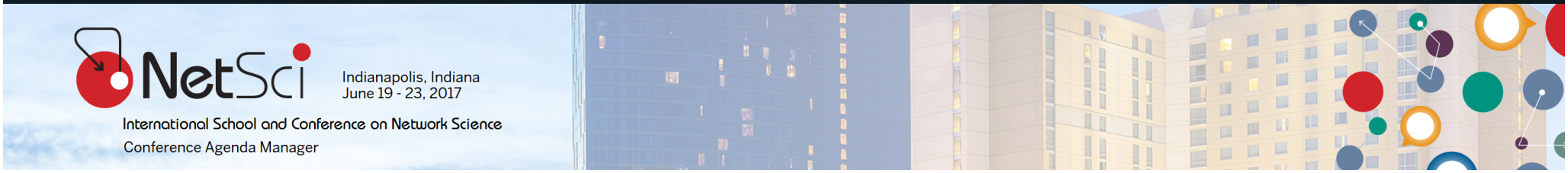
(Network) Data Visualization Literacy

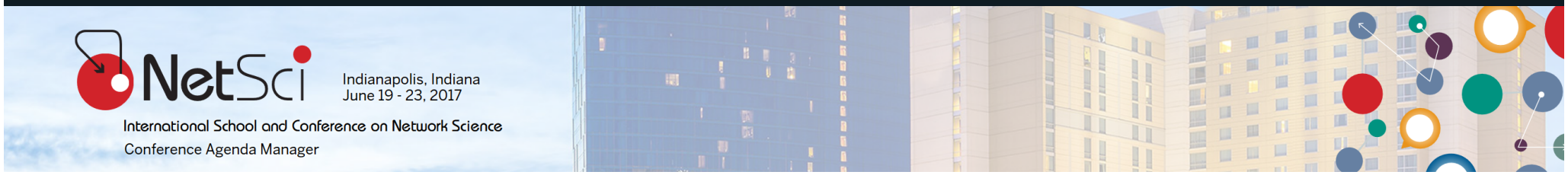
Katy Börner @katycns

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

NetSciEd6, NetSci Conferences, Indianapolis, IN
June 20, 2017



Data Visualization Literacy: Definition



Data Visualization Literacy

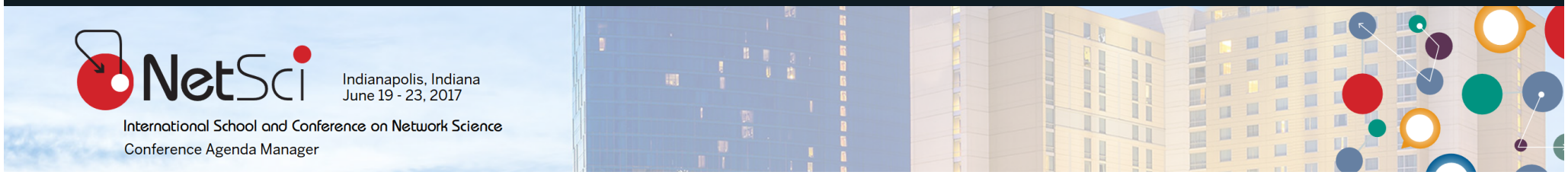
Data visualization literacy (ability to read, make, and explain data visualizations) requires

- *literacy* (ability to read and write text, e.g., in titles, axis labels, legend),
- *visual literacy* (ability to find, interpret, evaluate, use, and create images and visual media), and
- *data literacy* (ability to read, create, and communicate data).

Network Visualization Literacy: Novel Approaches to Measurement and Instruction

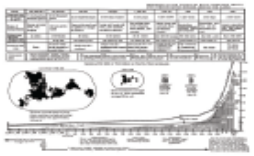
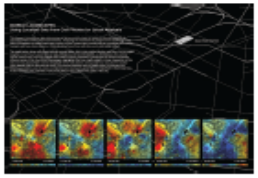
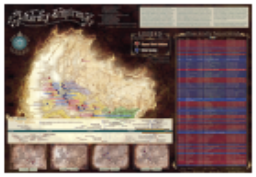
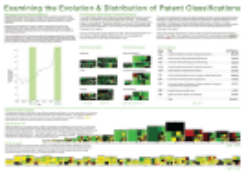
Angela Zoss, Duke University
Adam Maltese, Indiana University
Stephen Uzzo, New York Hall of Science
Katy Börner, Indiana University

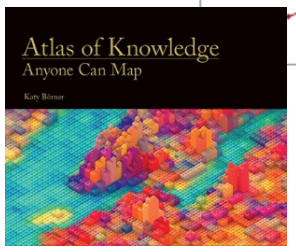
Data Visualization Literacy: Teaching



Register for free: http://ivmooc.cns.iu.edu. Class starts again Jan 9, 2017.

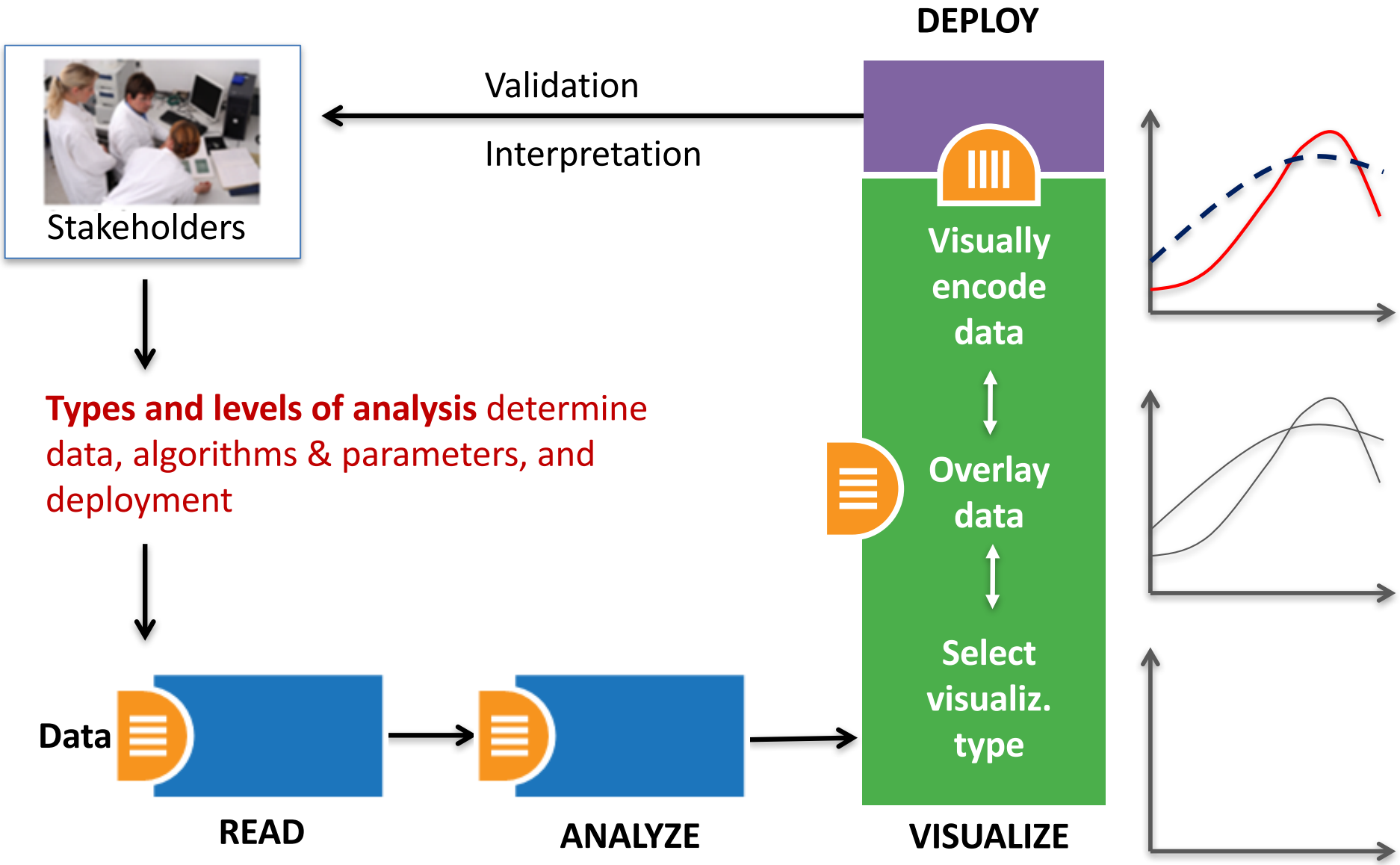
Tasks

TYPES	LEVELS 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 135	 Productivity of Russian life sciences research teams page 105	 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 93
WITH WHOM: Network Analysis page 60	 World Finance Corporation network page 87	 Electronic and new media art networks page 133	 World-wide scholarly collaboration networks page 157

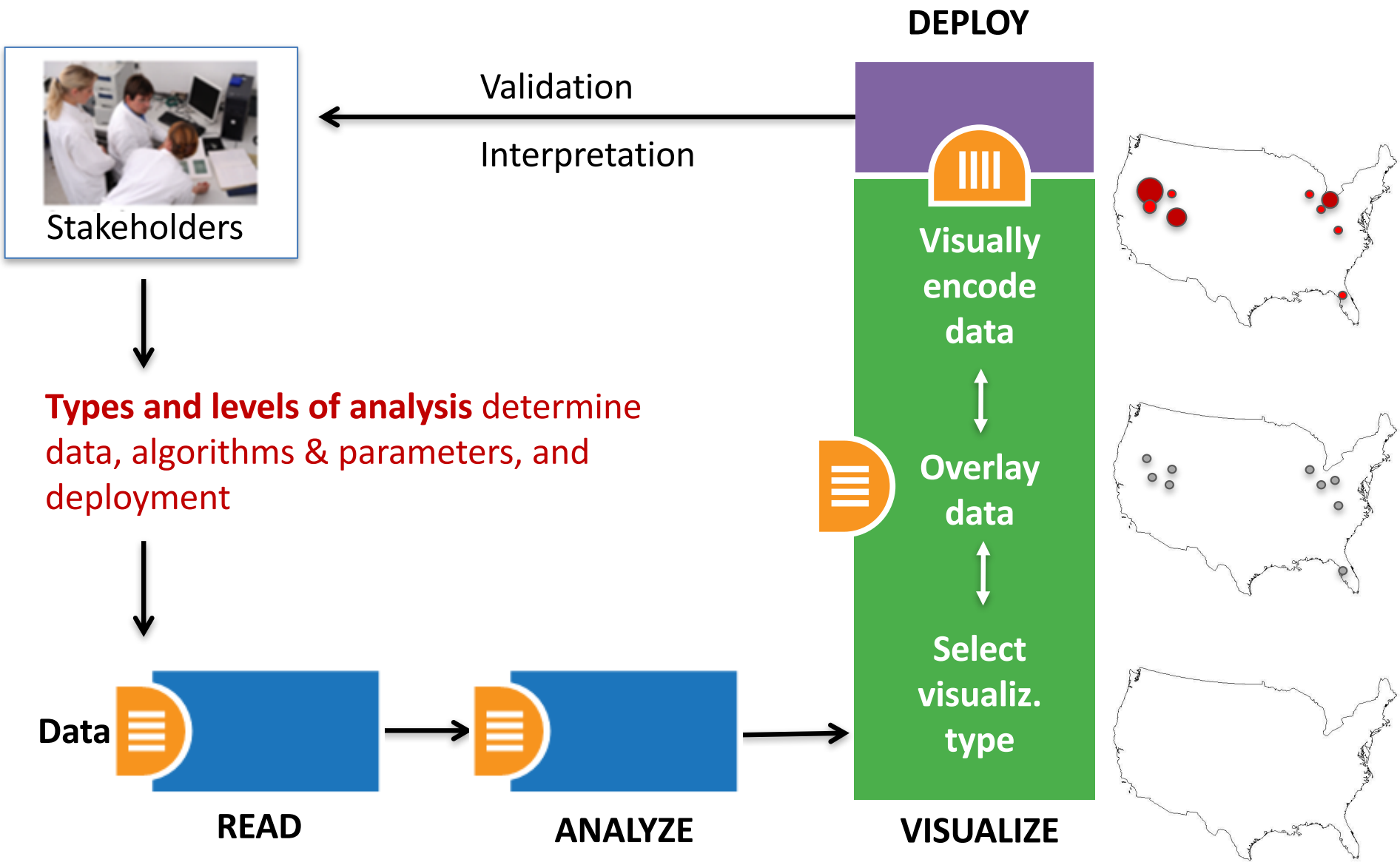


See page 5

Needs-Driven Workflow Design



Needs-Driven Workflow Design



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

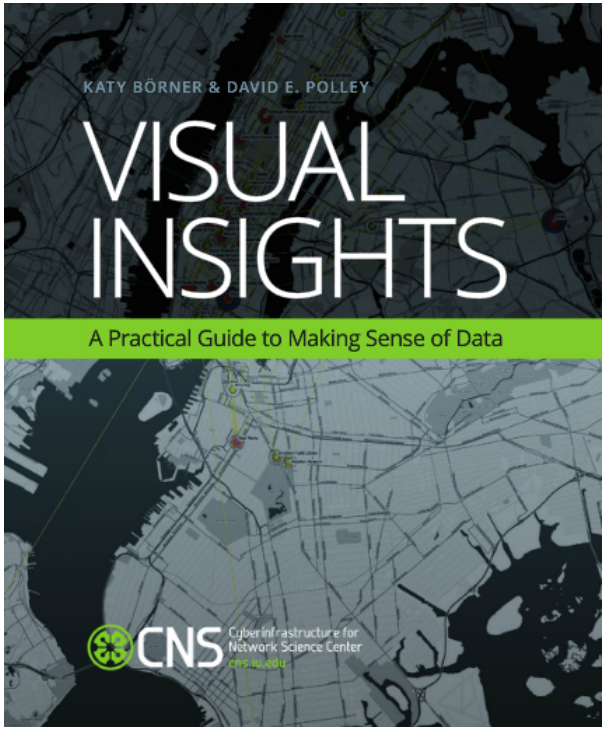
- **Session 5** – “With Whom:” Trees
- **Session 6** – “With Whom:” Networks
- **Session 7** – Dynamic Visualizations and Deployment

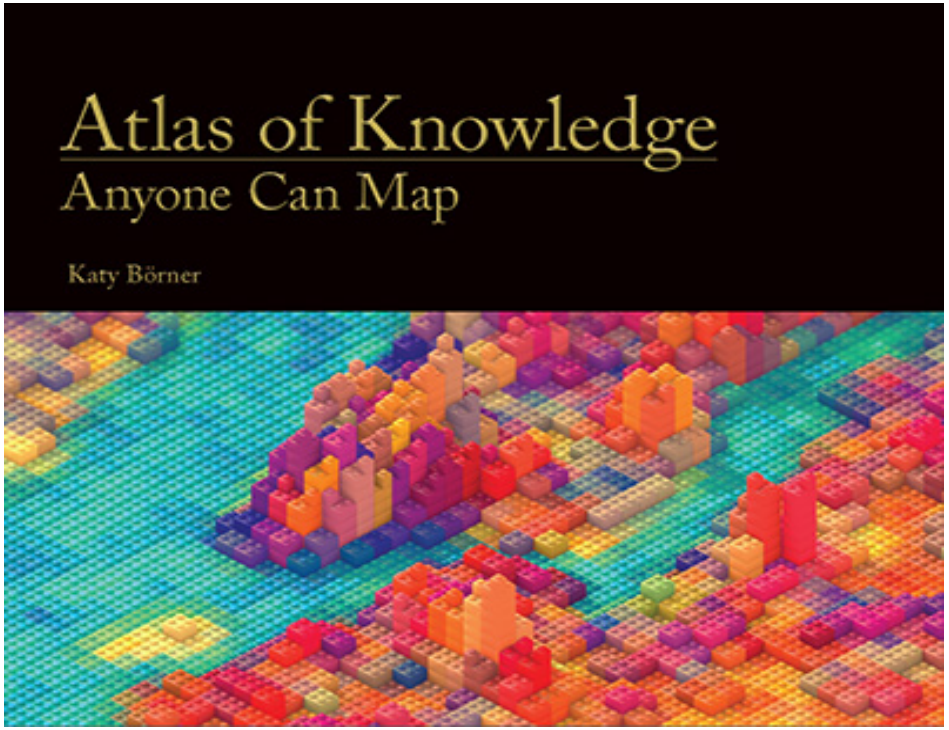
Final Exam

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%).

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.

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

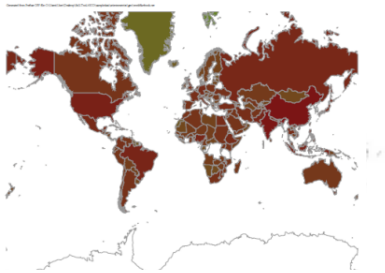
Statistical Analysis–p. 44

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

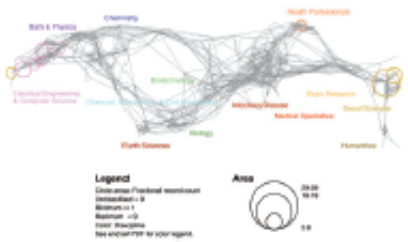
Geospatial Analysis–p. 52



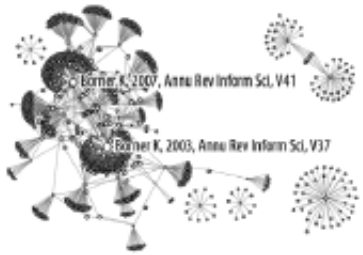
Load **One** File and Run **Many** Analyses and Visualizations

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

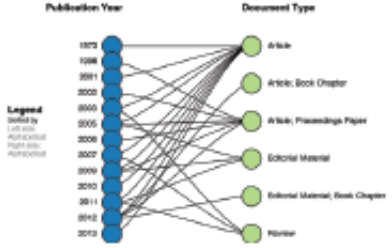
Topical Analysis–p. 56



Paper Citation Network–p. 60



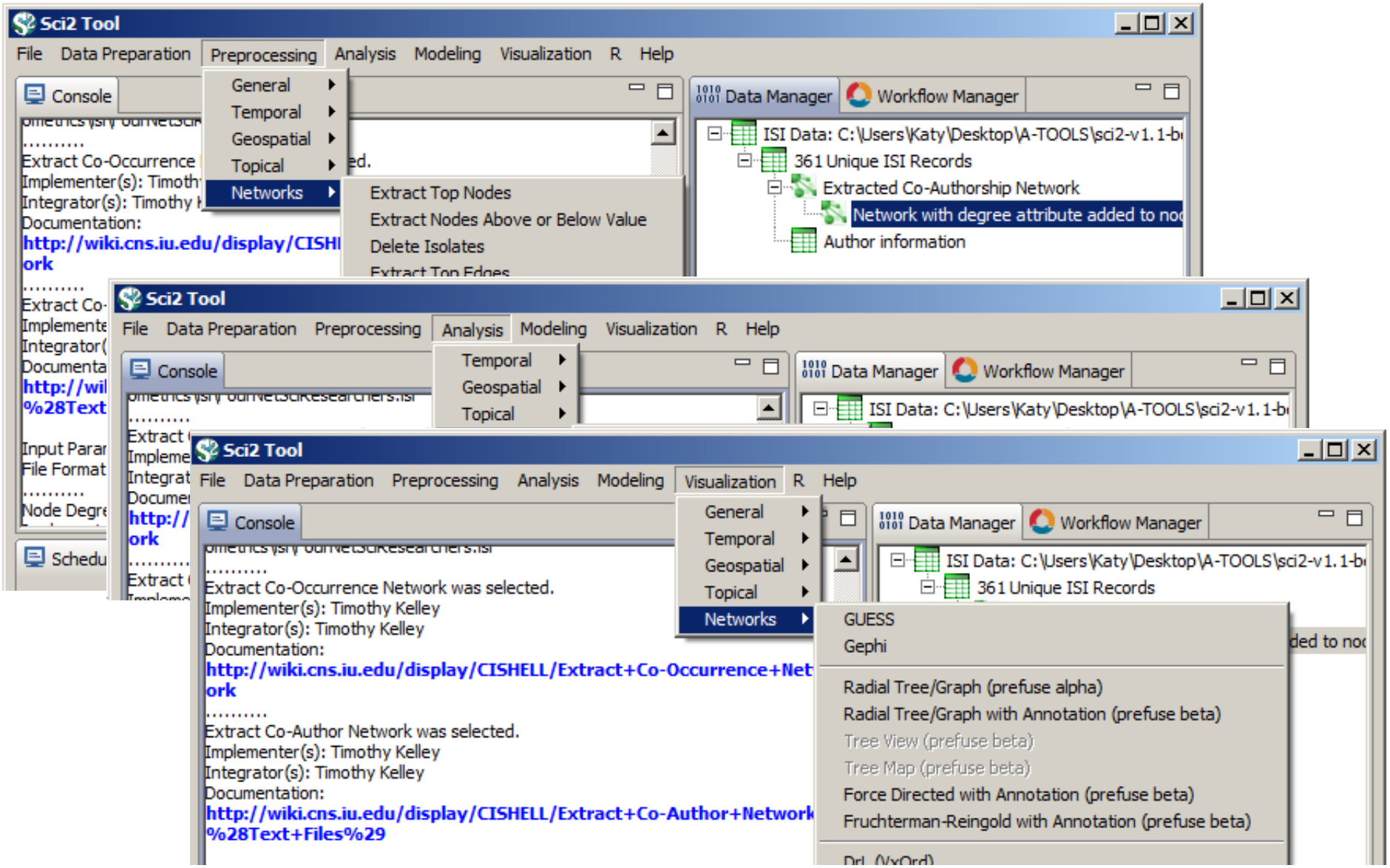
Bi-Modal Network–p. 60



Co-author and many other bi-modal networks.

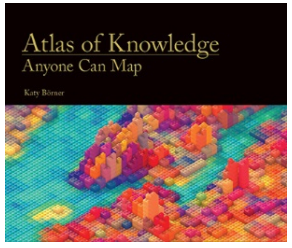
Sci2 Tool Interface Components

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



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
• 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	• table • chart • graph • map • network layout	• 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	• overview • zoom • search and locate • filter • details-on-demand • history • extract • link and brush • projection • distortion



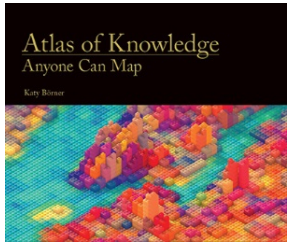
See page 24

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

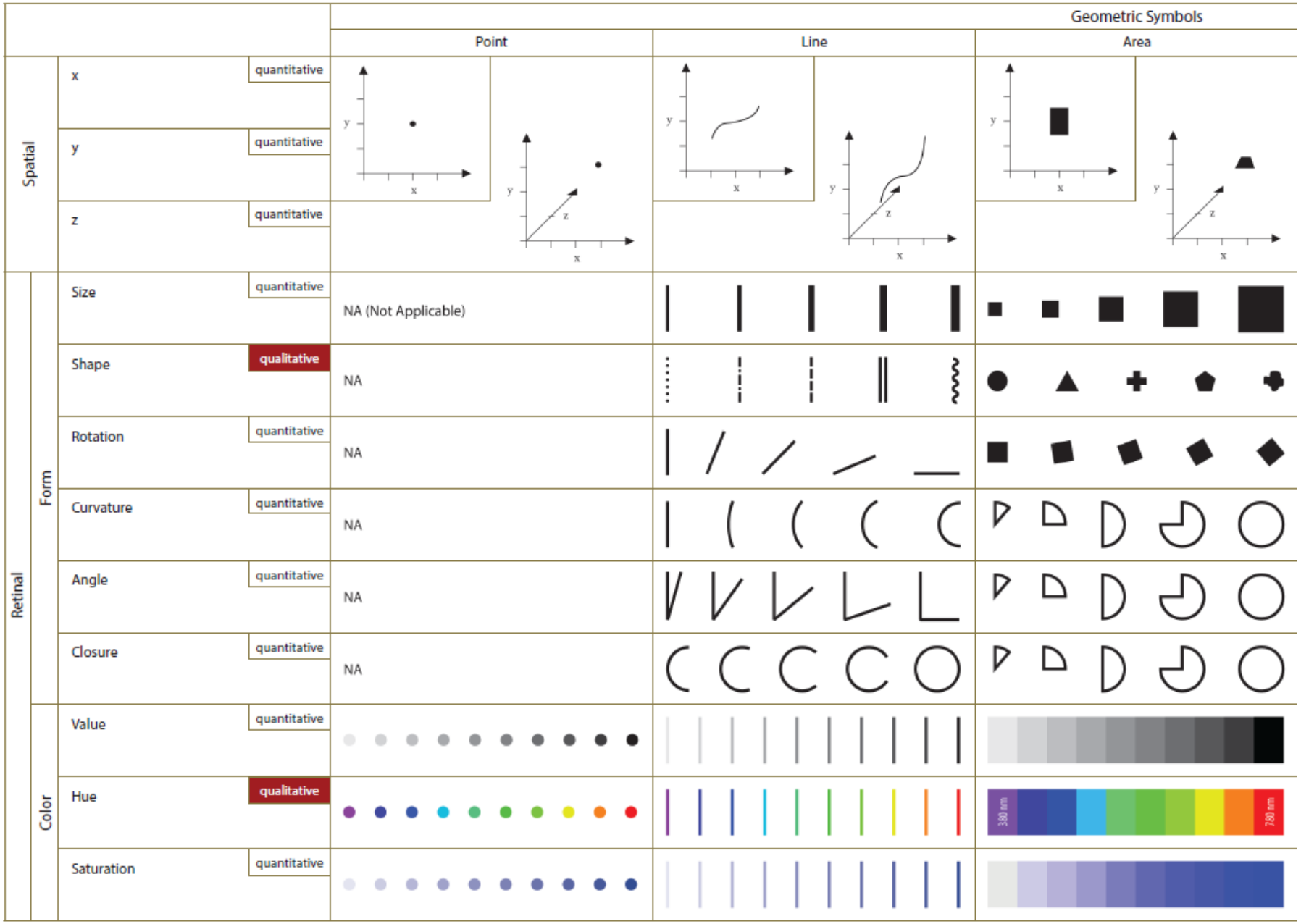
Visualization Framework

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
• 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	• table • chart • graph • map • network layout	• 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	• overview • zoom • search and locate • filter • details-on-demand • history • extract • link and brush • projection • distortion

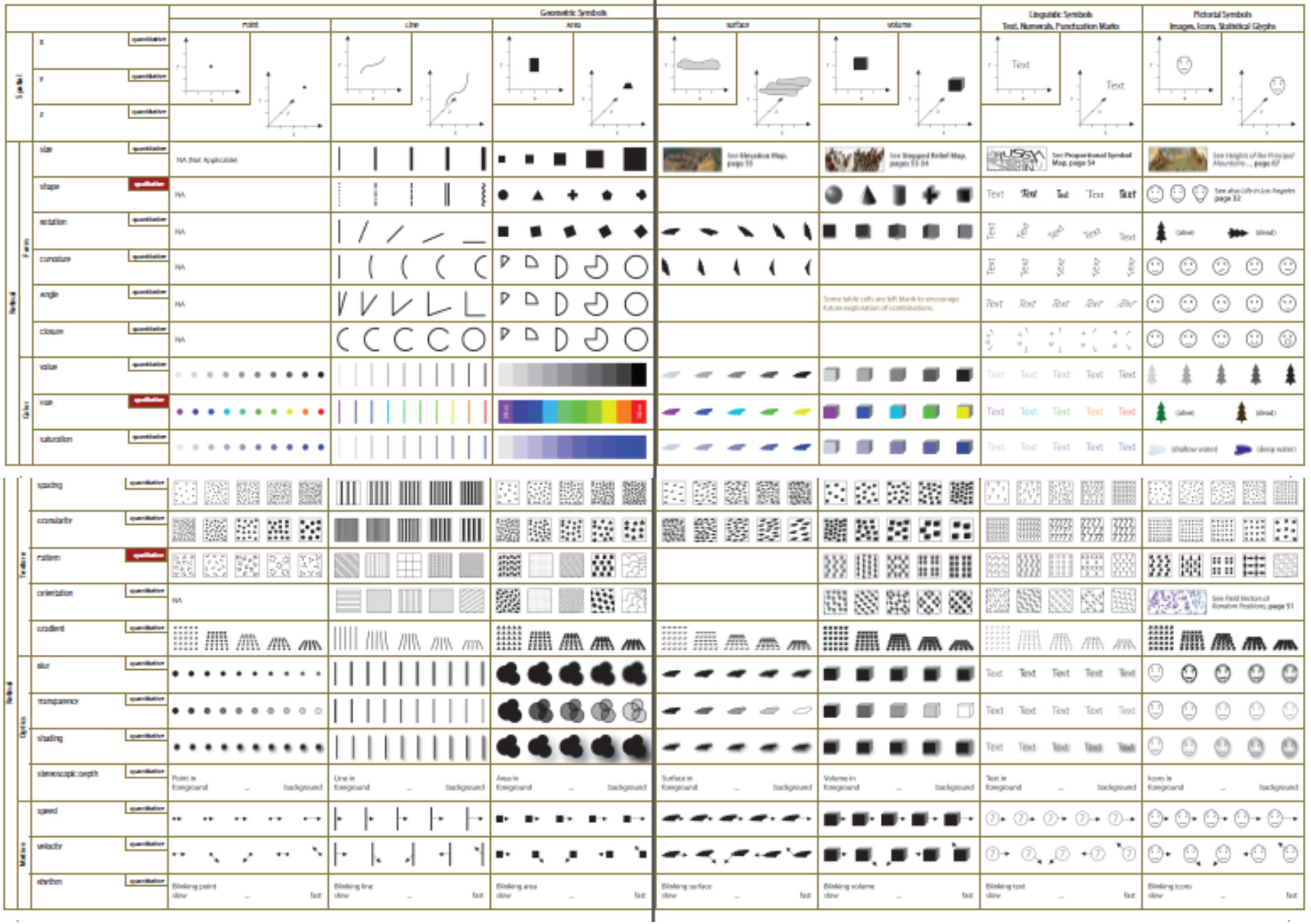


See page 24

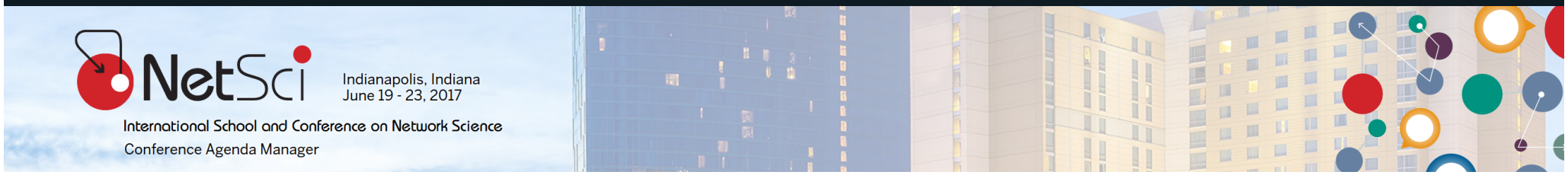
Graphic Variable Types Versus Graphic Symbol Types



Graphic Variable Types Versus Graphic Symbol Types



Data Visualization Literacy: Research



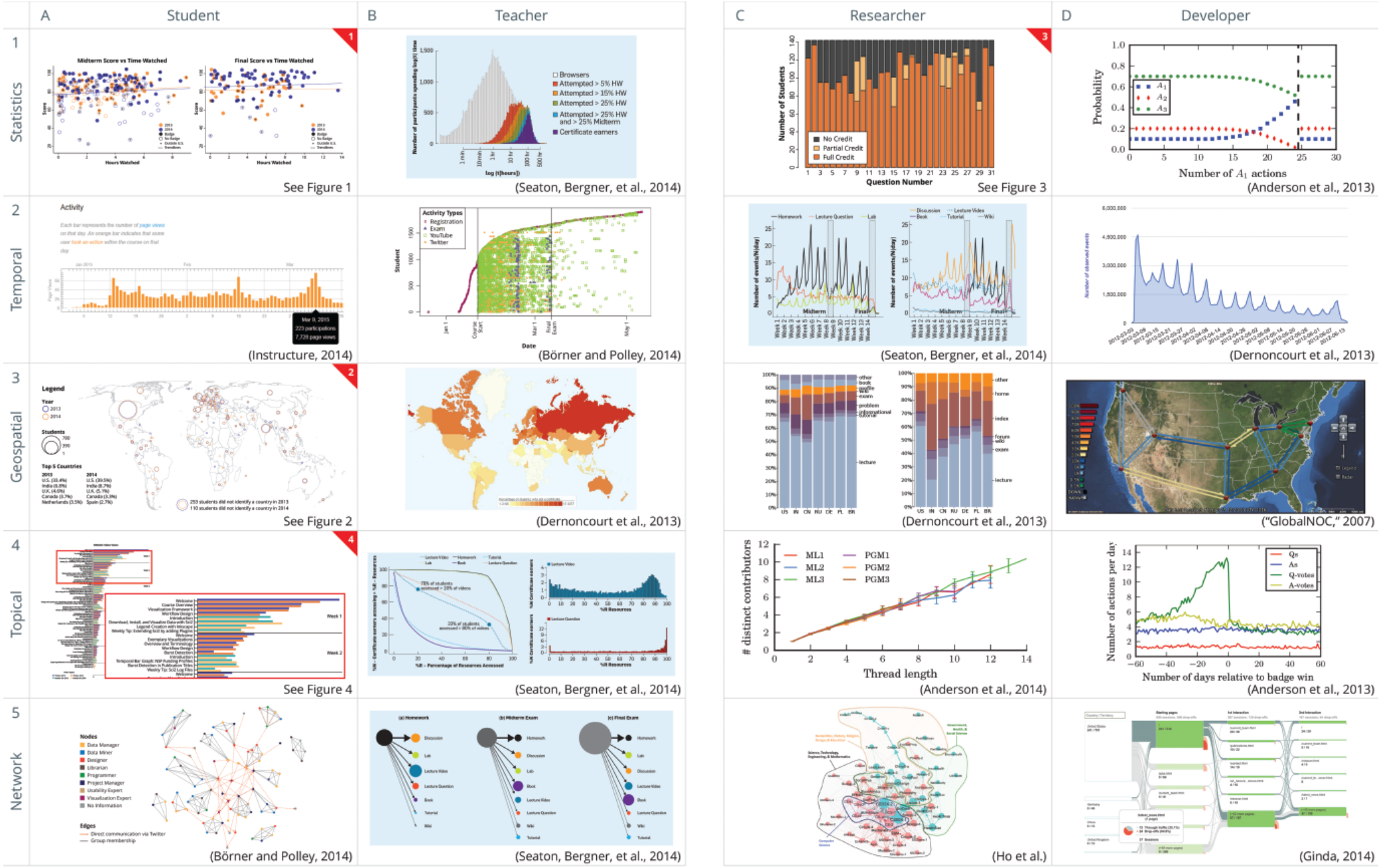
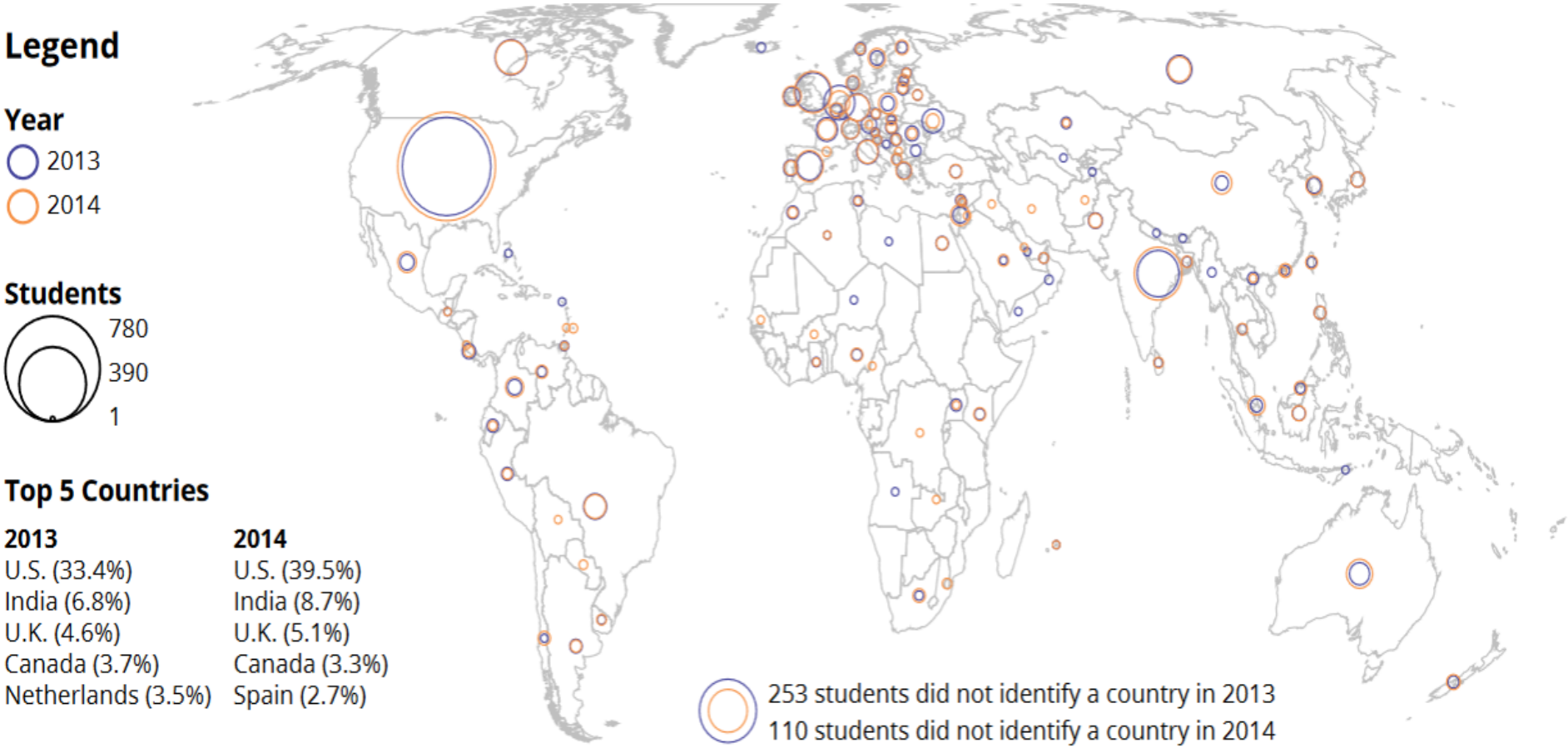


Figure 1: Analysis types vs. user needs, taken from

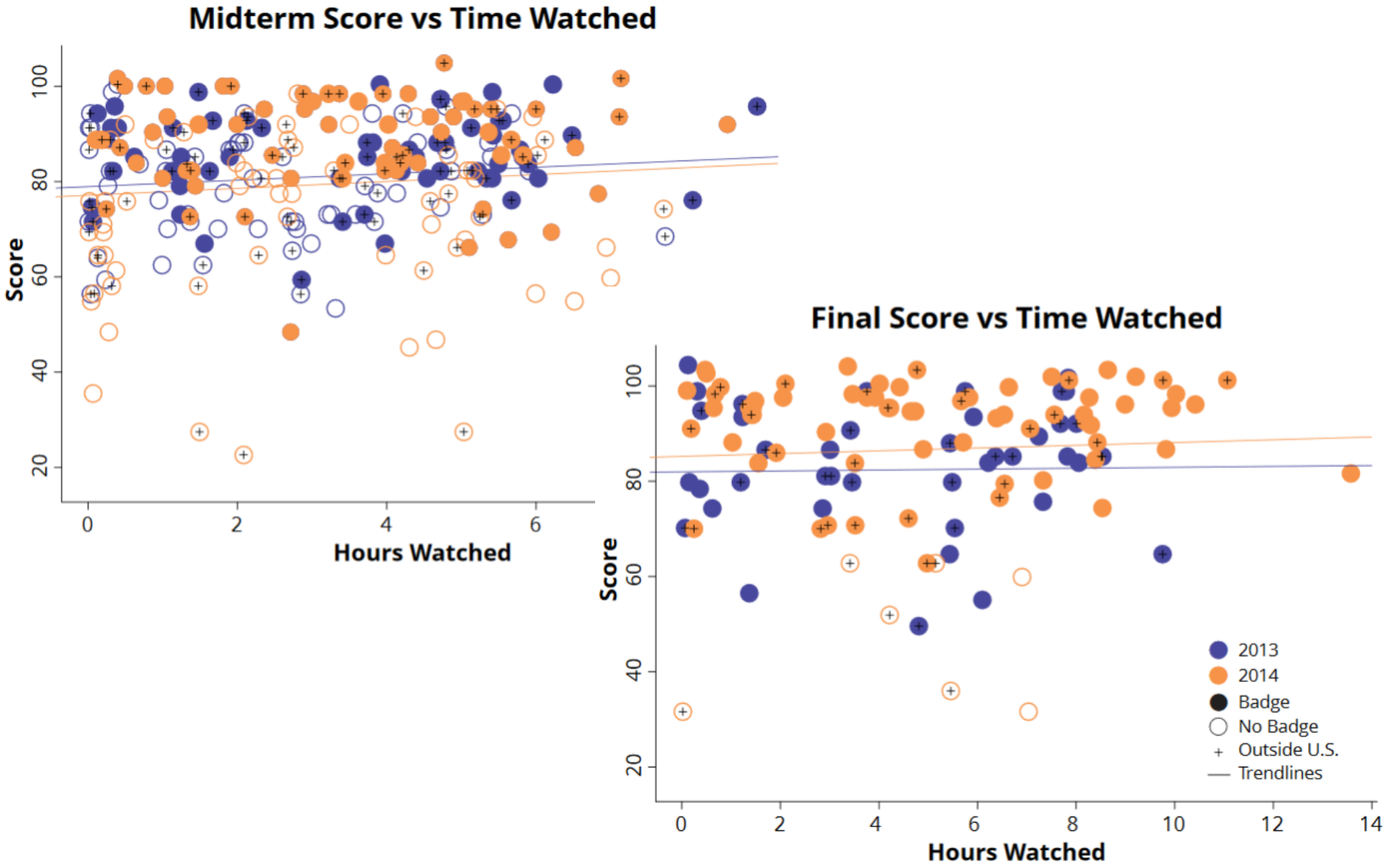
Emmons, Light, and Börner. "MOOC Visual Analytics: Empowering Teachers, Students, Researchers, and Developers of Massively Open Online Courses". *Journal of the Association for Information Science and Technology (in press)*.

Students' Countries



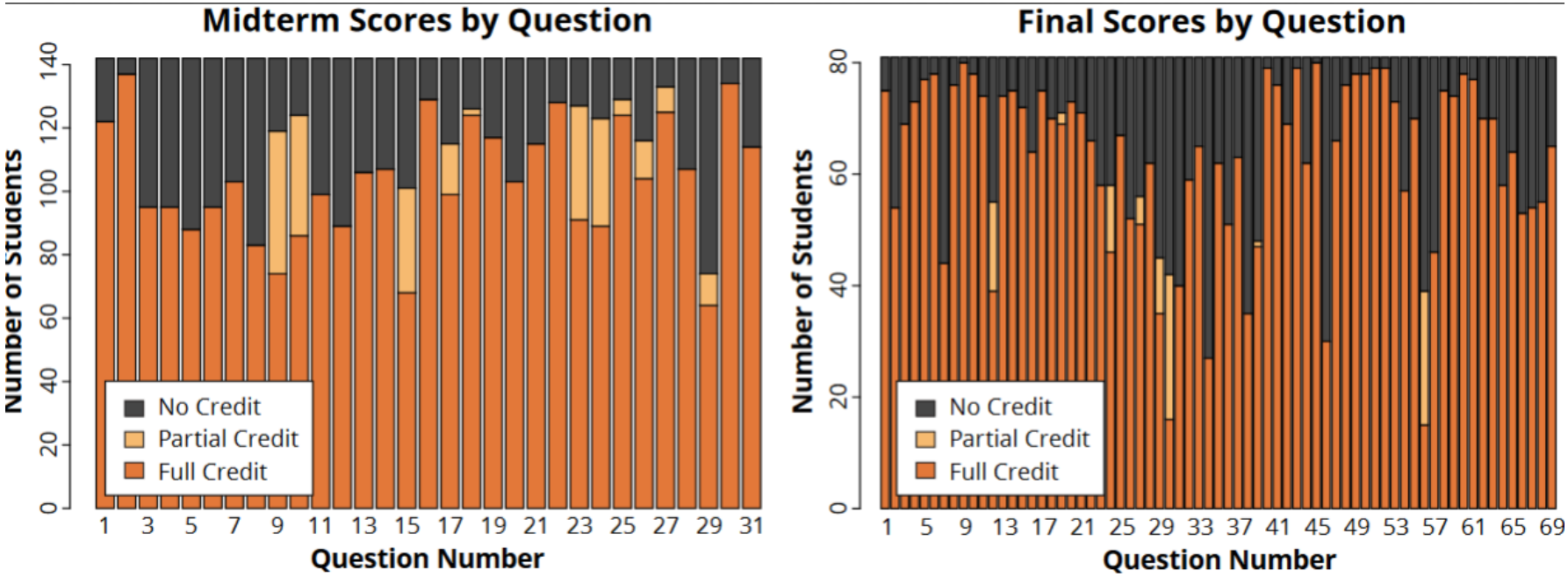
Proportional symbol map of the world showing the location of IVMOOC students from 2013 (blue) and 2014 (orange). Circles are area size coded by the number of students per country.

Student Final Score vs. Hours Watched

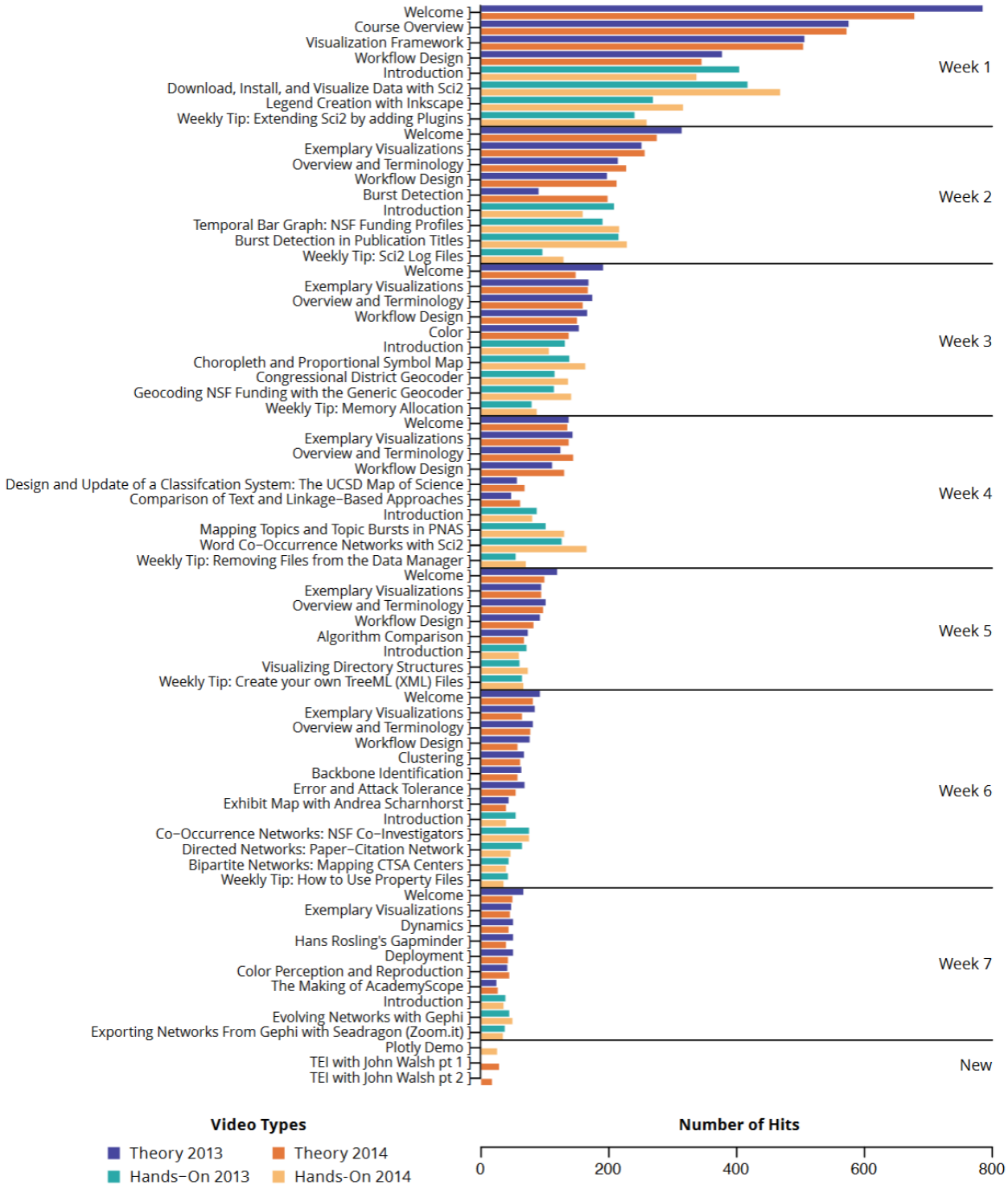


Scores vs. time invested watching course videos for students who took the 2013 (blue) and 2014 (orange) IVMOOC midterm (left) and final exam (right) and got at least 50% correct.

Exam Scores by Question



Student scores per question for midterm (left) and final exam (right) for IVMOOC 2014.



IVMOOC Video Views

IVMOOC video views in 2013 (blue) and 2014 (orange)

IVMOOC

Student Client Projects: All Interactions

Nodes

Previous experience with information visualization

- No response
- Very Low
- Low
- Medium
- High
- Instructors

Edges

- Direct communication via Twitter
- Group membership

Student Engagement and Performance

Learning Analytics

IVMOOC 2015 Student Group Engagement and Scores

	Pre-Course	Week 1	Week 2	Week 3	Week 4	Midterm	Week 5	Week 6	Week 7	Week 8	Week 9	Final	Curr. Score
IVMOOC	26.05%	38.32%	31.32%	29.96%	27.1%	28.34%	31.07%	24.28%	16.86%	18.23%	13.08%	13.41%	20.87%
Z637-29374	33.01%	52.91%	49.89%	59.22%	50.89%	82.56%	65.04%	49.99%	39.59%	61.63%	54.91%	82.25%	82.4%
Z637-32593	25.08%	54.54%	43.58%	50.67%	53.63%	77.67%	65.7%	59.48%	52.19%	65.71%	47.27%	72.59%	75.13%
Z637-33781	29.33%	55.38%	49.26%	62.18%	77.47%	85%	87.4%	69.8%	55.56%	57.6%	45.69%	70.89%	77.94%

IVMOOC 2015 Student Group Engagement for Midterm

	Midterm	Final	Curr. Score	Overall Engagemer
Student 198	100%	85.33%	92.67%	30.34%
Student 210	100%	84%	92%	33.91%
Student 242	97.14%	98.67%	97.9%	55.89%
Student 265	95.71%	92%	93.86%	82.64%
Student 216	95.71%	24%	59.86%	34.92%
Student 257	94.29%	98.67%	96.48%	68.25%
Student 264	94.29%	89.33%	91.81%	80.47%
Student 262	94.29%	85.33%	89.81%	79.65%

Legends

Engagement

Inactive Active Very Active

0% 25% 50% 75% 100%

Score

F D C B A

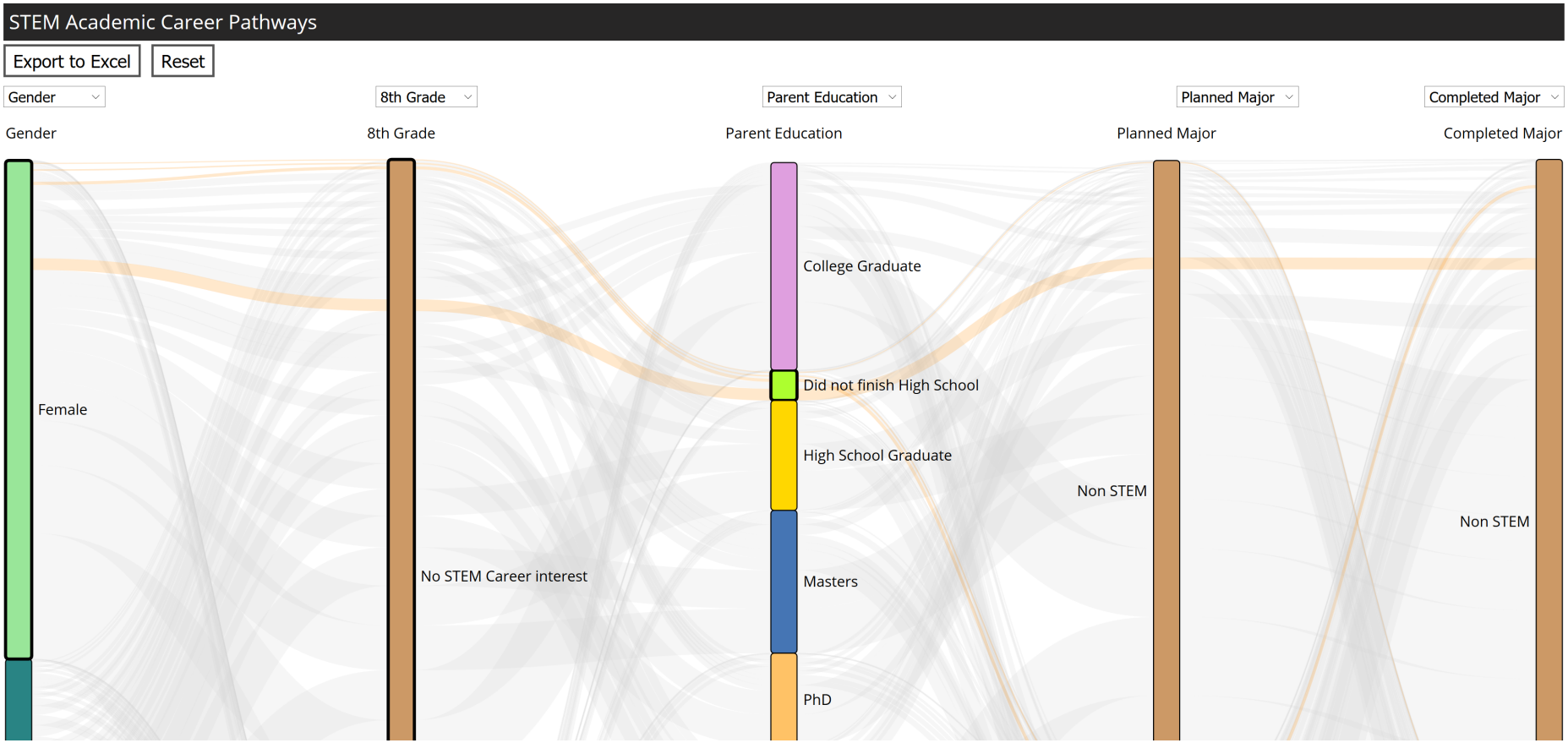
0% 60% 70% 80% 100%

Description

The heat map visualization is a representation of student engagement (magenta to blue color scale) and performance (red to green color scale) throughout a course. The visualization has two levels. The top level provides an overview of engagement and performance for groups of students, while the bottom level provides a detailed break out of student engagement statistics for individuals with an identified group.

Custom interactive visualizations of 2015 IVMOOC student engagement and performance data, explore functionality online at http://goo.gl/TYixCn

Student Flows – STEM Academic Career Pathways

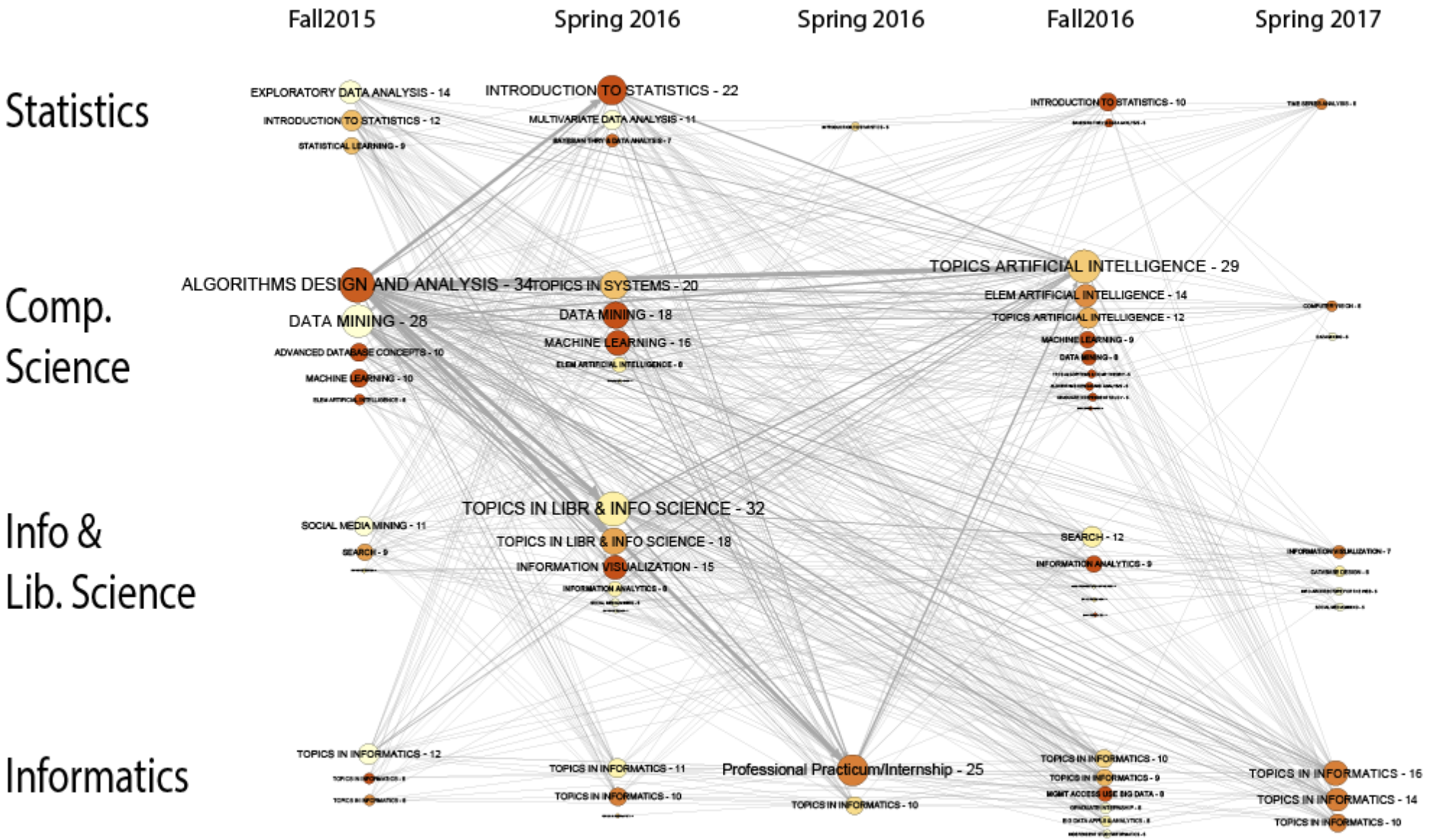


Measuring and Visualizing STEM Pathways. NSF NCSE-1538763 Award (Adam Maltese, Katy Börner) Aug. 15, 2015 - Jan. 2017.

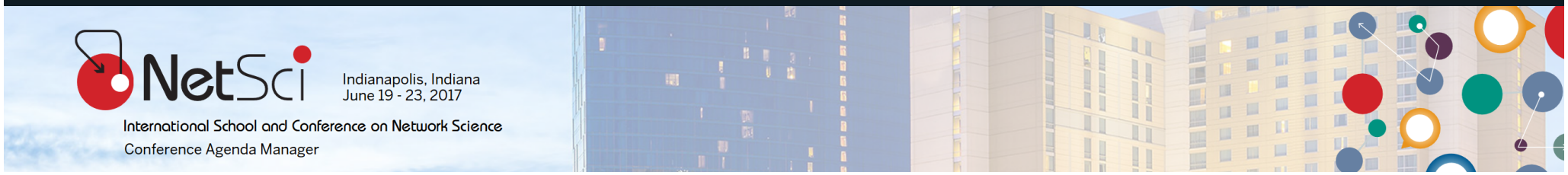
Interactive web site: http://demo.cns.iu.edu/client/stem

CNS Cyberinfrastructure for Network Science Center

Data Science Course Transition Matrix for Fall 2015 Cohort



Data Visualization Literacy: Outlook



IEEE
Advancing Technology for Humanity

PROGRAM SPEAKERS POSTER CONTEST VENUE REGISTER CONTACT

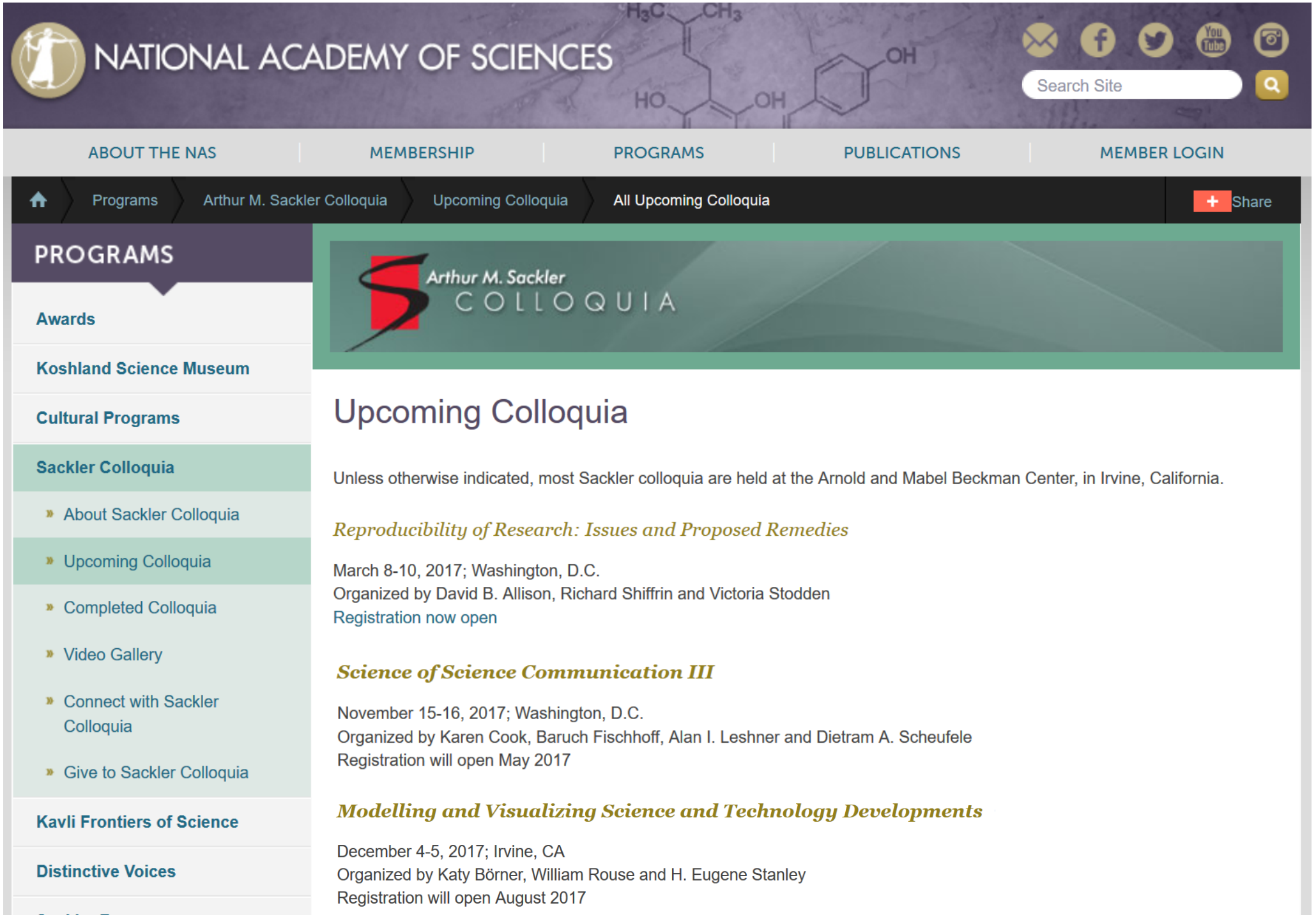
Friday, November 10

	Cyberinfrastructure Building	Innovation Center
7:45 AM	Shuttles leave hotel for campus	
8:00 AM	Continental Breakfast and Registration	
9:00 AM	Welcome	
9:30 AM	Keynote	
10:30 AM	Coffee Break	
11:00 AM	**S1**: Cybersecurity (3 speakers)	**S6**: Funders and Funding

IEEE EnCON Conference will take place at CIB, IUB on Nov 10-11, 2017.

One sessions is devoted to “Engineering Education.”

CSWS Session: Visualizing STEAM Data in Support of Smart Decision Making
November 15-17, 2017, Tokyo, Japan.



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

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

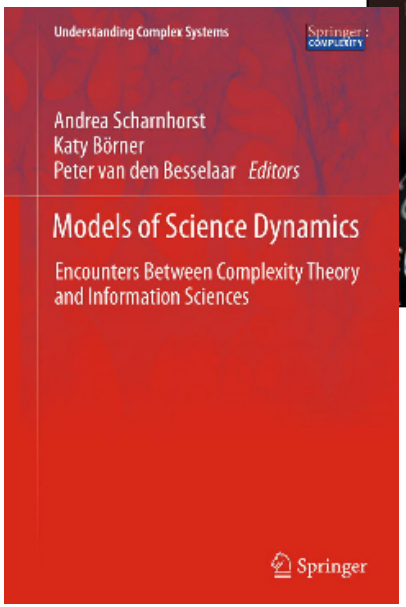
Börner, Katy (2010) **Atlas of Science: Visualizing What We Know**. The MIT Press. http://scimaps.org/atlas

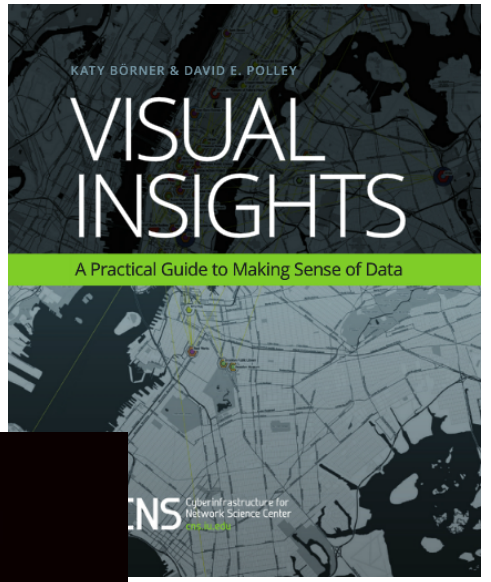
Scharnhorst, Andrea, Börner, Katy, van den Besselaar, Peter (2012) **Models of Science Dynamics**. Springer Verlag.

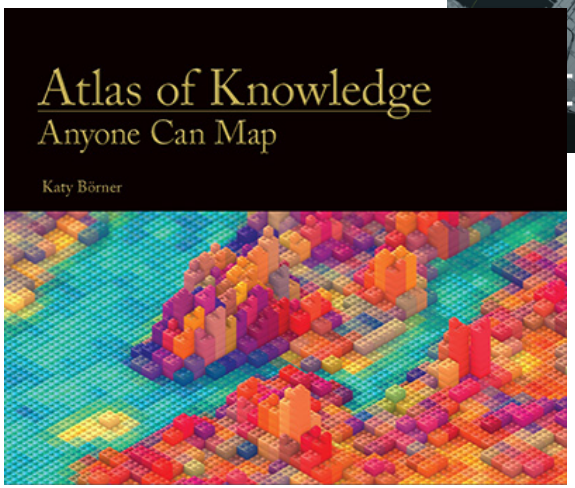
Katy Börner, Michael Conlon, Jon Corson-Rikert, Cornell, Ying Ding (2012) **VIVO: A Semantic Approach to Scholarly Networking and Discovery**. Morgan & Claypool.

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