Data-Driven Science Policy

Katy Börner
@katycns

Victor H. Yngve Distinguished Professor of
Engineering and Information Science
Director, Cyberinfrastructure for Network Science Center
Indiana University
United States

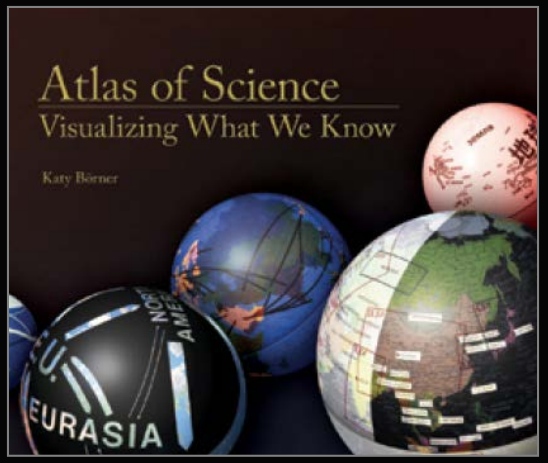
Reference:
Börner, Katy. 2016. "Data-Driven Science Policy". *Issues in Science and Technology* 33 (3): 26-28.

Science Center World Congress, Tokyo, Japan

November 16, 2017

Maps of Science & Technology

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

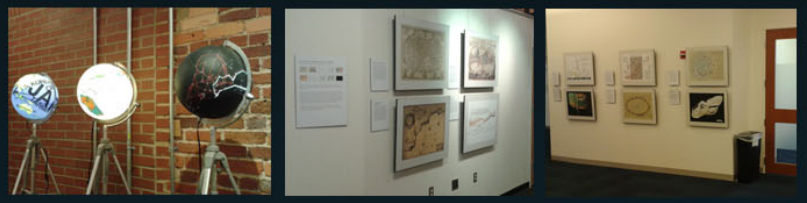


Maps of Science & Technology

http://scimaps.org

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

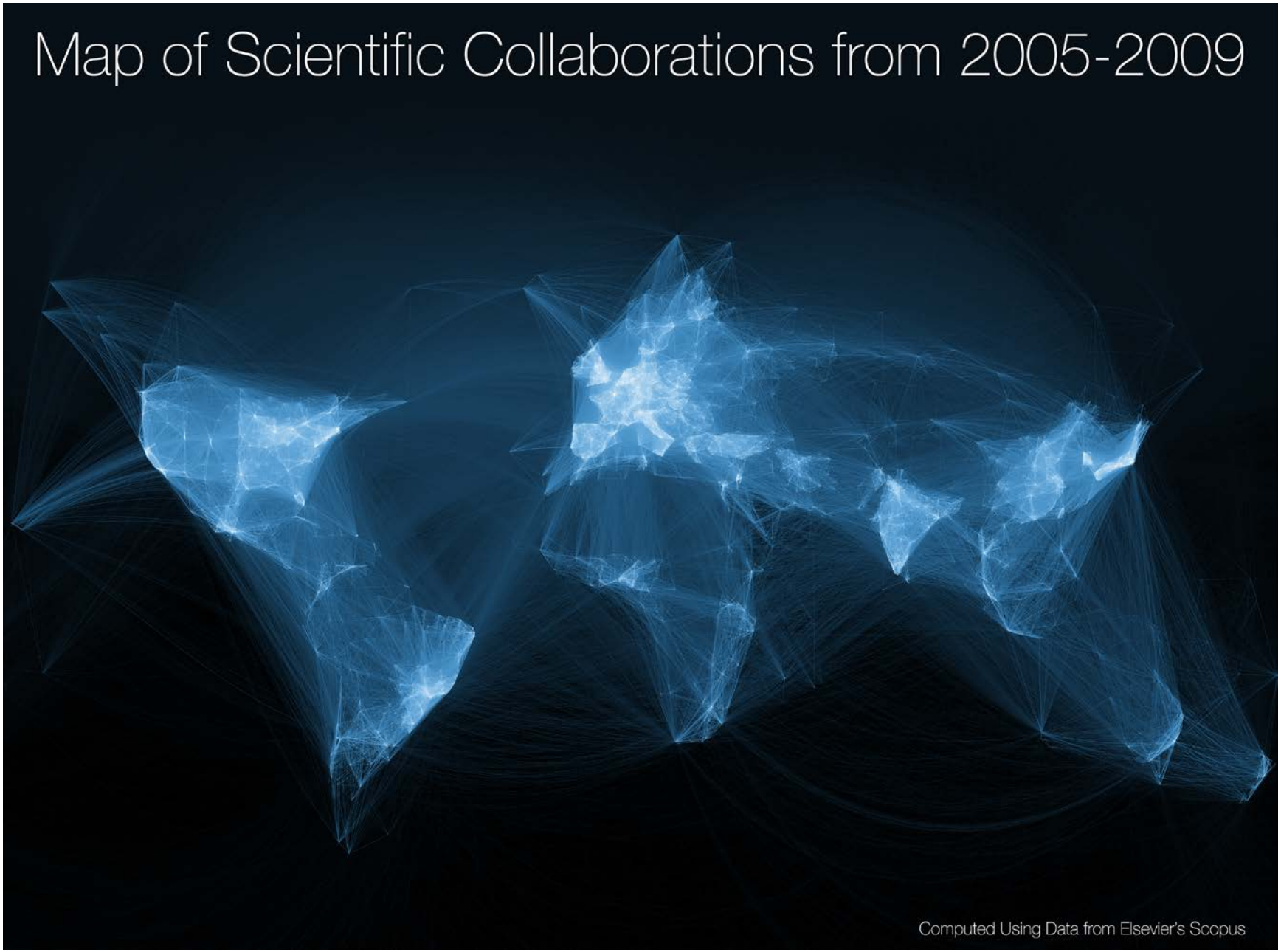
University of Miami, Miami, FL.
September 4 - December 11, 2014.



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.



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

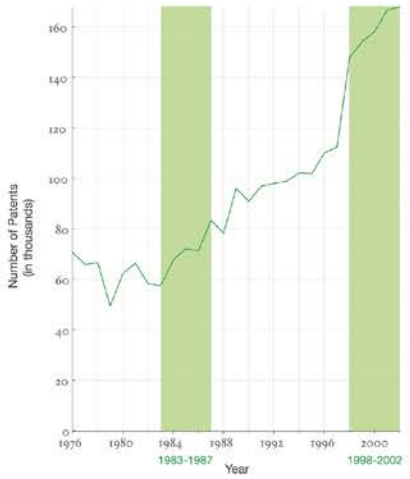
Examining the Evolution & Distribution of Patent Classifications

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."

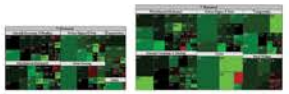
The top 10 fast growing patent classes for 1998–2002 are listed together with the number of patents granted. Most come from the 'Computer and Communications' and the 'Drugs and Medical' area.

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.

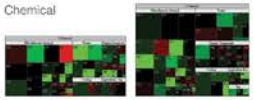
Slow Growing Classes

Mechanical



Others

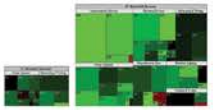
Chemical



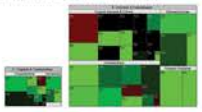
1983 - 1987 1998 - 2002

Fast Growing Classes

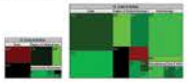
Electrical & Electronic



Computers & Communications



Drugs & Medical



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

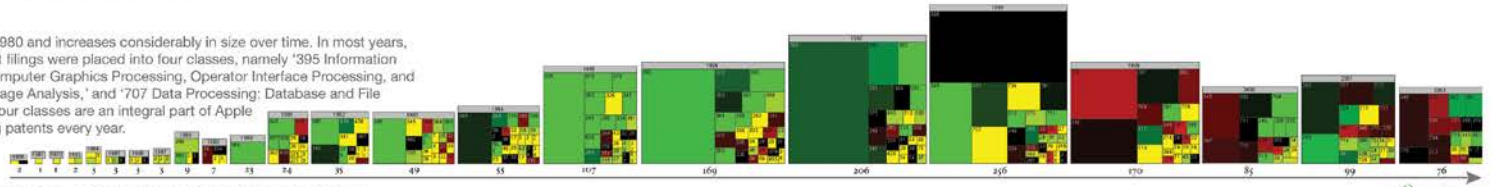
1998 - 2002

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.

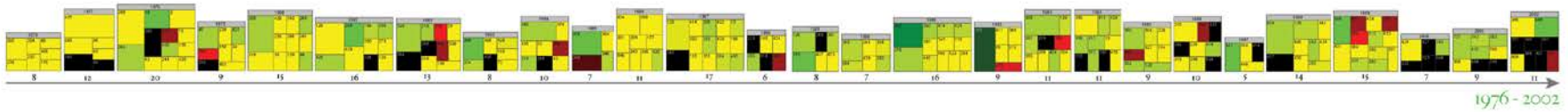
Apple Computer, Inc.

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.



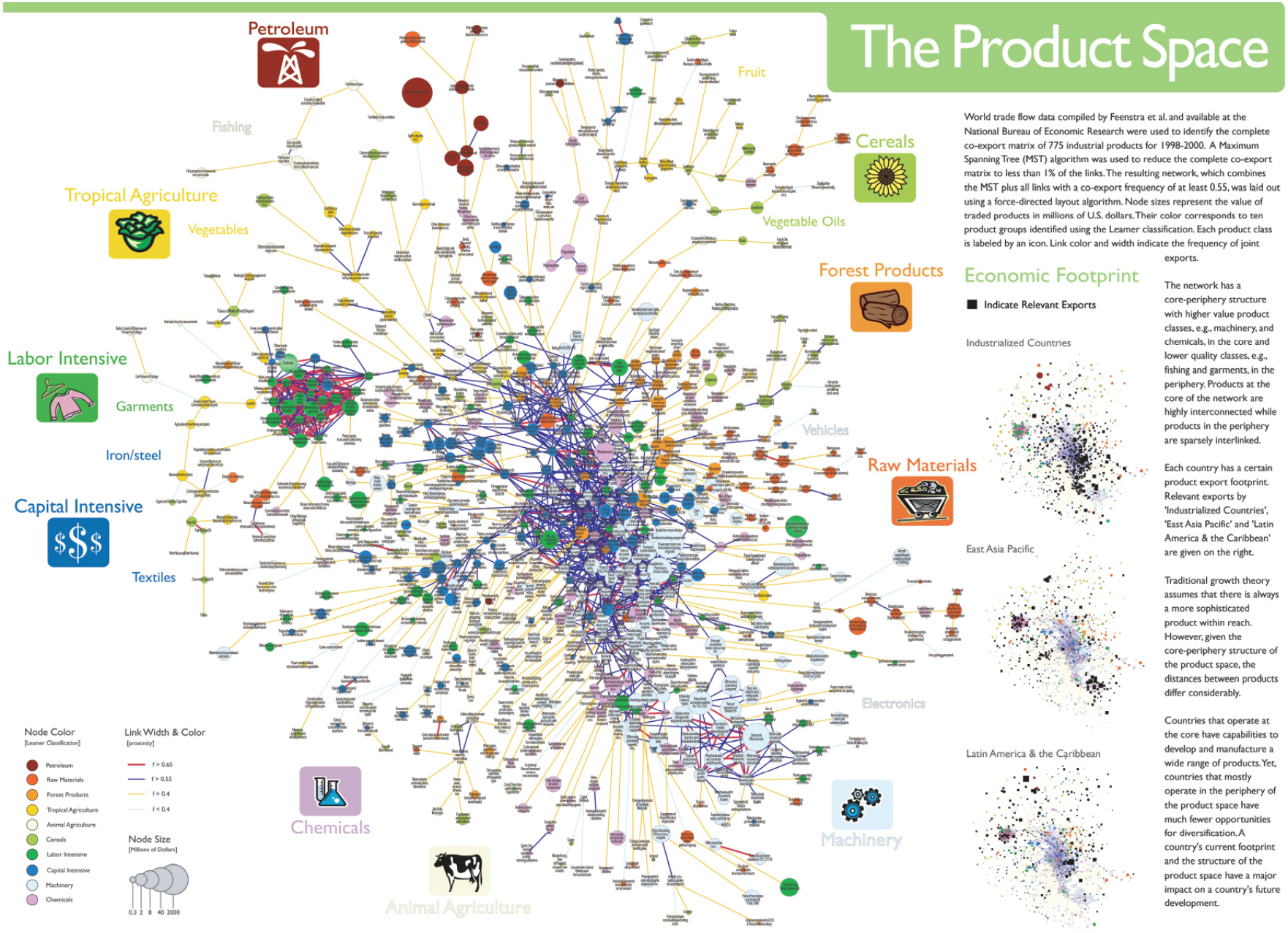
Jerome Lemelson

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

The Product Space



World trade flow data compiled by Feenstra et al. and available at the National Bureau of Economic Research were used to identify the complete co-export matrix of 775 industrial products for 1998-2000. A Maximum Spanning Tree (MST) algorithm was used to reduce the complete co-export matrix to less than 1% of the links. The resulting network, which combines the MST plus all links with a co-export frequency of at least 0.55, was laid out using a force-directed layout algorithm. Node sizes represent the value of traded products in millions of U.S. dollars. Their color corresponds to ten product groups identified using the Leamer classification. Each product class is labeled by an icon. Link color and width indicate the frequency of joint exports.

Economic Footprint

■ Indicate Relevant Exports

Industrialized Countries

East Asia Pacific

Latin America & the Caribbean

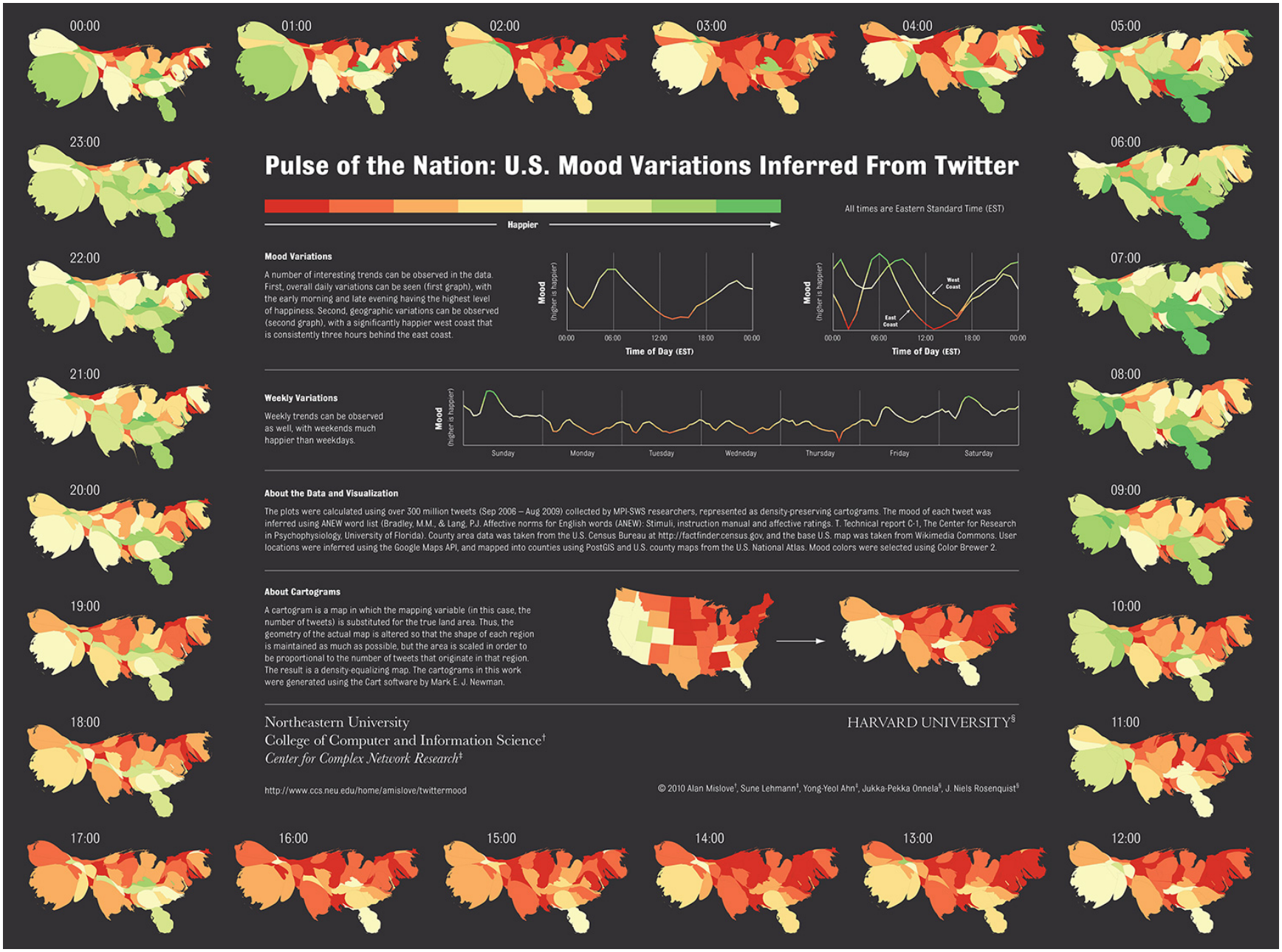
The network has a core-periphery structure with higher value product classes, e.g., machinery, and chemicals, in the core and lower quality classes, e.g., fishing and garments, in the periphery. Products at the core of the network are highly interconnected while products in the periphery are sparsely interlinked.

Each country has a certain product export footprint. Relevant exports by 'Industrialized Countries', 'East Asia Pacific' and 'Latin America & the Caribbean' are given on the right.

Traditional growth theory assumes that there is always a more sophisticated product within reach. However, given the core-periphery structure of the product space, the distances between products differ considerably.

Countries that operate at the core have capabilities to develop and manufacture a wide range of products. Yet, countries that mostly operate in the periphery of the product space have much fewer opportunities for diversification. A country's current footprint and the structure of the product space have a major impact on a country's future development.

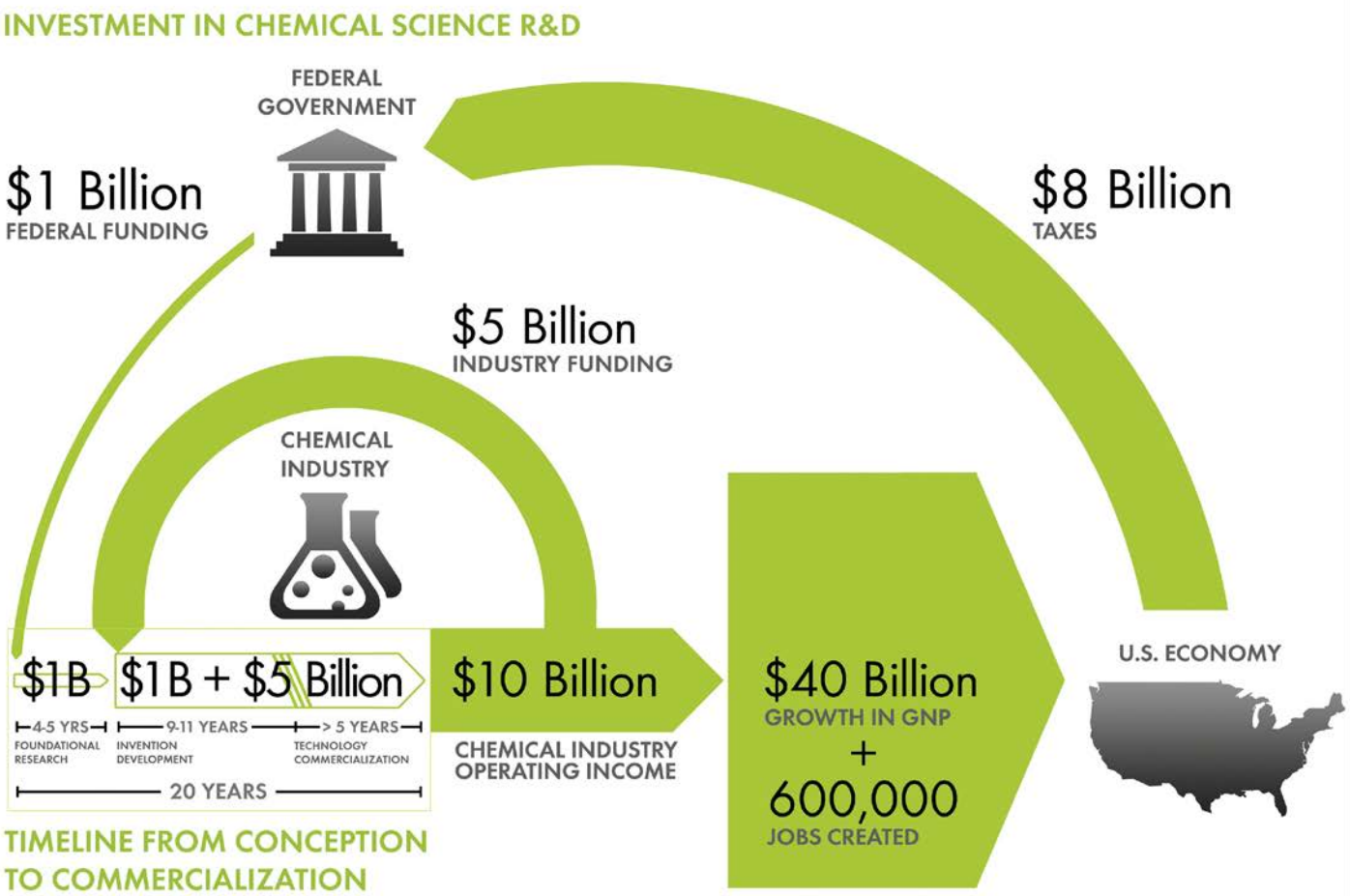
IV.7 *The Product Space* - César A. Hidalgo, Bailey Klinger, Albert-László Barabási, Ricardo Hausmann - 2007



IX.4 *Pulse of the Nation* - Alan Mislove, Sune Lehmann, Yong-Yeol Ahn, Jukka-Pekka Onnela, and James Niels Rosenquist - 2010

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

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

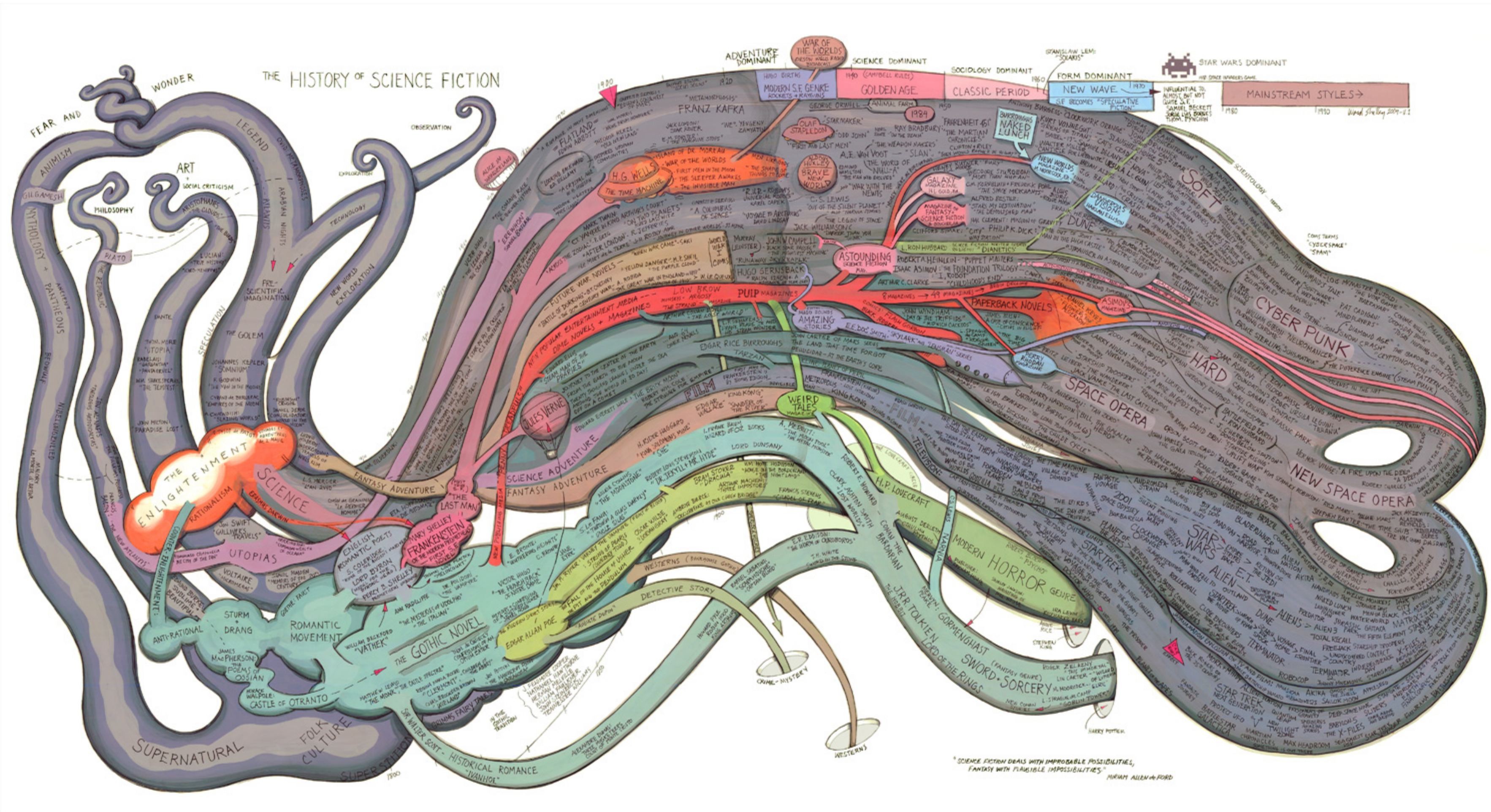


The Council for Chemical Research (CCR)

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.

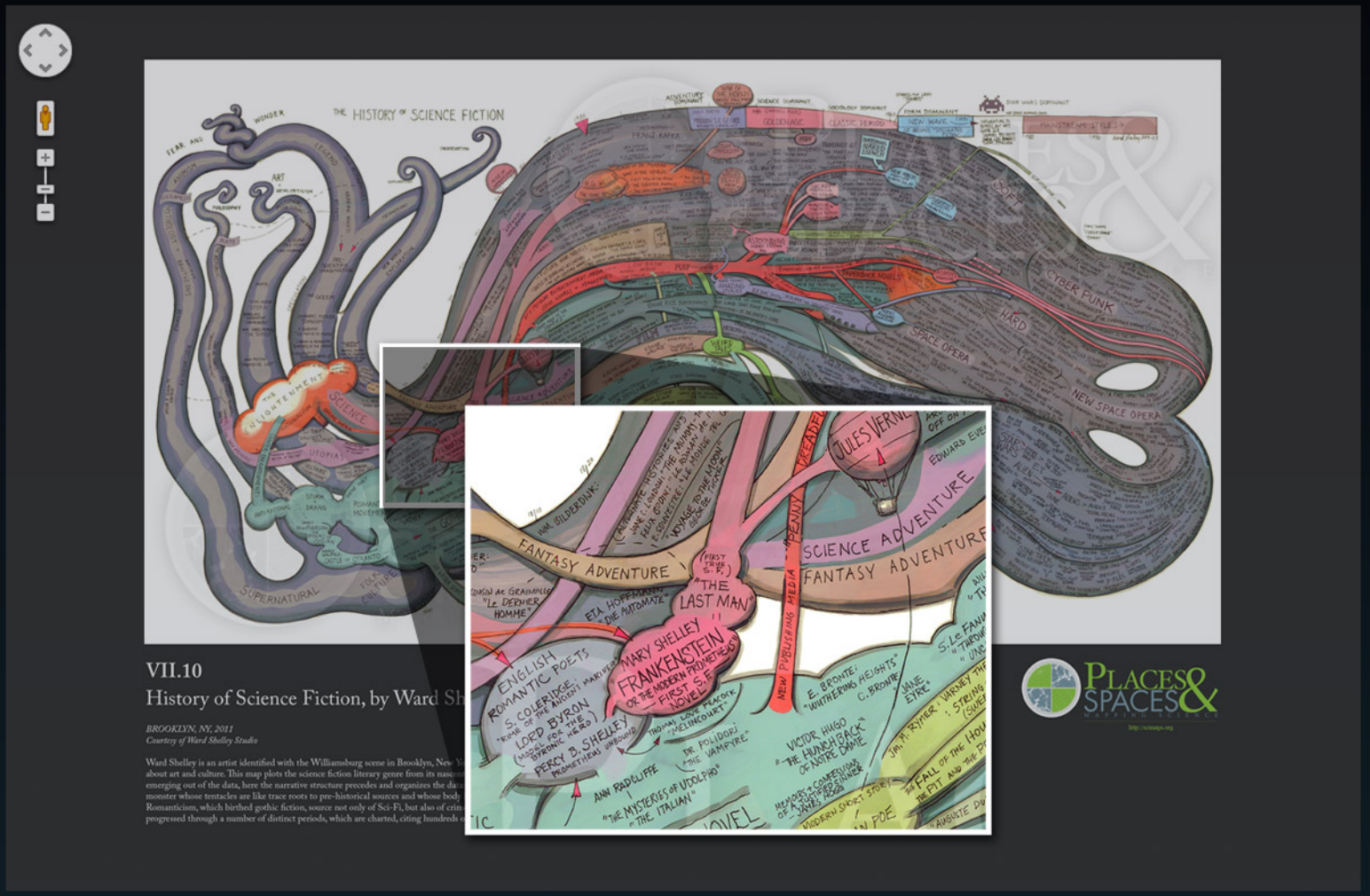
The design shows that an input of $1B in federal investment, leveraged by $5B in industry investment, brings new technologies to market and results in $10B of operating income for the chemical industry, $40B of growth in the Gross National Product (GNP) and further impacts the US economy by generating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



VII.10 *History of Science Fiction* - Ward Shelley - 2011

Check out our Zoom Maps online!



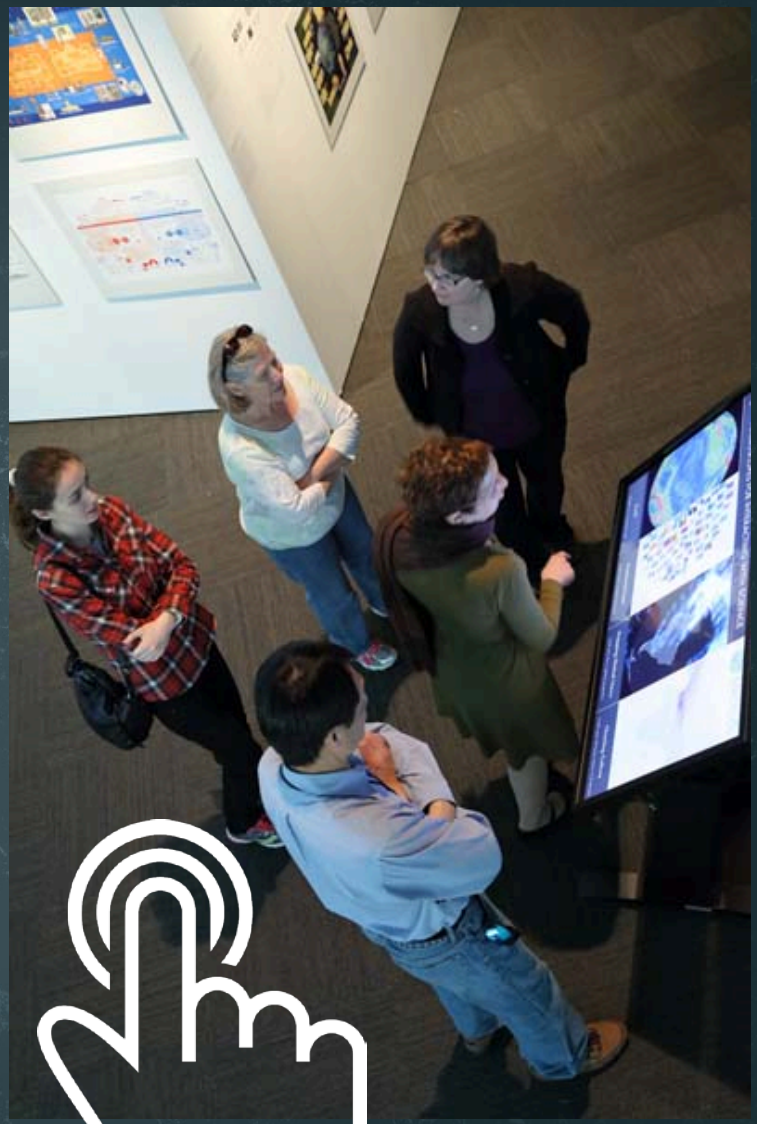
VII.10

History of Science Fiction, by Ward Sh

BROOKLYN, NY, 2011
Courtesy of Ward Shelley Studio

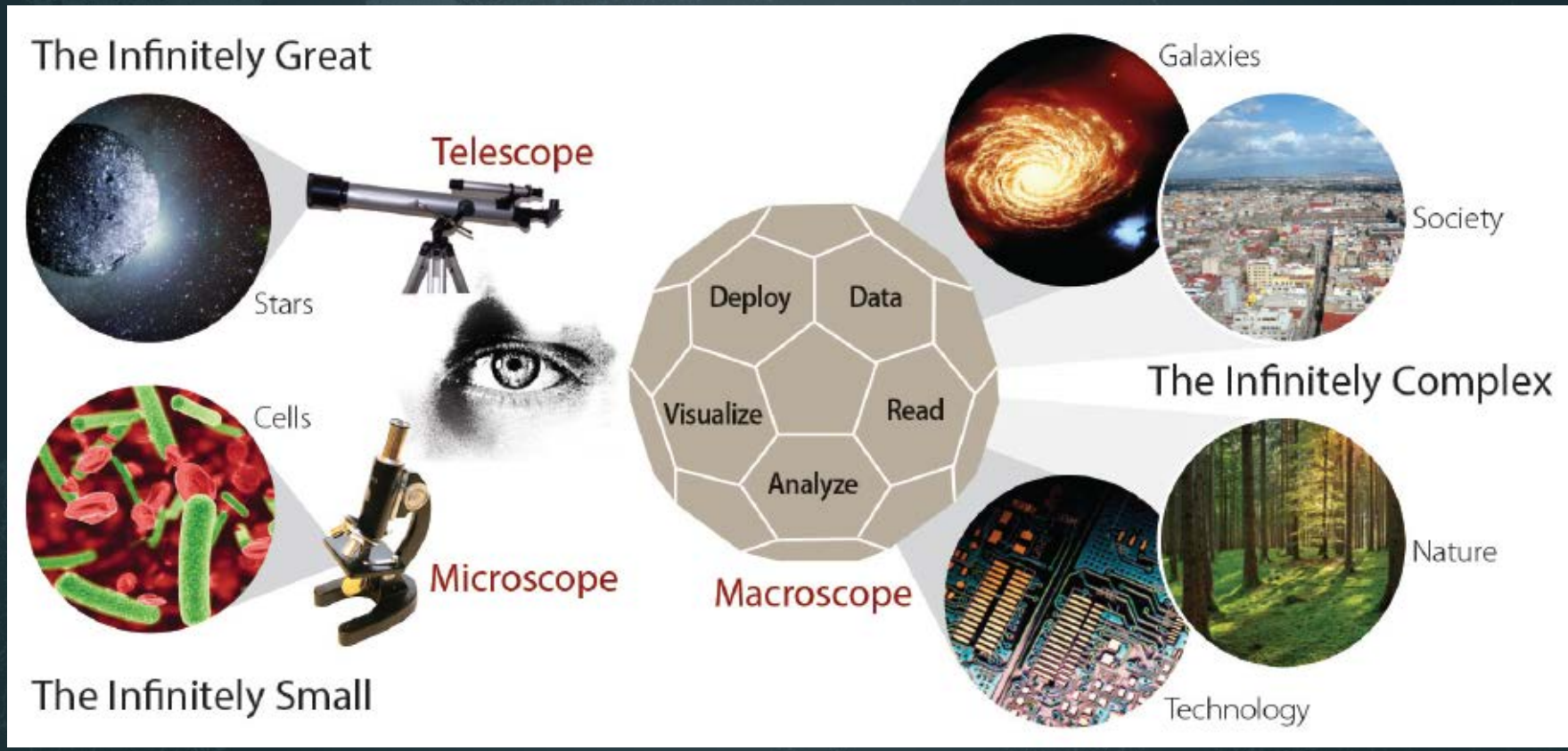
Ward Shelley is an artist identified with the Williamsburg scene in Brooklyn, New Y
about art and culture. This map plots the science fiction literary genre from its nasc
emerging out of the data, here the narrative structure precedes and organizes the da
monster whose tentacles are like trace roots to pre-historical sources and whose body
Romanticism, which birthed gothic fiction, source not only of Sci-Fi, but also of crim
progressed through a number of distinct periods, which are charted, citing hundreds

Visit scimaps.org and check out all our maps in stunning detail!



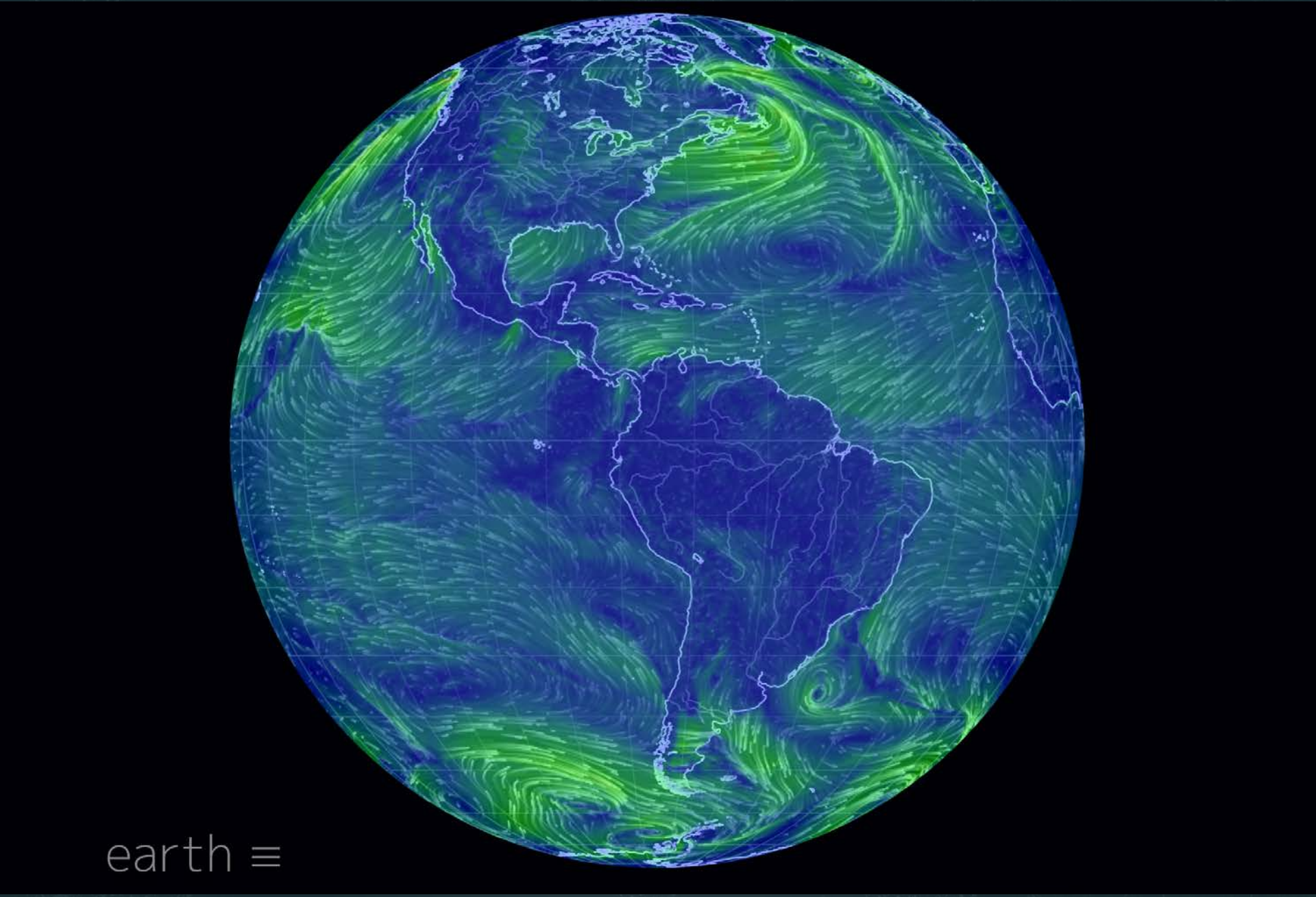
MAPS
vs.
MACROSCOPES

Microscopes & Telescopes vs. MACROSCOPES

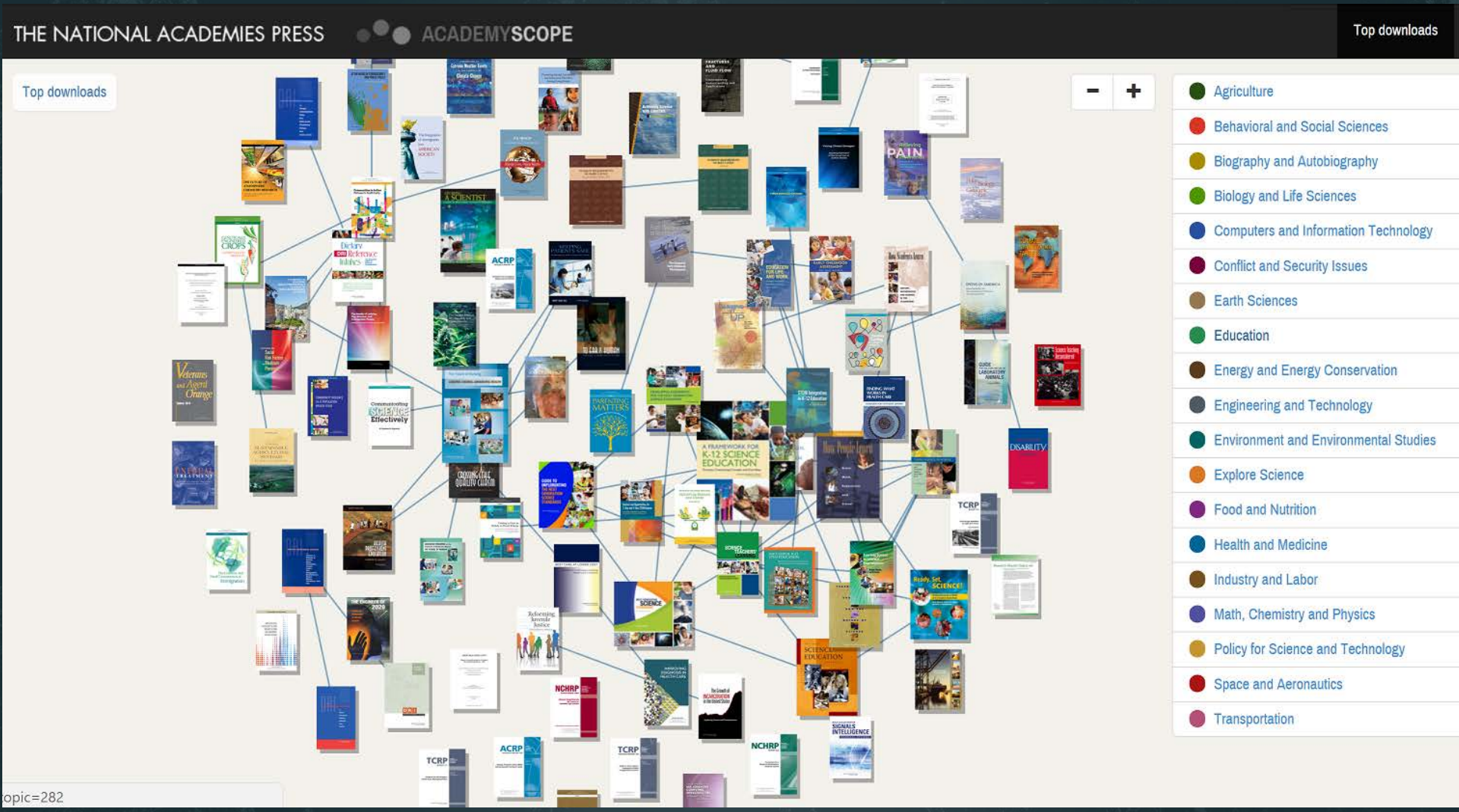


Iteration XI (2015): Macroscopes for Interacting with Science

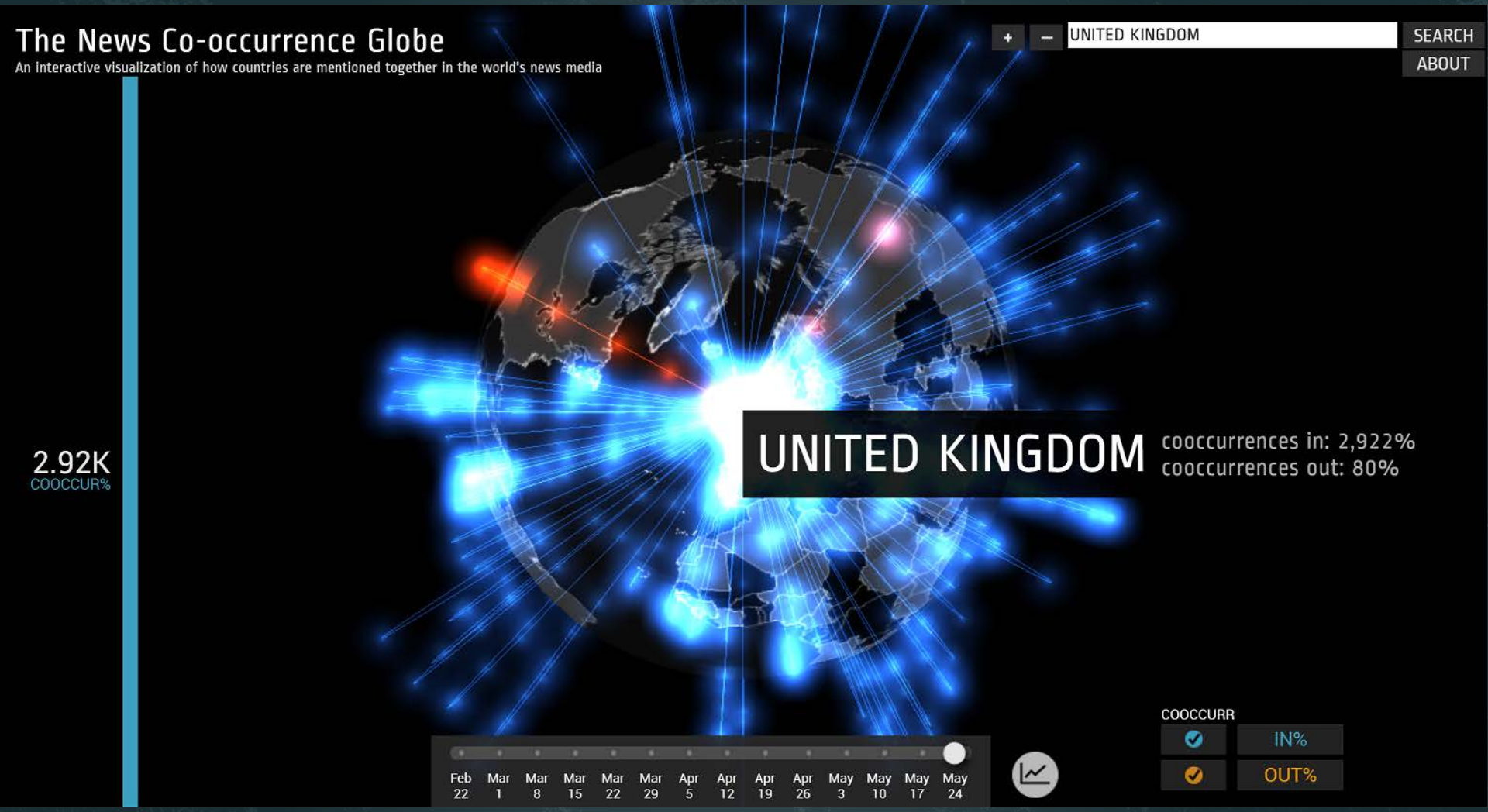
http://scimaps.org/iteration/11



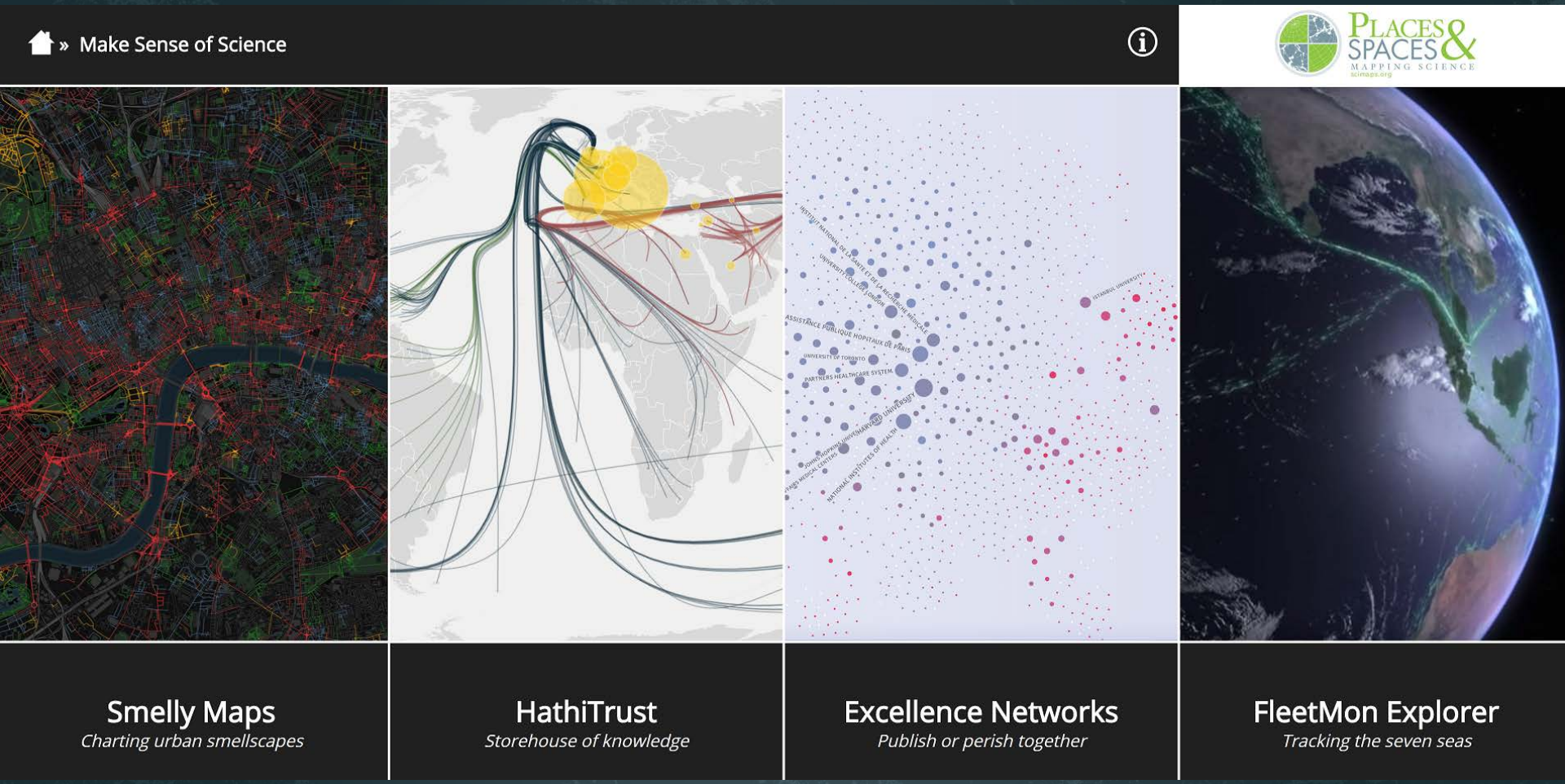
Earth – Cameron Beccario



AcademyScope – National Academy of the Sciences & CNS

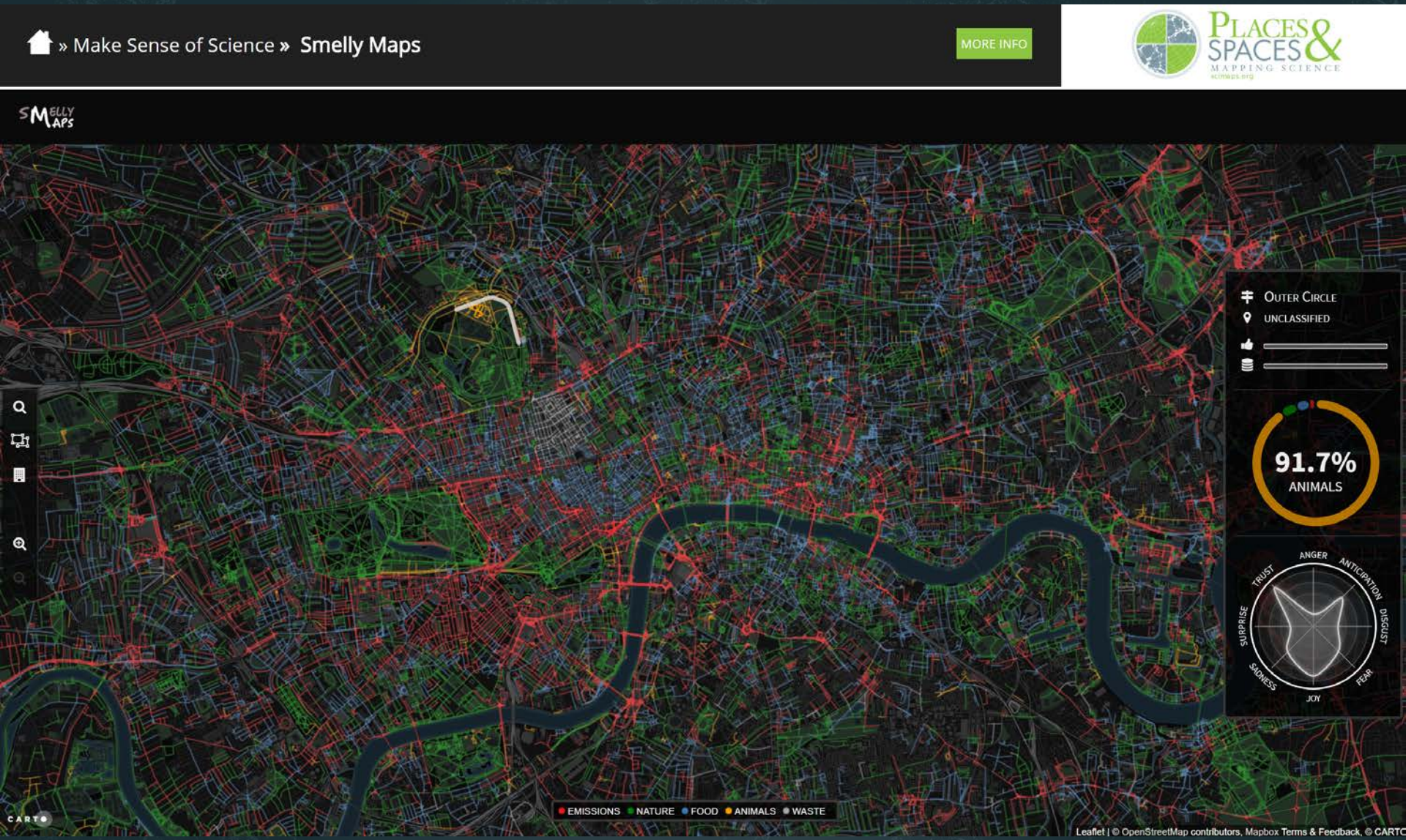


Mapping Global Society –Kalev Leetaru

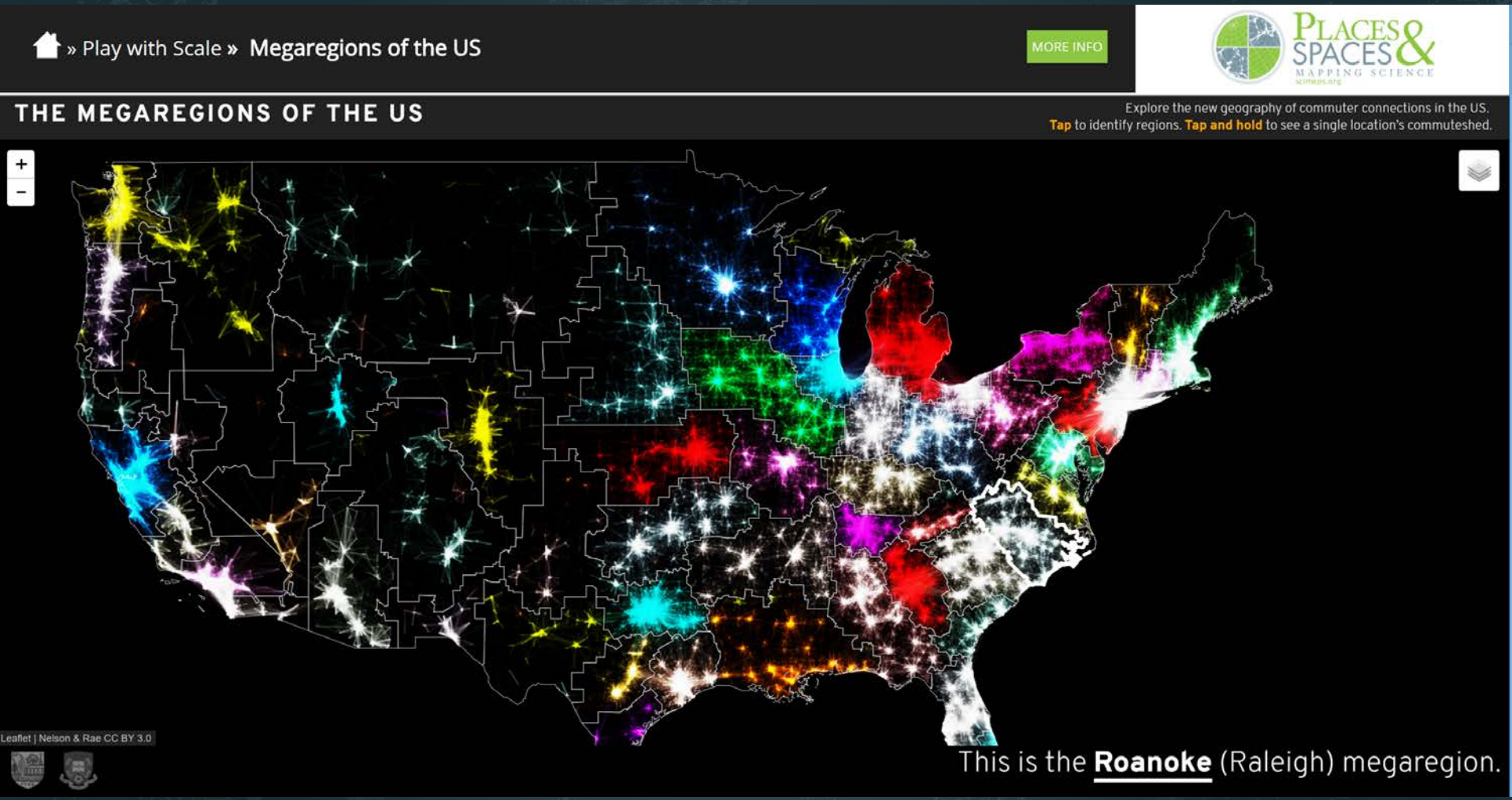


Iteration XII (2016): Macroscopes for Making Sense of Science

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

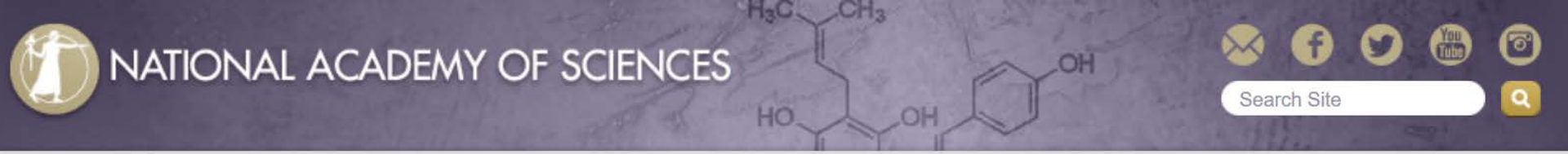
FleetMon Explorer – FleetMon – 2012

Models of Science & Technology

Using large scale datasets, advanced data mining, modeling, 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 report are available via http://modsti.cns.iu.edu/report



PROGRAMS

- Awards
- Koshland Science Museum
- Cultural Programs
- Sackler Colloquia
  - About Sackler Colloquia
  - Upcoming Colloquia
  - Completed Colloquia
  - Video Gallery
  - Connect with Sackler Colloquia
  - Give to Sackler Colloquia
- Kavli Frontiers of Science
- Distinctive Voices

Arthur M. Sackler COLLOQUIA

Upcoming Colloquia

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

Reproducibility of Research: Issues and Proposed Remedies

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

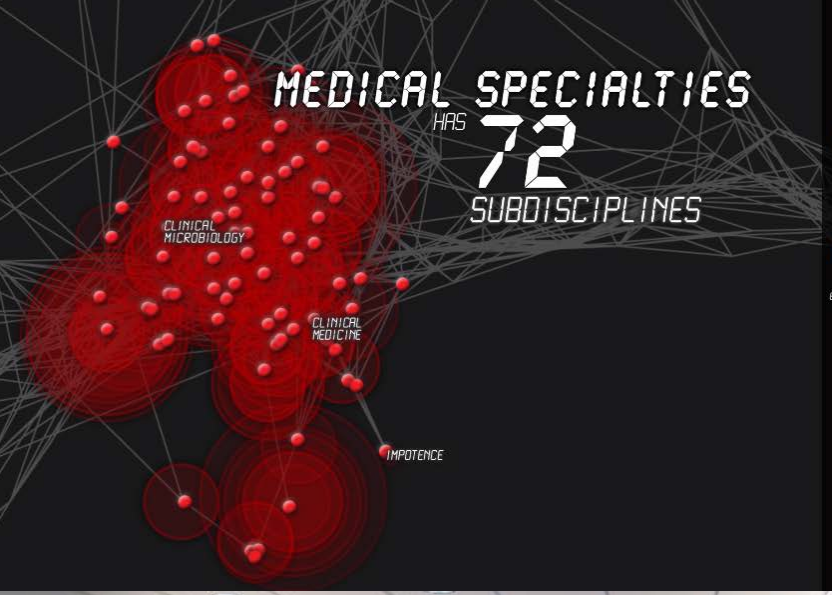
Science of Science Communication III

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

Modelling and Visualizing Science and Technology Developments

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

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



Science Forecast S1:E1

Science Forecast
S1:E1

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

Register for free: http://ivmooc.cns.iu.edu. Class restarts 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

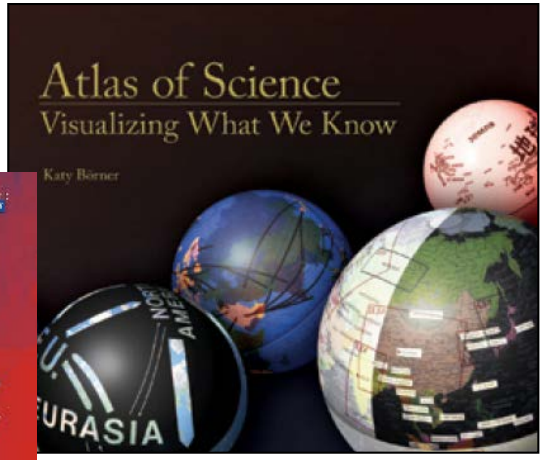
Börner, Katy (2010) **Atlas of Science: Visualizing What We Know**. The MIT Press. http://scimaps.org/atlas

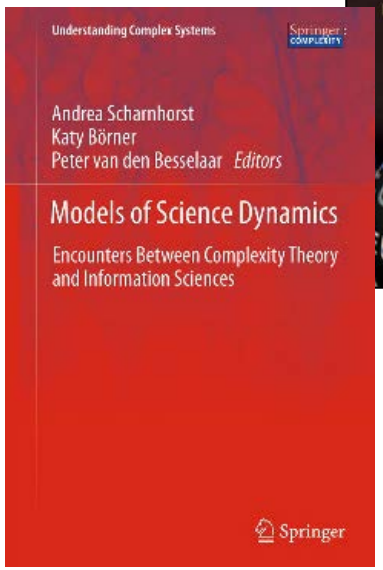
Scharnhorst, Andrea, Börner, Katy, van den Besselaar, Peter (2012) **Models of Science Dynamics**. Springer Verlag.

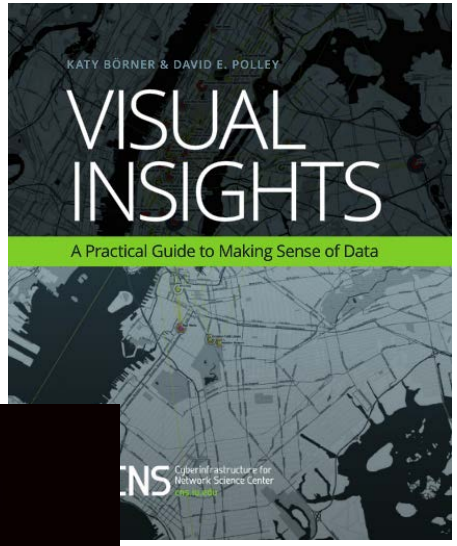
Katy Börner, Michael Conlon, Jon Corson-Rikert, Cornell, Ying Ding (2012) **VIVO: A Semantic Approach to Scholarly Networking and Discovery**. Morgan & Claypool.

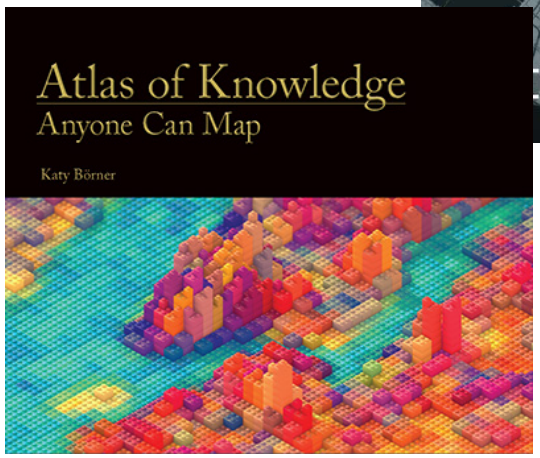
Katy Börner and David E Polley (2014) **Visual Insights: A Practical Guide to Making Sense of Data**. The MIT Press.

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









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