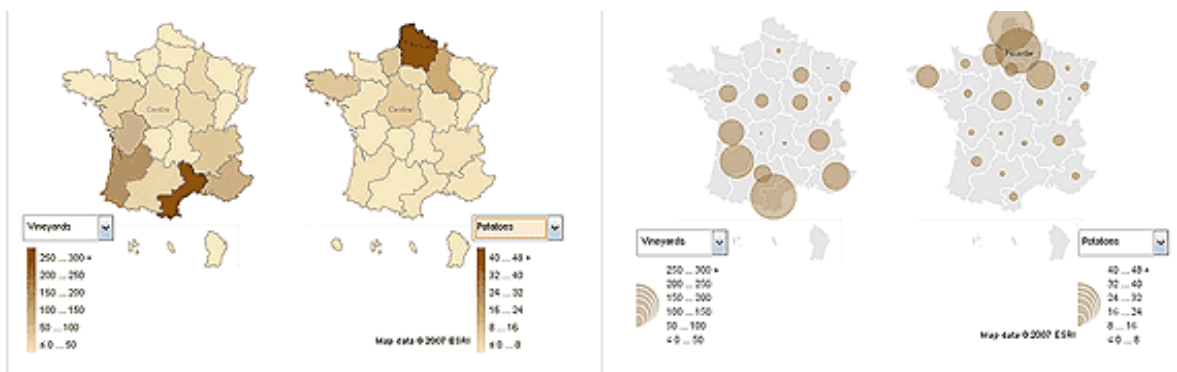
Side-by-side comparisons: on the left, two choropleth maps of France are being compared. On the right, the same maps are being compared in "bubble" mode.

16
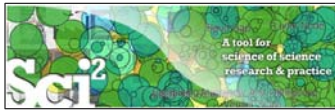
Geospatial Maps – Circle Annotation

USPTP Patent Influenza Data – Geo Map (circle)
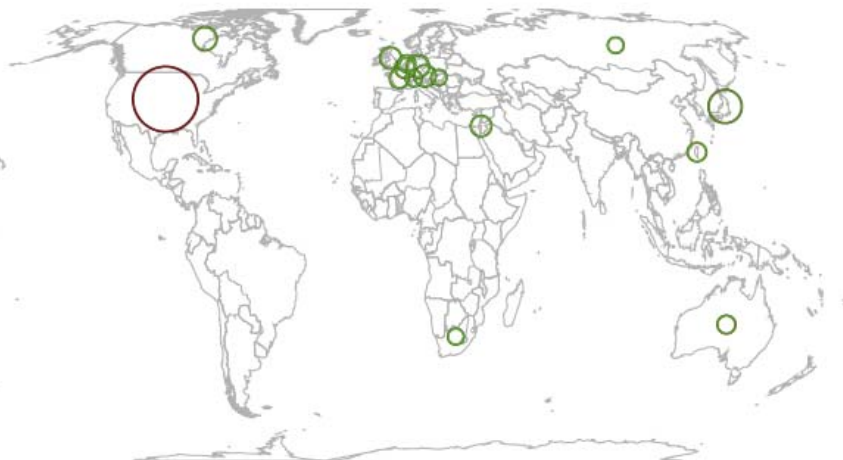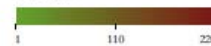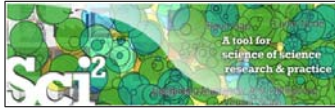
Geospatial Maps – Region Coding

USPTP Patent Influenza Data – Geo Map (region)
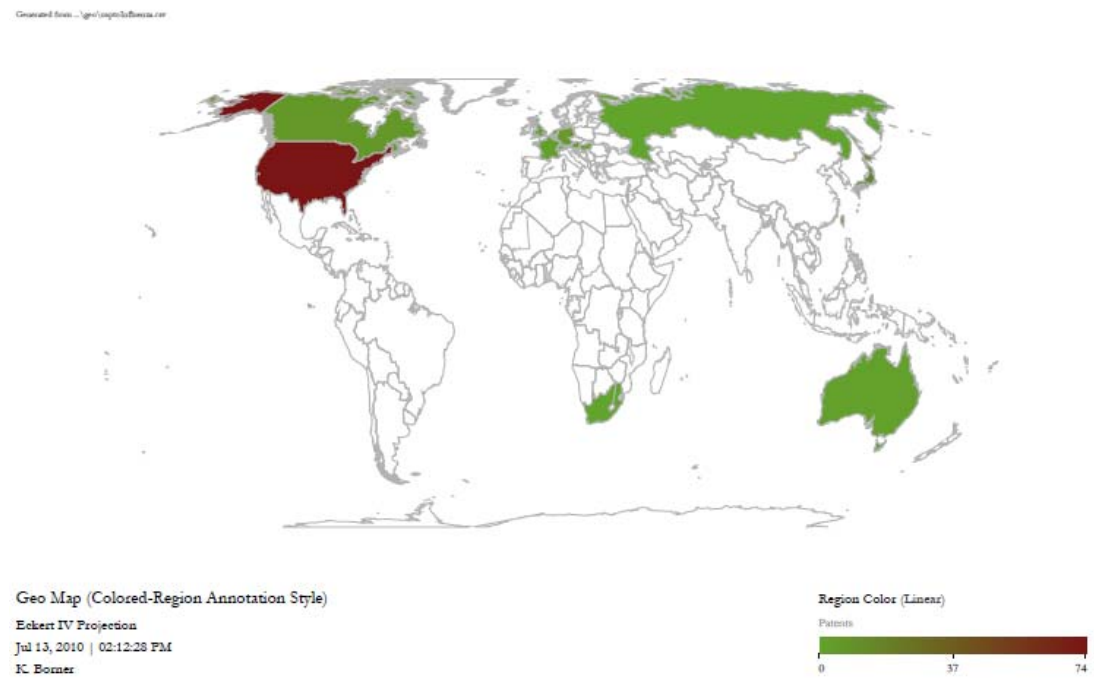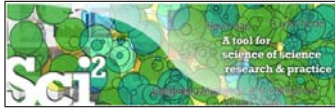
Generated from Prefuse CSV file: C:\Users\User\Desktop\Sci2-Tool-NICO\sampledata\scientometrics\geo\worldfactbook.csv
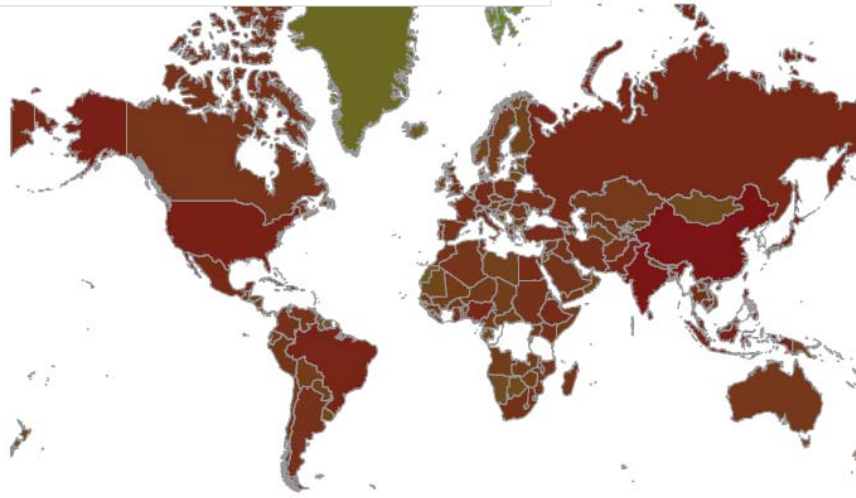
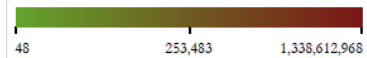Geo Map (Colored-Region Annotation Style)
Mercator Projection
Sep 03, 2009 | 12:35:50 AM
K. Borner

Created with Sci² Tool | Cyberinfrastructure for Network Science Center (http://cns.slis.indiana.edu)

Region Color (logarithmic)
Population

48   253,483   1,338,612,968

21

Geo Mapping: US Map, Data Aggregated Over States
MIDAS PI locations, use file NIH-MIDAS-Grants-Aggregated4GeoState.csv

Area color coding

Geo Map (Colored-Region Annotation Style)
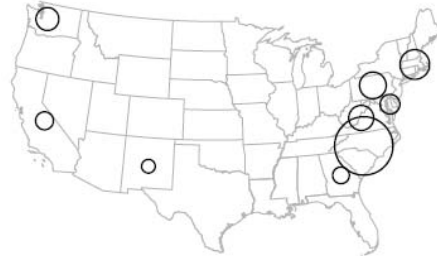Lambert Conformal Conic Projection
Oct 14, 2009 | 06:29:35 PM
Joseph Biberstine

Region Color (Linear)
Total Awd Tot $
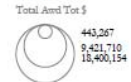
443,267   9,421,710   18,400,154

Circle coding

Geo Map (Circle Annotation Style)
Albers Equal-Area Conic Projection
Oct 14, 2009 | 06:24:56 PM
Joseph Biberstine

Area (Linear)
Total Awd Tot $

443,267
9,421,710
18,400,154

Inst St	Total Awd Tot $
New Mexico	443267
Georgia	960093
California	1193465
Maryland	1587622
Washington	2332220
Virginia	2880362
Pennsylvania	3370739

22

Geo Mapping: US Map, State Level Data
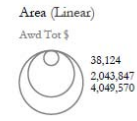*MIDAS PI locations*

Circle coding (Logarithmic)　　　　**Circle coding (Linear)**

Geo Map (Circle Annotation Style)
Albers Equal-Area Conic Projection
Oct 15, 2009 | 05:31:47 PM
K. Borner

Area (Logarithmic)
Awd Tot $
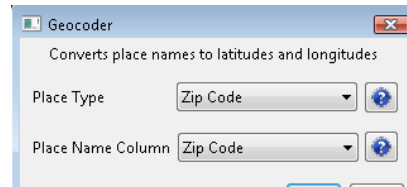
38,124
392,920
4,049,570

Geo Map (Circle Annotation Style)
Albers Equal-Area Conic Projection
Oct 15, 2009 | 05:30:59 PM
K. Borner

Area (Linear)
Awd Tot $

38,124
2,043,847
4,049,570

To convert US ZIP codes
into Latitude/Longitude run
Analysis > Geosaptial > Geocoder with parameter values:

Label	ZIP Code
J. Smith	47405

Geocoder
Converts place names to latitudes and longitudes

Place Type: Zip Code

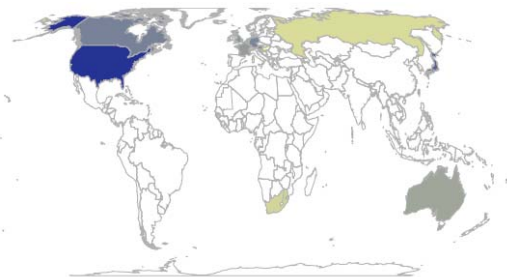Place Name Column: Zip Code

---

Geo Mapping: World Map, Aggregated per Country
Medline 2008 first author locations

Area color coding　　　　　　　　**Circle coding**
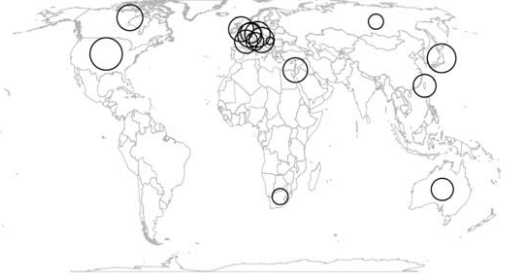
Geo Map (Colored-Region Annotation Style)
Eckert IV Projection
Oct 15, 2009 | 05:04:50 PM
K.Borner

Region Color (Logarithmic)
Patents

0　　　2　　　74

Geo Map (Circle Annotation Style)
Eckert IV Projection
Oct 15, 2009 | 05:30:49 PM
K.Borner

Area (Logarithmic)
Patents

0
2
74
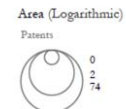
Use
- Area size coding when exactly one variable needs to be encoded.
- Circe coding when 100,000 of zip codes. Can encode 2 variables via area size and ring color.

[#05] Geospatial Analysis and Mapping

➢ **General Overview**
➢ **Designing Effective Geomaps**
➢ **Sci2-Geomaps With Circle and Colored Region Annotation**
➢ **Sci2-Animations**
➢ **Geographic Information Systems (GIS)**
➢ **Outlook**
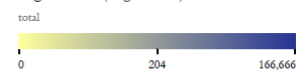➢ **Exercise: Identify Promising Geospatial Analyses of NIH Data**

Generated from Prefuse Table: C:\Documents and Settings\jcbiberz\Desktop\nianli-10-16-12\uspto_country_f_76-79.csv
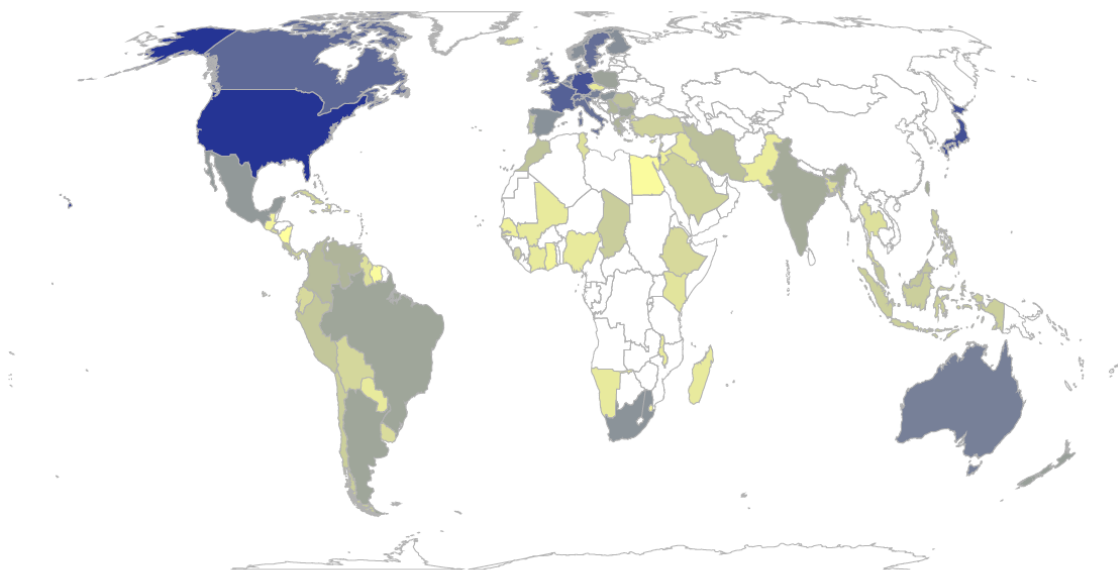
Geo Map (Colored-Region Annotation Style)
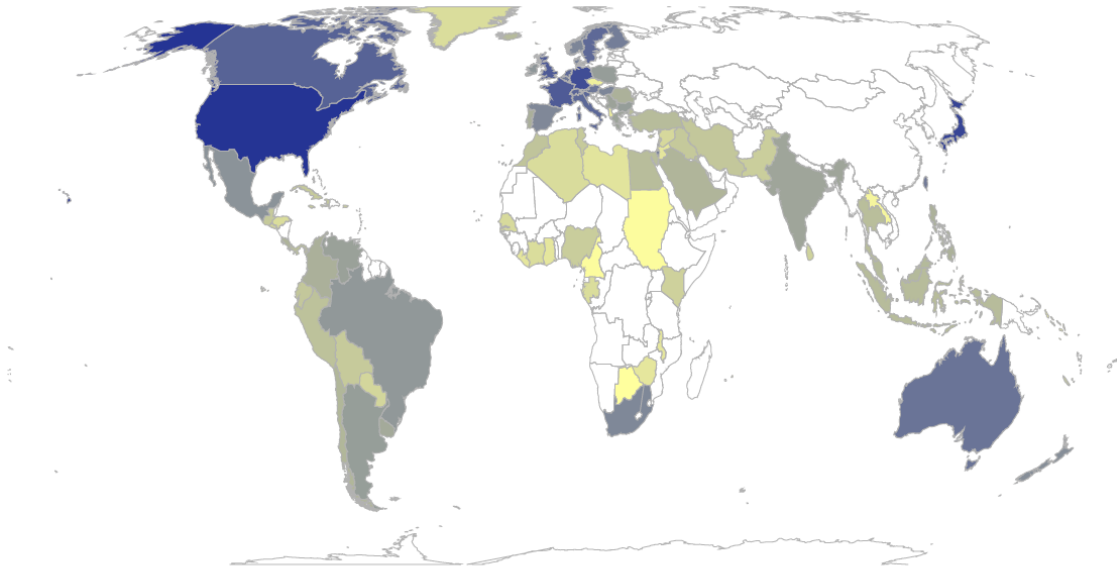Eckert IV Projection
Oct 16, 2009 | 03:03:32 PM
Joseph Biberstine

Region Color (Logarithmic)
total

0   204   166,666

Created with Sci² Tool | Cyberinfrastructure for Network Science Center (http://cns.slis.indiana.edu)
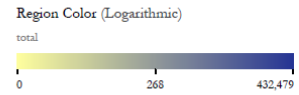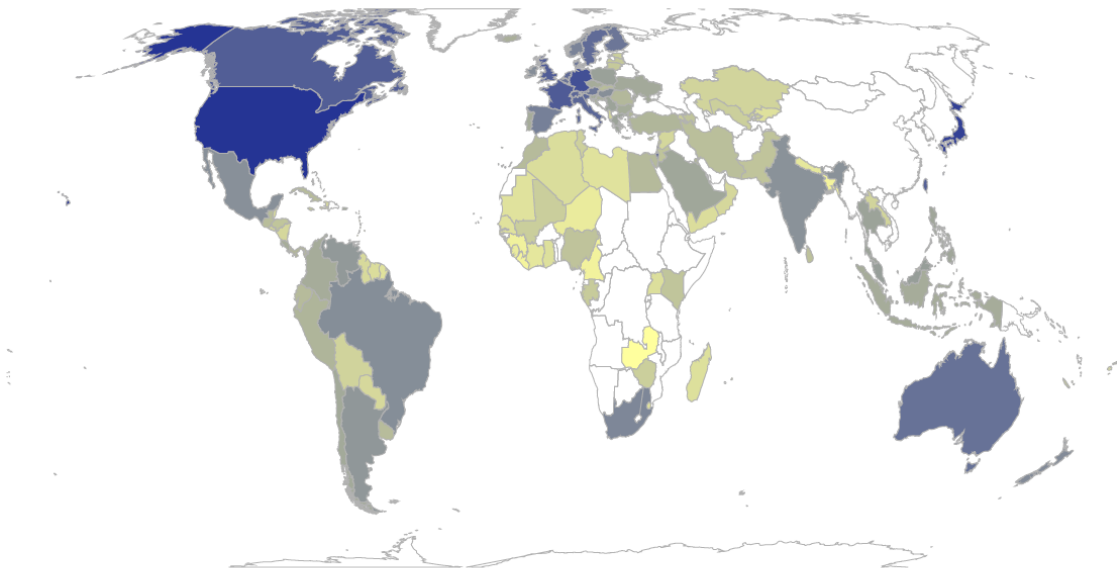
USPTO Patents published in 1976-1979
18 countries not found

Geo Map (Colored-Region Annotation Style)

Eckert IV Projection
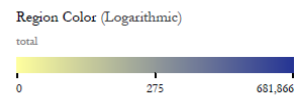
Oct 16, 2009 | 03:05:16 PM

Joseph Biberstine

Region Color (Logarithmic)

total

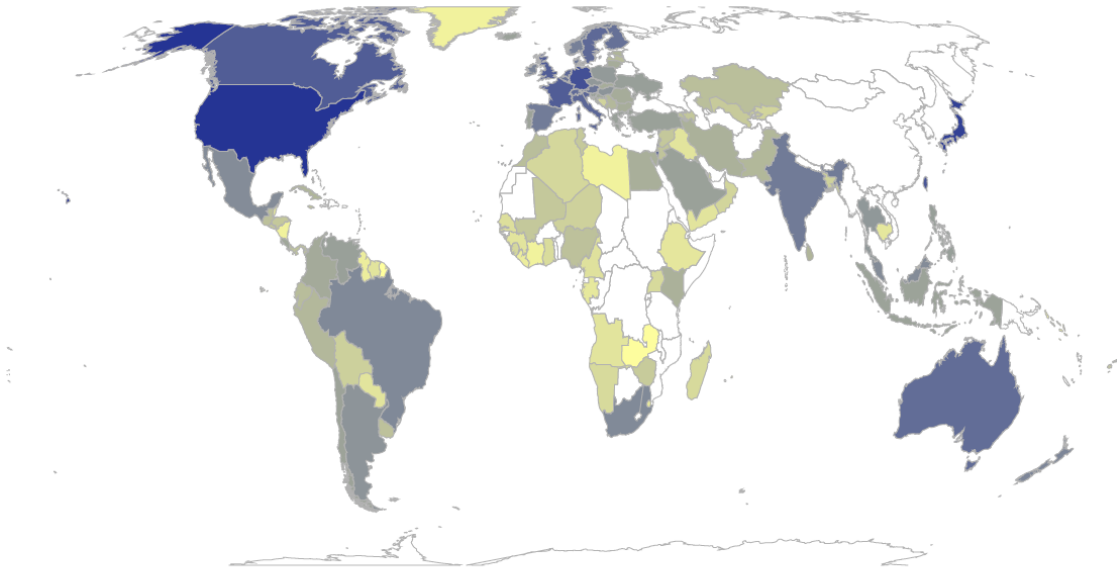| 0 | 268 | 432,479 |

USPTO Patents published in 1980-1989
20 countries not found

27

Geo Map (Colored-Region Annotation Style)

Eckert IV Projection

Oct 16, 2009 | 03:06:00 PM

Joseph Biberstine

Region Color (Logarithmic)
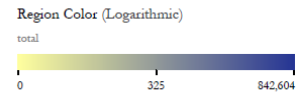
total

| 0 | 275 | 681,866 |

USPTO Patents published in 1990-1999
36 countries not found

28

Geo Map (Colored-Region Annotation Style)
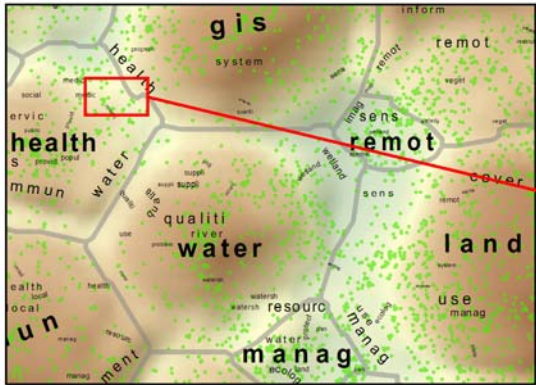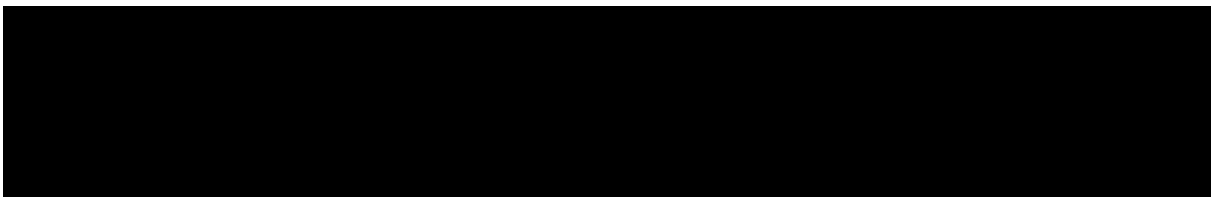Eckert IV Projection
Oct 16, 2009 | 03:06:44 PM
Joseph Biberstine

Region Color (Logarithmic)
total

0   325   842,604

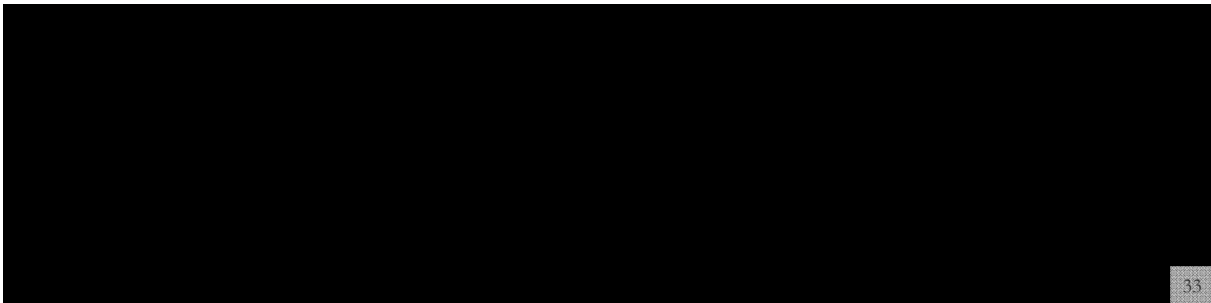**USPTO Patents published in 2000-2008**
33 countries not found

29

[#05] Geospatial Analysis and Mapping

➢ General Overview
➢ Designing Effective Geomaps
➢ Sci2-Geomaps With Circle and Colored Region Annotation
➢ Sci2-Animations
➢ **Geographic Information Systems (GIS)**
➢ Outlook
➢ Exercise: Identify Promising Geospatial Analyses of NIH Data

30

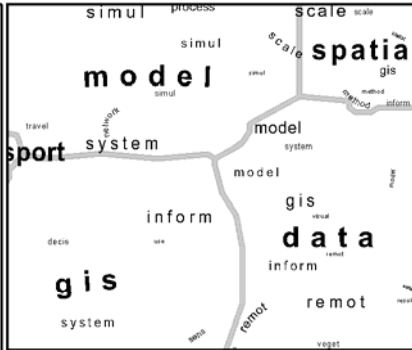*In Terms of Geography* - Andre Skupin - 2005

31

32

M.F. Goodchild and D.C. Ford (1971) Analysis of scallop patterns under controlled conditions. *Journal of Geology* 52–62.

M.F. Goodchild and B.H. Massam (1971) Some least-cost models of spatial administrative systems in Southern Ontario. In R.L. Gentilcore, editor, *Geographical Approaches to Canadian Problems*. Prentice-Hall, Canada, Ltd., 220–228.
...

D.R. Fesenmaier, M.F. Goodchild and S. Morrison (1979) The spatial structure of the rural urban fringe. *Canadian Geographer* 23: 255–265.

M.F. Goodchild (1979) Commentary: current issues in interaction. *Ontario Geography* 13: 85–89.

M.F. Goodchild and P.J. Booth (1976) Modelling human spatial behaviour in urban recreation facility site location., In *Research Program: Impact of the Public Sector on Local Economies*, Discussion Paper 7. London, Ontario: University of Western Ontario, Department of Economics
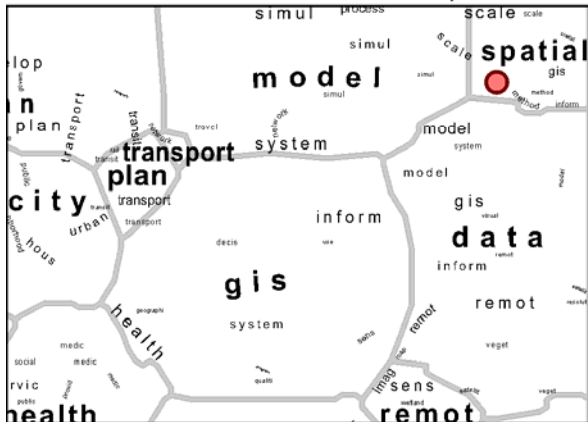...

R. Averack and M.F. Goodchild (1984) Methods and algorithms for boundary definition. In D.F. Marble, editor, *Proceedings of the International Symposium on Spatial Data Handling*, Zurich, 1: 238–250.
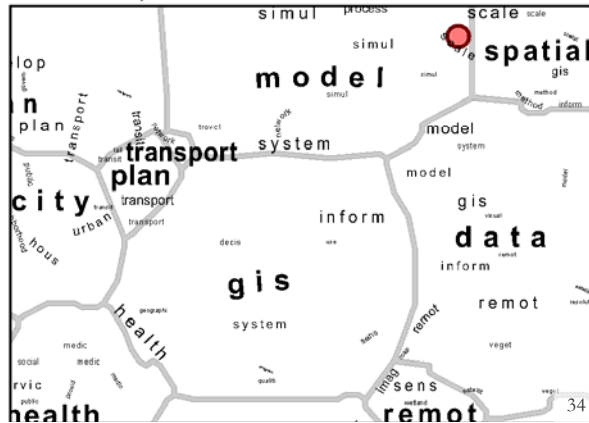
M.F. Goodchild and A.W. Grandfield (1984) Spatial aggregation and intransitivity in U.S. migration streams. *Modeling and Simulation (Proceedings of the 15th Annual Pittsburgh Conference)* 15: 501–505.
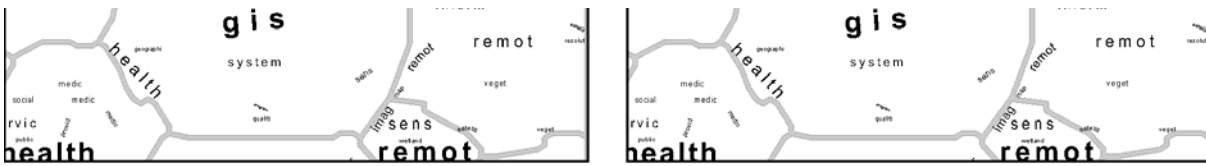
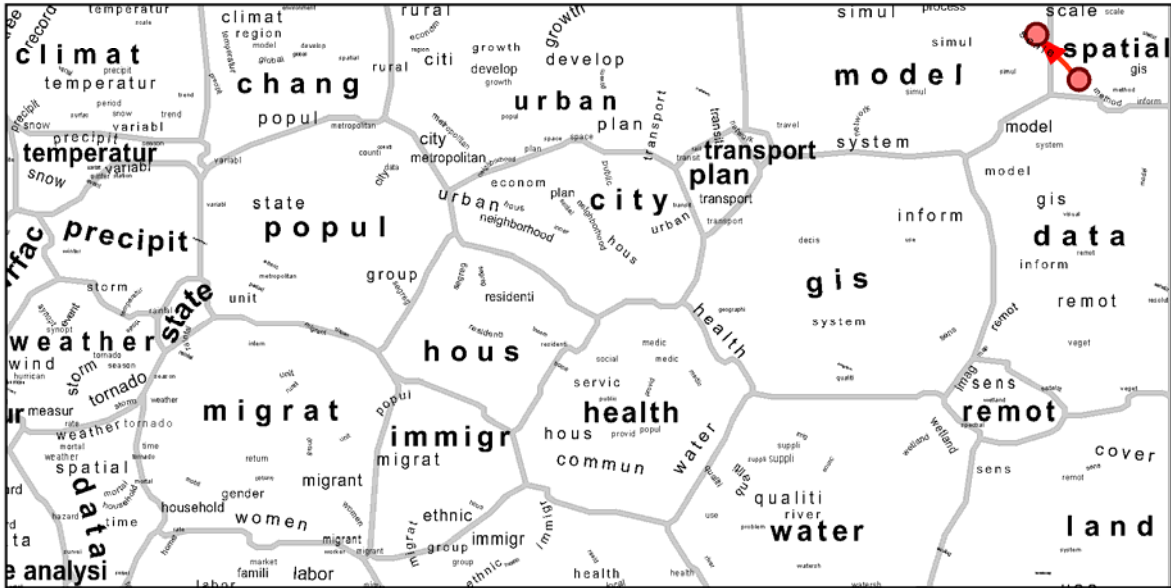Base Map

Author Location 1970-1979

Author Location 1975-1984

33

34

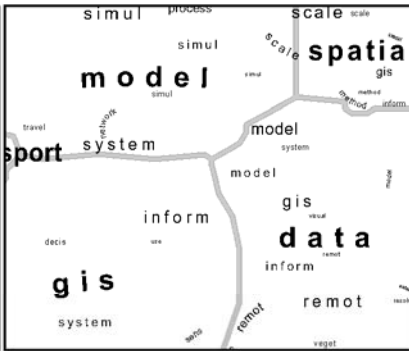Author Change Trajectory from 1970-79 to 1975-84

M.F. Goodchild and D.C. Ford (1971) Analysis of scallop patterns under controlled conditions. *Journal of Geology* 52–62.

M.F. Goodchild and B.H. Massam (1971) Some least-cost models of spatial administrative systems in Southern Ontario. In R.L. Gentilcore, editor, *Geographical Approaches to Canadian Problems*. Prentice-Hall, Canada, Ltd., 220–228.
...

D.R. Fesenmaier, M.F. Goodchild and S. Morrison (1979) The spatial structure of the rural urban fringe. *Canadian Geographer* 23: 255–265.

M.F. Goodchild (1979) Commentary: current issues in interaction. *Ontario Geography* 13: 85–89.

Base Map

M.F. Goodchild and P.J. Booth (1976) Modelling human spatial behaviour in urban recreation facility site location., In *Research Program: Impact of the Public Sector on Local Economies*, Discussion Paper 7. London, Ontario: University of Western Ontario, Department of Economics
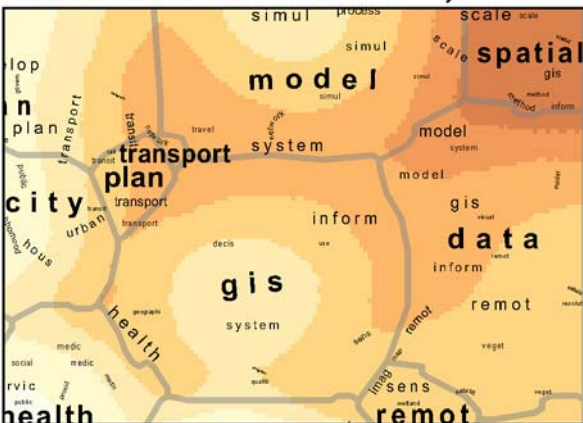...

R. Averack and M.F. Goodchild (1984) Methods and algorithms for boundary definition. In D.F. Marble, editor, *Proceedings of the International Symposium on Spatial Data Handling, Zurich*, 1: 238–250.
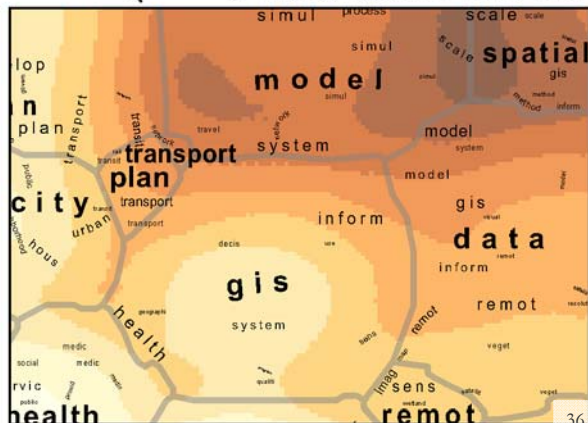
M.F. Goodchild and A.W. Grandfield (1984) Spatial aggregation and intransitivity in U.S. migration streams. *Modeling and Simulation (Proceedings of the 15th Annual Pittsburgh Conference)* 15: 501–505.
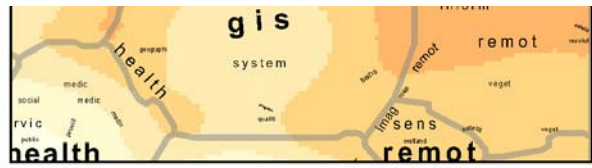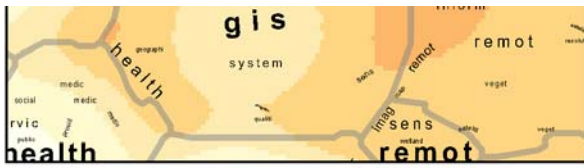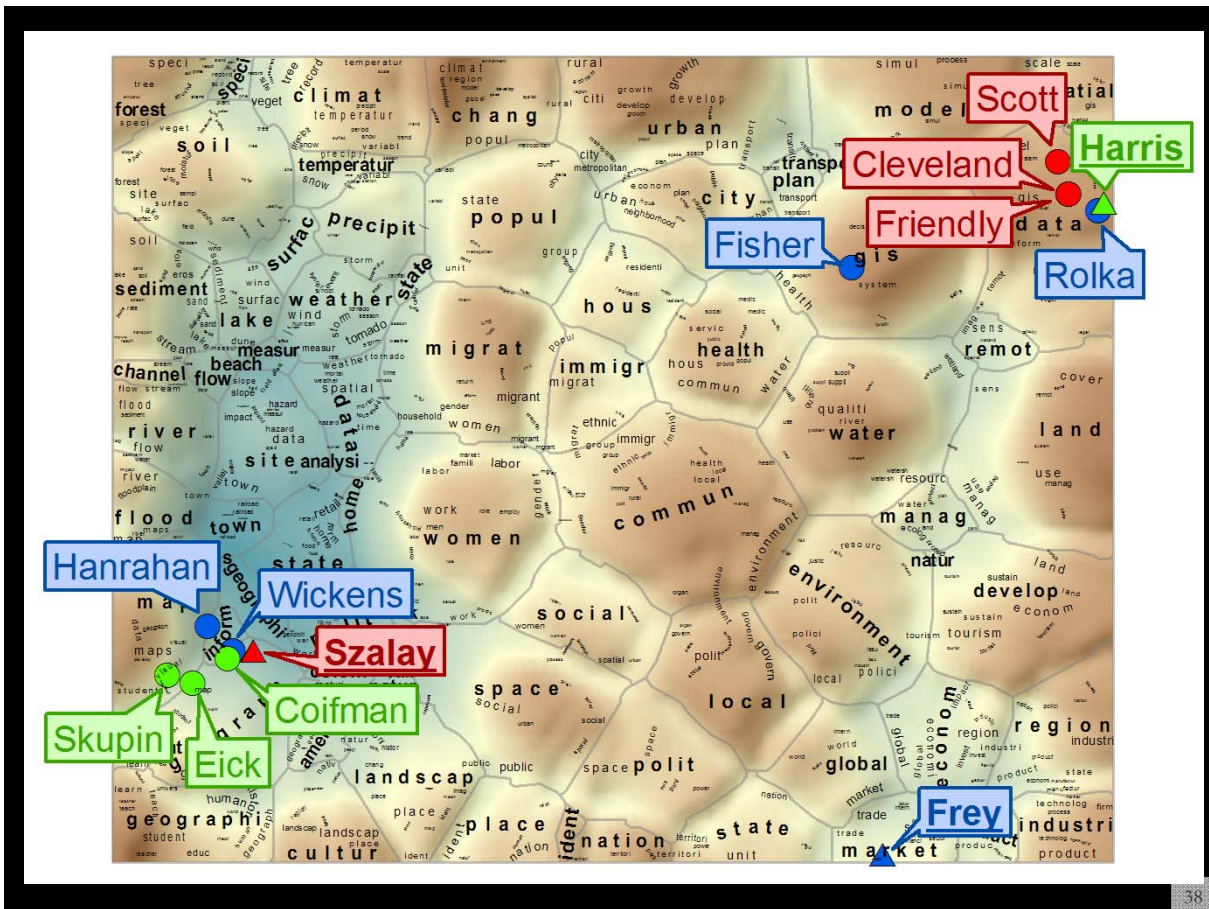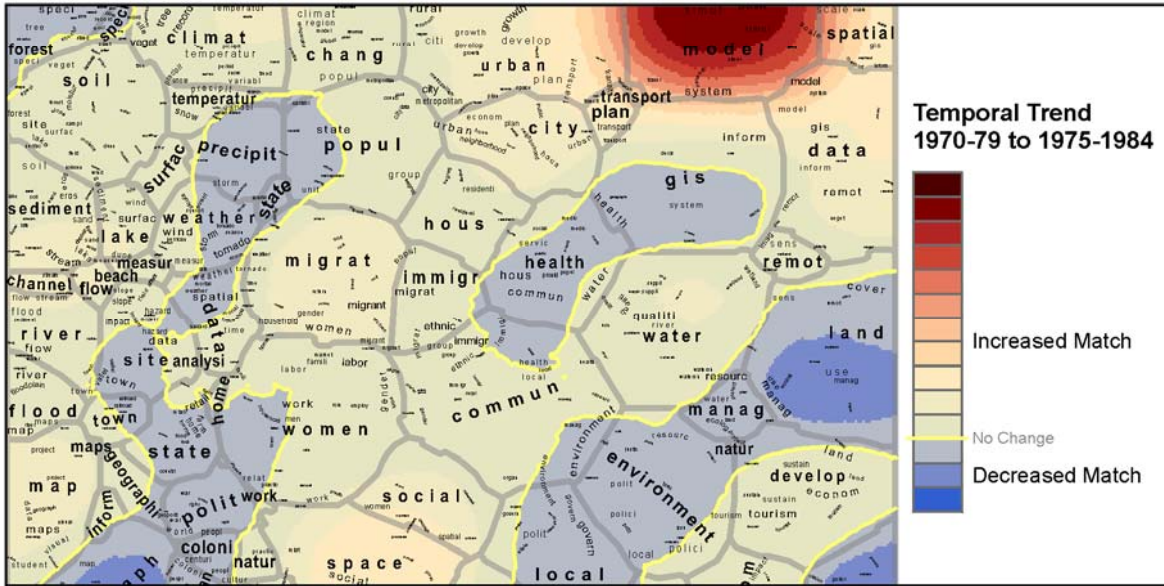
Author Location 1970-1979

Author Location 1975-1984

Author Change Surface from 1970-79 to 1975-84

Temporal Trend
1970-79 to 1975-1984

Increased Match

No Change

Decreased Match

A Semantic Landscape of the Last.fm Music Folksonomy using a Self-Organizing Map

Joseph Biberstine, Russell J. Duhon, Katy Börner, Elisha Hardy, CNS and André Skupin, San Diego State University
Sunbelt SNA Conference, 2010. http://cns.slis.indiana.edu/research/InTermsOfMusic_5Levels_Masked_a9.pdf

39

Visualization created by: Katy Börner (concept), Jeni Coffey (design), Kaveh Ekbia (ArcGIS) and Justin Peters (ArcGIS).

The National Research Network: VIVO: Enabling National Networking of Scientists NIH U24RR029822
Start: Sept 2009
PI: Michael Conlon, University of Florida
Award amount: $12,300,000

DRAFT

40

VIVO Enabling National Networking of Scientists

Legend	
VIVO Insallations	32,669
Email Contacts	95
Web Visits (from Top 25 Countries)	212,517

04/2010

Visualization created by: Katy Börner (concept), Jeni Coffey (design), Kaveh Ekbia (ArcGIS) and Justin Peters (ArcGIS).

DRAFT

Shown are the number of people profiles in the 7 different installation sites.
Email contacts by data and service providers as well as institutions interested to adopt VIVO.
The number of visitors on http://vivoweb.org

41

VIVO Enabling National Networking of Scientists

VIVO Download Counts	87

VIVO source code was registered on May 27 and released on June 3, 2010.
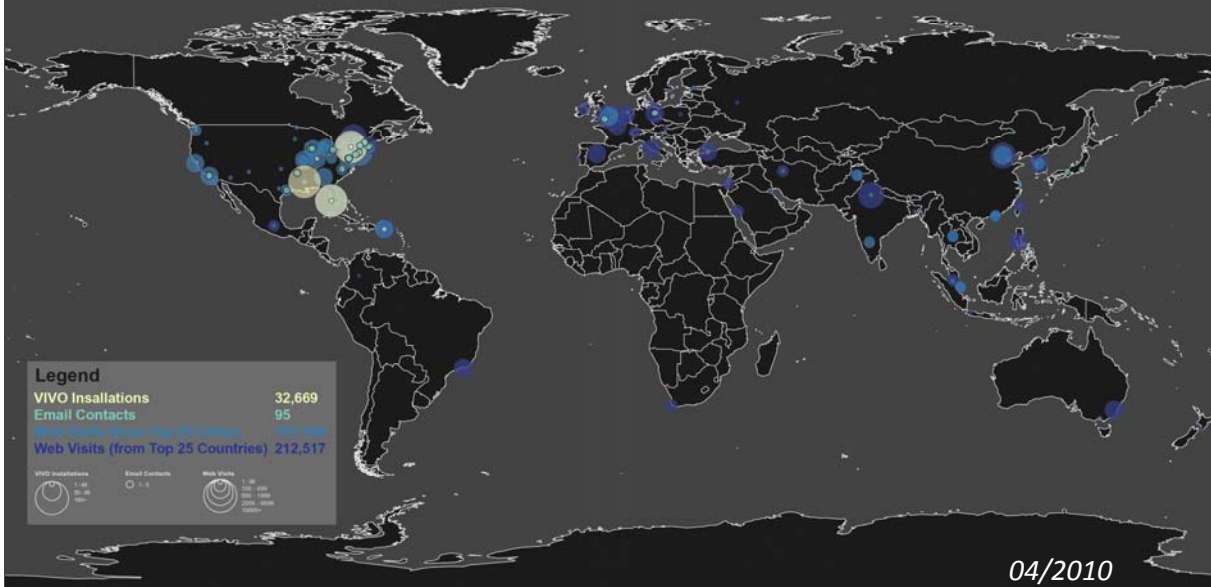
Legend	
VIVO Insallations	32,669
Email Contacts	95
Web Visits (from Top 25 Countries)	212,517

06/2010
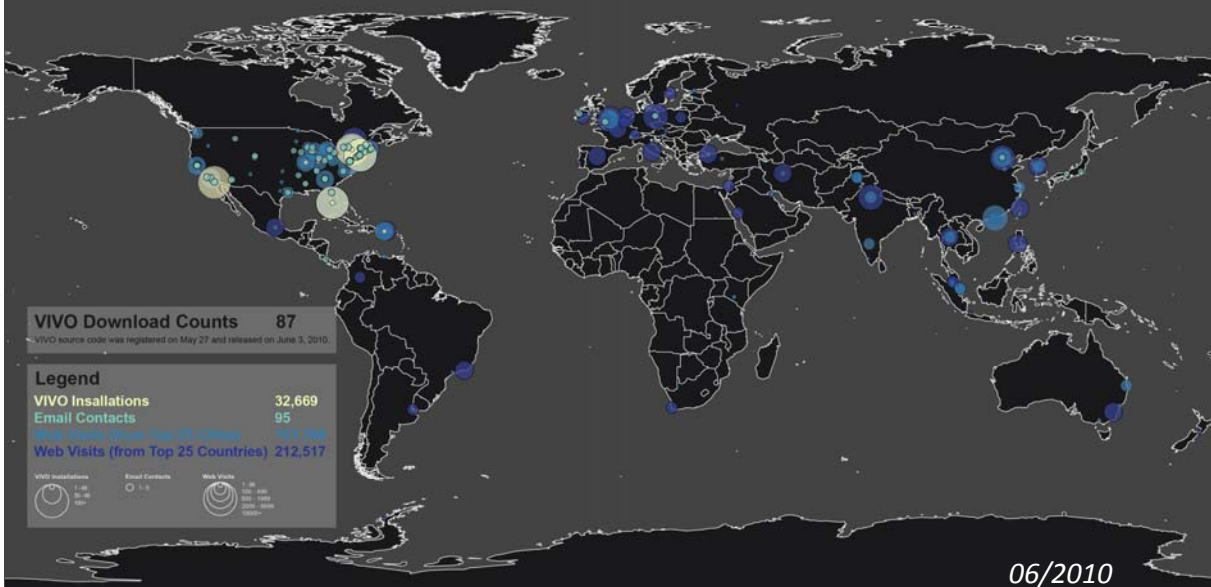
Visualization created by: Katy Börner (concept), Jeni Coffey (design), Kaveh Ekbia (ArcGIS) and Justin Peters (ArcGIS).

DRAFT

VIVO 1.0 source code was publicly released on April 14, 2010
87 Downloads by June 11, 2010
The more institutions adopt VIVO, the more high quality data will be available to understand, navigate, manage, utilize, and communicate progress in science and technology.

42

[#05] Geospatial Analysis and Mapping

➢ **General Overview**
➢ **Designing Effective Geomaps**
➢ **Sci2-Geomaps With Circle and Colored Region Annotation**
➢ **Sci2-Animations**
➢ **Geographic Information Systems (GIS)**
➢ **Outlook**
➢ Exercise: Identify Promising Geospatial Analyses of NIH Data

---

Outlook

Planned extensions of Sci2 Tool:

➢ (Flowmap) network overlays for geo maps and science maps.

➢ Easy means to render maps online.

Research Collaborations by the Chinese Academy of Sciences
By Weixia (Bonnie) Huang, Russell J. Duhon, Elisha F. Hardy, Katy Börner, Indiana University, USA

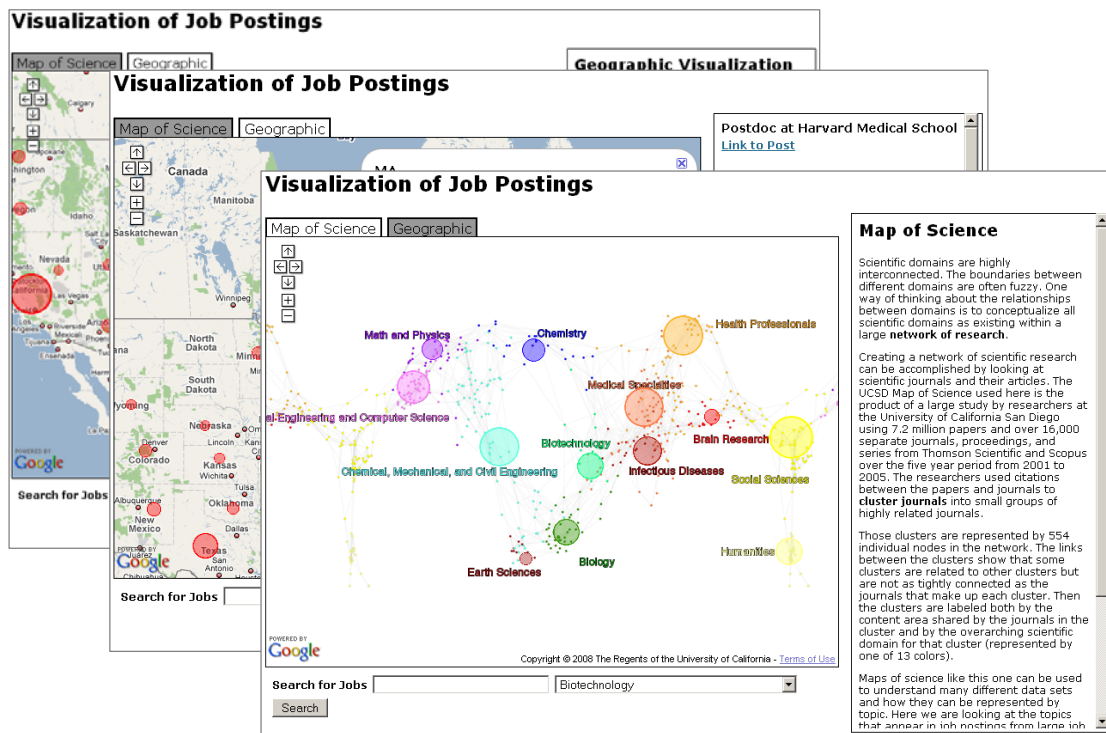## Interactive World and Science Map of S&T Jobs

Angela Zoss, Michael Connover, Katy Börner (2010).

Interactive World and Science Map of S&T Jobs

Angela Zoss, Michael Connover, Katy Börner (2010).

Interactive World and Science Map of S&T Jobs

Angela Zoss, Michael Connover, Katy Börner (2010).

[#05] Geospatial Analysis and Mapping

➢ General Overview
➢ Designing Effective Geomaps
➢ Sci2-Geomaps With Circle and Colored Region Annotation
➢ Sci2-Animations
➢ Geographic Information Systems (GIS)
➢ Outlook
➢ Exercise: Identify Promising Geospatial Analyses of NIH Data

Exercise

Please identify a promising geospatial analysis of NIH data.

Document it by listing

➢ Project title
➢ User, i.e., who would be most interested in the result?
➢ Insight need addressed, i.e., what would you/user like to understand?
➢ Data used, be as specific as possible.
➢ Analysis algorithms used.
➢ Visualization generated. Please make a sketch with legend.
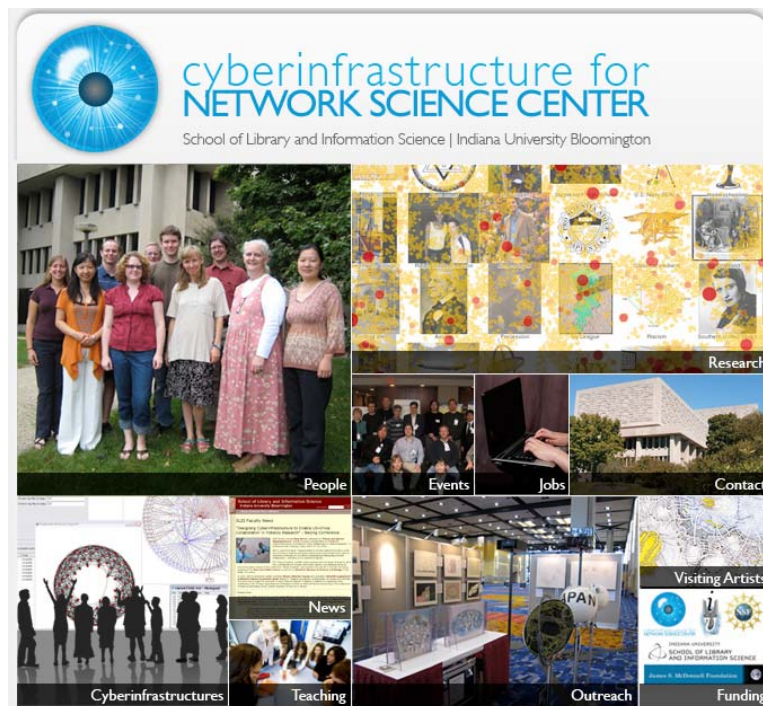
All papers, maps, cyberinfrastructures, talks, press are linked
from http://cns.slis.indiana.edu