Maps & Macroscopes: Envisioning Science, Technology, and Education

Katy Börner

Victor H. Yngve Distinguished Professor of Intelligent Systems Engineering & Information Science
Director, Cyberinfrastructure for Network Science Center
School of Informatics and Computing
Indiana University Network Science Institute
Indiana University, USA

Third Annual Innovation Conference by JCITR
Memorial Union, Indiana University, Bloomington
Thursday, April 6, 2017
How can we communicate the beauty, structure, and dynamics of science to a general audience?
The Structure of Science

The Social Sciences are the smallest and most diffuse of all the sciences. Psychology serves as the link between Medical Sciences (Psychiatry) and the Social Sciences. Statistics serves as the link with Computer Science and Mathematics.

Mathematics is our starting point, the parent of all sciences. It lies at the outer edge of the map. Computer Science, Electrical Engineering, and partly applied sciences that draw upon knowledge in Mathematics and Physics. These three disciplines provide a good example of a linear progression from one pure science (Mathematics) to another (Physics) through multiple disciplines. Although applied, these disciplines are highly concentrated with distinct bands of research communities that link them. Bands indicate interdisciplinary research.

Research is highly concentrated in Physics and Chemistry. These disciplines have few, but very distinct, bands of research communities that link them. The lack of these bands indicates an extensive amount of interdisciplinary research, which suggests that the boundaries between Physics and Chemistry are not as distinct as one might assume.

The Life Sciences, including Biology and Biochemistry, are less concentrated than Chemistry or Physics. Bands of linking research can be seen between the larger areas in the Life Sciences; for instance between Biology and Microbiology and between Biology and Environmental Science. Microbiology is very interesting as it is a large discipline that has visible links to disciplines in many areas of the map, including Biology, Chemistry, Neuroscience, and General Medicine. It is perhaps the most interdisciplinary of the sciences.

We are all familiar with traditional maps that show the relationships between countries, provinces, states, and cities. Similar relationships exist between the various disciplines and research topics in science. This allows us to map the structure of science.

One of the first maps of science was developed at the Institute for Scientific Information over 30 years ago. It identified 41 areas of science from the citation patterns in 17,000 scientific papers. That early map was intriguing, but it didn't capture enough of science to account for its structure.

Things are different today. We have enormous computing power and advanced visualization software that make mapping the structure of science possible. This galaxy-like map of science left was generated at Sandia National Laboratories using an advanced graph layout routine (YEd) from the citation patterns in 800,000 scientific papers published in 2011. Each dot in the galaxy represents one of the 96,000 research communities active in science in 2002. A research community is a group of papers (9 or average) that are written on the same research topic in a given year. Over time, communities can be born, continue, split, merge, or die.

The map of science can be used as a tool for science strategy. This is the terrain in which organizations and institutions locate their scientific capabilities. Additional information about the scientific and economic impact of each research community allows policy makers to decide which areas to explore, expand, eliminate, or ignore.

We also envision the maps as educational tools. For children, the theoretical relationship between areas of science can be replaced with a concrete map showing how math, physics, chemistry, biology, and social studies interact. For advanced students, areas of interest can be located and neighboring areas can be explored.

Nanotechnology
Most research communities in nanotechnology are concentrated in Physical Chemistry and Materials Science. However, many disciplines in the Life and Medical Sciences also have nanotechnology applications.

Proteomics
Research communities in proteomics are centered in Biochemistry. In addition, there is a heavy focus in the tools section of chemistry, such as Chromatography. The balance of the proteomics communities are widely dispersed among the Life and Medical Sciences.

Pharmacogenomics
Pharmacogenomics is a relatively new field, with much of its activity in Medicine. It also has many communities in genetic analysis and two communities in the Social Sciences.
Debut of 5th Iteration of the Mapping Science Exhibit at MEDIA X in 2009 at Wallenberg Hall, Stanford University.
Science Maps in “Expedition Zukunft” science train visited 62 cities in 7 months. Opening on April 23rd, 2009 by German Chancellor Merkel
Ingo Gunther's Worldprocessor globe design on display at the Museum of Emerging Science and Innovation in Tokyo, Japan.
Places & Spaces Digital Display in North Carolina State’s Immersion Theater
Exhibit Advisors and Ambassadors
Places & Spaces at Northwestern University
May 14 - September 23, 2015
Places & Spaces Exhibit at the David J. Sencer CDC Museum, Atlanta, GA

CDC Opening Event: Maps of Health
Tutorial and Symposium
February 4-5, 2016
Places & Spaces Exhibit at Vanderbilt University, Nashville, TN.
January 23-April 23, 2017  http://scimaps.org/vanderbilt
Maps

PLACES & SPACES
MAPPING SCIENCE

scimaps.org
10 iterations over 10 years

equal

$10 \times 10 = 100$ maps!
Maps that show STRUCTURE

scimaps.org
A New Map of the Whole World with Trade Winds According to the Latest and Most Exact Observations - Herman Moll - 1736
Maps that show FLOWS

scimaps.org
**Impact of Air Travel on Global Spread of Infectious Diseases**

14th Century: Black Death

Epidemic spreading pattern changed dramatically after the development of modern transportation systems.

In pre-industrial times disease spread was mainly a spatial diffusion phenomenon. During the spread of Black Death in the 14th century Europe, only few traveling means were available and typical trips were limited to relatively short distances on the time scale of one day. Historical studies confirm that the disease diffused smoothly generating an epidemic front traveling as a continuous wave through the continent at an approximate velocity of 200-400 miles per year.

21st Century: SARS

The SARS outbreak on the other hand was characterized by a patchy and heterogeneous spatio-temporal pattern mainly due to the air transportation network identified as the major channel of epidemic diffusion and ability to connect far apart regions in a short time period. The SARS maps are obtained with a data-driven stochastic computational model aimed at the study of the SARS epidemic pattern and analysis of the accuracy of the model’s predictions. Simulation results describe a spatio-temporal evolution of the disease (color coded countries) in agreement with the historical data. Analysis on the robustness of the model’s forecasts leads to the emergence and identification of epidemic pathways as the most probable routes of propagation of the disease. Only few preferential channels are selected (arrows, width indicates the probability of propagation along that path) out of the huge number of possible paths the infection could take following the complex nature of airline connections (light grey, source: IATA).

**Forecasts of the Next Pandemic Influenza**

Seasonal

Forecasts are obtained with a stochastic computational model which explicitly incorporates data on worldwide air travel and detailed census data to simulate the global spread of an influenza pandemic.

The modeling approach considers infection dynamics (i.e., virus transmission, onset of symptoms, infectiousness, recovery, etc.) among individuals living in urban areas around the world, and assumes that individuals are allowed to travel from one city to another by means of the airline transportation network.

Geographical

Numerical simulations provide results for the temporal and geographic evolution of the pandemic influenza in 3,100 urban areas located in 220 different countries. The model allows to study different spreading scenarios, characterized by different initial outbreak conditions, both geographical and seasonal.

The central map represents the cumulative number of cases in the world after the first year from the start of a pandemic influenza with R0 = 1.9 originating in Hanoi (Vietnam) in the Spring.

The US maps focus on the situation in the US after one year, and show the effect of changes in the original scenario analyzed. Different color coding is used for the sake of visualization.

The model includes the worldwide air transportation network (source: IATA) composed of 3,100 airports in 220 countries and 17,182 direct connections, each of them associated to the corresponding passenger flow. This dataset accounts for 99% of the worldwide traffic and is complemented by the census data of each large metropolitan area served by the corresponding airport.

Additional spreading scenarios can be obtained by modeling different levels of infectiousness of the virus, as expressed in terms of the reproductive number R0, representing the average number of infections generated by a sick person in a fully susceptible population.

**Intervention** strategies modeling the use of antiviral drugs can be considered. Two scenarios are compared: an uncooperative strategy in which countries only use their own stockpiles, and a cooperative intervention which envisages a limited worldwide sharing of the resources.

Impact of Air Travel on Global Spread of Infectious Diseases - Vittoria Colizza, Alessandro Vespignani - 2007
Ph.D. Thesis Map - Keith B. Nesbitt - 2004
Tracing of Key Events in the Development of the Video Tape Recorder - Mr. G. Benn, Francis Narin - 1968
Chemical Research & Development Powers the U.S. Innovation Engine

Macroeconomic Implications of Public and Private R&D Investments in Chemical Sciences

INVESTMENT IN CHEMICAL SCIENCE R&D

FEDERAL GOVERNMENT

$1 Billion
FEDERAL FUNDING

$5 Billion
INDUSTRY FUNDING

CHEMICAL INDUSTRY

$1B
4-5 YRS
FOUNDATION RESEARCH

$1B + $5 Billion
9-11 YEARS
INVENTION DEVELOPMENT

$10 Billion
> 5 YEARS
TECHNOLOGY COMMERCIALIZATION

TIMELINE FROM CONCEPTION TO COMMERCIALIZATION

$40 Billion
GROWTH IN GNP

+ 600,000
JOBS CREATED

U.S. ECONOMY

$8 Billion
TAXES

The design shows that an input of $1B in federal investment, leveraged by $5B industry investment, brings new technologies to market and results in $10B of operating income for the chemical industry, $40B growth in the Gross National Product (GNP) and further impacts the US economy by generating approximately 600,000 jobs, along with a return of $8B in taxes. Additional details, also reported in the CCR studies, are depicted in the map to the left. This map clearly shows the two R&D investment cycles: the shorter industry investment at the innovation stage to commercialization cycle; and the longer federal investment cycle which begins in basic research and culminates in national economic and job growth along with the increase tax base that in turn is available for investment in basic research.
Maps that show TRENDS

scimaps.org
Pulse of the Nation: U.S. Mood Variations Inferred From Twitter

Mood Variations
A number of interesting trends can be observed in the data. First, overall daily variations can be seen from the graphs, with the early morning and late evening having the highest level of happiness. Second, geographic variations can be observed from the graphs, with a region in the western US having consistently fewer hours behind the last coast.

Weekly Variations
Weekly trends can be observed as well, with weekdays seeing happier days than weekends.

About the Data and Visualization
The posts were collected using over 300 million tweets (Sep 2009 – Aug 2009) collected by MIRAGE researchers, represented as density-encoding cartograms. The mood of each tweet was inferred using LIWC software. MIRAGE, M. D. & Lang, J. M. Effective names for English words. LIWC (2001), instructions manual and software package. Technical report C-1. The Center for Research on Psychophysiologicals. University of Florida. Location data was taken from the U.S. Census Bureau at https://foreignarea.ssu.edu, and the base U.S. map data was taken from NASA/Geographic Information Systems. Location data were inferred using the Google Maps API and mapped into countries using PostGIS and U.S. country map from the U.S. National Atlas. Mood scores were related using color Brewer 2.

About Cartograms
A cartogram is a map in which the mapping variable is in scale; the area of the actual map is altered so that the area of each region is maintained as much as possible, but the area is scaled in order to be proportional to the number of tweets that indicate that mood. This result is a density-equalizing map. The cartograms in this work were generated using the Carto software by Mark E. J. Newman.

Northeastern University
College of Computer and Information Science
Center for Complex Network Research

http://www.ccs.neu.edu/home/amislw/twittermood

© 2010 Alan Mislove, Sune Lehmann, Yong-Yeol Ahn, Jukka-Pekka Onnela, and James Niels Rosenquist - 2010
The EMERGENCE of NANOTECHNOLOGY

MAPPING THE NANO REVOLUTION
The emergence of nanotechnology has been one of the major scientific-technological revolutions in the last decade and it led to a structural reorganization of major fields of science. Price (1965) showed that fields of science and their development can be mapped using aggregated citations among the journals in the fields and their relevant environments. The frames to the right show the evolving journal citation network for the years 1998-2003. Distances are proportional to cosine values between the citation patterns of the respective journals. Textual descriptions of key events during the development of Nanotechnology are given below each frame. Most notably, leading papers in Science and Nature catalyzed the breakthrough around 2000.

CHANGING ROLES OF DIFFERENT JOURNALS
The interdisciplinarity of a journal can be measured using betweeness centrality (BC) — journals that occupy many shortest paths between other journals in a network have higher BC values than those that do not. In the map, sizes of nodes are proportional to the betweenness centrality of the respective journal in the citation network. From being a specialist journal in applied physics, the journal Nanotechnology obtains a high BC value in the years of the transition, ca. 2001. This is preceded by the "intervention" of Science. After the transition, the new field of nanotechnology is established, new journals such as Nano Letters published by influential American Chemical Society join the lead, and a new specialty structure with low BC value journals results.

An animated sequence of this evolution is at: http://www.leydesdorff.net/journals/nanotech.

LEGEND
- Science
- Nature
- Nanotechnology
- Nano Letters

VALUES
- 0.82
- 0.72
- 0.53

References

1998
During the period 1998-2000, the journal Nanotechnology is part of a group of journals in applied physics.

1999
Increasingly, chemistry journals play a role in the citation impact environment of the journal Nanotechnology.

2000
The journal Science interfaces with relevant journals in both chemistry and applied physics. Nanotechnology emerges as core journal.

2001
The journal Nanotechnology now provides the interface between chemistry and physics. The "intervention" by Science is no longer needed.

2002
Other journals in nanoscience and technology begin to emerge, and the bridging role of the journal Nanotechnology gradually diminishes. Nano Letters and the Journal of Nanoscience and Nanotechnology join the new field of nanotechnology.

2003
The journal Science is relevant in the citation impact environment, but now functions as one of the specialist journals in nanotechnology. Nanoscience further develops as an increasingly integrated network of journals.
Macrosopes

PlACES & SPACES
MAPPING SCIENCE

scimaps.org
MAPS

vs.

MACROSCOPES
Microscopes & Telescopes vs. MACROSCOPES

The Infinitely Great
- Telescope
- Galaxies
- Society
- The Infinitely Complex

The Infinitely Small
- Microscope
- Cells
- Nature
- Technology

Deploy, Data, Visualize, Read, Analyze
Iteration XI (2015): Macrosopes for Interacting with Science
http://scimaps.org/iteration/11
Earth – Cameron Beccario
AcademyScope – National Academy of the Sciences & CNS
The News Co-occurrence Globe
An interactive visualization of how countries are mentioned together in the world's news media

2.92K

Mapping Global Society – Kalev Leetaru
Iteration XII (2016): Macroscopes for Making Sense of Science
http://scimaps.org/iteration/12
Four new macroscopes debut at Vanderbilt University:

1. **Smelly Maps**: Features a “smellscape” of 12 cities mapped by smell using social media.

2. **HathiTrust**: Highlights the diversity of publications collected in digital form by HathiTrust.

3. **Excellence Networks**: Compares how research institutions, such as Indiana and Vanderbilt universities, collaborate with one another.

4. **FleetMon**: Shows how the amount of shipping traffic that navigates the Strait of Malacca compared to other major shipping lanes of the world.

http://scimaps.org/vanderbilt
A visitor explores the macroscope kiosk at the Eskenazi Museum of Art at Indiana University.


**Background and Goals**

The Places & Spaces: Mapping Science exhibit is designed to open people's hearts and minds to the value, complexity, and beauty of maps of science and technology.

Drawing from across cultures and across scholarly disciplines, the Places & Spaces: Mapping Science exhibit demonstrates the power of accessible visualization tools to help everyone make sense of the world. Produced by SCIMaps, LLC.
IVMOOC.cns.iu.edu

Places & Spaces
Mapping Science

scimaps.org
Figure 1: Analysis types vs. user needs, taken from Emmons, Light, and Börner. "MOOC Visual Analytics: Empowering Teachers, Students, Researchers, and Developers of Massively Open Online Courses". Journal of the Association for Information Science and Technology (in press).
Proportional symbol map of the world showing the location of IVMOOC students from 2013 (blue) and 2014 (orange). Circles are area size coded by the number of students per country.
Scores vs. time invested watching course videos for students who took the 2013 (blue) and 2014 (orange) IVMOOC midterm (left) and final exam (right) and got at least 50% correct.
Exam Scores by Question

Student scores per question for midterm (left) and final exam (right) for IVMOOC 2014.
IVMOOC video views in 2013 (blue) and 2014 (orange)
Custom interactive visualizations of 2015 IVMOOC student engagement and performance data, explore functionality online at http://goo.gl/TYixCn
Student Flows – STEM Academic Career Pathways


Interactive web site:  http://demo.cns.iu.edu/client/stem
Forecasting S&T

places & spaces
mapping science

scimaps.org
Government, academic, and industry leaders discussed challenges and opportunities associated with using big data, visual analytics, and computational models in STI decision-making.

Conference slides, recordings, and report are available via [http://modsti.cns.iu.edu/report](http://modsti.cns.iu.edu/report)
Upcoming Colloquia

Unless otherwise indicated, most Sackler colloquia are held at the Arnold and Mabel Beckman Center, in Irvine, California.

**Reproducibility of Research: Issues and Proposed Remedies**

March 8-10, 2017; Washington, D.C.
Organized by David B. Allison, Richard Shiffrin and Victoria Stodden
Registration now open

**Science of Science Communication III**

November 15-16, 2017; Washington, D.C.
Organized by Karen Cook, Baruch Fischhoff, Alan I. Leshner and Dietram A. Scheufele
Registration will open May 2017

**Modelling and Visualizing Science and Technology Developments**

December 4-5, 2017; Irvine, CA
Organized by Katy Börner, William Rouse and H. Eugene Stanley
Registration will open August 2017

http://www.nasonline.org/programs/sackler-colloquia/upcoming-colloquia
References


All papers, maps, tools, talks, press are linked from http://cns.iu.edu
These slides are at http://cns.iu.edu/presentations.html
CNS Facebook: http://www.facebook.com/cnscenter
Mapping Science Exhibit Facebook: http://www.facebook.com/mappingscience