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, Computing, and Engineering
Indiana University Network Science Institute (IUNI)
Indiana University, Bloomington, IN, USA
+ 2018 Humboldt Fellow, TU Dresden, Germany

Data Visualization Session at the European Meeting on Applied Evolutionary Economics (EMAEE)
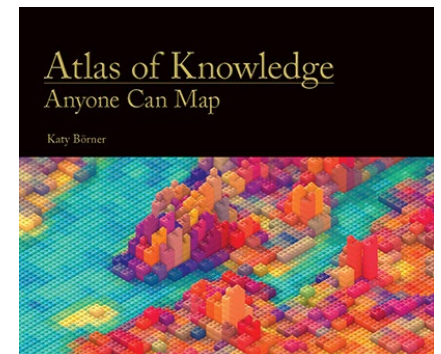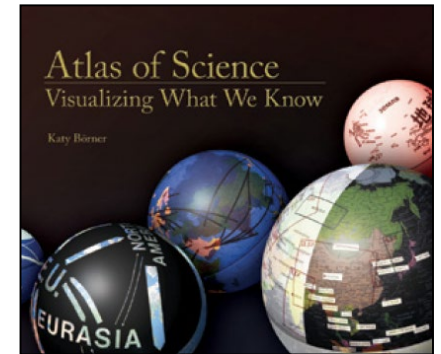
Brighton, UK | June 4, 2019

Overview

This talk will introduce a theoretical data visualization framework (DVL) meant to empower anyone to systematically render data into insights using temporal, geospatial, topical, and network analyses and visualizations.

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

Exemplarily, the DVL is applied to map science and technology, see interactive data visualizations from the *Places & Spaces: Mapping Science* exhibit (http://scimaps.org) and recent *PNAS* special issue on *Modelling and Visualizing Science and Technology Developments* (https://www.pnas.org/modeling)

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

Börner, Katy. 2010. *Atlas of Science: Visualizing What We Know*. Cambridge, MA: The MIT Press.

Atlas of Forecasts

Places & Spaces: Mapping Science Exhibit

1st **Decade** (2005-2014)

2nd **Decade** (2015-2024)

3rd **Decade** (2015-2034)

Maps

Macroscopes

http://scimaps.org

Data Visualization Literacy Framework

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.

Visualization Frameworks

MANY frameworks and taxonomies have been proposed to

- help organize and manage the evolving zoo of 500+ different data visualization types,

- provide guidance when designing data visualizations, and

- facilitate teaching.
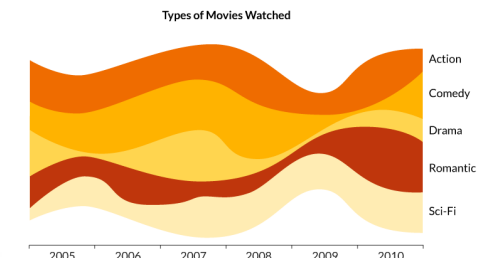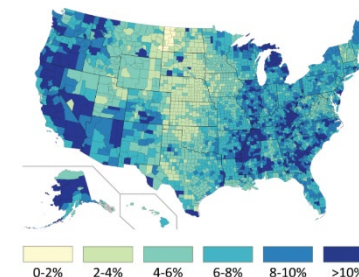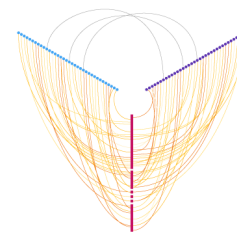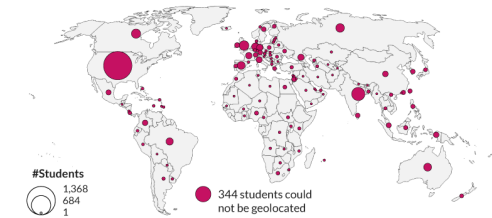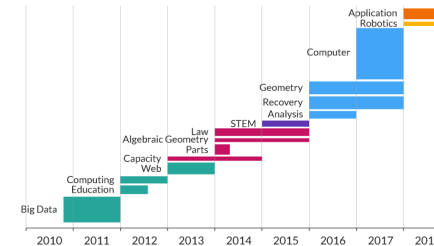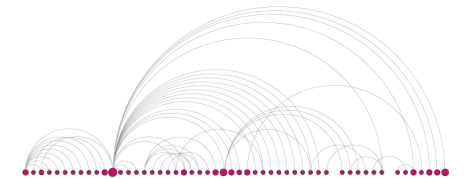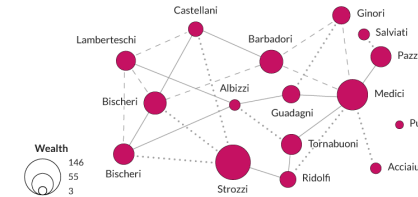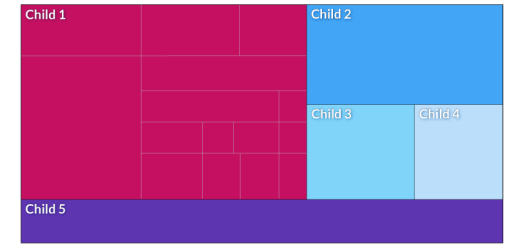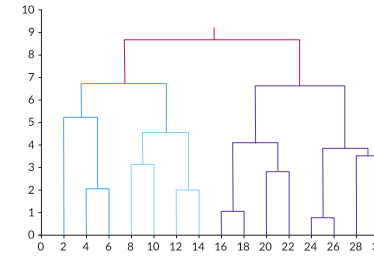
Existing Visualization Frameworks

Organize data visualizations by

- User insight needs

- User task types

- Data to be visualized

- Data transformations

- Visualization technique

- Visual mapping transformations

- Interaction techniques

- Deployment options

- and other features …
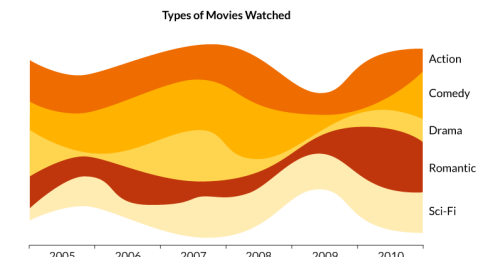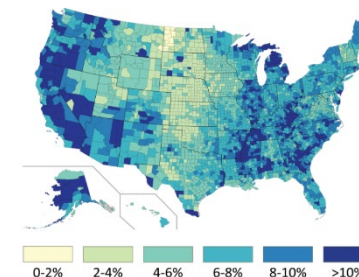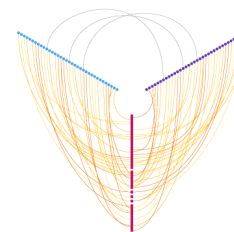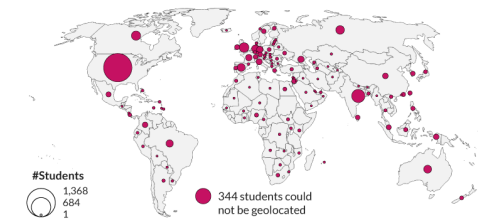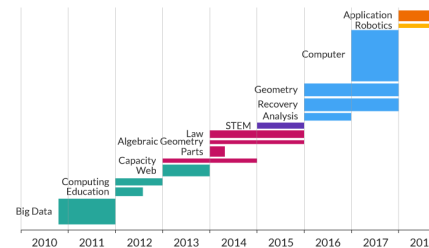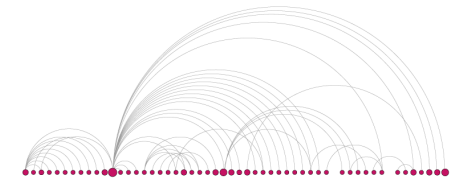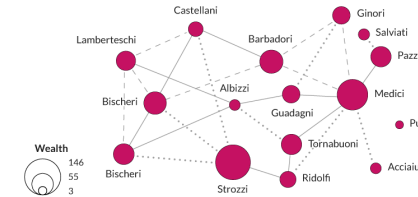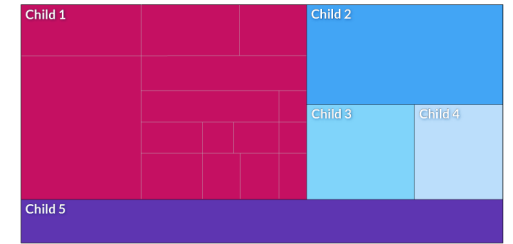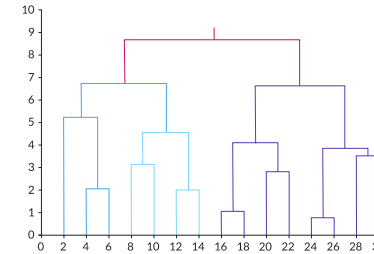
DVL Framework: Desirable Properties

- Most existing frameworks focus on **READING**. We believe that much expertise is gained from also **CONSTRUCTING** data visualizations.

- Reading and constructing data visualizations needs to take human perception and cognition into account.

- Frameworks should build on and consolidate prior work in cartography, psychology, cognitive science, statistics, scientific visualization, data visualization, learning sciences, etc. in support of a de facto standard.

- Theoretically grounded + practically useful + easy to learn/use.

- Highly modular and extendable.

DVL Framework: Development Process

- The initial DVL-FW was developed via an extensive literature review.

- The resulting DVL-FW typology, process model, exercises, and assessments were then tested in the *Information Visualization* course taught for more than 15 years at Indiana Unirsity. More than 8,500 students enrolled in the IVMOOC version (http://ivmooc.cns.iu.edu) over the last six years.

- The FW was further refined using feedback gained from constructing and interpreting data visualizations for 100+ real-world client projects.

- Data on student engagement, performance, and feedback guided the continuous improvement of the DVL-FW typology, process model, and exercises for defining, teaching, and assessing DVL.

- The DVL-FW used in this course supports the systematic construction and interpretation of data visualizations.

Data Visualization Literacy Framework (DVL-FW)

Consists of two parts:

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

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |

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

Data Scales
- nominal
- ordinal
- interval
- ratio

Analyses
- statistical
- temporal
- geospatial
- topical
- relational

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

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

Graphic Variables
- spatial
   position
- retinal
  form
  color
  optics
  motion

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.

Interpret · Deploy · Stakeholders · Acquire · Analyze · Visualize

CNS Cyberinfrastructure for Network Science Center
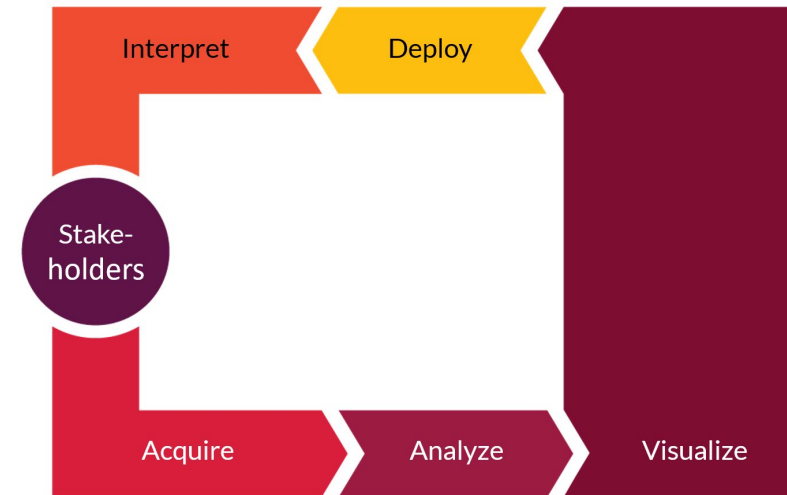
Data Visualization Literacy Framework (DVL-FW)

Consists of two parts that are interlinked:

**DVL Typology +
DVL Workflow Process**
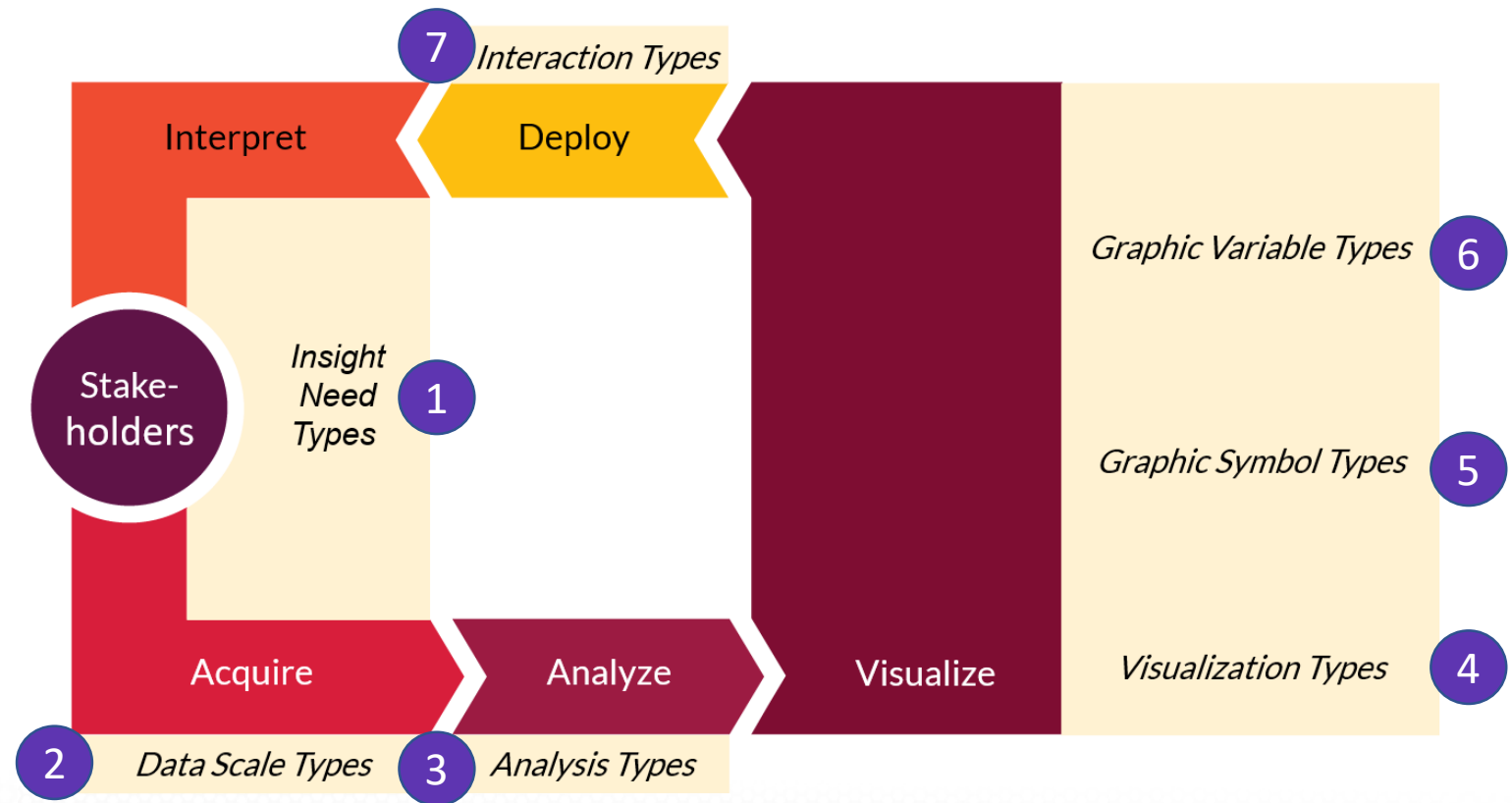
Data Visualization Literacy Framework (DVL-FW)

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

Typology of the Data Visualization Literacy Framework

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

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

Typology of the Data Visualization Literacy Framework

1

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

Börner, Katy. 2015. *Atlas of Knowledge: Anyone Can Map*. Cambridge, MA: The MIT Press. 26-27.

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

Typology of the Data Visualization Framework

2

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

Börner, Katy. 2015. *Atlas of Knowledge: Anyone Can Map*. Cambridge, MA: The MIT Press. 28-29.

Data Scale Types

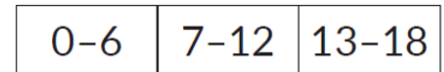**Nominal:** A categorical scale, also called a nominal or category scale, is qualitative. Categories are assumed to be non-overlapping.

Ordinal: An ordinal scale, also called sequence or ordered, is quantitative. It rank-orders values representing categories based on some intrinsic ranking, but not at measurable intervals.

Interval: An interval scale, also called a value scale, is a quantitative numerical scale of measurement where the distance between any two adjacent values (or intervals) is equal, but the zero point is arbitrary.

Ratio: A ratio scale, also called a proportional scale, is a quantitative numerical scale. It represents values organized as an ordered sequence, with meaningful uniform spacing, and a true zero point.
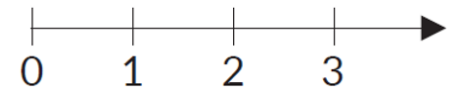
Data Scale Types - Examples

Nominal: Words or numbers constituting the "categorical" names and descriptions of people, places, things, or events.

Ordinal: Days of the week, degree of satisfaction and preference rating scores (e.g., using a Likert scale), or rankings such as low, medium, high.

Interval: Temperature in degrees or time in hours. Spatial variables such as latitude and longitude are interval.

Ratio: Physical measures such as height, weight, (reaction) time, or intensity of light; number of published papers, co-authors, citations.

Data Scale Types - Examples

Nominal: Words or numbers constituting the "categorical" names and descriptions of people, places, things, or events.

Qualitative

Ordinal: Days of the week, degree of satisfaction and preference rating scores (e.g., using a Likert scale), or rankings such as low, medium, high.
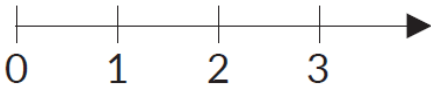
Quantitative

Interval: Temperature in degrees or time in hours. Spatial variables such as latitude and longitude are interval.

Ratio: Physical measures such as weight, height, (reaction) time, or intensity of light; number of published papers, co-authors, citations.

Data Scale Types - Mathematical Operations

This table shows the logical mathematical operations permissible, the measure of central tendency, and examples for the different data scale types.

Data Scale Types	Logical Mathematical Operations				Measure of Central Tendency	Examples
	= ≠	< >	+ -	x ÷		
Nominal	y				mode	🏠 🌲 🚗
Ordinal	y	y			median	😊 😐 ☹️
Interval	y	y	y		arithmetic mean	0–6 | 7–12 | 13–18
Ratio	y	y	y	y	geometric mean	0 1 2 3

Qualitative

Quantitative

Typology of the Data Visualization Literacy Framework

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

Data Scales
- nominal
- ordinal
- interval
- ratio

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

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

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

Graphic Variables
- spatial
  position
- retinal
  form
  color
  optics
  motion

Interactions
- 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.

Analysis Types

- When: Temporal Data Analysis + Statistical

- Where: Geospatial Data Analysis

- What: Topical Data Analysis

- With Whom: Network Analysis

Typology of the Data Visualization Literacy Framework

④

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

Data Scales
- nominal
- ordinal
- interval
- ratio

Analyses
- statistical
- temporal
- geospatial
- topical
- relational

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

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

Graphic Variables
- spatial
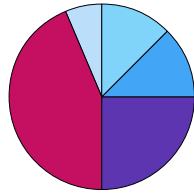 position
- retinal
  form
  color
  optics
  motion

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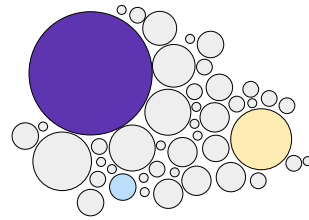- projection
- distortion

Börner, Katy. 2015. *Atlas of Knowledge: Anyone Can Map*. Cambridge, MA: The MIT Press. 30-31.
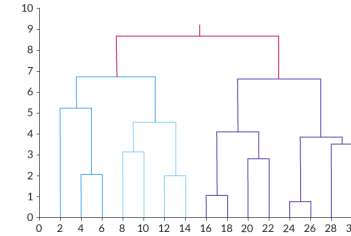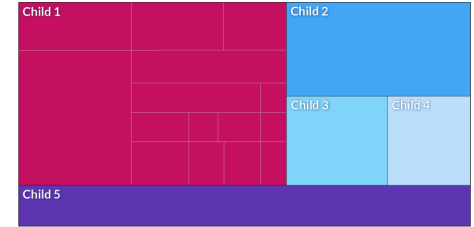
Visualization Types
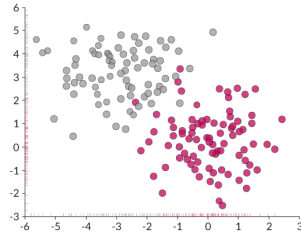
Chart

Pie Chart

Bubble Chart
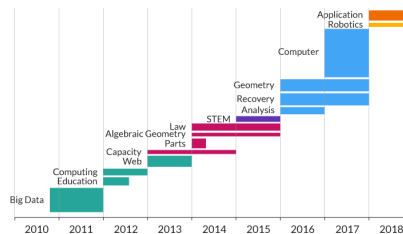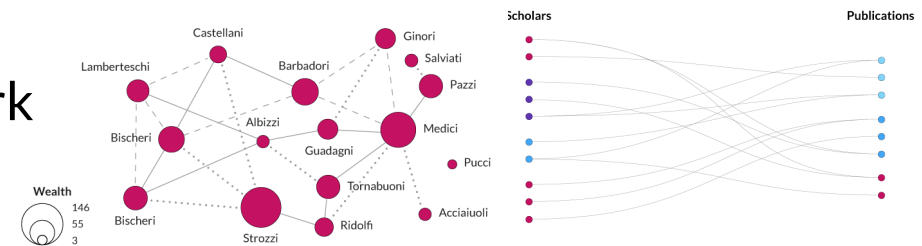
Tree

Dendrogram

Tree Map

Graph

Scatter Graph

Temporal Bar Graph

Network

Dendrogram

Tree Map

Map
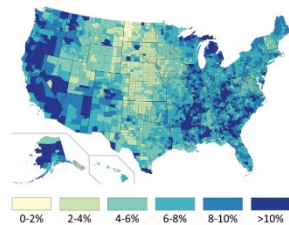
Choropleth Map
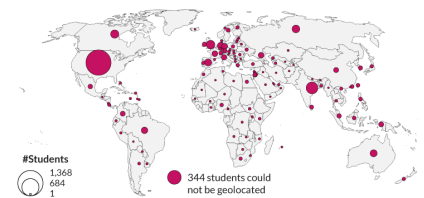
Proportional Symbol Map

Visualize: Reference Systems

4

Visualization Types
- table
- chart
- graph
- map
- network layout

Table
columns by rows
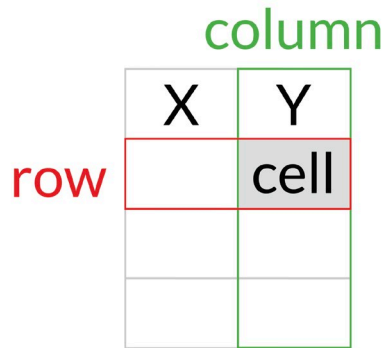
Graph
x-y coordinates

Map
latitude/
longitude

Network
local similarity

Visualize: Reference Systems, Graphic Symbols and Variables

Data Overlay

Graphic Variables 6

Graphic Symbols 5

Reference System

Visualizations 4

Scatter Graph   Geospatial Map   UCSD Science Map   Network

Typology of the Data Visualization Literacy Framework

(5)

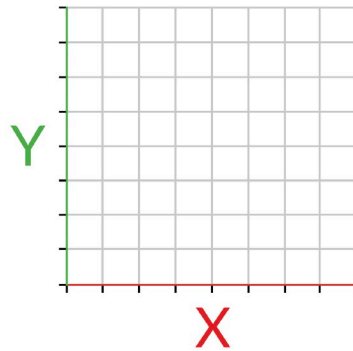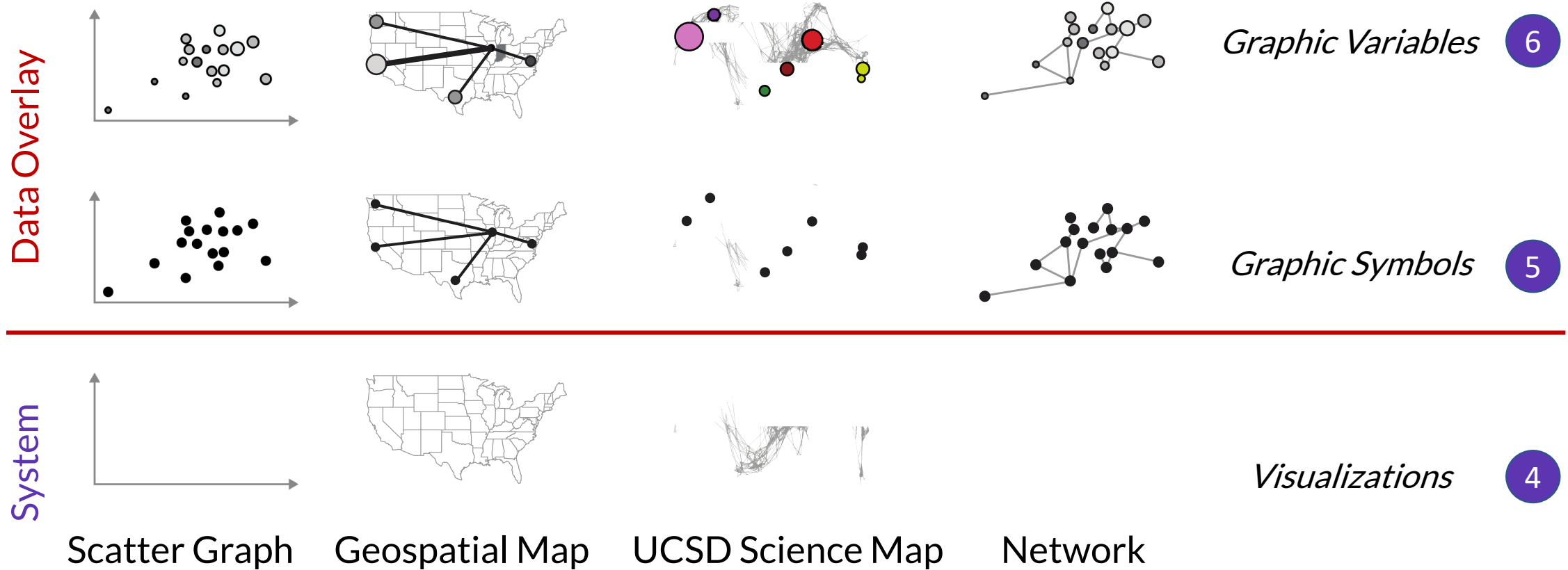| Insight Needs	Data Scales	Analyses	Visualizations	Graphic Symbols	Graphic Variables	Interactions
• categorize/cluster	• nominal	• statistical	• table	• geometric symbols	• spatial	• zoom
• order/rank/sort	• ordinal	• temporal	• chart	point	position	• search and locate
• distributions (also	• interval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

Börner, Katy. 2015. *Atlas of Knowledge: Anyone Can Map*. Cambridge, MA: The MIT Press. 32-33.

Typology of the Data Visualization Literacy Framework

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

Data Scales
- nominal
- ordinal
- interval
- ratio

Analyses
- statistical
- temporal
- geospatial
- topical
- relational

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

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

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

Interactions
- 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. 34-35.

Graphic Variable Types

Position: x, y; possibly z   Quantitative

Form:

• Size   Quantitative

• Shape   Qualitative

• Rotation (Orientation)   Quantitative

Color:

• Value (Lightness)   Quantitative

• Hue (Tint)   Qualitative

• Saturation (Intensity)   Quantitative

Optics: Blur, Transparency, Shading, Stereoscopic Depth

Texture: Spacing, Granularity, Pattern, Orientation, Gradient

Motion: Speed, Velocity, Rhythm

Graphic Variable Types

Position: x, y; possibly z

Form:

- Size

- Shape
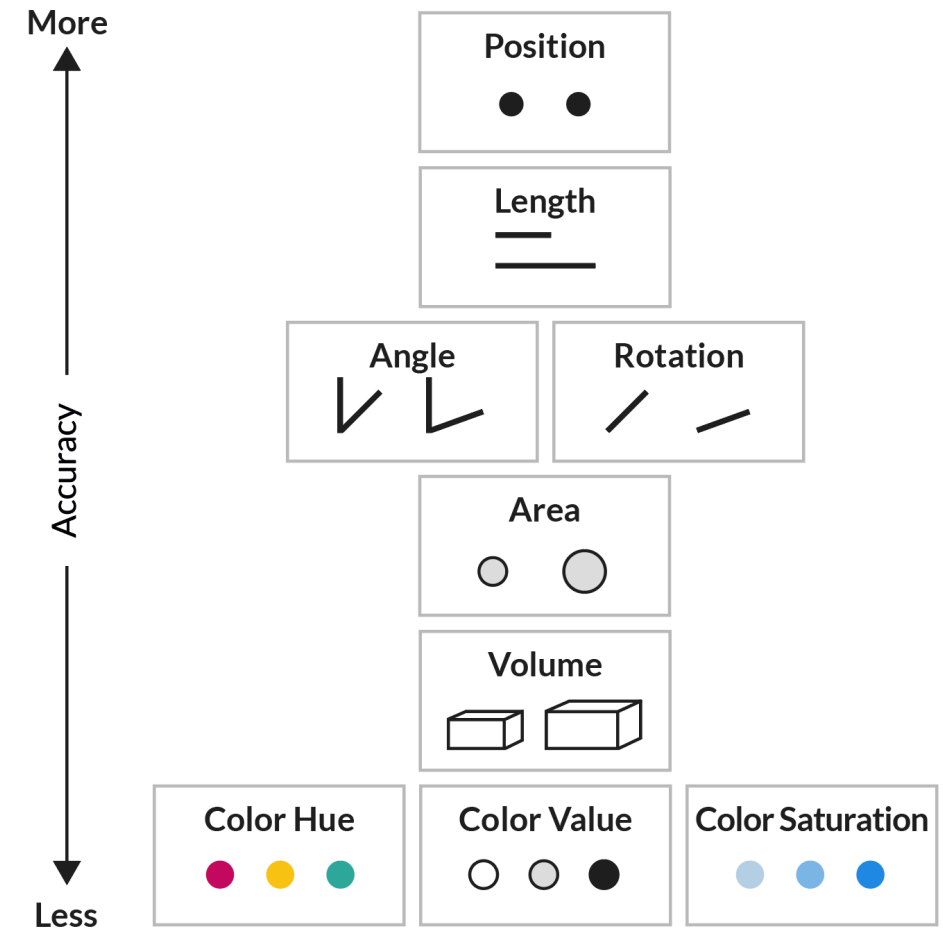
- Rotation (Orientation)

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

			Geometric Symbols		Linguistic Symbols	Pictorial Symbols
			Point	Line		
Spatial	Position	X Y				
Retinal	Form	Size				
		Shape				
	Color	Value				
		Hue			Text Text Text	(alive) (dead)
		Saturation				
	Texture	Granularity				
		Pattern				
	Optics	Blur				
	Motion	Speed				

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

▸ Qualitative

Also called:
Categorical Attributes
Identity Channels

Quantitative

Also called:
Ordered Attributes
Magnitude Channels

31

Graphic Variable Types Versus Graphic Symbol Types

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

Typology of the Data Visualization Literacy Framework

7

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

Börner, Katy. 2015. *Atlas of Knowledge: Anyone Can Map*. Cambridge, MA: The MIT Press. 26, 68-69.

https://visanalytics.cns.iu.edu

RESOURCES / TWITTER

Interdisciplinary Training in
Complex Networks and Systems

THE PROGRAM   RESEARCH   HOW TO APPLY   STUDENTS   COLLOQUIA   NEWS

Share   Tweet   Email

The program

Understanding complex networked systems is key to solving some of the most vexing problems confronting humankind, from discovering how dynamic brain connections give rise to thoughts and behaviors, to detecting and preventing the spread of misinformation or unhealthy behaviors across a population. Graduate training, however, typically occurs in one of two dimensions: experimental and observational methods in a specific area such as biology and sociology, or in general methodologies such as machine learning and data science.

CNS NRT

https://cns-nrt.indiana.edu

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