Visual Insights

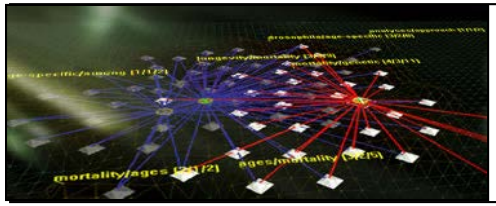
Katy Börner @katycns

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

E599 Lecture

October 15, 2017

Overview

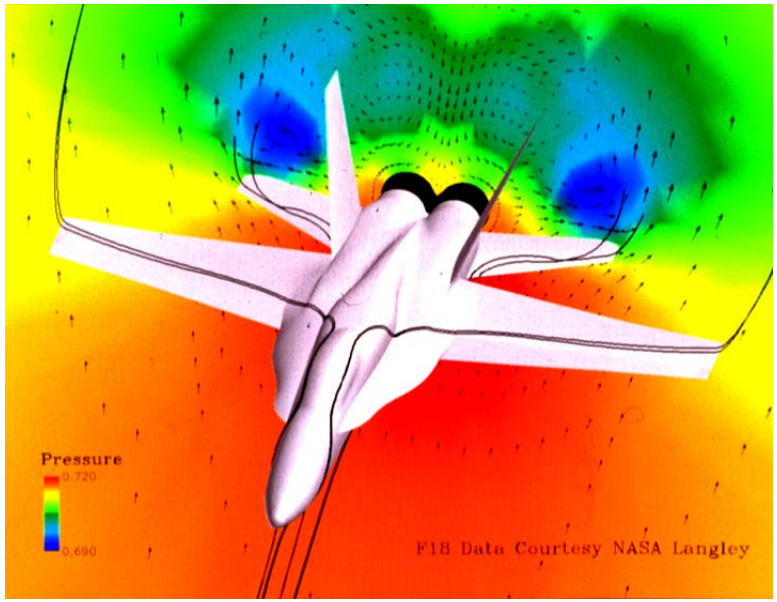


Information Visualization - Definition

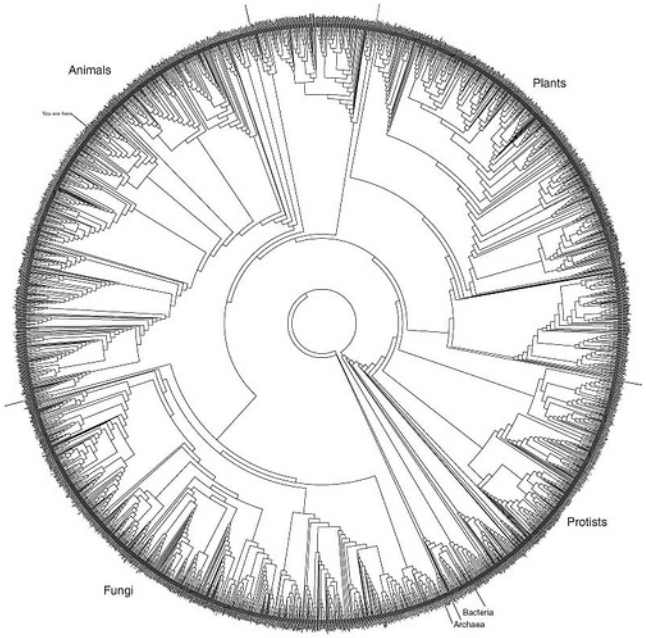
"Information Visualization is a process of transforming data and information that are not inherently spatial, into a visual form allowing the user to observe and understand the information."

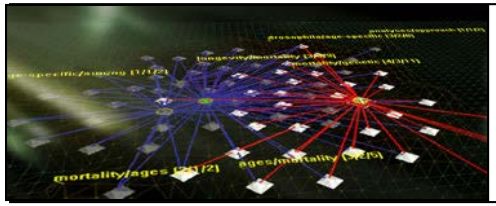
(Source: Gershon and Eick, First Symposium on Information Visualization)

Scientific Visualization



Information Visualization





Information Visualization – Human Advantage

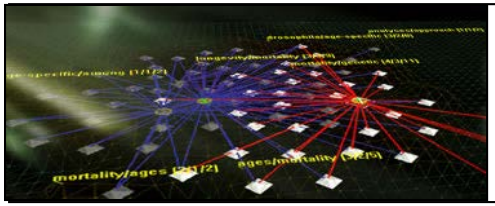
- Rooted in geography, scientific visualization.
- Not even 30 years old.

- Growing fast.
- Far reaching (IR, WWW, DL, HCI).
- Interdisciplinary nature: computer graphics, electronic engineering, information systems, geography, information science, …

- Tremendous potential.

  Humans can detect a single dark pixel in a 500 x 500 array of white pixels in less than a second. This screen can be replaced every second by another, enabling a search of 15 million pixels in a minute (Ware, 2000).

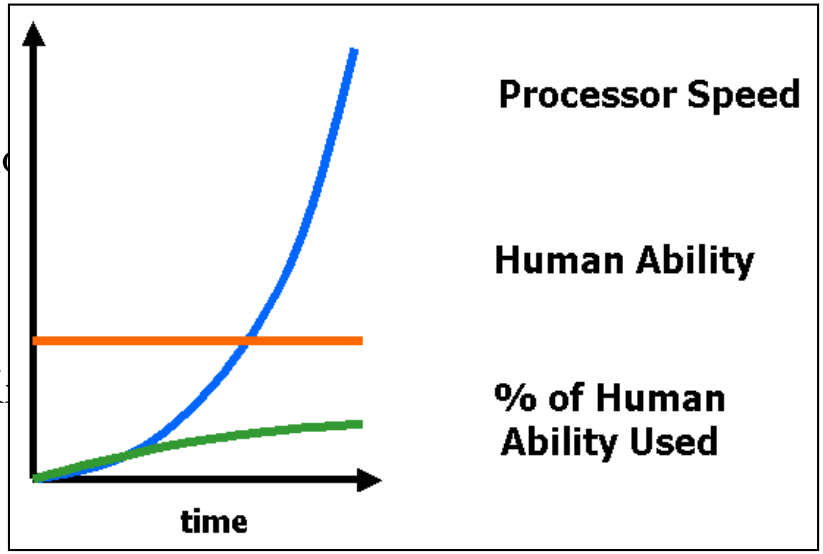
  Also, people have a truly remarkable ability to recall pictorial images. In one study, Standing, Conezio, & Haber (1970) showed S's 2560 pictures, each for 10 seconds over 7 hours, in a 4-day period. Afterwards, S's were asked to classify pictures presented at a rate of 16 pictures/minute and they achieved better than 90% accuracy.



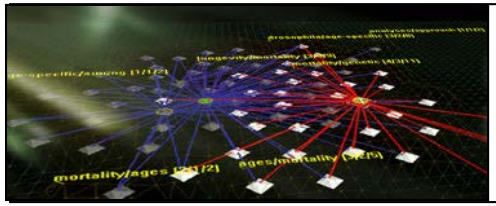
Information Visualization – Why Now?

- Information explosion (amount doubles every 18 months).
- Work is becoming more 'knowledge-oriented'.

- Increasing computing power (doubles every 18 months - Moore's Law).
- Decreasing cost of storage.
- Fast graphics processors.
- Larger hard disk sizes -> more informati
- High resolution color monitors.
- Alternative user interfaces Idesk, CAVE
- Connectivity between systems is expandi



- Increasing visual intelligence.
- There is a bad mismatch between computer displays and the human perceptual system and between computer controls and human motor functions.

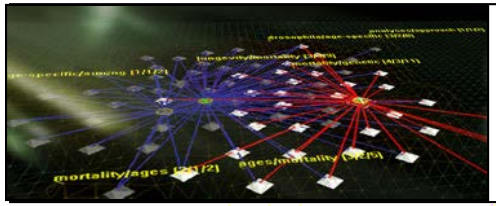


Information Visualization – Potential

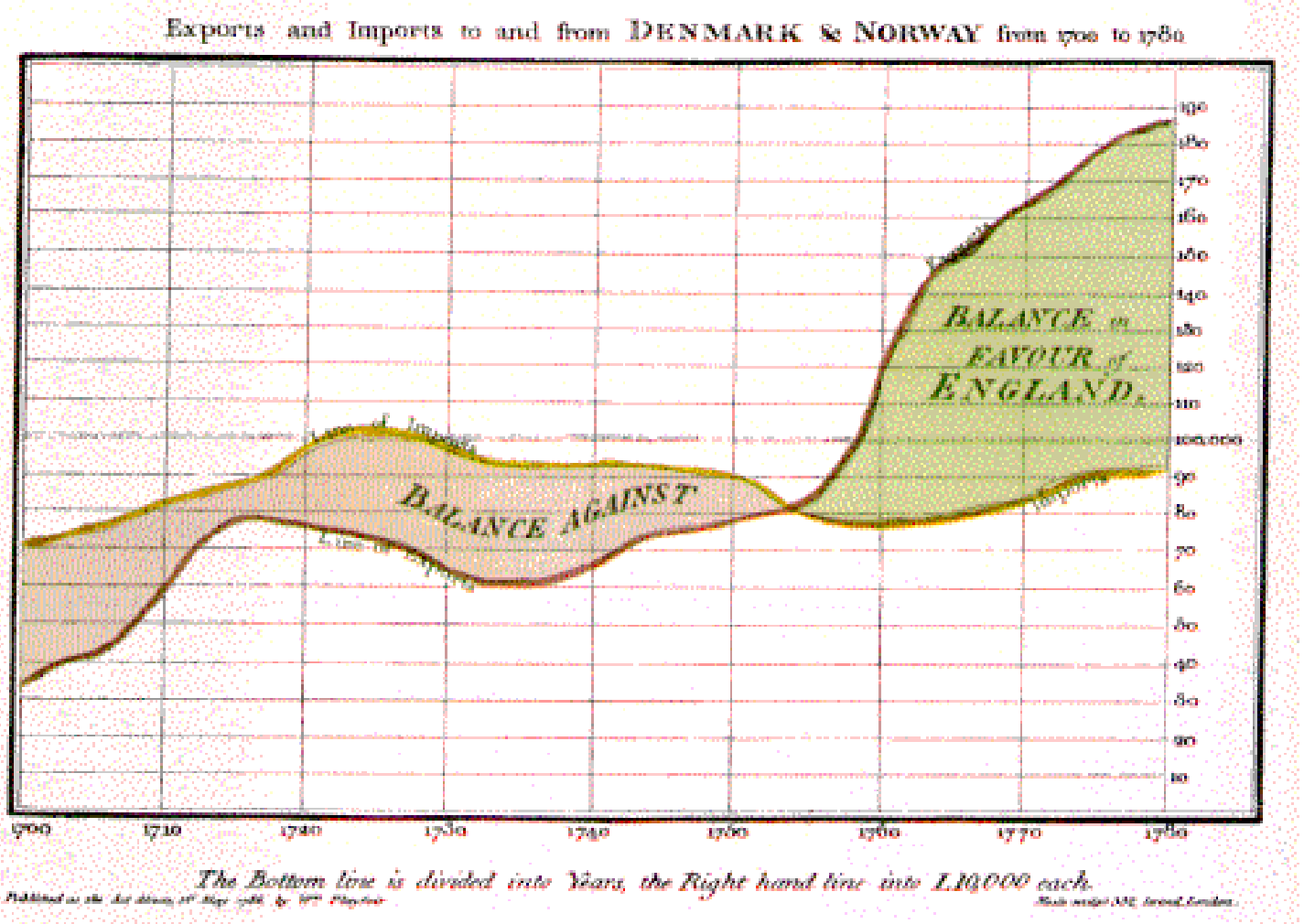
Well designed visualizations …

- Provide an ability to comprehend huge amounts of data.
- Reduce search time and reveal relations otherwise not being noticed (perception of emergent properties).
- Often reveal things not only about the data but how the data was collected - errors and artifacts jump out.
- Facilitate hypothesis formulation.
- Are effective sources of communication.

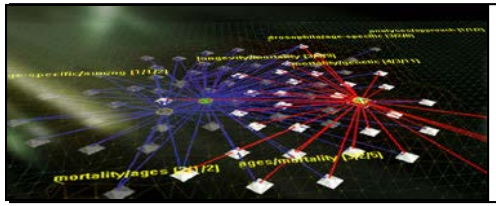
Let's Look at the History of Visualizations



Balance of Trade Chart by Playfair in 1786



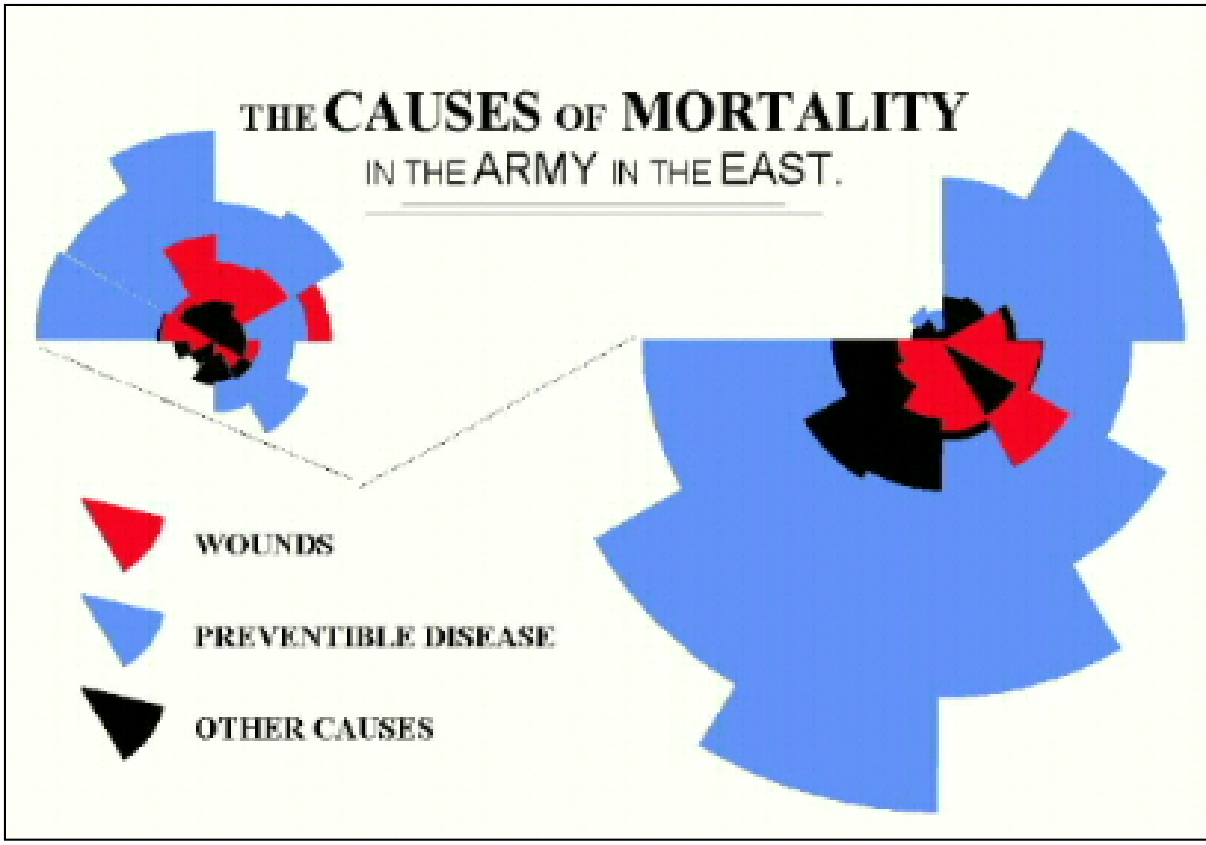
The area between two time-series curves was emphasized to show the difference between them, representing the balance of trade.

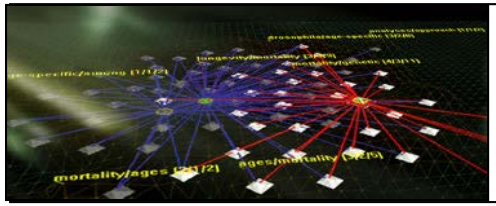


Coxcombs by Florence Nightingale in 1858

This figure (reproduced with SAS/GRAPH) shows that far more deaths were attributable to non-battle causes ("preventable causes") than to battle-related causes.

Nightingale's Coxcomb is notable for its display of frequency by area, like the pie chart. But, unlike the pie chart, the Coxcomb keeps angles constant and varies radius (proportional to sqrt(frequency)), a principle used in the FourFold Display for 2x2xk tables.



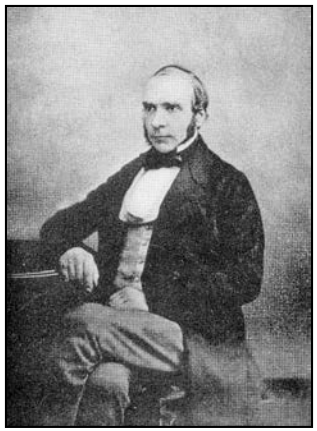


Map of Cholera by John Snow in 1855

John Snow use of statistical mapping to deduce the link between cholera and contaminated water.

Based on the map he observed that cholera occurred almost entirely among those who lived near (and drank from) the Broad Street water pump.

Removing the handle of the contaminated pump ended the neighborhood epidemic which had taken more than 500 lives.



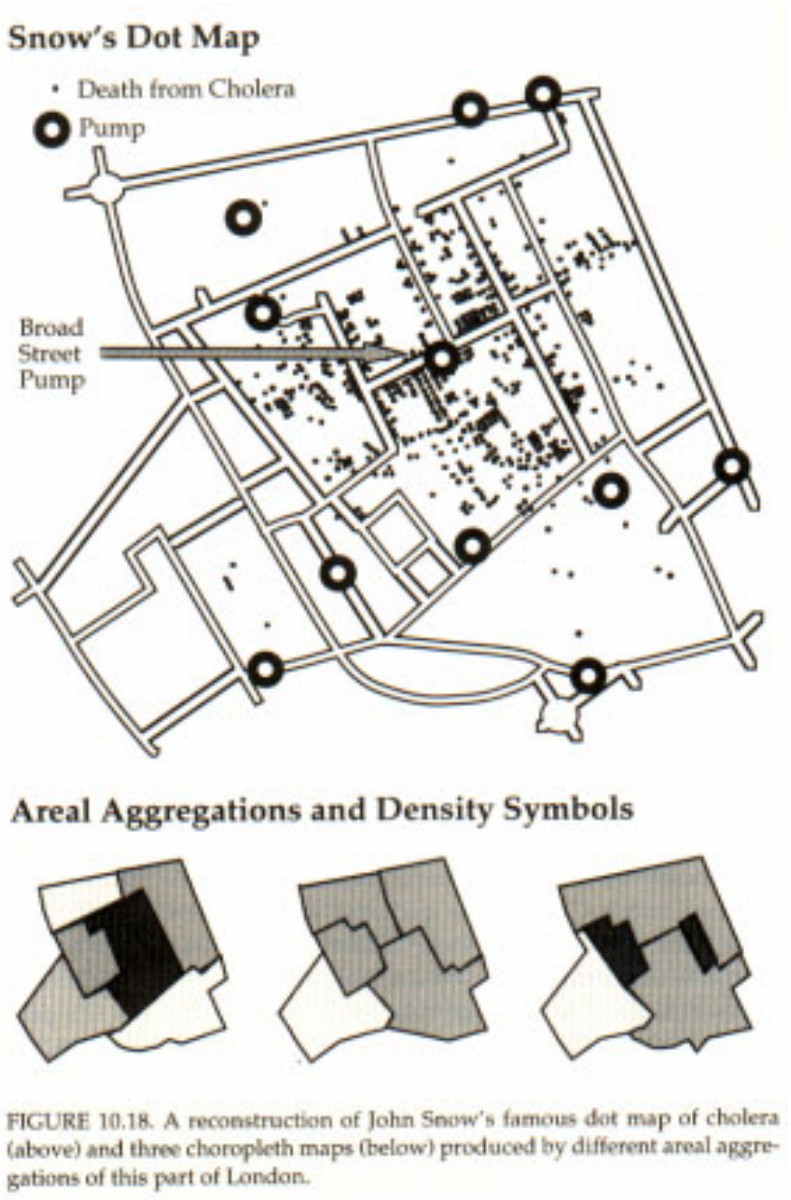
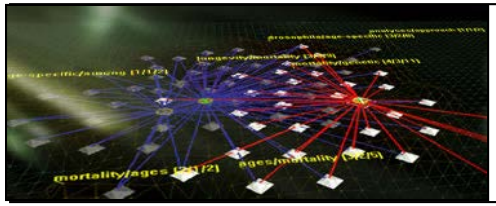
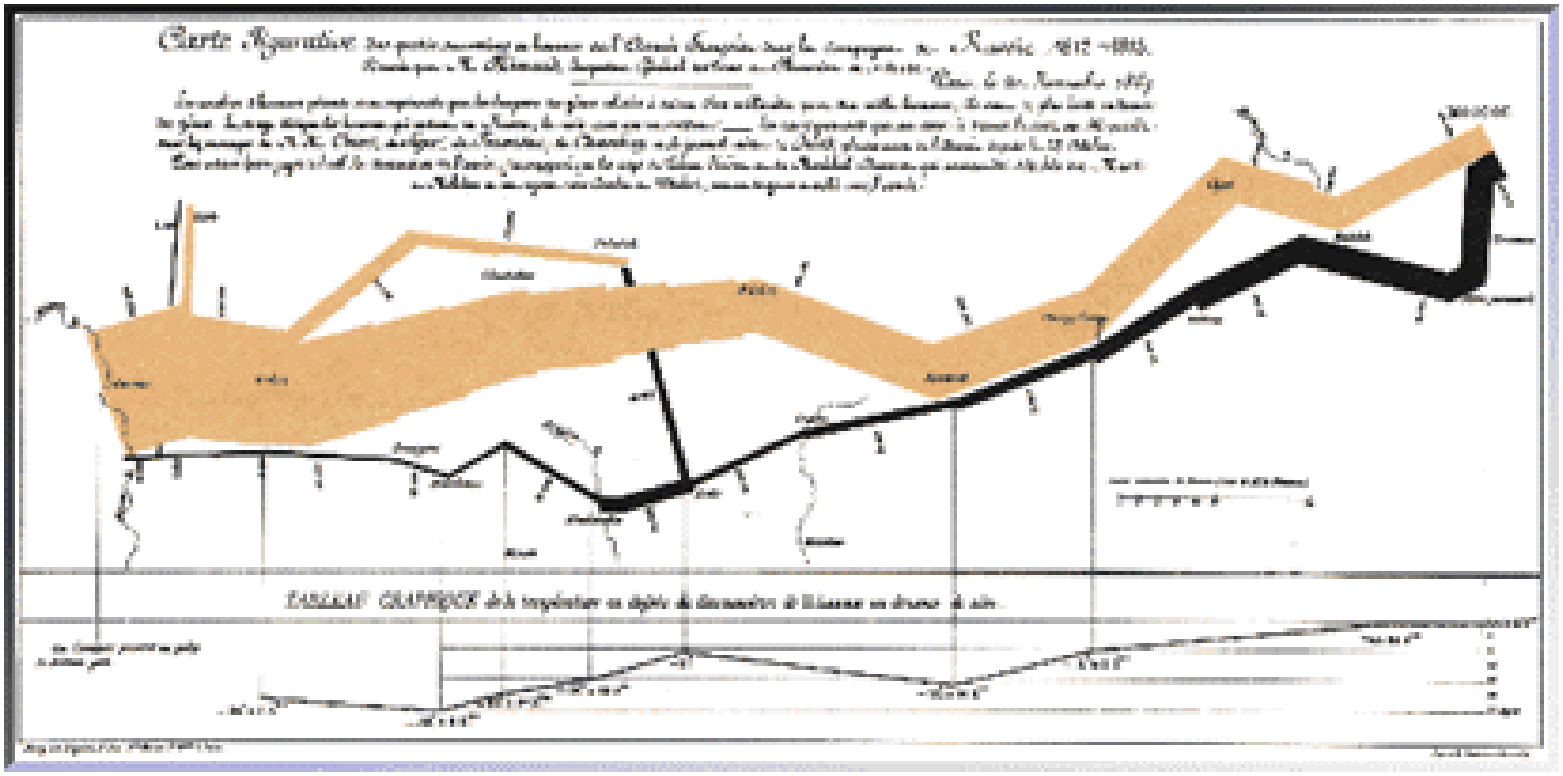


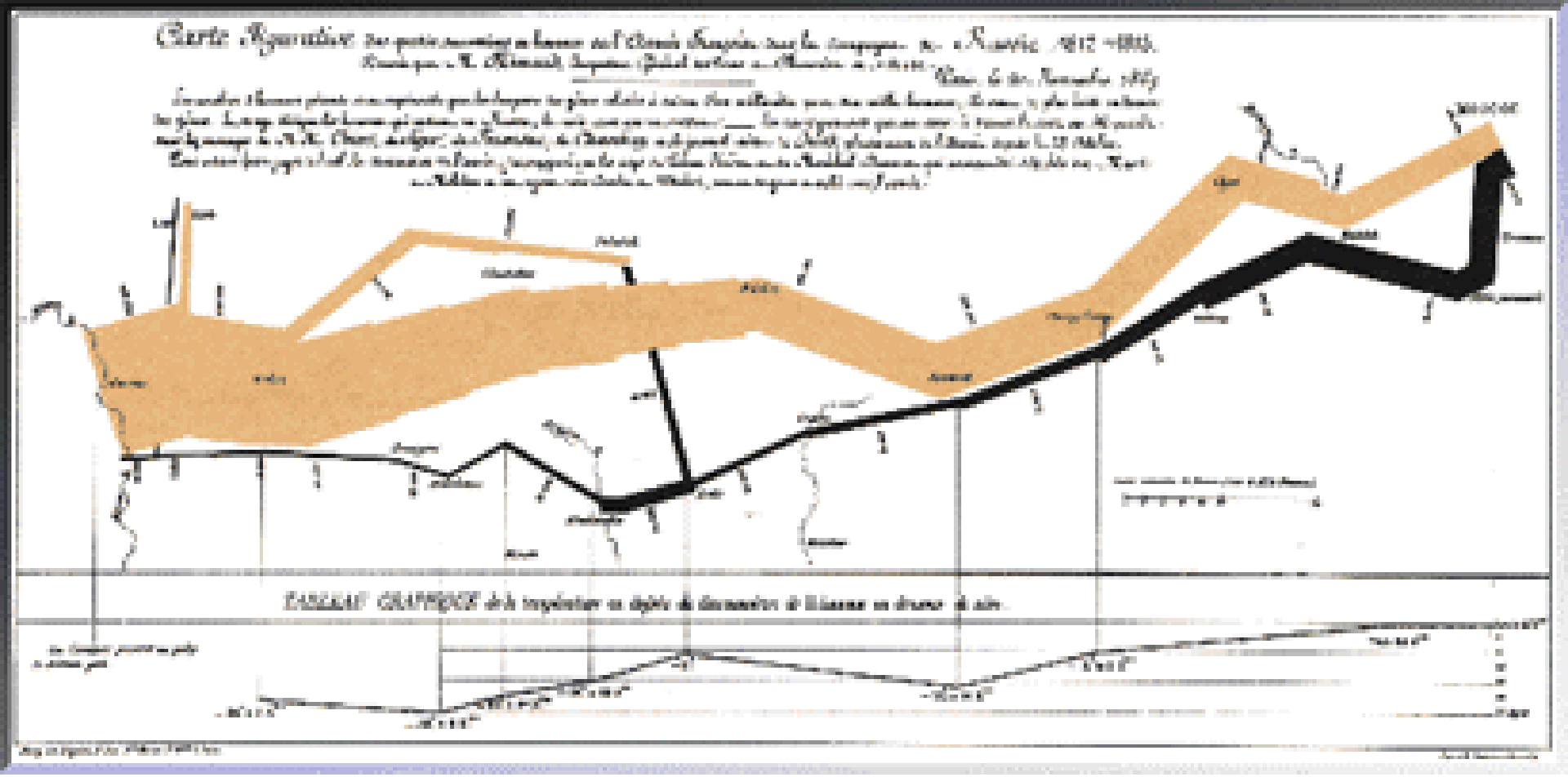
FIGURE 10.18. A reconstruction of John Snow's famous dot map of cholera (above) and three choropleth maps (below) produced by different areal aggregations of this part of London.



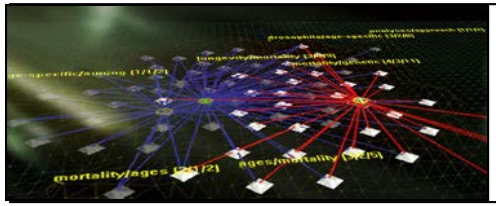
The Loss of Napoleons Army by Charles Minard in 1869

Minard (1781-1870), a French engineer, illustrated the disastrous result of Napoleon's failed Russian campaign of 1812. The graph shows the size of the army by the width of the band across the map of the campaign on its outward and return legs, with temperature on the retreat shown on the line graph at the bottom. Many consider Minard's original the best statistical graphic ever drawn.





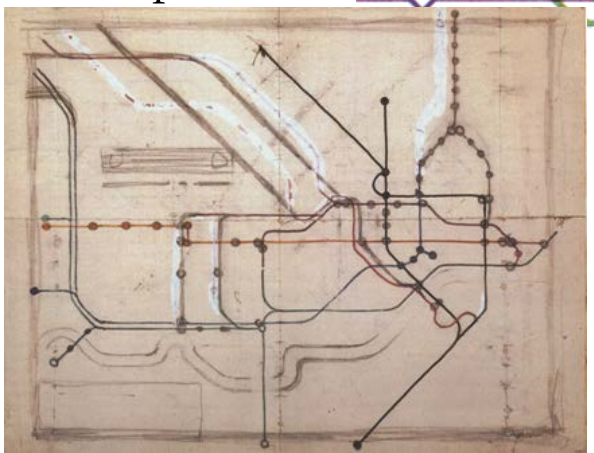
- Thickness of brown line is proportional to number of soldiers (422,00 soldiers started out).
- Black similarly encodes the retreat (10,000 returned).
- Crossing of Berezina river.
- Temperature plot at the bottom.



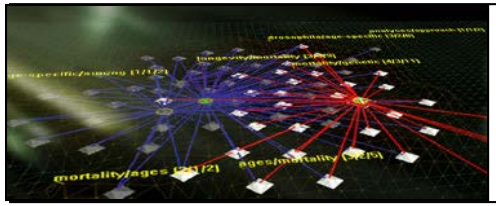
London Underground Map by Harry Beck in 1933

Harry Beck's London Underground's way finding system is an extraordinary example of directional signage.

Note the stylized angles (90 and 45 degrees) and the regular spacing between station stops.



Beck got 21 pounds for this work.

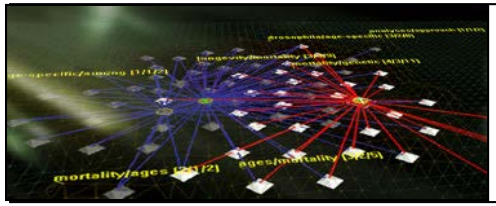


The Power of Mapping

- Visualizations are not objective, neutral artifacts.
- They are created. They include or leave out information.
- They communicate particular messages.
- Commonly, the messages are those of the powerful who pay for the visualizations.

Deconstruct Visualizations!

by questioning who the visualization was made for, by whom, why, and based on what data.



Edward Tufte's Principles Of Graphic Display

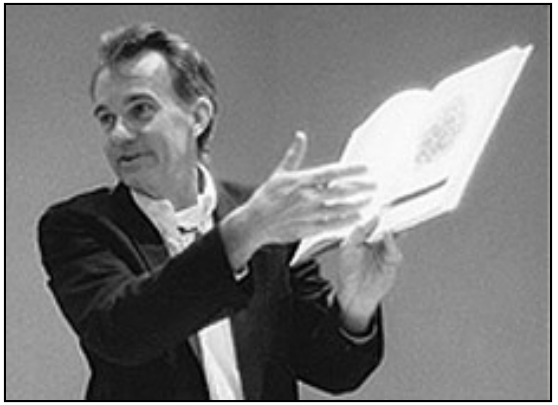
"The success of a visualization is based on deep knowledge and care about the substance, and the quality, relevance and integrity of the content."

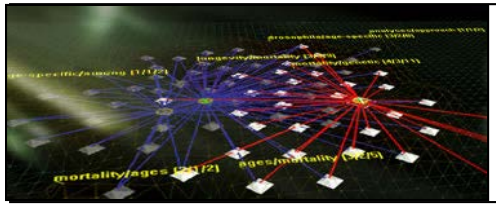
(Tufte, 1983)

Principles of Graphical Excellence

- Well-designed presentation of interesting data: substance, statistics, design.
- Complex ideas communicated with clarity, precision, and efficiency.
- Conveying the most knowledge in the shortest time with the least ink in the smallest space.
- It requires telling the truth about the data.
- It is nearly always multivariate.

(Tufte, 1983)



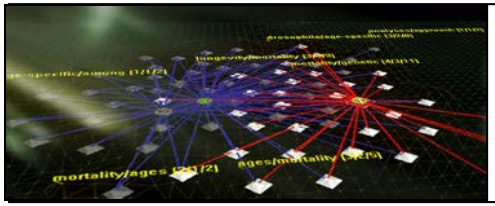


Five Principles in the Theory of Graphic Display

- Above all else show the data.
- Maximize the data-ink ratio, within reason.
- Erase non-data ink, within reason.
- Erase redundant data-ink.
- Revise and edit.

Visualizations should strive towards the following goals

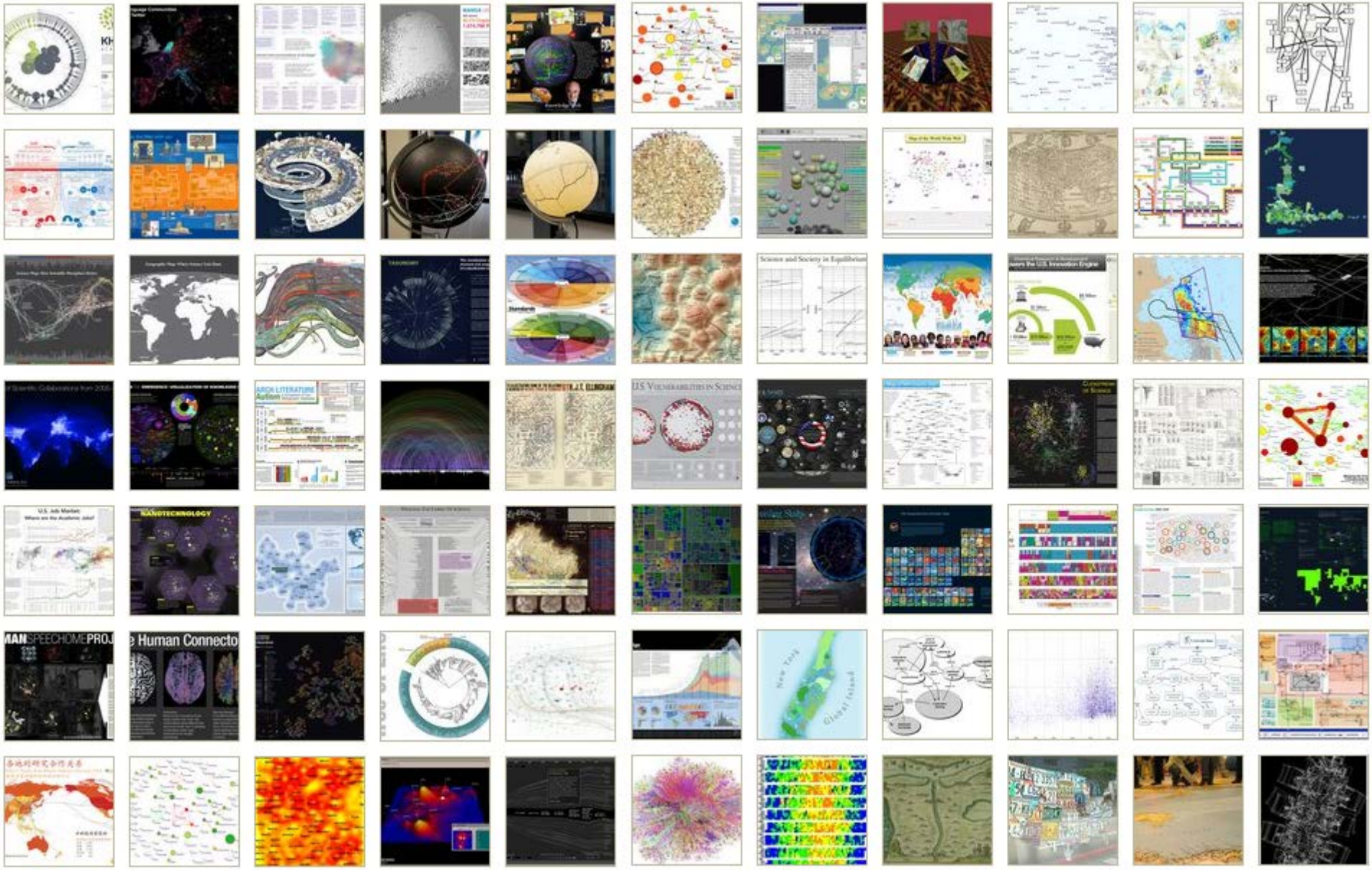
- Focus on content of data not the visualization technique.
- Strive for integrity.
- Utilize classic designs and concepts proven by time.
- Comparative rather than descriptive visualizations.
- High resolution.



Aesthetics

- Properly choose format and design
- Use words, numbers, drawings in close proximity
- Use lines of different weights as an attractive and compact way to display data.
- Reflect a balance, a proportion, a sense of relevant scale.
- Display an accessible complexity of detail.
- Let the graphics tell a story about the data.
- Avoid content-free decoration.
- Make use of symmetry to add beauty (although someone once said that "all true beauty requires some degree of asymmetry").
- Draw graphics an a professional manner, with the technical details of production down with care.

Visualization Frameworks


U.S. Job Market
NANOTECHNOLOGY
Human Connecto
Science and Society in Equilibrium
US Vulnerabilities in Science
New York
Global Island

How to Classify Different Visualizations?

By

- User insight needs?
- User task types?

- Data to be visualized?
- Data transformation?

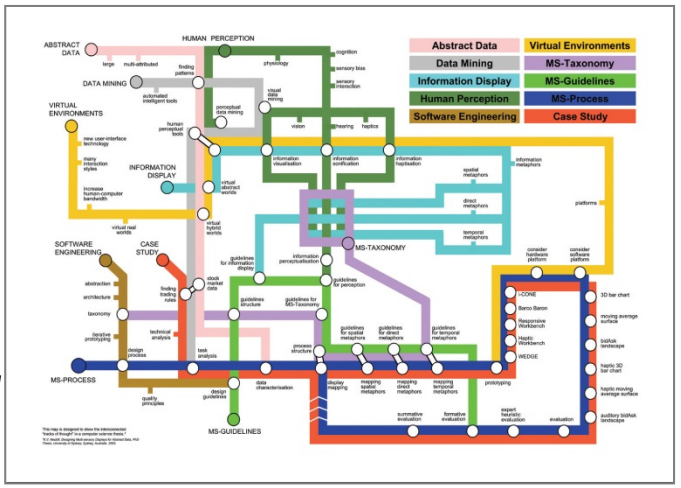
- Visualization technique?
- Visual mapping transformation?
- Interaction techniques?

- Or ?

Different Question Types

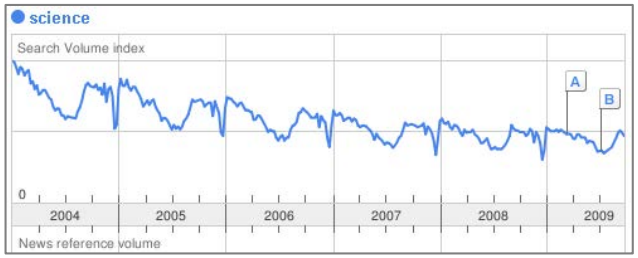
Terabytes of data

Descriptive & Predictive Models



Find your way

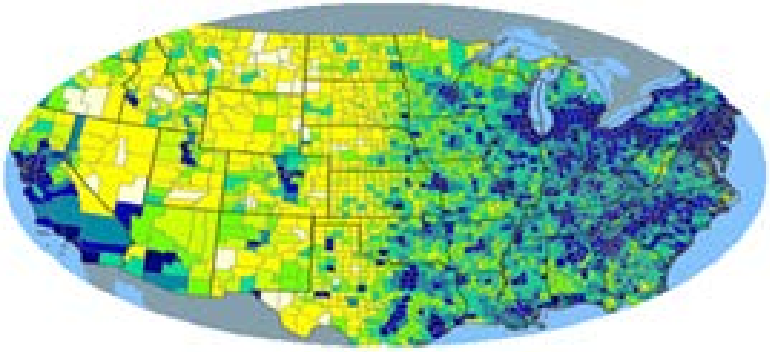
Find collaborators, friends



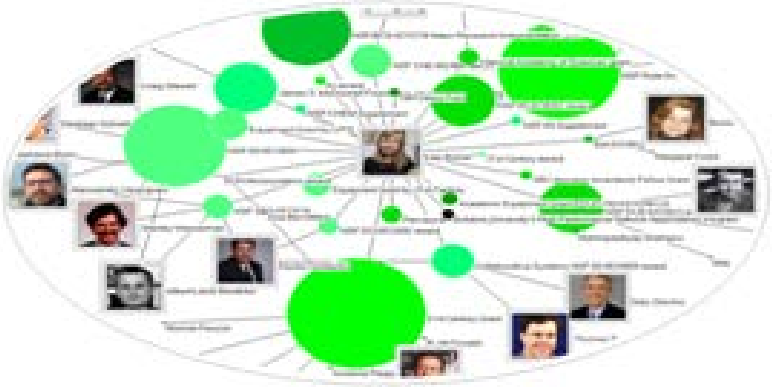
Identify trends

Different Levels of Abstraction/Analysis

Macro/Global
Population Level

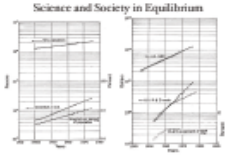
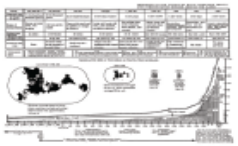
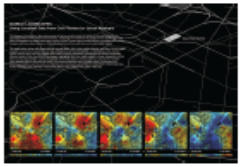
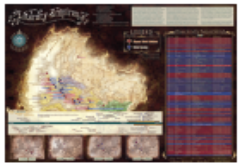
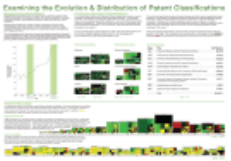


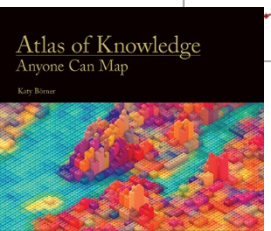
Meso/Local
Group Level



Micro
Individual Level

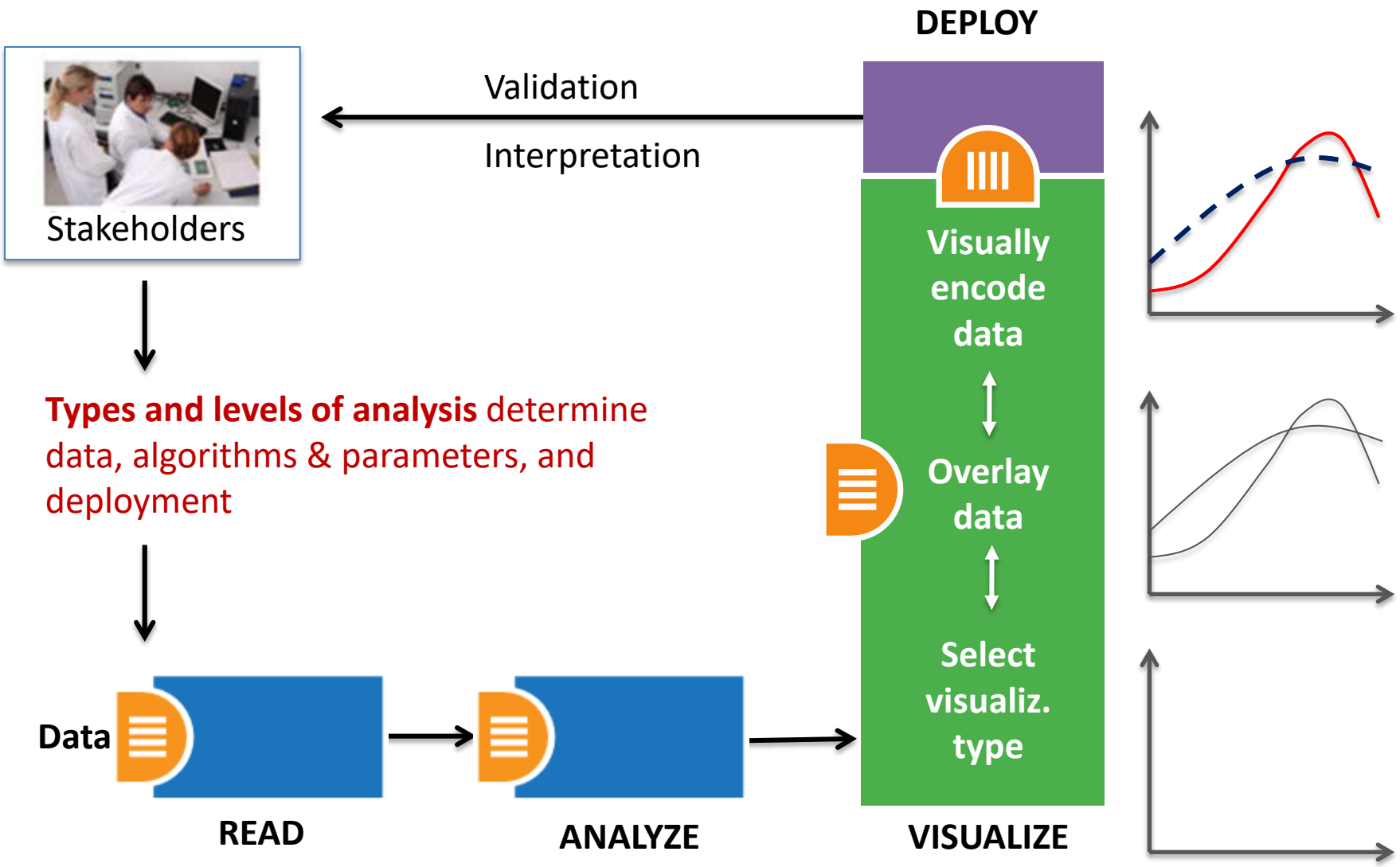
Tasks

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

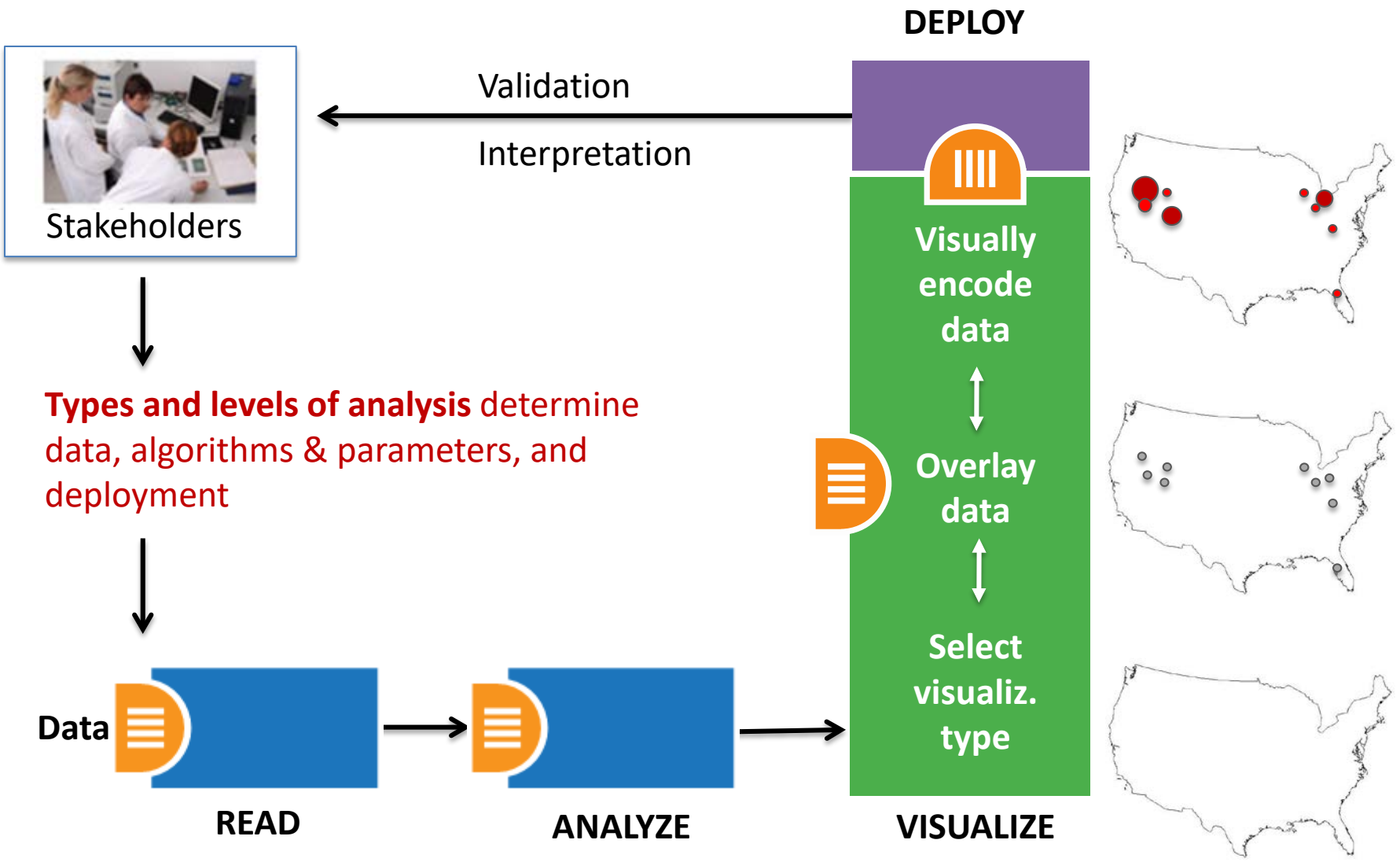


See page 5

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

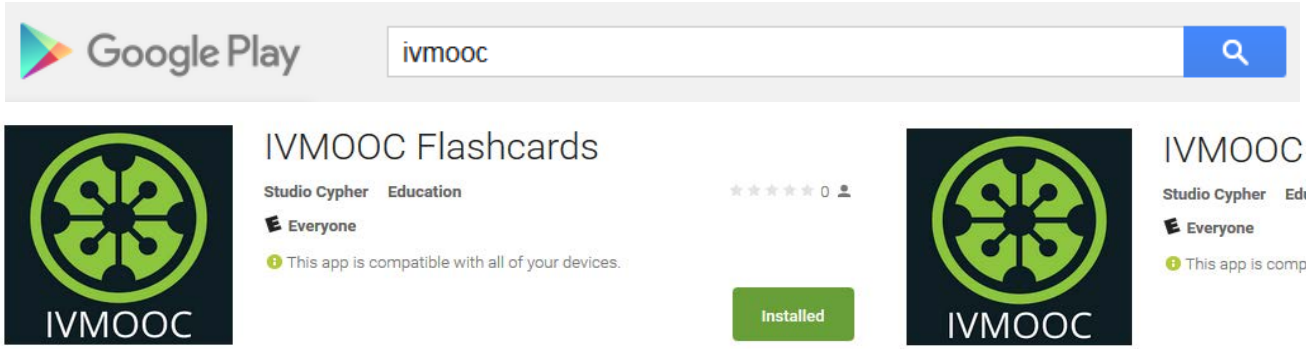
Final Exam

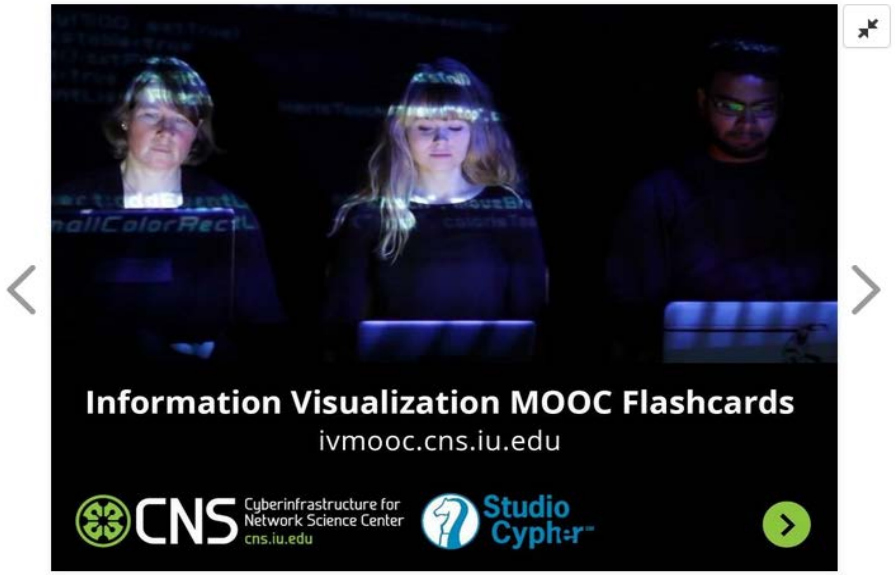
Part 2: Students work in teams on client projects.

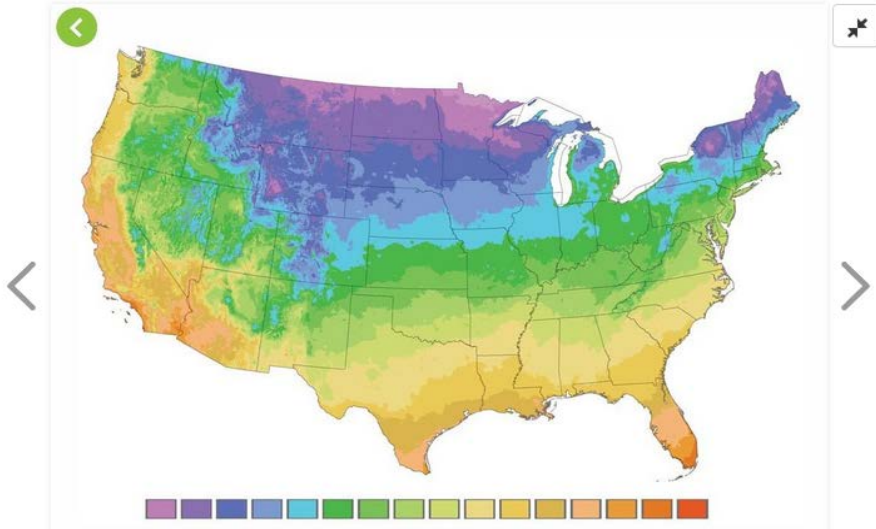
Final grade is based on Homework and Quizzes (**10%),** Midterm (**20%**), Final (**30%**), Client Project (**30%**), and Class Participation (**10%**).

IVMOOC App – More than 60 visualizations

The “IVMOOC Flashcards” app can be downloaded from Google Play and Apple iOS stores.







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

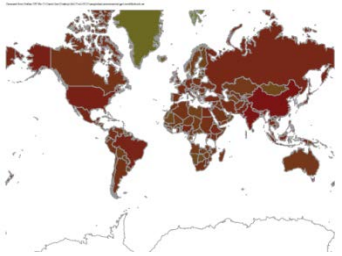
Statistical Analysis–p. 44

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

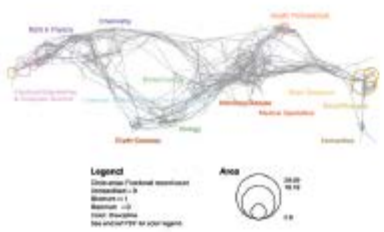
Geospatial Analysis–p. 52



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

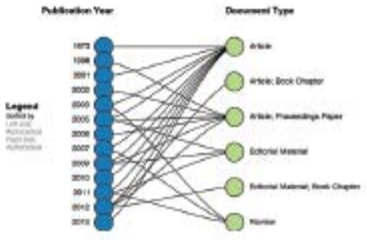
Times Cited	Publication Year	City of Publisher	Country	Journal Title (Full)	Title	Subject Category	Authors
12	2011	NEW YORK	USA	COMMUNICATIONS OF THE ACM	Plug-and-Play Macroscopes	Computer Science	Borner, K
18	2010	MALDEN	USA	CTS-CLINICAL AND TRANSLATIONAL SCIENCE	Advancing the Science of Team Science	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

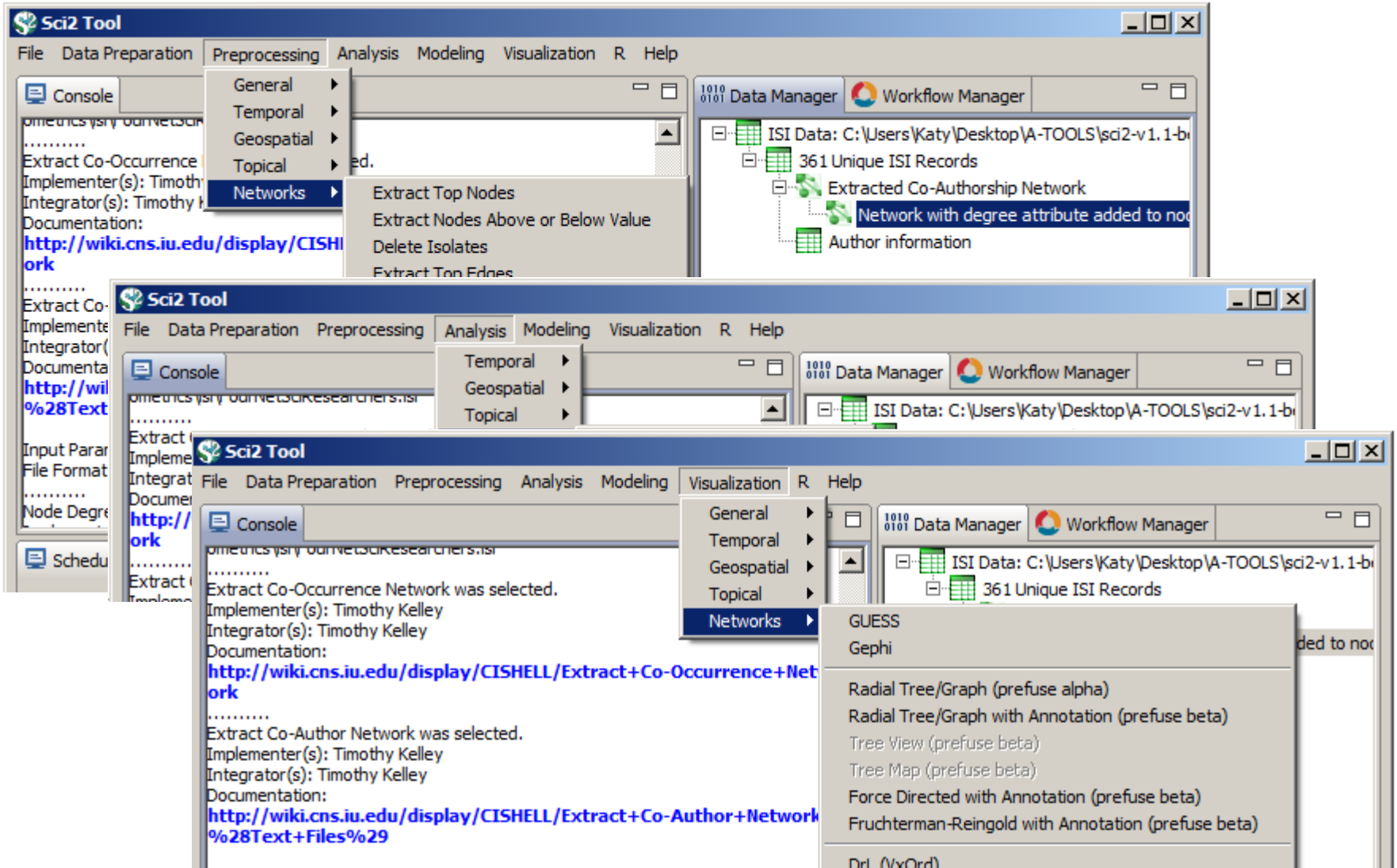
Bi-Modal Network–p. 60



Co-author and many other bi-modal networks.

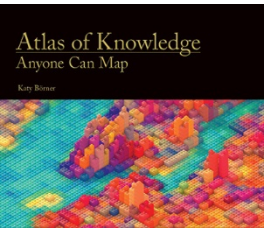
Sci2 Tool Interface Components Implement Vis Framework

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



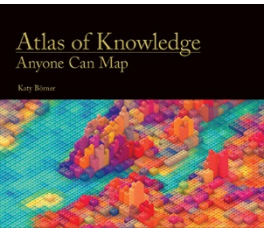
See page 24

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

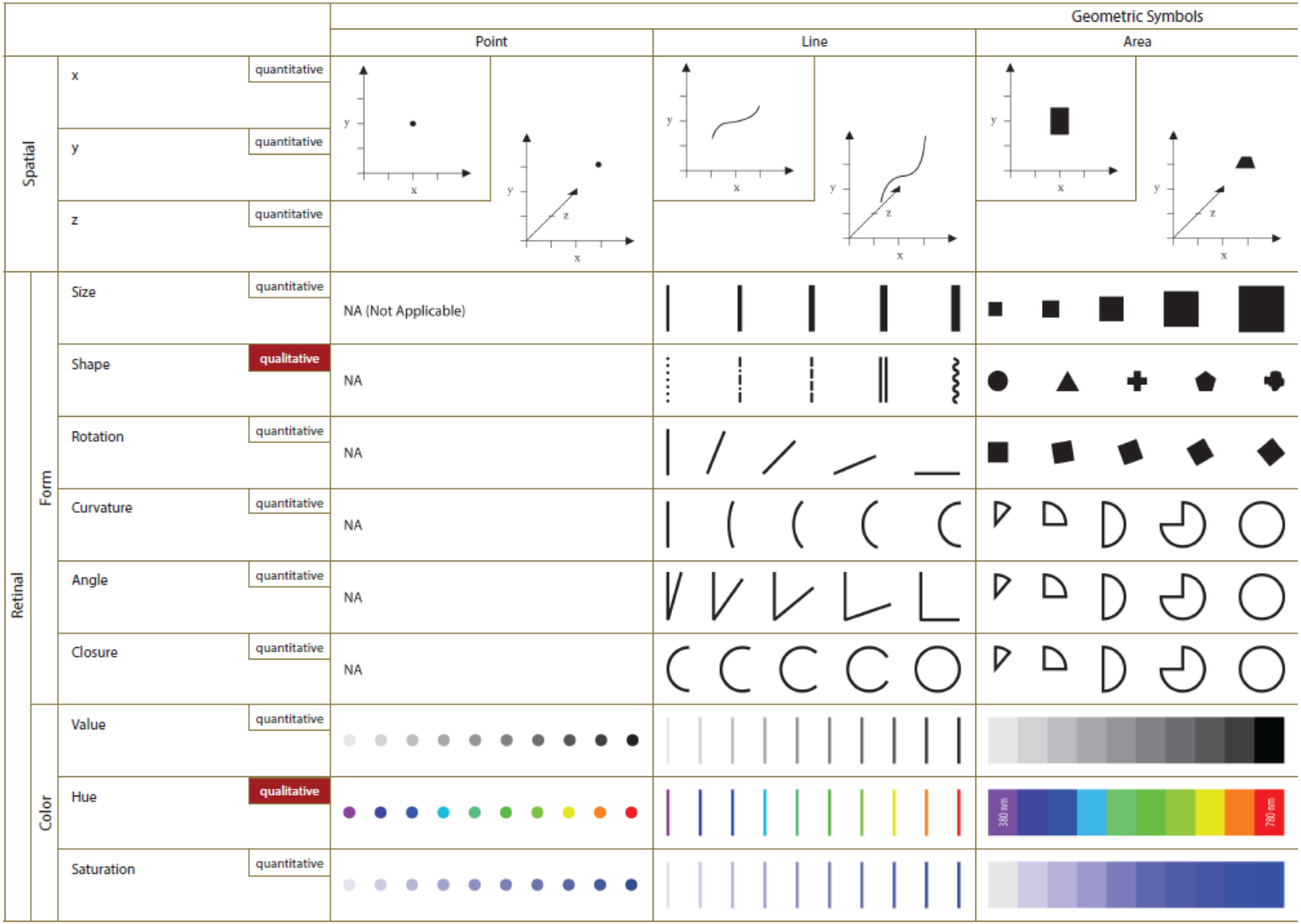
Visualization Framework

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

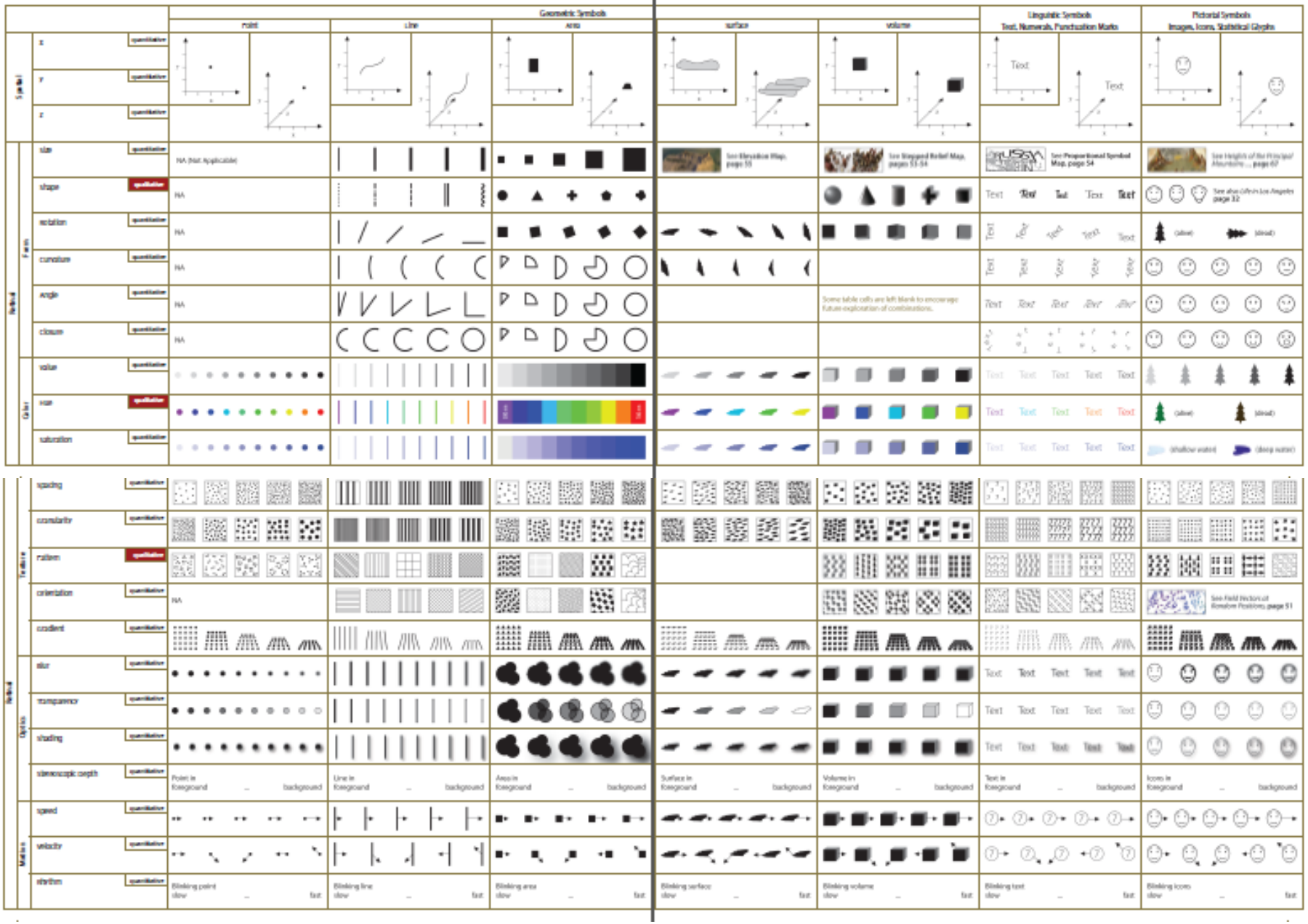


See page 24

Graphic Variable Types Versus Graphic Symbol Types



Graphic Variable Types Versus Graphic Symbol Types



Homework

Pick a visualization that provided you with new insights AND made you change your behavior.

Answer the following questions:

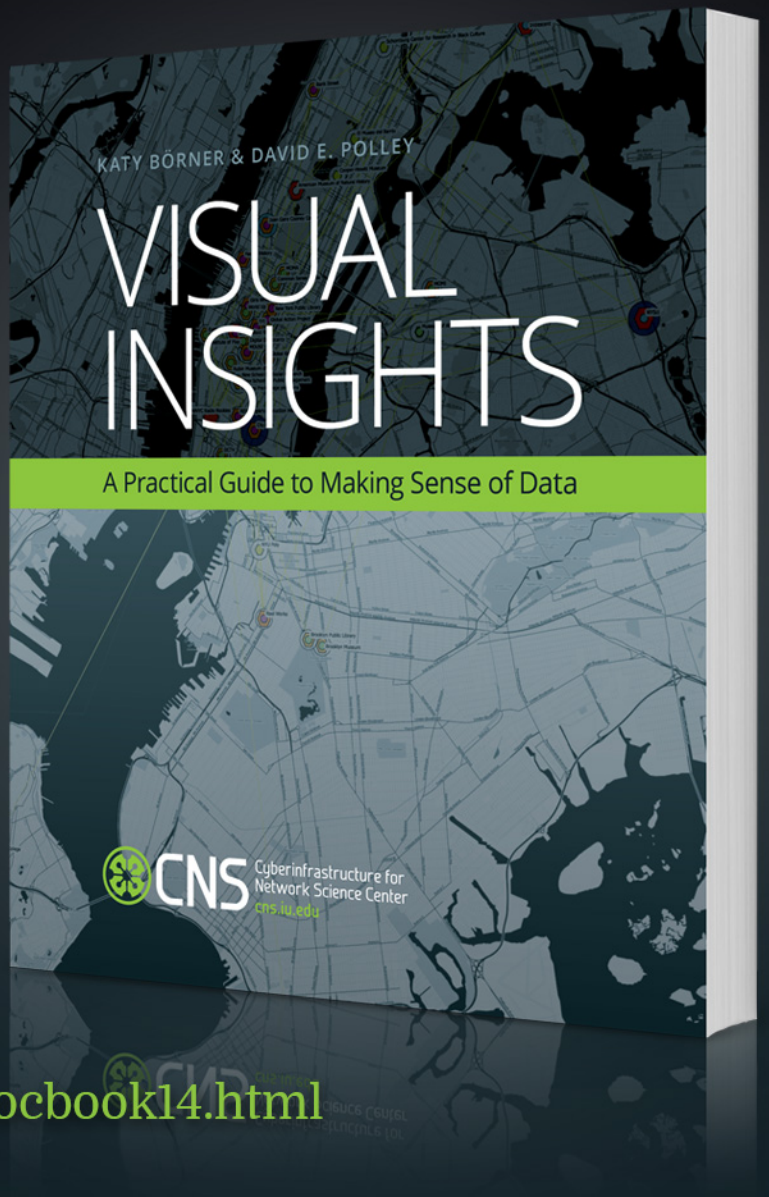
1. What tasks (temporal, geo, etc.) does the visualization support?
2. At what level of abstraction (micro, meso, macro)?
3. What dataset(s) were used? Cite sources.
4. What algorithms were used? Cite sources/provide Github links.
5. What visual mappings were used (graphic symbols and variables)?
6. How can the visualization be improved?

7. How did your behavior change?

The IVMOOC Companion Textbook

This textbook offers a gentle introduction to the design of insightful visualizations. It seamlessly blends theory and practice, giving readers both the theoretical foundation and the practical skills necessary to render data into insights.

The book accompanies the Information Visualization MOOC that attracted students, scholars, and practitioners from many fields of science and more than 100 different countries.



cns.iu.edu/ivmoocbook14.html

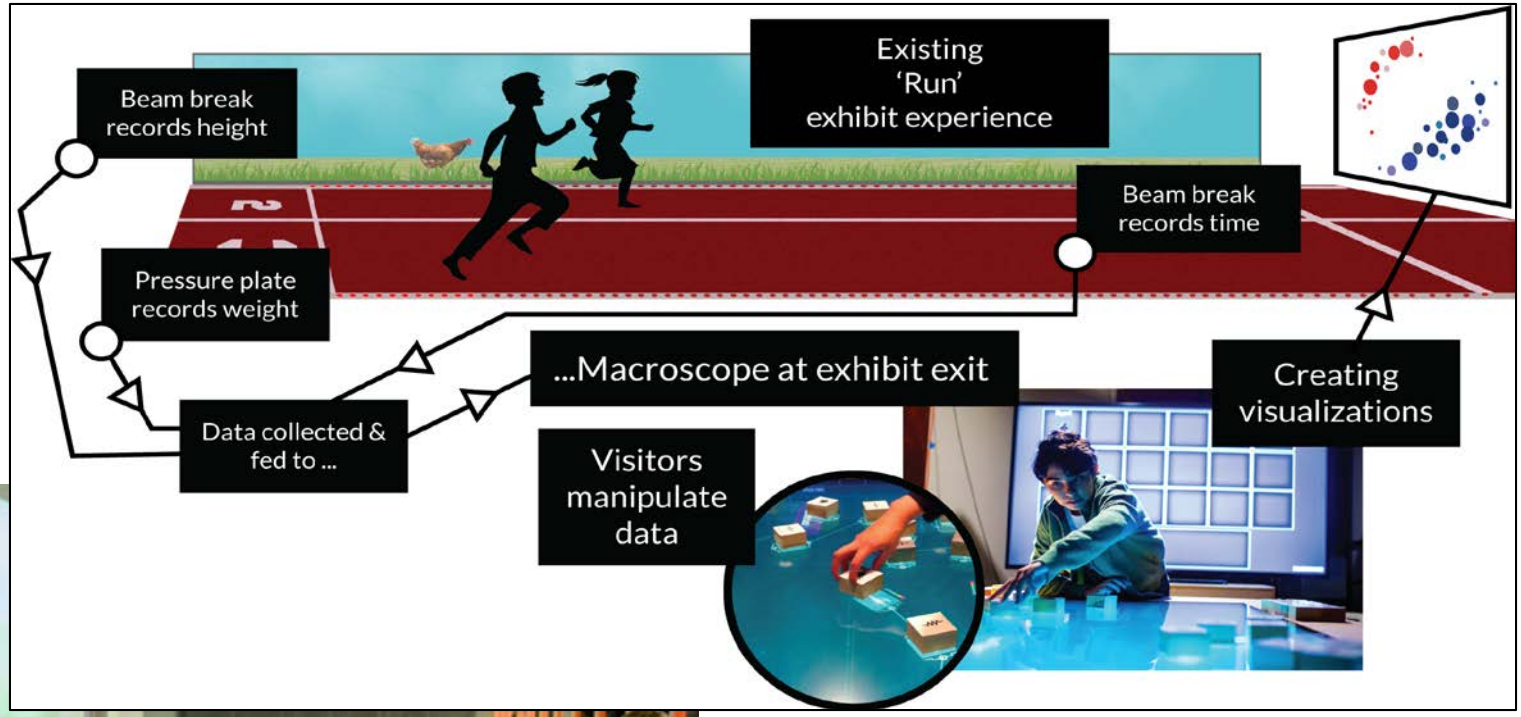
ENGR-E483/E583 Information Visualization, http://ivmooc.cns.iu.edu

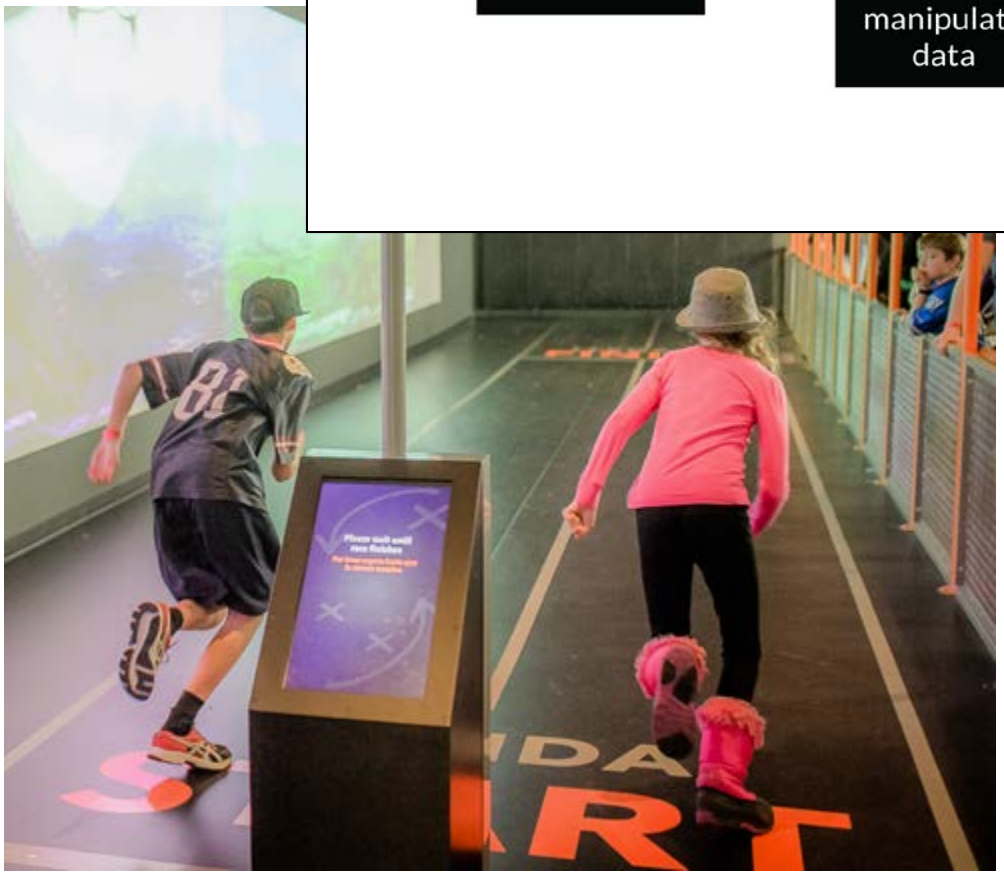
Data Visualization Literacy

Data Visualization Literacy: Research and Tools that Advance Public Understanding of Scientific Data

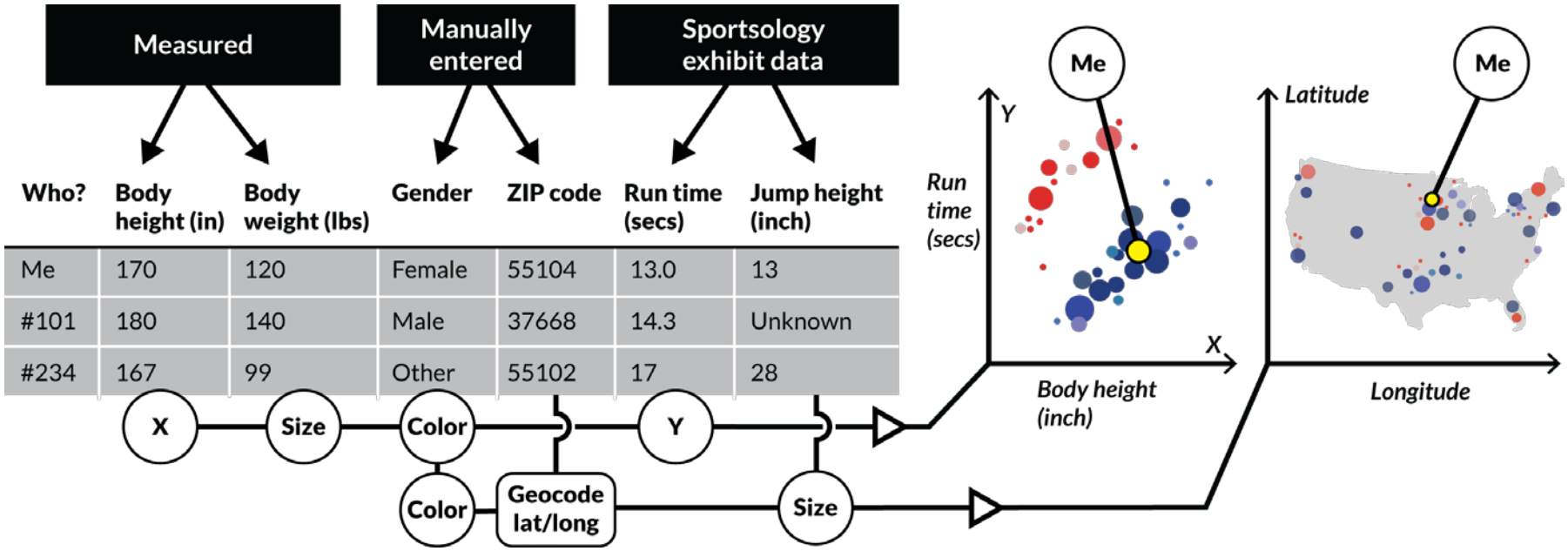
NSF Org:	DRL Division Of Research On Learning
Initial Amendment Date:	June 13, 2017
Latest Amendment Date:	June 13, 2017
Award Number:	1713567
Award Instrument:	Standard Grant
Program Manager:	Arlene M. de Strulle DRL Division Of Research On Learning EHR Direct For Education and Human Resources
Start Date:	August 1, 2017
End Date:	July 31, 2021 (Estimated)
Awarded Amount to Date:	$1,355,236.00
Investigator(s):	Katy Borner katy@indiana.edu (Principal Investigator) Kylie Peppler (Co-Principal Investigator) Bryan Kennedy (Co-Principal Investigator) Stephen Uzzo (Co-Principal Investigator) Joe Heimlich (Co-Principal Investigator)

https://www.youtube.com/watch?v=oy34R45EfBg

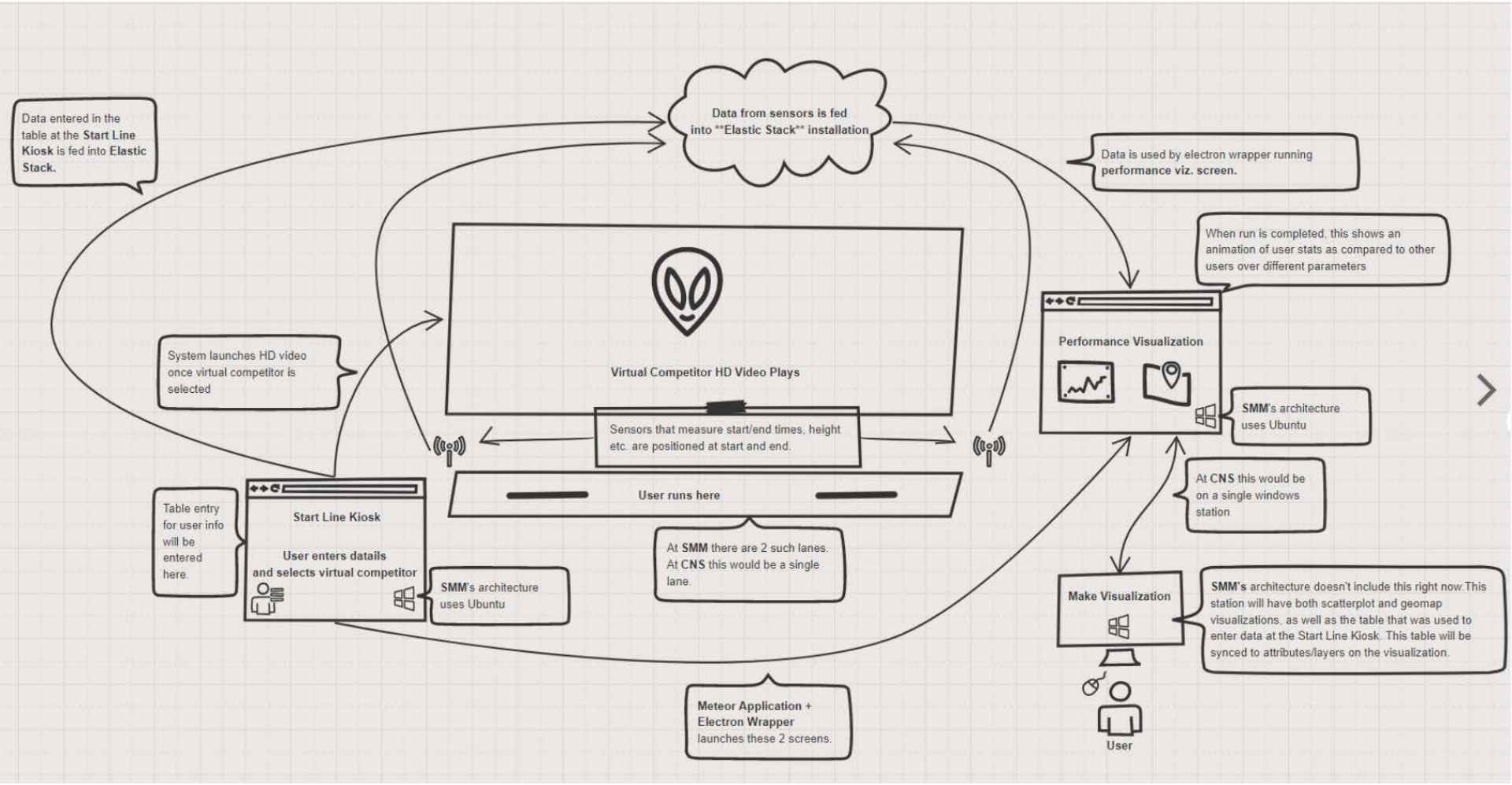




Sketch of the *Run* exhibit including data collection (top) and macroscope add-on that lets interested visitors explore more complex data visualizations using table-top displays.



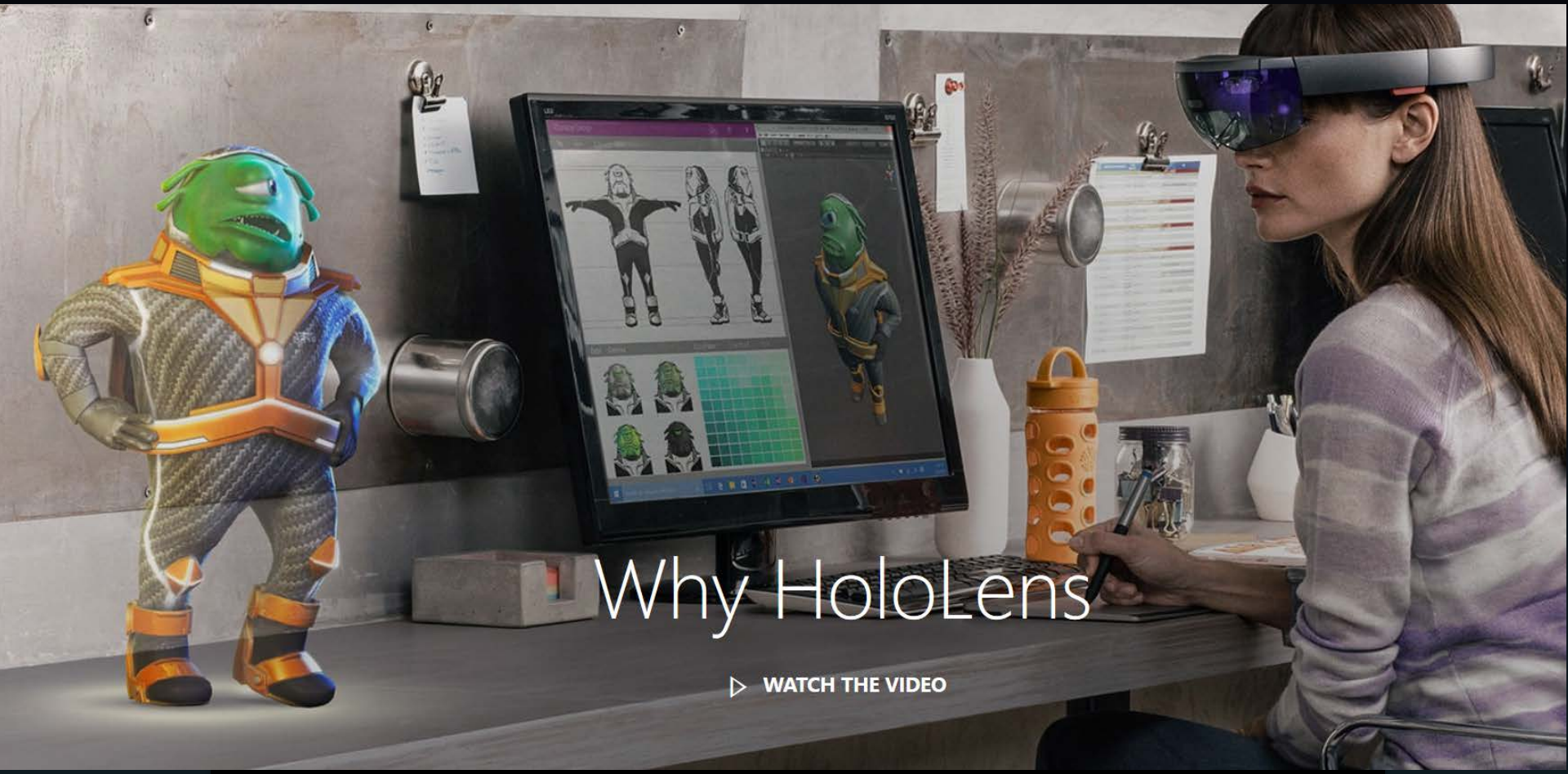
xMacroscope general setup and activity—Raw data on left is converted to visualization on right by dragging and dropping (or connecting) column headers to axes, paint buckets, size, and shape.



Data federation via Elastic Stack, Performance Visualization and Make Visualization

SCWS Session: Visualizing STEAM Data in Support of Smart Decision Making
November 15-17, 2017, Tokyo, Japan. http://scws2017.org

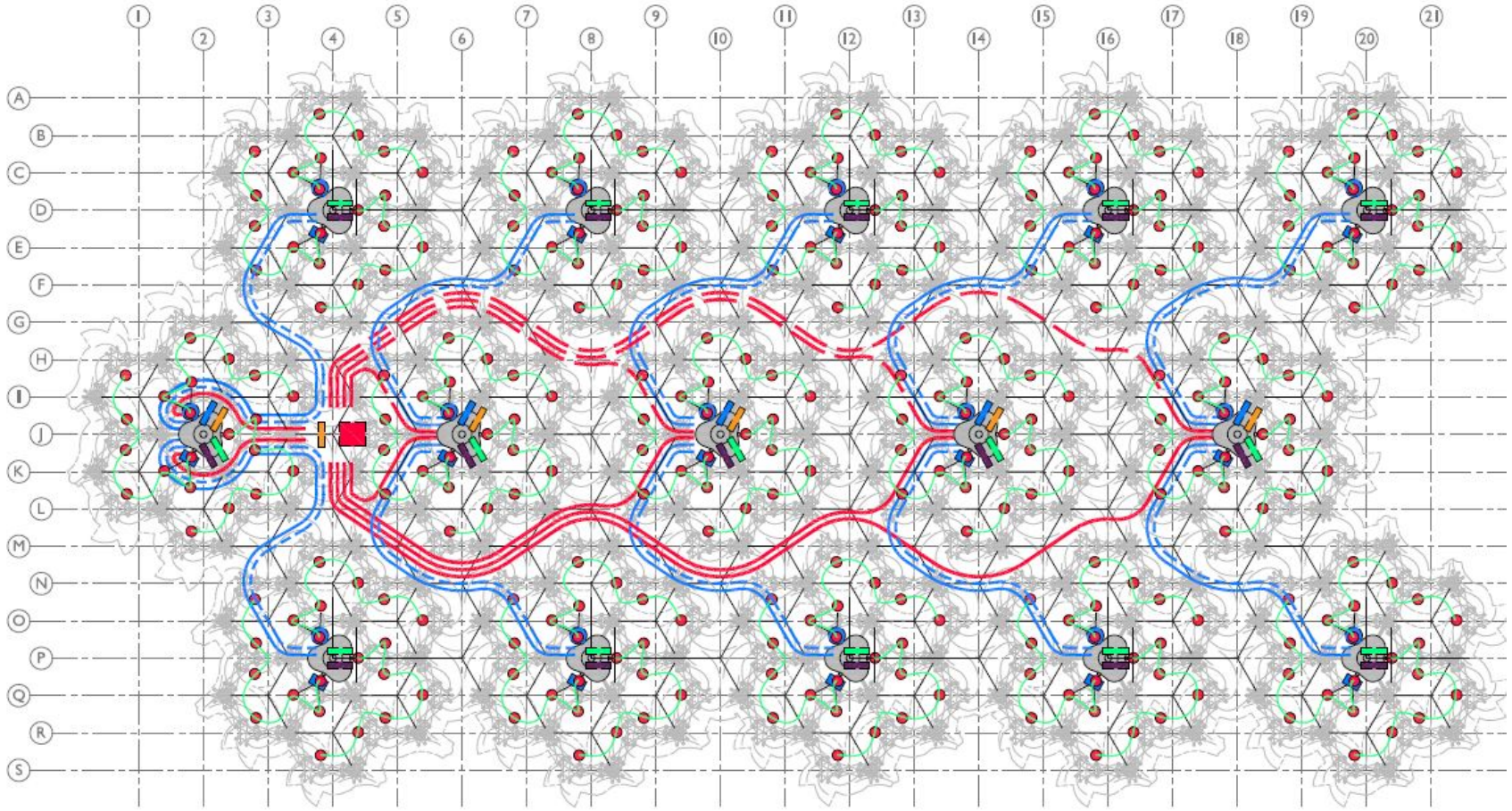
Augmented Reality Visualizations


Why HoloLens
WATCH THE VIDEO

Sentient Veil, Isabella Stewart Gardner Museum, Boston, MA (2017)

Sentient Veil, Isabella Stewart Gardner Museum, Boston, MA (2017)

Sentient Veil, Isabella Stewart Gardner Museum, Boston, MA (2017)



PCBs

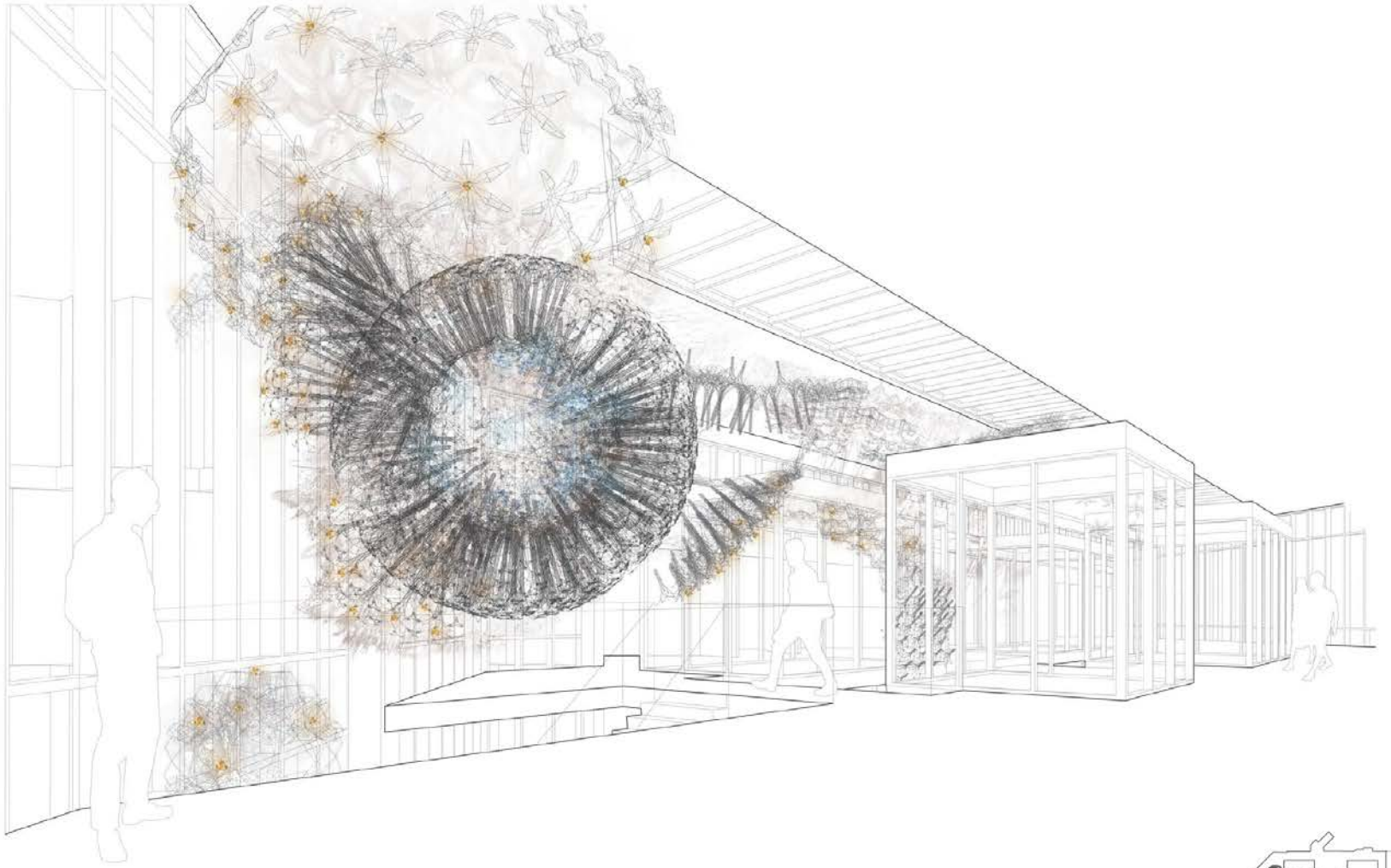
- Raspberry Pi 3.0 B
- 3.1 Control Node
- 3.1 Device Module
- Power Distribution board
- Mp3 Trigger board

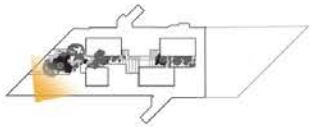
Interactive Devices

- Light Module Cluster
- Sound Unit
- Short Range IR Proximity Sensor
- Microphone Sensor

Mechanisms

- Main Power Trunk
- Main Communications Trunk
- Node to Device Module Power
- Node to Device Module Communications
- Neo Pixel Chain

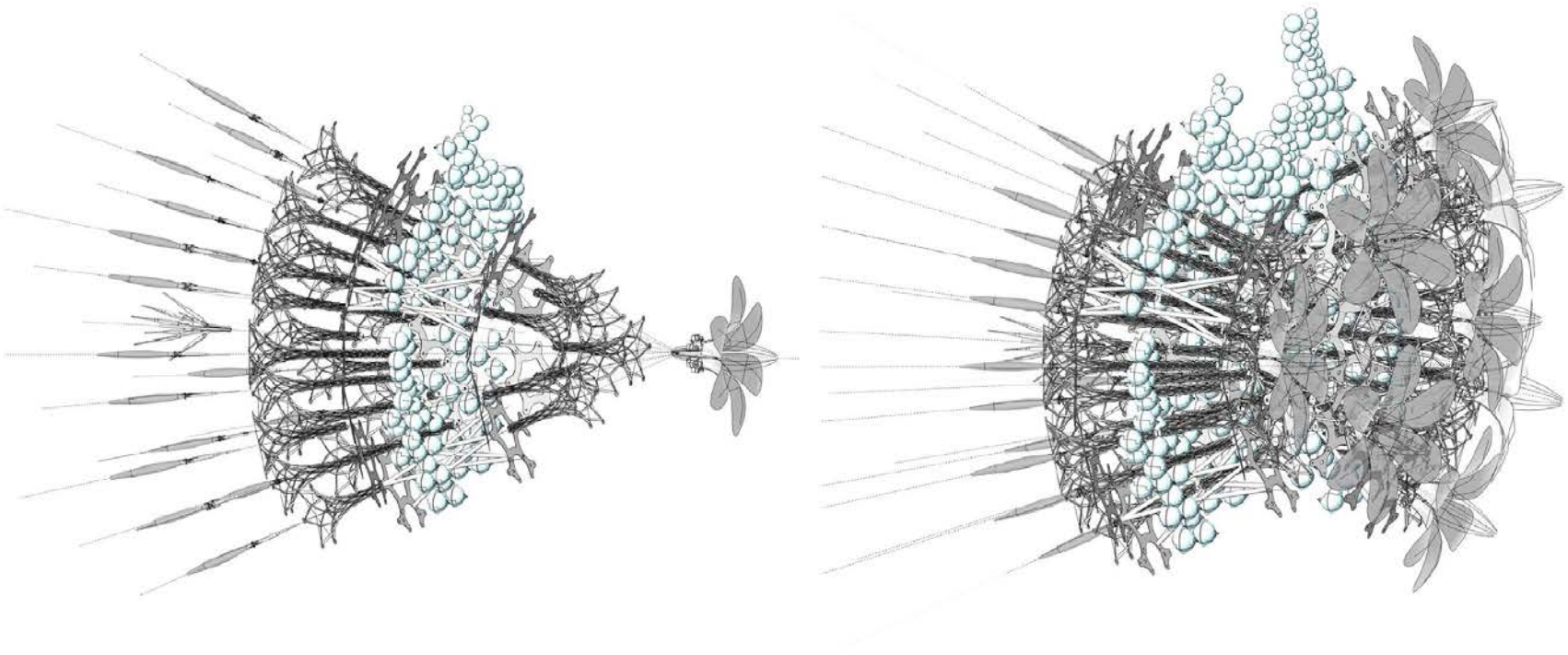




Luddy Hall Installation
Indiana University Bloomington
April 29 2017

UPPER ATRIUM

Philip Beesley • Living Architecture Systems



Luddy Hall Installation
Indiana University Bloomington
April 29 2017

ASSEMBLY SAMPLE

Philip Beesley • Living Architecture Systems

Upcoming Events

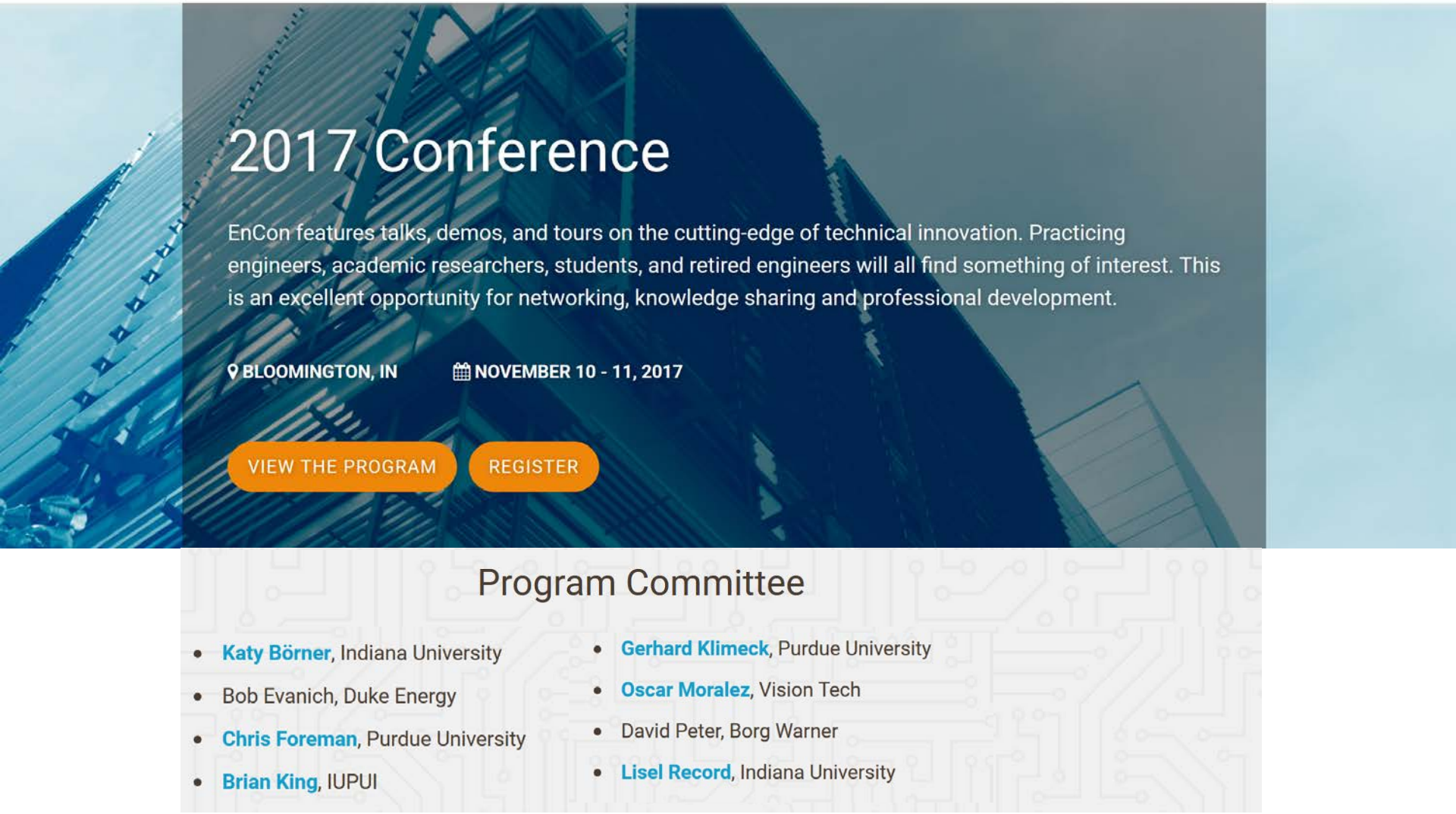
SCIENTIFIC VISUALIZATION

Description: SCIENTIFIC VISUALIZATION webinar track provides an overview of data visualization by exploring the underlying principles of visualization and set the stage for other introductory data visualization seminars. Webinars focus on individual tools and capabilibities for visualizing data and information.

The webinars offered include:

DATE	TITLE	SPEAKER	LENGTH (HOURS)
March 1, 2017	**Introduction to Data Visualization** Slides | Video | Survey	Vetria Byrd	1
March 15, 2017	**VisIT: Scalable HPC Visualization and Data Analysis** Slides | Video | Survey	Kevin Griffin, Eric Brugger, Cyrus Harrison	2
April 5, 2017	**Introduction to Scientific Visualization Using ParaView** Slides | Video | Survey	Vetria Byrd	1
June 7, 2017	**Visualizing non-Spatial Data** Slides | Video | Survey	Rob Sisneros, Mark Van Moer	1 – 2
June 21, 2017	Hyperglyphs	Jeff Sale	2
August 2, 2017	Badges for Visualization Micro-Certification	Jeff Sale	1
August 16, 2017	Introduction to the Eclipse Advanced Visualization Project	Robert Smith	1
October 25, 2017	Scientific Visualization in Houdini	Kalina Borkiewicz, A. J. Christensen	2

https://bluewaters.ncsa.illinois.edu/webinars/visualization



Program Committee

- Katy Börner, Indiana University
- Bob Evanich, Duke Energy
- Chris Foreman, Purdue University
- Brian King, IUPUI
- Gerhard Klimeck, Purdue University
- Oscar Moralez, Vision Tech
- David Peter, Borg Warner
- Lisel Record, Indiana University

http://www.cis-ieee.org/encon2017

CIS-IEEE EnCon 2017 Student Scholarship Application

Date:
November 10-11, 2017

Location:
Cyberinfrastructure Building
Indiana University
2709 East 10th Street
Bloomington, IN 47405

Twenty students will receive free admission to the full conference. Students who submit a poster will be given preference. For more information about the conference visit cis-ieee.org/EnCON2017 or contact Lisel Record at recorde@indiana.edu.

* Required

20 ISE students will receive free admission to CIS-IEEE EnCon.
To apply, please complete this web form.

Full name *

Your answer

INDIANA UNIVERSITY
INTELLIGENT SYSTEMS ENGINEERING

HOME CAMPS ABOUT ISE LOCATIONS CONTACT

Sentient Architecture — Sculptures that Listen and Talk

As the built environment becomes increasingly more complex and integrated with new technologies—including the emerging Internet of Things (IoT)—there is an urgent need to understand how embedded technologies affect the experience of individuals that inhabit these spaces.

View Sentient Architecture Camp

http://camps.engineering.indiana.edu

Join us at diverse Dendrite building sessions.

Join our annual Open House on Fri Oct 27, 4-7p, IUNI House.

CNS Tools and Services are used by hundreds of thousands around the globe.

Our mission is to advance datasets, tools, and services for the study of biomedical, social and behavioral science, physics, and other networks. A specific focus is research on the structure and evolution of science and technology (S&T) and the communication of results via static and interactive maps of science. Learn more at cishell.org.

CNS Cyberinfrastructure for Network Science Center
cns.iu.edu

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