CNS Cyberinfrastructure for Network Science Center

INDIANA UNIVERSITY
SCHOOL OF INFORMATICS, COMPUTING, AND ENGINEERING

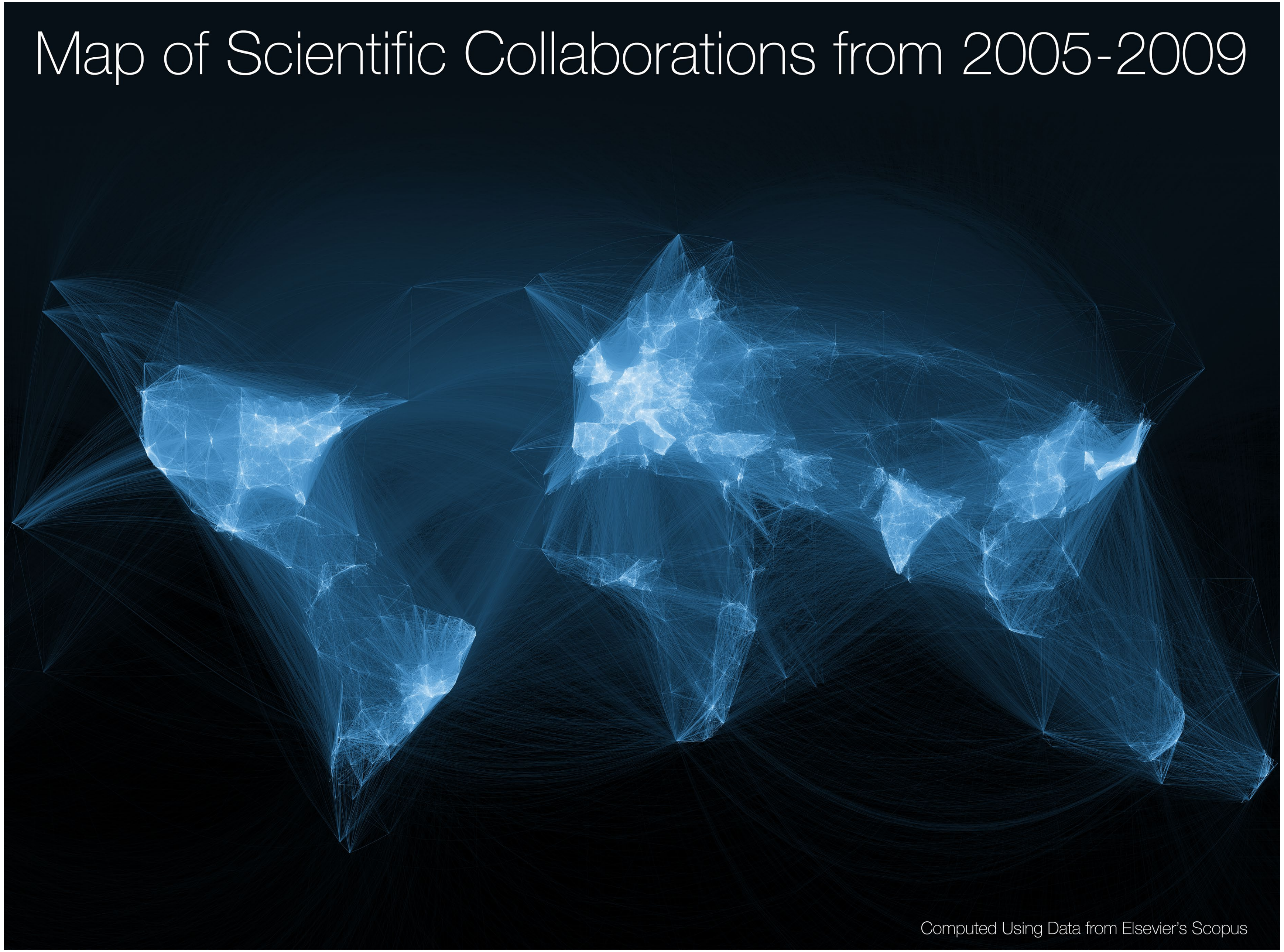
Modeling and Visualizing Complex Unifiable Systems

Katy Börner @katycns

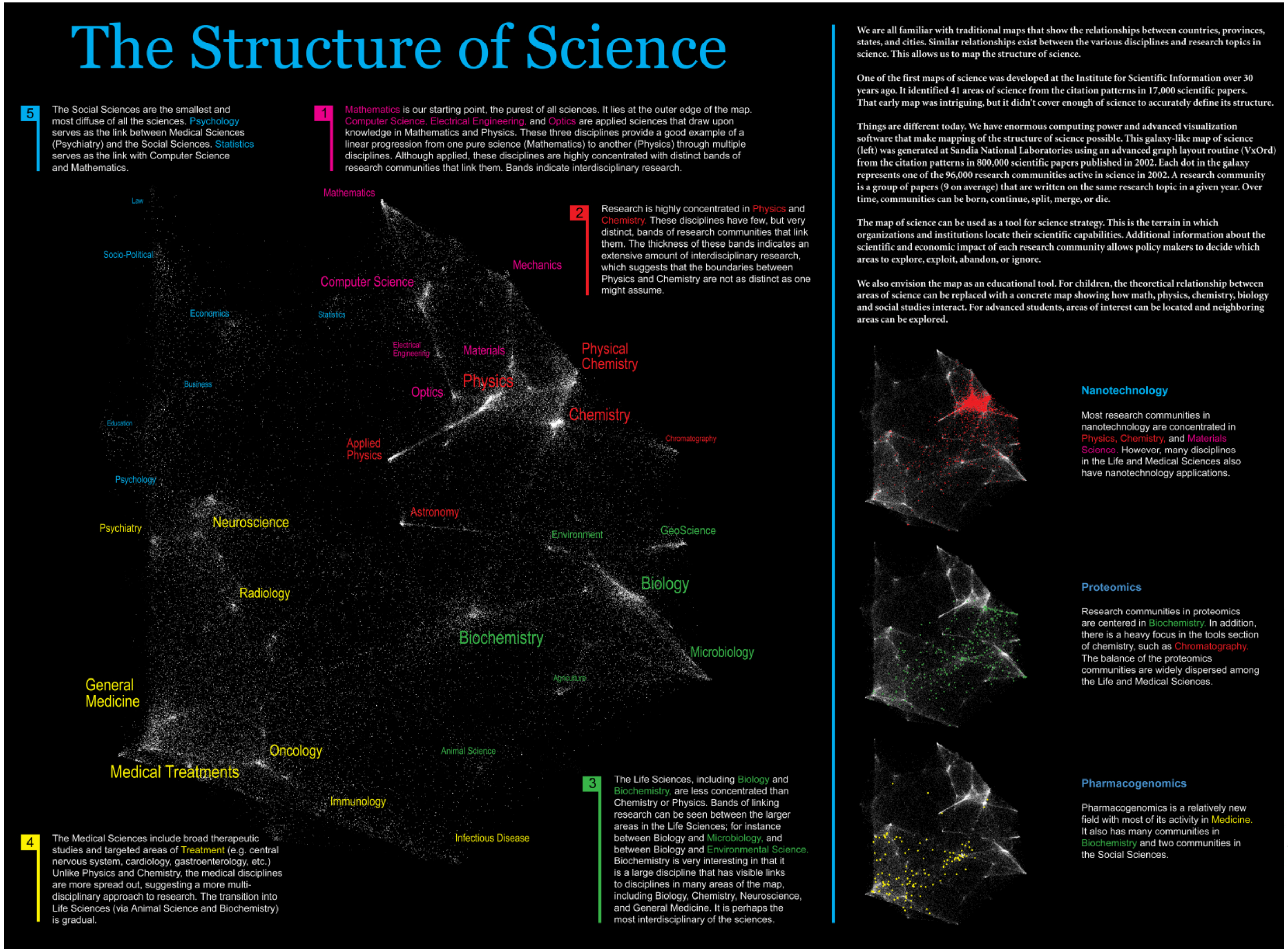
Victor H. Yngve Distinguished Professor of
Intelligent Systems Engineering & Information Science
Director, Cyberinfrastructure for Network Science Center
Luddy School of Informatics, Computing, and Engineering
Indiana University Network Science Institute (IUNI)
Indiana University, Bloomington, IN, USA

Forum on Complex Unifiable Systems: A Virtual Convocation
National Academy of Engineering

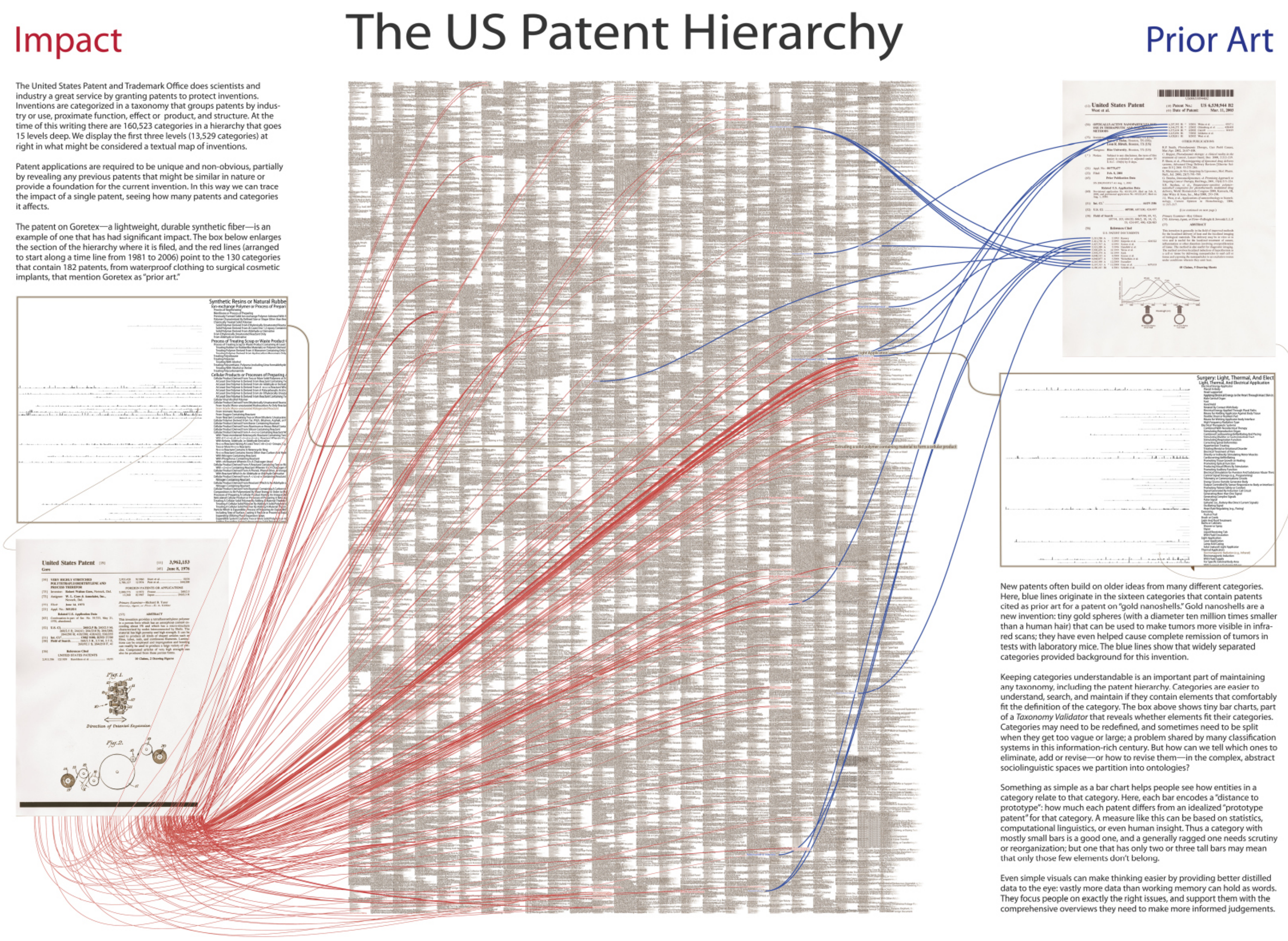
April 28, 2020



VII.6 Stream of Scientific Collaborations Between World Cities - Olivier H. Beauchesne - 2012



I.10 *The Structure of Science* - Kevin W. Boyack and Richard Klavans - 2005



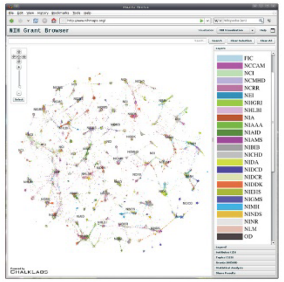
II.8 *Taxonomy Visualization of Patent Data* - Katy Börner, Elisha F. Hardy, Bruce W. Herr II, Todd Holloway, and W. Bradford Paley - 2006

A Topic Map of NIH Grants 2007

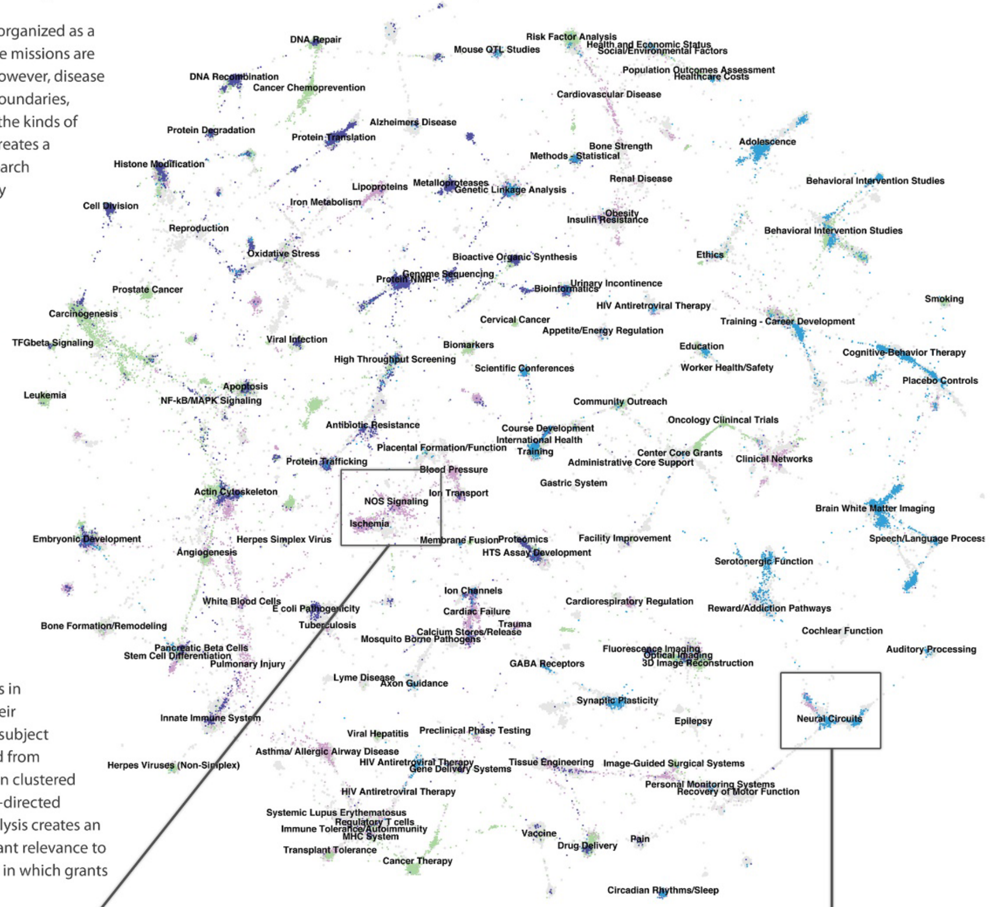
ChalkLabs UCIrvine

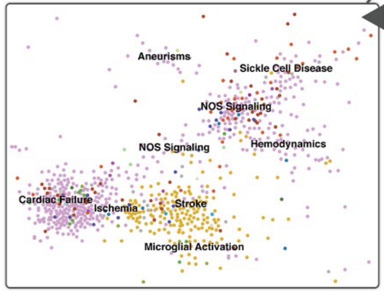
Bruce W. Herr II (Chalklabs & IU), Gully Burns (ISI), David Newman (UCI), Edmund Talley (NIH)

The National Institutes of Health (NIH) is organized as a multitude of Institutes and Centers whose missions are primarily focused on distinct diseases. However, disease etiologies and therapies flout scientific boundaries, and thus there is tremendous overlap in the kinds of research funded by each Institute. This creates a daunting landscape for decisions on research directions, funding allocations, and policy formulations. Shown here is devised an interactive topic map for navigating this landscape, online at www.nihmaps.org. Institute abbreviations can be found at www.nih.gov/icd.



Topic modeling, a statistical technique that automatically learns semantic categories, was applied to assess projects in terms used by researchers to describe their work, without the biases of keywords or subject headings. Grant similarities were derived from their topic mixtures, and grants were then clustered on a two-dimensional map using a force-directed simulated annealing algorithm. This analysis creates an interactive environment for assessing grant relevance to research categories and to NIH Institutes in which grants are localized.



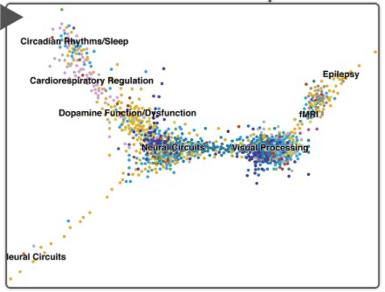


Cardiac Diseases Research

An area of the map focused on cardio-vascular function and dysfunction. Cardiac Failure (primarily funded by NHLBI) is topically clustered next to Stroke (NINDS), since these are the two major medical emergencies associated with ischemia, which results from a re-stricted blood supply. Also localized in this area are grants focused on Nitric Oxide (NOS) Signaling, a major biochem-ical pathway for vasodilation, and grants on Hemodynamics, Sickle Cell Disease, and Aneurysms.

Neural Circuits Research

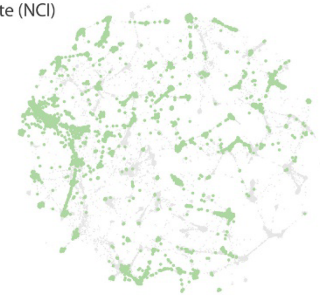
An area of the map focused on neural circuits, which shows the diversity of topics and NIH Institutes that fund research in this area, such as: Cardiorespiratory Regulation, primarily funded by NHLBI; Visual Processing, primarily funded by NEI; and Epilepsy, primarily funded by NINDS. For color coding, see legend in the upper-left inset.



National Cancer Institute (NCI)

TOP 10 TOPICS

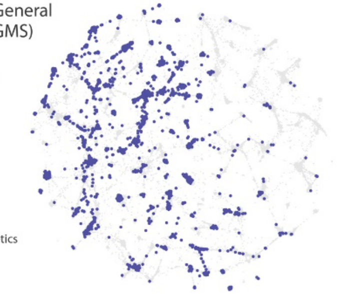
1. Oncology Clinical Trials
2. Cancer Treatment
3. Cancer Therapy
4. Carcinogenesis
5. Risk Factor Analysis
6. Cancer Chemotherapy
7. Metastasis
8. Leukemia
9. Prediction/Prognosis
10. Cancer Chemoprevention



National Institute of General Medical Sciences (NIGMS)

TOP 10 TOPICS

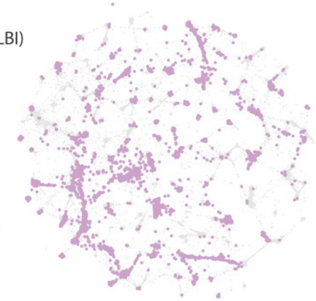
1. Bioactive Organic Synthesis
2. X-ray Crystallography
3. Protein NMR
4. Computational Models
5. Yeast Biology
6. Metalloproteases
7. Enzymatic Mechanisms
8. Protein Complexes
9. Invertebrate/Zebrafish Genetics
10. Cell Division



National Heart, Lung, and Blood Institute (NHLBI)

TOP 10 TOPICS

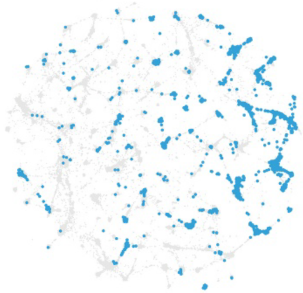
1. Cardiac Failure
2. Pulmonary Injury
3. Genetic Linkage Analysis
4. Cardiovascular Disease
5. Atherosclerosis
6. Hemostasis
7. Blood Pressure
8. Asthma/ Allergic Airway Disease
9. Gene Association
10. Lipoproteins



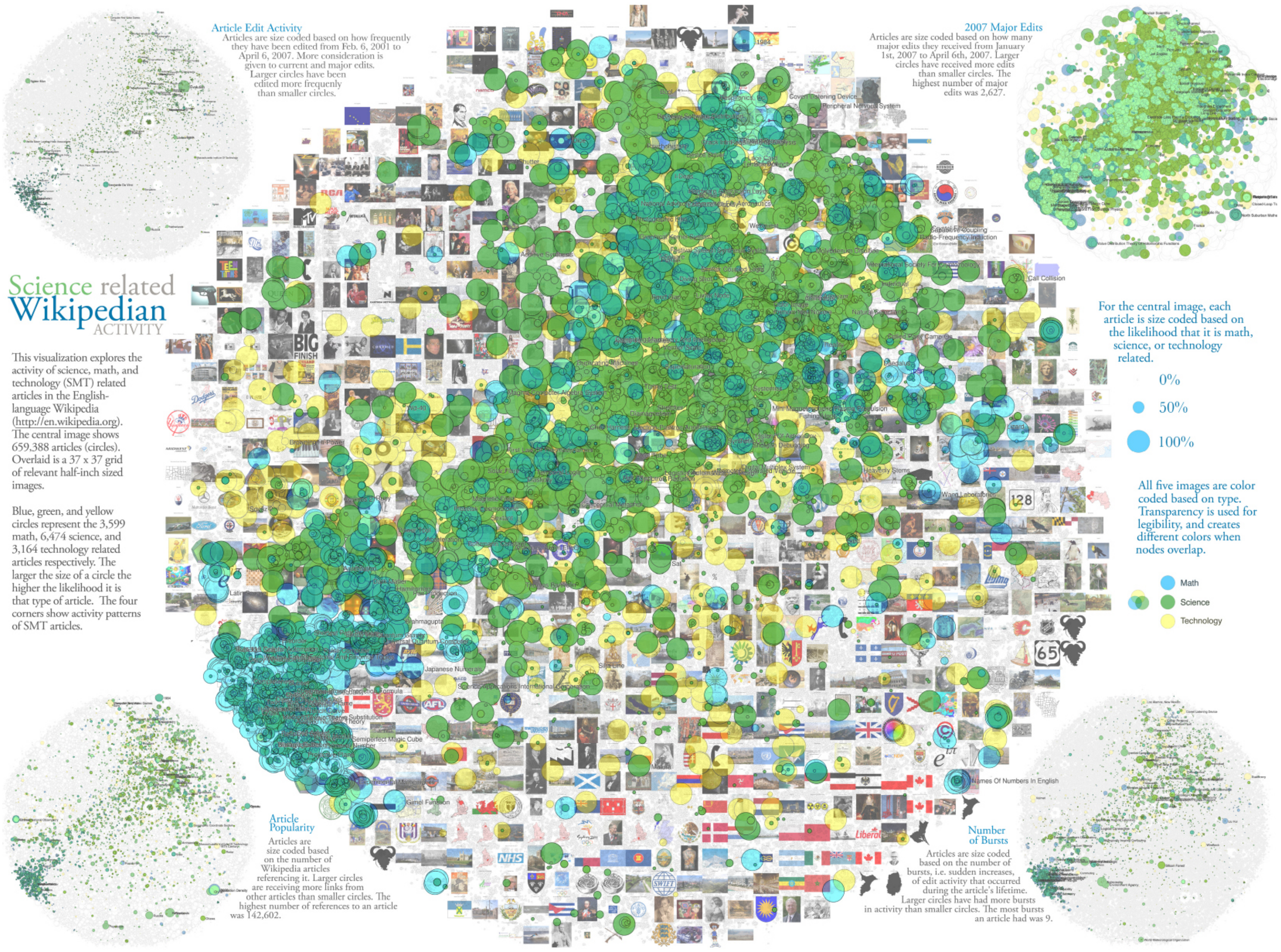
National Institute of Mental Health (NIMH)

TOP 10 TOPICS

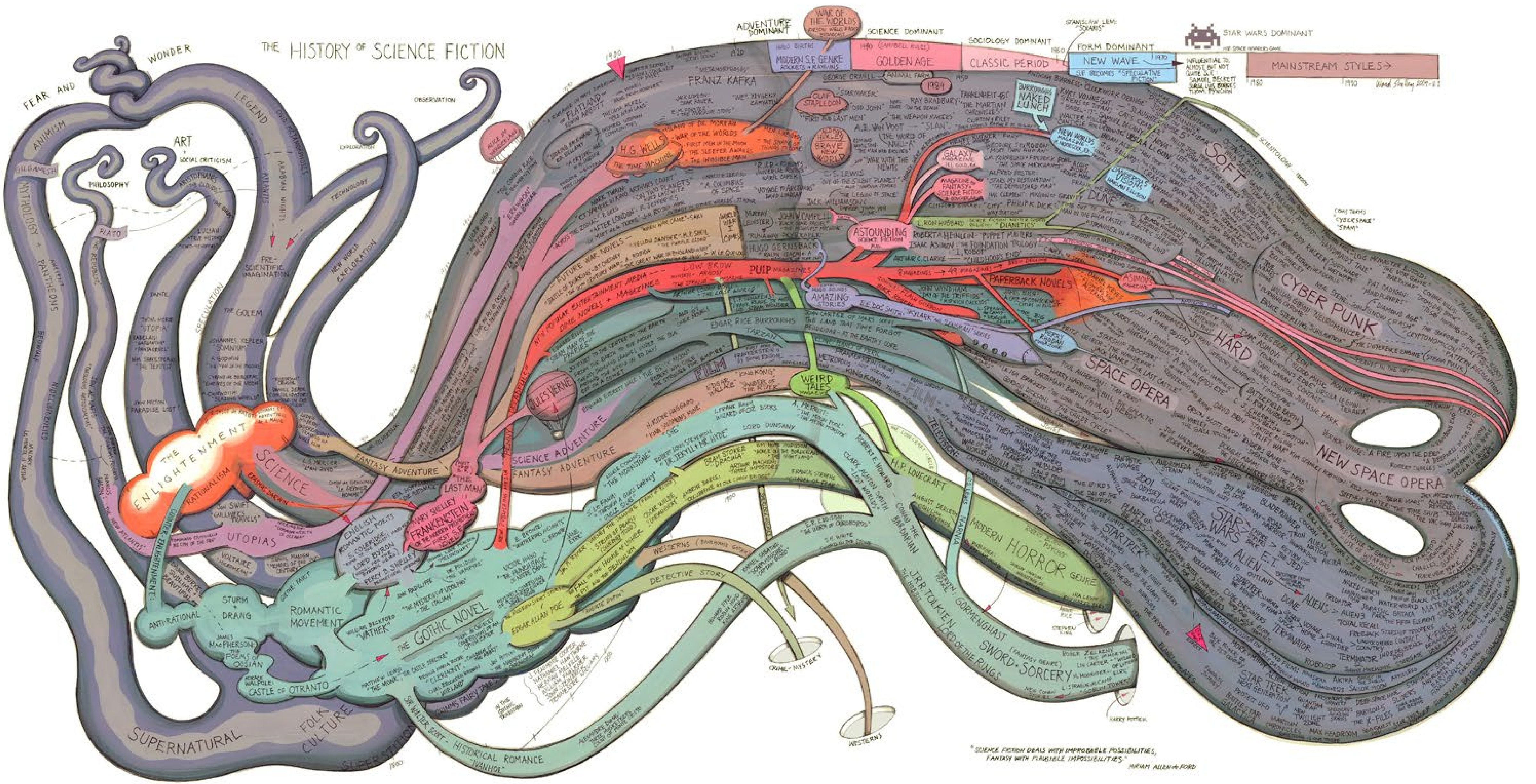
1. Mood Disorders
2. Schizophrenia
3. Behavioral Intervention Studies
4. Mental Health
5. Depression
6. Cognitive-Behavior Therapy
7. AIDS Prevention
8. Genetic Linkage Analysis
9. Adolescence
10. Childhood



V.7 *A Topic Map of NIH Grants 2007* - Bruce W. Herr II, Gully A.P.C. Burns, David Newman, and Edmund Talley - 2009

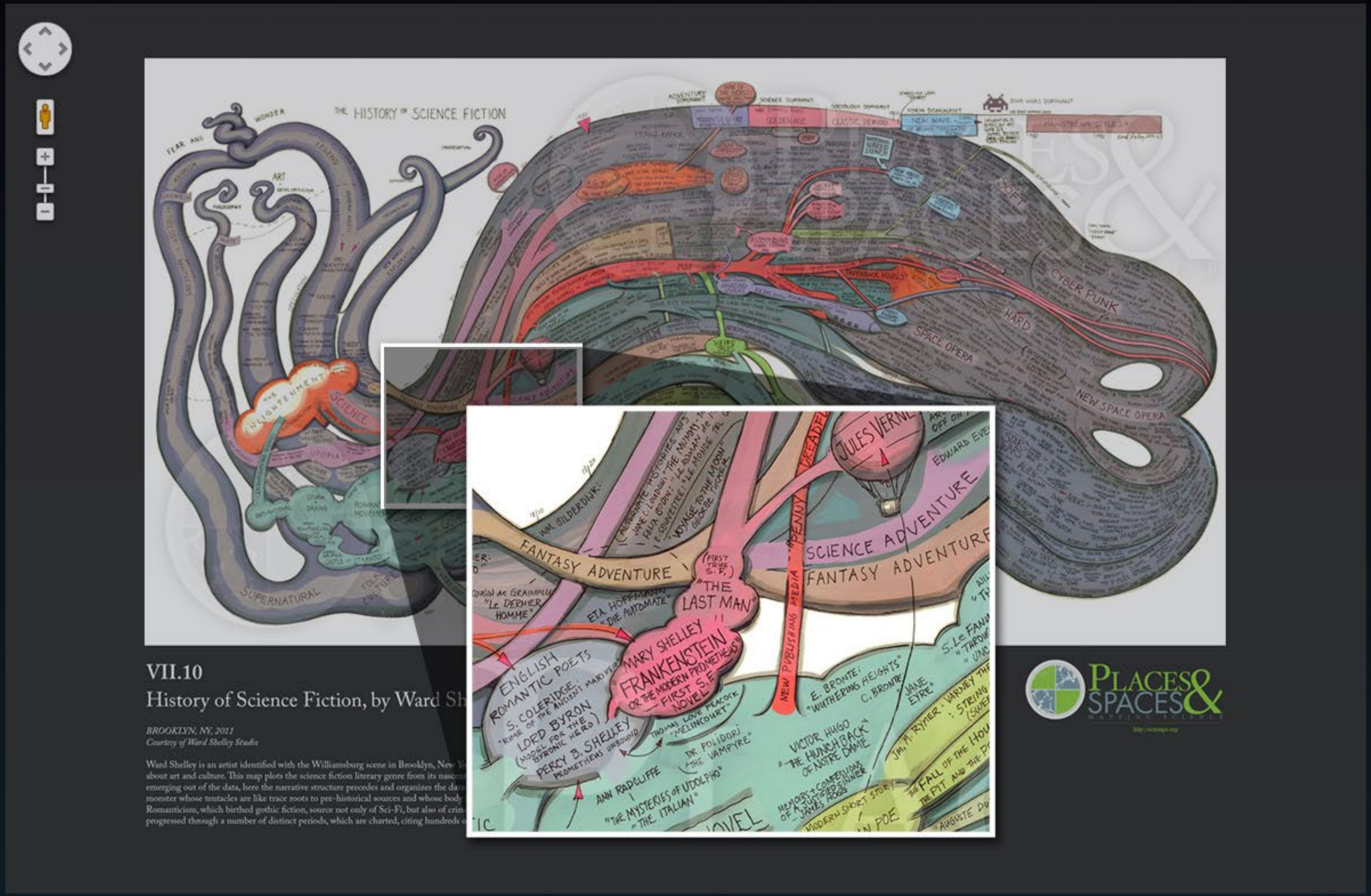


III.8 *Science-Related Wikipedian Activity* - Bruce W. Herr II, Todd M. Holloway, Elisha F. Hardy, Katy Börner, and Kevin Boyack - 2007



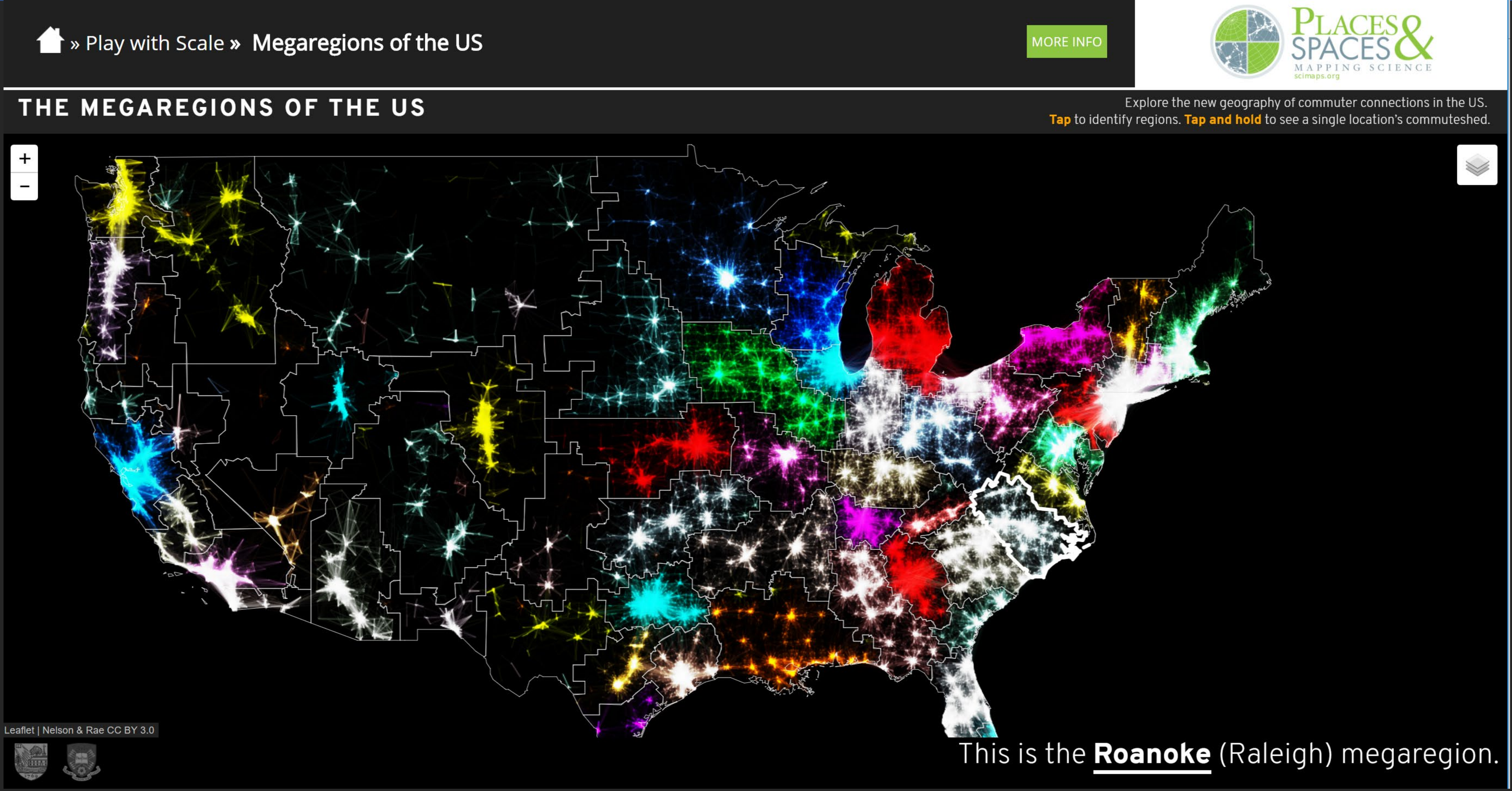
VII.10 *History of Science Fiction* - Ward Shelley - 2011

Check out our Zoom Maps online!



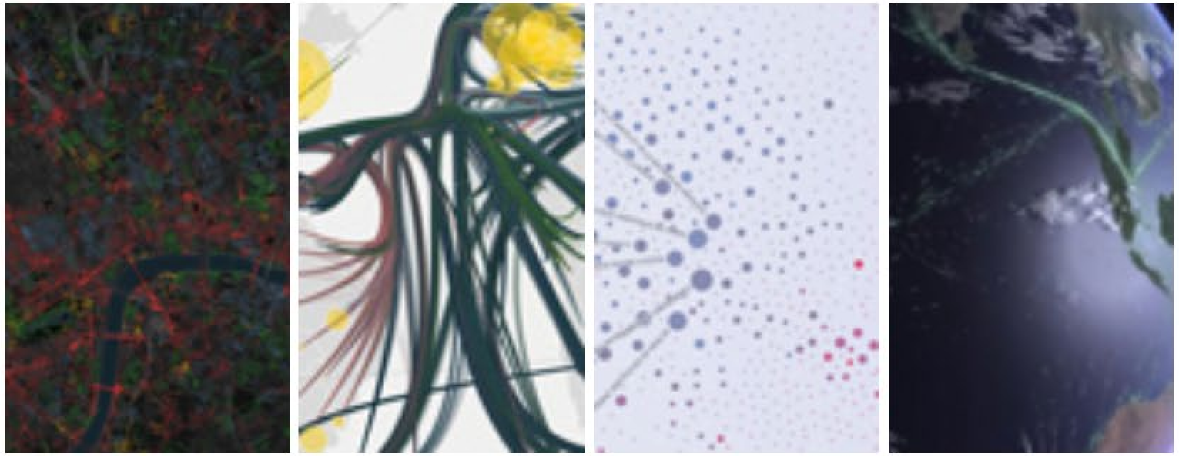
Visit scimaps.org and check out all our maps in stunning detail!

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



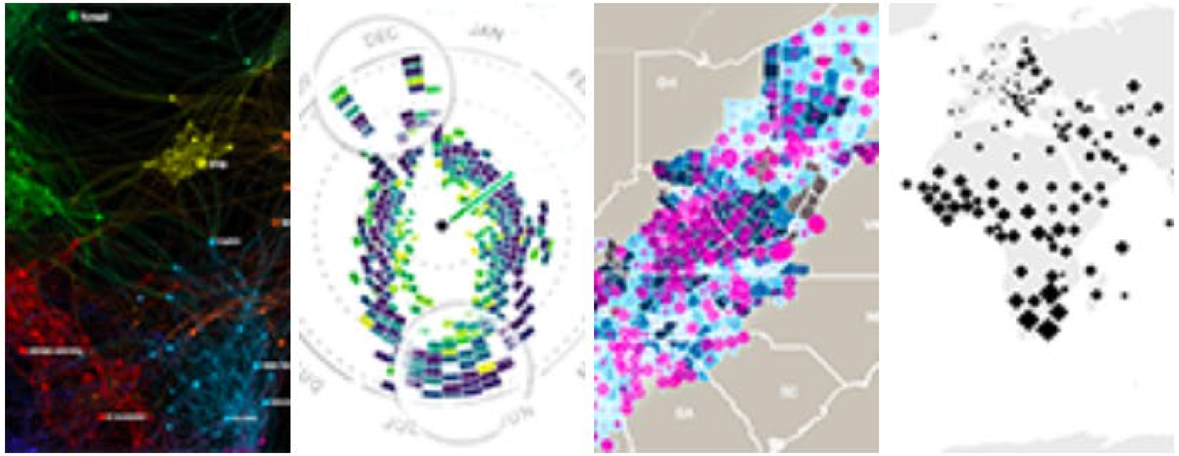
Megaregions of the US –Garrett Dash Nelson and Alasdair Rae – 2016

Iteration XII (2016)
Macroscopes for Making Sense of Science

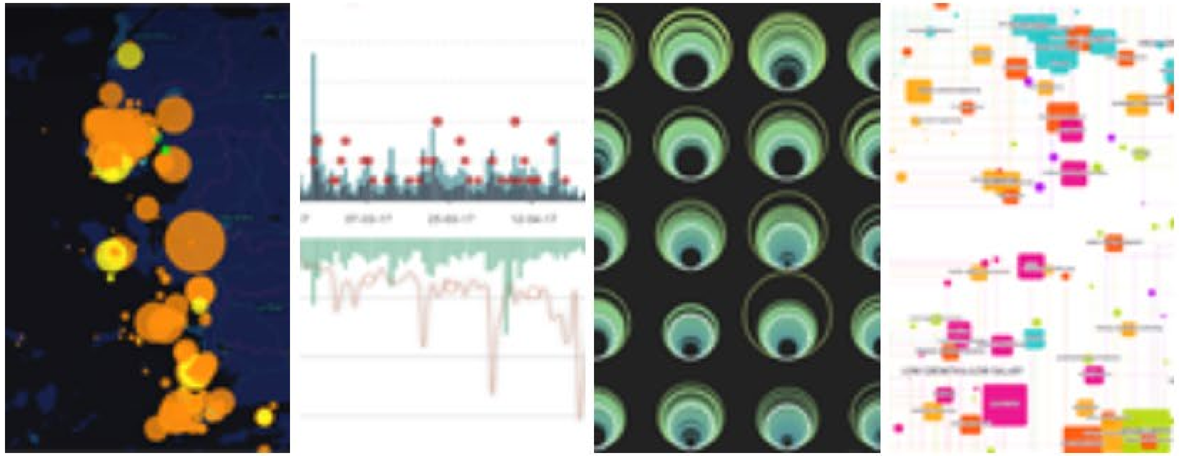


Iteration XIII (2017)
Macroscopes for Playing with Scale

Iteration XIV (2018)
Macroscopes for Ensuring our Well-being



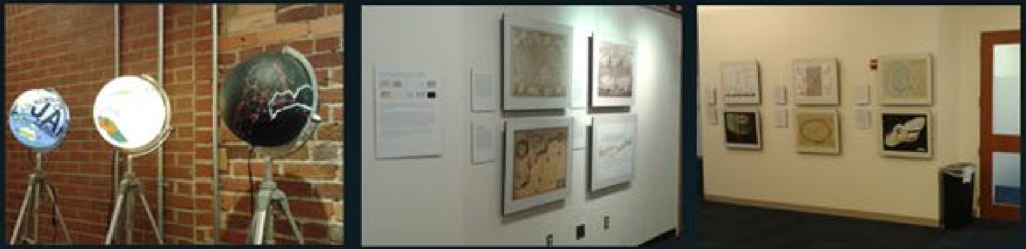
Iteration XV (2019)
Macroscopes for Tracking the Flow of Resources



http://scimaps.org

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

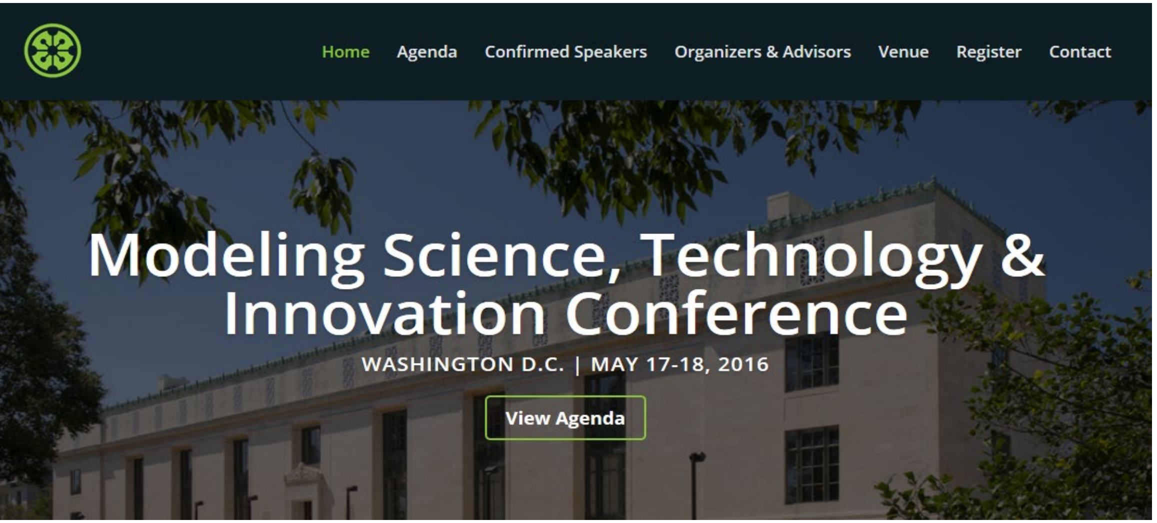
University of Miami, Miami, FL.
September 4 - December 11, 2014.



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

http://scimaps.org

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



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 at http://modsti.cns.iu.edu/report

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

Arthur M. Sackler COLLOQUIA OF THE NATIONAL ACADEMY OF SCIENCES Clarivate Microsoft

Arthur M. Sackler Colloquium on Modeling and Visualizing Science and Technology Developments

Twin-Win Model: A human-centered approach to research success

Ben Shneiderman

PNAS December 11, 2018 115 (50) 12590-12594; first published December 10, 2018. https://doi.org/10.1073/pnas.1802918115

Forecasting innovations in science, technology, and education

FROM THE COVER

Katy Börner, William B. Rouse, Paul Trunfio, and H. Eugene Stanley

PNAS December 11, 2018 115 (50) 12573-12581; first published December 10, 2018. https://doi.org/10.1073/pnas.1818750115

How science and technology developments impact employment and education

Wendy Martinez

PNAS December 11, 2018 115 (50) 12624-12629; first published December 10, 2018. https://doi.org/10.1073/pnas.1803216115

Scientific prize network predicts who pushes the boundaries of science

Yifang Ma and Brian Uzzi

PNAS December 11, 2018 115 (50) 12608-12615; first published December 10, 2018. https://doi.org/10.1073/pnas.1800485115

The role of industry-specific, occupation-specific, and location-specific knowledge in the growth and survival of new firms

C. Jara-Figueroa, Bogang Jun, Edward L. Glaeser, and Cesar A. Hidalgo

PNAS December 11, 2018 115 (50) 12646-12653; first published December 10, 2018. https://doi.org/10.1073/pnas.1800475115

https://www.pnas.org/modeling

Skill Discrepancies Between Research, Education, and Jobs Reveal the Critical Need to Supply Soft Skills for the Data Economy

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, robotics
- Social skills (project management, team leadership) become ever more important
- Increasing team size

Börner, Katy, Olga Scrivner, Mike Gallant, Shutian Ma, Xiaozhong Liu, Keith Chewning, Lingfei Wue, and James A. Evans. 2018. "Skill Discrepancies Between Research, Education, and Jobs Reveal the Critical Need to Supply Soft Skills for the Data Economy." *PNAS* 115(50): 12630-12637.

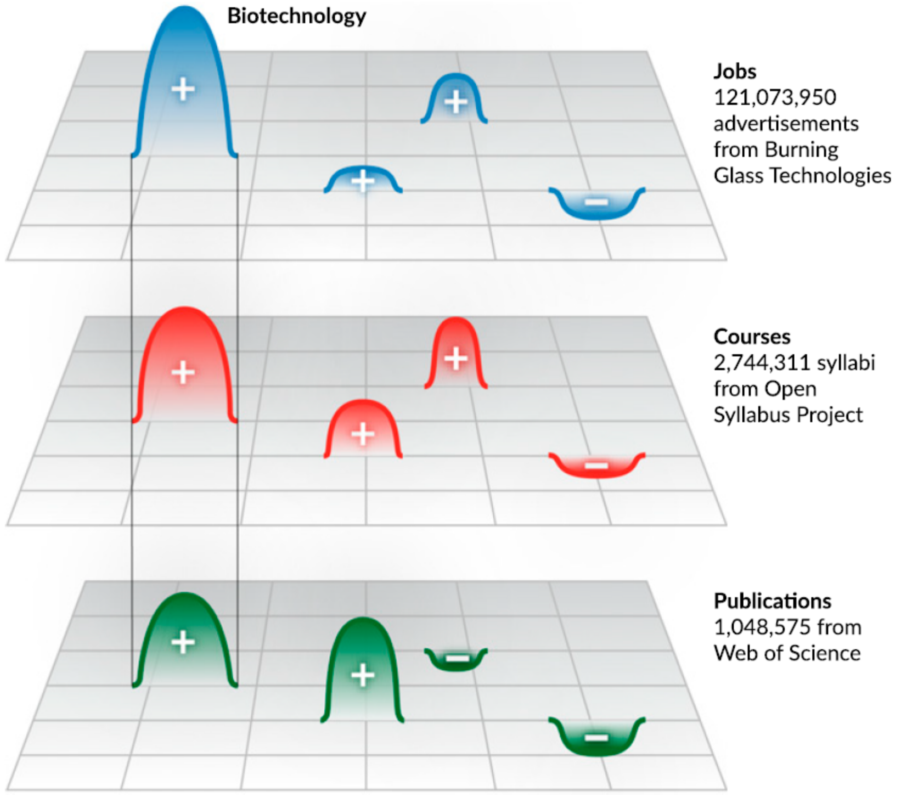


Fig. 1. The interplay of job market demands, educational course offerings, and progress in S&T as captured in publications. Color-coded mountains (+) and valleys (–) indicate different skill clusters. For example, skills related to Biotechnology might be mentioned frequently in job descriptions and taught in many courses, but they may not be as prevalent in academic publications. In other words, there are papers that mention these skills, but labor demand and commercial activity might be outstripping publication activity in this area. The numbers of jobs, courses, and publications that have skills associated and are used in this study are given on the right.

Stakeholders and Insight Needs

- **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 balance of timely and timeless knowledge (to get a job vs. learn how to learn) should I teach? How to innovate in teaching and maintain job security or tenure?
- **Universities:** What programs should be created? What is my competition doing? How do I tailor programs to fit local needs?
- **Science Funders:** How can S&T investments improve short- and long-term prosperity? Where will advances in knowledge also yield advances in skills and technology?
- **Employers:** What skills are needed next year and in 5 and 10 years? Which institutions produce the right talent? What skills does my competition list in job advertisements?
- **Economic Developers:** What critical skills are needed to improve business retention, expansion, and recruitment in a region?

What is the ROI of my time, money, compassion?

Urgency

- **35% of UK jobs, and 30% in London**, are at high risk from automation over the coming 20 years. https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/uk-futures/london-futures-agiletown.pdf
- The aerospace industry and NASA have a disproportionately large percentage of workers aged 50 and older compared to the national average, and up to **half of the current workforce** will be eligible for retirement within the coming five years. Astronautics AIAA (2012) Recruiting, retaining, and developing a world-class aerospace workforce. https://www.aiaa.org/uploadedFiles/Issues_and_Advocacy/Education_and_Workforce/Aerospace%20Workforce-%20030112.pdf
- The rise of artificial intelligence will lead to the displacement of **millions of blue-collar as well as white-collar jobs** in the coming decade. Beyer D (2016) The future of machine intelligence: Perspectives from leading practitioners ; Brynjolfsson E, McAfee A (2014) The second machine age: Work, progress, and prosperity in a time of brilliant technologies; Ford M (2015) Rise of the Robots: Technology and the Threat of a Jobless Future.
- COVID-19 economic freeze **could cost 47 million jobs** and send the **unemployment rate past 32%,** according to St. Louis Fed projections.

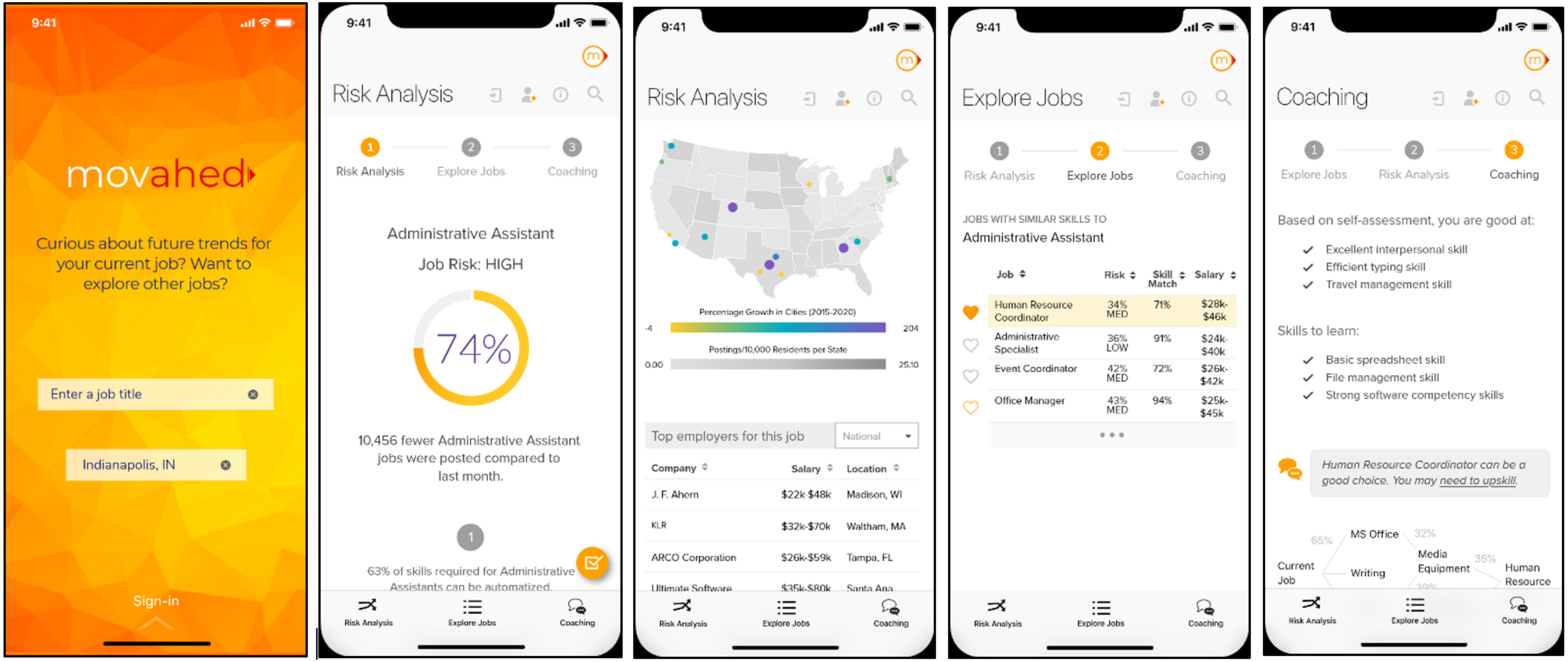
NSF RAISE: C-Accel Pilot: Analytics-Driven Accessible Pathways To Impacts-Validated Education (ADAPTIVE)

Goal: Development of data-driven tools to support the tens of millions of US workers whose jobs are being transformed by Artificial Intelligence (AI), automation, COVID-19, and other developments.

The project will demonstrate how labor market and course syllabi data, learning analytics, and insights on transferability of learned skills can be combined and visualized in novel ways to support a learner's decision-making about, sustained engagement in, and application to their job of professional skills acquired through education and job-related training.

Team B-6656: Katy Börner, Indiana University, Ariel Anbar, Arizona State University, Kemi Jona, Northeastern University, Martin Storksdieck and Heather Fischer, Oregon State University

https://www.nsf.gov/od/oia/convergence-accelerator/



Develop and deploy **socio-technical systems** that encourage US workers to explore the **evolving landscape of new jobs** and re/up-skilling opportunities—to not only recover from current risks/crisis but to leap forward into more resilient and more desirable futures.

CNS Cyberinfrastructure for Network Science Center

INDIANA UNIVERSITY
SCHOOL OF INFORMATICS, COMPUTING, AND ENGINEERING

Data Visualization Literacy

Börner, Katy, Andreas Bueckle, and Michael Ginda. 2019. Data visualization literacy: Definitions, conceptual frameworks, exercises, and assessments. *PNAS*, 116 (6) 1857-1864.

Data Visualization Literacy (DVL)

Data visualization literacy (ability to read, make, and explain data visualizations) requires:

- *literacy* (ability to read and write text in titles, axis labels, legends, etc.),
- *visual literacy* (ability to find, interpret, evaluate, use, and create images and visual media), and
- *mathematical literacy* (ability to formulate, employ, and interpret math in a variety of contexts).

Being able to "read and write" data visualizations is becoming as important as being able to read and write text. Understanding, measuring, and improving data and visualization literacy is important to strategically approach local and global issues.

Börner, Katy, Andreas Bueckle, and Michael Ginda. 2019. Data visualization literacy: Definitions, conceptual frameworks, exercises, and assessments. *PNAS*, 116 (6) 1857-1864.

Data Visualization Literacy Framework (DVL-FW)

Consists of two parts:

DVL Typology
Defines 7 types with 4-17 members each.

1 Insight Needs
- categorize/cluster
- order/rank/sort
- distributions (also outliers, gaps)
- comparisons
- trends (process and time)
- geospatial
- compositions (also of text)
- correlations/ relationships

2 Data Scales
- nominal
- ordinal
- interval
- ratio

3 Analyses
- statistical
- temporal
- geospatial
- topical
- relational

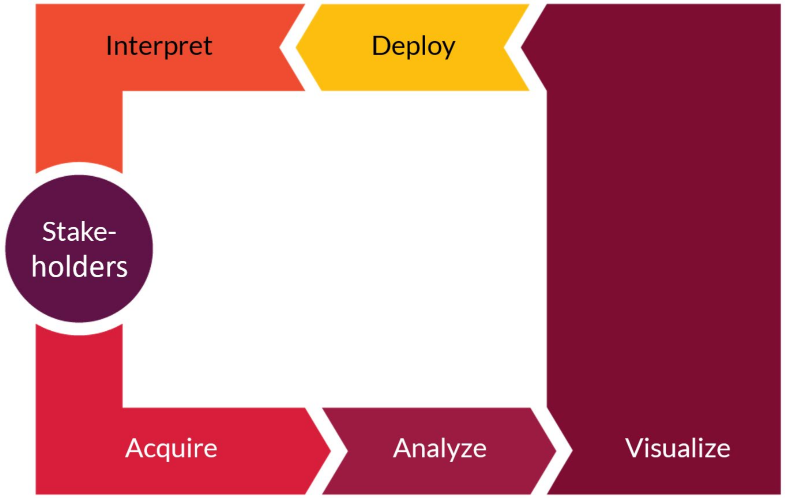
4 Visualizations
- table
- chart
- graph
- map
- tree
- network

5 Graphic Symbols
- geometric symbols
  - point
  - line
  - area
  - surface
  - volume
- linguistic symbols
  - text
  - numerals
  - punctuation marks
- pictorial symbols
  - images
  - icons
  - statistical glyphs

6 Graphic Variables
- spatial
  - position
- retinal
  - form
  - color
  - optics
  - motion

7 Interactions
- zoom
- search and locate
- filter
- details-on-demand
- history
- extract
- link and brush
- projection
- distortion

DVL Workflow Process
Defines 5 steps required to render data into insights.



Typology of the Data Visualization Literacy Framework

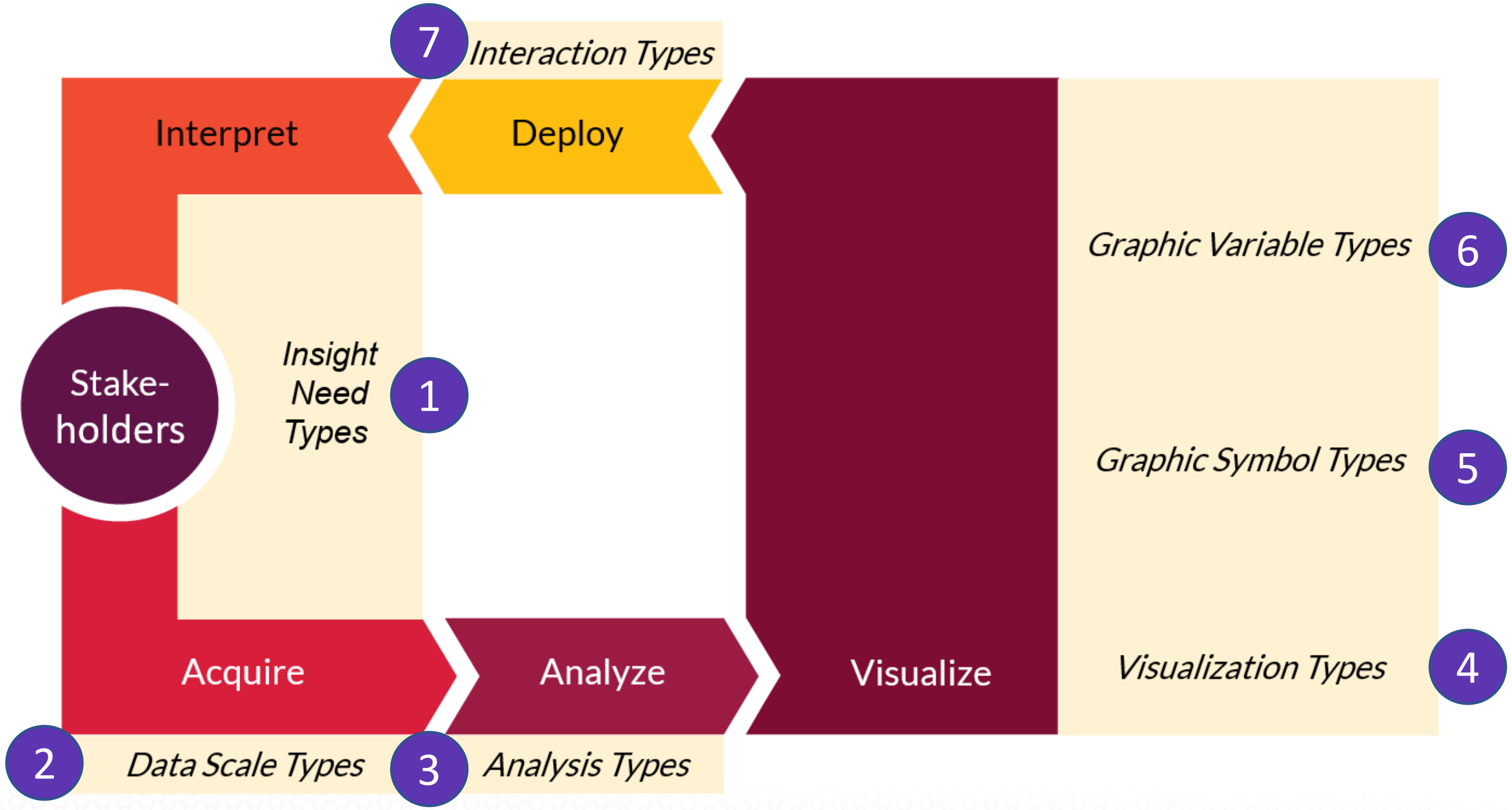
1	2	3	4	5	6	7
Insight Needs	**Data Scales**	**Analyses**	**Visualizations**	**Graphic Symbols**	**Graphic Variables**	**Interactions**
• categorize/cluster • order/rank/sort • distributions (also outliers, gaps) • comparisons • trends (process and time) • geospatial • compositions (also of text) • correlations/ relationships	• nominal • ordinal • interval • ratio	• statistical • temporal • geospatial • topical • relational	• table • chart • graph • map • tree • network	• geometric symbols point line area surface volume • linguistic symbols text numerals punctuation marks • pictorial symbols images icons statistical glyphs	• spatial position • retinal form color optics motion	• zoom • search and locate • filter • details-on-demand • history • extract • link and brush • projection • distortion

Börner, Katy. 2015. *Atlas of Knowledge: Anyone Can Map*. Cambridge, MA: The MIT Press. 25.

Data Visualization Literacy Framework (DVL-FW)

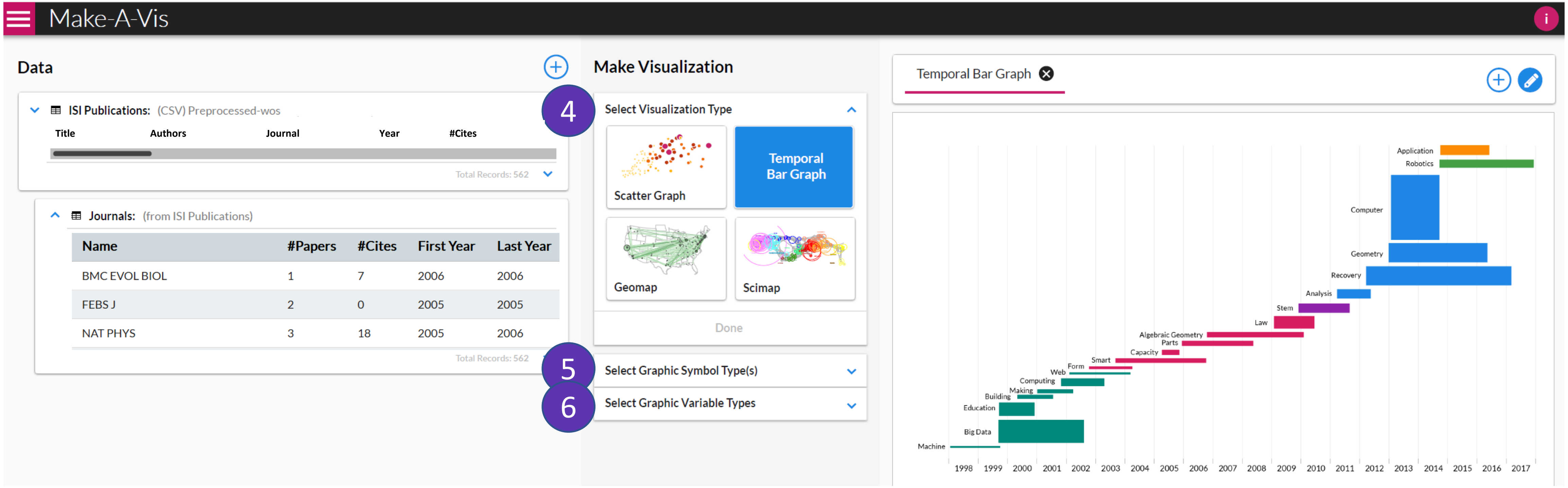
Consists of two parts that are interlinked:

DVL Typology +
DVL Workflow Process



Data Visualization Literacy Framework (DVL-FW)

Implemented in Make-A-Vis (MAV) to support learning via horizontal transfer, scaffolding, hands-on learning, etc.



CNS Cyberinfrastructure for Network Science Center

Graphic Variable Types

Position: x, y; possibly z

Form:

- Size
- Shape
- Rotation (Orientation)

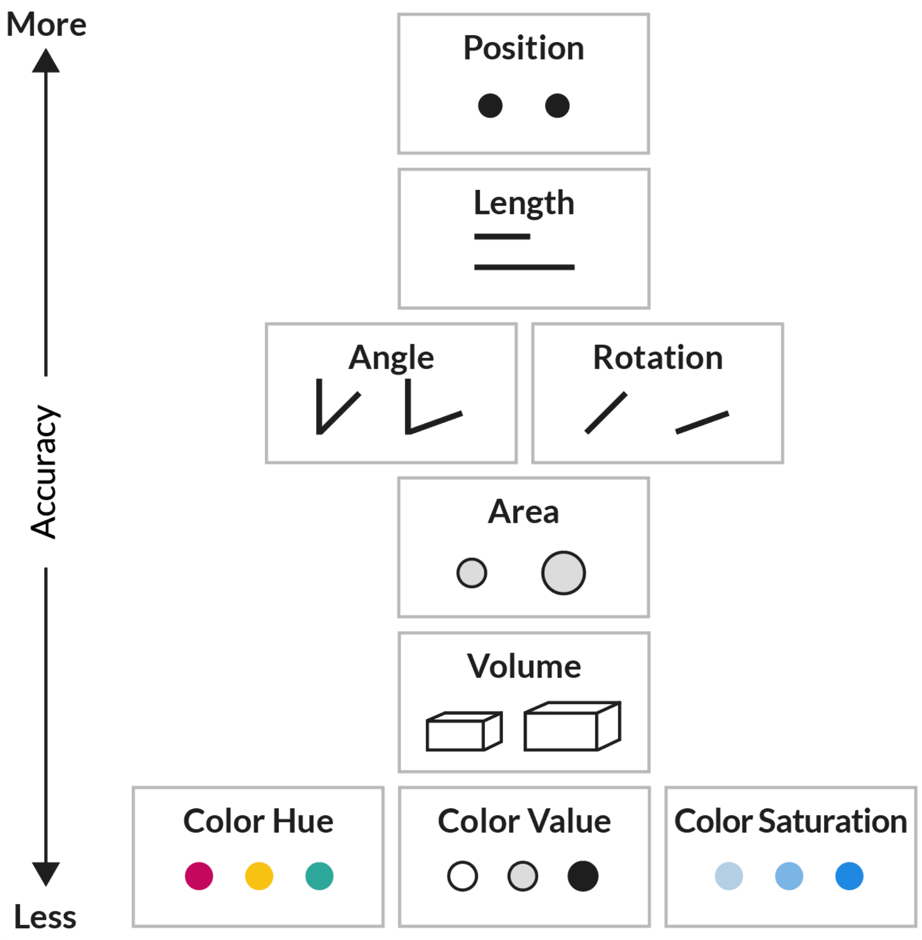
Color:

- Value (Lightness)
- Hue (Tint)
- Saturation (Intensity)

Optics: Blur, Transparency, Shading, Stereoscopic Depth

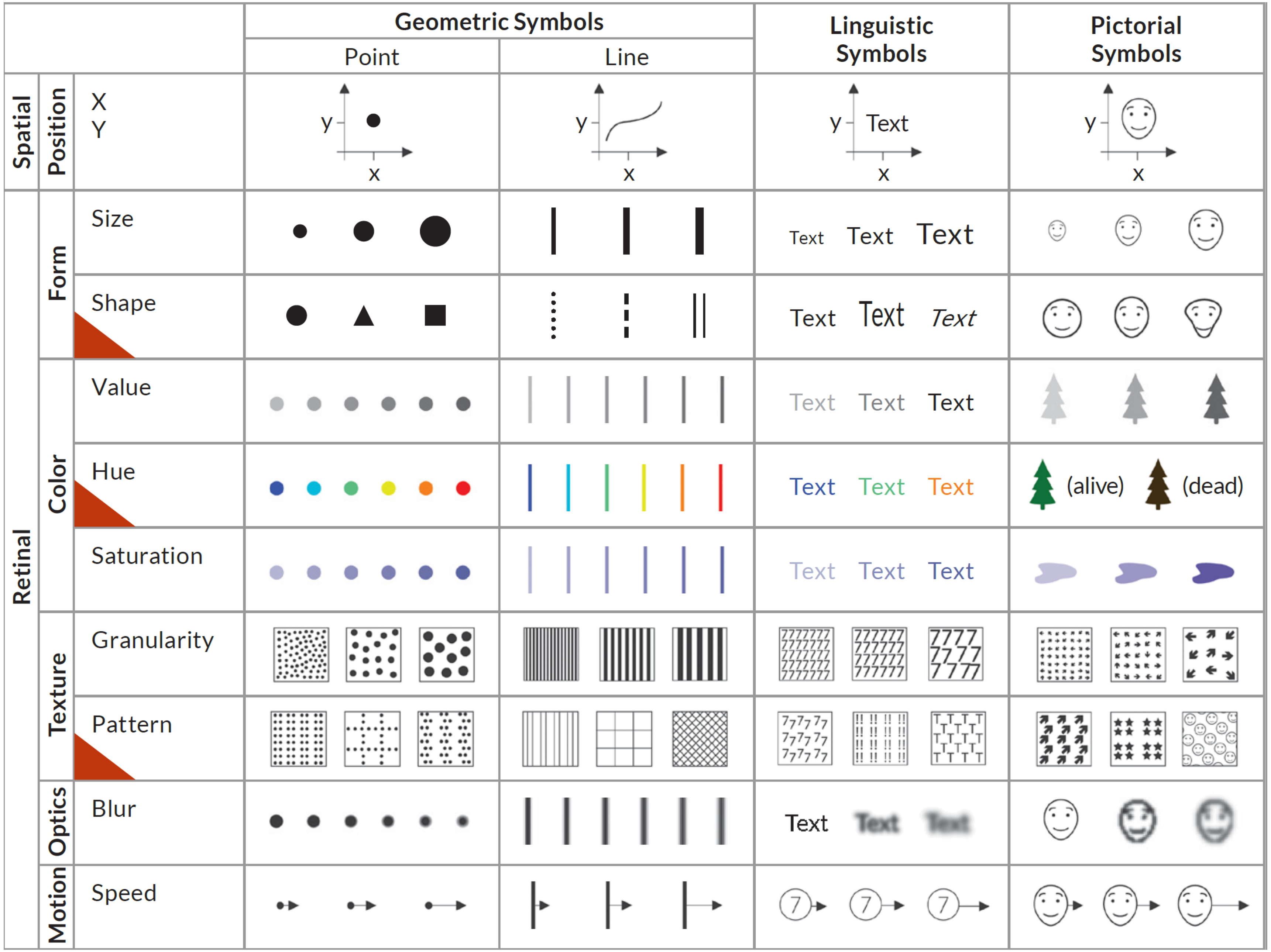
Texture: Spacing, Granularity, Pattern, Orientation, Gradient

Motion: Speed, Velocity, Rhythm



Graphic Symbol Types

Graphic Variable Types



See *Atlas of Knowledge* pages 36-39 for complete table.

Qualitative

Also called:

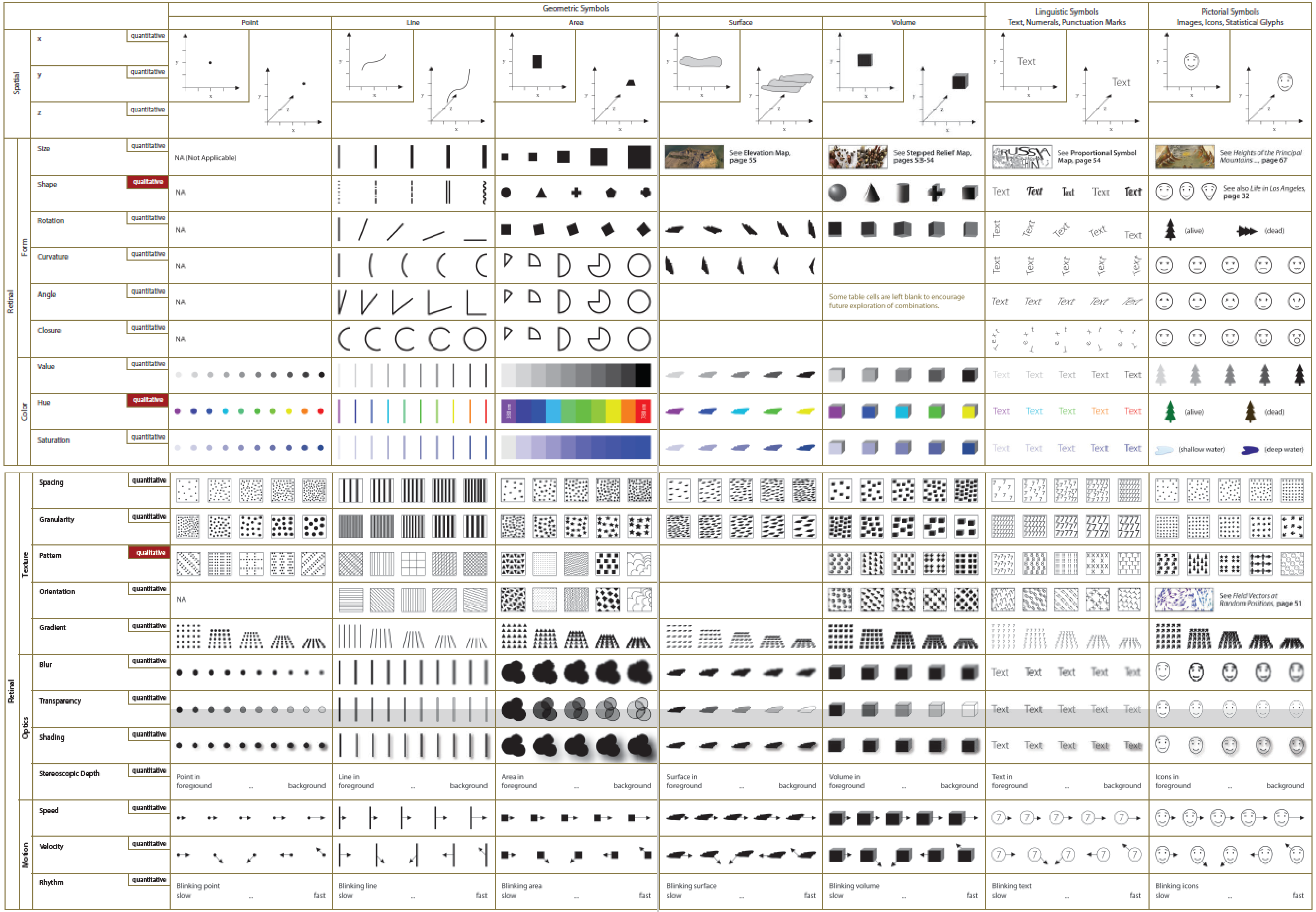
Categorical Attributes
Identity Channels

Quantitative

Also called:

Ordered Attributes
Magnitude Channels

Graphic Variable Types Versus Graphic Symbol Types



See *Atlas of Knowledge* pages 36-39 for complete table.

CNS Cyberinfrastructure for Network Science Center

INDIANA UNIVERSITY
SCHOOL OF INFORMATICS, COMPUTING, AND ENGINEERING

Visual Analytics Certificate

Advance your skills in one of the most in demand careers through this six-week (3 CEUs) online course focused on understanding and creating data visualizations that translate complex data into actionable insights.

DOWNLOAD FLYER

REGISTER FOR SEPT 7–OCT 18, 2020

Learn from Experts

Connect with industry professionals and leading researchers.

Evolve Yourself

Gain forever knowledge and skill-up in powerful data visualization tools.

Make a Difference

Embrace data-driven decision-making in your personal and professional life.

https://visanalytics.cns.iu.edu

Next course starts Sept 7, 2020

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 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

Börner, Katy, William Rouse, Paul Trunfio, and H. Eugene Stanley. 2018. **Forecasting Innovations in Science, Technology, and Education**. *PNAS* 115 (50): 12573-12581. https://www.pnas.org/modeling