

# Data Visualization

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



*Guest Lecture in T100 What is Data Science*

December 7, 2020



# Overview

Theoretical data visualization framework (DVL) meant to empower anyone to systematically render data into insights.

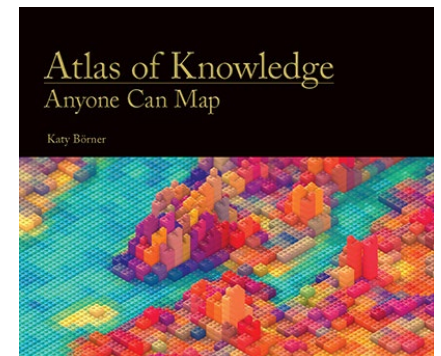
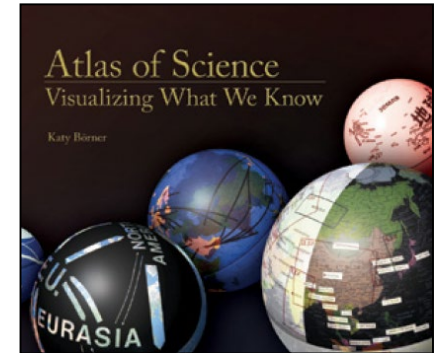
- Börner, Katy, Andreas Bueckle, and Michael Ginda. 2019. [Data visualization literacy: Definitions, conceptual frameworks, exercises, and assessments](#). *PNAS*, 116 (6) 1857-1864.
- Börner, Katy. 2015. [Atlas of Knowledge: Anyone Can Map](#). Cambridge, MA: The MIT Press.
- Börner, Katy. 2010. [Atlas of Science: Visualizing What We Know](#). Cambridge, MA: The MIT Press.

Scaling-Up: Increase global DVL via (in)formal education (AISL, <https://ivmooc.cns.iu.edu> & <https://visanalytics.cns.iu.edu>)

Opportunity: The Human BioMolecular Atlas Program (HuBMAP) (<https://hubmapconsortium.org>)

- Snyder, Michael P., et al. 2019. "[Mapping the Human Body at Cellular Resolution -- The NIH Common Fund Human BioMolecular Atlas Program](#)". *Nature*. 574, p. 187-192.

TOMORROW: Debut 16<sup>th</sup> iteration of the *Places & Spaces: Mapping Science* exhibit (<http://scimaps.org>) at The Mill!



Atlas of Forecasts

# 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.



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



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



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



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

<http://scimaps.org>



Oct 1-Nov 3, 2020: Exhibit on display at the Dimension Mill in Bloomington, IN on <https://dimensionmill.org>

# Places & Spaces: Mapping Science Exhibit

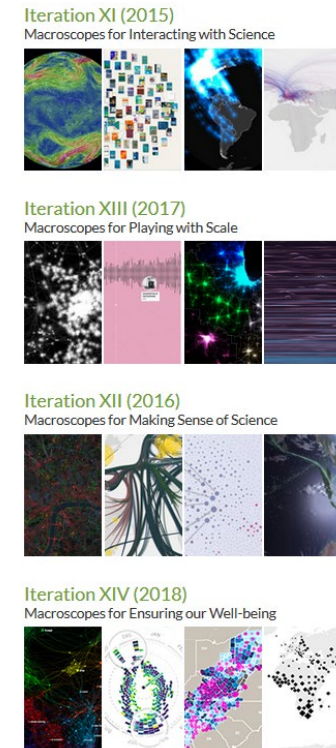
1<sup>st</sup> Decade (2005-2014)

## Maps



2<sup>nd</sup> Decade (2015-2024)

## Macroscopes

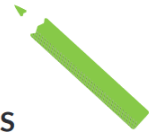


100

MAPS  
in large format, full color, and  
high resolution.

248

MAPMAKERS  
from fields as disparate as art,  
urban planning, engineering,  
and the history of science.



43



MACROSCOPE MAKERS  
including one whose job title is  
“Truth and Beauty Operator.”

20

MACROSCOPES  
for touching all kinds of data.

382

DISPLAY VENUES  
from the Cannes Film Festival  
to the World Economic Forum.

354



PRESS ITEMS  
including articles in *Nature*,  
*Science*, *USA Today*, and *Wired*.

<http://scimaps.org>

# Map of Scientific Collaborations from 2005-2009

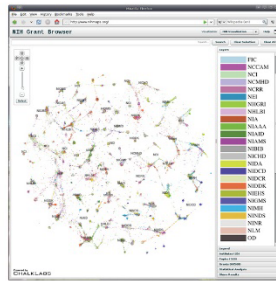


Computed Using Data from Elsevier's Scopus

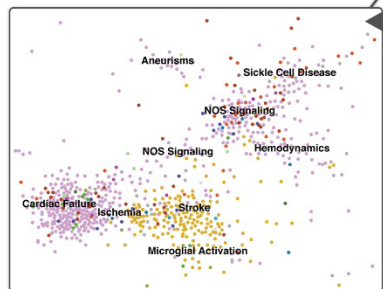
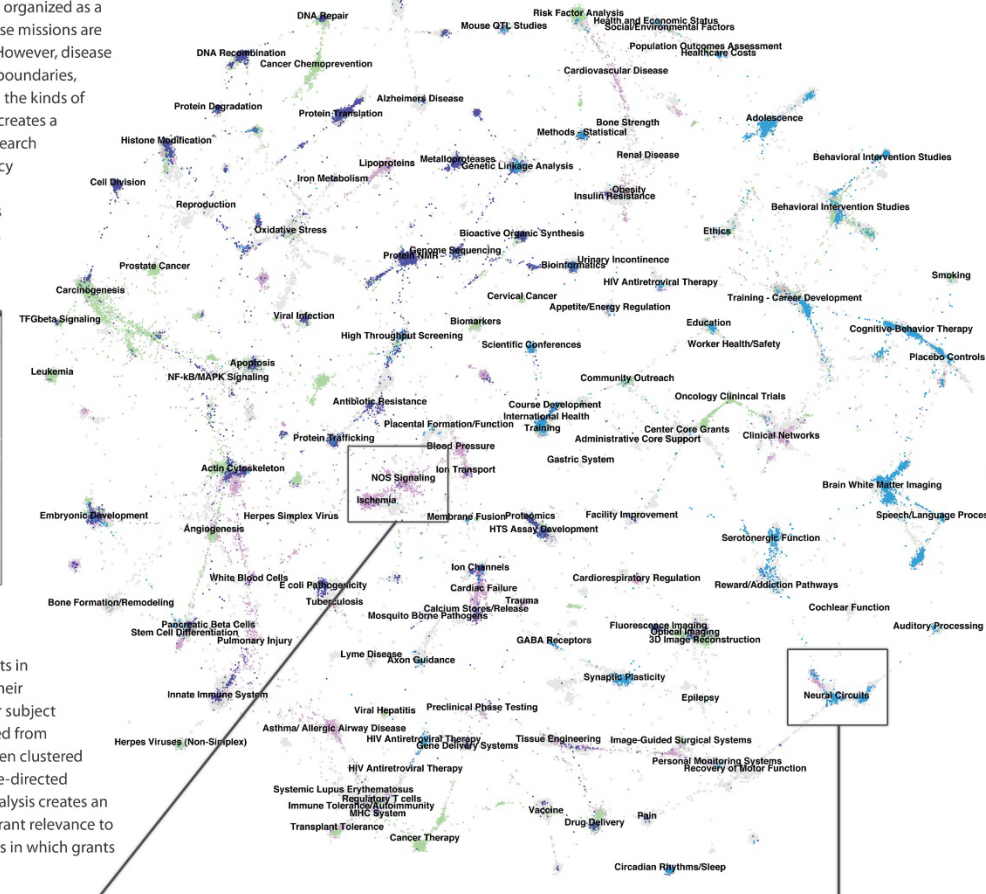
# A Topic Map of NIH Grants 2007

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

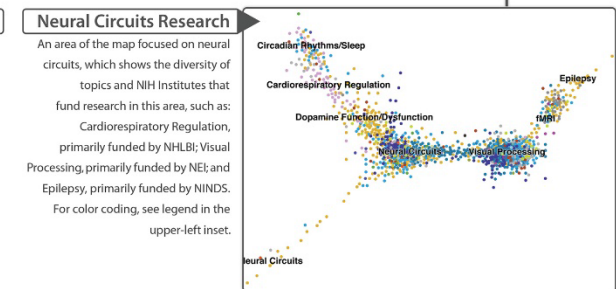
The National Institutes of Health (NIH) is organized as a multitude of Institutes and Centers whose missions are primarily focused on distinct diseases. However, disease etiologies and therapies flout scientific boundaries, and thus there is tremendous overlap in the kinds of research funded by each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



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



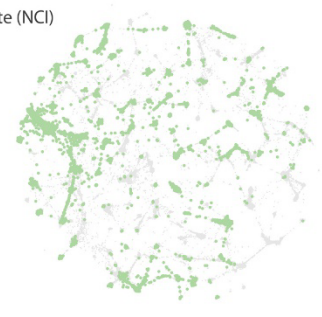
**Cardiac Diseases Research**  
 An area of the map focused on cardiovascular function and dysfunction. Cardiac Failure (primarily funded by NHLBI) is typically clustered next to Stroke (NINDS), since these are the two major medical emergencies associated with ischemia, which results from a restricted blood supply. Also localized in this area are grants focused on Nitric Oxide (NOS) Signaling, a major biochemical pathway for vasodilation, and grants on Hemodynamics, Sickle Cell Disease, and Aneurysms.



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

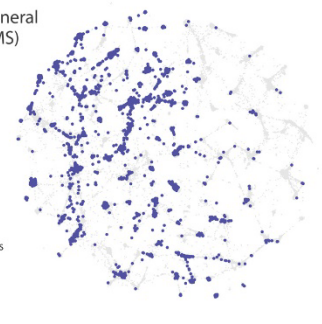
## National Cancer Institute (NCI)

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



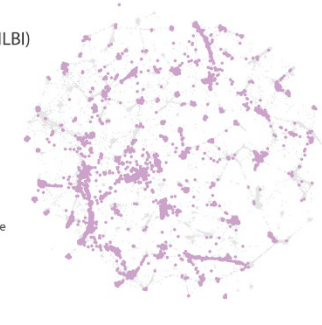
## National Institute of General Medical Sciences (NIGMS)

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



## National Heart, Lung, and Blood Institute (NHLBI)

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



## National Institute of Mental Health (NIMH)

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



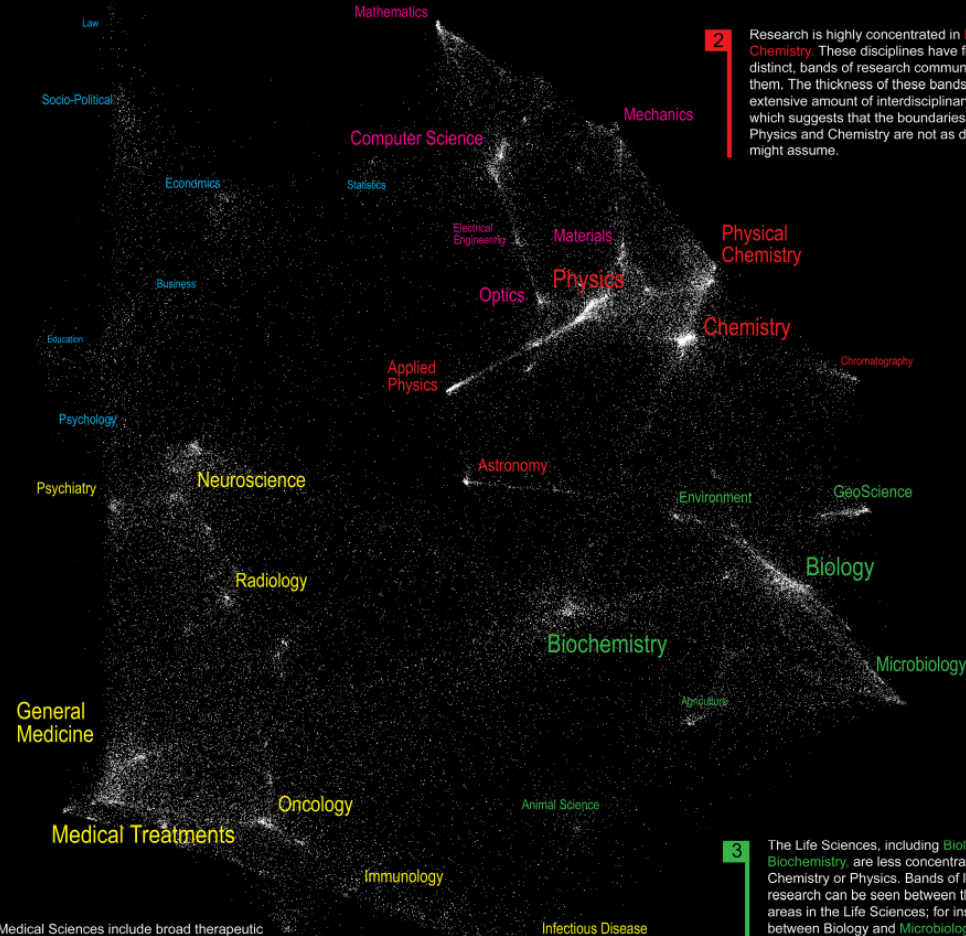


# The Structure of Science

**5** The Social Sciences are the smallest and most diffuse of all the sciences. **Psychology** serves as the link between Medical Sciences (Psychiatry) and the Social Sciences. **Statistics** serves as the link with Computer Science and Mathematics.

**1** **Mathematics** is our starting point, the purest of all sciences. It lies at the outer edge of the map. **Computer Science**, **Electrical Engineering**, and **Optics** are applied sciences that draw upon knowledge in Mathematics and Physics. These three disciplines provide a good example of a linear progression from one pure science (Mathematics) to another (Physics) through multiple disciplines. Although applied, these disciplines are highly concentrated with distinct bands of research communities that link them. Bands indicate interdisciplinary research.

**2** Research is highly concentrated in **Physics** and **Chemistry**. These disciplines have few, but very distinct, bands of research communities that link them. The thickness of these bands indicates an extensive amount of interdisciplinary research, which suggests that the boundaries between Physics and Chemistry are not as distinct as one might assume.



**4** The Medical Sciences include broad therapeutic studies and targeted areas of **Treatment** (e.g. central nervous system, cardiology, gastroenterology, etc.) Unlike Physics and Chemistry, the medical disciplines are more spread out, suggesting a more multi-disciplinary approach to research. The transition into Life Sciences (via Animal Science and Biochemistry) is gradual.

**3** The Life Sciences, including **Biology** and **Biochemistry**, are less concentrated than Chemistry or Physics. Bands of linking research can be seen between the larger areas in the Life Sciences; for instance between Biology and **Microbiology**, and between Biology and **Environmental Science**. Biochemistry is very interesting in that it is a large discipline that has visible links to disciplines in many areas of the map, including Biology, Chemistry, Neuroscience, and General Medicine. It is perhaps the most interdisciplinary of the sciences.

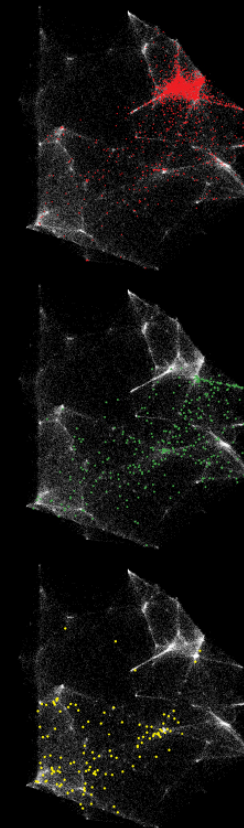
We are all familiar with traditional maps that show the relationships between countries, provinces, states, and cities. Similar relationships exist between the various disciplines and research topics in science. This allows us to map the structure of science.

One of the first maps of science was developed at the Institute for Scientific Information over 30 years ago. It identified 41 areas of science from the citation patterns in 17,000 scientific papers. That early map was intriguing, but it didn't cover enough of science to accurately define its structure.

Things are different today. We have enormous computing power and advanced visualization software that make mapping of the structure of science possible. This galaxy-like map of science (left) was generated at Sandia National Laboratories using an advanced graph layout routine (VxOrd) from the citation patterns in 800,000 scientific papers published in 2002. Each dot in the galaxy represents one of the 96,000 research communities active in science in 2002. A research community is a group of papers (9 on average) that are written on the same research topic in a given year. Over time, communities can be born, continue, split, merge, or die.

The map of science can be used as a tool for science strategy. This is the terrain in which organizations and institutions locate their scientific capabilities. Additional information about the scientific and economic impact of each research community allows policy makers to decide which areas to explore, exploit, abandon, or ignore.

We also envision the map as an educational tool. For children, the theoretical relationship between areas of science can be replaced with a concrete map showing how math, physics, chemistry, biology and social studies interact. For advanced students, areas of interest can be located and neighboring areas can be explored.



## Nanotechnology

Most research communities in nanotechnology are concentrated in **Physics**, **Chemistry**, and **Materials Science**. However, many disciplines in the Life and Medical Sciences also have nanotechnology applications.

## Proteomics

Research communities in proteomics are centered in **Biochemistry**. In addition, there is a heavy focus in the tools section of chemistry, such as **Chromatography**. The balance of the proteomics communities are widely dispersed among the Life and Medical Sciences.

## Pharmacogenomics

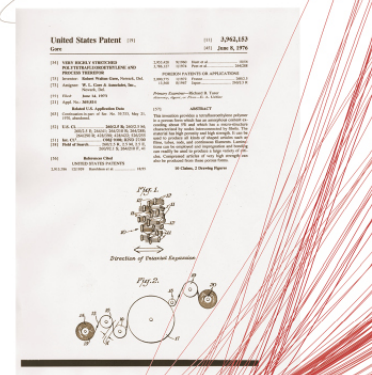
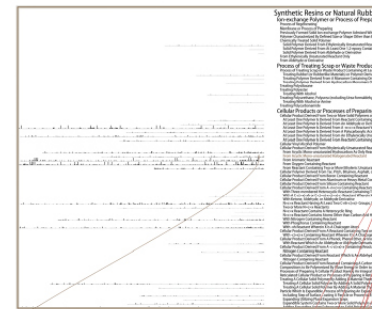
Pharmacogenomics is a relatively new field with most of its activity in **Medicine**. It also has many communities in **Biochemistry** and two communities in the Social Sciences.

# Impact

The United States Patent and Trademark Office does scientists and industry a great service by granting patents to protect inventions. Inventions are categorized in a taxonomy that groups patents by industry or use, proximate function, effect or product, and structure. At the time of this writing there are 160,523 categories in a hierarchy that goes 15 levels deep. We display the first three levels (13,529 categories) at right in what might be considered a textual map of inventions.

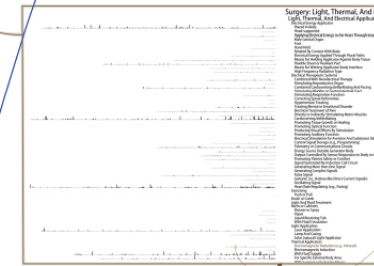
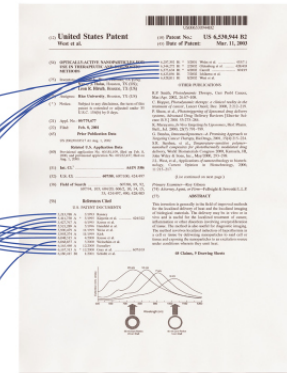
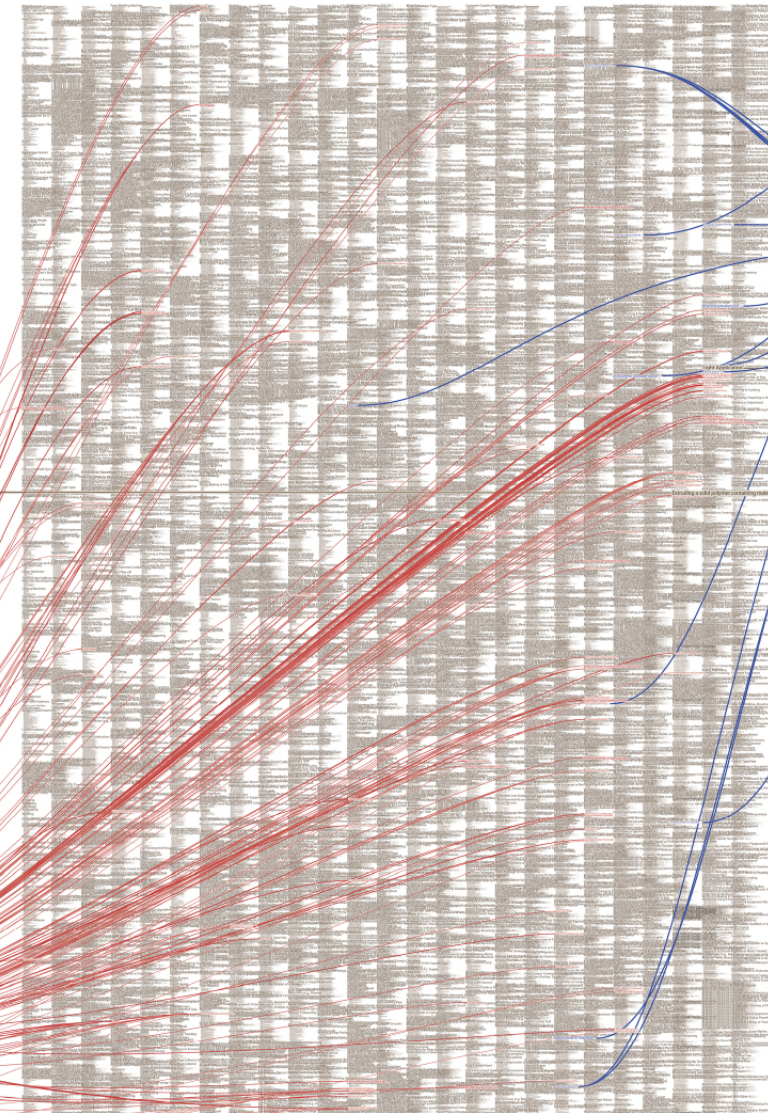
Patent applications are required to be unique and non-obvious, partially by revealing any previous patents that might be similar in nature or provide a foundation for the current invention. In this way we can trace the impact of a single patent, seeing how many patents and categories it affects.

The patent on Goretex—a lightweight, durable synthetic fiber—is an example of one that has had significant impact. The box below enlarges the section of the hierarchy where it is filed, and the red lines (arranged to start along a time line from 1981 to 2006) point to the 130 categories that contain 182 patents, from waterproof clothing to surgical cosmetic implants, that mention Goretex as "prior art."



# The US Patent Hierarchy

# Prior Art



New patents often build on older ideas from many different categories. Here, blue lines originate in the sixteen categories that contain patents cited as prior art for a patent on "gold nanoshells." Gold nanoshells are a new invention: tiny gold spheres (with a diameter ten million times smaller than a human hair) that can be used to make tumors more visible in infrared scans; they have even helped cause complete remission of tumors in tests with laboratory mice. The blue lines show that widely separated categories provided background for this invention.

Keeping categories understandable is an important part of maintaining any taxonomy, including the patent hierarchy. Categories are easier to understand, search, and maintain if they contain elements that comfortably fit the definition of the category. The box above shows tiny bar charts, part of a *Taxonomy Validator* that reveals whether elements fit their categories. Categories may need to be redefined, and sometimes need to be split when they get too vague or large; a problem shared by many classification systems in this information-rich century. But how can we tell which ones to eliminate, add or revise—or how to revise them—in the complex, abstract sociolinguistic spaces we partition into ontologies?

Something as simple as a bar chart helps people see how entities in a category relate to that category. Here, each bar encodes a "distance to prototype": how much each patent differs from an idealized "prototype patent" for that category. A measure like this can be based on statistics, computational linguistics, or even human insight. Thus a category with mostly small bars is a good one, and a generally ragged one needs scrutiny or reorganization; but one that has only two or three tall bars may mean that only those few elements don't belong.

Even simple visuals can make thinking easier by providing better distilled data to the eye: vastly more data than working memory can hold as words. They focus people on exactly the right issues, and support them with the comprehensive overviews they need to make more informed judgements.

# Science related Wikipedian ACTIVITY

This visualization explores the activity of science, math, and technology (SMT) related articles in the English-language Wikipedia (<http://en.wikipedia.org>). The central image shows 659,388 articles (circles). Overlaid is a 37 x 37 grid of relevant half-inch sized images.

Blue, green, and yellow circles represent the 3,599 math, 6,474 science, and 3,164 technology related articles respectively. The larger the size of a circle the higher the likelihood it is that type of article. The four corners show activity patterns of SMT articles.

**Article Edit Activity**  
Articles are size coded based on how frequently they have been edited from Feb. 6, 2001 to April 6, 2007. More consideration is given to current and major edits. Larger circles have been edited more frequently than smaller circles.

**2007 Major Edits**  
Articles are size coded based on how many major edits they received from January 1st, 2007 to April 6th, 2007. Larger circles have received more edits than smaller circles. The highest number of major edits was 2,627.

For the central image, each article is size coded based on the likelihood that it is math, science, or technology related.

- 0%
- 50%
- 100%

All five images are color coded based on type. Transparency is used for legibility, and creates different colors when nodes overlap.

- Math
- Science
- Technology

**Article Popularity**  
Articles are size coded based on the number of Wikipedia articles referencing it. Larger circles are receiving more links from other articles than smaller circles. The highest number of references to an article was 142,602.

**Number of Bursts**  
Articles are size coded based on the number of bursts, i.e. sudden increases, of edit activity that occurred during the article's lifetime. Larger circles have had more bursts in activity than smaller circles. The most bursts an article had was 9.



# Diseasome

## The Human Disease Network

Explore online at <http://diseasome.eu>

### Statistics

# of Nodes: 516  
 # of Edges: 1188  
 Density: 0,0089  
 Average Degree: 9,20  
 Diameter: 15  
 Average Shortest Path: 6,5

### Top 5 Diseases

1. Deafness
2. Leukemia
3. Colon Cancer
4. Retinitis Pigmentosa
5. Diabetes Mellitus

### Top 5 Genes

1. TP53
2. PAK6
3. FGFR2
4. RTN
5. MSH2

### Description

The map presents a network of 516 diseases linked by 1188 known disorder-gene associations, indicating the common genetic origin of many diseases.

#### GENE NETWORK CLUES

The map offers a rapid visual reference of the genetic links between disorders and a valuable global perspective for physicians, genetic counselors, and biomedical researchers alike. This view appears only when the network is zoomed, revealing to their associated genes, whether the understanding of the roots of disease, and the functions of particular genes.

#### NETWORK VISUALIZATION TECHNIQUES APPLIED

The map was done using the force-directed layout algorithm ForceAtlas in Gephi. Node sizes correspond to the number of genes to which the disease belongs, and the size is proportional to its node degree, the overall number of links. Link's width is proportional to the number of genes that are implicated in both diseases and colored with the average color between source and target nodes. Isolated diseases are not shown and only the giant component has been kept. The Clusters Mini-map shows most remarkable clusters and shows largest visual clusters.

The Disorder Class Interactivity graph below shows the interaction level between disorder classes, representing the number of shared genes, up to 80.

#### References

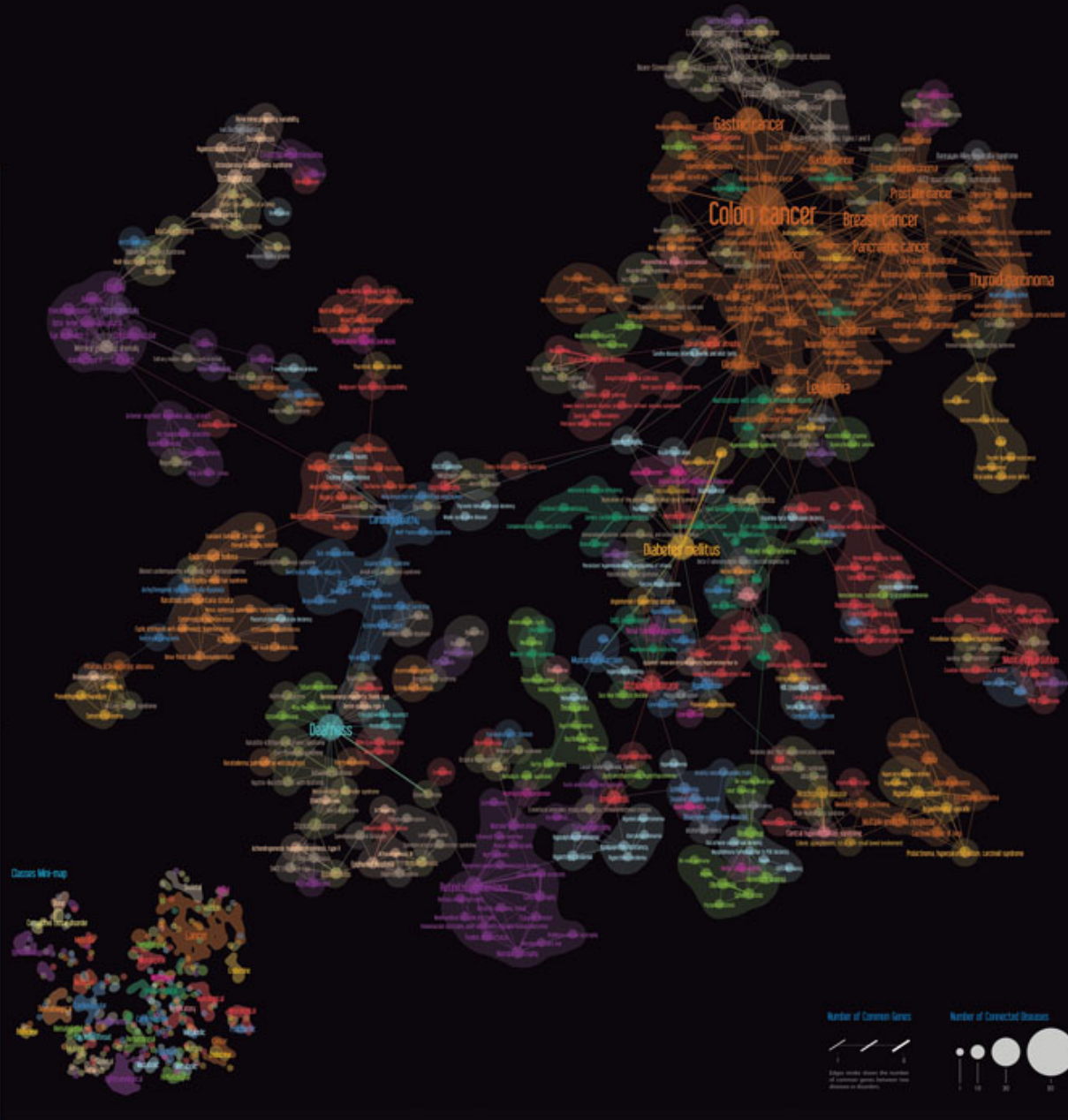
The Human Disease Network  
 Bastin & Heymann 2009, *PLoS ONE*, 4(10): e7000

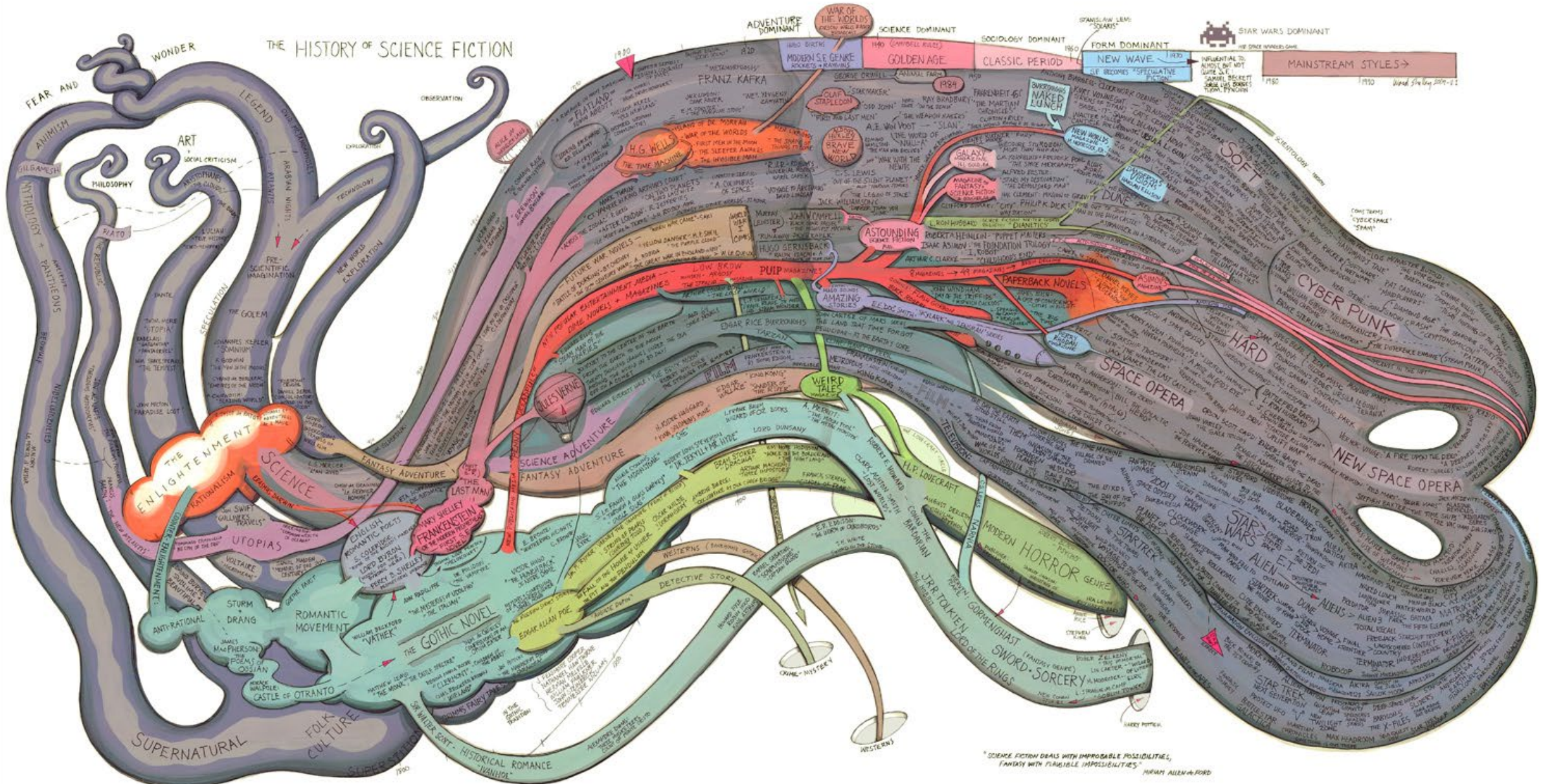
### Disorder Class Interactions



### Disorder Class

- Cancer
- Endocrine
- Ear, Nose, Throat
- Ophthalmological
- Neurological
- Hematological
- Cardiovascular
- Muscular
- Immunological
- Dermatological
- Nutritional
- Connective Tissue Disorder
- Renal
- Psychiatric
- Metabolic
- Bone
- Skeletal
- Developmental
- Gastrointestinal
- Respiratory
- Multiple
- Unclassified





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

# Check out our **Zoom Maps** online!



VII.10  
History of Science Fiction, by Ward Shulley

BROOKLYN, NY, 2011  
Courtesy of Ward Shulley Studio

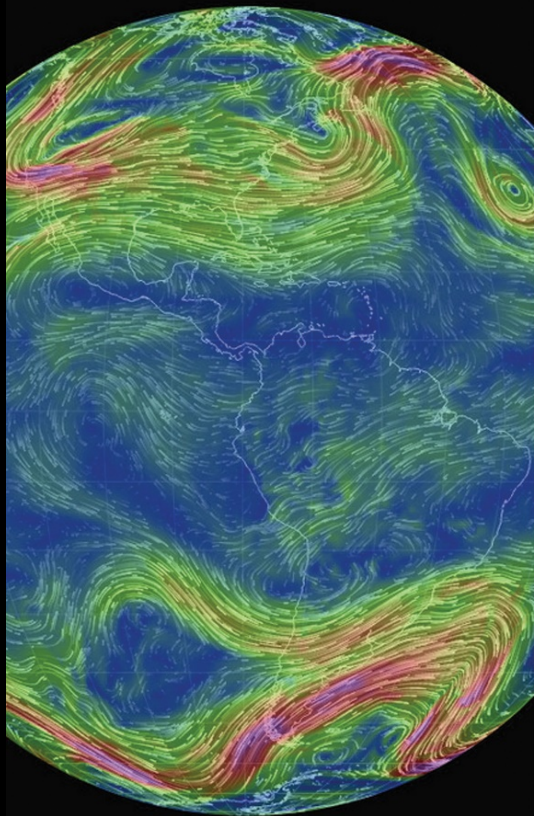
Ward Shulley is an artist identified with the Williamsburg scene in Brooklyn, New York. He is a writer, a painter, and a filmmaker. This map plots the science fiction literary genre from its nascent beginnings in the late 18th century, through the Victorian era, to the modernist and postmodernist eras. Emerging out of the data, here the narrative structure perceives and organizes the data into a form that is both informative and aesthetically pleasing. The map's structure is like a tree, with roots in the past and branches that spread out to the present. The map's structure is like a tree, with roots in the past and branches that spread out to the present. The map's structure is like a tree, with roots in the past and branches that spread out to the present.

PLACES & SPACES  
MAPPING & DESIGN

Visit [scimaps.org](http://scimaps.org) and check out all our maps in stunning detail!



# MACROSCOPES FOR INTERACTING WITH SCIENCE



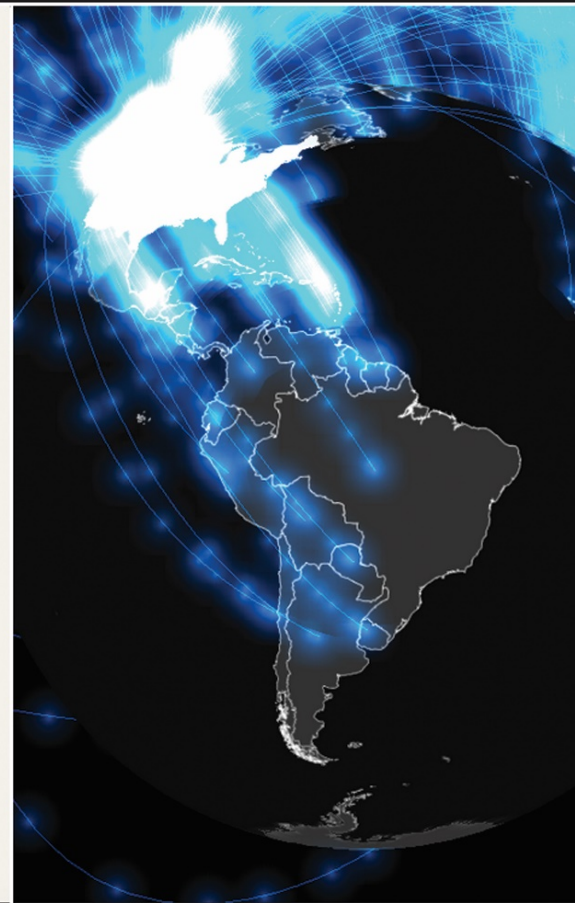
**Earth**

*Weather on a worldwide scale*



**AcademyScope**

*Exploring the scientific landscape*



**Mapping Global Society**

*Local news from a global perspective*

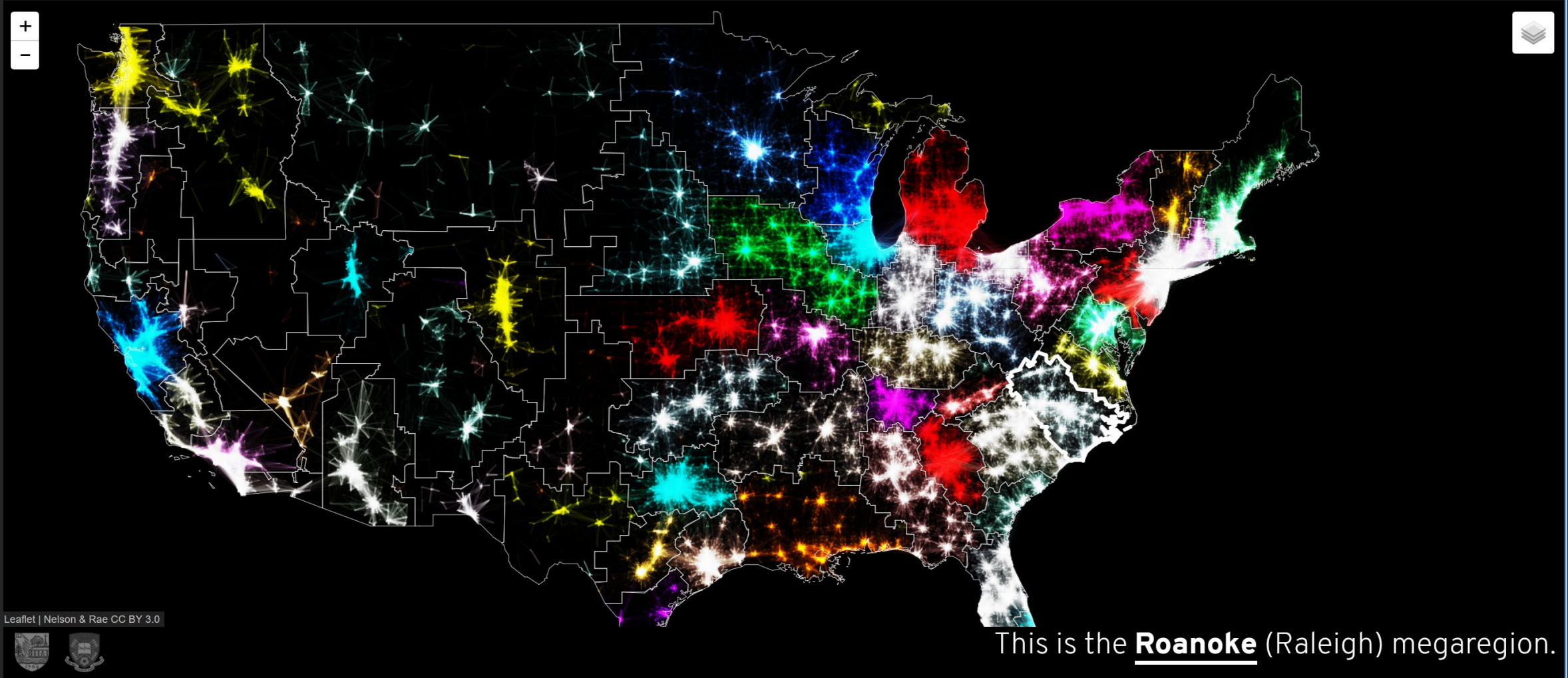


**Charting Culture**

*2,600 years of human history in 5 minutes*

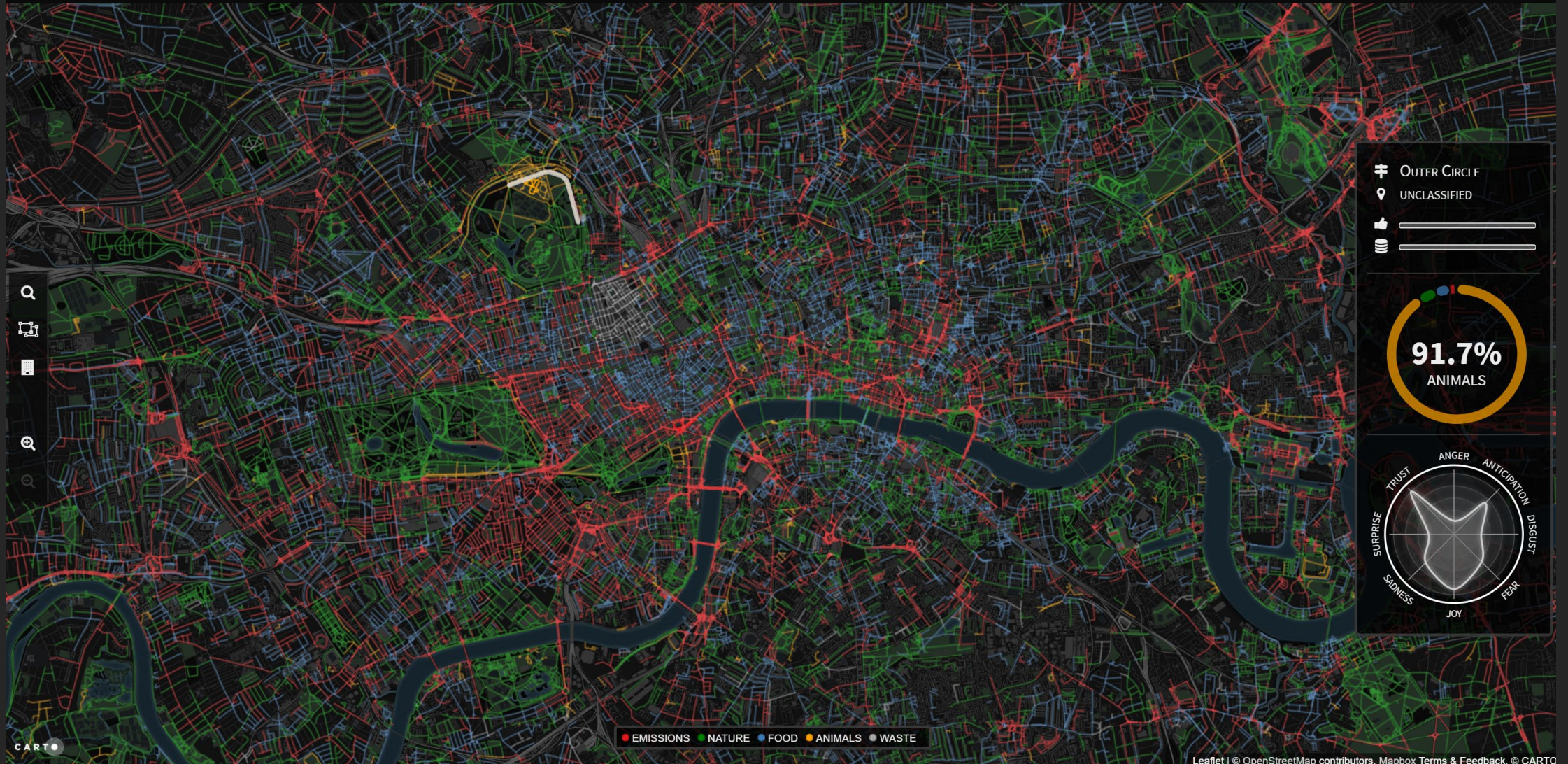
# THE MEGAREGIONS OF THE US

Explore the new geography of commuter connections in the US.  
Tap to identify regions. Tap and hold to see a single location's commuted.





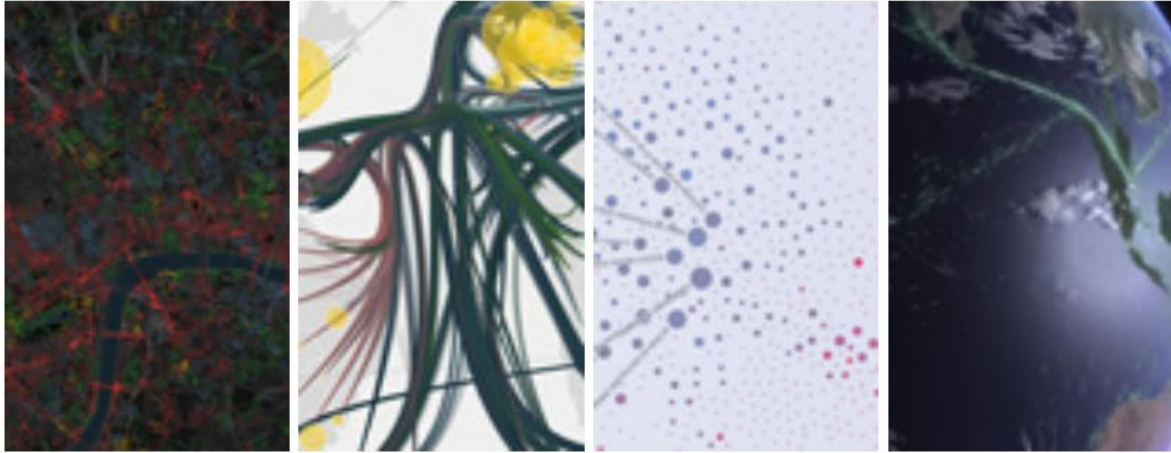
SMELLY  
MAPS



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

## Iteration XII (2016)

Macrosopes for Making Sense of Science



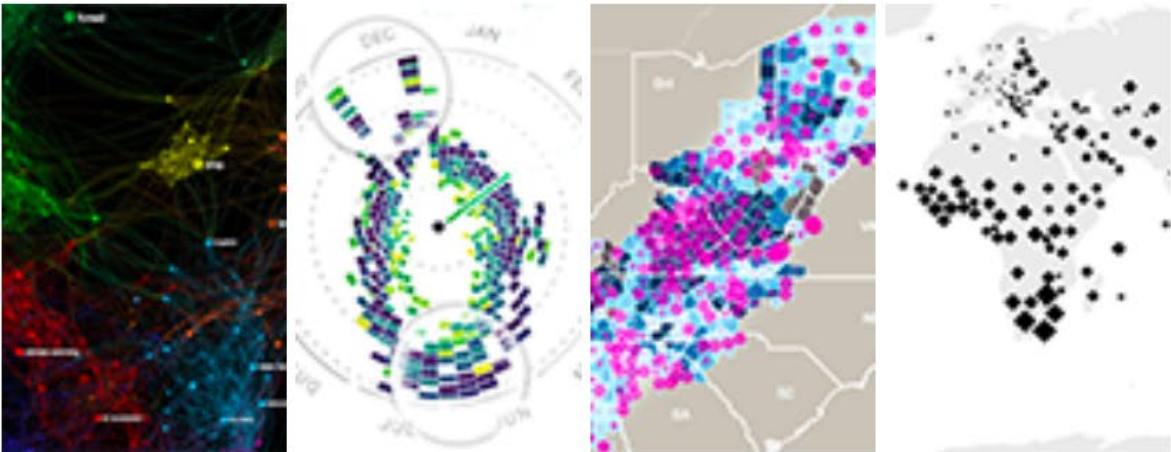
## Iteration XIII (2017)

Macrosopes for Playing with Scale



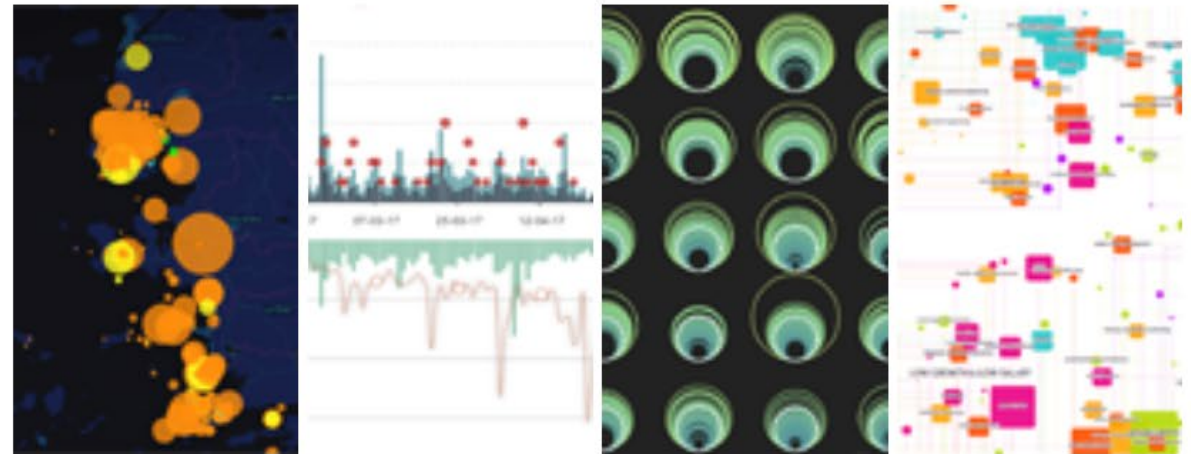
## Iteration XIV (2018)

Macrosopes for Ensuring our Well-being



## Iteration XV (2019)

Macrosopes for Tracking the Flow of Resources



# Acknowledgments

## Exhibit Curators



The exhibit team: Lisel Record, Katy Börner, and Todd Theriault.

<http://scimaps.org>

Plus, we thank the more than 250 authors of the 100 maps and 16 interactive macroscopes.

## Exhibit Advisory Board



**Gary Berg-Cross**

Cognitive psychologist (PhD, SUNY-Stony Brook). Potomac, MD, USA



**Donna J. Cox, MFA, Ph.D.**

Director of the **Advanced Visualization Laboratory** at the National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, IL, USA



**Bonnie DeVarco**

**Media X** Distinguished Visiting Scholar at Stanford University, Palo Alto, CA, USA



**Peter A. Hook**

Head of Digital and Scholarly Services and LawArXiv Administrator, Cornell Law Library, Ithaca, NY, USA



**Francis Harvey**

Professor of Visual Communication in Geography at the Leibnitz Institute for Regional Geography, Leipzig University, Germany



**Lev Manovich**

Professor, **The Graduate Center**, City University of New York; Director, **Software Studies Initiative** (big data, digital humanities, visualization)



**André Skupin**

Associate Professor of Geography at San Diego State University, California



**Moritz Stefaner**

Freelance designer on the crossroads of data visualization, information aesthetics, and user interface design in Germany



**Olga Subirós**

Curator of Big Bang Data and Founder of **Olga Subirós Studio** in Barcelona, Spain



**Stephen Uzzo**

Vice President of Science and Technology for the **New York Hall of Science**



**Benjamin Wiederkehr**

Founding Partner and Managing Director of **Interactive Things** in Zürich, Switzerland

# 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:

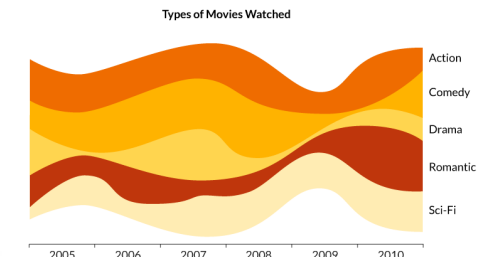
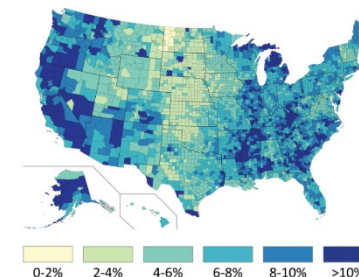
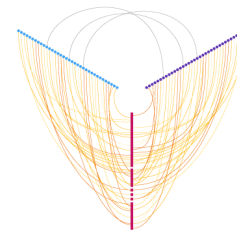
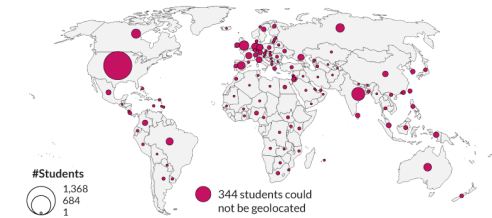
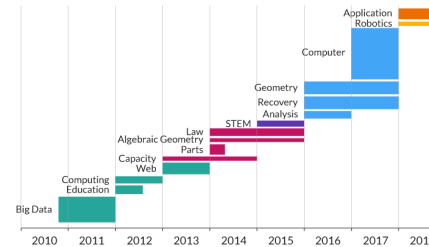
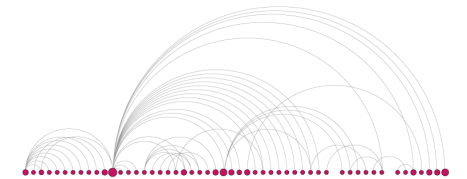
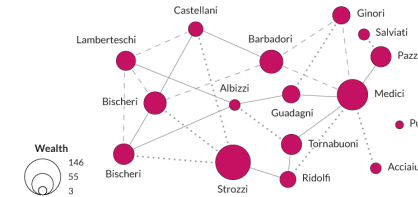
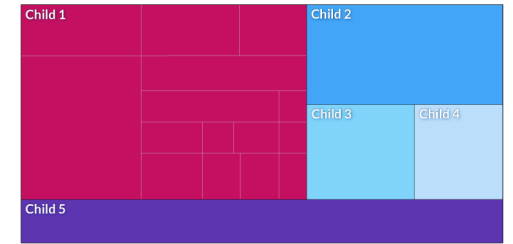
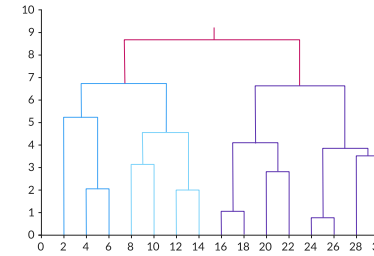
- literacy (ability to read and write text in titles, axis labels, legends, etc.),
- visual literacy (ability to find, interpret, evaluate, use, and create images and visual media), and
- mathematical literacy (ability to formulate, employ, and interpret math in a variety of contexts).

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

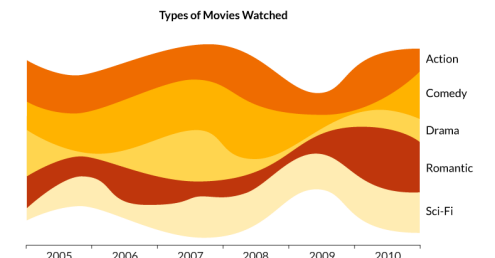
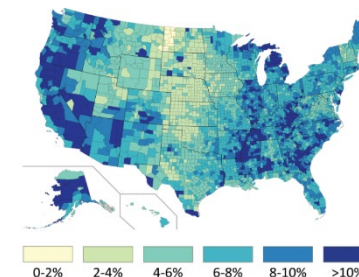
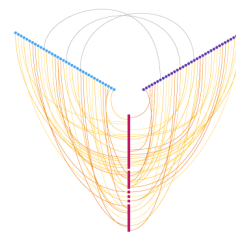
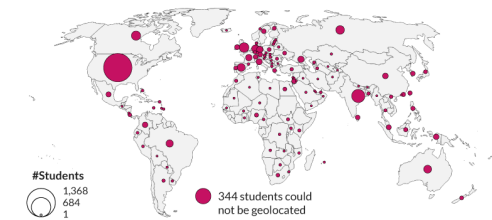
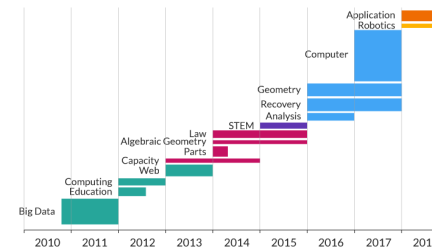
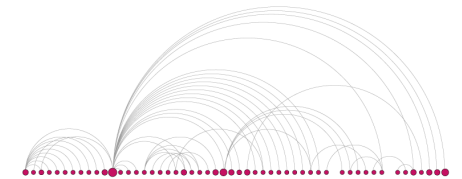
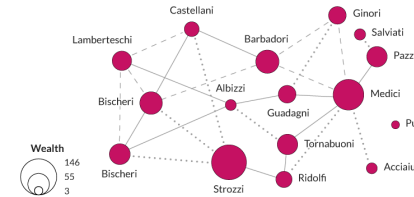
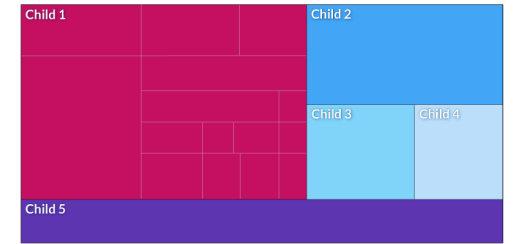
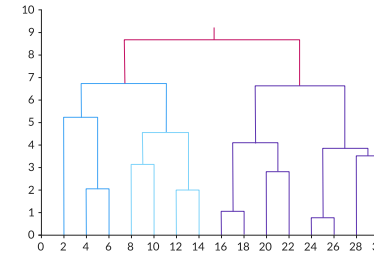
- help organize and manage the evolving zoo of 500+ different data visualization types,
- provide guidance when designing data visualizations, and
- facilitate teaching.



# 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

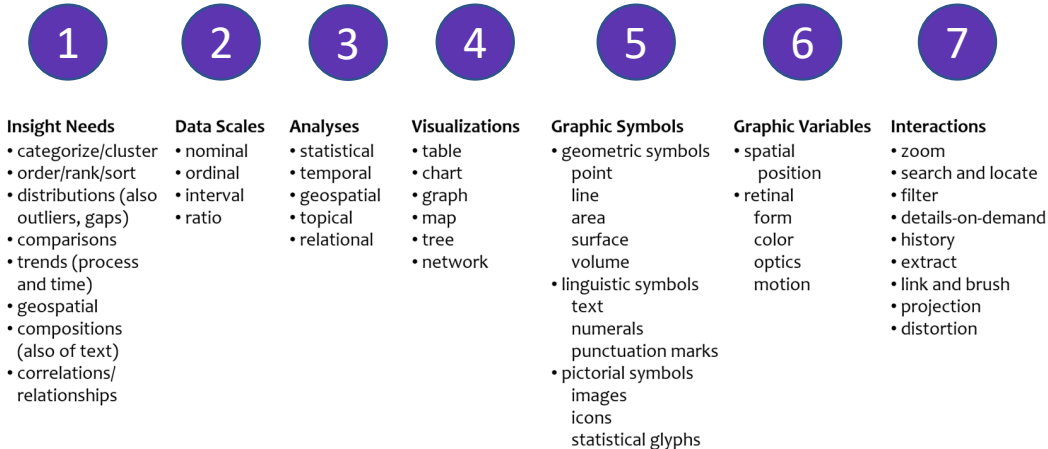
- The initial DVL-FW was developed via an extensive literature review.
- The resulting DVL-FW typology, process model, exercises, and assessments were then tested in the *Information Visualization* course taught for more than 17 years at Indiana University. More than 8,500 students enrolled in the IVMOOC 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:

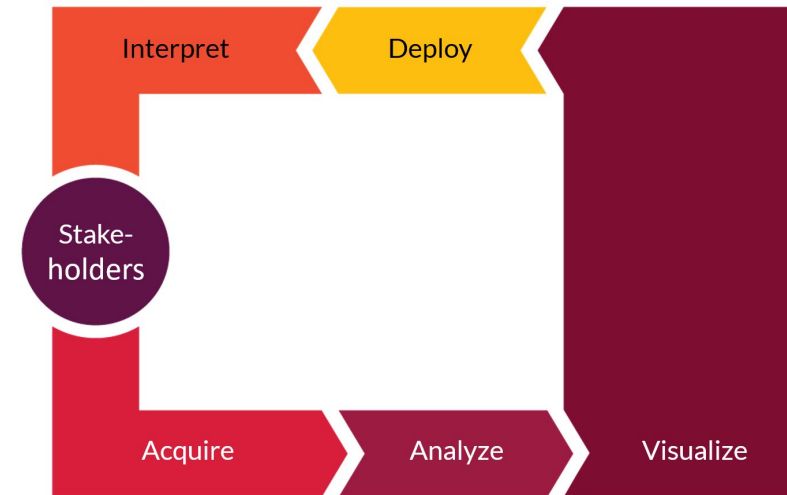
## DVL Typology

Defines 7 types with 4-17 members each.



## DVL Workflow Process

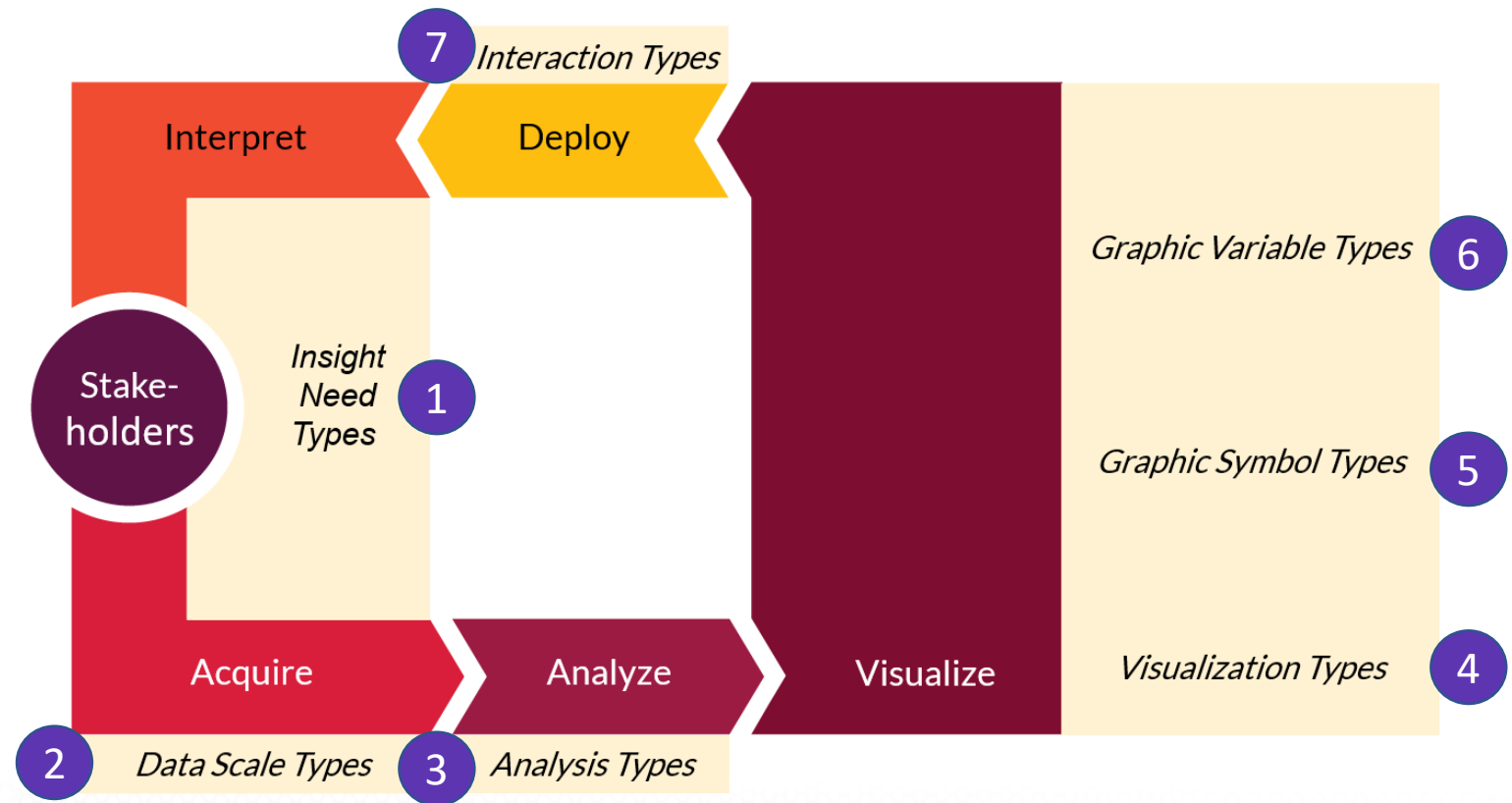
Defines 5 steps required to render data into insights.



# 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.

The screenshot shows the Make-A-Vis interface with three main sections: Data, Make Visualization, and a visualization preview.

**Data Section:**

- ISI Publications: (CSV) Preprocessed-wos**

Title	Authors	Journal	Year	#Cites
[Progress bar]				

Total Records: 562
- Journals: (from ISI Publications)**

Name	#Papers	#Cites	First Year	Last Year
BMC EVOL BIOL	1	7	2006	2006
FEBS J	2	0	2005	2005
NAT PHYS	3	18	2005	2006

Total Records: 562

**Make Visualization Section:**

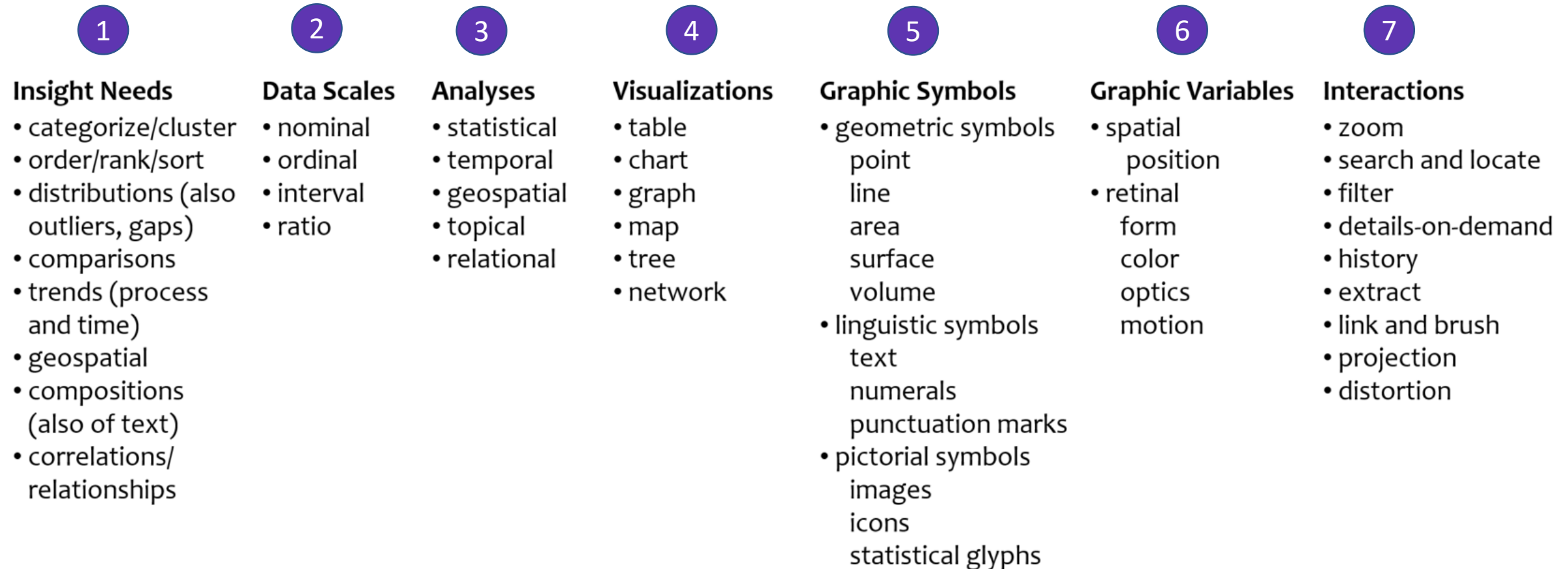
- Select Visualization Type:** Scatter Graph, **Temporal Bar Graph**, Geomap, Scimap. A blue circle with the number '4' is next to this section.
- Select Graphic Symbol Type(s):** (Dropdown menu)
- Select Graphic Variable Types:** (Dropdown menu)

**Temporal Bar Graph Preview:**

Temporal Bar Graph [Close] [Add] [Edit]

Year	Machine	Big Data	Education	Building	Making	Computing	Web	Form	Smart	Capacity	Algebraic Geometry	Parts	Law	Stem	Analysis	Recovery	Geometry	Computer	Application	Robotics	
1998	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
2000	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
2002	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
2004	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
2006	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
2008	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
2010	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
2012	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
2014	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
2016	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
2017	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low

# Typology of the Data Visualization Literacy Framework



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

- 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. 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 Literacy Framework

2

## 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. 28-29.



# Data Scale Types

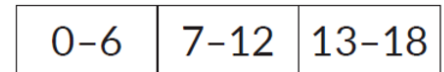
**Nominal:** A categorical scale, also called a nominal or category scale, is **qualitative**. Categories are assumed to be non-overlapping.



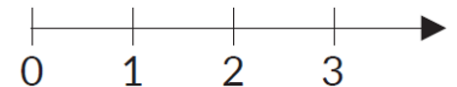
**Ordinal:** An ordinal scale, also called sequence or ordered, is **quantitative**. It rank-orders values representing categories based on some intrinsic ranking, but not at measurable intervals.



**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				
Stevens, 1946 <i>Scales of Measurement</i>	Bertin, 1967 <i>Level of Organization of the Components</i>	Harris, 1996 <i>Classification of Scales</i>	Munzner, 2011 <i>Visualization Principles</i>	Börner, 2014 <i>Data Scale Types</i>
nominal	quantitative	category	categorical/nominal	nominal
ordinal	ordered	sequence	ordinal	ordinal
interval	quantitative	quantitative	quantitative	interval
ratio	quantitative	quantitative	quantitative	ratio

# 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.







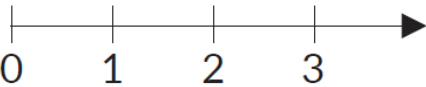
Quantitative

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

**Ratio:** Physical measures such as height, weight, (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	<table border="1" data-bbox="1396 1043 1819 1115"> <tr> <td>0-6</td> <td>7-12</td> <td>13-18</td> </tr> </table>	0-6	7-12	13-18
0-6	7-12	13-18							
Ratio	y	y	y	y	geometric mean				

Qualitative

Quantitative

# Typology of the Data Visualization Literacy Framework

3

## 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

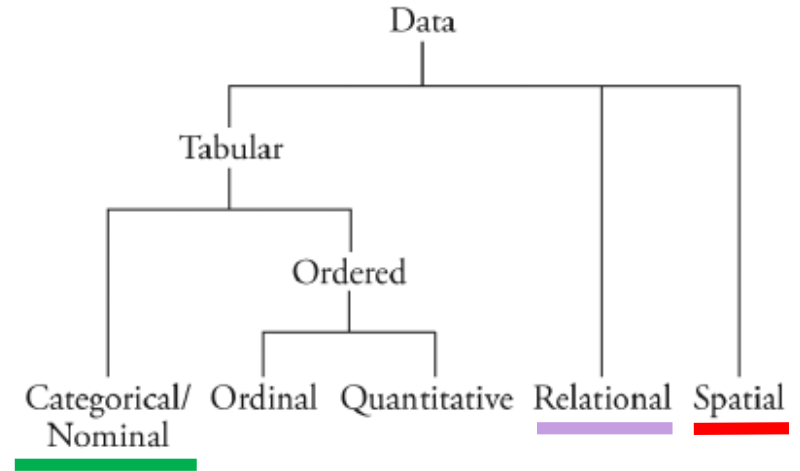
- zoom
- search and locate
- filter
- details-on-demand
- history
- extract
- link and brush
- projection
- distortion

Börner, Katy. 2015. [Atlas of Knowledge: Anyone Can Map](#). Cambridge, MA: The MIT Press. 25.

# Analysis Types

- When: Temporal Data Analysis + Statistical
- Where: Geospatial Data Analysis
- What: Topical Data Analysis
- With Whom: Network Analysis

*Data Hierarchy* by Tamara Munzner distinguishes tabular, relational, and spatial data.



# 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

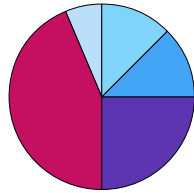
## Interactions

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

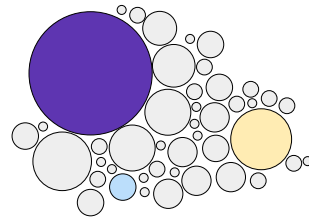
Börner, Katy. 2015. [Atlas of Knowledge: Anyone Can Map](#). Cambridge, MA: The MIT Press. 30-31.

# Visualization Types

Chart

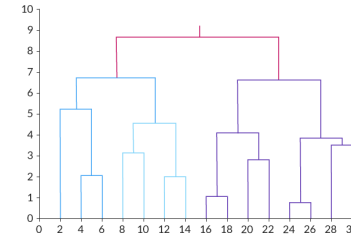


*Pie Chart*

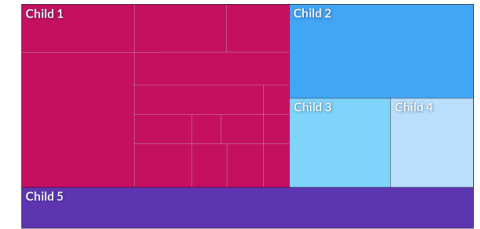


*Bubble Chart*

Tree

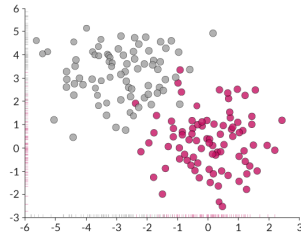


*Dendrogram*

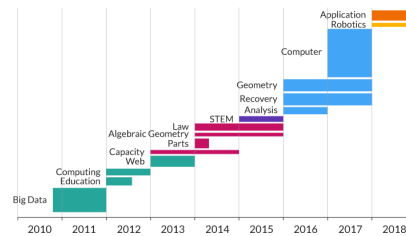


*Tree Map*

Graph

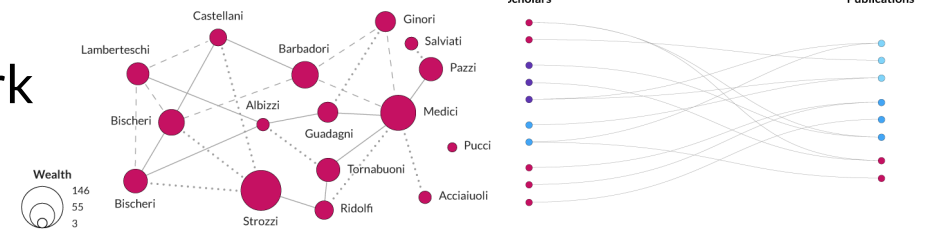


*Scatter Graph*



*Temporal Bar Graph*

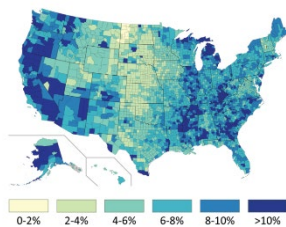
Network



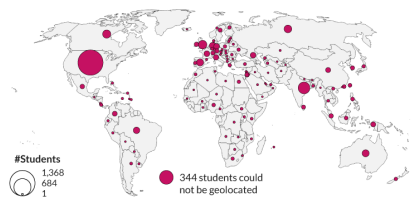
*Force-Directed Network Layout*

*Bimodal Network Layout*

Map



*Choropleth Map*

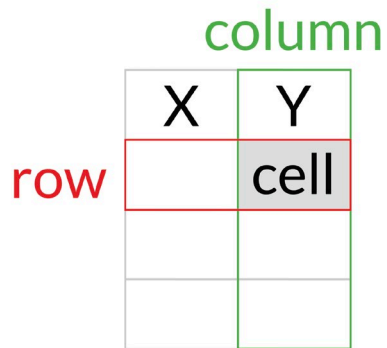


*Proportional Symbol Map*

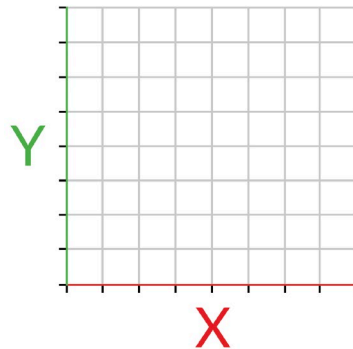


# Visualize: Reference Systems

**Table**  
columns by rows



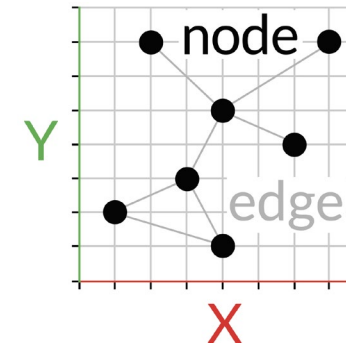
**Graph**  
x-y coordinates



**Map**  
latitude/  
longitude



**Network**  
local similarity

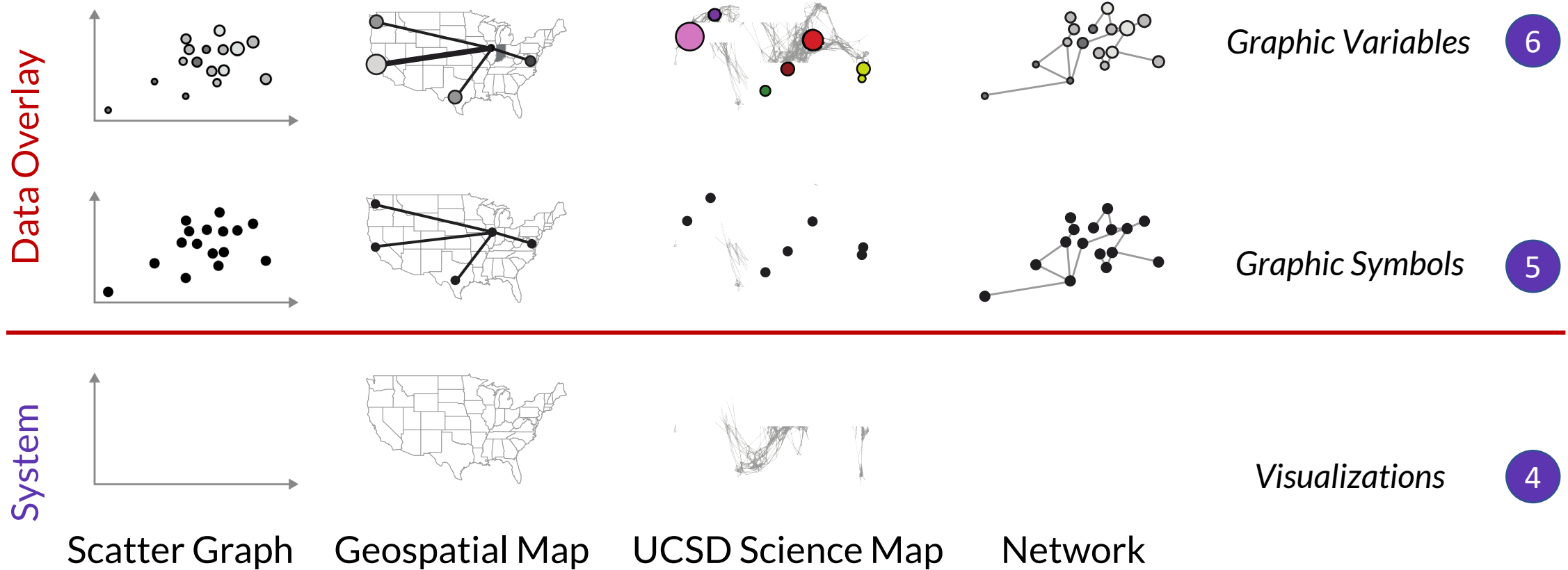


4

**Visualization Types**

- table
- chart
- graph
- map
- network layout

# Visualize: Reference Systems, Graphic Symbols and Variables



# Typology of the Data Visualization Literacy Framework

5

## 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. 32-33.

# Typology of the Data Visualization Literacy Framework

6

## 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
- Shape
- Rotation (Orientation)

Quantitative

Qualitative

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:**

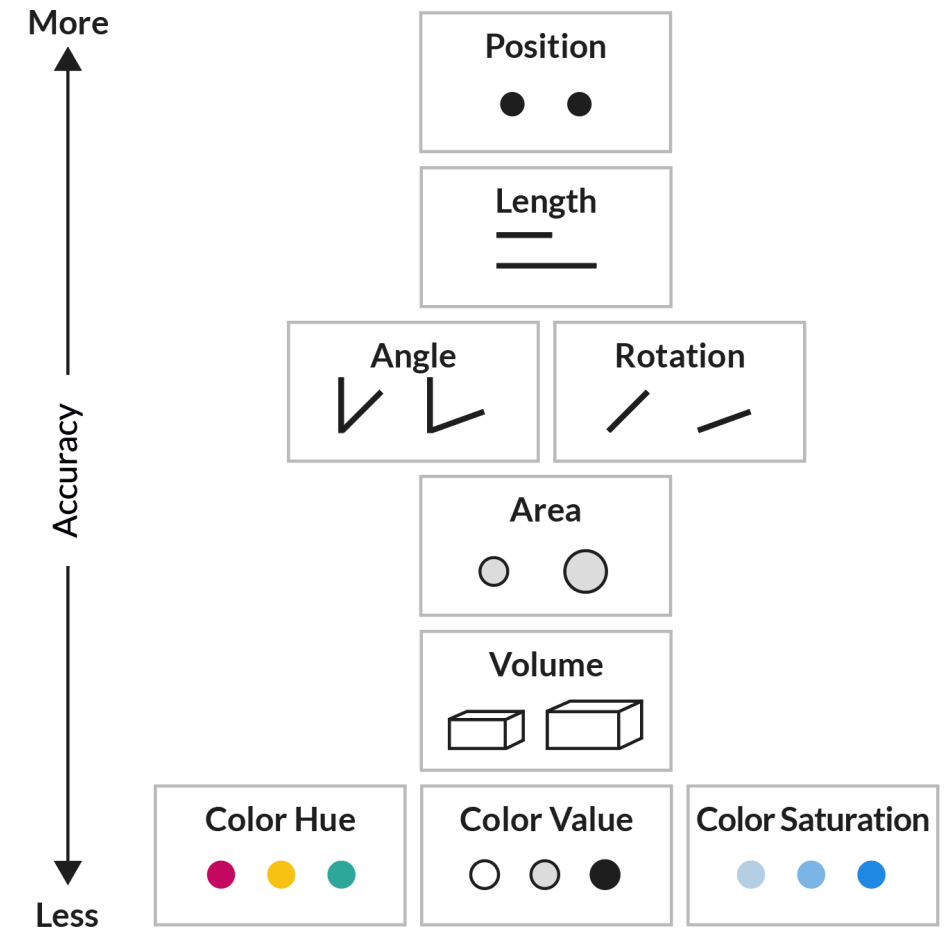
- Value (Lightness)
- Hue (Tint)
- Saturation (Intensity)



**Optics:** Blur, Transparency, Shading, Stereoscopic Depth

**Texture:** Spacing, Granularity, Pattern, Orientation, Gradient

**Motion:** Speed, Velocity, Rhythm



# Graphic Symbol Types

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

Graphic Variable Types

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

**Qualitative**

Also called:  
Categorical Attributes  
Identity Channels

**Quantitative**

Also called:  
Ordered Attributes  
Magnitude Channels

# Graphic Variable Types Versus Graphic Symbol Types

			Geometric Symbols					Linguistic Symbols Text, Numerals, Punctuation Marks					Pictorial Symbols Images, Icons, Statistical Glyphs					
			Point	Line	Area	Surface	Volume											
Spatial	x	quantitative																
	y	quantitative																
	z	quantitative																
Retinal	Form	Size	quantitative	NA (Not Applicable)														
		Shape	qualitative	NA														
		Rotation	quantitative	NA														
		Curvature	quantitative	NA														
	Angle	quantitative	NA															
	Closure	quantitative	NA															
	Value	quantitative																
	Color	Hue	qualitative															
Saturation	quantitative																	
Retinal	Texture	Spacing	quantitative															
		Granularity	quantitative															
		Pattern	qualitative															
		Orientation	quantitative	NA														
		Gradient	quantitative															
	Optics	Blur	quantitative															
		Transparency	quantitative															
		Shading	quantitative															
	Motion	Stereoscopic Depth	quantitative	Point in foreground .. background	Line in foreground .. background	Area in foreground .. background	Surface in foreground .. background	Volume in foreground .. background	Text in foreground .. background					Icons in foreground .. background				
		Speed	quantitative															
Velocity		quantitative																
Rhythm	quantitative	Blinking point slow .. fast	Blinking line slow .. fast	Blinking area slow .. fast	Blinking surface slow .. fast	Blinking volume slow .. fast	Blinking text slow .. fast					Blinking icons slow .. fast						

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



# Typology of the Data Visualization Literacy Framework

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

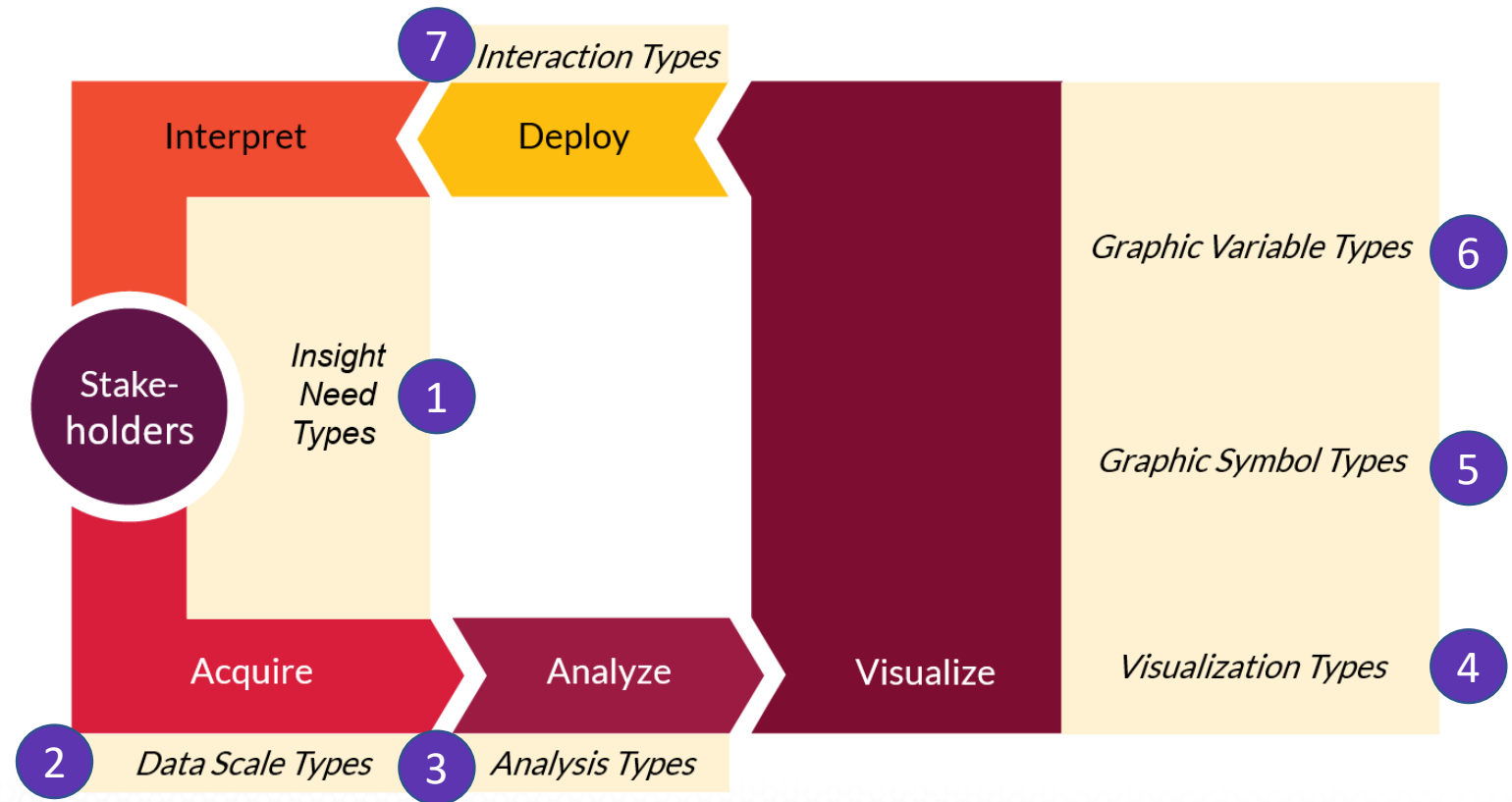
- zoom
- search and locate
- filter
- details-on-demand
- history
- extract
- link and brush
- projection
- distortion

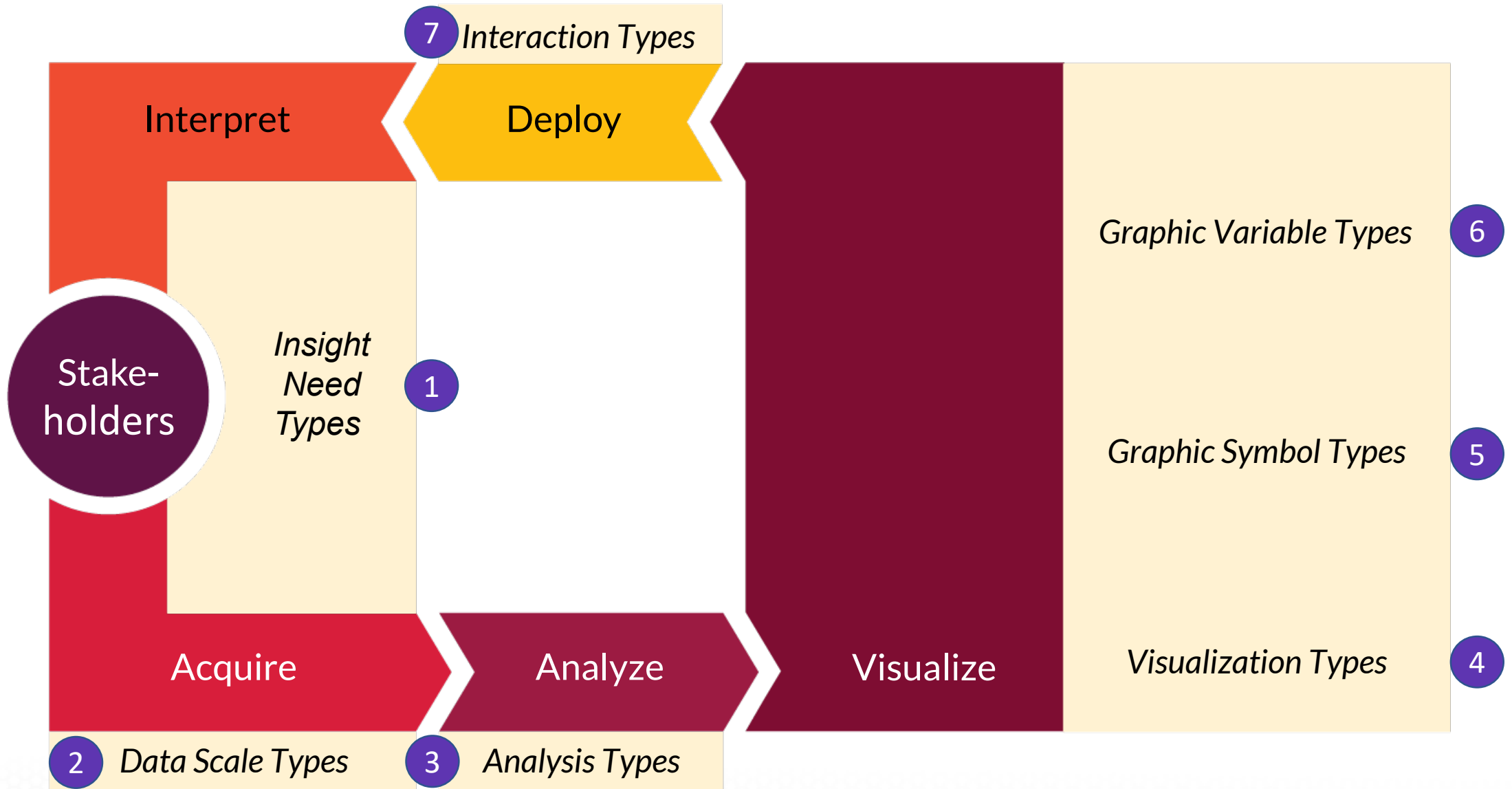
Börner, Katy. 2015. [Atlas of Knowledge: Anyone Can Map](#). Cambridge, MA: The MIT Press. 26, 68-69.

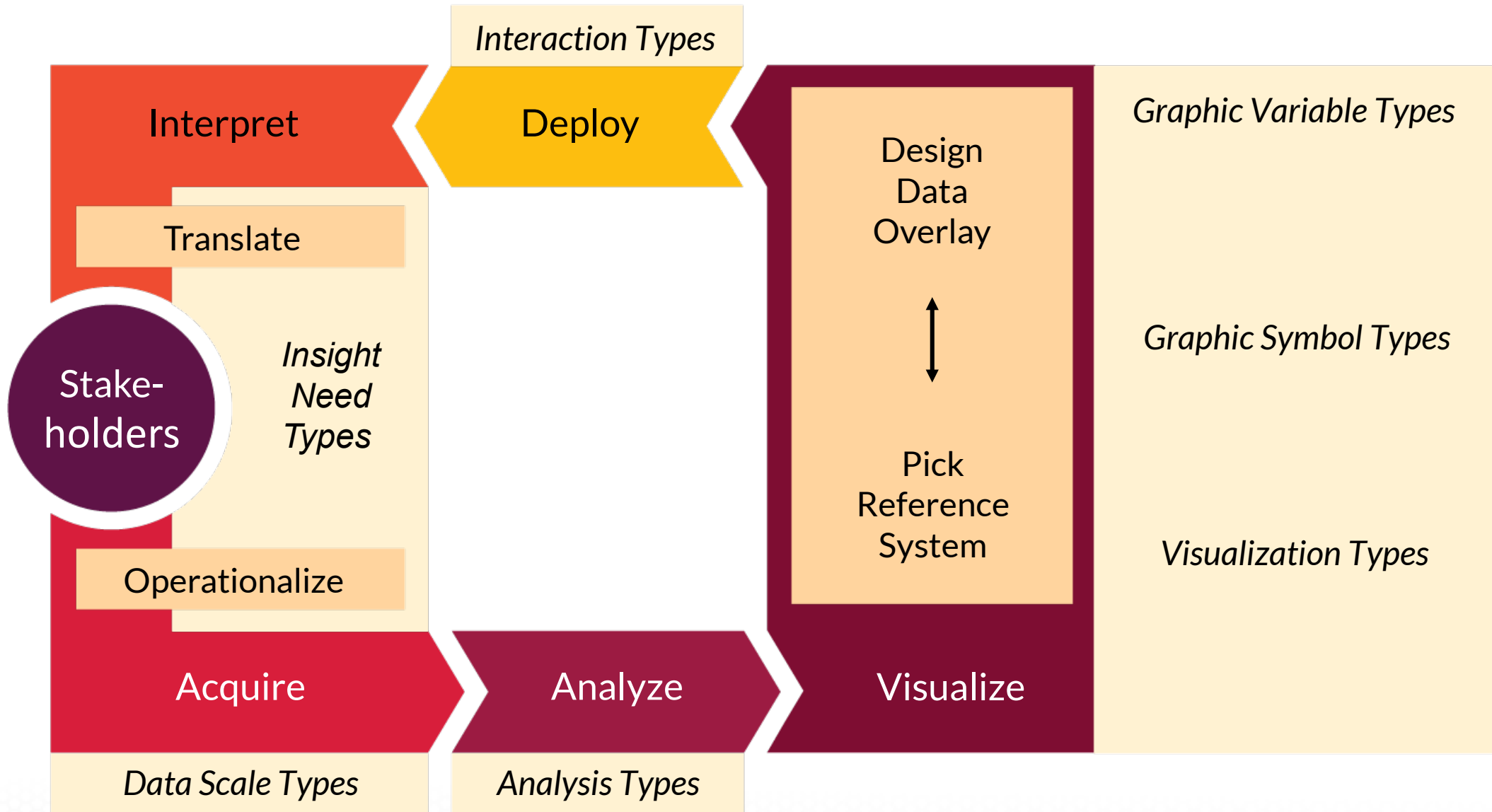
# Data Visualization Literacy Framework (DVL-FW)

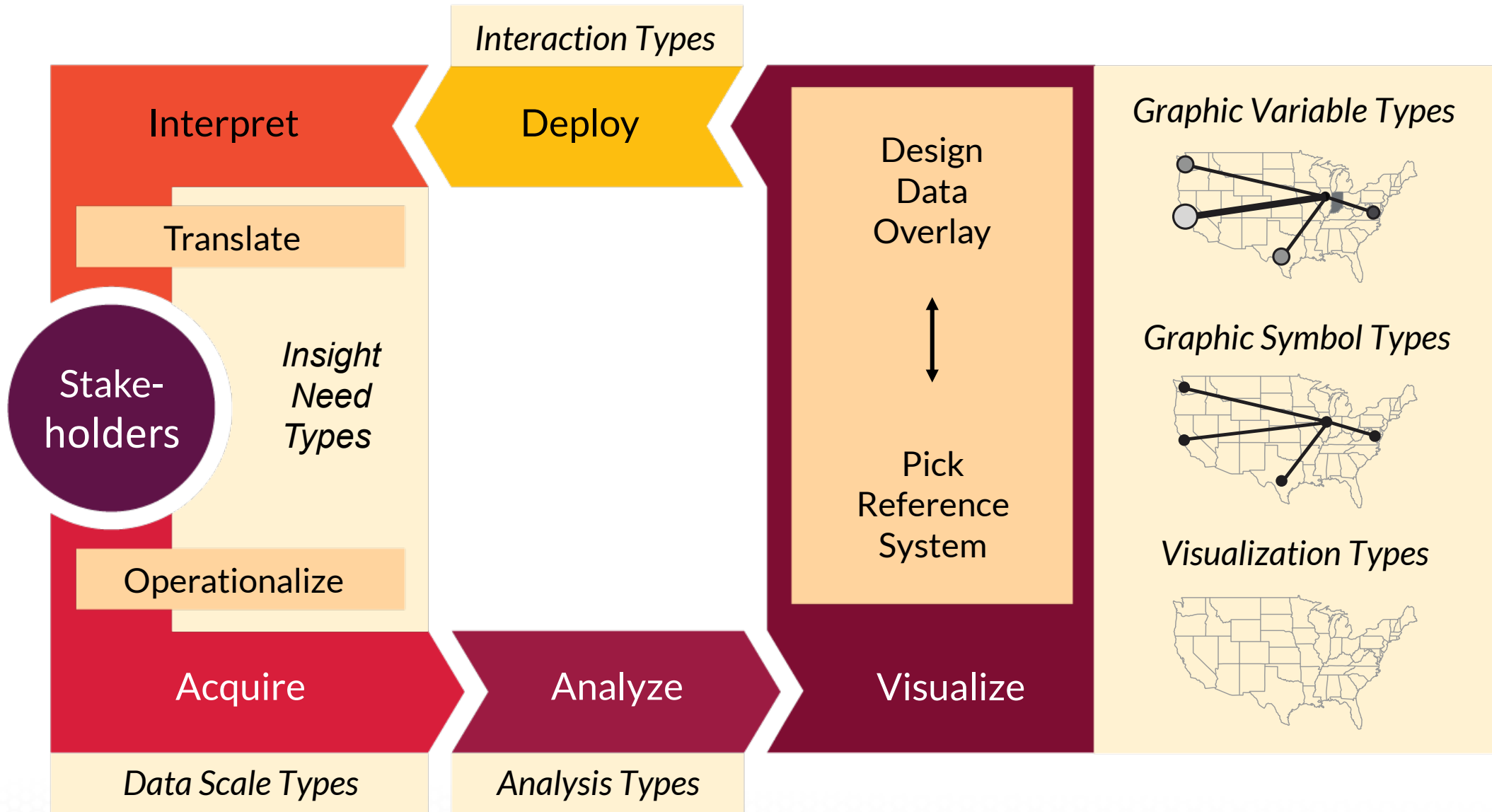
Consists of two parts *that are interlinked*:

**DVL Typology +  
DVL Workflow Process**









Scaling Up:

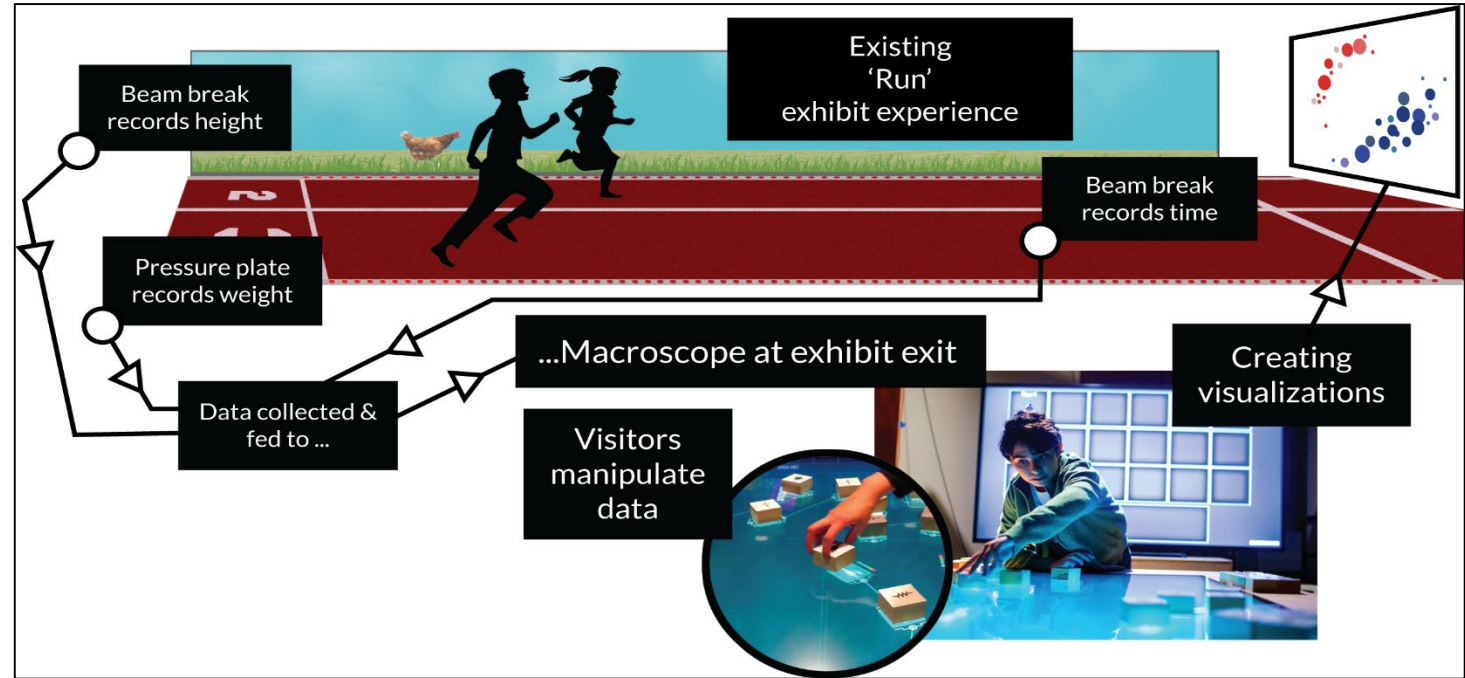
# Teaching Data Visualization Literacy

MAV in Science Museums

Information Visualization MOOC (IVMOOC) + Visual Analytics Certificate (VAC)

















# xMacrosopes in Science Museums



**Data Visualization Literacy:** Research and Tools that Advance Public Understanding of Scientific Data.  
NSF AISL #1713567

# DATA TABLE

	Age	Height (inches)	Time (seconds)	Opponent	Shoes	Zip
	2	32	.5	Visitor	Paws	47401
	7	6	136	Visitor	Paws	47402
	1					
	32					
	64					
	7					
	12					
	15					
	72					
	80					
	7					
	11					
	9					
	0					

Find and Select Your Data Record

Sort by

Age

## SCATTER GRAPH

Y Axis

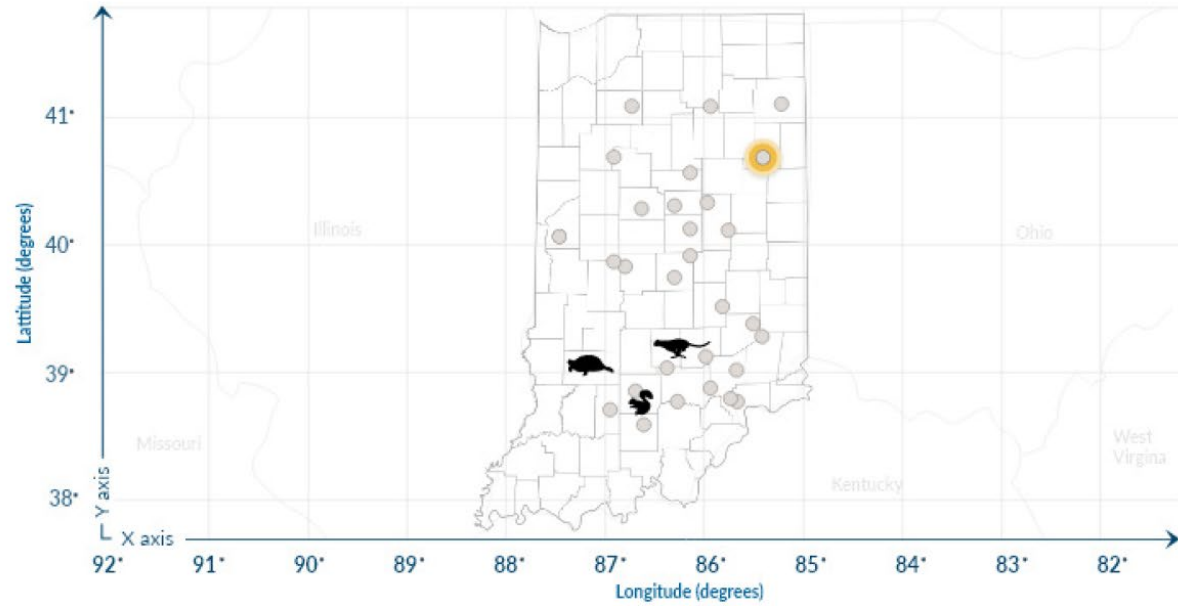
Age

Height

Time



## MAP



HOME



Data Table



HOME



HOME



Data Table



Scatter Graph



Map

Color: Opponent

- Turtle
- Squirrel
- Cheetah

Shape: Shoes

- Paws
- △ Wheels
- Sneakers
- ◇ Casual
- + Dress
- ✦ Heels

Size: Age

- 5
- 23
- 78



Data Visualization Literacy  
NSF AISL #1713567



[OVERVIEW](#)[INSTRUCTORS](#)[POLICIES ▾](#)[SCHEDULE](#)[READINGS](#)[FAQ](#)[CONTACT](#)[REGISTER](#)

## E583 | Z637 | Information Visualization MOOC

This graduate level course provides an overview of the state of the art in information visualization. The course teaches visualization theory and the process of producing effective and actionable visualizations that take the needs of users into account. Students apply the visualization knowledge and skills that they gain in the course by working in teams on real-world client projects.

SELF-PACED

### Among other topics, the course covers:

Stakeholder needs acquisition & project specification

Data mining algorithms and visualization tools

Temporal, geospatial, topical, and network visualization techniques

[REGISTER FOR THE COURSE](#)


## Data Visualization Literacy

In the information age, being able to create and interpret data visualizations is as important as being able to read and write text. This course introduces a theoretical visualization framework to define, measure, and advance student ability in data visualization literacy, discussed in part two in the *Atlas of Knowledge*, published by The MIT Press. The framework is used to organize course content and exams; support the design of effective workflows; to guide visual design, i.e., the mapping of data variables to graphic valuable types and graphic symbol types; and to effectively communicate using proper terminology.

<https://ivmooc.cns.iu.edu>

# Course Listing: INFORMATION VISUALIZATION

Spring 2021, Bloomington

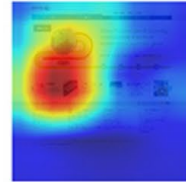
 Information on this report reflects data as of the end of the day Sunday, December 06, 2020

[Select another ENGR course](#) | [Select another department](#)

						<u>Seats</u>	<u>Avail</u>	<u>Wait</u>
<b>ENGR-E 583</b>	<b>INFORMATION VISUALIZATION (3 CR)</b>							
*****	ARR	ARR	WB WEB	Borner K		30	18	0
Above class open to graduates only								
<b>Discussion (DIS)</b>								
10331	ARR	ARR	WB WEB	Borner K		30	10	0
Above class meets with with ENGR-E 483								
Above class meets 100% Online through Asynchronous instruction. For more information visit <a href="https://covid.iu.edu/learning-modes/index.html">https://covid.iu.edu/learning-modes/index.html</a>								
<b>ENGR-E 583</b>	<b>INFORMATION VISUALIZATION (3 CR)</b>							
10379	RSTR	ARR	WB WEB	Borner K		50	29	0
This is a 100% online class taught by IU Bloomington. No on-campus class meetings are required. A distance education fee may apply; check the SCU website for more information at <a href="https://studentcentral.indiana.edu">https://studentcentral.indiana.edu</a>								
Above class for students not in residence on the Bloomington Campus								
Above class meets 100% Online through Asynchronous instruction. For more information visit <a href="https://covid.iu.edu/learning-modes/index.html">https://covid.iu.edu/learning-modes/index.html</a>								



## Client Projects



### Visualizing the Evolution of Website Design

With over 25 years of history, the web itself has become a significant cultural artifact. We are studying how website design has changed over time, and how these changes reflect changes

[Read more...](#)



### Visualizing Research Silos in Ecological Interaction datasets

Open access to high quality and interesting ecological datasets is important to better

[Read more...](#)



### ChaCha Menopause queries

The ChaCha menopause query data is the foundation for building intervention modules to improve people's knowledge and problem solving skills related to menopause. For this project,

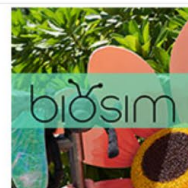
[Read more...](#)



### Text-Mining of User-Generated Queries on Menstrual Pain

Menstrual pain is highly prevalent across women of reproductive age. Distress signals of user

[Read more...](#)



### BioSimmer

BioSim is a participatory simulation where young students (grades K-3) enact the roles of ants and biological systems through the assistance of electronically-enhanced e-puppets. It is

[Read more...](#)

<https://ivmooc.cns.iu.edu/clients.html>

## Visual Analytics Certificate

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

FLYER

REGISTER FOR JAN 18–FEB 28, 2021

FAQS



### Learn from Experts

Connect with industry professionals and leading researchers.



### Evolve Yourself

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



### Make a Difference

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

<https://visanalytics.cns.iu.edu>

# Marvelous Visualization Opportunity: HuBMAP: Mapping 30+ Trillion Cells

Michael P. Snyder, et al. 2019. The human body at cellular resolution: The NIH Human Biomolecular Atlas Program. *Nature*. 574, p. 187-192.

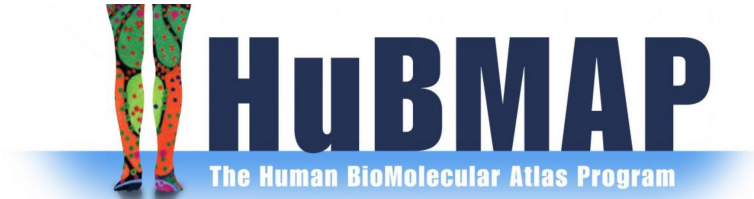
<https://www.nature.com/articles/s41586-019-1629-x.pdf>



# HuBMAP

## Vision

Catalyze the development of an open, global framework for comprehensively mapping the human body at cellular resolution.



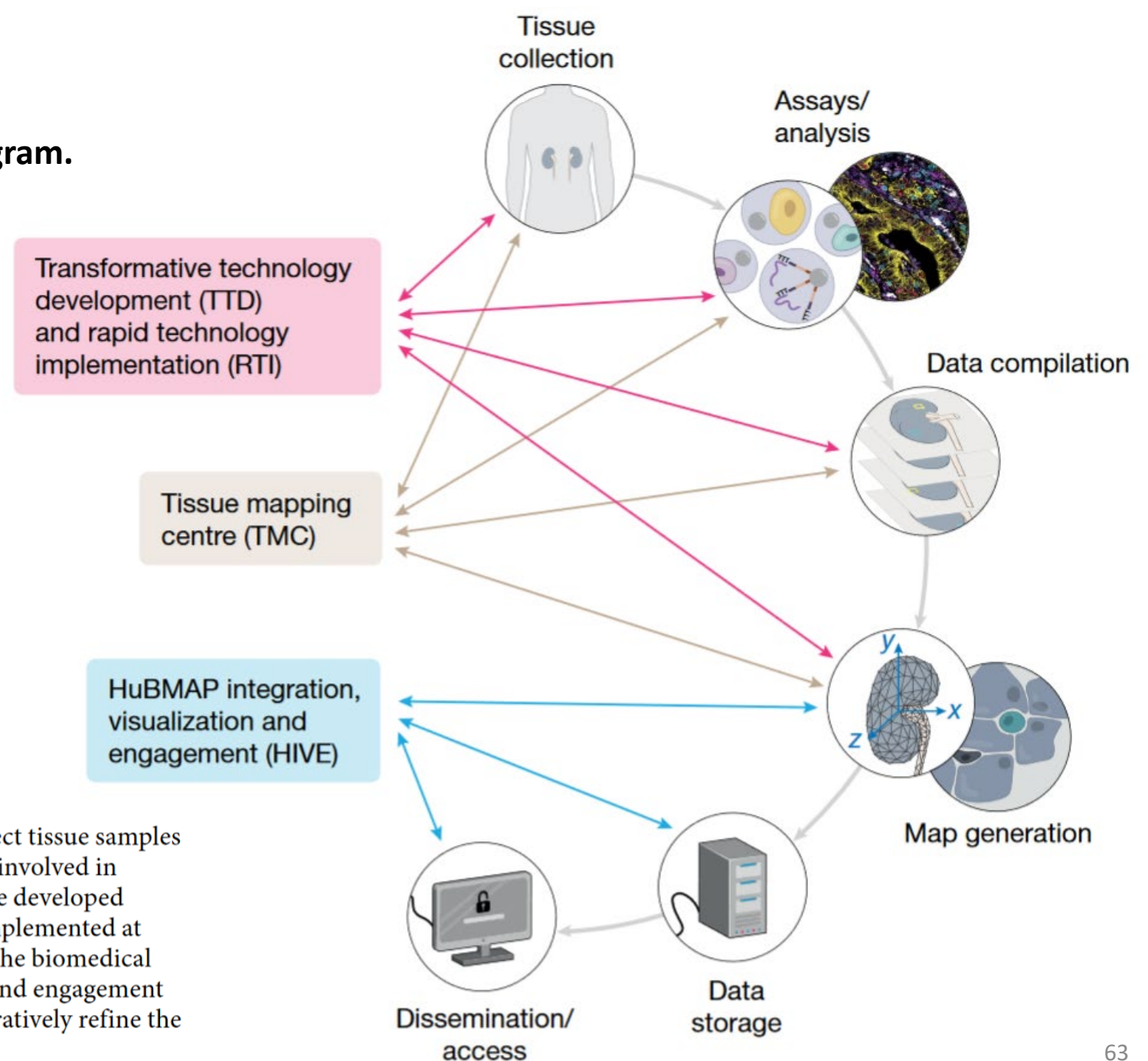
<https://commonfund.nih.gov/HuBMAP>

## Goals

1. Accelerate the development of the next generation of tools and techniques for constructing high resolution spatial tissue maps
2. Generate foundational 3D tissue maps
3. Establish an open data platform
4. Coordinate and collaborate with other funding agencies, programs, and the biomedical research community
5. Support projects that demonstrate the value of the resources developed by the program

# The Human Body at Cellular Resolution: The NIH Human Biomolecular Atlas Program.

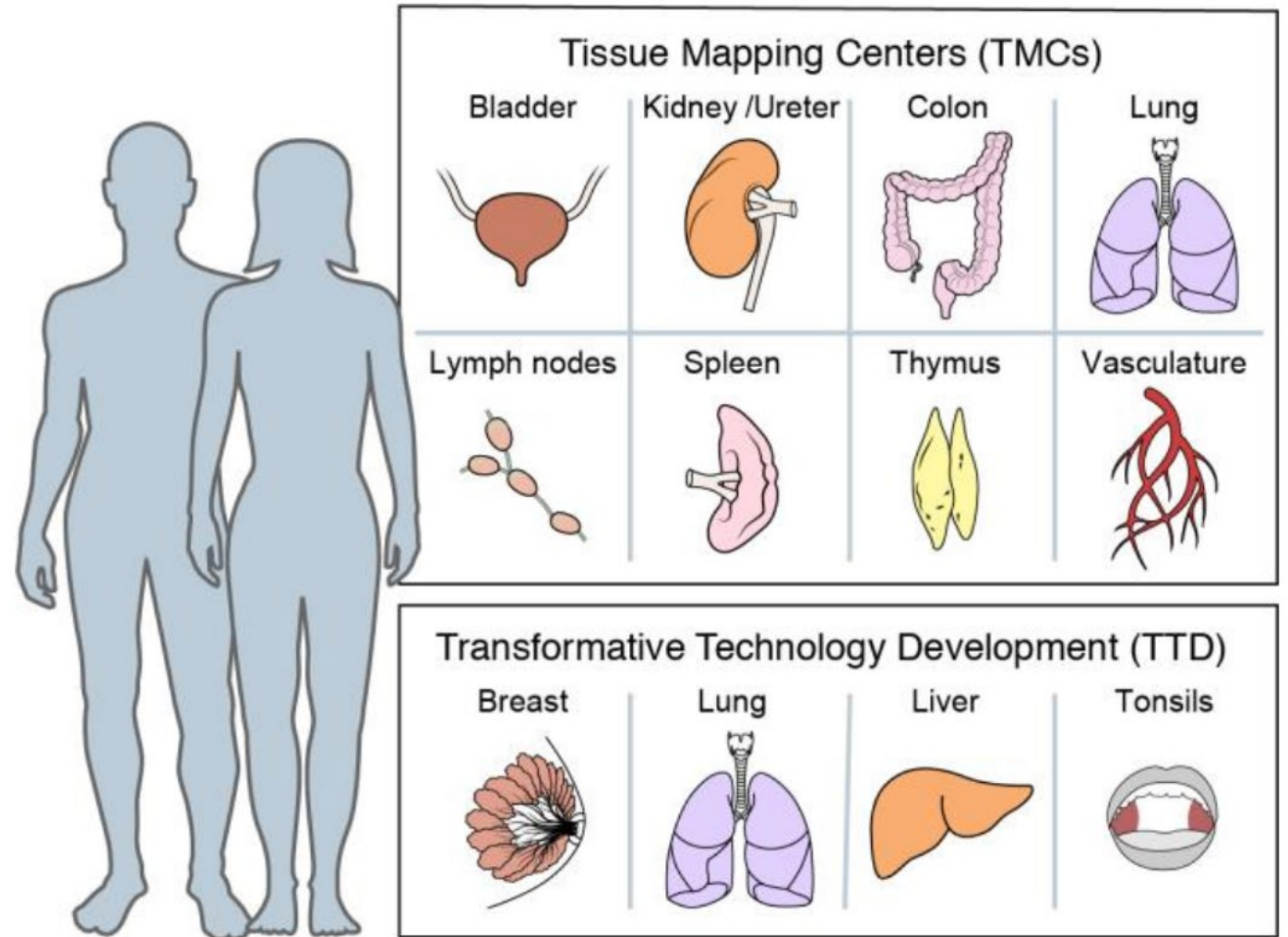
Snyder et al. *Nature*. 574, p. 187-192.



**Fig. 1 | The HubMAP consortium.** The TMCs will collect tissue samples and generate spatially resolved, single-cell data. Groups involved in TTD and RTI initiatives will develop emerging and more developed technologies, respectively; in later years, these will be implemented at scale. Data from all groups will be rendered useable for the biomedical community by the HuBMAP integration, visualization and engagement (HIVE) teams. The groups will collaborate closely to iteratively refine the atlas as it is gradually realized.

## The Human Body at Cellular Resolution: The NIH Human Biomolecular Atlas Program.

Snyder et al. *Nature*. 574, p. 187-192.

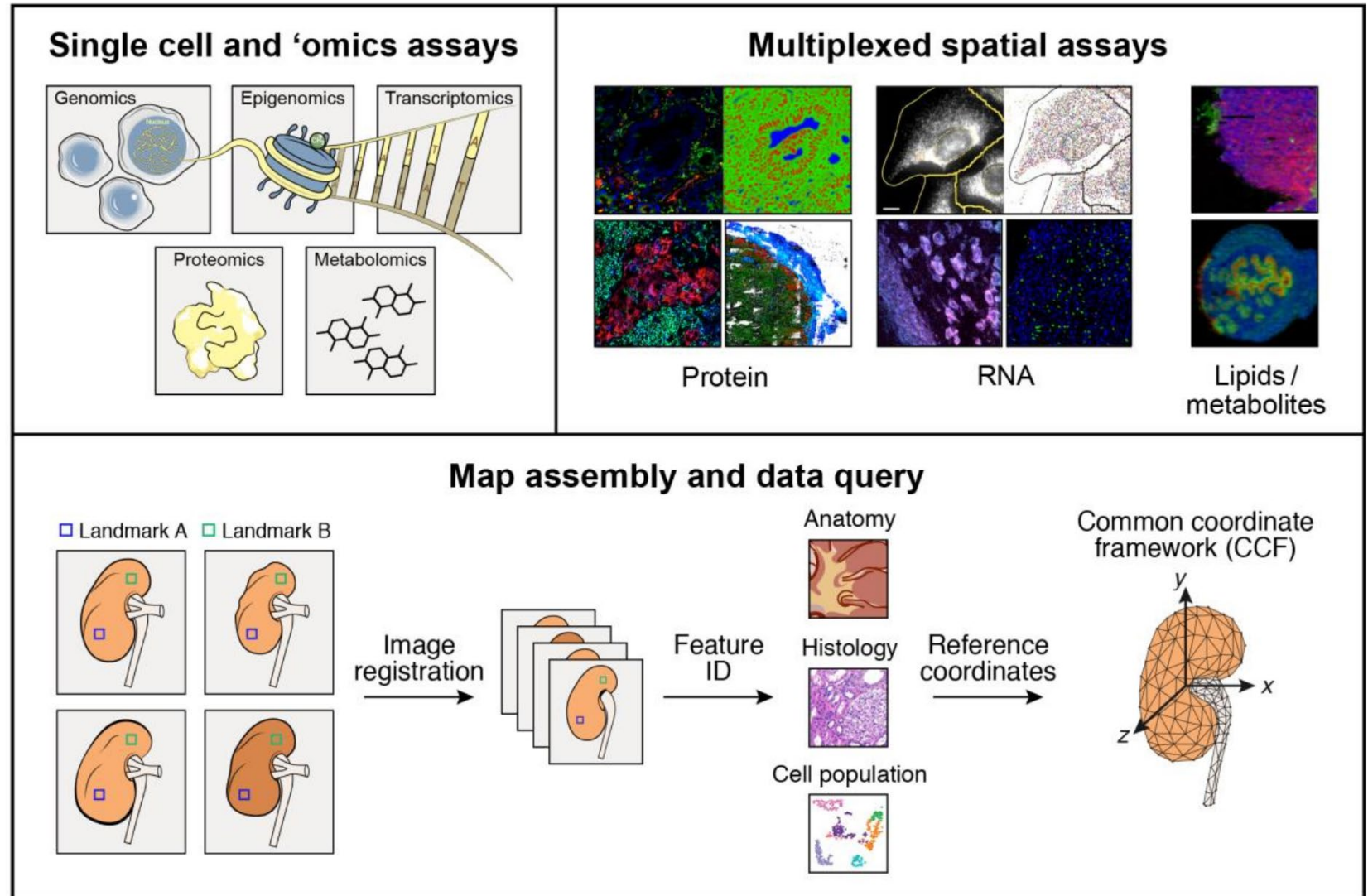


**Fig. 2 | Key tissues and organs initially analysed by the consortium.**

Using innovative, production-grade ('shovel ready') technologies, HuBMAP TMCs will generate data for single-cell, three-dimensional maps of various human tissues. In parallel, TTD projects (and later RTI projects) will refine assays and analysis tools on a largely distinct set of human tissues. Samples from individuals of both sexes and different ages will be studied. The range of tissues will be expanded throughout the program.



**The Human Body at Cellular Resolution: The NIH Human Biomolecular Atlas Program.**  
 Snyder et al. *Nature*. 574, p. 187-192.



**Fig. 3 | Map generation and assembly across cellular and spatial scales.** HuBMAP aims to produce an atlas in which users can refer to a histological slide from a specific part of an organ and, in any given cell, understand its contents on multiple 'omic levels—genomic, epigenomic, transcriptomic, proteomic, and/or metabolomic. To achieve these ends, centres will apply a combination of imaging, 'omics and mass spectrometry

techniques to specimens collected in a reproducible manner from specific sites in the body. These data will be then be integrated to arrive at a high-resolution, high-content three-dimensional map for any given tissue. To ensure inter-individual differences will not be confounded with collection heterogeneity, a robust CCF will be developed.

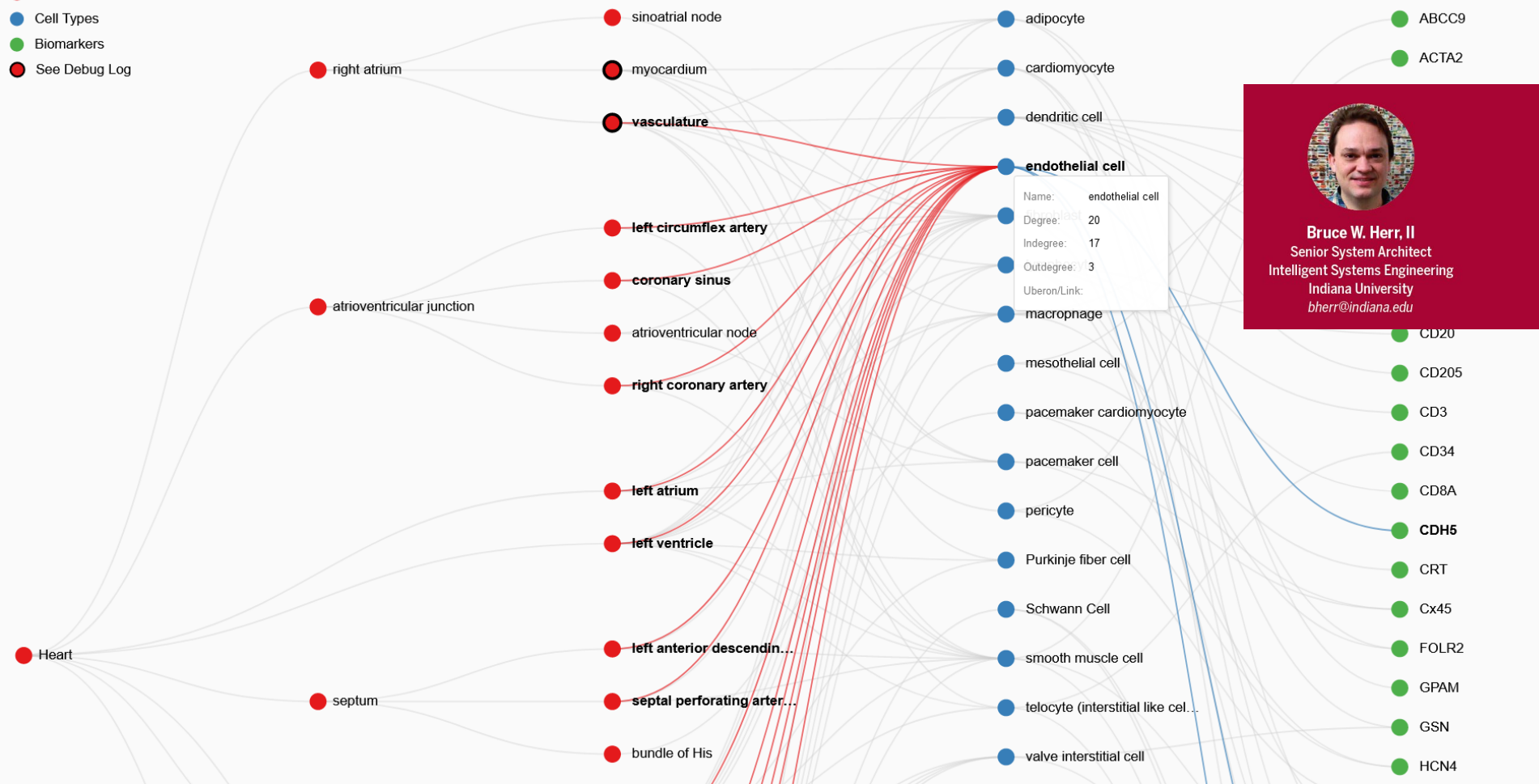
Anatomical Structures


Cell Types

Biomarkers


Legend

- Anatomical Structures
- Cell Types
- Biomarkers
- See Debug Log






**Bruce W. Herr, II**  
Senior System Architect  
Intelligent Systems Engineering  
Indiana University  
bherr@indiana.edu



**Hrishikesh Paul**  
Research Assistant  
Intelligent Systems Engineering  
Indiana University  
hrpaul@iu.edu

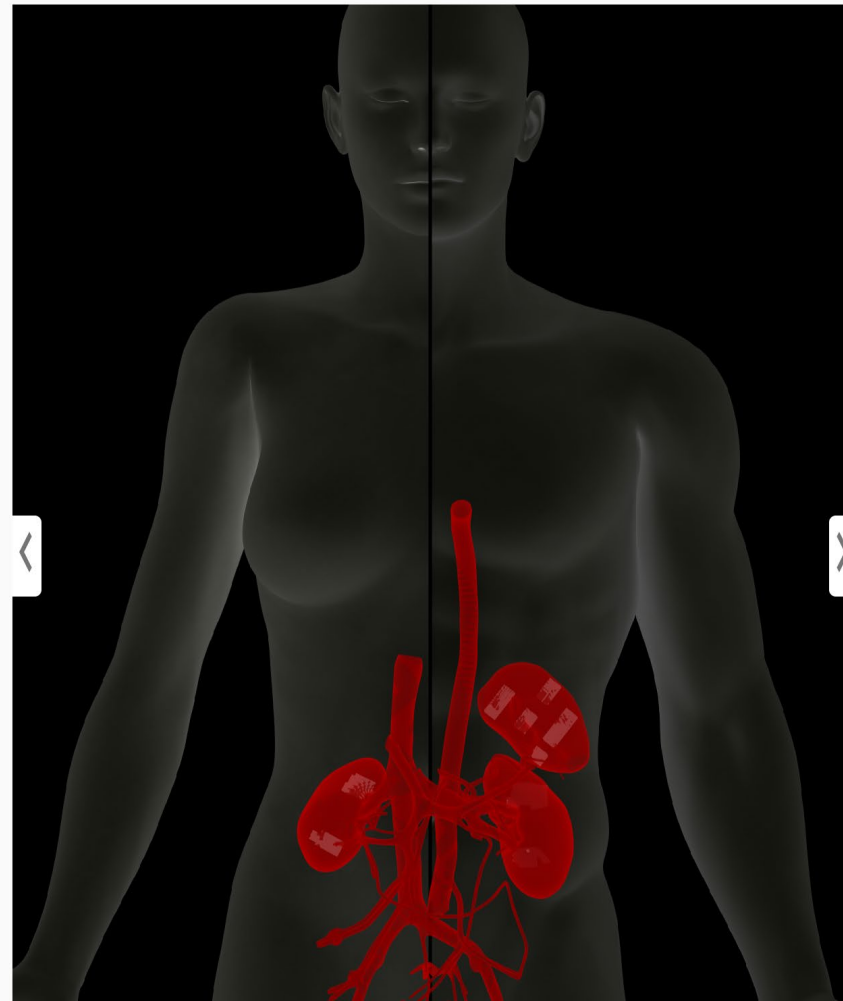


**Ellen M. Quardokus**  
Research Scientist  
Indiana University  
ellenmq@indiana.edu

Search ontology terms ...



- body
  - heart
  - lung
  - kidney
    - right kidney
    - left kidney
    - kidney capsule
    - cortex of kidney
    - renal medulla
    - renal column
    - renal pyramid
    - hilum of kidney
    - kidney interstitium
    - kidney calyx
    - renal pelvis
    - ureter
    - renal papilla
    - renal fat pad
    - nephron



body

2 Centers  
27 Donors  
41 Samples

	<b>Female, Age 14, BMI 14.7</b> HBM894.MPVN.828 TMC-Florida First case collected. Incomplete d...	
	<b>Male, Age 18, BMI 27.1</b> HBM436.GHWX.449 TMC-Florida section is 190um from block surface	
	<b>Male, Age 56, BMI 32.5</b> HBM696.XTVL.498 TMC-Vanderbilt Age 56, White Male	
	<b>Male, Age 53, BMI 26.5</b> HBM652.VRLD.292 TMC-Vanderbilt Age 53, Black Male	
	<b>Male, Age 58, BMI 22.0</b> HBM477.CJKM.888 TMC-Vanderbilt 107-111	
	<b>Male, Age 18, BMI 25.5</b> HBM473.VKCM.878 TMC-Florida section is 255um from block surface	
	<b>Male, Age 55, BMI 25.4</b> HBM824.BLXF.883 TMC-Vanderbilt 13-16	

# Acknowledgements

HuBMAP Consortium (<https://hubmapconsortium.org>)



Thanks go to all the **patients** that agreed to volunteer healthy tissue and open use of their data.



## TMCs



**Jeffrey Spraggins**  
TMC-Vanderbilt  
Vanderbilt University



**Sanjay Jain**  
TMC-UCSD  
Washington University,  
St. Louis



**Clive Wasserfall**  
TMC-UFL  
University of Florida



**Marda Jorgensen**  
TMC-UFL  
University of Florida



**Kristen Browne**  
Medical Imaging and  
3D Modeling Specialist  
NIAID

## 3D Models

## MC-IU HIVE Team



**Katy Börner**  
MC-IU PI  
CNS Director



**Griffin Weber**  
Assoc. Professor of Medicine  
Harvard Medical School



**Lisel Record**  
MC-IU PM  
CNS Associate Director



**Bruce Herr II**  
Sr. Systems Architect/PM



**Ellen Quardokus**  
Sr. Research Analyst



**Yingnan Ju**  
PhD Candidate



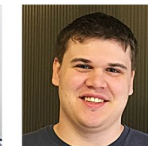
**Andreas Bueckle**  
PhD Candidate



**Leonard Cross**  
Sr. UX/UI Designer



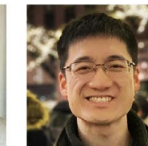
**Matthew Martindale**  
Center Assistant



**Daniel Bolin**  
Software Developer



**Adam Phillips**  
Software Developer



**Edward Lu**  
Software Developer



**Paul Hrishikesh**  
Research Assistant



**Leah Scherschel**  
Research Assistant



**Avinash Boppana**  
Research Consultant



# Interdisciplinary Training in Complex Networks and Systems

[RESOURCES](#) / [TWITTER](#)[THE PROGRAM](#)[RESEARCH](#)[HOW TO APPLY](#)[STUDENTS](#)[COLLOQUIA](#)[NEWS](#)

## The program

[Share](#) [Tweet](#) [Email](#)

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](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>

