

Mapping the Structure and Evolution of (Web) Science

Katy Börner

Victor H. Yngve Distinguished Professor of Information Science
Director, Cyberinfrastructure for Network Science Center
School of Informatics and Computing and IU Network Science Institute
Indiana University, USA

Saving the Web: The Ethics and Challenges of Preserving What's on the Internet
Library of Congress, Washington, D.C.

June 15, 2016



Visualizing (Web) Science



Map of Scientific Collaborations from 2005-2009



Computed Using Data from Elsevier's Scopus

Olivier H. Beauchesne, 2011. Map of Scientific Collaborations from 2005-2009.



1985 – Three first Domain Names: .us for the USA, .uk for the UK, and .il for Israel.

2010 – Astronaut TJ Creamer tweets from the International Space Station.

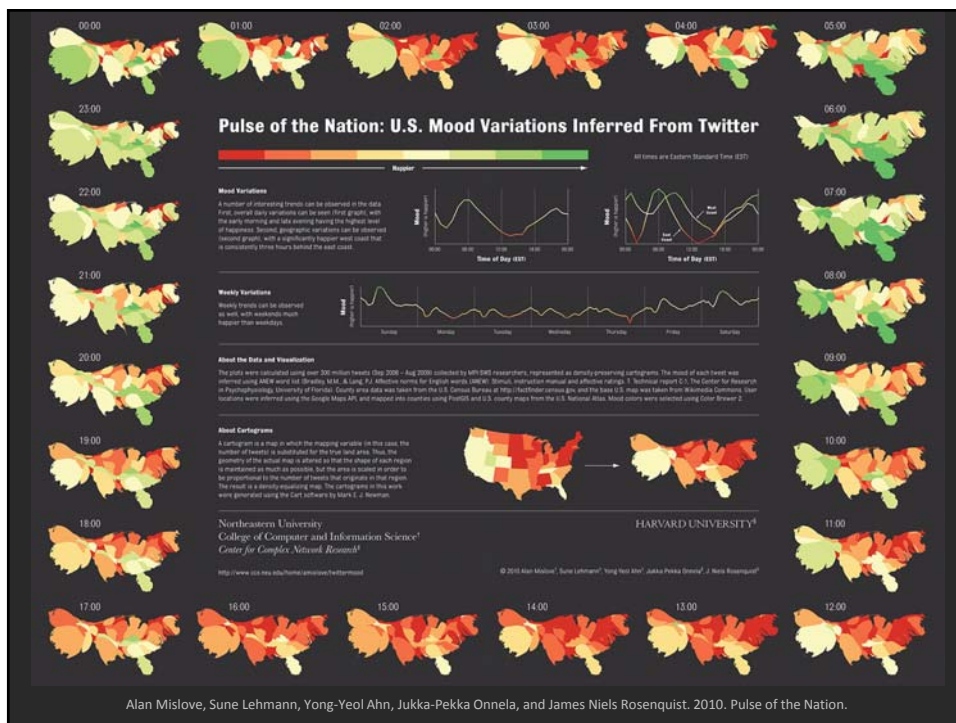
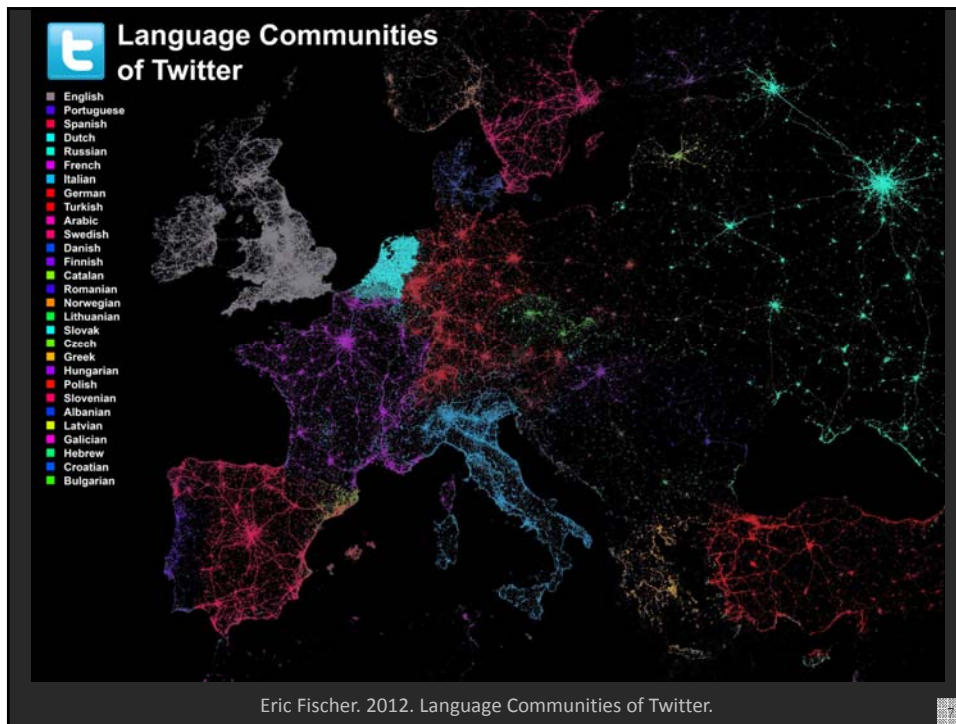
2015 – An estimated 3.2 billion internet users worldwide

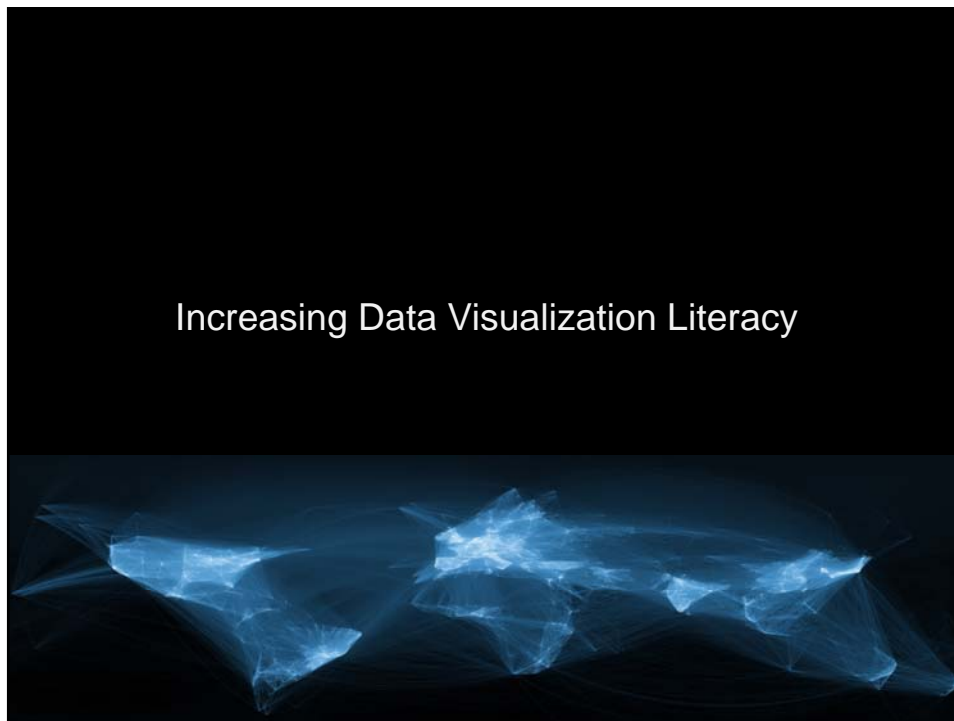
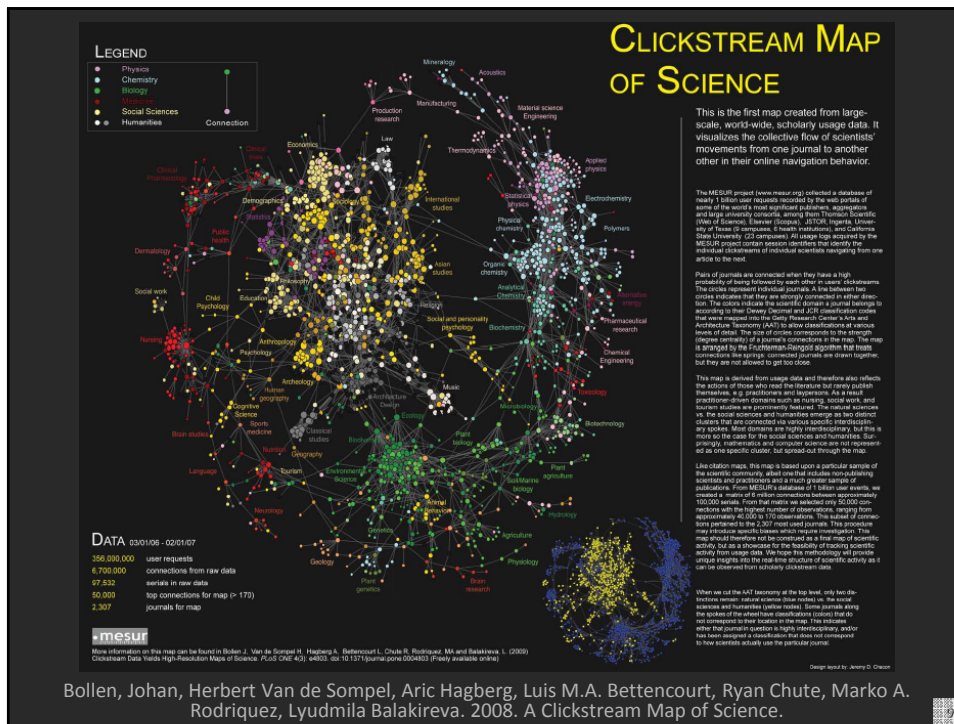
<http://www.nominet.uk/mapping-the-online-world/>



Martin Vargic. 2014. Map of the Internet.













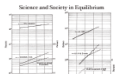





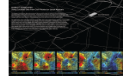












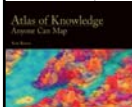
Register for free at <http://ivmooc.cns.iu.edu>

Tasks

LEVELS

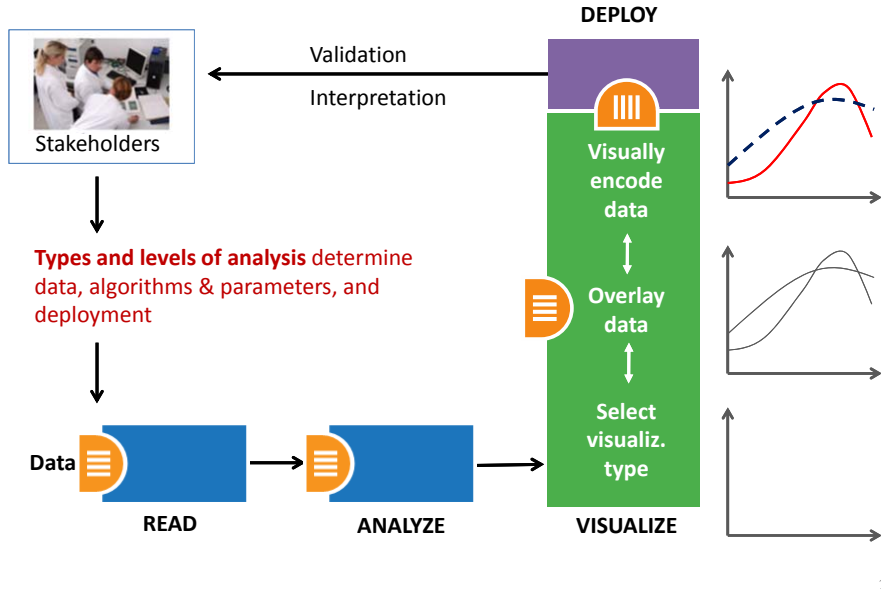
TYPES

	MICRO: Individual Level about 1-1,000 records page 6 	MESO: Local Level about 1,001-100,000 records page 8 	MACRO: Global Level more than 100,000 records page 10 
Statistical Analysis page 44 	 Knowledge Cartography page 110	 Productivity Russian life sciences research page 105	 Science and Society in Equalization Number of scientists versus population and R&D costs versus GNP page 103
WHEN: Temporal Analysis page 48 	 Visualizing decision making processes page 95	 Key events in the development of the video tape recorder page 85	 Increased travel and communication speeds page 83
WHERE: Geospatial Analysis page 52 	 Cell phone usage in Milan, Italy page 109	 Victorian poetry in Europe page 137	 Ecological footprint of countries page 99
WHAT: Topical Analysis page 56 	 Evolving patent holdings of Apple Computer, Inc. and Jerome Lemelson page 89	 Evolving journal networks in nanotechnology page 139	 Product space showing co-export patterns of countries page 95
WITH WHOM: Network Analysis page 60 	 World Finance Corporation network page 87	 Electronic and new media art networks page 153	 World-wide scholarship collaboration networks page 127



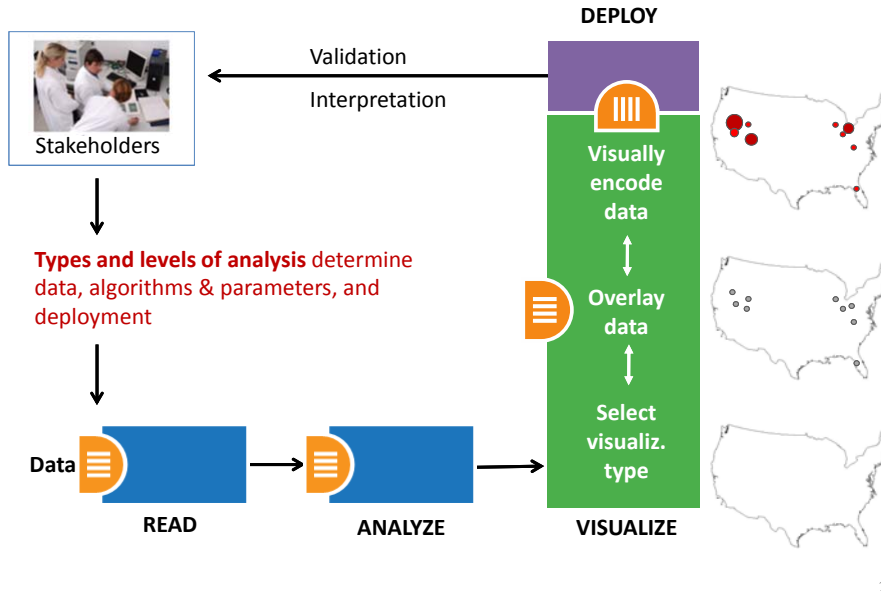
See page 5

Needs-Driven Workflow Design



13

Needs-Driven Workflow Design



14

Course Schedule

Part 1: Theory and Hands-On

- **Session 1** – Workflow Design and Visualization Framework
- **Session 2** – “When:” Temporal Data
- **Session 3** – “Where:” Geospatial Data
- **Session 4** – “What:” Topical Data

Mid-Term

- **Session 5** – “With Whom:” Trees
- **Session 6** – “With Whom:” Networks
- **Session 7** – Dynamic Visualizations and Deployment

Final Exam

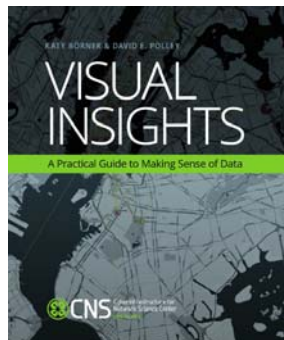
Part 2: Students work in teams on client projects.

Final grade is based on Class Participation (10%), Midterm (30%), Final Exam (30%), and Client Project(30%).



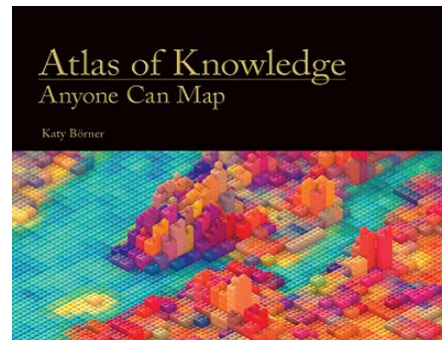
15

Books Used in the IVMOOC



Teaches timely knowledge:

Advanced algorithms, tools, and hands-on workflows.



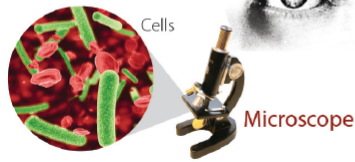
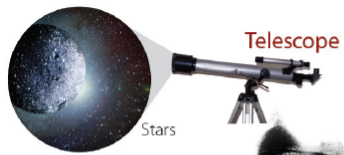
Teaches timeless knowledge:

Visualization framework—exemplified using generic visualization examples and pioneering visualizations.

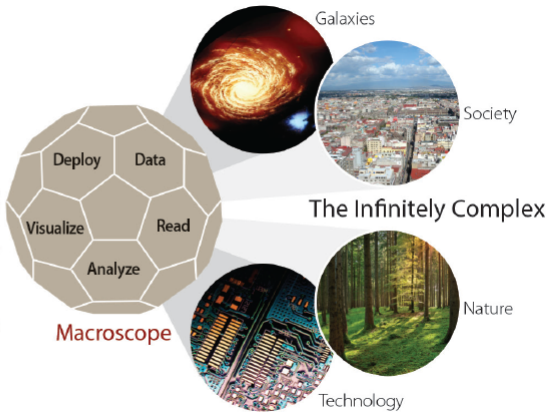
16

Microscopes, Telescopes, Macrosopes Plug-and-Play Macrosopes

The Infinitely Great



The Infinitely Small



17

Load **One** File and Run **Many** Analyses and Visualizations

Times Cited	Publication Year	City of Publisher	Country	Journal Title (Full)	Title	Subject Category	Authors
12	2011	NEW YORK	USA	COMMUNICATIONS OF THE ACM	Plug-and-Play Macrosopes	Computer Science	Borner, K
18	2010	MALDEN	USA	CTS-CLINICAL AND TRANSLATIONAL SCIENCE	Advancing the Science of Team Science	Research & Experimental Medicine	Falk-Krzesinski, HJ Borner, K Contractor, N Fiore, SM Hall, KL Keyton, J Spring, B Stokols, D Trochim, W Uzzi, B
13	2010	WASHINGTON	USA	SCIENCE TRANSLATIONAL MEDICINE	A Multi-Level Systems Perspective for the Science of Team Science	Cell Biology Research & Experimental Medicine	Borner, K Contractor, N Falk-Krzesinski, HJ Fiore, SM Hall, KL Keyton, J Spring, B Stokols, D Trochim, W Uzzi, B

Statistical Analysis—p. 44

Location	Count	# Citations
Netherlands	13	292
United States	9	318
Germany	11	36
United Kingdom	1	2

Temporal Burst Analysis—p. 48



Geospatial Analysis—p. 52



Geospatial Analysis—p. 52

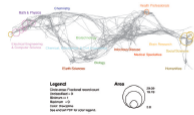


18

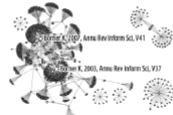
Load **One** File and Run **Many** Analyses and Visualizations

Times Cited	Publication Year	City of Publisher	Country	Journal Title (Full)	Title	Subject Category	Authors
12	2011	NEW YORK	USA	COMMUNICATIONS OF THE ACM	Plug-and-Play Macroscopes	Computer Science	Borner, K
18	2010	MALDEN	USA	CTS-CLINICAL AND TRANSLATIONAL SCIENCE	Advancing the Science of Team Science	Research & Experimental Medicine	Falk-Krzesinski, HJ Borner, K Contractor, N Fiore, SM Hall, KL Keyton, J Spring, B Stokols, D Trochim, W Uzzi, B
13	2010	WASHINGTON	USA	SCIENCE TRANSLATIONAL MEDICINE	A Multi-Level Systems Perspective for the Science of Team Science	Cell Biology Research & Experimental Medicine	Borner, K Contractor, N Falk-Krzesinski, HJ Fiore, SM Hall, KL Keyton, J Spring, B Stokols, D Trochim, W Uzzi, B

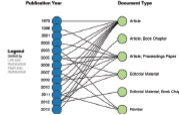
Topical Analysis—p. 56



Paper Citation Network—p. 60



Bi-Modal Network—p. 60



Co-author and many other bi-modal networks.

19



Sci2 Tool Interface Components

Download tool for free at <http://sci2.cns.iu.edu>

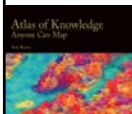
The screenshot displays the Sci2 Tool interface with several components visible:

- Console:** Shows command-line output for various operations like 'Extract Co-Occurrence Network' and 'Extract Co-Authorship Network'.
- Data Manager:** Lists data sources such as 'ISI Data' and '361 Unique ISI Records'.
- Workflow Manager:** Shows a sequence of operations performed, including 'Extracted Co-Authorship Network' and 'Network with degree attribute added to nodes'.
- Menu:** A 'Networks' menu is open, showing options like 'Extract Top Nodes', 'Extract Nodes Above or Below Value', and 'Delete Isolates'.
- Visualization Panel:** A 'Visualizations' menu is open, listing various network visualization techniques such as 'Radial Tree/Graph', 'Tree View', and 'Force Directed with Annotation'.

20

Visualization Framework

Insight Need Types page 26	Data Scale Types page 28	Visualization Types page 30	Graphic Symbol Types page 32	Graphic Variable Types page 34	Interaction Types page 26
<ul style="list-style-type: none"> categorize/cluster order/rank/sort distributions (also outliers, gaps) comparisons trends (process and time) geospatial compositions (also of text) correlations/relationships 	<ul style="list-style-type: none"> nominal ordinal interval ratio 	<ul style="list-style-type: none"> table chart graph map network layout 	<ul style="list-style-type: none"> geometric symbols <ul style="list-style-type: none"> point line area surface volume linguistic symbols <ul style="list-style-type: none"> text numerals punctuation marks pictorial symbols <ul style="list-style-type: none"> images icons statistical glyphs 	<ul style="list-style-type: none"> spatial <ul style="list-style-type: none"> position retinal <ul style="list-style-type: none"> form color optics motion 	<ul style="list-style-type: none"> overview zoom search and locate filter details-on-demand history extract link and brush projection distortion



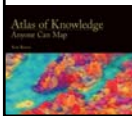
See page 24

Visualization Framework

Basic Task Types								
Bertin, 1967	Wehrend & Lewis, 1996	Few, 2004	Yau, 2011	Rendgen & Wiedemann, 2012	Frankel, 2012	Tool: Many Eyes	Tool: Chart Chooser	Börner, 2014
selection	categorize			category				categorize/cluster
order	rank	ranking					table	order/rank/sort
	distribution	distribution					distribution	distributions (also outliers, gaps)
	compare	nominal comparison & deviation	differences		compare and contrast	compare data values	comparison	comparisons
		time series	patterns over time	time	process and time	track rises and falls over time	trend	trends (process and time)
		geospatial	spatial relations	location		generate maps		geospatial
quantity		part-to-whole	proportions		form and structure	see parts of whole, analyze text	composition	compositions (also of text)
association	correlate	correlation	relationships	hierarchy		relations between data points	relationship	correlations/relationships

Visualization Framework

Insight Need Types page 26	Data Scale Types page 28	Visualization Types page 30	Graphic Symbol Types page 32	Graphic Variable Types page 34	Interaction Types page 26
<ul style="list-style-type: none"> • categorize/cluster • order/rank/sort • distributions (also outliers, gaps) • comparisons • trends (process and time) • geospatial • compositions (also of text) • correlations/relationships 	<ul style="list-style-type: none"> • nominal • ordinal • interval • ratio 	<ul style="list-style-type: none"> • table • chart • graph • map • network layout 	<ul style="list-style-type: none"> • geometric symbols <ul style="list-style-type: none"> • point • line • area • surface • volume • linguistic symbols <ul style="list-style-type: none"> • text • numerals • punctuation marks • pictorial symbols <ul style="list-style-type: none"> • images • icons • statistical glyphs 	<ul style="list-style-type: none"> • spatial <ul style="list-style-type: none"> • position • retinal <ul style="list-style-type: none"> • form • color • optics • motion 	<ul style="list-style-type: none"> • overview • zoom • search and locate • filter • details-on-demand • history • extract • link and brush • projection • distortion



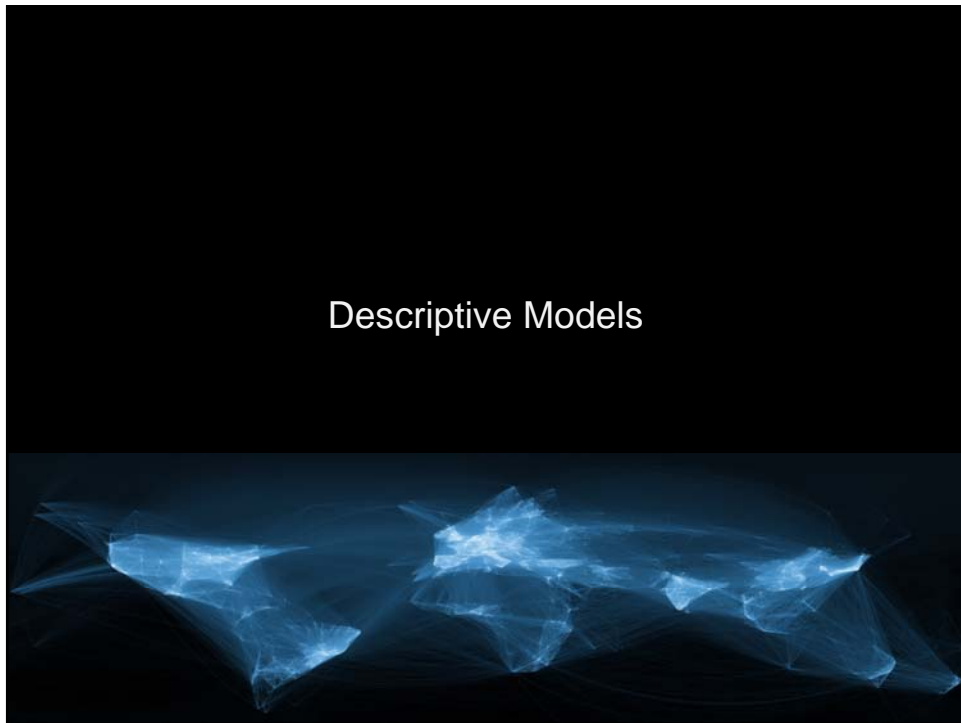
See page 24

Graphic Variable Types Versus Graphic Symbol Types

			Point	Line	Geometric Symbols Area
Spatial	x	quantitative			
	y	quantitative			
	z	quantitative			
Form	Size	quantitative	NA (Not Applicable)		
	Shape	qualitative	NA		
	Rotation	quantitative	NA		
	Curvature	quantitative	NA		
	Angle	quantitative	NA		
	Closure	quantitative	NA		
Color	Value	quantitative			
	Hue	qualitative			
	Saturation	quantitative			

Graphic Variable Types Versus Graphic Symbol Types

Type	Symbol	Geometric Symbols				Alphabetic Symbols				Pictorial Symbols			
		point	line	area	volume	text	text	text	text	image	image	image	image
POINT	dot	dot	dot	dot	dot	dot	dot	dot	dot	dot	dot	dot	dot
	dash	dash	dash	dash	dash	dash	dash	dash	dash	dash	dash	dash	dash
	stroke	stroke	stroke	stroke	stroke	stroke	stroke	stroke	stroke	stroke	stroke	stroke	stroke
	arc	arc	arc	arc	arc	arc	arc	arc	arc	arc	arc	arc	arc
	circle	circle	circle	circle	circle	circle	circle	circle	circle	circle	circle	circle	circle
	square	square	square	square	square	square	square	square	square	square	square	square	square
	rectangle	rectangle	rectangle	rectangle	rectangle	rectangle	rectangle	rectangle	rectangle	rectangle	rectangle	rectangle	rectangle
	triangle	triangle	triangle	triangle	triangle	triangle	triangle	triangle	triangle	triangle	triangle	triangle	triangle
	star	star	star	star	star	star	star	star	star	star	star	star	star
	cross	cross	cross	cross	cross	cross	cross	cross	cross	cross	cross	cross	cross
LINE	line	line	line	line	line	line	line	line	line	line	line	line	line
	curve	curve	curve	curve	curve	curve	curve	curve	curve	curve	curve	curve	curve
	arc	arc	arc	arc	arc	arc	arc	arc	arc	arc	arc	arc	arc
	circle	circle	circle	circle	circle	circle	circle	circle	circle	circle	circle	circle	circle
	square	square	square	square	square	square	square	square	square	square	square	square	square
	rectangle	rectangle	rectangle	rectangle	rectangle	rectangle	rectangle	rectangle	rectangle	rectangle	rectangle	rectangle	rectangle
	triangle	triangle	triangle	triangle	triangle	triangle	triangle	triangle	triangle	triangle	triangle	triangle	triangle
	star	star	star	star	star	star	star	star	star	star	star	star	star
	cross	cross	cross	cross	cross	cross	cross	cross	cross	cross	cross	cross	cross
	text	text	text	text	text	text	text	text	text	text	text	text	text
AREA	area	area	area	area	area	area	area	area	area	area	area	area	area
	curve	curve	curve	curve	curve	curve	curve	curve	curve	curve	curve	curve	curve
	arc	arc	arc	arc	arc	arc	arc	arc	arc	arc	arc	arc	arc
	circle	circle	circle	circle	circle	circle	circle	circle	circle	circle	circle	circle	circle
	square	square	square	square	square	square	square	square	square	square	square	square	square
	rectangle	rectangle	rectangle	rectangle	rectangle	rectangle	rectangle	rectangle	rectangle	rectangle	rectangle	rectangle	rectangle
	triangle	triangle	triangle	triangle	triangle	triangle	triangle	triangle	triangle	triangle	triangle	triangle	triangle
	star	star	star	star	star	star	star	star	star	star	star	star	star
	cross	cross	cross	cross	cross	cross	cross	cross	cross	cross	cross	cross	cross
	text	text	text	text	text	text	text	text	text	text	text	text	text
VOLUME	volume	volume	volume	volume	volume	volume	volume	volume	volume	volume	volume	volume	volume
	curve	curve	curve	curve	curve	curve	curve	curve	curve	curve	curve	curve	curve
	arc	arc	arc	arc	arc	arc	arc	arc	arc	arc	arc	arc	arc
	circle	circle	circle	circle	circle	circle	circle	circle	circle	circle	circle	circle	circle
	square	square	square	square	square	square	square	square	square	square	square	square	square
	rectangle	rectangle	rectangle	rectangle	rectangle	rectangle	rectangle	rectangle	rectangle	rectangle	rectangle	rectangle	rectangle
	triangle	triangle	triangle	triangle	triangle	triangle	triangle	triangle	triangle	triangle	triangle	triangle	triangle
	star	star	star	star	star	star	star	star	star	star	star	star	star
	cross	cross	cross	cross	cross	cross	cross	cross	cross	cross	cross	cross	cross
	text	text	text	text	text	text	text	text	text	text	text	text	text



The Global 'Scientific Food Web'

Mazloumian, Amin, Dirk Helbing, Sergi Lozano, Robert Light, and Katy Börner. 2013. "Global Multi-Level Analysis of the 'Scientific Food Web'". *Scientific Reports* 3, 1167.

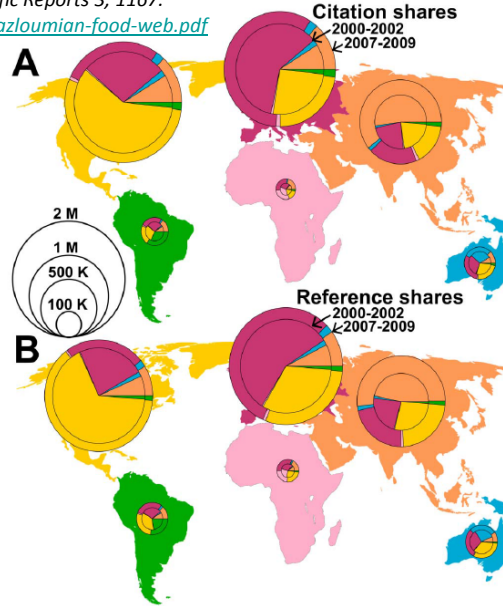
<http://cns.iu.edu/docs/publications/2013-mazloumian-food-web.pdf>

Contributions:

Comprehensive global analysis of scholarly knowledge production and diffusion on the level of continents, countries, and cities.

Quantifying knowledge flows between 2000 and 2009, we identify global sources and sinks of knowledge production. Our knowledge flow index reveals, where ideas are born and consumed, thereby defining a global 'scientific food web'.

While Asia is quickly catching up in terms of publications and citation rates, we find that its dependence on knowledge consumption has further increased.



27

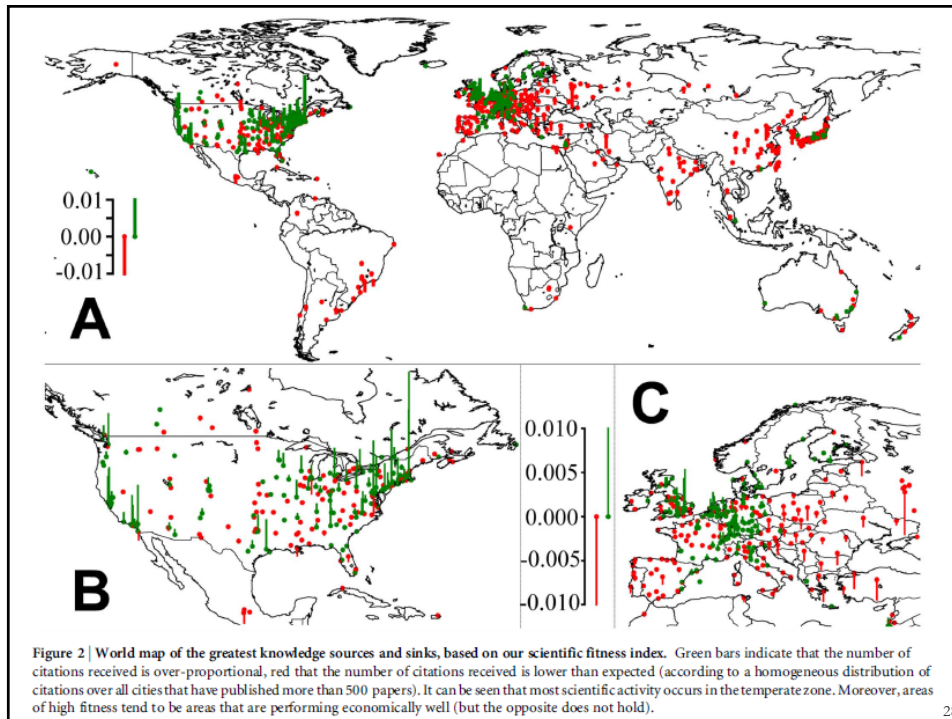


Figure 2 | World map of the greatest knowledge sources and sinks, based on our scientific fitness index. Green bars indicate that the number of citations received is over-proportional, red that the number of citations received is lower than expected (according to a homogeneous distribution of citations over all cities that have published more than 500 papers). It can be seen that most scientific activity occurs in the temperate zone. Moreover, areas of high fitness tend to be areas that are performing economically well (but the opposite does not hold).

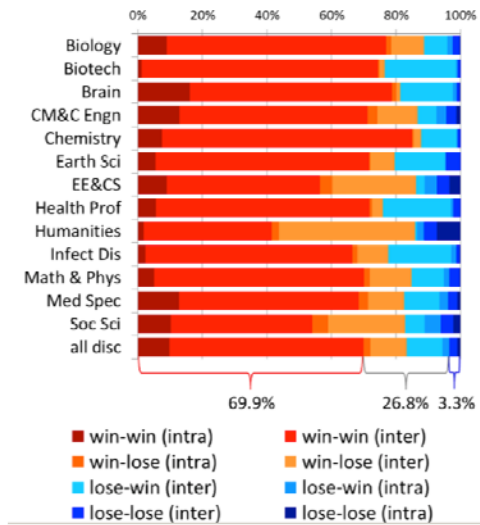
28

Long-Distance Interdisciplinarity Leads to Higher Scientific Impact

Larivière, Vincent, Stefanie Haustein, and Katy Börner. 2015. PLOS ONE DOI: 10.1371.

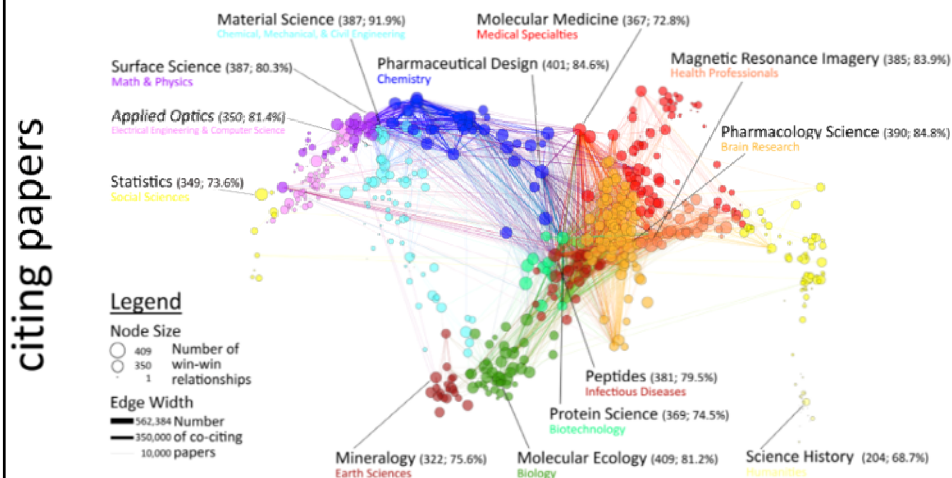
Data: 9.2 million interdisciplinary research papers published between 2000 and 2012.

Results: majority (69.9%) of co-cited interdisciplinary pairs are “win-win” relationships, i.e., papers that cite them have higher citation impact and there are as few as 3.3% “lose-lose” relationships. UCSD map of science is used to compute “distance.”



29

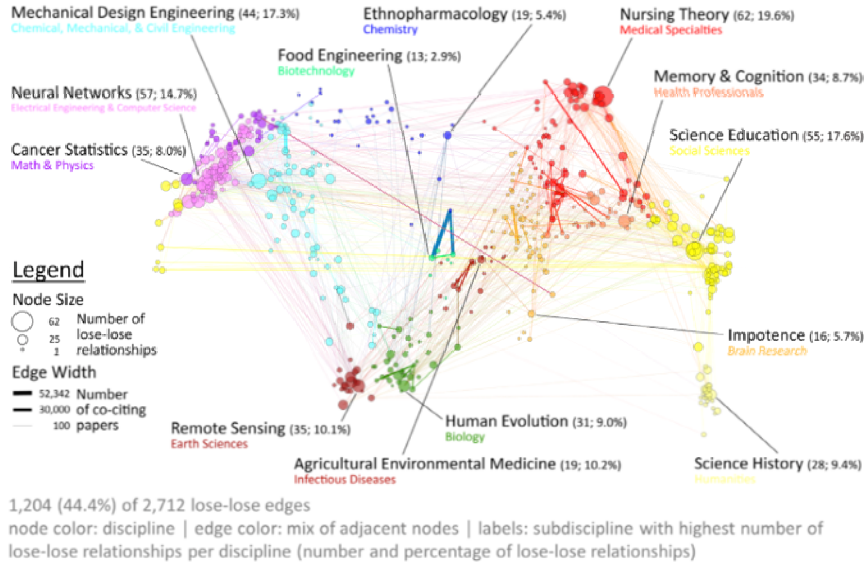
A1 Number of papers citing win-win relationships ($\geq 10,000$ citing papers)



2,940 (5.19%) of 56,614 win-win edges
 node color: discipline | edge color: mix of adjacent nodes | labels: subdiscipline with highest number of win-win relationships per discipline (number and percentage of win-win relationships)

30

B1 Number of papers citing lose-lose relationships (≥ 100 citing papers)



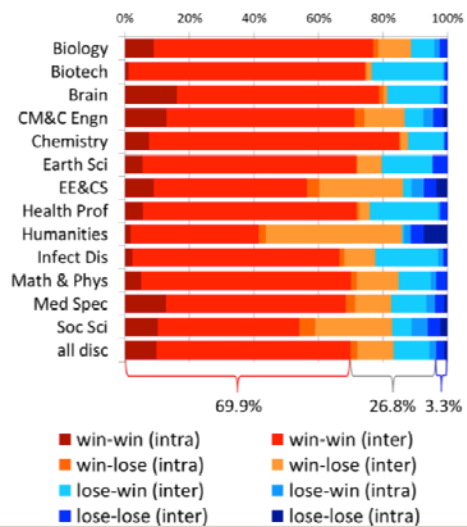
31

Long-Distance Interdisciplinarity Leads to Higher Scientific Impact

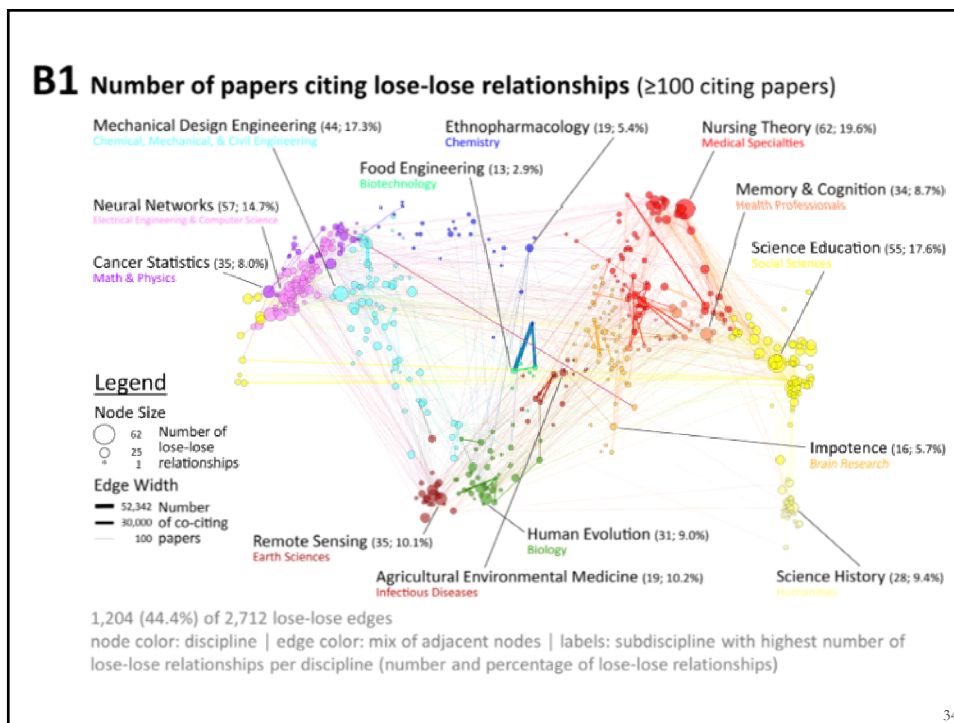
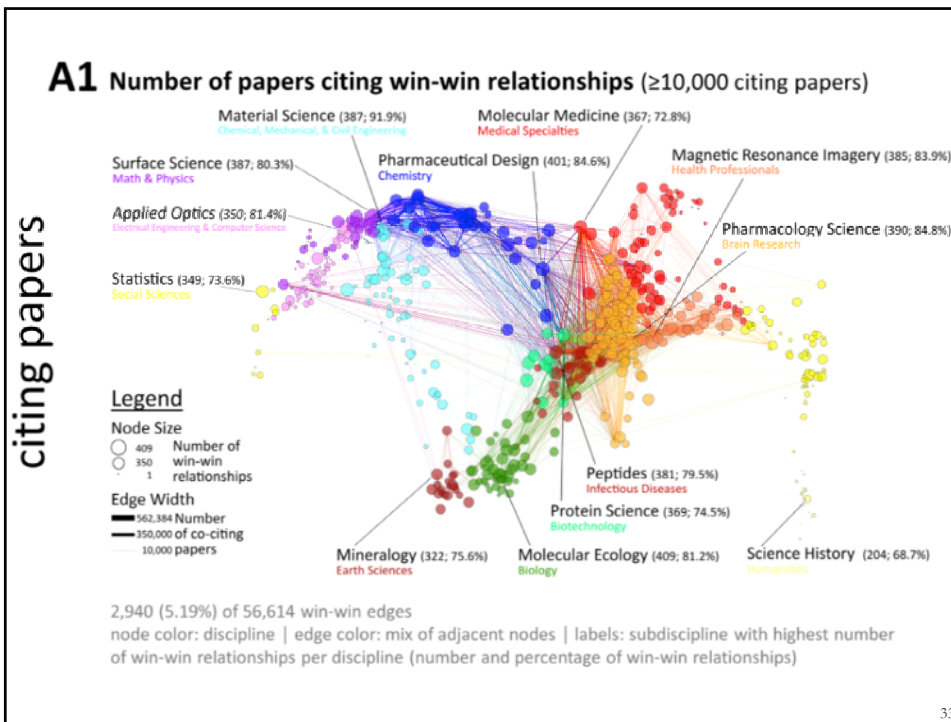
Larivière, Vincent, Stefanie Haustein, and Katy Börner. 2015. *PLOS ONE* DOI: 10.1371.

Data: 9.2 million interdisciplinary research papers published between 2000 and 2012.

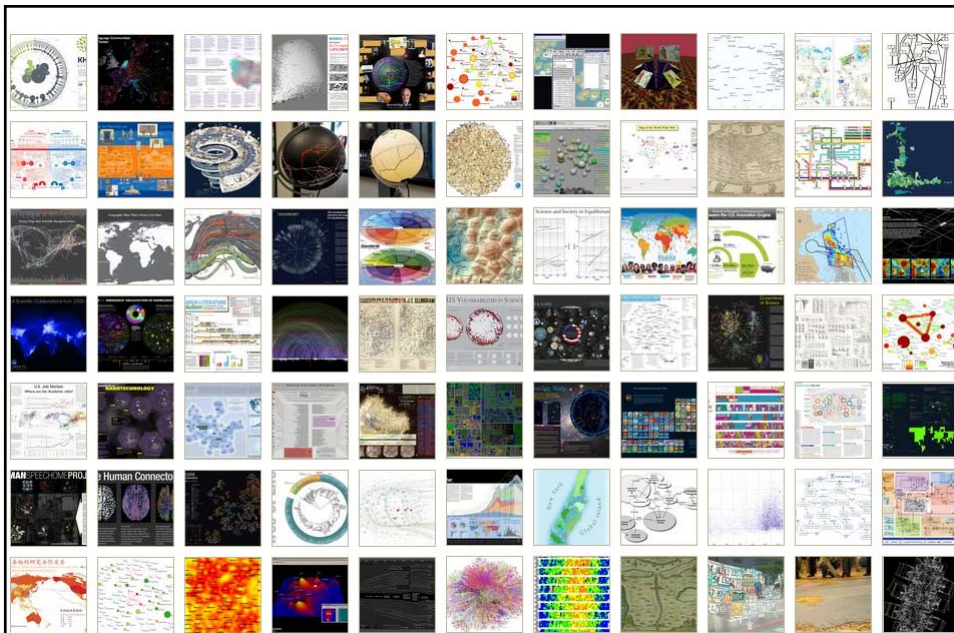
Results: majority (69.9%) of co-cited interdisciplinary pairs are “win-win” relationships, i.e., papers that cite them have higher citation impact and there are as few as 3.3% “lose-lose” relationships. UCSD map of science is used to compute “distance.”



32



Communicating (Web) Science to Different Audiences



Places & Spaces: Mapping Science Exhibit, online at <http://scimaps.org>

PLACES & SPACES & MAPPING SCIENCE

Curated by the Cyberinfrastructure for Network Science Center

search scimaps.org Search

About People Maps & More Exhibitions Hosting Publications Store News Contact

Hidalgo, César A., Bailey Klings, Albert-László Barabási, and Ricardo Hausmann, 2007. See also The Product Space map from Phase I of Places & Spaces.

Call for Macroscope Tools for the *Places & Spaces: Mapping Science* Exhibit (2016) <http://scimaps.org/call>

Background and Goals

The *Places & Spaces: Mapping Science* exhibit was created to in communicate human activity and scientific progress on a gloc that enable the close inspection of large-scale maps in public conferences; (2) novel, interactive macroscope tools that let

Themes for the upcoming iterations/years are:

- 11th Iteration (2015): Macroscopes for Interacting With Science
- 12th Iteration (2016): Macroscopes for Making Sense of Science
- 13th Iteration (2017): Macroscopes for Forecasting Science
- 14th Iteration (2018): Macroscopes for Economic Decision Makers
- 15th Iteration (2019): Macroscopes for Science Policy Makers

37

MACROSCOPES FOR INTERACTING WITH SCIENCE

PLACES & SPACES & MAPPING SCIENCE

Earth AcademyScope Mapping Global Society Charting Culture

<http://scimaps.org/iteration/11>

Macroscope Standard Setting

Plug-and-play system architectures—supporting workflow design.

- Oct 17, 2015, ASTC Panel on “Plug-and-Play Macroscopes: Modular Hardware and Software Platforms that Render Data into Insights” in Montreal, Canada
- See other workshops and slides at <http://cns.iu.edu/workshops>

Data-code-vis-expertise marketplaces—easy access to relevant datasets and tools.

- OSGI+CIShell, D3, ESRI, Plotly, many others

Visualization hardware—support existing de-facto display standards, envision novel interfaces.

- Science on a Sphere, augmented realities, IoT



Science on a Sphere by NOAA



Beesley's Living Architectures

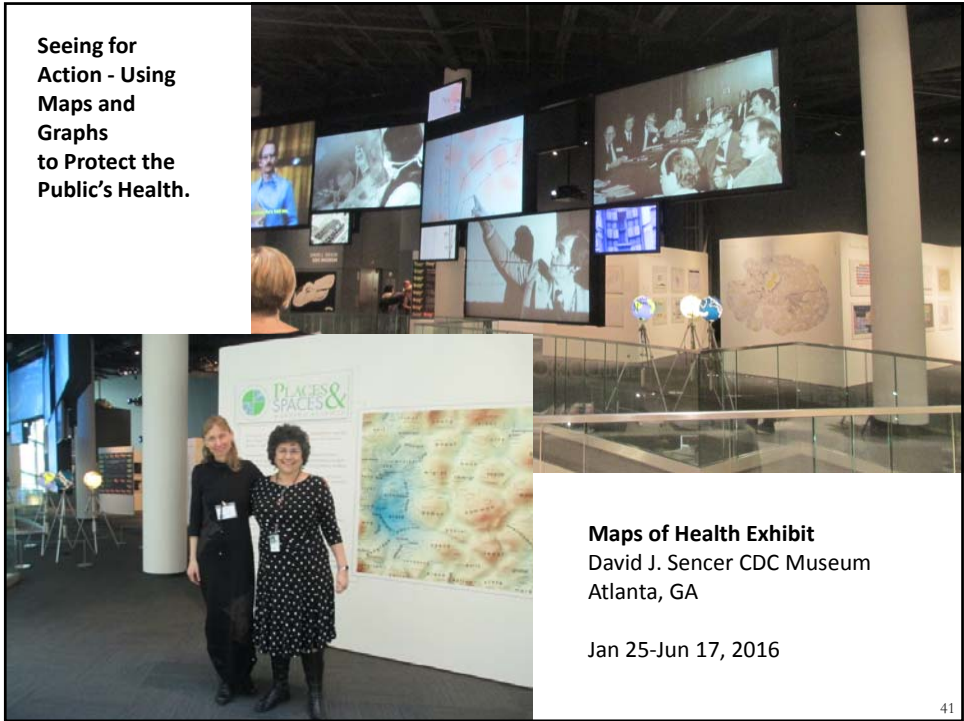
39



Places & Spaces Exhibit at the David J. Sencer CDC Museum, Atlanta, GA
January 25-June 17, 2016

40

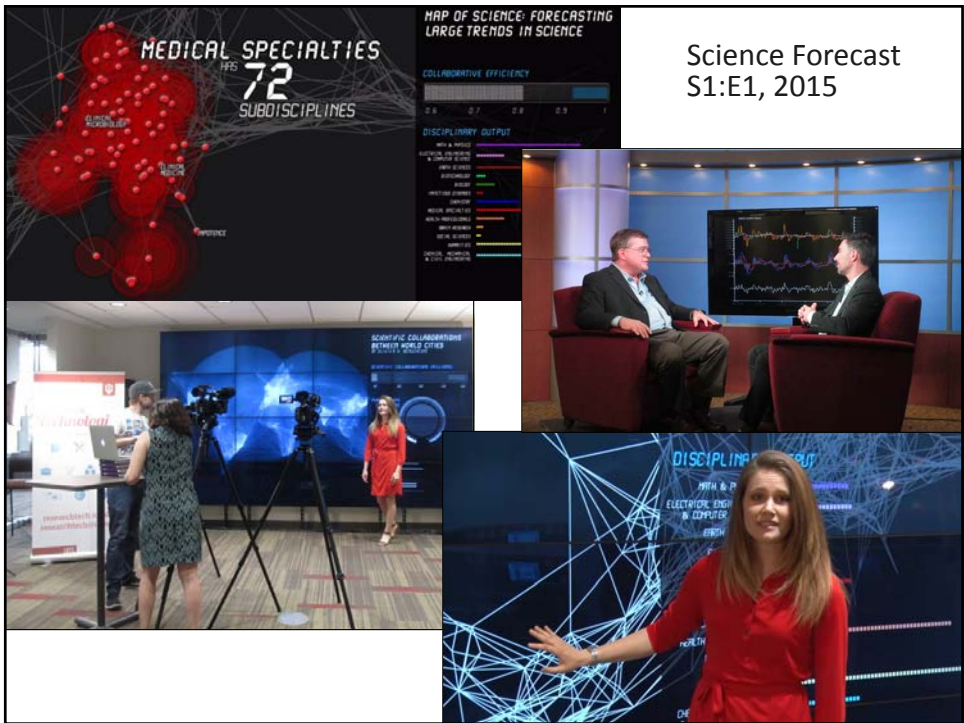
Seeing for Action - Using Maps and Graphs to Protect the Public's Health.



Maps of Health Exhibit
David J. Sencer CDC Museum
Atlanta, GA

Jan 25-Jun 17, 2016

41



Science Forecast
S1:E1, 2015

Predictive Models



[Home](#) [Agenda](#) [Confirmed Speakers](#) [Organizers & Advisors](#) [Venue](#) [Attend](#) [Contact](#)

Modeling Science, Technology & Innovation Conference

WASHINGTON D.C. | MAY 16-18, 2016

[View Agenda](#)

Co-funded by the NSF Science of Science and Innovation Policy (SciSIP) program. Conference brings together international experts and practitioners that develop and apply mathematical, statistical, and computational models to increase our understanding of the structure and dynamics of science, technology and innovation. Slides and recordings are available at <http://modsti.cns.iu.edu>.

44

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>



45

All papers, maps, tools, talks, press are linked from <http://cns.iu.edu>

These slides are at <http://cns.iu.edu/docs/presentations>

CNS Facebook: <http://www.facebook.com/cnscenter>

Mapping Science Exhibit Facebook: <http://www.facebook.com/mappingscience>

46