STEAM Observatory: Data Mining, Modeling, and Visualization in Support of Smart Decision Making

Katy Börner @katycns

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

NSF NSCI LECTURE SERIES

National Science Foundation, Room E3450

January 16, 10:30p, 2018

National Strategic Computing Initiative (NSCI)

NSCI is a whole-of-nation effort to **accelerate scientific discovery** and **economic competitiveness** by **maximizing the benefits of high-performance computing (HPC) research, development, and deployment.**

NSCI aims to address five strategic objectives:

1. Accelerate delivery of a capable exascale computing system,
2. Increase coherence between technology for modeling/simulation and data analytics,
3. Establish a viable path forward in the "post-Moore's Law" era,
4. Increase the capacity and capability of an enduring national HPC ecosystem, and
5. Develop U.S. government, industry, and academic collaborations to share the benefits.

The NSCI Strategic Plan was made public in July 2016:

https://www.whitehouse.gov/sites/whitehouse.gov/files/images/NSCI%20Strategic%20Plan.pdf

STEAM Observatory: Data Mining, Modeling, and Visualization in Support of Smart Decision Making

The STEAM Observatory will advance computational models and visualizations of STEM data (HPC usage, publications, patents, funding, clinical trials, stock market, social media) to help explore questions such as:

- What jobs will exist in ten years and what career paths lead to success?
- Which types of institutions will likely be most innovative in the future?
- How will the higher education cost bubble burst affect these institutions?
- What funding strategies have the highest return on investment?
- How will changing demographics, alternative economic growth trajectories, and relationships among nations impact answers to these and other questions?

STEAM Observatory cont.

The initial STEAM Observatory might have three major research thrusts:

- Modelling and Visualization of Complex Systems
- Communicating and Optimizing the Internet of Things
- STEAM Education and Workforce Development

All three are supported by a common theoretical and computational core

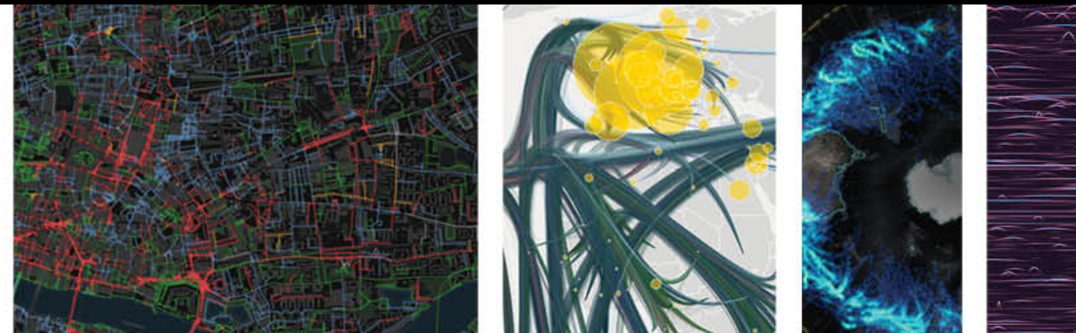
- Developing Enabling Theory and
- Cyber-Physical High-Performance Infrastructure.

Instead of using powerful telescopes for observing the planets and stars; STEAM Observatory will provide **easy access** to continuously evolving datasets and software artifacts running on HPC infrastructures that are optimized for analyzing, modelling, and visualizing STEAM data—transforming the way experts and the public see, learn, and act.

Next

- Current Work & Visions of The Future
- Maps & Macroscopes
- Interactive Visualizations of (NIH/NSF) Infrastructure Impact
- Visualizing the Internet of Things (IoT)
- Challenges & Opportunities

Arthur M. Sackler COLLOQUIA OF THE NATIONAL ACADEMY OF SCIENCES



Modeling and Visualizing Science and Technology Developments

National Academy of Sciences Sackler Colloquium, December 4-5, 2017, Irvine, CA

Rankings and the Efficiency of Institutions

H. Eugene Stanley | Albert-László Barabási | Lada Adamic | Marta González | Kaye Husbands Fealing | Brian Uzzi | John V. Lombardi

Higher Education and the Science & Technology Job Market

Katy Börner | Wendy L. Martinez | Michael Richey | William Rouse | Stasa Milojevic | Rob Rubin | David Krakauer

Innovation Diffusion and Technology Adoption

William Rouse | Donna Cox | Jeff Alstott | Ben Shneiderman | Rahul C. Basole | Scott Stern | Cesar Hidalgo

Modeling Needs, Infrastructures, Standards

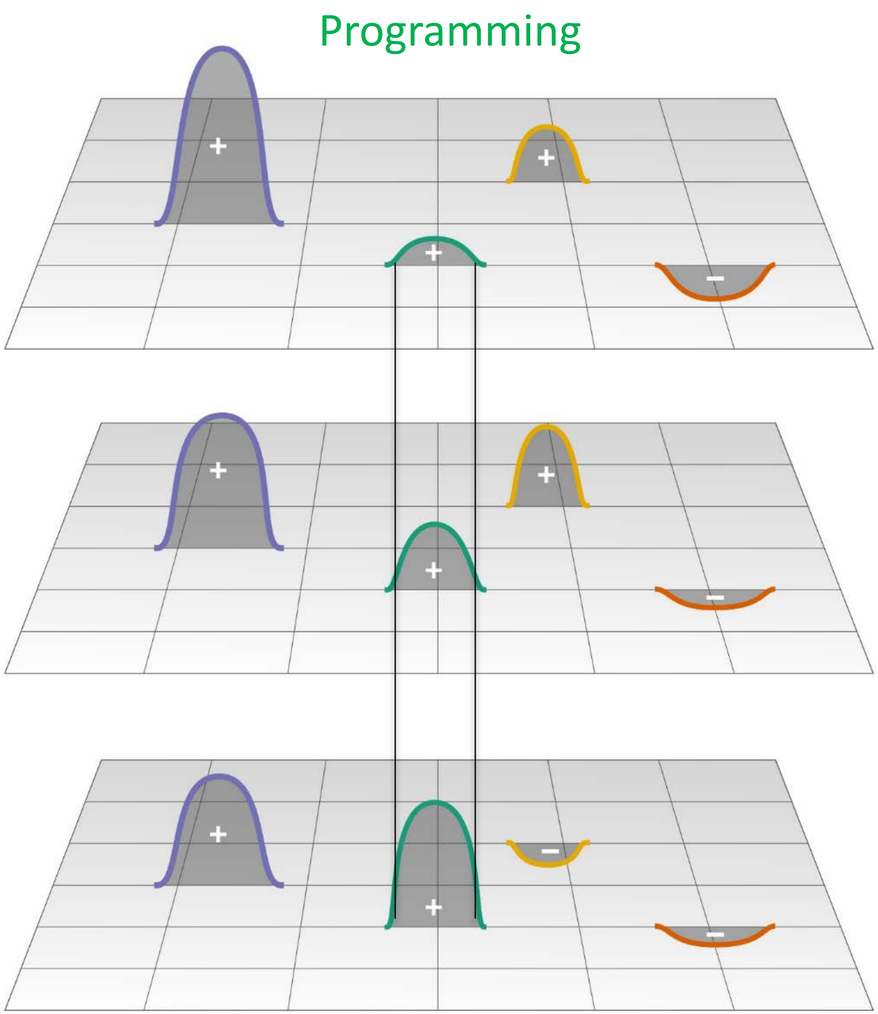
Paul Trunfio | Sallie Keller | Andrew L. Russell | Guru Madhavan | Azer Bestavros | Jason Owen-Smith

nasonline.org/Sackler-Visualizing-Science

Science & Technology vs. Education/Training vs. Jobs

Katy Börner, Olga B. Scrivner, Xiaozhong Liu, Indiana University

Need to study the **(mis)match** and **temporal dynamics** of S&T progress, education and workforce development options, and job requirements.



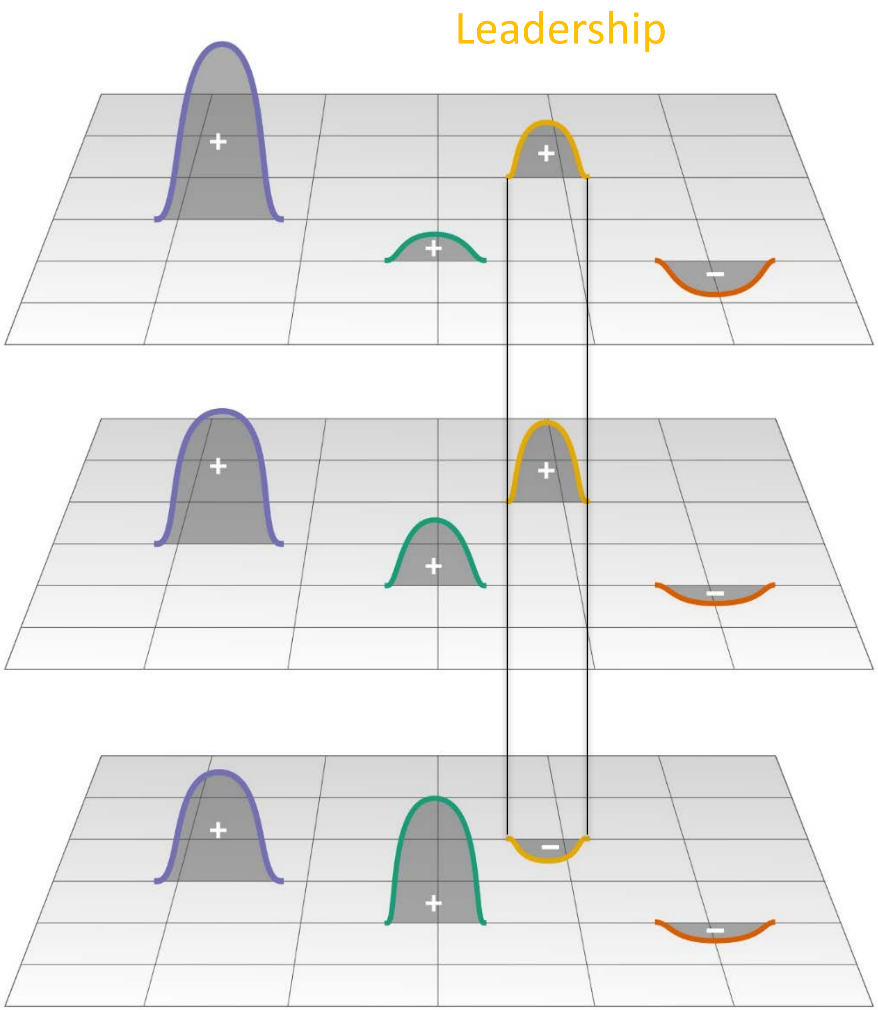
Challenges:

- Rapid change of STEM knowledge
- Increase in tools, AI
- Social skills (project management, team leadership)
- Increasing team size

Science & Technology vs. Education/Training vs. Jobs

Katy Börner, Olga B. Scrivner, Xiaozhong Liu, Indiana University

Need to study the **(mis)match** and **temporal dynamics** of S&T progress, education and workforce development options, and job requirements.



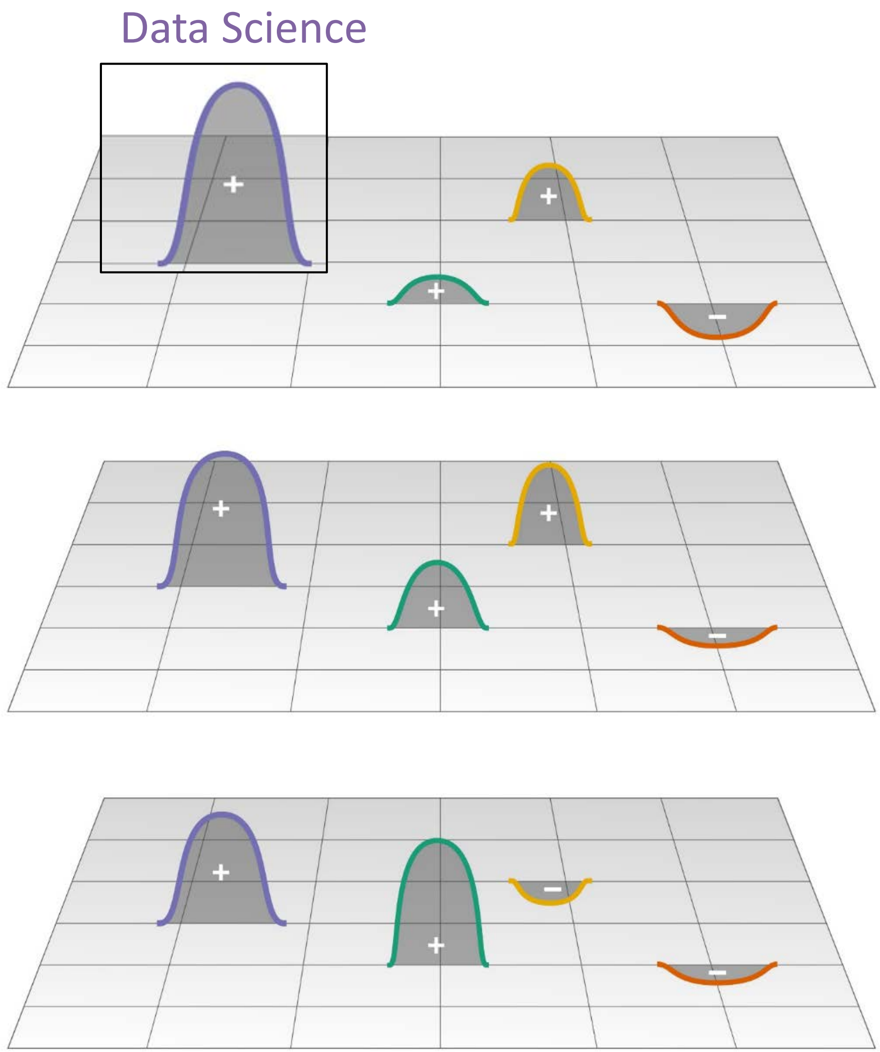
Challenges:

- Rapid change of STEM knowledge
- Increase in tools, AI
- Social skills (project management, team leadership)
- Increasing team size

Science & Technology vs. Education/Training vs. Jobs

Katy Börner, Olga B. Scrivner, Xiaozhong Liu, Indiana University

Need to study the **(mis)match** and **temporal dynamics** of S&T progress, education and workforce development options, and job requirements.



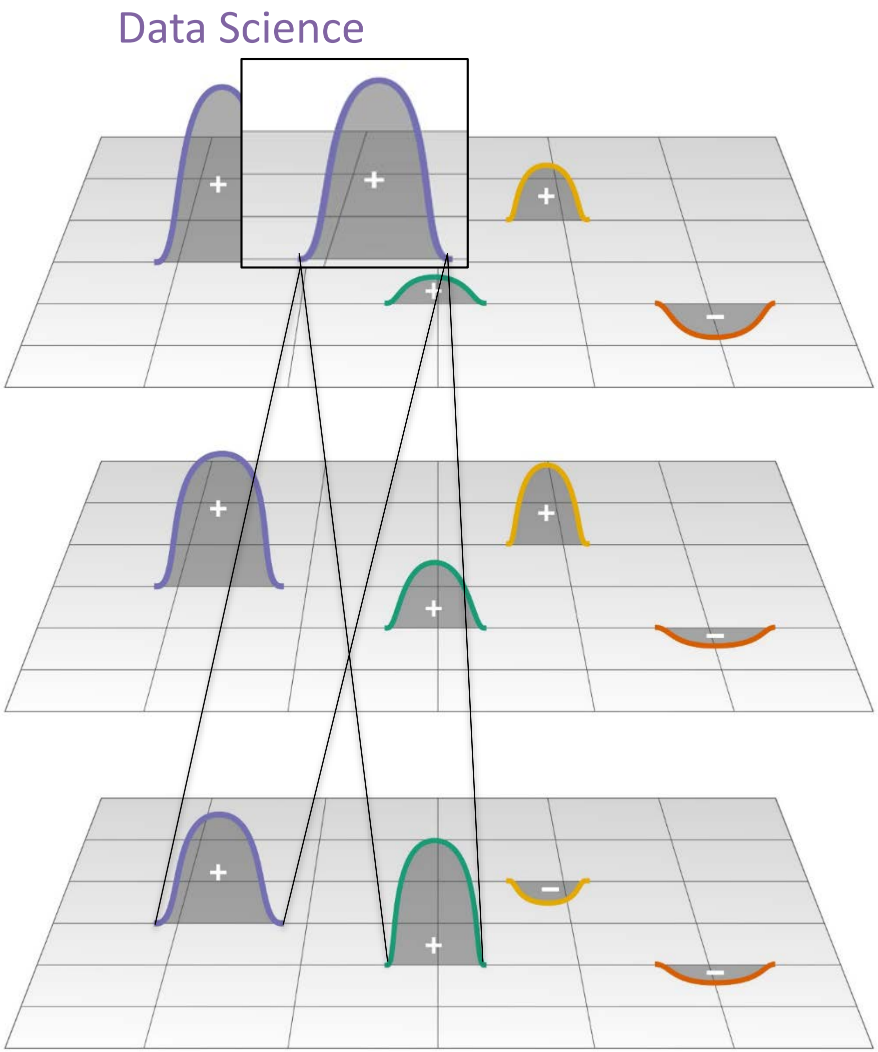
Challenges:

- Rapid change of STEM knowledge
- Increase in tools, AI
- Social skills (project management, team leadership)
- Increasing team size

Science & Technology vs. Education/Training vs. Jobs

Katy Börner, Olga B. Scrivner, Xiaozhong Liu, Indiana University

Need to study the **(mis)match** and **temporal dynamics** of S&T progress, education and workforce development options, and job requirements.



Challenges:

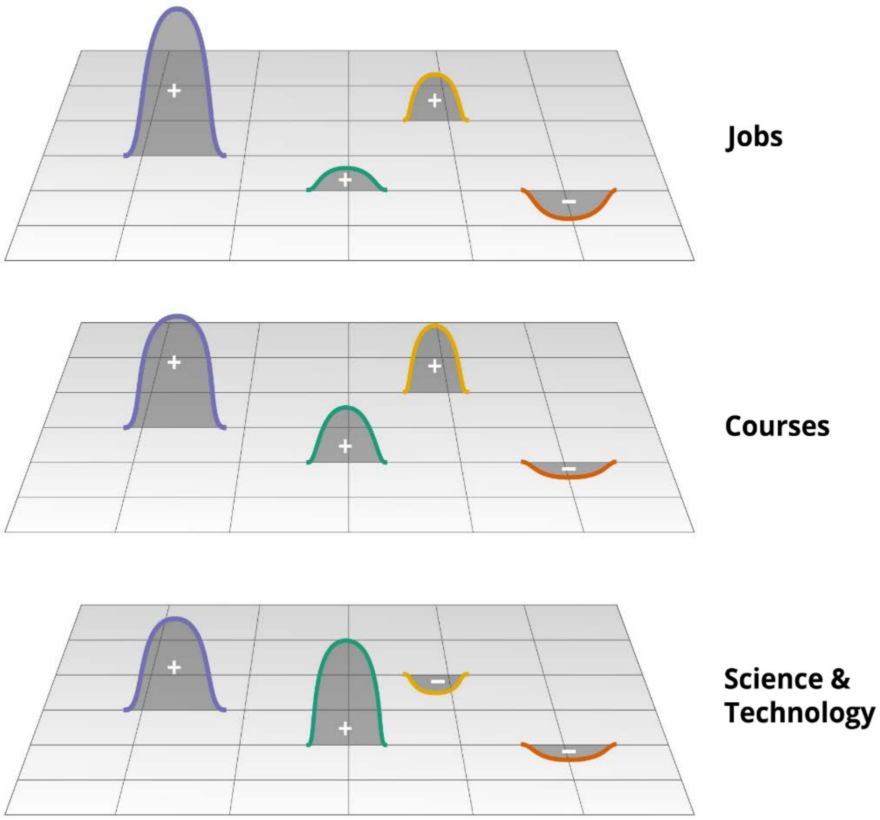
- Rapid change of STEM knowledge
- Increase in tools, AI
- Social skills (project management, team leadership)
- Increasing team size

Science & Technology vs. Education/Training vs. Jobs

Katy Börner, Olga B. Scrivner, Xiaozhong Liu, Indiana University

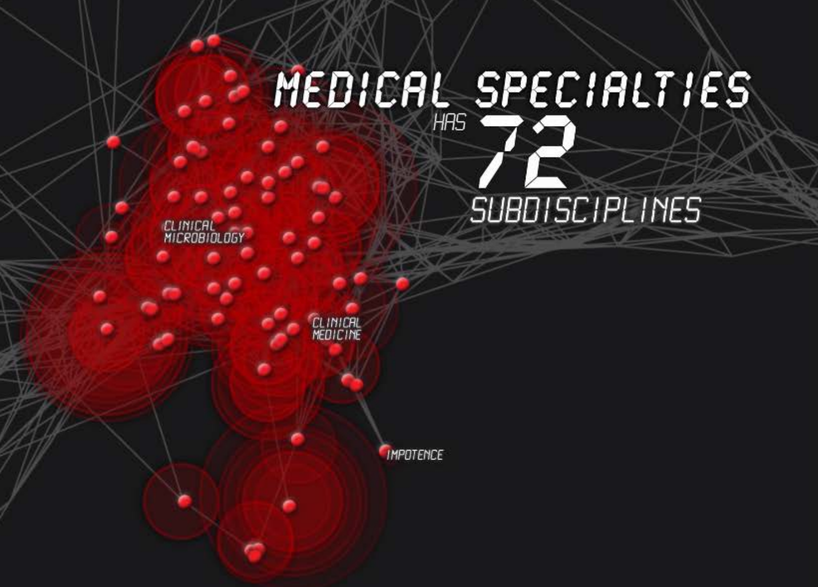
Study results are needed by:

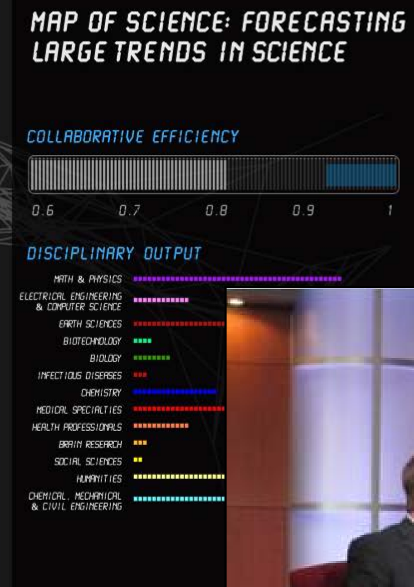
- **Students:** What jobs will exist in 1-4 years? What program/learning trajectory is best to get/keep my dream job?
- **Teachers:** What course updates are needed? What curriculum design is best? What is my competition doing? How much timely knowledge (to get a job) vs. forever knowledge (to be prepared for 80 productive years) should I teach? How to innovate in teaching and get tenure?
- **Employers:** What skills are needed next year, in 5 years? Who trains the best? What skills does my competition list in job advertisements? How to hire/train productive teams?

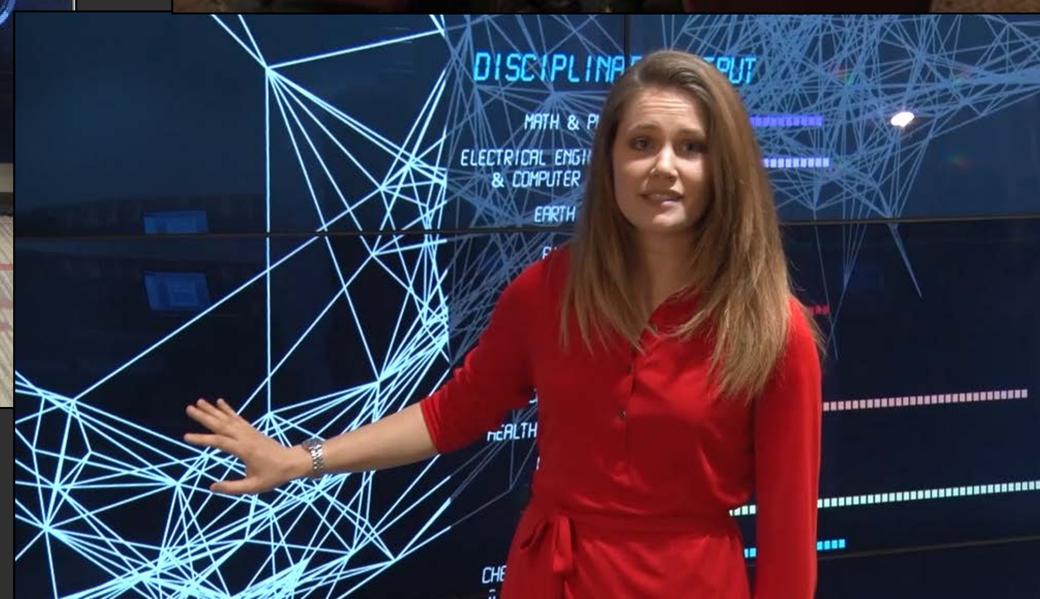


What is ROI of my time, money, compassion?

Science Forecast S1:E1





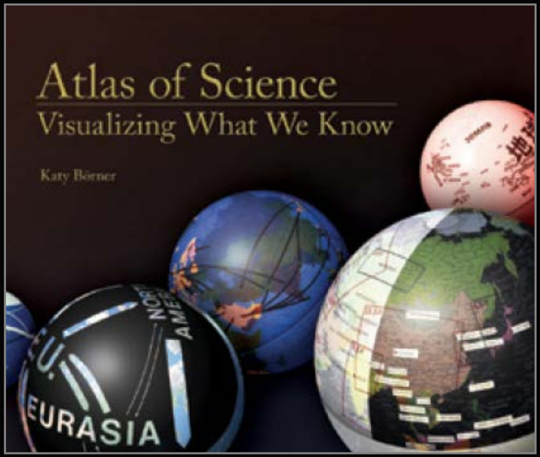


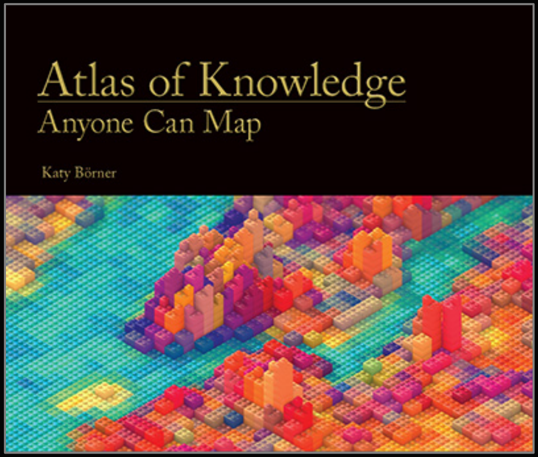
Science Forecast
S1:E1

https://www.youtube.com/watch?v=lByX2_eb_QQ

Maps of Science & Technology

Using large scale datasets, advanced data mining and visualization techniques, and substantial computing resources.

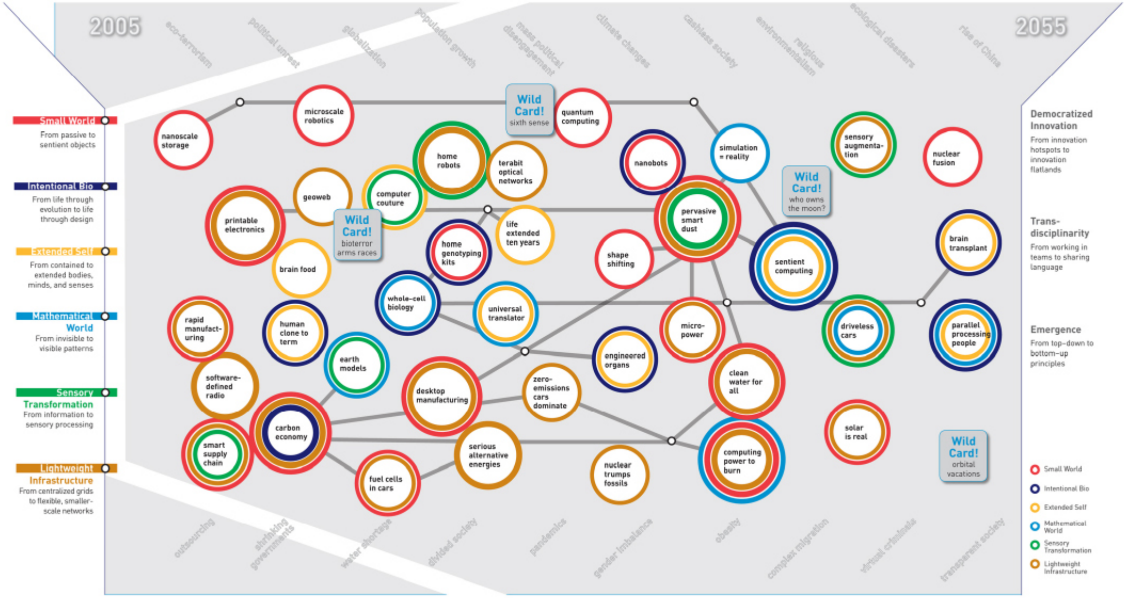




INSTITUTE FOR THE FUTURE

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

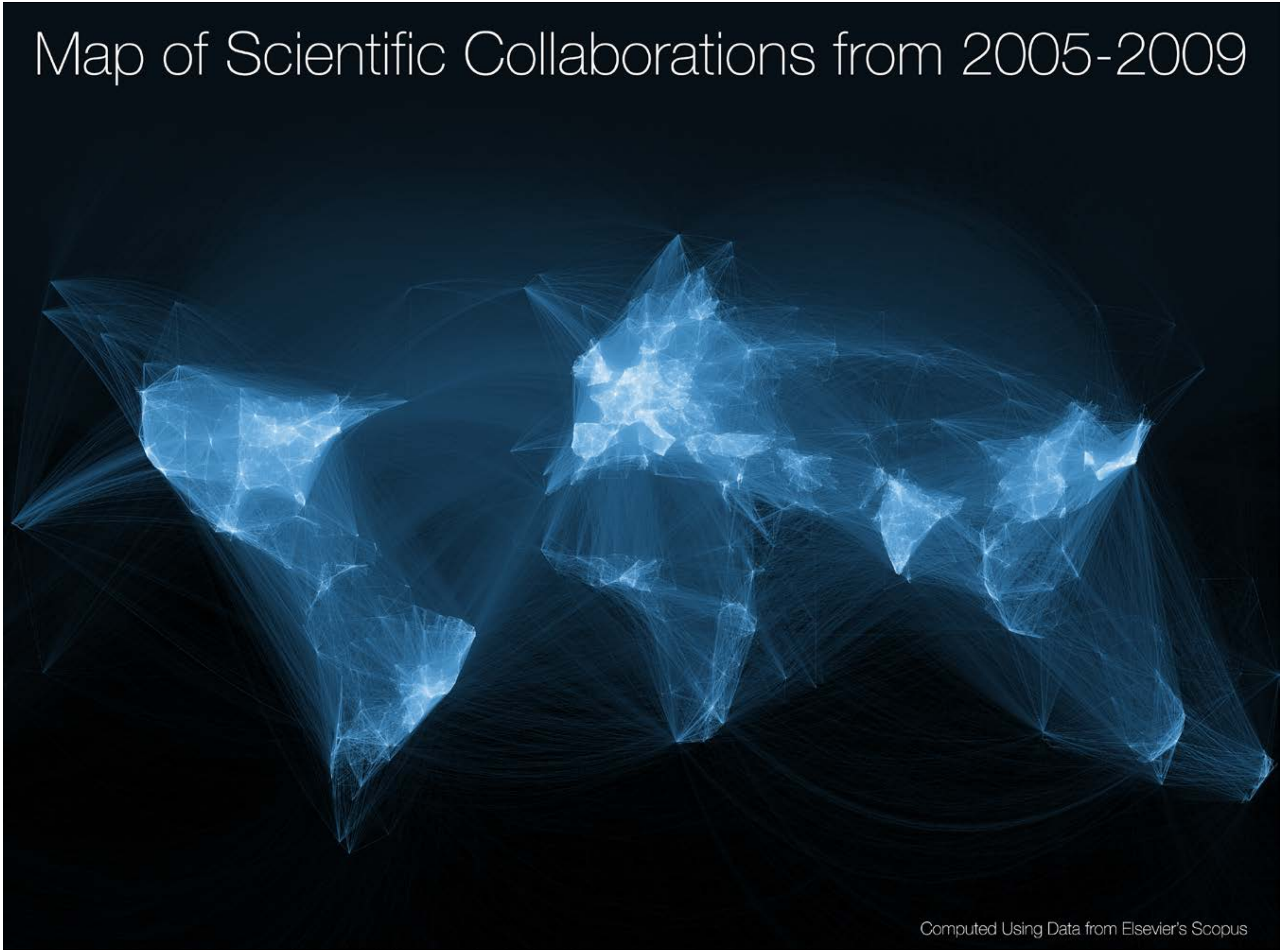
In the last two centuries, natural philosophy and natural history fractured into the now-familiar disciplines of physics, chemistry, biology, and so on. The sciences evolved into their current form in response to intellectual and professional opportunities, philanthropic priorities, and economic and 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.

III.5 *Science & Technology Outlook: 2005-2055* – Marina Gorbis, Jean Hagan, Alex Soojung-Kim Pang, and David Pescovitz - 2006

III.5 *Science & Technology Outlook: 2005-2055* – Marina Gorbis, Jean Hagan, Alex Soojung-Kim Pang, and David Pescovitz - 2006



Stream of Scientific Collaborations Between World Cities - Olivier H. Beauchesne - 2012

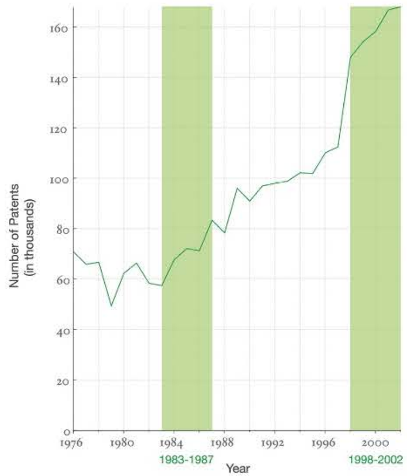
Examining the Evolution & Distribution of Patent Classifications

Managing Growing Patent Portfolios

Organizations, businesses, and individuals rely on patents to protect their intellectual property and business models. As market competition increases, patenting innovation and intellectual property rights becomes ever more important.

Managing the staggering number of patents demands new tools and methodologies. Grouping patents by their classifications offers an ideal resolution for better understanding how intellectual borders are established and change over time.

The charts below show the annual number of patents granted from January 1, 1976 to December 31, 2002 in the United States Patent and Trademark Office (USPTO) patent archive; slow and fast growing patent classes; the top 10 fast growing patent subclasses; and two evolving patent portfolios.



The Structure and Evolution of the Patent Space

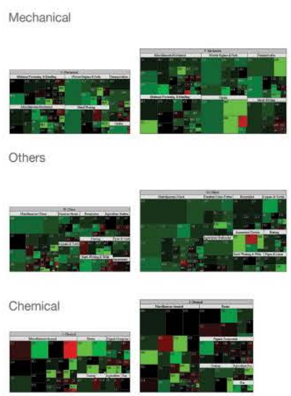
The United States Patent and Trademark Office assigns each patent to one of more than 450 classes covering broad application domains. For example, class 514 encompasses all patents dealing with 'Drug, Bio-Affecting and Body Treating Compositions.' Classes are further broken down by subclasses that have hierarchical associations. As one example, class 455 features subclass 99 entitled "with vehicle."

The top 10 fast growing patent classes for 1998–2002 are listed together with the number of patents granted. Most come from the 'Computer and Communications' and the 'Drugs and Medical' area.

The evolving hierarchical structure of patent classes and their sizes is represented using treemaps, a space-filling visualization technique developed by Ben Shneiderman at the University of Maryland. A treemap presents a hierarchy as a collection of nested rectangles—demarcating a parent-child relationship between nodes by nesting the child within the parent rectangle. The size and color of each rectangle represent certain attributes of the nodes.

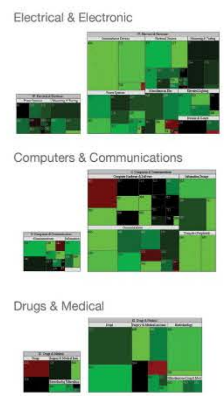
Here, each rectangle represents a class and the area size denotes the total number of patents in that class. The rectangle's color corresponds to percentage increase (**green**) or decrease (**red**) in the number of patents granted in that class from the previous interval.

Slow Growing Classes



1983 - 1987   1998 - 2002

Fast Growing Classes



1983 - 1987   1998 - 2002

Top-10 Subclasses

Class	Title	# of Patents
514	Drug, Bio-Affecting and Body Treating Compositions	18,778
438	Semiconductor Device Manufacturing:Process	17,775
435	Chemistry: Molecular Biology and Microbiology	17,474
424	Drug, Bio-Affecting and Body Treating Compositions	13,637
428	Stock Material or Miscellaneous Articles	13,314
257	Active Solid-State Devices (e.g., Transistors, Solid-State Diodes)	12,924
395	Information Processing System Organization	9,955
345	Computer Graphics Processing, Operator Interface Processing, and Selective Visual Display Systems	9,510
359	Optical: Systems and Elements	9,151
365	Static Information Storage and Retrieval	8,392
	Total	130,910

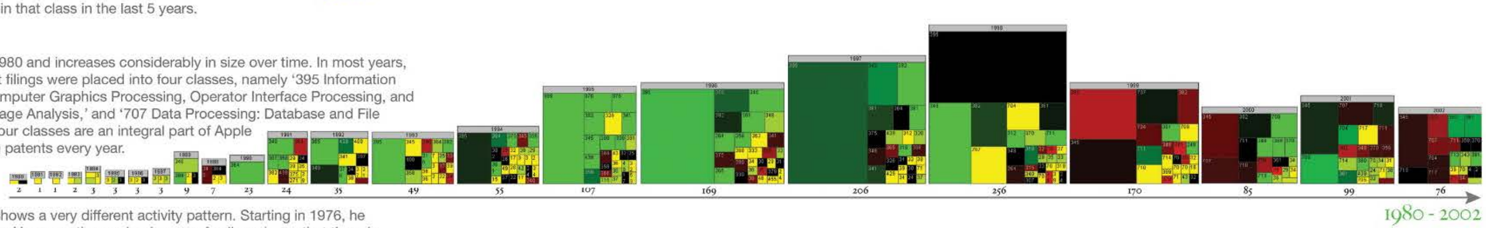
1998 - 2002

Patent Portfolio Analysis

A longitudinal analysis of portfolios reveals different patenting strategies. For each year (given in gray above each treemap), a treemap of all new patents granted to the assignee is shown. The number of patents is given below each treemap. The same size and color coding as above was used. In addition, yellow indicates that no patent has been granted in that class in the last 5 years.

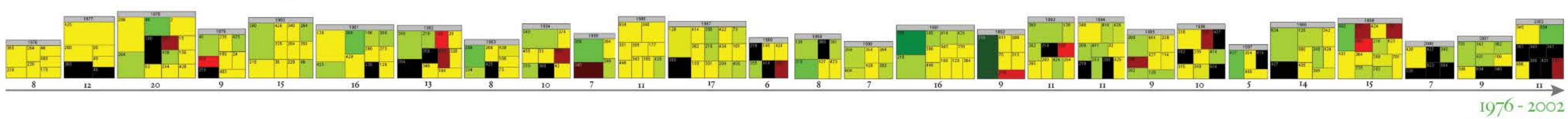
Apple Computer, Inc.

Apple Computer, Inc.'s portfolio starts in 1980 and increases considerably in size over time. In most years, more than half of Apple Computer's patent filings were placed into four classes, namely '395 Information Processing System Organization,' '345 Computer Graphics Processing, Operator Interface Processing, and Selective Visual Display Systems,' '382 Image Analysis,' and '707 Data Processing: Database and File Management or Data Structures.' These four classes are an integral part of Apple Computer, Inc.'s patent portfolio, receiving patents every year.

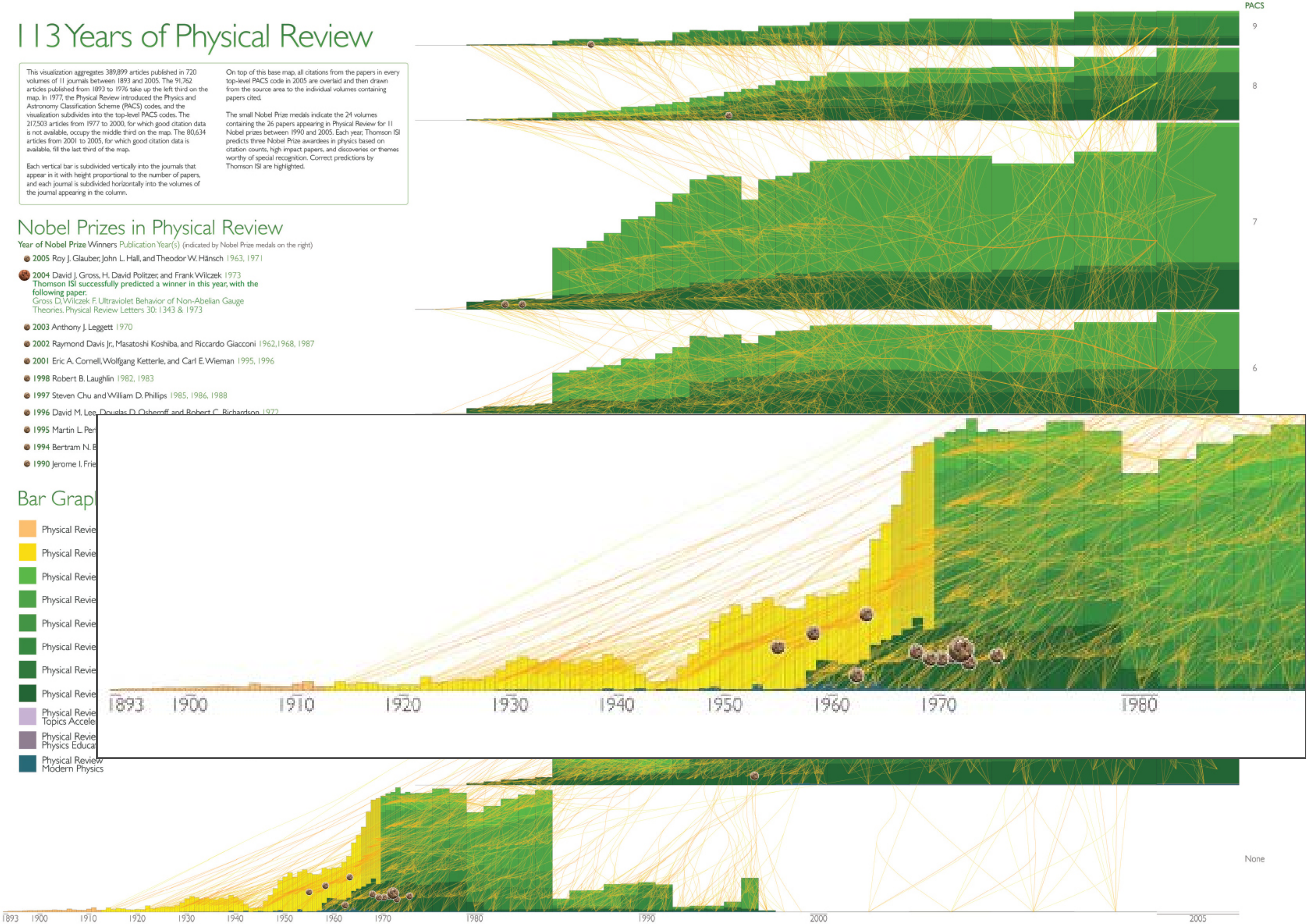


Jerome Lemelson

The patent portfolio of Jerome Lemelson shows a very different activity pattern. Starting in 1976, he publishes between 6–20 patents each year. However, the predominance of yellow shows that there is little continuity from previous years in regards to the classes into which patents are filed. No class dominates. Instead, more and more new intellectual space is claimed.



IV.5 *Examining the Evolution & Distribution of Patent Classifications* - Daniel O. Kutz, Katy Borner, and Elisha F. Hardy - 2004



III.6 *113 Years of Physical Review* - Bruce W. Herr II, Russell J. Duhon, Elisha F. Hardy, Shashikant Penumarthy, and Katy Börner - 2006

WEB OF SCIENCE™ 2016 CITATION LAUREATES

IMPACT OF SCIENTIFIC INNOVATIONS

CHEMISTRY

George Church and Feng Zhang developed application of CRISPR-cas9 gene editing in mouse and human cells.

Dennis Lo Yuk Ming detected cell-free fetal DNA in maternal plasma, a revolution in noninvasive prenatal testing.

Hiroshi Maeda and Yasuhiro Matsumura discovered the enhanced permeability and retention (EPR) effect of macromolecular drugs, a key finding for cancer therapeutics.

PHYSICS

Marvin L. Cohen for theoretical studies of solid materials, prediction of their properties, and especially for the empirical pseudopotential method.

Ronald W.P. Drever, Kip S. Thorne and Rainer Weiss developed the Laser Interferometer Gravitational-Wave Observatory (LIGO) that made possible the detection of gravitational waves.

Celso Grebogi, Edward Ott and James A. Yorke described a control theory of chaotic systems, the OGY method.

MEDICINE

James P. Allison, Jeffrey A. Bluestone and Craig B. Thompson explained how CD28 and CTLA-4 are regulators of T cell activation, modulating immune response.

Gordon J. Freeman, Tasuku Honjo and Arlene H. Sharpe elucidated programmed cell death-1 (PD-1) and its pathway, which has advanced cancer immunotherapy.

Michael N. Hall, David M. Sabatini and Stuart L. Schreiber discovered the growth regulator Target of Rapamycin (TOR) and the mechanistic Target of Rapamycin (mTOR).

Annually, Thomson Reuters analysts mine scientific literature citation data to identify the researchers whose work is worthy of Nobel recognition for induction into the Hall of Citation Laureates. They are the innovators responsible for the world's most influential scientific discoveries, with scholarly papers typically ranking in the top 0.1% by citations within their field. Many go on to win the Nobel Prize for their significant contributions toward the advancement of science.

To learn more visit: **stateofinnovation.com**

Source: Thomson Reuters Web of Science; InCites Essential Indicators.
Visit **stateofinnovation.com** to learn more about the 2016 Thomson Reuters Citation Laureates.

ECONOMICS

Olivier J. Blanchard contributed to macroeconomics, including determinants of economic fluctuations and employment.

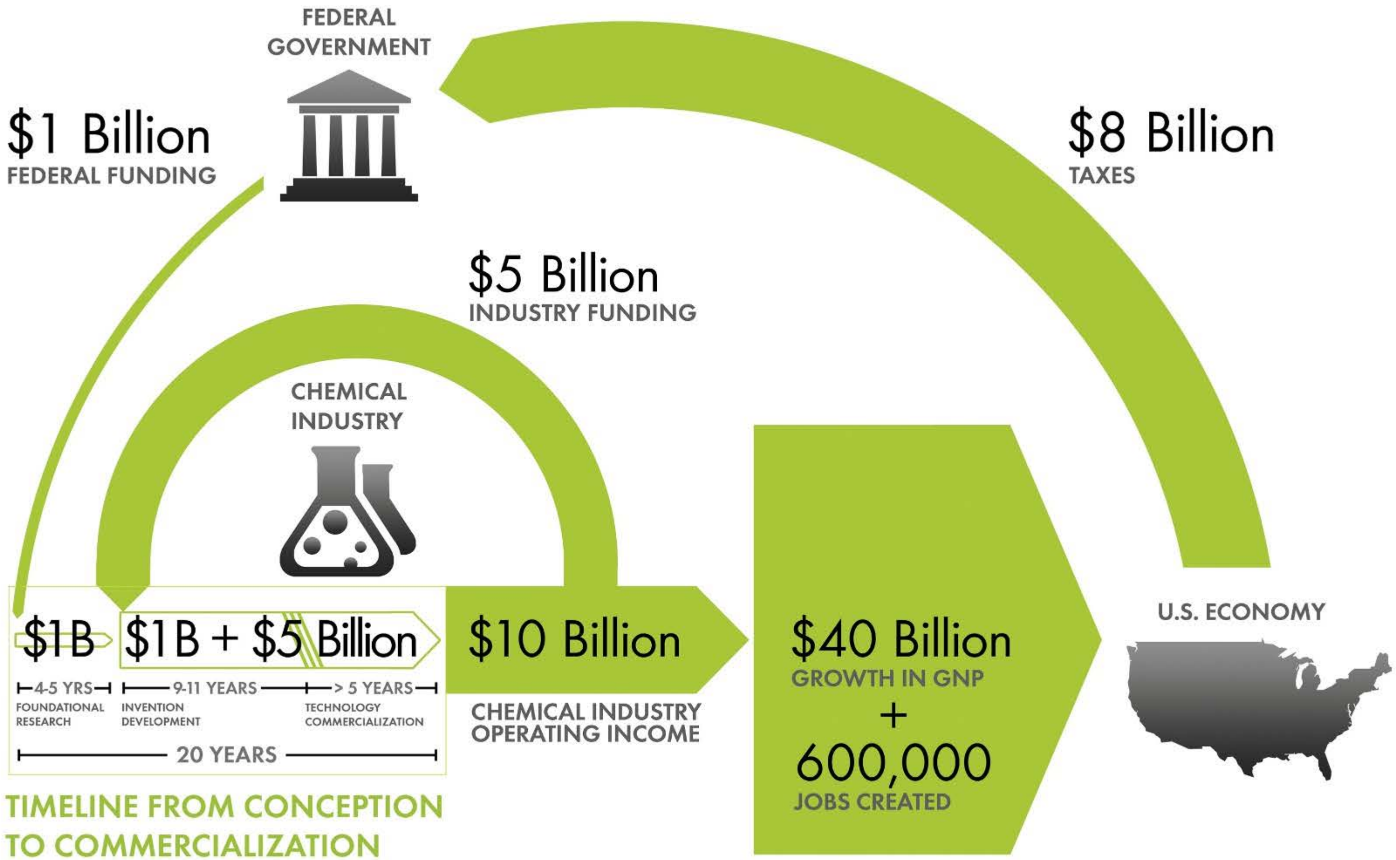
Edward P. Lazear developed the distinctive field of personnel economics.

Mark J. Melitz pioneered descriptions of firm heterogeneity and international trade.

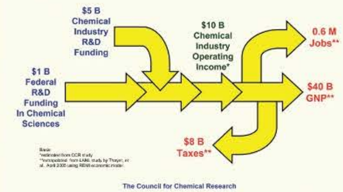
http://stateofinnovation.com/2016-citation-laureates

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.

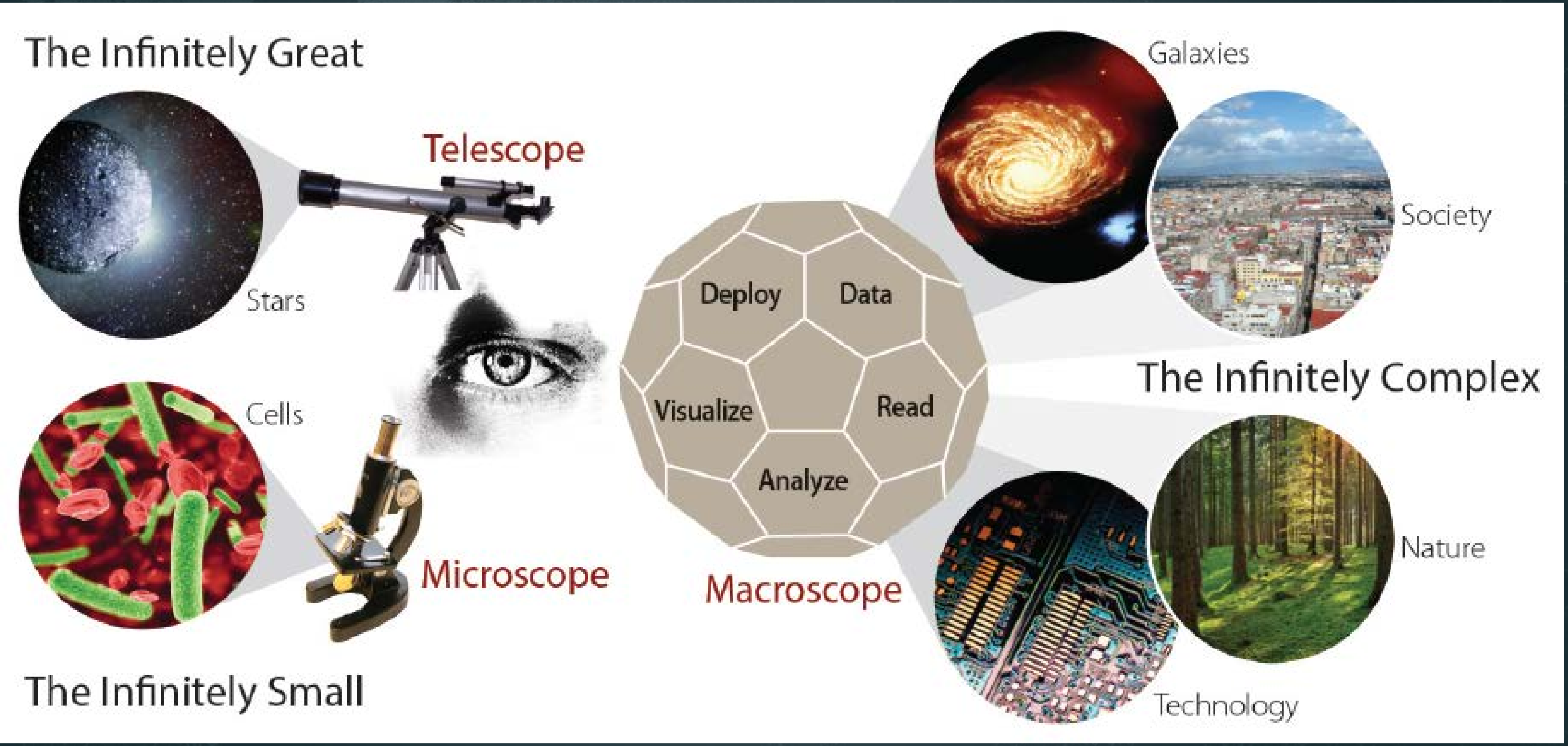


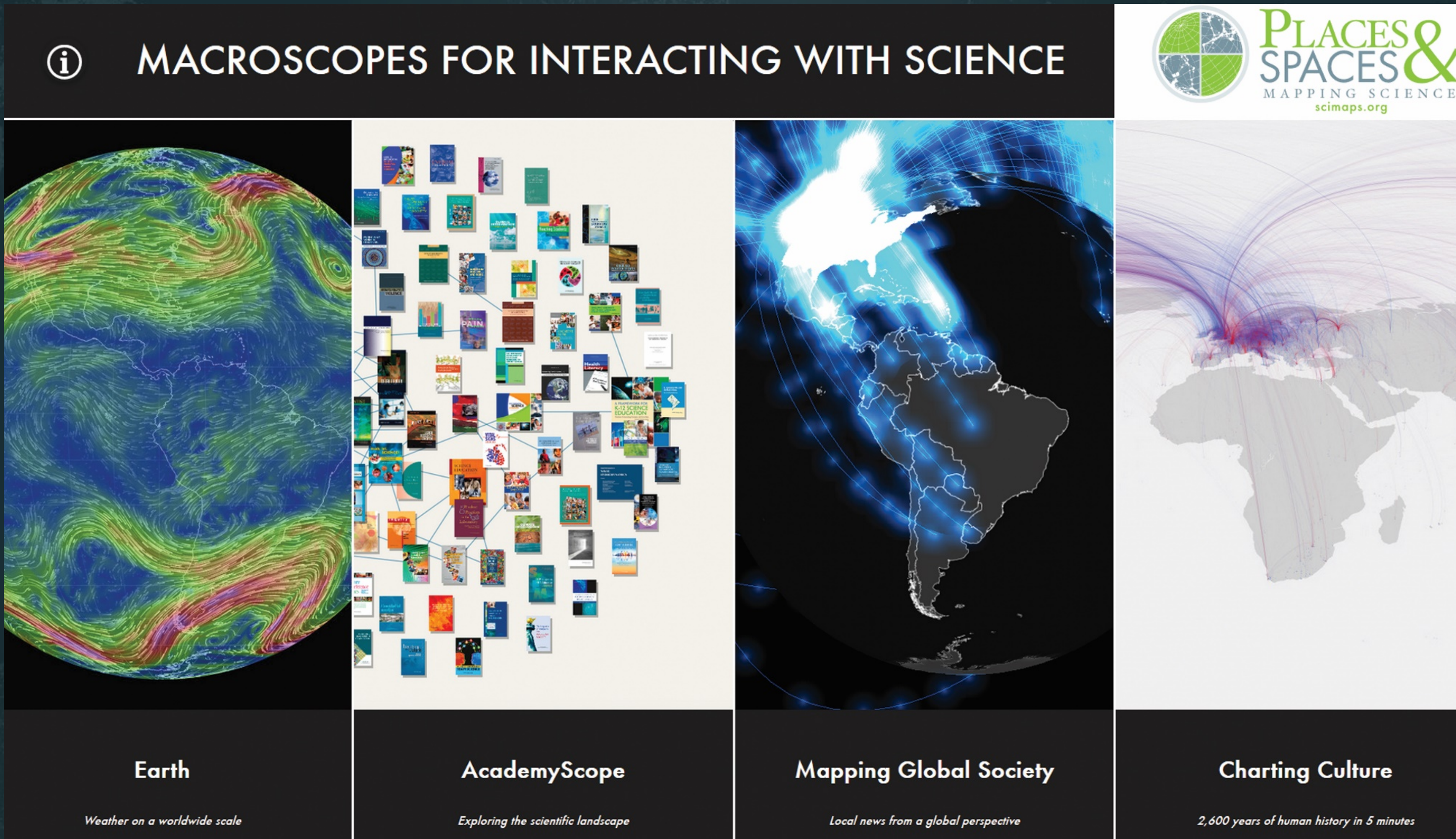
The design shows that an input of $1B in federal investment, leveraged by $5B in industry investment, brings new technologies to market and results in $10B of operating income for the chemical industry, $40B of growth in the Gross National Product (GNP) and further impacts the US economy by generating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

V.6 *Chemical R&D Powers the U.S. Innovation Engine* - The Council for Chemical Research - 2009

MAPS
vs.
MACROSCOPES

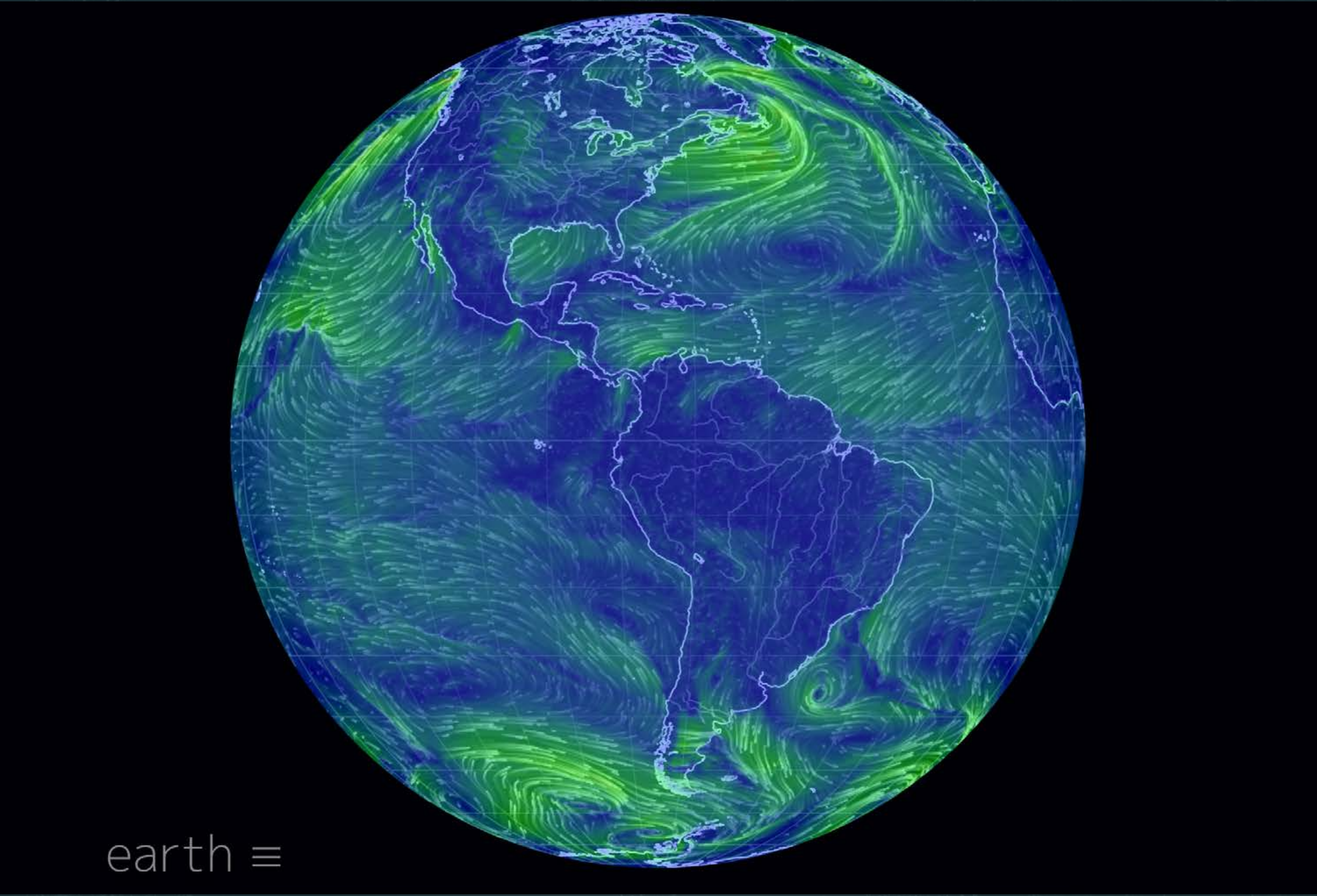
Microscopes & Telescopes vs. MACROSCOPES



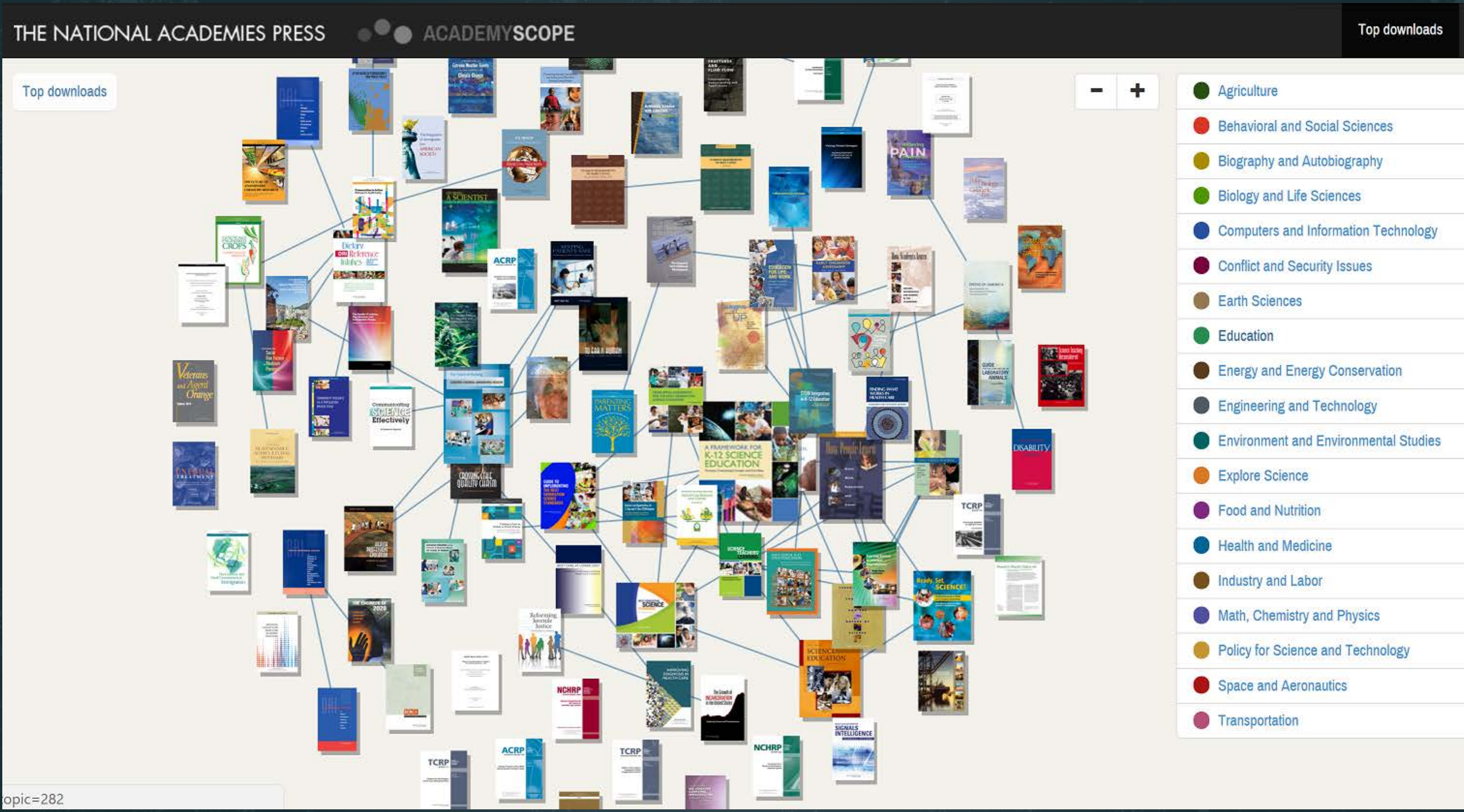


Iteration XI (2015): Macroscopes for Interacting with Science

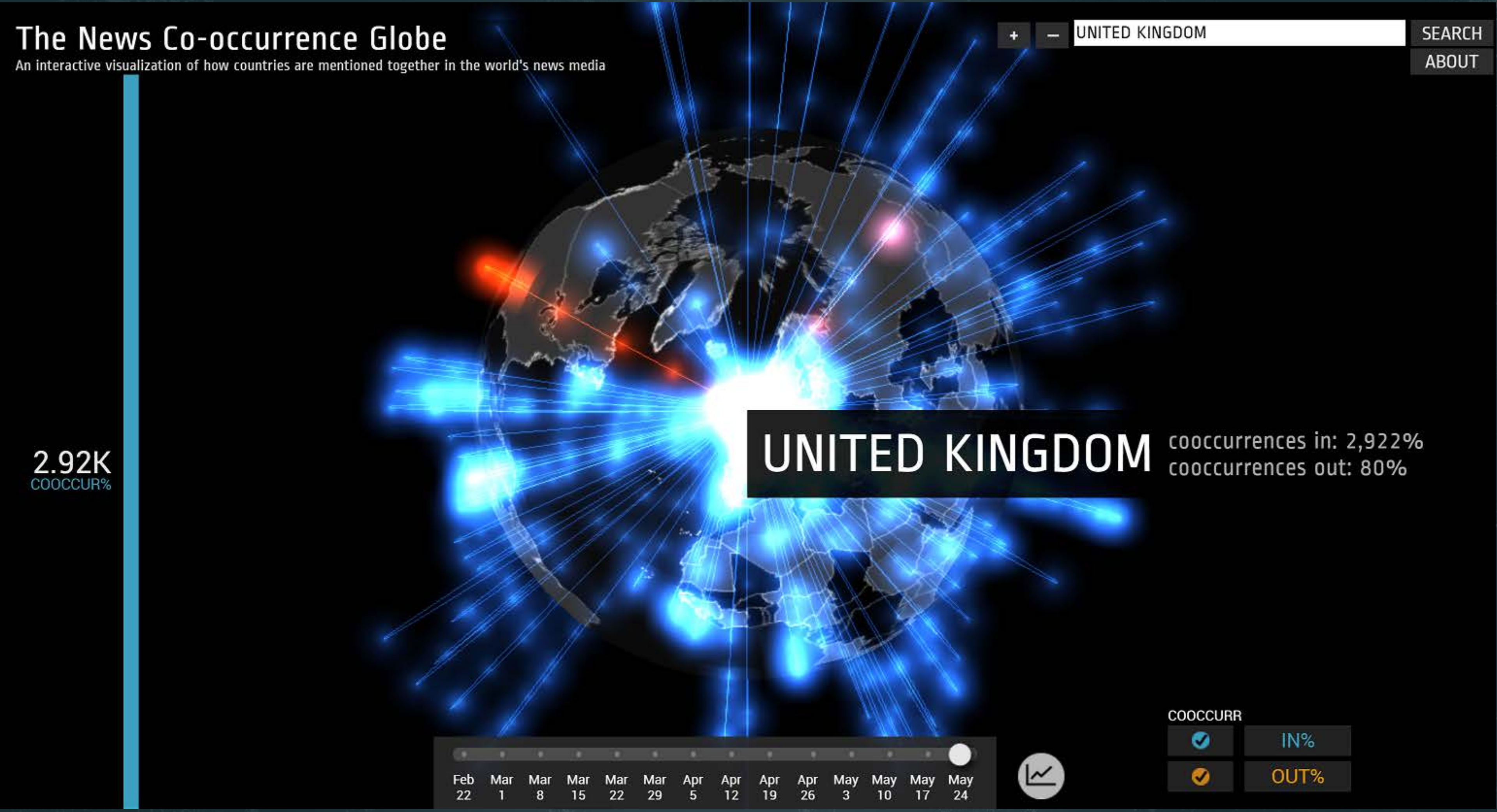
http://scimaps.org/iteration/11



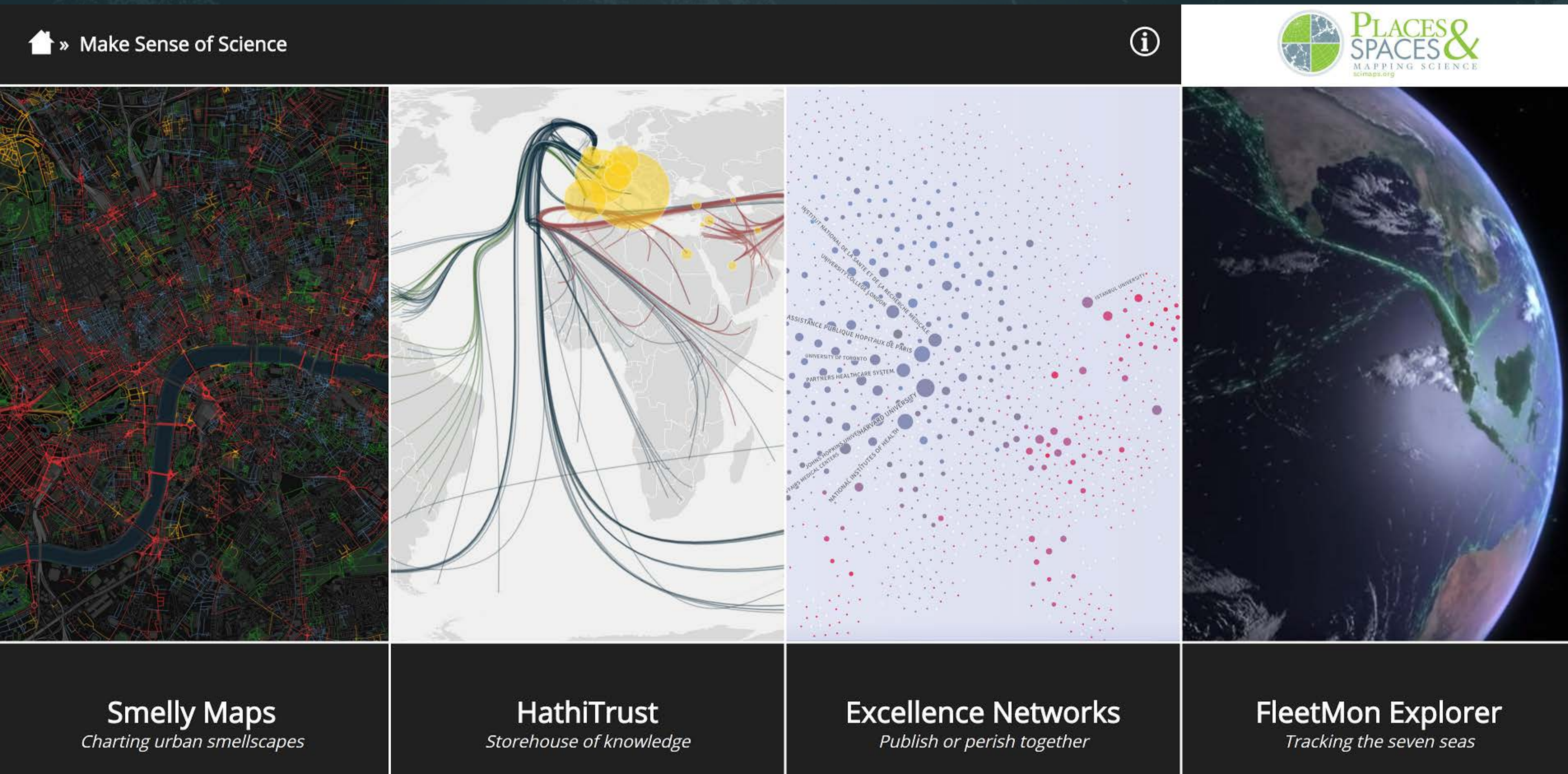
Earth – Cameron Beccario



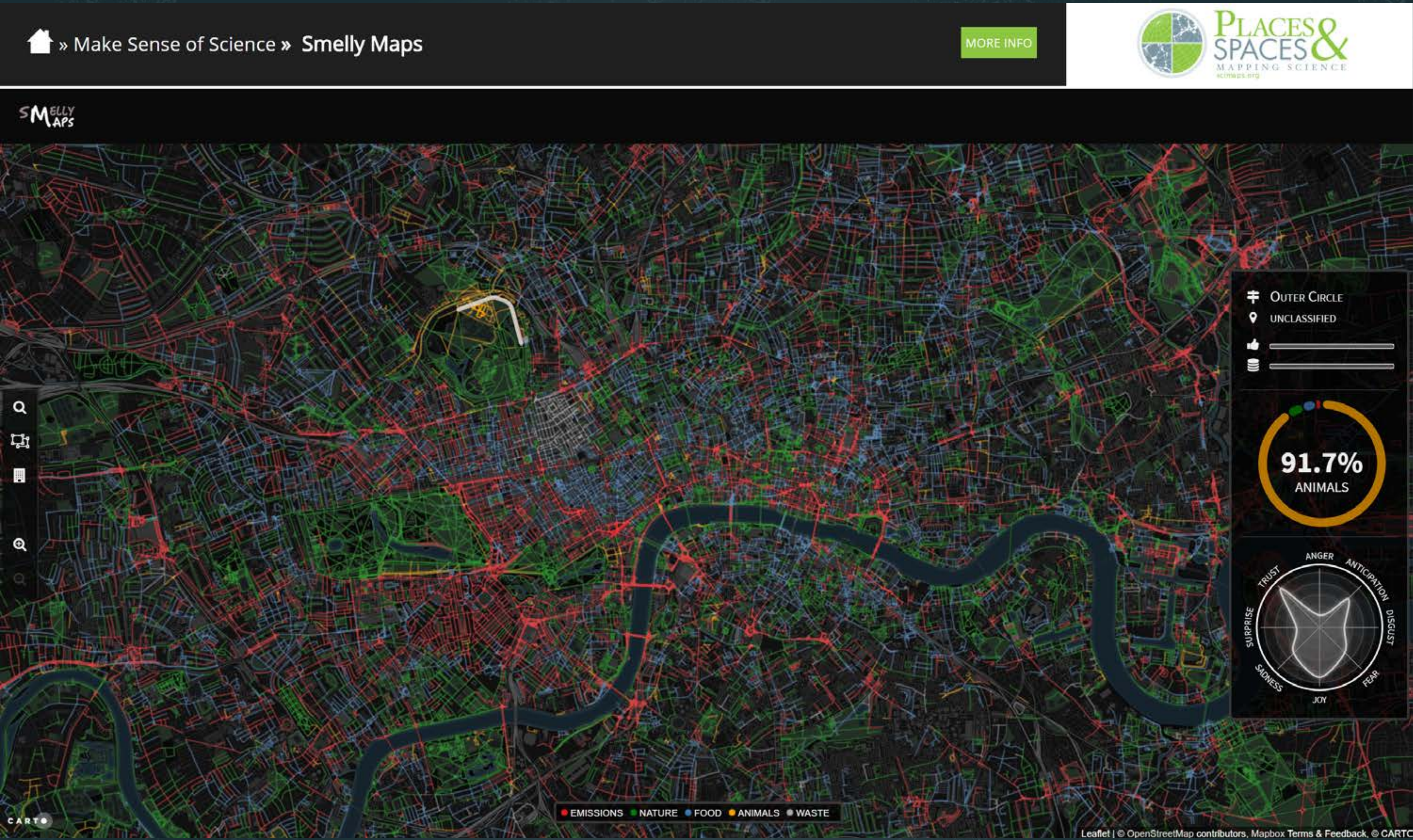
AcademyScope – National Academy of the Sciences & CNS



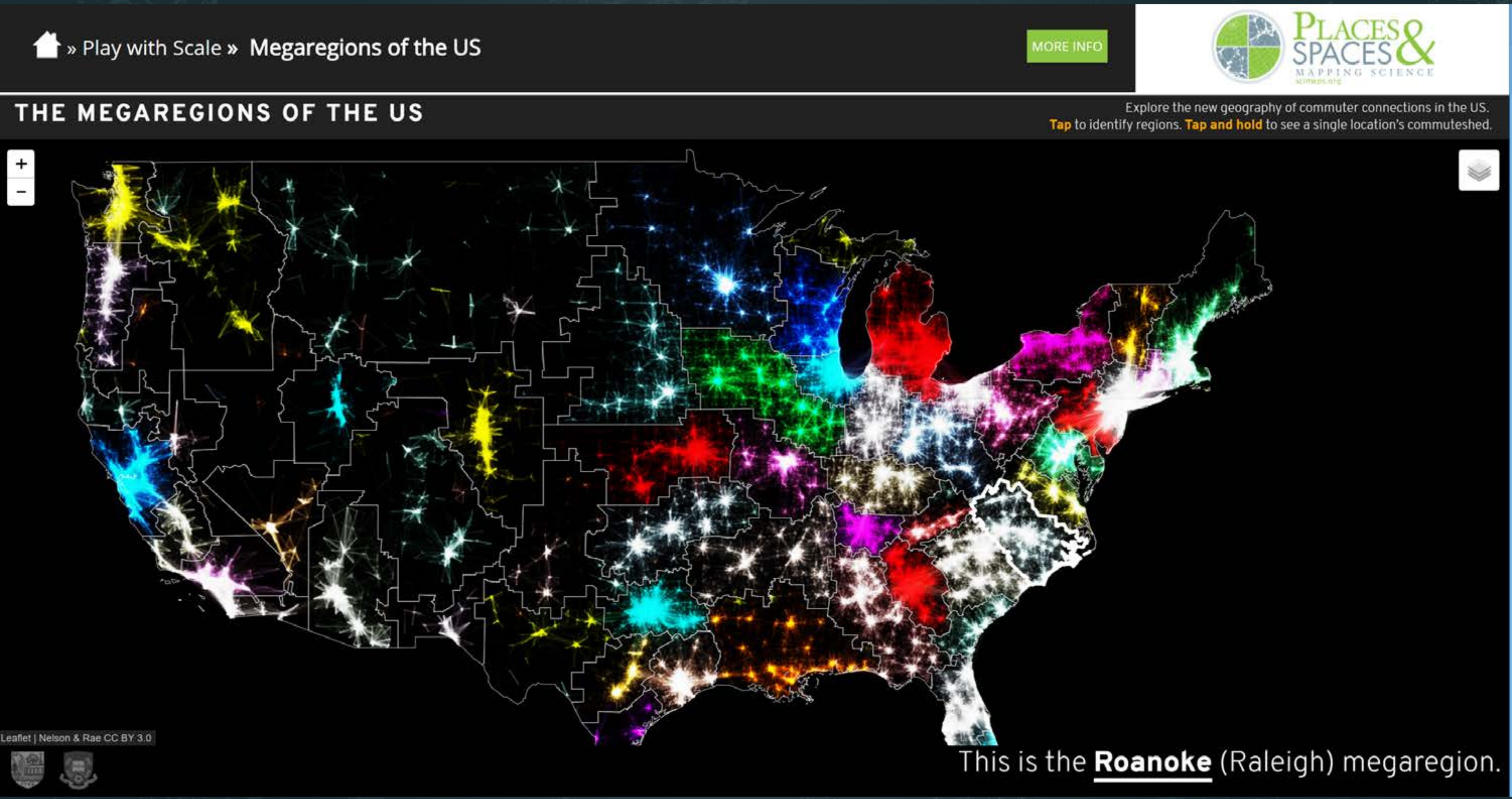
Mapping Global Society –Kalev Leetaru



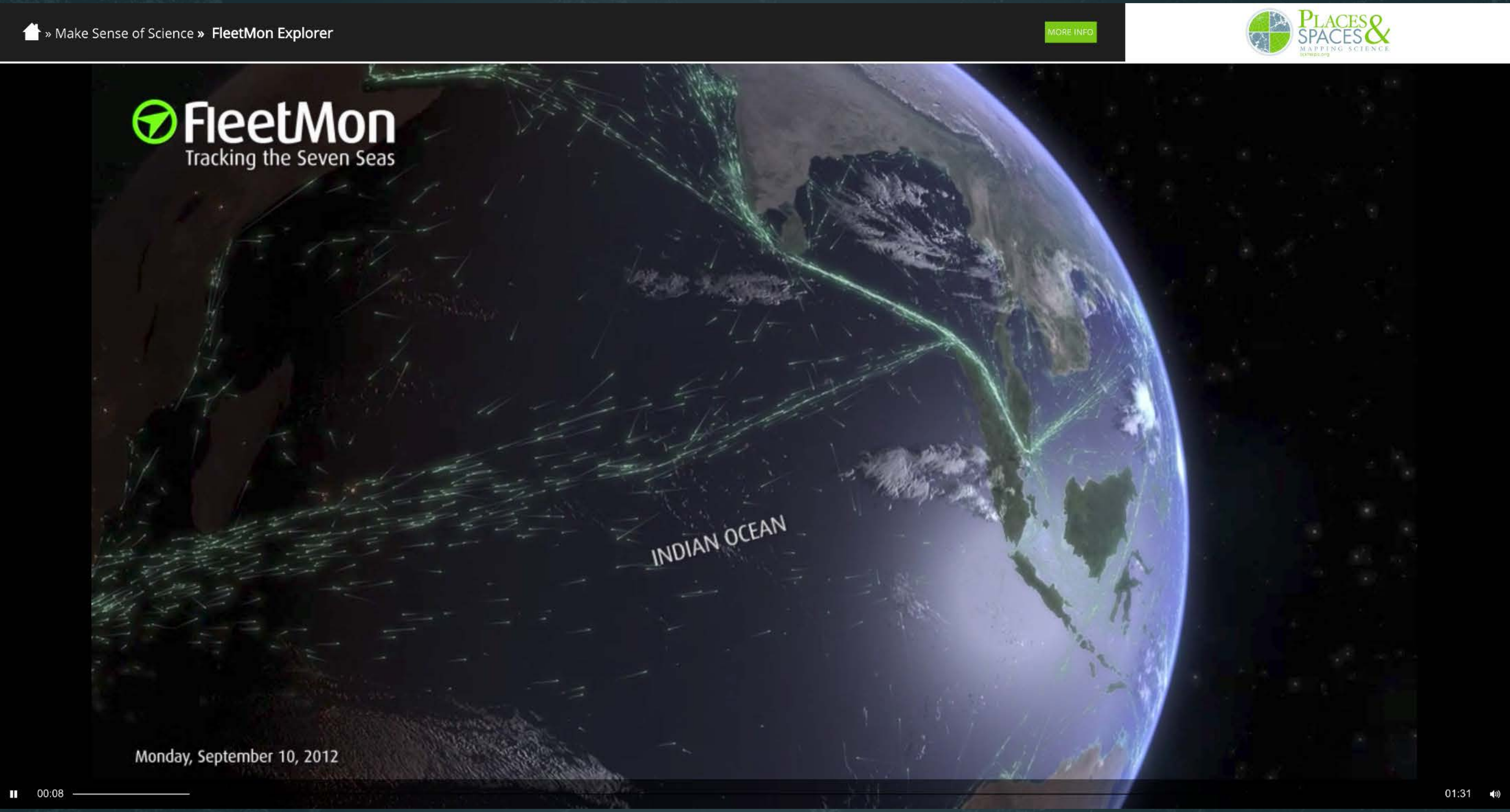
Iteration XII (2016): Macroscopes for Making Sense of Science
http://scimaps.org/iteration/12



Smelly Maps – Daniele Quercia, Rossano Schifanella, and Luca Maria Aiello – 2015



Megaregions of the US –Garrett Dash Nelson and Alasdair Rae – 2016



FleetMon Explorer – FleetMon – 2012

Maps of Science & Technology

http://scimaps.org

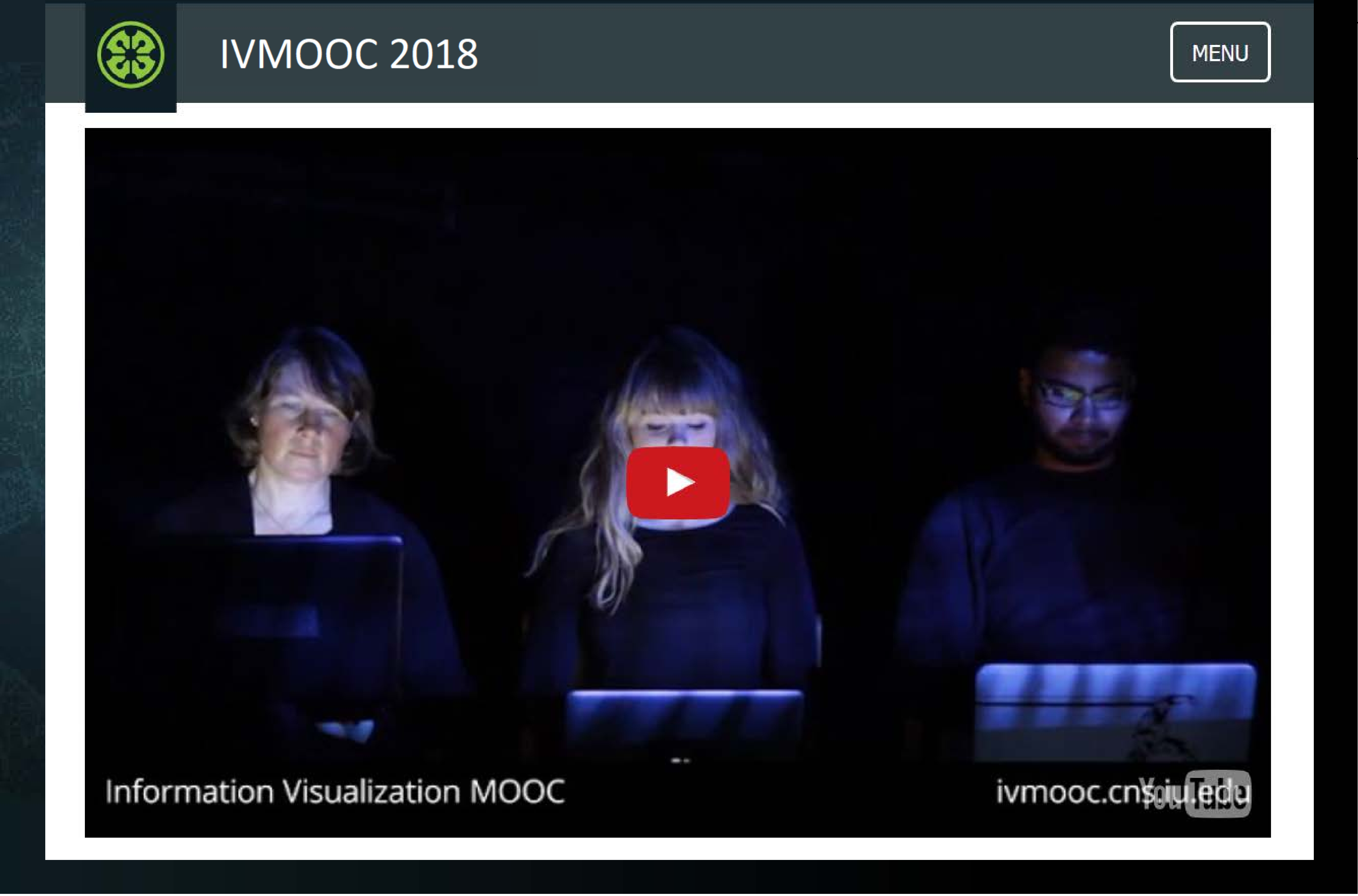
101st Annual Meeting of the Association of American Geographers, Denver, CO.
April 5th - 9th, 2005 (First showing of Places & Spaces)

University of Miami, Miami, FL.
September 4 - December 11, 2014.

Duke University, Durham, NC.
January 12 - April 10, 2015

The David J. Sencer CDC Museum, Atlanta, GA.
January 25 - June 17, 2016.

100 maps and 12 macroscopes by 215 experts on display at 354 venues in 28 countries.



Register for free: http://ivmooc.cns.iu.edu. Class started Jan 9, 2018.

References

Börner, Katy, Chen, Chaomei, and Boyack, Kevin. (2003). **Visualizing Knowledge Domains.** In Blaise Cronin (Ed.), *ARIST*, Medford, NJ: Information Today, Volume 37, Chapter 5, pp. 179-255. http://ivl.slis.indiana.edu/km/pub/2003-borner-arist.pdf

Shiffrin, Richard M. and Börner, Katy (Eds.) (2004). **Mapping Knowledge Domains**. *Proceedings of the National Academy of Sciences of the United States of America*, 101(Suppl_1). http://www.pnas.org/content/vol101/suppl_1

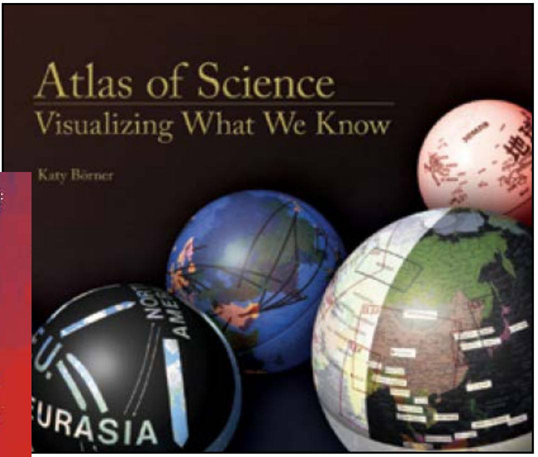
Börner, Katy (2010) **Atlas of Science: Visualizing What We Know**. The MIT Press. http://scimaps.org/atlas

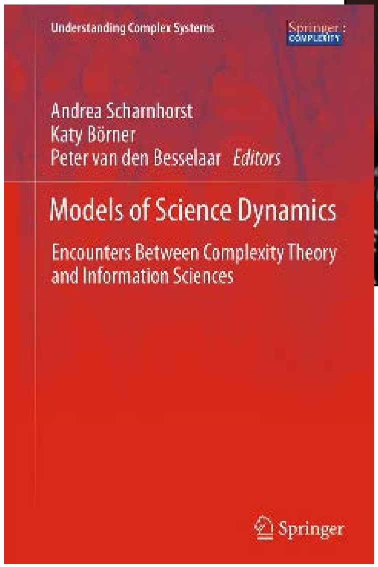
Scharnhorst, Andrea, Börner, Katy, van den Besselaar, Peter (2012) **Models of Science Dynamics**. Springer Verlag.

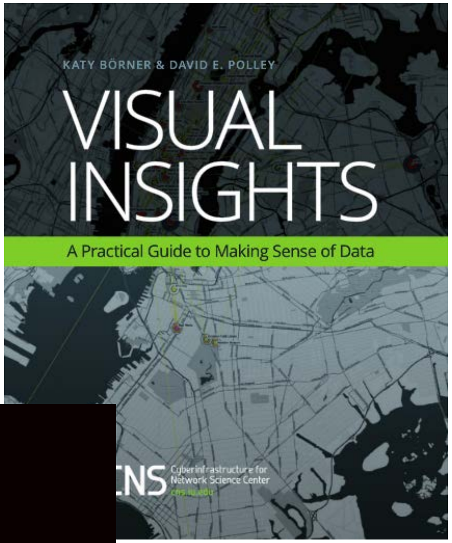
Katy Börner, Michael Conlon, Jon Corson-Rikert, Cornell, Ying Ding (2012) **VIVO: A Semantic Approach to Scholarly Networking and Discovery**. Morgan & Claypool.

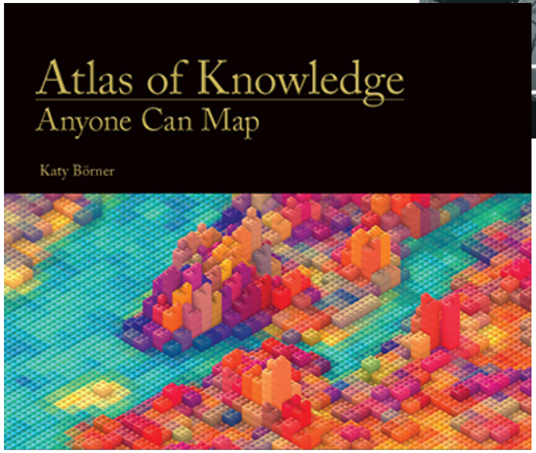
Katy Börner and David E Polley (2014) **Visual Insights: A Practical Guide to Making Sense of Data**. The MIT Press.

Börner, Katy (2015) **Atlas of Knowledge: Anyone Can Map**. The MIT Press. http://scimaps.org/atlas2



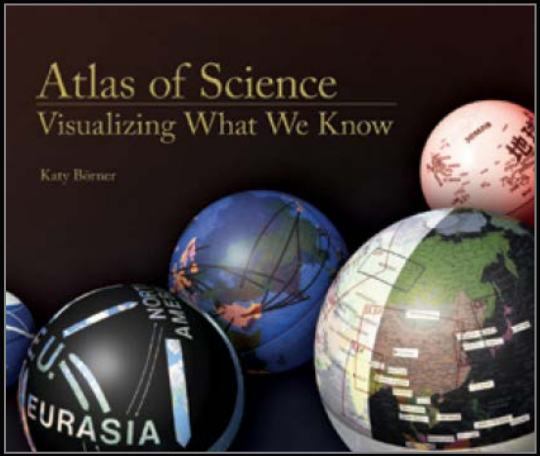


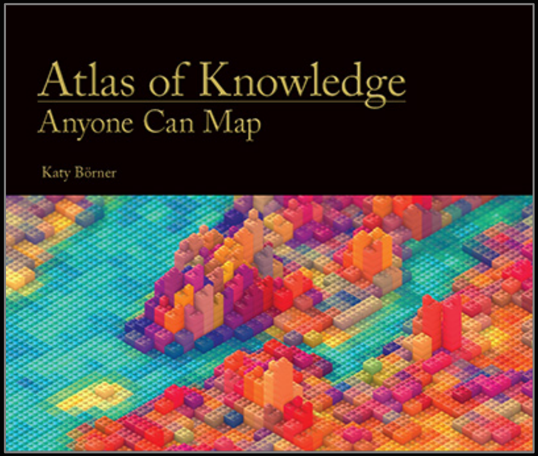




Interactive Visualizations of (NIH/NSF) Infrastructure Impact

Using large scale datasets, advanced data mining and visualization techniques, and substantial computing resources.





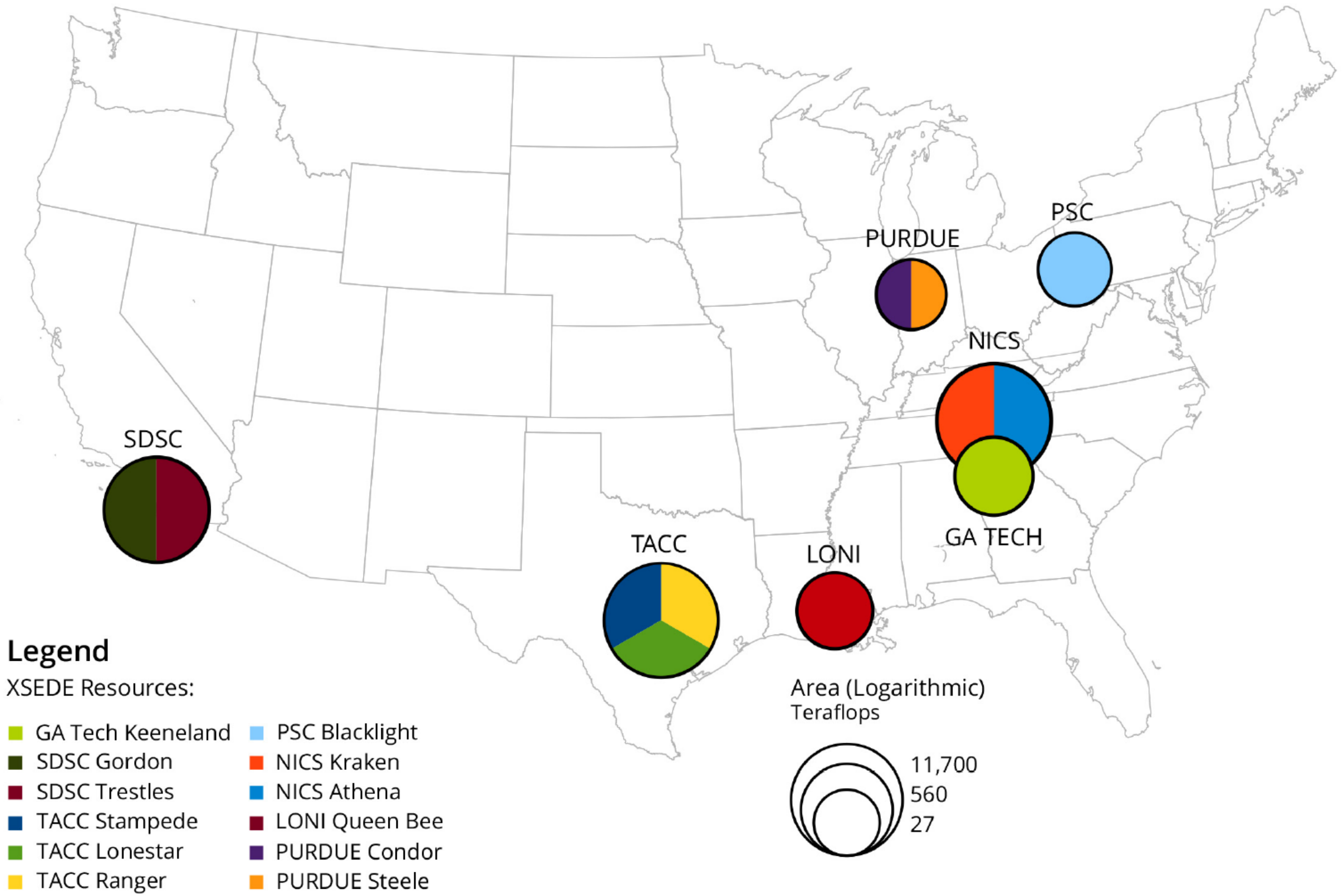


Fig 1. Service Provider Resources, sized by Teraflops.

doi:10.1371/journal.pone.0157628.g001

Knepper, Richard, and Katy Börner. 2016. "Comparing the Consumption of CPU Hours with Scientific Output for the Extreme Science and Engineering Discovery Environment (XSEDE)". *PLoS ONE* 11 (6): e0157628. doi: 10.1371/journal.pone.0157628.

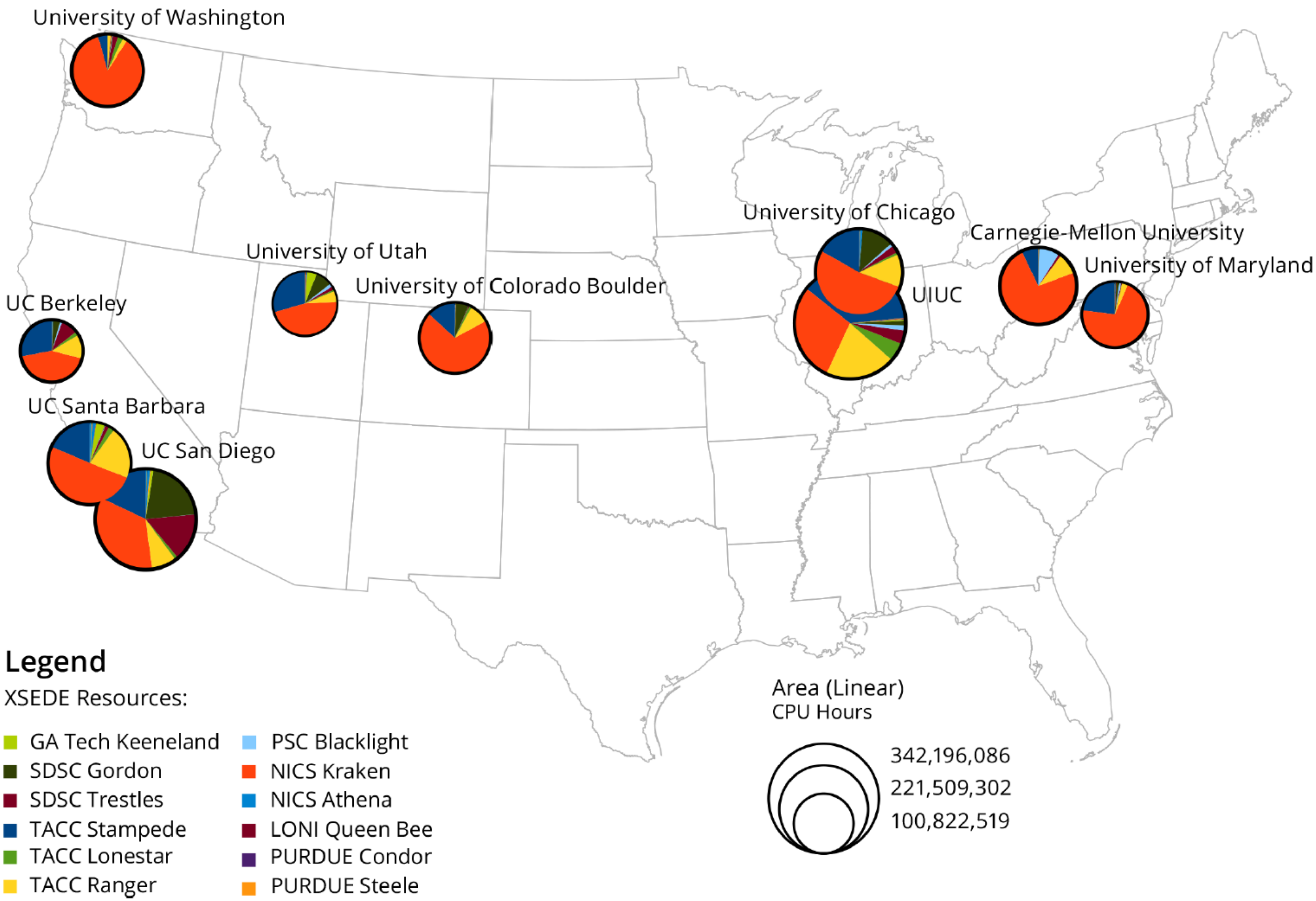


Fig 2. Top 10 Consumers of XSEDE Resources.

doi:10.1371/journal.pone.0157628.g002

Knepper, Richard, and Katy Börner. 2016. ["Comparing the Consumption of CPU Hours with Scientific Output for the Extreme Science and Engineering Discovery Environment (XSEDE)"](). *PLoS ONE* 11 (6): e0157628. doi: 10.1371/journal.pone.0157628.

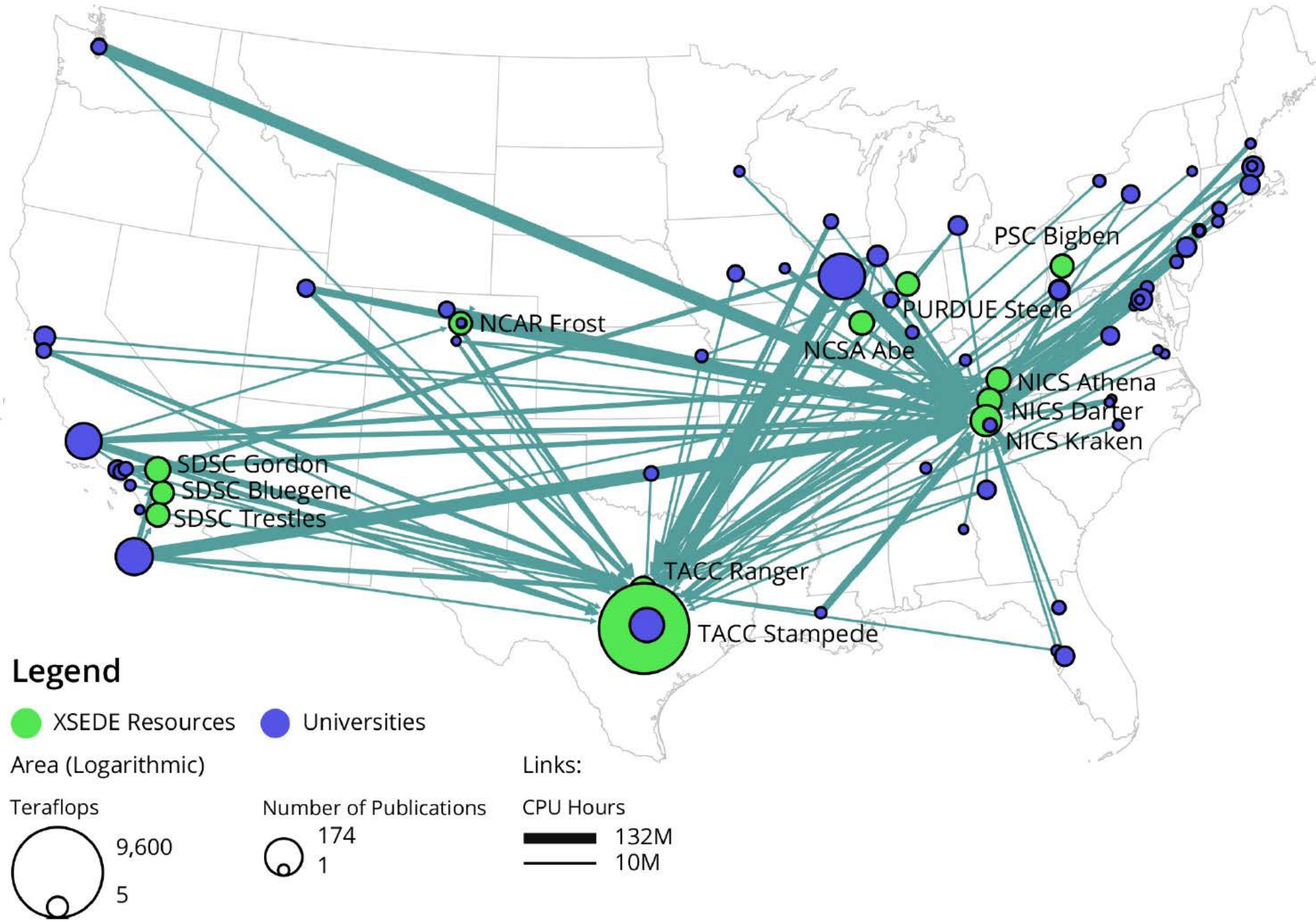
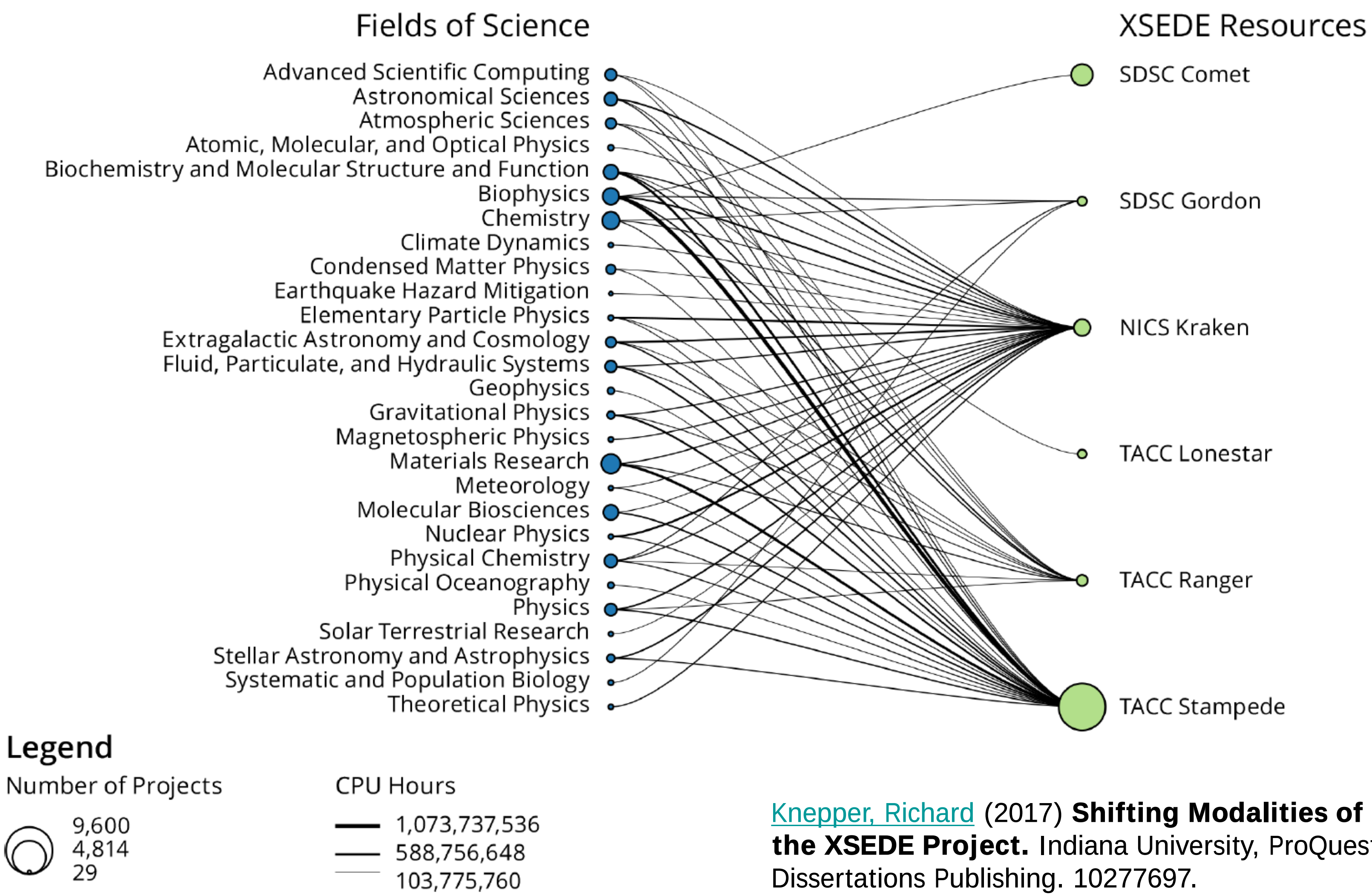


Fig 5. Geographic Layout of XSEDE Resources and University Consumers.

doi:10.1371/journal.pone.0157628.g005

Knepper, Richard, and Katy Börner. 2016. "Comparing the Consumption of CPU Hours with Scientific Output for the Extreme Science and Engineering Discovery Environment (XSEDE)". *PLoS ONE* 11 (6): e0157628. doi: 10.1371/journal.pone.0157628.



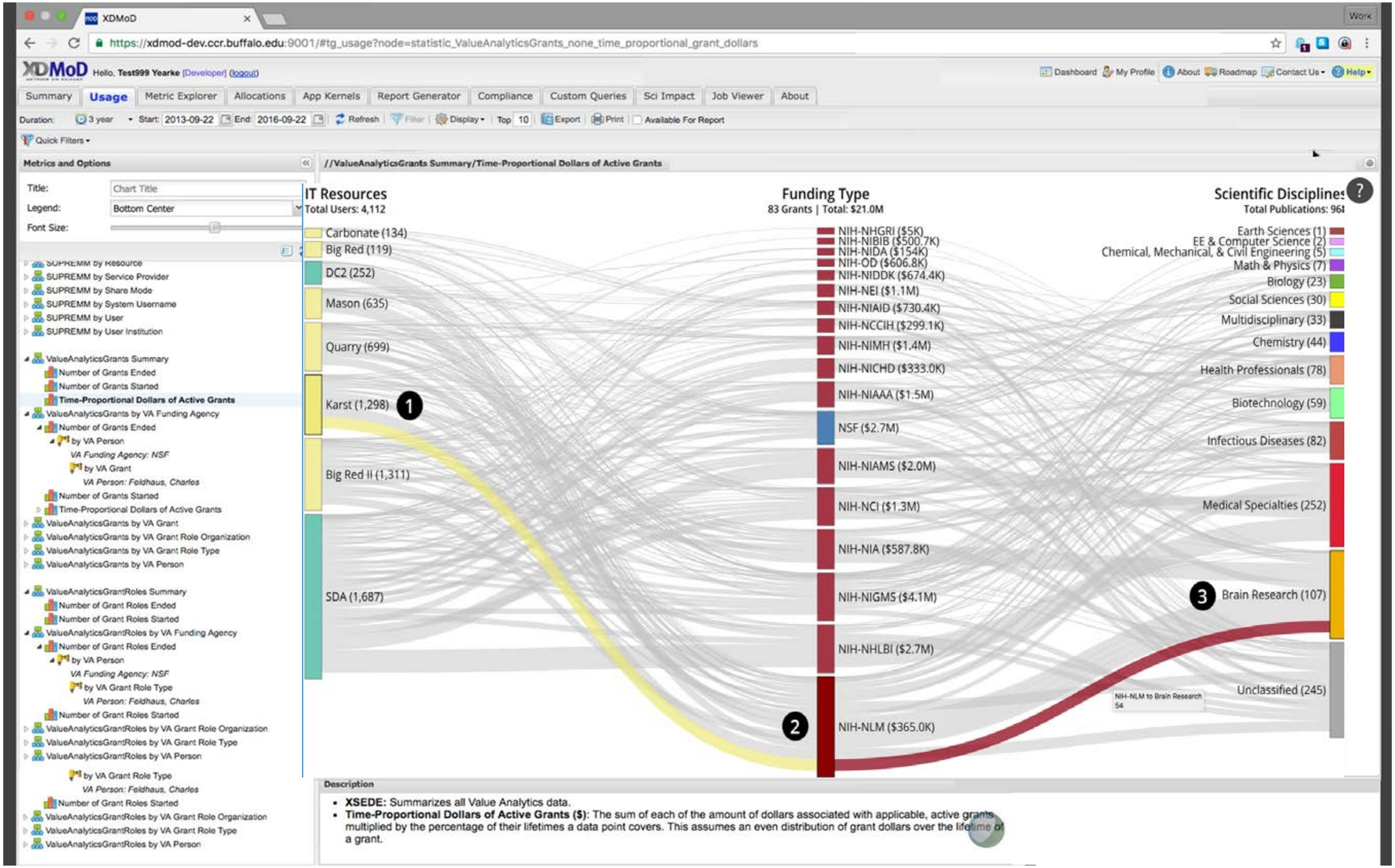
Knepper, Richard (2017) **Shifting Modalities of Use in the XSEDE Project.** Indiana University, ProQuest Dissertations Publishing. 10277697.

Fig 4. Bipartite field of science and resource network.

doi:10.1371/journal.pone.0157628.g004

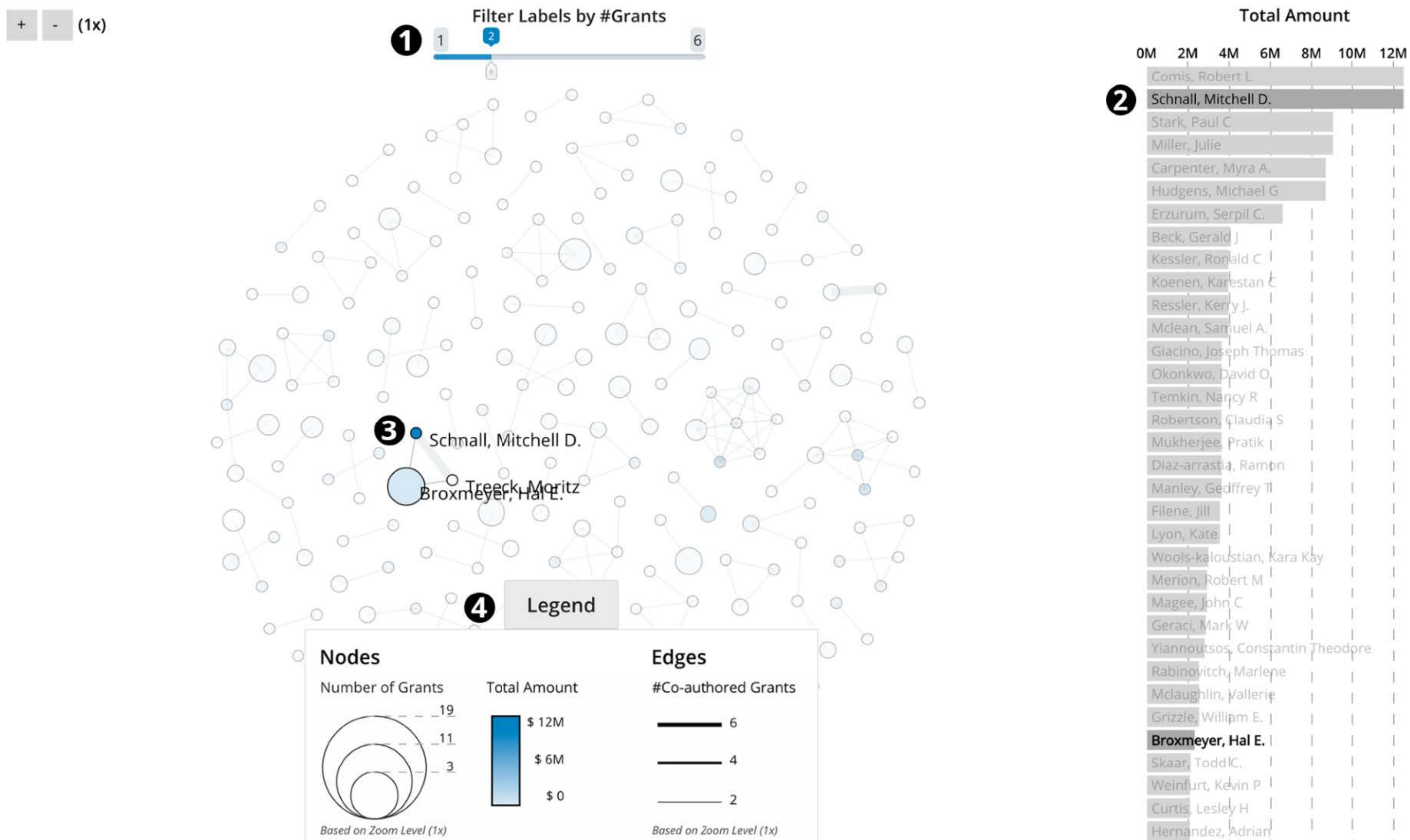
Knepper, Richard, and Katy Börner. 2016. "Comparing the Consumption of CPU Hours with Scientific Output for the Extreme Science and Engineering Discovery Environment (XSEDE)". *PLoS ONE* 11 (6): e0157628. doi: 10.1371/journal.pone.0157628.

XDMOD Sankey Diagram



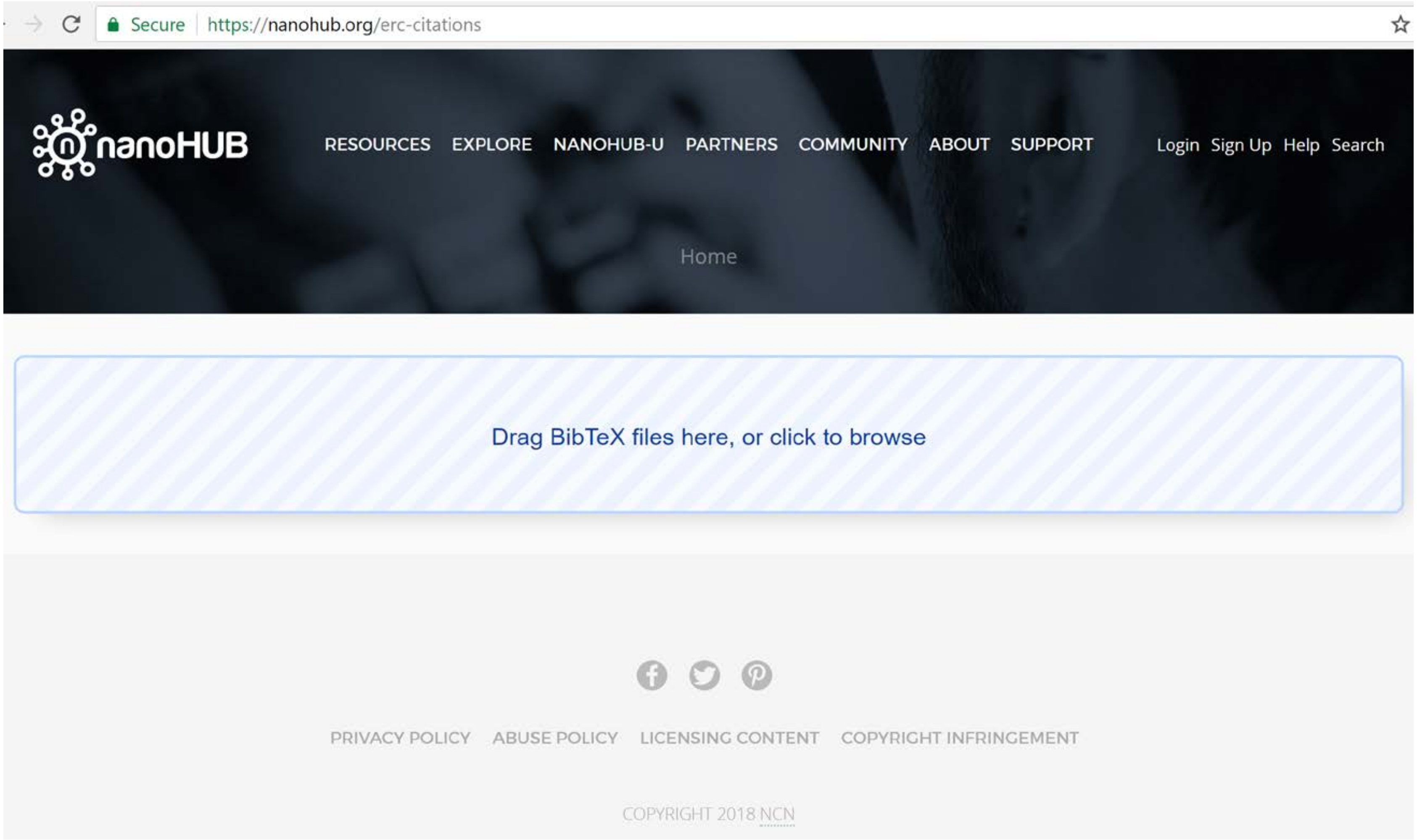
This Sankey diagram displays the relationship between IT resources, funding agencies, and publications. Line width indicates number of dollars awarded.

XDMOD Co-PI Visualization



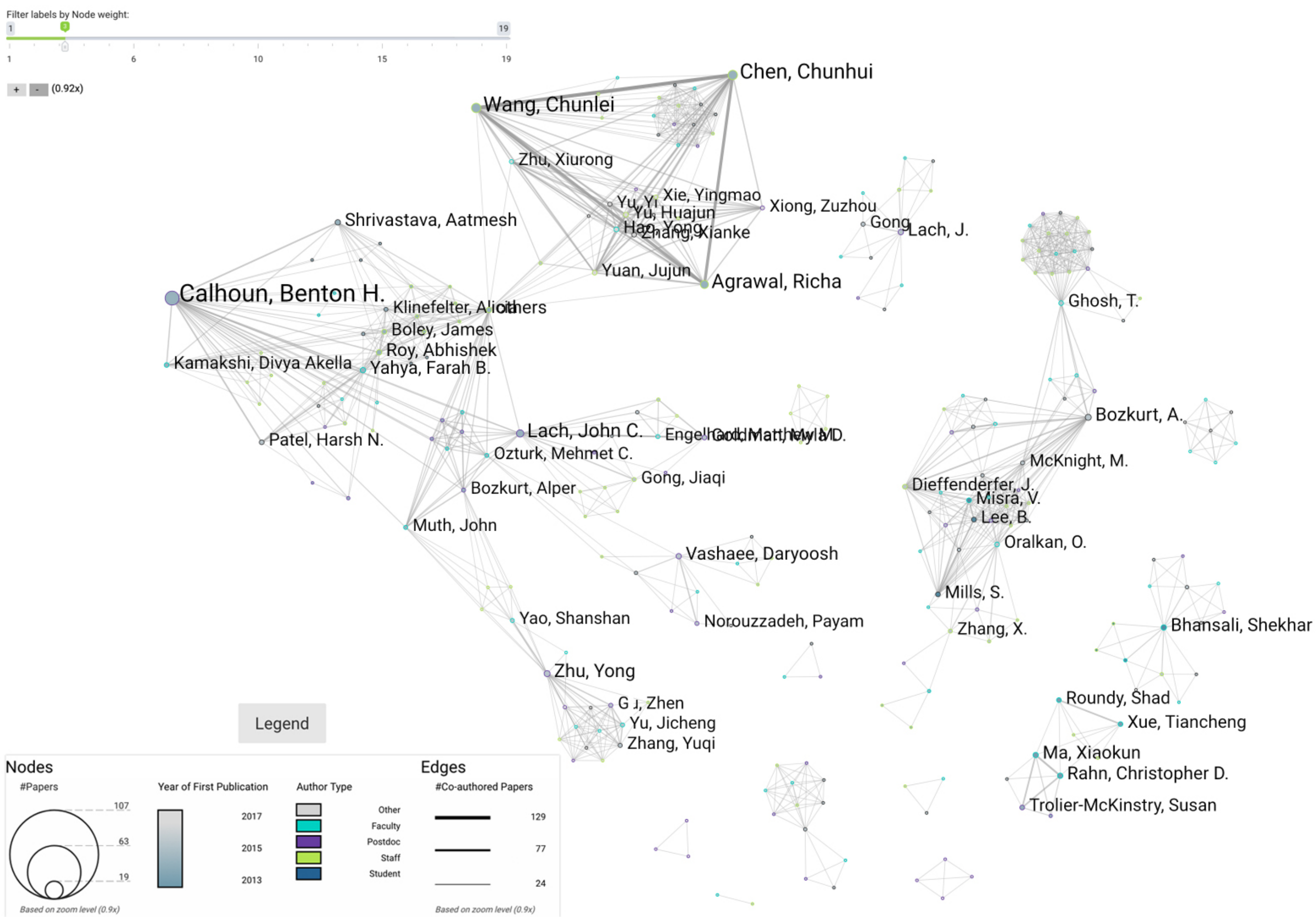
ERC Visualizations: Upload a Bibtex File

https://nanohub.org/erc-citations



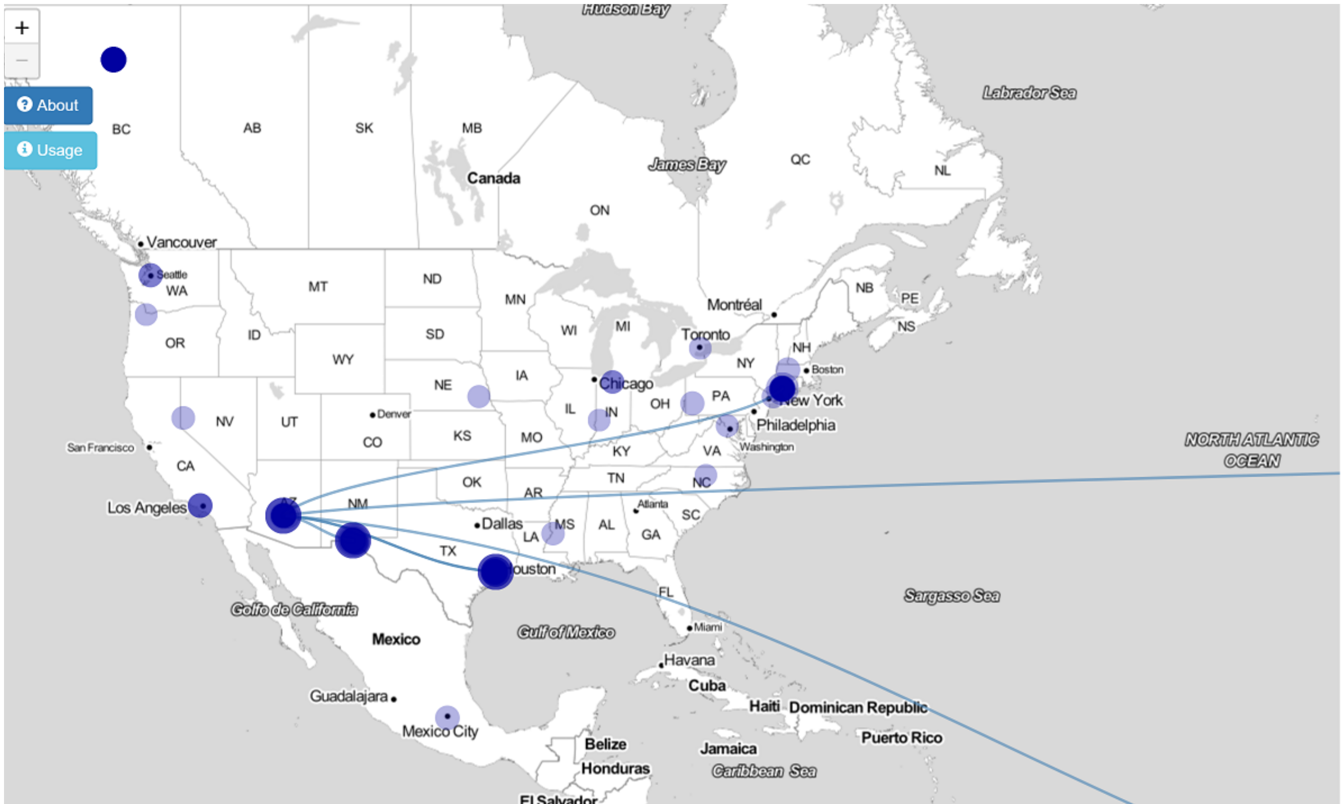
ERC Co-Author Networks

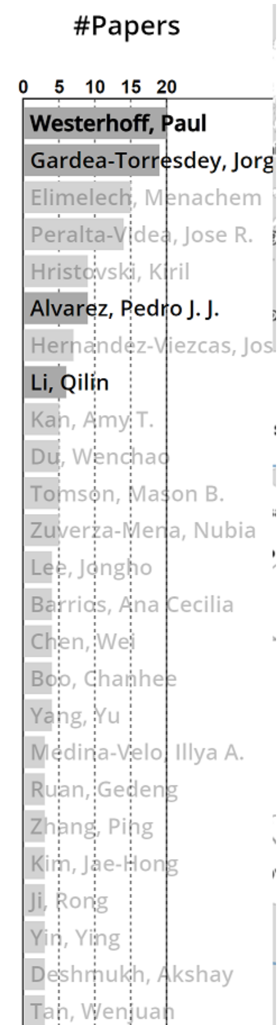
https://nanohub.org/erc-citations



ERC Co-Author Network Geomap

https://nanohub.org/erc-citations



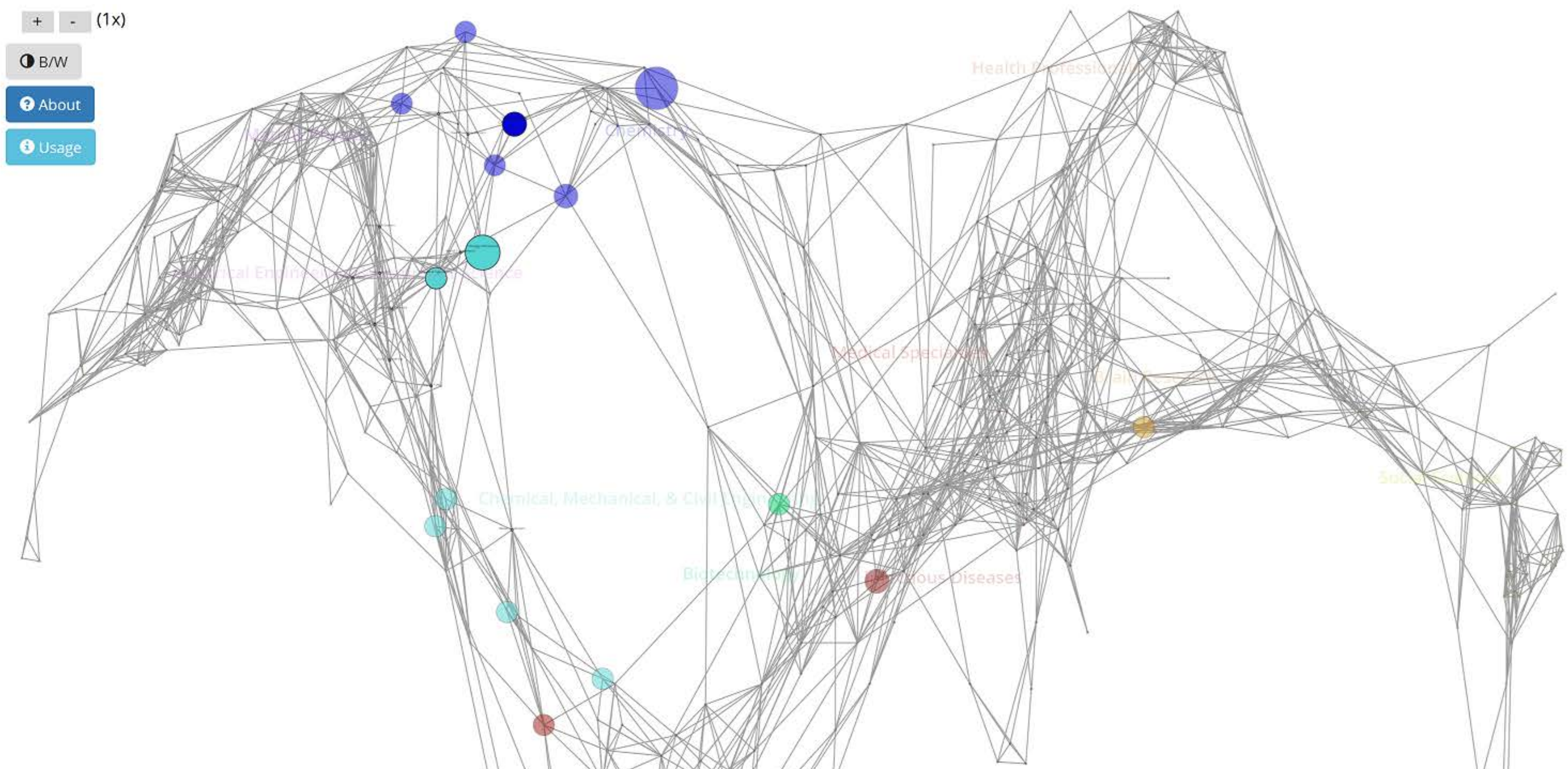


Zip code: 8528_

Authors	Year	Title	Journal
Chen, Si; Yuan, Ziwen; Hanigan, David; Westerhoff, Paul; Zhao, Huazhang; Ni, Jinren	2018	Coagulation behaviors of new covalently bound hybrid coagulants (CBHyC) in surface water treatment	Separation and Purification Technology
Rice, Jacelyn; Westerhoff, Paul	2017	High levels of endocrine pollutants in US	Nature Geoscience

ERC Science Map

https://nanohub.org/erc-citations



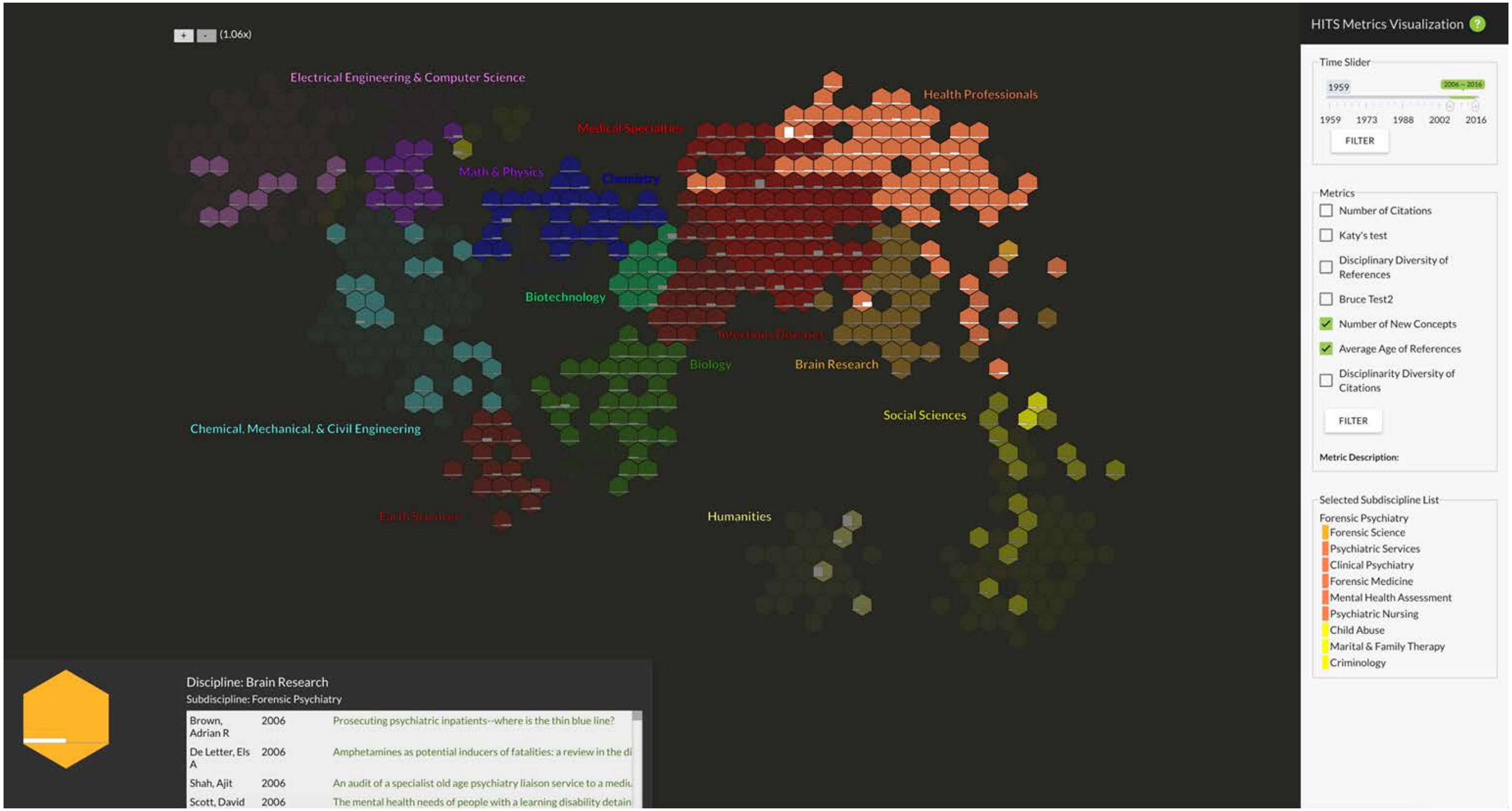
Discipline: Chemical, Mechanical, & Civil Engineering

Subdiscipline: Chemical Engineering

Authors	Year	Title	Journal
Jabbari, V.; Veleta, J. M.; Zarei-Chaleshtori, M.; Gardea-Torresdey, Jorge L.; Villagran, D.	2016	Green synthesis of magnetic MOF@GO and MOF@CNT hybrid nanocomposites with high adsorption capacity towards	Chemical Engineering Journal

ECON Hex Map of Science with HITS Metrix

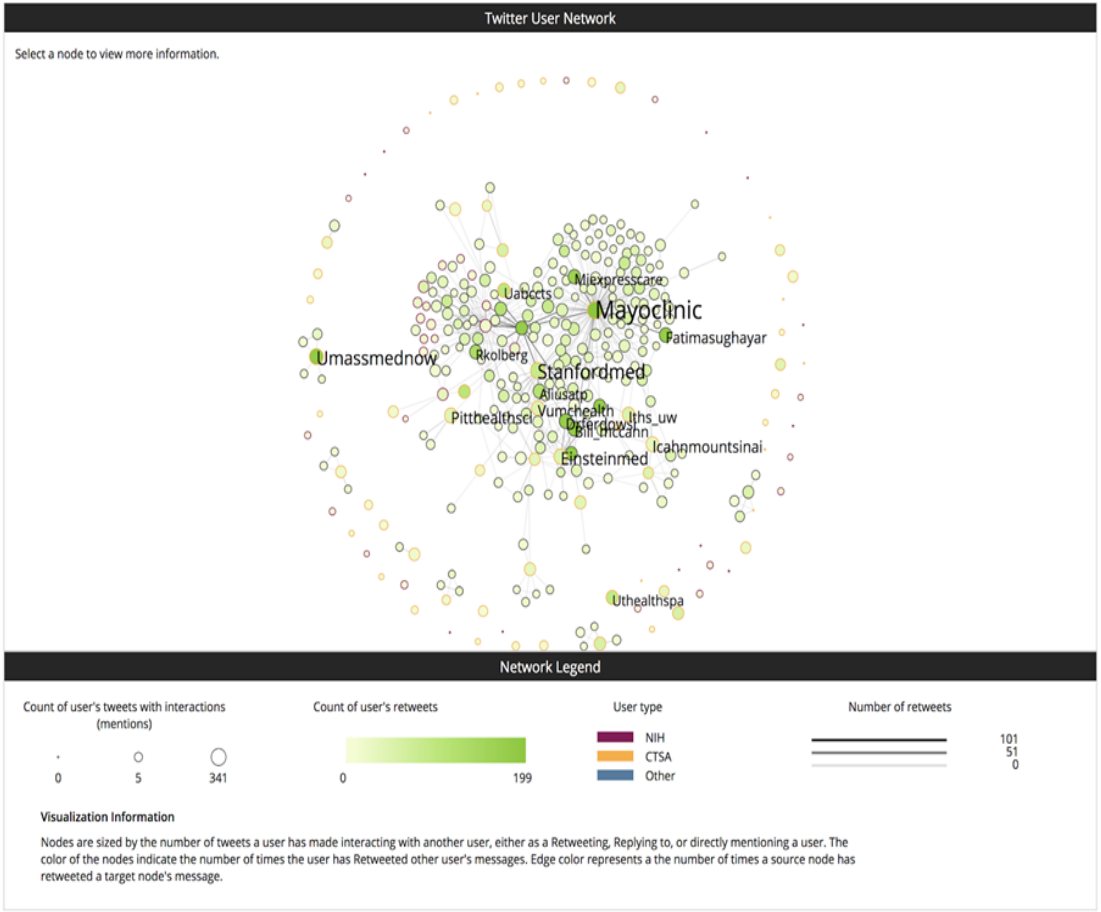
demo.cns.iu.edu/client/econ-hexmap

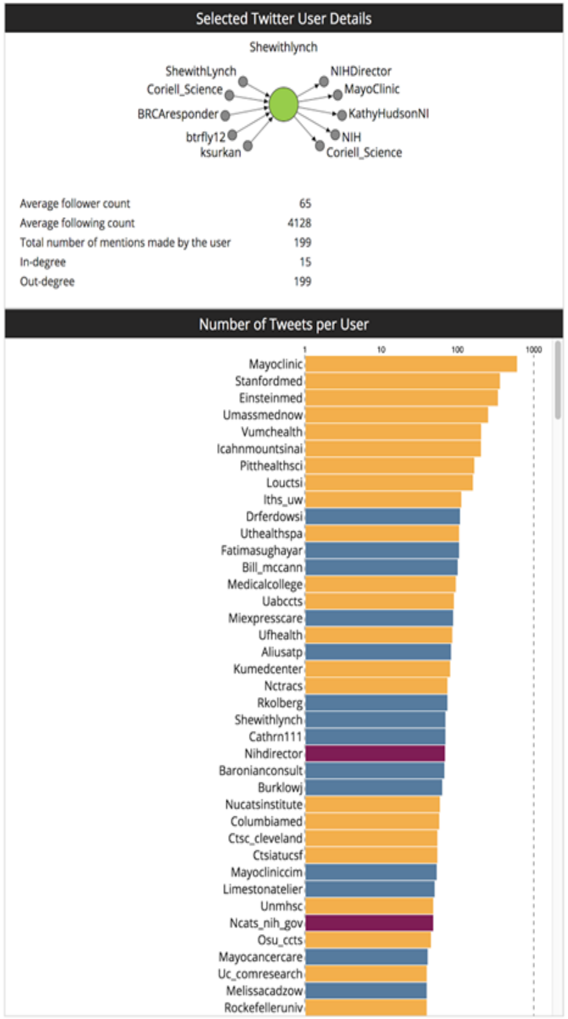


Hex-style rendering of the UCSD map of science allows for easier investigation of metric bars within each hexagon node. This Interactive map allows the user to hover over a discipline label to highlight all its subdisciplines, and also hover over a subdiscipline to see all subdisciplines it is connected to.

NIH Twitter Network

demo.cns.iu.edu/client/iai/twitter.html

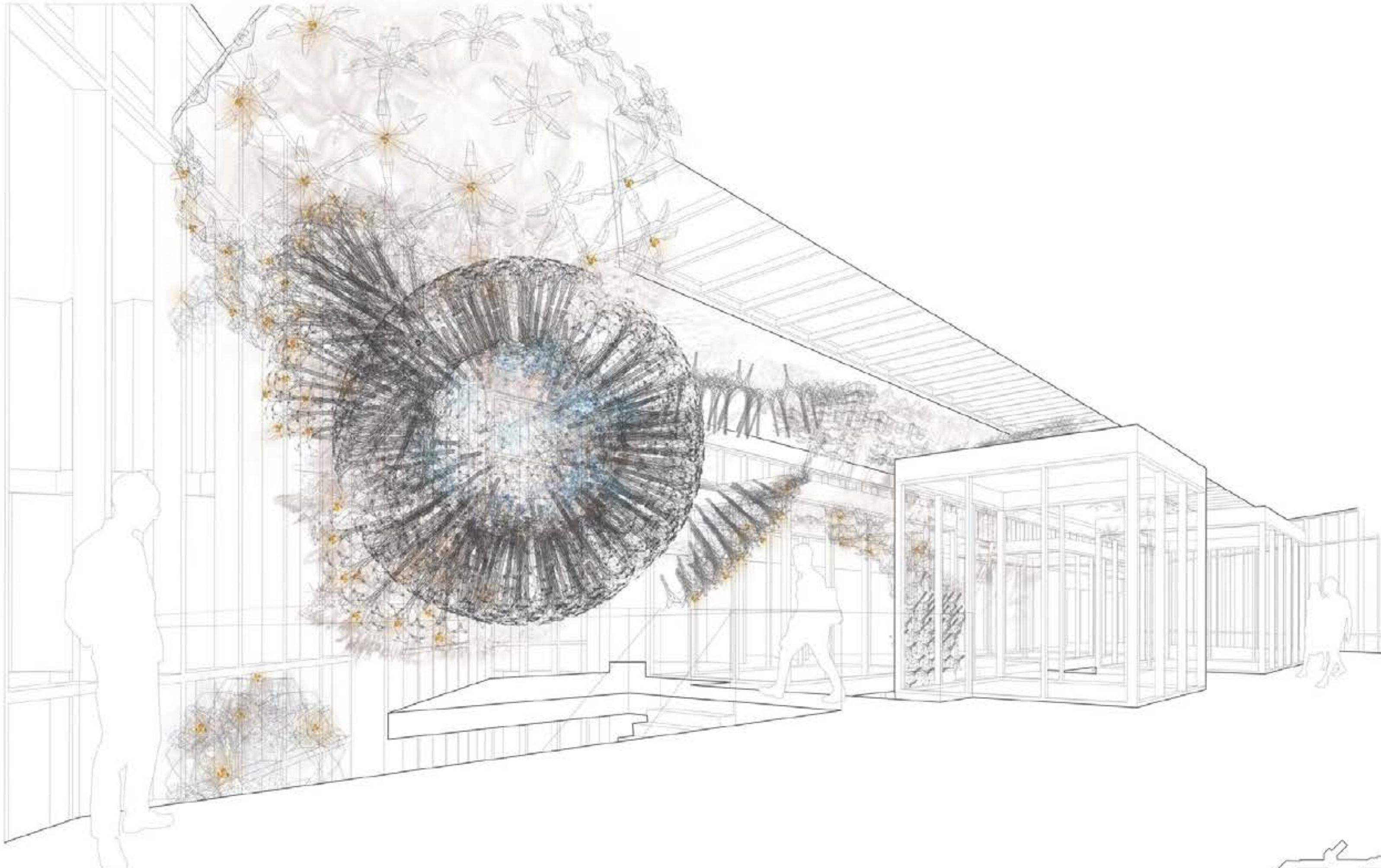


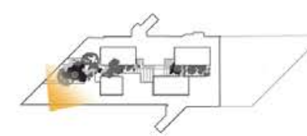


Visualizing the Internet of Things (IoT)

Using large scale datasets, advanced data mining and visualization techniques, and substantial computing resources.

Work by Philip Beesley | www.philipbeesley.ca | www.lasg.ca

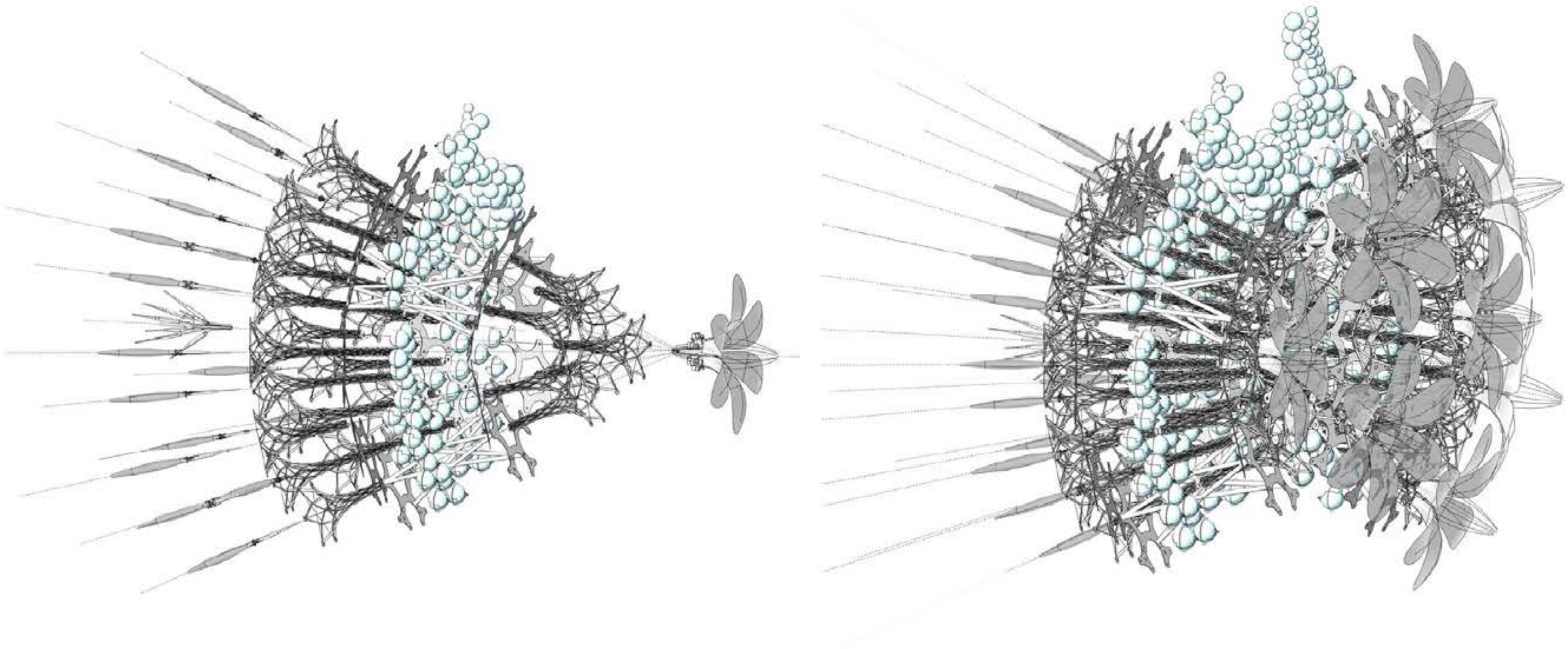


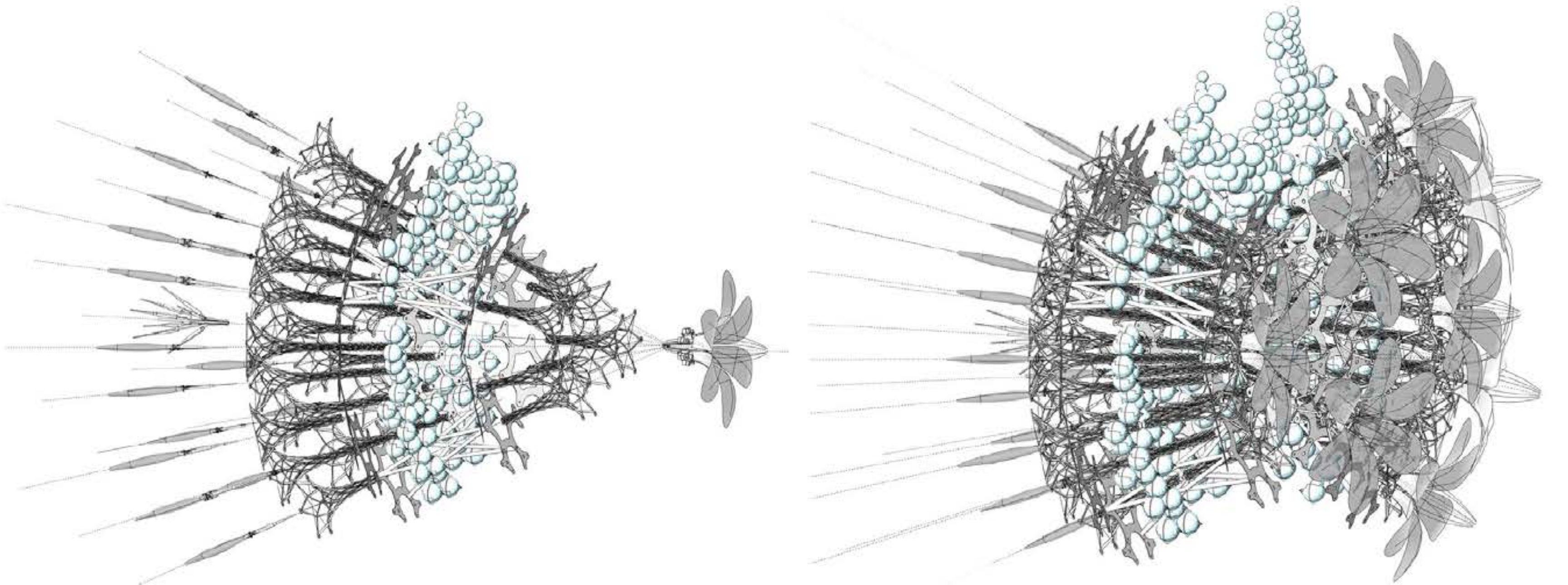


Luddy Hall Installation
Indiana University Bloomington
April 29 2017

UPPER ATRIUM

Philip Beesley • Living Architecture Systems





Come join us for

- April 9-12, Luddy Hall Opening
- April 12-13, Research Horizons.
  2016 pictures, slides, videos are at
  https://researchhorizons.soic.indiana.edu

Luddy Hall Installation
Indiana University Bloomington
April 29 2017

Philip Beesley • Living Architecture Systems

Intelligent Systems Engineering Summer Camps

2017 Info: http://camps.engineering.indiana.edu

Dendrites: http://cns.iu.edu/camps.html

Dendrite

2017 Student-Teacher Team

Challenges & Opportunities

For using large scale datasets, advanced data mining and visualization techniques, and substantial computing resources.

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

Conference slides, recordings, and **NSF Report** are available via http://modsti.cns.iu.edu/report

STEAM Observatory Challenges

- **Fundamental Research:** Few models scale and are validated using large-scale empirical data.
- **Applied Research:** Few models are production-strength, i.e., validated, well documented, 24/7 services. Few active partnerships among academia, government, and industry exist.
- **Cyberinfrastructure:** Unconnected silos of data and code repositories exist in different areas of science. Scholarly results are published in many different areas of science.
- **Education:** There is a need for improved “model literacy” and for a community of teachers/students that share data, code, results, training materials, etc.
- **Outreach:** Scientists need to be more actively involved in communicating with the public and engaging with other communities.

Details are in **NSF Report** at http://modsti.cns.iu.edu/report

STEAM Observatory Opportunities

- **Model Needs and Implementation:** Successful modeling requires close collaboration and active partnership between (policy) decision makers and researchers to ensure the usefulness of the models and increase the chances for their adoption.
- **Data Infrastructure:** Data quality, coverage, and richness are improving rapidly and support the design and validation of detailed models. Easy access to relevant data supports reproducibility. Many high-quality datasets are held by industry or government, hence close academia-industry-government partnerships are desirable.
- **Code Repository and Standards:** Need to define and adopt modeling guidelines and standards, create shared data and model code infrastructure.
- **Visualization and Communication of Modeling Results:** Communicate data quality, model complexity, and modeling results clearly to different stakeholders. Have decision makers “fly the future before writing a check.”
- **Funding:** Modeling needs increase with reduction of budgets, exponential growth of scientific productivity, larger team sizes, and higher interdisciplinarity.

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

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