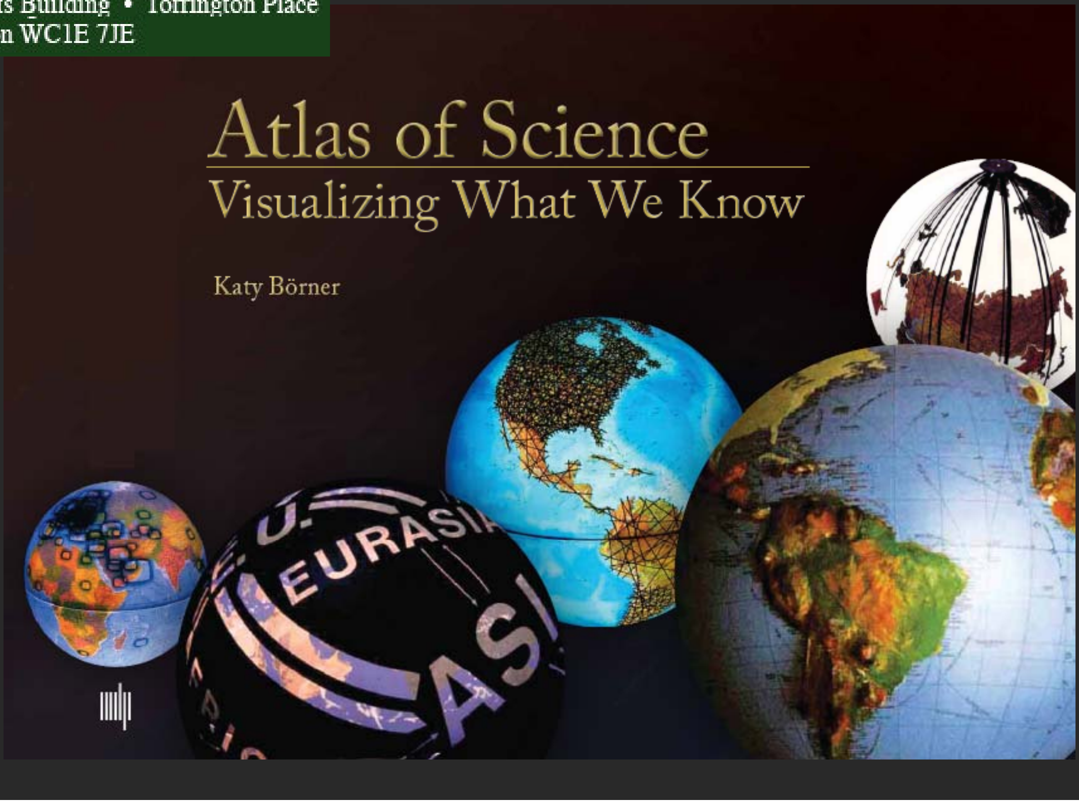
CASA Seminar Series

November 15th, 2010 • 6pm - 8pm
Roberts Building • Torrington Place
London WC1E 7JE



Atlas of Science
Visualizing What We Know

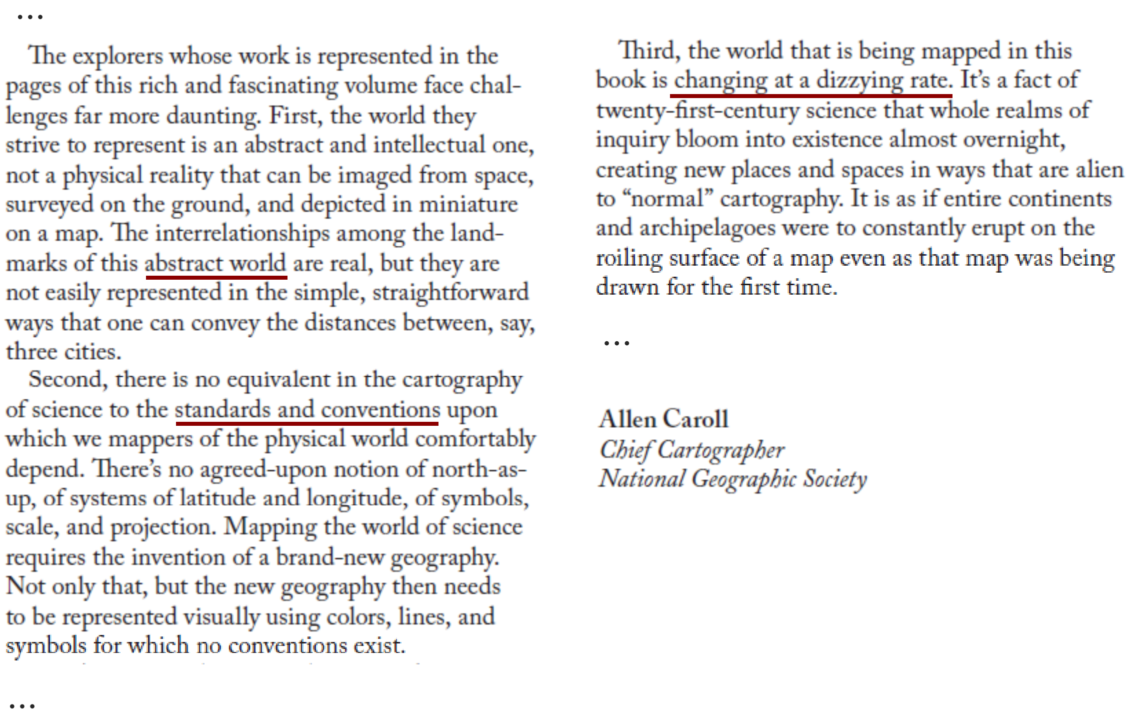
Katy Börner

EARTH
FROM ABOVE
Yann Arthus-Bertrand

Atlas of Science
Visualizing What We Know

Katy Börner

Foreword



…

The explorers whose work is represented in the pages of this rich and fascinating volume face challenges far more daunting. First, the world they strive to represent is an abstract and intellectual one, not a physical reality that can be imaged from space, surveyed on the ground, and depicted in miniature on a map. The interrelationships among the landmarks of this abstract world are real, but they are not easily represented in the simple, straightforward ways that one can convey the distances between, say, three cities.

Second, there is no equivalent in the cartography of science to the standards and conventions upon which we mappers of the physical world comfortably depend. There's no agreed-upon notion of north-as-up, of systems of latitude and longitude, of symbols, scale, and projection. Mapping the world of science requires the invention of a brand-new geography. Not only that, but the new geography then needs to be represented visually using colors, lines, and symbols for which no conventions exist.

…

Third, the world that is being mapped in this book is changing at a dizzying rate. It's a fact of twenty-first-century science that whole realms of inquiry bloom into existence almost overnight, creating new places and spaces in ways that are alien to "normal" cartography. It is as if entire continents and archipelagoes were to constantly erupt on the roiling surface of a map even as that map was being drawn for the first time.

…

Allen Caroll
Chief Cartographer
National Geographic Society

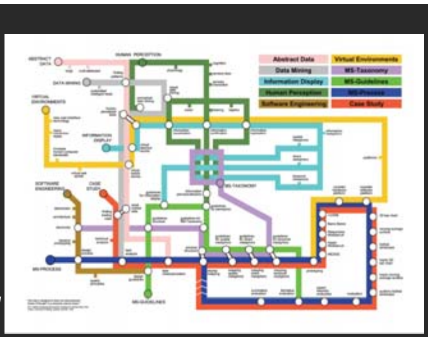
Early Maps of the World VERSUS Early Maps of Science

3D
Physically-based
Accuracy is measurable
Trade-offs have more to do with granularity
2-D projections are very accurate at local levels
Centuries of experience
Geo-maps can be a template for other data

n-D
Abstract space
Accuracy is difficult
Trade-offs indirectly affect accuracy
2-D projections neglect a great deal of data
Decades of experience
Science maps can be a template for other data

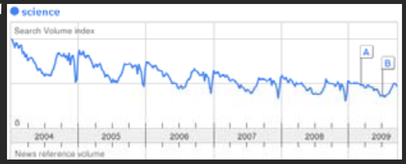
Take terra bytes of data

Black Box



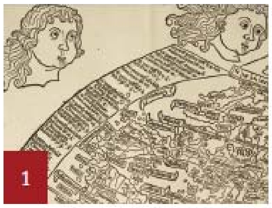
Find your way

Find collaborators, friends

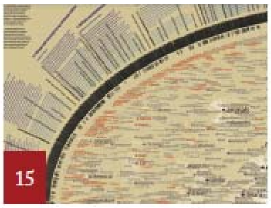


Identify trends

Contents

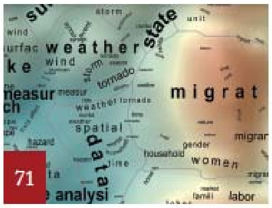


Part 1: Introduction



Part 2: The History of Science Maps

Part 3: Toward a Science of Science



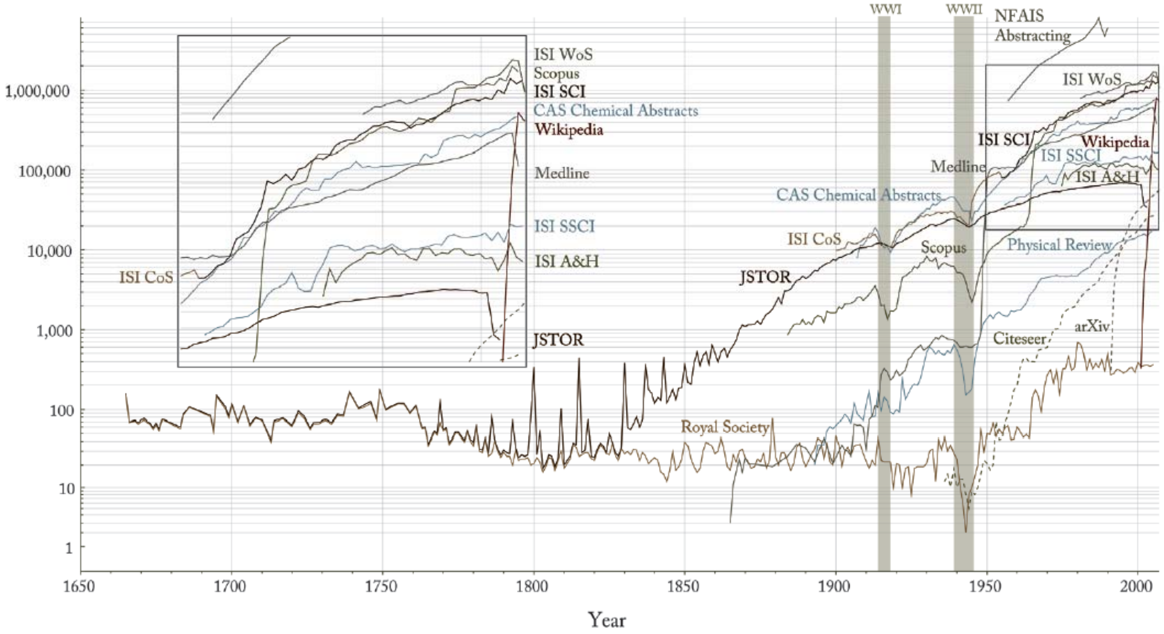
Part 1: Introduction

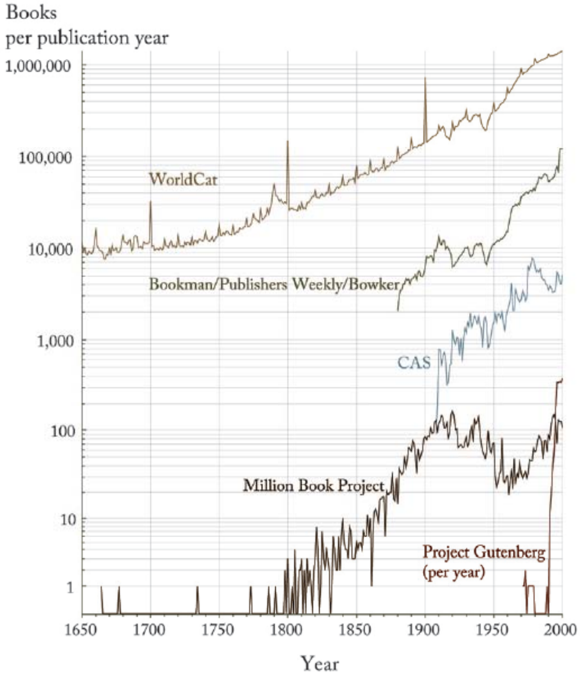
Because of the explosive power of exponential growth, the 21st century will be equivalent to 20,000 years of progress at today's rate of progress. The whole 20th century is equivalent to 20 years of progress at today's rate of progress. Organizations have to be able to redefine themselves at a faster and faster pace.

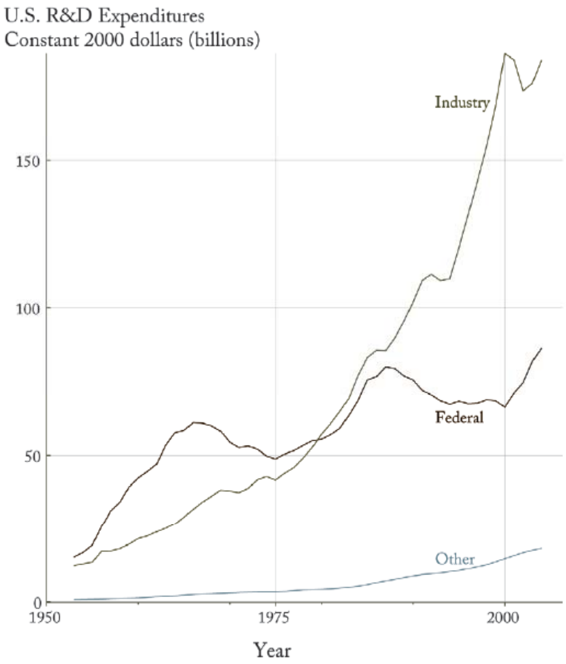
Ray Kurzweil

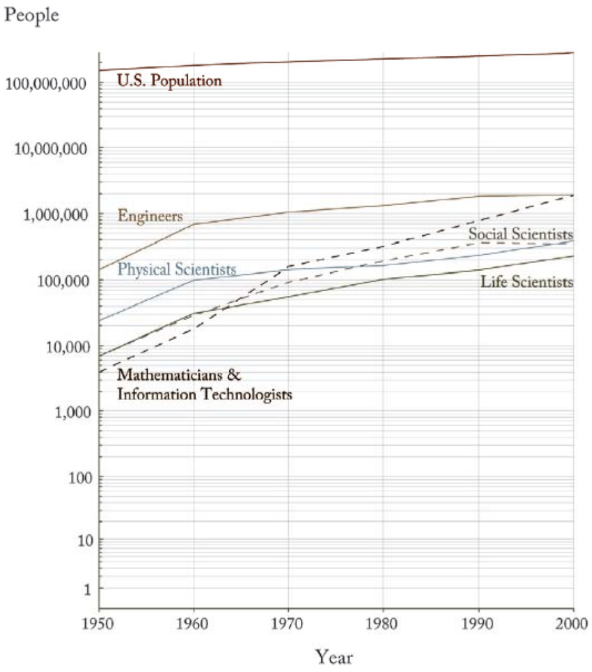
The Rise of Science and Technology

Papers & Wikipedia Entries



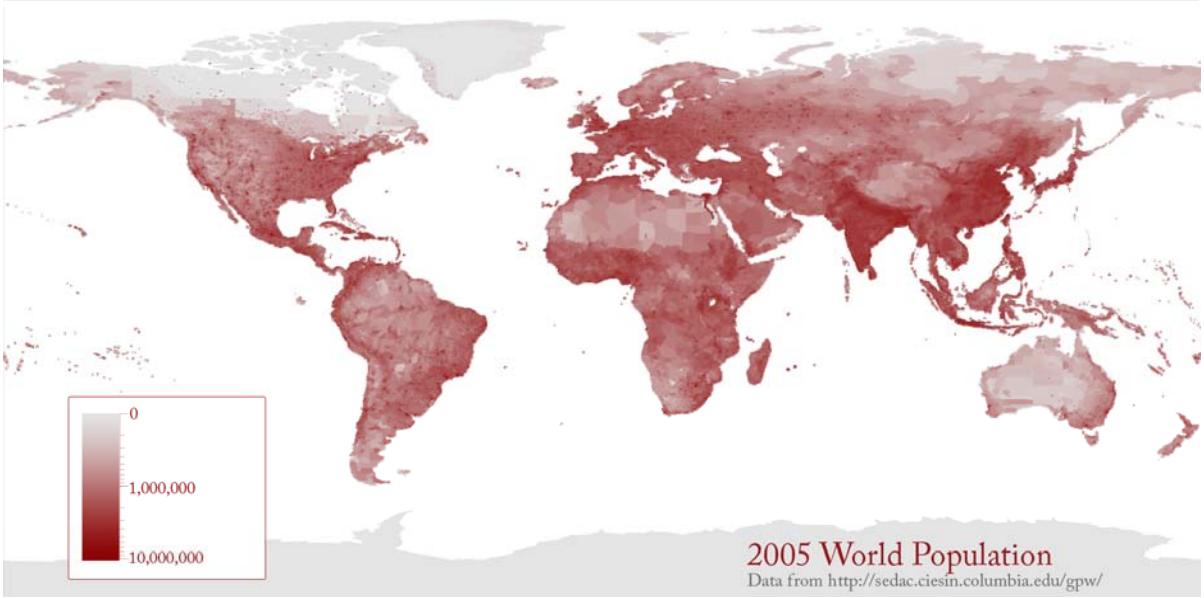






2005 World Population

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 per box using a logarithmic interpolation. The highest density boxes appear in Mumbai, with 11,687,850 people in the quarter degree block, Calcutta (10,816,010), and Shanghai (8,628,088).

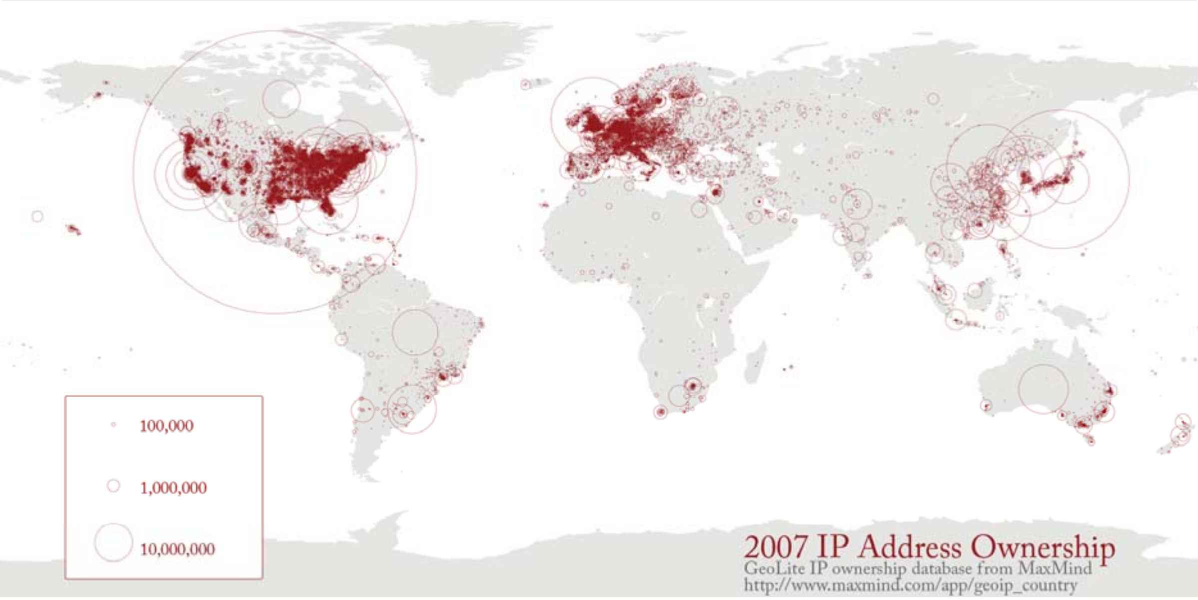


2005 World Population

Data from http://sedac.ciesin.columbia.edu/gpw/

2007 IP Address Ownership

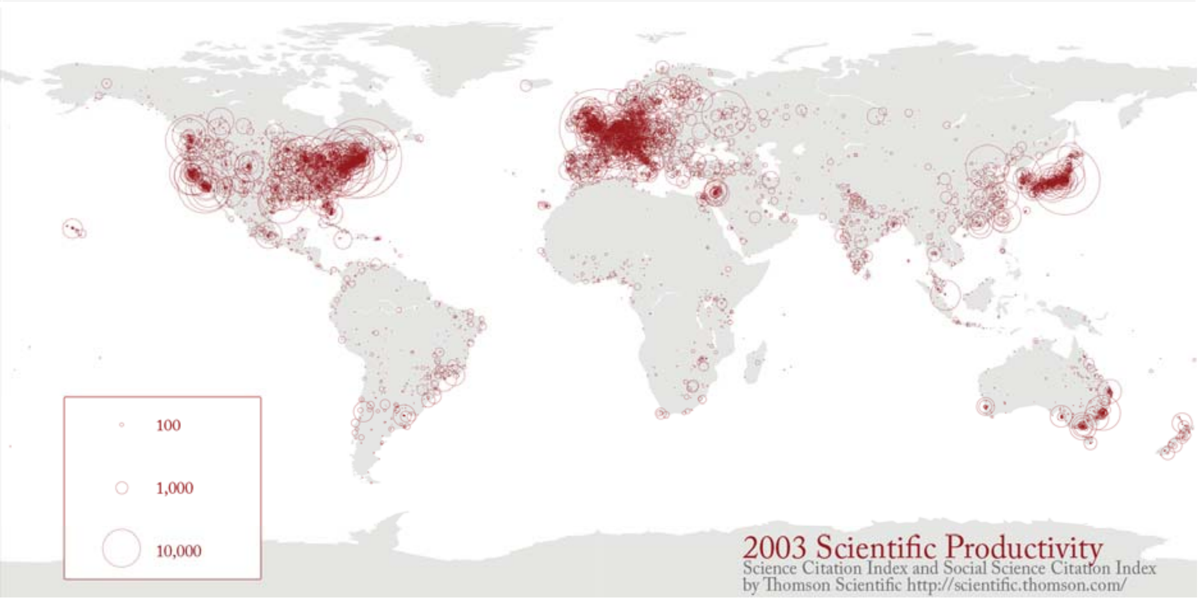
This map shows IP address ownership by location. Each owner is represented by a circle and the area size of the circle corresponds to the number of IP addresses owned. The larges circle denotes MIT's holdings of an entire class A subnet, which equates to 16,581,375 IP addresses. The countries that own the most IP addresses are US (560 million), Japan (130 million), Great Britain (47 million).



2007 IP Address Ownership
GeoLite IP ownership database from MaxMind
http://www.maxmind.com/app/geoip_country

2003 Scientific Productivity

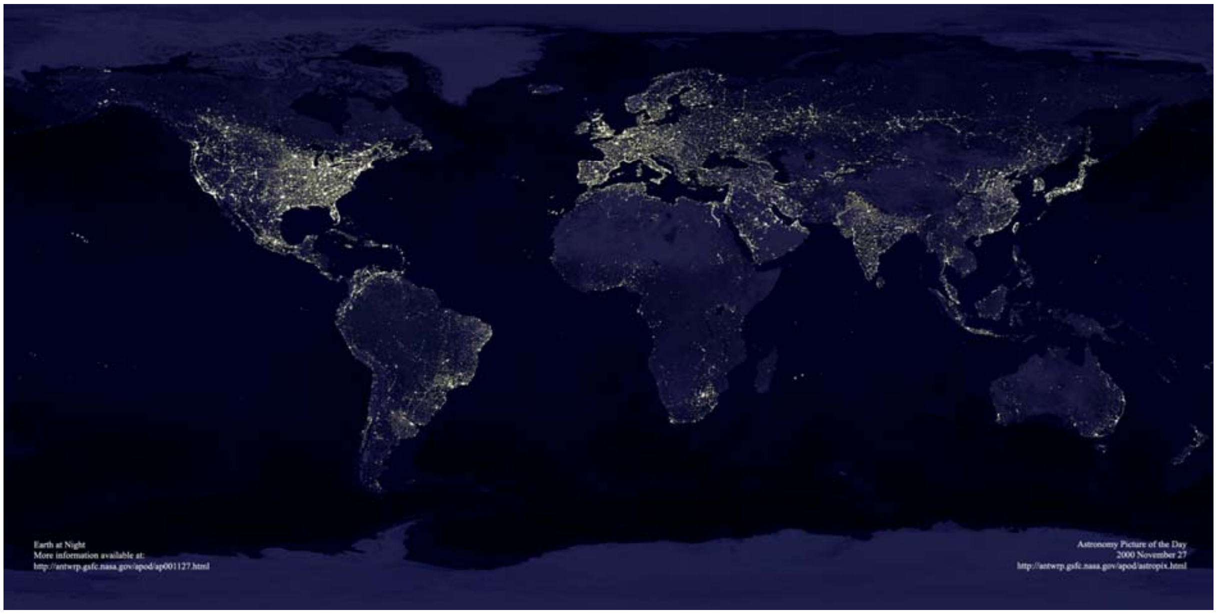
Shown is where science is performed today. Each circle indicates a geographic location at which scholarly papers are published. The larger the circle the more papers are produced. Boston, MA, London, England, and New York, NY are the top three paper production areas. Note the strong resemblance with the Night on Earth and the IP Ownership maps and the striking differences to the world population map.



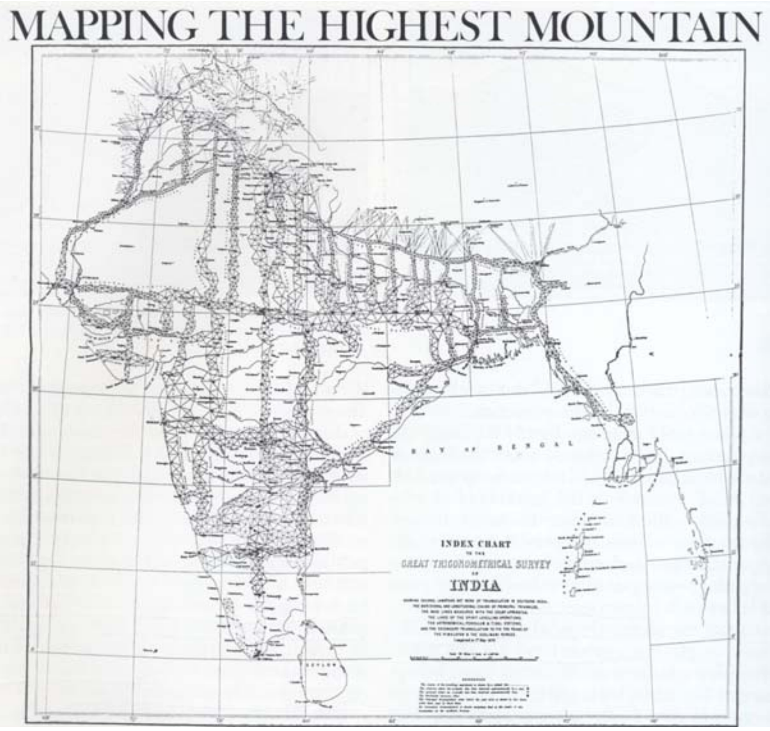
2003 Scientific Productivity
Science Citation Index and Social Science Citation Index
by Thomson Scientific http://scientific.thomson.com/

2000 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.



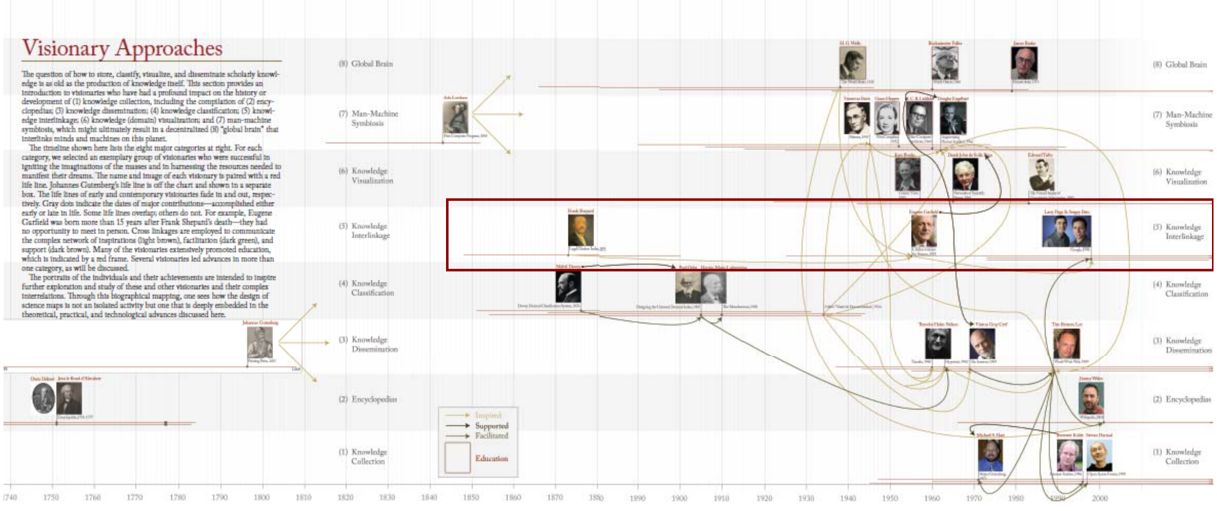
Part 1: Introduction



*In **1870**, Captain George Everest embarked to map India by triangulation. For generations, a vast network of repeating sightline triangles was meticulously measured and recorded (see map below). What resembles a pattern of eyelashes on the northern border represents the sightlines to stations built above treetops. While analyzing the triangles in the calculating offices of Calcutta, the mapmakers discovered the highest peak in the world: Mount Everest*

Part 2: The History of Science Maps

> *Noise becomes data when it has a cognitive pattern. Data becomes information when assembled into a coherent whole, which can be related to other information. Information becomes knowledge when integrated with other information in a form useful for making decisions and determining actions. Knowledge becomes understanding when related to other knowledge in a manner useful in anticipating, judging and acting. Understanding becomes wisdom when informed by purpose, ethics, principles, memory and projection.*
>
> George Santayana

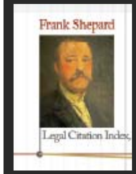


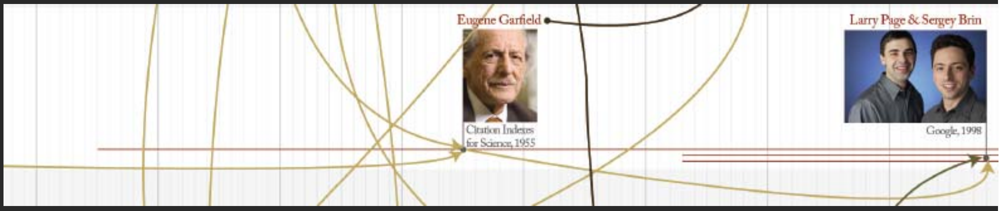
Visionary Approaches

The question of how to store, classify, visualize, and disseminate scholarly knowledge is as old as the production of knowledge itself. This section provides an introduction to visionaries who have had a profound impact on the history or development of (1) knowledge collection, including the compilation of (2) encyclopedias; (3) knowledge dissemination; (4) knowledge classification; (5) knowledge interlinkage; (6) knowledge (domain) visualization; and (7) man-machine symbiosis, which might ultimately result in a decentralized (8) "global brain" that interlinks minds and machines on this planet.

The timeline shown here lists the eight major categories at right. For each category, we selected an exemplary group of visionaries who were successful in igniting the imaginations of the masses and in harnessing the resources needed to manifest their dreams. The name and image of each visionary is paired with a red life line. Johannes Gutenberg's life line is off the chart and shown in a separate box. The life lines of early and contemporary visionaries fade in and out, respectively. Gray dots indicate the dates of major contributions—accomplished either early or late in life. Some life lines overlap; others do not. For example, Eugene Garfield was born more than 15 years after Frank Shepard's death—they had no opportunity to meet in person. Cross linkages are employed to communicate the complex network of inspirations (light brown), facilitation (dark green), and support (dark brown). Many of the visionaries extensively promoted education, which is indicated by a red frame. Several visionaries led advances in more than one category, as will be discussed.

The portraits of the individuals and their achievements are intended to inspire further exploration and study of these and other visionaries and their complex interrelations. Through this biographical mapping, one sees how the design of science maps is not an isolated activity but one that is deeply embedded in the theoretical, practical, and technological advances discussed here.



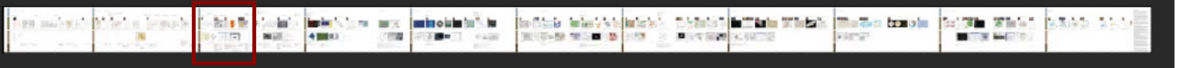


Legal Citation Index, 1873

Citation Indexes for Science, 1955

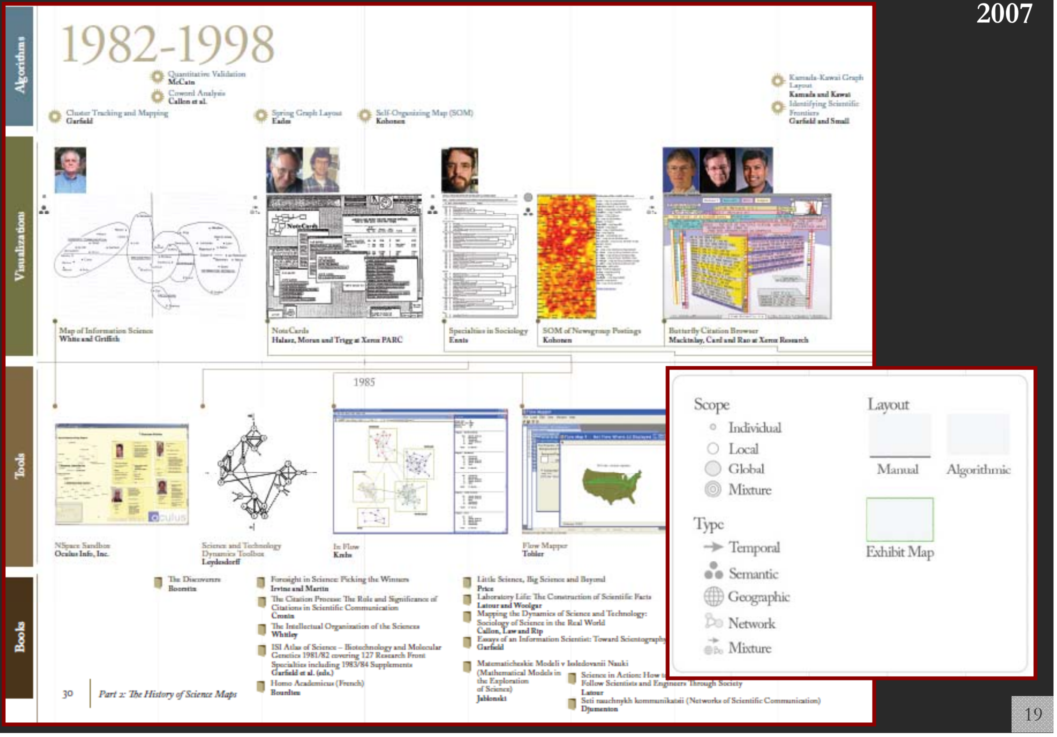
Google, 1998

Milestones in Mapping Science



1934

2007



1982-1998

Algorithms

Quantitative Validation
McCain

Coword Analysis
Callon et al.

Cluster Tracking and Mapping
Garfield

Spring Graph Layout
Eades

Self-Organizing Map (SOM)
Kohonen

Kamada-Kawai Graph Layout
Kamada and Kawai

Identifying Scientific Frontiers
Garfield and Small

Visualizations

Map of Information Science
White and Griffith

NoteCards
Halasz, Moran and Trigg at Xerox PARC

Specialties in Sociology
Ennis

SOM of Newsgroup Postings
Kohonen

Butterfly Citation Browser
Mackinlay, Card and Rao at Xerox Research

Tools

1985

NSpace Sandbox
Oculus Info, Inc.

Science and Technology Dynamics Toolbox
Leydesdorff

In Flow
Krebs

Flow Mapper
Tobler

Books

The Discoverers
Boorstin

Foresight in Science: Picking the Winners
Irvine and Martin

The Citation Process: The Role and Significance of Citations in Scientific Communication
Cronin

The Intellectual Organization of the Sciences
Whitley

ISI Atlas of Science – Biotechnology and Molecular Genetics 1981/82 covering 127 Research Front Specialties including 1983/84 Supplements
Garfield et al. (eds.)

Homo Academicus (French)
Bourdieu

Little Science, Big Science and Beyond
Price

Laboratory Life: The Construction of Scientific Facts
Latour and Woolgar

Mapping the Dynamics of Science and Technology: Sociology of Science in the Real World
Callon, Law and Rip

Essays of an Information Scientist: Toward Scientography
Garfield

Matematicheskie Modeli v Issledovanii Nauki (Mathematical Models in the Exploration of Science)
Jablonski

Science in Action: How t
Follow Scientists and Engineers Through Society
Latour

Seti nauchnykh kommunikatsii (Networks of Scientific Communication)
Djumenton

30 | *Part 2: The History of Science Maps*

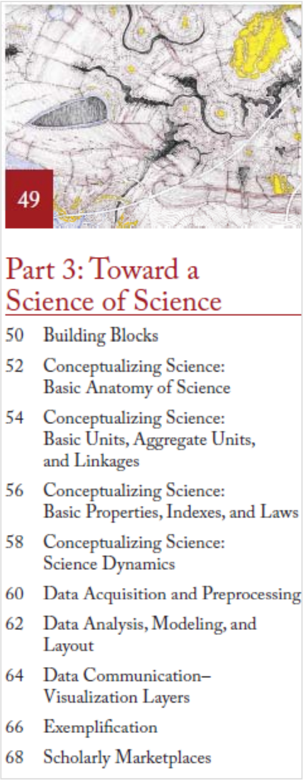
Scope
- ◦ Individual
- ○ Local
- ◎ Global
- ◎ Mixture

Type
- → Temporal
- ● Semantic
- 🌐 Geographic
- ○ Network
- ● Mixture

Layout
- Manual
- Algorithmic
- Exhibit Map

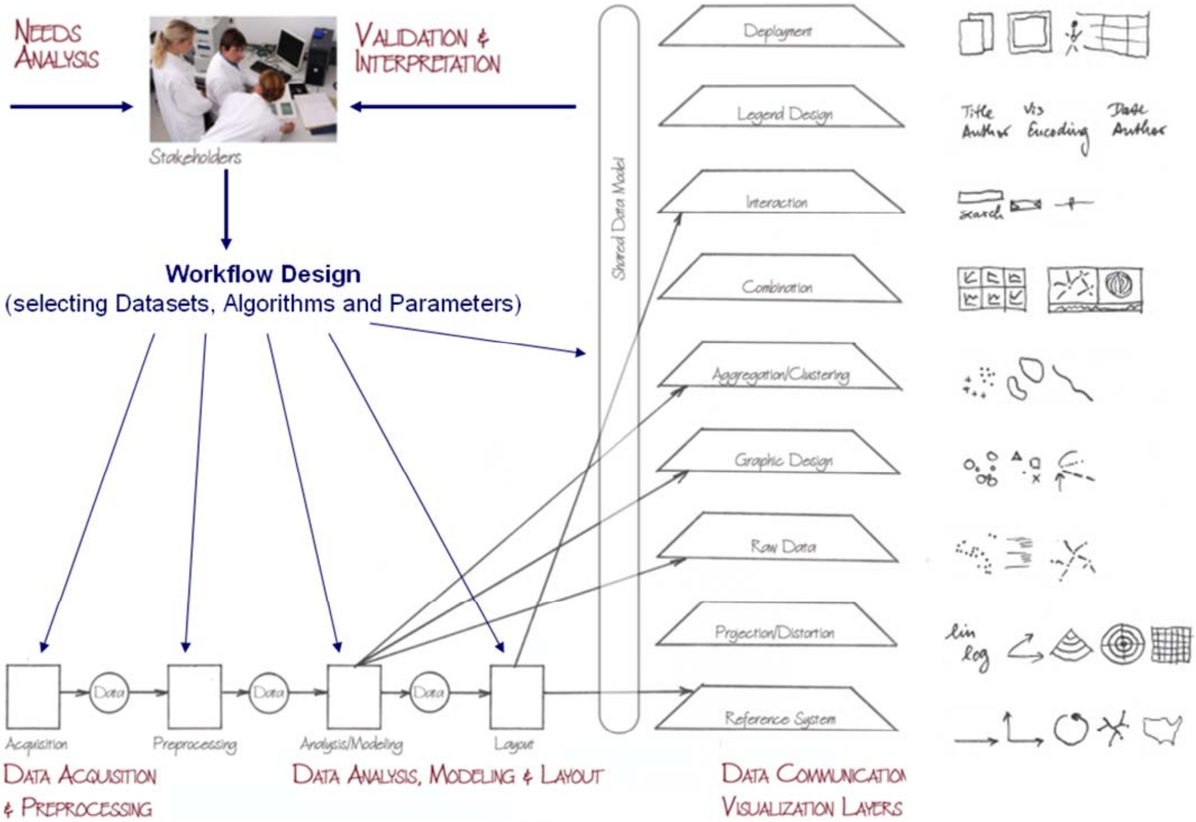
Part 3: Toward a Science of Science

Those who cannot remember the past are condemned to repeat it.
George Santayana



49

Part 3: Toward a Science of Science



Temporal Analysis

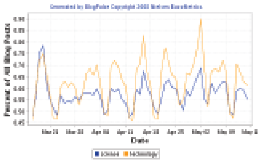
Science evolves over time. Attribute values of scholarly entities and their diverse aggregations increase and decrease at different rates and respond with different latency rates to internal and external events. Temporal analysis aims to identify the nature of phenomena represented by a sequence of observations such as patterns, trends, seasonality, outliers, and bursts of activity.

Data

A time series is a sequence of events or observations that are ordered in time. Time-series data can be continuous (there is an observation at every instant of time; see figure to the right) or discrete (observations exist for regularly or irregularly spaced intervals). Temporal aggregations—over journal volumes, years, or decades—are common.

Algorithms

Frequently, some form of filtering is applied to reduce noise and make patterns more salient. Smoothing (averaging using a smoothing window of a certain width) and curve approximation might be applied. The number of scholarly records is often plotted to get a first idea of the temporal distribution of a data set. It might be shown in total values or as a percentage of those. One may find out how long a scholarly entity was active; how old it was at a certain point; what growth, latency to peak, or decay rate it has; what correlations with other time series exist; or what trends are observable. Data models such as the least squares model (available in most statistical software packages) are applied to best fit a selected function to a data set and to determine if the trend is significant. Kleinburg's burst detection algorithm is commonly applied to identify words that have experienced a sudden change in frequency of occurrence.



Geographic Analysis

Geographic analysis aims to answer the question of where something happens and what impact it has on neighboring areas.

Data

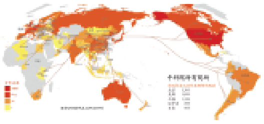
Geographic analysis requires spatial attribute values or geolocations for authors and their papers, extracted from affiliation data or spatial positions of nodes, generated from layout algorithms. Geographic data can be continuous (each record has a specific position) or discrete (a position or area exists for sets of records, like the number of papers per country). Spatial aggregations (for example, merging data via postal codes, counties, states, countries, and continents) are common (see page 66, Exemplification).

Algorithms

Cartographic generalization refers to the process of abstraction. This includes (1) graphic generalization: the simplification, enlargement, displacement, merging, or selection of entities without enhancement or effect to their symbology and (2) conceptual symbolization: the merging, selection, and symbolization of entities, including enhancement (such as representing high-density areas with a city symbol).

Geometric generalization aims to solve the conflict between the number of visualized features, the size of symbols, and the size of the display surface. Cartographers dealt with this conflict intuitively in part until researchers like Friedrich Töpfer attempted to solve them with quantifiable expressions.

Flow maps use line thickness and direction to show the number of tangible or intangible entities that diffuse over a geographic location or science space (see CAS coauthor network, below, and page 158, 113 Years of Physical Review).



Topical Analysis

The topic coverage and topical similarity of basic and aggregate units of science (authors or institutions) can be derived from the units associated with them (papers, patents, or grants).

Data

The topic or semantic coverage of a unit of science can be derived from the text associated with it. Topical aggregations (for example, over journal volumes, scientific disciplines, or institutions) are common.

Algorithms

Topic analysis extracts the set of unique words or word profiles and their frequency from a text corpus. Stop words, such as "the" and "of," are removed. Stemming can be applied. Coword analysis identifies the number of times two words are used in the title, keyword set, abstract, or full text of a paper. The space of co-occurring words can be mapped, providing a unique view of the topic coverage of a data set (see page 66, Exemplification). Similarly, units of science can be grouped according to the number of words they have in common.

Salton's term frequency inverse document frequency (TFIDF) is a statistical measure used to evaluate the importance of a word in a corpus. The importance increases proportionally to the number of times a word appears in the paper but is offset by the frequency of the word in the corpus.

Dimensionality reduction techniques (see table on opposite page) are commonly used to project high-dimensional information spaces (for example, the matrix of all unique papers multiplied by their unique terms) into a low, typically two-dimensional space.

The SOM map below shows the topic landscape of geography abstracts; see page 102, In Terms of Geography.

Network Analysis

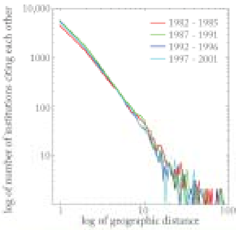
The study of networks aims to increase our understanding of natural and manmade networks. It builds on social network analysis, physics, information science, bibliometrics, scientometrics, informetrics, webometrics, communication theory, sociology of science, and several other disciplines.

Data

Authors, institutions, and countries, as well as words, papers, journals, patents, and funding, are represented as nodes and their complex interrelations as edges (see Part 3: Toward a Science of Science/Conceptualizing Science: Basic Anatomy of Science). Nodes and edges can have time-stamped attributes.

Algorithms

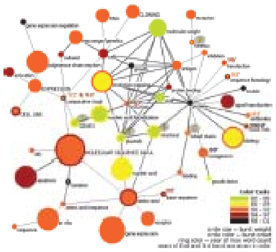
Diverse algorithms exist to calculate specific node, edge, and network properties (see "Network Science" review). Node properties include degree centrality, betweenness centrality, or hub and authority scores. Edge properties include durability, reciprocity, intensity (weak or strong), density (how many potential edges in a network actually exist), reachability (how many steps it takes to go from one "end" of a network to another), centrality (whether a network has a "center" point), quality (reliability or certainty), and strength. Network properties refer to the number of nodes and edges, network density, average path length, clustering coefficient, and distributions from which general properties such as "small-world," "scale-free," or "hierarchical" can be derived. Identifying major communities via community detection algorithms and calculating the "backbone" of a network via pathfinder network scaling or maximum flow algorithms helps to communicate and make sense of large-scale networks. See the coauthor network of information visualization researchers below.



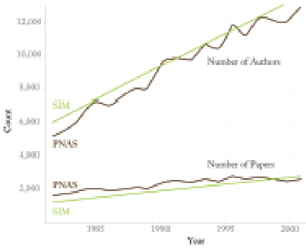
I(b) Log-log plot showing the variation of the number of institutions that cite one another over geographic distance between them for each of the four time slices

I(c) Geographic locations of the top 500 institutions with number of received citations indicated by height and color of bars

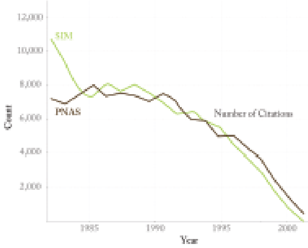
II(b) Fruchterman–Reingold layout of word co-occurrence matrix



II(c) Final layout with size- and color-coding, labels, and legend



III(b) Number of authors and papers over time



III(c) Number of received citations over time

Interpretation

Study I: Mapping Knowledge Diffusion and the Importance of Space

This study aimed to determine whether the Internet leads to more global citation patterns (that is, more citation links between papers produced at geographically distant research institutions). A novel approach to analyzing the dual role of institutions as information producers and consumers and to studying and visualizing the diffusion of information among them was developed. Surprisingly, the widespread adoption of the Internet does not seem to have affected the distance over which information diffuses as manifested by citation links. The citation linkages between institutions fall off with the distance between them, and there is a strong linear relationship between the log of the citation counts and the log of the distance that does not change over time. Reasons for local collaborations might include "winner takes all" funding schemes; the demands of complex, large-scale instrumentation; and the need to gain experience, train researchers, and sponsor protégés. . The social component of citation seems to become more important as researchers are flooded with information, and spatial proximity eases the creation and continuation of close personal relationships.

Study II: Identifying Research Topics and Trends

Scientific research is highly dynamic. New areas of science are continually evolving; some may shift in importance, others merge or split. Because of the steady increase in the number of scientific publications, it is challenging to keep abreast of the structure and dynamic development of one's own field, let alone all scientific domains. However, knowledge of "hot" topics, emergent research frontiers, and change of focus in certain areas is a critical component of resource allocation decisions in research laboratories, government institutions, and corporations. This study aimed to increase understanding of the topic coverage and activity bursts of words in highly cited PNAS papers. Interestingly, the burst of words seems to proceed their wide spread usage. . "Protein" and "model" were among the highly "bursty" terms between 1998 and 2001, and they have become important research topics since then.

Study III: Modeling the Coevolution of Author–Paper Networks

Models of scientific structure and evolution can help us understand the inner workings of science (see page 58, Conceptualizing Science: Science Dynamics). The TARL model (topics, aging, and recursive linking) describes the coevolution of coauthor and paper-citation networks. Using an agent-based approach, TARL simulates nodes (authors or papers), their edges (undirected coauthor, directed consumed, and directed paper-citation), and their attributes (time and topics). Topics cluster papers and authors *topically*. Aging is an antagonistic force to preferential attachment. Even highly connected nodes receive a decreasing number of links over time. Aging clusters papers and authors *temporally*. Recursive linking refers to the tendency of authors to cite papers referenced in material they are currently reading, which provides a grounded mechanism for the "rich get richer" phenomenon as an emergent property of the elementary activity of authors. According to this model, the number of topics is linearly related to the clustering coefficient of the simulated paper citation network.

Part 4: Science Maps in Action

If we ever get to the point of charting a whole city or a whole nation, we would have … a picture of a vast solar system of intangible structures, powerfully influencing conduct, as gravitation does in space. Such an invisible structure underlies society and has its influence in determining the conduct of society as a whole.

Jacob L. Moreno

First Iteration of Exhibit (2005): The Power of Maps

Four Early Maps of Our World
Versus
Six Early Maps of Science

The first exhibit iteration on *The Power of Maps* demonstrates how maps help us to understand, navigate, and manage both physical places and abstract knowledge spaces.

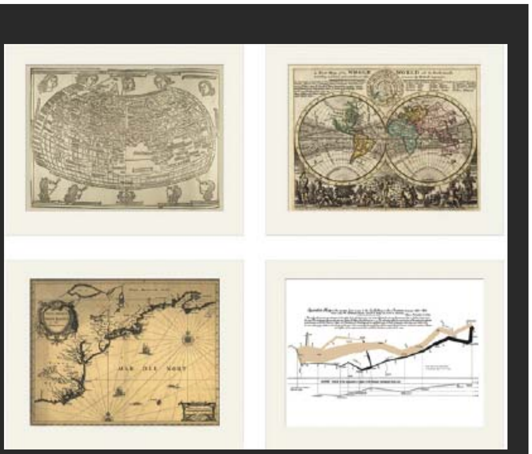
Early maps of our planet were certainly neither complete nor perfect, yet they proved invaluable for explorers. As keys to navigation, exploration, and communication, maps helped explorers find promising new lands while avoiding sea monsters.

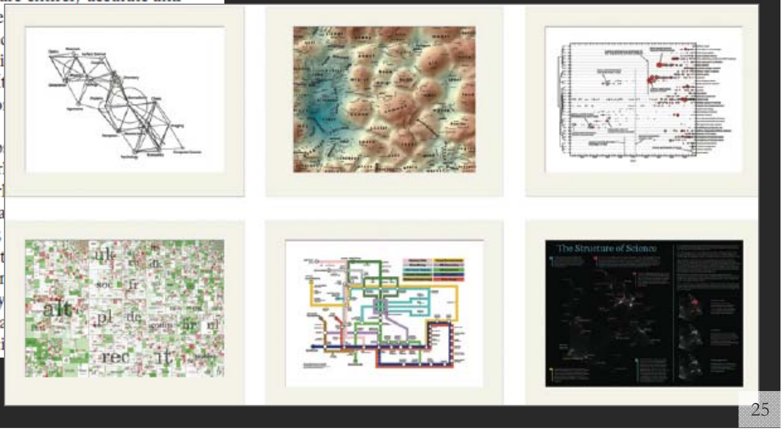
Maps of science today are based on limited knowledge and therefore imperfect. In order to generate comprehensive maps that are entirely accurate and reliable, we must first have proper coverage and inte
multidisciplinary, and multimedia scholarly knowle

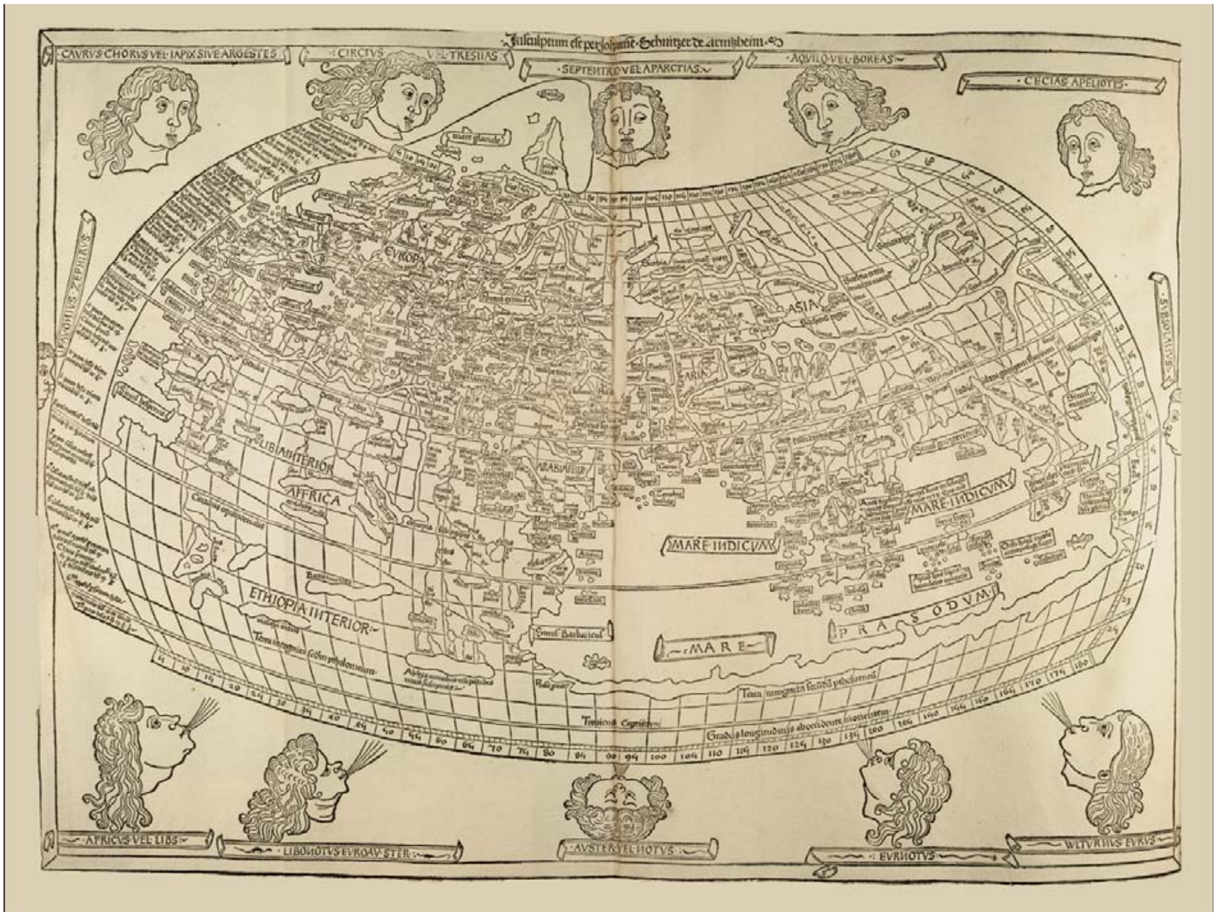
The first pictures of Earth from space were experi
mative of their perceptions of life and the cosmos. I
science will increase our appreciation and applicatio
serving as useful navigational tools.

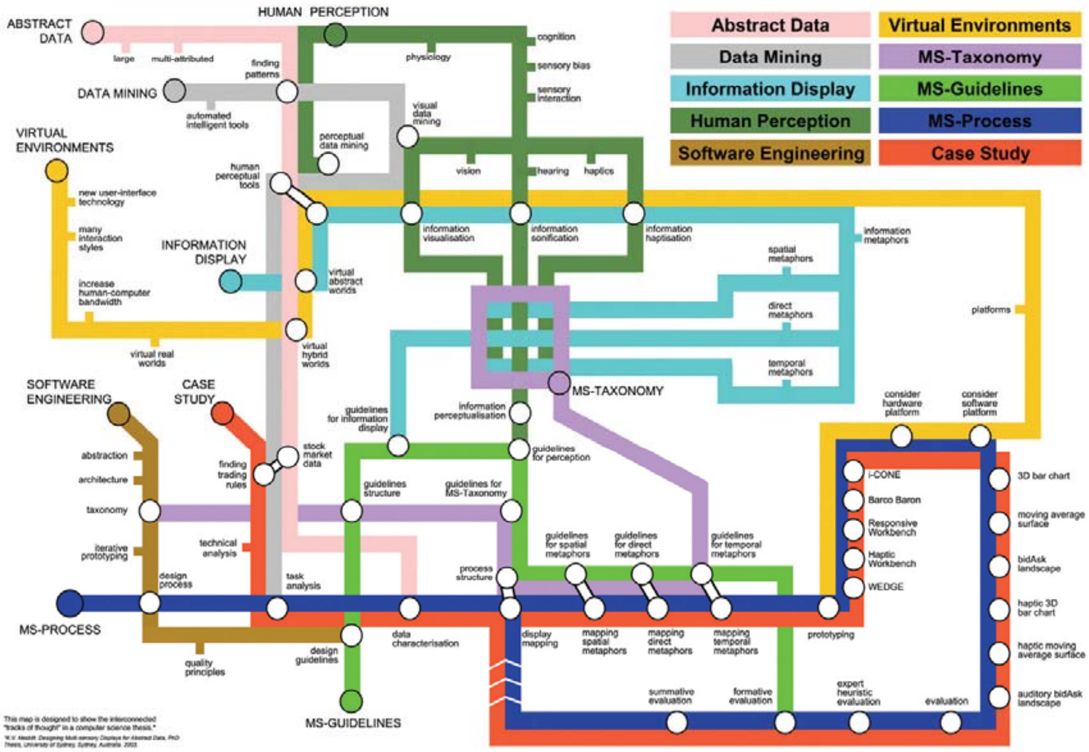
The Power of Maps features four cartographic map
earliest global maps of our world by Ptolemy, an ear
Johannes Janssonius, an early map of the whole wor
early statistical graph by Charles Joseph Minard. Ea
employs a different metaphor: a node-link diagram;
ing map rendered using geographic information syst
a crossmap; and a galaxy view. Which metaphor is
visual index of our collective science and technology

Note that the makers of the early cartographic ma
ing presses, while the makers of the first maps of sci



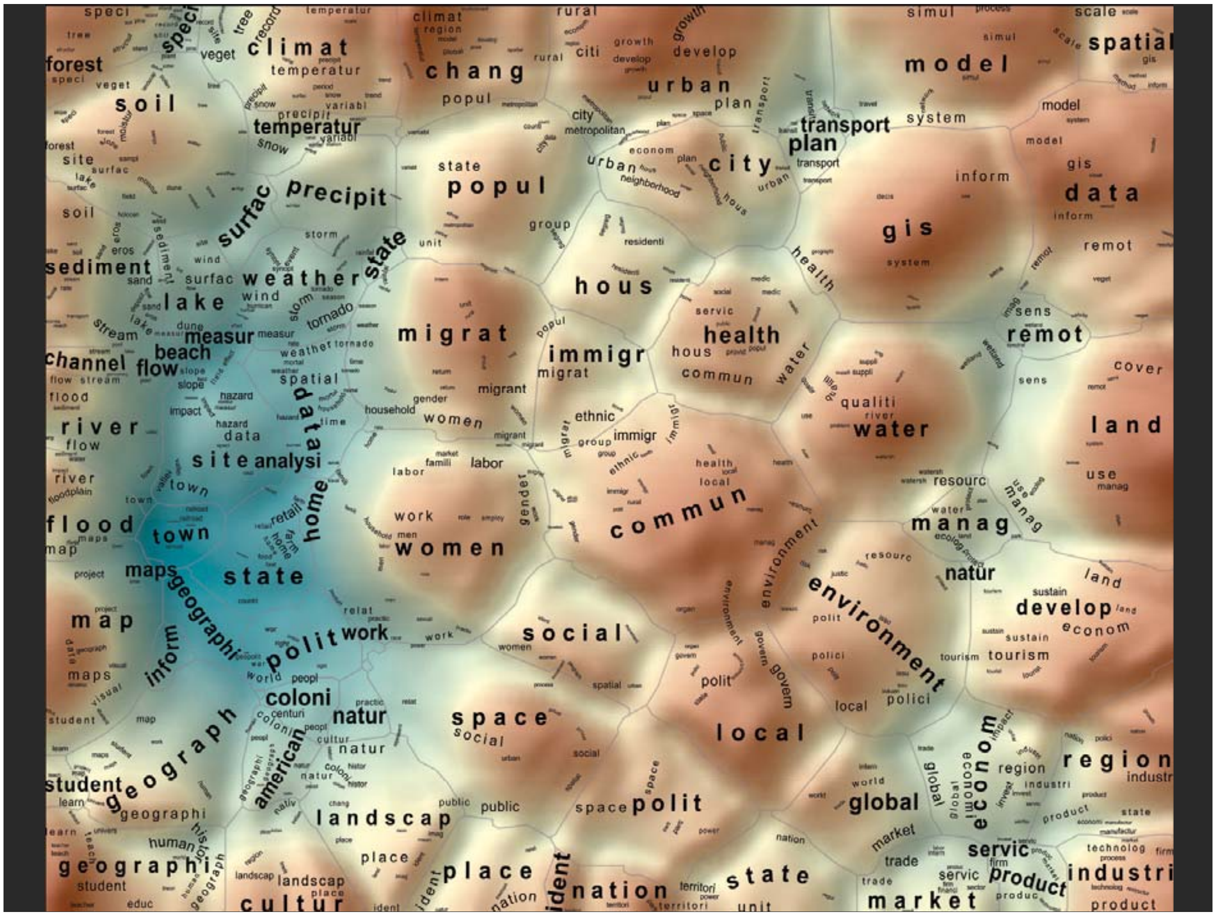




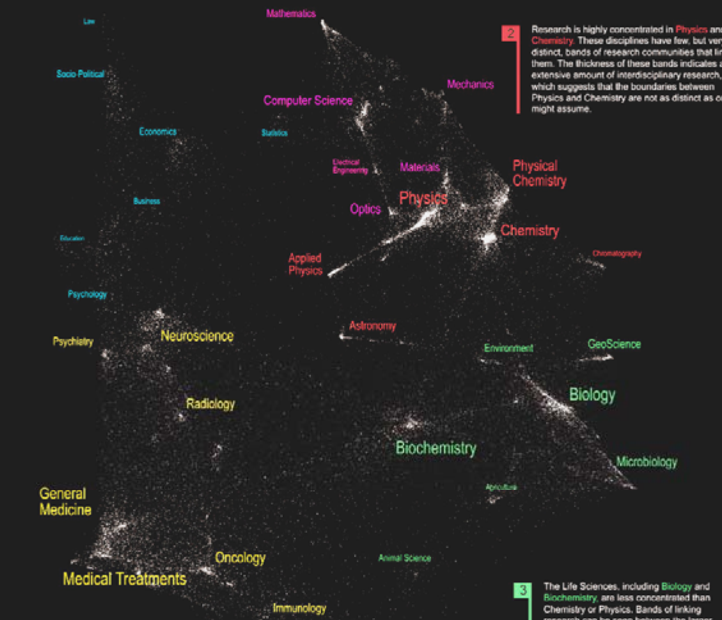


This map is designed to show the interconnected "tracks of thought" in a computer science thesis.*

*K.V. Nesbitt. *Designing Multi-sensory Displays for Abstract Data*, PhD Thesis, University of Sydney, Sydney, Australia. 2003.



The Structure of Science



5 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.

1 Mathematics is our starting point, the purest of all sciences. It lies at the outer edge of the map. Computer Science, Electrical Engineering, and Optics are 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.

2 Research is highly concentrated in Physics and Chemistry. These disciplines have few, but very distinct, bands of research communities that link them. The thickness 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.

4 The Medical Sciences include broad therapeutic studies and targeted areas of Treatment (e.g. central nervous system, cardiology, gastroenterology, etc.) Unlike Physics and Chemistry, the medical disciplines are more spread out, suggesting a more multi-disciplinary approach to research. The transition into Life Sciences (via Animal Science and Biochemistry) is gradual.

3 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. Biochemistry is very interesting in that 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.

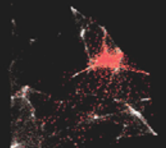
We are all familiar with traditional maps that show the relationships between countries, provinces, states, and cities. Similar relationships exist between the various disciplines and research topics in science. This allows us to map the structure of science.

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 cover enough of science to accurately define its structure.

Things are different today. We have enormous computing power and advanced visualization software that make mapping of the structure of science possible. This galaxy-like map of science (left) was generated at Sandia National Laboratories using an advanced graph layout routine (VxOrd) from the citation patterns in 800,000 scientific papers published in 2002. 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 on 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, exploit, abandon, or ignore.

We also envision the map as an educational tool. 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 Physics, 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 most of its activity in Medicine. It also has many communities in Biochemistry and two communities in the Social Sciences.

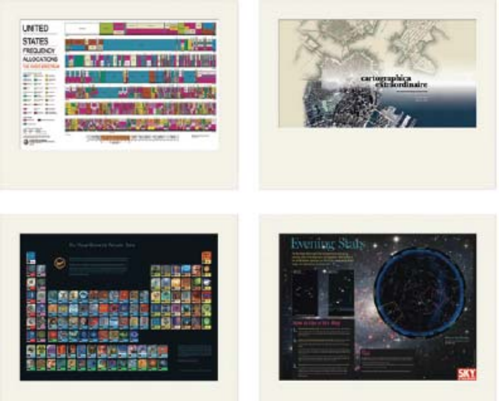
Second Iteration of Exhibit (2006): The Power of Reference Systems

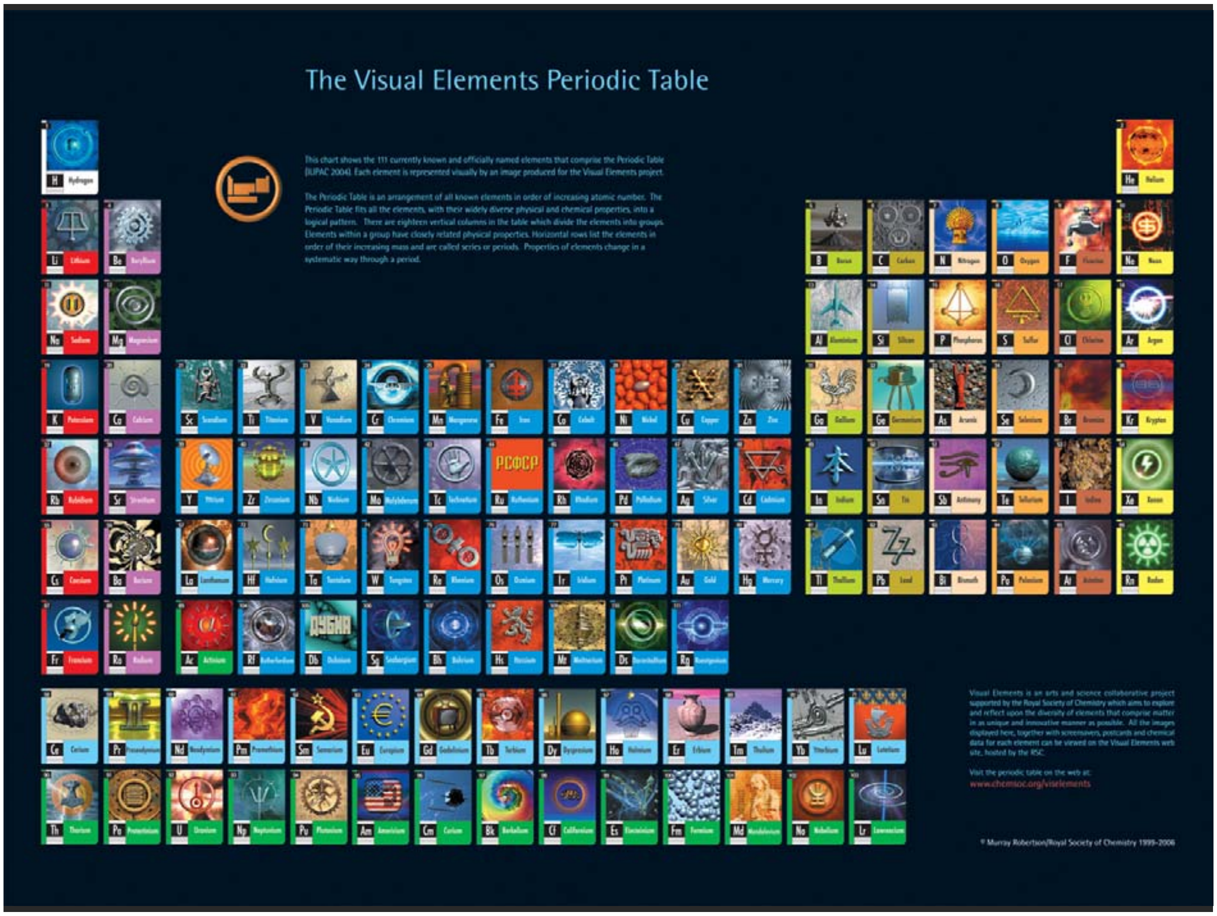
Four Existing Reference Systems
Versus
Six Potential Reference Systems

This iteration aims to inspire discussion about a common reference system for all existing scholarly knowledge. Throughout history, scientists have battled to agree on standardized reference systems for their respective fields of research. These standards are invaluable for indexing, storing, accessing, and managing scientific data efficiently.

Results include the description of the electromag
odic table of elements, geographic projections, and
systems, shown here. Note that the geographic ma
from paper to geographic information systems (GI
for public use and consumption.

In comparison to these four existing systems are
systems for scholarly knowledge. Each reference sy
sional timeline and the geographic system to the se
used to identify the location of an author, paper, pa
tory or contribution.





The Visual Elements Periodic Table

This chart shows the 111 currently known and officially named elements that comprise the Periodic Table (IUPAC 2004). Each element is represented visually by an image produced for the Visual Elements project.

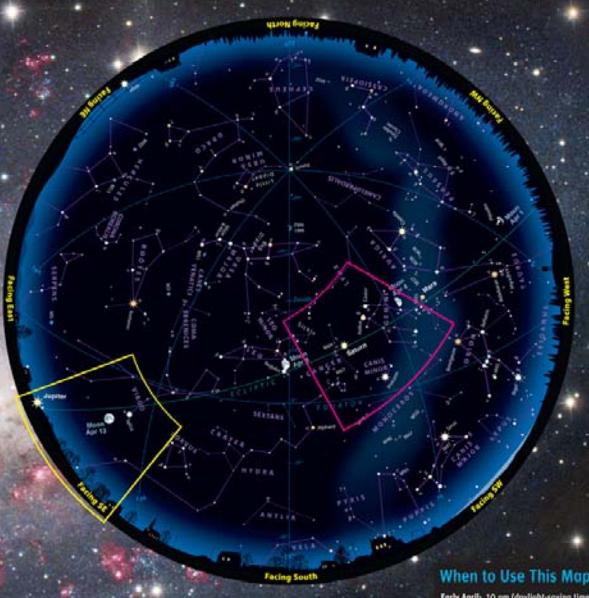
The Periodic Table is an arrangement of all known elements in order of increasing atomic number. The Periodic Table fits all the elements, with their widely diverse physical and chemical properties, into a logical pattern. There are eighteen vertical columns in the table which divide the elements into groups. Elements within a group have closely related physical properties. Horizontal rows list the elements in order of their increasing mass and are called series or periods. Properties of elements change in a systematic way through a period.

Visual Elements is an arts and science collaborative project supported by the Royal Society of Chemistry which aims to explore and reflect upon the diversity of elements that comprise matter in as unique and innovative manner as possible. All the images displayed here, together with screensavers, postcards and chemical data for each element can be viewed on the Visual Elements web site, hosted by the RSC.

Visit the periodic table on the web at:
www.chemsoc.org/viselements

Evening Stars

The Big Dipper floats high in the northeast these early spring evenings, while Orion sinks low in the southwest. These are just a few of the celestial sights you can find on any clear evening in April using a sky map like the one shown here.



How to Use a Sky Map

1. **Check the dates and times at right.** Take your map out under the night sky around the right time, and bring along a flashlight to read it by. It helps to attach a piece of red paper over the front or to use a flashlight with red LEDs; the dim red light won't spoil your night vision.

2. **Outside, you need to know which direction you're facing.** (If you're unsure, just note where the Sun sets; that's west.) Whichever way you're facing, make sure the corresponding yellow label along the curved edge of the map is at the bottom, right-side up.

This curved edge represents the horizon. The stars above it on the map match the stars in front of you. The farther up from the map's edge they appear, the higher they'll be in the sky.

The center of the map is the zenith (straight overhead). So a star halfway from the edge of the map to the center will appear halfway from straight ahead to straight up. Ignore all the parts of the map above horizons you're not facing.

3. **Let's give it a try!** Pretend you're facing the southwest horizon (labeled "Facing SW"). Just a little way up (that is, a little way in from the edge of the map) is Sirius, the brightest star in the night sky, in the constellation Canis Major. Farther up, nearly halfway overhead, is the star Procyon in Canis Minor. Still farther up is the ringed planet Saturn. Go out at the right time, face southwest, and look up into the sky — there they are!

When to Use This Map

Early April: 10 pm (daylight-saving time)
Late April: Dusk

Tips

A couple of tips: Look for the brightest stars and constellations first; light pollution or moonlight may wash out the fainter ones. And remember that star patterns in the sky will look a lot bigger than they do here on paper.

With a map like this, you can identify celestial sights all over the sky. Go out the next clear night and make some starry friends!

You can customize a night-sky map for any time and place at SkyandTelescope.com.

PHOTO COURTESY OF ROBERT GENDLER

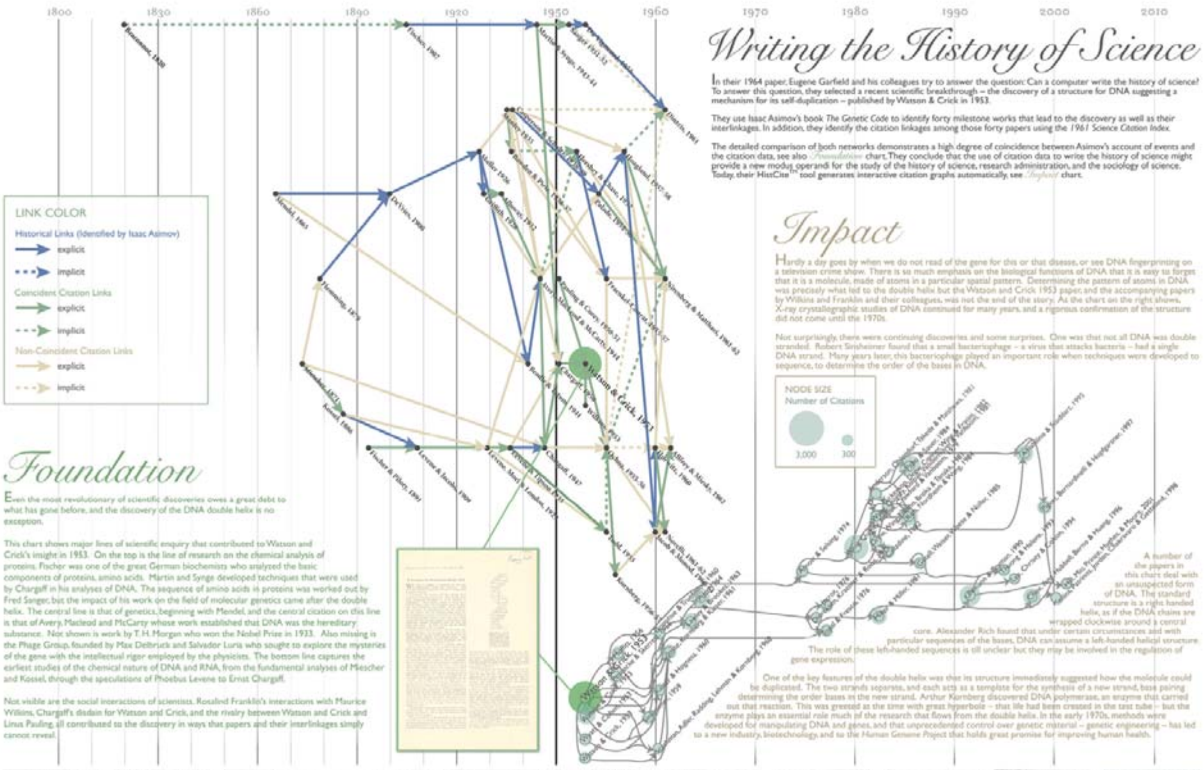
SKY® & TELESCOPE

Writing the History of Science

In their 1964 paper, Eugene Garfield and his colleagues try to answer the question: Can a computer write the history of science? To answer this question, they selected a recent scientific breakthrough – the discovery of a structure for DNA suggesting a mechanism for its self-duplication – published by Watson & Crick in 1953.

They use Isaac Asimov's book *The Genetic Code* to identify forty milestone works that lead to the discovery as well as their interlinkages. In addition, they identify the citation linkages among those forty papers using the *1961 Science Citation Index*.

The detailed comparison of both networks demonstrates a high degree of coincidence between Asimov's account of events and the citation data, see also *Foundation* chart. They conclude that the use of citation data to write the history of science might provide a new modus operandi for the study of the history of science, research administration, and the sociology of science. Today, their HistCite™ tool generates interactive citation graphs automatically, see *Impact* chart.



Impact

Hardly a day goes by when we do not read of the gene for this or that disease, or see DNA fingerprinting on a television crime show. There is so much emphasis on the biological functions of DNA that it is easy to forget that it is a molecule, made of atoms in a particular spatial pattern. Determining the pattern of atoms in DNA was precisely what led to the double helix but the Watson and Crick 1953 paper, and the accompanying papers by Wilkins and Franklin and their colleagues, was not the end of the story. As the chart on the right shows, X-ray crystallographic studies of DNA continued for many years, and a rigorous confirmation of the structure did not come until the 1970s.

Not surprisingly, there were continuing discoveries and some surprises. One was that not all DNA was double stranded. Robert Sinsheimer found that a small bacteriophage – a virus that attacks bacteria – had a single DNA strand. Many years later, this bacteriophage played an important role when techniques were developed to sequence, to determine the order of the bases in DNA.

A number of the papers in this chart deal with an unsuspected form of DNA. The standard structure is a right handed helix, as if the DNA chains are wrapped clockwise around a central core. Alexander Rich found that under certain circumstances and with particular sequences of the bases, DNA can assume a left-handed helical structure. The role of these left-handed sequences is still unclear but they may be involved in the regulation of gene expression.

One of the key features of the double helix was that its structure immediately suggested how the molecule could be duplicated. The two strands separate, and each acts as a template for the synthesis of a new strand, base pairing determining the order bases in the new strand. Arthur Kornberg discovered DNA polymerase, an enzyme that carried out that reaction. This was greeted at the time with great hyperbole – that life had been created in the test tube – but the enzyme plays an essential role much of the research that flows from the double helix. In the early 1970s, methods were developed for manipulating DNA and genes, and that unprecedented control over genetic material – genetic engineering – has led to a new industry, biotechnology, and to the *Human Genome Project* that holds great promise for improving human health.

Foundation

Even the most revolutionary of scientific discoveries owes a great debt to what has gone before, and the discovery of the DNA double helix is no exception.

This chart shows major lines of scientific enquiry that contributed to Watson and Crick's insight in 1953. On the top is the line of research on the chemical analysis of proteins. Fischer was one of the great German biochemists who analyzed the basic components of proteins, amino acids. Martin and Synge developed techniques that were used by Chargaff in his analyses of DNA. The sequence of amino acids in proteins was worked out by Fred Sanger, but the impact of his work on the field of molecular genetics came after the double helix. The central line is that of genetics, beginning with Mendel, and the central citation on this line is that of Avery, Macleod and McCarty whose work established that DNA was the hereditary substance. Not shown is work by T. H. Morgan who won the Nobel Prize in 1933. Also missing is the Phage Group, founded by Max Delbruck and Salvador Luria who sought to explore the mysteries of the gene with the intellectual rigor employed by the physicists. The bottom line captures the earliest studies of the chemical nature of DNA and RNA, from the fundamental analyses of Miescher and Kossel, through the speculations of Phoebus Levene to Ernst Chargaff.

Not visible are the social interactions of scientists. Rosalind Franklin's interactions with Maurice Wilkins, Chargaff's disdain for Watson and Crick, and the rivalry between Watson and Crick and Linus Pauling, all contributed to the discovery in ways that papers and their interlinkages simply cannot reveal.


1947 A. Mirsky & I. Goodman


1947 J. Monod


1947 E. Chargaff


1948 A. Mirsky & P.C. Koller


1950 R. Franklin


1953 J. D. Watson & F. H. C. Crick

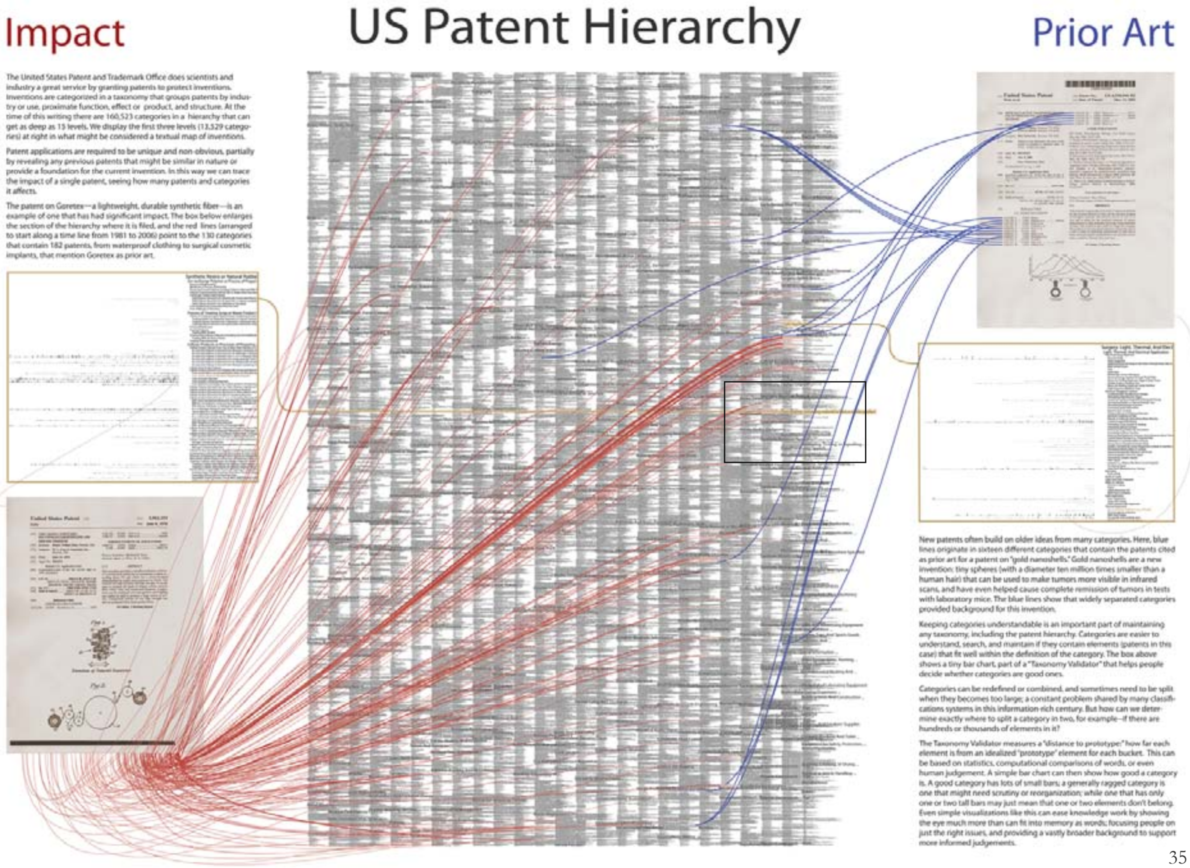

1963 V. Ingram, M. Nirenberg & H. Staehelin


1963 J. Speyer & M. Nirenberg


1968 C. Thomas & A. Kornberg


1974 Participants


1978 A. Kornberg



Impact

US Patent Hierarchy

Prior Art

The United States Patent and Trademark Office does scientists and industry a great service by granting patents to protect inventions. Inventions are categorized in a taxonomy that groups patents by industry or use, proximate function, effect or product, and structure. At the time of this writing there are 160,523 categories in a hierarchy that can get as deep as 15 levels. We display the first three levels (13,529 categories) at right in what might be considered a textual map of inventions.

Patent applications are required to be unique and non-obvious, partially by revealing any previous patents that might be similar in nature or provide a foundation for the current invention. In this way we can trace the impact of a single patent, seeing how many patents and categories it affects.

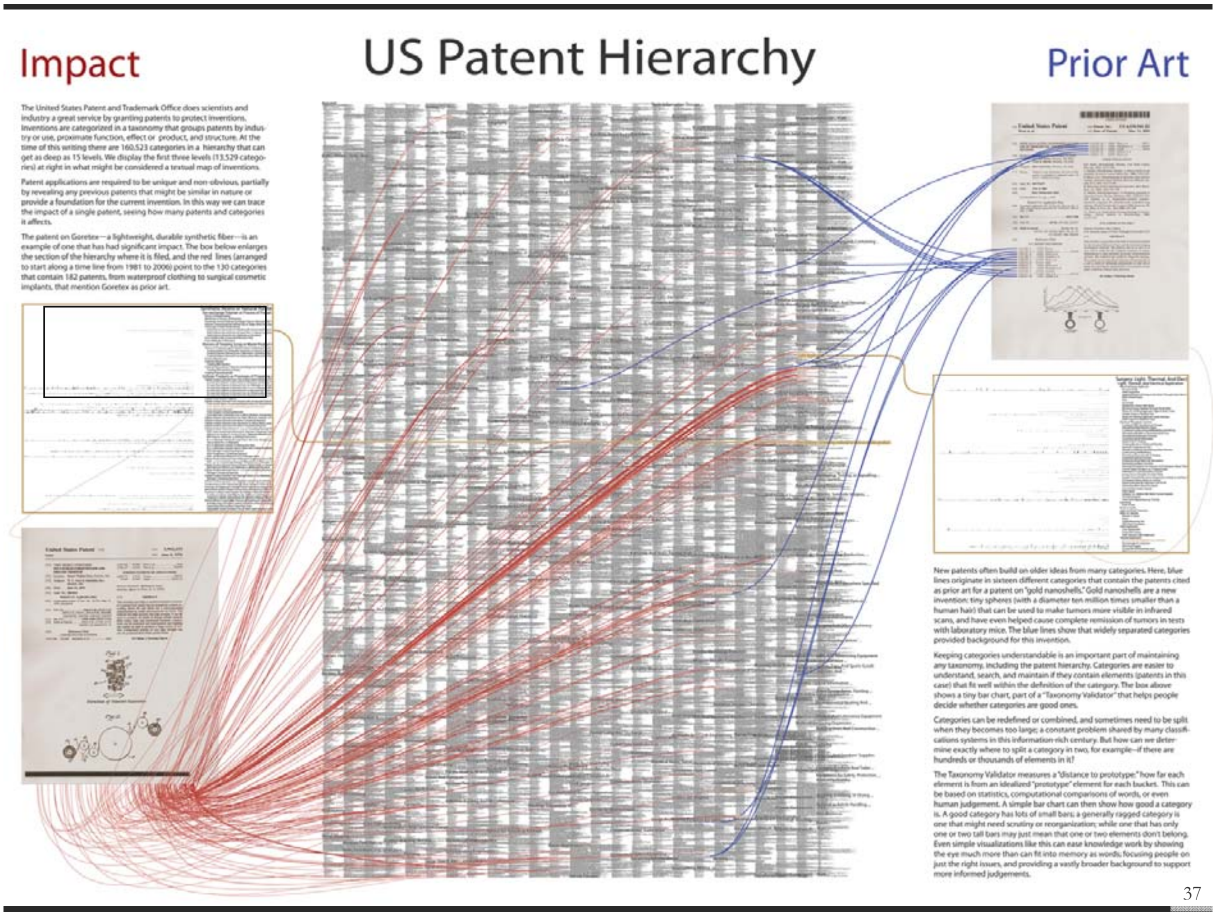
The patent on Goretex—a lightweight, durable synthetic fiber—is an example of one that has had significant impact. The box below enlarges the section of the hierarchy where it is filed, and the red lines (arranged to start along a time line from 1981 to 2006) point to the 130 categories that contain 182 patents, from waterproof clothing to surgical cosmetic implants, that mention Goretex as prior art.

New patents often build on older ideas from many categories. Here, blue lines originate in sixteen different categories that contain the patents cited as prior art for a patent on "gold nanoshells." Gold nanoshells are a new invention: tiny spheres (with a diameter ten million times smaller than a human hair) that can be used to make tumors more visible in infrared scans, and have even helped cause complete remission of tumors in tests with laboratory mice. The blue lines show that widely separated categories provided background for this invention.

Keeping categories understandable is an important part of maintaining any taxonomy, including the patent hierarchy. Categories are easier to understand, search, and maintain if they contain elements (patents in this case) that fit well within the definition of the category. The box above shows a tiny bar chart, part of a "Taxonomy Validator" that helps people decide whether categories are good ones.

Categories can be redefined or combined, and sometimes need to be split when they becomes too large; a constant problem shared by many classifications systems in this information-rich century. But how can we determine exactly where to split a category in two, for example—if there are hundreds or thousands of elements in it?

The Taxonomy Validator measures a "distance to prototype:" how far each element is from an idealized "prototype" element for each bucket. This can be based on statistics, computational comparisons of words, or even human judgement. A simple bar chart can then show how good a category is. A good category has lots of small bars; a generally ragged category is one that might need scrutiny or reorganization; while one that has only one or two tall bars may just mean that one or two elements don't belong. Even simple visualizations like this can ease knowledge work by showing the eye much more than can fit into memory as words; focusing people on just the right issues, and providing a vastly broader background to support more informed judgements.



US Patent Hierarchy

The United States Patent and Trademark Office does scientists and industry a great service by granting patents to protect inventions. Inventions are categorized in a taxonomy that groups patents by industry or use, proximate function, effect or product, and structure. At the time of this writing there are 160,523 categories in a hierarchy that can get as deep as 15 levels. We display the first three levels (13,529 categories) at right in what might be considered a textual map of inventions.

Patent applications are required to be unique and non-obvious, partially by revealing any previous patents that might be similar in nature or provide a foundation for the current invention. In this way we can trace the impact of a single patent, seeing how many patents and categories it affects.

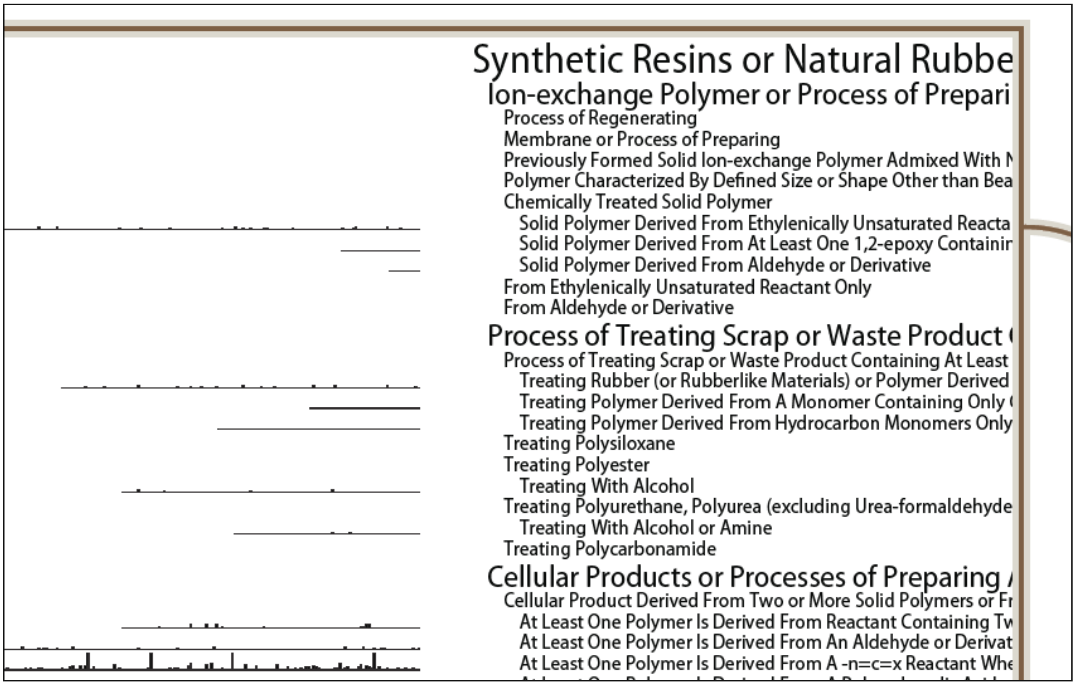
The patent on Goretex—a lightweight, durable synthetic fiber—is an example of one that has had significant impact. The box below enlarges the section of the hierarchy where it is filed, and the red lines (arranged to start along a time line from 1981 to 2006) point to the 130 categories that contain 182 patents, from waterproof clothing to surgical cosmetic implants, that mention Goretex as prior art.

New patents often build on older ideas from many categories. Here, blue lines originate in sixteen different categories that contain the patents cited as prior art for a patent on "gold nanoshells." Gold nanoshells are a new invention: tiny spheres (with a diameter ten million times smaller than a human hair) that can be used to make tumors more visible in infrared scans, and have even helped cause complete remission of tumors in tests with laboratory mice. The blue lines show that widely separated categories provided background for this invention.

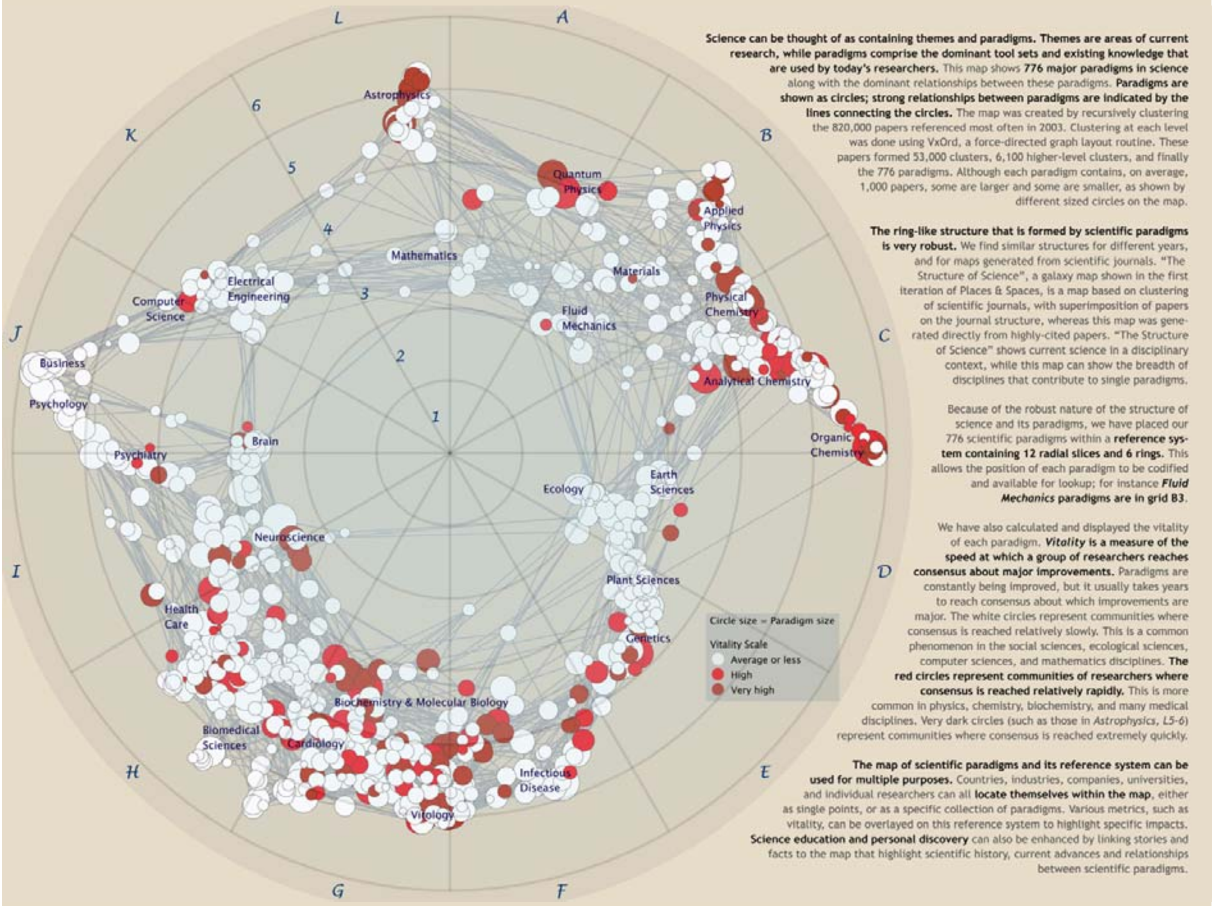
Keeping categories understandable is an important part of maintaining any taxonomy, including the patent hierarchy. Categories are easier to understand, search, and maintain if they contain elements (patents in this case) that fit well within the definition of the category. The box above shows a tiny bar chart, part of a "Taxonomy Validator" that helps people decide whether categories are good ones.

Categories can be redefined or combined, and sometimes need to be split when they becomes too large; a constant problem shared by many classifications systems in this information-rich century. But how can we determine exactly where to split a category in two, for example—if there are hundreds or thousands of elements in it?

The Taxonomy Validator measures a "distance to prototype:" how far each element is from an idealized "prototype" element for each bucket. This can be based on statistics, computational comparisons of words, or even human judgement. A simple bar chart can then show how good a category is. A good category has lots of small bars; a generally ragged category is one that might need scrutiny or reorganization; while one that has only one or two tall bars may just mean that one or two elements don't belong. Even simple visualizations like this can ease knowledge work by showing the eye much more than can fit into memory as words, focusing people on just the right issues, and providing a vastly broader background to support more informed judgements.



Synthetic Resins or Natural Rubbe
Ion-exchange Polymer or Process of Prepari
Process of Regenerating
Membrane or Process of Preparing
Previously Formed Solid Ion-exchange Polymer Admixed With N
Polymer Characterized By Defined Size or Shape Other than Bea
Chemically Treated Solid Polymer
Solid Polymer Derived From Ethylenically Unsaturated Reacta
Solid Polymer Derived From At Least One 1,2-epoxy Containir
Solid Polymer Derived From Aldehyde or Derivative
From Ethylenically Unsaturated Reactant Only
From Aldehyde or Derivative
Process of Treating Scrap or Waste Product (
Process of Treating Scrap or Waste Product Containing At Least
Treating Rubber (or Rubberlike Materials) or Polymer Derived
Treating Polymer Derived From A Monomer Containing Only (
Treating Polymer Derived From Hydrocarbon Monomers Only
Treating Polysiloxane
Treating Polyester
Treating With Alcohol
Treating Polyurethane, Polyurea (excluding Urea-formaldehyde
Treating With Alcohol or Amine
Treating Polycarbonamide
Cellular Products or Processes of Preparing /
Cellular Product Derived From Two or More Solid Polymers or Fr
At Least One Polymer Is Derived From Reactant Containing Tw
At Least One Polymer Is Derived From An Aldehyde or Derivat
At Least One Polymer Is Derived From A -n=c=x Reactant Whe



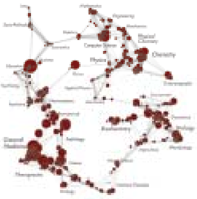
Science can be thought of as containing themes and paradigms. Themes are areas of current research, while paradigms comprise the dominant tool sets and existing knowledge that are used by today's researchers. This map shows **776 major paradigms in science** along with the dominant relationships between these paradigms. **Paradigms are shown as circles; strong relationships between paradigms are indicated by the lines connecting the circles.** The map was created by recursively clustering the 820,000 papers referenced most often in 2003. Clustering at each level was done using VxOrd, a force-directed graph layout routine. These papers formed 53,000 clusters, 6,100 higher-level clusters, and finally the 776 paradigms. Although each paradigm contains, on average, 1,000 papers, some are larger and some are smaller, as shown by different sized circles on the map.

The ring-like structure that is formed by scientific paradigms is very robust. We find similar structures for different years, and for maps generated from scientific journals. "The Structure of Science", a galaxy map shown in the first iteration of Places & Spaces, is a map based on clustering of scientific journals, with superimposition of papers on the journal structure, whereas this map was generated directly from highly-cited papers. "The Structure of Science" shows current science in a disciplinary context, while this map can show the breadth of disciplines that contribute to single paradigms.

Because of the robust nature of the structure of science and its paradigms, we have placed our 776 scientific paradigms within a **reference system containing 12 radial slices and 6 rings.** This allows the position of each paradigm to be codified and available for lookup; for instance ***Fluid Mechanics*** **paradigms are in grid B3**.

We have also calculated and displayed the vitality of each paradigm. ***Vitality*** **is a measure of the speed at which a group of researchers reaches consensus about major improvements.** Paradigms are constantly being improved, but it usually takes years to reach consensus about which improvements are major. The white circles represent communities where consensus is reached relatively slowly. This is a common phenomenon in the social sciences, ecological sciences, computer sciences, and mathematics disciplines. **The red circles represent communities of researchers where consensus is reached relatively rapidly.** This is more common in physics, chemistry, biochemistry, and many medical disciplines. Very dark circles (such as those in *Astrophysics, L5-6*) represent communities where consensus is reached extremely quickly.

The map of scientific paradigms and its reference system can be used for multiple purposes. Countries, industries, companies, universities, and individual researchers can all **locate themselves within the map**, either as single points, or as a specific collection of paradigms. Various metrics, such as vitality, can be overlayed on this reference system to highlight specific impacts. **Science education and personal discovery** can also be enhanced by linking stories and facts to the map that highlight scientific history, current advances and relationships between scientific paradigms.



2006

Map of Scientific Paradigms

By Kevin W. Boyack and Richard Klavans

ALBUQUERQUE, NEW MEXICO, AND BERWYN, PENNSYLVANIA, 2006

Courtesy of Kevin W. Boyack and Richard Klavans, SciTech Strategies, Inc.

Aim

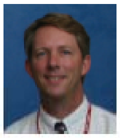
Science can be thought of as containing themes and paradigms; themes are current areas of research, while paradigms comprise the dominant tool sets and existing knowledge that are used by current researchers. What would a paradigm map of science look like? How many paradigms are currently active? How large and how vital are they?

Interpretation

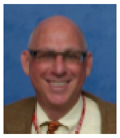
This map was generated by recursively clustering the 820,000 most important papers referenced in 2003 using the processing pipeline described on **page 12, Toward a Reference System for Science.** The result is a map of 776 paradigms, which are shown as circles on the map. Although each paradigm contains an average of 1,000 papers, they range in sizes, as shown by the variously sized circles on the map. The most dominant relationships between paradigms were also calculated and are shown as lines between paradigms. A reference system was added for means of navigation and communication.

Color-coding indicates the vitality of a research topic—the darker the red, the younger the average reference age and the more vital and faster moving the topic. The white circles represent paradigms where consensus is reached relatively slowly. This is a common phenomenon in the social sciences, ecological sciences, computer sciences, and mathematics disciplines. The red circles represent communities of researchers where consensus is reached relatively rapidly. This is more common in physics, chemistry, biochemistry, and many medical disciplines. Very dark circles (such as those in quantum physics) represent communities where consensus is reached most quickly.

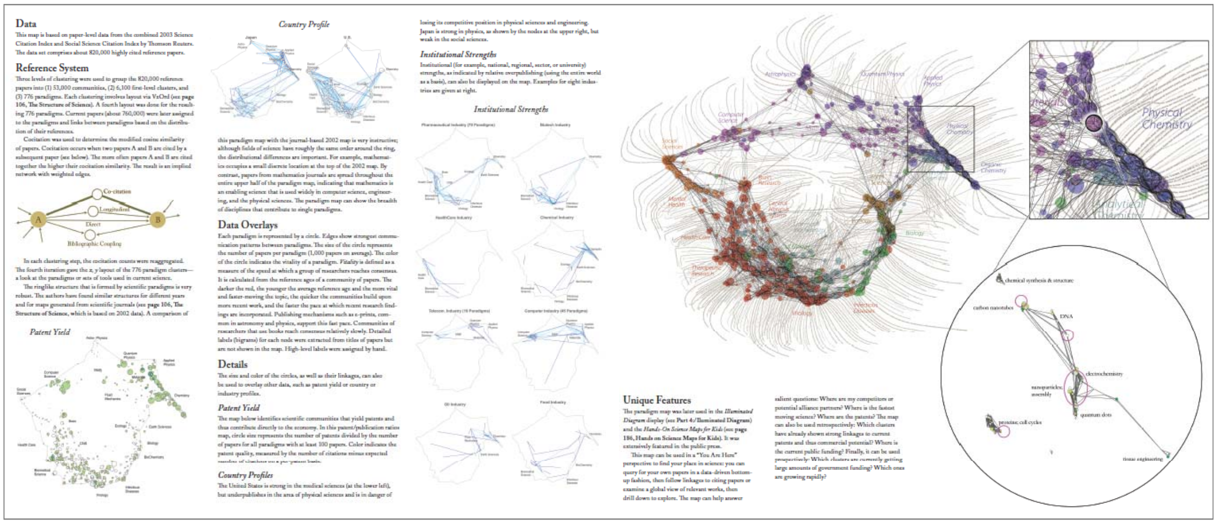
Countries, industries, companies, and individual researchers can all locate themselves within the map, either as single points or as a specific collection of paradigms. Science education and discovery can also be enhanced by linking to the map stories and facts that highlight content and relationships between scientific paradigms.



Kevin W. Boyack joined SciTech Strategies, Inc. in 2007 after working at Sandia National Laboratories, where he spent several years in the Computation, Computers, Information and Mathematics Center. He holds a PhD in chemical engineering from Brigham Young University. His current interests and work are related to information visualization, knowledge domains, science mapping with associated metrics and indicators, network analysis, and the integration and analysis of multiple data types.



Richard Klavans is the president of SciTech Strategies, Inc. He holds a PhD in management from the Wharton School of the University of Pennsylvania. His current work is related to the generation of highly accurate maps of science using multiple techniques, such as bibliographic coupling, cocitation, and coword, as well as the associated metrics and indicators that allow government and industry users to make more effective policy decisions. He is interested in semantics, augmented cognition, and the application of mathematical tools to information spaces.



Data

This map is based on paper-level data from the combined 2003 Science Citation Index and Social Science Citation Index by Thomson Reuters. The data set comprises about 820,000 highly cited reference papers.

Reference System

Three levels of clustering were used to group the 820,000 reference papers into (1) 53,000 communities, (2) 6,100 first-level clusters, and (3) 776 paradigms. Each clustering involves layout via VxOrd (see **page 106, The Structure of Science**). A fourth layout was done for the resulting 776 paradigms. Current papers (about 760,000) were later assigned to the paradigms and links between paradigms based on the distribution of their references.

Cocitation was used to determine the modified cosine similarity of papers. Cocitation occurs when two papers A and B are cited by a subsequent paper (see below). The more often papers A and B are cited together the higher their cocitation similarity. The result is an implied network with weighted edges.

In each clustering step, the cocitation counts were reaggregated. The fourth iteration gave the x, y layout of the 776 paradigm clusters—a look at the paradigms or sets of tools used in current science.

The ringlike structure that is formed by scientific paradigms is very robust. The authors have found similar structures for different years and for maps generated from scientific journals (see **page 106, The Structure of Science**, which is based on 2002 data). A comparison of this paradigm map with the journal-based 2002 map is very instructive; although fields of science have roughly the same order around the ring, the distributional differences are important. For example, mathematics occupies a small discrete location at the top of the 2002 map. By contrast, papers from mathematics journals are spread throughout the entire upper half of the paradigm map, indicating that mathematics is an enabling science that is used widely in computer science, engineering, and the physical sciences. The paradigm map can show the breadth of disciplines that contribute to single paradigms.

Data Overlays

Each paradigm is represented by a circle. Edges show strongest communication patterns between paradigms. The size of the circle represents the number of papers per paradigm (1,000 papers on average). The color of the circle indicates the vitality of a paradigm. *Vitality* is defined as a measure of the speed at which a group of researchers reaches consensus. It is calculated from the reference ages of a community of papers. The darker the red, the younger the average reference age and the more vital and faster-moving the topic, the quicker the communities build upon more recent work, and the faster the pace at which recent research findings are incorporated. Publishing mechanisms such as e-prints, common in astronomy and physics, support this fast pace. Communities of researchers that use books reach consensus relatively slowly. Detailed labels (bigrams) for each node were extracted from titles of papers but are not shown in the map. High-level labels were assigned by hand.

Details

The size and color of the circles, as well as their linkages, can also be used to overlay other data, such as patent yield or country or industry profiles.

Patent Yield

The map below identifies scientific communities that yield patents and thus contribute directly to the economy. In this patent/publication ratios map, circle size represents the number of patents divided by the number of papers for all paradigms with at least 100 papers. Color indicates the patent quality, measured by the number of citations minus expected number of citations on a per-patent basis.

Country Profiles

The United States is strong in the medical sciences (at the lower left), but underpublishes in the area of physical sciences and is in danger of losing its competitive position in physical sciences and engineering. Japan is strong in physics, as shown by the nodes at the upper right, but weak in the social sciences.

Institutional Strengths

Institutional (for example, national, regional, sector, or university) strengths, as indicated by relative overpublishing (using the entire world as a basis), can also be displayed on the map. Examples for eight industries are given at right.

Unique Features

The paradigm map was later used in the *Illuminated Diagram* display (see **Part 4:/Illuminated Diagram**) and the *Hands-On Science Maps for Kids* (see **page 186, Hands on Science Maps for Kids**). It was extensively featured in the public press.

This map can be used in a "You Are Here" perspective to find your place in science: you can query for your own papers in a data-driven bottom-up fashion, then follow linkages to citing papers or examine a global view of relevant works, then drill down to explore. The map can help answer salient questions: Where are my competitors or potential alliance partners? Where is the fastest moving science? Where are the patents? The map can also be used retrospectively: Which clusters have already shown strong linkages to current patents and thus commercial potential? Where is the current public funding? Finally, it can be used prospectively: Which clusters are currently getting large amounts of government funding? Which ones are growing rapidly?

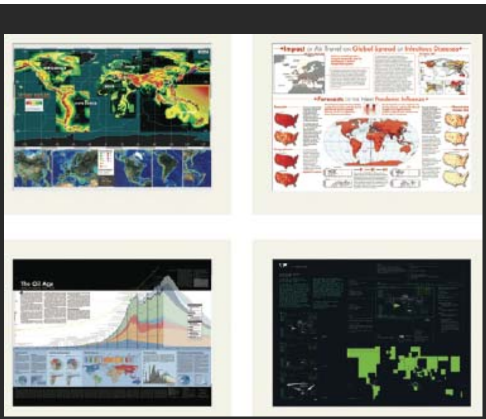
Third Iteration of Exhibit (2007): The Power of Forecasts

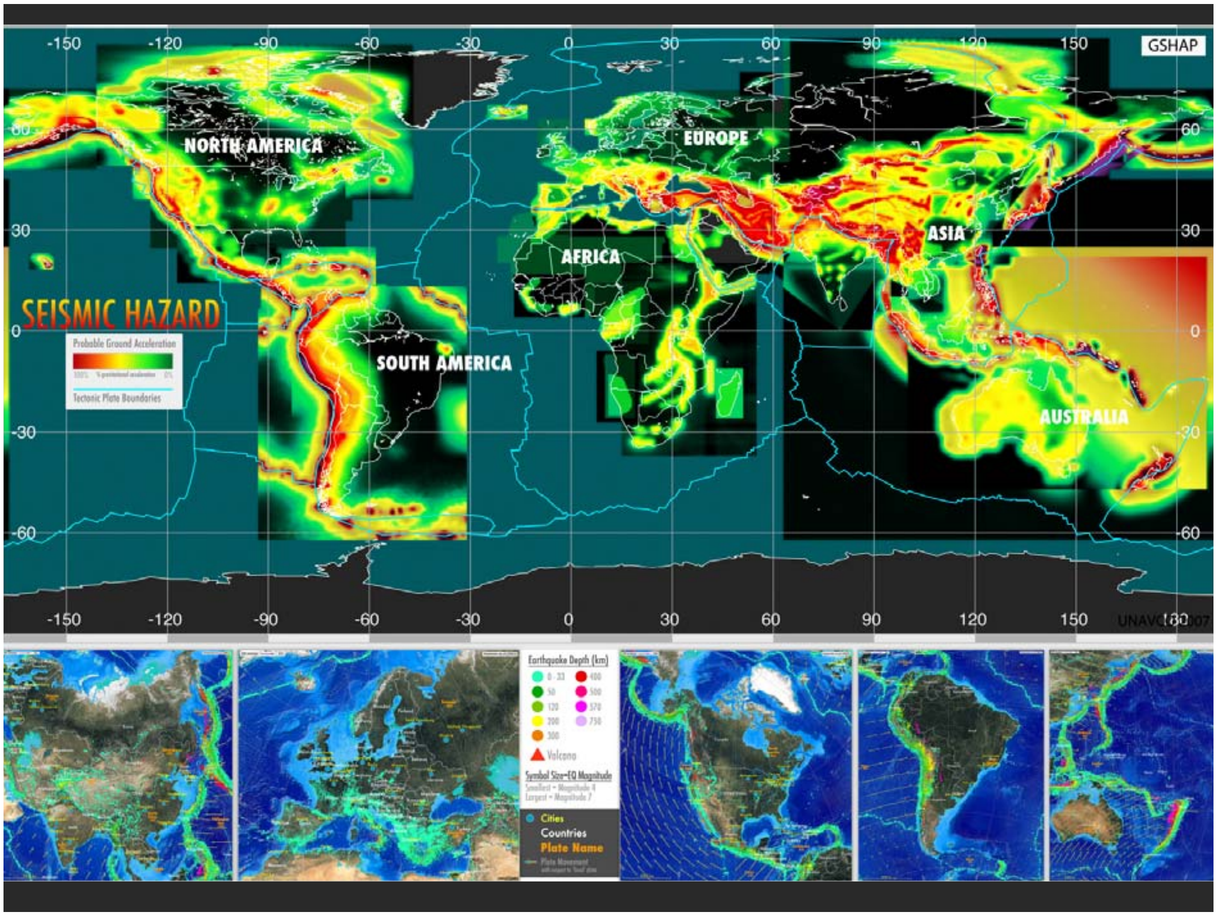
Four Existing Forecasts
Versus
Six Science Forecasts

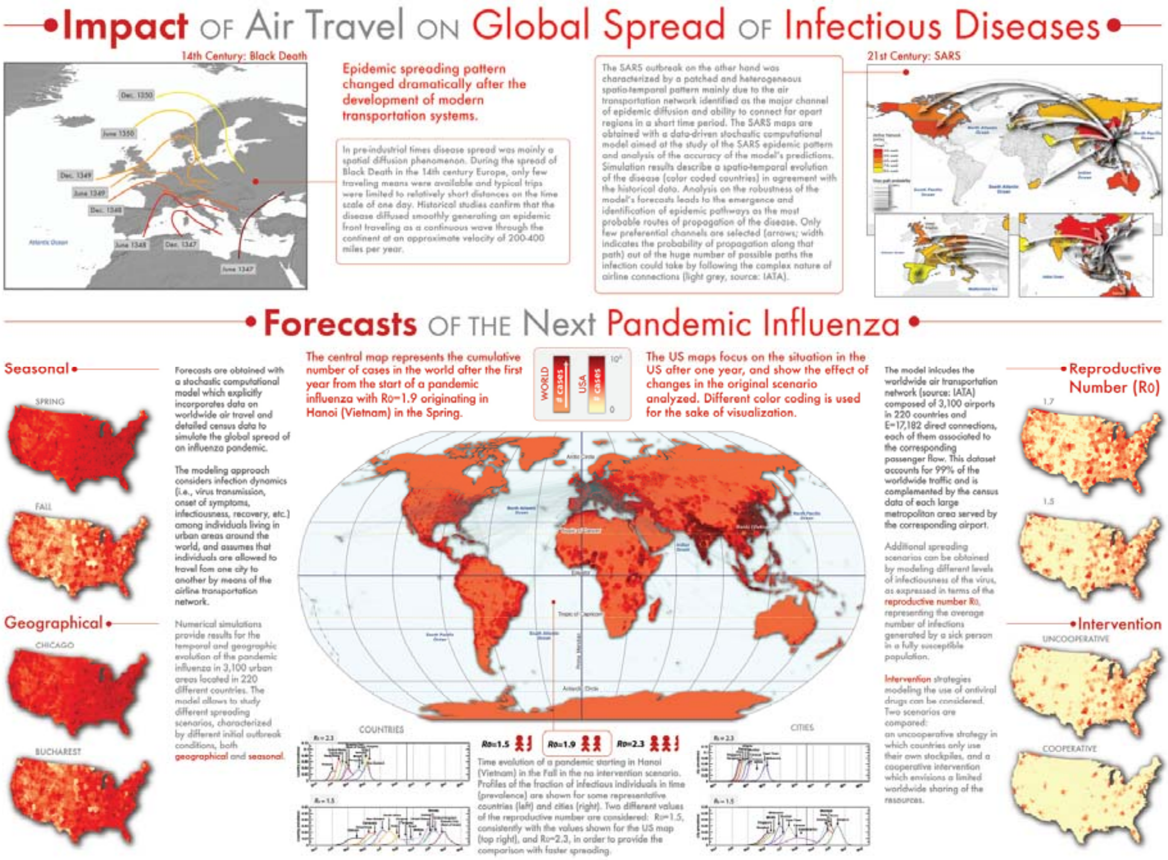
The third iteration of the exhibit compares and contrasts seismic hazard, economic, resource depletion, and epidemic forecast maps with maps forecasting the structure and evolution of science.

Real-time weather forecasts are served by the National Oceanic and Atmospheric Administration (NOAA) or the National Aeronautics and Space Administration (NASA). Computational models of the movements of tectonic plates help reduce losses due to earthqua tsunamis. Epidemic models make us understand and how actions far away affect us right here. Ec catastrophic and sustainable futures for mankind.

Daily science and technology forecasts would s of top experts/institutions/countries, major activit frontiers, augmenting our knowledge and decisio available on TV, in the press, and online?







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.

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).

21st Century: SARS

Forecasts OF THE Next Pandemic Influenza

Seasonal

SPRING

FALL

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 fom one city to another by means of the airline transportation network.

Geographical

CHICAGO

BUCHAREST

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.

R_0=1.5 **R_0=1.9** **R_0=2.3**

Time evolution of a pandemic starting in Hanoi (Vietnam) in the Fall in the no intervention scenario. Profiles of the fraction of infectious individuals in time (prevalence) are shown for some representative countries (left) and cities (right). Two different values of the reproductive number are considered: R_0=1.5, consistently with the values shown for the US map (top right), and R_0=2.3, in order to provide the comparison with faster spreading.

The model inlcudes 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)

1.7

1.5

Intervention

UNCOOPERATIVE

COOPERATIVE

INSTITUTE FOR THE FUTURE

Science & Technology Outlook: 2005–2055

Technology Horizons Program
Institute for the Future
124 University Avenue, 2nd Floor, Palo Alto, CA 94301
t 650.854.6322 f 650.854.7850 www.iftf.org

A map is a tool for navigating an unknown terrain. In the case of this map, **Science & Technology Outlook: 2005–2055**, the terrain we're navigating is the uncharted territory of science and technology (S&T) in the next 50 years. However, the map of the future is not a tool for prediction or, for that matter, the product of predictions. Nor is it comparable to modern navigation techniques in which we rely on a shrinking number of strong signals, like GPS coordinates, to show the right path. Rather, it's more akin to classical low-tech navigational techniques with their reliance on an array of weak signals such as wind direction, the look and feel of the water, and the shape of cloud formations. Taken together, these signals often prove more useful for navigation than high-tech methods because, in addition to aiding travelers in selecting the "right" path, the signals contextualize information and reveal interdependencies and connections between seemingly unrelated events, thus enriching our understanding of the landscape. That's precisely the intention of this map of the future of S&T—to give the reader a deeper contextual understanding of the landscape and to point to the intricacies and interdependencies between trends.

While developing the map, the **Institute for the Future (IFTF)** team listened for and connected a variety of weak signals, including those generated during interviews and workshop conversations involving more than 100 eminent U.K. and U.S. experts in S&T—academicians, policymakers, journalists, and corporate researchers. The IFTF team also compiled a database of outlooks on developments that are likely to impact the full range of S&T disciplines and practice areas over the next 50 years. We also relied on IFTF's 40 years of experience in forecasting S&T developments to create the map and an accompanying set of **S&T Perspectives** that discuss issues emerging on the S&T horizon and are important for organizations, policymakers, and society-at-large to understand.

On this map, six themes are woven together across the 50-year horizon, often resulting in important breakthroughs. These are supported by key technolgies, innovations, and discoveries. In addition to the six themes, three meta-themes—democratized innovation, transdisciplinarity, and emergence—will overlay the future S&T landscape influencing how we think about, learn about, and practice science. Finally, S&T trends won't operate in a vacuum. Wider social, demographic, political, economic, and environmental trends will both influence S&T trends and will be influenced by them. Some of these wider trends surround the map to remind us of the larger picture.

MAP THEMES

Small World

After 20 years of basic research and development at the 100-nanometer scale, the importance of nanotechnology as a source of innovations and new capabilities in everything from materials science to medicine is already well-understood. Three trends, however, will define how nanotechnology will unfold, and what impacts it will have. First, nanotechnology is not a single field with a coherent intellectual program; it's an opportunistic hybrid, shaped by a combination of fundamental research questions, promising technical applications, and venture and state capital. Second, nanotechnology is moving away from the original vision of small-scale mechanical engineering—in which assemblers build mechanical systems from individual atoms—toward one in which molecular biology and biochemistry contribute essential tools (such as proteins that build nanowires). Finally, nanotechnology will also serve as a model for transdisciplinary science. It will support both fundamental research and commercially oriented innovation; and it will be conducted not within the boundaries of conventional academic or corporate research departments, but in institutional and social milieux that emphasize heterogeneity.

Intentional Biology

For 3.6 billion years, evolution has governed biology on this planet. But today, Mother Nature has a collaborator. Inexpensive tools to read and rewrite the genetic code of life will bootstrap our ability to manipulate biology from the bottom up. We'll not only genetically re-engineer existing life but actually create new life forms with purpose. Still, we will not be blind to what nature has to teach us. Evolution's elegant engineering at the smallest scales will be a rich source of inspiration as we build the bio-nanotechnology of the next 50 years.

Extended Self

In the next 50 years, we will be faced with broad opportunities to remake our minds and bodies in profoundly different ways. Advances in biotechnology, brain science, information technology, and robotics will result in an array of methods to dramatically alter, enhance, and extend the mental and physical hand that nature has dealt us. Wielding these tools on ourselves, humans will begin to define a variety of different "transhumanist" paths—that is, ways of being and living that extend beyond what we today consider natural for our species. In the very long term, following these paths could someday lead to an evolutionary leap for humanity.

Mathematical World

The ability to process, manipulate, and ultimately understand patterns in enormous amounts of data will allow decoding of previously mysterious processes in everything from biological to social systems. Scientists are learning that at the core of many biological phenomena—reproduction, growth, repair, and others—are computational processes that can be decoded and simulated. Using techniques of combinatorial science to uncover such patterns—whether these are physical, biological, or social—will likely occupy an increasing share of computing cycles in the next 50 years. Such massive computation will also make simulation widespread. Computer simulation will be used not only to help make decisions about large complex scientific and social problems but also to help individuals make better choices in their daily lives.

Sensory Transformation

In the next ten years, physical objects, places, and even human beings themselves will increasingly become embedded with computational devices that can sense, understand, and act upon their environment. They will be able to react to contextual clues about the physical, social, and even emotional state of people and things in their surroundings. As a result, increasing demands will be placed on our visual, auditory, and other sensory abilities. Information previously encoded as text and numbers will be displayed in richer sensory formats—as graphics, pictures, patterns, sounds, smells, and tactile experiences. This enriched sensory environment will coincide with major breakthroughs in our understanding of the brain—in how we process sensory information and connect various sensory functions. Humans will become much more sophisticated in their ability to understand, create, and manage sensory information and ability to perform such tasks will become keys to success.

Lightweight Infrastructure

A confluence of new materials and distributed intelligence is pointing the way toward a new kind of infrastructure that will dramatically reshape the economics of moving people, goods, energy, and information. From the molecular level to the macroeconomic level, these new infrastructure designs will emphasize smaller, smarter, more independent components. These components will be organized into more efficient, more flexible, and more secure ways than the capital-intensive networks of the 20th century. These lightweight infrastructures have the potential to boost emerging economies, improve social connectivity, mitigate the environmental impacts of rapid global urbanization, and offer new future paths in energy.

META-THEMES

Democratized Innovation

Before the 20th century, many of the greatest scientific discoveries and technical inventions were made by amateur scientists and independent inventors. In the last 100 years, a professional class of scientists and engineers, supported by universities, industry, and the state, pushed amateurs aside as a creative force. At the national scale, the capital-intensive character of scientific research made world-class research the property of prosperous advanced nations. In the new century, a number of trends and technologies will lower the barriers to participation in science and technology again, both for individuals and for emerging countries. The result with be a renaissance of the serious amateurs, the growth of new scientific and technical centres of excellence in developing countries, and a more global distribution of world-class scientists and technologists.

Transdisciplinarity

In the last two centuries, natural philosophy and natural history fractured into the now-familiar disciplines of physics, chemistry, biology, and so on. The sciences evolved into their current form in response to intellectual and professional opportunities, philanthropic priorities, and economic and state needs. Through most of the 20th century, the growth of the sciences, and academic and career pressures, encouraged ever-greater specialization. In the coming decades, transdisciplinary research will become an imperative. According to Howard Rheingold, a prominent forecaster and author, "transdisciplinarity goes beyond bringing together researchers from different disciplines to work in multidisciplinary teams. It means educating researchers who can speak languages of multiple disciplines—biologists who have understanding of mathematics, mathematicians who understand biology."

Emergence

The phenomenon of self-organizing swarms that generate complex behavior by following simple rules—will likely become an important research area, and an important model for understanding how the natural world works and how artificial worlds can be designed. Emergent phenomena have been observed across a variety of natural phenomena, from physics to biology to sociology. The concept has broad appeal due to the diversity of fields and problems to which it can be applied. It is proving useful for making sense of a very wide range of phenomena. Meanwhile, emergence can be modeled using relatively simple computational tools, although those models often require substantial processing power. More generally, it is a richly suggestive as a way of thinking about designing complex, robust technological systems. Finally, emergence is an accessible and vivid a metaphor for understanding nature. Just as classical physics profited from popular treatments of Newtonian mechanics, so too will scientific study and technical reproductions of emergent phenomena likely draw benefits from the popularization of its underlying concepts.

---

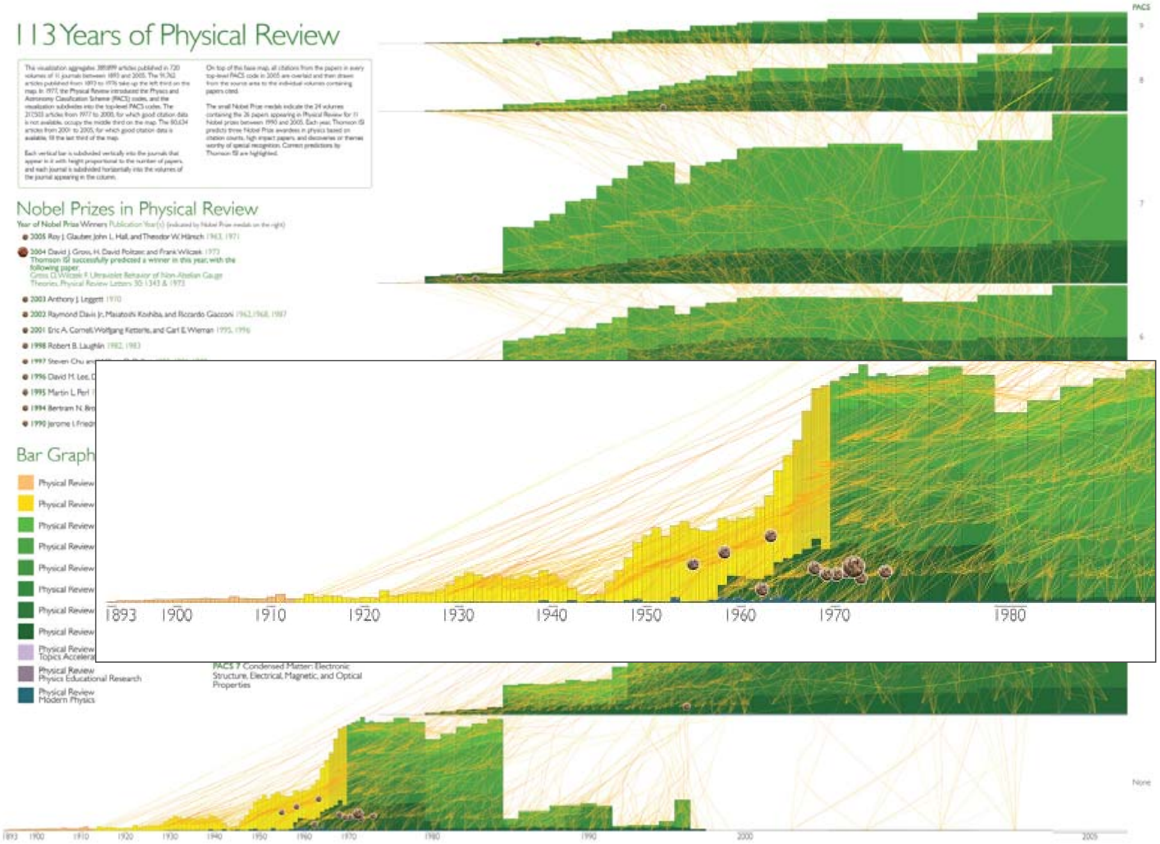
113 Years of Physical Review

This visualization aggregates 389,899 articles published in 720 volumes of 11 journals between 1893 and 2005. The 91,762 articles published from 1893 to 1976 take up the left third on the map. In 1977, the Physical Review introduced the Physics and Astronomy Classification Scheme (PACS) codes, and the visualization subdivides into the top-level PACS codes. The 217,503 articles from 1977 to 2000, for which good citation data is not available, occupy the middle third on the map. The 80,634 articles from 2001 to 2005, for which good citation data is available, fill the last third of the map.

Each vertical bar is subdivided vertically into the journals that appear in it with height proportional to the number of papers, and each journal is subdivided horizontally into the volumes of the journal appearing in the column.

On top of this base map, all citations from the papers in every top-level PACS code in 2005 are overlaid and then drawn from the source area to the individual volumes containing papers cited.

The small Nobel Prize medals indicate the 24 volumes containing the 26 papers appearing in Physical Review for 11 Nobel prizes between 1990 and 2005. Each year, Thomson ISI predicts three Nobel Prize awardees in physics based on citation counts, high impact papers, and discoveries or themes worthy of special recognition. Correct predictions by Thomson ISI are highlighted.



Nobel Prizes in Physical Review

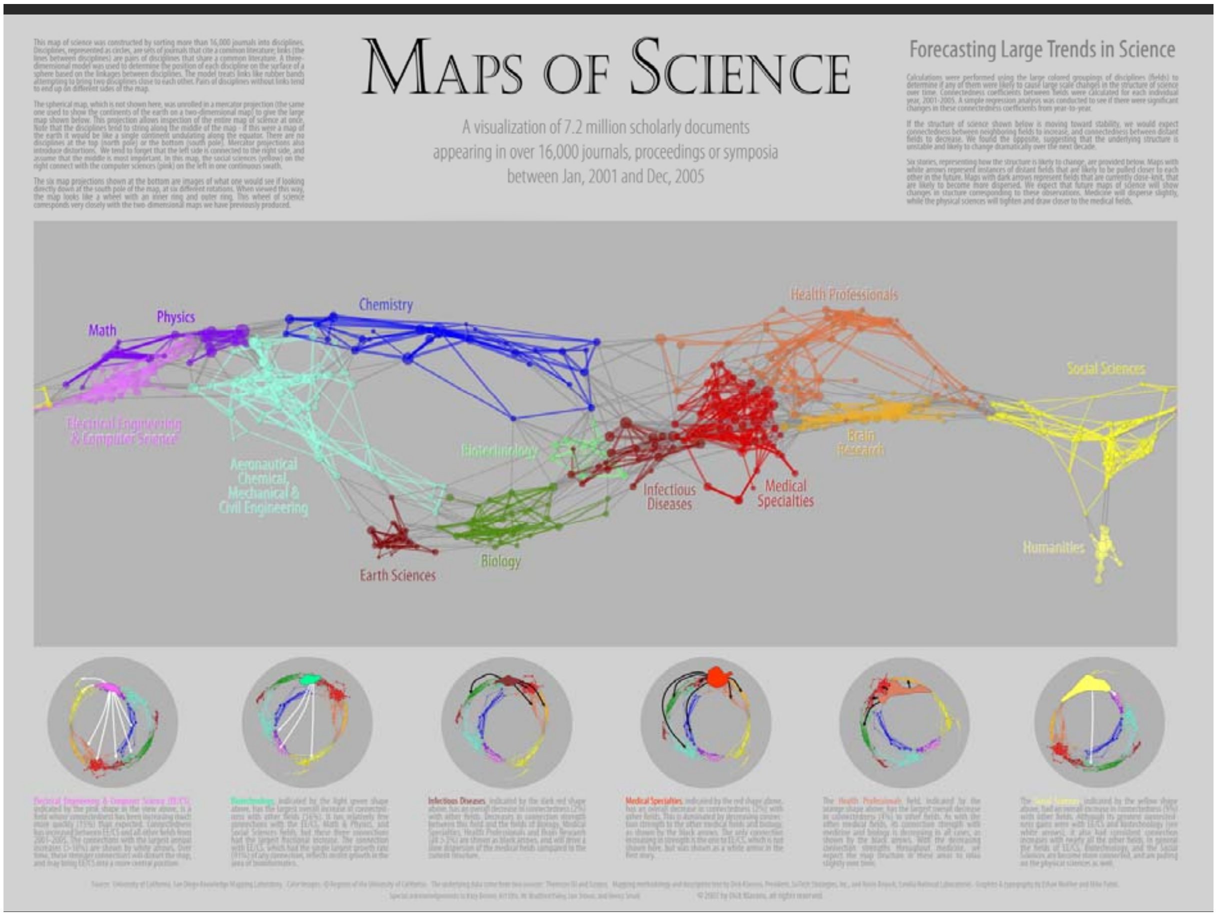
Year of Nobel Prize Winners Publication Year(s) (indicated by Nobel Prize medals on the right)

- 2005 Roy J. Glauber, John L. Hall, and Theodor W. Hänsch 1963, 1971
- 2004 David J. Gross, H. David Politzer, and Frank Wilczek 1973
  Thomson ISI successfully predicted a winner in this year, with the following paper:
  Gross D, Wilczek F. Ultraviolet Behavior of Non-Abelian Gauge Theories. Physical Review Letters 30: 1343 & 1973
- 2003 Anthony J. Leggett 1970
- 2002 Raymond Davis Jr., Masatoshi Koshiba, and Riccardo Giacconi 1962, 1968, 1987
- 2001 Eric A. Cornell, Wolfgang Ketterle, and Carl E. Wieman 1995, 1996
- 1998 Robert B. Laughlin 1982, 1983
- 1997 Steven Chu an
- 1996 David M. Lee, D
- 1995 Martin L. Perl
- 1994 Bertram N. Bro
- 1990 Jerome I. Fried

Bar Graph

- 🟧 Physical Review
- 🟨 Physical Review
- 🟩 Physical Review
- 🟩 Physical Review
- 🟩 Physical Review
- 🟩 Physical Review
- 🟩 Physical Review
- 🟩 Physical Review
- 🟪 Physical Review Topics Accelera
- 🟪 Physical Review Physics Educational Research
- 🟦 Physical Review Modern Physics

PACS 7 Condensed Matter: Electronic Structure, Electrical, Magnetic, and Optical Properties



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

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 (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 string along the middle of the map - if this were a map of the earth it would be like a single continent undulating along the equator. There are no 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 swath.

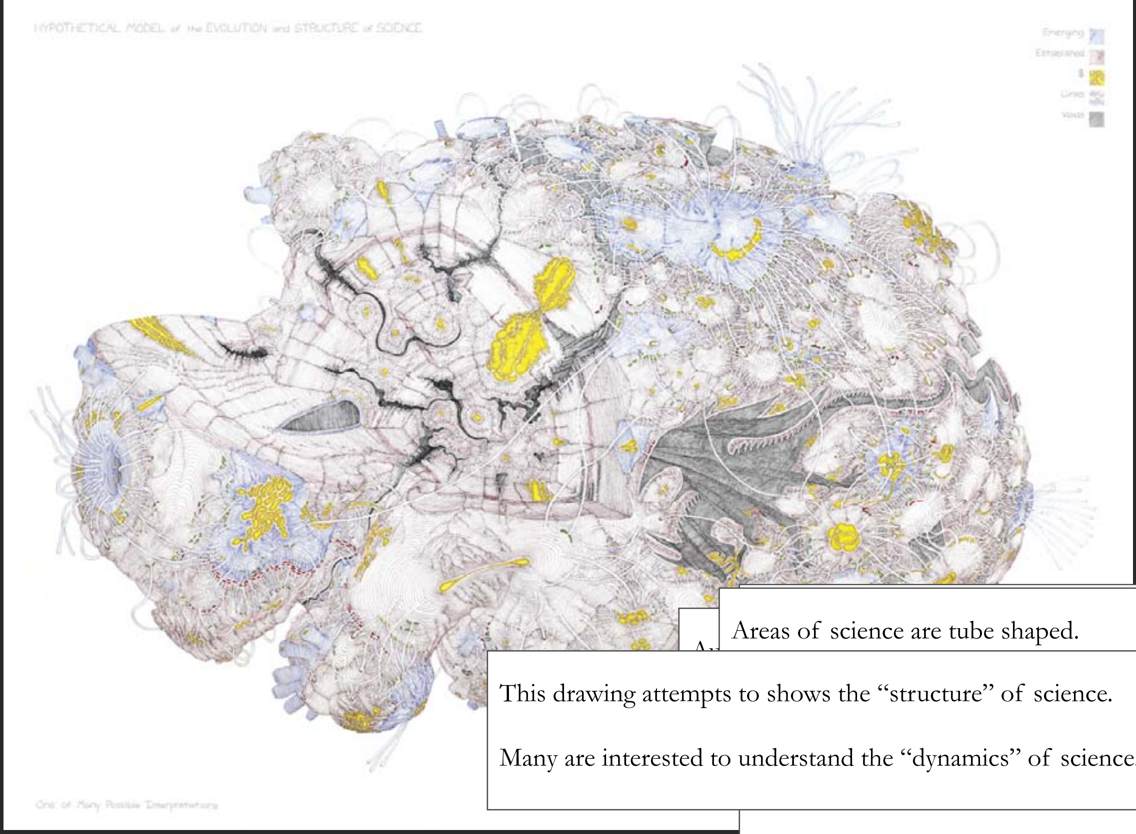
The six map projections shown at the bottom are images of what one would see if looking directly down at the south pole of the map, at six different rotations. When viewed this way, the map looks like a wheel with an inner ring and outer ring. This wheel of science corresponds very closely with the two-dimensional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

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 shown below is moving toward stability, we would expect connectedness 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.

Six stories, 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-knit, that are likely to become more dispersed. We expect that future maps of science will show changes in stucture corresponding to these observations. Medicine will disperse slightly, while the physical sciences will tighten and draw closer to the medical fields.

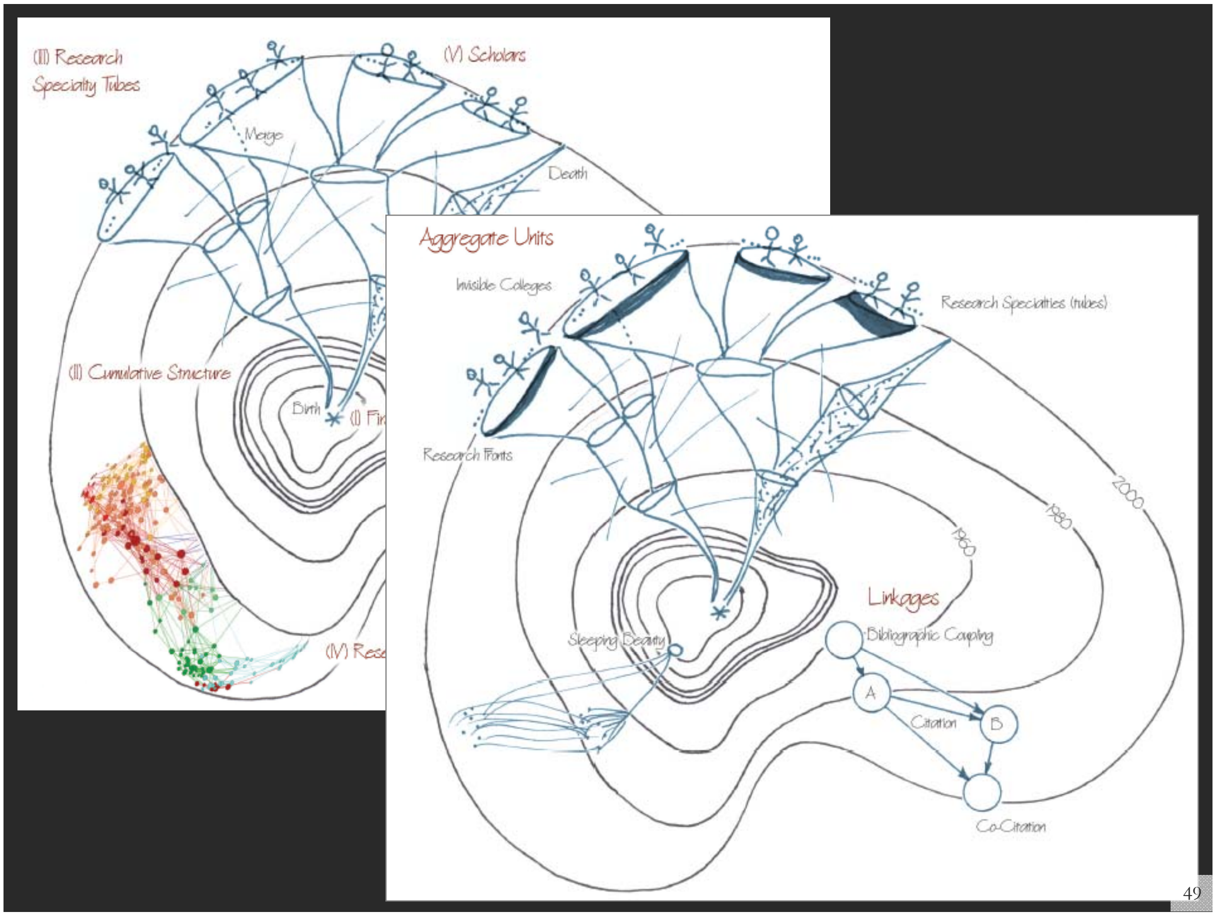


Areas of science are tube shaped.

This drawing attempts to shows the “structure” of science.

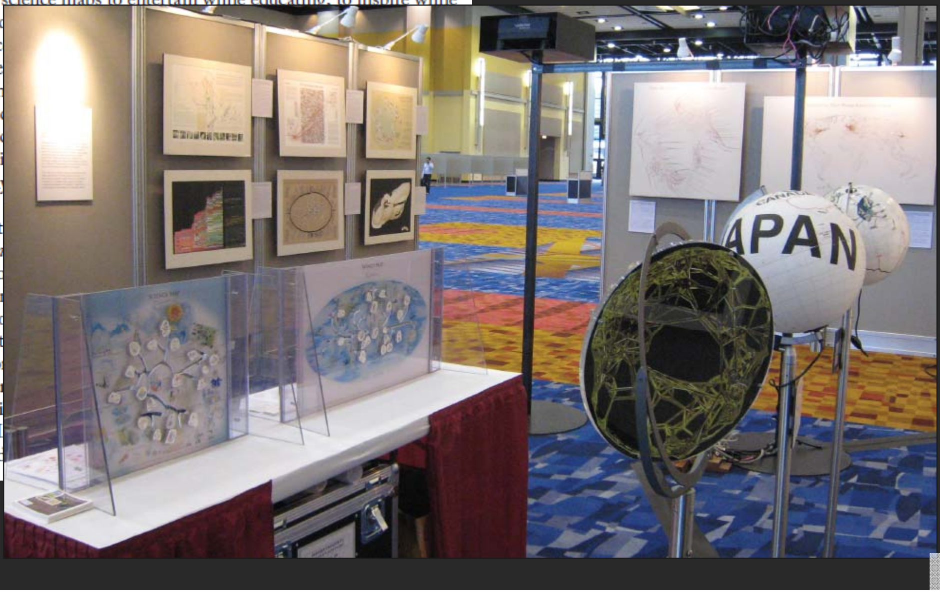
Many are interested to understand the “dynamics” of science.

Hypothetical Model of the Evolution of Science - Daniel Zeller - 2007



Additional Elements of the Exhibit

Certainly science maps and data graphs work to engage viewers intellectually—
but can they also capture the imagination, as did the early maps of the world? Is
it possible to involve viewers in a more dynamic way that heightens both their
awareness and appreciation of data, information, and knowledge? What can be
learned from theater, movies, and art exhibits—as well as science displays—to
improve the ability of science maps to entertain while educating, to inspire while
being true to facts, and
ers to engage in scienc
Additional exhibit e
and interact with scien
exceptional high data
and a map of today's s
drives a touch panel di
the touch panel display
on any given topic are
given geographic locat
The *Hands-On Scien*
stand science from abo
color drawings. Childr
placing images of maj
appropriate places on t
of various countries fo
patents. **Shape of Scie**
The **Video of the Exhi**
Public Library (NYPI
NYPL officials, who c



Illuminated Diagram Display

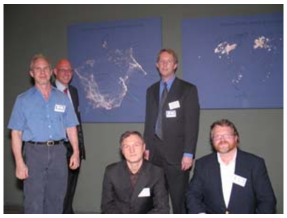
W. Bradford Paley, Kevin W. Boyack, Richard Kalvans, and Katy Börner (2007) Mapping, Illuminating, and Interacting with Science. SIGGRAPH 2007.

Questions:

- Who is doing research on what topic and where?
- What is the 'footprint' of interdisciplinary research fields?
- What impact have scientists?

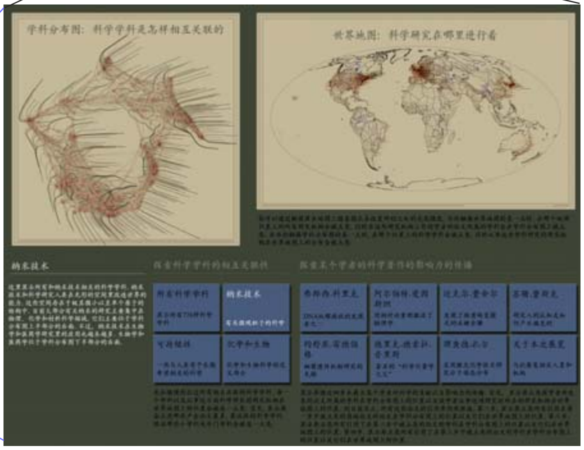
Contributions:

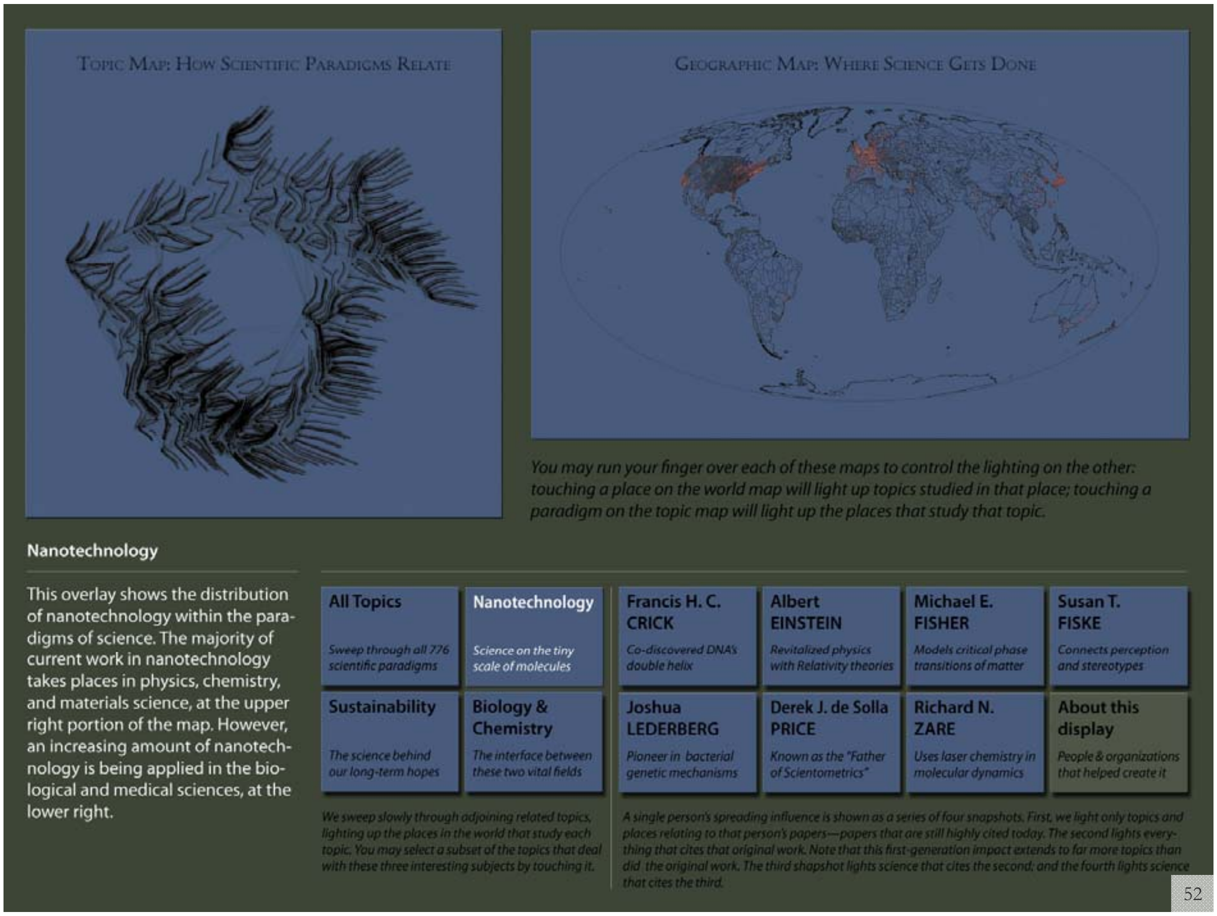
- Interactive, high resolution interface to access and make sense of data about scholarly activity.



Large-scale, high resolution prints illuminated via projector or screen.

Interactive touch panel.





Topic Map: How Scientific Paradigms Relate

Geographic Map: Where Science Gets Done

You may run your finger over each of these maps to control the lighting on the other: touching a place on the world map will light up topics studied in that place; touching a paradigm on the topic map will light up the places that study that topic.

Nanotechnology

This overlay shows the distribution of nanotechnology within the paradigms of science. The majority of current work in nanotechnology takes places in physics, chemistry, and materials science, at the upper right portion of the map. However, an increasing amount of nanotechnology is being applied in the biological and medical sciences, at the lower right.

All Topics *Sweep through all 776 scientific paradigms*

Nanotechnology *Science on the tiny scale of molecules*

Sustainability *The science behind our long-term hopes*

Biology & Chemistry *The interface between these two vital fields*

We sweep slowly through adjoining related topics, lighting up the places in the world that study each topic. You may select a subset of the topics that deal with these three interesting subjects by touching it.

Francis H. C. CRICK *Co-discovered DNA's double helix*

Albert EINSTEIN *Revitalized physics with Relativity theories*

Michael E. FISHER *Models critical phase transitions of matter*

Susan T. FISKE *Connects perception and stereotypes*

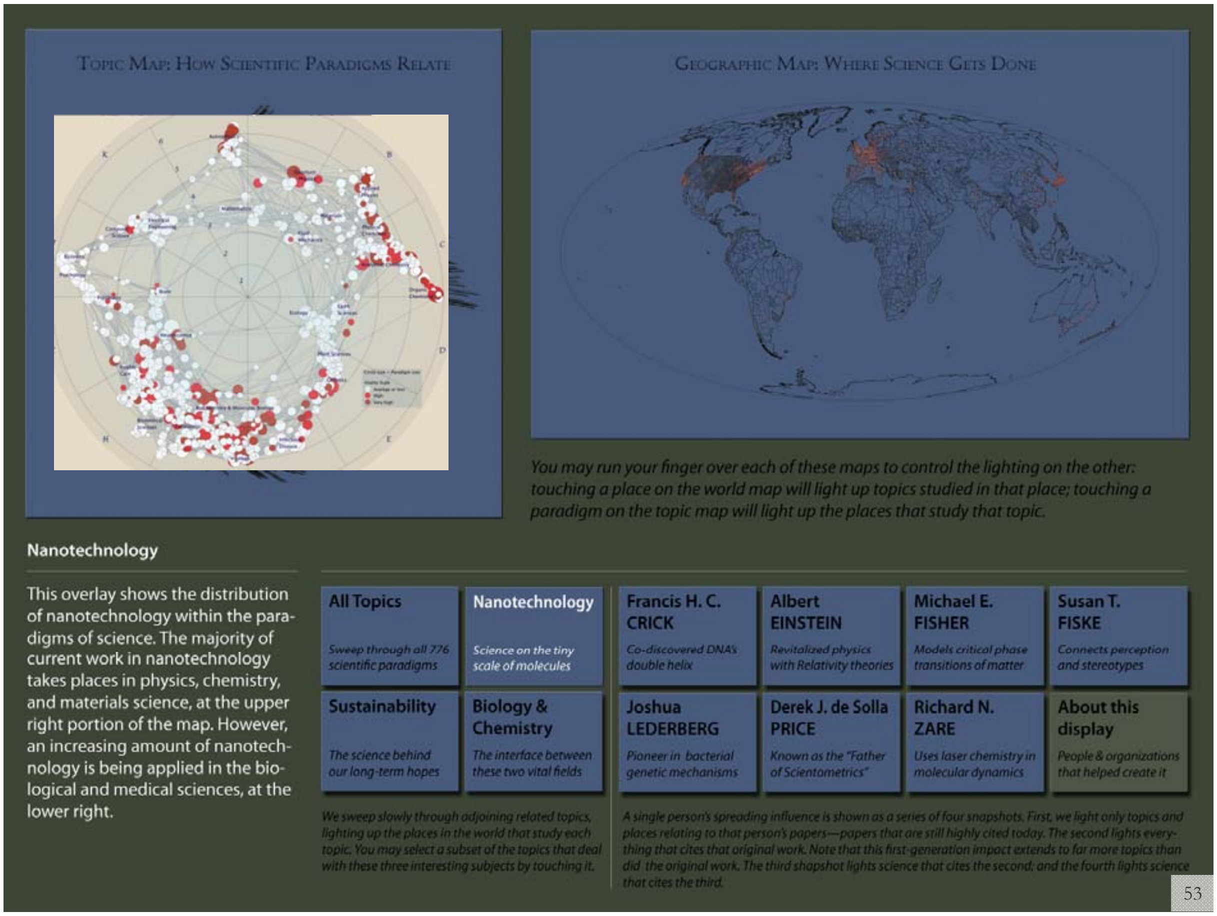
Joshua LEDERBERG *Pioneer in bacterial genetic mechanisms*

Derek J. de Solla PRICE *Known as the "Father of Scientometrics"*

Richard N. ZARE *Uses laser chemistry in molecular dynamics*

About this display *People & organizations that helped create it*

A single person's spreading influence is shown as a series of four snapshots. First, we light only topics and places relating to that person's papers—papers that are still highly cited today. The second lights everything that cites that original work. Note that this first-generation impact extends to far more topics than did the original work. The third shapshot lights science that cites the second; and the fourth lights science that cites the third.



Topic Map: How Scientific Paradigms Relate

Geographic Map: Where Science Gets Done

You may run your finger over each of these maps to control the lighting on the other: touching a place on the world map will light up topics studied in that place; touching a paradigm on the topic map will light up the places that study that topic.

Nanotechnology

This overlay shows the distribution of nanotechnology within the paradigms of science. The majority of current work in nanotechnology takes places in physics, chemistry, and materials science, at the upper right portion of the map. However, an increasing amount of nanotechnology is being applied in the biological and medical sciences, at the lower right.

All Topics
Sweep through all 776 scientific paradigms

Nanotechnology
Science on the tiny scale of molecules

Sustainability
The science behind our long-term hopes

Biology & Chemistry
The interface between these two vital fields

We sweep slowly through adjoining related topics, lighting up the places in the world that study each topic. You may select a subset of the topics that deal with these three interesting subjects by touching it.

Francis H. C. CRICK
Co-discovered DNA's double helix

Albert EINSTEIN
Revitalized physics with Relativity theories

Michael E. FISHER
Models critical phase transitions of matter

Susan T. FISKE
Connects perception and stereotypes

Joshua LEDERBERG
Pioneer in bacterial genetic mechanisms

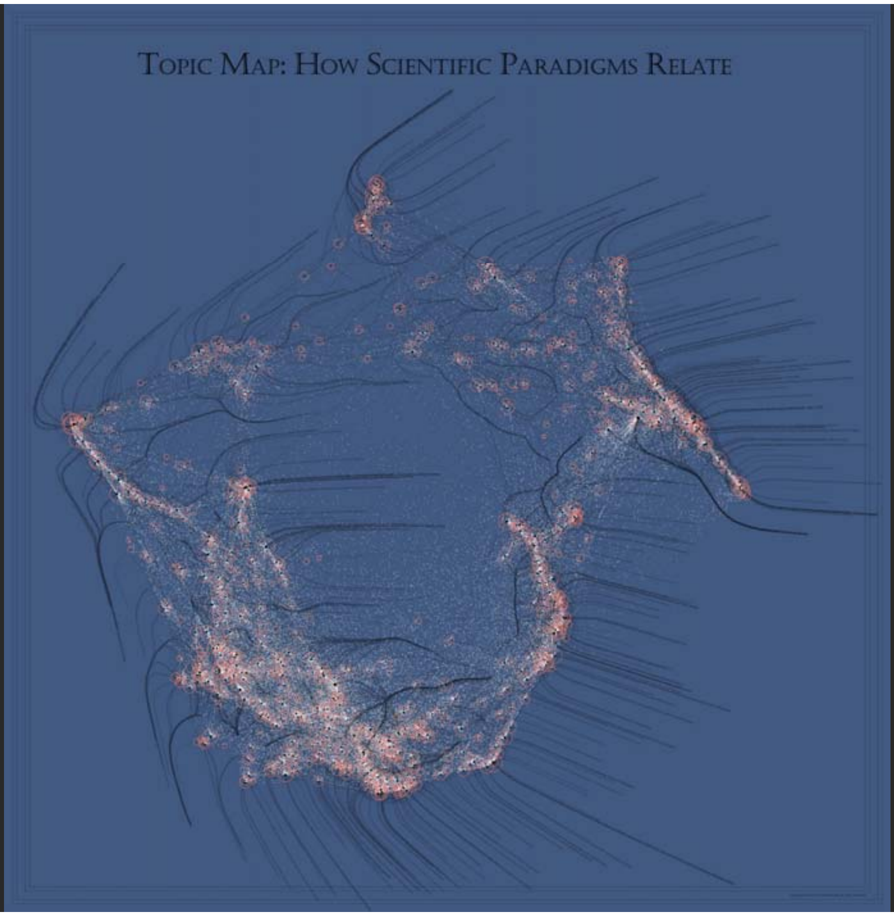
Derek J. de Solla PRICE
Known as the "Father of Scientometrics"

Richard N. ZARE
Uses laser chemistry in molecular dynamics

About this display
People & organizations that helped create it

A single person's spreading influence is shown as a series of four snapshots. First, we light only topics and places relating to that person's papers—papers that are still highly cited today. The second lights everything that cites that original work. Note that this first-generation impact extends to far more topics than did the original work. The third shapshot lights science that cites the second; and the fourth lights science that cites the third.

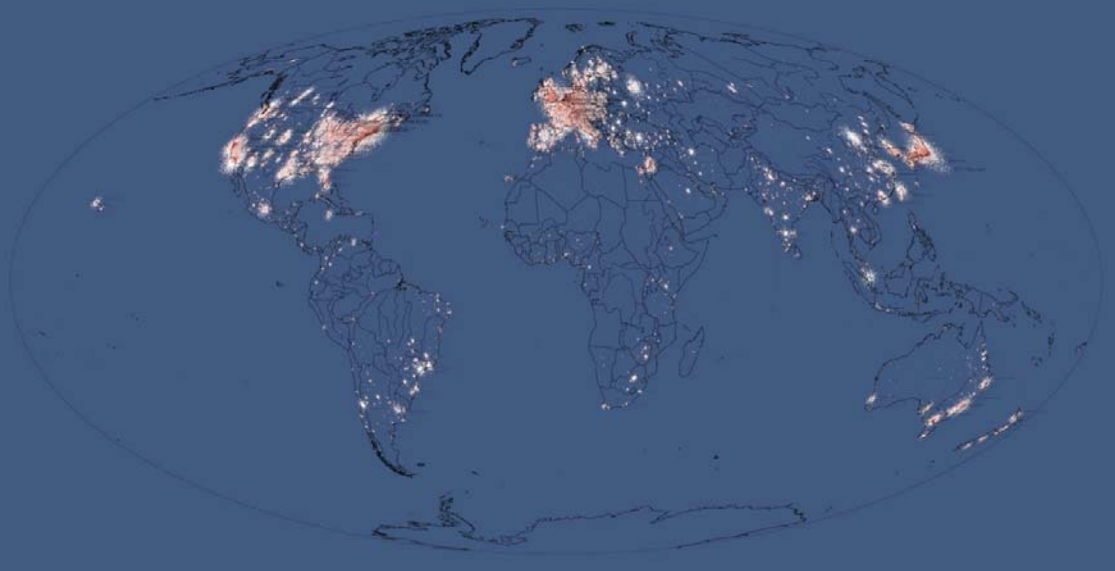
Science Maps in "Expedition Zukunft" science train visiting 62 cities in 7 months 12 coaches, 300 m long
Opening was on April 23rd, 2009 by German Chancellor Merkel
http://www.expedition-zukunft.de



Hands-On Science Maps for Kids, by Fileve Palmer (Painting), Julie Smith (Data Acquisition), Elisha Hardy and Katy Börner (Graphic Design), BLOOMINGTON, IN, 2006. Courtesy of Indiana University. Learn more at www.scimaps.org. This map plots the locations of where scientific papers were published: each light green dot represents 10 or fewer papers; they are scattered around the exact location for visibility, within a labelled green circle whose size is proportional to the number of papers published in that place. The base map is part of an "illuminated diagram" display which used a computer and two projectors, projecting spots of light on the prints to highlight different kinds of scientific research (on a sibling map of scientific paradigms) and the areas in the world where such

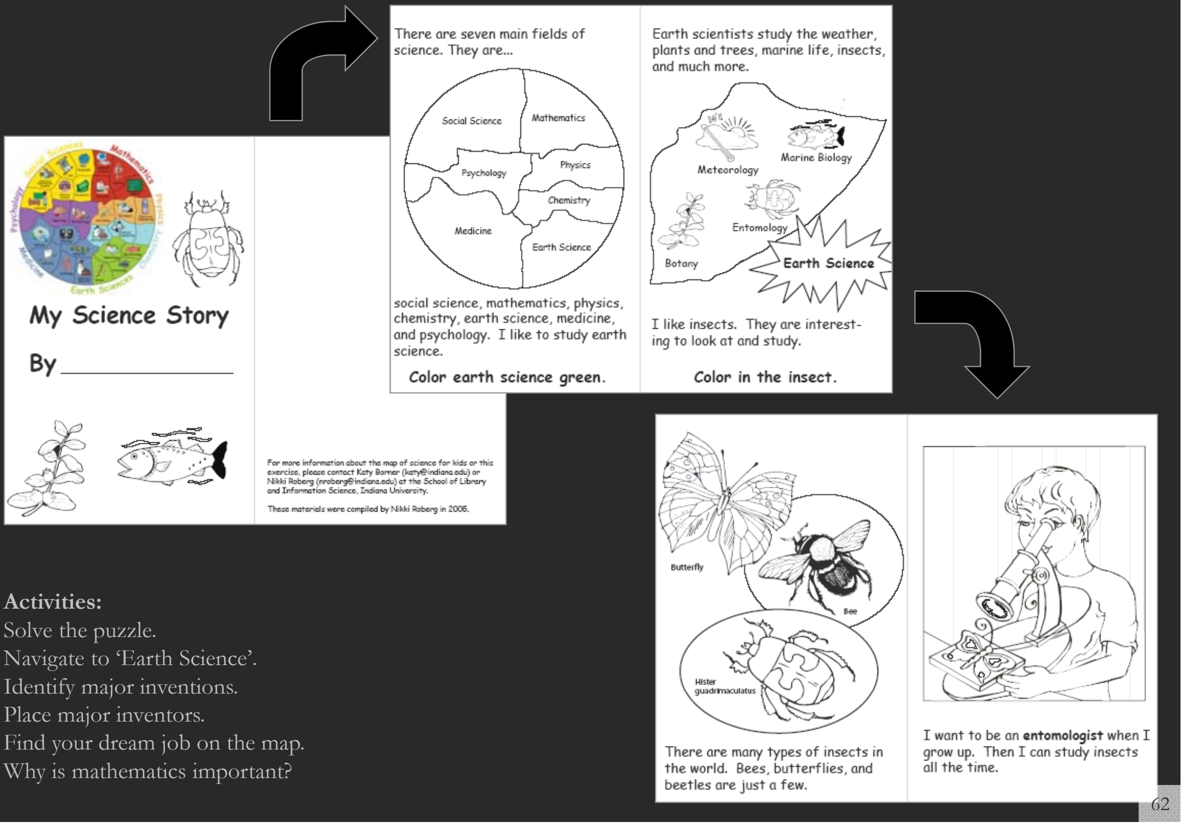
Hands-On Science Maps for Kids, by Fileve Palmer (Painting), Julie Smith (Data Acquisition), Elisha Hardy and Katy Börner (Graphic Design), BLOOMINGTON, IN, 2006. Courtesy of Indiana University. Learn more at www.scimaps.org. This map plots the locations of where scientific papers were published: each light green dot represents 10 or fewer papers; they are scattered around the exact location for visibility, within a labelled green circle whose size is proportional to the number of papers published in that place. The base map is part of an "illuminated diagram" display which used a computer and two projectors, projecting spots of light on the prints to highlight different kinds of scientific research (on a sibling map of scientific paradigms) and the areas in the world where such science was performed. Base map research by Kevin Boyack and Dick Klavans, cartography by John Burgoon, data from Thompson ISI, graphics and typography by W. Bradford Paley. Copyright © 2006 W. Bradford Paley, all rights reserved.

GEOGRAPHIC MAP: WHERE SCIENCE GETS DONE



Science Puzzle Map for Kids by Fileve Palmer, Julie Smith, Elisha Hardy and Katy Börner, Indiana University, 2006.
(Base map taken from Illuminated Diagram display by Kevin Boyack, Richard Klavans, and W. Bradford Paley.)

Hands-On Science Maps for Kids, by Fileve Palmer (Painting), Julie Smith (Data Acquisition), Elisha Hardy and Katy Börner (Graphic Design), BLOOMINGTON, IN, 2006, Courtesy of Indiana University. Learn more at www.scimaps.org. This map plots the locations of where scientific papers were published: each light green dot represents 10 or fewer papers; they are scattered around the exact location for visibility, within a labelled green circle whose size is proportional to the number of papers published in that place. The base map is part of an "Illuminated diagram" display which used a computer and two projectors, projecting spots of light on the prints to highlight different kinds of scientific research (on a sibling map of scientific paradigms) and the areas in the world where such science was performed. Base map research by Kevin Boyack and Dick Klavans, cartography by John Burgoon, data from Thompson ISI, graphics and typography by W. Bradford Paley. Copyright © 2006 W. Bradford Paley, all rights reserved.



My Science Story

By ______________

For more information about the map of science for kids or this exercise, please contact Katy Borner (katy@indiana.edu) or Nikki Roberg (nroberg@indiana.edu) at the School of Library and Information Science, Indiana University.

These materials were compiled by Nikki Roberg in 2006.

There are seven main fields of science. They are...

Social Science
Mathematics
Psychology
Physics
Chemistry
Medicine
Earth Science

social science, mathematics, physics, chemistry, earth science, medicine, and psychology. I like to study earth science.

Color earth science green.

Earth scientists study the weather, plants and trees, marine life, insects, and much more.

Meteorology
Marine Biology
Entomology
Botany
Earth Science

I like insects. They are interesting to look at and study.

Color in the insect.

Butterfly
Bee
Hister quadrimaculatus

There are many types of insects in the world. Bees, butterflies, and beetles are just a few.

I want to be an **entomologist** when I grow up. Then I can study insects all the time.

Activities:
Solve the puzzle.
Navigate to 'Earth Science'.
Identify major inventions.
Place major inventors.
Find your dream job on the map.
Why is mathematics important?

Part 5: The Future of Science Maps

The inspiration of a noble cause involving human interests wide and far, enables men to do things they did not dream themselves capable of before, and which they were not capable of alone. The consciousness of belonging, vitally, to something beyond individuality; of being part of a personality that reaches we know not where, in space and time, greatens the heart to the limit of the soul's ideal, and builds out the supreme of character.

Joshua L. Chamberlain

Part 5: The Future of Science Maps

Mapping Science Exhibit – 10 Iterations in 10 years

http://scimaps.org/

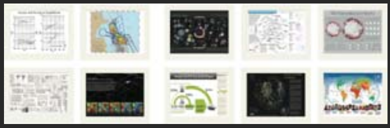
The Power of Maps (2005)

The Power of Reference Systems (2006)

The Power of Forecasts (2007)

Science Maps for Economic Decision Makers (2008)

Science Maps for Science Policy Makers (2009)



Science Maps for Scholars (2010)
Science Maps as Visual Interfaces to Digital Libraries (2011)
Science Maps for Kids (2012)
Science Forecasts (2013)
How to Lie with Science Maps (2014)

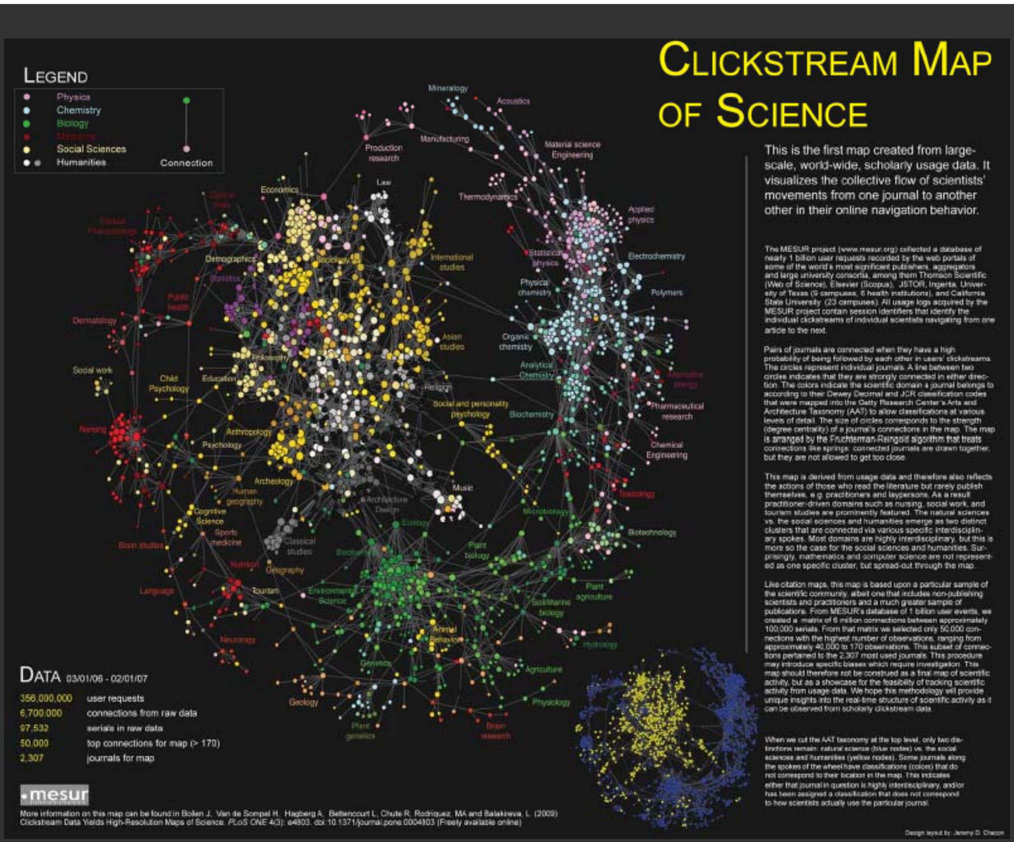
Exhibit has been shown in 72 venues on four continents. Currently at
- NSF, 10th Floor, 4201 Wilson Boulevard, Arlington, VA
- Center of Advanced European Studies and Research, Bonn, Germany
- Science Train, Germany
- Cultural Dimensions of Innovation, UCD Conference, Dublin, Ireland

Debut of 5th Iteration of Mapping Science Exhibit at MEDIA X was on May 18, 2009 at Wallenberg Hall, Stanford University, http://mediax.stanford.edu, http://scaleindependentthought.typepad.com/photos/scimaps

Part 5: The Future of Science Maps

This is the only mockup in this slide show.

Everything else is available today.



Clickstream Map of Science

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 other 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 (9 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. This 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 science (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 that 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

mesur

More information on this map can be found in Bollen J., Van de Sompel H., Hagberg A., Bettencourt L, Chute R, Rodriquez, 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

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.

Interactive Maps of Science – Philanthropy

About this Map

- 7,908 Grantmakers
- 93,091 Recipients
- 418,555 Grants

Data View

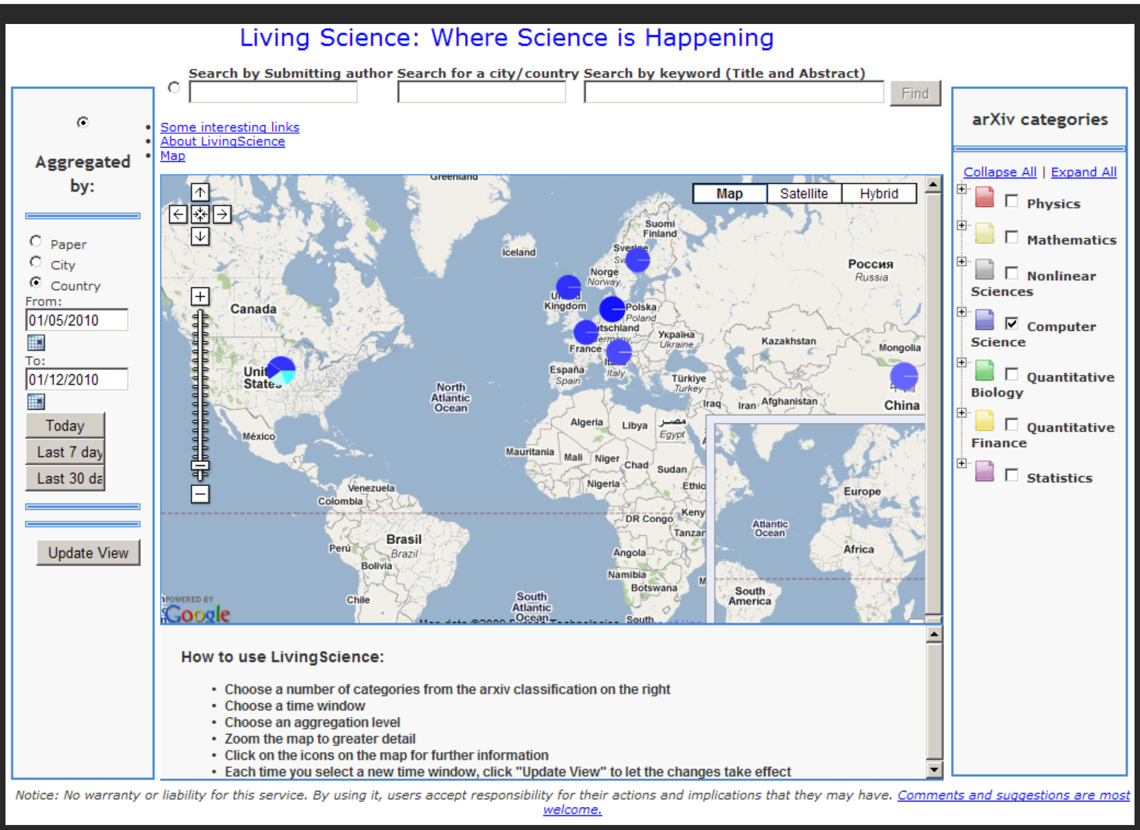
Current Parameters

Geographic Display: State
Grant Indicator: None
Thematic Indicator: None
Grant Years: 2009, 2008, 2007
Grantmaker Filters: None

Print/Save Map
Copy Map

Grantmaker/Grant data ©2009 The Foundation Center

http://www.philanthropyinsight.org



Living Science: Where Science is Happening

How to use LivingScience:

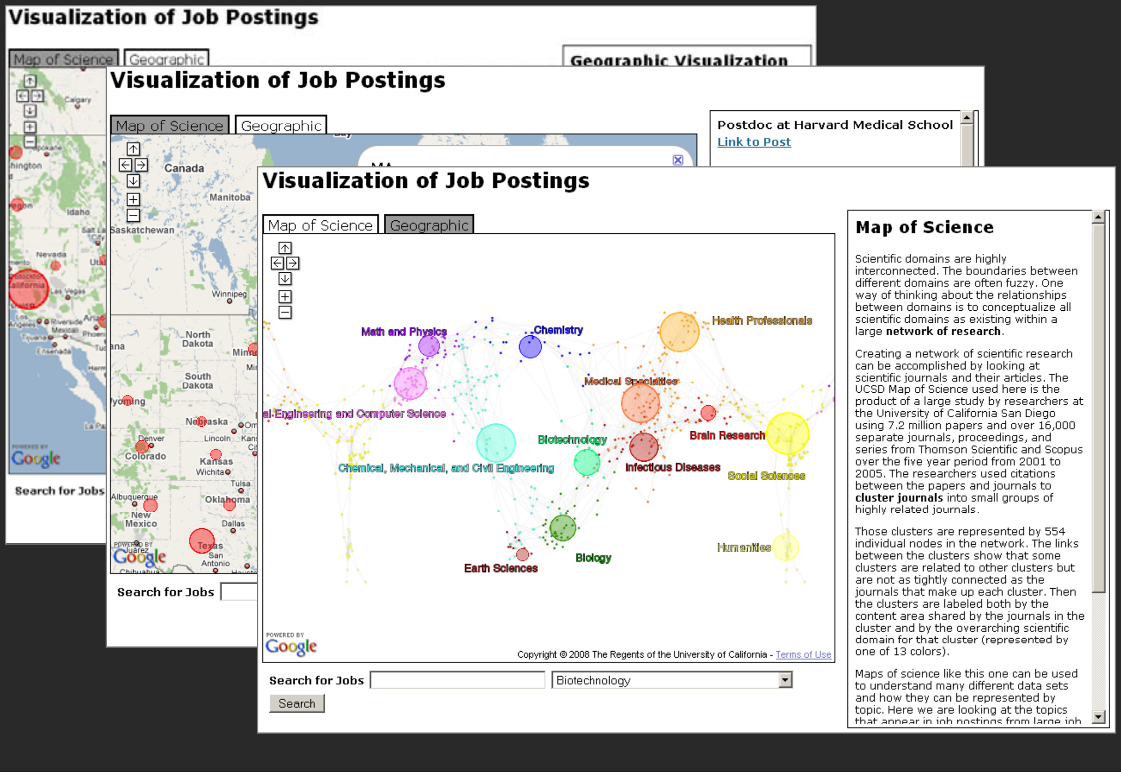
- Choose a number of categories from the arxiv classification on the right
- Choose a time window
- Choose an aggregation level
- Zoom the map to greater detail
- Click on the icons on the map for further information
- Each time you select a new time window, click "Update View" to let the changes take effect

Notice: No warranty or liability for this service. By using it, users accept responsibility for their actions and implications that they may have. Comments and suggestions are most welcome.

http://www.livingscience.ethz.ch

Interactive World and Science Map of S&T Jobs

Angela Zoss, Michael Connover, Katy Börner (2010)



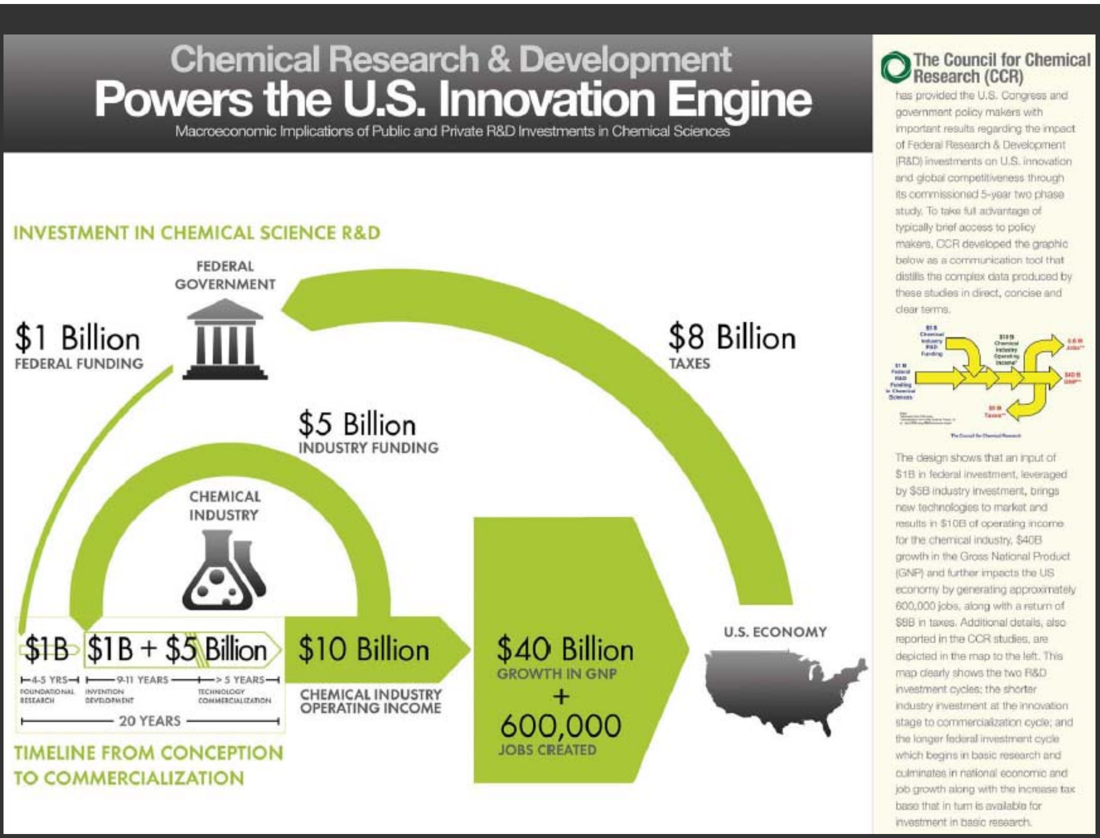
Map of Science

Scientific domains are highly interconnected. The boundaries between different domains are often fuzzy. One way of thinking about the relationships between domains is to conceptualize all scientific domains as existing within a large **network of research**.

Creating a network of scientific research can be accomplished by looking at scientific journals and their articles. The UCSD Map of Science used here is the product of a large study by researchers at the University of California San Diego using 7.2 million papers and over 16,000 separate journals, proceedings, and series from Thomson Scientific and Scopus over the five year period from 2001 to 2005. The researchers used citations between the papers and journals to **cluster journals** into small groups of highly related journals.

Those clusters are represented by 554 individual nodes in the network. The links between the clusters show that some clusters are related to other clusters but are not as tightly connected as the journals that make up each cluster. Then the clusters are labeled both by the content area shared by the journals in the cluster and by the overarching scientific domain for that cluster (represented by one of 13 colors).

Maps of science like this one can be used to understand many different data sets and how they can be represented by topic. Here we are looking at the topics that appear in job postings from large job



Chemical Research & Development

Powers the U.S. Innovation Engine

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

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 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.

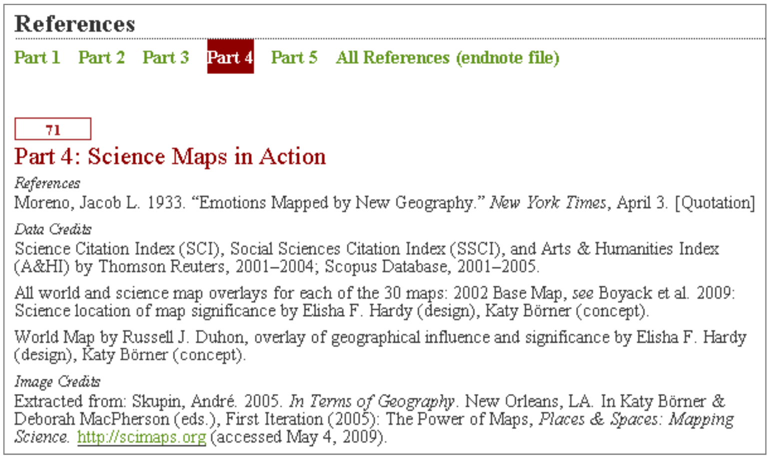
Council for Chemical Research. 2009. Chemical R&D Powers the U.S. Innovation Engine. Washington, DC. Courtesy of the Council for Chemical Research.

References & Credits

This section lists 1,650 citation references, more than 580 image credits, 80 data credits, and 60 software credits. More than 150 scholars provided input on the material presented in the atlas, and their contributions are acknowledged here.

As some spreads have up to 80 references and adding 80 parenthetical references or four-digit numbers to the page layout would considerably hurt readability, the references and credits are not given in the text. Instead, they are listed here by section. References and credits are ordered alphabetically except for those for **Part 2/Timeline,** which are ordered chronologically.

The Web site for the atlas (**http://scimaps.org**) supports pinpoint citations (that is, references and credits are associated with the specific text they support). In addition, the site will make available EndNote and bibtex files containing all the references.



References

Part 1 Part 2 Part 3 Part 4 Part 5 All References (endnote file)

71

Part 4: Science Maps in Action

References

Moreno, Jacob L. 1933. “Emotions Mapped by New Geography.” *New York Times*, April 3. [Quotation]

Data Credits

Science Citation Index (SCI), Social Sciences Citation Index (SSCI), and Arts & Humanities Index (A&HI) by Thomson Reuters, 2001–2004; Scopus Database, 2001–2005.

All world and science map overlays for each of the 30 maps: 2002 Base Map, *see* Boyack et al. 2009: Science location of map significance by Elisha F. Hardy (design), Katy Börner (concept).

World Map by Russell J. Duhon, overlay of geographical influence and significance by Elisha F. Hardy (design), Katy Börner (concept).

Image Credits

Extracted from: Skupin, André. 2005. *In Terms of Geography*. New Orleans, LA. In Katy Börner & Deborah MacPherson (eds.), First Iteration (2005): The Power of Maps, *Places & Spaces: Mapping Science*. http://scimaps.org (accessed May 4, 2009).

Introduction
Overview, promotion, and how to order.

Images
Access references and download high resolution images

References
Search all references and download in fielded format

Author
About the author

History
History of the Atlas

Questions & Answers
You ask, the author answers.

Vendors & Events
Where to buy

Press

Atlas of Science

Katy Börner
Published by MIT Press
Pre-order via Amazon, Barnes & Nobles, or MIT Press.

$29.95T/£22.95 cloth
13 x 11, 288 pp.
500 color illus.
978-0-262-01445-8

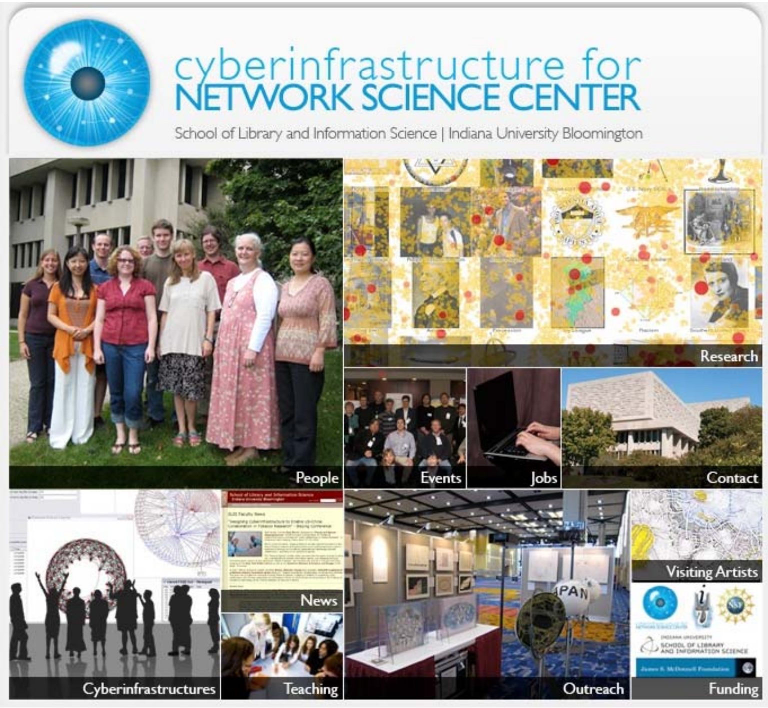
Download brochure
with full size sample pages

Cartographic maps have guided our explorations for centuries, allowing us to navigate the world. Science maps have the potential to guide our search for knowledge in the same way, allowing us to visualize scientific results. Science maps help us navigate, understand, and communicate the dynamic and changing structure of science and technology—help us make sense of the avalanche of data generated by scientific research today. 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.

Atlas of Science, based on the popular exhibit, "Places & Spaces: Mapping Science," describes and displays successful mapping techniques. The heart of the book is a visual feast: Claudius Ptolemy's Cosmographia World Map from 1482; a guide to a PhD thesis that resembles a subway map; "the structure of science" as revealed in a map of citation relationships in papers published in 2002; a visual

http://scimaps.org

We would like to thank the map makers



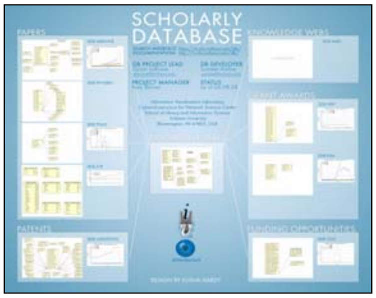
http://cns.slis.indiana.edu

Computational Scientometrics Cyberinfrastructures

Scholarly Database: 25 million scholarly records
http://sdb.slis.indiana.edu

James S. McDonnell Foundation

VIVO Research Networking
http://vivoweb.org



Information Visualization Cyberinfrastructure
http://iv.slis.indiana.edu

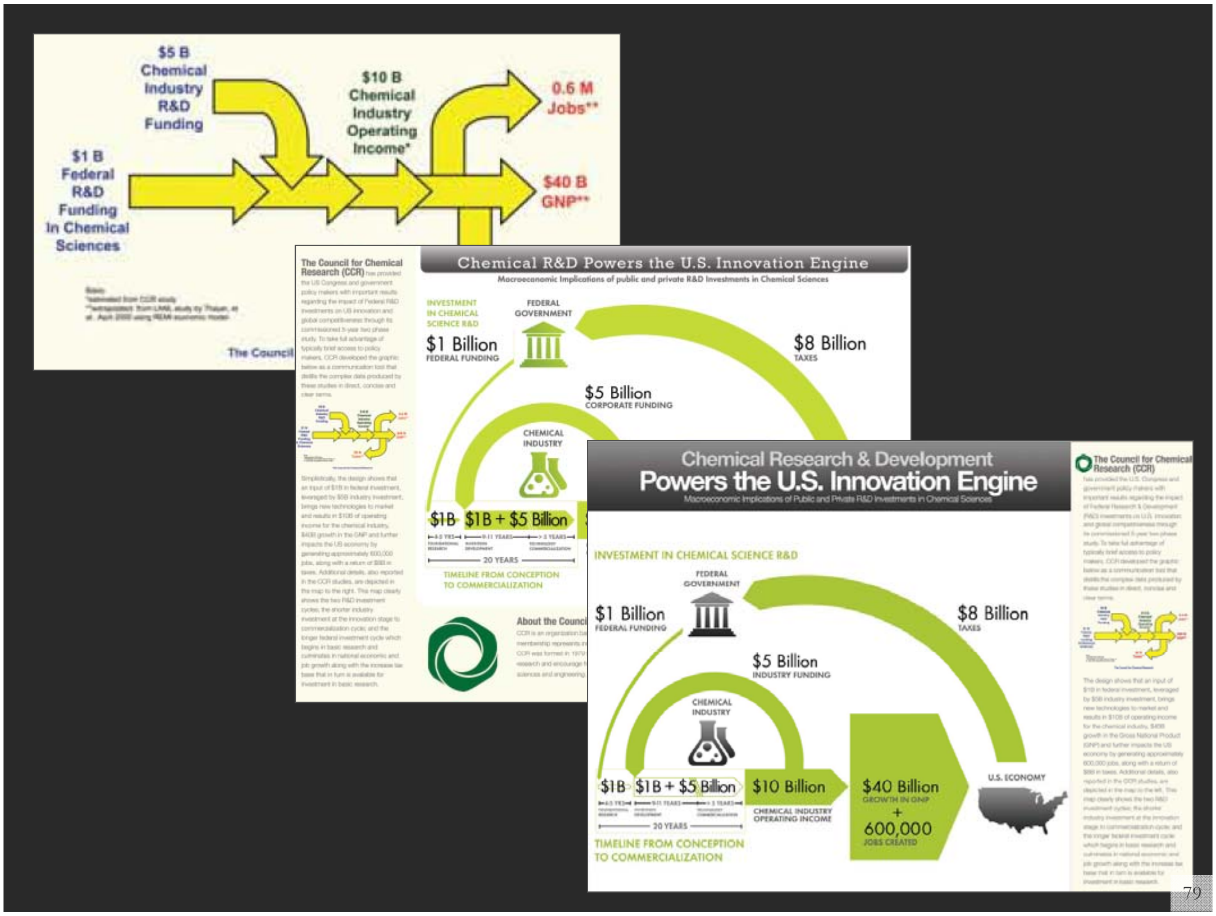
Network Workbench Tool & Community Wiki
http://nwb.slis.indiana.edu

Science of Science (Sci^2) Tool and CI Portal
http://sci.slis.indiana.edu

Epidemics Cyberinfrastructure
http://epic.slis.indiana.edu/

OSGi Alliance

CIShell
cyberinfrastructure shell



The Council for Chemical Research (CCR) has provided the US Congress and government policy makers with important results regarding the impact of Federal R&D investments on US innovation and global competitiveness through its commissioned 3-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.

Simplistically, 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 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 right. 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.

Chemical R&D Powers the U.S. Innovation Engine

Macroeconomic Implications of public and private R&D Investments in Chemical Sciences

Chemical Research & Development Powers the U.S. Innovation Engine

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

The Council for Chemical Research (CCR)