

Visualizing What We Know

Noodles and Co.

February 2017



PLACES & SPACES &

SPACES &

MAPPING SCIENCE





CNS

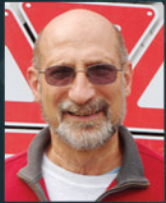
Cyberinfrastructure for
Network Science Center
cns.iu.edu



Places & Spaces On Display



Meet the international advisory board that helps select the maps that make up the exhibit



Gary
Berg-Cross



Bob
Bishop



Kevin W.
Boyack



Donna J.
Cox



Bonnie
DeVarco



Sara Irina
Fabrikant



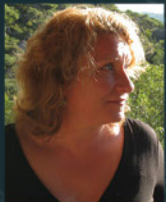
Marjorie M.K.
Hlava



Peter A.
Hook



Manuel
Lima



Deborah
MacPherson



Lev
Manovich



Carlo
Ratti



Eric
Rodenbeck



André
Skupin



Moritz
Stefaner



Stephen
Uzzo



Caroline
Wagner



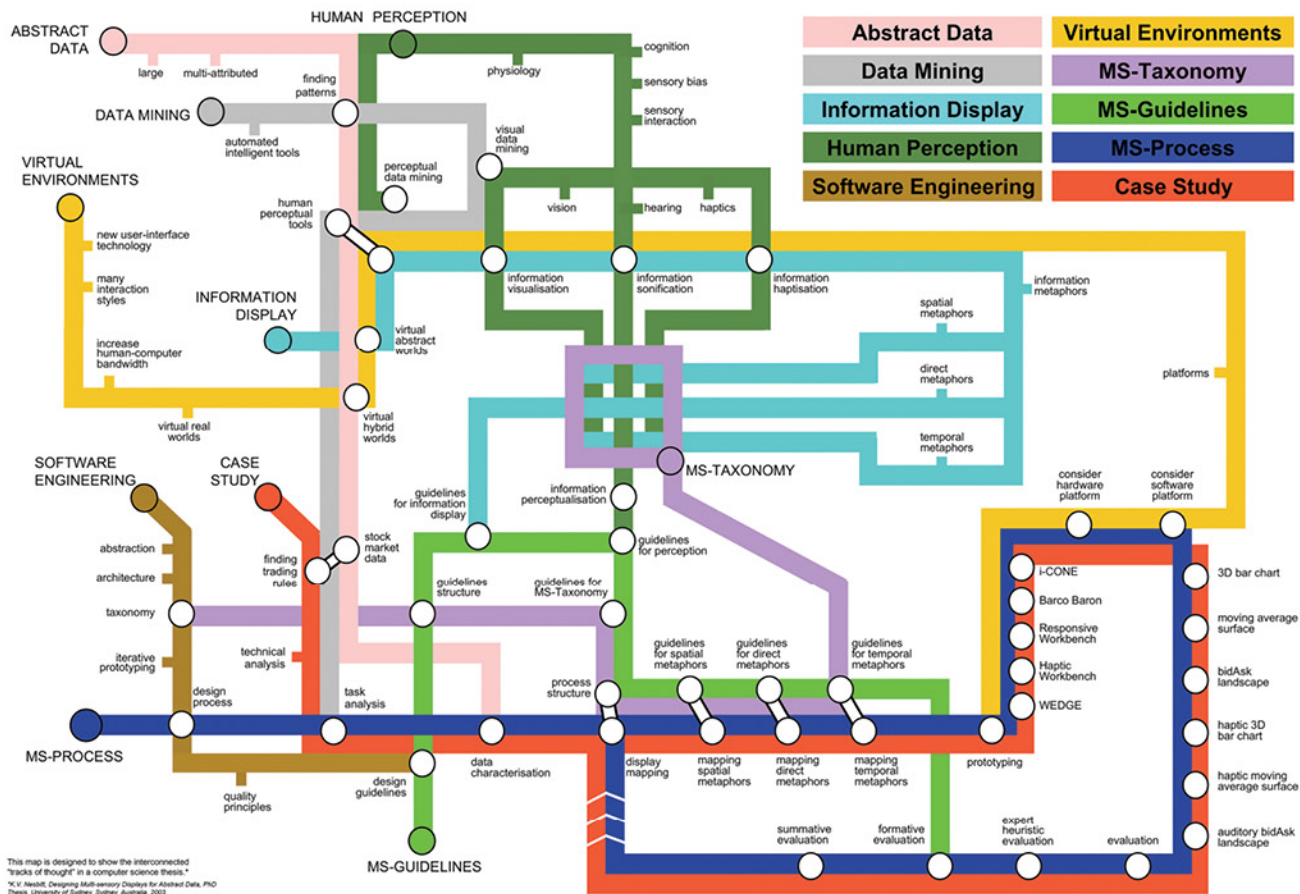
Benjamin
Wiederkehr

The process of selecting the exhibit's pieces begins each year with a call for maps corresponding to a particular theme or addressing the needs of a particular audience. Once the submissions have been gathered, a team of international reviewers and exhibit advisors select the ten most articulate and innovative maps for entry into *Places & Spaces*.

The *Places & Spaces* Exhibit Ambassadors

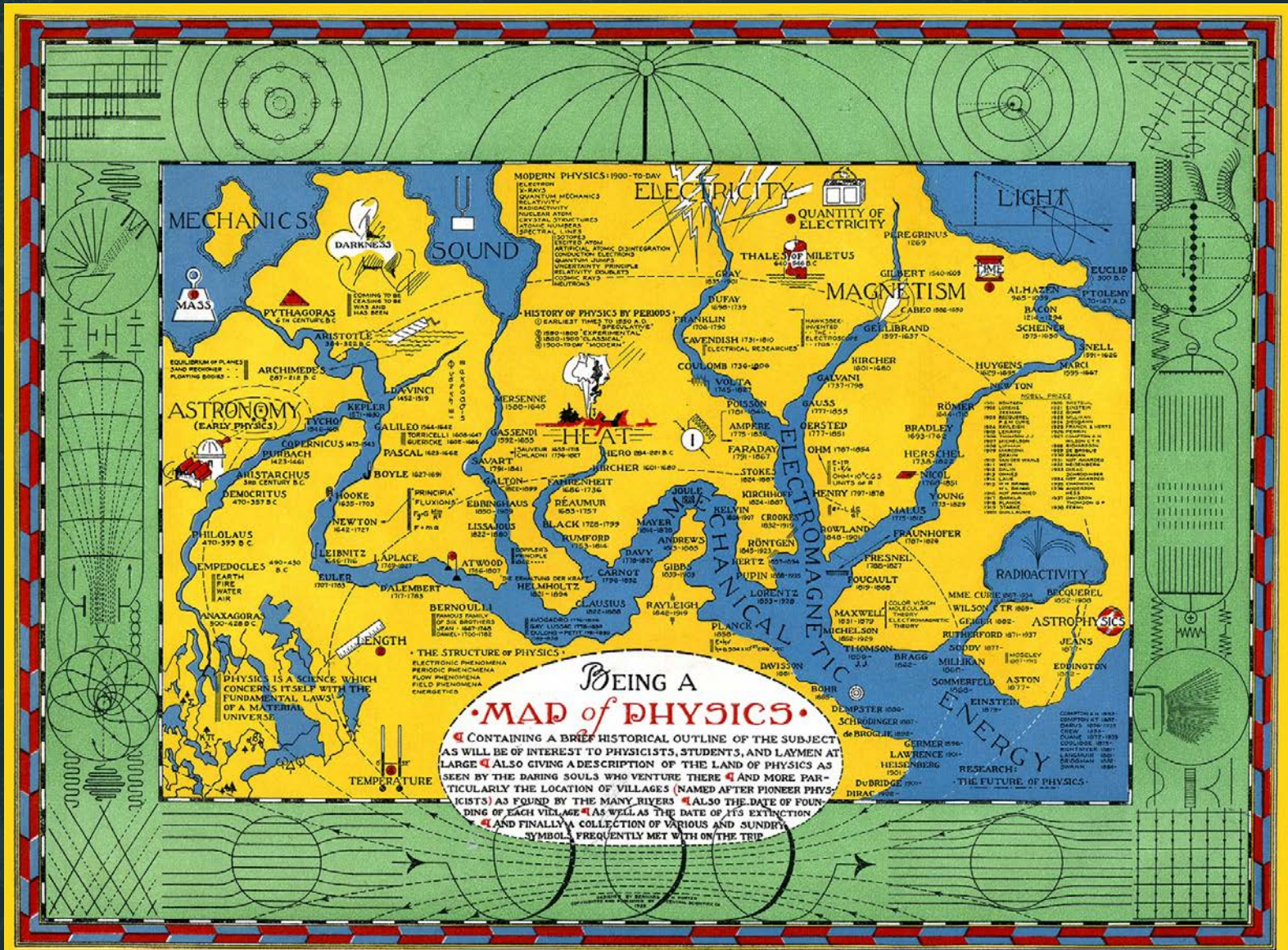


These men and women from around the globe work selflessly to make the exhibit a success. Their intellectual guidance and commitment to promoting science mapping are what has made *Places & Spaces* the vital exhibit it is today.



This map is designed to show the interconnected "tracks of thought" in a computer science thesis.
 ©V. Nesbitt, Designing Multi-sensory Displays for Abstract Data, PhD Thesis, University of Sydney, Sydney, Australia, 2003.

Ph.D. Thesis Map, by Keith V. Nesbitt

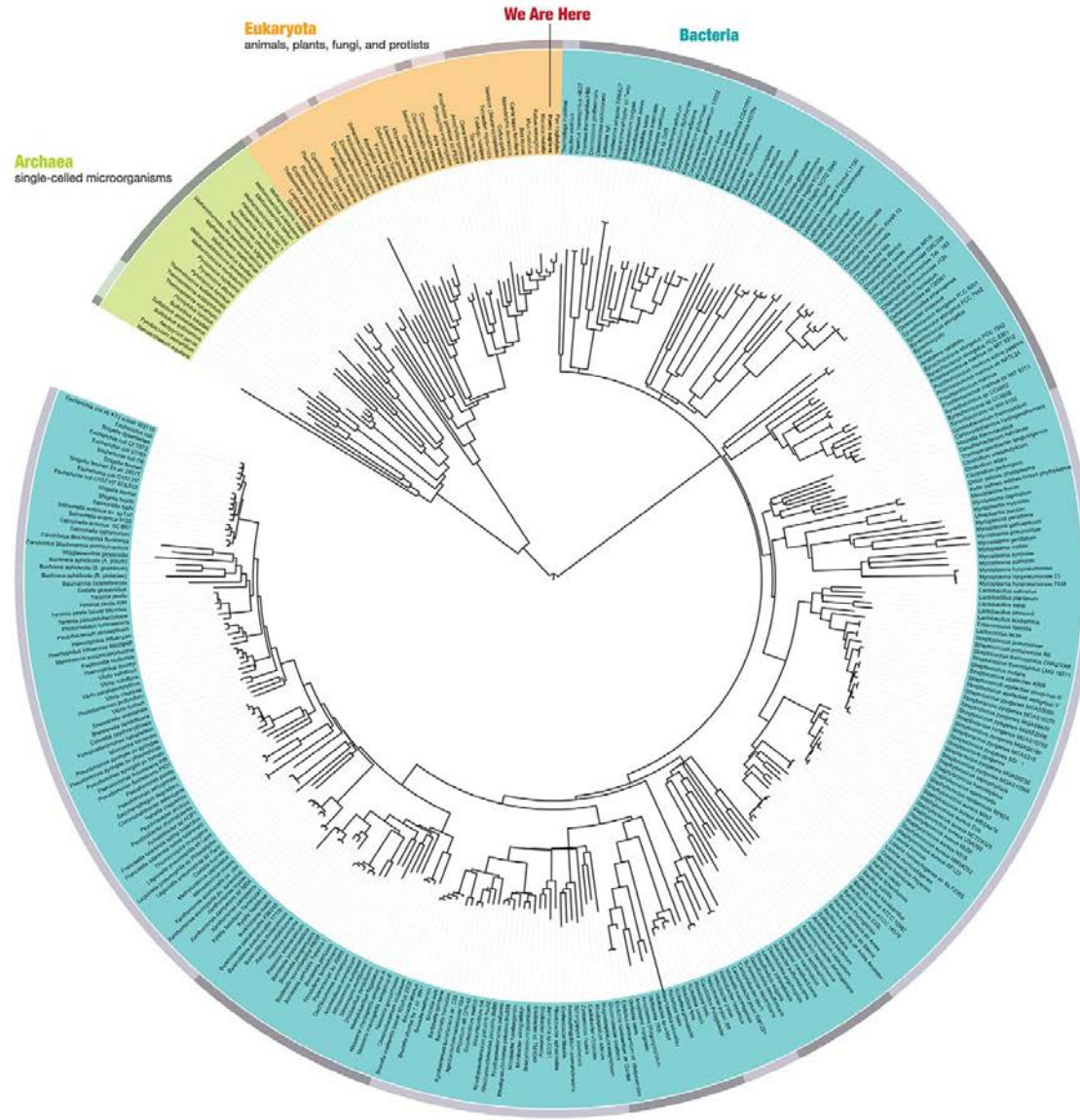


Being a Map of Physics, by Bernard Porter

“When you’re lost in information,
an information map is kind of
useful.”

-David McCandless,
data journalist and information designer

Tree of Life



Tree of Life, by Peer Bork, Francesca Ciccarelli, Berend Snel, Chris Creevey, Christian Von Mering

MAPS OF SCIENCE

A visualization of 7.2 million scholarly documents
appearing in over 16,000 journals, proceedings or symposia
between Jan, 2001 and Dec, 2005

Forecasting Large Trends in Science

Calculations were performed using the large colored groupings of disciplines (fields) to determine if any of them were likely to cause large-scale changes in the structure of science over time. Connectedness coefficients between fields were calculated for each individual year, 2001-2005. A simple regression analysis was conducted to see if there were significant changes in these connectedness coefficients from year-to-year.

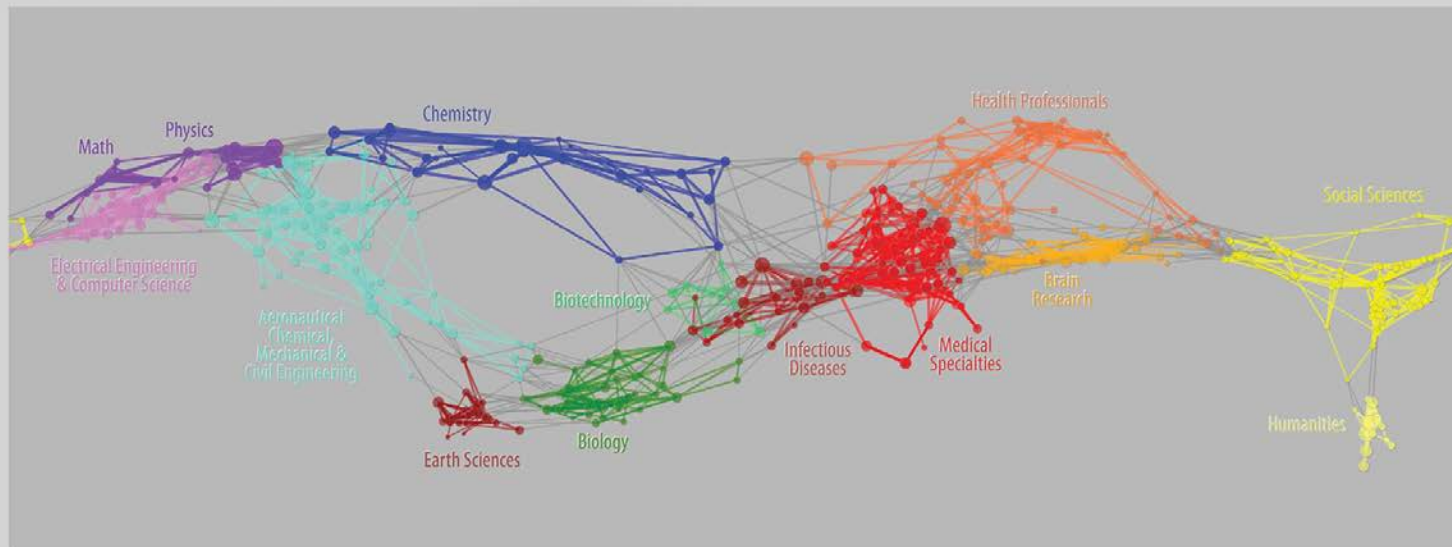
If the structure of science shows below is moving toward stability, we would expect connections between neighboring fields to increase, and connectedness between distant fields to decrease. We found the opposite, suggesting that the underlying structure is unstable and likely to change dramatically over the next decade.

So, again, representing how the structure is likely to change, are provided below. Maps with white arrows represent instances of distant fields that are likely to be pulled closer to each other in the future. Maps with dark arrows represent fields that are currently close-but that are likely to become more dispersed. We expect that future maps of science will show changes in structure corresponding to these observations. Medicine will disperse slightly, while the physical sciences will tighten and draw closer to the medical fields.

This map of science was constructed by sorting more than 16,000 journals into disciplines. Disciplines, represented as circles, are sets of journals that cite a common literature. Links (the lines between disciplines) are pairs of disciplines that share a common literature. A three-dimensional model was used to determine the position of each discipline on the surface of a sphere based on the linkages between disciplines. The model treats links like rubber bands attempting to bring two disciplines close to each other. Pairs of disciplines without links tend to end up on different sides of the map.

The spherical map, which is not shown here, was unrolled in a Mercator projection (on the same one used to show the continents of the earth on a two-dimensional map) to give the large map shown below. This projection allows inspection of the entire map of science at once. Note that the disciplines tend to group along the middle of the map - if this were a map of the earth it would be like a single continent extending along the equator. There are two disciplines at the top (north pole) or the bottom (south pole). Mercator projections also introduce distortions. We tend to forget that the left side is connected to the right side, and assume that the middle is most important. In this map, the social sciences (yellow) on the right connect with the computer sciences (pink) on the left in one continuous swirl.

The six map projections shown at the bottom are images of what one would see if looking directly down at the south pole of the map, at six different rotations. When viewed this way, the map looks like a wheel with an inner ring and outer ring. This wheel of science corresponds very closely with the two-dimensional maps we have previously produced.



Electrical Engineering & Computer Science (EECS), indicated by the pink shape above, has the largest overall increase in connectedness with other fields (19%). It has relatively few connections with the EECS, Math & Physics, and Social Sciences fields, but these three connections had the largest fractional increase. The connection with EECS, which had the single largest growth rate (19%) of any connection, reflects recent growth in the area of bioinformatics.



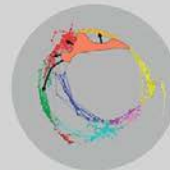
Biotechnology, indicated by the light green shape above, has the largest overall increase in connectedness with other fields (16%). It has relatively few connections with the EECS, Math & Physics, and Social Sciences fields, but these three connections had the largest fractional increase. The connection with EECS, which had the single largest growth rate (19%) of any connection, reflects recent growth in the area of bioinformatics.



Infectious Diseases, indicated by the dark red shape above, has an overall decrease in connectedness (2%) with other fields. Decreases in connection strength between this field and the fields of Biology, Medical Specialties, Health Professionals and Brain Research (all > 5%) are shown as black arrows, and will drive a more dispersion of the medical fields compared to the current structure.



Medical Specialties, indicated by the red shape above, has an overall decrease in connectedness (2%) with other fields. This is dominated by decreasing connection strength to the other medical fields and biology, as shown by the black arrows. The only connection increasing in strength is the one to EECS, which is not shown here, but is shown as a white arrow in the first story.



The **Health Professionals** field, indicated by the orange shape above, has the largest overall decrease in connectedness (9%) to other fields. As with the other medical fields, its connection strength with medicine and biology is decreasing as all cases, as shown by the black arrows. With the decreasing connection strength throughout medicine, we expect the map structure in these areas to relax slightly over time.



The **Social Sciences**, indicated by the yellow shape above, has an overall increase in connectedness (9%) with other fields. Although its greatest connectedness gains were with EECS and Biotechnology (see white arrows), it also had consistent connection increases with nearly all the other fields. In general the fields of EECS, Biotechnology, and the Social Sciences are becoming more connected, and are pulling on the physical sciences as well.

Source: University of California, San Diego Knowledge Mapping Laboratory - Gene Hahnig, © Program of the University of California. The underlying data came from two sources: Thomson ISI and Scopus. Mapping methodology and description by Dick Klavans, President, SciTech Strategies, Inc., and Kevin Boyack, Santa Monica Laboratories. Graphics & typography by Brian Hoffman and Mike Pate. Special acknowledgments to Gary Dorso, Art Ellis, W. Bradford Falck, Lee Silver, and Henry Small. © 2007 by Dick Klavans, all rights reserved.

Maps of Science, by Kevin W. Boyack and Richard Klavans

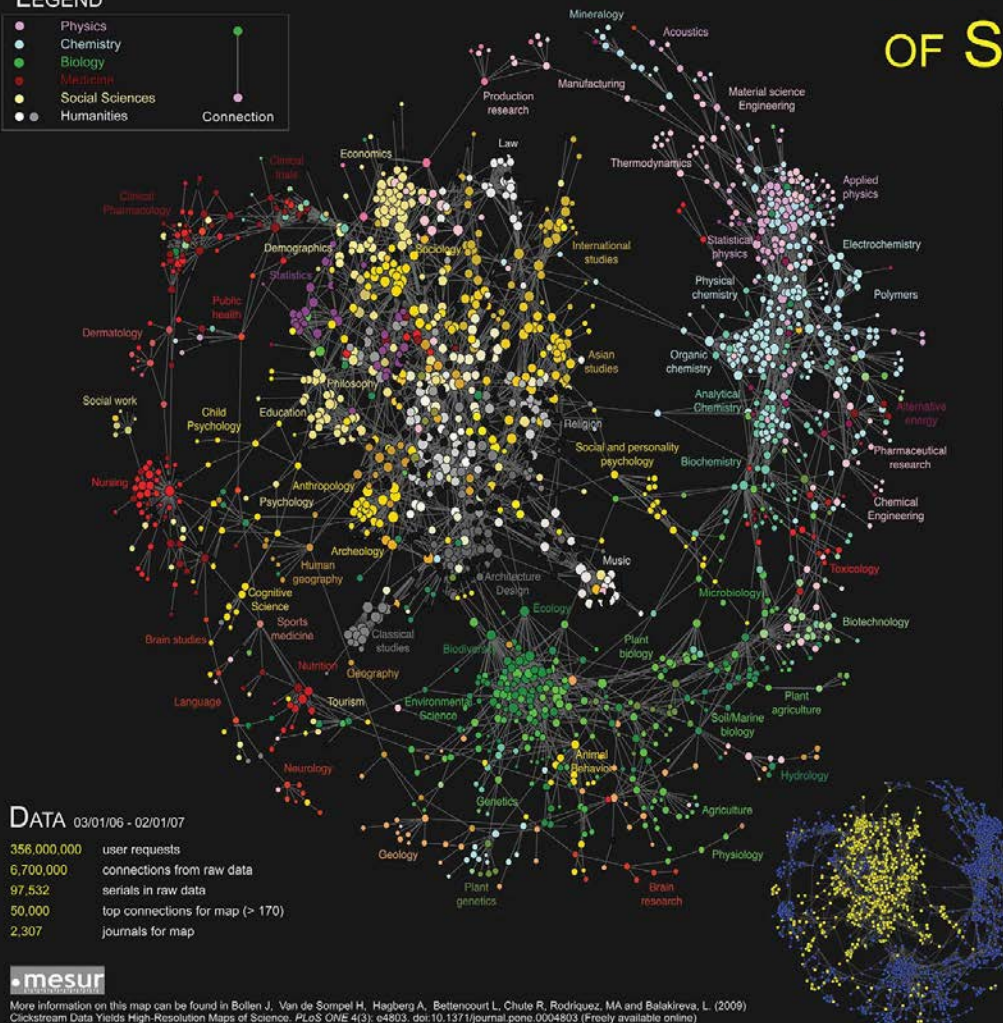
Map of Scientific Collaborations from 2005-2009



Scientific Collaborations between World Cities, by Olivier H. Beauchesne

CLICKSTREAM MAP OF SCIENCE

LEGEND



This is the first map created from large-scale, world-wide, scholarly usage data. It visualizes the collective flow of scientists' movements from one journal to another in their online navigation behavior.

The MESUR project (www.mesur.org) collected a database of nearly 1 billion user requests recorded by the web portals of some of the world's most significant publishers, aggregators and large university consortia, among them Thomson Scientific (Web of Science), Elsevier (Scopus), JSTOR, Ingenta, University of Texas (8 campuses, 6 health institutions), and California State University (23 campuses). All usage logs acquired by the MESUR project contain session identifiers that identify the individual clickstreams of individual scientists navigating from one article to the next.

Pairs of journals are connected when they have a high probability of being followed by each other in users' clickstreams. The circles represent individual journals. A line between two circles indicates that they are strongly connected in either direction. The colors indicate the scientific domain a journal belongs to according to their Dewey Decimal and JCR classification codes that were mapped into the Getty Research Center's Arts and Architecture Taxonomy (AAT) to allow classifications at various levels of detail. The size of circles corresponds to the strength (degree centrality) of a journal's connections in the map. The map is arranged by the Fruchterman-Reingold algorithm that treats connections like springs: connected journals are drawn together, but they are not allowed to get too close.

This map is derived from usage data and therefore also reflects the actions of those who read the literature but rarely publish themselves, e.g., practitioners and laypersons. As a result practitioner-driven domains such as nursing, social work, and tourism studies are prominently featured. The natural sciences vs. the social sciences and humanities emerge as two distinct clusters that are connected via various specific interdisciplinary spokes. Most domains are highly interdisciplinary, but this is more so the case for the social sciences and humanities. Surprisingly, mathematics and computer science are not represented as one specific cluster, but spread out through the map.

Like citation maps, this map is based upon a particular sample of the scientific community, albeit one that includes non-publishing scientists and practitioners and a much greater sample of publications. From MESUR's database of 1 billion user events, we created a matrix of 6 million connections between approximately 100,000 serials. From that matrix we selected only 50,000 connections with the highest number of observations, ranging from approximately 40,000 to 170 observations. This subset of connections pertained to the 2,307 most used journals. This procedure may introduce specific biases which require investigation. The map should therefore not be construed as a final map of scientific activity, but as a showcase for the feasibility of tracking scientific activity from usage data. We hope this methodology will provide unique insights into the real-time structure of scientific activity as it can be observed from scholarly clickstream data.

When we cut the AAT taxonomy at the top level, only two distinctions remain: natural sciences (blue nodes) vs. the social sciences and humanities (yellow nodes). Some journals along the spokes of the wheel have classifications (colors) that do not correspond to their location in the map. This indicates either the journal in question is highly interdisciplinary and/or has been assigned a classification that does not correspond to how scientists actually use the particular journal.

DATA 03/01/06 - 02/01/07

356,000,000	user requests
6,700,000	connections from raw data
97,532	serials in raw data
50,000	top connections for map (> 170)
2,307	journals for map



More information on this map can be found in Bollen J, Van de Sompel H, Hagberg A, Bettencourt L, Chute R, Rodriguez, MA and Balakireva, L. (2009) Clickstream Data Yields High-Resolution Maps of Science. *PLoS ONE* 4(3): e4803. doi:10.1371/journal.pone.0004803 (Freely available online)

Design layout by: Jeremy D. Chacon

Clickstream Map of Science, by Johan Bollen, Herbert Van de Sompel, Aric Hagberg, Luis M.A. Bettencourt, Ryan Chute, Marko A. Rodriguez, and Ludmila Balakireva



Mapping a Stream of Data

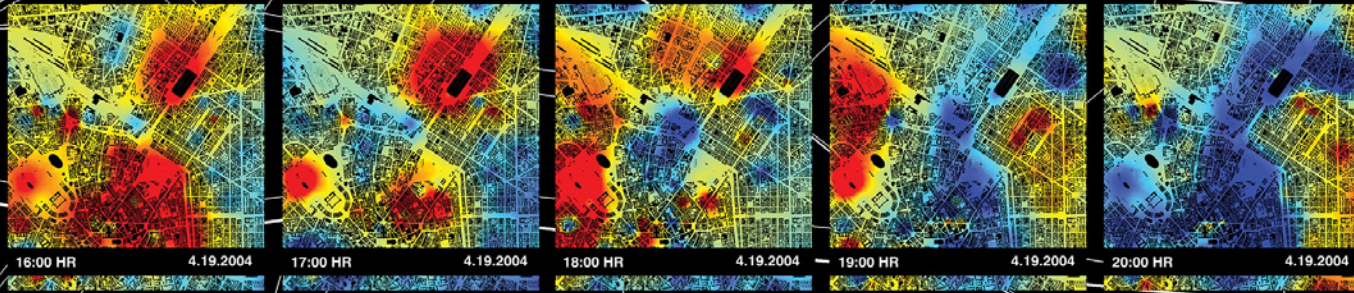
MOBILE LANDSCAPES Using Location Data from Cell Phones for Urban Analysis

This research shows how geo-referencing cell phone activity data can allow for hourly estimates of population flows within urban environments. Analyzing the movement of people on this scale had never been investigated in depth before this project and the results allow researchers to investigate how people navigate and use urban systems. Understanding these flows will allow us to plan better cities.

The maps below show cell phone activity around Milan, Italy's train station during rush hour. As one might imagine at rush hour's peak there is high activity around the train station, as time passes this activity moves further away from this transit node. Milan's urban population is beginning to inhabit completely different parts of the city. While this analysis highlights what one might expect, a closer inspection of the data not only shows high volumes of people at the train station during rush hour, but also the smaller urban plazas that are activated at dusk. The contrast between day and night helps to illustrate how Milan's population uses its urban environment and what parts of the city are important to their daily flow. It also illustrates the potential of cell phone data to tell us about the pulse of the city.

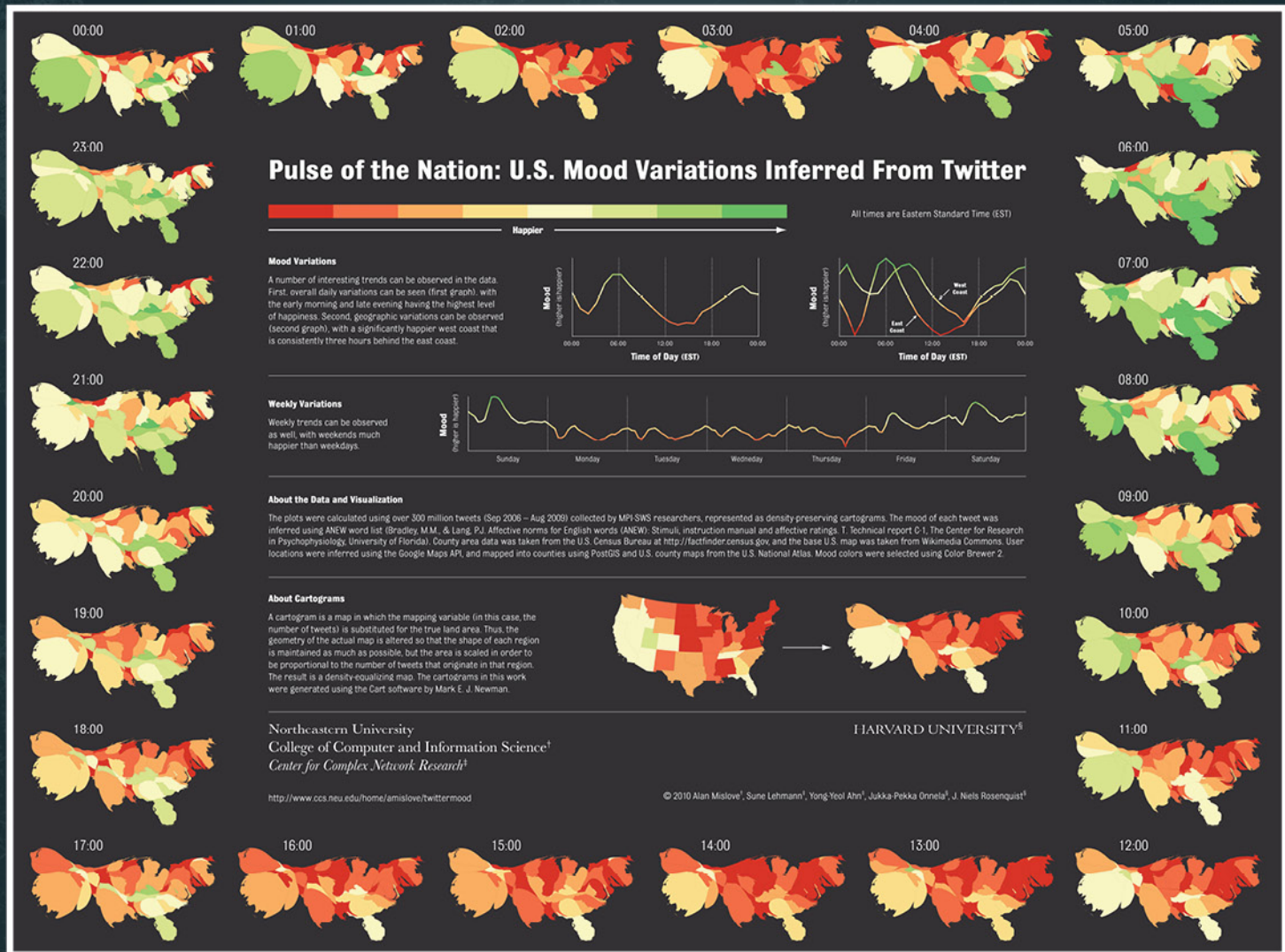


MILAN TRAIN STATION

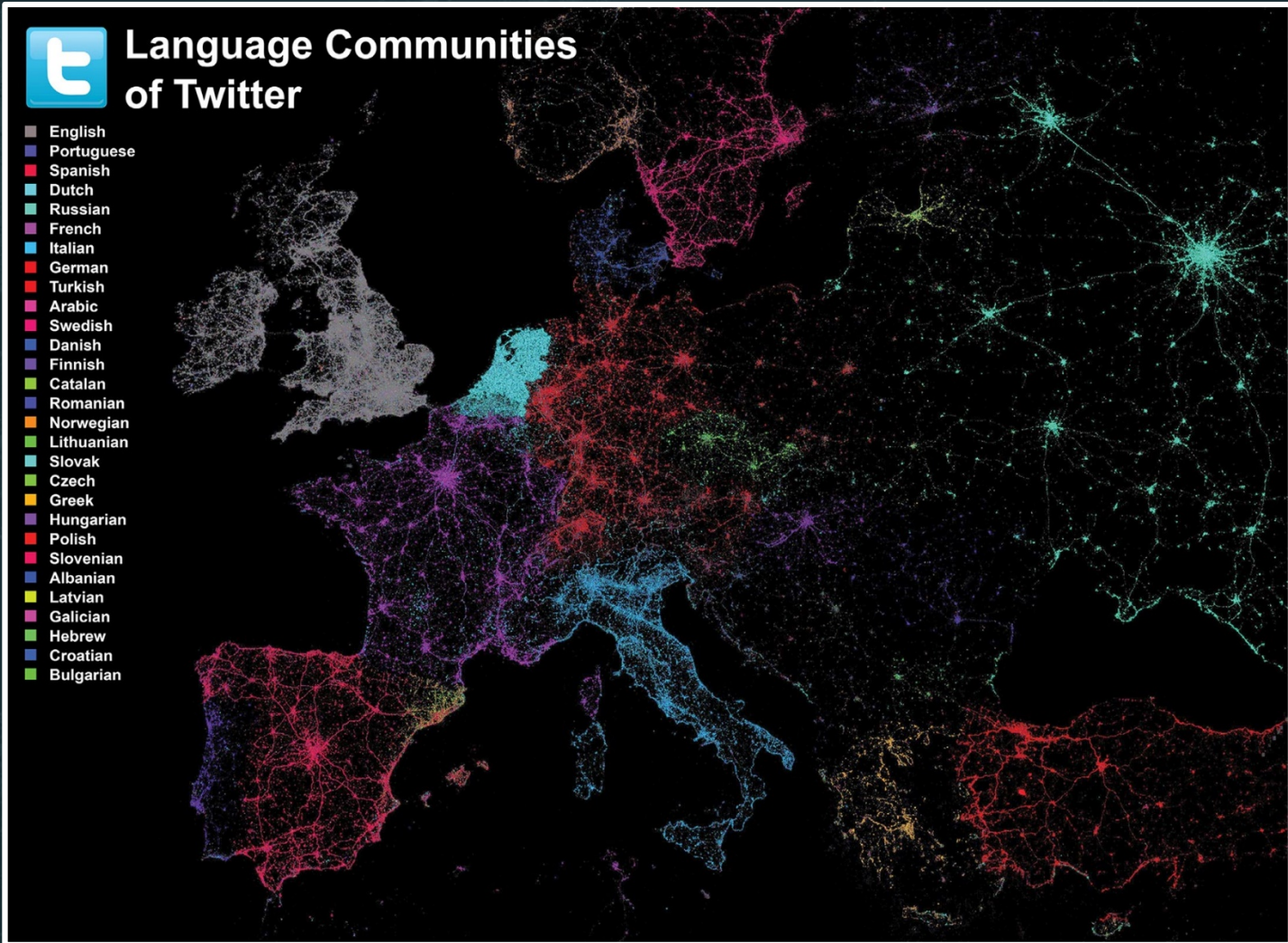


Mobile Landscapes, by Carlo Ratti, Riccardo Maria Pulselli, Sarah Williams

Could you convey the collective “mood” of Twitter users with just text? Maybe—but it would take pages and pages to convey the same insights this map does in seconds.



Pulse of the Nation, by Alan Mislove, Sune Lehmann, Yong-Yeol Ahn, Jukka-Pekka Onnela, and James Niels Rosenquist



Language Communities of Twitter, by Eric Fischer

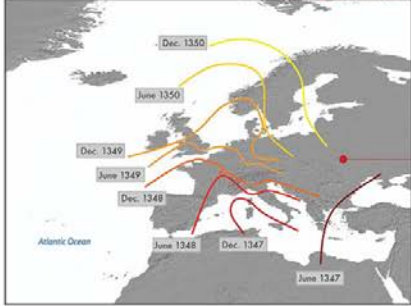


<https://spreecommerce.com/blog/data-driven-decisions>

Maps as Tools for Decision-making

Impact of Air Travel on Global Spread of Infectious Diseases

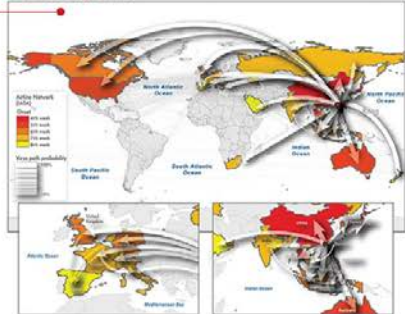
14th Century: Black Death



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.

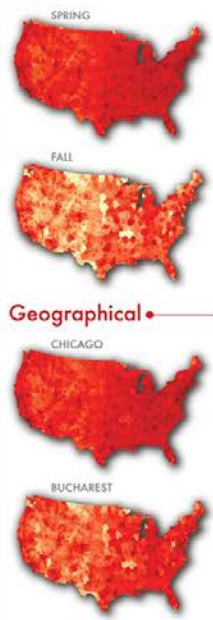
21st Century: SARS



The SARS outbreak on the other hand was characterized by a patched 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 by 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.

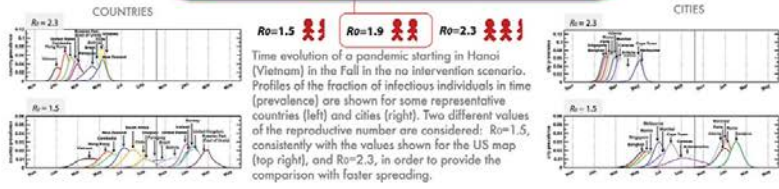
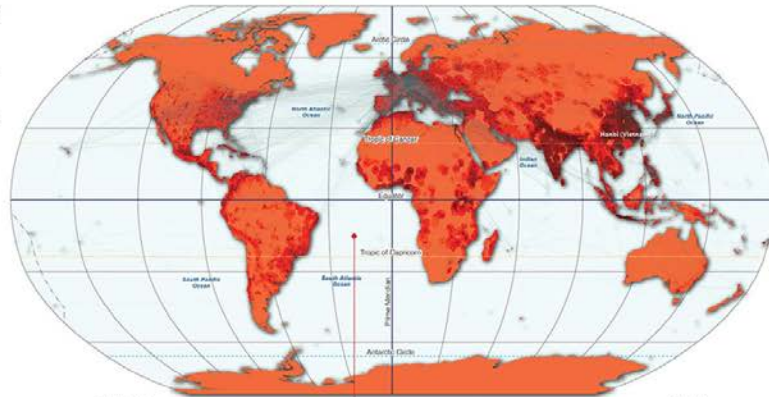
Geographical

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 $R_0=1.9$ 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 $E=17,182$ 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 R_0 , representing the average number of infections generated by a sick person in a fully susceptible population.

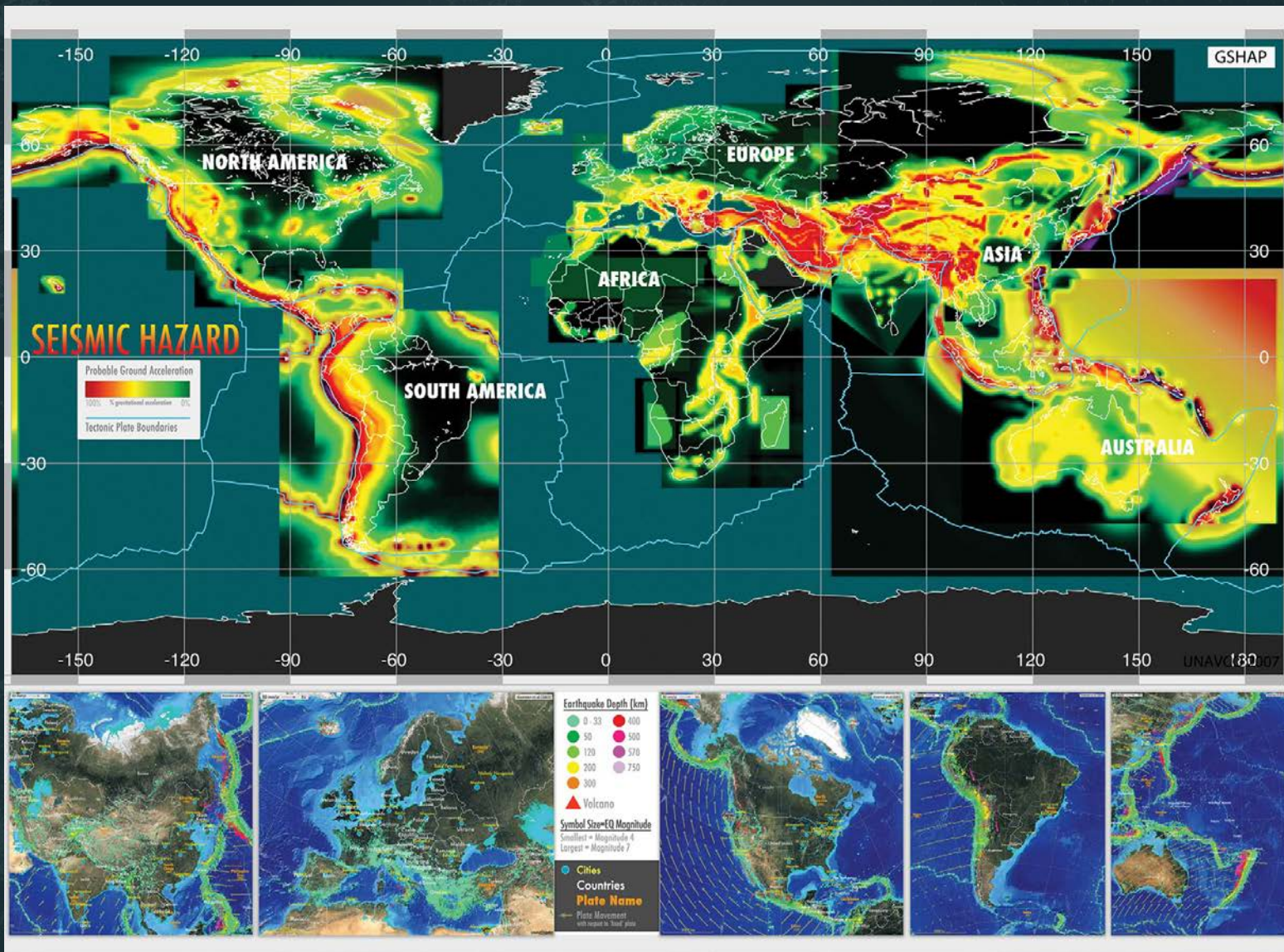
Intervention strategies modeling the use of antiviral drugs can be considered. Two scenarios are compared: an uncooperative strategy in which countries only use their own stockpiles, and a cooperative intervention which envisions a limited worldwide sharing of the resources.

Reproductive Number (R_0)



Intervention





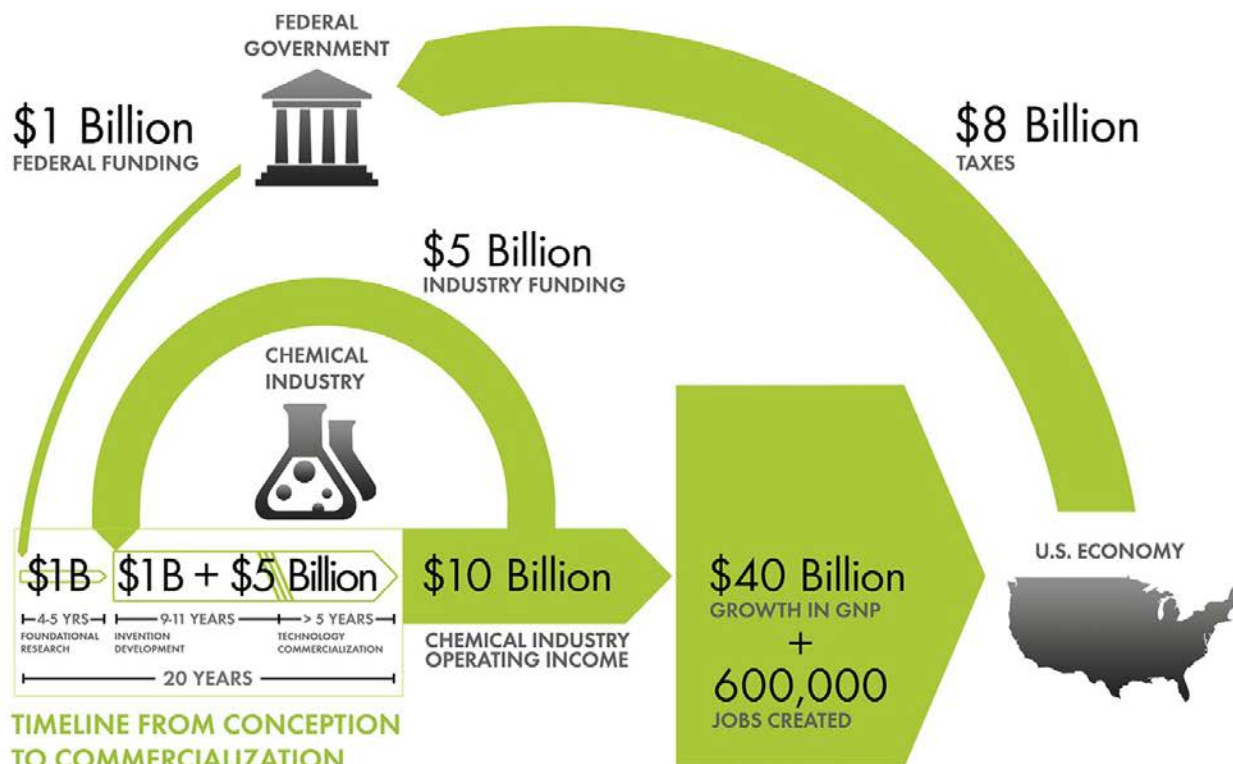
Tectonic Movements and Earthquake Hazard Predictions

by Chuck Meertens, Elisha F. H. Allgood, Michael W. Hamburger, and Lou Estey

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



The Council for Chemical Research (CCR)

has provided the U.S. Congress and government policy makers with important results regarding the impact of Federal Research & Development (R&D) investments on U.S. innovation and global competitiveness through its commissioned 5-year two phase study. To take full advantage of typically brief access to policy makers, CCR developed the graphic below as a communication tool that distills the complex data produced by these studies in direct, concise, and clear terms.

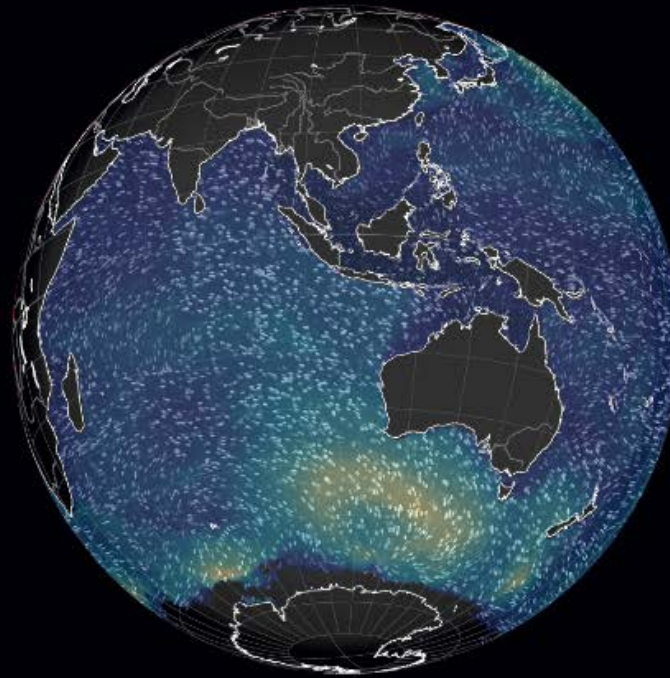


The design shows that an input of \$1B in federal investment, leveraged by \$5B in industry investment, brings new technologies to market and results in \$10B of operating income for the chemical industry, \$40B of growth in the Gross National Product (GNP) and further impacts the US economy by generating 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 in tax base that in turn is available for investment in basic research.

Interactive Maps as Tools for Discovery



Interactive visualization on display at the CDC Museum in Atlanta.
Photo courtesy of Mike Jensen.



earth ≡

Earth

Cameron Beccario

earth.nullschool.net

The News Co-occurrence Globe

an interactive visualization of how countries are mentioned together in the world's news media

2.37K
COOCCURRIN

136
COOCCURROUT

2.37K 136



RUSSIAN FEDERATION

cooccurrences in: 2,374%
cooccurrences out: 136%

Feb 22 Mar 1 Mar 8 Mar 15 Mar 22 Mar 29 Apr 5 Apr 12 Apr 19 Apr 26 May 3 May 10 May 17 May 24

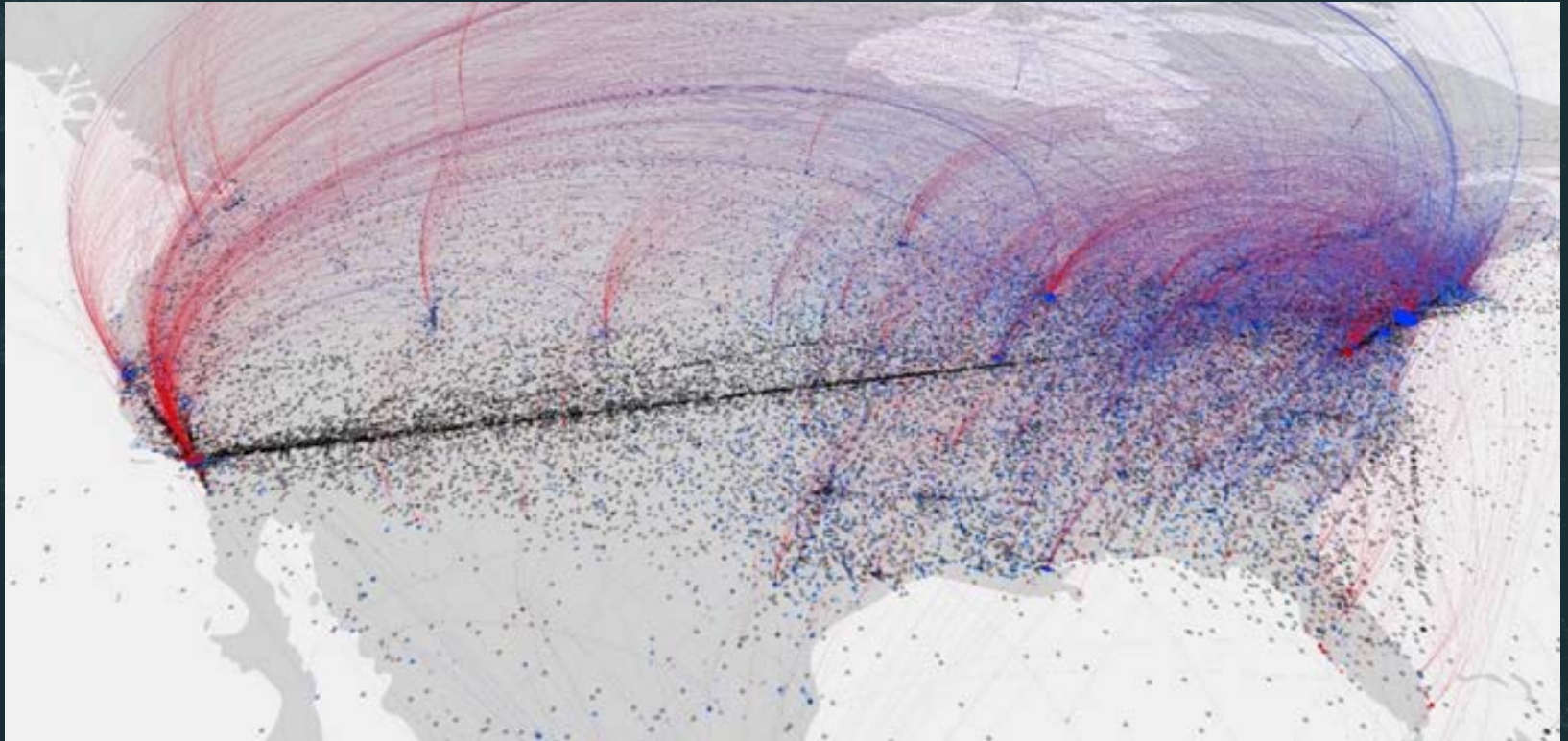
COOCCURRIN

IN%
OUT%

Mapping Global News

Kalev Leetaru

gdeltproject.org



Charting Culture

Maximilian Schich and Mauro Martino

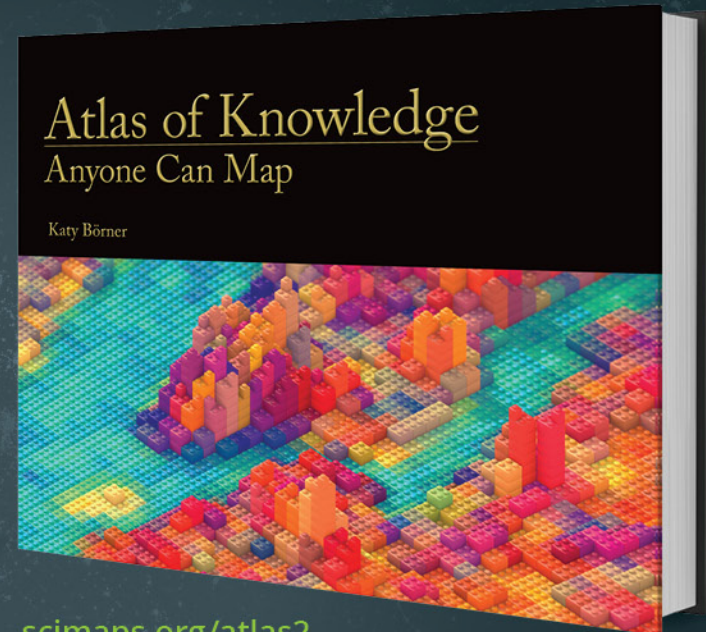
cultsci.net

Enjoy the first two books in Katy Börner's 3-Part *Atlas* series



scimaps.org/atlas1

Atlas of Science, featuring more than thirty full-page science maps, fifty data charts, a timeline of science-mapping milestones, and 500 color images, serves as a sumptuous visual index to the evolution of modern science and as an introduction to “the science of science”—charting the trajectory from scientific concept to published results.



scimaps.org/atlas2

The *Atlas of Knowledge* introduces a theoretical visualization framework meant to empower anyone to systematically render data into insights. It aims to teach “timeless” knowledge that holds true over a lifetime while referring to an extensive set of references for “timely” advice on what tool and workflow is currently the best for answering a specific question.



facebook.com/cnscenter



[@cnscenter](https://twitter.com/cnscenter)

IVMOOC

Information Visualization MOOC

The Information Visualization MOOC provides an overview about the state of the art in information visualization, teaching the process of producing effective visualizations that take the needs of users into account.

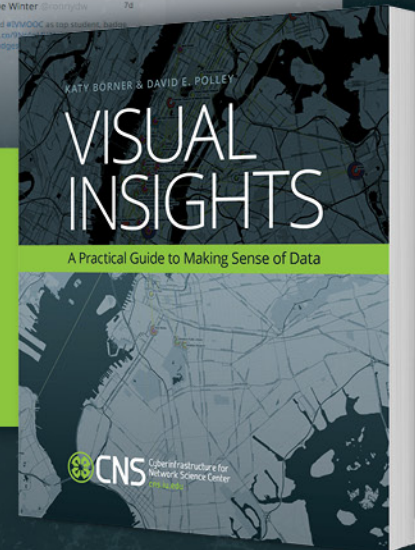
The inaugural IVMOOC, which launched in January 2013, attracted participants from more than 100 countries. It is one of the first MOOCs offered by IU and the first to offer an opportunity for students to work in teams with real clients. All registrants gain free access to the Scholarly Database and the Sci2 Tool.

The course can be taken for three Indiana University credits as part of the Online Data Science Program offered by the School of Informatics and Computing.

The course will return in January 2016. Learn more at ivmooc.cns.iu.edu.

The screenshot shows the Canvas LMS interface for the Information Visualization MOOC 2015. The main content area includes a welcome message, a course schedule, and a list of video resources. The sidebar on the right contains navigation links and a 'Coming Up' section. A social media feed for #ivmooc is visible on the right side of the main content area.

This IVMOOC companion textbook offers a gentle introduction to the design of insightful visualizations. It seamlessly blends theory and practice, giving readers both the theoretical foundation and the practical skills necessary to render data into insights.





Lisel Record
recorde@indiana.edu
www.scimaps.org