Atlas of Forecasts: Predicting and Broadcasting Science, Technology, and Innovation

Katy Börner  @katycns

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

Rotary Talk

Noon ET on March 7, 2023
Overview

- Data Visualizations of Science
- The Science of Data Visualization
- Open Challenges
Atlas Trilogy

Atlas of Science
Visualizing What We Know
Katy Börner
2010

Atlas of Knowledge
Anyone Can Map
Katy Börner
2015

Atlas of Forecasts
Modeling and Mapping Desirable Futures
Katy Börner
2021

https://mitpress.mit.edu/books/atlas-forecasts
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.

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

http://scimaps.org

The David J. Sencer CDC Museum, Atlanta, GA.
Places & Spaces: Mapping Science Exhibit

1st Decade (2005-2014)

Maps

2nd Decade (2015-2024)

Macrosopes

http://scimaps.org
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.
The Structure of Science

The Social Sciences are the smallest and least obvious of the three areas. Psychology serves as the tie between Medical Sciences and the other Sciences. Sociology serves as the tie with Computer Science and Mathematics.

Mathematics is our starting point, the purest of all sciences. It lies on the outer edge of the map. Computational Science, Quantum Engineering and Physics are applied disciplines that form a new science domain in mathematics and science. These three disciplines provide a glimpse of the entire structure of science. Physics through include disciplines. Although applied, these disciplines are highly concentrated with distinct bands of research connecting them. Bands indicate interdisciplinary research.

Research is highly concentrated in Physics and Chemistry. These disciplines have few, but very distinct, bands of disciplines connected to them. The richness of these bands indicates an extensive amount of interdisciplinary research. Mathematics and Chemistry are not as distinct as one might assume.

The Life Sciences are including Biology and Biochemistry, are less concentrated than Chemistry or Physics. Many links between the larger areas of the Life Sciences, for instance between Biology, and Biochemistry, and between Biology and Environmental Science. Biochemistry is very interrelated in the Life Sciences. It is a large discipline that has visible links to disciplines across 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 11 areas of science from the citation patterns of 1,500 scientific topics. That early map was fascinating, but it didn’t cover enough of science to accurately define its structure.

Things are different today. We have enormous computer power and advanced visualization software that make mapping the entire structure of science possible. This paper (see map of science) was generated at ScioNet National Laboratory using an advanced graph-theoretic algorithm (VOS). The data is taken from the citation patterns of 80,000 scientific papers published in 1998. This graph represents one of the 97000 research communities active in science in 2002. A research community for each group of papers. Each group has 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 an educational tool. This is the terrain in which organizations and laboratories create their scientific capability. Additional information about the scientific and economic impact of each research community allows policy makers to decide which areas to expand, evolve, shrink, or ignore.

We also calculate the maps as educational tools for children. The theoretical relationship between areas of science can be explored with a concrete map. Children observe how math, physics, chemistry, biology and other social sciences inform each other. Areas of science can be learned and new learning areas can be explored.
Check out our Zoom Maps online!

Visit scimaps.org and check out all our maps in stunning detail!
Iteration XI (2015)
Macrosopes for Interacting with Science

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

Iteration XVI (2020)
Macrosopes for Harnessing the Power of Data

http://idemo.cns.iu.edu/macroscopic-kiosk
This is the Roanoke (Raleigh) megaregion.
Acknowledgements

Exhibit Curators

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

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

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

Bonnie DeVarco
MediaX Distinguished Visiting Scholar at Stanford University, Palo Alto, CA, USA

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

Peter A. Hook
Head of Digital and Scholarly Services and LawArXiv Administrator, Cornell Law Library, Ithaca, NY, USA

Stephen Uzzo
Vice President of Science and Technology for the New York Hall of Science

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

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

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

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

http://scimaps.org
Call for Macroscopes: 19\textsuperscript{th} Iteration

What to Submit

• Each entry needs to include:
• Title of macroscope
• Author(s) name, email address, affiliation, mailing address
• Link to online site that features the macroscope tool or to executable code
• Macroscope tool description (300 words max): user group and needs served, data used, data analysis performed, visualization techniques applied, and main insights gained
• References to relevant publications or online sites that should be cited, links to related projects or works
• Tell us about the impact your data visualization has had on public awareness, social policy, or political action.

Review Process

Submissions will be reviewed and evaluated by the exhibit advisory board (listed below) in terms of their:
• Scientific rigor
• Value as a tool for data exploration
• Ability to provide new, actionable insights
• Relevance for a general audience

Important Dates

• Submissions due: Feb 15, 2023
• Notification to mapmakers: April 1, 2023
• Submit final entries: May 30, 2023
• Iteration ready for display: August 31, 2023

https://scimaps.org/call
Atlas of Forecasts
Modeling and Mapping Desirable Futures
Katy Börner
I am deeply grateful to all those who helped to make this atlas and the exhibit maps it features. Financial support came from the National Science Foundation under Grants No. DRL-1226098, OCE-0948264, SSE-0738111, and CBET-0831366; the National Institutes of Health under Grants No. U54 GM078890, R21-DM042379, and U54-RR029822; the John S. McDonnell Foundation; the Bill & Melinda Gates Foundation; Indiana’s 21st Century Fund; Thomson Reuters; Elsevier; the Cyberinfrastructure for Network Science Center, University Information Technology Services, and the former School of Library and Information Science—all those at Indiana University. Some of the data used to generate the science maps is from the Web of Science by Thomson Reuters and Scopus by Elsevier. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

A substantial part of the source review and initial writing was completed while I was a visiting professor at the Royal Netherlands Academy of Arts and Sciences (KNAW) in the spring of 2012. I would like to thank Paul Weathers of CTVS, Andrea Scharnhorst and Jeannette Haugan of Mementos, and Peter Deves, Linda Reijndoudt, and Lucas Puida of DANS for their support.

Part 2, "Evolutionizing Science," benefited deeply from any teaching of relevant courses at Indiana University over the last 15 years, including teaching Information Visualization MOOC (IVMOOC) to students from more than 100 countries in the spring of 2013. The Places & Spaces: Mapping Science exhibit would not have been possible without the expertise and professional excellence of the more than 150 mapmakers and the 42 exhibit ambassadors around the globe. Exhibit advisors for the maps featured in this book included: Deborah MacManus (Accuracy/Aesthetics), Kevin W. Boyack (SciTech Strategies, Inc.), Sara Irina Fabrikant (Geography Department, University of Zurich, Switzerland), Peter A. Hooke (La Trobe University, Victoria), Andri Stapić (Geography, San Diego State University), Bonita DeGraaf (Rensselaer, NY), and Drew Whitley (Geography and Oceanography, Oregon State University). External experts that reviewed iterations 4 through 7 included: John R. Hobert (Chief of the Geography and Map Division, Library of Congress), Thomas B. Hickey (OCLC), Michael Kurtz (Harvard-Smithsonian Center for Astrophysics), Denise A. Bedford (World Bank), William Ying (ISO ArtSTOR), Michael Kost (JSTOR), Carl Logue (Cornell University), Richard Forrester (Linus & MDiv University), Vincent Larivière (Université de Quèbec à Montréal), Canada), Adam Bh (CIES of WEED), Alex Wright (author of Guide: Mastering Information Through the Ages), and Mills Dodge (ProjectSNIC.org).

Focused brainstorming workshops, organized with colleagues between 2008 and 2012, contributed greatly to the discussion of research and development (R&D) work that is continued in these pages. A total of 16 such workshops were held on a range of topics, including "How to Measure, Map, and Dramatize Science," "Mapping the History and Philosophy of Science," "Modeling Knowledge Dynamics," "Atmospheric Science & Technology," and "Plug-and-Play Microscopes" (see group photos below).

It may seem strange to devote a major part of one's research time to writing a series of books for readers who are unlikely to read papers or otherwise cite these books in academic circles. And yet it seems quite on target to enable those who finance science via tax dollars to benefit from the research results—for bettering the maximization of citation counts via the production of research papers. Many others have taken this route, including the following luminaries who have inspired my own journey: Jacques-Yves Cousteau, the French explorer and researcher of the sea; David Attenborough, especially with his Life on Earth and Living Planet series; Paul Orter, with his Universal Atlas or Encyclopaedia Universalis Mundanae; Stephen Brand, author of The Whole World Catalog; Richard Dawkins, famed for his "Growing Up in the Universe" lecture; Al Gore for his environmental efforts, as featured in the Academy Award documentary and Hans Rosling, whose Gapminder effort gave rise to the motto, "Let your dataset change your mindset." It is my hope that this atlas series helps in giving both inspiration and encouragement to future science communicators.

Coping with the atlas was performed by Cornelia Jeliničić, atlas layout and design by Tracy Tharartik, with many of the images specifically created for this book by Petra Manor-Lajoie and Samuel T. Mills; reference checks and formatting by Todd N. Tharartik; and copyright acquisition by Samantha Halk, Brunnah Marshall, Joseph Stankiewicz, and Michael P. Oden. Other valued contributions are acknowledged in the References & Credits (page 178).

This atlas was influenced by research and developments in many areas of science; it also benefited from countless discussions and brainstorming sessions with extended colleagues. And yet the browser-first decision making regarding content, format, structure, and design at every stage was mine alone to make.

I am indebted to family and friends for providing much inspiration, energy, and loving support. This book benefited deeply from nurturing and thoughtful provoking family dinner discussions and powerlifting girl’s nights out. My gratitude also goes with our cat, Jiji, who kept me company through the many long periods of writing.

http://mitpress.mit.edu/books/atlas-forecasts

Acknowledgments
Atlas of Forecasts: Models of (Desirable) Futures

Model Classes
Many different modeling approaches exist. The table below by William B. Rouse shows exemplary levels of modeling, issues needing to be addressed, and models that have been successfully applied to support decision-making.

<table>
<thead>
<tr>
<th>Level</th>
<th>Concern</th>
<th>Models</th>
</tr>
</thead>
<tbody>
<tr>
<td>Society</td>
<td>GDP, Supply/Demand, Policy</td>
<td>Macroeconomic</td>
</tr>
<tr>
<td></td>
<td>Economic Cycles</td>
<td>System Dynamics</td>
</tr>
<tr>
<td></td>
<td>Intra-Firm Relations, Competition</td>
<td>Network Models</td>
</tr>
<tr>
<td>Organizations</td>
<td>Profit Maximization</td>
<td>Microeconomic</td>
</tr>
<tr>
<td></td>
<td>Competition</td>
<td>Game Theory</td>
</tr>
<tr>
<td></td>
<td>Investment</td>
<td>DCF, Options</td>
</tr>
<tr>
<td>Processes</td>
<td>Patient, Material Flow</td>
<td>Discrete-Event Models</td>
</tr>
<tr>
<td></td>
<td>Process Efficiency</td>
<td>Learning Models</td>
</tr>
<tr>
<td></td>
<td>Workflow</td>
<td>Network Models</td>
</tr>
<tr>
<td>People</td>
<td>Patient Behavior</td>
<td>Agent-Based Models</td>
</tr>
<tr>
<td></td>
<td>Risk Aversion</td>
<td>Utility Models</td>
</tr>
<tr>
<td></td>
<td>Discourse Progression</td>
<td>Markov, Bayes Models</td>
</tr>
</tbody>
</table>
**Phenomena of Interest**

- **Oscillation**
  
  A term that refers to the periodic or cyclic behavior of a system, often observed in electrical circuits or biological processes.

- **Tipping Point**
  
  A critical threshold where a system's behavior changes suddenly, often used to describe environmental or social systems.

- **Reentrance**
  
  A phenomenon where a system returns to a previously observed state or state space after a perturbation or change in parameters.

**Seasonality**

- **Synchronization**
  
  The process by which oscillating systems adjust their frequency and phase to a common value, often observed in biological rhythms or financial markets.

- **Phase Transition**
  
  The transformation of a thermodynamic system from one phase to another, often accompanied by a change in properties such as volume or temperature.

**Perspective**

- **Modeling Goals**
  
  Models aim to capture key phenomena at the levels that are most relevant for the understanding, communication, and management of systems. This broad-spectrum perspective comprises key phenomena that are commonly studied when aiming to understand complex systems. Phenomena are roughly organized by question type (temporal, geographic, topical, and network) and complexity. Models that use static reference systems and no feedback cycles are introduced first, followed by phenomena that aim to capture evolving networks and activity patterns unfolding over time, including feedback or causal loops.

**Phenomena**

- **Oscillation**
  
  A term that refers to the periodic or cyclic behavior of a system, often observed in electrical circuits or biological processes.

- **Tipping Point**
  
  A critical threshold where a system's behavior changes suddenly, often used to describe environmental or social systems.

- **Reentrance**
  
  A phenomenon where a system returns to a previously observed state or state space after a perturbation or change in parameters.

**Seasonality**

- **Synchronization**
  
  The process by which oscillating systems adjust their frequency and phase to a common value, often observed in biological rhythms or financial markets.

- **Phase Transition**
  
  The transformation of a thermodynamic system from one phase to another, often accompanied by a change in properties such as volume or temperature.

**Perspective**

- **Modeling Goals**
  
  Models aim to capture key phenomena at the levels that are most relevant for the understanding, communication, and management of systems. This broad-spectrum perspective comprises key phenomena that are commonly studied when aiming to understand complex systems. Phenomena are roughly organized by question type (temporal, geographic, topical, and network) and complexity. Models that use static reference systems and no feedback cycles are introduced first, followed by phenomena that aim to capture evolving networks and activity patterns unfolding over time, including feedback or causal loops.
Modeling Framework

When developing a model of a real-world system, many critical decisions must be made regarding model components, their behavior, the environment, and system dynamics evolving over time. Any model designed should start with a specification of stakeholders and their needs, followed by phenomena of interest, and finally the success criteria that define when a model is fit-for-purpose. Model validation and communication must be elaborated. Various approaches have been proposed to provide templates and standards for systematic model development and documentation—important for ensuring replicability of results. This paper reviews prior work on modeling frameworks and then introduces and explains the data visualization approach in a new Atlas of Iconology, Part 2, to cover the emergent phenomena discussed in the previous section, as well as the atlas, descriptive, predictive, and definitive models throughout the Atlas of Iconology. We cannot stop the march of history, but we can influence its direction.

Prior Work

There exist many frameworks that aim to guide researchers and experts in the design, use, evaluation, and validation of models. While many experts, researchers, and practitioners have developed various frameworks, a few main themes can be identified: (1) The ISO/IEC COCOMO II model has developed and documented a process model used by software developers that describes phases in projects (4); (2) Initials (common definition, (5) website dedicated to modeling, (6) COCOMO II; (7) development of COCOMO II; (8) and (9) mandatory models of model-based (10) processes and (11). A common theme among these frameworks is the qualitative and quantitative relationship between different stakeholders and model components, as well as the communication and documentation of models. (12) It is clear that these frameworks provide a good starting point for understanding the relationship of visual and descriptive models in data visualization systems. However, these frameworks do not address the specific needs of the current project and do not provide a comprehensive framework for the design and implementation of models. (13)

Modeling Framework

Modeling Framework

When developing a model of a real-world system, many critical decisions must be made regarding model components, their behavior, the environment, and system dynamics evolving over time. Any model designed should start with a specification of stakeholders and their needs, followed by phenomena of interest, and finally the success criteria that define when a model is fit-for-purpose. Model validation and communication must be elaborated. Various approaches have been proposed to provide templates and standards for systematic model development and documentation—important for ensuring replicability of results. This paper reviews prior work on modeling frameworks and then introduces and explains the data visualization approach in a new Atlas of Iconology, Part 2, to cover the emergent phenomena discussed in the previous section, as well as the atlas, descriptive, predictive, and definitive models throughout the Atlas of Iconology. We cannot stop the march of history, but we can influence its direction.

Prior Work

There exist many frameworks that aim to guide researchers and experts in the design, use, evaluation, and validation of models. While many experts, researchers, and practitioners have developed various frameworks, a few main themes can be identified: (1) The ISO/IEC COCOMO II model has developed and documented a process model used by software developers that describes phases in projects (4); (2) Initials (common definition, (5) website dedicated to modeling, (6) COCOMO II; (7) development of COCOMO II; (8) and (9) mandatory models of model-based (10) processes and (11). A common theme among these frameworks is the qualitative and quantitative relationship between different stakeholders and model components, as well as the communication and documentation of models. (12) It is clear that these frameworks provide a good starting point for understanding the relationship of visual and descriptive models in data visualization systems. However, these frameworks do not address the specific needs of the current project and do not provide a comprehensive framework for the design and implementation of models. (13)

Modeling Framework

Modeling Framework

When developing a model of a real-world system, many critical decisions must be made regarding model components, their behavior, the environment, and system dynamics evolving over time. Any model designed should start with a specification of stakeholders and their needs, followed by phenomena of interest, and finally the success criteria that define when a model is fit-for-purpose. Model validation and communication must be elaborated. Various approaches have been proposed to provide templates and standards for systematic model development and documentation—important for ensuring replicability of results. This paper reviews prior work on modeling frameworks and then introduces and explains the data visualization approach in a new Atlas of Iconology, Part 2, to cover the emergent phenomena discussed in the previous section, as well as the atlas, descriptive, predictive, and definitive models throughout the Atlas of Iconology. We cannot stop the march of history, but we can influence its direction.

Prior Work

There exist many frameworks that aim to guide researchers and experts in the design, use, evaluation, and validation of models. While many experts, researchers, and practitioners have developed various frameworks, a few main themes can be identified: (1) The ISO/IEC COCOMO II model has developed and documented a process model used by software developers that describes phases in projects (4); (2) Initials (common definition, (5) website dedicated to modeling, (6) COCOMO II; (7) development of COCOMO II; (8) and (9) mandatory models of model-based (10) processes and (11). A common theme among these frameworks is the qualitative and quantitative relationship between different stakeholders and model components, as well as the communication and documentation of models. (12) It is clear that these frameworks provide a good starting point for understanding the relationship of visual and descriptive models in data visualization systems. However, these frameworks do not address the specific needs of the current project and do not provide a comprehensive framework for the design and implementation of models. (13)

Modeling Framework

Modeling Framework

When developing a model of a real-world system, many critical decisions must be made regarding model components, their behavior, the environment, and system dynamics evolving over time. Any model designed should start with a specification of stakeholders and their needs, followed by phenomena of interest, and finally the success criteria that define when a model is fit-for-purpose. Model validation and communication must be elaborated. Various approaches have been proposed to provide templates and standards for systematic model development and documentation—important for ensuring replicability of results. This paper reviews prior work on modeling frameworks and then introduces and explains the data visualization approach in a new Atlas of Iconology, Part 2, to cover the emergent phenomena discussed in the previous section, as well as the atlas, descriptive, predictive, and definitive models throughout the Atlas of Iconology. We cannot stop the march of history, but we can influence its direction.

Prior Work

There exist many frameworks that aim to guide researchers and experts in the design, use, evaluation, and validation of models. While many experts, researchers, and practitioners have developed various frameworks, a few main themes can be identified: (1) The ISO/IEC COCOMO II model has developed and documented a process model used by software developers that describes phases in projects (4); (2) Initials (common definition, (5) website dedicated to modeling, (6) COCOMO II; (7) development of COCOMO II; (8) and (9) mandatory models of model-based (10) processes and (11). A common theme among these frameworks is the qualitative and quantitative relationship between different stakeholders and model components, as well as the communication and documentation of models. (12) It is clear that these frameworks provide a good starting point for understanding the relationship of visual and descriptive models in data visualization systems. However, these frameworks do not address the specific needs of the current project and do not provide a comprehensive framework for the design and implementation of models. (13)

Modeling Framework

Modeling Framework

When developing a model of a real-world system, many critical decisions must be made regarding model components, their behavior, the environment, and system dynamics evolving over time. Any model designed should start with a specification of stakeholders and their needs, followed by phenomena of interest, and finally the success criteria that define when a model is fit-for-purpose. Model validation and communication must be elaborated. Various approaches have been proposed to provide templates and standards for systematic model development and documentation—important for ensuring replicability of results. This paper reviews prior work on modeling frameworks and then introduces and explains the data visualization approach in a new Atlas of Iconology, Part 2, to cover the emergent phenomena discussed in the previous section, as well as the atlas, descriptive, predictive, and definitive models throughout the Atlas of Iconology. We cannot stop the march of history, but we can influence its direction.

Prior Work

There exist many frameworks that aim to guide researchers and experts in the design, use, evaluation, and validation of models. While many experts, researchers, and practitioners have developed various frameworks, a few main themes can be identified: (1) The ISO/IEC COCOMO II model has developed and documented a process model used by software developers that describes phases in projects (4); (2) Initials (common definition, (5) website dedicated to modeling, (6) COCOMO II; (7) development of COCOMO II; (8) and (9) mandatory models of model-based (10) processes and (11). A common theme among these frameworks is the qualitative and quantitative relationship between different stakeholders and model components, as well as the communication and documentation of models. (12) It is clear that these frameworks provide a good starting point for understanding the relationship of visual and descriptive models in data visualization systems. However, these frameworks do not address the specific needs of the current project and do not provide a comprehensive framework for the design and implementation of models. (13)

Modeling Framework

Modeling Framework

When developing a model of a real-world system, many critical decisions must be made regarding model components, their behavior, the environment, and system dynamics evolving over time. Any model designed should start with a specification of stakeholders and their needs, followed by phenomena of interest, and finally the success criteria that define when a model is fit-for-purpose. Model validation and communication must be elaborated. Various approaches have been proposed to provide templates and standards for systematic model development and documentation—important for ensuring replicability of results. This paper reviews prior work on modeling frameworks and then introduces and explains the data visualization approach in a new Atlas of Iconology, Part 2, to cover the emergent phenomena discussed in the previous section, as well as the atlas, descriptive, predictive, and definitive models throughout the Atlas of Iconology. We cannot stop the march of history, but we can influence its direction.

Prior Work

There exist many frameworks that aim to guide researchers and experts in the design, use, evaluation, and validation of models. While many experts, researchers, and practitioners have developed various frameworks, a few main themes can be identified: (1) The ISO/IEC COCOMO II model has developed and documented a process model used by software developers that describes phases in projects (4); (2) Initials (common definition, (5) website dedicated to modeling, (6) COCOMO II; (7) development of COCOMO II; (8) and (9) mandatory models of model-based (10) processes and (11). A common theme among these frameworks is the qualitative and quantitative relationship between different stakeholders and model components, as well as the communication and documentation of models. (12) It is clear that these frameworks provide a good starting point for understanding the relationship of visual and descriptive models in data visualization systems. However, these frameworks do not address the specific needs of the current project and do not provide a comprehensive framework for the design and implementation of models. (13)

Modeling Framework

Modeling Framework

When developing a model of a real-world system, many critical decisions must be made regarding model components, their behavior, the environment, and system dynamics evolving over time. Any model designed should start with a specification of stakeholders and their needs, followed by phenomena of interest, and finally the success criteria that define when a model is fit-for-purpose. Model validation and communication must be elaborated. Various approaches have been proposed to provide templates and standards for systematic model development and documentation—important for ensuring replicability of results. This paper reviews prior work on modeling frameworks and then introduces and explains the data visualization approach in a new Atlas of Iconology, Part 2, to cover the emergent phenomena discussed in the previous section, as well as the atlas, descriptive, predictive, and definitive models throughout the Atlas of Iconology. We cannot stop the march of history, but we can influence its direction.
Model Visualization

Model assumptions, design, and results should together be communicated in a format that is appropriate for a wide range of modeling stakeholders and experts. Visualizations can help domain, modeling, and programming experts to collaborate closely in the conceptualization and development of model setup and run, the impact of different parameter values on model results—

including emergent phenomena—can be visually explored. Furthermore, such visualizations may help stakeholders compare and interpret model results, and then communicate them to experts in general audiences. Visualizations can be static, dynamic, or interactive.

The linkage of sophistication is simplicity. 

Charles B. Todd

Visualizations Types

The design of effective visualizations requires identifying image needs, and prioritizing, selecting the appropriate data, model, visual design, and visualization types, performing an analysis of alternative visual tactics in graphic symbols, as well as to choose specific designs. Visualizations tend to involve detection, communication, and interpretation of the data, and are expanded to Modeling Framework (page 8).

In Model Impact Strategies, 2015, modeling when applied to solve complex problems involving big data, visual modeling, algorithms, data science, and machine learning. It is important to note that all models have some form of understanding of model quality, and the accuracy of the data. Visualizations vary in their capacity to help communicate model assumptions, model design, simulation results, or model comparison results. They may be a positive or negative function of a considerable number of large-scale model components and state variables, variable interaction, and hierarchies of understanding, and to compare multiple model views or model types. Hence, the specific visualizations we choose to represent the statistical model in the first place should be identified and explained.

This general process of visualizing types and prediction, modeling, or model comparison, visualization to model, or model comparison, visualization to model design, and model visualization to model composition, design, and visualization is represented on model factors are listed on page 37 of 2015.

Model Conceptualization

The concept model diagram depicts the model conceptualization that makes model conceptualization more visible. There are different types of visualizations that can be used to support this task.

Model Visualization

This is a set of LTEV variables at one level, then we are to define the mean, and standard deviation, in order to sample the values and the distribution of values. The model conceptualization that makes model conceptualization more visible. There are different types of visualizations that can be used to support this task.

Model Visualization

This is a set of LTEV variables at one level, then we are to define the mean, and standard deviation, in order to sample the values and the distribution of values. The model conceptualization that makes model conceptualization more visible. There are different types of visualizations that can be used to support this task.

Model Visualization

This is a set of LTEV variables at one level, then we are to define the mean, and standard deviation, in order to sample the values and the distribution of values. The model conceptualization that makes model conceptualization more visible. There are different types of visualizations that can be used to support this task.

Conceptual Circles

This model includes and illustrates the major components of a target system using a paper model or digital maps. The model example shows how major components, written in small pieces of paper, are placed around the outside of a ring-like shape according to their similarity. Similar components can be identified in similar ones to recognize structural and relationship diagrams. Particularly important concepts can be highlighted or emphasized. Lines of different colors can be used to display different types of relationship components.

Model Design Visualizations

The structures and functions of models can be characterized using conceptual models (cause model loops), mathematical functions, and input-output sub-processes. Graphical symbols include the rectangles that represent the mechanism of logical operations, with arrows showing the relationships between blocks. Each block is linked to a single input, output, and transfer function, for example the input of the process and transfer functions. A take of power output is a single output, or the power of multiple blocks or meaning transfer functions. Each removing input has one more impact and a single output, or it produces the deviation of the input or output. Inputs and outputs here are due to be identified as the input or output. Each removing input has one more impact and a single output, or it produces the deviation of the input or output. Inputs and outputs here are due to be identified as the input or output.

Behavior-Over-Time Graphs

The behavior-Over-Time Graphs show how the components of a target system change over time, such as the example in 1999 to 2015 (page 37) and the graph below.

Linear and Circular Structures

The linear model allows for processes, patterns, and structures to be identified, and for changes in mental models, or, conceptual structure, to be productively discussed.

MLV Diagrams

MLV diagrams show the main components of a target system using a paper model or digital maps. Each block is connected to a single input, output, and transfer function, for example the input of the process and transfer functions. A take of power output is a single output, or the power of multiple blocks or meaning transfer functions. Each removing input has one more impact and a single output, or it produces the deviation of the input or output. Inputs and outputs here are due to be identified as the input or output. Each removing input has one more impact and a single output, or it produces the deviation of the input or output. Inputs and outputs here are due to be identified as the input or output. Each removing block is connected to a single output. Each removing input has one more impact and a single output, or it produces the deviation of the input or output. Inputs and outputs here are due to be identified as the input or output.

Linear and Circular Structures

The linear model allows for processes, patterns, and structures to be identified, and for changes in mental models, or, conceptual structure, to be productively discussed.
Model Validation

Models should aim to capture the behavior of real-world systems in a simple yet usable manner that can be validated across scales. At the micro level, the type and behavior of individual components (e.g., agents for agent-based models or nodes for network models) need to match up with their real-world counterparts. At the macro level, the aggregate, emergent properties of the models, oscillations or adjectives, must reflect the phenomenon observed in the real world. Models must be evaluated based on the accuracy and generality of their predictions. Evaluation metrics need to be used to assess the accuracy, specificity, or generality of the models, or to make model results easier to understand and use by decision-makers.

The more any quantitative social indicator is used for social decision-making, the more subject it is to corruption, particularly the more its use is to be evaded and covert the social interest it is intended to monitor.

Donald E. Campbell

Quality Assurance Framework

In a complex system (e.g., a policy, a market, or a network), every component (e.g., a node, a decision, or a transaction) interacts with others and can be monitored. The model is robust if its outward behavior matches its intended behavior. Robustness is a property that can be quantified in the form of confidence intervals or worst-case scenarios. Robustness is a measure of the ability of a system to maintain its performance under varying conditions.

Model Robustness

The examination of a model's sensitivity to changes in the model's parameters and assumptions is a critical step in model validation. Robustness analysis helps identify the critical parameters and assumptions that affect the model's predictions and can be used to improve the model's accuracy and reliability.

For simple models with low parameter sensitivity, one can use simulation modeling techniques, such as Monte Carlo simulations, to assess the model's performance across a range of parameter values. For more complex models, higher-order methods, such as sensitivity analysis and Monte Carlo simulations, can be used to assess the model's performance under varying conditions.

Model Replication

Reproduction of the model is achieved by reproducing the output of the model under the same conditions. If the output is the same as the original output, the model is considered to be reproducible. Reproducibility is a critical aspect of model validation, as it allows other researchers to verify the results and reproduce the findings.

Model Replication Test

The model replication test is conducted by running the model under the same conditions as the original test and comparing the results. If the output is the same, the model is considered to be reproducible. If the output differs, the model is considered to be non-reproducible.

Model Replication Test

The model replication test is conducted by running the model under the same conditions as the original test and comparing the results. If the output is the same, the model is considered to be reproducible. If the output differs, the model is considered to be non-reproducible.

Model Replication Test

The model replication test is conducted by running the model under the same conditions as the original test and comparing the results. If the output is the same, the model is considered to be reproducible. If the output differs, the model is considered to be non-reproducible.

Model Robustness

The examination of a model's sensitivity to changes in the model's parameters and assumptions is a critical step in model validation. Robustness analysis helps identify the critical parameters and assumptions that affect the model's predictions and can be used to improve the model's accuracy and reliability.
Cellular Automata (1940s)

Cellular automata (CA) are mathematical models that can be used to simulate complex systems or processes. CAs are applied in several fields—including biology, physics, and computer science—to analyze artificial systems such as artificial life, plant growth, or epithelial cells. CAs consist of elements called cells. Each cell has a value, or state. Cells are connected to certain neighboring cells to form a one- or multidimensional lattice. Cell state changes at discrete times using a set of predefined rules that take the previous states of connected neighboring cells into account.

Brief History

Cellular automata were developed by John von Neumann and Stanislaw Ulam in the 1940s. They were initially used in self-replicating machines, as in the Brac-Ballistic Model of Ulam's Game of Life. The game was first proposed in the 1950s. Later, Neumann automata became a popular modeling framework for exploring emergent behaviors and the dynamics underlying complex systems in domains such as computer science, biology, and physics.

Terminology

The term "cellular automaton" is a general term for a dynamical system with a discrete number of states and a discrete time evolution. The states are distributed on a regular grid. The grid is usually two-dimensional, but it can also be one-dimensional. The states of the cells can be observed in real time, and the system can evolve according to predefined rules.

Key Insights

CAs are used extensively for modeling phenomena with a cellular or discrete structure. They are useful for modeling systems with spatial and temporal evolution, such as biological systems, physical processes, and social systems. CAs are also used in computer science, artificial intelligence, and game theory.

Phenomena

- Oscillation
- Synchronization
- Propagation
- Phase Transition
- Self-Organized Criticality (SOC)
- Pattern Formation
- Adaptation & Learning
- Reaction-Diffusion Dynamics
- Network Growth
- Network Geometers
- Network Attack and Error
- Diffusions Spreading

Model Classes

- Predator-Prey Model
- Troyeber's Game Model
- Moran Chaos Model
- McKendrick-Pettigrew Model
- Eden Growth Model
- Schelling's Segregation Model
- Prisoner's Dilemma Model
- Buckley-Landau Model
- Axiter-Richardson Model
- Keller-Levine Model
- Machine Learning Models

Target System Models

- Predator-Prey Model (1920s)
- Troyeber's Game Model (1960s)
- Moran Chaos Model (1970s)
- McKendrick-Pettigrew Model (1927)
- Eden Growth Model (1960s)
- Schelling's Segregation Model (1971)
- Prisoner's Dilemma Model (1980s)
- Buckley-Landau Model (1978)
- Axiter-Richardson Model (1979)
- Keller-Levine Model (1980s)
- Machine Learning Models (1980s)

In 1971, the economist Thomas C. Schelling showed that individual bids can lead to collective behavior. He was awarded the Nobel Prize in 2005 for the Nobel Prize in Economics for his work on dynamic models of conflict and cooperation in the context of the spread of ideas and of international conflict. Schelling's model has been used to study the spread of diseases, the evolution of language, and the dynamics of social segregation.
Geospatial models—"Where"

Geospatial position and context are significant. Location matters to biodiversity, with the choice of model and methodology mattering to how a model approaches using location in model prediction. Models can vary by the data used, model approach, and reference data. Many models use a more specific area or space, with geospatial data used in models and references often in contexts and comparison to others using the same data.

Topical Models—"What"

Topical models are specific to a domain. They are specific to a field of study and are more focused in their application. These models are specific to a specific area or topic, such as climate change, biodiversity, or urban planning. They are specific to a specific set of data and are more focused in their application. These models are specific to a specific area or topic, such as climate change, biodiversity, or urban planning. They are specific to a specific set of data and are more focused in their application.

Network Models—"With Whom"

Network models are used to study relationships and interactions between entities. They are used to study systems that are made up of interacting parts. These models are specific to a specific set of data and are more focused in their application. These models are specific to a specific area or topic, such as climate change, biodiversity, or urban planning. They are specific to a specific set of data and are more focused in their application.

Historical Progression

Communicating Models

- Network models
- Geospatial models
- Topical models

Timeline

- 1973: Introduction of geographical data
- 1984: Development of geographic information systems
- 1990s: Emergence of spatial modeling
- 2000s: Integration of network models

Historical Milestones

- 1973: Introduction of geographical data
- 1984: Development of geographic information systems
- 1990s: Emergence of spatial modeling
- 2000s: Integration of network models

Future Directions

- Integration of social and environmental factors
- Development of advanced computational techniques
- Use of big data and artificial intelligence

References


Appendix

- Table of network models
- Glossary of terms
- Further reading
Reducing Human Bias

Humans tend to be subjective, often acting according to biased opinions rather than objective judgment. Cognitive biases are systematic deviations from normative rationality and self-regulation in judgment, as well as field of psychology and behavioral economics. While many such biases have been confirmed in independently reproducible research, concerns about their prevalence abound in the field of psychology. In order to make objective, well-informed decisions, we need to understand and proactively resist these biases. Reducing bias in science, business, and society requires, with behaviors, and suggestions, with origins for how to correct them. Ultimately, biases and behaviors may have a major impact on life satisfaction. Understanding our own biases is an essential step in correcting a fulfilling present and future.

All models are wrong, but some are useful.
George Box

To Erze Is Human

Though human beings have the capacity for understanding and self-control, human errors are inherent in our ability to understand, interpret, and communicate with the world. The level of error is frequently, specifically, individualized, and culturally-mediated (a high level of individualism can greatly influence the structure and style of the data). While perfect data is not possible, bias in research can be managed by using statistical techniques. To understand the data, biases must be understood. The level of error is frequently, specifically, individualized, and culturally-mediated. The level of error is frequently, specifically, individualized, and culturally-mediated. The level of error is frequently, specifically, individualized, and culturally-mediated.

Data Bias

Every system receives certain data, which is gathered by sensors that monitor, record, and store data on various systems. This data is then analyzed and interpreted by the researchers. However, the data is not always accurate or complete. Therefore, it is important to understand and interpret the data correctly. The level of error is frequently, specifically, individualized, and culturally-mediated. The level of error is frequently, specifically, individualized, and culturally-mediated. The level of error is frequently, specifically, individualized, and culturally-mediated. The level of error is frequently, specifically, individualized, and culturally-mediated.

Behavioral Bias

Behavioral bias is a type of error that occurs when people do not follow the instructions provided. Behavioral bias can occur when people do not follow the instructions provided. Behavioral bias can occur when people do not follow the instructions provided. Behavioral bias can occur when people do not follow the instructions provided. Behavioral bias can occur when people do not follow the instructions provided.

Self-Perpetuating Bias

As described in the previous section, self-perpetuating bias is a phenomenon that occurs when people do not follow the instructions provided. Behavioral bias can occur when people do not follow the instructions provided. Behavioral bias can occur when people do not follow the instructions provided. Behavioral bias can occur when people do not follow the instructions provided. Behavioral bias can occur when people do not follow the instructions provided.

Exposing Biases

People tend to resist changes in their own biases and behaviors. This can lead to a cycle of self-perpetuating bias, where people do not change their behavior because they do not want to admit that they have a bias. This can lead to a cycle of self-perpetuating bias, where people do not change their behavior because they do not want to admit that they have a bias. This can lead to a cycle of self-perpetuating bias, where people do not change their behavior because they do not want to admit that they have a bias. This can lead to a cycle of self-perpetuating bias, where people do not change their behavior because they do not want to admit that they have a bias.
The Future of Learning & Work Workshop

Open Digital Future. Perspectives on data at the intersection of education and job markets. Toward a new role of visual and learning analytics.

https://cns-iu.github.io/workshops/2022-03-14-futurium
Visualizing big science projects

Katy Borner, Filip Nascimento Silva and Staša Milojević

Abstract | The number, size and complexity of 'big science' projects are growing—and are increasingly recognized as major contributors to basic research and innovation. It is therefore essential to develop effective ways of visualizing the complex networks and collaborations that underpin these projects. This is especially true for large-scale projects, such as the Large Hadron Collider, which require the integration of data from many different sources. The authors describe a tool called bigscience.js, which is designed to help researchers visualize and understand the complex relationships within big science projects.

Big science is the development of large-scale scientific projects, such as the Large Hadron Collider, which require the integration of data from many different sources. The authors describe a tool called bigscience.js, which is designed to help researchers visualize and understand the complex relationships within big science projects.

Big science is the development of large-scale scientific projects, such as the Large Hadron Collider, which require the integration of data from many different sources. The authors describe a tool called bigscience.js, which is designed to help researchers visualize and understand the complex relationships within big science projects.

Big science is the development of large-scale scientific projects, such as the Large Hadron Collider, which require the integration of data from many different sources. The authors describe a tool called bigscience.js, which is designed to help researchers visualize and understand the complex relationships within big science projects.

Big science is the development of large-scale scientific projects, such as the Large Hadron Collider, which require the integration of data from many different sources. The authors describe a tool called bigscience.js, which is designed to help researchers visualize and understand the complex relationships within big science projects.

Big science is the development of large-scale scientific projects, such as the Large Hadron Collider, which require the integration of data from many different sources. The authors describe a tool called bigscience.js, which is designed to help researchers visualize and understand the complex relationships within big science projects.

Big science is the development of large-scale scientific projects, such as the Large Hadron Collider, which require the integration of data from many different sources. The authors describe a tool called bigscience.js, which is designed to help researchers visualize and understand the complex relationships within big science projects.
Indiana University Bloomington will host the International Society of Scientometrics & Informetrics Conference (ISSI) July 2-5, 2023
https://cns-iu.github.io/workshops/2023-07-02_issi/
24 Hour Science Map Event

https://cns-iu.github.io/workshops/2021-12-10_24hour_science_map

Dec 11, noon - Dec 12, noon ET, 2021
24 Hour Human Reference Atlas Event
Let's map the human body at single-cell resolution!

Dec 10, noon – Dec 11, noon ET, 2022

https://humanatlas.io/events/2022-24h
Overview

- Data Visualizations of Science
- The Science of Data Visualization
- Open Challenges
Data Visualization Literacy


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.
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.

---

### DVL Typology

- **Insight Needs**
  - categorize/cluster
  - order/arrange/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
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

## Typology of the Data Visualization Literacy Framework

<table>
<thead>
<tr>
<th>Insight Needs</th>
<th>Data Scales</th>
<th>Analyses</th>
<th>Visualizations</th>
<th>Graphic Symbols</th>
<th>Graphic Variables</th>
<th>Interactions</th>
</tr>
</thead>
<tbody>
<tr>
<td>• categorize/cluster</td>
<td>• nominal</td>
<td>• statistical</td>
<td>• table</td>
<td>• geometric symbols</td>
<td>• spatial</td>
<td>• zoom</td>
</tr>
<tr>
<td>• order/rank/sort</td>
<td>• ordinal</td>
<td>• temporal</td>
<td>• chart</td>
<td>point</td>
<td>position</td>
<td>• search and locate</td>
</tr>
<tr>
<td>• distributions (also outliers, gaps)</td>
<td>• interval</td>
<td>• geospatial</td>
<td>• graph</td>
<td>line</td>
<td></td>
<td>• filter</td>
</tr>
<tr>
<td>• comparisons</td>
<td>• ratio</td>
<td>• topical</td>
<td>• map</td>
<td>area</td>
<td></td>
<td>• details-on-demand</td>
</tr>
<tr>
<td>• trends (process and time)</td>
<td></td>
<td>• relational</td>
<td>• tree</td>
<td>surface</td>
<td></td>
<td>• history</td>
</tr>
<tr>
<td>• geospatial</td>
<td></td>
<td></td>
<td>• network</td>
<td>volume</td>
<td></td>
<td>• extract</td>
</tr>
<tr>
<td>• compositions (also of text)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• link and brush</td>
</tr>
<tr>
<td>• correlations/relationships</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• projection</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>selection</td>
<td>categorize</td>
<td>category</td>
<td></td>
<td></td>
<td></td>
<td>categorize/cluster</td>
<td>order/rank/sort</td>
<td>table</td>
</tr>
<tr>
<td>order</td>
<td>rank</td>
<td>ranking</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>distributions (also outliers, gaps)</td>
<td>order/rank/sort</td>
</tr>
<tr>
<td>distribution</td>
<td>distribution</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>distribution</td>
<td></td>
<td></td>
</tr>
<tr>
<td>compare</td>
<td>nominal comparison &amp; deviation</td>
<td>differences</td>
<td>compare and contrast</td>
<td>compare data values</td>
<td>comparison</td>
<td>comparisons</td>
<td></td>
<td></td>
</tr>
<tr>
<td>time series</td>
<td>patterns over time</td>
<td>time</td>
<td>process and time</td>
<td>track rises and falls over time</td>
<td>trend</td>
<td>trends (process and time)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>geospatial</td>
<td>spatial relations</td>
<td>location</td>
<td></td>
<td>generate maps</td>
<td>geospatial</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>quantity</td>
<td>part-to-whole</td>
<td>proportions</td>
<td>form and structure</td>
<td>see parts of whole, analyze text</td>
<td>composition</td>
<td>compositions (also of text)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>association</td>
<td>correlate</td>
<td>correlation</td>
<td>relationships</td>
<td>hierarchy</td>
<td>relationships between data points</td>
<td>relationship</td>
<td>correlations/relationships</td>
<td></td>
</tr>
</tbody>
</table>
## Typology of the Data Visualization Literacy Framework

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

Visualization Types

Chart
- **Pie Chart**
- **Bubble Chart**

Graph
- **Scatter Graph**
- **Temporal Bar Graph**

Map
- **Choropleth Map**
- **Proportional Symbol Map**

Tree
- **Dendrogram**
- **Tree Map**

Network
- **Force-Directed Network Layout**
- **Bimodal Network Layout**
Visualize: Reference Systems

**Table**
columns by rows

**Graph**
x-y coordinates

**Map**
latitude/longitude

**Network**
local similarity

**Visualization Types**
- table
- chart
- graph
- map
- network layout
Visualize: Reference Systems, Graphic Symbols and Variables

Data Overlay

Reference System

Graphic Variables

Graphic Symbols

Visualizations

Scatter Graph Geospatial Map UCSD Science Map Network
### Typology of the Data Visualization Literacy Framework

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

---

Typology of the Data Visualization Literacy Framework

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

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

<table>
<thead>
<tr>
<th>Geometric Symbols</th>
<th>Linguistic Symbols</th>
<th>Pictorial Symbols</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Point</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Line</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Spatial Position
- **X**
- **Y**

#### Form
- **Size**: 
  - ![Symbol](image)
  - ![Symbol](image)
  - ![Symbol](image)
- **Shape**: 
  - ![Symbol](image)
  - ![Symbol](image)
  - ![Symbol](image)

#### Quantitative
- **Value**: 
  - ![Symbol](image)
  - ![Symbol](image)
  - ![Symbol](image)

#### Qualitative
- **Hue**: 
  - ![Symbol](image)
  - ![Symbol](image)
  - ![Symbol](image)
  - ![Symbol](image)
  - ![Symbol](image)
- **Saturation**
  - ![Symbol](image)
  - ![Symbol](image)
  - ![Symbol](image)

#### Graphic Variable Types
- **Color**: 
  - ![Symbol](image)
- **Granularity**
  - ![Symbol](image)
  - ![Symbol](image)
  - ![Symbol](image)
- **Texture**: 
  - ![Symbol](image)
  - ![Symbol](image)
  - ![Symbol](image)
- **Pattern**
  - ![Symbol](image)
  - ![Symbol](image)
  - ![Symbol](image)

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

- **Qualitative**
  - Also called: Categorical Attributes
  - Identity Channels

- **Quantitative**
  - Also called: Ordered Attributes
  - Magnitude Channels
<table>
<thead>
<tr>
<th>Graphic Variable Types</th>
<th>Vowel</th>
<th>Consonant</th>
<th>Geometric Symbol</th>
<th>Surface</th>
<th>Volume</th>
<th>Linguistic Symbol</th>
<th>Picture Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graphic Variable Types</td>
<td>Vowel</td>
<td>Consonant</td>
<td>Geometric Symbol</td>
<td>Surface</td>
<td>Volume</td>
<td>Linguistic Symbol</td>
<td>Picture Symbol</td>
</tr>
</tbody>
</table>

US Employers which have sent students include The Boeing Company, Eli Lilly, DOE, CDC, NSWC Crane.
Overview

- Data Visualizations of Science
- The Science of Data Visualization
- Open Challenges
Scientific ontologies are systems and/or knowledge structures that specify concepts of science with agreed-upon labels and definitions and provide a framework for complex relationships among the concepts. Ontologies support efficient knowledge generation, organization, reuse, integration, and analysis. The goal of this consensus study is to review the role of ontologies in the behavioral sciences, assess their potential to accelerate behavioral science research, and identify gaps and challenges, and offer conclusions and recommendations for strengthening behavioral ontologies.
Envisioning SPOKE: 3M Nodes and 30M Edges

The Scalable Precision Medicine Oriented Knowledge Engine (SPOKE) graph federates about 19 open datasets into a public data commons of health-relevant knowledge. This site lets users explore the massive SPOKE knowledge graph.

The site was designed for two user groups: (1) novice users interested in understanding the coverage and quality of SPOKE data and (2) expert users interested in analyzing and optimizing the interlinked knowledge graphs in SPOKE. The overview visualization shows the different entity types and their diverse interrelations.

This project is funded by NSF award 2033093.

https://cns-iu.github.io/spoke-vis
Anatomical structures, cell types and biomarkers of the Human Reference Atlas


The Human Reference Atlas (HRA) aims to map all of the cells of the human body to advance biomedical research and clinical practice. This Perspective presents collaborative work by members of 16 international consortia on two essential and interconnected parts of the HRA: (1) three-dimensional representations of anatomy that are linked to (2) tables that name and interlink major anatomical structures, cell types, plus biomarkers (ASCT-B). We discuss four examples that demonstrate the practical utility of the HRA.

With developments in massively parallel sequencing in bulk and at the single-cell level, researchers can now detect genomic features and genome expression with great precision. Profiling single cells within tissues and organs enables researchers to map the distribution of cells and their developmental trajectories across organs and gives indications as to their functions. In 2021, there are several ongoing, ambitious efforts to map all of the cells in the human body and to create a digital reference atlas of the human body. The final atlas will encompass the threedimensional (3D) organization of whole organs and all of the anatomical structures, the interdependencies between trillions of cells, and the biomarkers that characterize and distinguish cell types. It will make the human body computable, supporting spatial and semantic queries run over 3D structures linked to their scientific terminology and existing ontologies. It will establish a benchmark reference that helps us to understand how the healthy human body works and what changes during ageing or disease.

A network of 16 consortia is contributing to the construction of the HRA based on studies of 30 organs (Fig. 1a) with four

The 16 consortia include the Allen Brain Atlas, the Brain Research through Advancing Innovative Neurotechnologies Initiative—Cell Census Network Initiative, the Chan Zuckerberg Initiative Seed Networks for HCA, HCA awards by the EU’s Horizon 2020 program, the Genotype-Tissue Expression project, the Genitourinary Developmental Molecular Anatomy Project, the Wellcome Charitable Trust: Gut Cell Atlas, the Human Tumor Atlas Network, the Human Biomolecular Atlas Program (HuBMAP), the Kidney Precision Medicine Project (KPMP), LungMAP, HCA grants from the United Kingdom Research and Innovation Medical Research Council (https://mrc.ukri.org), (Re)building the Kidney, Stimulating Peripheral Activity to Relieve Conditions, The Cancer Genome Atlas, and Wellcome funding for HCA pilot projects.Z

In total, more than 2,600 experts from around the globe are working together to construct an open-source and free-to-use digital HRA using a wide variety of single or multimodal spatially resolved and bulk tissue assays. Imaging methods for anatomical structure segmentation include computed tomography, magnetic resonance imaging or optical coherence tomography (OCT).

Fig. 1 | Components and construction of the HRA. a, Alphabetical listing of 16 HRA construction efforts (left) linked to the 30 human organs that they study (right). The lungs are studied by ten consortia (orange links). This review focuses on ten organs (bold) plus vasculature. BCCN, Brain Research through Advancing Innovative Neurotechnologies Initiative—Cell Census Network Initiative; C2I, Chan Zuckerberg Initiative—H2C2I; Horizon 2020; GUTX, Genotypos-Tissue Expression project; GUDMAP, Genito-Urinary Developmental Molecular Anatomy Project; HTAAN, Human Tumor Atlas Network; MRC, Medical Research Council-RIK; (Re)building the Kidney; SPARC, Stimulating Peripheral Activity to Relieve Conditions; TCGA, The Cancer Genome Atlas. b, The 3D reference objects for major anatomical structures were jointly developed for 31 organs. c, An example ASCT-B+ table showing anatomical structures (AS) and cell types (CT) and some biomarkers (B) for the kidney, together with the names of the three entity types (anatomical structures, cell types and biomarkers) and four relationship types (part_of, is_a, located_in and characterizes). Note that the is_a relationship exists for cell types and biomarkers.

https://www.nature.com/articles/s41556-021-00788-6
Indiana University Bloomington will host the **International Society of Scientometrics & Informetrics Conference (ISSI)**
July 2-5, 2023

https://cns-iu.github.io/workshops/2023-07-02_issi/