NSF WORKSHOP on
Knowledge Management & Visualization Tools
IN SUPPORT OF DISCOVERY

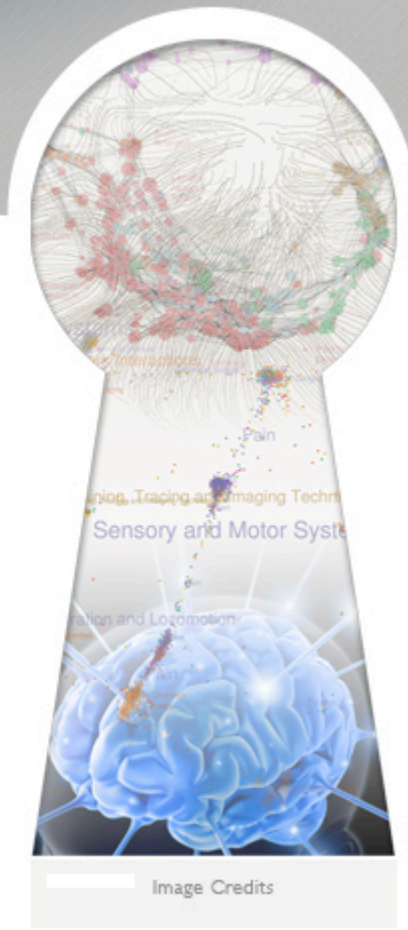


Image Credits

Welcome Everybody!

Date
March 10 & 11, 2008

Meeting Place
NSF, Room II-555, 4201 Wilson Boulevard, Arlington, VA

Organizers


PR^2

Katy Börner
Indiana University
Mapping Science (PNAS Issue, Exhibit), CI Design (IVC, NWB Tool)


PR^2

Luis M. A. Bettencourt
Los Alamos National Laboratory
Social Dynamics and Organization, Information systems for streaming data, Innovation and Development


PR^2

Mark Gerstein
Yale University
Genomics, Proteomics, Structural Genomics, Computational Biophysics


PR^2

Stephen Miles Uzzo
New York Hall of Science
Ecology, Scientific Visualization, Cybernetics, Education and Epistemology


PR^2

Weixia (Bonnie) Huang
Indiana University (NWB Tool System Architect)
note-taking

Workshop (I) Goals

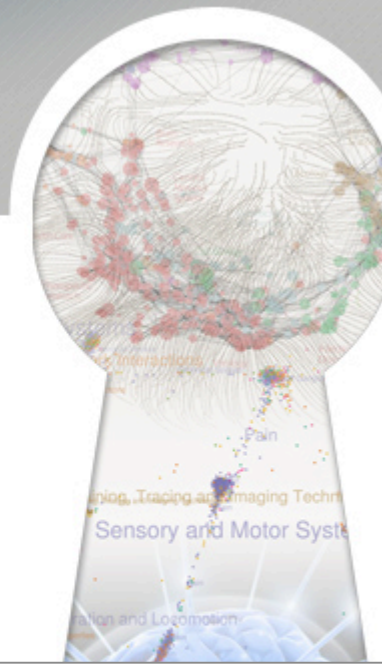
- Identify major needs that can be solved in the next 10 years.
- Provide detailed use scenarios.

- List existing approaches, partial solutions.

- Incentive structures (learn from eBay, Wikiedia, etc.)
- Sustainability.

as input to Workshop (II).

The envisioned approaches/tools must be so desirable, lucrative, (and feasible) that people start throwing money at it.



Why Biomedicine and Science of Science?

Both domains face serious data (streams) integration, mining, visualization problems and both promise major solutions to important challenges.

They are sufficiently different – solutions that work might transfer to other domains.

Workshop organizers have expertise in these domains.

Science Policy Challenges relate to the study of:

- The evolution of scientific communities/fields – birth, growth, maturation, decline.
- Interactions among fields. Who 'eats' who's papers?
- Trends, patterns, or emergent research frontiers, feedback loops, etc.
- Interplay of competition and collaboration.
- Diffusion of people, ideas, skills, etc. in geospatial space and topic space.
- Effects of different funding models, e.g., few large vs. many small grants.

Biomedical challenges comprise:

- Build the Encyclopedia of Life, Watch E. O. Wilson's TED Prize wish
- Prevent the Next Pandemic, Watch Larry Brilliant's TED Prize wish
- Creating an Inventory of Genotypes and Using it for Clean Energy and Nutrition. Watch Craig Venter's TED talk.

The Power of Space and Visuals

- Humans evolved being immersed in a 3D world.
- Location serves as a powerful index into human memory, e.g., memory palaces used in classical times.
- Capitalize on spatial processing capacities of human brain.

Well designed visualizations

- Provide the ability to comprehend huge amounts of data.
- Reduce search time and reveal relations otherwise not being noticed.
- Facilitate hypothesis formulation.
- Are effective sources of communication.

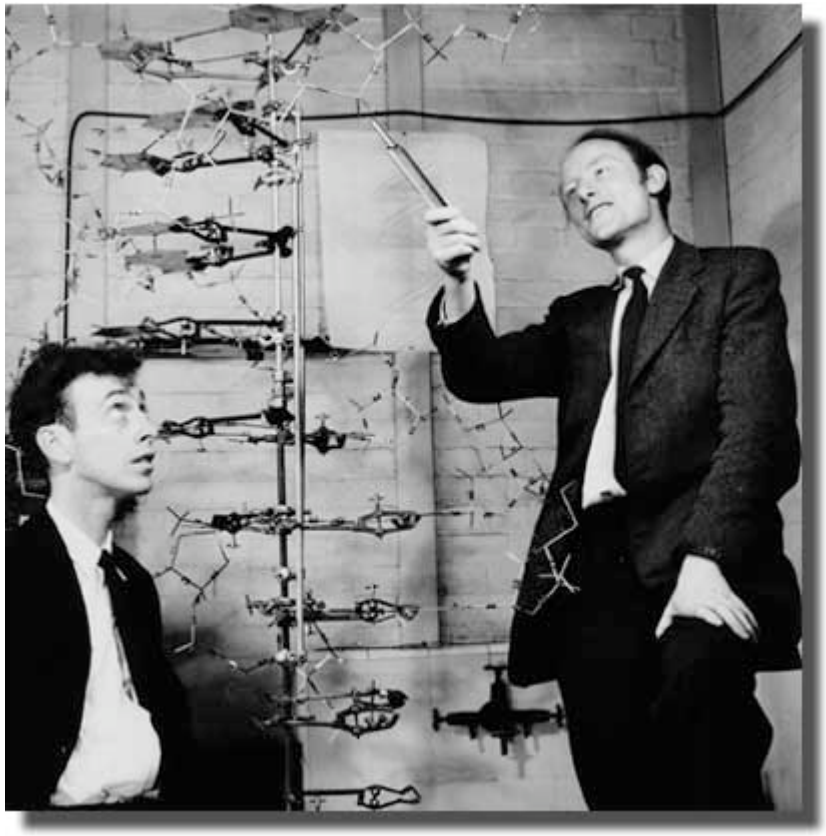
Ease data/information/knowledge/expertise access & management by supporting:

- Spatial navigation: mimics our experiences in physical world
- Semantic navigation: driven by semantic relationships or underlying logic.
- Social navigation: takes advantage of the behavior of like-minded people.

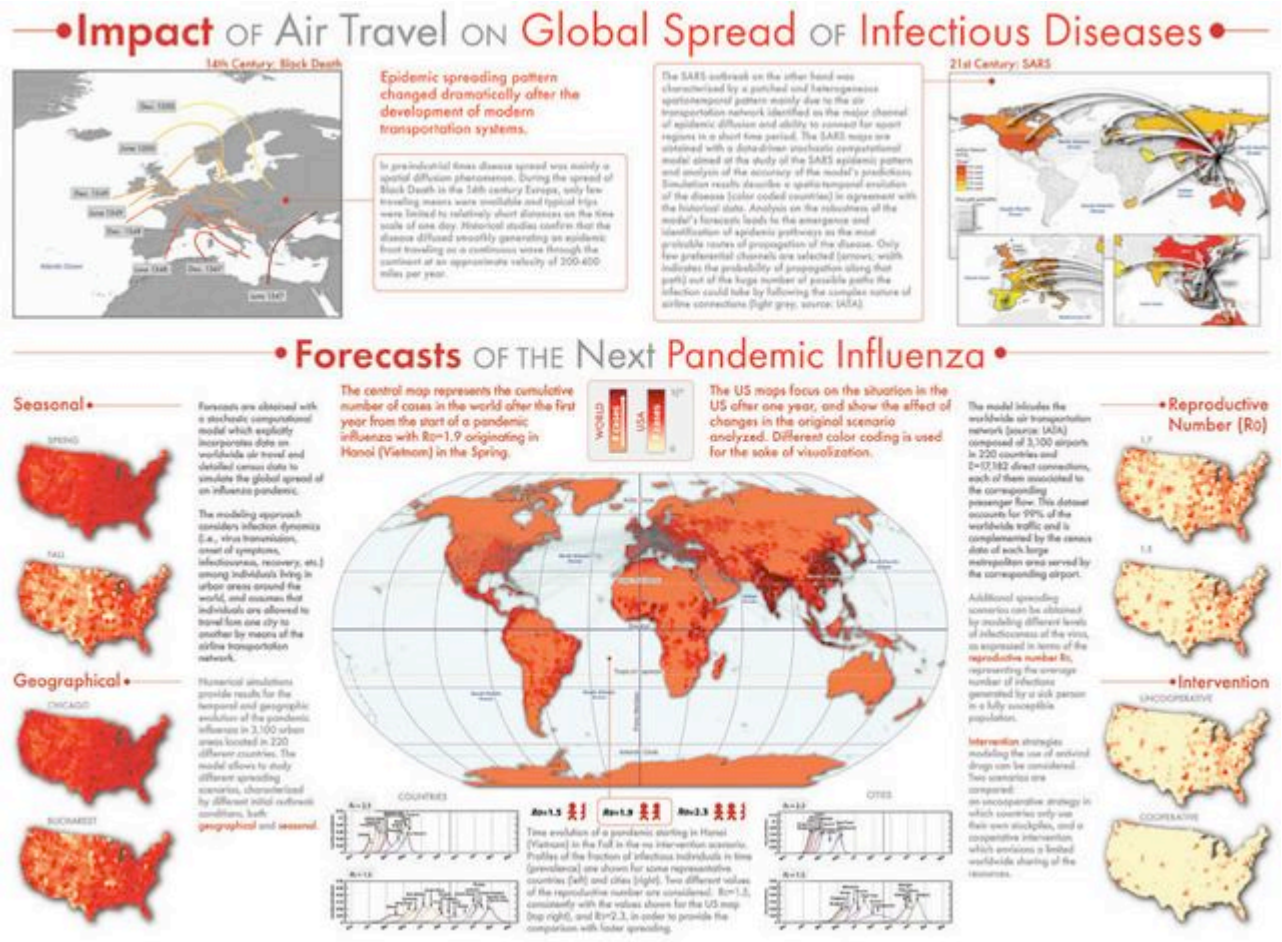
The Power of Space and Visuals – Biomedicine

Mark Gerstein: *"The most insightful visualization that I know about was, of course, Watson & Crick's synthesis of Franklin and Wilkins' diffraction data into an easy to comprehend 3D model and then going on to show how this could explain many of the fundamental processes of genetics. The power of this visualization, of course, stems from the fact that the 3D structure represented a real chemical entity.*

It also connected two disciplines – chemistry and genetics -- with a central visual metaphor that proves useful to this day."



The Power of Space and Visuals – Biomedicine



Epidemic Modeling in Complex realities, V. Colizza, A. Barrat, M. Barthelemy, A.Vespignani, Comptes Rendus Biologie, 330, 364-374 (2007).

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

The Power of Space and Visuals – Science Studies

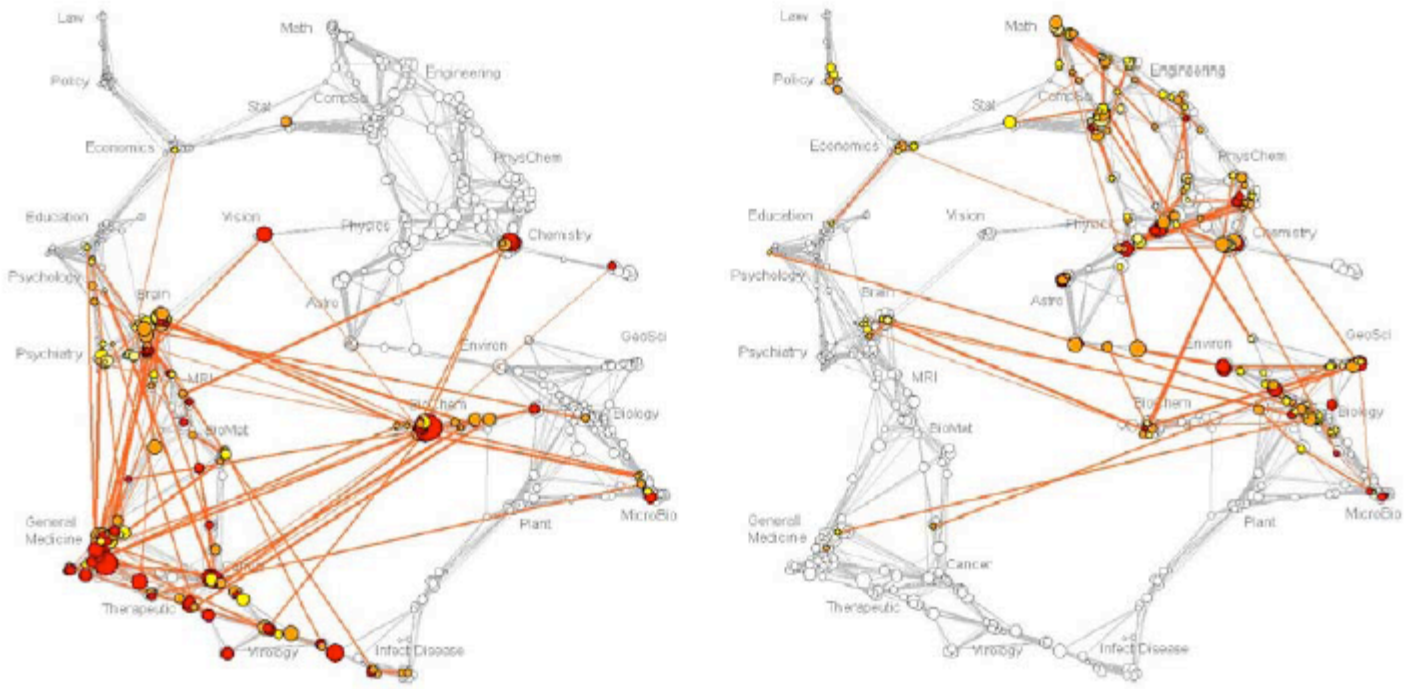
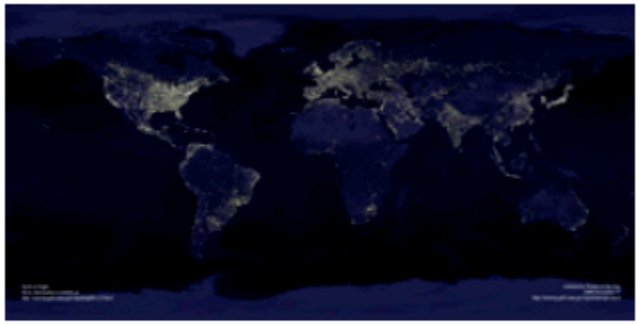


Figure 1: 2002 base map (top). Each node is a cluster of journals, and is sized to show numbers of papers in the journal cluster. NIH (bottom left) and NSF (bottom right) funding profile overlays on the 2002 base map. Colored nodes show the distribution and numbers of papers tied to grants; red nodes indicate faster moving science than yellow nodes; colored edges show linkages in the funding profiles that are stronger than the corresponding linkages in the base map.

Boyack, Kevin W, Börner, Katy & Klavans, Richard. (2007). Mapping the Structure and Evolution of Chemistry Research. Proceedings of the 11th International Conference on Scientometrics and Informetrics (ISSI 2007), Madrid, Spain, June 25-27, pp. 112-123.

The Power of Space and Visuals – Global Challenges

2005 World Population
The population map uses a quarter degree box resolution. Boxes with zero people are given in white. Darker shades of red indicate higher population counts up to 11,687,850 people per box using a logarithmic interpolation. The highest density boxes appear in ???. The People's Republic of China and India are the only two countries to have a more than a billion inhabitants.



Night on Earth
This image shows city lights at night. It was composed from hundreds of pictures made by orbiting satellites. The seaboards of Europe, the eastern United States, and Japan are particularly well lit. Many cities exist near rivers or oceans so that goods can be exchanged cheaply by boat. The central parts of South America, Africa, Asia, and Australia are rather dark despite their high population density, see map to the left.

2003 Scientific Productivity
Shown is where science is performed today, see >>> Part IV: Illuminated Diagram -- GeoMap for details. Each circle indicates a geographic location at which scholarly papers are published. The larger the circle the more papers are produced. Boston, MA, London, England, and New York, NY are the top three paper production areas. Note the strong resemblance with the Night on Earth and the IP Ownership maps and the striking differences to the world population map.

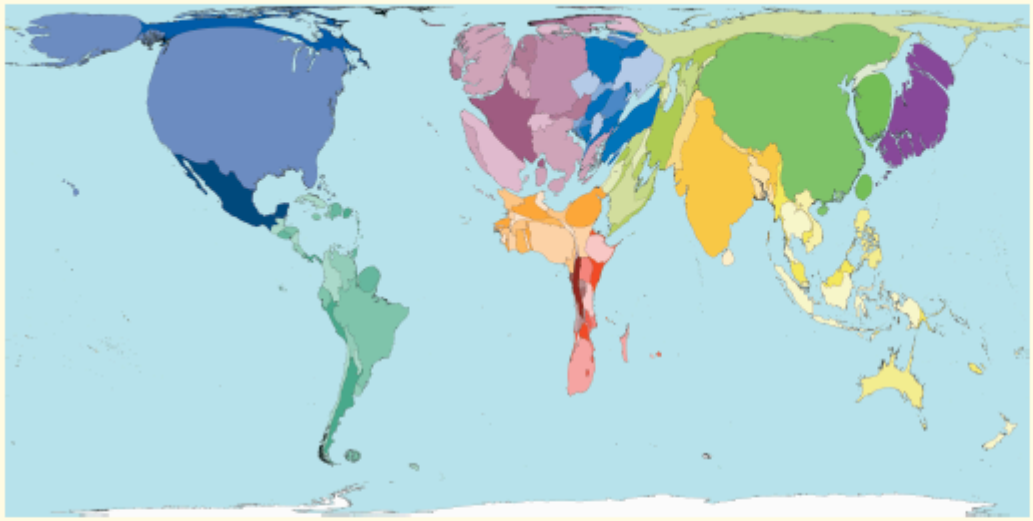
2007 IP Address Ownership
This map shows IP address ownership by location. Each owner is represented by a circle and the area size of the circle corresponds to the number of IP addresses owned. The larges circle denotes MIT's holdings of 563,796,332 IP addresses. The countries that own the most IP addresses are US (560 million), Japan (130 million), Great Britain (47 million). There are strong correlations between population and IP address ownership in economically developed countries. In other areas of the world, Internet access is sparse and often limited to urban areas.

Katy Börner. Atlas of Science: Guiding the Navigation and Management of Scholarly Knowledge, ESRI Press.

The Power of Space and Visuals – Global Challenges

Ecological Footprint

Produced by the SASI group (Sheffield) and Mark Newman (Michigan)



The ecological footprint is a measure of the area needed to support a population's lifestyle. This includes the consumption of food, fuel, wood, and fibres. Pollution, such as carbon dioxide emissions, is also counted as part of the footprint.

The United States, China and India have the largest ecological footprints. Without knowing population size we cannot understand what this means about individuals' ecological demands. Large populations live in China and India. In both territories resource use is below the world average. The per person footprint in the United States is almost five times the world average, and almost ten times what would be sustainable.

Territory size shows the proportion of the worldwide ecological footprint which is made there.

Land area

Technical notes

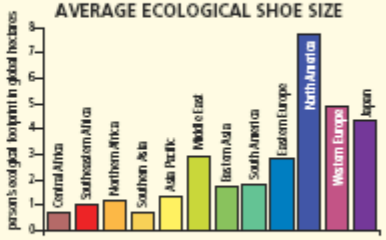
- Data are from the WWF (Worldwide Fund for Nature) International and Institute of Zoology.
- Ecological footprint is measured in global hectares. One global hectare is an area that has the world average biological productivity of one hectare.
- See website for further information.

LARGEST AND SMALLEST ECOLOGICAL SHOE SIZES

Rank	Territory	Value	Rank	Territory	Value
1	United Arab Emirates	10.6	191	Nepal	0.61
2	United States	9.7	192	Democratic Republic of Congo	0.58
3	Greenland	7.7	193	Zambia	0.58
3	Bahamas	7.7	194	Congo	0.58
5	Canada	7.5	195	Malawi	0.57
6	Kuwait	7.4	196	Haiti	0.57
7	Australia	7.0	197	Cambodia	0.55
8	Finland	6.8	198	Bangladesh	0.47
9	Estonia	6.1	199	Somalia	0.23
10	New Zealand	6.1	200	Afghanistan	0.11

ecological footprint in global hectares per person, 2002

AVERAGE ECOLOGICAL SHOE SIZE



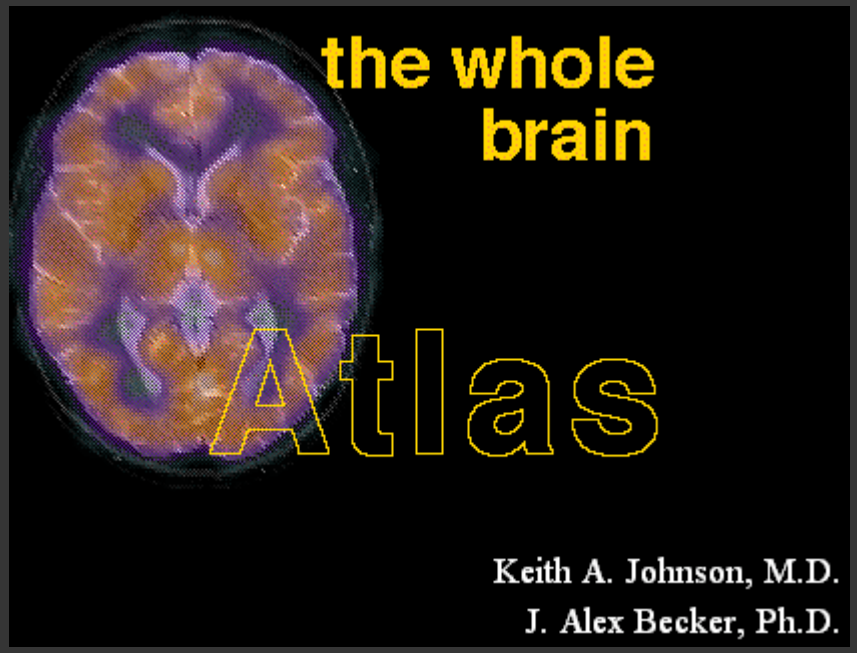
"People consume resources and ecological services from all over the world, so their footprint is the sum of these areas, wherever they may be on the planet."

The Living Planet Report, 2006

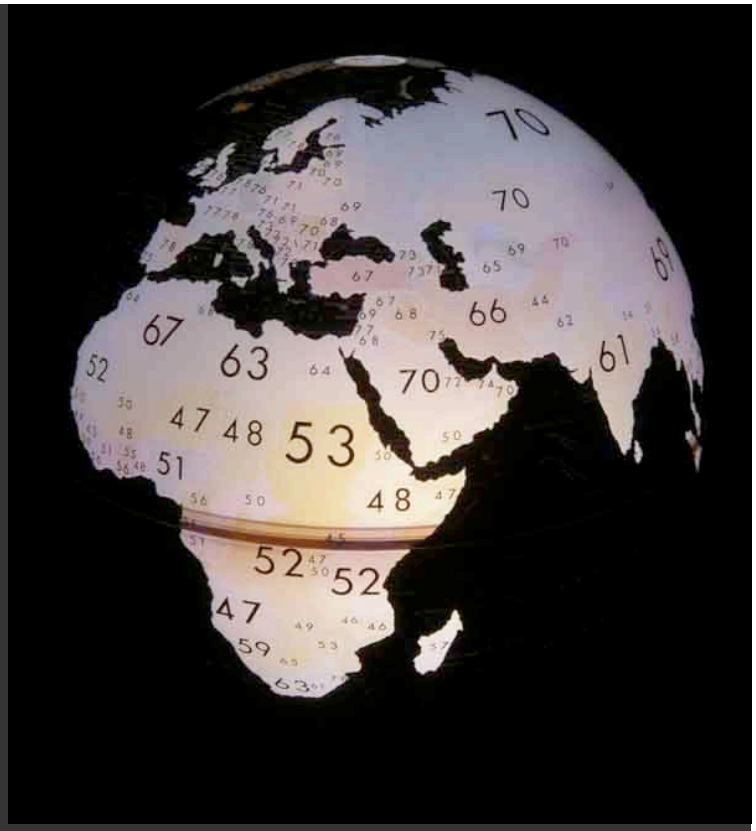
Map 322

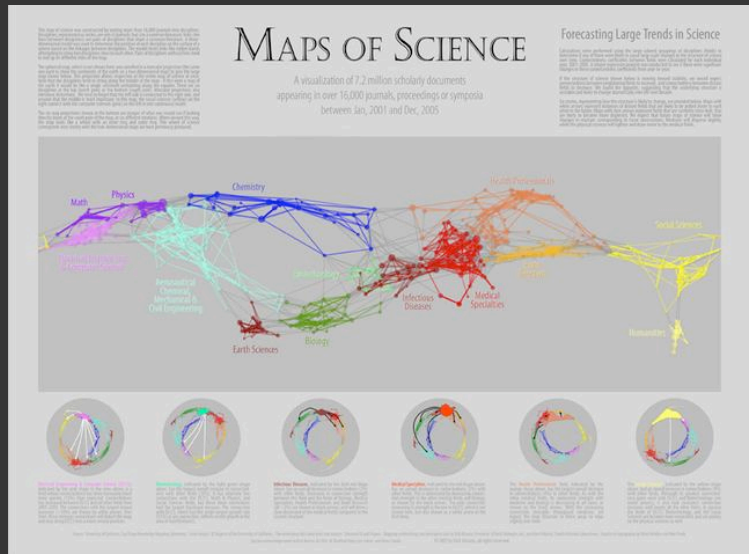
http://worldprocessor.com/
http://www.kodomo-project.org/worldprocessor
By Ingo Gunther

Atlas of the Body



http://www.med.harvard.edu/AANLIB





What reference systems are useful?
How to interlink/navigate them?

NSF WORKSHOP on
Knowledge Management & Visualization Tools
IN SUPPORT OF DISCOVERY

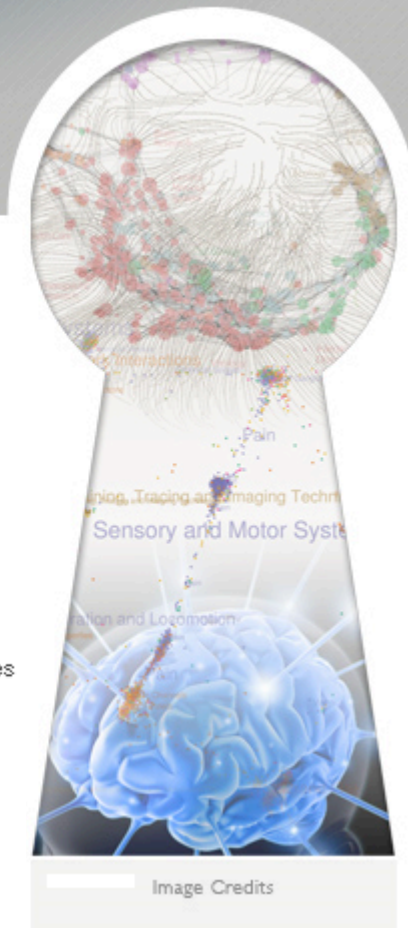


Image Credits

White Paper & Workshop (II)

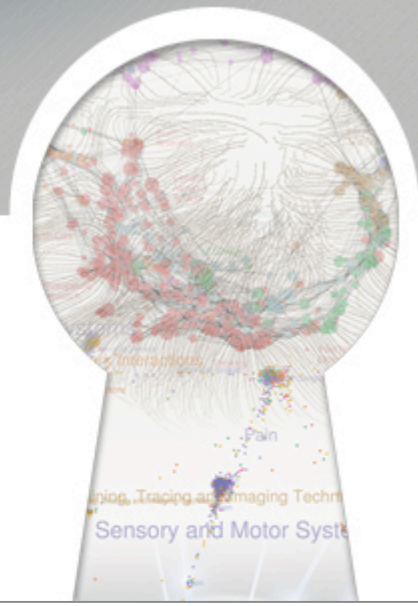
Results from Workshop (I) will be shared
within 2 weeks with participants of Workshop (II).

Date
April 7 & 8, 2008

Meeting Place
New York Hall of Science, Queens,
Lower Level Boardroom
5 minutes from La Guardia airport.

- **Kevin W. Boyack**, SciTech Strategies
  Science Indicators and Maps
- **Noshir Contractor**, UIUC & NCSA
  Professor of Speech Communication, Psychology, Networks in Communities
- **Anthony Diaz**, The Weather Channel, Inc.*
  Meteorologist
- **Ingo Gunther**, Artist
  Worlprocessor
- **Tony Hay**, Microsoft Research*
  Corporate Vice President for Technical Computing
- **Sharon Jordan**, U.S. Department of Energy
  Office of Scientific and Technical Information
- **Hiroaki Kitano**, Sony, Japan*
  Symbiotic Systems
- **David Lazar**, Harvard University
  Social Science Sociology
- **Barend Mons**, Erasmus University, Netherlands & KnewCo, Inc.
  Wikiproteins
- **Avi Silberschatz**, Yale University*
  Real-time Database Systems, Multidatabase Transaction Management, Knowledge Discovery
- **Ben Shneiderman**, University of Maryland
  Interface Design, Visualization, Creativity Support Tools
- **Michael Stonebraker**, MIT*
  Database Research and Technology
- **Walter L. Warnick**, U.S. Department of Energy
  Director of the Office of Scientific and Technical Information

NSF WORKSHOP on
Knowledge Management & Visualization Tools
IN SUPPORT OF DISCOVERY



Agenda

Day 1:

12:00pm	Welcome by Organizers by Katy Borner
12:15pm	Introduction by Participants (5 min per person/organization). Led by Stephen Uzzo
2:00pm	*Break*
2:15pm	Workshop Goals. Presentation by NSF Officer
2:30pm	Challenges and Opportunities by Luis Bettencourt
3:00pm	Breakout Sessions on "$10 Million SciPolicy and Bio Challenge". Intro by Stephen Uzzo
4:00pm	Breakout Session Reports
4:30pm	Interactive Timeline Assembly - see connections and on them. Led by Alex Pang
6:30pm	*Adjourn*
7:00pm	*Joint dinner*

Day 2:

9:00am	*Light Breakfast*
9:30am	Invited Presentation/Inspiration by Mark Gerstein
10:00am	Breakout Sessions on "Envision and Draw your Dream Tool" Intro by Katy Borner
11:00am	Breakout Session Reports
11:30am	Invited Presentation/Inspiration by Kevin Boyack
12:00pm	*Joint Lunch*
1:00pm	Write Description of 2nd Best Idea for CDI Grant Proposal. Led by Alex Pang
2:00pm	Presentation to Group
2:45pm	*Break*
3:00pm	Collective Exercise on "Who would like to collaborate with whom on what?" Lead by Katy Borner
4:00pm	Discussion of Next Steps, Funding Opportunities, etc.
5:00pm	*Adjourn*