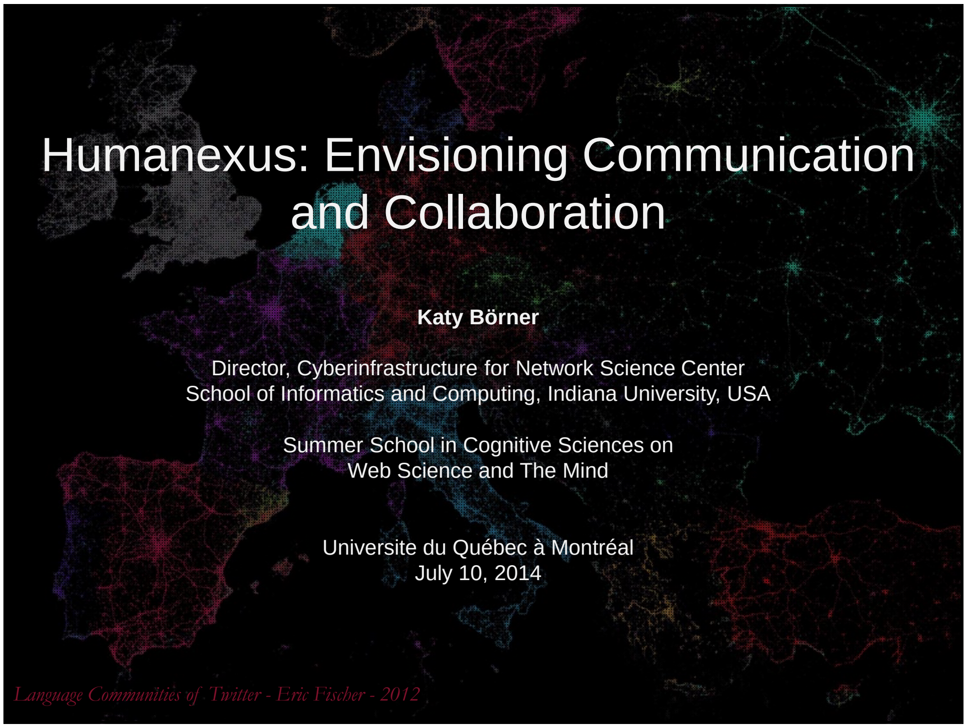


READINGS

Papers

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

Books

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

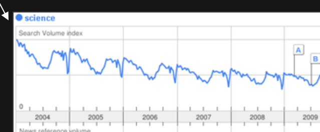
2

Terra bytes of data

Descriptive & Predictive Models

Find your way

Find collaborators, friends



Identify trends

Descriptive Models

Multiple levels: Micro … Macro

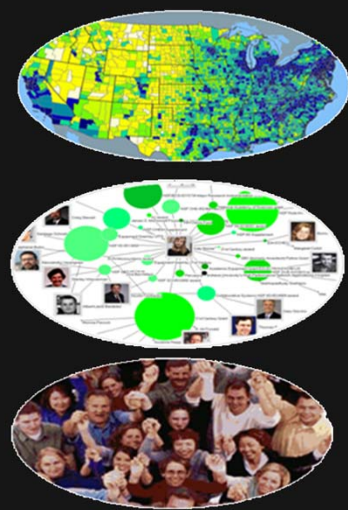
Answering: When? Where? What? With Whom?

Different Levels of Abstraction/Analysis

Macro/Global
Population Level

Meso/Local
Group Level

Micro
Individual Level

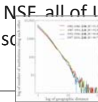
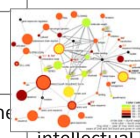
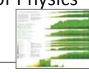


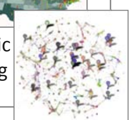
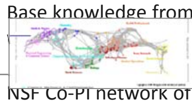
Type of Analysis vs. Level of Analysis

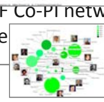
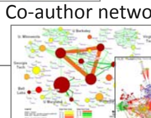
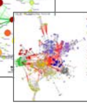
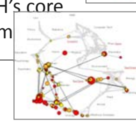
	Micro/Individual (1-100 records)	*Meso/Local (101–100,000 records)*	*Macro/Global (100,000 < records)*
Statistical Analysis/Profiling	Individual person and their expertise profiles	Larger labs, centers, universities, research domains, or states	All of NSF, all of USA, all of science.
Temporal Analysis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

Type of Analysis vs. Level of Analysis

	Micro/Individual (1-100 records)	*Meso/Local (101–100,000 records)*	*Macro/Global (100,000 < records)*
Statistical Analysis/Profiling	Individual person and their expertise profiles	Larger labs, centers, universities, research domains, or states	All of NSF, all of USA, all of sc
Temporal Analysis (When?)	Funding portfolio of one individual	ic bursts in PNAS	113 Years of Physics Research
Geospatial Analysis (Where?)	Career trajectory of one individual	intellectual la	PNAS publications
Topical Analysis (What?)	Base knowledge from s.	f research	VxOrd/Topic NIH funding
Network Analysis (With Whom?)	NSF Co-PI network of one	Co-author network	NIH's core com

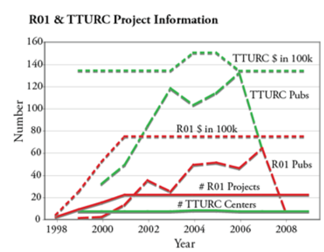
7

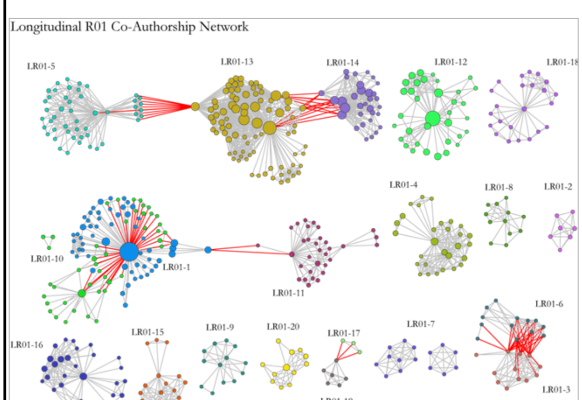
Mapping Transdisciplinary Tobacco Use Research Centers Publications

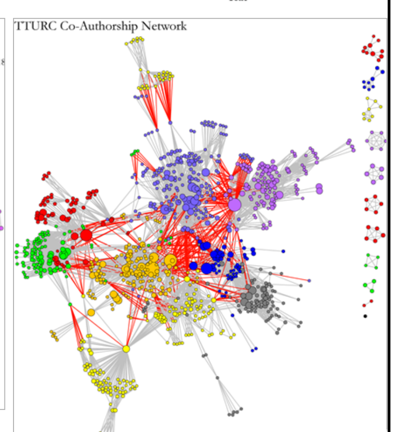
Compare R01 investigator based funding with TTURC Center awards in terms of number of publications and evolving co-author networks.

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

Supported by NIH/NCI Contract HHSN261200800812







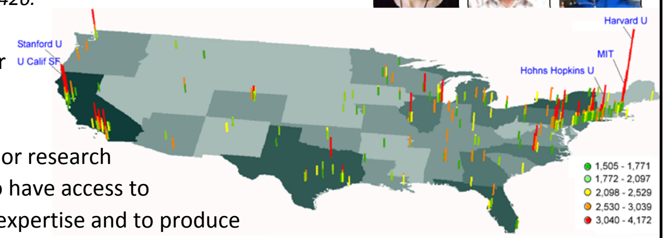
8

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

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

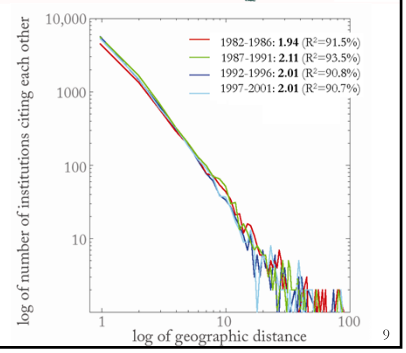
Research questions:

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.



The Global 'Scientific Food Web'

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

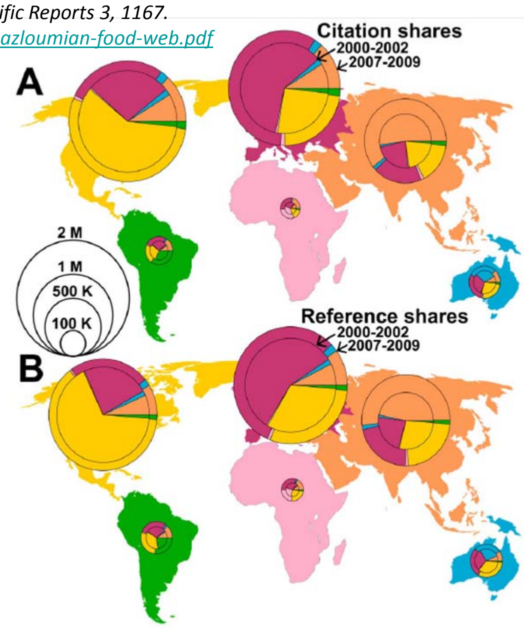
http://cns.iu.edu/docs/publications/2013-mazloumian-food-web.pdf

Contributions:

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

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

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



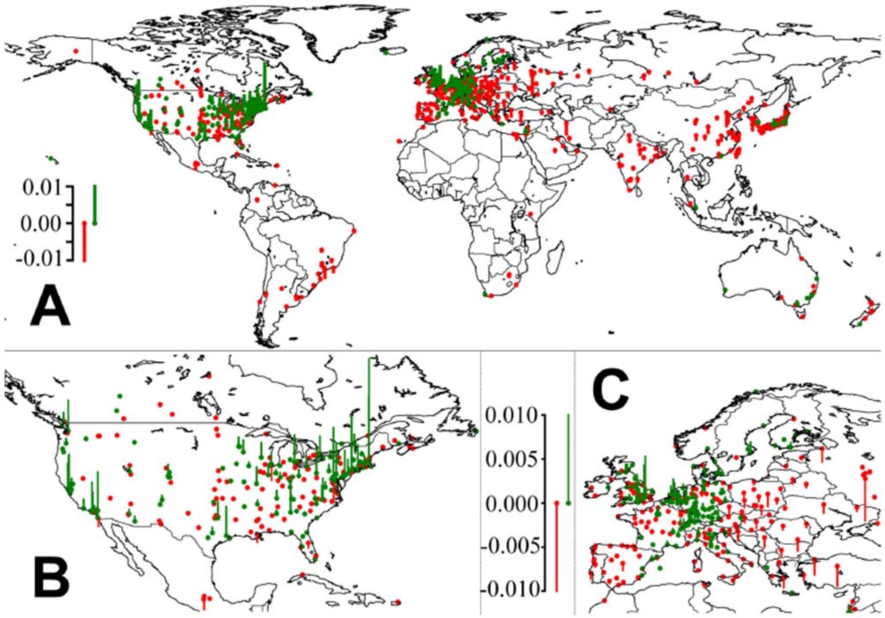
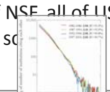
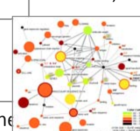
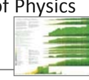
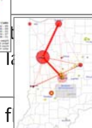
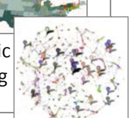


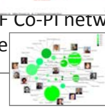
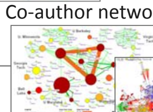
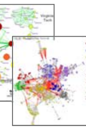
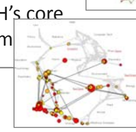
Figure 2 | World map of the greatest knowledge sources and sinks, based on our scientific fitness index. Green bars indicate that the number of citations received is over-proportional, red that the number of citations received is lower than expected (according to a homogeneous distribution of citations over all cities that have published more than 500 papers). It can be seen that most scientific activity occurs in the temperate zone. Moreover, areas of high fitness tend to be areas that are performing economically well (but the opposite does not hold).

Type of Analysis vs. Level of Analysis

	Micro/Individual (1-100 records)	*Meso/Local (101–100,000 records)*	*Macro/Global (100,000 < records)*
Statistical Analysis/Profiling	Individual person and their expertise profiles	Larger labs, centers, universities, research domains, or states	All of NSF, all of USA, all of sc
Temporal Analysis (When?)	Funding portfolio of one individual	ic bursts in PNAS	113 Years of Physics Research
Geospatial Analysis (Where?)	Career trajectory of one individual	intellectual la	PNAS publications
Topical Analysis (What?)	Base knowledge from s.	f esearch	VxOrd/Topic NIH funding
Network Analysis (With Whom?)	NSF Co-PI network of one	Co-author network	NIH's core com



12

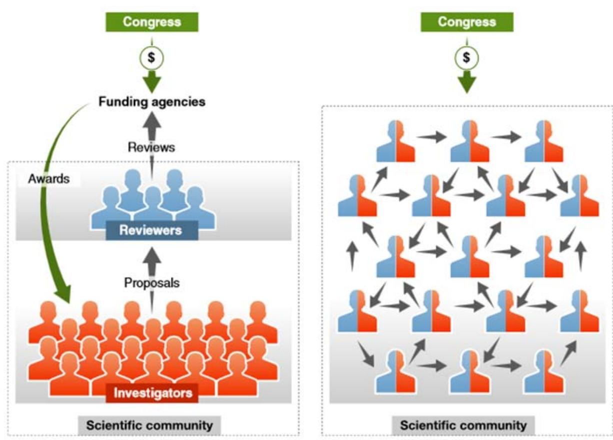
Predictive Models (Why?)

Example: Collective allocation of science funding as an alternative to peer review

13

From funding agencies to scientific agency: Collective allocation of science funding as an alternative to peer review

Bollen, Crandall, Junk, Ding & Börner. 2014. EMBO Reports 15 (1): 1-121.



Existing (left) and proposed (right) funding systems. Reviewers in blue; investigators in red.

In the proposed system, all scientists are both investigators and reviewers: every scientist receives a fixed amount of funding from the government and discretionary distributions from other scientists, but each is required in turn to redistribute some fraction of the total they received to other investigators.

14

From funding agencies to scientific agency: Collective allocation of science funding as an alternative to peer review

Bollen, Johan, David Crandall, Damion Junk, Ying Ding & Katy Börner. 2014. EMBO Reports 15 (1): 1-121.

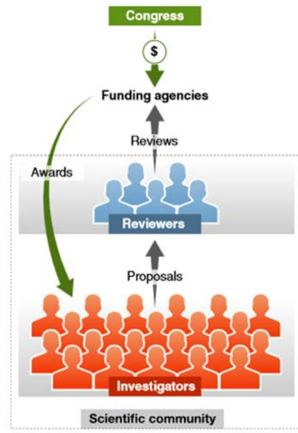
Current Model is Expensive:

If four professors work four weeks full-time on a proposal submission, labor costs are about $30k [1]. With typical funding rates below 20%, about five submission-review cycles might be needed resulting in a total expected labor cost of **$150k**. The average NSF grant is **$128k** per year.

U.S. universities charge about 50% overhead (ca. $42k), leaving about **$86k**.

In other words, the four professors lose **$150k-$86k= - $64k** of paid research time by **obtaining** a grant to perform the proposed research.

To add: Time spent by researchers to review proposals. In 2012 alone, NSF convened more than 17,000 scientists to review 53,556 proposals.



[1] Taulbee Survey of Salaries Computer Science , http://cra.org/resources/taulbee

15

From funding agencies to scientific agency: Collective allocation of science funding as an alternative to peer review

Bollen, Crandall, Junk, Ding & Börner. 2014. EMBO Reports 15 (1): 1-121.

Assume

Total funding budget in year y is t_y

Number of qualified scientists is n

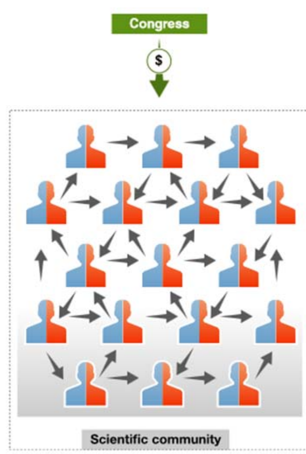
Each year,

the funding agency deposits a fixed amount into each account, equal to the total funding budget divided by the total number of scientists: t_y/n.

Each scientist must distribute a fixed fraction, e.g., 50%, of received funding to other scientists (no self-funding, COIs respected).

Result

Scientists collectively assess each others' merit based on different criteria; they "fund-rank" scientists; highly ranked scientists have to distribute more money.



16

From funding agencies to scientific agency: Collective allocation of science funding as an alternative to peer review

Bollen, Crandall, Junk, Ding & Börner. 2014. EMBO Reports 15 (1): 1-121.

Example:

Total funding budget per year is 2012 NSF budget

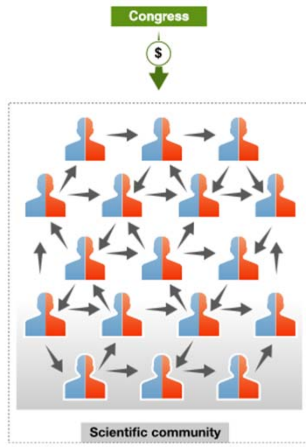
Given the number of NSF funded scientists, each receives a $100,000 basic grant.

Fraction is set to 50%

In 2013, scientist *S* receives a basic grant of $100,000 plus $200,000 from her peers, i.e., a total of $300,000.

In 2013, *S* can spend 50% of that total sum, $150,000, on her own research program, but must donate 50% to other scientists for their 2014 budget.

Rather than submitting and reviewing project proposals, *S* donates directly to other scientists by logging into a centralized website and entering the names of the scientists to donate to and how much each should receive.



17

From funding agencies to scientific agency: Collective allocation of science funding as an alternative to peer review

Bollen, Crandall, Junk, Ding & Börner. 2014. EMBO Reports 15 (1): 1-121.

Model Run and Validation:

Model is presented in http://arxiv.org/abs/1304.1067

It uses **citations as a proxy** for how each scientist might distribute funds in the proposed system.

Dataset: 37M articles from TR 1992 to 2010 Web of Science (WoS) database with **770M citations** and 4,195,734 unique author names. The **867,872 names** who had authored at least one paper per year in any five years of the period 2000–2010 were used in validation.

For each pair of authors we determined the number of times one had cited the other in each year of our citation data (1992–2010).

NIH and NSF funding records from IU's Scholarly Database provided 347,364 grant amounts for 109,919 unique scientists for that time period.

Simulation run begins in year 2000, in which every scientist was given a fixed budget of B = $100k. In subsequent years, scientists distribute their funding in proportion to their citations over the prior 5 years.

The model yields funding patterns similar to existing NIH and NSF distributions.

18

NEWSFOCUS

Making Every Scientist a Research Funder

When it comes to using peer review to distribute research dollars, Johan Bollen favors radical simplicity.

Over the years, many scientists have suggested that the current system could be improved by changing the composition of the review panels, tweaking the interactions among reviewers, or revising how the proposals are scored. But Bollen, a computer scientist at Indiana University, Bloomington, would simply award all eligible researchers a block grant—and then require them to give some of it away to colleagues they judge most deserving.

That radical step, described in a paper Bollen and four Indiana colleagues recently posted on *EMBO Reports*, retains peer review's core concept of tapping into the views of the most knowledgeable researchers. But it would eliminate the huge investment in time and money required to submit proposals and assemble panels to judge them.

Bollen's process would be almost instantaneous: In a version of expert-directed crowdsourcing, scientists would fill out a form once a year listing their favored researchers, and a predetermined portion of their annual grant money—a total of, say, 50%—would then be transferred to their choices.

"So many scientists spend so much time on peer review, and there's a high level of frustration," Bollen explains. "We already know who the best people are. And if you're doing good work, then you deserve to receive support."

Others are skeptical. "I've known Johan for a long time and have the highest regard for his ability as an out-of-the-box thinker," says Stephen Griffin, a retired National Science Foundation (NSF) program manager who's now a visiting professor of information sciences at the University of Pittsburgh in Pennsylvania. "But there are a number of issues he doesn't address."

Those sticking points include the likely mismatch between what researchers need and what their colleagues give them; the absence of any replacement for the overhead payments in today's grants, which support infrastructure at host institutions; and the dearth of public accountability for the billions of dollars that would flow from public coffers to individuals. "Scientists aren't really equipped to be a funding agency," Griffin notes.

Bollen acknowledges that the process would need safeguards to ensure that scientists don't reward their friends or punish their enemies. But his analysis suggests that the U.S. research landscape would not look all that different if his radical proposal were adopted.

Drawing upon citation data in 37 million papers over 20 years, the Indiana researchers conducted a simulation premised on the idea that scientists would reallocate their federal dollars according to how often they cited their peers. The simulation, he says, yielded a funding pattern "similar in shape to the actual distribution" at NSF and the National Institutes of Health for the past decade—at a fraction of the overhead required by the current system.

–JDM

February 7, 2014

Science 7 February 2014: Vol. 343 no. 6171 p. 598
DOI: 10.1126/science.343.6171.598
http://www.sciencemag.org/content/343/6171/598.full?sid=4f40a7f0-6ba2-4ad8-a181-7ab394fe2178

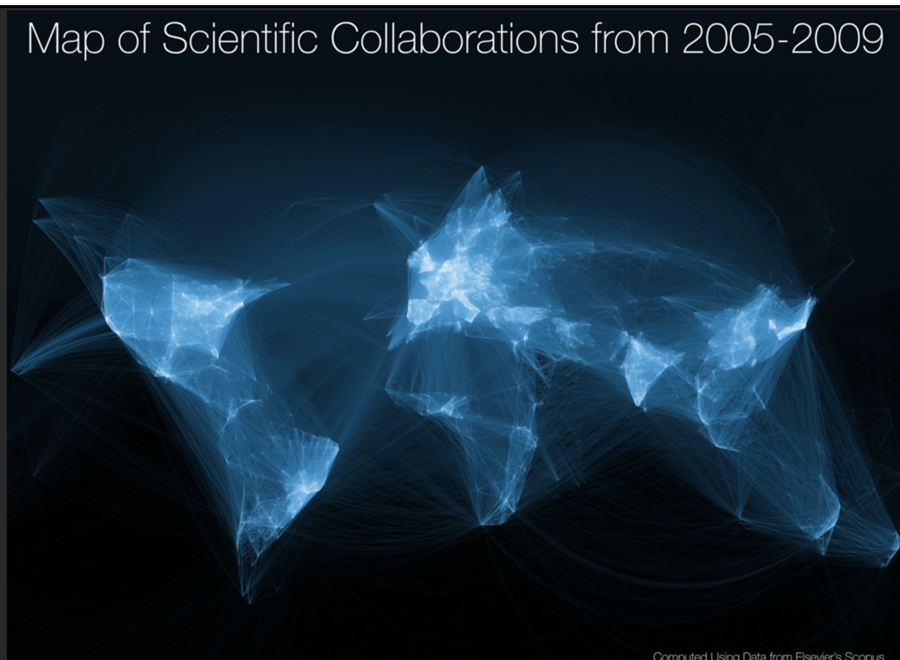
Visualizing STI Model Results

Example: Places & Spaces: Mapping Science Exhibit

20

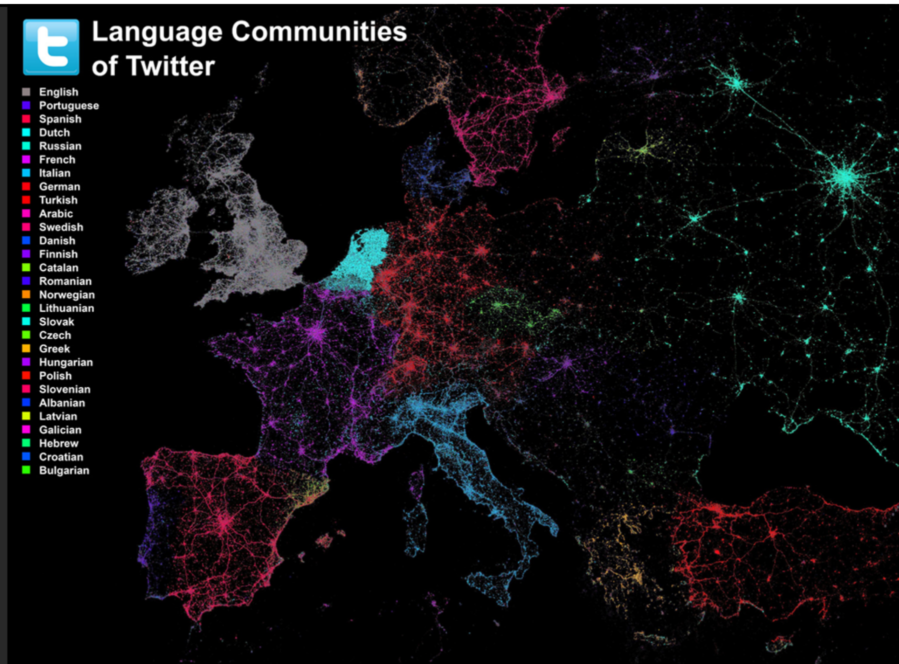
Mapping Science Exhibit on display at MEDIA X, Stanford University
http://mediax.stanford.edu, *http://scaleindependentthought.typepad.com/photos/scimaps*

21



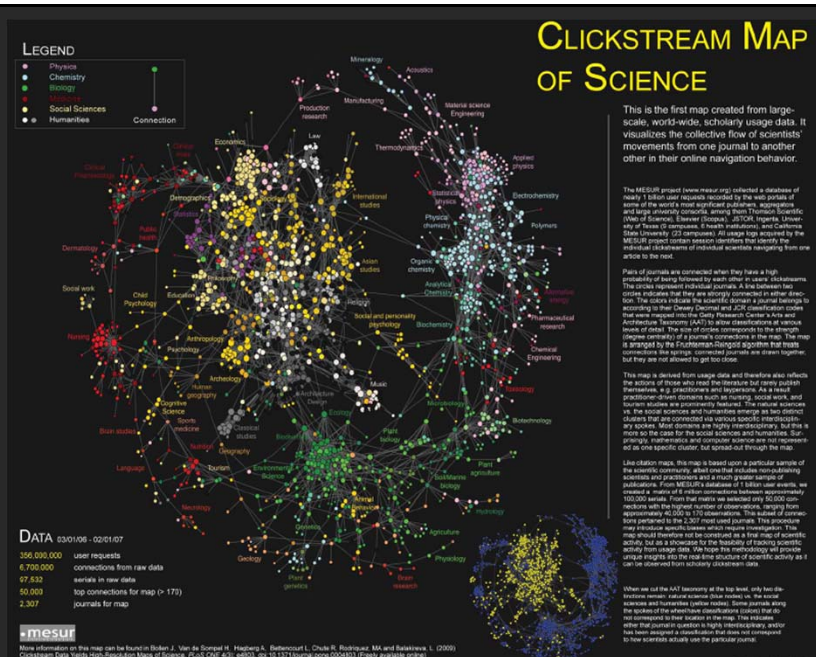
Olivier H. Beauchesne, 2011. Map of Scientific Collaborations from 2005-2009.

22



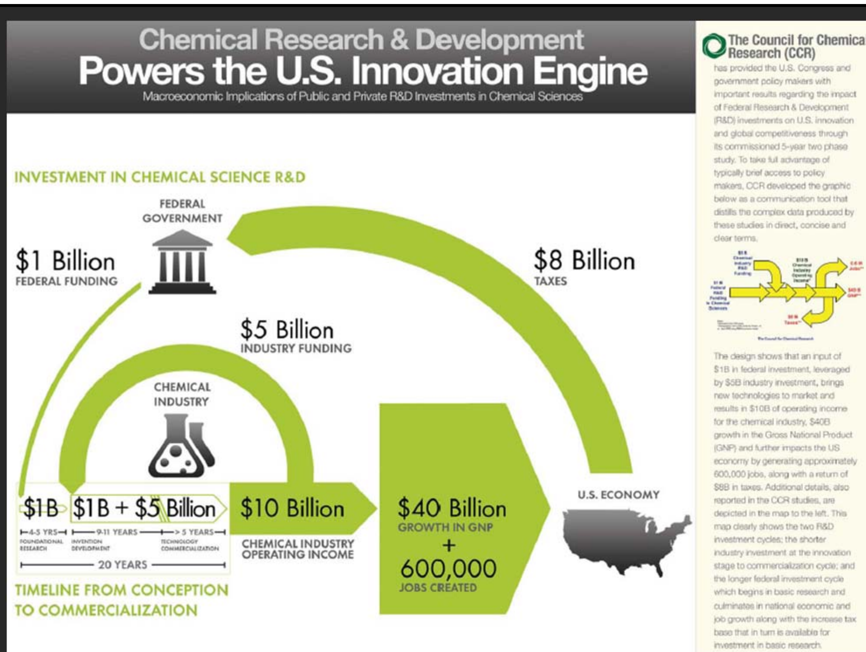
Language Communities of Twitter - Eric Fischer - 2012

23



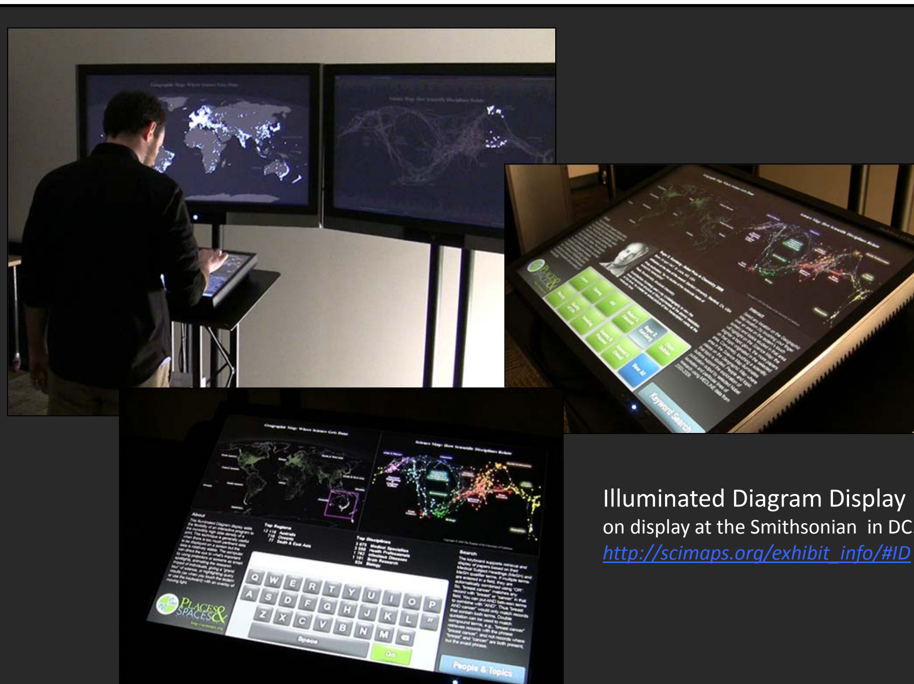
Bollen, Johan, Herbert Van de Sompel, Aric Hagberg, Luis M.A. Bettencourt, Ryan Chute, Marko A. Rodriquez, Lyudmila Balakireva. 2008. A Clickstream Map of Science.

24



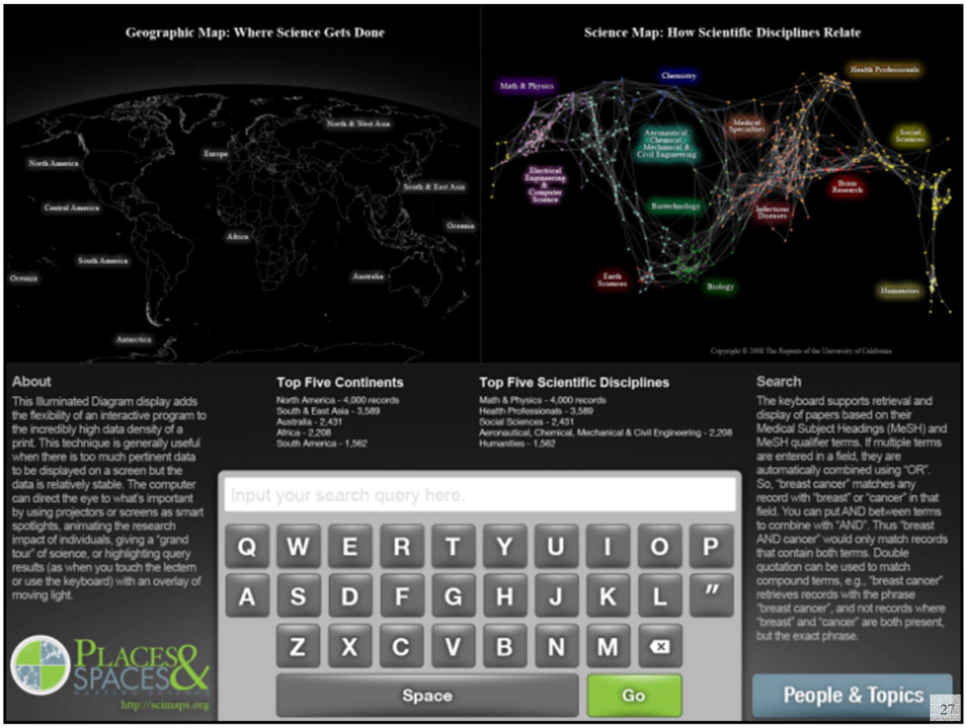
Council for Chemical Research. 2009. Chemical R&D Powers the U.S. Innovation Engine. Washington, DC. Courtesy of the Council for Chemical Research.

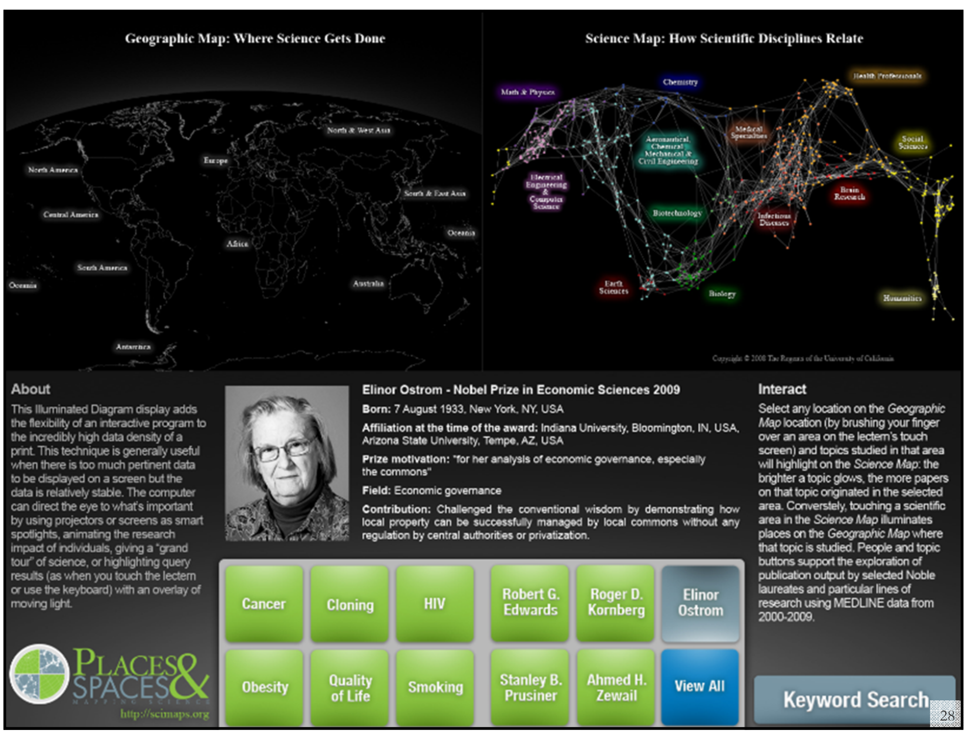
25



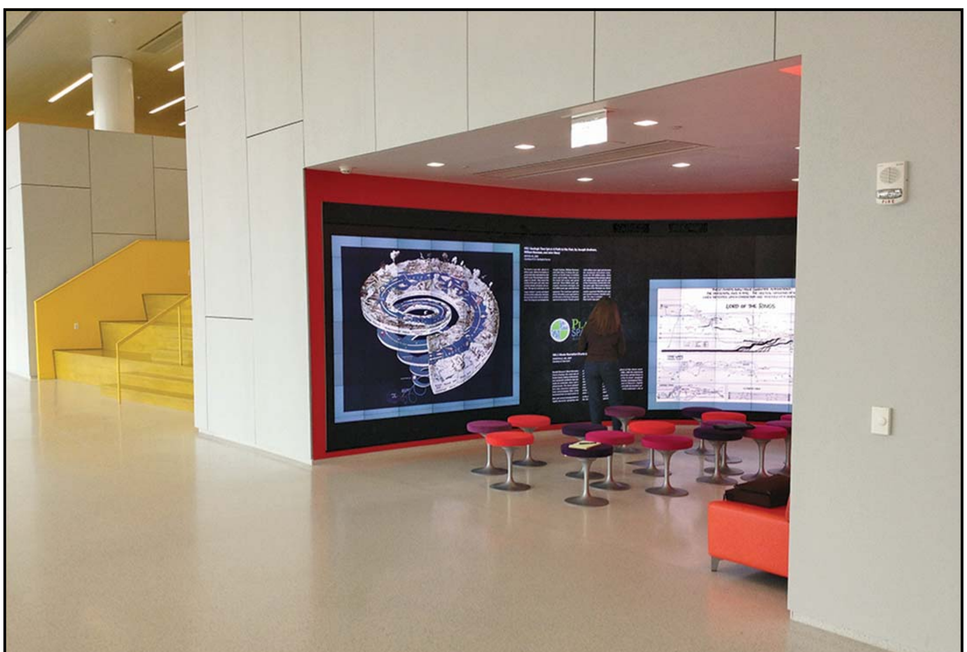
Illuminated Diagram Display
on display at the Smithsonian in DC.
http://scimaps.org/exhibit_info/#ID

26


Geographic Map: Where Science Gets Done
Science Map: How Scientific Disciplines Relate
North America
Central America
South America
Europe
Africa
North & West Asia
South & East Asia
Australia
Oceania
Antarctica
Math & Physics
Chemistry
Health Professionals
Medical Specialties
Social Sciences
Aeronautical, Chemical, Mechanical & Civil Engineering
Electrical Engineering & Computer Science
Biotechnology
Brain Research
Infectious Diseases
Earth Sciences
Biology
Humanities
About
This Illuminated Diagram display adds the flexibility of an interactive program to the incredibly high data density of a print. This technique is generally useful when there is too much pertinent data to be displayed on a screen but the data is relatively stable. The computer can direct the eye to what's important by using projectors or screens as smart spotlights, animating the research impact of individuals, giving a "grand tour" of science, or highlighting query results (as when you touch the lectern or use the keyboard) with an overlay of moving light.
Top Five Continents
North America - 4,000 records
South & East Asia - 3,589
Australia - 2,431
Africa - 2,208
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
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.
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
People & Topics
PLACES & SPACES
MAPPING SCIENCE
http://scimaps.org
27


Geographic Map: Where Science Gets Done
Science Map: How Scientific Disciplines Relate
North America
Central America
South America
Europe
Africa
North & West Asia
South & East Asia
Australia
Oceania
Antarctica
Math & Physics
Chemistry
Health Professionals
Medical Specialties
Social Sciences
Aeronautical, Chemical, Mechanical & Civil Engineering
Electrical Engineering & Computer Science
Biotechnology
Brain Research
Infectious Diseases
Earth Sciences
Biology
Humanities
About
This Illuminated Diagram display adds the flexibility of an interactive program to the incredibly high data density of a print. This technique is generally useful when there is too much pertinent data to be displayed on a screen but the data is relatively stable. The computer can direct the eye to what's important by using projectors or screens as smart spotlights, animating the research impact of individuals, giving a "grand tour" of science, or highlighting query results (as when you touch the lectern or use the keyboard) with an overlay of moving light.
Elinor Ostrom - Nobel Prize in Economic Sciences 2009
Born: 7 August 1933, New York, NY, USA
Affiliation at the time of the award: Indiana University, Bloomington, IN, USA, Arizona State University, Tempe, AZ, USA
Prize motivation: "for her analysis of economic governance, especially the commons"
Field: Economic governance
Contribution: Challenged the conventional wisdom by demonstrating how local property can be successfully managed by local commons without any regulation by central authorities or privatization.
Interact
Select any location on the Geographic Map location (by brushing your finger over an area on the lectern's touch screen) and topics studied in that area will highlight on the Science Map: the brighter a topic glows, the more papers on that topic originated in the selected area. Conversely, touching a scientific area in the Science Map illuminates places on the Geographic Map where that topic is studied. People and topic buttons support the exploration of publication output by selected Noble laureates and particular lines of research using MEDLINE data from 2000-2009.
Cancer
Cloning
HIV
Robert G. Edwards
Roger D. Kornberg
Elinor Ostrom
Obesity
Quality of Life
Smoking
Stanley B. Prusiner
Ahmed H. Zewail
View All
Keyword Search
PLACES & SPACES
MAPPING SCIENCE
http://scimaps.org
28

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



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

Places & Spaces: Mapping Science Exhibit

http://scimaps.org

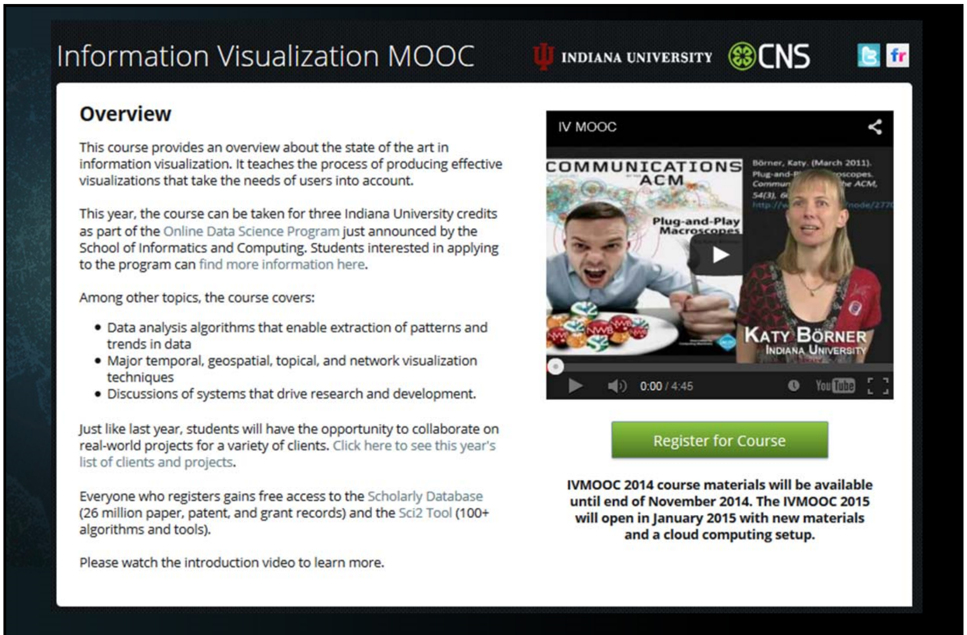
Maps are available for sale and the exhibit can be hosted by anyone.

31

Visualizing STI Model Results

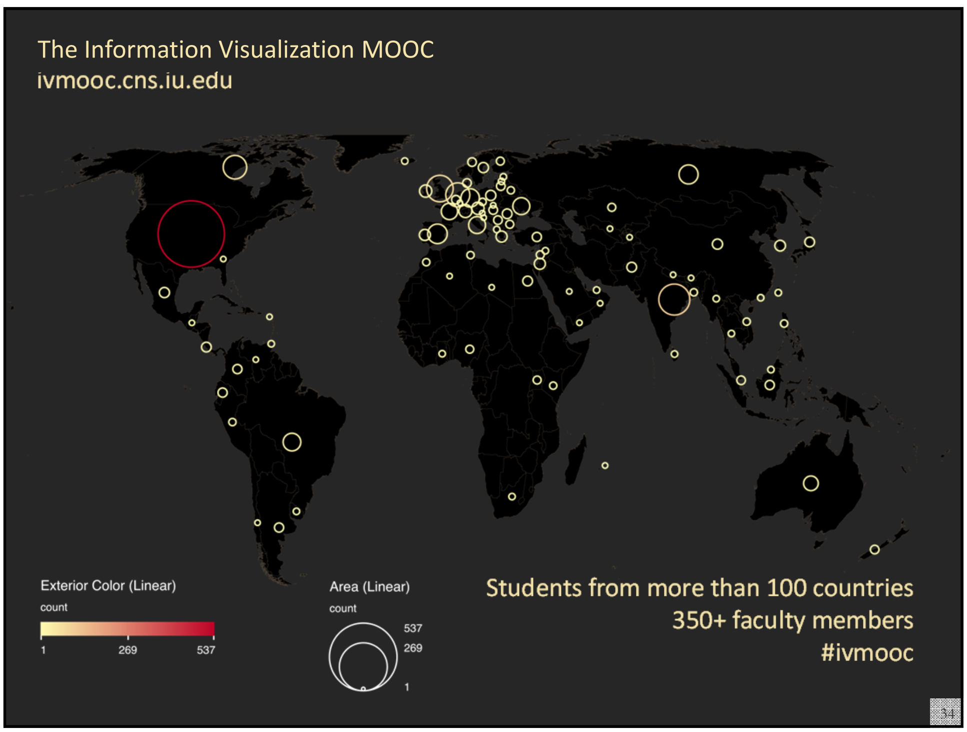
Example: The Information Visualization MOOC

32



Register for free at http://ivmooc.cns.iu.edu. Class will restart in January 2015.

33



Course Schedule

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

Mid-Term

Students work in teams with clients.

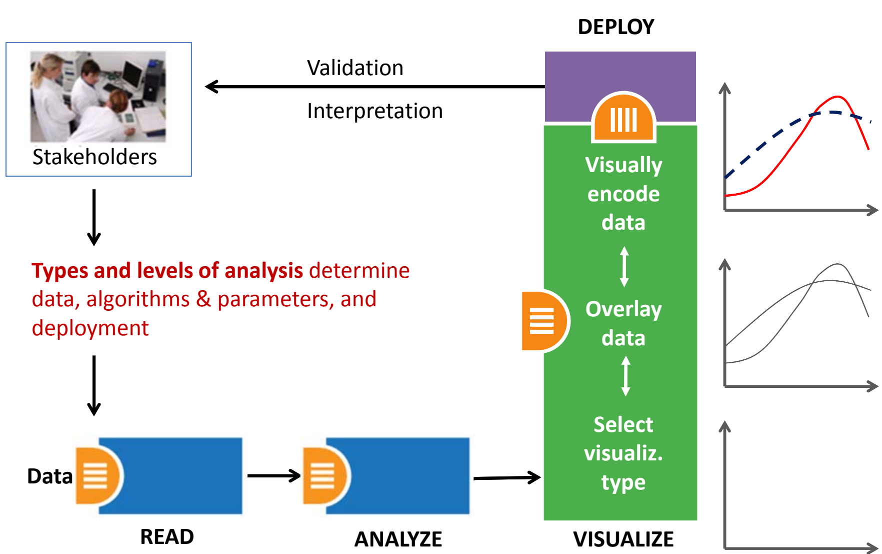
- **Session 5** – “With Whom:” Trees
- **Session 6** – “With Whom:” Networks
- **Session 7** – Dynamic Visualizations and Deployment

Final Exam

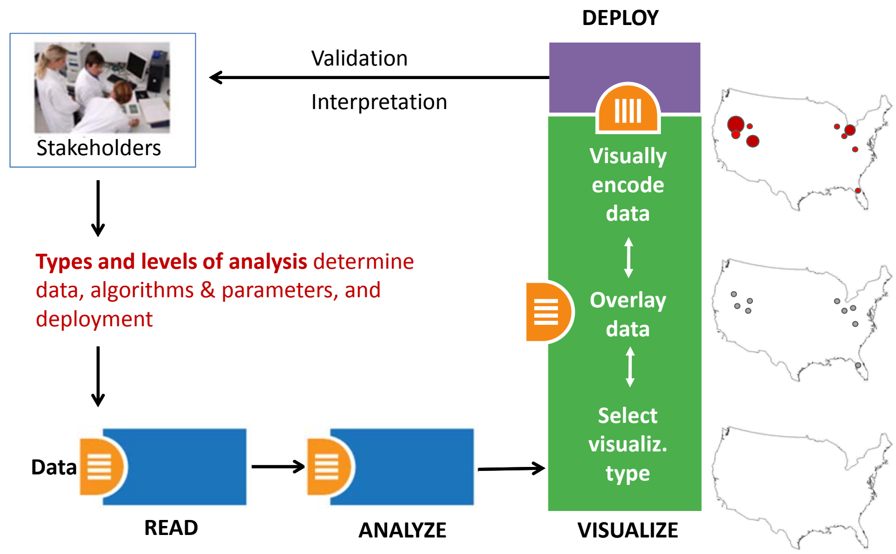
Final grade is based on Midterm (**30%**), Final (**40%**), Client Project (**30%**).

35

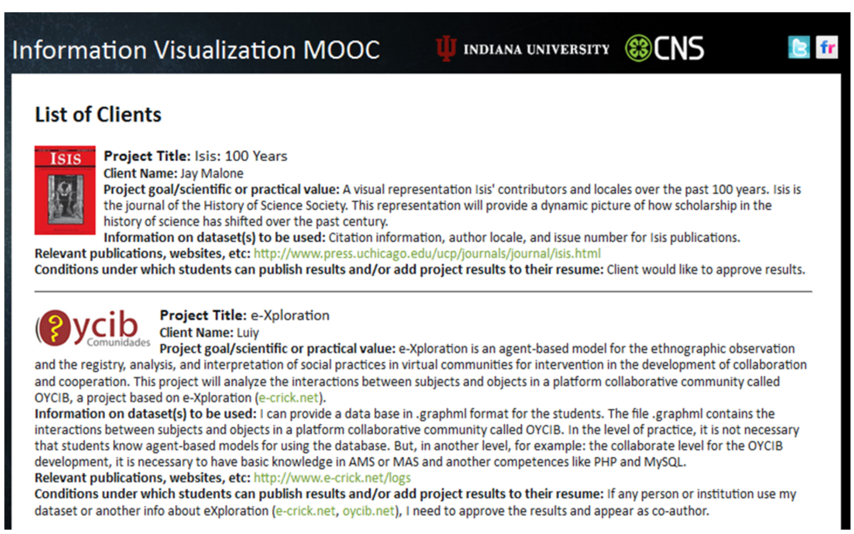
Needs-Driven Workflow Design



Needs-Driven Workflow Design

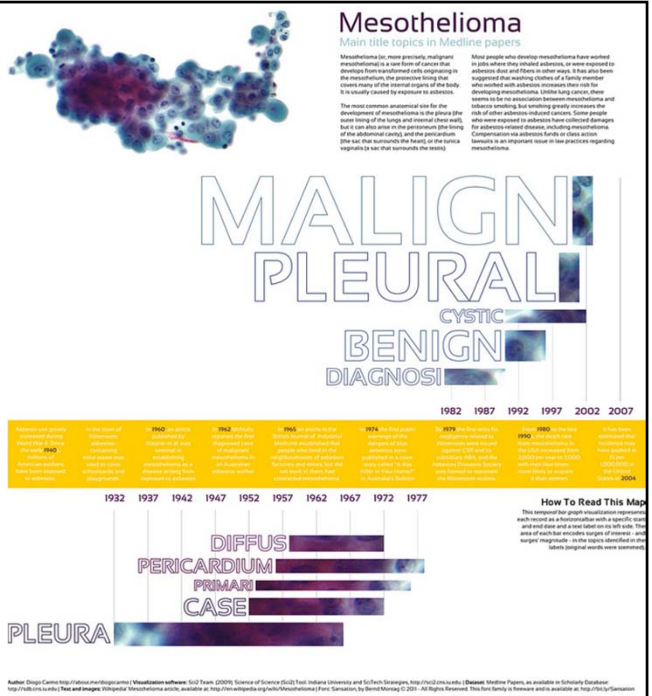


Clients

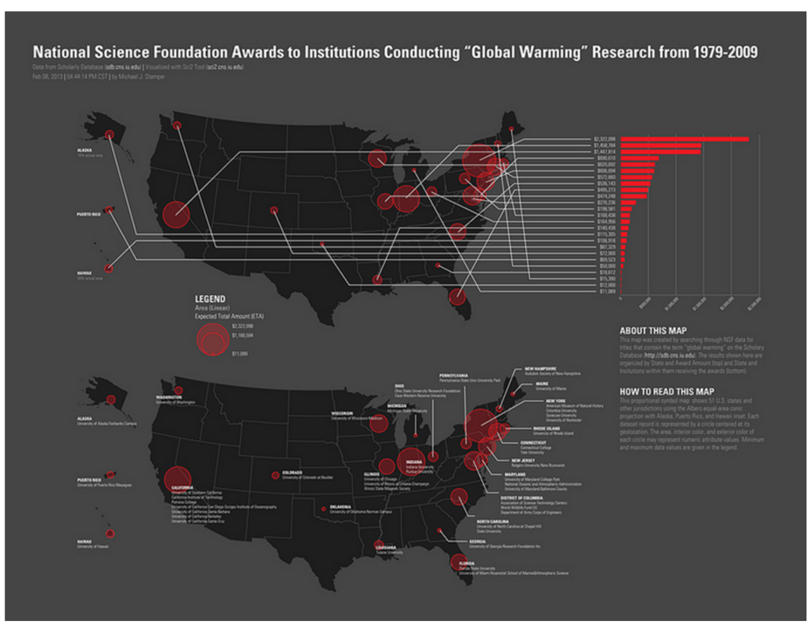


http://ivmooc.cns.iu.edu/clients.html

Diogo Carmo



39



mjstamper_ivmooc

40

Humanexus

Watch the official trailer »

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

41

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.

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

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

42

All papers, maps, tools, talks, press are linked from http://cns.iu.edu
These slides will soon be at http://cns.iu.edu/docs/presentations

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

43